

Tuesday, June 6

AASP-L1: Audio for Multimedia and Multimodal Processing

Room: Salon de Roses A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Mark Plumbley, Lin Wang

10:50 AM

647 (AASP-L1.1): Diverse and Vivid Sound Generation from Text Descriptions

Guangwei Li (Shanghai Jiao Tong University); Xuenan Xu (Shanghai Jiao Tong University); Lingfeng Dai (Shanghai Jiao Tong University); Mengyue Wu (Shanghai Jiao Tong University); Kai Yu (Shanghai Jiao Tong University)

11:05 AM

2248 (AASP-L1.2): EPIC-SOUNDS : A LARGE-SCALE DATASET OF ACTIONS THAT SOUND

Jaesung Huh (University of Oxford); Jacob Chalk (University of Bristol); Evangelos Kazakos (Dept. of Computer Science and Engineering - University of Ioannina); Dima Damen (University of Bristol); Andrew Zisserman (University of Oxford)

11:20 AM

784 (AASP-L1.3): I SEE WHAT YOU HEAR: A VISION-INSPIRED METHOD TO LOCALIZE WORDS

Mohammad Samragh (Apple); Arnav Kundu (Apple); Ting-Yao Hu (Carnegie Mellon University); Aman Chadha (Stanford University/Amazon Inc.); Ashish Shrivastava (Apple); Minsik Cho (Apple); Oncel Tuzel (Apple); Devang Naik (Apple)

11:35 AM

6119 (AASP-L1.4): Incorporating lip features into audio-visual multi-speaker DOA estimation by gated fusion

Ya Jiang (University of Science and Technology of China); Hang Chen (USTC); Jun Du (University of Science and Technology of China); Qing Wang (University of Science and Technology of China); Chin-Hui Lee (Georgia Institute of Technology)

11:50 AM

6787 (AASP-L1.5): UAVM: Towards Unifying Audio and Visual Models (SPS Journal Paper)*

Yuan Gong (Massachusetts Institute of Technology); Alexander H Liu (MIT); Andrew Rouditchenko (MIT CSAIL); James Glass (Massachusetts Institute of Technology)

GC-1: Drone-vs-Bird Detection Grand Challenge at ICASSP23

Room: Nefeli B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): TBA

10:50 AM

6617 (GC-L1.1): Introduction

Angelo Coluccia (University of Salento); Alessio Fascista (University of Salento); Arne Schumann (Fraunhofer IOSB); Lars Sommer (Fraunhofer IOSB, Karlsruhe, Germany); Anastasios Dimou (Information Technologies Institute / Centre for Research and Technology Hellas); Dimitrios Zarpalas (CERTH / CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS); Nabin Sharma (University of Technology, Sydney)

11:10 AM

6834 (GC-L1.2): HIGH-SPEED DRONE DETECTION BASED ON YOLO-V8

JUN-HWA KIM (Dongguk University); Namho KIM (Dongguk University); Chee Sun Won (Dongguk University)

11:22 AM

6863 (GC-L1.3): S-FEATURE PYRAMID NETWORK AND ATTENTION MODEL FOR DRONE DETECTION

Pengcheng Dong (Shandong Normal University); Chuntao Wang (Shandong Normal University); Zhenyong Lu (Shandong Normal University); Kai Zhang (Shandong Normal University); Wenbo Wan (Shandong Normal University); Jiande Sun (Shandong Normal University)

11:34 AM

6881 (GC-L1.04): DRONE-VS-BIRD: DRONE DETECTION USING YOLOV7 WITH CSRT TRACKER

Sahaj K Mistry (Indian Institute of Technology Jammu); Shreyas Chatterjee (Indian Institute of Technology Jammu); Ajeet Kumar Verma (Indian Institute of Technology Jammu); Vinit Jakhetiya (IIT JAMMU); Badri Subudhi (Indian Institute of Technology, Jammu); Sunil Jaiswal (K|Lens GmbH)

IVMSP-L1: Human Identification and Face Recognition

Room: Athina

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Mang Ye, Lizhuang Ma

10:50 AM

530 (IVMSP-L1.1): EMCLR: Expectation Maximization Contrastive Learning Representations

Meng Liu (Shanghai Jiao Tong University); Ran Yi (Shanghai Jiao Tong University); Lizhuang Ma (Shanghai Jiao Tong University)

Tuesday, June 6

11:05 AM

711 (IVMSP-L1.2): BOOSTING PERSON RE-IDENTIFICATION WITH VIEWPOINT CONTRASTIVE LEARNING AND ADVERSARIAL TRAINING

Xingyue Shi (Peking University Shenzhen Graduate School); Hong Liu (Peking University Shenzhen Graduate School); Wei Shi (Peking University Shenzhen Graduate School); Zihui Zhou (Peking University, Shenzhen Graduate School); Yidi Li (Peking University Shenzhen Graduate School)

11:20 AM

812 (IVMSP-L1.3): TOP-K VISUAL TOKENS TRANSFORMER: SELECTING TOKENS FOR VISIBLE-INFRARED PERSON RE-IDENTIFICATION

Bin Yang (Wuhan University); Jun Chen (Wuhan University); Mang Ye (Wuhan University)

11:35 AM

2531 (IVMSP-L1.4): FREQUENCY-AWARE ATTENTIONAL FEATURE FUSION FOR DEEPFAKE DETECTION

Cheng Tian (Xiamen University); Zhiming Luo (Xiamen University); Guimin Shi (Wuyi University); Shaozi Li (Xiamen University, China)

11:50 AM

5309 (IVMSP-L1.5): RECURSIVE JOINT ATTENTION FOR AUDIO-VISUAL FUSION IN REGRESSION BASED EMOTION RECOGNITION

Gnana Praveen Rajasekhar (Ecole Technologie Supérieure); Eric Granger (ETS Montreal); Patrick Cardinal (École de technologie supérieure)

12:05 PM

3475 (IVMSP-L1.06): Multi-Stream Facial Adaptive Network for Expression Recognition from A Single Image

Baichuan Zhang (Sun Yat-sen University); Fanyang Meng (Peng Cheng Laboratory); Runwei Ding (Peking University Shenzhen Graduate School); Mengyuan Liu (Peking University, Shenzhen Graduate School)

MLSP-L1: Self-supervised Learning Methods I

Room: Salon de Roses B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Zheng-Hua Tan,

10:50 AM

429 (MLSP-L1.1): PointACL:Adversarial Contrastive Learning for Robust Point Clouds Representation under Adversarial Attack

Junxuan Huang (University at Buffalo); Junsong Yuan ("State University of New York at Buffalo, USA"); Chunming Qiao (University at Buffalo); yatong an (xmotors); Cheng Lu (Xiaopeng); Chen Bai (Xpeng Motors)

11:05 AM

2579 (MLSP-L1.2): Enhancing Representation Learning with Deep Classifiers in Presence of Shortcut#

Amirhossein Ahmadian (Linköping University); Fredrik Lindsten (Linköping University)

11:20 AM

730 (MLSP-L1.3): K2NN: Self-supervised Learning with Hierarchical Nearest Neighbors for Remote Sensing

Jianlong Yuan (Alibaba Group); Yuanhong Xu (Alibaba Group); Zhibin Wang (Alibaba Group)

11:35 AM

4453 (MLSP-L1.4): TRINET: STABILIZING SELF-SUPERVISED LEARNING FROM COMPLETE OR SLOW COLLAPSE

Lixin Cao (Tencent); Jun Wang (Tencent); ben yang (Peking University); Dan Su (Tencent); Dong Yu (Tencent AI Lab)

11:50 AM

1629 (MLSP-L1.5): On minimal variations for unsupervised representation learning

Vivien A Cabannes (FAIR); Alberto Bietti (Inria); Randall Balestriero (Facebook AI Research)

12:05 PM

740 (MLSP-L1.6): ADAPTIVE DATA AUGMENTATION FOR CONTRASTIVE LEARNING

Yuhan Zhang (Brainnetome Center and NLPR, Institute of Automation, Chinese Academy of Sciences; School of Artificial Intelligence, University of Chinese Academy of Sciences(UCAS)); He Zhu (Brainnetome Center and NLPR; School of Future Technology, UCAS; University of Chinese Academy of Sciences; Institute of Automation, Chinese Academy of Sciences); Shan Yu (Brainnetome Center and NLPR;University of Chinese Academy of Sciences;CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences;)

Tuesday, June 6

SLT-L1: ASR with Constrained Resource

Room: Jupiter

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Hung-yi Lee, Marco Siniscalchi

10:50 AM

690 (SLT-L1.1): DE'HUBERT: DISENTANGLING NOISE IN A SELF-SUPERVISED MODEL FOR ROBUST SPEECH RECOGNITION

Dianwen Ng (Alibaba Group/Nanyang Technological University); Ruixi Zhang (National University of Singapore); Jia Qi Yip (Alibaba Group); Zhao Yang (Xi'an Jiaotong University); Jinjie Ni (Nanyang Technological University); Chong Zhang (Alibaba Group); Yukun Ma (Alibaba Group); Chongjia Ni (Alibaba); Eng Siong Chng (Nanyang Technological University); Bin Ma ("Alibaba, Singapore R&D Center")

11:05 AM

1948 (SLT-L1.2): Masked Token Similarity Transfer for Compressing Transformer-Based ASR Models

Euntae Choi (Seoul National University); Youshin Lim (42dot); Byeong-Yeol Kim (42dot); Hyung Yong Kim (42dot); Hanbin Lee (42dot); Yunkyoo Lim (42dot); Seung Woo Yu (42dot); Sungjoo Yoo (Seoul National University)

11:20 AM

2888 (SLT-L1.3): Unsupervised Fine-Tuning Data Selection for ASR Using Self-Supervised Speech Models

Reem A Gody (The University of Texas at Austin); David Harwath (The University of Texas at Austin)

11:35 AM

3250 (SLT-L1.4): CB-Conformer: Contextual Biasing Conformer for Biased Word Recognition

Yaoxun Xu (Tsinghua University); 刘 裕基 (XVerse); Qiaochu Huang (Tsinghua University); Xingchen Song (Tsinghua University); Zhiyong Wu (Tsinghua University); Shiyin Kang (XVerse Inc.); Helen Meng (The Chinese University of Hong Kong)

11:50 AM

3712 (SLT-L1.5): Context-aware Fine-tuning of Self-supervised speech models

Suwon Shon (ASAPP); Felix Wu (ASAPP); Kwangyoum Kim (ASAPP); Prashant Sridhar (ASAPP); Karen Livescu (TTI-Chicago); Shinji Watanabe (Carnegie Mellon University)

12:05 PM

6449 (SLT-L1.6): DATA2VEC-AQC: SEARCH FOR THE RIGHT TEACHING ASSISTANT IN THE TEACHER-STUDENT TRAINING SETUP

Vasista Sai Lodagala (Indian Institute of Technology, Madras); Sreyan Ghosh (University of Maryland, College Park); S Umesh (IIT Chennai)

SLT-L2: ASR: Multilingual Speech Recognition

Room: Delphi

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Michael Picheny, Tara Sainath

10:50 AM

2417 (SLT-L2.1): Hierarchical Softmax for End-to-End Low-resource Multilingual Speech Recognition

Qianying Liu (Kyoto University); Zhuo Gong (The University of Tokyo); Zhengdong Yang (Kyoto University); Yuhang Yang (School of Information Science and Engineering, Xinjiang University, China); Sheng Li (National Institute of Information & Communications Technology (NICT)); Chenchen Ding (); Nobuaki Minematsu (The University of Tokyo); Hao Huang (Xinjiang University); Fei Cheng (Kyoto University); Chenhui Chu (Kyoto University); Sadao Kurohashi (Kyoto University)

11:05 AM

4510 (SLT-L2.2): Improving Massively Multilingual ASR With Auxiliary CTC Objectives

William Chen (Carnegie Mellon University); Brian Yan (Carnegie Mellon University); Jiatong Shi (Carnegie Mellon University); Yifan Peng (Carnegie Mellon University); Soumi Maiti (CMU); Shinji Watanabe (Carnegie Mellon University)

11:20 AM

4777 (SLT-L2.3): Massively Multilingual Shallow Fusion with Large Language Models

Ke Hu (Google); Tara Sainath (Google); Bo Li (Google); Nan Du (Google Brain); Yanping Huang (Google Brain); Andrew M Dai (Google Brain); Yu Zhang (Google); Rodrigo Cabrera (Google); Zhifeng Chen (Google); Trevor Strohman (Google)

11:35 AM

5465 (SLT-L2.4): UML: A Universal Monolingual Output Layer for Multilingual ASR

Chao Zhang (Tsinghua University); Bo Li (Google); Tara Sainath (Google); Trevor Strohman (Google); Shuo-yiin Chang (Google)

Tuesday, June 6

11:50 AM

5744 (SLT-L2.5): Investigation into phone-based subword units for Multilingual end-to-end speech recognition
Saierdaer Yusuyin (Xinjiang University); Hao Huang (Xinjiang University); Junhua Liu (University of Science and Technology of China); Cong Liu (iFLYTEK Research)

12:05 PM

6221 (SLT-L2.6): Massively Multilingual ASR on 70 Languages: Tokenization, Architecture, and Generalization Capabilities
Andros Tjandra (Meta AI); Nayan Singhal (Facebook); David Zhang (Meta AI); Ozlem Kalinli (Meta AI); Abdelrahman Mohamed (Rembrand Inc); Duc Le (Meta); Michael L. Seltzer (Meta)

SPTM-L1: Adaptive Signal Processing

Room: Nafsika A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Victor Solo, Konstantinos Slavakis

10:50 AM

1224 (SPTM-L1.1): A Compensated Shrinkage Affine Projection Algorithm for Debaised Sparse Adaptive Filtering
Yi Zhang (Tokyo Institute of Technology); Isao Yamada (Tokyo Institute of Technology)

11:05 AM

1761 (SPTM-L1.2): Dynamic Selection of p-Norm in Linear Adaptive Filtering via Online Kernel-Based Reinforcement Learning
Minh Vu (Tokyo Institute of Technology); Yuki Akiyama (Tokyo Institute of Technology); Konstantinos Slavakis (Tokyo Institute of Technology)

11:20 AM

2511 (SPTM-L1.3): Neural Network Models with Integrated Training and Adaptation for Nonlinear Acoustic System Identification
Svantje Voit (Carl von Ossietzky University of Oldenburg); Gerald Enzner (Carl von Ossietzky University Oldenburg)

11:35 AM

3895 (SPTM-L1.4): NEURAL MODE ESTIMATION
peng sun (Zhejiang University of Technology); Zhenyu Wen (Zhejiang University of Technology); Yejian Zhou (Zhejiang University of Technology); Zhen Hong (Zhejiang University of Technology); Tao Lin (Westlake University)

11:50 AM

5352 (SPTM-L1.5): Adaptive ECCM for Mitigating Smart Jammers
Shashwat Jain (Cornell University); Kunal Pattanayak (Cornell University); Vikram Krishnamurthy (Cornell University); Christopher Berry (Lockheed Martin Advanced Technology Labs)

12:05 PM

6529 (SPTM-L1.6): Differentiable adaptive short-time Fourier transform with respect to the window length
Maxime Leiber (INRIA); Yosra Marnissi (SAFRAN TECH); Axel Barrau (Offroad); Mohammed El Badaoui (Safran Tech)

SS-L1: 6G Integrated Sensing and Communication (ISAC) from Theory to Practice – A Signal Processing Perspective

Room: Nefeli A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Jia He,

10:50 AM

3049 (SS-L1.1): 6G integrated sensing and communication - Sensing assisted environmental reconstruction and communication
Zhi Zhou (Huawei Technologies Co., Ltd., Chengdu 610000, China); Xianjin Li (Huawei Technologies Co., Ltd., Chengdu 610000, China); Jia He (HUAWEI); Xiaoyan Bi (Huawei Technologies Canada Co., Ltd., Ottawa K2K 3J1, Canada); Yan Chen (Huawei Technologies); Guangjian Wang (Huawei Technologies Co., Ltd., Chengdu 610000, China); peiyong zhu (Huawei Technologies Canada)

11:05 AM

3325 (SS-L1.2): Neurally Augmented State Space Model for Simultaneous Communication and Tracking with Low Complexity Receivers
Fernando Pedraza (Technische Universität Berlin); Giuseppe Caire (Technische Universität Berlin)

11:20 AM

3456 (SS-L1.3): Multi-View Millimeter-Wave Imaging Over Wireless Cellular Network
Xin Tong (Zhejiang University); Zhaoyang Zhang (Zhejiang University); Zhaohui Yang (Zhejiang University)

Tuesday, June 6

11:35 AM

3803 (SS-L1.4): Joint Data Association, NLOS Mitigation, and Clutter Suppression for Networked Device-Free Sensing in 6G Cellular Network

Qin Shi (The Hong Kong Polytechnic University); Liang Liu (The Hong Kong Polytechnic University); Shuowen Zhang (The Hong Kong Polytechnic University)

11:50 AM

4255 (SS-L1.5): INTEGRATING THE SENSING AND RADIO COMMUNICATIONS CHANNEL MODELLING FROM RADAR MUTUAL INTERFERENCE

Narcis Cardona (iTEAM Research Institute, Universitat Politècnica de València); Jhoan Samuel Romero (iTEAM Research Institute, Universitat Politècnica de València); Wenfei Yang (Huawei Technologies); Jian Li (Huawei Technologies)

12:05 PM

5326 (SS-L1.6): Active Beam Tracking With Reconfigurable Intelligent Surface

Han Han (University of Toronto); Tao Jiang (University of Toronto); Wei Yu (University of Toronto)

ASPS-P1: Applications to Physiological Signals, Audio, and Speech

Room: Dome A

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Chenshu Wu, Robin Rajamaki

5872 (ASPS-P1.1): ClassA Entropy for the analysis of structural complexity of physiological signals

Hongjian Xiao (Imperial College London); Ling Li (City, University of London); Danilo P. Mandic (Imperial College of London, UK)

1034 (ASPS-P1.2): UNOBTRUSIVE RESPIRATORY MONITORING SYSTEM FOR INTENSIVE CARE

Xudong Tan (East China Normal University); Menghan Hu (East China Normal University); Guangtao Zhai (Shanghai Jiao Tong University); Yan Zhu (Shanghai Changzheng Hospital); Wenfang Li (Shanghai Changzheng Hospital); Xiao-Ping Zhang (Ryerson University)

4381 (ASPS-P1.3): IMPROVED WIFI-BASED RESPIRATION TRACKING VIA CONTRAST ENHANCEMENT

Wei-Hsiang Wang (University of Maryland, College Park); Xiaolu Zeng (Beijing Institute of Technology); Beibei Wang (Origin Wireless Inc.); K. J. Ray Liu (Origin Wireless Inc.)

4851 (ASPS-P1.4): Joint Angle and Respiration Estimation for Passive and Device-Free Respiration Monitoring

Gerrit Maus (University of Wuppertal); Dieter Brückmann (University of Wuppertal)

3418 (ASPS-P1.5): Implementing Continuous HRTF Measurement in Near-Field

Ee-Leng Tan (Nanyang Technological University); Santi Peksi (NTU Singapore); Woon Seng Gan (NTU)

5094 (ASPS-P1.6): SeliNet: A Lightweight Model for Single Channel Speech Separation

Ha Minh Tan (National Central University); Duc-Quang Vu (Thai Nguyen University of Education); Jia-Ching Wang (National Central University)

5196 (ASPS-P1.7): ADAPTIVE TIME-SCALE MODIFICATION FOR IMPROVING SPEECH INTELLIGIBILITY BASED ON PHONEME CLUSTERING FOR STREAMING SERVICES

Sohee Jang (Hanyang University); Jiye Kim (Hanyang University); Yeon-Ju Kim (Hanyang University); Joon-Hyuk Chang (Hanyang University)

3109 (ASPS-P1.8): CUTTING THROUGH THE NOISE: AN EMPIRICAL COMPARISON OF PSYCHOACOUSTIC AND ENVELOPE-BASED FEATURES FOR MACHINERY FAULT DETECTION

Peter Wißbrock (Lenze SE); Yvonne Richter (FH Bielefeld); David Pelkmann (Fachhochschule Bielefeld); Zhao Ren (L3S Research Center); Gregory Palmer (L3S Research Center)

4835 (ASPS-P1.9): Cochlear Decomposition: A Novel Bio-Inspired Multiscale Analysis Framework

Hessa Alfalahi (Khalifa University of Science and Technology); Ahsan Khandoker (Khalifa University); Ghada Alhussein (Khalifa University of Science and Technology); Leontios Hadjileontiadis (Khalifa University of Science and Technology)

2458 (ASPS-P1.10): DESIGN AND PERFORMANCE OF THE LOW-POWER NOISE REDUCTION ALGORITHM OF THE MED-EL SONNET 2 COCHLEAR IMPLANT AUDIO PROCESSOR

Ernst Aschbacher (MED-EL); Florian Fruehauf (MED-EL); Anja Kurz (University Hospital of Würzburg); Peter Nopp (MED-EL)

6491 (ASPS-P1.11): Modulo EEG Signal Recovery using Transformers

Tianyu Geng (Nanyang Technological University); Feng Ji (Nanyang Technological University); Pratibha Rana (Agency for Science, Technology and Research); Wee Peng Tay (Nanyang Technological University)

454 (ASPS-P1.12): Knowledge-graph Augmented Music Representation for Genre Classification

Han Ding (Xi'an Jiaotong University); Wenjing Song (Xi'an Jiaotong University); Cui Zhao (Xi'an Jiaotong University); Fei Wang (Xi'an Jiaotong University); Ge Wang (Xi'an Jiaotong University); Wei Xi (Xi'an Jiaotong University); Jizhong Zhao (Xi'an Jiaotong University)

Tuesday, June 6

IVMSP-P1: Super Resolution

Room: Dome J

Type: Poster

10:50 AM to 12:10 PM

Chair(s): Chun Yuan, Xinbo Gao

275 (IVMSP-P1.1): PFT-SSR: Parallax Fusion Transformer for Stereo Image Super-Resolution

Hansheng GUO (The Chinese University of Hong Kong); Juncheng Li (The Chinese University of Hong Kong); Guangwei Gao (Nanjing University of Posts and Telecommunications); Zhi Li (East China Normal University); Tiejong Zeng (The Chinese University of Hong Kong)

326 (IVMSP-P1.2): RAISING THE LIMIT OF IMAGE RESCALING USING AUXILIARY ENCODING

Chenzhong Yin (University of Southern California); Zhihong Pan (Baidu Research (USA)); Xin Zhou (Baidu USA); Le Kang (Baidu Research); Paul Bogdan (USC)

1431 (IVMSP-P1.3): Kernel estimation and deconvolution for blind image super-resolution

Jiali Gong (East China Normal University); Hongfan Gao (East China Normal University); Jiahao Chao (East China Normal University); Zhou Zhou (East China Normal University); Zhengfeng Yang (East China Normal University); Zhenbing Zeng (Shanghai University)

1555 (IVMSP-P1.4): A Comprehensive Comparison of Projections in Omnidirectional Super-Resolution

Huicheng Pi (Beijing Jiaotong University); Ming Lu (Intel Labs China); Senmao Tian (Beijing Jiaotong University); Jiaming Liu (Peking University); Yandong Guo (OPPO Research Institute); Shunli Zhang (Beijing Jiaotong University)

1900 (IVMSP-P1.5): LONG-SHORT ATTENTION NETWORK FOR THE SPECTRAL SUPER-RESOLUTION OF MULTISPECTRAL IMAGES

Kai Zhang (Shandong Normal University); Tian Jin (Shandong Normal University); Feng Zhang (Shandong Normal University); Jiande Sun (Shandong Normal University)

2363 (IVMSP-P1.6): Multi-level fusion for burst super-resolution with deep permutation-invariant conditioning

Martina Cilia (Politecnico di Torino); Diego Valsesia (Politecnico di Torino); Giulia Fracastoro (Polito); Enrico Magli (POLITO)

2684 (IVMSP-P1.7): Frequency Reciprocal Action and Fusion for Single Image Super-Resolution

Shuting Dong (Tsinghua University); Feng Lu (Tsinghua University); Chun Yuan (Graduate school at ShenZhen, Tsinghua university)

2777 (IVMSP-P1.8): FCIR: RETHINK AERIAL IMAGE SUPER RESOLUTION WITH FOURIER ANALYSIS

Yan Zhang (Chongqing University of Posts and Telecommunications); Pengcheng Zheng (Chongqing University of Posts and Telecommunications); Jianan Jiang (Chongqing University Of Posts And Telecommunications); Xiao PU (Chongqing University of Posts and Telecommunications); Xinbo Gao (Chongqing University of Posts and Telecommunications)

2962 (IVMSP-P1.9): A content-based multi-scale network for single image super-resolution

Jiahuan Ji (College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics); Baojiang Zhong (School of Computer Science and Technology, Soochow University); Kai-Kuang Ma (Nanyang Technological University, Singapore)

3053 (IVMSP-P1.10): LEARNING TO EXPLAIN: A GRADIENT-BASED ATTRIBUTION METHOD FOR INTERPRETING SUPER-RESOLUTION NETWORKS

Anni Yu (State Key Laboratory for Novel Software Technology, Nanjing University); Yu-Bin Yang (State Key Laboratory for Novel Software Technology, Nanjing University)

3140 (IVMSP-P1.11): CNN Filter for RPR-Based SR in VVC with Wavelet Decomposition

Hui Lan (Xidian University); Cheolkon Jung (Xidian University); Yang Liu (OPPO Mobile); Ming Li (OPPO)

3555 (IVMSP-P1.12): LOCAL TO GLOBAL PRIOR LEARNING FOR BLIND UNSUPERVISED IMAGE SUPER RESOLUTION

Kazuhiro Yamawaki (Yamaguchi University); Xian-Hua Han (Yamaguchi University)

IVMSP-P2: Denoising

Room: Dome K

Type: Poster

10:50 AM to 12:10 PM

Chair(s): Qiang Shen, Yijin Liu

5974 (IVMSP-P2.1): Rain2Avoid: Self-supervised Single Image Deraining

Yan-Tsung Peng (National Chengchi University); Wei Hua Li (National Chengchi University)

5479 (IVMSP-P2.2): A PROGRESSIVE IMAGE DEHAZING FRAMEWORK WITH INTER AND INTRA CONTRASTIVE LEARNING

honglei xu (Harbin Institute of Technology); Shaohui Liu (Harbin Institute of Technology); Yan Shu (State Key Laboratory of Communication Content Cognition, People's Daily Online, Beijing, China; Harbin Institute of Technology; Institute of Information Engineering, CAS); Feng Jiang (Harbin Institute of Technology, Harbin)

Tuesday, June 6

5267 (IVMSP-P2.3): GRAPH-BASED POINT CLOUD COLOR DENOISING WITH 3-DIMENSIONAL PATCH-BASED SIMILARITY
Ryosuke Watanabe (KDDI Research, Inc.); Keisuke Nonaka (KDDI Research Inc.); Eduardo Pavez (University of Southern California); Tatsuya Kobayashi (KDDI Research Inc.); Antonio Ortega (University of Southern California)

2310 (IVMSP-P2.4): CAENet: Using Collaborative Attention Transformer and Add-Boost Strategy for Single Image Deraining
Shengdi Qin (Beijing Jiaotong University); Shunli Zhang (Beijing Jiaotong University); Yu Zhang (Beihang University); Haoyu Gao (Beijing Jiaotong University)

1791 (IVMSP-P2.5): SFEMGN: IMAGE DENOISING WITH SHALLOW FEATURE ENHANCEMENT NETWORK AND MULTI-SCALE CONVGRU
Qidong Wang (China University of Mining and Technology); Lili Guo (China University of Mining and Technology); Shifei Ding (China University of Mining and Technology); Jian Zhang (china university of mining and technology); xiao xu (China University of Mining and Technology)

1554 (IVMSP-P2.6): Affinity Learning with Blind-spot Self-Supervision for Image Denoising
Yuhongze Zhou (McGill University); Liguang Zhou (The Chinese University of Hong Kong, Shenzhen); Issam Hadj Laradji (ServiceNow); Tin Lun Lam (The Chinese University of Hong Kong, Shenzhen); Yangsheng Xu (Shenzhen Institute of Artificial Intelligence and Robotics for Society)

1473 (IVMSP-P2.7): SAR IMAGE DESPECKLING WITH RESIDUAL-IN-RESIDUAL DENSE GENERATIVE ADVERSARIAL NETWORK
Yunpeng Bai (Aberystwyth University); Yayuan Xiao (Northwestern Polytechnical University); Xuan Hou (aberystwyth university); Ying Li (Northwestern Polytechnical University); Chaangjing Shang (Aberystwyth University); Qiang Shen (Aberystwyth University)

1211 (IVMSP-P2.8): UNCER2NATURAL: UNCERTAINTY-AWARE UNSUPERVISED IMAGE DENOISING
Chenyu Huang (Fudan University); Weimin Tan (Fudan University); Jiaying Shi (Fudan University); Zhen Xing (Fudan University); Bo Yan (Fudan University)

553 (IVMSP-P2.9): HPFTN: Hierarchical Progressive Fusion Transformer Network for Video Denoising
Shuaitao Zhang (Hikvision Research Institute); Yuan Zhang (Hikvision Research Institute); Zheng Zhao (Hikvision Research Institute); Di Xie (Hikvision Research Institute); Shiliang Pu (Hikvision Research Institute)

398 (IVMSP-P2.10): Subspace Modeling enabled High-sensitivity X-ray Chemical Imaging
Jizhou Li (City University of Hong Kong); Bin Chen (Max-Planck-Institut für Informatik); Guibin Zan (Stanford University); Guannan Qian (Stanford University); Piero Pianetta (Stanford University); Yijin Liu (SLAC National Accelerator Laboratory)

274 (IVMSP-P2.11): MSP-Former: Multi-Scale Projection Transformer for Single Image Desnowing
Sixiang Chen (Jimei University); Tian Ye (Jimei University); Yun Liu (Southwest University); Tao Dong Liao (JiMei University); JingXia Jiang (jimei university); Erkang Chen (Jimei University); Peng Chen (Jimei University)

117 (IVMSP-P2.12): Hyperspectral Image Denoising via Nonlocal Rank Residual Modeling
Zhiyuan Zha (Nanyang Technological University); Bihan Wen (Nanyang Technological University); Xin Yuan (Westlake University); Jiantao Zhou (University of Macau); Ce Zhu (University of Electronic Science & Technology of China)

IVMSP-P3: Semantic Segmentation

Room: Dome L

Type: Poster

10:50 AM to 12:10 PM

Chair(s): Tae-Kyun Kim, Xue Li

190 (IVMSP-P3.1): LOG-CAN: LOCAL-GLOBAL CLASS-AWARE NETWORK FOR SEMANTIC SEGMENTATION OF REMOTE SENSING IMAGES

Xiaowen Ma (Zhejiang University); Mengting Ma (Zhejiang University); Chenlu Hu (Zhejiang University); Zhiyuan Song (Zhejiang University); Ziyan Zhao (Zhejiang University); Tian Feng (Zhejiang University; Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies); Wei Zhang (Zhejiang University)

406 (IVMSP-P3.2): WUDA: Unsupervised Domain Adaptation Based on Weak Source Domain Labels
Shengjie Liu (Beijing University of Posts and Telecommunications); Chuang Zhu (Beijing University of Posts and Telecommunications); Yuan Li (Peking University); Wenqi Tang (Beijing University of Posts and Telecommunications)

555 (IVMSP-P3.3): Class-Aware Contextual Information for Semantic Segmentation
Huadong Tang (University of Technology Sydney); Youpeng Zhao (University of Central Florida); yingying jiang (Samsung Research China,Beijing); Zhuoxin Gan (Samsung Research Institute China-Beijing (SRC-B)); Qiang Wu (University of Technology Sydney)

1132 (IVMSP-P3.4): SEMI-SUPERVISED SEMANTIC SEGMENTATION WITH STRUCTURED OUTPUT SPACE ADAPTION
Weiquan Huang (Northeastern University(China)); Fu Zhang (Northeastern University)

Tuesday, June 6

1170 (IVMSP-P3.5): PRRD: PIXEL-REGION RELATION DISTILLATION FOR EFFICIENT SEMANTIC SEGMENTATION
Chen Wang (Chongqing University); Jiang Zhong (); Qizhu Dai (Chongqing University); yafei qi (Central South University); Rongzhen Li (Chongqing University); Qin Lei (Chongqing University); BIN FANG (Chongqing University); Xue Li (University of Queensland)

2521 (IVMSP-P3.6): SPATIAL CORRELATION FUSION NETWORK FOR FEW-SHOT SEGMENTATION
Xueliang Wang (Tsinghua University); Wenqi Huang (China southern power grid); Wenming Yang (Tsinghua University); Qingmin Liao (Tsinghua University)

3306 (IVMSP-P3.7): Exploring vision transformer layer choosing for semantic segmentation
Fangjian Lin (alibaba-inc); Yizhe Ma (Xinjiang University); Shengwei Tian (Xinjiang University)

3941 (IVMSP-P3.8): JOINT TRAINING OF HIERARCHICAL GANS AND SEMANTIC SEGMENTATION FOR EXPRESSION TRANSLATION
Rumeysa Bodur (Imperial College London); Binod Bhattarai (University of Aberdeen); Tae-Kyun Kim (Imperial College London)

6357 (IVMSP-P3.9): PROGRESSIVE REFINEMENT LEARNING BASED ON FEATURE CROSS PERCEPTION FOR RESIDENTIAL AREAS SEMANTIC SEGMENTATION
Xinran Lyu (Beijing Normal University); Libao Zhang (Beijing Normal University)

1599 (IVMSP-P3.10): Lightweight Portrait Segmentation via Edge-optimized Attention
Xinyue Zhang (Qingdao university); Guodong Wang (Qingdao University); Lijuan Yang (Hisense Visual Technology Co., Ltd); Chenglizhao Chen (China University of Petroleum (East China))

3857 (IVMSP-P3.11): A Dynamic Cross-scale Transformer with Dual-compound Representation for 3D Medical Image Segmentation
Ruixia Zhang (Northeastern University); Zhiqiong Wang (Northeastern University); Zhongyang Wang (Northeastern University); Junchang Xin (Northeastern University)

3793 (IVMSP-P3.12): LABANet: Lead-Assisting Backbone Attention Network for oral multi-pathology segmentation
Huabao Chen (Hohai University); Xiaolong Huang (Chongqing University of Technology); Qiankun Li (USTC); Jianqing Wang (Shanghai International Studies University); bo fang (Northeastern University); Junxin Chen (Dalian University of Technology)

IVMSP-P4: Object Segmentation
Room: Dome M
Type: Poster
10:50 AM to 12:10 PM
Chair(s): Truong Nguyen, Tiejun Huang

3473 (IVMSP-P4.1): Robust Video Object Segmentation With Restricted Attention
Huaizheng Zhang (Fudan University); Pinxue Guo (Fudan University); Zhongwen Le (Fudan University); Wenqiang Zhang (Fudan University)

3501 (IVMSP-P4.2): STACKING-BASED ATTENTION TEMPORAL CONVOLUTIONAL NETWORK FOR ACTION SEGMENTATION
Liu Yang (School of Computer Science and Engineering, Central South University); Yu Jiang (School of Computer Science and Engineering, Central South University); Junkun Hong (School of Computer Science and Engineering, Central South University); Zhenjie Wu (School of Computer Science and Engineering, Central South University); Zhan Yang (Big Data Institute, Central South University); Jun Long (Central South University)

2436 (IVMSP-P4.3): VLKP:VIDEO INSTANCE SEGMENTATION WITH VISUAL-LINGUISTIC KNOWLEDGE
ruixiang chen (Zhejiang University of Technology); Sheng Liu (Zhejiang University of Technology); Junhao Chen (Zhejiang University of Technology); Blngran Guo (Zhejiang University of Technology); Feng Zhang (Zhejiang University of Technology)

4867 (IVMSP-P4.4): Automatic Error Detection in Integrated Circuits Image Segmentation: A Data-driven Approach
Zhikang Zhang (Arizona State University); Bruno Trindade (TechInsights Inc.); Michael Green (TechInsights Inc.); Zifan Yu (Arizona State University); Christopher Pawlowicz (TechInsights Inc.); Fengbo Ren (Arizona State University)

3745 (IVMSP-P4.5): TransWnet: Integrating Transformers into CNNs via Row and Column Attention for Abdominal Multi-organ Segmentation
Yazhen Xie (Xiangtan University); Yanglin Huang (Xiangtan University); Yuan Zhang (Xiangtan University); Xuanya Li (Baidu); Xiongjun Ye (Xiangtan University); Kai Hu (Xiangtan University)

5844 (IVMSP-P4.6): ACTIVE PERCEPTION SYSTEM FOR ENHANCED VISUAL SIGNAL RECOVERY USING DEEP REINFORCEMENT LEARNING
Gaurav Chaudhary (Indian Institute of Technology Kanpur, India); Prof Laxmidhar Behera (IIT Kanpur); Tushar Sandhan (Indian Institute of Technology Kanpur)

302 (IVMSP-P4.7): OAFFormer: Learning Occlusion Distinguishable Feature for Amodal Instance Segmentation
Zhixuan Li (Peking University); Ruohua Shi (Peking University); Tiejun Huang (Peking University); Tingting Jiang (Peking University)

Tuesday, June 6

698 (IVMSP-P4.8): Encoder-Decoder Graph Convolutional Network for Automatic Timed-Up-and-Go and Sit-to-Stand Segmentation

Bo Wen (University of California, San Diego); Chen Du (University of California, San Diego); Truong Nguyen (UC San Diego)

758 (IVMSP-P4.9): Meta++ Network for Few-shot Aerospace Crack Segmentation

Chengyuan Xu (Northwestern Polytechnical University); Kang Liu (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

1764 (IVMSP-P4.10): IAST: Instance Association Relying on Spatio-temporal Features for Video Instance Segmentation

Junhao Chen (Zhejiang University of Technology); Sheng Liu (Zhejiang University of Technology); ruixiang chen (Zhejiang University of Technology); Blngnan Guo (Zhejiang University of Technology); Feng Zhang (Zhejiang University of Technology)

2469 (IVMSP-P4.11): Continual Cell Instance Segmentation of Microscopy Images

Tzu-Ting Chuang (National Sun Yat-sen University); Ting-Yun Wei (National Taiwan University); Yu-Hsing Hsieh (National Taiwan University); Chu-Song Chen (National Taiwan University); Huei-Fang Yang (National Sun Yat-sen University)

IVMSP-P4: Object Sementation

Room: Dome M

Type: Poster

10:50 AM to 12:10 PM

Chair(s): Truong Nguyen, Tiejun Huang

3473 (IVMSP-P4.1): Robust Video Object Segmentation With Restricted Attention

Huaizheng Zhang (Fudan University); Pinxue Guo (Fudan University); Zhongwen Le (Fudan University); Wenqiang Zhang (Fudan University)

3501 (IVMSP-P4.2): STACKING-BASED ATTENTION TEMPORAL CONVOLUTIONAL NETWORK FOR ACTION SEGMENTATION

Liu Yang (School of Computer Science and Engineering, Central South University); Yu Jiang (School of Computer Science and Engineering, Central South University); Junkun Hong (School of Computer Science and Engineering, Central South University); Zhenjie Wu (School of Computer Science and Engineering, Central South University); Zhan Yang (Big Data Institute, Central South University); Jun Long (Central South University)

2436 (IVMSP-P4.3): VLKP:VIDEO INSTANCE SEGMENTATION WITH VISUAL-LINGUISTIC KNOWLEDGE

ruixiang chen (Zhejiang University of Technology); Sheng Liu (Zhejiang University of Technology); Junhao Chen (Zhejiang University of Technology); Blngnan Guo (Zhejiang University of Technology); Feng Zhang (Zhejiang University of Technology)

4867 (IVMSP-P4.4): Automatic Error Detection in Integrated Circuits Image Segmentation: A Data-driven Approach

Zhikang Zhang (Arizona State University); Bruno Trindade (TechInsights Inc.); Michael Green (TechInsights Inc.); Zifan Yu (Arizona State University); Christopher Pawlowicz (TechInsights Inc.); Fengbo Ren (Arizona State University)

3745 (IVMSP-P4.5): TransWnet: Integrating Transformers into CNNs via Row and Column Attention for Abdominal Multi-organ Segmentation

Yazhen Xie (Xiangtan University); Yanglin Huang (Xiangtan University); Yuan Zhang (Xiangtan University); Xuanya Li (Baidu); Xiongjun Ye (Xiangtan University); Kai Hu (Xiangtan University)

5844 (IVMSP-P4.6): ACTIVE PERCEPTION SYSTEM FOR ENHANCED VISUAL SIGNAL RECOVERY USING DEEP REINFORCEMENT LEARNING

Gaurav Chaudhary (Indian Institute of Technology Kanpur, India); Prof Laxmidhar Behera (IIT Kanpur); Tushar Sandhan (Indian Institute of Technology Kanpur)

302 (IVMSP-P4.7): OAFFormer: Learning Occlusion Distinguishable Feature for Amodal Instance Segmentation

Zhixuan Li (Peking University); Ruohua Shi (Peking University); Tiejun Huang (Peking University); Tingting Jiang (Peking University)

698 (IVMSP-P4.8): Encoder-Decoder Graph Convolutional Network for Automatic Timed-Up-and-Go and Sit-to-Stand Segmentation

Bo Wen (University of California, San Diego); Chen Du (University of California, San Diego); Truong Nguyen (UC San Diego)

758 (IVMSP-P4.9): Meta++ Network for Few-shot Aerospace Crack Segmentation

Chengyuan Xu (Northwestern Polytechnical University); Kang Liu (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

1764 (IVMSP-P4.10): IAST: Instance Association Relying on Spatio-temporal Features for Video Instance Segmentation

Junhao Chen (Zhejiang University of Technology); Sheng Liu (Zhejiang University of Technology); ruixiang chen (Zhejiang University of Technology); Blngnan Guo (Zhejiang University of Technology); Feng Zhang (Zhejiang University of Technology)

2469 (IVMSP-P4.11): Continual Cell Instance Segmentation of Microscopy Images

Tzu-Ting Chuang (National Sun Yat-sen University); Ting-Yun Wei (National Taiwan University); Yu-Hsing Hsieh (National Taiwan University); Chu-Song Chen (National Taiwan University); Huei-Fang Yang (National Sun Yat-sen University)

Tuesday, June 6

MLSP-P1: Deep Learning for Image and Video Processing I

Room: Dome F

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

5397 (MLSP-P1.1): Spammer Detection on Short Video Applications: A New Challenge and Baselines

Muyang Yi (Shanghai Jiao Tong University); Dong Liang (ByteDance); Rui Wang (Bytedance AI Lab); Yue Ding (Shanghai Jiao Tong University); Hongtao Lu (Shanghai Jiao Tong University)

814 (MLSP-P1.2): Weakly- and Semi-Supervised Object Localization

Zhen-Tang Huang (National Taiwan Normal University); Yan-He Chen (National Taiwan Normal University); Mei-Chen Yeh (National Taiwan Normal University)

2503 (MLSP-P1.3): Balanced Mixup Loss for Long-tailed Visual Recognition

Haibo Ye (Nanjing University of Aeronautics and Astronautics); Fangyu Zhou (Nanjing University of Aeronautics and Astronautics); Xinjie Li (Nanjing University of Aeronautics and Astronautics); Qingheng Zhang (Nanjing University of Aeronautics and Astronautics)

4130 (MLSP-P1.4): On Cross-Layer Alignment for Model Fusion of Heterogeneous Neural Networks

Dang Nguyen (VinAI); Thien Trang Nguyen Vu (Hanoi University of Science and Technology); Khai Nguyen (University of Texas at Austin); Dinh Q Phung (Monash University); Hung Bui (VinAI Research); Nhat Ho (University of Texas at Austin)

2813 (MLSP-P1.5): Invariant Adversarial Imitation Learning from Visual Inputs

Haoran Zhang (East China Normal University); Yinghong Tian (East China Normal University); Liang Yuan (Beijing University of Chemical Technology); Yue Lu (East China Normal University)

6423 (MLSP-P1.6): SpectraNet-SO(3): Learning Satellite Orientation from Optical Spectra by Implicitly Modeling Mutually Exclusive Probability Distributions on the Rotation Manifold

Matthew Phelps (Odyssey Systems); Ryan Swindle (Odyssey Systems); Zack Gazak (Odyssey Systems); Andrew Vandenberg (AFRL); Justin Fletcher (Odyssey Systems)

3097 (MLSP-P1.7): STRUCTURED-ANCHOR PROJECTED CLUSTERING FOR HYPERSPECTRAL IMAGES

Guozhu Jiang (China University of Geosciences); jie zhang (University of Macau); Yongshan Zhang (China University of Geosciences); Xinwei Jiang (China University of Geosciences); Zhihua Cai (China University of Geosciences)

140 (MLSP-P1.8): Learning sparse auto-encoders for green AI image coding

Cyprien Gille (UMONS); Frederic Guyard (Orange Labs); Marc Antonini (Universite Nice Sophia Antipolis); Michel Barlaud (University of Nice)

643 (MLSP-P1.9): Learning to Generate 3D Representations of Building Roofs Using Single-View Aerial Imagery

Maxim Khomiakov (Technical University of Denmark); Alejandro Valverde Mahou (Technical University of Denmark); Alba Reinders Sánchez (Technical University of Denmark); Jes Frelsen (Technical University of Denmark); Michael Andersen (Technical University of Denmark)

4843 (MLSP-P1.10): Robust Monocular Localization of Drones by Adapting Domain Maps to Depth Prediction Inaccuracies

Priyesh Shukla (University of Illinois Chicago); Sureshkumar Senthilkumar (University of Illinois at Chicago); Alex C Stutts (University of Illinois Chicago); Sathya Ravi (University of Illinois at Chicago); Theja Tulabandhula (UIC); Amit R Trivedi (University of Illinois at Chicago)

5940 (MLSP-P1.11): Large dimensional analysis of LS-SVM transfer learning: Application to POLSAR classification

Cyprien DOZ (Sondra - Centrale Supélec (University Paris Saclay)); Chengfang Ren (Sondra - CentraleSupélec); Jean-Philippe Ovarlez (ONERA, CentraleSupélec, SONDRRA, Université Paris-Saclay); Romain COUILLET (CentraleSupélec, GIPSA-lab @ Université Grenoble-Alpes)

5062 (MLSP-P1.12): SMUG: Towards robust MRI reconstruction by smoothed unrolling

Hui Li (Huazhong University of Science and Technology); jinghan jia (Michigan state university); Shijun Liang (michigan state university); Yuguang Yao (Michigan State University); Saiprasad Ravishankar (Michigan State University); Sijia Liu (Michigan State University)

MLSP-P2: Graph based Learning II

Room: Dome G

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

715 (MLSP-P2.1): Graph-Graph Context Dependency Attention for Graph Edit Distance

Ruiqi Jia (Wangxuan Institute of Computer Technology, Peking University); xianbing feng (peking university); Xiaoqing Lyu (Peking University); Zhi Tang (Peking University)

Tuesday, June 6

3882 (MLSP-P2.2): Topology Uncertainty Modeling For Imbalanced Node Classification on Graphs

Jiayi Gao (Southeast University); Jiaying Li (Southeast University); Ke Zhang (Southeast University); Youyong Kong (Southeast University)

589 (MLSP-P2.3): CPD-GAN: Cascaded Pyramid Deformation GAN for Pose Transfer

Yuan Huang (Nanjing University); Yuting Tang (Nanjing University); Xiu Zheng (Nanjing University); Jie Tang (Nanjing University)

5321 (MLSP-P2.4): Space-Time Graph Neural Networks with Stochastic Graph Perturbations

Samar Hadou (University of Pennsylvania); Charilaos Kanatsoulis (University of Pennsylvania); Alejandro Ribeiro (University of Pennsylvania)

6793 (MLSP-P2.5): Untrained Graph Neural Networks for Denoising (SPS Journal Paper)*

Samuel Rey (King Juan Carlos University); Santiago Segarra (Rice University); Reinhard Heckel (TUM); Antonio G. Marques (King Juan Carlos University)

5846 (MLSP-P2.6): Learning on Graphs under Label Noise

Jingyang Yuan (Peking University); Xiao Luo (UCLA); Yifang Qin (Peking University); Yusheng Zhao (Peking University); Wei Ju (Peking University); Ming Zhang (Peking University)

2906 (MLSP-P2.7): Select the Best: Enhancing Graph Representation with Adaptive Negative Sample Selection

Xiangping Zheng (Renmin University of China); Xun Liang (Renmin University of China); Bo Wu (Renmin University of China)

2586 (MLSP-P2.8): Learning with Multigraph Convolutional Filters

Landon G Butler (University of California, Berkeley); Alejandro Parada-Mayorga (University of Pennsylvania); Alejandro Ribeiro (University of Pennsylvania)

2164 (MLSP-P2.9): Self-supervised Guided Hypergraph Feature Propagation for Semi-supervised Classification with Missing Node Features

chengxiang lei (Huazhong University of Science and Technology); Sichao Fu (Huazhong University of Science and Technology); Yuetian Wang (Huazhong University of Science and Technology); Wenhao Qiu (Huazhong University of Science and Technology); Yachen Hu (Huazhong University of Science and Technology); Qinmu Peng (Huazhong University of Science and Technology); Xinge YOU (Huazhong University of Science and Technology)

3752 (MLSP-P2.10): Incorporating reliability in graph information propagation by fluid dynamics diffusion: a case of multimodal semisupervised deep learning

Andrea Marinoni (UiT the Arctic University of Norway); Marine Mercier (University of Cambridge); Qian Shi (Sun Yat-sen University); Sivasakthy Selvakumaran (University of Cambridge); Mark Girolami (University of Cambridge)

5159 (MLSP-P2.11): GraphMAD: Graph Mixup for Data Augmentation using Data-Driven Convex Clustering

Madeline Navarro (Rice University); Santiago Segarra (Rice University)

3724 (MLSP-P2.12): Time-varying Signals Recovery via Graph Neural Networks

Jhon A Castro Correa (University of Delaware); Jhony H. Giraldo (Télécom Paris); Anindya Mondal (Jadavpur University); Mohsen Badiey (University of Delaware); Thierry BOUWMANS (Univ. La Rochelle); Fragkiskos Malliaros (CentraleSupélec)

MLSP-P3: Learning from Multimodal Data

Room: Dome H

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

3546 (MLSP-P3.1): Multimodal Knowledge Distillation for Arbitrary-Oriented Object Detection in Aerial Images

Zhanchao Huang (Beijing Institute of Technology); Wei Li (Beijing Institute of Technology, Beijing, China); Ran Tao (Beijing Institute of Technology)

1234 (MLSP-P3.2): HIERARCHICAL SPATIAL-TEMPORAL TRANSFORMER WITH MOTION TRAJECTORY FOR INDIVIDUAL ACTION AND GROUP ACTIVITY RECOGNITION

Xiaolin Zhu (Xiangtan University); Dongli Wang (Xiangtan University); Yan ZHOU (Xiangtan University)

693 (MLSP-P3.3): Autonomous Soundscape Augmentation with Multimodal Fusion of Visual and Participant-linked Inputs

Kenneth Ooi (Nanyang Technological University); Karn N Watcharasupat (Georgia Institute of Technology); Bhan Lam (NTU); Zhen-Ting Ong (Nanyang Technological University); Woon Seng Gan (NTU)

1571 (MLSP-P3.4): TOWARDS ROBUST AUDIO-BASED VEHICLE DETECTION VIA IMPORTANCE-AWARE AUDIO-VISUAL LEARNING

Jung Uk Kim (Kyung Hee University); Seong Tae Kim (Kyung Hee University)

841 (MLSP-P3.5): Hierarchical Multi-Task Learning for Fabric Component Analysis Based on NIR Spectral Signals

Joseph Kim (Fudan University); Dong Wu (Fudan University); mingmin Chi (Fudan university); Gaoqi Xu (Zhongshan PoolNet Technology Co. Ltd.)

Tuesday, June 6

1706 (MLSP-P3.6): Cross Modality Knowledge Distillation for Robust Pedestrian Detection in Low Light and Adverse Weather Conditions

Mazin Hnewa (Michigan State University); Alireza Rahimpour (Ford Motor Company- Palo Alto); Justin Miller (Ford); Devesh Upadhyay (Ford Motor Co.); Hayder Radha (Michigan State University)

6375 (MLSP-P3.7): Data leakage in cross-modal retrieval training: A case study

Benno Weck (Music Technology Group, Universitat Pompeu Fabra (UPF)); Xavier Serra (Universitat Pompeu Fabra)

5825 (MLSP-P3.8): DIFFICULTY-AWARE DATA AUGMENTOR FOR SCENE TEXT RECOGNITION

Guanghao Meng (Tsinghua University); Tao Dai (Shenzhen University); Bin Chen (Harbin Institute of Technology, Shenzhen); Naiqi Li (Tsinghua-Berkeley Shenzhen Institute); Yong Jiang (Tsinghua University); Shu-Tao Xia (Tsinghua University)

461 (MLSP-P3.9): TinyOOD: Effective Out-of-Distribution Detection for TinyML

Yongchang Li (Soochow University); Juncheng Jia (Soochow University); Yan Zuo (Jiangsu New Hope Technology Co., Ltd); Weipeng Zhu (SOOCHOW UNIVERSITY)

4211 (MLSP-P3.10): A principled approach to model validation in domain generalization

Boyang Lyu (Tufts University); Thuan Nguyen (Tufts University); Matthias Scheutz (Tufts University); Prakash Ishwar (Boston University); Shuchin Aeron (Tufts University)

4220 (MLSP-P3.11): Scale-adaptive tiny object detection enhanced by across-scale and shape-preserved semantic location

Yuting He (Southwest University); Renjie Huang (Southwest University); Yangguang Shi (Southwest University); Guoqiang Xiao (College of Computer and Information Science, Southwest University, Chongqing, China); Bin Yang (Southwest University); Yuqi Li (Southwest University)

3735 (MLSP-P3.12): Audio-Visual Inpainting: Reconstructing Missing Visual Information with Sound

Valentina Sanguineti (Istituto Italiano di Tecnologia); Sanket Thakur (Istituto Italiano di Tecnologia); Pietro Morerio (Istituto Italiano di Tecnologia); Alessio Del Bue (Istituto Italiano di Tecnologia (IIT)); Vittorio Murino (Istituto Italiano di Tecnologia)

MLSP-P4: Matrix/Tensor Factorization and Completion

Room: Dome I

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

507 (MLSP-P4.1): Learn Topological Representation with Flexible Manifold Layer

Ziheng Jiao (Northwestern Polytechnical University.); Hongyuan Zhang (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

1438 (MLSP-P4.2): Tensorized LSSVMs for Multitask Regression

Jiani Liu (University of Electronic Science and Technology of China); Qinghua Tao (KU Leuven); Ce Zhu (University of Electronic Science & Technology of China); Yipeng Liu (University of Electronic Science and Technology of China); Johan Suykens (KU Leuven)

3571 (MLSP-P4.3): A Bayesian Perspective for Determinant Minimization Based Robust Structured Matrix Factorization

Gokcan Tatli (University of Wisconsin-Madison); Alper Erdogan (Koc University)

5045 (MLSP-P4.4): Volume-regularized Nonnegative Tucker Decomposition with Identifiability Guarantees

Yuchen Sun (University of Florida); Kejun Huang (University of Florida)

687 (MLSP-P4.5): Transductive Matrix Completion with Calibration for Multi-Task Learning

Hengfang Wang (Fujian Normal University); Yasi Zhang (University of California, Los Angeles); Xiaojun Mao (Shanghai Jiao Tong University); Zhonglei Wang (Xiamen University)

1668 (MLSP-P4.6): Projected Hierarchical ALS for generalized Boolean matrix factorization

Rodrigo Cabral Farias (Université Côte d'Azur, CNRS, I3S Laboratory); Sebastian Miron (University of Lorraine)

2934 (MLSP-P4.7): ROBUST BINARY COMPONENT DECOMPOSITIONS

Christos Kolomvakis (University of Mons); Nicolas Gillis (Université de Mons)

3897 (MLSP-P4.8): MULTI-RESOLUTION CONVOLUTIONAL DICTIONARY LEARNING FOR RIVERBED DYNAMICS MODELING

Eisuke Kobayashi (Niigata Univ.); Hiroyasu Yasuda (Niigata Univ.); Kiyoshi Hayasaka (Niigata Univ.); Yu Otake (Tohoku Univ.); Shunsuke Ono (Tokyo Institute of Technology); Shogo Muramatsu (Niigata Univ.)

2388 (MLSP-P4.9): PARAFAC2-based Coupled Matrix and Tensor Factorizations

Carla Schenker (Simula Metropolitan Center for Digital Engineering); Xiulin Wang (Affiliated Zhongshan Hospital of Dalian University); Evrim Acar (Simula Metropolitan Center for Digital Engineering)

Tuesday, June 6

6088 (MLSP-P4.10): Deep plug-and-play for tensor robust principal component analysis

Hao Tan (Southwest University); Jianjun Wang (Southwest University); Weichao Kong (Southwest University)

6125 (MLSP-P4.11): Geometric Matrix Completion with Collaborative Routing between Capsules

Xuan Li (School of Software Tsinghua University); Li Zhang (School of Software Tsinghua University)

3256 (MLSP-P4.12): Enrollment Rate Prediction in Clinical Trials based on CDF Sketching and Tensor Factorization tools

Magda Amiridi (University of Virginia); Cheng Qian (IQVIA); Nicholas D Sidiropoulos (University of Virginia); Lucas Glass (IQVIA)

SLT-P1: ASR - Improve Latency, Efficiency, and Accuracy

Room: Dome B

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Yossi Adi,

900 (SLT-P1.1): Multi-blank Transducers for Speech Recognition

Hainan Xu (NVIDIA); Fei Jia (NVIDIA Corporation); Somshubra Majumdar (NVIDIA); Shinji Watanabe (Carnegie Mellon University); Boris Ginsburg (NVIDIA)

1642 (SLT-P1.2): DIAGONAL STATE SPACE AUGMENTED TRANSFORMERS FOR SPEECH RECOGNITION

George Saon (IBM); Ankit Gupta (IBM Research); Xiaodong Cui (IBM T. J. Watson Research Center)

1661 (SLT-P1.3): TrimTail: Low-Latency Streaming ASR with Simple but Effective Spectrogram-Level Length Penalty

Xingchen Song (Tsinghua University); Di Wu (horizon); Zhiyong Wu (Tsinghua University); Binbin Zhang (horizon); Yuekai Zhang (Wenet Open Source Community); Zhendong Peng (horizon); Wenpeng Li (horizon); Fuping Pan (horizon); Changbao Zhu (horizon)

3385 (SLT-P1.4): Towards Accurate and Real-time End-of-speech Estimation

Yifeng Fan (University of Illinois at Urbana-Champaign); Colin Vaz (Amazon); Di He (Amazon); Jahn Heymann (Amazon); Viet Anh Trinh (Amazon); Zhe Zhang (Amazon); Venkatesh Ravichandran (Amazon)

3999 (SLT-P1.5): Peak-First CTC: Reducing the Peak Latency of CTC Models by Applying Peak-First Regularization

Zhengkun Tian (Meituan Inc.); Hongyu Xiang (Meituan Inc.); Min Li (Meituan Inc.); Feifei Lin (Meituan Inc.); Ke Ding (Meituan Inc.); Guanglu Wan (Meituan)

4330 (SLT-P1.6): EVALUATING PARAMETER-EFFICIENT TRANSFER LEARNING APPROACHES ON SURE BENCHMARK FOR SPEECH UNDERSTANDING

Li Yingting (Beijing University of Posts and Telecommunications); Ambuj Mehrish (SUTD); RISHABH BHARDWAJ (Singapore University of Technology and Design); Navonil Majumder (SUTD); Bo Cheng (Beijing University of Posts and Telecommunications); Shuai Zhao (Beijing University of Posts and Telecommunications); Amri Zadeh (Amazon Science); Rada Mihalcea (University of Michigan); Soujanya Poria (Singapore University of Technology and Design)

5058 (SLT-P1.7): Powerful and Extensible WFST Framework for RNN-Transducer Losses

Aleksandr Laptev (NVIDIA, ITMO University); Vladimir Bataev (NVIDIA); Igor Gitman (NVIDIA); Boris Ginsburg (NVIDIA)

5337 (SLT-P1.8): PREDICTING MULTI-CODEBOOK VECTOR QUANTIZATION INDEXES FOR KNOWLEDGE DISTILLATION

Liyong Guo (Northwestern Polytechnical University); Xiaoyu Yang (Xiaomi Corp., Beijing); Quandong Wang (Xiaomi Corp., Beijing); Yuxiang Kong (Xiaomi Corp., Beijing); Zengwei Yao (Xiaomi Corp., Beijing); fan cui (xiaomi); Fangjun Kuang (Xiaomi Corp., Beijing); Wei Kang (Xiaomi Corp., Beijing, China); Long Lin (Xiaomi Corp., Beijing); Mingshuang Luo (Xiaomi Corp., Beijing); Piotr Żelasko (Johns Hopkins University); Daniel Povey (Johns Hopkins University)

5434 (SLT-P1.9): IMPROVING NON-AUTOREGRESSIVE SPEECH RECOGNITION WITH AUTOREGRESSIVE PRETRAINING

Yanjia Li (Fano Labs); Lahiru T Samarakoon (Fano Labs, Hong Kong); Ivan Fung (Fano Labs)

5558 (SLT-P1.10): Conversation-oriented ASR with multi-look-ahead CBS architecture

Huaibo Zhao (Waseda University); Shinya Fujie (Waseda University); Tetsuji Ogawa (Waseda University); Jin Sakuma (Waseda University); Yusuke Kida (LINE Corp); Tetsunori Kobayashi (Waseda University)

5607 (SLT-P1.11): Using Adapters to Overcome Catastrophic Forgetting in End-to-End Automatic Speech Recognition

Steven Vander Eeck (KU Leuven); Hugo Van hamme (KU LEUVEN)

5824 (SLT-P1.12): FAST AND PARALLEL DECODING FOR TRANSDUCER

Wei Kang (Xiaomi Corp., Beijing, China); Liyong Guo (Xiaomi Corp.); Fangjun Kuang (Xiaomi Corp.); Long Lin (Xiaomi Corp., Beijing, China); Mingshuang Luo (Xiaomi Corp., Beijing, China); Zengwei Yao (Xiaomi Corp., Beijing, China); Xiaoyu Yang (Xiaomi Corp., Beijing, China); Piotr Żelasko (Johns Hopkins University); Daniel Povey (Johns Hopkins University)

Tuesday, June 6

SLT-P2: ASR: Domain Adaptation and Robust Training

Room: Dome C

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Masakiyo Fujimoto, Catalin Zorila

505 (SLT-P2.1): SAN: a robust end-to-end ASR model architecture

Zeping Min (Peking University); Qian Ge (Peking University); Guanhua Huang (USTC)

1604 (SLT-P2.2): Explanations for Automatic Speech Recognition

Xiaoliang Wu (University of Edinburgh); Peter Bell (University of Edinburgh); Ajitha Rajan (University of Edinburgh)

1674 (SLT-P2.3): On-the-fly Text Retrieval for End-to-End ASR Adaptation

Bolaji Yusuf (Bogazici University); Aditya Gourav (Amazon); Ankur Gandhe (Amazon Alexa); Ivan Bulyko (Amazon)

2397 (SLT-P2.4): Unsupervised model-based speaker adaptation of end-to-end lattice-free MMI model for speech recognition

Xurong Xie (Institute of Software, Chinese Academy of Sciences); Xunying Liu (The Chinese University of Hong Kong); Hui Chen (Institute of Software, Chinese Academy of Sciences); Hongan Wang (Institute of Software, Chinese Academy of Sciences)

3258 (SLT-P2.5): Domain Adaptation with External Off-Policy Acoustic Catalogs for Scalable Contextual End-To-End Automated Speech Recognition

David Chan (University of California, Berkeley); Shalini Ghosh (Amazon Alexa AI); Ariya Rastrow (Amazon Alexa); Bjorn Hoffmeister (Amazon)

3600 (SLT-P2.6): Comparison of Soft and Hard Target RNN-T Distillation for Large-scale ASR

Dongseong Hwang (Google); Khe C Sim (Google Inc.); Yu Zhang (Google); Trevor Strohman (Google)

3973 (SLT-P2.7): WeavSpeech: Data Augmentation Strategy for Automatic Speech Recognition via Semantic-Aware Weaving

Kyusung Seo (KAIST); Joonhyung Park (KAIST); Jaeyun Song (KAIST); Eunho Yang (KAIST)

4139 (SLT-P2.8): Joint Discriminator and Transfer Based Fast Domain Adaptation for End-to-End Speech Recognition

Hang Shao (Shanghai Jiao Tong University); Tian Tan (Aispeech Ltd.); wei wang (Shanghai Jiao Tong University); Xun Gong (Shanghai Jiaotong University); Yanmin Qian (Shanghai Jiao Tong University)

5424 (SLT-P2.9): IMPROVING FAIRNESS AND ROBUSTNESS IN END-TO-END SPEECH RECOGNITION THROUGH UNSUPERVISED CLUSTERING

Irina-Elena Veliche (Meta); Pascale Fung (Hong Kong University of Science and Technology)

5491 (SLT-P2.10): Improving Fast-slow Encoder based Transducer with Streaming Deliberation

Ke Li (Meta AI); Jay Mahadeokar (Meta AI); Jinxi Guo (Meta); Yangyang Shi (Meta AI); Gil Keren (Meta AI); Ozlem Kalinli (Meta AI); Michael Seltzer (Meta AI); Duc Le (Meta AI)

5496 (SLT-P2.11): Dynamic Alignment Mask CTC: Improved Mask-CTC with Aligned Cross Entropy

Xulong Zhang (Ping An Technology (Shenzhen) Co., Ltd.); Haobin Tang (USTC); Jianzong Wang (Ping An Technology (Shenzhen) Co., Ltd); Ning Cheng (Ping An Technology (Shenzhen) Co., Ltd); Jian Luo (Ping An Technology (Shenzhen) Co., Ltd); Jing Xiao (Ping An Insurance (Group) Company of China)

5902 (SLT-P2.12): Improving Accented Speech Recognition with Multi-Domain Training

Lucas Maison (Laboratoire Informatique d'Avignon); Yannick Estève (LIA - Avignon University)

SLT-P3: ASR: New Models

Room: Dome D

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Bhuvana Ramabhadran, Xie Chen

179 (SLT-P3.1): UCONV-CONFORMER: HIGH REDUCTION OF INPUT SEQUENCE LENGTH FOR END-TO-END SPEECH RECOGNITION

Andrei Andrusenko (ITMO University); Rauf Nasretdinov (STC); Aleksei Romanenko (STC-innovations Ltd)

876 (SLT-P3.2): A Comparison of Semi-Supervised Learning Techniques for Streaming ASR at Scale

Charles C Peyser (Google Inc.); Michael Picheny (NYU); Kyunghyun Cho (New York University); Tara Sainath (Google); W. Ronny Huang (Google); Rohit Prabhavalkar (Google)

1356 (SLT-P3.3): Improving Contextual Biasing with Text Injection

Tara Sainath (Google); Rohit Prabhavalkar (Google); Diamantino Caseiro (Google, Inc.); Pat Rondon (Google, Inc.); Cyril Allauzen (Google)

Tuesday, June 6

1655 (SLT-P3.4): STRUCTURED STATE SPACE DECODER FOR SPEECH RECOGNITION AND SYNTHESIS

Koichi Miyazaki (CyberAgent, Inc.); Masato Murata (CyberAgent, Inc.); Tomoki Koriyama (CyberAgent, Inc.)

3365 (SLT-P3.5): JEIT: JOINT END-TO-END MODEL AND INTERNAL LANGUAGE MODEL TRAINING FOR SPEECH RECOGNITION

Zhong Meng (Google LLC); Weiran Wang (Google); Rohit Prabhavalkar (Google); Tara Sainath (Google); Tongzhou Chen (Google); Ehsan Variani (Google); Yu Zhang (Google); Bo Li (Google); Andrew Rosenberg (Google LLC); Bhuvana Ramabhadran (Google)

3368 (SLT-P3.6): Variable Attention Masking for Configurable Transformer Transducer Speech Recognition

Pawel Swietojanski (Apple); Stefan Braun (Apple); Dogan Can (Apple); Thiago Fraga da Silva (Apple); Arnab Ghoshal (Apple); Takaaki Hori (Apple); Roger Hsiao (Apple); Henry Mason (Apple); Erik McDermott (Apple); Jan Silovsky (Apple); Ruchir Travadi (Apple); Xiaodan Zhuang (Apple)

3499 (SLT-P3.7): Factorized Blank Thresholding for Improved Runtime Efficiency of Neural Transducers

Duc Le (Meta); Frank Seide (Meta); Yuhao Wang (Meta); Yang Li (Meta); Kjell Schubert (Meta); Ozlem Kalinli (Meta); Mike Seltzer (Meta)

3926 (SLT-P3.8): Fast-U2++: Fast and Accurate End-to-End Speech Recognition in Joint CTC/Attention Frames

Chengdong Liang (Northwestern Polytechnical University); Zhang XiaoLei (Northwestern Polytechnical University); Binbin Zhang (Horizon Robotics); Di Wu (Horizon Robotics); Shengqiang Li (Horizon Robotics); Xingchen Song (Horizon Robotics); Zhendong Peng (Horizon Robotics); Fuping Pan (Horizon Robotics)

4365 (SLT-P3.9): Understanding Shared Speech-Text Representations

Yuan Wang (Google); Kyle Kastner (Google); Zhehuai Chen (Google); Ankur Bapna (Google Research); Andrew Rosenberg (Google LLC); Bhuvana Ramabhadran (Google); Yu Zhang (Google)

4534 (SLT-P3.10): Front-End Adapter: Adapting Front-End Input of Speech based Self-Supervised Learning for Speech Recognition

Xie Chen (Shanghai Jiaotong University); Ziyang Ma (Shanghai Jiao Tong University); Changli Tang (Tsinghua University); Yujin Wang (Tsinghua University); Zhisheng Zheng (Shanghai Jiao Tong University)

2237 (SLT-P3.11): Lego-Features: Exporting modular encoder features for streaming and deliberation ASR

Rami Botros (Google); Rohit Prabhavalkar (Google); Johan Schalkwyk (Google); Ciprian Chelba (Google Research); Tara Sainath (Google); Françoise Beaufays (Google)

5384 (SLT-P3.12): Modular Conformer Training for Flexible End-to-End ASR

Kartik Audhkhasi (Google); Brian Farris (Google); Bhuvana Ramabhadran (Google); Pedro J Moreno (Google)

SLT-P4: ASR: Noise Robustness

Room: Dome E

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Marc Delcroix, Soumi Maiti

1897 (SLT-P4.1): ON WORD ERROR RATE DEFINITIONS AND THEIR EFFICIENT COMPUTATION FOR MULTI-SPEAKER SPEECH RECOGNITION SYSTEMS

Thilo von Neumann (Paderborn University); Christoph B Boeddeker (Paderborn University); Keisuke Kinoshita (Google); Marc Delcroix (NTT); Reinhold Haeb-Umbach (University of Paderborn)

1919 (SLT-P4.2): Gradient Remedy for Multi-Task Learning in End-to-End Noise-Robust Speech Recognition

Yuchen Hu (Nanyang Technological University); Chen Chen (Nanyang Technological University); Ruizhe Li (University of Aberdeen); Qiu-Shi Zhu (University of Science and Technology of China); Eng Siong Chng (Nanyang Technological University)

1929 (SLT-P4.3): MAD: Inter-domain Matching and Intra-domain Discrimination for Cross-domain Speech Recognition

Jiaming Zhou (Nankai University); Shiwan Zhao (Independent Researcher); Ning Jiang (Mashang Consumer Finance Co., Ltd.); Guoqing Zhao (Mashang Consumer Finance Co., Ltd.); Yong Qin (Nankai University)

1971 (SLT-P4.4): ROBUST DATA2VEC: NOISE-ROBUST SPEECH REPRESENTATION LEARNING FOR ASR BY COMBINING REGRESSION AND IMPROVED CONTRASTIVE LEARNING

Qiu-Shi Zhu (University of Science and Technology of China); Long Zhou (Microsoft Research Asia); Jie Zhang (University of Science and Technology of China); Shujie Liu (Microsoft Research Asia); Yuchen Hu (Nanyang Technological University); Lirong Dai (University of Science and Technology of China)

2040 (SLT-P4.5): Robust Audio-Visual ASR with Unified Cross-modal Attention

Jiahong Li (Shanghai Jiao Tong University); Chenda Li (Shanghai Jiao Tong University); Yifei Wu (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)

3292 (SLT-P4.6): HuBERT-AGG: Aggregated Representation Distillation of Hidden-unit BERT for Robust Speech Recognition

wei wang (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)

Tuesday, June 6

4124 (SLT-P4.7): SPEECH AND NOISE DUAL-STREAM SPECTROGRAM REFINE NETWORK WITH SPEECH DISTORTION LOSS FOR ROBUST SPEECH RECOGNITION

Haoyu Lu (Tianjin University); Nan Li (Tianjin University); Tongtong Song (Tianjin University); Longbiao Wang (Tianjin University); Jianwu Dang (Tianjin University); Xiaobao Wang (Tianjin University); Shiliang Zhang (Alibaba Group)

4680 (SLT-P4.8): RobustDistiller: Compressing Universal Speech Representations for Enhanced Environment Robustness

Heitor R Guimarães (Institut National de la Recherche Scientifique); Arthur S Pimentel (Institut National de la Recherche Scientifique (INRS)); Anderson R Avila (INRS-EMT); Mehdi Rezagholizadeh (Huawei Technologies); Boxing Chen (Huawei Technologies); Tiago H Falk (INRS-EMT)

5455 (SLT-P4.9): Cleanformer: A Multichannel Array Configuration-Invariant Neural Enhancement Frontend for ASR in Smart Speakers

Joseph P Caroselli (Google); Arun Narayanan (Google Inc.); Nathan Howard (Google); Tom O'Malley (Google)

5504 (SLT-P4.10): On the effectiveness of monoaural target source extraction for distant end-to-end automatic speech recognition

Catalin Zorila (Toshiba Cambridge Research Laboratory); Rama S Doddipatla (Toshiba Europe LTD)

6389 (SLT-P4.11): Noise-aware target extension with self-distillation for robust speech recognition

Ju-seok Seong (Hanyang University); Jeong-Hwan Choi (Hanyang University); Jehyun Kyung (Hanyang University); Ye-Rin Jeoung (Hanyang University); Joon-Hyuk Chang (Hanyang University)

7177 (SLT-P4.12): Leveraging Domain Features for Detecting Adversarial Attacks Against Deep Speech Recognition in Noise

Christian H Lindbjerg (MapsPeople); Zheng-Hua Tan (Aalborg University)

AASP-L2: Audio Signal Restoration and Editing

Room: Salon de Roses A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Timo Gerkmann, Tomohiro Nakatani

02:00 PM

5003 (AASP-L2.1): AERO: AUDIO SUPER RESOLUTION IN THE SPECTRAL DOMAIN

Moshe Mandel (Hebrew University of Jerusalem); Or Tal (Hebrew University of Jerusalem); Yossi Adi (Bar-Ilan University)

02:15 PM

1768 (AASP-L2.2): UPGLADE: Unplugged Plug-and-Play audio declipper based on consensus equilibrium of DNN and sparse optimization

Tomoro Tanaka (Waseda University); Kohei Yatabe (Tokyo University of Agriculture and Technology); Yasuhiro Oikawa (Waseda University)

02:30 PM

2121 (AASP-L2.3): Improving performance of real-time full-band blind packet-loss concealment with predictive network

Nguyen Viet Anh (NamiTech JSC); Anh Nguyen (NamiTech JSC); Andy W H Khong (Nanyang Technological University)

02:45 PM

4388 (AASP-L2.4): Faster Than Fast: Accelerating The Griffin-Lim Algorithm

Rossen Nenov (Austrian Academy of Sciences - Acoustics Research Institute); Dang-Khoa Nguyen (University of Vienna); Peter Balazs (Acoustics Research Institute, Austrian Academy of Sciences)

03:00 PM

3726 (AASP-L2.5): Improving phase-vocoder-based time stretching by time-directional spectrogram squeezing

Natsuki Akaishi (Waseda University); Kohei Yatabe (Tokyo University of Agriculture and Technology); Yasuhiro Oikawa (Waseda University)

03:15 PM

6288 (AASP-L2.6): Extreme Audio Time Stretching using Neural Synthesis

Leonardo Fierro (Aalto University); Alec P Wright (Aalto University); Vesa Valimäki (Aalto University); Matti Hämäläinen (Nokia Technologies)

Tuesday, June 6

GC-2: Epilepsy Detection Grand Challenge

Room: Nefeli B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): TBA

02:00 PM

6635 (GC-L2.1): Introduction

Christos Chatzichristos (KU Leuven); Miguel C Bhagubai (KU Leuven); Maarten De Vos (KU Leuven); Wim Van Paesschen (UZ Leuven)

02:20 PM

7015 (GC-L2.2): Lightweight Machine Learning for Seizure Detection on Wearable Devices

Baichuan Huang (Lund University); Azra Abtahi (Lund University); Amir Aminifar (Lund University)

02:32 PM

7021 (GC-L2.3): Pretrained Transformers for Seizure Detection

Saarang Panchavati (UCLA); Samuel Vander Dussen (UCLA); Hemal Semwal (UCLA); Ahmed Ali (UCLA); Justin Chen (UCLA); Haoran Li (UCLA); Corey Arnold (UCLA); William Speier (UCLA)

02:44 PM

7022 (GC-L2.4): Towards Interpretable Seizure Detection Using Wearables

Irfan Al-Hussaini (Georgia Institute of Technology); Cassie S Mitchell (Georgia Institute of Technology)

02:56 PM

7033 (GC-L2.5): OPTIMIZATION OF THE DEEP NEURAL NETWORKS FOR SEIZURE DETECTION

Andrey Kiryasov (Brainify.AI); Aleksei Shovkun (Brainify.AI); Ilya Zakharov (Brainify.AI)

IVMSP-L2: Deep Learning Theory

Room: Athina

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Adrian Bors, Xiaoheng Deng

02:00 PM

2465 (IVMSP-L2.1): MSFormer: Multi-Scale Transformer with Neighborhood Consensus for Feature Matching

Dongyue Li (Southeast University); Yaping Yan (Southeast University); Dong Liang (Nanjing University of Aeronautics and Astronautics); Songlin Du (Southeast University)

02:15 PM

3498 (IVMSP-L2.2): Decoupled Visual Causality for Robust Detection

Ping Jiang (Central South University); Xiaoheng Deng (Central South University); Shichao Zhang (Central South University)

02:30 PM

2500 (IVMSP-L2.3): SEMANTICS-DISENTANGLED CONTRASTIVE EMBEDDING FOR GENERALIZED ZERO-SHOT LEARNING

Jian Ni (University of Science and Technology of China); Yong Liao (University of Sciences and Technology of China)

02:45 PM

4730 (IVMSP-L2.4): Dynamic Scalable Self-Attention Ensemble for Task-Free Continual Learning

Fei Ye (University of York); Adrian Bors (University of York)

03:00 PM

2125 (IVMSP-L2.5): Ultimate Negative Sampling For Contrastive Learning

Huijie Guo (Beihang University); Lei Shi (Beihang University)

03:15 PM

3936 (IVMSP-L2.6): An application of quantum mechanics to attention methods in computer vision

Juntao Zhang (Institute of System Engineering, AMS); Yihao Luo (Yichang Testing Technique R&D Institute); Peng Cheng (Coolanyp LLC); Zehan Li (University of Electronic Science and Technology of China); Hao Wu (Institute of System Engineering, AMS); Kun Yu (Institute of System Engineering, AMS); Wenbo An (Institute of System Engineering, AMS); Jun Zhou (Institute of System Engineering, AMS)

MLSP-L2: Neural Architecture Search

Room: Salon de Roses B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Zheng-Hua Tan,

Tuesday, June 6

02:00 PM

3492 (MLSP-L2.1): Search for efficient deep visual-inertial odometry through neural architecture search

Yu Chen (University of Michigan); Mingyu Yang (University of Michigan); Hun Seok Kim (Nil)

02:15 PM

4072 (MLSP-L2.2): Receptive Field Reliant Zero-Cost Proxies for Neural Architecture Search

Prateek Keserwani (Samsung Research Institute Bangalore); Srinivas S Miriyala (Samsung Research Institute Bangalore); Vikram Nelvoy Rajendiran (samsung Research Institute Bangalore); Pradeep Nelahonne Shivamurthappa (Samsung R & D Institute Bangalore)

02:30 PM

4346 (MLSP-L2.3): ZO-DARTS: DIFFERENTIABLE ARCHITECTURE SEARCH WITH ZERO-TH-ORDER APPROXIMATION

Lunchen Xie (Tongji University); Kaiyu Huang (Tongji University); Fan Xu (Peng Cheng Laboratory); Qingjiang Shi (Tongji University)

02:45 PM

2675 (MLSP-L2.4): Performing Neural Architecture Search Without Gradients

Pavel Rumiantsev (McGill University); Mark Coates (McGill University)

03:00 PM

796 (MLSP-L2.5): Neural Architecture of Speech

Subba Reddy Oota (IIIT Hyderabad); Khushbu Pahwa (University of California Los Angeles); Mounika Marreddy (IIIT Hyderabad); Manish Gupta (Microsoft); Raju Surampudi Bapi (International Institute of Information Technology Hyderabad)

03:15 PM

1461 (MLSP-L2.6): BHE-DARTS: Bilevel Optimization based on Hypergradient Estimation for Differentiable Architecture Search

Zicheng Cai (Guangdong University of Technology); Lei Chen (Guangdong University of Technology); Hai-Lin Liu (Guangdong University of Technology)

SLT-L3: Expressive and Controllable TTS I

Room: Jupiter

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Haizhou Li, Junichi Yamagishi

02:00 PM

2625 (SLT-L3.1): Improving Speech Prosody of Audiobook Text-to-Speech Synthesis with Acoustic and Textual Contexts

Detai Xin (The University of Tokyo); Sharath Adavanne (Rakuten Inc.); Federico Ang (Rakuten Inc.); Ashish Kulkarni (Rakuten); Shinnosuke Takamichi (The University of Tokyo); Hiroshi Saruwatari (The University of Tokyo)

02:15 PM

4768 (SLT-L3.2): Context-aware Coherent Speaking Style Prediction with Hierarchical Transformers for Audiobook Speech Synthesis

Shun Lei (Tsinghua University); Yixuan Zhou (Tsinghua University); Liyang Chen (Tsinghua University); Zhiyong Wu (Tsinghua University); Shiyin Kang (XVerse Inc.); Helen Meng (The Chinese University of Hong Kong)

02:30 PM

4776 (SLT-L3.3): Ensemble prosody prediction for expressive speech synthesis

Tian Huey Teh (Papercup); Vivian Hu (Papercup); Devang Mohan (Papercup); Zack Hodari (Papercup); Christopher Wallis (Papercup); Tomás Gómez Ibarrondo (Papercup); Alexandra Torresquintero (Papercup); James Leoni (Papercup); Mark Gales (University of Cambridge); Simon King (University of Edinburgh)

02:45 PM

5782 (SLT-L3.4): EXPRESSIVE-VC: HIGHLY EXPRESSIVE VOICE CONVERSION WITH ATTENTION FUSION OF BOTTLENECK AND PERTURBATION FEATURES

Ziqian Ning (Northwestern Polytechnical University); Qicong Xie (Northwestern Polytechnical University); Pengcheng Zhu (Fuxi AI Lab, NetEase Inc.); Zhichao Wang (Northwestern Polytechnical University); Liუმeng Xue (Northwestern Polytechnical University); Jixun Yao (Northwestern Polytechnical University); Lei Xie (NWPUI); Mengxiao Bi (Netease Fuxi AI Lab)

03:00 PM

5970 (SLT-L3.5): HIGH-ACOUSTIC FIDELITY TEXT TO SPEECH SYNTHESIS WITH FINE-GRAINED CONTROL OF SPEECH ATTRIBUTES

Rafael Valle (NVIDIA); João Felipe Santos (NVIDIA); Kevin Shih (NVIDIA); Rohan Badlani (NVIDIA); Bryan Catanzaro (NVIDIA)

03:15 PM

6203 (SLT-L3.6): Embedding a differentiable mel-cepstral synthesis filter to a neural speech synthesis system

Takenori Yoshimura (Nagoya Institute of Technology); Shinji Takaki (Nagoya Institute of Technology); Kazuhiro Nakamura (Techno-Speech, Inc.); Keiichiro Oura (Techno-Speech, Inc.); Yukiya Hono (Nagoya Institute of Technology); Kei Hashimoto (Nagoya Institute of Technology); Yoshihiko Nankaku (Nagoya Institute of Technology); Keichi Tokuda (Department of Computer Science and Engineering, Nagoya Institute of Technology)

Tuesday, June 6

SLT-L4: Keyword Spotting

Room: Delphi

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Dimitris Dimitriadis, Arun Narayanan

02:00 PM

1848 (SLT-L4.1): Disentangled Training with Adversarial Examples For Robust Small-footprint Keyword Spotting

ZHENYU WANG (UTD); Li Wan (Meta); Biqiao Zhang (Meta); Yiteng Huang (Meta Platforms); Shang-Wen Li (Meta); Ming Sun (Meta); Xin Lei (Meta); Zhaojun Yang (Meta)

02:15 PM

3578 (SLT-L4.2): Dual-Attention Neural Transducers for Efficient Wake Word Spotting in Speech Recognition

Saumya Yashmohini Sahai (Amazon); Jing Liu (Amazon.com); Thejaswi Muniyappa (Amazon); Kanthashree Mysore Sathyendra (Amazon); Anastasios Alexandridis (Amazon.com); Grant Strimel (Amazon); Ross McGowan (Amazon); Ariya Rastrow (Amazon Alexa); Athanasios Mouchtaris (Amazon Alexa); Feng-Ju Chang (Amazon); Siegfried Kunzmann (Amazon)

02:30 PM

5025 (SLT-L4.3): FIXED-POINT QUANTIZATION AWARE TRAINING FOR ON-DEVICE KEYWORD-SPOTTING

Sashank Kumar Macha (Amazon); Om Oza (Amazon); Alex Escott (Amazon); Francesco Caliva (Amazon); Robbie Armitano (Amazon); Santosh Kumar Cheekatmalla (Amazon); Sree Hari Krishnan Parthasarathi (Amazon); Yuzong Liu (Amazon)

02:45 PM

5106 (SLT-L4.4): To Wake-up or Not to Wake-up: Reducing Keyword False Alarm by Successive Refinement

Yashas Malur Saidutta (Samsung Research America); Rakshith Sharma Srinivasa (Samsung Research America); Ching-Hua Lee (Samsung Research America); Chouchang Yang (Samsung Research America); Yilin Shen (Samsung Research America); Hongxia Jin (Samsung Research America)

03:00 PM

5584 (SLT-L4.5): Transcription free filler word detection with Neural semi-CRFs

Ge Zhu (University of Rochester); Yujia Yan (University of Rochester); Juan-Pablo Caceres (Stanford); Zhiyao Duan (University of Rochester)

03:15 PM

6078 (SLT-L4.6): The DKU Post-Challenge Audio-Visual Wake Word Spotting System for the 2021 MISP Challenge: Deep Analysis

Haoxu Wang (Wuhan University); Ming Cheng (Duke Kunshan University); Qiang Fu (Alibaba Group); Ming Li (Duke Kunshan University)

SPTM-L2: Detection and Classification

Room: Nafsika A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Ignacio Santamaria, Jean-Philippe Ovarlez

02:00 PM

657 (SPTM-L2.1): Passive detection of rank-one Gaussian signals for known channel subspaces and arbitrary noise

David Ramirez (Universidad Carlos III de Madrid); Ignacio Santamaria (University of Cantabria); Louis Scharf (University of Colorado)

02:15 PM

2389 (SPTM-L2.2): False alarm regulation for off-grid target detection with the Matched Filter

Pierre Develter (ONERA; SONDR, CentraleSupélec, Université Paris-Saclay); Jonathan Bosse (ONERA); Olivier Rabaste (ONERA); Philippe Forster (ENS Paris-Saclay, CNRS, Université Paris-Saclay); Jean-Philippe Ovarlez (ONERA; SONDR, CentraleSupélec, Université Paris-Saclay)

02:30 PM

2536 (SPTM-L2.3): Data-Driven Quickest Change Detection in Markov Models

Qi Zhang (University at Buffalo); Zhongchang Sun (University at Buffalo, the State University of New York); Luis Herrera (University at Buffalo); Shaofeng Zou (University at Buffalo, the State University of New York)

02:45 PM

3510 (SPTM-L2.4): Quickest Change Detection with Leave-one-out Density Estimation

Yuchen Liang (UIUC); Venugopal V. Veeravalli (University of Illinois at Urbana Champaign)

03:00 PM

4778 (SPTM-L2.5): Identifying Coordination in a Cognitive Radar Network - A Multi-Objective Inverse Reinforcement Learning Approach

Luke Snow (Cornell University); Vikram Krishnamurthy (Cornell University); Brian M Sadler (Army Research Laboratory, USA)

Tuesday, June 6

03:15 PM

4815 (SPTM-L2.6): Improved Small Sample Hypothesis Testing using the Uncertain Likelihood Ratio
James Z Hare (DEVCOM Army Research Lab); Lance Kaplan (DEVCOM Army Research Laboratory)

SS-L2: Advances in Signal Processing and Machine Learning for Non-Intrusive Load Monitoring

Room: Nefeli A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Iosif Mporas,

02:00 PM

2170 (SS-L2.1): A Wavelet Scattering Approach For Load Identification with Limited Amount of Training Data
Pascal A Schirmer (University of Hertfordshire); Iosif Mporas (University of Hertfordshire)

02:15 PM

2653 (SS-L2.2): Applying Symmetrical Component Transform for Industrial Appliance Classification in Non-Intrusive Load Monitoring

Anthony Faustine (Imr); Lucas Pereira (ITI, LARSyS, Técnico Lisboa)

02:30 PM

3326 (SS-L2.3): CONTINILM: A Continual Learning Scheme for Non-Intrusive Load Monitoring

Stavros Sykiotis (National Technical University of Athens); Maria Kaselimi (National Technical University of Athens); Anastasios Doulamis (Technical University of Crete); Nikolaos Doulamis (National Technical University of Athens)

02:45 PM

5853 (SS-L2.4): Improving Knowledge Distillation for Non-Intrusive Load Monitoring through Explainability Guided Learning

Djordje Batic (University of Strathclyde); Giulia Tanoni (Università Politecnica delle Marche); Lina Stankovic (University of Strathclyde); Vladimir Stankovic (University of Strathclyde); Emanuele Principi (Università Politecnica delle Marche)

03:00 PM

6414 (SS-L2.5): IMPROVED APPLIANCE TRANSIENT FEATURE EXTRACTION VIA TEMPLATE MATCHING

Bo Liu (Tianjin University); Fenglei Chang (Tianjin University); Wenpeng Luan (Tianjin University); Bochao Zhao (Tianjin University)

ASPS-P2: Machine Learning Applications

Room: Dome A

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Robin Scheibler, Ryan Corey

6355 (ASPS-P2.1): Causal discovery and causal inference based counterfactual fairness in machine learning

Yajing Wang (BNU-HKBU United International College); Zongwei Luo (BNU ZH)

4965 (ASPS-P2.2): Benchmarking Convolutional Neural Network Inference on Low-Power Edge Devices

Oscar Ferraz (IT, Dep. of Electrical and Computer Engineering, University of Coimbra, Portugal); Helder Araujo (University of Coimbra); Vítor Silva (IT, Dep. of Electrical and Computer Engineering, University of Coimbra, Portugal); Gabriel Falcao (IT, University of Coimbra, Portugal)

1115 (ASPS-P2.3): Code-Enhanced Fine-Grained Semantic Matching for Tag Recommendation in Software Information Sites

Lin Li (Wuhan University of Technology); Peipei Wang (Wuhan University of Technology); Xinhao Zheng (Wuhan University of Technology); Qing Xie (Wuhan University of Technology)

394 (ASPS-P2.4): Robust Dominant Periodicity Detection for Time Series with Missing Data

Qingsong Wen (Alibaba DAMO Academy); Linxiao Yang (Machine Intelligence Technology, Alibaba Group, Hangzhou, China); Liang Sun (Alibaba Group)

3994 (ASPS-P2.5): Dynamic Split Computing for Efficient Deep Edge Intelligence

Arian Bakhtiarnia (Aarhus University); Nemanja B Milosevic (UNSPMF); Qi Zhang (Aarhus University); Dragana Bajovic (University of Novi Sad, Serbia); Alexandros Iosifidis (Aarhus University)

5723 (ASPS-P2.6): DENSE ADVERSARIAL TRANSFER LEARNING BASED ON CLASS-INVARIANCE

Bach-Tung Pham (National Central University); Ting-Yu Wang (National Central University); Le Phuong (National Central University); Khai-Thinh Nguyen (National Central University); Yuan-Shan Lee (National Central University); Tzu-Chiang Tai (Providence University); Jia-Ching Wang (National Central University)

4620 (ASPS-P2.7): VAN-ICP: GPU-Accelerated Approximate Nearest Neighbor Search for ICP Registration via Voxel Dilation

Weimin Wang (Dalian University of Technology); Qiong Chang (Tokyo Institute of Technology)

Tuesday, June 6

5776 (ASPS-P2.8): Clustering-based Supervised Contrastive Learning for Identifying Risk Items on Heterogeneous Graph
Ao Li (Alibaba Group); Yugang Ji (Alibaba Group); Guanyi Chu (Alibaba Group); Xiao Wang (Beijing University of Posts and Telecommunications); Dong Li (Alibaba Group); Chuan Shi (Beijing University of Posts and Telecommunications)

4052 (ASPS-P2.9): MULTIREOLUTION SIGNAL PROCESSING OF FINANCIAL MARKET OBJECTS
Ioana Boier (Nvidia)

1752 (ASPS-P2.10): HIERARCHICAL MULTI-AGENT REINFORCEMENT LEARNING WITH INTRINSIC REWARD RECTIFICATION
Zhihao Liu (Institute of Automation, Chinese Academy of Sciences); Zhiwei Xu (Institute of Automation, Chinese Academy of Sciences); Guoliang Fan (Institute of Automation, Chinese Academy of Sciences)

3493 (ASPS-P2.11): AN ANTISPOOFING APPROACH IN BIOMETRIC AUTHENTICATION SYSTEM FOR A SMARTCARD
Han-Sol Lee (Samsung Electronics); Moon-Kyu Song (Samsung Electronics); Junseo Lee (Samsung Electronics); Yeolmin Seong (Samsung Electronics); Ducksoo Kim (Samsung Electronics); Kwanghyuk Bae (Samsung Electronics); Seongwook Song (Samsung Electronics)

3576 (ASPS-P2.12): UNSUPERVISED DOMAIN ADAPTATION VIA SUBSPACE INTERPOLATING DEEP DICTIONARY LEARNING: A CASE STUDY IN MACHINE INSPECTION
Kriti Kumar (TCS Research and Innovation); Angshul Majumdar (IIIT Delhi); Achanna Anil Kumar (Tata Consultancy Services); Mariswamy Girish Chandra (Tata Consultancy Services)

IVMSP-P5: Classification
Room: Dome J
Type: Poster
02:00 PM to 03:30 PM
Chair(s): C.-C. Jay Kuo, Ran Tao

283 (IVMSP-P5.1): Multi-modal domain generalization for Cross-Scene Hyperspectral Image Classification
Yuxiang Zhang (Beijing Institute of Technology); Mengmeng Zhang (Beijing Institute of Technology); Wei Li (Beijing Institute of Technology, Beijing, China); Ran Tao (Beijing Institute of Technology)

1056 (IVMSP-P5.2): HIERARCHICAL TRANSFORMER FOR MULTI-LABEL TRAILER GENRE CLASSIFICATION
Zihui Cai (School of Cyber Science and Engineering, Wuhan University); Hongwei Ding (School of Cyber Science and Engineering, Wuhan University); Xuemeng Wu (School of Cyber Science and Engineering, Wuhan University); Mohan Xu (School of Cyber Science and Engineering, Wuhan University); Xiaohui Cui (School of Cyber Science and Engineering, Wuhan University)

1236 (IVMSP-P5.3): S3I-POINTHOP: SO(3)-INVARIANT POINTHOP FOR 3D POINT CLOUD CLASSIFICATION
Pranav A Kadam (University of Southern California); Handrik Prajapati (University of Southern California); Min Zhang (University of Southern California); Jinjang Xue (University of Southern California); Shan Liu (Tencent America); C.-C. Jay Kuo (USC)

1302 (IVMSP-P5.4): Sample-aware Knowledge Distillation for Long-tailed Learning
Shanshan Zheng (Xiamen University); Yachao Zhang (Tsinghua University); Yanyun Qu (XMU); hongyi huang (XMU)

1562 (IVMSP-P5.5): Laryngeal Leukoplakia Classification via Dense Multiscale Feature Extraction in White Light Endoscopy Images
Zhenzhen You (Xi'an University of Technology); Yan Yan (Second Affiliated Hospital of Medical College, Xi'an Jiaotong University); Zhenghao Shi (School of Computer Science and Engineering, Xi'an University of Technology); Minghua Zhao (Xi'an University of Technology); Jing Yan (Second Affiliated Hospital of Medical College, Xi'an Jiaotong University); Haiqin Liu (Second Affiliated Hospital of Medical College, Xi'an Jiaotong University); Xinhong Hei (Xi'an University of Technology); Xiaoyong Ren (Second Affiliated Hospital of Medical College, Xi'an Jiaotong University)

1904 (IVMSP-P5.6): Long-tailed Recognition with Causal Invariant Transformation
Yahong zhang (Lenovo); Sheng Shi (Lenovo Research); Chenchen Fan (Lenovo Research); Yixin Wang (Lenovo Research); Wenli Ouyang (Lenovo AI lab); Wei Fan (Lenovo); Jianping Fan (Lenovo)

2199 (IVMSP-P5.7): StackMaps: A Visualization Technique for Diabetic Retinopathy Grading
Ismail M El-Yamany (Alexandria University); Abdelrahman Wael (Faculty of Engineering, University of Alexandria); Noha Adly (MCT); Marwan Torki (Alexandria University)

2904 (IVMSP-P5.8): Gender-Cartoon: Image cartoonization method based on gender classification
Long Feng (Northwest University); Xingrui Ma (Northwest University); Chen Guo (Shaanxi Normal University); longquan yan (Northwest University); Guohua Geng (Northwest University); Zhan Li (Northwest University); Kang Li (Northwest University)

3167 (IVMSP-P5.9): Extracting the Brain-like Representation by an Improved Self-Organizing Map for Image Classification
Jiahong Zhang (Communication University of China); Lihong Cao (Communication University of China); Moning Zhang (Communication University of China); Wenlong Fu (Communication University of China)

Tuesday, June 6

3888 (IVMSP-P5.10): DDN: Dynamic Aggregation Enhanced Dual-stream Network for Medical Image Classification
Lang Wang (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Juan Liu (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Peng Jiang (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Dehua Cao (Landing Artificial Intelligence Center for Pathological Diagnosis); Baochuan Pang (Landing Artificial Intelligence Center for Pathological Diagnosis)

4696 (IVMSP-P5.11): LGViT: Local-Global Vision Transformer for Breast Cancer Histopathological Image Classification
Lang Wang (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Juan Liu (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Peng Jiang (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Dehua Cao (Landing Artificial Intelligence Center for Pathological Diagnosis); Baochuan Pang (Landing Artificial Intelligence Center for Pathological Diagnosis)

5583 (IVMSP-P5.12): LEARNING A WEIGHT MAP FOR WEAKLY-SUPERVISED LOCALIZATION
Tal Shaharbany (Tel Aviv University); Lior Wolf (Tel Aviv University, Israel)

IVMSP-P6: Human Posture Estimation

Room: Dome K

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Nenghai Yu, Ce Zhu

301 (IVMSP-P6.1): INTERWEAVED GRAPH AND ATTENTION NETWORK FOR 3D HUMAN POSE ESTIMATION
Ti Wang (Peking University Shenzhen Graduate School); Hong Liu (Peking University Shenzhen Graduate School); Runwei Ding (Peking University Shenzhen Graduate School); Wenhao Li (Peking University); Yingxuan You (Peking University); Xia Li (ETH Zurich)

3696 (IVMSP-P6.2): Learning 3D Human Pose and Shape Estimation Using Uncertainty-Aware Body Part Segmentation
Ziming Wang (Fudan University); Han Yu (Fudan University); Xiaoguang Zhu (Shanghai Jiao Tong University); Zengwen Li (Chongqing Changan Automobile Co., Ltd.); Changxue Chen (Chongqing Changan Automobile Co., Ltd.); Liang Song (Fudan University)

3841 (IVMSP-P6.3): Monocular 3D Human Pose Estimation Based on Global Temporal-Attentive and Joints-Attention in Video
ruhan He (Wuhan Textile University); shanshan xiang (Wuhan Textile University); Tao Peng (Wuhan Textile University); Yongsheng Yu (武汉理工大学)

4380 (IVMSP-P6.4): EVOPOSE: A RECURSIVE TRANSFORMER FOR 3D HUMAN POSE ESTIMATION WITH KINEMATIC STRUCTURE PRIORS
Yaqi Zhang (University of Science and Technology of China); Yan Lu (University of Sydney); Bin Liu (University of Science and Technology of China); Zhiwei Zhao (University of Science and Technology of China); Qi Chu (University of Science and Technology of China); Nenghai Yu (University of Science and Technology of China)

142 (IVMSP-P6.5): HTNET: HUMAN TOPOLOGY AWARE NETWORK FOR 3D HUMAN POSE ESTIMATION
Jialun Cai (Peking university); Hong Liu (Peking University Shenzhen Graduate School); Runwei Ding (Peking University Shenzhen Graduate School); Wenhao Li (Peking University); Jianbing Wu (Peking University); Miaoju Ban (Peking University)

1107 (IVMSP-P6.6): Improving Occluded Human Pose Estimation via Linked Joints
Suhang Ye (Tsinghua University); Zebo Hong (Xiamen University); Jiawen Zheng (Xiamen University); ShengChuan Zhang (Xiamen University)

5121 (IVMSP-P6.7): Efficient and Effective Multi-Camera Pose Estimation with Weighted M-Estimate Sample Consensus
Xinyu Lin (University of Electronic Science and Technology of China); Yingjie Zhou (Sichuan University); Xun Zhang (Institut superieur d'electronique de Paris - ISEP); Yipeng Liu (University of Electronic Science and Technology of China); Ce Zhu (University of Electronic Science & Technology of China)

5668 (IVMSP-P6.8): AMPose: Alternately Mixed Global-Local Attention Model for 3D Human Pose Estimation
Hong-Xin Lin (National Taiwan University); Yun-Wei Chiu (National Taiwan University); Pei-Yuan Wu (National Taiwan University)

5750 (IVMSP-P6.9): FLOWPOSE: CONDITIONAL NORMALIZING FLOWS FOR 3D HUMAN POSE AND SHAPE ESTIMATION FROM MONOCULAR VIDEOS
Yaoyao Du (Tsinghua University); Zixiao Zhang (Huawei); Zhihao Li (Huawei Noah's Ark Lab); Peng Wei (Huawei Device BG); Qingmin Liao (Tsinghua University); Wenming Yang (Tsinghua University)

6050 (IVMSP-P6.10): Animal Re-identification Algorithm for Posture Diversity
zhimin he (Ningbo University); Jiangbo Qian (Ningbo University); Yan Diqun (Ningbo University); Chong Wang (Ningbo University); Yu Xin (Ningbo University)

Tuesday, June 6

6322 (IVMSP-P6.11): RETRIEVAL-BASED NATURAL 3D HUMAN MOTION GENERATION
Zehan Tan (Fudan University); Weidong Yang (Fudan University); Shuai Wu (Fudan University)

2453 (IVMSP-P6.12): Human Pose Estimation from Ambiguous Pressure Recordings with Spatio-temporal Masked Transformers
Vandad Davoodnia (Queen's University); Ali Etemad (Queen's University)

IVMSP-P7: Human Reconstruction
Room: Dome L
Type: Poster
02:00 PM to 03:30 PM
Chair(s): Soo-Chang Pei, Kin-Man Lam

4237 (IVMSP-P7.1): Time-Frequency Awareness Network for Human Mesh Recovery from Videos
Boyang Zhang (Ningxia University); Suping Wu (Ningxia University); Meining Jia (NingXia University)

2028 (IVMSP-P7.2): Diffusion Motion: Generate Text-Guided 3D Human Motion by Diffusion Model
Zhiyuan Ren (Michigan State University); Zhihong Pan (Baidu Research (USA)); Xin Zhou (Baidu USA); Le Kang (Baidu Research)

4667 (IVMSP-P7.3): GATOR: GRAPH-AWARE TRANSFORMER WITH MOTION-DISENTANGLED REGRESSION FOR HUMAN MESH RECOVERY FROM A 2D POSE
Yingxuan You (Peking University); Hong Liu (Peking University Shenzhen Graduate School); Xia Li (ETH Zurich); Wenhao Li (Peking University); Ti Wang (Peking University Shenzhen Graduate School); Runwei Ding (Peking University Shenzhen Graduate School)

5538 (IVMSP-P7.4): real-time Human reconstruction based on human pose prior and epipolar refinement
Kuncheng Luo (Tsinghua University); Zhiheng Li (Tsinghua University)

642 (IVMSP-P7.5): Efficient Feature Fusion for Learning-based Photometric Stereo
Yakun Ju (The Hong Kong Polytechnic University); Kin-Man Lam (The Hong Kong Polytechnic University); Jun Xiao (The Hong Kong Polytechnic University); Cong Zhang (The Hong Kong Polytechnic University); Cuixin Yang (The Hong Kong Polytechnic University); Junyu Dong (Ocean University of China)

2442 (IVMSP-P7.6): Volumetric 3D Reconstruction with Window-wise Global Feature Aggregation
Shihao Ren (Tsinghua University); Yikang Ding (Tsinghua University); Jinli Liao (Tsinghua University); Xinghui Li (Tsinghua University); Jia Guo (None); Wensen Feng (the Shenzhen Graduate School, Tsinghua University, Shenzhen 518071, China); Xueqian Wang (Tsinghua University)

4008 (IVMSP-P7.7): STEREOSCOPIC VIDEO RETARGETING BASED ON CAMERA MOTION CLASSIFICATION
Linghui Cai (Guangxi University); Zhenhua Tang (Guangxi University)

4893 (IVMSP-P7.8): Detail-aware Uncalibrated Photometric Stereo
Antonio Agudo (Institut de Robotica i Informatica Industrial, CSIC-UPC)

5712 (IVMSP-P7.9): SDRNet: Shape Decoupled Regression Network for 3D Face Reconstruction
Shikun Zhang (Nanjing Normal University); Fengyi Song (Nanjing Normal University); GE SONG (Nanjing Normal University); Ming Yang (Nanjing Normal University)

1119 (IVMSP-P7.10): Binary Image Fast Perfect Recovery From Sparse 2D-DFT Coefficients
Soo-Chang Pei (Department of Electrical Engineering, National Taiwan University); Kuo-Wei Chang (Chunghwa Telecom)

1175 (IVMSP-P7.11): HQP-MVS:A HIGH-QUALITY PLANE PRIOR ASSISTED MULTI-VIEW STEREO FOR LOW-TEXTURED AREA
zefan tian (peking university); Rongjie Wang (PCL); Zhenyu Wang (Shenzhen Graduate School, Peking University); Ronggang Wang (Peking University)

3183 (IVMSP-P7.12): Dynamic Multi-View Scene Reconstruction Using Neural Implicit Surface
Decai Chen (Fraunhofer Heinrich Hertz Institute); Haofei Lu (The Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute); Ingo Feldmann (Fraunhofer HHI); Oliver Schreer (Fraunhofer Heinrich-Hertz-Institute); Peter Eisert (Fraunhofer HHI / Humboldt University Berlin)

Tuesday, June 6

IVMSP-P8: Face Recognition

Room: Dome M

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Bin Sun, Marios Savvides

3959 (IVMSP-P8.1): LOGO-Former: Local-Global Spatio-Temporal Transformer for Dynamic Facial Expression Recognition
Fuyan Ma (Hunan University); Bin Sun (Hunan University); Shutao Li (Hunan University)

4254 (IVMSP-P8.2): QUATERNION ORTHOGONAL TRANSFORMER FOR FACIAL EXPRESSION RECOGNITION IN THE WILD
Yu Zhou (Huazhong University Of Science And Technology); Liyuan Guo (Huazhong University of Science and Technology); Lianghai Jin (Huazhong University of Science and Technology)

3490 (IVMSP-P8.3): Privacy Preserving Face Recognition with Lensless Camera
Chris Henry (University of Missouri-Kansas City); M. Salman Asif (University of California, Riverside); Zhu Li (university of missouri-kansas city)

3649 (IVMSP-P8.4): MaskDUL: Data Uncertainty Learning in Masked Face Recognition
Libo Zhang (Southwest University); Weiming Xiong (Southwest University); Ku Zhao (SWU); Kehan Chen (Southwest University); Mingyang Zhong (Southwest University)

4814 (IVMSP-P8.5): Cov loss: Covariance-based Loss for Deep Face Recognition
Ibrahim Alkanhal (Carnegie Mellon University); Abdullah Almansour (National Center for Artificial Intelligence); Lamia Alsalloom (National Center for Artificial Intelligence); Raied Aljadaany (National Center for Artificial Intelligence); Marios Savvides (Carnegie Mellon University)

5674 (IVMSP-P8.6): Boosting Face Recognition Performance with Synthetic Data and Limited Real Data
Wenqing Wang (University of Macau); Lingqing Zhang (University of Macau); Chi-Man Pun (University of Macau); Jiucheng Xie (Nanjing University of Posts and Telecommunications)

2762 (IVMSP-P8.7): A Dual-branch Adaptive Distribution Fusion Framework for Real-world Facial Expression Recognition
Shu Liu (Central South University); Yan Xu (Central South University); Tongming Wan (Central South University); Xiaoyan Kui (Central South University)

4199 (IVMSP-P8.8): Efficient Practices for Profile-to-Frontal Face Synthesis and Recognition
Huijiao Wang (Wuhan University); Xulei Yang (Institute for Infocomm Research (I2R), ASTAR)

4208 (IVMSP-P8.9): LEARNING CAUSAL REPRESENTATIONS FOR GENERALIZABLE FACE ANTI SPOOFING
Guanghao Zheng (Shanghai Jiao Tong University); Yuchen Liu (Shanghai Jiao Tong university); Wenrui Dai (Shanghai Jiao Tong University); Chenglin Li (Shanghai Jiao Tong University); Junni Zou (Shanghai Jiao Tong University); Hongkai Xiong (Shanghai Jiao Tong University)

2767 (IVMSP-P8.10): Self-paced Partial Domain-Aware Learning for Face Anti-spoofing
Zhiyi Chen (Xiamen University); Yao Lu (Xiamen University); Xinzhe Deng (Tencent); Jia Meng (Tencent); ShengChuan Zhang (Xiamen University); Liujuan Cao (Xiamen University)

746 (IVMSP-P8.11): Context-Aware Face Clustering with Graph Convolutional Networks
dafeng zhang (Samsung Research China – Beijing (SRCB)); Jiangbo Guo (Samsung Research China – Beijing (SRCB)); Zhezhu Jin (Samsung Research Institute China – Beijing (SRC-B))

MLSP-P5: Source Separation, ICA, and Sparsity

Room: Dome F

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

193 (MLSP-P5.1): A Multi-Stage Triple-Path Method for Speech Separation in Noisy and Reverberant Environments
Zhaoxi Mu (Xi'an Jiaotong University); Xinyu Yang (Xi'an Jiaotong University); Xiangyuan Yang (Xi'an Jiaotong University); WenJing Zhu (DXM)

524 (MLSP-P5.2): On the minimum perimeter criterion for bounded component analysis
Sergio Cruces (Universidad de Sevilla)

4129 (MLSP-P5.3): Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images
Kinan ABBAS (Univ. Littoral Côte d'Opale , LISIC); Matthieu PUIGT (Univ. Littoral Côte d'Opale, LISIC); Gilles Delmaire (LISIC); Gilles Roussel (Univ. Littoral Côte d'Opale)

5036 (MLSP-P5.4): Identifiable Bounded Component Analysis via Minimum Volume Enclosing Parallelotope
Jingzhou Hu (University of Florida); Kejun Huang (University of Florida)

Tuesday, June 6

5587 (MLSP-P5.5): BALANCED DEEP CCA FOR BIRD VOCALIZATION DETECTION

SUMIT KUMAR (IIT Kanpur); B Anshuman (IIT Kanpur); Linus Ruettimann (University of Zurich and ETH Zurich); Richard Hahnloser (University of Zurich and ETH Zurich); Vipul Arora (IIT Kanpur)

1692 (MLSP-P5.6): Independent Vector Analysis with multivariate Gaussian model: a scalable method by multilinear regression

Ben Gabrielson (University of Maryland, Baltimore County); Mingyu Sun (University of Maryland, Baltimore County); Mohammad Akhonda (UMBC); Vince Calhoun (TReNDS); Tulay Adali (University of Maryland, Baltimore County)

3184 (MLSP-P5.7): ACTIVITY-INFORMED INDUSTRIAL AUDIO ANOMALY DETECTION VIA SOURCE SEPARATION

Jaechang Kim (POSTECH); YUNJOO LEE (POSTECH); Hyun Mi Cho (POSCO ICT); Dong Woo Kim (POSCO ICT); Chi Hoon Song (POSCO ICT); Jungseul Ok (POSTECH)

6717 (MLSP-P5.8): Double Nonstationarity: Blind Extraction of Independent Nonstationary Vector/Component from Nonstationary Mixtures – Algorithms (SPS Journal Paper)*

Zbynek Koldovsky (Technical University of Liberec)

6798 (MLSP-P5.9): Towards Flexible Sparsity-Aware Modeling: Automatic Tensor Rank Learning Using the Generalized Hyperbolic Prior (SPS Journal Paper)*

Zhongtao Chen (The University of Hong Kong); Lei Cheng (Zhejiang University)

5426 (MLSP-P5.10): MedleyVox: An Evaluation Dataset for Multiple Singing Voices Separation

Chang-Bin Jeon (Seoul National University); Hyeonggi Moon (Gaudio Lab.); Keunwoo Choi (Gaudio Lab); Ben Sangbae Chon (Gaudio Lab); Kyogu Lee (Seoul National University)

674 (MLSP-P5.11): Hybrid Transformers For Music Source Separation

Simon Rouard (Meta AI Research); Francisco Massa (Facebook AI Research); Alexandre Défossez (Meta AI Research)

5141 (MLSP-P5.12): DICTIONARY LEARNING ON GRAPH DATA WITH WEISFIELER-LEHMAN SUB-TREE KERNEL AND KSVD

Kaveen G Liyanage (Montana State University); Reese Pearsall (Montana State University); Clemente Izurieta (Montana State University); Bradley M Whitaker (Montana State University)

MLSP-P6: Neural Sound Synthesis and Representation

Room: Dome G

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Danilo Comminiello,

2678 (MLSP-P6.1): GANStrument: Adversarial Instrument Sound Synthesis with Pitch-invariant Instance Conditioning

Gaku Narita (Sony Computer Science Laboratories); Junichi Shimizu (Sony Computer Science Laboratories); Taketo Akama (Sony CSL)

2555 (MLSP-P6.2): I hear your true colors: Image Guided Audio Generation

Roy Sheffer (The Hebrew University of Jerusalem, Israel); Yossi Adi (Facebook AI Research)

1261 (MLSP-P6.3): Grad-StyleSpeech: Any-speaker Adaptive Text-to-Speech Synthesis with Diffusion Models

Minki Kang (AITRICS, KAIST); Dongchan Min (KAIST); Sung Ju Hwang (KAIST, AITRICS)

3085 (MLSP-P6.4): Voice Conversion Using Feature Specific Loss Function based Self-Attentive Generative Adversarial Network

Sandipan Dhar (National Institute of Technology Durgapur); Padmanabha Banerjee (Jalpaiguri Engineering College); Dr. Nanda Dulal Jana (NIT Durgapur); Swagatam Das (Indian Statistical Institute)

1268 (MLSP-P6.5): TRIAAN-VC: TRIPLE ADAPTIVE ATTENTION NORMALIZATION FOR ANY-TO-ANY VOICE CONVERSION

Hyun Joon Park (Korea University); Seok Woo Yang (Korea University); Jin Sob Kim (Korea University); Wooseok Shin (Korea University); Sung Won Han (Korea University)

6748 (MLSP-P6.6): Decorrelating Feature Spaces for Learning General-Purpose Audio Representations (SPS Journal Paper)*

Sreyan Ghosh (University of Maryland, College Park); Ashish Seth (IIT Madras); S Umesh (IIT Chennai)

4904 (MLSP-P6.7): Continuous descriptor-based control for deep audio synthesis

Ninon Devis (IRCAM); Nils Demerlé (IRCAM); Sarah Nabi (IRCAM); David Genova (IRCAM); Philippe Esling (IRCAM)

5786 (MLSP-P6.8): Rigid-Body Sound Synthesis with Differentiable Modal Resonators

Rodrigo Diaz (Queen Mary University of London); Ben Hayes (Queen Mary University of London); Charalampos Saitis (Queen Mary University of London); Gyorgy Fazekas (Queen Mary University of London); Mark Sandler (Queen Mary University of London)

Tuesday, June 6

5349 (MLSP-P6.9): Exploring Approaches to Multi-Task Automatic Synthesizer Programming

Daniel A Faronbi (New York University); Iran R Roman (NYU); Juan P Bello (New York University)

6710 (MLSP-P6.10): Speech Time-Scale Modification With GANs (SPS Journal Paper)*

Eyal Cohen (Technion); Joseph Keshet (Technion - Israel Institute of Technology); Felix Kreuk (Bar-Ilan University)

4339 (MLSP-P6.11): Full-band General Audio Synthesis with Score-based Diffusion

Santiago Pascual (Dolby Laboratories); Gautam Bhattacharya (Dolby Laboratories); Chunghsin Yeh (Dolby Laboratories); Jordi Pons (Dolby Laboratories); Joan Serra (Dolby Laboratories)

4443 (MLSP-P6.12): Is Quality Enough? Integrating Energy Consumption in a Large-Scale Evaluation of Neural Audio Synthesis Models

Constance Douwes (IRCAM); Giovanni Bindi (IRCAM); Antoine CAILLON (IRCAM); Philippe Esling (IRCAM); Jean-Pierre Briot (CNRS)

MLSP-P7: Deep Learning for Audio and Music Applications

Room: Dome H

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Paola Garcia

896 (MLSP-P7.1): Controllable music inpainting with mixed-level and disentangled representation

Shiqi Wei (Fudan University); Ziyu Wang (NYU Shanghai); Weiguo Gao (Fudan University); Gus Xia (New York University Shanghai)

1991 (MLSP-P7.2): HIPI: A Hierarchical Performer Identification model based on Symbolic Representation of Music.

Syed RM Rafee (QUEEN MARY UNIVERSITY OF LONDON); George Fazekas (QMUL); Geraint A. Wiggins (Vrije Universiteit Brussel)

207 (MLSP-P7.3): Chord-Conditioned Melody Harmonization with Controllable Harmonicity

Shangda Wu (Central Conservatory of Music); Xiaobing Li (Central Conservatory of Music); Maosong Sun (Tsinghua University)

1878 (MLSP-P7.4): Jazznet: A Dataset of Fundamental Piano Patterns for Music Audio Machine Learning Research

Tosiron Adegbija (University of Arizona)

5273 (MLSP-P7.5): Music Mixing Style Transfer: A Contrastive Learning Approach to Disentangle Audio Effects

Junghyun Koo (Seoul National University); Marco A Martinez Ramirez (Sony Group Corporation); WeiHsiang Liao (Sony Group Corporation); Stefan Uhlich (Sony European Technology Center); Kyogu Lee (Seoul National University); Yuki Mitsufuji (Sony Group Corporation)

1442 (MLSP-P7.6): An improved optimal transport kernel embedding method with gating mechanism for singing voice separation and speaker identification

Weitao Yuan (Tiangong University); Yuren Bian (Tiangong University); Shengbei Wang (Tiangong University); Masashi Unoki (JAIST); Wenwu Wang (University of Surrey)

3448 (MLSP-P7.7): Tempo vs. Pitch: understanding self-supervised tempo estimation

Giovana V Morais (University of São Paulo); Matthew Davies (INESTEC); Marcelo Queiroz (University of São Paulo); Magdalena Fuentes (New York University)

1995 (MLSP-P7.8): Adversarial Permutation Invariant Training for Universal Sound Separation

Emilian Postolache (Sapienza University of Rome); Jordi Pons (Dolby Laboratories); Santiago Pascual (Dolby Laboratories); Joan Serra (Dolby Laboratories)

1379 (MLSP-P7.9): Anomalous Sound Detection using Audio Representation with Machine ID based Contrastive Learning Pretraining

Jian Guan (Harbin Engineering University); Feiyang Xiao (Harbin Engineering University); Youde Liu (Harbin Institute of Technology); Qiaoxi Zhu (University of Technology Sydney); Wenwu Wang (University of Surrey)

4727 (MLSP-P7.10): Low-Resource Music Genre Classification with Cross-Modal Neural Model Reprogramming

Yun-Ning Hung (TikTok); Chao-Han Huck Yang (Georgia Institute of Technology); Pin-Yu Chen (IBM Research); Alexander Lerch (Georgia Institute of Technology)

1375 (MLSP-P7.11): SPADE: SELF-SUPERVISED PRETRAINING FOR ACOUSTIC DISENTANGLEMENT

John Harvill (University of Illinois at Urbana-Champaign); Jarred Barber (Google); Arun A Nair (Amazon Inc.); Ramin Pishehvar (Amazon)

1615 (MLSP-P7.12): On Out-of-Distribution Detection for Audio with Deep Nearest Neighbors

Zaharah Bukhsh (Eindhoven University of Technology); Aaqib Saeed (Eindhoven University of Technology)

Tuesday, June 6

MLSP-P8: Machine Learning for Image and Video Processing

Room: Dome I

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Qing Liu,

1011 (MLSP-P8.1): IoU-Aware Multi-Expert Cascade Network via Dynamic Ensemble for Long-tailed Object Detection

Wan-Cyuan Fan (National Taiwan University); Cheng-Yao Hong (Academia Sinica); Yen-Chi Hsu (Academia Sinica); Tyng-Luh Liu (Academia Sinica)

1622 (MLSP-P8.2): Efficient Compressed Video Action Recognition via Late Fusion with a Single Network

Hayato Terao (Hokkaido University); Wataru Noguchi (Hokkaido University); Hiroyuki Iizuka (Hokkaido University); Masahito Yamamoto (Hokkaido University)

1649 (MLSP-P8.3): Amicable Aid: Perturbing Images to Improve Classification Performance

Juyeop Kim (Yonsei University); Jun-Ho Choi (Yonsei University); Soobeom Jang (Yonsei University); Jong-Seok Lee ("Yonsei University, Korea")

3861 (MLSP-P8.4): Spatial Cross-Attention for Transformer-based Image Captioning

Khoa Anh Ngo (Seoul National University); Kyuhong Shim (Seoul National University); Byonghyo Shim (Seoul National University)

3879 (MLSP-P8.5): Towards hyperbolic regularizers for point cloud part segmentation

Antonio Montanaro (Politecnico di Torino); Diego Valsesia (Politecnico di Torino); Enrico Magli (POLITO)

5265 (MLSP-P8.6): Clip4VideoCap: Rethinking CLIP for Video Captioning with Multiscale Temporal Fusion and Commonsense Knowledge

Tanvir Mahmud (The University of Texas at Austin); Feng Liang (The University of Texas at Austin); Yaling Qing (University of Texas at Austin); Diana Marculescu (The University of Texas at Austin)

6356 (MLSP-P8.7): Learning silhouettes with group sparse autoencoders

Emmanouil Theodosis (Harvard University); Demba Ba (Harvard)

5042 (MLSP-P8.8): DEEP LEARNING FOR LAGRANGIAN DRIFT SIMULATION AT THE SEA SURFACE

Daria Botvynko (ENIB); Carlos Granero-Belinchon (IMT Atlantique); Simon van Gennip (Mercator Ocean International); Abdesslam BENZINO (ENIB); ronan fablet (IMT Atlantique)

2382 (MLSP-P8.9): DIFFERENCE GUIDED VHR REMOTE SENSING IMAGE CHANGE DETECTION

Jiukai Sun (Northwestern Polytechnical University); Ganchao Liu (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University); Yuan Yuan (Northwestern Polytechnical University)

2696 (MLSP-P8.10): ADAPTIVE SUBMANIFOLD-PRESERVING SPARSE REGRESSION FOR FEATURE SELECTION AND MULTICLASS CLASSIFICATION

Rui Xu (Renmin University of China); Xun Liang (Renmin University of China)

6814 (MLSP-P8.11): Learning Multiscale Convolutional Dictionaries for Image Reconstruction (SPS Journal Paper)*

Tianlin Liu (University of Basel); Anadi Chaman (University of Illinois at Urbana-Champaign); David Belius (University of Basel); Ivan Dokmanic (University of Basel)

7162 (MLSP-P8.12): Impact of PolSAR pre-processing and balancing methods on complex-valued neural networks segmentation tasks

Jose Agustin Barrachina (ONERA & Centralesupelec); Chengfang Ren (Sondra - CentraleSupélec); Christele Morisseau (ONERA); Gilles Vieillard (ONERA); Jean-Philippe Ovarlez (ONERA, CentraleSupélec, SONDRRA, Université Paris-Saclay)

SLT-P5: ASR: Text Adaptation

Room: Dome B

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Jasha Droppo, Yanmin Qian

209 (SLT-P5.1): Improving Contextual Spelling Correction by External Acoustics Attention and Semantic Aware Data Augmentation

Xiaoqiang Wang (Microsoft); Yanqing Liu (Microsoft); Jinyu Li (Microsoft); sheng zhao (Microsoft)

1007 (SLT-P5.2): AdapITN: A FAST, RELIABLE, AND DYNAMIC ADAPTIVE INVERSE TEXT NORMALIZATION

Binh Thai Nguyen (Karlsruhe Institute of Technology); Duc Minh Nhat Le (Vietnam Artificial Intelligence Solutions); Quang Minh Nguyen (Vietnam Artificial Intelligence Solutions); Quoc Truong Do (Vietnam Artificial Intelligence Solutions); Chi-Mai Luong (ICTLab, University of Science and Technology of Hanoi, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam.); Alexander Waibel (Karlsruhe Institute of Technology)

Tuesday, June 6

1373 (SLT-P5.3): Fast and accurate factorized neural transducer for text adaption of end-to-end speech recognition models
Rui Zhao (Microsoft); JIAN XUE (Microsoft Corporation); Partha Parthasarathy (Microsoft); Veljko Miljanic (Microsoft); Jinyu Li (Microsoft)

1628 (SLT-P5.4): EFFECTIVE TRAINING OF RNN TRANSDUCER MODELS ON DIVERSE SOURCES OF SPEECH AND TEXT DATA
Takashi Fukuda (IBM Research); Samuel Thomas (IBM Research AI)

1672 (SLT-P5.5): Text is All You Need: Personalizing ASR Models using Controllable Speech Synthesis
Karren D Yang (Apple); Ting-Yao Hu (Carnegie Mellon University); Jen-Hao Rick Chang (Apple); Hema Koppula (Apple); Oncel Tuzel (Apple)

2409 (SLT-P5.6): SLOT-TRIGGERED CONTEXTUAL BIASING FOR PERSONALIZED SPEECH RECOGNITION USING NEURAL TRANSDUCERS
Sibo Tong (Amazon); Philip Harding (Amazon Alexa); Simon Wiesler (Amazon)

3355 (SLT-P5.7): Fine-grained Textual Knowledge Transfer to Improve RNN Transducers for Speech Recognition and Understanding
Vishal Sunder (The Ohio State University); Samuel Thomas (IBM Research AI); Jeff Kuo (IBM); Brian Kingsbury (IBM Research); Eric Fosler-Lussier (Ohio State)

4612 (SLT-P5.8): Gated contextual adapters for selective contextual biasing in neural transducers
Anastasios Alexandridis (Amazon.com); Kanthashree Mysore Sathyendra (Amazon); Grant Strimel (Amazon.com); Feng-Ju Chang (Amazon); Ariya Rastrow (Amazon Alexa); Nathan Susanj (Amazon.com); Athanasios Mouchtaris (Amazon Alexa)

4830 (SLT-P5.9): Internal Language Model Estimation based Adaptive Language Model Fusion for Domain Adaptation
Rao Ma (University of Cambridge); Xiaobo Wu (ByteDance); Jin Qiu (ByteDance); Yanan Qin (ByteDance); Haihua Xu (ByteDance); Peihao Wu (Bytedance); Zejun Ma (Bytedance)

4970 (SLT-P5.10): Adaptable End-to-End ASR Models using Replaceable Internal LMs and Residual Softmax
Keqi Deng (University of Cambridge); Phil Woodland (Machine Intelligence Laboratory, Cambridge University Department of Engineering)

5596 (SLT-P5.11): Mask The Bias: Improving Domain-Adaptive Generalization of CTC-based ASR with Internal Language Model Estimation
Nilaksh Das (AWS AI Labs, Amazon); Monica Sunkara (Amazon); Sravan Babu Bodapati (Amazon); Jinglun Cai (Amazon); Devang Kulshreshtha (Amazon); Jeff Farris (Amazon); Katrin Kirchhoff (Amazon)

6116 (SLT-P5.12): Factorized AED: Factorized Attention-based Encoder-Decoder for Text-only Domain Adaptive ASR
Xun Gong (Shanghai Jiaotong University); wei wang (Shanghai Jiao Tong University); Hang Shao (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)

SLT-P6: ASR: Training Methods

Room: Dome C

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Jinyu Li, Duc Le

3731 (SLT-P6.1): Weight Averaging: A Simple Yet Effective Method to Overcome Catastrophic Forgetting in Automatic Speech Recognition
Steven Vander Eeck (KU Leuven); Hugo Van hamme (KU LEUVEN)

112 (SLT-P6.2): Reducing the gap between streaming and non-streaming Transducer-based ASR models by adaptive two-stage knowledge distillation
Haitao Tang (iFlytek Research); Yu Fu (Zhejiang University); Lei Sun (iFlytek Research); Jiabin Xue (Harbin Institute of Technology); Dan Liu (iFLYTEK Co., LTD.); Yongchao Li (iFlytek Research); Zhiqiang Ma (iFlytek Research); Minghui Wu (iFlytek Research); Jia Pan (iFlytek Research); Genshun Wan (iFlytek Research); Ming'en Zhao (iFlytek Research)

164 (SLT-P6.3): Alignment Entropy Regularization
Ehsan Variani (Google); Ke Wu (Google); David Rybach (Google); Cyril Allauzen (Google); Michael Riley (Google)

392 (SLT-P6.4): From English to More Languages: Parameter-Efficient Model Reprogramming for Cross-Lingual Speech Recognition
Chao-Han Huck Yang (Georgia Institute of Technology); Bo Li (Google); Yu Zhang (Google); Nanxin Chen (John Hopkins University); Rohit Prabhavalkar (Google); Tara Sainath (Google); Trevor Strohman (Google)

1499 (SLT-P6.5): Neural Transducer Training: Reduced Memory Consumption with Sample-wise Computation
Stefan Braun (Apple); Erik McDermott (Apple); Roger Hsiao (Apple)

Tuesday, June 6

2433 (SLT-P6.6): TOWARDS DOMAIN GENERALISATION IN ASR WITH ELITIST SAMPLING AND ENSEMBLE KNOWLEDGE DISTILLATION

Rehan Ahmad (University of Sheffield); Md Asif Jalal (Samsung Research UK); Muhammad Umar Farooq (University of Sheffield); Anna L Ollerenshaw (University of Sheffield); Thomas Hain (University of Sheffield)

2677 (SLT-P6.7): Accelerating RNN-T Training and Inference Using CTC guidance

Yongqiang Wang (Google); Zhehuai Chen (Google); Chengjian Zheng (Google); Yu Zhang (Google); Wei Han (Google); Parisa Haghani (Google)

3382 (SLT-P6.8): Resource-Efficient Transfer Learning From Speech Foundation Model Using Hierarchical Feature Fusion

Zhouyuan Huo (Google); Khe C Sim (Google Inc.); Bo Li (Google); Dongseong Hwang (Google); Tara Sainath (Google); Trevor Strohman (Google)

3917 (SLT-P6.9): Robust Knowledge Distillation from RNN-T Models With Noisy Training Labels Using Full-Sum Loss

Mohammad Zeineldien (RWTH Aachen University / AppTek); Kartik Audhkhasi (Google); Murali Karthick Baskar (Google); Bhuvana Ramabhadran (Google)

5520 (SLT-P6.10): More Speaking or More Speakers?

Dan Berrebbi (Carnegie Mellon University); Ronan Collobert (Apple); Navdeep Jaitly (Apple); Tatiana Likhomanenko (Apple)

5845 (SLT-P6.11): FEDERATED LEARNING FOR ASR BASED ON WAV2VEC 2.0

Tuan Manh Nguyen (LIA, Avignon University); Salima Mdhaffar (LIA - University of Avignon); Natalia Tomashenko (LIA, University of Avignon); Jean-Francois Bonastre (Université d'Avignon); Yannick Estève (LIA - Avignon University)

6343 (SLT-P6.12): Estimating Shapley Values of Training Utterances for Automatic Speech Recognition Models

Ali Raza Syed (The Graduate Center, CUNY); Michael I Mandel (Brooklyn College, CUNY)

SLT-P7: ASR: VAD and Other Topics I

Room: Dome D

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Takashi Fuguda, Yongqiang Wang

691 (SLT-P7.1): Real-time Speech Interruption Analysis: From Cloud to Client Deployment

Quchen Fu (Vanderbilt University); Szu-Wei Fu (Microsoft Corporation); Yaran Fan (Microsoft Corporation); Yu Wu (Microsoft Research Asia); Zhuo Chen (Microsoft); Jayant Gupchup (Microsoft); Ross Cutler (Microsoft Corporation)

2005 (SLT-P7.2): Audio-to-Intent Using Acoustic-Textual Subword Representations from End-to-End ASR

Pranay Dighe (Apple); Prateeth Nayak (Apple); Ognjen Rudovic (Apple); Erik Marchi (Apple); Xiaochuan Niu (Apple); Ahmed Tewfik (Apple)

2615 (SLT-P7.3): Adaptive End-pointing with Deep Contextual Multi-armed Bandits

Do June Min (University of Michigan); Andreas Stolcke (Amazon); Anirudh Raju (Amazon Alexa); Colin Vaz (Amazon); Di He (Amazon Alexa); Venkatesh Ravichandran (Amazon); Viet Anh Trinh (Amazon)

2616 (SLT-P7.4): DYNAMIC SPEECH ENDPOINT DETECTION WITH REGRESSION TARGETS

Dawei Liang (UT Austin); Hang Su (Meta Platforms Inc); Tarun Singh (Meta Platforms Inc); Jay Mahadeokar (Meta Platforms Inc); Shanil Puri (Meta Platforms Inc); Jiedan Zhu (Meta Platforms Inc); Edison Thomaz (The University of Texas at Austin); Mike Seltzer (Meta Platforms Inc)

2665 (SLT-P7.5): Speaker Change Detection for Transformer Transducer ASR

Jian Wu (Microsoft); Zhuo Chen (Microsoft); Min Hu (Microsoft); Xiong Xiao (Microsoft); Jinyu Li (Microsoft)

4769 (SLT-P7.6): Less is more: A unified architecture for device-directed speech detection with multiple invocation types

Ognjen Rudovic (Apple); Wonil Chang (Apple); Vineet Garg (Apple); Pranay Dighe (Apple); Pramod Jaya Simha (Apple Inc); John Berkowitz (Apple); Ahmed Hussen Abdelaziz (Apple); Erik Marchi (Apple); Sachin Kajarekar (Apple); Saurabh Adya (Apple)

4865 (SLT-P7.7): SG-VAD: STOCHASTIC GATES BASED SPEECH ACTIVITY DETECTION

Jonathan Svirsky (Bar Ilan University); Ofir Lindenbaum (Yale)

5523 (SLT-P7.8): FILLER WORDS DETECTION WITH HARD CATEGORY MINING AND INTER-CATEGORY FOCAL LOSS

Zhiyuan Zhao (MSRA); Lijun Wu (Microsoft Research); Chuanxin Tang (Microsoft); Dacheng Yin (University of Science and Technology of China); Yucheng Zhao (University of Science and Technology of China); Chong Luo (MSRA)

5787 (SLT-P7.9): Unsupervised Voice Type Discrimination Score Adaptation Using X-vector Clusters

Mark R Lindsey (Carnegie Mellon University); Tyler Vuong (Carnegie Mellon University); Richard M Stern (Carnegie Mellon University)

Tuesday, June 6

6269 (SLT-P7.10): Multilingual Word Error Rate Estimation: e-WER3

Shammur Chowdhury (QCRI); Ahmed Ali (Qatar Computing Research Institute, HBKU)

5792 (SLT-P7.11): Multilingual Query-by-Example Keyword Spotting with Metric Learning and Phoneme-to-Embedding Mapping

Paul M Reuter (Fraunhofer IDMT - HSA); Christian Rollwage (Fraunhofer IDMT - HSA); Bernd Meyer (Carl von Ossietzky University Oldenburg)

SLT-P8: ASR: VAD and Other Topics II

Room: Dome E

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Jintao Jiang, Xiaohui Zhang

836 (SLT-P8.1): Keyword-Specific Acoustic Model Pruning for Open Vocabulary Keyword Spotting

Yujie Yang (Tsinghua University); Kun Zhang (The Chinese University of Hong Kong); Zhiyong Wu (Tsinghua University); Helen Meng (The Chinese University of Hong Kong)

5030 (SLT-P8.10): Self-supervised speech representation learning for keyword-spotting with light-weight transformers

Chenyang Gao (Rutgers University); Yue Gu (Amazon); Francesco Caliva (Amazon); Yuzong Liu (Amazon)

5579 (SLT-P8.11): Lightweight feature encoder for wake-up word detection based on self-supervised speech representation

Hyungjun Lim (LG AI Research); Younggwan Kim (LG AI Research); Kiho Yeom (LG AI Research); Eunjoo Seo (LG AI Research); Hoodong Lee (LG AI Research); Stanley Jungkyu Choi (LG AI Research); Honglak Lee (LG AI Research)

5649 (SLT-P8.12): VE-KWS: VISUAL MODALITY ENHANCED END-TO-END KEYWORD SPOTTING

Ao Zhang (Northwestern Polytechnical University); He Wang (NWPU); Pengcheng Guo (Northwestern Polytechnical University); Yihui Fu (Northwestern Polytechnical University); Lei Xie (NWPU); Yingying Gao (China Mobile Research Institute); Shilei Zhang (China Mobile Research Institute); Junlan Feng (China Mobile Research)

1378 (SLT-P8.2): Unified Keyword Spotting and Audio Tagging on Mobile Devices with Transformers

Heinrich Dinkel (Xiaomi Technology); Yongqing Wang (Xiaomi); Zhiyong Yan (Xiaomi); Junbo Zhang (Xiaomi); Yujun Wang (xiaomi)

1518 (SLT-P8.3): Continual Learning for On-Device Speech Recognition using Disentangled Conformers

Anuj Diwan (University of Texas at Austin); Ching-Feng Yeh (Facebook); Wei-Ning Hsu (Massachusetts Institute of Technology); Paden Tomasello (Meta); Eunsol Choi (University of Texas at Austin); David Harwath (The University of Texas at Austin); Abdelrahman Mohamed (Rembrand Inc)

1986 (SLT-P8.4): Filterbank Learning for Noise-Robust Small-Footprint Keyword Spotting

Iván López-Espejo (Aalborg University); RAM CHARAN M CHANDRA SHEKAR (University of Texas at Dallas); Zheng-Hua Tan (Aalborg University); Jesper Jensen (Aalborg University); John H Hansen (Univ. of Texas at Dallas)

3390 (SLT-P8.5): Locale Encoding for scalable multilingual keyword spotting models

Pai Zhu (Google); Hyun Jin Park (Google Inc.); Alex Park (Google); Angelo Scorza Scarpati (Google); Ignacio Lopez Moreno (Google)

3531 (SLT-P8.6): SMALL-FOOTPRINT SLIMMABLE NETWORKS FOR KEYWORD SPOTTING

Zuhaib Akhtar (Amazon); Mohammad Omar Khursheed (Amazon); Dongsu Du (AMAZON); Yuzong Liu (Amazon)

3615 (SLT-P8.7): METRIC LEARNING FOR USER-DEFINED KEYWORD SPOTTING

Jaemin Jung (KAIST); Youkyum Kim (KAIST); Jihwan Park (42dot Inc.); Youshin Lim (42dot); Byeong-Yeol Kim (42dot); Youngjoon Jang (KAIST); Joon Son Chung (KAIST)

3928 (SLT-P8.8): WeKws: A production first small-footprint end-to-end Keyword Spotting Toolkit

Jie Wang (School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an, China); Menglong Xu (Horizon Robotics); Jingyong Hou (Northwestern Polytechnical University); Binbin Zhang (Horizon Robotics); Zhang XiaoLei (Northwestern Polytechnical University); Lei Xie (NWPU); Fuping Pan (Horizon Robotics)

4822 (SLT-P8.9): Exploring Sequence-to-Sequence Transformer-Transducer Models for Keyword Spotting

Beltrán Labrador (Audias - Universidad Autónoma de Madrid); Guanlong Zhao (Google); Ignacio Lopez Moreno (Google); Angelo Scorza Scarpati (Google); Liam Fowl (Google); Quan Wang (Google)

Tuesday, June 6

AASP-L3: Automatic Audio Captioning and Retrieval

Room: Salon de Roses A

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Romain Serizel, Mark Cartwright

03:40 PM

662 (AASP-L3.1): A NOVEL METRIC FOR EVALUATING AUDIO CAPTION SIMILARITY

Swapnil P Bhosale (TCS Research and Innovation); Rupayan Chakraborty (TCS Research); Sunil Kumar Kopparapu (TCS Research)

03:55 PM

5376 (AASP-L3.2): On Negative Sampling for Contrastive Audio-Text Retrieval

Huang Xie (Tampere University); Okko Räsänen (Tampere University); Tuomas Virtanen (Tampere University)

04:10 PM

2001 (AASP-L3.3): Audio-Text Models Do Not Yet Leverage Natural Language

Ho-Hsiang Wu (New York University); Oriol Nieto (Pandora); Juan P Bello (New York University); Justin Salamon (Adobe Research)

04:25 PM

4981 (AASP-L3.4): IMPROVING AUDIO CAPTIONING USING SEMANTIC SIMILARITY METRICS

Rehana Mahfuz (Qualcomm); Yinyi Guo (Qualcomm); Erik Visser (Qualcomm)

04:40 PM

4900 (AASP-L3.5): SPICE+: Evaluation of Automatic Audio Captioning Systems with Pre-trained Language Models

Felix Gontier (INRIA); romain serizel (Université de Lorraine); Christophe Cerisara (CNRS)

04:55 PM

6766 (AASP-L3.6): Local Information Assisted Attention-Free Decoder for Audio Captioning (SPS Journal Paper)*

Feiyang Xiao (Harbin Engineering University); Jian Guan (Harbin Engineering University); Haiyan Lan (Harbin Engineering University); Qiaoxi Zhu (University of Technology Sydney); Wenwu Wang (University of Surrey)

GC-3: Auditory EEG Decoding Challenge

Room: Nefeli B

Type: Oral

03:40 PM to 05:10 PM

Chair(s): TBA

03:40 PM

6636 (GC-L3.1): Introduction

Lies Bollens (KU Leuven); Hugo Van hamme (KU LEUVEN); Mohammad Jalilpour Monesi (KU Leuven); Tom Francart (KU Leuven); Bernd Accou (ExpORL, KU Leuven); Jonas Vanthornhout (KU Leuven)

04:00 PM

6832 (GC-L3.2): HappyQuokka system for ICASSP 2023 Auditory EEG challenge

Zhenyu Piao (Yonsei University); Miseul Kim (Yonsei University); Hyungchan Yoon (Yonsei University); Hong-Goo Kang (Yonsei University)

04:12 PM

6855 (GC-L3.3): Relate Auditory Speech to EEG by Shallow-Deep Attention-based Network

Fan Cui (Mi); Liyong Guo (Xiaomi Corp.); Lang He (XUPT); Jiyao Liu (NWPU); Ercheng Pei (XUPT); Yujun Wang (Mi); Dongmei Jiang (Northwestern Polytechnical University \ Peng Cheng Laboratory)

04:24 PM

6859 (GC-L3.4): Multi-Head Attention and GRU for Improved Match-Mismatch Classification of Speech Stimulus and EEG Response

Marvin Borsdorf (University of Bremen); Saurav Pahuja (University of Bremen); Gabriel Ivucic (University of Bremen); Siqi Cai (National University of Singapore); Haizhou Li (The Chinese University of Hong Kong, Shenzhen); Tanja Schultz (University of Bremen)

04:36 PM

6861 (GC-L3.5): RELATING EEG RECORDINGS TO SPEECH USING ENVELOPE TRACKING AND THE SPEECH-FFR

Michael D Thornton (Imperial College London); Danilo Mandic (Imperial College London); Tobias Reichenbach (FAU)

04:48 PM

6882 (GC-L3.6): Decoding Auditory EEG Responses using an Adapted WaveNet

Bob M.S.L. Van Dyck (KU Leuven); Liuyin Yang (KU Leuven); Marc Van Hulle (KU Leuven)

Tuesday, June 6

IVMSP-L3: Image Restoration

Room: Athina

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Christine Guillemot, Edward Choi

03:40 PM

564 (IVMSP-L3.1): MRNet: Multi-Refinement Network for Dual-pixel Images Defocus Deblurring

dafeng zhang (Samsung Research China – Beijing (SRCB)); Xiaobing Wang (Samsung Research China-Beijing); Zhezhu Jin (Samsung Research Institute China – Beijing (SRC-B))

03:55 PM

5802 (IVMSP-L3.2): JOINT COMPRESSION AND DEMOSAICKING FOR SATELLITE IMAGES

Pascal Bacchus (INRIA); Renaud Fraisse (Airbus); Aline Roumy (INRIA); Christine Guillemot (INRIA)

04:10 PM

1157 (IVMSP-L3.3): Decontamination Transformer for Blind Image Inpainting

Chun-Yi Li (National Chiao Tung University); Yen-Yu Lin (National Yang Ming Chiao Tung University); Wei-Chen Chiu (National Chiao Tung University)

04:25 PM

658 (IVMSP-L3.4): EXPLORATION INTO TRANSLATION-EQUIVARIANT IMAGE QUANTIZATION

Woncheol Shin (Korea Advanced Institute of Science and Technology, KAIST); Gyubok Lee (KAIST); Jiyoung Lee (KAIST); Eunyi Lyou (Seoul national university); Joonseok Lee (Google Research & Seoul National University); Edward Choi (KAIST)

04:40 PM

2562 (IVMSP-L3.5): TENSOR DECOMPOSITION BASED LATENT FEATURE CLUSTERING FOR HYPERSPECTRAL BAND SELECTION

Jianwen Qi (China University of Geosciences); jie zhang (University of Macau); Yongshan Zhang (China University of Geosciences); Xinwei Jiang (China University of Geosciences); Zhihua Cai (China University of Geosciences)

MLSP-L3: Interpretable and Explainable Machine Learning

Room: Salon de Roses B

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Che Lin,

03:40 PM

1497 (MLSP-L3.1): Interpretability in the Context of Sequential Cost-Sensitive Feature Acquisition

Yasitha Warahena Liyanage (Microsoft); Daphney-Stavroula Zois (University at Albany)

03:55 PM

4603 (MLSP-L3.2): Understandable ReLU Neural Network for Signal Classification

Marie Guyomard (Université Côte d'Azur, CNRS, I3S); Susana Barbosa (Université Côte d'Azur, CNRS, IPMC); Lionel Fillatre (Université Côte d'Azur, CNRS, I3S)

04:10 PM

5323 (MLSP-L3.3): Interpretation of Neural Networks is Susceptible to Universal Adversarial Perturbations

Haniyeh Ehsani Oskouie (Sharif University of Technology); Farzan Farnia (The Chinese University of Hong Kong)

04:25 PM

6701 (MLSP-L3.4): Interpreting Intermediate Convolutional Layers of Generative CNNs Trained on Waveforms (SPS Journal Paper)*

Gasper Begus (UC Berkeley); Alan Zhou (Johns Hopkins University)

04:40 PM

6718 (MLSP-L3.5): Statistical Hypothesis Testing Based on Machine Learning: Large Deviations Analysis (SPS Journal Paper)*

Paolo Braca (NATO STO CMRE); Leonardo Millefiori (NATO STO CMRE); Augusto Aubry (University of Naples "Federico II"); Stefano Marano (University of Salerno); Antonio De Maio (University of Naples "Federico II"); Peter Willett (University of Connecticut)

04:55 PM

2051 (MLSP-L3.6): Interpretable Multi-scale Neural Network for Granger Causality Discovery

Chenchen Fan (Lenovo Research); Yixin Wang (Lenovo Research); Yahong zhang (lenovo); Wenli Ouyang (Lenovo AI lab)

Tuesday, June 6

SLT-L5: Language Modeling

Room: Jupiter

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Michael Seltzer, Hainan Xu

03:40 PM

971 (SLT-L5.1): Mitigating Unintended Memorization in Language Models via Alternating Teaching

Zhe Liu (Meta); Xuedong Zhang (Meta); Fuchun Peng (Facebook)

03:55 PM

2352 (SLT-L5.2): ITERATIVE SHALLOW FUSION OF BACKWARD LANGUAGE MODEL FOR END-TO-END SPEECH RECOGNITION

Atsunori Ogawa (NTT Corporation); Takafumi Moriya (NTT); Naoyuki Kamo (NTT Corporation); Naohiro Tawara (NTT); Marc Delcroix (NTT)

04:10 PM

2840 (SLT-L5.3): PROCTER: PRONUNCIATION-AWARE CONTEXTUAL ADAPTER FOR PERSONALIZED SPEECH RECOGNITION IN NEURAL TRANSDUCERS

Rahul Pandey (George Mason University); Roger Ren (Amazon); Qi Luo (Amazon.com Inc.); Jing Liu (Amazon.com); Ariya Rastrow (Amazon Alexa); Ankur Gandhe (Amazon Alexa); Denis Filimonov (Amazon); Grant Strimel (Amazon); Andreas Stolcke (Amazon); Ivan Bulyko (Amazon)

04:25 PM

3275 (SLT-L5.4): Adaptive Multi-Corpora Language Model Training for Speech Recognition

Yingyi Ma (Meta); Zhe Liu (Meta); Xuedong Zhang (Meta)

04:40 PM

4371 (SLT-L5.5): SUFFIX RETRIEVAL-AUGMENTED LANGUAGE MODELING

Ze Cheng Wang (New York University Shanghai); Yik-Cheung Tam (NYU Shanghai)

04:55 PM

4458 (SLT-L5.6): Large-scale Language Model Rescoring on Long-form Data

Tongzhou Chen (Google); Cyril Allauzen (Google); Yinghui Huang (Google); Daniel S Park (Google Brain); David Rybach (Google); W. Ronny Huang (Google); Rodrigo Cabrera (Google); Kartik Audhkhasi (Google); Bhuvana Ramabhadran (Google); Pedro J Moreno (Google); Michael Riley (Google)

SLT-L6: Language Modeling and Spoken Language Understanding

Room: Delphi

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Wen Wang, Ahmed Hussien Abdelaziz

03:40 PM

3463 (SLT-L6.1): Speech summarization of long spoken document: Improving memory efficiency of speech/text encoders

Takatomo Kano (NTT Corporation); Atsunori Ogawa (NTT Corporation); Marc Delcroix (NTT); Roshan S Sharma (Carnegie Mellon University); Kohei Matsuura (NTT); Shinji Watanabe (Carnegie Mellon University)

03:55 PM

4197 (SLT-L6.2): STREAMING JOINT SPEECH RECOGNITION AND DISFLUENCY DETECTION

Hayato Futami (Sony Group Corporation); Emiru Tsunoo (Sony Group Corporation); Kentaro Shibata (Sony); Yosuke Kashiwagi (Sony); Takao Okuda (Sony); Siddhant Arora (Carnegie Mellon University); Shinji Watanabe (Carnegie Mellon University)

04:10 PM

4584 (SLT-L6.3): Improved Training of Mixture-of-Experts Language GANs

Yekun Chai (Baidu Inc.); Qiyue Yin (Institute of Automation, Chinese Academy of Sciences); Junge Zhang (CASIA)

04:25 PM

5069 (SLT-L6.4): JOINT MODELLING OF SPOKEN LANGUAGE UNDERSTANDING TASKS WITH INTEGRATED DIALOG HISTORY

Siddhant Arora (Carnegie Mellon University); Hayato Futami (Sony Group Corporation); Emiru Tsunoo (Sony Group Corporation); Brian Yan (Carnegie Mellon University); Shinji Watanabe (Carnegie Mellon University)

04:40 PM

5242 (SLT-L6.5): WEIGHTED SAMPLING FOR MASKED LANGUAGE MODELING

Linhan Zhang (University of New South Wales); Qian Chen (Speech Lab, DAMO Academy, Alibaba Group); Wen Wang (Alibaba Group); Chong Deng (Alibaba inc); Xin Cao (University of New South Wales); Kongzhang Hao (UNSW); Yuxin Jiang (HKUST); Wei Wang (Hong Kong University of Science and Technology (Guangzhou))

Tuesday, June 6

04:55 PM

5445 (SLT-L6.6): COMPENSATORY DEBIASING FOR GENDER IMBALANCES IN LANGUAGE MODELS

Tae-Jin Woo (Korea University); Woo-Jeoung Nam (Kyungpook National University); Yeong-Joon Ju (Korea University); Seong-Whan Lee (Korea University)

SPTM-L3: Estimation Theory and Methods

Room: Nafsika A

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Arie Yeredor, Frédéric Pascal

03:40 PM

2120 (SPTM-L3.1): Large Covariance Matrix Estimation With Oracle Statistical Rate

Quan Wei (ShanghaiTech University); Ziping Zhao (ShanghaiTech University)

03:55 PM

2600 (SPTM-L3.2): Cramér-Rao bound on Lie groups with observations on Lie groups: application to $\$SE(2)\$$

Samy LABSIR (IPSA); Alexandre Renaux (Université Paris Saclay); Jordi Vilà-Valls (ISAE-SUPAERO); Eric Chaumette (ISAE-SUPAERO)

04:10 PM

2611 (SPTM-L3.3): Various Performance Bounds on the Estimation of Low-Rank Probability Mass Function Tensors from Partial Observations

Tomer Hershkovitz (Tel Aviv University); Martin Haardt (Ilmenau University of Technology); Arie Yeredor (Tel Aviv University)

04:25 PM

4611 (SPTM-L3.4): Radar Clutter Covariance Estimation: A Nonlinear Spectral Shrinkage Approach

Shashwat Jain (Cornell University); Vikram Krishnamurthy (Cornell University); Muralidhar Rangaswamy (AFRL); Bosung Kang (University of Dayton Research Institute); Sandeep Gogineni (Information Systems Laboratories Inc.)

04:40 PM

4639 (SPTM-L3.5): Elliptical Wishart distribution: maximum likelihood estimator from information geometry

Imen AYADI (université Paris Saclay); Florent Bouchard (L2S); Frédéric Pascal (CentraleSupélec)

04:55 PM

4746 (SPTM-L3.6): Consistent estimators of a new class of covariance matrix distances in the large dimensional regime

Roberto Pereira (Centre Tecnològic de Telecomunicacions de Catalunya); Xavier Mestre (Centre Tecnològic de Telecomunicacions de Catalunya); David Gregoratti (SRS)

SS-L3: AI Security and Privacy in Speech and Audio Processing

Room: Nefeli A

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Zhao Ren,

03:40 PM

673 (SS-L3.1): PRIVACY-ENHANCED FEDERATED LEARNING AGAINST ATTRIBUTE INFERENCE ATTACK FOR SPEECH EMOTION RECOGNITION

Huan Zhao (Hunan University); Haijiao Chen (Hunan University); Yufeng Xiao (Hunan University); Zixing Zhang (Hunan University)

03:55 PM

2009 (SS-L3.2): Privacy-Preserving Occupancy Estimation

Jennifer Williams (University of Southampton); Vahid Yazdanpanah (University of Southampton); Sebastian Stein (University of Southampton)

04:10 PM

3761 (SS-L3.3): FEDERATED INTELLIGENT TERMINALS FACILITATE STUTTERING MONITORING

Yongzi Yu (Beijing institute of technology); Wanyong Qiu (Beijing Institute of Technology); Chen Quan (Beijing Institute of Technology); Kun Qian (Beijing Institute of Technology); Zihua Wang (The University of Tokyo); Yu Ma (Beijing Institute of Technology); Bin Hu (Beijing Institute of Technology); Bjorn W. Schuller (Imperial College London); Yoshiharu Yamamoto (The University of Tokyo)

04:25 PM

4942 (SS-L3.4): Beyond Neural-on-Neural Approaches to Speaker Gender Protection

Loes van Bommel (Radboud University); Zhuoran Liu (Radboud University); Nik Vaessen (Radboud University); Martha Larson (Radboud University)

Tuesday, June 6

04:40 PM

6129 (SS-L3.5): Distinguishable Speaker Anonymization based on Formant and Fundamental Frequency Scaling

Jixun Yao (Northwestern Polytechnical University); Qing Wang (Northwestern Polytechnical University); Yi Lei (Northwestern Polytechnical University); Pengcheng Guo (Northwestern Polytechnical University); Lei Xie (NWPU); Namin Wang (Huawei Cloud); Jie Liu (Huawei Cloud)

AASP-P1: Binaural Audio; Multichannel Source Separation

Room: Dome A

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Zbynek Koldovsky, Shoko Araki

1755 (AASP-P1.1): SPATIALLY INFORMED INDEPENDENT VECTOR ANALYSIS FOR SOURCE EXTRACTION BASED ON THE CONVOLUTIVE TRANSFER FUNCTION MODEL

Xianrui Wang (Northwestern Polytechnical University); Andreas Brendel (Friedrich-Alexander-University Erlangen-Nürnberg); Gongping Huang (University of Erlangen-Nuremberg); Yichen Yang (Northwestern Polytechnical University); Walter Kellermann (Friedrich-Alexander-University Erlangen-Nürnberg); Jingdong Chen (Northwestern Polytechnical University)

2514 (AASP-P1.2): Fast Online Source Steering Algorithm for Tracking Single Moving Source Using Online Independent Vector Analysis

Taishi Nakashima (Tokyo Metropolitan University); Rintaro Ikeshita (NTT); Nobutaka Ono (Tokyo Metropolitan University); Shoko Araki (NTT Corporation); Tomohiro Nakatani (NTT Communication Science Laboratories)

4589 (AASP-P1.3): Online Binaural Speech Separation of Moving Speakers With a Wavesplit Network

Cong Han (Columbia University); Nima Mesgarani (Columbia University)

5759 (AASP-P1.4): Convolutional NTF for Ambisonic Source Separation Under Reverberant Conditions

Mateusz Guzik (AGH University of Science and Technology); Konrad Kowalczyk (AGH University of Science and Technology)

4677 (AASP-P1.5): On the relevance of the differences between HRTF measurement setups for machine learning

Johan Pauwels (Queen Mary University of London); Lorenzo Picinali (Imperial College London)

6362 (AASP-P1.6): Neural Fourier Shift for Binaural Speech Rendering

Jin Woo Lee (Seoul National University); Kyogu Lee (Seoul National University)

1620 (AASP-P1.7): Global HRTF Interpolation via Learned Affine Transformation of Hyper-conditioned Features

Jin Woo Lee (Seoul National University); Sungho Lee (Seoul National University); Kyogu Lee (Seoul National University)

4790 (AASP-P1.8): HRTF Field: Unifying Measured HRTF Magnitude Representation with Neural Fields

You Zhang (University of Rochester); Yuxiang Wang (University of Rochester); Zhiyao Duan (University of Rochester)

5041 (AASP-P1.9): Learning to Personalize Equalization for High-Fidelity Spatial Audio Reproduction

Arjun Gupta (Meta); Pablo Hoffmann (Meta); Sebastian Prepelitã (Meta); Philip Robinson (Meta); Vamsi Krishna Ithapu (Meta); David Alon (Meta)

6719 (AASP-P1.10): A Data-Driven Approach to Audio Decorrelation (SPS Journal Paper)*

Carlotta Anemüller (AudioLabs Erlangen); Oliver Thiergart (International Audio Laboratories Erlangen); Emanuel Habets (AudioLabs Erlangen)

6777 (AASP-P1.11): Switching Independent Vector Analysis and Its Extension to Blind and Spatially Guided Convolutional Beamforming Algorithms (SPS Journal Paper)*

Tomohiro Nakatani (NTT Communication Science Laboratories); Rintaro Ikeshita (NTT); Keisuke Kinoshita (Google); Hiroshi Sawada (NTT); Naoyuki Kamo (NTT); Shoko Araki (NTT Corporation)

IVMSP-P10: Image/Video Caption Generation

Room: Dome K

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Jing Ma, Xuequan Lu

6029 (IVMSP-P10.1): End-to-End Non-Autoregressive Image Captioning

Hong Yu (Dalian University of Technology); Yuanqiu Liu (Dalian University of Technology); BaoKun Qi (Dalian University of Technology); Zhaolong Hu (Dalian University of Technology); Han Liu (Dalian University of Technology)

337 (IVMSP-P10.2): Enhancing Multimodal Alignment with Momentum Augmentation for Dense Video Captioning

yiwei wei (Tianjin university); Shaozu Yuan (JD AI); Meng Chen (JD AI); Longbiao Wang (Tianjin University)

Tuesday, June 6

450 (IVMSP-P10.3): I-Tuning: Tuning Frozen Language Models with Image for Lightweight Image Captioning

Ziyang Luo (Hong Kong Baptist University); Zhipeng Hu (NetEase Fuxi AI Lab); Yadong Xi (Fuxi AI Lab, Netease Inc.); Rongsheng Zhang (Fuxi AI Lab, Netease Inc.); Jing Ma (Hong Kong Baptist University)

972 (IVMSP-P10.4): VIDEO CAPTIONING VIA RELATION-AWARE GRAPH LEARNING

Yi Zheng (Fudan University); Heming Jing (Fudan University); Qiuji Xie (School of Computer Science, Fudan University); Yuejie Zhang (Fudan University); Rui Feng (Fudan University); Tao Zhang (Shanghai University of Finance and Economics); Shang Gao (Deakin University)

1192 (IVMSP-P10.5): Improving Image Captioning with Control Signal of Sentence Quality

Zhangzi Zhu (University of Electronic Science and Technology of China); shuai Wang (University of Electronic Science and Technology of China); Hong Qu (University of Electronic Science and Technology of China)

5827 (IVMSP-P10.6): Background Disturbance Mitigation for Video Captioning via Entity-Action Relocation

Zipeng Li (Wuhan University of Technology); Xian Zhong (Wuhan University of Technology); Shuqin Chen (Hubei University of Education); Wenxuan Liu (Wuhan University of Technology); Wenxin Huang (Hubei University); Lin Li (Wuhan University of Technology)

5304 (IVMSP-P10.7): MOTION-AWARE VIDEO PARAGRAPH CAPTIONING VIA EXPLORING OBJECT-CENTERED INTERNAL KNOWLEDGE

hu yimin (Fudan University); Guorui Yu (Fudan University); Yuejie Zhang (Fudan University); Rui Feng (Fudan University); Tao Zhang (Shanghai University of Finance and Economics); Xuequan Lu (Deakin University); Shang Gao (Deakin University)

2203 (IVMSP-P10.8): Associative Learning Network for Coherent Visual Storytelling

Xin Li (School of Computer Science & Technology, Soochow University); Chunping Liu (School of Computer Science and Technology, Soochow University); Yi Ji (School of Computer Science and Technology, Soochow University)

6772 (IVMSP-P10.9): Shot noise analysis for differential sampling in Indirect Time of Flight cameras (SPS Journal Paper)*

Nofre Sanmartin-Vich (Analog Devices Inc); Amina Achaibou (Analog devices); Javier Calpe (Analog Devices, Inc); Filiberto Pla (University Jaume I)

IVMSP-P11: Flow Estimation

Room: Dome L

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Franz Hlawatsch, Mahesan Niranjan

377 (IVMSP-P11.1): LEARNING SCENE FLOW FROM 3D POINT CLOUDS WITH CROSS-TRANSFORMER AND GLOBAL MOTION CUES

Mingliang Zhai (Nanjing University of Posts and Telecommunications); Kang Ni (Nanjing University of Posts and Telecommunications); Jiucheng Xie (Nanjing University of Posts and Telecommunications); Hao Gao (Nanjing University of Posts and Telecommunications)

379 (IVMSP-P11.0): SPIKE-BASED OPTICAL FLOW ESTIMATION VIA CONTRASTIVE LEARNING

Mingliang Zhai (Nanjing University of Posts and Telecommunications); Kang Ni (Nanjing University of Posts and Telecommunications); Jiucheng Xie (Nanjing University of Posts and Telecommunications); Hao Gao (Nanjing University of Posts and Telecommunications)

380 (IVMSP-P11.3): CROSS-MODAL OPTICAL FLOW ESTIMATION VIA MODALITY COMPENSATION AND ALIGNMENT

Mingliang Zhai (Nanjing University of Posts and Telecommunications); Kang Ni (Nanjing University of Posts and Telecommunications); Jiucheng Xie (Nanjing University of Posts and Telecommunications); Hao Gao (Nanjing University of Posts and Telecommunications)

577 (IVMSP-P11.4): Deformable cross attention for learning optical flow

Rokia Mohsen Abdein (Harbin Engineering University); Xuezhi Xiang (Harbin Engineering University); Ning Lv (Harbin Engineering University); Abdulmotaleb El Saddik (University of Ottawa)

3366 (IVMSP-P11.5): Bayesian Methods for Optical Flow Estimation Using a Variational Approximation, With Applications to Ultrasound

Jan Dorazil (TU Wien); Bernard H. Fleury (TU Wien); Franz Hlawatsch (TU Wien)

138 (IVMSP-P11.6): Cross-Modality Depth Estimation via Unsupervised Stereo RGB-to-Infrared Translation

Shi Tang (Tsinghua University); Xinchen Ye (Dalian University of Technology); Fei Xue (Dalian University of Technology); Rui Xu (Dalian University of Technology)

273 (IVMSP-P11.7): DEHRFormer: Real-time Transformer for Depth Estimation and Haze Removal from Varicolored Haze Scenes

Sixiang Chen (Jimei University); Tian Ye (Jimei University); Shi Jun (Xinjiang University); Yun Liu (Southwest University); JingXia Jiang (jimei university); Erkang Chen (Jimei University); Peng Chen (Jimei University)

Tuesday, June 6

914 (IVMSP-P11.8): Depth Estimation for a Single Omnidirectional Image with Reversed-gradient Warming-up Thresholds Discriminator

Yihong Wu (University of Southampton); Yuwen Heng (University of Southampton); Mahesan Niranjan (University of Southampton); Hansung Kim (University Of Southampton)

3327 (IVMSP-P11.9): Efficiently fusing sparse LiDAR for enhanced Self-supervised Monocular Depth Estimation

Yue Wang (University College London); Mingrong Gong (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences); Lei Xia (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences); Qieshi Zhang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences); Jun Cheng (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences)

3432 (IVMSP-P11.10): GLOBAL MATCHING-OPTIMIZATION NETWORK FOR STEREO DEPTH ESTIMATION

Yidi Zhang (Tsinghua University); Wenqi Huang (China southern power grid); Wenming Yang (Tsinghua University)

6147 (IVMSP-P11.11): ADAPTIVE SEMANTIC FUSION FRAMEWORK FOR UNSUPERVISED MONOCULAR DEPTH ESTIMATION

Ruoqi Li (University of Electronic Science and Technology of China); huimin yu (uestc); du kaiyang (uestc); Zhuoling Xiao (University of Electronic Science and Technology of China); Bo Yan (University of Electronic Science and Technology of China); zhengxi yuan (university of electronic science and technology of china)

IVMSP-P9: Image/Video Retrieval

Room: Dome J

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Bingbing Ni, Pan Zhou

375 (IVMSP-P9.1): Tracking Objects and Activities with Attention for Temporal Sentence Grounding

Zeyu Xiong (Huazhong University of Science and Technology); Daizong Liu (Peking University); Pan Zhou (Huazhong University of Science and Technology); Jiahao Zhu (Huazhong University of Science and Technology)

4291 (IVMSP-P9.2): Ontology-Aware Network for Zero-Shot Sketch-based Image Retrieval

Haoxiang Zhang (School of Information and Control, China University of Mining and Technology); He Jiang (School of Information and Control, China University of Mining and Technology); Ziqiang Wang (School of Information and Control, China University of Mining and Technology); Deqiang Cheng (School of Information and Control, China University of Mining and Technology)

382 (IVMSP-P9.3): MovieNet-PS: A Large-Scale Person Search Dataset in the Wild

Jie Qin (Nanjing University of Aeronautics and Astronautics); Peng Zheng (NUAA, MBZUAI, Aalto University); Yichao Yan (Shanghai Jiao Tong University); Rong Quan (Nanjing University of Aeronautics and Astronautics); Xiaogang CHENG (Nanjing University of Posts and Telecommunications); Bingbing Ni (Shanghai Jiao Tong University)

620 (IVMSP-P9.4): Semantic-Preserving Augmentation For Robust Image-Text Retrieval

Sunwoo Kim (Seoul National University); Kyuhong Shim (Seoul National University); Luong Trung Nguyen (Seoul National University); Byonghyo Shim (Seoul National University)

2055 (IVMSP-P9.5): Trust Your Partner's Friends: Hierarchical Cross-modal Contrastive Pre-training for Video-Text Retrieval

Yuhan Xiang (Xiamen University); Kaijian Liu (SenseTime Group Limited); Shixiang Tang (The University of Sydney); Lei Bai (Shanghai AI Laboratory); Feng Zhu (University of Science and Technology of China); Rui Zhao (SenseTime Group Limited); Xianming Lin (Xiamen University)

3145 (IVMSP-P9.6): Sketch Less Face Image Retrieval: A New Challenge

Dawei Dai (Chongqing Key Laboratory of Computational Intelligence, College of Computer Science and Technology, Chongqing University of Posts and Telecommunications); YuTang Li (Chongqing Key Laboratory of Computational Intelligence, College of Computer Science and Technology, Chongqing University of Posts and Telecommunications); liang wang (Chongqing Key Laboratory of Computational Intelligence, College of Computer Science and Technology, Chongqing University of Posts and Telecommunications); shiyu fu (Chongqing Key Laboratory of Computational Intelligence, College of Computer Science and Technology, Chongqing University of Posts and Telecommunications); Shuyin Xia (Chongqing University of Posts and Telecommunications); Guoyin Wang (Chongqing Key Laboratory of Computational Intelligence; Chongqing University of Posts and Telecommunications)

4708 (IVMSP-P9.7): ENABLING LARGE-SCALE IMAGE SEARCH WITH CO-ATTENTION MECHANISM

Zechao Hu (University of York); Adrian Bors (University of York)

4719 (IVMSP-P9.8): FEW BUT INFORMATIVE LOCAL HASH CODE MATCHING FOR IMAGE RETRIEVAL

Zechao Hu (University of York); Adrian Bors (University of York)

1291 (IVMSP-P9.9): Towards Making a Trojan-horse Attack on Text-to-Image Retrieval

Fan Hu (Renmin University of China); Aozhu Chen (Renmin University of China); Xirong Li (Renmin University of China)

Tuesday, June 6

335 (IVMSP-P9.10): Jointly Visual- and Semantic-Aware Graph Memory Networks for Temporal Sentence Localization in Videos

Daizong Liu (Peking University); Pan Zhou (Huazhong University of Science and Technology)

803 (IVMSP-P9.11): Image Generation is MAY All You Need for VQA

Kyung Ho Kim (ActionPower); Junseo Lee (ActionPower); Jihwa Lee (ActionPower)

MLSP-P10: Transfer Learning II

Room: Dome G

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Alexandros Iosifidis,

1397 (MLSP-P10.1): Rethinking Implicit Neural Representations for Vision Learners

Yiran Song (Shanghai Jiao Tong University); Qianyu Zhou (Shanghai Jiao Tong University); Lizhuang Ma (Shanghai Jiao Tong University)

1401 (MLSP-P10.2): Efficient Multi-Scale Attention Module with Cross-Spatial Learning

Daliang Ouyang (AEROSPACE SCIENCE & INDUSTRY SHENZHEN(GROUP)CO.,LTD.); Su He (AEROSPACE SCIENCE & INDUSTRY SHENZHEN(GROUP)CO.,LTD.); Guozhong Zhang (AEROSPACE SCIENCE & INDUSTRY SHENZHEN(GROUP)CO.,LTD.); Mingzhu Luo (AEROSPACE SCIENCE & INDUSTRY SHENZHEN(GROUP)CO.,LTD.); Huaiyong Guo (AEROSPACE SCIENCE & INDUSTRY SHENZHEN(GROUP)CO.,LTD.); Jian Zhan (AEROSPACE SCIENCE & INDUSTRY SHENZHEN(GROUP)CO.,LTD.); Zhijie Huang (AEROSPACE SCIENCE & INDUSTRY SHENZHEN(GROUP)CO.,LTD.)

3394 (MLSP-P10.3): Physics-Informed Transfer Learning for Voltage Stability Margin Prediction

Manish K Singh (University of Minnesota); Konstantinos D. Polyzos (University of Minnesota); Panagiotis Traganitis (Michigan State University); Sairaj Dhople (University of Minnesota); Georgios B. Giannakis (University of Minnesota)

4407 (MLSP-P10.4): Weight-based Mask for Domain Adaptation

EunSeop Lee (POSTECH); Inhan Kim (POSTECH); Daijin Kim (Pohang University of Science and Technology)

944 (MLSP-P10.5): TransAdapt: A Transformative Framework for Online Test Time Adaptive Semantic Segmentation

Debasmit Das (Qualcomm AI Research); Shubhankar Borse (Qualcomm AI Research); Hyojin Park (Qualcomm AI Research); Kambiz Azarian (Qualcomm AI Research); Hong Cai (Qualcomm AI Research); Risheek Garrepalli (Qualcomm AI Research); Fatih Porikli (Qualcomm AI Research)

2054 (MLSP-P10.6): Reliable Cluster-based Framework for Open Set Domain Adaptation

Xiu Zheng (Nanjing University); Yuan Huang (Nanjing University); Jie Tang (Nanjing University)

1837 (MLSP-P10.7): Refined Pseudo Labeling for Source-free Domain Adaptive Object Detection

Siqi Zhang (Institute of Automation, Chinese Academy of Sciences); Lu Zhang (CASIA); Zhiyong Liu (State Key Lab of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences)

1002 (MLSP-P10.8): TEST-TIME TRAINING-FREE DOMAIN ADAPTATION

Yongxiang Feng (Huawei Technologies Co., Ltd); Weihua He (Tsinghua University); Kaichao You (Huawei Technologies Co., Ltd); Bing Liu (Peking University); Ziyang Zhang (HUAWEI TECHNOLOGIES CO.LTD); Yaoyuan Wang (Huawei Technologies Co., Ltd.); Minglei Li (Huawei Technologies Co., Ltd.); yihang lou (huawei); Jiawei Li (Huawei Technologies Co., Ltd.); Guoqi Li (Tsinghua University); Jianxing Liao (HUAWEI TECHNOLOGIES CO.LTD)

5348 (MLSP-P10.9): GENERALIZED INVARIANT MATCHING PROPERTY VIA LASSO

Kang Du (University of Utah); Yu Xiang (University of Utah)

2728 (MLSP-P10.10): Cross-Domain Object Classification via Successive Subspace Alignment

Kecheng Chen (City University of Hong Kong); Haoliang Li (CityU); Hong Yan (City University of Hong Kong)

3052 (MLSP-P10.11): DTR: DETECTING TEXT WITH TRANSFORMERS

Jing Yang (Hunan University); Zhiqiang You (Hunan University); Zhiwei Zhong (Hunan University); peng liu (Guangdong university of technology); Langqi Mei (npic); Shenguang Huang (Ningbo Port Information Communication Co., Ltd.)

4225 (MLSP-P10.12): D-Conformer: Deformable Sparse Transformer Augmented Convolution for Voxel-based 3D Object Detection

Xiao Zhao (Fudan University); Liuzhen Su (Fudan University); Xukun Zhang (Fudan University); Dingkang Yang (Fudan University); Mingyang Sun (Fudan University); Shunli Wang (Fudan University); Peng Zhai (Fudan university); Lihua Zhang (Fudan University)

MLSP-P11: Learning Theory and Algorithms II

Room: Dome H

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Weiyu Xu,

Tuesday, June 6

6297 (MLSP-P11.1): The Uniqueness Problem of Physical Law Learning

Philipp Scholl (Ludwig Maximilian University of Munich); Aras Bacho (Ludwig Maximilian University of Munich); Holger Boche (Technische Universität München); Gitta Kutyniok (Ludwig Maximilian University of Munich)

4763 (MLSP-P11.2): Asymptotically Optimal Nonparametric Classification Rules for Spike Train Data

Mirosław Pawlak (University of Manitoba); Mateusz Pabian (AGH UST); Dominik Rzepka (AGH University of Science and Technology)

3439 (MLSP-P11.3): Estimation of High-Dimensional Differential Graphs from Multi-Attribute Data

Jitendra K Tugnait (Auburn University)

2501 (MLSP-P11.4): Learnable frontends that do not learn: Quantifying sensitivity to filterbank initialisation

Mark Anderson (Trinity College Dublin); Tomi H. Kinnunen (University of Eastern Finland); Naomi Harte (Trinity College Dublin)

5885 (MLSP-P11.5): Differential Analysis for Networks Obeying Conservation Laws

Anirudh Rayas (Arizona State University); Rajasekhar Anguluri (Arizona State University); Jiajun Cheng (Arizona State University); Gautam Dasarathy (Arizona State University)

6041 (MLSP-P11.6): On the Fairness of Multitask Representation Learning

Yingcong Li (University of California, Riverside); Samet Oymak (University of California, Riverside)

2583 (MLSP-P11.7): AN ONLINE ALGORITHM FOR CHANCE CONSTRAINED RESOURCE ALLOCATION

Yuwei Chen (Cainiao Network); Zengde Deng (Cainiao Network); Yinzhi Zhou (Cainiao Network); Zaiyi Chen (Cainiao Network); yujie chen (Cainiao network); Haoyuan Hu (Cainiao Network)

6144 (MLSP-P11.8): Sequential Invariant Information Bottleneck

Yichen Zhang (Xi'an Jiaotong University, China); Shujian Yu (Vrije Universiteit Amsterdam); Badong Chen ("Xi'an Jiaotong University, China")

4321 (MLSP-P11.9): ICEL: Learning with Inconsistent Explanations

Biao Liu (Southern University of Science and Technology); xiaoyu wu (Huawei); Bo Yuan (Southern University of Science and Technology)

3508 (MLSP-P11.10): Optimal Compression for Minimizing Classification Error Probability: an Information-Theoretic Approach

Jingchao Gao (the University of Iowa); Ao Tang (Cornell University); Weiyu Xu (University of Iowa)

1679 (MLSP-P11.11): PROVABLE COMPUTATIONAL AND STATISTICAL GUARANTEES FOR EFFICIENT LEARNING OF CONTINUOUS-ACTION GRAPHICAL GAMES

Adarsh Barik (Purdue University); Jean Honorio (Purdue)

6417 (MLSP-P11.12): Asymptotic Distribution of Stochastic Mirror Descent Iterates in Average Ensemble Models

Taylan Kargin (California Institute of Technology); Fariborz Salehi (California Institute of Technology); Babak Hassibi (Caltech)

MLSP-P12: Distributed and Federated Learning II

Room: Dome I

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Tommy S. Alstrøm,

5168 (MLSP-P12.1): M22: RATE-DISTORTION INSPIRED GRADIENT COMPRESSION

Yangyi Liu (McMaster University); Sadaf Dr Salehkalaibar (McMaster university); stefano rini (nycu); Jun Chen (McMaster University)

627 (MLSP-P12.2): Scalable and Secure Federated XGBoost

Quang M Nguyen (Massachusetts Institute of Technology); Nhan Khanh Le (TUM); Lam M Nguyen (IBM Research, Thomas J. Watson Research Center)

1977 (MLSP-P12.3): Batch Normalization damages Federated Learning on Non-IID data: Analysis and Remedy

Yanmeng Wang (The Chinese University of Hong Kong, Shenzhen); Qingjiang Shi (Tongji University); Tsung-Hui Chang ("The Chinese University of Hong Kong.")

4997 (MLSP-P12.4): Fully Distributed Federated Learning with Efficient Local Cooperations

Evangelos Georgatos (Computer Engineering and Informatics Dept., University of Patras); Christos Mavrokefalidis (Computer Engineering and Informatics Dept., University of Patras, Greece); Kostas Berberidis (University of Patras)

5118 (MLSP-P12.5): Efficient personalized federated learning on selective model training

Yeting Guo (College of Computer, National University of Defense Technology); Liu Fang (Hunan University); Tongqing Zhou (National University of Defense Technology); Zhiping Cai (NUDT); Nong Xiao (N)

Tuesday, June 6

6278 (MLSP-P12.6): DPP-based Client Selection for Federated Learning with Non-IID Data

Yuxuan Zhang (Northwest A&F University); chao xu (Northwest A&F University); Howard H. Yang (ZJU-UIUC Institute); Xijun Wang (Sun Yat-sen University); Tony Quek (Singapore University of Technology and Design)

2070 (MLSP-P12.7): Multilayer Subspace Learning with Self-sparse Robustness for Two-dimensional Feature Extraction

Han Zhang (Xidian University); Maoguo Gong (Xidian University); Feiping Nie (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

1658 (MLSP-P12.8): Channel-driven decentralized Bayesian federated learning for trustworthy decision making in D2D networks

Luca Barbieri (Politecnico di Milano); Osvaldo Simeone (King's College London); Monica Nicoli (Politecnico di Milano University)

1662 (MLSP-P12.9): Cross-device Federated Learning for Mobile Health Diagnostics: A First Study on COVID-19 Detection

Tong Xia (University of Cambridge); Jing Han (); Abhirup Ghosh (University of Cambridge); Cecilia Mascolo (University of Cambridge)

2398 (MLSP-P12.10): A BANDIT ONLINE CONVEX OPTIMIZATION APPROACH TO DISTRIBUTED ENERGY MANAGEMENT IN NETWORKED SYSTEMS

Ioannis Tsetis (University of Tübingen); Xiaotong Cheng (University Tübingen); Setareh Maghsudi (University of Tübingen)

6747 (MLSP-P12.11): Robust Aggregation for Federated Learning (SPS Journal Paper)*

Krishna Pillutla (University of Washington); Sham Kakade (Harvard University); Zaid Harchaoui (University of Washington)

6430 (MLSP-P12.12): FedSD: A New Federated Learning Structure Used in Non-iid Data

Minmin Yi (Tsinghua University); Houchun Ning (Tsinghua University); Peng Liu (PingAn Tech/Hong Kong Polytechnic University)

MLSP-P9: Machine Learning for Telecommunications

Room: Dome F

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Chang D. Yoo,

4142 (MLSP-P9.1): Neural Source Coding for bandwidth-efficient brain-computer interfacing with wireless neuro-sensor networks

Thomas Strypsteen (KU Leuven); Alexander Bertrand (KU Leuven)

5019 (MLSP-P9.2): Jamming Source Localization Using Augmented Physics-based Model

Andrea Nardin (Politecnico di Torino); Tales Imbiriba (Northeastern University); Pau Closas (Northeastern University)

5072 (MLSP-P9.3): Towards a Robust and Efficient Classifier for Real World Radio Signal Modulation Classification

Dancheng Liu (University of California, San Diego); Kazim Ergun (University of California San Diego); Tajana S Rosing (University of California, San Diego)

2848 (MLSP-P9.4): Regularized Deep Generative Model Learning for Real-time Massive MIMO Channel Tracking

Lixiang Lian (ShanghaiTech University); Ben Wang (ShanghaiTech University)

4047 (MLSP-P9.5): SIGNAL RECONSTRUCTION FOR FMCW RADAR INTERFERENCE MITIGATION USING DEEP UNFOLDING

Jeroen Overvest (NXP Semiconductors, Technical University of Eindhoven); Arie G.C. Koppelaar (NXP Semiconductors); Marco J.G. Bekooij (NXP Semiconductors); Jihwan Youn (Technical University of Eindhoven); Ruud J. G. van Sloun (Technical university of Eindhoven)

599 (MLSP-P9.6): JOINT HUMAN ORIENTATION-ACTIVITY RECOGNITION USING WIFI SIGNALS FOR HUMAN-MACHINE INTERACTION

Hojjat Salehinejad (Mayo Clinic); Navid Hasanzadeh (University of Toronto); Radomir Djogo (University of Toronto); Shahrokh Valaee (University of Toronto)

999 (MLSP-P9.7): High-resolution neural network processing of LFM radar pulses

Jabran Akhtar (FFI)

2230 (MLSP-P9.8): Deep Root MUSIC Algorithm for Data-Driven DoA Estimation

Dor Haim Shmuel (Ben-Gurion University of the Negev); Julian P. Merkofer (TU Eindhoven); Guy Revach (ETH Zürich); Ruud J. G. van Sloun (Technical university of Eindhoven); Nir Shlezinger (Ben-Gurion University)

2275 (MLSP-P9.9): Removing Radio Frequency Interference from Auroral Kilometric Radiation with Stacked Autoencoders

Allen Chang (University of Southern California); Mary Knapp (Massachusetts Institute of Technology Haystack Observatory); James LaBelle (Dartmouth College); John Swoboda (Massachusetts Institute of Technology Haystack Observatory); Ryan Volz (Massachusetts Institute of Technology Haystack Observatory); Philip Erickson (Massachusetts Institute of Technology Haystack Observatory)

Tuesday, June 6

2794 (MLSP-P9.10): Strategies for Enhanced Signal Modulation Classifications Under Unknown Symbol Rates and Noise Conditions

Ruixuan Wang (Villanova University); Yue Qi (villanova university); Mojtaba Vaezi (Villanova University); Xun Jiao (Villanova University); Moeness Amin (Villanova University)

4276 (MLSP-P9.11): On Neural Architectures for Deep Learning-based Source Separation of Co-Channel OFDM Signals

Gary CF Lee (MIT); Amir Weiss (Massachusetts Institute of Technology); Alejandro Lancho (MIT); Yury Polyanskiy (MIT); Gregory W Wornell (MIT)

SLT-P10: Dialog and Multimodal Processing of Language

Room: Dome C

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Samuel Thomas, Gerasimos Potamianos

1859 (SLT-P10.1): Identifying Entrainment in Task-oriented Conversations

Run Chen (Columbia University); Seokhwan Kim (Amazon Alexa AI); Alexandros Papangelis (Amazon Alexa AI); Julia Hirschberg (Columbia University); Yang Liu (Amazon, Alexa AI); Dilek Z Hakkani-Tur (Amazon Alexa AI)

1943 (SLT-P10.2): Empathetic Response Generation via Emotion Cause Transition Graph

Yushan Qian (Tianjin University); Bo Wang (Tianjin University); Ting-En Lin (Alibaba Group); Yinhe Zheng (Lingxin AI); Ying Zhu (Tianjin University); Dongming Zhao (China Mobile Communication Group Tianjin Co., Ltd); Yuexian Hou (Tianjin University); Yuchuan Wu (Alibaba); Yongbin Li (Alibaba Group)

2108 (SLT-P10.3): PAGE: A POSITION-AWARE GRAPH-BASED FRAMEWORK FOR EMOTION CAUSE ENTAILMENT IN CONVERSATION

Xiaojie Gu (Hangzhou City University); Renze Lou (Pennsylvania State University); Lin Sun (Hangzhou City University); Shangxin Li (Hangzhou City University)

2201 (SLT-P10.4): Towards Zero-Shot Personalized Table-to-Text Generation with Contrastive Persona Distillation

Haolan Zhan (Monash University); Shaobo Cui (Tsinghua University); Xuming Lin (Alibaba Group); Zhongzhou Zhao (Alibaba Group); Wei Zhou (Alibaba Group); Haiqing Chen (Alibaba Inc.)

2400 (SLT-P10.5): Query-Utterance Attention with Joint modeling for Query-Focused Meeting Summarization

Xingxian Liu (Beijing University of Posts and Telecommunications); Bin Duan (Beijing University of Posts and Telecommunications); Bo Xiao (Beijing University of Posts and Telecommunications); Yajing Xu (Beijing University of Posts and Telecommunications)

2788 (SLT-P10.6): CLICKER: Attention-Based Cross-Lingual Commonsense Knowledge Transfer

Ruolin Su (Georgia Institute of Technology); Zhongkai Sun (Amazon Alexa AI); Sixing Lu (Amazon); chengyuan ma (amazon); Chenlei Guo (Amazon)

3695 (SLT-P10.7): A Topic-Enhanced Approach for Emotion Distribution Forecasting in Conversations

Xin Lu (Harbin Institute of Technology); Weixiang Zhao (Harbin Institute of Technology); Yanyan Zhao (Harbin Institute of Technology); Bing Qin (Harbin Institute of Technology); Zhentao Zhang (CMB NT); wen junjie (China Merchants Bank)

4016 (SLT-P10.8): Adapted Multimodal BERT with Layer-wise Fusion for Sentiment Analysis

Odysseas S Chlapanis (National Technical University of Athens); Georgios Paraskevopoulos (National Technical University of Athens); Alexandros Potamianos (National Technical University of Athens)

4508 (SLT-P10.9): Improving Disfluency Detection with Multi-scale Self Attention and Contrastive Learning

Peiyang Wang (JD AI); Chaoqun Duan (JD AI Research); Meng Chen (JD AI); Xiaodong He (JDT)

4711 (SLT-P10.10): Dialog act guided contextual adapter for personalized speech recognition

Feng-Ju Chang (Amazon); Thejaswi Muniyappa (Amazon); Kanthashree Mysore Sathyendra (Amazon); Kai Wei (Amazon); Grant Strimel (Amazon); Ross McGowan (Amazon)

4975 (SLT-P10.11): OUTSIDE KNOWLEDGE VISUAL QUESTION ANSWERING VERSION 2.0

Benjamin Reichman (Georgia Institute of Technology); Anirudh S Sundar (Georgia Institute of Technology); Christopher G Richardson (Georgia Institute of Technology); Tamara Zubatiy (Georgia Institute of Technology); Prithwijiit Chowdhury (Georgia Institute of Technology); Aaryan Shah (Georgia Institute of Technology); Jack Truxal (Georgia Institute of Technology); Micah Grimes (Georgia Institute of Technology); Dristi Shah (Georgia Institute of Technology); Woo Ju Chee (Georgia Institute of Technology); Saif Punjwani (Georgia Institute of Technology); Atishay Jain (Georgia Institute of Technology); Larry Heck (Georgia Institute of Technology)

5238 (SLT-P10.12): SERI: SkEtching-Reasoning-Integrating Progressive Workflow for Empathetic Response Generation

Guanqun Bi (Institute of Information Engineering, Chinese Academy of Sciences; School of Cyber Security, University of Chinese Academy of Sciences); Yanan Cao (Institute of Information Engineering, Chinese Academy of Sciences); Piji Li (Nanjing University of Aeronautics and Astronautics); Yuqiang Xie (Institute of Information Engineering, Chinese Academy of Sciences); Fang Fang (Institute of Information Engineering, Chinese Academy of Sciences); Zheng Lin (iie)

Tuesday, June 6

SLT-P11: Discourse and Dialog

Room: Dome D

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Chiori Hori, Jan Trmal

821 (SLT-P11.1): Coarse-to-Fine Knowledge Selection for Document Grounded Dialogs

Yeqin Zhang (Nanjing University); Haomin Fu (Nanjing University); Cheng Fu (Alibaba); Haiyang Yu (Alibaba); Yongbin Li (Alibaba Group); Cam-Tu Nguyen (Nanjing University)

2181 (SLT-P11.2): A Slot-shared Span Prediction-based Neural Network for Multi-Domain Dialogue State Tracking

Abibulla Atawulla (University of Chinese Academy of Sciences); Xi Zhou (Xinjiang Technical Institute of Physics & Chemistry, Chinese Academy of Sciences); Yating Yang (Xinjiang Technical Institute of Physics & Chemistry, Chinese Academy of Sciences); Bo Ma (Xinjiang Technical Institute of Physics & Chemistry, Chinese Academy of Sciences); Fengyi Yang (University of Chinese Academy of Sciences)

3648 (SLT-P11.3): {NASE: A Chinese Benchmark for Evaluating Robustness of Spoken Language Understanding Models in Slot Filling

Meizheng Peng (Wuhan University); Xu Jia (Wuhan University); Min Peng (Wuhan University)

3933 (SLT-P11.4): An Asynchronous Updating Reinforcement Learning Framework for Task-oriented Dialog System

Sai Zhang (Beijing University of Posts and Telecommunications); Yuwei Hu (Beijing University of Posts and Telecommunications); Xiaojie Wang (Beijing University of Posts and Telecommunications); Caixia Yuan (Beijing University of Posts and Telecommunications)

4097 (SLT-P11.5): DialogMI: A Dialogue Model Based on Enhancing Dialogue Mutual Information

Yibo Zhang (Beijing University of Posts and Telecommunications); Ping Gong (Beijing University of Posts and Telecommunications); Zelin Wang (Beijing University of Posts and Telecommunications); Zhe Li (Beijing University of Posts and Telecommunications); Xuanyuan Yang (Beijing University of Posts and Telecommunications)

4281 (SLT-P11.6): LABEL-GUIDED CONTRASTIVE LEARNING FOR OUT-OF-DOMAIN DETECTION

Shun Zhang (Beihang University); Tongliang Li (Beihang University); Jiaqi Bai (Beihang University); Zhoujun Li (Beihang University)

4548 (SLT-P11.7): Improving Retrieval-based Dialogue System via Syntax-Informed Attention

Tengtao Song (Peking University); Nuo Chen (Peking University); Ji Jiang (Peking University); Zhihong Zhu (Peking University); Yuexian Zou (Peking University)

4646 (SLT-P11.8): Modeling Global Latent Semantic in Multi-Turn Conversations with Random Context Reconstruction

Chengwen Zhang (Beijing University of Posts & Telecommunications); Danqin Wu (Beijing University of Posts & Telecommunications)

5331 (SLT-P11.9): MODELING TURN-TAKING IN HUMAN-TO-HUMAN SPOKEN DIALOGUE DATASETS USING SELF-SUPERVISED FEATURES

Edmilson da Silva Morais (IBM Research Brazil); Matheus Damasceno (IBM Research); Hagai Aronowitz (IBM Research - AI); Aharon Satt (IBM Research); Ron Hoory (IBM Research)

5451 (SLT-P11.10): Choice Fusion as Knowledge for Zero-Shot Dialogue State Tracking

Ruolin Su (Georgia Institute of Technology); Jingfeng Yang (Amazon); Ting-Wei Wu (Georgia Institute of Technology); Biing-Hwang Juang (Georgia Institute of Technology)

5806 (SLT-P11.11): Think before you speak: Concept-guided Explicit Persona Reasoning for Personalized Dialogue Generation

Yunpeng Li (Institute of Information Engineering, Chinese Academy of Sciences); Yue Hu (Institute of Information Engineering, Chinese Academy of Sciences); Wei Peng (Institute of Information Engineering, Chinese Academy of Sciences); Yuqiang Xie (Institute of Information Engineering, Chinese Academy of Sciences)

5921 (SLT-P11.12): Efficient Uncertainty Estimation with Gaussian Process for Reliable Dialog Response Retrieval

Tong Ye (Ping An Technology (Shenzhen) Co., Ltd.; University of Science and Technology of China); Zhitao Li (Ping An Technology (Shenzhen) Co., Ltd.); Jianzong Wang (Ping An Technology (Shenzhen) Co., Ltd.); Ning Cheng (Ping An Technology (Shenzhen) Co., Ltd.); Jing Xiao (Ping An Insurance (Group) Company of China)

Tuesday, June 6

SLT-P12: Emerging Topics in Speech Synthesis

Room: Dome E

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Berrak Sisman,

975 (SLT-P12.1): Acoustically-Driven Phoneme Removal That Preserves Vocal Affect Cues

Camille Noufi (Stanford University); Jonathan Berger (Stanford University); Karen Parker (Stanford University); Daniel L Bowling (Stanford University)

1484 (SLT-P12.2): Mid-attribute Speaker Generation using Optimal-Transport-based Interpolation of Gaussian Mixture Models

Aya Watanabe (The University of Tokyo); Shinnosuke Takamichi (The University of Tokyo); Yuki Saito ("The University of Tokyo, Japan"); Detai Xin (The University of Tokyo); Hiroshi Saruwatari (The University of Tokyo)

1487 (SLT-P12.3): M2-CTTS: End-to-End Multi-scale Multi-modal Conversational Text-to-Speech Synthesis

Jinlong Xue (Beijing University of Posts and Telecommunications); Yayue Deng (Beijing University of Posts and Telecommunications); Fengping Wang (Beijing University of Posts and Telecommunications); Ya Li (Beijing University of Posts and Telecommunications); Yingming Gao (Beijing University of Posts and Telecommunications); Jianhua Tao ("National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences"); Jianqing Sun (Unisound AI Technology Co.,Ltd); Jiaen Liang (Unisound)

1593 (SLT-P12.4): CrossSpeech: Speaker-independent Acoustic Representation for Cross-lingual Speech Synthesis

Ji-Hoon Kim (42dot); Hong-Sun Yang (42dot Inc); Yooncheol Ju (AIRS Company, Hyundai Motor Group, Seoul, Republic of Korea); ILHWAN KIM (42dot); Byeong-Yeol Kim (42dot)

1993 (SLT-P12.5): NEURAL SPEECH PHASE PREDICTION BASED ON PARALLEL ESTIMATION ARCHITECTURE AND ANTI-WRAPPING LOSSES

Yang Ai (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China)

2715 (SLT-P12.6): ZERO-SHOT PERSONALIZED LIP-TO-SPEECH SYNTHESIS WITH FACE IMAGE BASED VOICE CONTROL

Zheng-Yan Sheng (University of Science and Technology of China); Yang Ai (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China)

2721 (SLT-P12.7): Speech reconstruction from silent tongue and lip articulation by pseudo target generation and domain adversarial training

Rui-Chen Zheng (University of Science and Technology of China); Yang Ai (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China)

3239 (SLT-P12.8): Multi-speaker Speech Synthesis from Electromyographic Signals by Soft Speech Unit Prediction

Kevin Scheck (University of Bremen); Tanja Schultz (University of Bremen)

3286 (SLT-P12.9): An End-to-End Neural Network for Image-to-Audio Transformation

Chen Liu (Oregon Health & Science University); Michael Deisher (Intel Corporation); Munir Georges (Intel Corporation); Munir Georges (THI)

4058 (SLT-P12.10): Duration-aware pause insertion using pre-trained language model for multi-speaker text-to-speech

Dong Yang (The University of Tokyo); Tomoki Koriyama (CyberAgent, Inc.); Yuki Saito ("The University of Tokyo, Japan"); Takaaki Saeki (The University of Tokyo); Detai Xin (The University of Tokyo); Hiroshi Saruwatari (The University of Tokyo)

4378 (SLT-P12.11): Low-latency electrolaryngeal speech enhancement based on FastSpeech2-based voice conversion and self-supervised speech representation

Kazuhiro Kobayashi (Nagoya University); Tomoki Hayashi (Nagoya University); Tomoki Toda (Nagoya University)

6371 (SLT-P12.12): Towards Building Text-To-Speech Systems for the Next Billion Users

Gokul Karthik Kumar (Mohamed Bin Zayed University of Artificial Intelligence); Praveen S V (Indian Institute of Technology Madras); Pratyush Kumar (Indian Institute of Technology Madras); Mitesh M. Khapra (Indian Institute of Technology Madras); Karthik Nandakumar (Mohamed Bin Zayed University of Artificial Intelligence)

SLT-P9: Audio and Text Segmentation, Tagging and Parsing

Room: Dome B

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Hung-yi Lee, Atsunori Ogawa

901 (SLT-P9.1): E2E Segmentation in a Two-Pass Cascaded Encoder ASR Model

W. Ronny Huang (Google); Shuo-yiin Chang (Google); Tara Sainath (Google); Yanzhang He (Google); David Rybach (Google); Robert David (Google); Rohit Prabhavalkar (Google); Cyril Allauzen (Google); Charles C Peysner (Google Inc.); Trevor Strohman (Google)

Tuesday, June 6

1687 (SLT-P9.2): Analyzing Acoustic Word Embeddings from Pre-trained Self-supervised Models

Ramon R Sanabria (*The University Of Edinburgh*); Hao Tang (*The University of Edinburgh*); Sharon Goldwater (*University of Edinburgh*)

1916 (SLT-P9.3): Unsupervised Word Segmentation Using Temporal Gradient Pseudo-Labels

Tzeviya S Fuchs (*Bar-Ilan University*); Yedid Hoshen (*The Hebrew University of Jerusalem*)

2243 (SLT-P9.4): Cascading and Direct Approaches to Unsupervised Constituency Parsing on Spoken Sentences

Yuan Tseng (*National Taiwan University*); Cheng-I Lai (*MIT*); Hung-yi Lee (*National Taiwan University*)

2267 (SLT-P9.5): Towards trustworthy phoneme boundary detection with autoregressive model and improved evaluation metric

Hyeongju Kim (*Supertone, Inc.*); Hyeong-Seok Choi (*Seoul National University*)

2977 (SLT-P9.6): Integrating Syntactic and Semantic Knowledge in AMR Parsing with Heterogeneous Graph Attention Network

Yikemaiti Sataer (*Southeast University*); Chuanqi Shi (*Southeast University*); Miao Gao (*Southeast University*); Yunlong Fan (*Southeast University*); Bin Li (*Southeast University*); Zhiqiang Gao (*Southeast University*)

3872 (SLT-P9.7): Lexicon-injected Semantic Parsing for Task-Oriented Dialog

Xiaojun Meng (*Noah's Ark Lab, Huawei Technologies*); Wenlin Dai (*Tsinghua University*); Yasheng Wang (*NoahArk Lab, Huawei*); Baojun wang (*Noah's Ark Lab of Huawei*); Zhiyong Wu (*Tsinghua University*); Xin Jiang (*Huawei Noah's Ark Lab*); Qun Liu (*Huawei Noah's Ark Lab*)

4451 (SLT-P9.8): A Prototypical Semantic Decoupling Method via Joint Contrastive Learning for Few-Shot Named Entity Recognition

Guanting Dong (*Beijing University of Posts and Telecommunications*); Zechen Wang (*Beijing University of Posts and Telecommunications*); Liwen Wang (*Beijing University of Posts and Telecommunications*); Daichi Guo (*Beijing University of Posts and Telecommunications*); Dayuan Fu (*Beijing University of Posts and Telecommunications*); yuxiang wu (*Beijing University of Posts and Telecommunications*); Chen Zeng (*Beijing University of Posts and Telecommunications*); Xuefeng Li (*Beijing University of Posts and Telecommunications*); Tingfeng Hui (*Beijing University of Posts and Telecommunications*); Keqing He (*Beijing University of Posts and Telecommunications*); Xinyue Cui (*Beijing University of Posts and Telecommunications*); QiXiang Gao (*Beijing University of Posts and Telecommunications*); Weiran Xu (*Beijing University of Posts and Telecommunications*)

4452 (SLT-P9.9): Revisit Out-of-vocabulary Problem for Slot Filling: A Unified Contrastive Framework with Multi-level Data Augmentations

Daichi Guo (*Beijing University of Posts and Telecommunications*); Guanting Dong (*Beijing University of Posts and Telecommunications*); Dayuan Fu (*Beijing University of Posts and Telecommunications*); yuxiang wu (*Beijing University of Posts and Telecommunications*); Chen Zeng (*Beijing University of Posts and Telecommunications*); Tingfeng Hui (*Beijing University of Posts and Telecommunications*); Liwen Wang (*Beijing University of Posts and Telecommunications*); Xuefeng Li (*Beijing University of Posts and Telecommunications*); Zechen Wang (*Beijing University of Posts and Telecommunications*); Keqing He (*Beijing University of Posts and Telecommunications*); Xinyue Cui (*Beijing University of Posts and Telecommunications*); Weiran Xu (*Beijing University of Posts and Telecommunications*)

5249 (SLT-P9.10): ZEPHYR: ZERO-SHOT PUNCTUATION RESTORATION

Minghan Wang (*Huawei*); Yinglu Li (*HUAWEI TECHNOLOGIES CO., LTD.*); Jiaxin GUO (*Huawei*); Xiaosong Qiao (*Huawei*); Chang Su (*Huawei*); Min Zhang (*Huawei*); Shimin Tao (*Huawei*); Hao Yang (*Huawei*)

6022 (SLT-P9.11): ANCIENT CHINESE WORD SEGMENTATION AND PART-OF-SPEECH TAGGING USING DISTANT SUPERVISION

Shuo Feng (*Nanjing University of Aeronautics and Astronautics*); Piji Li (*Nanjing University of Aeronautics and Astronautics*)

Wednesday, June 7

AASP-L4: Diffusion-Based Generative Models for Audio and Speech

Room: Salon de Roses A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Yuki Mitsufuji, Francois Germain

08:15 AM

5245 (AASP-L4.1): Cold Diffusion for Speech Enhancement

Hao Yen (Georgia Institute of Technology); François G Germain (Mitsubishi Electric Research Laboratories (MERL)); Gordon Wichern (Mitsubishi Electric Research Laboratories (MERL)); Jonathan LeRoux (Mitsubishi Electric Research Laboratories (MERL))

08:30 AM

5709 (AASP-L4.2): Analysing Diffusion-based Generative Approaches versus Discriminative Approaches for Speech Restoration

Jean-Marie Lemercier (Universität Hamburg); Julius Richter (Universität Hamburg); Simon Welker (Universität Hamburg); Timo Gerkmann (Universität Hamburg)

08:45 AM

2264 (AASP-L4.3): Unsupervised vocal dereverberation with diffusion-based generative models

Koichi Saito (Sony Group Corporation); Naoki Murata (Sony Group Corporation); Toshimitsu Uesaka (Sony Group Corporation); Chieh-Hsin Lai (Sony Group Corporation); Yuhta Takida (Sony Group Corporation); Takao Fukui (Sony Group Corporation); Yuki Mitsufuji (Sony Group Corporation)

09:00 AM

5637 (AASP-L4.4): Solving audio inverse problems with a diffusion model

Eloi Moliner (Aalto University); Jaakko Lehtinen (NVIDIA & Aalto University); Vesa Valimaki (Aalto University)

09:15 AM

5778 (AASP-L4.5): DiffPhase: Generative Diffusion-based STFT Phase Retrieval

Tal Peer (Universität Hamburg); Simon Welker (Universität Hamburg); Timo Gerkmann (Universität Hamburg)

09:30 AM

3196 (AASP-L4.6): Optimal Transport in Diffusion Modeling for Conversion Tasks in Audio Domain

Vadim Popov (Huawei Noah's Ark Lab); Amantur Amatov (Huawei); Mikhail Kudinov (Huawei Noah's Ark Lab); Vladimir Gogoryan (Huawei Noah's Ark Lab); Tasnima Sadekova (Huawei Noah's Ark Lab); Ivan Vovk (Huawei Noah's Ark Lab)

GC-4: Multilingual Alzheimer's Dementia Recognition through Spontaneous Speech: a Signal Processing Grand Challenge

Room: Nefeli B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): TBA

08:15 AM

6634 (GC-L4.1): Introduction

Saturnino Luz (University of Edinburgh); Fasih Haider (The University of Edinburgh); Davida Fromm Fromm (Carnegie Mellon University); Brian MacWhinney (CMU)

08:35 AM

6866 (GC-L4.2): CONSEN: Complementary and Simultaneous Ensemble for Alzheimer's Disease Detection and MMSE Score Prediction

LONGBIN JIN (Konkuk University); Yealim Oh (Konkuk University); Hyunseo Kim (Konkuk University); Hyuntaek Jung (Konkuk University); Hyo Jin Jon (Konkuk University); Jung Eun Shin (Voinosis Inc.); Eun Yi Kim (Konkuk University)

08:47 AM

6893 (GC-L4.3): Exploring Language-Agnostic Speech Representations using Domain Knowledge for Detecting Alzheimer's Dementia

Zehra Shah (University of Alberta); Shi-ang Qi (University of Alberta); Fei Wang (University of Alberta); Mahtab Farrokh (University of Alberta); Mashrura Tasnim (University of Alberta); Eleni Stroulia (University of Alberta); Russell Greiner (U Alberta); Manos Plitsis (Athena Research Center); Athanasios Katsamanis ("ATHENA R.C., Behavioral Signal Technologies")

08:59 AM

6896 (GC-L4.4): The USTC System for ADReSS-M Challenge

Kangdi Mei (University of Science and Technology of China); Xinyun Ding (iFlytek Research); yinlong liu (USTC); Zhiqiang Guo (University of Science and Technology of China); Feiyang Xu (iFlytek Co.Ltd); Xin Li (University of Science and Technology of China); Tuya Naren (University of Science and Technology of China); Jiahong Yuan (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China)

Wednesday, June 7

09:11 AM

6898 (GC-L4.5): Cross-lingual Alzheimer's disease detection based on paralinguistic and pre-trained features

Chen Xuchu (Tsinghua University); Yu Pu (Tsinghua University); Jinpeng Li (Tsinghua University); Wei-Qiang Zhang (Tsinghua University)

09:23 AM

6930 (GC-L4.6): Cross-Lingual Transfer Learning for Alzheimer's Detection From Spontaneous Speech

Bastiaan Tamm (KU Leuven); Rik Vandenberghe (University of Leuven); Hugo Van hamme (KU Leuven)

IVMSP-L4: Image Segmentation

Room: Athina

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Liujuan Cao, Kai Hu

08:15 AM

1887 (IVMSP-L4.1): CANDY: CAtegory-kerNelized DYnamic Convolution for Instance Segmentation

Yao Lu (Xiamen University); Zhiyi Chen (XiaMen University); Zehui Chen (University of Science and Technology of China); Jie Hu (Xiamen University); Liujuan Cao (Xiamen University); ShengChuan Zhang (Xiamen University)

08:30 AM

2221 (IVMSP-L4.2): SPATIAL SIMILARITY GUIDANCE FOR FEW-SHOT SEGMENTATION

Xiaoliu Luo (Chongqing University); Zhao Duan (Chongqing University); Taiping Zhang (Chongqing University)

08:45 AM

2360 (IVMSP-L4.3): Learning Task-aligned Mask Query for Instance Segmentation

Bin Fu (School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School); Hongliang He (School of Electronic and Computer Engineering, Peking University); Pengxu Wei (Sun Yat-sen University); Jie Chen (Peking University)

09:00 AM

4590 (IVMSP-L4.4): IMAGE SEGMENTATION FOR IMPROVED LOSSLESS SCREEN CONTENT COMPRESSION

Shabhrish Reddy Reddy Uddehal (Coburg University); Tilo Strutz (Coburg University); Hannah Och (Friedrich-Alexander University Erlangen-Nürnberg); Andre Kaup (Friedrich-Alexander-Universität Erlangen-Nürnberg)

09:15 AM

2799 (IVMSP-L4.5): Knowledge Distillation with Active Exploration and Self-attention based Inter-Class Variation Transfer For Image Segmentation

Yifan Zhang (Shenzhen University); Shaojie Li (Shenzhen University); Xuan Yang (Shenzhen University)

09:30 AM

5735 (IVMSP-L4.6): Unsupervised Action Segmentation of Untrimmed Egocentric Videos

Sam Perochon (Ecole Normale Supérieure Paris-Saclay); Laurent Oudre (ENS Paris-Saclay)

MLSP-L4: Model Pruning and Compression

Room: Salon de Roses B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Sijia Liu,

08:15 AM

2195 (MLSP-L4.1): WHC: Weighted Hybrid Criterion for Filter Pruning on Convolutional Neural Networks

Shaowu Chen (Shenzhen University); Weize Sun (Shenzhen University); Lei Huang (Shenzhen University)

08:30 AM

818 (MLSP-L4.2): Filter Pruning via Filters Similarity in Consecutive Layers

Xiaorui Wang (Ping An Technology (Shenzhen) Co., Ltd.); Jun Wang (Ping An Technology (Shenzhen) Co. Ltd.); xin tang (Ping An property&casualty insurance company of China.LTD.); Peng Gao (Ping An Technology); Rui Fang (Ping An property&casualty insurance company of China.LTD.); Guotong Xie (Ping An Technology (Shenzhen) Co. Ltd.)

08:45 AM

3267 (MLSP-L4.3): Prune then Distill: Dataset Distillation with Importance Sampling

Anirudh S Sundar (Georgia Institute of Technology); Gokce Keskin (Amazon Inc.); Chander Chandak (Amazon Inc.); I-Fan Chen (Amazon Inc.); Pegah Ghahremani (Amazon Inc.); Shalini Ghosh (Amazon Alexa AI)

09:00 AM

5620 (MLSP-L4.4): A Probabilistic Framework for Pruning Transformers via a Finite Admixture of Keys

Tan Minh Nguyen (University of California, Los Angeles); Tam Minh Nguyen (FPT Software); Long Minh Bui (FPT Software); Hai Do (FPT Software); Duy Khuong Nguyen (FPT Software Ltd. - FPT Corporation); Dung D. D. Le (College of Engineering and Computer Science, VinUniversity); Hung Tran-The (Deakin University); Nhat Ho (University of Texas at Austin); Stanley Osher (UCLA); Richard Baraniuk (Rice University)

Wednesday, June 7

09:15 AM

5578 (MLSP-L4.5): A Contrastive Knowledge Transfer Framework for Model Compression and Transfer Learning
kaiqi zhao (Arizona State University); Yitao Chen (Arizona State University); Ming Zhao (Arizona State University)

09:30 AM

5081 (MLSP-L4.6): Online Model Compression for Federated Learning with Large Models
Tien-Ju Yang (Google); Yonghui Xiao (Google); Giovanni Motta (Google, Inc.); Françoise Beaufays (Google); Rajiv Mathews (Google); Mingqing Chen (Google Inc.)

SLT-L7: Machine Learning Methods for Language I

Room: Jupiter

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Wen Wang, Zhehuai Chen

08:15 AM

442 (SLT-L7.1): Improving the Out-Of-Distribution Generalization Capability of Language Models: Counterfactually-Augmented Data is not Enough

Caoyun Fan (Shanghai Jiao Tong University); Wenqing Chen (Sun Yat-sen University); Jidong Tian (Shanghai Jiao Tong University); Yitian Li (Shanghai Jiao Tong University); Hao He (Shanghai Jiao Tong University); Yaohui Jin (Shanghai Jiao Tong University)

08:30 AM

2168 (SLT-L7.2): Boosting BERT Subnets with Neural Grafting

Ting Hu (Hasso Plattner Institute); Christoph Meinel (Hasso Plattner Institute); Haojin Yang (Hasso-Plattner-Institut für Digital Engineering gGmbH)

08:45 AM

2255 (SLT-L7.3): TERMINOLOGY-AWARE MEDICAL DIALOGUE GENERATION

Chen Tang (University of Surrey); Hongbo Zhang (University of Sheffield); Tyler Loakman (The University of Sheffield); Chenghua Lin (University of Sheffield); Frank Guerin (University of Surrey)

09:00 AM

4657 (SLT-L7.4): Targeted Adversarial Attacks against Neural Machine Translation

Sahar Sadrizadeh (EPFL); AmirHossein Dabiri Aghdam (University of Tehran); Ljiljana Dolamic (armasuisse); Pascal Frossard (EPFL)

09:15 AM

5653 (SLT-L7.5): ACF: Aligned Contrastive Finetuning for Language and Vision Tasks

Wei Zhu (East China Normal University); Peng Wang (Northwestern Normal Univ); Xiaoling Wang (East China Normal University); Yuan Ni (Ping An Technology); Guotong Xie (Ping An Technology (Shenzhen) Co. Ltd.)

SLT-L8: Machine Translation and Dialog System

Room: Delphi

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Satoshi Nakamura, Tanja Schultz

08:15 AM

1647 (SLT-L8.1): Named Entity Detection and Injection for Direct Speech Translation

Marco Gaido (Fondazione Bruno Kessler); Yun Tang (Meta); Iliia Kulikov (Meta); Rongqing Huang (Meta); Hongyu Gong (Meta); Hirofumi Inaguma (Meta)

08:30 AM

1733 (SLT-L8.2): CONTRASTIVE LEARNING WITH DIALOGUE ATTRIBUTES FOR NEURAL DIALOGUE GENERATION

Jie Tan (The Chinese University of Hong Kong); Hengyi Cai (Baidu Inc.); Hongshen Chen (JD.com); Hong Cheng (Chinese University of Hong Kong); Helen Meng (The Chinese University of Hong Kong); Zhuoye Ding (JD.com)

08:45 AM

1875 (SLT-L8.3): Role of Bias Terms in Dot-Product Attention

Mahdi Namazifar (Amazon Alexa AI); Devamanyu Hazarika (Amazon Alexa AI); Dilek Z Hakkani-Tur (Amazon Alexa AI)

09:00 AM

3633 (SLT-L8.4): Generic Dependency Modeling for Multi-Party Conversation

Weizhou Shen (Sun Yat-sen University); Xiaojun Quan (Sun Yat-sen University); Ke Yang (Sun Yat-sen University)

09:15 AM

4760 (SLT-L8.5): TEXTLESS DIRECT SPEECH-TO-SPEECH TRANSLATION WITH DISCRETE SPEECH REPRESENTATION

Xinjian Li (Carnegie Mellon University); Ye Jia (Tomato AI); Chung-Cheng Chiu (Google)

Wednesday, June 7

09:30 AM

5882 (SLT-L8.6): LEARNING TO BALANCE THE GLOBAL COHERENCE AND INFORMATIVENESS IN KNOWLEDGE-GROUNDED DIALOGUE GENERATION

Chenxu Niu (Institute of Information Engineering, Chinese Academy of Sciences); Yue Hu (Institute of Information Engineering, Chinese Academy of Sciences); Wei Peng (Institute of Information Engineering, Chinese Academy of Sciences); Yuqiang Xie (Institute of Information Engineering, Chinese Academy of Sciences)

SS-L.26: Radar Waveform Design: Recent Advances and New Emerging Applications

Room: Nafsika A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Mohammad Alaei,

08:15 AM

1475 (SS-L26.1): Co-Design for MIMO radar and MIMO communication aided by reconfigurable intelligent surface

Da Li (National University of Defense Technology); Bo Tang (National University of Defense Technology); Lei Xue (National University of Defense Technology)

08:30 AM

2263 (SS-L26.2): Interpretable, Unrolled Deep Radar Beampattern Design

Xue Yao (Southeast University); Cui Guolong (UESTC); Xianxiang Yu (UESTC)

08:45 AM

3271 (SS-L26.3): Interpretable, Unrolled Deep Radar Beampattern Design

Kareem M Metwaly (The Pennsylvania State University); Junho Kweon (The Pennsylvania State University); Khaled Alhujaili (The Taibah University); Maria S. Greco (University of Pisa); Fulvio Gini (University of Pisa); Vishal Monga (The Pennsylvania State University)

09:00 AM

3358 (SS-L26.4): RIS-Aided Wideband DFRC with Reconfigurable Holographic Surface

Tong Wei (Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg); Linlong Wu (University of Luxembourg); Kumar Vijay Mishra (United States DEVCOM Army Research Laboratory); Bhavani Shankar Mysore Ramarao (University of Luxembourg)

09:15 AM

3548 (SS-L26.5): Joint Waveform and Passive Beamformer Design in Multi-IRS Aided Radar

Zahra Esmailbeig (University of Illinois at Chicago); Arian Eamaz (University of Illinois - Chicago, IL); Kumar Vijay Mishra (United States DEVCOM Army Research Laboratory); Mojtaba Soltanalian (University of Illinois)

09:30 AM

4515 (SS-L26.6): RESOURCE ALLOCATION FOR UAV-ENABLED INTEGRATED SENSING AND COMMUNICATION (ISAC) VIA MULTI-OBJECTIVE OPTIMIZATION

Omid Rezaei (Sharif University of Technology); Mohammad Mahdi Naghsh (Isfahan University of Technology); Seyed Mohammad Karbasi (Sharif University of Technology); Mohammad Mahdi Nayebi (Sharif University of Technology)

SS-L4: Conversational Healthcare Interfaces

Room: Nefeli A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Aki Harma,

08:15 AM

1644 (SS-L4.1): HEALTHCALL CORPUS AND TRANSFORMER EMBEDDINGS FROM HEALTHCARE CUSTOMER-AGENT CONVERSATIONS

Nikola Lackovic (Malakoff Humanis); Montacié Claude (Sorbonne Université); Cédric Lequilliec (Malakoff Humanis); marie-josé Caraty (Sorbonne Université)

08:30 AM

3195 (SS-L4.2): Forecasting of breathing events from speech for respiratory support

Aki Harma (Philips); Ulf Grosseckthofer (Philips Research); Okke Ouweltjes (Philips Research); Venkata Srikanth Nallanthighal (Philips Research)

08:45 AM

5137 (SS-L4.3): Navigating and Reaching Therapeutic Goals with Dynamical Systems in Conversation-based Interventions

Victor Ardulov (Amazon); Shrikanth Narayanan (USC)

Wednesday, June 7

09:00 AM

5359 (SS-L4.4): Exploiting prompt learning with pre-trained language models for Alzheimer's Disease detection

Yi Wang (The Chinese University of Hong Kong); Jiajun Deng (The Chinese University of Hong Kong); Tianzi Wang (The Chinese University of Hong Kong); Bo ZHENG (the Chinese University of Hong Kong); Shoukang Hu (Nanyang Technological University); Xunying Liu (The Chinese University of Hong Kong); Helen Meng (The Chinese University of Hong Kong)

09:15 AM

6377 (SS-L4.5): Egocentric Action Anticipation for Personal Health

Ivan Rodin (University of Catania); Antonino Furnari (University of Catania); Dimitrios Mavroeidis (Philips Research); Giovanni Maria Farinella (University of Catania, Italy)

09:30 AM

6428 (SS-L4.6): A Controllable Lifestyle Simulator for use in Deep Reinforcement Learning Algorithms

Libio Gonçalves Braz (UPSSITECH); Allmin Susaiyah (Philips)

SS-L5: Quantum Computing for Machine Learning and Signal Processing

Room: Nafsika B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Vassilis Kekatos

08:15 AM

184 (SS-L5.1): Learning Quantum Entanglement Distillation with Noisy Classical Communications

HARI HARA SUTHAN CHITTOOR (Kings College London); Osvaldo Simeone (King's College London)

08:30 AM

6391 (SS-L5.2): A Quantum Approach for Stochastic Constrained Binary Optimization

Sarthak Gupta (Virginia Tech); Vassilis Kekatos (Virginia Tech)

08:45 AM

5284 (SS-L5.3): DISTRIBUTED QUANTUM SENSING NETWORK WITH GEOGRAPHICALLY CONSTRAINED MEASUREMENT STRATEGIES

Yingkang Cao (University of Maryland-College Park); Xiaodi Wu (University of Maryland)

09:00 AM

5405 (SS-L5.4): Quantum Graph Transformers

Georgios Kollias (IBM Research); Vasileios Kalantzis (IBM Research); Theodoros Salonidis (IBM T.J. Watson Research Center); Shashanka Ubaru (IBM Research)

09:15 AM

6386 (SS-L5.5): FINER-GRAINED DECOMPOSITION FOR PARALLEL QUANTUM MIMO PROCESSING

Minsung Kim (Princeton University); Kyle Jamieson (Princeton University)

09:30 AM

4852 (SS-L5.6): The role of initial entanglement in adaptive Gibbs state preparation on quantum computers

Sophia Economou (Virginia Tech); Ada Warren (Virginia Tech); Edwin Barnes (Virginia Tech)

ASPS-P3: Computer Vision Applications

Room: Dome A

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Dong Hye Ye, Anthony Vetro

6551 (ASPS-P3.1): ON THE QUANTIZATION OF RECURRENT NEURAL NETWORKS FOR SMILES GENERATION

Adriano Duroao (IT / Dep. of Electrical and Computer Engineering, University of Coimbra, Portugal); Joel Arrais (CISUC, University of Coimbra); Bernardete Ribeiro (CISUC, University of Coimbra); Gabriel Falcao (IT, University of Coimbra, Portugal)

4821 (ASPS-P3.2): WIFI-BASED ROBUST CHILD PRESENCE DETECTION FOR SMART CARS

Sakila S Jayaweera (University of Maryland, College Park); Beibei Wang (Origin Wireless Inc.); Xiaolu Zeng (Beijing Institute of Technology); Wei-Hsiang Wang (University of Maryland, College Park); K. J. Ray Liu (Origin Wireless Inc.)

6365 (ASPS-P3.3): CAN2V: CAN-BUS DATA-BASED SEQ2SEQ MODEL FOR VEHICLE VELOCITY PREDICTION

Jae-Heung Cho (Hanyang University); Joon-Hyuk Chang (Hanyang University)

246 (ASPS-P3.4): An Evaluation Platform to Scope Performance of Synthetic Environments in Autonomous Ground Vehicles Simulation

Xiangyu Bai (Northeastern University); Jiang Le (Northeastern University); Yedi Luo (Northeastern University); Aniket Gupta (Northeastern University); Pushyami Kaveti (Northeastern University); Hanumant Singh (Northeastern University); Sarah Ostadabbas (Northeastern University)

Wednesday, June 7

3000 (ASPS-P3.5): PREFALLKD: PRE-IMPACT FALL DETECTION VIA CNN-VIT KNOWLEDGE DISTILLATION

Tin-Han Chi (Department of Biomedical Engineering, National Yang Ming Chiao Tung University); Kai-Chun Liu (Academia Sinica); Chia-Yeh Hsieh (Bachelor's Program in Medical Informatics and Innovative Applications, Fu Jen Catholic University); Yu Tsao (Academia Sinica); Chia-Tai Chan (Department of Biomedical Engineering, National Yang Ming Chiao Tung University)

3733 (ASPS-P3.6): Finding Optimal Numerical Format for Sub-8-bit Post-Training Quantization of Vision Transformers

Janghwan Lee (Hanyang University); Youngdeok Hwang (Baruch College - The City University of New York (CUNY)); Jungwook Choi (Hanyang University)

3961 (ASPS-P3.7): A Multi-Channel Aggregation Framework for Object Detection in Large-Scale SAR Image

Chule Yang (Defense Innovation Institute (DII)); Chao Zhang (College of Computer Science and Technology, Harbin Engineering University); Zunlin Fan (National Innovation Institute of Defense Technology, China); Zeting Yu (Defense Innovation Institute (DII)); Qianchong Sun (Defense Innovation Institute (DII)); Mengyuan Dai (Defense Innovation Institute (DII))

3136 (ASPS-P3.8): Tracking Targets in Hyper-scale Cameras using Movement Predication

Jiaping Yu (National University of Defense Technology); Tongqing Zhou (National University of Defense Technology); Zhiping Cai (NUDT); Wenyuan Kuang (360 Digital Security Group)

2421 (ASPS-P3.9): RGB-D BASED POSE-INVARIANT FACE RECOGNITION VIA ATTENTION DECOMPOSITION MODULE

Wei-Chen Lin (Department of Computer Science, National Tsing Hua University); Ching-Te Chiu (National Tsing Hua University); Kuan-Chang Shih (Department of Computer Science, National Tsing Hua University)

256 (ASPS-P3.10): NL-DSE: Non-Local Neural Network with Decoder-Squeeze-and-Excitation for Monocular Depth Estimation

Tsung-Han Tsai (National Central University); Wei-Chung Wan (NCU)

3137 (ASPS-P3.11): Real-time modelling of observation filter in the Remote Microphone Technique for an Active Noise Control application

Chung Kwan Lai (Nanyang Technological University); Bhan Lam (NTU); Dongyuan Shi (NTU); Woon Seng Gan (NTU)

1054 (ASPS-P3.12): An Adaptive DFE Using Light-Pattern-Protection Algorithm in 12 nm CMOS Technology

Shiyuan Xing (Institute of Computing Technology, Chinese Academy of Sciences ; University of Chinese Academy of Sciences); Changlong Lin (Loongson Technology Corporation); Yuchen Li (Loongson Technology Corporation); Huandong Wang (Loongson Technology Corporation)

IVMSP-P13: Image Restoration

Room: Dome J

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Xuelong Li, Ju Sun

772 (IVMSP-P13.1): HYPERNETWORK-BASED ADAPTIVE IMAGE RESTORATION

Gil Ben-Artzi (Ariel University); Shai S.Y Aharon (Ariel University)

1015 (IVMSP-P13.2): Burst Perception-Distortion Tradeoff: Analysis and Evaluation

Danna Xue (Northwestern Polytechnical University); Luis Herranz (Computer Vision Center); Javier Vazquez-Corral (Autonomous University of Barcelona); Yanning Zhang (Northwestern Polytechnical University)

1412 (IVMSP-P13.3): ONE-SHOT NEURAL BAND SELECTION FOR SPECTRAL RECOVERY

Hai-Miao Hu (Beihang Univeristy); Zhenbo Xu (Hangzhou Innovation Institute, Beihang University, Hangzhou, China); Wenshuai Xu (School of Software, Beihang University); You Song (Beihang University); YiTao Zhang (Hangzhou Innovation Institute, Beihang University, Hangzhou, China); Liu Liu (Hangzhou ShiFang Technology Inc.); Zhilin Han (shifang); Ajin Meng (ShiFang Technology Inc.)

1721 (IVMSP-P13.4): Facial Texture Perceiver: Towards High-Fidelity Facial Texture Recovery with Input-Level Inductive Biased Perceiver IO

Seungeun Lee (UNIST)

1984 (IVMSP-P13.5): SANDFORMER: CNN and Transformer under Gated Fusion for Sand Dust Image Restoration

Shi Jun (Xinjiang University); Bingcai Wei (Shandong University of Technology); Gang Zhou (Xinjiang University); Liye Zhang (Shandong university of technology)

3073 (IVMSP-P13.6): Matrix Recovery using Deep Generative Priors with Low-Rank Deviations

Pengbin Yu (Southwest University); Jianjun Wang (Southwest University); Chen Xu (University of Ottawa)

3701 (IVMSP-P13.7): Underwater Image Restoration With Light-Aware Progressive Network

Jian Yang (Northwestern Polytechnical University); Chen Li (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

Wednesday, June 7

3929 (IVMSP-P13.8): Image Inpainting with Semantic-aware Transformer

Shiyu Chen (Southwest University of Science and Technology); Wenxin Yu (Southwest University of Science and Technology); Qi Wang (Southwest University of Science and Technology); Jun Gong (Beijing Institute of Technology); Peng Chen (Chengdu Hongchengyun Technology Co., Ltd)

4314 (IVMSP-P13.9): Multi-dimensional Signal Recovery using Low-rank Deconvolution

David Reixach (Universitat Politècnica de Catalunya, BarcelonaTech)

5753 (IVMSP-P13.10): Random Projector: Efficient Deep Image Prior

Taihui Li (University of Minnesota); Zhong Zhuang (University of Minnesota); Hengkang Wang (University of Minnesota); Ju Sun (University of Minnesota)

3094 (IVMSP-P13.11): Optimising Different Feature Types for Inpainting-based Image Representations

Ferdinand Jost (Saarland University); Vassillen Chizhov (Saarland University); Joachim Weickert (Saarland University)

3216 (IVMSP-P13.12): Non-convex approaches for low-rank tensor completion under tubal sampling

Zheng Tan (University of California, Los Angeles); Longxiu Huang (Michigan State University); HanQin Cai (University of Central Florida); Yifei Lou (University of Texas at Dallas)

IVMSP-P14: Domain-Specific Detection

Room: Dome K

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Hanzi Wang, F Richard Yu

396 (IVMSP-P14.1): DPTNet: A Dual-Path Transformer Architecture for Scene Text Detection

Jingyu Lin (厦门大学); Yan Yan (Xiamen University); Hanzi Wang (Xiamen University)

1591 (IVMSP-P14.2): DQFORMER: DYNAMIC QUERY TRANSFORMER FOR LANE DETECTION

Hao Yang (Xiamen University); Shuyuan Lin (Jinan University); Runqing Jiang (Xiamen University); Yang Lu (Xiamen University); Hanzi Wang (Xiamen University)

1780 (IVMSP-P14.3): CANet: Curved Guide Line Network with Adaptive Decoder for Lane Detection

Zhongyu Yang (University of Electronic Science and Technology of China); Chen Shen (Didi chuxing); Wei Shao (Didi Chuxing); Tengfei Xing (Didi chuxing); Runbo Hu (DiDi Chuxing); Pengfei Xu (Didi Chuxing); Hua Chai (Didi Chuxing); Ruini Xue (University of Electronic Science and Technology of China)

2139 (IVMSP-P14.4): Sample-Adapt Fusion Network for RGB-D Hand Detection in the Wild

Xingyu Liu (Beijing University of Posts and Telecommunications); Pengfei Ren (Beijing University of Posts and Telecommunications); Yuchen Chen (Beijing University of Posts and Telecommunications); Cong Liu (China Mobile); Jing Wang (Beijing University of Posts and Telecommunications); Haifeng Sun (Beijing university of posts and telecommunications); Qi Qi (Beijing University of Posts and Telecommunications); Jingyu Wang (Beijing University of Posts and Telecommunications)

2538 (IVMSP-P14.5): 2DSBG: A 2D SEMI BI-GAUSSIAN FILTER ADAPTED FOR ADJACENT AND MULTI-SCALE LINE FEATURE DETECTION

Baptiste Magnier (IMT Mines Ales CERIS); Ghulam Sakhi Shokouh (IMT Mines Ales); Louis Berthier (IMT Mines Ales CERIS); Marcel Pie-Tapia (IMT Mines Ales CERIS); Adrien Ruggiero (IMT Mines Ales CERIS)

3537 (IVMSP-P14.6): BAGGING R-CNN: ENSEMBLE FOR OBJECT DETECTION IN COMPLEX TRAFFIC SCENES

Pengteng Li (Shenzhen University); Ying He (Shenzhen University); Dongfu Yin (Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ)); F Richard Yu (Shenzhen University); Pinhao Song (KU Leuven)

4036 (IVMSP-P14.7): D-3DLD: Depth-aware Voxel Space Mapping for Monocular 3D Lane Detection with Uncertainty

Nayeon Kim (Samsung Electronics); Moonsub Byeon (Samsung Electronics); Daehyun Ji (Samsung Electronics); Dokwan Oh (Samsung Electronics)

4755 (IVMSP-P14.8): SELF-SIMILARITY IS ALL YOU NEED FOR FAST AND LIGHT-WEIGHT GENERIC EVENT BOUNDARY DETECTION

Sourabh Vasant Gothe (SAMSUNG R&D INSTITUTE BANGALORE, KARNATAKA, INDIA); Jayesh Rajkumar Vachhani (Samsung R&D Institute Bengaluru); Rishabh Khurana (Samsung Research, Bangalore); Pranay Kashyap (Samsung Research Institute Bangalore)

5144 (IVMSP-P14.9): Level-line Guided Edge Drawing for Robust Line Segment Detection

Xinyu Lin (University of Electronic Science and Technology of China); Yingjie Zhou (Sichuan University); Yipeng Liu (University of Electronic Science and Technology of China); Ce Zhu (University of Electronic Science & Technology of China)

Wednesday, June 7

5283 (IVMSP-P14.10): Dynamic Local and Global Context Exploration For Small Object Detection

Ziji Zhang (Beijing University of Posts and Telecommunications); Ping Gong (Beijing University of Posts and Telecommunications); Haotian Sun (Beijing University of Posts and Telecommunications); Pingping Wu (Beijing University of Posts and Telecommunications); Xuanyuan Yang (Beijing University of Posts and Telecommunications)

5567 (IVMSP-P14.11): INFORMATION EXTRACTION FROM PILL BOTTLE IMAGES VIA TEXT STITCHING

Rahul Kumar Gupta (Walmart Global Tech); Shilka Roy (Walmart); Sujit Jos (Walmart Global Tech); Unni V.S. (Walmart Global Tech); Lauren Lavoie (Walmart Global Tech); Frederic Medous (Walmart Global Tech); Walter Smith (Walmart Global Tech)

5663 (IVMSP-P14.12): EI2SR: LEARNING AN ENHANCED INTRA-INSTANCE SEMANTIC RELATIONSHIP FOR ARBITRARY-SHAPED SCENE TEXT DETECTION

Yan Shu (State Key Laboratory of Communication Content Cognition, People's Daily Online, Beijing, China; Harbin Institute of Technology; Institute of Information Engineering, CAS); Shaohui Liu (Harbin Institute of Technology); Yu Zhou (Institute of Information Engineering, CAS; Also with University of Chinese Academy of Sciences); honglei xu (Harbin Institute of Technology); Feng Jiang (Harbin Institute of Technology, Harbin)

IVMSP-P15: Temporal Video Analysis and Detection

Room: Dome L

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Tyng-Luh Liu, Lijun Yin

613 (IVMSP-P15.1): One-shot Action Detection via Attention Zooming In

He-Yen Hsieh (Academia Sinica); Ding-Jie Chen (Academia Sinica); Cheng-Wei Chang (Academia Sinica); Tyng-Luh Liu (Academia Sinica)

619 (IVMSP-P15.2): ScaleMix: Intra- and inter-layer multiscale feature combination for change detection

Rui Huang (Civil Aviation University of China); Qingyi Zhao (Civil Aviation University of China); Ruofei Wang (Civil Aviation University of China); Caihua Liu (College of Computer Science and Technology, Civil Aviation University of China); Sihua Gao (Civil Aviation University of China); yuxiang zhang (Civil Aviation University of China); Wei Fan (Civil Aviation University of China)

1470 (IVMSP-P15.3): SEMI-SUPERVISED REMOTE SENSING IMAGE CHANGE DETECTION USING MEAN TEACHER MODEL FOR CONSTRUCTING PSEUDO-LABELS

mao zan (ucas); xinyu tong (Computer Network Information Center); Ze Luo (Computer Network Information Center, Chinese Academy of Sciences)

1575 (IVMSP-P15.4): MODULATION-BASED CENTER ALIGNMENT AND MOTION MINING FOR SPATIAL TEMPORAL ACTION DETECTION

Weiji Zhao (Shanghai Jiao Tong University); KeFeng Huang (Shanghai Jianke Engineering Consulting Co.,Ltd); Chongyang Zhang (Shanghai Jiao Tong University)

2692 (IVMSP-P15.5): DL-NET: DILATION LOCATION NETWORK FOR TEMPORAL ACTION DETECTION

Dianlong You (yanshan university); Houlin Wang (yanshan university); Bingxin Liu (yanshan university); Yang Yu (yanshan university); Zhiming Li (yanshan university)

2873 (IVMSP-P15.6): SEMI-SUPERVISED CONTRASTIVE LEARNING WITH SOFT MASK ATTENTION FOR FACIAL ACTION UNIT DETECTION

Zhongling Liu (Fujitsu Research and Development Center); Rujie Liu (Fujitsu Research & Development Center Co., Ltd.); Ziqiang Shi (Fujitsu Research & Development Center); Liu Liu (Fujitsu Research & Development Center); Xiaoyu Mi (Fujitsu Laboratories Ltd.); Kento Murase (Fujitsu Laboratories Ltd.)

4046 (IVMSP-P15.7): LOCAL-GLOBAL SIAMESE NETWORK WITH EFFICIENT INTER-SCALE FEATURE LEARNING FOR CHANGE DETECTION IN VHR REMOTE SENSING IMAGES

Yue Zhang (Shaanxi University of Science and Technology); Tao Lei (Shaanxi University of Science and Technology); Shaoxiang Han (Norinco Group Testing And Research institute); Yetong Xu (Shaanxi University of Science and Technology); Asoke K Nandi (Brunel University London)

4951 (IVMSP-P15.8): Multimodal Facial Action Unit Detection with Physiological Signals

Zhihua Li (Binghamton University); Lijun Yin (State University of New York at Binghamton)

5755 (IVMSP-P15.9): Background-Weakening Consistency Regularization for Semi-Supervised Video Action Detection

Xian Zhong (Wuhan University of Technology); Aoyu Yi (Wuhan University of Technology); Wenxuan Liu (Wuhan University of Technology); Wenxin Huang (Hubei University); Chengming Zou (Wuhan University of Technology); Zheng Wang (Wuhan University)

5713 (IVMSP-P15.10): Low in Resolution, High in Precision: UAV Detection with Super-Resolution and Motion Information Extraction

Hanzhuo Wang (Zhejiang University); Xingjian Wang (Zhejiang University); Chengwei Zhou (Zhejiang University); Wenchao Meng (Zhejiang University); Zhiguo Shi (Zhejiang University)

Wednesday, June 7

3350 (IVMSP-P15.11): Temporal Contrastive Learning with Curriculum
Shuvendu Roy (Queen's University); Ali Etemad (Queen's University)

320 (IVMSP-P15.12): LongShortNet: Exploring Temporal and Semantic Features Fusion in Streaming Perception
Chenyang Li (DAMO Academy, Alibaba Group); Zhi-Qi Cheng (Carnegie Mellon University); Jun-Yan He (DAMO Academy, Alibaba Group); Pengyu Li (Alibaba Group); Bin Luo (DAMO Academy, Alibaba Group); Hanyuan Chen (Alibaba); Yifeng Geng (Alibaba Group); Jin-Peng Lan (DAMO Academy, Alibaba Group); Xuansong Xie (DAMO Academy, Alibaba Group)

IVMSP-P16: Object Detection
Room: Dome M
Type: Poster
08:15 AM to 09:45 AM
Chair(s): Kaihua Zhang, Haoran Xie

296 (IVMSP-P16.1): Group-wise Co-salient Object Detection with Siamese Transformers via Brownian Distance Covariance Matching
Yang Wu (nuist); Hao Zhang (Nuist); lingyan liang (inspur); Yaqian Zhao (Inspur); Kaihua Zhang (Inspur, NUIST)

4090 (IVMSP-P16.2): Gated Enhanced RPN and Hybrid-View for Few-Shot Object Detection
Xujun Wei (Fudan University); Zechu Zhou (Academy of Engineering and Technology, Fudan University); Pinxue Guo (Fudan University); Wenqiang Zhang (Fudan University)

1966 (IVMSP-P16.3): D2Q-DETR: Decoupling and Dynamic Queries for Oriented Object Detection with Transformers
Qiang Zhou (Alibaba Group); Chaohui Yu (Alibaba Group); Zhibin Wang (Alibaba Group); Fan Wang (Alibaba Group)

2710 (IVMSP-P16.4): iSmallNet: Densely Nested Network with Label Decoupling for Infrared Small Target Detection
Zhiheng Hu (Nanjing University of Aeronautics and Astronautics); Yongzhen Wang (Nanjing University of Aeronautics and Astronautics); Peng Li (Nanjing University of Aeronautics and Astronautics); Jie Qin (Nanjing University of Aeronautics and Astronautics); Haoran Xie (Lingnan University); Mingqiang Wei (Nanjing University of Aeronautics and Astronautics)

2858 (IVMSP-P16.5): IFUNET++: ITERATIVE FEEDBACK UNET++ FOR INFRARED SMALL TARGET DETECTION
Zhangying Weng (Nanjing University of Aeronautics and Astronautics); Peng Li (Nanjing University of Aeronautics and Astronautics); Xin Zhuang (BeijingAerospaceIntelligentManufacturingTechnologyDevelopmentCo.,Ltd); Xuefeng Yan (Nanjing University of Aeronautics and Astronautics); Lina Gong (Nanjing University of Aeronautics and Astronautics); Haoran Xie (Lingnan University); Mingqiang Wei (Nanjing University of Aeronautics and Astronautics)

3032 (IVMSP-P16.6): YOLO-Based Lightweight Object Detection with Structure Simplification and Attention Enhancement
Shuqi Sun (University of Jinan); Xiaohui Yang (University of Jinan); Jingliang Peng (University of Jinan)

575 (IVMSP-P16.7): Building Change Detection using Cross-temporal Feature Interaction Network
Yuchao Feng (Zhejiang University of Technology); jiawei jiang (zhejiang university of technology); Honghui Xu (Zhejiang University of Technology); Jianwei Zheng (Zhejiang University of Technology)

3477 (IVMSP-P16.8): TINYCOD: TINY AND EFFECTIVE MODEL FOR CAMOUFLAGED OBJECT DETECTION
Haozhe Xing (Fudan University); Shuyong Gao (Fudan University); Hao Tang (ETH Zurich); Tsui Qin Mok (Fudan University); Yanlan Kang (Fudan University); Wenqiang Zhang (Fudan University)

3532 (IVMSP-P16.9): BAUENet: Boundary-Aware Uncertainty Enhanced Network for Infrared Small Target Detection
Tianxiang Chen (University of Science and Technology of China); Qi Chu (University of Science and Technology of China); Zhentao Tan (Alibaba DAMO Academy); Bin Liu (University of Science and Technology of China); Nenghai Yu (University of Science and Technology of China)

5955 (IVMSP-P16.10): YOLOX-B: A BETTER YOLOX MODEL FOR REAL-TIME DRIVER BEHAVIOR DETECTION
Xu Guo (Inner Mongolia University); Ming Ma (Inner Mongolia University); Jiaqiang Zhang (Inner Mongolia University); Shaojie Li (Inner Mongolia University)

5699 (IVMSP-P16.11): STATIC-SCENE CONSTRAINED OPTIMIZATION FOR MATRIX/TENSOR-DECOMPOSITION-FREE FOREGROUND-BACKGROUND SEPARATION
Kazuki Naganuma (Tokyo Institute of Technology); Shunsuke Ono (Tokyo Institute of Technology)

MLSP-P13: Graph based Learning III
Room: Dome F
Type: Poster
08:15 AM to 09:45 AM
Chair(s): TBA

4742 (MLSP-P13.1): Data-Driven Graph Convolutional Neural Networks for Power System Contingency Analysis
Valentin Bolz (DigSILENT GmbH & University of Tuebingen); Johannes Ruess (DigSILENT GmbH); Andreas Zell (University of Tuebingen)

Wednesday, June 7

4563 (MLSP-P13.2): Hierarchical Graph Learning for Stock Market Prediction via a Domain-Aware Graph Pooling Operator
Arie N Arya (Imperial College London); Yao Lei Xu (Imperial College London); Ljubisa Stankovic (University of Montenegro); Danilo P. Mandic (Imperial College of London, UK)

5963 (MLSP-P13.3): LE-DTA: Local Extrema convolution for Drug Target Affinity Prediction
Tanoj Langore (National Taiwan University); Te-Cheng Hsu (National Tsing Hua University); Yi Hsien Hsieh (National Taiwan University); Che Lin (National Taiwan University)

5761 (MLSP-P13.4): RETHINKING RANDOM WALK IN GRAPH REPRESENTATION LEARNING
DingYi Zeng (University of Electronic Science and Technology of China); Wenyu Chen (University of Electronic Science and Technology of China); Wanlong Liu (University of Electronic Science and Technology of China); Li Zhou (University of Electronic Science and Technology of China); Hong Qu (University of Electronic Science and Technology of China)

2547 (MLSP-P13.5): TREEGNN: CAN GRADIENT-BOOSTED DECISION TREES HELP BOOST HETEROGENEOUS GRAPH NEURAL NETWORKS?
Ming-Yi Hong (National Taiwan University); Shih-Yen Chang (National Taiwan University); Hao-Wei Hsu (National Taiwan University); Yi-Hsiang Huang (National Taiwan University); Chih-Yu Wang (Academia Sinica); Che Lin (National Taiwan University)

5288 (MLSP-P13.6): SEMI-SUPERVISED GRAPH ULTRA-SPARSIFIERS USING REWEIGHTED L1 OPTIMIZATION
Jiayu Li (Syracuse University); Tianyun Zhang (Cleveland State University); Shengmin Jin (Syracuse University); Reza Zafarani (Syracuse University)

982 (MLSP-P13.7): Higher-order Sparse Convolutions in Graph Neural Networks
Jhony H. Giraldo (Télécom Paris); Sajid Javed (Khalifa University of Science and Technology); Arif Mahmood (Information Technology University); Fragkiskos Malliaros (CentraleSupélec); Thierry BOUWMANS (Univ. La Rochelle)

1023 (MLSP-P13.8): Deep Manifold Graph Auto-Encoder for Attributed Graph Embedding
Bozhen Hu (Zhejiang University & Westlake University); Zelin Zang (Zhejiang University & Westlake University); Jun Xia (Westlake University); Lirong Wu (Westlake University); Cheng Tan (Zhejiang University & Westlake University); Stan Z. Li (Westlake University)

6178 (MLSP-P13.9): Relevance Propagation through Deep Conditional Random Fields
Xiangyu Yang (Vrije Universiteit Brussel); Boris Joukovsky (Vrije Universiteit Brussel - imec); Nikos Deligiannis (Vrije Universiteit Brussel - imec)

5246 (MLSP-P13.10): DIRECTION AWARE POSITIONAL AND STRUCTURAL ENCODING FOR DIRECTED GRAPH NEURAL NETWORKS
Yonas A Sium (Iowa State University); Georgios Kollias (IBM Research); Tsuyoshi Ide (IBM Research, T. J. Watson Research Center); Payel Das (IBM Research); Naoki Abe (IBM Research); Aurelie Lozano (IBM Research); Qi Li (Iowa State University)

5310 (MLSP-P13.11): Graph Representation Learning for Stroke Recurrence Prediction
Nicholas Glaze (Rice University); Artun Bayer (Rice University); Xiaoqian Jiang (The University of Texas Health Science Center at Houston); Sean Savitz (University of Texas Health Science Center at Houston); Santiago Segarra (Rice University)

5551 (MLSP-P13.12): Dual-Stage Graph Convolution Network With Graph Learning For Traffic Prediction
Li Zilong (Heilongjiang University); Qianqian Ren (Heilongjiang University); Long Chen (Heilongjiang University); jianguo sun (xidian university)

MLSP-P14: Deep Learning for Image and Video Processing II

Room: Dome G

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Alexandros Iosifidis,

5437 (MLSP-P14.1): POSITION-AWARE GRAPH-BASED LEARNING OF WHOLE SLIDE IMAGES
Milan Aryal (Marquette University); Nasim Yahyasoltani (Marquette university)

560 (MLSP-P14.2): DepthFormer: Multimodal Positional Encodings and Cross-Input Attention for Transformer-Based Segmentation Networks
Francesco Barbato (University of Padova); Giulia Rizzoli (University of Padova); Pietro Zanuttigh (University of Padova)

1735 (MLSP-P14.3): OPEN-SET AUTOMATIC TARGET RECOGNITION
Bardia Safaei (Johns Hopkins University); Vibashan VS (Johns Hopkins University); Celso M. de Melo (DEVCOM Army Research Laboratory); Shuowen Hu (US Army Research Laboratory); Vishal Patel (Johns Hopkins University)

2180 (MLSP-P14.4): Ternary Weight Networks
Bin Liu (Shanghai Jiao Tong University); Fengfu Li (UCAS); Xiaoxing Wang (Shanghai Jiao Tong University); Bo Zhang (Institute of Applied Mathematics, AMSS, Chinese Academy of Sciences); Junchi Yan (Shanghai Jiao Tong University)

Wednesday, June 7

5109 (MLSP-P14.5): gSwin: Gated MLP Vision Model with Hierarchical Structure of Shifted Window
Mocho Go (PKSHA Technology Inc.); Hideyuki Tachibana (PKSHA Technology)

2853 (MLSP-P14.6): PU-EdgeFormer: Edge Transformer for Dense Prediction in Point Cloud Upsampling
Dohoon Kim (Chung-Ang University); Minwoo Shin (Chungang University); Joonki Paik (Chungang University)

1065 (MLSP-P14.7): Self-Attention based Action Segmentation using Intra-and Inter-segment Representations
Constantin Patsch (Technical University of Munich); Eckehard Steinbach (TUM)

1463 (MLSP-P14.8): Not All Classes are Equal: Adaptively Focus-Aware Confidence for Semi-Supervised Object Detection
Hui Zhu (Institute of Computing Technology, Chinese Academy of Sciences); Yongchun Lu (Mashang Consumer Finance Co., Ltd.); hongyu zhao (Mashang Consumer Finance Company Ltd.); Guoqing Zhao (Mashang Consumer Finance Co., Ltd); Xiaofang Zhao (Institute of Computing Technology, Chinese Academy of Sciences; Institute of Intelligent Computing Technology, Suzhou, CAS)

2434 (MLSP-P14.9): Structural Reparameterization Lightweight Network for Video Action Recognition
AnLei Zhu (Jiangnan University); Wang Yinghui (Jiangnan University); Wei Li (Jiangnan University); Pengjiang Qian (Jiangnan University)

3207 (MLSP-P14.10): JOINT ANN-SNN CO-TRAINING FOR OBJECT DETECTION AND IMAGE SEGMENTATION
Marc J Baltes (Ohio University); Nidal Abuhajar (Ohio University); Ye Yue (Ohio University); Charles Smith (University of Kentucky); Jundong Liu (Ohio University)

624 (MLSP-P14.11): Enhanced Low-resolution LiDAR-Camera Calibration Via Depth Interpolation and Supervised Contrastive Learning
Zhikang Zhang (Arizona State University); Zifan Yu (Arizona State University); Suya You (U.S. Army Research Laboratory); Raghuvveer Rao (Army Research Laboratory); Sanjeev Agarwal (U.S. Army DEVCOM C5ISR Center); Fengbo Ren (Arizona State University)

5544 (MLSP-P14.12): Meta-Learning for Image-Guided Millimeter-Wave Beam Selection in Unseen Environments
Jerry Gu (Northeastern University); Liam Collins (University of Texas at Austin); Debashri Roy (Northeastern University); Aryan Mokhtari (UT Austin); Sanjay Shakkottai (University of Texas at Austin); Kaushik Chowdhury (Northeastern University)

MLSP-P15: Deep Learning for Speech and Audio Processing I

Room: Dome H

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Paola Garcia,

1515 (MLSP-P15.1): HEIMDaL: Highly Efficient Method for Detection and Localization of wake-words
Arnav Kundu (Apple); Mohammad Samragh (Apple); Minsik Cho (Apple); Priyanka Padmanabhan (Apple); Devang Naik (Apple)

1961 (MLSP-P15.2): The MBSTOI Binaural Intelligibility Metric Using a Close-Talking Microphone Reference
Pierre Guiraud (Imperial College London); Alastair H Moore (Imperial College London); Rebecca Vos (Imperial College London); Patrick A. Naylor (Imperial College London); Mike Brookes (Imperial College London)

6760 (MLSP-P15.3): Learning Representations for New Sound Classes With Continual Self-Supervised Learning (SPS Journal Paper)*
Zhepei Wang (University of Illinois at Urbana Champaign); Cem Subakan (Mila); Xilin Jiang (University of Illinois at Urbana Champaign); Junkai Wu (University of Illinois at Urbana Champaign); Efthymios Tzinis (University of Illinois at Urbana-Champaign); Mirco Ravanelli (Université de Montréal); Paris Smaragdīs (University of Illinois at Urbana-Champaign)

3798 (MLSP-P15.4): Feature-Rich Audio Model Inversion for Data-Free Knowledge Distillation Towards General Sound Classification
Zuheng Kang (Ping An Technology (Shenzhen) Co., Ltd); Yayun He (Ping An Technology (Shenzhen) Co., Ltd); Jianzong Wang (Ping An Technology (Shenzhen) Co., Ltd); Junqing Peng (Ping An Technology (Shenzhen) Co., Ltd); Xiaoyang Qu (Ping An Technology (Shenzhen) Co., Ltd); Jing Xiao (Ping An Insurance (Group) Company of China)

5108 (MLSP-P15.5): A Content Adaptive Learnable "Time-Frequency" Representation For Audio Signal Processing
Prateek Verma (Stanford University); Chris Chafe (organization)

5055 (MLSP-P15.6): HeartToHeart: The Arts of Infant Versus Adult-Directed Speech Classification
Najla D Al Futaisi (Imperial College London); Alejandrina Cristia (PSL Research University); Bjoern W. Schuller (Imperial College London)

5063 (MLSP-P15.7): RepackagingAugment: Overcoming Prediction Error Amplification in Weight-averaged Speech Recognition Models Subject to Self-training
Jae-Hong Lee (Hanyang University); Dong-Hyun Kim (Hanyang University); Joon-Hyuk Chang (Hanyang University)

Wednesday, June 7

4729 (MLSP-P15.8): MAST: Multiscale Audio Spectrogram Transformers

Sreyan Ghosh (University of Maryland, College Park); Ashish Seth (IIT Madras); S Umesh (IIT Chennai); Dinesh Manocha (University of Maryland at College Park)

5294 (MLSP-P15.10): ERSAM: Neural Architecture Search For Energy-Efficient and Real-Time Social Ambiance Measurement

Chaojian Li (Georgia Institute of Technology); Wenwan Chen (Rice University); Jiayi Yuan (Rice University); Yingyan (Celine) Lin (Georgia Tech); Ashutosh Sabharwal (Rice University)

4736 (MLSP-P15.11): SLICER: Learning universal audio representations using low-resource self-supervised pre-training

Ashish Seth (IIT Madras); Sreyan Ghosh (University of Maryland, College Park); S Umesh (IIT Chennai); Dinesh Manocha (University of Maryland at College Park)

3690 (MLSP-P15.12): On Adversarial Robustness of Audio Classifiers

Kangkang Lu (A-STAR); Cuong Nguyen (Institute for Infocomm Research, ASTAR); Xun Xu (Institute for Infocomm Research, ASTAR); Chuan Sheng Foo (Institute for Infocomm Research, ASTAR)

MLSP-P16: Deep Learning for Speech and Language Processing

Room: Dome I

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Iván López-Espejo,

2024 (MLSP-P16.1): Wav2Seq: Pre-training Speech-to-Text Encoder-Decoder Models Using Pseudo Languages

Felix Wu (ASAPP); Kwangyoum Kim (ASAPP); Shinji Watanabe (Carnegie Mellon University); Kyu Jeong Han (ASAPP); Ryan Mcdonald (ASAPP); Kilian Weinberger (Cornell University); Yoav Artzi (Cornell University)

5031 (MLSP-P16.2): Towards Dialogue Modeling Beyond Text

Tongzi Wu (University of Toronto); Yuhao Zhou (Talka AI); Wang Ling (Talka AI); Hojin Yang (talka ai); Joana Veloso (Talka AI); Lin Sun (Talka AI); Ruixin Huang (Talka AI); Norberto Guimaraes (Talka AI); Scott Sanner (University of Toronto)

152 (MLSP-P16.3): Active Learning of Non-semantic Speech Tasks with Pretrained Models

Harlin Lee (University of California Los Angeles); Aaqib Saeed (Eindhoven University of Technology); Andrea L. Bertozzi (UCLA)

5012 (MLSP-P16.4): ANALYSING THE MASKED PREDICTIVE CODING TRAINING CRITERION FOR PRE-TRAINING A SPEECH REPRESENTATION MODEL

Hemant Yadav (MIDAS); Sunayana Sitaram (Microsoft Research); Rajiv Ratn Shah (IIIT Delhi)

2655 (MLSP-P16.5): AURA: PRIVACY-PRESERVING AUGMENTATION TO IMPROVE TEST SET DIVERSITY IN SPEECH ENHANCEMENT

xavier gitiaux (Microsoft); Aditya Khant (Microsoft); Ross Cutler (Microsoft Corporation); Chandan Reddy (Google); Ebrahim Beyrami (Microsoft); Jayant Gupchup (Microsoft)

5021 (MLSP-P16.6): Audio-visual speaker diarization in the framework of multi-user human-robot interaction

Timothée Dhaussy (Université Avignon); Bassam Jabaian (LIA - Avignon university); Fabrice Lefevre (Univ. Avignon); Radu Horaud (Inria)

3450 (MLSP-P16.7): Leveraging Language Embeddings for Cross-lingual Self-supervised Speech Representation Learning

Tomohiro Tanaka (NTT); Ryo Masumura (NTT Corporation); Mana Ithori (NTT); Hiroshi Sato (NTT Corporation); Taiga Yamane (NTT); Takanori Ashihara (NTT Corp.); Kohei Matsuura (NTT); Takafumi Moriya (NTT)

4048 (MLSP-P16.8): SHUFFLEAUGMENT: A DATA AUGMENTATION METHOD USING TIME SHUFFLING

Yoshinao Sato (Fairy Devices Inc.); Narumitsu Ikeda (Graduate School of Information Science and Technology, The University of Tokyo); Hirokazu Takahashi (Graduate School of Information Science and Technology, The University of Tokyo)

1282 (MLSP-P16.9): PRECOGNITION IN CONTEXTUAL SPOKEN LANGUAGE UNDERSTANDING VIA KNOWLEDGE DISTILLATION

Nan Su (Ant Group); Bingzhu Du (Ant Group); Yuchi Zhang (Ant Financial Services Group); Chao Liu (Ant Group); Yongliang Wang (Ant Group); Hong Chen (Ant Group); xin lu (ant group)

3090 (MLSP-P16.10): Exploration of Language Dependency for Japanese Self-Supervised Speech Representation Models

Takanori Ashihara (NTT Corp.); Takafumi Moriya (NTT); Kohei Matsuura (NTT); Tomohiro Tanaka (NTT Corp.)

3590 (MLSP-P16.11): Large-Scale Contrastive Language-Audio Pretraining with Feature Fusion and Keyword-to-Caption Augmentation

Yusong Wu (Mila, University of Montreal); Ke Chen (University of California San Diego); Tianyu Zhang (Mila, Université de Montréal); Yuchen Hui (Université de Montréal); Taylor Berg-Kirkpatrick (UCSD); Shlomo Dubnov (UC San Diego)

Wednesday, June 7

5910 (MLSP-P16.12): HINDI AS A SECOND LANGUAGE: IMPROVING VISUALLY GROUNDED SPEECH WITH SEMANTICALLY SIMILAR SAMPLES

Hyeonggon Ryu (KAIST); Arda Senocak (KAIST); In So Kweon (KAIST); Joon Son Chung (KAIST)

SLT-P13: Expressive and Controllable TTS II

Room: Dome B

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Wang Xin,

1241 (SLT-P13.1): PERIOD VITS: VARIATIONAL INFERENCE WITH EXPLICIT PITCH MODELING FOR END-TO-END EMOTIONAL SPEECH SYNTHESIS

Yuma Shirahata (LINE Corp.); Ryuichi Yamamoto (LINE Corp.); Eunwoo Song (Naver Corporation); Ryo Terashima (LINE Corp.); Jae-Min Kim (NAVER Cloud Corp.); Kentaro Tachibana (LINE Corp.)

2233 (SLT-P13.2): EmoDiff: Intensity Controllable Emotional Text-to-Speech with Soft-Label Guidance

Yiwei Guo (Shanghai Jiao Tong University); Chenpeng Du (Shanghai Jiao Tong University); Xie Chen (Shanghai Jiaotong University); Kai Yu (Shanghai Jiao Tong University)

2729 (SLT-P13.3): PROSODY-AWARE SPEECH5 FOR EXPRESSIVE NEURAL TTS

Yan Deng (Microsoft); Long Zhou (Microsoft Research Asia); Yuanhao Yi (Microsoft); Shujie Liu (Microsoft Research Asia); Lei He (Microsoft Cloud and AI)

2832 (SLT-P13.4): MULTI-SPEAKER EXPRESSIVE SPEECH SYNTHESIS VIA MULTIPLE FACTORS DECOUPLING

Xinfa Zhu (Northwestern Polytechnical University); Yi Lei (Northwestern Polytechnical University); Kun Song (Northwestern Polytechnical University); yongmao zhang (Audio, Speech and Language Processing Group (ASLP@NPU), School of Computer Science, Northwestern Polytechnical University, Xi'an, China); Tao Li (School of Computer Science, Northwestern Polytechnical University, Xi'an); Lei Xie (NWPU)

3038 (SLT-P13.5): Self-adaptive Incremental Machine Speech Chain for Lombard TTS with High-granularity ASR Feedback in Dynamic Noise Condition

Sashi Novitasari (Nara Institute of Science and Technology); Sakriani Sakti (Japan Advanced Institute of Science and Technology); Satoshi Nakamura (Nara Institute of Science and Technology, Japan)

3151 (SLT-P13.6): FINE-GRAINED EMOTIONAL CONTROL OF TEXT-TO-SPEECH: LEARNING TO RANK INTER- AND INTRA-CLASS EMOTION INTENSITIES

Shijun Wang (University of St. Gallen); Jon Gudnason (Reykjavik University); Damian Borth (University of St. Gallen)

3487 (SLT-P13.7): PROMPTTTS: CONTROLLABLE TEXT-TO-SPEECH WITH TEXT DESCRIPTIONS

Zhifang Guo (University of Chinese Academy of Sciences); Yichong Leng (University of Science and Technology of China); Yihan Wu (Renmin University of China); sheng zhao (Microsoft); Xu Tan (Microsoft Research Asia)

3687 (SLT-P13.8): NSV-TTS: NON-SPEECH VOCALIZATION MODELING AND TRANSFER IN EMOTIONAL TEXT-TO-SPEECH

Haitong Zhang (Netease Games AI Lab); Xinyuan Yu (Netease Games AI Lab); Yue Lin (NetEase Games AI Lab)

4083 (SLT-P13.9): Prosody-controllable spontaneous TTS with neural HMMs

Harm Lameris (KTH Royal Institute of Technology); Shivam Mehta (KTH Royal Institute of Technology); Gustav Eje Henter (KTH Royal Institute of Technology); Joakim Gustafson (KTH Royal Institute of Technology); Eva Szekely (KTH Royal Institute of Technology)

4189 (SLT-P13.10): Source-Filter HiFi-GAN: Fast and Pitch Controllable High-Fidelity Neural Vocoder

Reo Yoneyama (Nagoya University); Yi-Chiao Wu (META); Tomoki Toda (Nagoya University)

5362 (SLT-P13.11): Enhancement of text-predicting style token with generative adversarial network for expressive speech synthesis

Hiroki Kanagawa (NTT Corporation); Yusuke Ijima (NTT Corporation)

5605 (SLT-P13.12): QI-TTS: Questioning Intonation Control for Emotional Speech Synthesis

Haobin Tang (USTC); Xulong Zhang (Ping An Technology (Shenzhen) Co., Ltd.); Jianzong Wang (Ping An Technology (Shenzhen) Co., Ltd); Ning Cheng (Ping An Technology (Shenzhen) Co., Ltd); Jing Xiao (Ping An Insurance (Group) Company of China)

SLT-P14: Language Modeling and Representation Learning

Room: Dome C

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Samuel Thomas, Atsunori Ogawa

188 (SLT-P14.1): Heuristic Masking for Text Representation Pretraining

Yimeng Zhuang (Samsung Research China - Beijing (SRC-B))

Wednesday, June 7

1059 (SLT-P14.3): A CONTRASTIVE FRAMEWORK TO ENHANCE UNSUPERVISED SENTENCE REPRESENTATION LEARNING

Haoyang Ma (North China Institute of Computing Technology); Zeyu Li (Communication university of China); Hongyu Guo (North China Institute of Computing Technology)

1313 (SLT-P14.4): ESCL: EQUIVARIANT SELF-CONTRASTIVE LEARNING FOR SENTENCE REPRESENTATIONS

Jie Liu (China Mobile Research); Yixuan Liu (Beijing University of Posts and Telecommunications); Xue Han (China Mobile Research); Chao Deng (China Mobile Research Institute); Junlan Feng (China Mobile Research)

2621 (SLT-P14.5): Training Large-Vocabulary Neural Language Models by Private Federated Learning for Resource-Constrained Devices

Mingbin Xu (Apple); Congzheng Song (Apple); Ye Tian (Apple); Neha Agrawal (Apple); Filip Granqvist (Apple); Rogier C van Dalen (Samsung AI Center, Cambridge, UK); Xiao Zhang (Apple); Arturo Argueta (Apple); Shiyi Han (Apple); Yaqiao Deng (Apple); Leo Liu (Apple); Anmol Walia (Apple); Alex Jin (Apple)

2864 (SLT-P14.6): ADAPTER TUNING WITH TASK-AWARE ATTENTION MECHANISM

Jinliang Lu (Institute of Automation, Chinese Academy of Sciences); Feihu Jin (Institute of Automation, Chinese Academy of Sciences); Jiajun Zhang (Institute of Automation Chinese Academy of Sciences)

3099 (SLT-P14.7): UNIFIED PROMPT LEARNING MAKES PRE-TRAINED LANGUAGE MODELS BETTER FEW-SHOT LEARNERS

Feihu Jin (Institute of Automation, Chinese Academy of Sciences); Jinliang Lu (Institute of Automation, Chinese Academy of Sciences); Jiajun Zhang (Institute of Automation Chinese Academy of Sciences)

3656 (SLT-P14.8): Selecting Language Models Features via Software-Hardware Co-Design

Vlad Pandealea (Nanyang Technological University); Edoardo Ragusa (University of Genova); Paolo Gastaldo (University of Genova); Erik Cambria (Nanyang Technological University, Singapore)

4855 (SLT-P14.9): SUPERVISED CONTRASTIVE LEARNING AS MULTI-OBJECTIVE OPTIMIZATION FOR FINE-TUNING LARGE PRE-TRAINED LANGUAGE MODELS

youness moukafih (International University of Rabat); Mounir Ghogho (Université Internationale de Rabat); Kamel Smaïli (University of Lorraine)

5093 (SLT-P14.10): Meta Learning for Domain Agnostic Soft Prompt

Ming-Yen Chen (National Yang Ming Chiao Tung University); Mahdin Rohmatillah (National Yang Ming Chiao Tung University); Ching-hsien Lee (Industrial Technology Research Institute); Jen-Tzung Chien (National Yang Ming Chiao Tung University)

5616 (SLT-P14.11): URM4DMU: An User Representation Model for Darknet Markets Users

Hongmeng Liu (Beijing University of Posts and Telecommunications); zhao jiapeng (Beijing University of Posts and Telecommunications); Yixuan Huo (Beijing University of Posts and Telecommunications); Wang Yuyan (Beijing University of Posts and Telecommunications); Chun Liao (Institute of Information Engineering, CAS); Liyan Shen (Beijing University of Posts and Telecommunications); Shiyao Cui (Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China); Jinqiao Shi (Beijing University of Posts and Telecommunications)

531 (SLT-P14.2): A Reality Check and A Practical Baseline for Semantic Speech Embedding

Guangyu Chen (Renmin University of China); Yuanyuan Cao (Renmin University of China)

SLT-P15: Lightweight TTS and TTS Analysis

Room: Dome D

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Andros Tjandra,

462 (SLT-P15.1): LightGrad: Lightweight Diffusion Probabilistic Model for Text-to-Speech

Jie Chen (Shenzhen International Graduate School, Tsinghua University); Xingchen Song (Horizon Robotics, Beijing, China); Zhendong Peng (Horizon Robotics, Beijing, China); Binbin Zhang (Horizon Robotics, Beijing, China); Fuping Pan (Horizon Robotics, Beijing, China); Zhiyong Wu (Tsinghua University)

565 (SLT-P15.2): StreamSpeech: Low-Latency Neural Architecture for High-Quality On-Device Speech Synthesis

Georgi S Shopov (IICT-BAS); Stefan Gerdjikov (FMI, Sofia University); Stoyan Mihov (IICT-BAS)

578 (SLT-P15.3): LiteG2P: A fast, light and high accuracy model for Grapheme-to-Phoneme conversion

Chunfeng Wang (Bytedance Inc); Peisong Huang (ByteDance Inc.); Yuxiang Zou (Bytedance); Haoyu Zhang (Bytedance); Shichao Liu (ByteDance); Xiang Yin (ByteDance AI LAB); Zejun Ma (Bytedance)

663 (SLT-P15.4): Evaluating Speech-Phoneme Alignment and Its Impact on Neural Text-To-Speech Synthesis

Frank Zalkow (Fraunhofer IIS); Prachi Govalkar (Fraunhofer IIS); Meinard Müller (International Audio Laboratories Erlangen); Emanuel Habets (Fraunhofer IIS); Christian Dittmar (Fraunhofer IIS)

Wednesday, June 7

1713 (SLT-P15.5): SQuid: Measuring Speech Naturalness in Many Languages

Thibault Sellam (Google); Ankur Bapna (Google Research); Joshua Camp (Google); Diana Mackinnon (Google); Ankur Parikh (Google); Jason Riesa (Google)

2642 (SLT-P15.6): Phoneme-Level BERT for Enhanced Prosody of Text-to-Speech with Grapheme Predictions

Yinghao A Li (Columbia University); Cong Han (Columbia University); Xilin Jiang (Columbia University); Nima Mesgarani (Columbia University)

3035 (SLT-P15.7): LIGHTWEIGHT AND HIGH-FIDELITY END-TO-END TEXT-TO-SPEECH WITH MULTI-BAND GENERATION AND INVERSE SHORT-TIME FOURIER TRANSFORM

Masaya Kawamura (The University of Tokyo); Yuma Shirahata (LINE Corp.); Ryuichi Yamamoto (LINE Corp.); Kentaro Tachibana (LINE Corp.)

3971 (SLT-P15.8): HOW TO PUSH THE FASTEST MODEL 50X FASTER: STREAMING NON-AUTOREGRESSIVE SPEECH SYNTHESIS ON RESOURCE-LIMITED DEVICES

Thinh Van Nguyen (VinBigdata); Cuong H Pham (VinBigdata JSC); Dang-Khoa MAC (VinBigdata)

4328 (SLT-P15.9): VF-TACO2: TOWARDS FAST AND LIGHTWEIGHT SYNTHESIS FOR AUTOREGRESSIVE MODELS WITH VARIATION AUTOENCODER AND FEATURE DISTILLATION

Yuhao Liu (Tianjin University); Cheng Gong (Tianjin University); Longbiao Wang (Tianjin University); Xixin Wu (The Chinese University of Hong Kong); Qiuyu Liu (Tianjin University); Jianwu Dang (Tianjin University)

4850 (SLT-P15.10): Investigating Content-Aware Neural Text-To-Speech MOS Prediction Using Prosodic and Linguistic Features

Alexandra Vioni (Innoetics, Samsung Electronics); Georgia Maniati (Samsung Electronics); Nikolaos Ellinas (Innoetics, Samsung Electronics); June Sig Sung (Samsung Electronics); Inchul Hwang (Samsung Research); Aimilios Chalamandaris (Samsung Electronics); Pirros Tsiakoulis (Samsung)

5204 (SLT-P15.11): A Fast and Accurate Pitch Estimation Algorithm Based on the Pseudo Wigner-Ville Distribution

Yisi Liu (University of Chinese Academy of Sciences); Peter Wu (UC Berkeley); Alan Black (CMU); Gopala Krishna Anumanchipalli (UC Berkeley)

5603 (SLT-P15.12): Personalized Lightweight Text-to-Speech: Voice Cloning with Adaptive Structured Pruning

Sung-Feng Huang (National Taiwan University); Chia-ping Chen (Intelligo Technology Inc); Zhi-Sheng Chen (Intelligo Technology Inc); Yu-Pao Tsai (Intelligo Technology Inc); Hung-yi Lee (National Taiwan University)

SLT-P16: Machine Translation for Spoken and Written Language

Room: Dome E

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Markus Mueller, Sakriani Sakti

683 (SLT-P16.1): Improving Speech-to-Speech Translation Through Unlabeled Text

Xuan-Phi Nguyen (Nanyang Technological University); Sravya Popuri (Facebook Inc); Changhan Wang (Facebook AI Research); Yun Tang (Facebook); Iliia Kulikov (Meta AI); Hongyu Gong (Meta AI)

1867 (SLT-P16.2): A Holistic Cascade System, Benchmark, and Human Evaluation Protocol for Expressive Speech-to-Speech Translation

Wen-Chin Huang (Nagoya University); Benjamin Peloquin (Meta AI); Justine Kao (Meta AI); Changhan Wang (Facebook AI Research); Hongyu Gong (Meta AI); Elizabeth Salesky (Johns Hopkins University); Yossi Adi (Facebook AI Research); Ann Lee (Facebook, Inc.); Peng-Jen Chen (Meta AI)

3026 (SLT-P16.3): DECOUPLED NON-PARAMETRIC KNOWLEDGE DISTILLATION FOR END-TO-END SPEECH TRANSLATION

Hao Zhang (University of Information Engineering); Nianwen Si (University of Information Engineering); Yaqi Chen (Information Engineering University); Wen-Lin Zhang (National Digital Switching System Engineering and Technological R&D Center); Xukui Yang (ZZ Institute of Advance Technology); Dan Qu (National Digital Switching System Engineering and Technological R&D Center); Zhen Li (University of Information Engineering)

3135 (SLT-P16.4): JOINT PRE-TRAINING WITH SPEECH AND BILINGUAL TEXT FOR DIRECT SPEECH TO SPEECH TRANSLATION

Kun Wei (School of Computer Science, Northwestern Polytechnical University); Long Zhou (Microsoft Research Asia); Ziqiang Zhang (University of Science and Technology of China); LIPING CHEN (Microsoft); Shujie Liu (Microsoft Research Asia); Lei He (Microsoft Cloud and AI); Jinyu Li (Microsoft); Furu Wei (Microsoft Research Asia)

3822 (SLT-P16.5): LEAPT: Learning Adaptive Prefix-to-prefix Translation For Simultaneous Machine Translation

Lei Lin (Xiamen University); Shuangtao Li (Xiamen University); xiaodong shi (xiamen university)

Wednesday, June 7

3889 (SLT-P16.6): ENHANCING SPEECH-TO-SPEECH TRANSLATION WITH MULTIPLE TTS TARGETS

Jiatong Shi (Carnegie Mellon University); Yun Tang (Facebook); Ann Lee (Facebook, Inc.); Hirofumi Inaguma (Meta AI); Changhan Wang (Facebook AI Research); Juan Miguel Pino (Facebook); Shinji Watanabe (Carnegie Mellon University)

4196 (SLT-P16.7): Rethinking the Reasonability of the Test Set for Simultaneous Machine Translation

Mengge Liu (Beijing Institute of Technology); Wen Zhang (Xiaomi AI Lab); Xiang Li (Xiaomi AI Lab); Jian Luan (Xiaomi AI Lab); Bin Wang (Xiaomi AI Lab); Yuhang Guo (Beijing Engineering Research Center of High Volume Language Information Processing and Cloud Computing Applications, Department of Computer Science and Technology, Beijing Institute of Technology); Shuoying Chen (Beijing Institute of Technology)

4387 (SLT-P16.8): Align, Write, Re-order: Explainable End-to-End Speech Translation via Operation Sequence Generation

Motoi Omachi (Yahoo Japan Corporation); Brian Yan (Carnegie Mellon University); Siddharth Dalmia (Carnegie Mellon University); Yuya Fujita (Yahoo Japan Corporation); Shinji Watanabe (Carnegie Mellon University)

4983 (SLT-P16.9): Efficient Speech Translation with Dynamic Latent Perceivers

Ioannis Tsiamas (Universitat Politècnica de Catalunya (UPC)); Gerard Ion Gállego (Universitat Politècnica de Catalunya); José A. R. Fonollosa (Universitat Politècnica de Catalunya); Marta R. Costa-jussà (Meta AI)

5169 (SLT-P16.10): Joint Training And Decoding for Multilingual End-to-End Simultaneous Speech Translation

Wuwe Huang (Xiaomi Corporation); Renren Jin (Tianjin University); Wen Zhang (Xiaomi AI Lab); Jian Luan (Xiaomi AI Lab); Bin Wang (Xiaomi AI Lab); Deyi Xiong (Tianjin University)

5381 (SLT-P16.11): Enhancing Ontology Translation through Cross-Lingual Agreement

Mingjie Tian (School of Artificial Intelligence, Jilin University); Fausto Giunchiglia (University of Trento); Rui Song (School of Artificial Intelligence, Jilin University); Xing chen (Jilin University); Hao Xu (Jilin University)

6523 (SLT-P16.12): M3ST: MIX AT THREE LEVELS FOR SPEECH TRANSLATION

Xuxin Cheng (Peking University); Qianqian Dong (ByteDance); fengpeng yue (ByteDance); Tom Ko (Bytedance); Mingxuan Wang (Bytedance); Yuexian Zou (Peking University)

AASP-L5: Music Audio Synthesis and Modeling

Room: Salon de Roses A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Magdalena Fuentes, Kazuyoshi Yoshii

10:50 AM

6350 (AASP-L5.1): TransPlayer: Timbre Style Transfer with Flexible Timbre Control

Yuxuan Wu (Carnegie Mellon University); Yifan He (Carnegie Mellon University); Xinlu Liu (Carnegie Mellon University); Yi Wang (Carnegie Mellon University); Roger B. Dannenberg (School of Computer Science, Carnegie Mellon University)

11:05 AM

927 (AASP-L5.2): SYNTHESIZER PRESET INTERPOLATION USING TRANSFORMER AUTO-ENCODERS

Gwendal Le Vaillant (University of Mons / IRISIB (HE2B-ISIB)); Thierry Dutoit (University of Mons)

11:20 AM

1851 (AASP-L5.3): Can Knowledge of End-to-End Text-to-Speech Models Improve Neural MIDI-to-Audio Synthesis Systems?

Xuan Shi (University of Southern California); Erica Cooper (); Xin Wang (National Institute of Informatics); Junichi Yamagishi (National Institute of Informatics); Shrikanth Narayanan (USC)

11:35 AM

6209 (AASP-L5.4): Modelling black-box audio effects with time-varying feature modulation

Marco Comunita (Queen Mary University of London); Christian J. Steinmetz (Queen Mary University of London); Huy Phan (Amazon Alexa); Joshua D. Reiss (Queen Mary University of London)

11:50 AM

4516 (AASP-L5.5): Disentangling the Horowitz factor: Learning content and style from expressive piano performance

Huan Zhang (Queen Mary University of London); Simon Dixon (Queen Mary University of London)

12:05 PM

249 (AASP-L5.6): Compose & Embellish: Well-Structured Piano Performance Generation via A Two-Stage Approach

Shih-Lun Wu (National Taiwan University); Yi-Hsuan Yang (Academia Sinica)

10:50 AM

3890 (IVMSP-L5.1): CFFMixer: Multi-dimensional Feature Fusion For Object Detection

Hao Xie (Southeast University); weizhe yuan (Southeast University); Bin Kang (Nanjing University of Posts and Telecommunication); Songlin Du (Southeast University)

Wednesday, June 7

GC-5: Spoken Language Understanding Grand Challenge

Room: Nefeli B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): TBA

10:50 AM

6648 (GC-L5.1): Introduction

Akshat Shrivastava (Facebook); Suyoun Kim (Meta); Paden P Tomasello (Meta); Ali Elkahky (Meta); Daniel A Lazar (Meta); Trang Le (Meta); Shan Jiang (Meta); Duc Le (Meta); Aleksandr Livshits (Meta); Ahmed Aly (Meta)

11:10 AM

6864 (GC-L5.2): E-BRANCHFORMER-BASED E2E SLU TOWARD STOP ON-DEVICE CHALLENGE

Yosuke Kashiwagi (Sony); Siddhant Arora (Carnegie Mellon University); Hayato Futami (Sony Group Corporation); Jessica Huynh (Carnegie Mellon University); Shih-Lun Wu (Carnegie Mellon University); Yifan Peng (Carnegie Mellon University); Brian Yan (Carnegie Mellon University); Emiru Tsunoo (Sony Group Corporation); Shinji Watanabe (Carnegie Mellon University)

11:22 AM

6868 (GC-L5.3): A TWO-STAGE SYSTEM for SPOKEN LANGUAGE UNDERSTANDING

zhang gaosheng (transsion.com); shilei miao (传音控股); tang linghui (Transsion); qian peijia (Transsion)

11:34 AM

6879 (GC-L5.4): A Transformer-Based E2E SLU model for Improved Semantic Parsing

Othman Istaiteh (Samsung Research Jordan); Yasmeen Kussad (Samsung Research Jordan); Yahya Daqour (Samsung Research Jordan); Maria Habib (Samsung); Mohammad Habash (Samsung Research Jordan); Dhananjaya Gowda (Samsung Electronics)

11:46 AM

6885 (GC-L5.5): A Study on the Integration of Pipeline and E2E SLU systems for Spoken Semantic Parsing toward STOP Quality Challenge

Siddhant Arora (Carnegie Mellon University); Hayato Futami (Sony Group Corporation); Shih-Lun Wu (Carnegie Mellon University); Jessica Huynh (Carnegie Mellon University); Yifan Peng (Carnegie Mellon University); Yosuke Kashiwagi (Sony); Emiru Tsunoo (Sony Group Corporation); Brian Yan (Carnegie Mellon University); Shinji Watanabe (Carnegie Mellon University)

11:58 AM

6892 (GC-L5.6): The pipeline system of ASR and NLU with MLM-based data augmentation toward STOP low-resource challenge

Hayato Futami (Sony Group Corporation); Jessica Huynh (Carnegie Mellon University); Siddhant Arora (Carnegie Mellon University); Shih-Lun Wu (Carnegie Mellon University); Yosuke Kashiwagi (Sony); Yifan Peng (Carnegie Mellon University); Brian Yan (Carnegie Mellon University); Emiru Tsunoo (Sony Group Corporation); Shinji Watanabe (Carnegie Mellon University)

IVMSP-L5: Image Recognition and Detection

Room: Athina

Type: Oral

10:50 AM to 12:10 PM

Chair(s): Sheng Liu, Yu-Bin Yang

10:50 AM

3890 (IVMSP-L5.1): CFFMixer: Multi-dimensional Feature Fusion For Object Detection

Hao Xie (Southeast University); weizhe yuan (Southeast University); Bin Kang (Nanjing University of Posts and Telecommunication); Songlin Du (Southeast University)

11:05 AM

1242 (IVMSP-L5.2): SANet: Spatial Attention Network with Global Average Contrast Learning for Infrared Small Target Detection

Jiewen Zhu (UESTC); Shengjia Chen (University of Electronic Science and Technology of China); lexiao li (UESTC); Luping Ji (UESTC)

11:20 AM

736 (IVMSP-L5.3): LOGOVIT: LOCAL-GLOBAL VISION TRANSFORMER FOR OBJECT RE-IDENTIFICATION

Phan Nguyen (VinBrain); Ta Duc Huy (Vinbrain); Soan T. M. Duong (Le Quy Don Technical University); Nguyen Hoang Tran (VinBrain); Sam Bao Tran (Vinbrain); Dao Huu Hung (VinBrain); Chanh D Tr Nguyen (VinBrain); Trung Bui (Individual); QUOC HUNG TRUONG (VINBRAIN)

11:35 AM

319 (IVMSP-L5.4): ProContEXT: Exploring Progressive Context Transformer for Tracking

Jin-Peng Lan (DAMO Academy, Alibaba Group); Zhi-Qi Cheng (Carnegie Mellon University); Jun-Yan He (DAMO Academy, Alibaba Group); Chenyang Li (DAMO Academy, Alibaba Group); Bin Luo (DAMO Academy, Alibaba Group); Xu Bao (DAMO Academy, Alibaba Group); Wangmeng Xiang (DAMO Academy, Alibaba Group); Yifeng Geng (Alibaba Group); Xuansong Xie (DAMO Academy, Alibaba Group)

Wednesday, June 7

11:50 AM

3268 (IVMSP-L5.5): PAIR DETR: TOWARD FASTER CONVERGENT DETR

Seyed mehdi Iranmanesh (Amazon); Sherry X Chen (University of California, Santa Barbara); Kuo-Chin Lien (Appen)

12:05 PM

907 (IVMSP-L5.6): Data-aware Zero-shot Neural Architecture Search for Image Recognition

Yi Fan (State Key Laboratory for Novel Software Technology, Nanjing University); Zhong-Han Niu (State Key Laboratory for Novel Software Technology, Nanjing University); Yu-Bin Yang (State Key Laboratory for Novel Software Technology, Nanjing University)

SLT-L10: Multi-speaker ASR

Room: Delphi

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Reinhold Haeb-Umbach, Ozlem Kalinli

10:50 AM

393 (SLT-L10.1): VarArray Meets t-SOT: Advancing the State of the Art of Streaming Distant Conversational Speech Recognition

Naoyuki Kanda (Microsoft); Jian Wu (Microsoft); Xiaofei Wang (Microsoft); Zhuo Chen (Microsoft); Jinyu Li (Microsoft); Takuya Yoshioka (Microsoft)

11:05 AM

934 (SLT-L10.2): Simulating realistic speech overlaps improves multi-talker ASR

Muqiao Yang (Carnegie Mellon University); Naoyuki Kanda (Microsoft); Xiaofei Wang (Microsoft Corp.); Jian Wu (Microsoft); Sunit Sivasankaran (Microsoft); Zhuo Chen (Microsoft); Jinyu Li (Microsoft); Takuya Yoshioka (Microsoft)

11:20 AM

1639 (SLT-L10.3): Multi-speaker Data Augmentation for Improved End-to-end Automatic Speech Recognition

Samuel Thomas (IBM Research AI); Jeff Kuo (IBM); George Saon (IBM); Brian Kingsbury (IBM Research)

11:35 AM

1850 (SLT-L10.4): Anchored Speech Recognition with Neural Transducers

Desh Raj (Johns Hopkins University); Junteng Jia (Meta AI); Jay Mahadeokar (Meta AI); Chunyang Wu (Meta AI); Niko Moritz (Meta); Xiaohui Zhang (Meta); Ozlem Kalinli (Meta AI)

11:50 AM

4278 (SLT-L10.5): A Sidecar Separator Can Convert a Single-Talker Speech Recognition System to a Multi-Talker One

Lingwei Meng (The Chinese University of Hong Kong); Jiawen Kang (The Chinese University of Hong Kong); Mingyu Cui (The Chinese University of Hong Kong); Yuejiao Wang (The Chinese University of Hong Kong); Xixin Wu (The Chinese University of Hong Kong); Helen Meng (The Chinese University of Hong Kong)

12:05 PM

5295 (SLT-L10.6): Conditional Conformer: Improving Speaker Modulation for Single and Multi-User Speech Enhancement

Tom O'Malley (Google); Shaojin Ding (Google); Arun Narayanan (Google Inc.); Quan Wang (Google); Rajeev Rikhye (Google); Qiao Liang (Google Inc.); Yanzhang He (Google); Ian McGraw ()

SLT-L9: Multimodal Processing of Language and Language Systems II

Room: Jupiter

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Jan Yenda Trmal, Paola Garcia

10:50 AM

1158 (SLT-L9.1): PREFIX TUNING FOR AUTOMATED AUDIO CAPTIONING

Minkyu Kim (POSTECH); Kim Sung-Bin (POSTECH); Tae-Hyun Oh (POSTECH)

11:05 AM

1648 (SLT-L9.2): C2KD: Cross-Lingual Cross-Modal Knowledge Distillation for Multilingual Text-Video Retrieval

Andrew Rouditchenko (MIT CSAIL); Yung-Sung Chuang (MIT); Nina Shvetsova (Goethe University Frankfurt); Samuel Thomas (IBM Research AI); Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research); Brian Kingsbury (IBM Research); Leonid Karlinsky (IBM-Research); David Harwath (The University of Texas at Austin); Hilde Kuehne (Goethe University Frankfurt); James Glass (Massachusetts Institute of Technology)

11:20 AM

2096 (SLT-L9.3): The Edinburgh International Accents of English Corpus: Towards the Democratization of English ASR

Ramon R Sanabria (The University Of Edinburgh); Nikolay Bogoychev (The University Of Edinburgh); Nina Markl (University of Edinburgh); Andrea Carmantini (University of Edinburgh); Ondrej Klejch (University of Edinburgh); Peter Bell (University of Edinburgh)

Wednesday, June 7

11:35 AM

2768 (SLT-L9.4): Adaptive Knowledge Distillation between Text and Speech Pre-trained Models

Jinjie Ni (Nanyang Technological University); Yukun Ma (Alibaba Group); Wen Wang (Alibaba Group); Qian Chen (Speech Lab, DAMO Academy, Alibaba Group); Dianwen Ng (Alibaba Group/Nanyang Technological University); HAN LEI (Nanyang Technological University); Trung Hieu Nguyen (Alibaba Group); Chong Zhang (Alibaba Group); Bin Ma ("Alibaba, Singapore R&D Center"); Erik Cambria (Nanyang Technological University, Singapore)

11:50 AM

6140 (SLT-L9.5): A processing framework to access large quantities of whispered speech found in ASMR

Pablo Pérez Zarazaga (KTH Royal Institute of Technology); Gustav Eje Henter (KTH Royal Institute of Technology); Zofia Malisz (KTH Royal Institute of Technology)

SPTM-L5: Tracking

Room: Nafsika A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Yunpeng Li, Wenjing Yang

10:50 AM

2380 (SPTM-L5.1): ON TRACKING A STOCHASTICALLY TIME-VARYING SUBSPACE

Victor Solo (University of New South Wales)

11:05 AM

1006 (SPTM-L5.2): ROBUST MULTI-OBJECT TRACKING WITH SPATIAL UNCERTAINTY

Pin-Jie Liao (National Tsing Hua university); Yu-Cheng Huang (National Tsing Hua University); Chen-Kuo Chiang (National Chung Cheng University); Shang-Hong Lai (National Tsing Hua University)

11:20 AM

1284 (SPTM-L5.3): GaPP: Multi-Target Tracking with Gaussian Processes

Alexander F Goodyer (University of Cambridge); Bashar I. Ahmad (University of Cambridge); Simon Godsill (Department of Engineering, University of Cambridge)

11:35 AM

3538 (SPTM-L5.4): Possibilistic Bernoulli Filter for Extended Target Tracking

Zhijin Chen (RMIT University); Branko Ristic (RMIT University); Du Yong Kim (RMIT University)

11:50 AM

5686 (SPTM-L5.5): Particle Flow Gaussian Sum Particle Filter

Karthik Comandur (Signal Processing and Communication Research Centre, IIT Hyderabad); Yunpeng Li (University of Surrey); Santosh Nannuru (IIT Hyderabad)

12:05 PM

5912 (SPTM-L5.6): Progressive Perception Learning for Distribution Modulation in Siamese Tracking

Kun Hu (National University of Defense Technology); Xianchen Zhou (National University of Defense Technology); Mingyu Cao (NUDT); Mengzhu Wang (NUDT); Guangjie Gao (NUDT); Wenjing Yang (National University of Defense Technology); Huibin Tan (NUDT)

SS-L25: Radar-Assisted Perception (RAP)

Room: Salon de Roses B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Pu Wang,

10:50 AM

291 (SS-L25.1): Exploiting Virtual Array Diversity For Accurate Radar Detection

Junfeng Guan (UIUC); Sohrab Madani (UIUC); Waleed Ahmed (UIUC); Samah Ahmed Hussein (EPFL); Saurabh Gupta (UIUC); Haitham Z Alhassanieh (EPFL)

11:05 AM

609 (SS-L25.2): ST-MVDNet++: Improve Vehicle Detection with Lidar-Radar Geometrical Augmentation via Self-Training

Yu-Jhe Li (Carnegie Mellon University); Matthew O'Toole (Carnegie Mellon University); Kris Kitani (Carnegie Mellon University)

11:20 AM

1203 (SS-L25.3): Graph Neural Networks for Object Type Classification Based on Automotive Radar Point Clouds and Spectra

Loveneet Saini (Room 28); Axel Acosta (Bosch); Gor Hakobyan (Bosch)

Wednesday, June 7

11:35 AM

1422 (SS-L25.4): Fast 3D Human Pose Estimation Using RF Signals

Cong Yu (University of Electronic Science and Technology of China); Dongheng Zhang (University of Science and Technology of China); Zhi Wu (University Of Science And Technology Of China); Chunyang Xie (University of Electronic Science and Technology of China); Zhi Lu (University of Science and Technology of China); Yang Hu (University of Science and Technology of China); Yan Chen (University of Science and Technology of China)

11:50 AM

3467 (SS-L25.5): SPATIAL-DOMAIN OBJECT DETECTION UNDER MIMO-FMCW AUTOMOTIVE RADAR INTERFERENCE

Sian Jin (Princeton University); Pu Wang (MERL); Petros Boufounos (Mitsubishi Electric Research Laboratories); Ryuhei Takahashi (Mitsubishi Electric Information Technology R&D Center); Sumit Roy (University of Washington)

12:05 PM

4666 (SS-L25.6): Online Learning-based Waveform Selection for Improved Vehicle Recognition in Automotive Radar

Charles E Thornton (Virginia Tech); William Howard (Virginia Tech); Michael R. Buehrer (Virginia Tech, USA)

SS-L6: Data Driven and Machine Learning based Room Acoustic Modeling

Room: Nefeli A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Enzo De Sena, Paul Calamia, Wenyu Jin

10:50 AM

2532 (SS-L6.1): Towards Improved Room Impulse Response Estimation for Speech Recognition

Anton J Ratnarajah (University of Maryland, College Park); Ishwarya Ananthabhotla (Reality Labs Research at Meta, Redmond, WA); Vamsi Krishna Ithapu (Reality Labs Research at Meta, Redmond, WA); Pablo Hoffmann (Reality Labs Research at Meta, Redmond, WA); Dinesh Manocha (University of Maryland at College Park); Paul Calamia (Reality Labs Research at Meta, Redmond, WA)

11:05 AM

4315 (SS-L6.2): Room Impulse Response Reconstruction Based on Spatio-Temporal-Spectral Features Learned from a Spherical Microphone Array Measurement

AMY BASTINE (THE AUSTRALIAN NATIONAL UNIVERSITY); thushara abhayapala (The Australian National University); Jihui (Aimee) Zhang (University of Southampton)

11:20 AM

4415 (SS-L6.3): Contrastive Representation Learning for Acoustic Parameter Estimation

Philipp Goetz (International Audio Laboratories Erlangen); Cagdas Tuna (Fraunhofer Institute for Integrated Circuits IIS); Andreas Walther (Fraunhofer Institute for Integrated Circuits IIS); Emanuel Habets (AudioLabs Erlangen)

11:35 AM

4799 (SS-L6.4): Interpolation of spatial room impulse responses using partial optimal transport

Aaron Geldert (Aalto University); Niils Meyer-Kahlen (Aalto University); Sebastian J Schlecht (Aalto University)

11:50 AM

4884 (SS-L6.5): Simultaneous Acoustic Echo Sorting and 3-D Room Geometry Inference

Kathleen C MacWilliam (Department of Electrical Engineering (ESAT-STADIUS/ETC)); Filip Elvander (Aalto University); Toon van Waterschoot (Department of Electrical Engineering (ESAT-STADIUS/ETC))

12:05 PM

4967 (SS-L6.6): Blind Acoustic Room Parameter Estimation Using Phase Features

Christopher A Ick (New York University); Adib Mehrabi (Sonos Experience Limited); Wenyu Jin (Sonos, Inc.)

ASPS-P4: Sensing Applications

Room: Dome A

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Ranadip Pal,

5600 (ASPS-P4.1): Improved Belief Propagation Decoding of Turbo Codes

Yifei Shen (EPFL); Yuqing Ren (EPFL); Andreas Toftegaard Kristensen (Ecole Polytechnique Federale de Lausanne (EPFL)); Xiaohu You (Southeast University); Chuan Zhang (Southeast University); Andreas Burg (EPFL)

286 (ASPS-P4.2): mmSense: Detecting Concealed Weapons with a Miniature Radar Sensor

Kevin Mitchell (University of Glasgow); Khaled Kasseem (University of Glasgow); Chaitanya Kaul (University of Glasgow); Valentin Kapitany (University of Glasgow); Philip Binner (University of Glasgow); Andrew Ramsay (University of Glasgow); Daniele Faccio (University of Glasgow); Roderick Murray-Smith (University of Glasgow)

Wednesday, June 7

6301 (ASPS-P4.3): Multiple Target Measurements: Bayesian Framework for Moving Object Detection in MIMO Radar
Bastian Eisele (Friedrich-Alexander-Universität Erlangen-Nürnberg); Ali Bereyhi (Friedrich-Alexander-Universität Erlangen-Nürnberg); Ralf Müller (Friedrich-Alexander-Universität Erlangen-Nürnberg)

4116 (ASPS-P4.4): A Momentum Two-gradient Direction Algorithm with Variable Step Size Applied to Solve Practical Output Constraint Issue for Active Noise Control

Xiaoyi Shen (Nanyang Technological University); Dongyuan Shi (Nanyang Technological University); Zhengding Luo (Nanyang Technological University); Junwei Ji (Nanyang Technological University); Woon Seng Gan (NTU)

5231 (ASPS-P4.5): TEFISTA-NET: GTD PARAMETER ESTIMATION OF LOW-FREQUENCY ULTRA-WIDEBAND RADAR VIA MODEL-BASED DEEP LEARNING

Rui Li (Tsinghua University); Xueqian Wang (Tsinghua University); Gang Li (Tsinghua University); Xiao-Ping Zhang (Toronto Metropolitan University)

5441 (ASPS-P4.6): Enhancing the Accuracy of Resistive In-memory Architectures using Adaptive Signal Processing
Han-Mo Ou (University of Illinois Urbana-Champaign); Naresh Shanbhag (University of Illinois at Urbana-Champaign)

1171 (ASPS-P4.7): Improved Indoor Localization With NLOS Signal Propagations

Wei Huang (Southwest University); Yixin Zhao (Southwest University); Xuechao Wu (Southwest University); Le Yin (Southwest University)

2355 (ASPS-P4.8): Parallel 2D Seismic Ray Tracing using CUDA on a Jetson Nano

Ban-Sok Shin (German Aerospace Center); Luis Wientgens (German Aerospace Center); Dmitriy Shutin (DLR)

5009 (ASPS-P4.9): Unlimited Sampling Radar: Life Below the Quantization Noise

Thomas Feuillen (Imperial College London); Bhavani Shankar Mysore Ramarao (University of Luxembourg); Ayush Bhandari (Imperial College London)

4820 (ASPS-P4.10): Recursive/Iterative unique Projection-Aggregation decoding of Reed-Muller codes

Marzieh Hashemipour-Nazari (Eindhoven University of Technology); Renate Debets (Eindhoven University of Technology); Kees Goossens (Eindhoven University of Technology); Alexios Balatsoukas-Stimming (Eindhoven University of Technology)

6094 (ASPS-P4.11): Optimization of Sensor Configurations for Fault Identification in Smart Buildings

Naveed Ahmad (INSA Lyon); Malcolm Egan (INRIA); Jean-Marie Gorce (INSA Lyon); Jilles Steeve Dibangoye (INSA Lyon, INRIA); Frederic Le-Mouel (INSA Lyon)

3320 (ASPS-P4.12): FedAudio: A Federated Learning Benchmark for Audio Tasks

Tuo Zhang (University of Southern California); Tiantian Feng (University of Southern California); Samiul Alam (Michigan State University); Sunwoo Lee (Inha University); Mi Zhang (The Ohio State University); Shrikanth Narayanan (University of Southern California); Salman Avestimehr (University of Southern California)

CI-P1: Computational Imaging II

Room: Dome L

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Bo Zhao,

1042 (CI-P1.1): Unrolled Fourier Disparity Layer optimization for scene reconstruction from few-shots focal stacks

Brandon Le Bon (Centre INRIA de l'Université de Rennes); Mikaël Le Pendu (InterDigital, Rennes); Christine Guillemot (INRIA)

1641 (CI-P1.2): A Targeted Sampling Strategy for Compressive Cryo Focused Ion Beam Scanning Electron Microscopy

Daniel Nicholls (University of Liverpool); Jack Wells (University of Liverpool); Alex W Robinson (University of Liverpool); Amirafshar Moshtaghpour (Rosalind Franklin Institute); Maryna Kobylenska (King's College London); Roland Fleck (King's College London); Angus Kirkland (University of Oxford); Nigel Browning (University of Liverpool)

1811 (CI-P1.3): Hardware Friendly Spline Sketched Lidar

Michael Sheehan (University of Edinburgh); Julián Tachella (CNRS & ENS de Lyon); Mike Davies (University of Edinburgh)

2419 (CI-P1.4): DMSA: DYNAMIC MULTI-SCALE UNSUPERVISED SEMANTIC SEGMENTATION BASED ON ADAPTIVE AFFINITY

Kun Yang (Heilongjiang University); Jun Lu (Heilongjiang University)

3104 (CI-P1.5): Self-Supervised Learning with Explorative knowledge Distillation

Tongtong Su (Nankai University); Jinsong Zhang (Nankai University); Wang Gang (Nankai University); Liu Xiaoguang (Nankai University)

3245 (CI-P1.6): Joint Neural Representation for Multiple Light Fields

Guillaume Le Guludec (Inria); Christine Guillemot (INRIA)

Wednesday, June 7

4937 (CI-P1.7): An Edge Alignment-based Orientation Selection Method for Neutron Tomography

Diyu Yang (Purdue University); Shimin Tang (Oak Ridge National Laboratory); Singanallur Venkatakrishnan (Oak Ridge National Laboratory); Mohammad Samin Nur Chowdhury (Purdue University); Yuxuan Zhang (Oak Ridge National Laboratory); Hassina Bilheux (Oak Ridge National Laboratory); Gregory T Buzzard (Purdue University); Charles Bouman (Purdue University)

5111 (CI-P1.8): FACTORIZED PROJECTION-DOMAIN SPATIO-TEMPORAL REGULARIZATION FOR DYNAMIC TOMOGRAPHY

Berk Iskender (University of Illinois at Urbana-Champaign); Marc L Klasky (Los Alamos National Laboratory); Brian M Patterson (Los Alamos National Laboratory); Yoram Bresler (UIUC)

5629 (CI-P1.9): Alternating Phase Langevin Sampling with Implicit Denoiser Priors for Phase Retrieval

Rohun Agrawal (California Institute of Technology); Oscar Leong (California Institute of Technology)

435 (CI-P1.10): Single-Shot Fractional Fourier Phase Retrieval

Yixiao Yang (Beijing Institute of Technology); Ran Tao (Beijing Institute of Technology)

6695 (CI-P1.11): Stability of image-reconstruction algorithms (SPS Journal Paper)*

Pol del Aguila Pla (EPFL - CIBM); Sebastian Neumayer (EPFL); Michael Unser (EPFL)

6697 (CI-P1.12): Coded Illumination for Improved Lensless Imaging (SPS Journal Paper)*

Yucheng Zheng (University of California, Riverside); M. Salman Asif (University of California, Riverside)

IVMSP-P17: Anomaly Detection

Room: Dome J

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Zhenhua Guo, Zhiqiang Wu

998 (IVMSP-P17.1): A Physically Explainable Framework for Human-Related Anomaly Detection

Yalong Jiang (Beihang University); huining Li (Beihang University); changkang li (Beihang University)

1267 (IVMSP-P17.2): Two-stream Decoder Feature Normality Estimating Network for Industrial Anomaly Detection

Chaewon Park (Yonsei University); Minhyeok Lee (Yonsei University); Suhwan Cho (Yonsei University); Donghyeong Kim (Yonsei University); Sangyoun Lee (Yonsei University)

1794 (IVMSP-P17.3): FAPM: Fast Adaptive Patch Memory for Real-time Industrial Anomaly Detection

Donghyeong Kim (Yonsei University); Chaewon Park (Yonsei University); Suhwan Cho (Yonsei University); Sangyoun Lee (Yonsei University)

2270 (IVMSP-P17.4): Spatial-Temporal Graph Convolutional Network boosted Flow-Frame Prediction for Video Anomaly Detection

Kai Cheng (Fudan University); Xinhua Zeng (Fudan University); Yang Liu (Fudan University); Mengyang Zhao (FUDAN University); pang chengxin (Shanghai University of Electric Power); Xing Hu (university of shanghai for science and technology)

2446 (IVMSP-P17.5): LOW-RANK CONSTRAINED MEMORY AUTOENCODER FOR HYPERSPECTRAL ANOMALY DETECTION

yuyun lian (China University of Geosciences); Yongshan Zhang (China University of Geosciences); Xuxiang Feng (Chinese Academy of Sciences); Xinwei Jiang (China University of Geosciences); Zhihua Cai (China University of Geosciences)

2714 (IVMSP-P17.6): A Two-branch Network for Video Anomaly Detection with Spatio-temporal Feature Learning

Guoqiu Li (Tsinghua Shenzhen International Graduate School, Tsinghua University); Shengjie Chen (Tsinghua University); Yujiu Yang (Tsinghua University); Zhenhua Guo (Tianyi Traffic Technology)

3014 (IVMSP-P17.7): SCOREFORMER: SCORE FUSION-BASED TRANSFORMERS FOR WEAKLY-SUPERVISED VIOLENCE DETECTION

Yang Xiao (Xinjiang University); Liejun Wang (Xinjiang University); Tongguan Wang (Xinjiang University); Huicheng Lai (Xinjiang University)

4544 (IVMSP-P17.8): Synthetic Pseudo Anomalies for Unsupervised Video Anomaly Detection: A Simple yet Efficient Framework based on Masked Autoencoder

Xiangyu Huang (School of informatics Xiamen University); Caidan Zhao (School of Informatics Xiamen University); Chenxing Gao (xiamen university); Chen Lvdong (xiamen university); Zhiqiang Wu (Wright State University)

4607 (IVMSP-P17.9): A Video Anomaly Detection Framework based on Appearance-Motion Semantics Representation Consistency

Xiangyu Huang (School of informatics Xiamen University); Caidan Zhao (School of Informatics Xiamen University); Zhiqiang Wu (Wright State University)

6383 (IVMSP-P17.10): SSGD: A smartphone screen glass dataset for defect detection

Haonan Han (Tsinghua University); Rui Yang (Tsinghua University); SHUYAN LI (University of Cambridge); Runze Hu (Beijing Institute of Technology); Xiu Li (Tsinghua University)

Wednesday, June 7

107 (IVMSP-P17.11): LEARNABLE FLOW MODEL CONDITIONED ON GRAPH REPRESENTATION MEMORY FOR ANOMALY DETECTION

ziyu zhu (Tsinghua University); Wenlei Liu (Tsinghua University); ZHIDONG Deng (Tsinghua University)

498 (IVMSP-P17.12): A parallel attention mechanism for image manipulation detection and localization

Qiang Zeng (Sichuan University); Hongxia Wang (Sichuan University); yang zhou (sichuan university); Rui Zhang (Sichuan University); Sijiang Meng (Sichuan University)

IVMSP-P18: Deep Neural Network

Room: Dome K

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Yu-Bin Yang, Lizhuang Ma

748 (IVMSP-P18.1): A Lightweight Convolutional Neural Network Using Feature Filtering Module

Nan Jing (Inner Mongolia University); Yu Zhang (Inner Mongolia University)

1478 (IVMSP-P18.2): CLMAE: a lighter and faster Masked Autoencoders

Yiran Song (Shanghai Jiao Tong University); Lizhuang Ma (Shanghai Jiao Tong University)

3192 (IVMSP-P18.3): IMPLICITLY ROTATION EQUIVARIANT NEURAL NETWORKS

Naman Khetan (IIT (ISM) Dhanbad); Tushar Arora (IIT (ISM) Dhanbad); Samee Ur Rehman (Transmute AI); Deepak K Gupta (UiT The Arctic University of Norway)

4440 (IVMSP-P18.4): Learning how to learn domain-invariant parameters for domain generalization

Feng Hou (University of Chinese Academy of Sciences); Yao Zhang (Shanghai AI Lab); Yang Liu (Institute of Computing Technology, University of Chinese Academy of Sciences, Lenovo AI Lab); Jin Yuan (Southeast University); Cheng Zhong (Lenovo Research, AI Lab); Yang Zhang (Lenovo Ltd); zhongchao shi (lenovo company); Jianping Fan (Lenovo); Zhiqiang He (Lenovo Ltd.)

4464 (IVMSP-P18.5): LEARNING ON ENTROPY CODED IMAGES WITH CNN

Rémi Piau (INRIA); Thomas Maugey (INRIA); Aline Roumy (INRIA)

1983 (IVMSP-P18.6): Look and Think: Intrinsic Unification of Self-attention and Convolution for Spatial-Channel Specificity

Xiang Gao (South China University of Technology); Honghui Lin (South China University of Technology); Yu Li (South China University of Technology); Ruiyan Fang (South China University of Technology); Xin Zhang (South China University of Technology)

2489 (IVMSP-P18.7): MULTIPLE DOMAIN-ADVERSARIAL ENSEMBLE LEARNING FOR DOMAIN GENERALIZATION

Ze-Yu Mi (Nanjing university); Kun Long (State Key Laboratory for Novel Software Technology, Nanjing University); Yu-Bin Yang (State Key Laboratory for Novel Software Technology, Nanjing University)

1568 (IVMSP-P18.8): Contrastive Domain Adaptation via Delimitation Discriminator

Xing Wei (Hefei University of Technology); bin wen (Hefei University of Technology); Lei Chen (Institute of Intelligent Machines, HFIPS, Chinese Academy of Sciences); Yujie Liu (Hefei University of Technology); Chong Zhao (HeFei University of Technology); Yang Lu (Hefei University of Technology)

2273 (IVMSP-P18.9): Nasty-SFDA: Source Free Domain Adaptation from A Nasty Model

Jiajiong Cao (Ant Financial Service Group); Yufan Liu (Institute of Automation, Chinese Academy of Sciences); Weiming Bai (Chinese Academy of Sciences); Jingting Ding (Ant Financial); Liang Li (Ant Financial Service Group)

7163 (IVMSP-P18.10): On The Relationship Between Universal Adversarial Attacks And Sparse Representations

Dana Weitzner (Tel Aviv University); Raja Giryes (Tel Aviv University)

MLSP-P17: Deep Learning II

Room: Dome F

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

866 (MLSP-P17.1): Surrogate Based Post-hoc Calibration for Distributional Shift

Jun Zhang (Tsinghua University; National Innovation Institute of Defense Technology, Chinese Academy of Military Science)

5891 (MLSP-P17.2): Training Robust Spiking Neural Networks with ViewPoint Transform and SpatioTemporal Stretching

Haibo Shen (Huazhong University of Science and Technology); Juyi Xiao (Huazhong University of Science and Technology); Yihao Luo (Yichang Testing Technique R&D Institute); Xiang Cao (School of Computer Science and Technology, Huazhong University of Science and Technology); Liangqi Zhang (Huazhong University of Science and Technology); Tianjiang Wang (School of Computer Science and Technology, Huazhong University of Science and Technology)

Wednesday, June 7

3380 (MLSP-P17.3): Improving Electric Load Demand Forecasting with Anchor-based Forecasting Method

Maria Tzelepi (Aristotle University of Thessaloniki); Paraskevi Nousi (Aristotle University of Thessaloniki); ANASTASIOS TEFAS (Aristotle University of Thessaloniki)

4117 (MLSP-P17.4): GENERAL CATEGORY NETWORK: HANDWRITTEN MATHEMATICAL EXPRESSION RECOGNITION WITH COARSE-GRAINED RECOGNITION TASK

Xinyu Zhang (Nanjing University); Han Ying (Nanjing University); Ye Tao (Nanjing University); Youlu Xing (Nanjing University); Guihuan Feng (Nanjing University)

6021 (MLSP-P17.5): Deep architecture for doa trajectory localization.

Shreyas Jaiswal (SPCRC, IIT Hyderabad); Ruchi Pandey (IIT Hyderabad); Santosh Nannuru (IIT Hyderabad)

1135 (MLSP-P17.6): Conditional LS-GAN based Skyliht Polarization Image Restoration and Application in Meridian Localization

Tian Yang (Hefei University of Technology); Hongbo Bo (University of Bristol); Xinyu Yang (Lancaster University); Jun Gao (Hefei University of Technology); Zijian Shi (University of Bristol)

1864 (MLSP-P17.7): A Perturbation-based Policy Distillation Framework with Generative Adversarial Nets

LiHua Zhang (School of Computer Science and Technology, Soochow University); Quan Liu (School of Computer Science and Technology, Soochow University); Zhang Xiongzheng (School of Computer Science and Technology, Soochow University, Suzhou, China); Yapeng Xu (School of Computer Science and Technology, Soochow University)

2375 (MLSP-P17.8): Maximum Likelihood Distillation for Robust Modulation Classification

Javier J Maroto (EPFL); G r me Bovet (armasuisse Science & Technology); Pascal Frossard (EPFL)

2117 (MLSP-P17.9): Batch-Ensemble Stochastic Neural Networks for Out-of-Distribution Detection

Xiongjie Chen (University of Surrey); Yunpeng Li (University of Surrey); Yongxin Yang (Queen Mary University of London)

1076 (MLSP-P17.10): Building Blocks for a Complex-Valued Transformer Architecture

Florian Eilers (University of M nster); Xiaoyi Jiang (University of M nster)

204 (MLSP-P17.11): Explicit and Implicit Knowledge Distillation via Unlabeled Data

Yuzheng Wang (Fudan University); zuhao ge (fudan university); Zhaoyu Chen (Fudan University); Xian Liu (Fudan University); Chuangjia Ma (Fudan University); Yunquan Sun (Fudan University); Lizhe Qi (Fudan University)

1701 (MLSP-P17.12): Sparse Mixture Once-for-all Adversarial Training for Efficient In-Situ Trade-Off Between Accuracy and Robustness of DNNs

Souvik Kundu (University of Southern California); Sairam Sundaresan (Intel AI Lab); Sharath Nittur Sridhar (Intel AI Lab); SHUNLIN LU (The Chinese University of Hong Kong); han Tang (University of Southern California); Peter A. Beerel (University of Southern California)

MLSP-P18: Deep and Sequential Learning

Room: Dome G

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

4224 (MLSP-P18.1): ENLIGHTENING THE STUDENT IN KNOWLEDGE DISTILLATION

Yujie Zheng (Ningbo University); Chong Wang (Ningbo University); Yi Chen (Ningbo University); Jiangbo Qian (Ningbo University); Jun Wang (China University of Mining and Technology); JIAFEI WU (SenseTime Research)

103 (MLSP-P18.2): Overcoming the Seesaw in Monocular 3D Object Detection via Language Knowledge Transferring

Weichen Xu (Peking University); Tianhao Fu (Peking University)

1597 (MLSP-P18.3): Focusing On Targets For Improving Weakly Supervised Visual Grounding

Viet-Quoc Pham (Toshiba Research and Development Center); Nao Mishima (Toshiba Research and Development Center)

167 (MLSP-P18.4): Energy Regularized RNNs for Solving Non-Stationary Bandit Problems

Michael Rotman (Tel Aviv University); Lior Wolf (Tel Aviv University, Israel)

2285 (MLSP-P18.5): Transformer-based tracking Network for Maneuvering Targets

yushu zhang (Tsinghua University); Gang Li (Tsinghua University); Xiao-Ping Zhang (Toronto Metropolitan University); You He (Tsinghua University)

2598 (MLSP-P18.6): MENDAM: Multi-Expert Network with Distribution-Aware Momentum for Long-Tailed Recognition

Qingheng Zhang (Nanjing University of Aeronautics and Astronautics); Haibo Ye (Nanjing University of Aeronautics and Astronautics); Kaicheng Yu (Alibaba Inc.)

Wednesday, June 7

4064 (MLSP-P18.7): Frequency and Scale Perspectives of Feature Extraction

Liangqi Zhang (Huazhong University of Science and Technology); Yihao Luo (Yichang Testing Technique R&D Institute); Xiang Cao (School of Computer Science and Technology, Huazhong University of Science and Technology); Haibo Shen (Huazhong University of Science and Technology); Tianjiang Wang (School of Computer Science and Technology, Huazhong University of Science and Technology)

4915 (MLSP-P18.8): GAITMIXER: SKELETON-BASED GAIT REPRESENTATION LEARNING VIA WIDE-SPECTRUM MULTI-AXIAL MIXER

Ekkasit Pinyoanuntapong (University of North Carolina at Charlotte); Ayman Ali (UNCC); Pu Wang (UNCC); Minwoo Lee (University of North Carolina at Charlotte); Chen Chen (University of Central Florida)

815 (MLSP-P18.9): KalmanBOT: KalmanNet and Bollinger Bands based Learned Trader for Pairs Trading

Haoran Deng (ETH Zürich); Guy Revach (ETH Zürich); Hai Morgenstern (BeyondMinds); Nir Shlezinger (Ben-Gurion University)

3875 (MLSP-P18.10): On the Value of Stochastic Side Information in Online Learning

Junzhang Jia (University of Melbourne); Xuetong Wu (University of Melbourne); Jamie S Evans (University of Melbourne); Jingge Zhu (University of Melbourne)

861 (MLSP-P18.11): BIOLOGICALLY-INSPIRED CONTINUAL LEARNING OF HUMAN MOTION SEQUENCES

Joachim C Ott (ETH Zurich); Shih-Chii Liu (Institute of Neuroinformatics)

2065 (MLSP-P18.12): TriCL: Triplet Continual Learning

Xianchao Zhang (Dalian University of Technology); Guanglu Wang (Dalian University of Technology); Xiaotong Zhang (School of Software, Dalian University of Technology); Han Liu (Dalian University of Technology); Zhengxi Yin (Huawei Technologies Co. Ltd); Wentao Yang (Dalian University of Technology)

MLSP-P19: Machine learning for time series analysis II

Room: Dome H

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Tommy S. Alstrøm,

386 (MLSP-P19.1): Preformer: Predictive Transformer with Multi-Scale Segment-wise Correlations for Long-Term Time Series Forecasting

Dazhao Du (Institute of Software Chinese Academy of Sciences); Bing Su (Renmin University of China); Zhewei Wei (Renmin University of China)

3902 (MLSP-P19.2): Smoothing Point Adjustment-based Evaluation of Time Series Anomaly Detection

Mingyu Liu ("National University of Defense Technology, China"); Yijie Wang ("National University of Defense Technology, China"); Hongzuo Xu (National University of Defense Technology); Xiaohui Zhou (National University of Defense Technology); Bin Li (National University of Defense Technology); Yongjun Wang (College of Computer, National University of Defense Technology)

4384 (MLSP-P19.3): Investigating SINDy As a Tool For Causal Discovery In Time Series Signals

Andrew O'Brien (Drexel University); Rosina Weber (Drexel University); Edward Kim (Drexel University)

6300 (MLSP-P19.4): SyncNet: correlating objective for time delay estimation in audio signals

Akshay Raina (Indian Institute of Technology Kanpur); Vipul Arora (IIT Kanpur)

1698 (MLSP-P19.5): Towards Diverse and Coherent Augmentation for Time-Series Forecasting

Xiyuan Zhang (University of California, San Diego); Ranak Roy Chowdhury (University of California, San Diego); Jingbo Shang (University of California, San Diego); Rajesh Gupta (UC San Diego); Dezhi Hong (UC San Diego)

2747 (MLSP-P19.6): CoRe: Transferable Long-Range Time Series Forecasting Enhanced by Covariates-Guided Representation

Xin-Yi Li (State Key Laboratory for Novel Software Technology, Nanjing University); Pei-Nan Zhong (General Development Dept, Huawei Technologies Co. Ltd.); Di Chen (State Key Laboratory for Novel Software Technology, Nanjing University); Yu-Bin Yang (State Key Laboratory for Novel Software Technology, Nanjing University)

3750 (MLSP-P19.7): Does Your Model Think Like an Engineer? Explainable AI for Bearing Fault Detection with Deep Learning

Thomas Decker (Siemens AG and Ludwig Maximilians University); Michael Lebacher (Siemens AG); Volker Tresp (Siemens AG and Ludwig Maximilian University of Munich)

3759 (MLSP-P19.8): Multi-layer Seasonal Perception Network for Time Series Forecasting

Ruoshu Wang (Engineering Research Center of Cyberspace, Yunnan University); Shengfa Miao (Yunnan University); Di Liu (Yunnan University); Xin Jin (Yunnan University); Weisheng Zhang (Yunnan University)

Wednesday, June 7

1810 (MLSP-P19.9): Aleatoric Uncertainty Estimation of Overnight Sleep Statistics through Posterior Sampling using Conditional Normalizing Flows

Hans van Gorp (Eindhoven University of Technology); Merel M. van Gilst (Eindhoven University of Technology); Pedro Fonseca (Philips Research); Sebastiaan Overeem (Eindhoven University of Technology); Ruud J. G. van Sloun (Technical university of Eindhoven)

1517 (MLSP-P19.10): SEQUENTIAL DATUM-WISE JOINT FEATURE SELECTION AND CLASSIFICATION IN THE PRESENCE OF EXTERNAL CLASSIFIER

Sachini Piyoni Ekanayake (University at Albany SUNY); Daphney-Stavroula Zois (University at Albany); Charalampos Chelmis (University at Albany)

4044 (MLSP-P19.11): DEEP AUTOENCODING ONE-CLASS TIME SERIES ANOMALY DETECTION

Xudong Mou (Beihang University); Rui Wang (Beihang University); tiejun wang (BeiHang University); Jie Sun (Beihang University); Bo Li (Beihang University); Tianyu Wo (Beihang University); Xudong Liu (Beihang University)

4400 (MLSP-P19.12): MULTI-RESOLUTION SEQUENCE AGGREGATION AND MODEL AGNOSTIC FRAMEWORK FOR TIME SERIES FORECASTING

Juhyun Lyu (LG AI Research); Jinseok Yang (LG AI Research); Junghee Kim (LG AI Research); Woohyung Lim (LG AI Research); Wonbin Ahn (LG AI Research); Dongwan Kang (LG AI Research); Minjae Kim (LG AI Research); Nam Soo Kim (Seoul National University)

MLSP-P20: Machine learning for time series analysis III

Room: Dome M

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Ercan E Kuruoglu,

5498 (MLSP-P20.1): ONLINE CACHING WITH FETCHING COST FOR ARBITRARY DEMAND PATTERN: A DRIFT-PLUS-PENALTY APPROACH

Shashank P (IIT Dharwad); Bharath Bettagere (IIT Dharwad)

5828 (MLSP-P20.2): Sinusoidal Frequency Estimation by Gradient Descent

Ben Hayes (Queen Mary University of London); Charalampos Saitis (Queen Mary University of London); Gyorgy Fazekas (Queen Mary University of London)

2745 (MLSP-P20.3): Robust Time Series Recovery and Classification Using Test-Time Noise Simulator Networks

Eun Som Jeon (Arizona State University); Suhas Lohit (Mitsubishi Electric Research Laboratories); Rushil Anirudh (Lawrence Livermore National Laboratory); Pavan Turaga (Arizona State University)

1868 (MLSP-P20.4): Class-incremental learning on multivariate time series via shape-aligned temporal distillation

Zhongzheng Qiao (Nanyang Technological University); Minghui Hu (Nanyang Technological University); Xudong Jiang (Nanyang Technological University); Ponnuthurai Suganthan (Nanyang Technological University); Ramasamy Savitha (I2R ASTAR)

1254 (MLSP-P20.5): Memory-Augmented U-Transformer for Multivariate Time Series Anomaly Detection

Shuxin Qin (Purple Mountain Laboratories); Yongcan Luo (Purple Mountain Laboratories); Gaofeng Tao (Purple Mountain Laboratories)

4871 (MLSP-P20.6): Anomalous signal detection for cyber-physical systems using interpretable causal neural network

Shuo Zhang (East China Normal University); Jing Liu (East China Normal University)

2149 (MLSP-P20.7): Dynamic Vehicle Graph Interaction for Trajectory Prediction based on Video Signals

Jian Chen (Sun Yat-sen University); Wei Wang (Shenzhen MSU-BIT University); Junxin Chen (Dalian University of Technology); Ming Cai (School of Engineering, Sun Yat-sen University)

2231 (MLSP-P20.8): SDG-L: A Semiparametric Deep Gaussian Process based Framework for Battery Capacity Prediction

Hanbing Liu (Tsinghua University); Yanru Wu (Tsinghua University); Yang Li (Tsinghua-Berkeley Shenzhen Institute, Tsinghua University); Ercan E Kuruoglu (Tsinghua-Berkeley Shenzhen Institute); Xuan Zhang (Tsinghua University)

2596 (MLSP-P20.9): Learned Kalman Filtering in Latent Space with High-Dimensional Data

Itay Buchnik (Ben Gurion University); Damiano Steger (ETH Zurich); Guy Revach (ETH Zürich); Ruud J. G. van Sloun (Technical university of Eindhoven); Tirza S Routtenberg (Ben Gurion University of the Negev); Nir Shlezinger (Ben-Gurion University)

3028 (MLSP-P20.10): Constrained Dynamical Neural ODE for Time Series Modelling: A Case Study on Continuous Emotion Prediction

Ting Dang (University of Cambridge); Antoni Dimitriadis (University of New South Wales); Jingyao Wu (University of New South Wales); Vidhyasaharan Sethu (University of New South Wales); Eliathamby Ambikairajah (The University of New South Wales)

Wednesday, June 7

3468 (MLSP-P20.11): SADI: A SELF-ADAPTIVE DECOMPOSED INTERPRETABLE FRAMEWORK FOR ELECTRICITY LOAD FORECASTING UNDER EXTREME EVENTS

Hengbo LIU (Alibaba DAMO Academy); Ziqing MA (Alibaba); Linxiao Yang (Machine Intelligence Technology, Alibaba Group, Hangzhou, China); Tian Zhou (Alibaba DAMO Academy); Rui Xia (University of Cambridge); Yi Wang (The University of Hong Kong); Qingsong Wen (Alibaba Group U.S.); Liang Sun (Alibaba Group)

4914 (MLSP-P20.12): Counterfactual explanation for multivariate times series using a contrastive variational autoencoder

William Todo (Liebherr aerospace); Merwann Selmani (Liebherr Aerospace Toulouse); Béatrice Laurent (Institut de Mathématiques de Toulouse (UMR 5219), Université de Toulouse, INSA de Toulouse); Jean-Michel Loubes (Université Toulouse Paul Sabatier Institut de Mathématiques de Toulouse)

SLT-P17: Machine Learning Methods for Language II

Room: Dome B

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Wen Wang,

298 (SLT-P17.1): Contrastive Learning at the Relation and Event Level for Rumor Detection

Yingrui Xu (Institute of Information Engineering, Chinese Academy of Sciences; School of Cyber Security, University of Chinese Academy of Sciences); Jingyuan Hu (Institute of Information Engineering, Chinese Academy of Sciences); jingguo ge (ie, cas); Yulei Wu (University Of Exeter); Hui Li (Institute of Information Engineering, Chinese Academy of Sciences); Tong Li (Institute of Information Engineering, Chinese Academy of Sciences)

610 (SLT-P17.2): SELF-HEALING THROUGH ERROR DETECTION, ATTRIBUTION, AND RETRAINING

Ansel MacLaughlin (Amazon); Anna Rumshisky (University of Massachusetts Lowell); Rinat Khaziev (Amazon Alexa AI); Anil K Ramakrishna (Amazon); Yuval Merhav (Amazon); Rahul Gupta (Amazon)

865 (SLT-P17.3): Towards A Unified Training for Levenshtein Transformer

Kangjie Zheng (Peking University); Longyue Wang (Tencent AI Lab); Zhihao Wang (Xiamen University); Chen Binqi (Peking University); Ming Zhang (Peking University); Zhaopeng Tu (Tencent AI Lab)

1081 (SLT-P17.4): Database-Aware ASR Error Correction for Speech-to-SQL Parsing

Yutong Shao (University of California San Diego); Arun Kumar (University of California, San Diego); Ndapa Nakashole (University of California, San Diego)

2772 (SLT-P17.5): Picking the Underused Heads: A Network Pruning Perspective of Attention Head Selection for Fusing Dialogue Coreference Information

Zhengyuan Liu (ASTAR); Nancy Chen (Institute for Infocomm Research)

3066 (SLT-P17.6): Multi-lingual pronunciation assessment with unified phoneme set and language-specific embeddings

Binghui Lin (MIG, Tencent Science and Technology Ltd.); Liyuan wang (Tencent Technology Co., Ltd)

3092 (SLT-P17.7): Multi-modal ASR error correction with joint ASR error detection

Binghui Lin (MIG, Tencent Science and Technology Ltd.); Liyuan wang (Tencent Technology Co., Ltd)

4780 (SLT-P17.8): Pyramid Dynamic Inference: Encouraging Faster Inference via Early Exit Boosting

Ershad Banijamali (Amazon Inc.); Pegah Kharazmi (Amazon); Sepehr Eghbali (Amazon); Jixuan Wang (Amazon); Clement Chung (Amazon); Samridhi Choudhary (Amazon)

5117 (SLT-P17.9): Self-Supervised Adversarial Training for Contrastive Sentence Embedding

Jen-Tzung Chien (National Yang Ming Chiao Tung University); Yuan-An Chen (National Yang Ming Chiao Tung University)

5905 (SLT-P17.10): FULLY UNSUPERVISED TOPIC CLUSTERING OF UNLABELLED SPOKEN AUDIO USING SELF-SUPERVISED REPRESENTATION LEARNING AND TOPIC MODEL

Takashi Maekaku (Yahoo Japan Corporation); Yuya Fujita (Yahoo Japan Corporation); Xuankai Chang (Carnegie Mellon University); Shinji Watanabe (Carnegie Mellon University)

6298 (SLT-P17.11): Visual Information Matters for ASR Error Correction

Vanya BK (Indian Institute Of Technology, Madras); Shanbo Cheng (ByteDance); Ningxin Peng (ByteDance); Yuchen Zhang (ByteDance)

6348 (SLT-P17.12): UCorrect: An Unsupervised Framework for Automatic Speech Recognition Error Correction

Jiaxin GUO (Huawei); Minghan Wang (Huawei); Xiaosong Qiao (Huawei); Daimeng Wei (Huawei); Hengchao Shang (HW-TSC); ZongYao LI (HW-TSC); Zhengzhe YU (HW-TSC); Yinglu Li (HUAWEI TECHNOLOGIES CO., LTD.); Chang Su (Huawei); Min Zhang (Huawei); Shimin Tao (Huawei); Hao Yang (Huawei)

Wednesday, June 7

SLT-P18: Machine Learning Methods for Language III

Room: Dome C

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Leibny Garcia,

277 (SLT-P18.1): A Simple yet Effective Approach to Structured Knowledge Distillation

Wenye Lin (Tsinghua Shenzhen International Graduate School, Tsinghua University); Yangming Li (Tencent AI Lab); Lemao Liu (Tencent AI Lab); Shuming Shi (Tsinghua University); Hai-Tao Zheng (Tsinghua University)

611 (SLT-P18.2): Question Answering system with Sparse and Noisy Feedback

Djallel Bouneffouf (IBM); Ozgur Alkan (Optum); Raphael Feraud (Orange Labs); Baihan Lin (Columbia University)

629 (SLT-P18.3): Group Personalized Federated Learning

Zhe Liu (Meta); Yue Hui (Meta); Fuchun Peng (Facebook)

779 (SLT-P18.4): Dialogue System with Missing Observation

Djallel Bouneffouf (IBM); mayank agarwal (ibm); Irina Rish (university of montreal)

845 (SLT-P18.5): Boosting Prompt-based Few-shot Learners through Out-of-domain Knowledge Distillation

Xiaoqing Chen (Chongqing University); Chengyu Wang (Alibaba); junwei dong (Chongqing university); Minghui Qiu (Alibaba); Liang Feng (Chongqing University, China); Jun Huang (Alibaba Group)

1457 (SLT-P18.6): Optimal Transport with a Diversified Memory Bank for Cross-Domain Speaker Verification

Ruiteng Zhang (Tianjin University); Jianguo Wei (School of Computer Software, Tianjin University, Tianjin, China); Xugang Lu (NICT); Wenhuan Lu (Tianjin University); Di Jin (Tianjin University); Lin Zhang (National Institute of Informatics); Junhui Xu (Tianjin Key Laboratory of Cognitive Computing and Application, College of Intelligence and Computing, Tianjin University)

2372 (SLT-P18.7): SOURCE-FREE UNSUPERVISED DOMAIN ADAPTATION FOR QUESTION ANSWERING

Zishuo Zhao (Sun Yat-Sen University); Yuexiang Xie (Alibaba Group); Jingyou Xie (Sun Yat-sen University); Zhenzhou Lin (Sun Yat-sen University); Yaliang Li (Alibaba Group); Ying Shen (Sun Yat-Sen University)

2676 (SLT-P18.8): Improving BERT Fine-tuning via Stabilizing Cross-layer Mutual Information

Jicun Li (1. Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences (ICT/CAS) 2. University of Chinese Academy of Sciences, Beijing, China); Xingjian Li (1. Big Data Lab, Baidu Research; 2. State Key Lab of IOTSC, University of Macau); Tianyang Wang (University of Alabama at Birmingham); Shi Wang (1. Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences (ICT/CAS) 2. University of Chinese Academy of Sciences, Beijing, China); Yanan Cao (Institute of Information Engineering, Chinese Academy of Sciences); Cheng-Zhong Xu (University of Macau); Dejing Dou (Baidu)

2850 (SLT-P18.9): ATTENTION LOCALNESS IN SHARED ENCODER-DECODER MODEL FOR TEXT SUMMARIZATION

Li Huang (Southwestern University of Finance and Economics); Hongmei Wu (Southwestern University of Finance and Economics); Qiang Gao (Southwestern University of Finance and Economics); Guisong Liu (Southwestern University of Finance and Economics)

3017 (SLT-P18.10): SkillNet-NLG: General-Purpose Natural Language Generation with a Sparsely Activated Approach

Junwei Liao (University of Electronic Science and Technology of China); Duyu Tang (Tencent); Fan Zhang (Tianjin University); Shuming Shi (Tsinghua University)

3169 (SLT-P18.11): Improving Spoken Language Identification with Map-Mix

Shangeth Rajaa (skit.ai); Kriti Anandan (skit.ai); Swaraj Dalmia (skit.ai); Tarun Gupta (IIT Indore); Eng Siong Chng (Nanyang Technological University)

3461 (SLT-P18.12): A mutual implicit sentiment analysis model with bundle-aware contrastive learning

siqi cai (Wuhan University of Technology); Jingling Yuan (Wuhan University of Technology); Lin Li (Wuhan University of Technology)

SLT-P19: Machine Learning Methods for Language IV

Room: Dome D

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Jan Yenda Trmal,

3725 (SLT-P19.1): MHLAT: Multi-hop Label-wise Attention Model for Automatic ICD Coding

Junwen Duan (Central South University); Han Jiang (Central South University); Ying Yu (Central South University)

3866 (SLT-P19.2): Ensemble knowledge distillation of self-supervised speech models

Kuan-Po Huang (National Taiwan University); Tzu-hsun Feng (National Taiwan University); YU-KUAN FU (NTU); Tsu-Yuan Hsu (National Taiwan University); Po-Chieh Yen (National Taiwan University); Wei-Cheng Tseng (National Taiwan University); Kai-Wei Chang (National Taiwan University); Hung-yi Lee (National Taiwan University)

Wednesday, June 7

4066 (SLT-P19.3): Once-for-All Sequence Compression for Self-Supervised Speech Models

Hsuan-Jui Chen (National Taiwan University); Yen Meng (National Taiwan University); Hung-yi Lee (National Taiwan University)

4155 (SLT-P19.4): Mutually Guided Few-shot Learning for Relational Triple Extraction

Chengmei Yang (Tongji University); shuai jiang (tongji university); Bowei He (City University of Hong Kong); Chen Ma (City University of Hong Kong); 何良华 (同济大学)

4178 (SLT-P19.5): F-PABEE: Flexible-patience-based Early Exiting for Single-label and Multi-label text Classification Tasks

Xiangxiang Gao (Shanghai Jiaotong University); Wei Zhu (East China Normal University); Jiasheng Gao (Shenzhen University); Congrui Yin (Nanchang University)

4673 (SLT-P19.6): Lost In Translation: Generating Adversarial Examples Robust to Round-Trip Translation

Neel Bhandari (RV College of Engineering); Pin-Yu Chen (IBM Research)

4818 (SLT-P19.7): Analysing Discrete Self Supervised Speech Representation for Spoken Language Modeling

Amitay Sicherman (The Hebrew University of Jerusalem); Yossi Adi (Facebook AI Research)

4854 (SLT-P19.8): LEARNING DEPENDENCIES OF DISCRETE SPEECH REPRESENTATIONS WITH NEURAL HIDDEN MARKOV MODELS

Sung-Lin Yeh (University of Edinburgh); Hao Tang (The University of Edinburgh)

5011 (SLT-P19.9): LEVERAGING LABEL CORRELATIONS IN A MULTI-LABEL SETTING: A CASE STUDY IN EMOTION

Georgios Chochlakis (University of Southern California); Girish M Mahajan (Microsoft); Sabyasachee Baruah (University of Southern California); Keith Burghardt (ISI, University of Southern California); Kristina Lerman (USC Information Sciences Institute); Shrikanth Narayanan (USC)

5597 (SLT-P19.10): Parallel Sentence-Level Explanation Generation for Real-World Low-Resource Scenarios

Yan Liu (Microsoft Research); Xiaokang Chen (Peking University); Qi Dai (Microsoft Research)

5598 (SLT-P19.11): COMPARATIVE LAYER-WISE ANALYSIS OF SELF-SUPERVISED SPEECH MODELS

Ankita Pasad (Toyota Technological Institute at Chicago); Bowen Shi (Toyota Technological Institute at Chicago); Karen Livescu (TTI-Chicago)

6369 (SLT-P19.12): Unsupervised Out-of-Distribution Detection Using Few In-Distribution Samples

Chandan Gautam (ASTAR (Institute for Infocomm Research)); Aditya Kane (Pune Institute of Computer Technology); Ramasamy Savitha (I2R ASTAR); Suresh Sundaram (Indian Institute of Science)

SLT-P20: Multilingual Speech Recognition and Identification

Room: Dome E

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Andreas Stolcke, Peter Bell

199 (SLT-P20.1): Learning ASR pathways: A sparse multilingual ASR model

Mu Yang (University of Texas at Dallas); Andros Tjandra (Meta Platforms, Inc); Chunxi Liu (Two Sigma); David Zhang (Meta); Duc Le (Meta); Ozlem Kalinli (Meta)

401 (SLT-P20.2): Joint unsupervised and supervised learning for context-aware language identification

Jinseok Park (42dot); Hyung Yong Kim (42dot); Jihwan Park (42dot Inc.); Byeong-Yeol Kim (42dot); Shukjae Choi (Hyundai Motor Company); Yunkyu Lim (42dot)

2702 (SLT-P20.3): Reducing Language Confusion for Code-switching Speech Recognition with Token-level Language Diarization

Hexin Liu (Nanyang Technological University); Haihua Xu (Temasek Laboratories, Nanyang Technological University, Singapore); Paola Garcia (Johns Hopkins University); Andy W H Khong (Nanyang Technological University); Yi He (Bytedance); Sanjeev Khudanpur (Johns Hopkins University)

3417 (SLT-P20.4): Accidental Learners: Spoken Language Identification in Multilingual Self-Supervised Models

Travis M Bartley (NVIDIA; CUNY); Fei Jia (NVIDIA Corporation); Krishna C Puvvada (NVIDIA); Samuel Krizan (NVIDIA); Boris Ginsburg (NVIDIA)

3852 (SLT-P20.5): DOMAIN AND LANGUAGE ADAPTATION USING HETEROGENEOUS DATASETS FOR WAV2VEC2.0-BASED SPEECH RECOGNITION OF LOW-RESOURCE LANGUAGE

Kak Soky (Kyoto University); Sheng Li (National Institute of Information & Communications Technology (NICT)); Chenhui Chu (Kyoto University); Tatsuya Kawahara (Kyoto University)

Wednesday, June 7

3896 (SLT-P20.6): Dynamic TF-TDNN: Dynamic Time Delay Neural Network based on Temporal-Frequency Attention for Dialect Recognition

Chao Liao (Kuaishou); Jinwen Huang (Kuaishou Technology); Huan Yuan (Kuaishou Technology); Peng Yao (Kuaishou Inc.); Jianchao Tan (Kwai Inc.); zhang dawei (Kuaishou Technology); Feng Deng (Kuaishou); Xiaorui Wang (Kwai); Chengru Song (Kuaishou)

4186 (SLT-P20.7): MoLE : MIXTURE OF LANGUAGE EXPERTS FOR MULTI-LINGUAL AUTOMATIC SPEECH RECOGNITION
YooHwan Kwon (Naver cooperation); Soo-Whan Chung (Naver Corporation)

4355 (SLT-P20.8): TOWARDS ZERO-SHOT CODE-SWITCHED SPEECH RECOGNITION

Brian Yan (Carnegie Mellon University); Matthew S Wiesner (Johns Hopkins University); Ondrej Klejch (University of Edinburgh); Preethi Jyothi (Indian Institute of Technology Bombay); Shinji Watanabe (Carnegie Mellon University)

5226 (SLT-P20.9): Learning Cross-lingual Visual Speech Representations

Andreas Zinonos (Imperial College London); Alexandros Haliassos (Imperial College London); Pingchuan Ma (Meta); Stavros Petridis (Imperial College London); Maja Pantic (Imperial College London)

5503 (SLT-P20.10): Code-Switching Text Generation and Injection in Mandarin-English ASR

Haibin Yu (Shanghai Jiao Tong University); Yuxuan Hu (Microsoft); Yao Qian (Microsoft); Ma Jin (Microsoft); Linqun Liu (Microsoft); Shujie Liu (Microsoft Research Asia); Yu Shi (Microsoft); Yanmin Qian (Shanghai Jiao Tong University); Ed C Lin (Microsoft); Michael Zeng (Microsoft)

5700 (SLT-P20.11): Exploring universal singing speech language identification using self-supervised learning based front-end features

Xingming Wang (Wuhan University); Hao Wu (Speech, Audio and Music Intelligence (SAMI) group, ByteDance); Chen Ding (Speech, Audio and Music Intelligence (SAMI) group, ByteDance); Chuanzeng Huang (Speech, Audio and Music Intelligence (SAMI) group, ByteDance); Ming Li (Duke Kunshan University)

AASP-L6: Sound Event Detection

Room: Salon de Roses A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Gaël Richard, Jonathan Le Roux

02:00 PM

4341 (AASP-L6.1): SUBBAND DEPENDENCY MODELING FOR SOUND EVENT DETECTION

Yadong Guan (Harbin Institute of Technology); 贵滨 郑 (哈尔滨工业大学计算机科学与技术学院); jiqing Han (Harbin Institute of Technology); huanliang wang (Qdreamer)

02:15 PM

5545 (AASP-L6.2): NAS-DYMC: NAS-based Dynamic Multi-Scale Convolutional Neural Network for Sound Event Detection

Wang Jun (Kuaishou Technology); Peng Yao (Kuaishou Inc.); Feng Deng (Kuaishou); Jianchao Tan (Kwai Inc.); Chengru Song (Kuaishou); Xiaorui Wang (Kwai)

02:30 PM

1350 (AASP-L6.3): Visual onoma-to-wave: environmental sound synthesis from visual onomatopoeias and sound-source images

Hien Ohnaka (National Institute of Technology, Tokuyama College); Shinnosuke Takamichi (The University of Tokyo); Keisuke Imoto (Doshisha University); Yuki Okamoto (Ritsumeikan University); Kazuki Fujii (The University of Tokyo); Hiroshi Saruwatari (The University of Tokyo)

02:45 PM

2895 (AASP-L6.4): AST-SED: an Effective Sound Event Detection Method Based on Audio Spectrogram Transformer

Kang Li (University of Science and Technology of China, National Engineering Research Center of Speech and Language Information Processing.); Yan Song (USTC); Lirong Dai (University of Science and Technology of China); Ian McLoughlin (Singapore Institute of Technology); Xin Fang (iFlytek Research); Lin Liu (iFlytek Research)

03:00 PM

4702 (AASP-L6.5): HiSSNet: Sound Event Detection and Speaker Identification via Hierarchical Prototypical Networks for Low-Resource Headphones

N Shashaank (Columbia University); Berker Banar (Queen Mary University of London); Mohammad Izadi (BOSE); Jeremy Kemmerer (BOSE); Shuo Zhang (Bose); Chuan-Che Huang (BOSE)

03:15 PM

5734 (AASP-L6.6): Performance above all ? Energy consumption vs. performance, a study on sound event detection with heterogeneous data

romain serizel (Université de Lorraine); Samuele Cornell (Università Politecnica delle Marche); Nicolas Turpault (Inria)

Wednesday, June 7

BISP-L1: Brain Connectivity

Room: Nafsika A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Tulay Adali,

02:00 PM

788 (BISP-L1.01): CONSTRAINED INDEPENDENT COMPONENT ANALYSIS BASED ON ENTROPY BOUND MINIMIZATION FOR SUBGROUP IDENTIFICATION FROM MULTISUBJECT FMRI DATA

Hanlu Yang (University of Maryland, Baltimore County); Fateme Ghayem (University of Maryland, Baltimore County); Ben Gabrielson (University of Maryland, Baltimore County); Mohammad Akhonda (UMBC); Vince Calhoun (TReNDS); Tulay Adali (University of Maryland, Baltimore County)

02:15 PM

5102 (BISP-L1.02): Spatio-Temporal Attention in Multi-Granular Brain Chroconnectomes for Detection of Autism Spectrum Disorder

James Orme-Rogers (University of Southern California); Ajitesh Srivastava (University of Southern California)

02:30 PM

4982 (BISP-L1.03): Glacier: Glass-box Transformer for Interpretable Dynamic Neuroimaging

Usman Mahmood (Georgia State University); Zening Fu (Georgia State University); Vince Calhoun (TReNDS); Sergey Plis (Georgia State University)

02:45 PM

4185 (BISP-L1.04): TOPGFORMER: TOPOLOGICAL-BASED GRAPH TRANSFORMER FOR MAPPING BRAIN STRUCTURAL CONNECTIVITY TO FUNCTIONAL CONNECTIVITY

Dalu Guo (Southeast University); Ke Zhang (Southeast University); Jiaying Li (Southeast University); Youyong Kong (Southeast University)

03:00 PM

3445 (BISP-L1.05): ADHD Classification with biomarker identification using a triplet loss attention auto-encoding network

Yibin Tang (Hohai University); Ying Chen (Changzhou University); Yuan Gao (Hohai University); Aimin Jiang (Hohai University); Lin Zhou (Southeast University)

03:15 PM

6402 (BISP-L1.06): New Interpretable Patterns and Discriminative Features from Brain Functional Network Connectivity Using Dictionary Learning

Fateme Ghayem (UMBC); Hanlu Yang (University of Maryland, Baltimore County); Furkan Kantar (UMBC); Seung-Jun Kim (University of Maryland, Baltimore County); Vince Calhoun (TReNDS); Tulay Adali (University of Maryland, Baltimore County)

GC-6: Speech Signal Improvement Signal Processing Grand Challenge 2023

Room: Nefeli B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): TBA

02:00 PM

6652 (GC-L6.1): Introduction

Ross Cutler (Microsoft Corporation); Ando Saabas (Microsoft); Babak Naderi (Microsoft); Ristea N Catalin (Microsoft); Sebastian Braun (Microsoft); Robert Aichner (Microsoft)

02:20 PM

6878 (GC-L6.2): Speech Signal Improvement Using Causal Generative Diffusion Models

Julius Richter (Universität Hamburg); Simon Welker (Universität Hamburg); Jean-Marie Lemercier (Universität Hamburg); Bunlong Lay (Universität Hamburg); Tal Peer (Universität Hamburg); Timo Gerkmann (Universität Hamburg)

02:32 PM

6880 (GC-L6.3): Half-temporal and half-frequency attention U2Net for speech signal improvement

Zehua Zhang (Harbin Institute of Technology(Shenzhen)); Shiyun Xu (Harbin Institute of Technology(Shenzhen)); Xuyi Zhuang (Harbin Institute of Technology(Shenzhen)); Yukun Qian (Harbin Institute of Technology (Shenzhen)); Liyanu Zhou (Harbin Institute of Technology(Shenzhen)); Mingjiang Wang (Harbin Institute of Technology Shenzhen)

02:44 PM

6906 (GC-L6.4): Gesper: A Unified Framework for General Speech Restoration

Jun Chen (Tsinghua University); yupeng shi (tencent); wenzhe liu (Tencent); Wei Rao (Tencent); shulin 柯 (Tencent); Andong Li (Institute of Acoustics, Chinese Academy of Sciences); Yannan Wang (Tencent); Zhiyong Wu (Tsinghua University); Shi-dong Shang (tencent); Chengshi Zheng (Chinese Academy of Science)

Wednesday, June 7

02:56 PM

6911 (GC-L6.5): SSI-Net: A MULTI-STAGE SPEECH SIGNAL IMPROVEMENT SYSTEM FOR ICASSP 2023 SSI CHALLENGE
weixin zhu (tencent); Zilin Wang (Tsinghua University); Jiuxin Lin (Tsinghua University); Chang Zeng (National Institute of Informatics); Tao Yu (Tencent)

03:08 PM

6919 (GC-L6.6): TWO-STAGE NEURAL NETWORK FOR ICASSP 2023 SPEECH SIGNAL IMPROVEMENT CHALLENGE
Mingshuai Liu (NWPUP); Shubo Lv (Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, School of Computer Science, Northwestern Polytechnical University); Zihan Zhang (Northwestern Polytechnical University); Runduo Han (Northwestern Polytechnical University); Xiang Hao (NWPUP); Xianjun Xia (ByteDance); Li Chen (ByteDance); Yijian Xiao (ByteDance); Lei Xie (NWPUP)

IFS-L1: Anonymization and Data Privacy

Room: Salon de Roses B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): TBA

02:00 PM

1556 (IFS-L1.01): APGP: ACCURACY-PRESERVING GENERATIVE PERTURBATION FOR DEFENDING AGAINST MODEL CLONING ATTACKS

Anda Cheng (CASIA); jian cheng (casia)

02:15 PM

3608 (IFS-L1.02): Towards Practical Edge Inference Attacks against Graph Neural Networks

Kailai Li (Shanghai Jiao Tong University); Jiawei Sun (Shanghai Jiao Tong University); Ruoxin Chen (Shanghai Jiao Tong University); Wei Ding (Shanghai Jiao Tong University); Kexue Yu (Shanghai Jiao Tong University); Jie Li (Shanghai Jiao Tong University); Chentao Wu (Shanghai Jiao Tong University)

02:30 PM

344 (IFS-L1.03): GAPter: Gray-box Data Protector for Deep Learning Inference Services at User Side

Hao Wu (Nanjing University); Bo Yang (Nanjing University); Xiaopeng Ke (Nanjing University); Siyi He (Nanjing University); Fengyuan Xu (Nanjing University); Sheng Zhong (Nanjing University)

02:45 PM

4204 (IFS-L1.04): DISTANCE-BASED ONLINE LABEL INFERENCE ATTACKS AGAINST SPLIT LEARNING

Junlin Liu (Beijing University of Posts and Telecommunications); Xinchen Lyu (Beijing University of Posts and Telecommunications)

03:00 PM

4920 (IFS-L1.05): Prosody is Not Identity: A Speaker Anonymization Approach Using Prosody Cloning

Sarina Meyer (University of Stuttgart); Florian Lux (University of Stuttgart); Julia Koch (University of Stuttgart); Pavel Denisov (University of Stuttgart); Pascal Tilli (University of Stuttgart); Ngoc Thang Vu (University of Stuttgart)

03:15 PM

6006 (IFS-L1.06): A PRIVACY-PRESERVING TRAJECTORY MINING MODEL

Ziyang Wang (Shenzhen University); Xiaoxiao Wu (Shenzhen University); Junjie Zhu (Shenzhen University); Yingying Zhu (University of Texas Arlington)

MLSP-L6: Machine Learning for Time Series Analysis I

Room: Jupiter

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Che Lin,

02:00 PM

931 (MLSP-L6.1): Wassertein GAN synthesis for time series with complex temporal dynamics: Frugal architectures and arbitrary sample-size generation

Thomas Beroud (Ecole Centrale Nantes); Patrice Abry (CNRS, Physics Department, Ecole Normale Supérieure de Lyon); Yannick Malevergne (Univ. Paris1); Marc Senneret (Vivienne Investissement); Gerald Perrin (Vivienne Investissement); Johan Macq (Vivienne Investissement)

02:15 PM

3732 (MLSP-L6.2): Change Point Detection with Neural Online Density-ratio Estimator

Xiuheng Wang (Université Côte d'Azur, CNRS, OCA); Ricardo Borsoi (UL); Cédric Richard (University Nice Sophia Antipolis); Jie Chen (Northwestern Polytechnical University)

Wednesday, June 7

02:30 PM

423 (MLSP-L6.3): NRTSI: NON-RECURRENT TIME SERIES IMPUTATION

Siyuan Shan (Department of Computer Science, University of North Carolina at Chapel Hill); Yang Li (Department of Computer Science, University of North Carolina at Chapel Hill); Junier Oliva (UNC-Chapel Hill)

02:45 PM

2828 (MLSP-L6.4): MEASURING DEVIATION FROM STOCHASTICITY IN TIME-SERIES USING AUTOENCODER BASED TIME-INVARIANT REPRESENTATION: APPLICATION TO BLACK HOLE DATA

Sai Pradeep Chakka (IIIT Bangalore); Neelam Sinha (IIIT Bangalore); Banibrata Mukhopadhyay (Indian Institute of Science)

03:00 PM

2925 (MLSP-L6.5): Leveraging neural koopman operators to learn continuous representations of dynamical systems from scarce data

Anthony Frion (IMT Atlantique); Lucas Drumetz (IMT Atlantique); Mauro Dalla Mura (Grenoble INP); Guillaume Tochon (EPITA Research and Development Laboratory (LRDE)); Abdeldjalil Aissa-El-Bey (France)

03:15 PM

3080 (MLSP-L6.6): OTW: Optimal Transport Warping for Time Series

Fabian R Latorre (EPFL); Chenghao Liu (Salesforce); Doyen Sahoo (Salesforce); Steven Hoi (Salesforce)

SLT-L11: Natural Language Processing I

Room: Delphi

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Zhengyuan Liu,

02:00 PM

1114 (SLT-L11.01): Permutation Invariant Training for Paraphrase Identification

Jun Bai (Beihang University); Chuantao Yin (Beihang University); Hanhua Hong (Beihang University); Jianfei zhang (Beihang University); Chen Li (Beihang University); Yanmeng Wang (Ping An Technology); Wenge Rong (Beihang University)

02:15 PM

1213 (SLT-L11.02): UNSUPERVISED EXTRACTIVE SUMMARIZATION WITH HETEROGENEOUS GRAPH EMBEDDINGS FOR CHINESE DOCUMENTS

Chen Lin (Tencent); Ye Liu (Tencent); Siyu An (Tencent); Di Yin (Tencent)

02:30 PM

2426 (SLT-L11.03): A Dynamic Graph Interactive Framework with Label-Semantic Injection for Spoken Language Understanding

Zhihong Zhu (Peking University); Weiyuan Xu (Peking University); Xuxin Cheng (Peking University); Tengtao Song (Peking University); Yuexian Zou (Peking University)

02:45 PM

2865 (SLT-L11.04): Twitter Stance Detection via Neural Production Systems

Bowen Zhang (Shenzhen Technology University); Daijun Ding (Shenzhen Technology University); Guangning Xu (Harbin Institute of Technology, Shenzhen ▲); Jinjin Guo (JD Intelligent Cities Research); Zhichao Huang (JD Intelligent Cities Research); Xu Huang (Harbin Institute of Technology, Shenzhen)

03:00 PM

4312 (SLT-L11.05): An Interpretable model using evidence information for Multi-hop Question Answering over Long texts

Yanyi Chen (Beijing University of Posts and Telecommunications); Ruifang Liu (Beijing University of Posts and Telecommunications); Xiyan Liu (Beijing University of Posts and Telecommunications); Yidong Shi (Beijing University of Posts and Telecommunications); Ge Bai (Beijing University of Posts and Telecommunications)

03:15 PM

6027 (SLT-L11.06): SIAST: A Slot Imbalance-Aware Self-Training Scheme for Semi-Supervised Slot Filling

Jiachi Liu (Beijing University of Posts and Telecommunications); Sishi Xiong (Beijing University of Posts and Telecommunications); Yuehuan He (University of Toronto); tong zhou (Beijing University of Posts and Telecommunications); Liwen Wang (Beijing University of Posts and Telecommunications); Xuefeng Li (Beijing University of Posts and Telecommunications); Bo Xiao (Beijing University of Posts and Telecommunications)

Wednesday, June 7

SLT-L12: Pronunciation and Fluency Assessment

Room: Athina

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Eric Fosler-Lussier, Yossi Keshet

02:00 PM

597 (SLT-L12.1): Phonetic RNN-Transducer for Mispronunciation Diagnosis

Daniel Yue Zhang (Amazon); Soumya Saha (Amazon); Sarah Campbell (Amazon)

02:15 PM

1253 (SLT-L12.2): AN ASR-FREE FLUENCY SCORING APPROACH WITH SELF-SUPERVISED LEARNING

Wei Liu (The Chinese University of Hong Kong); Kaiqi Fu (Bytedance); Xiaohai Tian (ByteDance); Shuju Shi (ByteDance); Wei Li (Bytedance); Zejun Ma (Bytedance); Tan Lee (The Chinese University of Hong Kong)

02:30 PM

1286 (SLT-L12.3): LEVERAGING PHONE-LEVEL LINGUISTIC-ACOUSTIC SIMILARITY FOR UTTERANCE-LEVEL PRONUNCIATION SCORING

Wei Liu (The Chinese University of Hong Kong); Kaiqi Fu (Bytedance); Xiaohai Tian (ByteDance); Shuju Shi (ByteDance); Wei Li (Bytedance); Zejun Ma (Bytedance); Tan Lee (The Chinese University of Hong Kong)

02:45 PM

3632 (SLT-L12.4): Hierarchical Pronunciation Assessment with Multi-Aspect Attention

Heejin Do (POSTECH); Yunsu Kim (POSTECH); Gary Geunbae Lee (Postech)

03:00 PM

4427 (SLT-L12.5): End-to-End word-level disfluency detection and classification in children's reading assessment

Lavanya Venkatasubramaniam (Ohio State University); Vishal Sunder (The Ohio State University); Eric Fosler-Lussier (Ohio State)

03:15 PM

6367 (SLT-L12.6): Relative dynamic time warping comparison for pronunciation errors

Caitlin Richter (Reykjavik University); Jon Gudnason (Reykjavik University)

SS-L7: Edge Learning for Emerging Wireless Technologies

Room: Nefeli A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Paolo Di Lorenzo, George Alexandropoulos, Mattia Merluzzi

02:00 PM

576 (SS-L7.1): Calibrating AI Models for Few-Shot Demodulation via Conformal Prediction

Kfir Cohen (KCL); Sangwoo Park (King's College London); Osvaldo Simeone (King's College London); Shlomo Shamai (The Technion)

02:15 PM

2905 (SS-L7.2): CADET: Control-Aware Dynamic Edge Computing for Real-Time Target Tracking in UAV Systems

Luis Felipe Florenzan Reyes (University of L'Aquila); Francesco Smarra (University of L'Aquila); Alessandro D'Innocenzo (University of L'Aquila); marco levorato (University of California, Irvine)

02:30 PM

3457 (SS-L7.3): RELIABLE BEAMFORMING AT TERAHERTZ BANDS: ARE CAUSAL REPRESENTATIONS THE WAY FORWARD?

Christo Kurisummoottil Thomas (Virginia Tech); Walid Saad (Virginia Tech)

02:45 PM

4520 (SS-L7.4): Personalizing Federated Learning with Over-the-Air Computations

Zihan Chen (Singapore University of Technology and Design); Zeshen Li (Zhejiang University); Howard H. Yang (ZJU-UIUC Institute); Tony Quek (Singapore University of Technology and Design)

03:00 PM

4969 (SS-L7.5): BER-aware dynamic resource management for edge-assisted goal-oriented communications

Francesco F Binucci (University of Perugia); Paolo Banelli (University of Perugia)

03:15 PM

6336 (SS-L7.6): Lyapunov-driven deep reinforcement learning for edge inference empowered by Reconfigurable Intelligent Surfaces

Kyriakos Stylianopoulos (National and Kapodistrian University of Athens); Mattia Merluzzi (CEA-Leti); Paolo Di Lorenzo (Sapienza University of Rome); George Alexandropoulos (National and Kapodistrian University of Athens)

Wednesday, June 7

AASP-P2: Acoustic Sensor Array Processing and Sound Source Localization

Room: Dome B

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Jesper Rindom Jensen,

4227 (AASP-P2.1): ROBUST BINAURAL SOUND LOCALISATION WITH TEMPORAL ATTENTION

Qi Hu (*Institute of Acoustics of Chinese Academy of Sciences*); Ning Ma (*University of Sheffield*); Guy J. Brown (*University of Sheffield*)

4303 (AASP-P2.2): Geometry-aware DoA Estimation using a Deep Neural Network with mixed-data input features

Ulrik Kowalk (*Institute of Hearing Technology and Audiology, Jade University of Applied Sciences, Oldenburg*); Simon Doclo (*University of Oldenburg*); Joerg Bitzer (*Institute of Hearing Technology and Audiology, Jade University of Applied Sciences, Oldenburg*)

5153 (AASP-P2.3): ESTIMATING ACOUSTIC DIRECTION OF ARRIVAL USING A SINGLE STRUCTURAL SENSOR ON A RESONANT SURFACE

Tre DiPassio (*University of Rochester*); Michael Heilemann (*University of Rochester*); Benjamin` Thompson (*University of Rochester*); Mark Bocko (*University of Rochester*)

4542 (AASP-P2.4): Assisted RTF-Vector-Based Binaural Direction of Arrival Estimation Exploiting a Calibrated External Microphone Array

Daniel Feigin (*University of Oldenburg*); Simon Doclo (*University of Oldenburg*)

4670 (AASP-P2.5): Distributed Adaptive Norm Estimation for Blind System Identification in Wireless Sensor Networks

Matthias Blochberger (*KU Leuven*); Filip Elvander (*Aalto University*); Randall Ali (*KU Leuven*); Jan Ostergaard (*Aalborg University*); Jesper Jensen (*Aalborg University*); Marc Moonen (*KU Leuven*); Toon van Waterschoot (*Department of Electrical Engineering (ESAT-STADIUS/ETC)*)

4848 (AASP-P2.6): Acoustic source localization in the spherical harmonics domain exploiting low-rank approximations

Maximo Cobos (*Universitat de Valencia*); Mirco Pezzoli (*Politecnico di Milano*); Fabio Antonacci (*Politecnico di Milano*); Augusto Sarti (*Politecnico di Milano*)

969 (AASP-P2.7): Long-term Synchronization of Wireless Acoustic Sensor Networks with Nonpersistent Acoustic Activity using Coherence State

Aleksej Chinaev (*Carl-von-Ossietzky University of Oldenburg*); Niklas Knaepper (*Carl-von-Ossietzky University of Oldenburg*); Gerald Enzner (*Carl von Ossietzky University Oldenburg*)

1066 (AASP-P2.8): Graph neural networks for sound source localization on distributed microphone networks

Eric Grinstein (*Imperial College London*); Mike Brookes (*Imperial College London*); Patrick A. Naylor (*Imperial College London*)

1359 (AASP-P2.9): Noise PSD Insensitive RTF Estimation in a Reverberant and Noisy Environment

Changheng Li (*Delft University of Technology*); Richard Hendriks (*TU Delft*)

2356 (AASP-P2.10): Generalized Relative Harmonic Coefficients

Yonggang Hu (*Australian National University*); Sharon Gannot (*Bar-Ilan University*); thushara abhayapala (*The Australian National University*)

2751 (AASP-P2.11): NEURAL OPTIMIZATION OF GEOMETRY AND FIXED BEAMFORMER FOR LINEAR MICROPHONE ARRAYS

Longfei Yan (*Victoria University of Wellington*); Weilong Huang (*Alibaba Group*); W. Bastiaan Kleijn (*Victoria University of Wellington*); thushara abhayapala (*The Australian National University*)

CI-P2: Computational Imaging III

Room: Dome L

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Chandra Sekhar Seelamantula,

4478 (CI-P2.1): DEEP NETWORK SERIES FOR LARGE-SCALE HIGH-DYNAMIC RANGE IMAGING

Amir Aghabiglou (*Heriot Watt university*); Matthieu Terris (*Heriot-Watt University*); Adrian Jackson (*EPCC, University of Edinburgh*); Yves Wiaux (*Heriot-Watt University*)

5061 (CI-P2.2): Hadamard Layer to Improve Semantic Segmentation

Angello Hoyos (*Centro de Investigación en Matemáticas, A.C.*); Mariano Rivera (*Centro de Investigación en Matemáticas AC*)

Wednesday, June 7

6370 (CI-P2.3): G2CNN: GEOMETRIC PRIOR BASED GCNN FOR SINGLE-VIEW 3D RECONSTRUCTION WITH LOOP SUBDIVISION

Kun Cao (Beijing University of Technology); Na Qi (Beijing University of Technology); Wei Xu (Faculty of Information Technology, Beijing University of Technology); Qing Zhu (Beijing University of Technology); Shibo Xu (Beijing University of Technology); Changxin Pan (Beijing University of Technology)

6475 (CI-P2.4): Model-based spectral reconstruction of interferometric acquisitions

Mohamad Jouni (Grenoble INP); Daniele Picone (Grenoble INP); Mauro Dalla Mura (Grenoble INP)

1996 (CI-P2.5): Super-Resolution for Macro X-ray Fluorescence Data Collected from Old Master Paintings

Su Yan (Imperial College London); Herman Jadan (Imperial College London); Jun-Jie Huang (National University of Defense Technology); Nathan S Daly (The Fitzwilliam Museum); Catherine Higgitt (The National Gallery); Pier Luigi Dragotti (Imperial College London)

385 (CI-P2.6): Event-Based Visual Microphone

Matthew D Howard (Air Force Research Laboratory); Keigo Hirakawa (University of Dayton)

441 (CI-P2.7): Ultra Real-Time Portrait Matting via Parallel Semantic Guidance

Xin Huang (University of Maryland, Baltimore County); Jiake Xie (PicUP.Ai); Bo Xu (OPPO Research Institute); Han Huang (OPPO Research Institute); Ziwen Li (OPPO Research Institute); Cheng Lu (XPENG); Yandong Guo (OPPO Research Institute); Yong Tang (PicUP.Ai)

860 (CI-P2.8): Transient Dictionary Learning for Compressed Time-of-Flight Imaging

Miguel Heredia Conde (University of Siegen)

4292 (CI-P2.9): Fast Multiscale 3D Reconstruction Using Single-Photon LiDaR Data

Sandor Plosz (Heriot-Watt University); Istvan Gyongy (University of Edinburgh); Jonathan Leach (Heriot-Watt University); Stephen McLaughlin (School of Engineering, Heriot-Watt University); Gerald S. Buller (Heriot-Watt University); Abderrahim Halimi (Heriot-Watt university)

5082 (CI-P2.10): DEEP BORN OPERATOR LEARNING FOR REFLECTION TOMOGRAPHIC IMAGING

Qingqing Zhao (Stanford University); Yanting Ma (Mitsubishi Electric Research Laboratories, USA); Petros Boufounos (Mitsubishi Electric Research Laboratories); Saleh Nabi (); Hassan Mansour (Mitsubishi Electric Research Laboratories (MERL))

6713 (CI-P2.11): Annotated Pedestrians: A Dataset for Soft Biometrics Estimation for Varying Distances (SPS Journal Paper)*

BILAL HASSAN (Northumbria University UK London campus); Muhammad Fiaz (Superior University, Lahore); Husnain Sherazi (University of West London); Usman Butt (Brunel University London)

IVMSP-P19: Deep Learning

Room: Dome J

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Byonghyo Shim, Sergiy A. Vorobyov

908 (IVMSP-P19.1): STRUCTURE-PRESERVING AND REDUNDANCY-FREE FEATURES REFINEMENT FOR GENERALIZED ZERO-SHOT LEARNING

Jian Ni (University of Science and Technology of China); Yong Liao (University of Sciences and Technology of China)

932 (IVMSP-P19.2): SINE: SIMILARITY-REGULARIZED INTRA-CLASS EXPLOITATION FOR CROSS-GRANULARITY FEW-SHOT LEARNING

Jinhai Yang (Shanghai Jiao Tong University); Hua Yang (Shanghai Jiao Tong University)

1322 (IVMSP-P19.3): Vision Transformer-based Feature Extraction for Generalized Zero-Shot Learning

Jiseob Kim (Seoul National University); Kyuhong Shim (Seoul National University); Junhan Kim (Seoul National University); Byonghyo Shim (Seoul National University)

3451 (IVMSP-P19.4): Progressive Meta-Pooling Learning for Lightweight Image Classification Model

Peijie Dong (School of Computer Science, National University of Defense Technology); Xin Niu (NUDT); ZHILIANG TIAN (National University of Defense Technology); Lujun Li (Chinese Academy of Sciences); Xiaodong Wang (National University of Defense Technology); Zimian Wei (School of Computer Science, National University of Defense Technology); Hengyue Pan (National University of Defense Technology); Dongsheng Li (School of Computer Science, National University of Defense Technology)

1561 (IVMSP-P19.5): Efficient Online Convolutional Dictionary Learning Using Approximate Sparse Components

Farshad G Veshki (Aalto university); Sergiy A. Vorobyov (Aalto University)

5313 (IVMSP-P19.6): Toward Auto-evaluation with Confidence-based Category Relation-aware Regression

Jiexin Wang (Renmin University of China); Jiahao Chen (Renmin University of China); Bing Su (Renmin University of China)

Wednesday, June 7

2325 (IVMSP-P19.7): Deep Double Self-expressive Subspace Clustering

zhao ling (Southwest University); Ma Yunpeng (Southwest University); Shanxiang Chen (southwest university); Jun Zhou (Southwest University)

4420 (IVMSP-P19.8): Towards Realizing the Value of Labeled Target Samples: a Two-Stage Approach for Semi-Supervised Domain Adaptation

Mengqun Jin (Tsinghua University); Kai Li (NEC LABORATORIES AMERICA, INC); SHUYAN LI (University of Cambridge); Chunming He (Tsinghua University); Xiu Li (Tsinghua University)

5489 (IVMSP-P19.9): DMFormer: Closing the Gap between CNN and Vision Transformers

Zimian Wei (School of Computer Science, National University of Defense Technology); Hengyue Pan (National University of Defense Technology); Lujun Li (Chinese Academy of Sciences); MengLong Lu (National University of Defense Technology); Xin Niu (NUDT); Peijie Dong (School of Computer Science, National University of Defense Technology); Dongsheng Li (School of Computer Science, National University of Defense Technology)

4370 (IVMSP-P19.10): SEMANTIC CENTRALIZED CONTRASTIVE LEARNING FOR UNSUPERVISED HASHING

Fengming Liang (Beijing University of Posts and Telecommunications); Changlin Fan (Beijing University of Posts and Telecommunications); Bo Xiao (Beijing University of Posts and Telecommunications); Kongming Liang (Beijing University of Posts and Telecommunications)

6706 (IVMSP-P19.11): Transformed Gaussian random fields for unsupervised image deconvolution (SPS Journal Paper)*

Jean-Baptiste Courbot (IRIMAS); Bruno Colicchio (IRIMAS - Université de Haute-Alsace)

IVMSP-P20: Representation Learning

Room: Dome K

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Ju Sun, Ye Peng

5941 (IVMSP-P20.1): Robust Autoencoders for Collective Corruption Removal

Taihui Li (University of Minnesota); Hengkang Wang (University of Minnesota); Le Peng (University of Minnesota); XianE Tang (University of Minnesota Duluth); Ju Sun (University of Minnesota)

6536 (IVMSP-P20.2): Learning Supervised Covariation Projection Through General Covariance

Xiangze Bao (Yangzhou University); Yunhao Yuan (Yangzhou University); Yun Li (Yangzhou University); Jipeng Qiang (Yangzhou University); Yi Zhu (Yangzhou University)

670 (IVMSP-P20.3): Composition of Motion From Video Animation Through Learning Local Transformations

Michalis Vrigkas (University of Western Macedonia); Virginia Tagka (University of Ioannina); Marina Plissiti (University of Ioannina); Christophoros Nikou (University of Ioannina)

776 (IVMSP-P20.4): VISION2TOUCH: IMAGING ESTIMATION OF SURFACE TACTILE PHYSICAL PROPERTIES

Jie Chen (Hunan University); ZHOU SHIZHE (Hunan University)

1959 (IVMSP-P20.5): Decaying Contrast for Fine-grained Video Representation Learning

Heng Zhang (Gaoling School of Artificial Intelligence, Renmin University of China); Bing Su (Renmin University of China)

2502 (IVMSP-P20.6): MMCOSINE: MULTI-MODAL COSINE LOSS TOWARDS BALANCED AUDIO-VISUAL FINE-GRAINED LEARNING

Ruize Xu (Renmin University of China); Ruoxuan Feng (Renmin University of China); Shixiong Zhang (Tencent); Di Hu (Renmin University of China)

1946 (IVMSP-P20.7): CDHD: CONTRASTIVE DREAMER FOR HINT DISTILLATION

yu le (Tsinghua University); Hua TongYan (Guangdong Bright Dream Robotics Co., Ltd.); Wenming Yang (Tsinghua University); Ye Peng (Guangdong Bright Dream Robotics Co., Ltd.); Qingmin Liao (Tsinghua University)

3444 (IVMSP-P20.8): RD-NAS: Enhancing One-shot Supernet Ranking Ability via Ranking Distillation from Zero-cost Proxies

Peijie Dong (School of Computer Science, National University of Defense Technology); Xin Niu (NUDT); Lujun Li (Chinese Academy of Sciences); ZHILIANG TIAN (National University of Defense Technology); Xiaodong Wang (National University of Defense Technology); Zimian Wei (School of Computer Science, National University of Defense Technology); Hengyue Pan (National University of Defense Technology); Dongsheng Li (School of Computer Science, National University of Defense Technology)

3597 (IVMSP-P20.9): MTFD : Multi-teacher Fusion Distillation For Compressed Video Action Recognition

Jinxin Guo (Inner Mongolia University); Jiaqiang Zhang (Inner Mongolia University); Shaojie Li (Inner Mongolia University); Xiaojing Zhang (Inner Mongolia University); Ming Ma (Inner Mongolia University)

4693 (IVMSP-P20.10): Clean Sample Guided Self-Knowledge Distillation For Image Classification

Jiyue Wang (South China University of Technology); Yanxiong Li (South China University of Technology); Qianhua He (SOUTH CHINA UNIVERSITY OF TECHNOLOGY); Wei Xie (South China University of Technology)

Wednesday, June 7

438 (IVMSP-P20.11): CORSD: Class-Oriented Relational Self Distillation

Muzhou Yu (Xi'an Jiaotong University); Sia Huat Tan (Tsinghua University); Kailu Wu (Tsinghua University); Runpei Dong (Xi'an Jiaotong University); Linfeng Zhang (Tsinghua University); Kaisheng Ma (Tsinghua University)

660 (IVMSP-P20.12): SELF-SUPERVISED AUDIO-VISUAL SPEECH REPRESENTATIONS LEARNING BY MULTIMODAL SELF-DISTILLATION

Jing-Xuan Zhang (University of Science and Technology of China); Genshun Wan (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China); Jia Pan (iFlytek Research); Jianqing Gao (iFLYTEK); Cong Liu (iFLYTEK Research)

MLSP-P21: Adversarial machine learning II

Room: Dome F

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

287 (MLSP-P21.1): Optimization for Robustness Evaluation beyond Lp Metrics

Hengyue Liang (University of Minnesota); Buyun Liang (University of Minnesota); Ying Cui (University of Minnesota); Tim Mitchell (Queens College / CUNY); Ju Sun (University of Minnesota)

5854 (MLSP-P21.2): Learning Unbiased Rewards with Mutual Information in Adversarial Imitation Learning

LiHua Zhang (School of Computer Science and Technology, Soochow University); Quan Liu (School of Computer Science and Technology, Soochow University); Zhigang Huang (School of Computer Science and Technology, Soochow University, Suzhou, China); Lan Wu (School of Computer Science and Technology, Soochow University)

2184 (MLSP-P21.3): BATT: Backdoor Attack with Transformation-based Triggers

Tong Xu (Tsinghua University); Yiming Li (Tsinghua University); Yong Jiang (Tsinghua University); Shu-Tao Xia (Tsinghua University)

987 (MLSP-P21.4): Backdoor Defense via Suppressing Model Shortcuts

Sheng Yang (Tsinghua University); Yiming Li (Tsinghua University); Yong Jiang (Tsinghua University); Shu-Tao Xia (Tsinghua University)

1036 (MLSP-P21.5): Exploiting One-class classification optimization objectives for increasing adversarial robustness

Vasileios Mygdalis (Aristotle University of Thessaloniki); Ioannis Pitas (Aristotle University of Thessaloniki)

1354 (MLSP-P21.6): Untargeted Backdoor Attack against Object Detection

Chengxiao Luo (Tsinghua University); Yiming Li (Tsinghua University); Yong Jiang (Tsinghua University); Shu-Tao Xia (Tsinghua University)

1581 (MLSP-P21.7): DEFENDING AGAINST UNIVERSAL PATCH ATTACKS BY RESTRICTING TOKEN ATTENTION IN VISION TRANSFORMERS

Hongwei Yu (University of Science and Technology Beijing); Jiansheng Chen (University of Science and Technology Beijing); Huimin Ma (University of Science and Technology Beijing); Cheng Yu (Tsinghua University); Xinlong Ding (University of Science and Technology Beijing)

2041 (MLSP-P21.8): SIMILARITY RELATION PRESERVING CROSS-MODAL LEARNING FOR MULTISPECTRAL PEDESTRIAN DETECTION AGAINST ADVERSARIAL ATTACKS

Jung Uk Kim (Kyung Hee University); Yong Man Ro (KAIST)

3996 (MLSP-P21.9): Multi-Agent Adversarial Training Using Diffusion Learning

Ying Cao (École polytechnique fédérale de Lausanne - EPFL); Elsa Rizk (EPFL); Stefan Vlaski (Imperial College London); Ali H. Sayed (Ecole Polytechnique Fédérale de Lausanne)

205 (MLSP-P21.10): Adversarial Contrastive Distillation with Adaptive Denoising

Yuzheng Wang (Fudan University); Zhaoyu Chen (Fudan University); Dingkang Yang (Fudan University); Yang Liu (Fudan University); Siao Liu (Fudan University); Wenqiang Zhang (Fudan University); Lizhe Qi (Fudan University)

4530 (MLSP-P21.11): SafeDeep: A Scalable Robustness Verification Framework for Deep Neural Networks

Anahita Baninajjar (Lund University); Kamran Hosseini (Linköping University); Ahmed Rezine (Linköping University); Amir Aminifar (Lund University)

3335 (MLSP-P21.12): Visual Prompting for Adversarial Robustness

Aochuan Chen (Michigan State University); Peter Lorenz (Fraunhofer); Yuguang Yao (Michigan State University); Pin-Yu Chen (IBM Research); Sijia Liu (Michigan State University)

Wednesday, June 7

MLSP-P22: Self-supervised Learning Methods II

Room: Dome G

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Thuan Nguyen,

1374 (MLSP-P22.1): Augmentation Robust Self-Supervised Learning for Human Activity Recognition

Cong Xu (Amazon Inc.); Yuhang Li (Yale University); Dae Lee (Amazon); Dae Hoon Park (Amazon); Hongda Mao (Amazon Inc.); Huyen Do (Amazon Inc.); Jonathan Chung (Amazon); Dinesh Nair (Amazon Inc.)

2946 (MLSP-P22.2): SEMI-SUPERVISED LEARNING WITH PER-CLASS ADAPTIVE CONFIDENCE SCORES FOR ACOUSTIC ENVIRONMENT CLASSIFICATION WITH IMBALANCED DATA

Luan V. Fiorio (Eindhoven University of Technology); Boris Karanov (Eindhoven University of Technology); Johan David (NXP Semiconductors); Wim van Houtum (NXP Semiconductors); Frans Widdershoven (NXP Semiconductors); Ronald Aarts (Eindhoven University of Technology)

3100 (MLSP-P22.3): Self-supervised Facial Action Unit Detection with Region and Relation Learning

Juan Song (Tianjin University); Zhilei Liu (Tianjin University)

1828 (MLSP-P22.4): Training set cleansing of backdoor poisoning by self-supervised representation learning

Hang Wang (The Pennsylvania State University); Sahar Karimi (Meta); Ousmane A Dia (Meta); Hippolyt Ritter (Meta); Ehsan Emamjomeh-Zadeh (Meta); Jiahui Chen (Meta); Zhen Xiang (University of Illinois Urbana-Champaign); David Miller (Pennsylvania State University); George Kesidis (Penn State University)

1189 (MLSP-P22.5): NC-WAMKD: Neighborhood Correction Weight-Adaptive Multi-teacher Knowledge Distillation For Graph-based Semi-supervised Node Classification

Jiahao Liu (Xi'an Jiaotong University); pengcheng guo (Xi'an Jiaotong University); Yonghong Song (Xi'an Jiaotong University)

3665 (MLSP-P22.6): Robust Log-based Anomaly Detection with Hierarchical Contrastive Learning

Yuhui Zhao (Sichuan University); Ruichun Yang (The Chinese University of Hong Kong, Shenzhen); Ning Yang (Sichuan University); Tao LIN (Sichuan University); Qiuai Fu (HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.); YUCHI MA (HUAWEI CLOUD)

3315 (MLSP-P22.7): SuperCM: Revisiting Clustering for Semi-Supervised Learning

Durgesh K. Singh (UiT The Arctic University of Norway); Ahcène Boubekki (UiT - The Arctic University of Norway); Robert Jenssen (UiT - The Arctic University of Norway); Michael C. Kampffmeyer (UiT The Arctic University of Norway)

1833 (MLSP-P22.8): SUVR: A Search-based Approach to Unsupervised Visual Representation Learning

Yizhan Xu (National Cheng Kung University); Chih-Yao Chen (Academia Sinica); Cheng-Te Li (National Cheng Kung University)

4762 (MLSP-P22.9): Improving Self-Supervised Learning for Audio Representations by Feature Diversity and Decorrelation

Bac Nguyen (Sony Europe B.V.); Stefan Uhlich (Sony European Technology Center); Fabien Cardinaux (Sony European Technology Center)

3270 (MLSP-P22.10): Audio Barlow Twins: Self-Supervised Audio Representation Learning

Jonah Anton (Imperial College London); Harry Coppock (Imperial College London); Panoram Shukla (Imperial College London); Bjoern W. Schuller (Imperial College London)

6204 (MLSP-P22.11): Boosting Semi-Supervised Federated Learning with Model Personalization and Client-Variance-Reduction

Shuai Wang (Singapore University of Technology and Design); Yanqing Xu (The Chinese University of HongKong, Shenzhen); Yanli Yuan (Singapore University of Technology and Design); Xiuhua Wang (Huazhong University of Science and Technology); Tony Quek (Singapore University of Technology and Design)

615 (MLSP-P22.12): FFedCL: Fair Federated Learning with Contrastive Learning

xiaorong shi (nankai university); Liping Yi (Nankai University); Liu Xiaoguang (Nankai University); Wang Gang (Nankai University)

MLSP-P23: Learning Theory and Algorithms III

Room: Dome H

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

1538 (MLSP-P23.1): Output-Dependent Gaussian Process State-Space Model

Zhidi Lin (The Chinese University of Hong Kong, Shenzhen); Lei Cheng (Zhejiang University); Feng Yin (The Chinese University of Hong Kong, Shenzhen); Lexi Xu (China United Network Communications Corporation); Shuguang Cui (The Chinese University of Hong Kong, Shenzhen)

Wednesday, June 7

3717 (MLSP-P23.2): Diversifying Message Aggregation in Multi-Agent Communication via Normalized Tensor Nuclear Norm Regularization

Yuanzhao Zhai (National University of Defense Technology); Kele Xu (National Key Laboratory of Parallel and Distributed Processing (PDL)); Ding Bo (National University of Defense Technology); Dawei Feng (National University of Defense Technology); Zijian Gao (National University of Defense Technology); Huaimin Wang (National University of Defense Technology)

6527 (MLSP-P23.3): On weighted cross-entropy for label-imbalanced separable data: An algorithmic-stability study

Puneesh Deora (University of British Columbia); Christos Thrampoulidis (University of British Columbia)

2707 (MLSP-P23.4): Training Stronger Spiking Neural Networks with Biomimetic Adaptive Internal Association Neurons

Haibo Shen (Huazhong University of Science and Technology); Yihao Luo (Yichang Testing Technique R&D Institute); Xiang Cao (School of Computer Science and Technology, Huazhong University of Science and Technology); Liangqi Zhang (Huazhong University of Science and Technology); Juyu Xiao (Huazhong University of Science and Technology); Tianjiang Wang (School of Computer Science and Technology, Huazhong University of Science and Technology)

2911 (MLSP-P23.5): NOWCASTING OF EXTREME PRECIPITATION USING DEEP GENERATIVE MODELS

Haoran Bin (TU Delft); Max Kyryliuk (TU Delft); Zhiyi Wang (TU Delft); Cristian Meo (TUDelft); Yanbo Wang (TU Delft); Ruben Imhoff (Deltares); Remko Uijlenhoet (TU Delft); Justin Dauwels (TU Delft)

3159 (MLSP-P23.6): FAST SINGLE-PERSON 2D HUMAN POSE ESTIMATION USING MULTI-TASK CONVOLUTIONAL NEURAL NETWORKS

Christos Papaioannidis (Aristotle University of Thessaloniki); Ioannis Mademlis (Department of Informatics, Aristotle University of Thessaloniki); Ioannis Pitas (Aristotle University of Thessaloniki)

3766 (MLSP-P23.7): Intent Does Matter! Propagating High-order Relations for Exploring Interest Preferences

Xiangping Zheng (Renmin University of China); Xun Liang (Renmin University of China); Bo Wu (Renmin University of China); Junlan Feng (China Mobile Research Institute); Yuhui Guo (Renmin University of China); Sensen Zhang (Renmin University of China)

4120 (MLSP-P23.8): MCNet: Measurement-Consistent Networks via a Deep Implicit Layer for Solving Inverse Problems

Rahul Mourya (Heriot-Watt University); Joao F.C. Mota (Heriot-Watt University)

6214 (MLSP-P23.9): RUNTIME PREDICTION OF MACHINE LEARNING ALGORITHMS IN AUTOML SYSTEMS

Parijat Dube (IBM Research); Theodoros Salonidis (IBM T.J. Watson Research Center); Parikshit Ram (IBM Research); Ashish Verma (Amazon)

2244 (MLSP-P23.10): A Nested Ensemble Method to Bilevel Machine Learning

Lisha Chen (RENSSELAER POLYTECHNIC INST); Momin Abbas (Rensselaer Polytechnic Institute); Tianyi Chen (Rensselaer Polytechnic Institute)

885 (MLSP-P23.11): EXTENDED EXPECTATION MAXIMIZATION FOR UNDER-FITTED MODELS

Aref Miri Rekavandi (University of Western Australia); Abd-Krim Seghouane (University of Mebourne); Farid Boussaid (University of Western Australia); Mohammed Bennamoun (University of Western Australia)

2161 (MLSP-P23.12): LQGNet: Hybrid Model-Based and Data-Driven Linear Quadratic Stochastic Control

Solomon Goldgraber Casspi (Ben-Gurion University of the Negev); Oliver Husser (ETH Zurich); Guy Revach (ETH Zürich); Nir Shlezinger (Ben-Gurion University)

SAM-P2: Target Detection and Classification

Room: Dome M

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

910 (SAM-P2.1): Variational Message Passing-based Respiratory Motion Estimation and Detection Using Radar Signals

Jakob Möderl (Graz University of Technology); Erik Leitinger (Graz University of Technology); Franz Pernkopf (Graz University of Technology); Klaus Witrisal (Graz University of Technology, Austria)

1372 (SAM-P2.2): CD-FSOD: A Benchmark for Cross-domain Few-shot Object Detection

Wuti Xiong (University of Oulu, Finland)

2330 (SAM-P2.3): Dual-Stream Siamese Vision Transformer with Mutual Attention for Radar Gait Verification

Ran Ji (School of Computer Science, University of Nottingham Ningbo China); Jiarui Li (School of Computer Science, University of Nottingham Ningbo China); Wentao He (University of Nottingham Ningbo China); Jianfeng Ren (University of Nottingham Ningbo China); Xudong Jiang (Nanyang Technological University)

2404 (SAM-P2.4): ONE-SHOT MEDICAL ACTION RECOGNITION WITH A CROSS-ATTENTION MECHANISM AND DYNAMIC TIME WARPING

Leiyu Xie (Newcastle University); Yuxing Yang (Newcastle University); Zeyu Fu (University of Exeter); Syed Mohsen Naqvi (Newcastle University)

Wednesday, June 7

2839 (SAM-P2.5): Hypothesis test for leakage detection in water pipelines with high-dimensional sensor signals
Liusha Yang (Shenzhen Technology University); Matthew McKay (University of Melbourne); Xun Wang (Beihang University)

3474 (SAM-P2.6): MMWAVE WI-FI TRAJECTORY ESTIMATION WITH CONTINUOUS-TIME NEURAL DYNAMIC LEARNING
Cristian J Vaca Rubio (Aalborg University); Pu Wang (MERL); Toshiaki Koike-Akino (Mitsubishi Electric Research Laboratories); Ye Wang (Mitsubishi Electric Research Laboratories); Petros Boufounos (Mitsubishi Electric Research Laboratories); Petar Popovski (Aalborg University)

3975 (SAM-P2.7): AN AUTOMOTIVE RADAR DATASET FOR OBJECT CLASSIFICATION
Akshad Shyam (Indian Institute of Technology Hyderabad); Kusum K (IIT Hyderabad); Monika Gautam (Indian Institute of Technology); Vamshi Krishna Kancharla (IIIT Bangalore college); vennela gudisa (IIT HYDERABAD); Virendra Patil (Indian Institute of Technology Hyderabad); Aanandh S Balasubramanian (Intel); Sumohana S. Channappayya (IIT Hyderabad)

4572 (SAM-P2.8): A Radar-Jammer Zero-Sum Repeated Bayesian Game
Sofia Suvorova (The University of Melbourne); Ali Pezeshki (Colorado State University); Ross Kyprianou (Defence Science and Technology Group); William Moran (The University of Melbourne)

5266 (SAM-P2.9): DIRECT POSITION DETERMINATION WITH ONE-BIT SIGNAL FOR MULTIPLE TARGETS
Lihua Ni (University of Electronic Science and Technology of China); Di Zhang (University of Electronic Science and Technology of China); Tianyi Xing (University of Electronic Science and Technology of China); Maoyan Ran (University of Electronic and Technology of China); Ning Liu (Northern Institute of Electronic Equipment); Qun Wan (University of Electronic Science and Technology of China)

6754 (SAM-P2.10): Maximum Likelihood Algorithm for Time-Delay Based Multistatic Target Localization (SPS Journal Paper)*
Kuntal Panwar (Indian Institute of technology, Delhi); Prabhu Babu (IIT Delhi); Petre Stoica (Uppsala University)

6821 (SAM-P2.12): Bayesian Quickest Detection of Propagating Spatial Events (SPS Journal Paper)*
Tope Halme (Aalto University); Eyal Nitzan (Aalto University); Visa Koivunen (Aalto university)

SLT-P21: Multimodal Processing of Language and Language Systems I

Room: Dome C

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Jan Yenda Trmal,

567 (SLT-P21.1): Cross-Modal Mutual Learning for Cued Speech Recognition
Lei Liu (The Chinese University of Hong Kong, Shenzhen); Li Liu (Shenzhen Research Institute of Big Data, the chinese university of hong kong shenzhen)

1886 (SLT-P21.2): SLBERT: A NOVEL PRE-TRAINING FRAMEWORK FOR JOINT SPEECH AND LANGUAGE MODELING
Onkar Susladkar (Natter Labs); Prajwal Gatti (Dayananda Sagar College of Engineering); Santosh Kumar Yadav (Natter Labs)

2190 (SLT-P21.3): CROSS-MODAL ADVERSARIAL CONTRASTIVE LEARNING FOR MULTI-MODAL RUMOR DETECTION
Ting Zou (Soochow University); Zhong Qian (Soochow University); Peifeng Li (Soochow University); Qiaoming Zhu (Soochow University)

2884 (SLT-P21.4): MULTIPLE CONTRASTIVE LEARNING FOR MULTIMODAL SENTIMENT ANALYSIS
Xiaocui Yang (Northeastern University); Shi Feng (Northeastern University); Daling Wang (Northeastern University); Pengfei Hong (Singapore University of Technology and Design); Soujanya Poria (Singapore University of Technology and Design)

3666 (SLT-P21.0): token2vec: A Joint Self-Supervised Pre-training Framework Using Unpaired Speech and Text
Xianghu Yue (National University of Singapore); Junyi Ao (The Chinese University of Hong Kong (Shenzhen)); Xiaoxue Gao (National University of Singapore); Haizhou Li (The Chinese University of Hong Kong (Shenzhen))

3714 (SLT-P21.6): DAIS: THE DELFT DATABASE OF EEG RECORDINGS OF DUTCH ARTICULATED AND IMAGINED SPEECH
Bo Dekker (Department of Biomechanical Engineering, Delft University of Technology); Alfred Schouten (Department of Biomechanical Engineering, Delft University of Technology); Odette Scharenborg (Multimedia Computing Group, Delft University of Technology)

4409 (SLT-P21.7): A Token-level Contrastive Framework for Sign Language Translation
Biao Fu (Xiamen University); Peigen Ye (Xiamen University); liang zhang (Xiamen University); Pei Yu (Xiamen University); Cong Hu (Xiamen University); xiaodong shi (xiamen university); Yidong Chen (Xiamen University)

4801 (SLT-P21.8): SIGN LANGUAGE RECOGNITION VIA DEFORMABLE 3D CONVOLUTIONS AND MODULATED GRAPH CONVOLUTIONAL NETWORKS
Katerina Papadimitriou (University of Thessaly); Gerasimos Potamianos (ECE, University of Thessaly)

Wednesday, June 7

4837 (SLT-P21.9): LAST: Scalable Lattice-Based Speech Modelling in JAX

Ke Wu (Google); Ehsan Variansi (Google); Tom Bagby (Google); Michael Riley (Google)

4989 (SLT-P21.10): M-SpeechCLIP: Leveraging Large-Scale, Pre-Trained Models for Multilingual Speech to Image Retrieval

Layne Berry (University of Texas at Austin); Yi-Jen Shih (National Taiwan University); Hsuan-Fu Wang (Institute of Information Science, Academia Sinica; National Taiwan University); Heng-Jui Chang (Massachusetts Institute of Technology); Hung-yi Lee (National Taiwan University); David Harwath (The University of Texas at Austin)

5014 (SLT-P21.11): USING EMOTION EMBEDDINGS TO TRANSFER KNOWLEDGE BETWEEN EMOTIONS, LANGUAGES, AND ANNOTATION FORMATS

Georgios Chochlakis (University of Southern California); Girish M Mahajan (Microsoft); Sabyasachee Baruah (University of Southern California); Keith Burghardt (ISI, University of Southern California); Kristina Lerman (USC Information Sciences Institute); Shrikanth Narayanan (USC)

5146 (SLT-P21.12): SPEECH-TEXT BASED MULTI-MODAL TRAINING WITH BIDIRECTIONAL ATTENTION FOR IMPROVED SPEECH RECOGNITION

Yuhang Yang (School of Information Science and Engineering, Xinjiang University, China); Haihua Xu (Temasek Laboratories, Nanyang Technological University, Singapore); Hao Huang (Xinjiang University); Eng Siong Chng (Nanyang Technological University); Sheng Li (National Institute of Information & Communications Technology (NICT))

SLT-P22: Natural Language Processing II

Room: Dome D

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Ji Wu,

720 (SLT-P22.1): Improving Sentence Similarity Estimation for Unsupervised Extractive Summarization

Shichao Sun (The Hong Kong Polytechnic University); Ruifeng Yuan (The Hong Kong Polytechnic University); Wenjie Li (Department of Computing, the Hong Kong Polytechnic University); Sujian Li (Peking University)

767 (SLT-P22.2): Contextual Similarity is More Valuable than Character Similarity: An Empirical Study for Chinese Spell Checking

Ding Zhang (Tsinghua University); Yinghui Li (Tsinghua University); Qingyu Zhou (OPPO Research Institute); Shirong Ma (Tsinghua University); Li Yangning (Tsinghua Shenzhen International Graduate School); Yunbo Cao (Tencent); Hai-Tao Zheng (Tsinghua University)

770 (SLT-P22.3): Prompt-Distiller: Few-shot Knowledge Distillation for Prompt-based Language Learners with Dual Contrastive Learning

Boyu Hou (Chongqing University); Chengyu Wang (Alibaba); Xiaoqing Chen (Chongqing University); Minghui Qiu (Alibaba); Liang Feng (Chongqing University, China); Jun Huang (Alibaba Group)

831 (SLT-P22.4): Matching-based Term Semantics Pre-training for Spoken Patient Query Understanding

Zefa Hu (Institute of Automation, Chinese Academy of Sciences); Xiuyi Chen (Institute of Automation, Chinese Academy of Science); Haoran Wu (Institute of Automation, Chinese Academy of Sciences); Minglun Han (Institute of Automation, Chinese Academy of Sciences); Ni Ziyi (CASIA); Jing Shi (Institute of Automation Chinese Academy of Sciences); Shuang Xu (casia); Bo Xu (Institute of Automation, Chinese Academy of Sciences)

2167 (SLT-P22.5): Contrastive Learning of Sentence Embeddings in Product Search

Bo-Wen Zhang (Beijing Academy of Artificial Intelligence); Yan Yan (CUMTB); Jiawei Yu (Alibaba Group)

2342 (SLT-P22.6): Dual Path Modeling for Semantic Matching by Perceiving Subtle Conflicts

Chao Xue (Beihang University); Di Liang (Centre for Natural Language Processing, Meituan Inc., Beijing, China); Sirui Wang (Centre for Natural Language Processing, Meituan Inc., Beijing, China); Jing Zhang (Beihang University); Wei Wu (Centre for Natural Language Processing, Meituan Inc., Beijing, China)

2552 (SLT-P22.7): PROMPT MAKES MASK LANGUAGE MODELS BETTER ADVERSARIAL ATTACKERS

He Zhu (Institute of Information Engineering, Chinese Academy of Sciences); Ce Li (Institute of Information Engineering, Chinese Academy of Sciences); Haitian Yang (Institute of Information Engineering, Chinese Academy of Sciences); Yan Wang (Institute of Information Engineering, Chinese Academy of Sciences); Weiqing Huang (Institute of Information Engineering, Chinese Academy of Sciences)

2566 (SLT-P22.8): Towards Polymorphic Adversarial Examples Generation for Short Text

Yuhang Liang (University of Chinese Academy of Science); Zheng Lin (iie); Fengcheng Yuan (UCAS, IIE); Hanwen Zhang (UCAS, IIE); Lei Wang (Institute of Information Engineering, Chinese Academy of Sciences); Weiping Wang (Institute of Information Engineering, CAS, China)

4017 (SLT-P22.9): Knowledge-augmented Frame Semantic Parsing with Hybrid Prompt-tuning

Rui Zhang (Artificial Intelligence Application Research Center, Huawei Technologies); yajing sun (huawei); JINGYUAN YANG (Artificial Intelligence Application Research Center, Huawei Technologies); Wei Peng (Huawei Technologies)

Wednesday, June 7

5339 (SLT-P22.10): BERT is Robust! A Case Against Word Substitution-based Adversarial Attacks

Jens Hauser (ETHZ); Zhao Meng (ETHZ); Damian Pascual (ETHZ); Roger Wattenhofer (ETH Zurich)

5415 (SLT-P22.11): LEARNING TO BUILD REASONING CHAINS BY RELIABLE PATH RETRIEVAL

Minjun Zhu (CASIA); Yixuan Weng (CASIA); Shizhu He (Institute of Automation, Chinese Academy of Sciences); Kang Liu (Institute of Automation, Chinese Academy of Sciences); Jun Zhao (Institute of Automation, Chinese Academy of Sciences)

6512 (SLT-P22.12): G2PL: Lexicon Enhanced Chinese Polyphone Disambiguation using BERT Adapter with a New Dataset

Haifeng Zhao (Anhui University); Hongzhi Wan (Anhui University); Lili Huang (Anhui University; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center); Mingwei Cao (Anhui University)

SLT-P23: Natural Language Processing III

Room: Dome E

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Samridhi Choudhary,

4798 (SLT-P23.1): Conversational Text-to-SQL: An Odyssey into State-of-the-Art and Challenges Ahead

Sree Hari Krishnan Parthasarathi (Amazon); Lu Zeng (Amazon); Dilek Z Hakkani-Tur (Amazon Alexa AI)

381 (SLT-P23.2): CommDRE: Document-level Relation Extraction with Self-supervised Commonsense Learning

Rongzhen Li (Chongqing University); Jiang Zhong (); Zhongxuan Xue (Chongqing University); Qizhu Dai (Chongqing University); Chen Wang (Chongqing University); Xue Li (University of Queensland)

2056 (SLT-P23.3): Relational Representation Learning for Zero-shot Relation Extraction with Instance Prompting and Prototype Rectification

Bin Duan (Beijing University of Posts and Telecommunications); Xingxian Liu (Beijing University of Posts and Telecommunications); Shusen Wang (Beijing University of Posts and Telecommunications); Yajing Xu (Beijing University of Posts and Telecommunications); Bo Xiao (Beijing University of Posts and Telecommunications)

2793 (SLT-P23.4): SPTEAE: A SOFT PROMPT TRANSFER MODEL FOR ZERO-SHOT CROSS-LINGUAL EVENT ARGUMENT EXTRACTION

Huipeng Ma (National Computer System Engineering Research Institute of China); qiu tang (National Computer System Engineering Research Institute of China); ni zhang (National Computer System Engineering Research Institute of China); Rui Xu (National Computer System Engineering Research Institute of China); Yanhua Shao (National Computer System Engineering Research Institute of China); Wei Yan (National Computer System Engineering Research Institute of China); Yaojun Wang (China Agricultural University)

3023 (SLT-P23.5): Multi-task Transformer with Relation-attention and Type-attention for Named Entity Recognition

Ying Mo (Beihang University); Hongyin Tang (Meituan); Jiahao Liu (Meituan); Qifan Wang (Meta AI); Zenglin Xu (Harbin Institute of Technology, Shenzhen); Jingang Wang (Meituan); Wei Wu (Meituan); Zhoujun Li (Beihang University)

3189 (SLT-P23.6): TableE: Capturing the Interactions among Sub-tasks in Information Extraction via Double Tables

Jiaxing Lin (Peking University); Runxin Xu (Peking University); Baobao Chang (Peking University)

3252 (SLT-P23.7): SENER: Sentiment Element Named Entity Recognition for Aspect-based Sentiment Analysis

Sun-Kyung Lee (KAIST); Jong-Hwan Kim (KAIST)

4115 (SLT-P23.8): Meeting Action Item Detection with Regularized Context Modeling

Jiaqing Liu (Speech Lab, Alibaba Group); Chong Deng (Alibaba Inc); Qinglin Zhang (Speech Lab, Alibaba Group); Qian Chen (Speech Lab, DAMO Academy, Alibaba Group); Wen Wang (Alibaba Group)

4348 (SLT-P23.9): Gaussian Prior Reinforcement Learning for Nested Named Entity Recognition

Yawen Yang (Tsinghua University); Xuming Hu (Tsinghua University); Fukun Ma (Tsinghua University); Shuang Li (Tsinghua University); Aiwei Liu (Tsinghua University); Lijie Wen (Tsinghua University); Philip S Yu (UIC)

5290 (SLT-P23.10): Knowledge-aware Few Shot Learning for Event Detection from Short Texts

Jinjin Guo (JD Intelligent Cities Research); Zhichao Huang (JD Intelligent Cities Research); Guangning Xu (Harbin Institute of Technology, Shenzhen ▲); Bowen Zhang (Shenzhen Technology University); Chaoqun Duan (JD AI Research)

5784 (SLT-P23.11): RECOUPLE EVENT FIELD VIA PROBABILISTIC BIAS FOR EVENT EXTRACTION

Xingyu Bai (Tsinghua University); Taiqiang Wu (Tsinghua University); Han Guo (Tencent); Zhe Zhao (Tencent); Xuefeng Yang (Tencent); Jiayi Li (Tsinghua University); Weijie Liu (Tencent Inc.); Qi JU (Tencent); weigang guo (Tencent); Yujiu Yang (Tsinghua University)

6081 (SLT-P23.12): DOCDRE-FE: A DOCUMENT-LEVEL FINE-GRAINED ENTITY AND RELATION EXTRACTION DATASET

Hongbo Wang (Peking University); Weimin Xiong (Peking University); Yifan Song (Peking University); Dawei Zhu (Peking University); Yu Xia (Peking University); Sujian Li (Peking University)

Wednesday, June 7

SS-P1: Spatial Processing for Audio and Speech

Room: Dome A

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Walter Kellermann, Timo Gerkman

1753 (SS-P1.1): ON MULTIPLE-INPUT/BINAURAL-OUTPUT ANTIPHASIC SPEAKER SIGNAL EXTRACTION

Xianrui Wang (Northwestern Polytechnical University); Ningning Pan (Northwestern Polytechnical University); Jacob Benesty (INRS); Jingdong Chen (Northwestern Polytechnical University)

2357 (SS-P1.2): MODEL-MATCHING PRINCIPLE APPLIED TO THE DESIGN OF AN ARRAY-BASED ALL-NEURAL BINAURAL RENDERING SYSTEM FOR AUDIO TELEPRESENCE

Yicheng Hsu (National Tsing Hua University); Chenghung Ma (National Tsing Hua University); Mingsian Bai (National Tsing Hua University)

4766 (SS-P1.3): Beamformer-Guided Target Speaker Extraction

Mohamed Elminshawi (International Audio Laboratories Erlangen); Srikanth Raj Chetupalli (Fraunhofer IIS); Emanuel Habets (AudioLabs Erlangen)

2807 (SS-P1.04): Streaming Multi-channel Speech Separation with Online Time-domain Generalized Wiener Filter

Yi Luo (Tencent AI Lab)

3112 (SS-P1.5): Multi-Microphone Speaker Separation by Spatial Regions

Julian Wechsler (AudioLabs Erlangen); Srikanth Raj Chetupalli (Fraunhofer IIS); Wolfgang Mack (AudioLabs Erlangen); Emanuel Habets (AudioLabs Erlangen)

3253 (SS-P1.6): Exploiting spatial information with the informed complex-valued spatial autoencoder for target speaker extraction

Annika Briegleb (Friedrich-Alexander-University Erlangen-Nürnberg); Mhd Modar Halimeh (Friedrich-Alexander-University Erlangen-Nürnberg); Walter Kellermann (Friedrich-Alexander-University Erlangen-Nürnberg)

3949 (SS-P1.7): McNet: Fuse Multiple Cues for Multichannel Speech Enhancement

Yujie Yang (Westlake University); Changsheng Quan (Westlake University); Xiaofei Li (Westlake University)

4787 (SS-P1.8): Spatially Selective Deep Non-linear Filters for Speaker Extraction

Kristina Tesch (Universität Hamburg); Timo Gerkman (Universität Hamburg)

3300 (SS-P1.9): Learning Audio-Visual Dereverberation

Changan Chen (University of Texas at Austin); Wei Sun (University of Texas at Austin); David Harwath (The University of Texas at Austin); Kristen Grauman (Facebook AI Research & UT Austin)

2094 (SS-P1.10): Decorrelating language model embeddings for speech-based prediction of cognitive impairment

Lingfeng Xu (Arizona State University); Kimberly D. Mueller (University of Wisconsin-Madison); Julie Liss (Arizona State University); Visar Berisha (Arizona State University)

BISP-L2: Brain Computer Interfaces

Room: Nafsika A

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Toshihasa Tanaka,

03:40 PM

511 (BISP-L2.1): Two-Phase Prototypical Contrastive Domain Generalization for Cross-Subject EEG-Based Emotion Recognition

Honghua Cai (South China Normal University); Jiahui Pan (South China Normal University)

03:55 PM

1049 (BISP-L2.2): EEG2IMAGE: Image Reconstruction from EEG Brain Signals

Prajwal Singh (Indian Institute of Technology Gandhinagar, Gujarat, India); Pankaj Pandey (Indian Institute of Technology Gandhinagar); Krishna P Miyapuram (Indian Institute of Technology, Gandhinagar, India); Shanmuganathan Raman (Indian Institute of Technology (IIT) Gandhinagar)

04:10 PM

4918 (BISP-L2.3): Applying Independent Vector Analysis on EEG-based motor imagery classification

Caroline P. A. Moraes (Federal University of ABC (UFABC)); Bruno Aristimunha (Federal University of ABC); Lucas Heck dos Santos (UFABC); Walter Hugo Lopez Pinaya (King's College London); Raphael Y de Camargo (UFABC); Denis Fantinato (Federal University of ABC); Aline Neves (Federal University of ABC)

Wednesday, June 7

04:25 PM

6193 (BISP-L2.4): Synthesizing Speech from ECoG with a Combination of Transformer-based Encoder and Neural Vocoder
Kai Shigemori (Tokyo University of Agriculture and Technology); Shuji Komeiji (Tokyo University of Agriculture and Technology); Takumi Mitsuhashi (Juntendo University School of Medicine); Yasushi Jimura (Juntendo University School of Medicine); Hiroharu Suzuki (Juntendo University School of Medicine); Hidenori Sugano (Juntendo University School of Medicine); Koichi SHINODA (Tokyo Institute of Technology); Kohei Yatabe (Tokyo University of Agriculture and Technology); Toshihisa Tanaka (Tokyo University of Agriculture and Technology)

04:40 PM

1667 (BISP-L2.5): Unbiased unsupervised stimulus reconstruction for EEG-based auditory attention decoding
Nicolas Heintz (KU Leuven); Simon Geirnaert (KU Leuven); Tom Francart (KU Leuven); Alexander Bertrand (KU Leuven)

04:55 PM

5180 (BISP-L2.6): A LEARNABLE SPATIAL MAPPING FOR DECODING THE DIRECTIONAL FOCUS OF AUDITORY ATTENTION USING EEG

Yuanming Zhang (Nanjing University); Haoxin Ruan (Nanjing University); Ziyang Yuan (Nanjing University); Haoliang Du (Nanjing Drum Tower Hospital); Xia Gao (Nanjing Drum Tower Hospital); Jing Lu (Nanjing University)

GC-7: Acoustic Echo Cancellation Signal Processing Grand Challenge 2023

Room: Nefeli B

Type: Oral

03:40 PM to 05:10 PM

Chair(s): TBA

03:40 PM

6653 (GC-L7.1): Introduction

Ross Cutler (Microsoft Corporation); Ando Saabas (Microsoft); Tanel Parnamaa (Microsoft); Marju Purin (Microsoft); Evgenii Indenbom (Microsoft); Ristea N Catalin (Microsoft); Jegor Guzhvin (Microsoft); Hannes Gamper (Microsoft); Sebastian Braun (Microsoft); Robert Aichner (Microsoft)

04:00 PM

6852 (GC-L7.2): MULTI-TASK SUB-BAND NETWORK FOR DEEP RESIDUAL ECHO SUPPRESSION

Jiayao Sun (Northwestern Polytechnical University); Dawei Luo (Li Auto); Zhaoxia Li (Li Auto); Jingdong Li (Tencent); Yukai Jv (Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, School of Computer Science, Northwestern Polytechnical University); Yang Li (Li Auto)

04:12 PM

6856 (GC-L7.3): A Progressive Neural Network for Acoustic Echo Cancellation

Zhuangqi Chen (South China University of Technology); Xianjun Xia (RTC Lab, ByteDance); Siyu Sun (Wuhan University); Ziqian Wang (Northwestern Polytechnical University); Cheng Chen (ByteDance); Guoliang Xie (ByteDance); Pingjiang Zhang (South China University of Technology); Yijian Xiao (ByteDance)

04:24 PM

6858 (GC-L7.4): A TAYLOR_STYLE_NEURAL_NETWORK_IN_FULLBAND_ECHO_CANCELLATION

Xu Weiming (Northwest Polytechnic University); Guo Zhihao (elevoc)

04:36 PM

6869 (GC-L7.5): A Low-Latency Deep Hierarchical Fusion Network for Fullband Acoustic Echo Cancellation

Haoran Zhao (Kuaishou Technology); Nan Li (北京达佳互联信息技术有限公司); Runqiang Han (kuaishou); Xiguang Zheng (北京达佳互联信息技术有限公司); Chen Zhang (北京达佳互联信息技术有限公司)

04:48 PM

6899 (GC-L7.6): TWO-STEP BAND-SPLIT NEURAL NETWORK APPROACH FOR FULL-BAND RESIDUAL ECHO SUPPRESSION

Zihan Zhang (Northwestern Polytechnical University); Shimin Zhang (Northwestern Polytechnical University); Mingshuai Liu (NWPU); Yanhong Leng (Bytedance Inc); Zhe Han (ByteDance); Li Chen (ByteDance); Lei Xie (NWPU)

MLSP-L7: Learning Theory and Algorithms I

Room: Jupiter

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Qing Qu,

03:40 PM

226 (MLSP-L7.1): SD-PINN: Physics informed neural networks for spatially dependent PDEs

Ruixian Liu (University of California, San Diego); Peter Gerstoft (University of California San Diego)

Wednesday, June 7

03:55 PM

4484 (MLSP-L7.2): Learning Properties of Holomorphic Neural Networks of Dual Variables

Dmitry Kozlov (Huawei RRI); Mikhail V Bakulin (Huawei RRI); Stanislav Pavlov (HSE); Aleksandr Zuev (Huawei); Maria Krylova (Huawei RRI); Igor Kharchikov (Huawei)

04:10 PM

3402 (MLSP-L7.3): Active Subsampling Using Deep Generative Models by Maximizing Expected Information Gain

Koen van de Camp (Eindhoven University of Technology); Hamdi Joudeh (Eindhoven University of Technology); Duarte Antunes (Eindhoven University of Technology); Ruud J. G. van Sloun (Technical university of Eindhoven)

04:25 PM

3374 (MLSP-L7.4): ADAPTIVE STEP-SIZE METHODS FOR COMPRESSED SGD

Adarsh Muthuveeru-Subramaniam (University of Illinois at Urbana-Champaign); Akshayaa Magesh (University of Illinois at Urbana-Champaign); Venugopal V. Veeravalli (University of Illinois at Urbana Champaign)

04:40 PM

1421 (MLSP-L7.5): Designing Transformer networks for sparse recovery of sequential data using deep unfolding

Brent De Weerd (Vrije Universiteit Brussel); Yonina Eldar (); Nikos Deligiannis (Vrije Universiteit Brussel - imec)

04:55 PM

1600 (MLSP-L7.6): LEVERAGING SPARSITY WITH SPIKING RECURRENT NEURAL NETWORKS FOR ENERGY-EFFICIENT KEYWORD SPOTTING

Manon Dampfhofer (SPINTEC University Grenoble Alpes); Thomas Mesquida (CEA LIST); Emmanuel Hardy (CEA-Leti); Alexandre Valentin (CEA-List); Lorena Anghel (SPINTEC University Grenoble Alpes)

SAM-L1: DoA Estimation

Room: Salon de Roses A

Type: Oral

03:40 PM to 05:10 PM

Chair(s): TBA

03:40 PM

1362 (SAM-L1.01): UNITARY ESPRIT FOR COPRIME ARRAYS

Po-Chih Chen (California Institute of Technology); Dr.P P Vaidyanathan (California Institute of Technology)

03:55 PM

6750 (SAM-L1.02): Coprime Nested Arrays for DOA Estimation: Exploiting the Nesting Property of Coprime Array (SPS Journal Paper)*

Shiwei Ren (Beijing Institute of Technology); Zhe Peng (Beijing Institute of Technology)

04:10 PM

1813 (SAM-L1.03): EXPLICIT ZIV-ZAKAI BOUND FOR MULTIPLE SOURCES DOA ESTIMATION

Zongyu Zhang (Zhejiang University); Yujie Gu (Aptiv); Zhiguo Shi (Zhejiang University)

04:25 PM

3214 (SAM-L1.04): Tensorized Neural Layer Decomposition for 2-D DOA Estimation

Hang Zheng (Zhejiang University); Chengwei Zhou (Zhejiang University); Sergiy A. Vorobyov (Aalto University); Zhiguo Shi (Zhejiang University)

04:40 PM

6272 (SAM-L1.05): A DNN BASED NORMALIZED TIME-FREQUENCY WEIGHTED CRITERION FOR ROBUST WIDEBAND DOA ESTIMATION

Kuan-Lin Chen (University of California San Diego); Ching-Hua Lee (University of California, San Diego); Bhaskar Rao (UC San Diego); Harinath Garudadri (University of California, San Diego)

04:55 PM

1709 (SAM-L1.06): Direction-of-arrival estimation using Gaussian process interpolation

Ishan D Khurjekar (Scripps Institute of Oceanography); Peter Gerstoft (UCSD); Christoph F Mecklenbräuer (Technische Universität Wien); Zoi-Heleni Michalopoulou (New Jersey Institute of Technology)

SLT-L13: Speaker Recognition I: Scoring, Fairness, Privacy

Room: Delphi

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Jahangir Alam, Man-Wai MaK

03:40 PM

1890 (SLT-L13.1): A Study on Bias and Fairness In Deep Speaker Recognition

Amirhossein Hajavi (Queen's University); Ali Etemad (Queen's University)

Wednesday, June 7

03:55 PM

2522 (SLT-L13.2): Privacy-preserving Automatic Speaker Diarization

Francisco Teixeira (INESC-ID/IST, University of Lisbon); Alberto Abad (INESC-ID); Bhiksha Raj (Carnegie Mellon University); Isabel Trancoso (INESC ID)

04:10 PM

2862 (SLT-L13.3): Improving learning objectives for speaker verification from the perspective of score comparison

Min Hyun Han (Seoul National University); Sung Hwan Mun (Seoul National University); Minchan Kim (Seoul National University); Myeonghun Jeong (Seoul National University); Sunghwan Ahn (Seoul National University); Nam Soo Kim (Seoul National University)

04:25 PM

4499 (SLT-L13.4): Covariance Regularization for Probabilistic Linear Discriminant Analysis

ZHIYUAN PENG (CUHK); Mingjie Shao (The Chinese University of Hong Kong, Shandong University); Xuanji He (meituan); Xu Li (ARC Lab, Tencent); Tan Lee (The Chinese University of Hong Kong); Ke Ding (meituan); Guanglu Wan (Meituan)

04:40 PM

5788 (SLT-L13.5): Toroidal Probabilistic Spherical Discriminant Analysis

Anna Silnova (Brno University of Technology); Niko Brummer (Amazon); Albert DP Swart (Speechly); Lukáš Burget (Brno University of Technology)

04:55 PM

6080 (SLT-L13.6): INCORPORATING UNCERTAINTY FROM SPEAKER EMBEDDING ESTIMATION TO SPEAKER VERIFICATION

Qiongqiong Wang (ASTAR); Kong Aik Lee (Institute for Infocomm Research, ASTAR); Tianchi Liu (Institute for Infocomm Research, ASTAR)

SLT-L14: Speaker Recognition II: Verification, Diarization, Anti-spoofing

Room: Athina

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Massimiliano Todisco, Jenthe Thienpondt

03:40 PM

826 (SLT-L14.1): Advancing the dimensionality reduction of speaker embeddings for speaker diarisation: disentangling noise and informing speech activity

You Jin Kim (Naver Corporation); Heesoo Heo (Naver Corp.); Jee-weon Jung (Naver Corporation); Youngki Kwon (Naver Corporation); Bong-Jin Lee (Naver Corporation); Joon Son Chung (KAIST)

03:55 PM

2833 (SLT-L14.2): MULTI-SPEAKER AND WIDE-BAND SIMULATED CONVERSATIONS AS TRAINING DATA FOR END-TO-END NEURAL DIARIZATION

Federico Landini (Brno University of Technology); Mireia Diez (Brno University of Technology); Alicia Lozano-Diez (Universidad Autonoma de Madrid); Lukáš Burget (Brno University of Technology)

04:10 PM

3059 (SLT-L14.3): PUSHING THE LIMITS OF SELF-SUPERVISED SPEAKER VERIFICATION USING REGULARIZED DISTILLATION FRAMEWORK

Yafeng Chen (Speech Lab, Alibaba Group); Siqi Zheng (Alibaba Group); Hui Wang (Speech Lab, Alibaba Group); Luyao Cheng (Speech Lab, Alibaba Group); Qian Chen (Speech Lab, DAMO Academy, Alibaba Group)

04:25 PM

3303 (SLT-L14.4): Can spoofing countermeasure and speaker verification systems be jointly optimised?

Wanying Ge (EURECOM); Hemlata Tak (EURECOM); Massimiliano Todisco (EURECOM); Nicholas Evans (EURECOM)

04:40 PM

4991 (SLT-L14.5): ASSD: SYNTHETIC SPEECH DETECTION IN THE AAC COMPRESSED DOMAIN

Amit Kumar Singh Yadav (Purdue University); Ziyue Xiang (Purdue University); Emily Bartusiak (Purdue University); Paolo Bestagini (Politecnico di Milano); Stefano Tubaro (Politecnico di Milano, Italy); Edward Delp (Purdue University)

04:55 PM

5661 (SLT-L14.6): MARGIN-MIXUP: A METHOD FOR ROBUST SPEAKER VERIFICATION IN MULTI-SPEAKER AUDIO

Jenthe Thienpondt (IDLab, Ghent University); Nilesh Madhu (IDLab, Ghent University - imec); Kris Demuynck (Ghent University)

SS-L12: Recent Advances in Robust Learning for Modern Computational Imaging

Room: Nefeli A

Type: Oral

03:40 PM to 05:10 PM

Wednesday, June 7

Chair(s): M. Salman Asif,

03:40 PM

1037 (SS-L12.1): Provably Convergent Plug & Play Linearized ADMM, applied to Deblurring Spatially Varying Kernels
Charles Laroche (GoPro); Andres Almansa (CNRS & Université Paris Cité, MAP5); Eva Coupeté (GoPro); Matias Tassano (Meta Inc)

03:55 PM

2014 (SS-L12.2): Robust Data-Driven Accelerated Mirror Descent
Hong Ye Tan (University of Cambridge); Subhadip Mukherjee (University of Cambridge); Junqi Tang (University of Cambridge); Andreas Hauptmann (University of Oulu); Carola-Bibiane B Schönlieb (Cambridge University)

04:10 PM

3187 (SS-L12.3): Robustness of Deep Equilibrium Architectures to Changes in the Measurement Model
Junhao Hu (Wustl); Shirin Shoushtari (Washington University in St. Louis); Zihao Zou (Washington University in St. Louis); Jiaming Liu (Washington University in St. Louis); Zhixin Sun (Washington University in St. Louis); Ulugbek S. Kamilov (Washington University in St. Louis)

04:25 PM

3345 (SS-L12.4): A Variational Inequality Model for Learning Neural Networks
Patrick Combettes (); Jean-Christophe Pesquet (); Audrey Repetti (Heriot Watt University)

04:40 PM

3641 (SS-L12.5): Compressive Sensing with Tensorized Autoencoder
Rakib Hyder (University of California, Riverside); M. Salman Asif (University of California, Riverside)

04:55 PM

5624 (SS-L12.6): Image Reconstruction Without Explicit Priors
Angela F Gao (Caltech); Oscar Leong (California Institute of Technology); He Sun (Peking University); Katherine Bouman (Caltech)

SS-L16: Signal Processing and Machine Learning for Networked Autonomous Agents

Room: Salon de Roses B

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Siwei Zhang, Francesco Guidi, Anna Guerra

03:40 PM

1017 (SS-L16.1): Approximation Error Backtracking for Q-function in Scalable Reinforcement Learning with Tree Dependence Structure
Yuzi Yan (Tsinghua University); Yu Dong (Tsinghua University); Kai Ma (Tsinghua University); Yuan Shen (Tsinghua University)

03:55 PM

1656 (SS-L16.2): Implicit vehicle positioning with cooperative lidar sensing
Luca Barbieri (Politecnico di Milano); Bernardo Camajori Tedeschini (Politecnico di Milano); Mattia Brambilla (Politecnico di Milano); Monica Nicoli (Politecnico di Milano University)

04:10 PM

2153 (SS-L16.3): Distributed ADMM with Limited Communications via Deep Unfolding
Yoav Noah (Ben-Gurion University of the Negev); Nir Shlezinger (Ben-Gurion University)

04:25 PM

3232 (SS-L16.4): DRL Path Planning For UAV-Aided V2X Networks: Comparing Discrete To Continuous Action Spaces
Leonardo Spampinato (WiLab, CNIT / DEI, University of Bologna); Alessia Tarozzi (WiLab, CNIT / DEI, University of Bologna); Chiara Buratti (WiLab, CNIT / DEI, University of Bologna); Riccardo Marini (WiLab, CNIT / DEI, University of Bologna)

04:40 PM

4082 (SS-L16.5): Autonomous Navigation of a Robotic Swarm in Space Exploration Missions
Siwei Zhang (German Aerospace Center (DLR)); Tobias Baumgartner (German Aerospace Center (DLR)); Emanuel Staudinger (German Aerospace Center (DLR) e.V.); Robert Pöhlmann (DLR); Fabio Brogghammer (German Aerospace Center (DLR)); Armin Dammann (German Aerospace Center (DLR) e.V.)

04:55 PM

5010 (SS-L16.6): UWB Localization-of-Things via Soft Information: Network Experimentation in Indoor Environment
Carlos Antonio Gomez Vega (University of Ferrara); Moe Win (Massachusetts Institute of Technology, USA); Andrea Conti (University of Ferrara)

Wednesday, June 7

AASP-P3: Active Noise Control, Echo Reduction and Feedback Reduction

Room: Dome A

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Gerald Enzner,

4903 (AASP-P3.1): Neural-AFC: Learning-Based Step-Size Control for Adaptive Feedback Cancellation with Closed-loop Model Training

Behrad Soleimani (Starkey Hearing Technologies); Henning Schepker (Starkey Hearing Technologies); Majid Mirbagheri (Starkey Hearing Technologies)

2376 (AASP-P3.2): Speaker Diaphragm Excursion Prediction: deep attention and online adaptation

Yuwei Ren (Qualcomm AI Research, QUALCOMM Wireless Communication Technologies (China) Limited); Matt Zivney (Qualcomm AI Research, Qualcomm Technologies, Inc.); Yin Huang (Qualcomm); Eddie Choy (Qualcomm AI Research, Qualcomm Technologies, Inc.); Chirag Patel (Qualcomm); Hao Xu (Qualcomm AI Research, Qualcomm Technologies, Inc.)

2956 (AASP-P3.3): A practical distributed active noise control algorithm overcoming communication restrictions

Junwei Ji (Nanyang Technological University); Dongyuan Shi (Nanyang Technological University); Zhengding Luo (Nanyang Technological University); Xiaoyi Shen (Nanyang Technological University); Woon Seng Gan (NTU)

3595 (AASP-P3.4): Active Noise control over 3D space: A realistic error microphone geometry design

Huiyuan Sun (The Australian National University); Prasanga Samarasinghe (Australian National University); thushara abhayapala (The Australian National University)

267 (AASP-P3.5): RNN-based step-size estimation for the RLS algorithm with application to acoustic echo cancellation

Ofer Schwartz (CEVA Inc.); Ayal Schwartz (BIU)

4462 (AASP-P3.6): Low-Complexity Acoustic Echo Cancellation with Neural Kalman Filtering

Dong Yang (Tencent); Fei Jiang (Tencent); Wei Wu (Tencent); Xuefei Fang (Tencent); Muyong Cao (Tencent)

2656 (AASP-P3.7): SCA: STREAMING CROSS-ATTENTION ALIGNMENT FOR ECHO CANCELLATION

Yang Liu (Meta); Yangyang Shi (Facebook); Yun Li (Meta); Kaustubh Kalgaonkar (Meta); Sriram Srinivasan (Meta); Xin Lei (Meta)

5908 (AASP-P3.8): CENTRALIZED CASCADE MULTI-CHANNEL NOISE REDUCTION AND ACOUSTIC FEEDBACK CANCELLATION IN A WIRELESS ACOUSTIC SENSOR AND ACTUATOR NETWORK

Santiago Ruiz (KU Leuven); Toon van Waterschoot (Department of Electrical Engineering (ESAT-STADIUS/ETC)); Marc Moonen (Department of Electrical Engineering (ESAT-STADIUS), KU Leuven)

2920 (AASP-P3.9): Deep Generative Fixed-filter Active Noise Control

Zhengding Luo (Nanyang Technological University); Dongyuan Shi (Nanyang Technological University); Xiaoyi Shen (Nanyang Technological University); Junwei Ji (Nanyang Technological University); Woon Seng Gan (NTU)

4575 (AASP-P3.10): A Frequency-weighted Leaky FxLMS Algorithm with Application to Feedback Active Noise Control Systems

Yu Tang (Southwest Jiaotong University); Hongwei Zhang (Harbin Institute of Tech. Shenzhen)

5056 (AASP-P3.11): Deep AHS: A Deep Learning Approach to Acoustic Howling Suppression

Hao Zhang (Tencent AI Lab); Meng Yu (Tencent); Dong Yu (Tencent AI Lab)

5328 (AASP-P3.12): IMPROVING ACOUSTIC ECHO CANCELLATION BY MIXING SPEECH LOCAL AND GLOBAL FEATURES WITH TRANSFORMER

yajie liu (School of Computer Science, Wuhan University); Xinmeng Xu (Wuhan University); Weiping Tu (Wuhan University); Yuhong Yang (Wuhan University); Li Xiao (School of Computer Science, Wuhan University)

AASP-P4: Anomaly Detection and Representation Learning for Audio Classification

Room: Dome B

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Toni Heittola,

3680 (AASP-P4.1): Self-supervised learning of audio representations using angular contrastive loss

Shanshan Wang (Tampere University); Soumya Tripathy (Tampere University of Technology); Annamaria Mesaros (Tampere University)

5566 (AASP-P4.2): Unsupervised Anomaly Detection and Localization of Machine Audio: A GAN-based Approach

Anbai Jiang (Tsinghua University); Wei-Qiang Zhang (Tsinghua University); Yufeng Deng (Tsinghua University); Pingyi Fan (Tsinghua University); Jia Liu (Tsinghua University)

Wednesday, June 7

1773 (AASP-P4.3): MASKED MODELING DUO: LEARNING REPRESENTATIONS BY ENCOURAGING BOTH NETWORKS TO MODEL THE INPUT

Daisuke Niizumi (NTT Corporation); Daiki Takeuchi (NTT Corporation); Yasunori Ohishi (NTT Corporation); Noboru Harada (NTT); Kunio Kashino (NTT Communication Science Laboratories)

4215 (AASP-P4.4): SW-WaveNet: Learning Representation from Spectrogram and Wavegram Using WaveNet for Anomalous Sound Detection

Haihui Chen (Huazhong University of Science and Technology); Likai Ran (Huazhong University of Science and Technology); Xixia Sun (Nanjing University of Posts and Telecommunications); Chao Cai (Huazhong University of Science and Technology)

224 (AASP-P4.5): CLAP Learning Audio Concepts From Natural Language Supervision

Benjamin Elizalde (Microsoft); Soham Deshmukh (Microsoft); Mahmoud Al Ismail (Microsoft); Huaming Wang (Microsoft)

1329 (AASP-P4.6): Design Choices for Learning Embeddings from Auxiliary Tasks for Domain Generalization in Anomalous Sound Detection

Kevin Wilkinghoff (Fraunhofer FKIE)

3674 (AASP-P4.8): Joint Generative-Contrastive Representation Learning for Anomalous Sound Detection

Xiaomin Zeng (University of Science and Technology of China); Yan Song (USTC); ZHU ZHUO (alibaba); Yu Zhou (alibaba); Yuhong Li (Alibaba); hui xue (Alibaba); Lirong Dai (University of Science and Technology of China); Ian McLoughlin (Singapore Institute of Technology)

4002 (AASP-P4.9): Time-weighted Frequency Domain Audio Representation with GMM Estimator for Anomalous Sound Detection

Jian Guan (Harbin Engineering University); Youde Liu (Harbin Institute of Technology); Qiaoxi Zhu (University of Technology Sydney); 铁然 郑 (哈尔滨工业大学); jiqing Han (Harbin Institute of Technology); Wenwu Wang (University of Surrey)

4485 (AASP-P4.10): AN EFFECTIVE ANOMALOUS SOUND DETECTION METHOD BASED ON REPRESENTATION LEARNING WITH SIMULATED ANOMALIES

Han Chen (University of Science and Technology of China); Yan Song (USTC); ZHU ZHUO (alibaba); Yu Zhou (alibaba); Yuhong Li (Alibaba); hui xue (Alibaba); Ian McLoughlin (Singapore Institute of Technology)

IVMSP-P21: Data Processing

Room: Dome J

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Xuelong Li, Heinz Handels

4028 (IVMSP-P21.1): Unsupervised Feature Selection with Self-Weighted and L2,0-norm Constraint

Yongjin Yuan (Northwestern Polytechnical University); Zheng Wang (Xi'an Jiaotong University); Feiping Nie (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

4553 (IVMSP-P21.2): Semi-Supervised Domain Generalization with Graph-based Classifier

Minxiang Ye (Zhejianglab); Yifei Zhang (Zhejiang lab); Shiqiang Zhu (Zhejiang Lab); Anhuan Xie (ZhejiangLab, Zhejiang University); Senwei Xiang (ZhejiangLab)

6235 (IVMSP-P21.3): Towards Privacy and Utility in Tourette Tic Detection Through Pretraining Based on Publicly Available Video Data of Healthy Subjects

Nele Sophie Brügge (Universität zu Lübeck); Esfandiar Mohammadi (Universität zu Lübeck); Alexander Münchau (Universität zu Lübeck); Tobias Bäumer (Universität zu Lübeck); Christian Frings (Universität Trier); Christian Beste (Technische Universität Dresden); Veit Roessner (Technische Universität Dresden); Heinz Handels (University of Lübeck)

1729 (IVMSP-P21.4): Exploring instance relation for decentralized multi-source domain adaptation

Yikang Wei (Tianjin University); Yahong Han (Tianjin University)

2940 (IVMSP-P21.5): SL-MOE: A TWO-STAGE MIXTURE-OF-EXPERTS SEQUENCE LEARNING FRAMEWORK FOR FORECASTING RAPID INTENSIFICATION OF TROPICAL CYCLONE

Jian Xu (Beijing University of Posts and Telecommunications); Yang Lei (Beijing University of Posts and Telecommunications); Guangqi Zhu (Beijing University of Posts and Telecommunications); Yunling Feng (Beijing University of Posts and Telecommunications); Bo Xiao (Beijing University of Posts and Telecommunications); Qifeng Qian (National Meteorological Center of China); Yajing Xu (Beijing University of Posts and Telecommunications)

4301 (IVMSP-P21.6): COMBINING LOSS REWEIGHTING AND SAMPLE RESAMPLING FOR LONG-TAILED INSTANCE SEGMENTATION

Yaochi Zhao (Hainan University); Sen Chen (Hainan University); Qiong Chen (Hainan University); Zhuhua Hu (Hainan University)

Wednesday, June 7

2059 (IVMSP-P21.7): BiSVP: Building Footprint Extraction via Bidirectional Serialized Vertex Prediction

Mingming Zhang (Beihang University); Ye Du (Beihang University); Zhenghui Hu (Hangzhou Innovation Institute, Beihang University); Qingjie Liu (State Key Laboratory of Virtual Reality Technology and System, Beihang University, Beijing 100191, China); Yunhong Wang (State Key Laboratory of Virtual Reality Technology and System, Beihang University, Beijing 100191, China)

1580 (IVMSP-P21.8): MHSCNet: A Multimodal Hierarchical Shot-aware Convolutional Network for Video Summarization

Wujiang Xu (Xi'an Jiaotong University); Runzhong Wang (Shanghai Jiao Tong University); xiaobo guo (antgroup); Shaoshuai Li (Ant Group); Qiongxiu Ma (Ant Group); Yunan Zhao (Ant Group); Sheng Guo (Ant Group); Zhenfeng Zhu (bjtu); Junchi Yan (Shanghai Jiao Tong University)

1651 (IVMSP-P21.9): Mutual Information based Reweighting for Precipitation Nowcasting

Yuan Cao (Fudan University); Danchen Zhang (Pittsburgh University); Xin Zheng (Fudan University); Hongming Shan (Fudan University); Junping Zhang (Fudan University)

1380 (IVMSP-P21.10): Solving Jigsaw Puzzle of Large Eroded Gaps Using Puzzlet Discriminant Network

Xingke Song (University of Nottingham Ningbo China); Xiaoying YANG (University of Nottingham Ningbo China); Jianfeng Ren (University of Nottingham Ningbo China); RUIBIN BAI (University of Nottingham); Xudong Jiang (Nanyang Technological University)

6810 (IVMSP-P21.11): Block-Term Tensor Decomposition Model Selection and Computation: The Bayesian Way (SPS Journal Paper)*

Paris Giampouras (Johns Hopkins University); Athanasios Rontogiannis (National Observatory of Athens); Eleftherios Kofidis (University of Piraeus)

IVMSP-P22: Perceptual Assessment

Room: Dome K

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Jinli Suo, Wenming Yang

3108 (IVMSP-P22.1): Boosting No-Reference Super-Resolution Image Quality Assessment with Knowledge Distillation and Extension

Haiyu Zhang (Northwestern Polytechnical University); Shaolin Su (Northwestern Polytechnical University); Yu Zhu (Northwestern Polytechnical University); Jinqiu Sun (Northwestern Polytechnical University); Yanning Zhang (Northwestern Polytechnical University)

3262 (IVMSP-P22.2): LSTM-based Video Quality Prediction Accounting for Temporal Distortions in Videoconferencing Calls

Gabriel Mittag (Microsoft Corporation); Babak Naderi (Technical University of Berlin); Vishak Gopal (Microsoft Corporation); Ross Cutler (Microsoft Corporation)

2505 (IVMSP-P22.3): LOCAL FEATURE ENHANCED ADVERSARIAL NETWORK FOR THE BLIND IMAGE QUALITY ASSESSMENT

Xiaomei Shi (Northwest University); Min Zhang (Northwest University); Shou Hai Xia (Northwest University); Ru Xue Zhang (Northwest University); Jun Feng (Northwest University)

3583 (IVMSP-P22.4): LiNuQA: Lightweight No-Reference Image Quality Assessment based on Non-Uniform Weighting

Wook-Hyung Kim (Samsung Electronics Co. Ltd); Cheul-Hee Hahm (Samsung Electronics Co. Ltd); Anant Bajjal (Samsung Electronics Co. Ltd); Namuk Kim (Samsung Electronics Co. Ltd); Ilhyun Cho (Samsung Electronics Co. Ltd); Jayoon Koo (Samsung Electronics Co. Ltd)

5771 (IVMSP-P22.5): OPTIMIZED QUALITY FEATURE LEARNING FOR VIDEO QUALITY ASSESSMENT

Ngai-Wing Kwong (The Hong Kong Polytechnic University); Yui-Lam Chan (The Hong Kong Polytechnic University); Sik-Ho Tsang (Centre for Advances in Reliability and Safety (CAiRS)); Daniel P.K. Lun (The Hong Kong Polytechnic University)

6013 (IVMSP-P22.6): ROBUST CONTENT-VARIANT REFERENCE IMAGE QUALITY ASSESSMENT VIA SIMILAR PATCH MATCHING

Wenbo Shi (Tsinghua University); Wenming Yang (Tsinghua University); Qingmin Liao (Tsinghua University)

5569 (IVMSP-P22.7): No reference quality assessment for screen content images based on entire and high-influence regions

Zhuoran Xu (Anhui University); Yang Yang (Anhui University); Zhixiang Zhang (Hefei High-Dimensional Data Technology Co.,Ltd); Weiming Zhang (University of Science and Technology of China)

6056 (IVMSP-P22.8): PCQA-GRAPHPOINT: EFFICIENT DEEP-BASED GRAPH METRIC FOR POINT CLOUD QUALITY ASSESSMENT

Marouane Tliba (University of Orleans); Aladine Chetouani (Université d'Orléans, France); Giuseppe Valenzise (CNRS); Frederic Dufaux (CNRS)

Wednesday, June 7

2898 (IVMSP-P22.9): Ψ -Net: Point Structural Information Network for No-reference Point Cloud Quality Assessment
Jian Xiong (Nanjing University of Posts and Telecommunications); Sifan Wu (Nanjing University of Posts and Telecommunications); Wang Luo (Nanjing University of Posts and Telecommunications); Jinli Suo (Tsinghua University); Hao Gao (Nanjing University of Posts and Telecommunications)

4359 (IVMSP-P22.10): Test your samples jointly: Pseudo-reference for image quality evaluation
Marcelin Tworski (Telecom Paris); Stéphane Lathuilière (Telecom-Paris)

MLSP-P25: Machine Learning for Recommendation, Search and other Applications

Room: Dome F

Type: Poster

03:40 PM to 05:10 PM

Chair(s): TBA

6262 (MLSP-P25.1): Rethinking Rule-based Approaches in Session-based Recommendation
Liuyin Wang (Tsinghua University); Mingchao Li (Tsinghua University); Hai-Tao Zheng (Tsinghua University)

6242 (MLSP-P25.2): High-level Feature Fusion Network for Session-based Social Recommendation
Liuyin Wang (Tsinghua University); Mingchao Li (Tsinghua University); Hai-Tao Zheng (Tsinghua University)

2855 (MLSP-P25.3): Tree-like Interaction Learning for Bundle Recommendation
Haole Ke (Wuhan University of Technology); Lin Li (Wuhan University of Technology); Peipei Wang (Wuhan University of Technology); Jingling Yuan (Wuhan University of Technology); Xiaohui Tao (The University of Southern Queensland)

1474 (MLSP-P25.4): Int-GNN: a User Intention Aware Graph Neural Network for Session-Based Recommendation
Guangning Xu (Harbin Institute of Technology, Shenzhen Δ); Jinyang Yang (Harbin Institute of Technology, Shenzhen); Jinjin Guo (JD Intelligent Cities Research); Zhichao Huang (JD Intelligent Cities Research); Bowen Zhang (Shenzhen Technology University)

2443 (MLSP-P25.5): FFFN: Fashion Feature Fusion Network by Co-attention Model for Fashion Recommendation
Zhantu Lin (College of Computer Science and Software Engineering, Shenzhen University); Xiaoyan Zhang (College of Computer Science and Software Engineering, Shenzhen University)

2451 (MLSP-P25.6): RØROS: Building a Responsive Online Recommender System via Meta-Gradients Updating
Xudong Pan (Fudan University); Mi Zhang (Fudan University); Duocai Wu (Ant Group)

6100 (MLSP-P25.7): Multi-aspect Interest Neighbor-augmented Network for Next-basket Recommendation
Zhiying Deng (Huazhong University of Science and Technology); Jianjun Li (School of Computer Science and Technology, Huazhong University of Science and Technology); Zhiqiang Guo (School of Computer Science and Technology, Huazhong University of Science and Technology); Guohui Li (School of Computer Science and Technology, Huazhong University of Science and Technology)

4402 (MLSP-P25.8): DUAL-GRAPH CO-REPRESENTATION LEARNING FOR KNOWLEDGE-GRAPH ENHANCED RECOMMENDATION
Xinbiao Liu (Fudan University); Bin Liang (Fudan University); JunYu Niu ("Fudan University, China"); Chaofeng Sha (Fudan University); Dong Wu (Fudan University)

2886 (MLSP-P25.9): Towards Real-Time Person Search with Invariant Feature Learning
Chengyou Jia (Xi'an Jiaotong University); Minnan Luo (School of Electronic and Information Engineering, Xi'an Jiaotong University); Zhuohang Dang (Xi'an Jiaotong University); Xiaojun Chang (University of Technology Sydney); Qinghua Zheng (Xi'an Jiaotong University)

4668 (MLSP-P25.10): Bayesian Network Modeling and Prediction of Transitions within the Homelessness System
Khandker Sadia Rahman (University at Albany); Daphney-Stavroula Zois (University at Albany); Charalampos Chelms (University at Albany)

4350 (MLSP-P25.11): A Deep Temporal Factor Analysis Method for Large Scale Financial Portfolio Selection
Yao Zhou (Shanghai JiaoTong University); Ruidan Su (Shanghai Jiao Tong University); Shikui Tu (Shanghai Jiao Tong University); Lei Xu (Shanghai Jiao Tong University)

MLSP-P26: Reinforcement Learning I

Room: Dome G

Type: Poster

03:40 PM to 05:10 PM

Chair(s): TBA

3947 (MLSP-P26.1): CLASS-GUIDED TRIPLE HEAD PREDICTION NETWORK FOR LONG-TAIL OBJECT DETECTION
xuyang liu (Inner Mongolia University); Yuan Zheng (Inner Mongolia University)

Wednesday, June 7

4074 (MLSP-P26.2): SMCL: SALIENCY MASKED CONTRASTIVE LEARNING FOR LONG-TAILED VISUAL RECOGNITION
Sanglee Park (Sogang University, LG Electronics); Seung-won Hwang (Seoul National University); Jungmin So (Sogang University)

816 (MLSP-P26.3): COMPLEMENTARY LEARNING SYSTEM BASED INTRINSIC REWARD IN REINFORCEMENT LEARNING
Zijian Gao (National University of Defense Technology); Kele Xu (National Key Laboratory of Parallel and Distributed Processing (PDL)); Hongda Jia (National University of Defense Technology); Tianjiao Wan (National University of Defense Technology); Ding Bo (National University of Defense Technology); Dawei Feng (National University of Defense Technology); Xinjun Mao (National University of Defense Technology); Huaimin Wang (National University of Defense Technology)

1130 (MLSP-P26.4): Promoting Cooperation in Multi-Agent Reinforcement Learning via Mutual Help
Yunbo Qiu (Tsinghua University); Yue Jin (University of Warwick); Lebin Yu (Tsinghua University); Jian Wang (Tsinghua university); Xudong Zhang (Tsinghua university)

1691 (MLSP-P26.5): A UNIFIED UNCERTAINTY-AWARE EXPLORATION: COMBINING EPISTEMIC AND ALEATORY UNCERTAINTY
Parvin Malekzadeh (university of Toronto); Ming Hou (Department of National Defence's Innovation for Defence Excellence and Security (IDEaS) program, Canada); Konstantinos N Plataniotis (UofT)

2234 (MLSP-P26.6): MEET: A Monte Carlo Exploration-Exploitation Trade-off for Buffer Sampling
Julius Ott (Infineon Technologies AG / Technical University Munich); Lorenzo Servadei (Infineon Technologies AG); Jose Arjona-Medina (Johannes Kepler University Linz); Enrico Rinaldi (University of Michigan); Gianfranco Mauro (Infineon Technologies AG); Daniela Sanchez Lopera (Infineon Technologies AG / Technical University Munich); Michael Stephan (Infineon Technologies AG); Thomas Stadelmayer (Infineon Technologies AG); Avik Santra (Infineon Technologies AG); Robert Wille (Technical University of Munich)

3667 (MLSP-P26.7): CONVERGENCE ANALYSIS OF GRAPHICAL GAME-BASED NASH Q -LEARNING USING THE INTERACTION DETECTION SIGNAL OF N -STEP RETURN
Yunkai Zhuang (Nanjing University); Shangdong Yang (Nanjing University of Posts and Telecommunications); Wenbin Li (Nanjing University); Yang Gao (Nanjing University)

3942 (MLSP-P26.8): DEEP REINFORCEMENT LEARNING FOR GREEN UAV-ASSISTED DATA COLLECTION
Abhishek Mondal (National Institute of Technology Silchar); Deepak Mishra (University of New South Wales, Sydney); Ganesh Prasad (National Institute of Technology Silchar); Ashraf Hossain (National Institute of Technology Silchar)

4956 (MLSP-P26.9): MATRIX LOW-RANK APPROXIMATION FOR POLICY GRADIENT METHODS
Sergio Rozada (King Juan Carlos University); Antonio G. Marques (King Juan Carlos University)

5143 (MLSP-P26.10): NETWORKED POLICY GRADIENT PLAY IN MARKOV POTENTIAL GAMES
Sarper Aydin (Texas A&M University); Ceyhun A Eksin (Texas A&M University)

3414 (MLSP-P26.11): PROGRESSIVE DIVERSIFYING POLICY FOR MULTI-AGENT REINFORCEMENT LEARNING
Shaoqi Sun (National University of Defense Technology); Yuanzhao Zhai (National University of Defense Technology); Kele Xu (National Key Laboratory of Parallel and Distributed Processing (PDL)); Dawei Feng (National University of Defense Technology); Ding Bo (National University of Defense Technology)

MLSP-P27: Deep Learning for Speech and Audio Processing II

Room: Dome H

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Yang Liu,

626 (MLSP-P27.1): SEQUENCE-BASED DEVICE-FREE GESTURE RECOGNITION FRAMEWORK FOR MULTI-CHANNEL ACOUSTIC SIGNALS

Zhizheng Yang (Nanjing University); Xun Wang (Nanjing University); Dongyu Xia (Nanjing University); Wei Wang (Nanjing University); Haipeng Dai (Nanjing University)

4627 (MLSP-P27.2): CAT: Causal Audio Transformer for Audio Classification

Xiaoyu Liu (University of Maryland, College Park); Hanlin Lu (ByteDance Inc.); Jianbo Yuan (Bytedance); Xinyu Li (Amazon)

5329 (MLSP-P27.3): CNEG-VC: Contrastive Learning using Hard Negative Example in Non-parallel Voice Conversion

Bima Prihasto (National Central University); YiXing Lin (National Central University); Le Phuong (National Central University); CHIEN-LIN HUANG (NCKU); Jia-Ching Wang (National Central University)

3259 (MLSP-P27.4): TOWARDS ROBUST DATA-DRIVEN UNDERWATER ACOUSTIC LOCALIZATION: A DEEP CNN SOLUTION WITH PERFORMANCE GUARANTEES FOR MODEL MISMATCH

Amir Weiss (Massachusetts Institute of Technology); Andrew C Singer (University of Illinois); Gregory W Wornell (MIT)

Wednesday, June 7

1269 (MLSP-P27.5): AD-YOLO: YOU LOOK ONLY ONCE IN TRAINING MULTIPLE SOUND EVENT LOCALIZATION AND DETECTION

Jin Sob Kim (Korea University); Hyun Joon Park (Korea University); Wooseok Shin (Korea University); Sung Won Han (Korea University)

1429 (MLSP-P27.6): Framewise multiple sound source localization and counting using binaural spatial audio signals

Lei Wang (Shanghai Jiao Tong University); Zhibin Jiao (Huawei Technologies Co., Ltd.); Qiyong Zhao (Huawei Technologies Co., Ltd.); jie zhu (Shanghai Jiao Tong University); Yang Fu (Huawei Technologies Co., Ltd.)

5601 (MLSP-P27.7): Learning Speech Representations with Flexible Hidden Feature Dimensions

Huaizhen Tang (University of Science and Technology of China); Xulong Zhang (Ping An Technology (Shenzhen) Co., Ltd.); Jianzong Wang (Ping An Technology (Shenzhen) Co., Ltd); Ning Cheng (Ping An Technology (Shenzhen) Co., Ltd); Jing Xiao (Ping An Insurance (Group) Company of China)

3364 (MLSP-P27.8): Guided Speech Enhancement Network

Yang Yang (Google LLC); Shao-Fu Shih (Google LLC); Hakan Erdogan (Google); Jamie Menjay Lin (Google); Chehung Lee (Google LLC); Yunpeng Li (Google); George Sung (Google LLC); Matthias Grundmann (Google Research)

3818 (MLSP-P27.9): Blind Estimation of Audio Processing Graph

Sungho Lee (Seoul National University); Jaehyun Park (Seoul National University); Seungryeol Paik (Seoul National University); Kyogu Lee (Seoul National University)

4309 (MLSP-P27.10): MarginNCE: Robust Sound Localization with a Negative Margin

Sooyoung Park (ETRI); Arda Senocak (KAIST); Joon Son Chung (KAIST)

2785 (MLSP-P27.11): Improving Noisy Student Training on Non-target Domain Data for Automatic Speech Recognition

YU CHEN (University of Hong Kong); Wen Ding (NVIDIA); Junjie Lai (NVIDIA)

5791 (MLSP-P27.12): Large-Scale Nonverbal Vocalization Detection Using Transformers

Panagiotis Tzirakis (Hume AI); Alice Baird (Hume AI); Jeff Brooks (Hume AI); Chris Gagne (Hume.ai); Lauren Kim (Hume AI); Michael Opara (Hume AI); Christopher Gregory (Hume AI); Jacob Metrick (Hume AI); Garrett Boseck (Hume AI); Vineet Tiruvadi (Hume AI); Bjoern W. Schuller (Imperial College London); Dacher Keltner (UC Berkeley); Alan S Cowen (Hume AI)

MLSP-P28: Pattern Recognition and Classification I

Room: Dome L

Type: Poster

03:40 PM to 05:10 PM

Chair(s): TBA

1137 (MLSP-P28.1): Towards Trustworthy Multi-label Sewer Defect Classification via Evidential Deep Learning

Chenyang Zhao (School of Software, Southeast University); Chuanfei Hu (University of Shanghai for Science and Technology); Hang Shao (University of Shanghai for Science and Technology); Zhe Wang (University of Shanghai for Science and Technology); Yongxiang Wang (University of Shanghai for Science and Technology)

6138 (MLSP-P28.2): Inv-SENet: Invariant Self Expression Network for clustering under biased data

Ashutosh Singh (Northeastern University); Ashish Singh (University of Massachusetts Amherst); Aria Masoomi (Northeastern University); Tales Imbiriba (Northeastern University); Erik Learned-Miller (University of Massachusetts, Amherst); Deniz Erdogmus (Northeastern University)

4442 (MLSP-P28.3): Multi-Label Temporal Evidential Neural Networks for Early Event Detection

Xujiang Zhao (NEC Lab America); Xuchao Zhang (Microsoft); Chen Zhao (Kitware Inc.); Jin-Hee Cho (Virginia Tech); Lance Kaplan (DEVCOM Army Research Laboratory); DONG HYUN JEONG (University of the District of Columbia); Audun Jøsang (University of Oslo); Haifeng Chen (NEC Labs); Feng Chen (UT Dallas)

3520 (MLSP-P28.4): DECOMFORMER: DECOMPOSE SELF-ATTENTION VIA FOURIER TRANSFORM FOR VHR AERIAL IMAGE SCENE CLASSIFICATION

Yan Zhang (Chongqing University of Posts and Telecommunications); Xiyuan Gao (Chongqing University of Posts and Telecommunications); Xiao PU (Chongqing University of Posts and Telecommunications); Tao Wang (Chongqing University of Posts and Telecommunications); Xinbo Gao (Chongqing University of Posts and Telecommunications)

2019 (MLSP-P28.5): Accelerating matrix trace estimation by Aitken's Δ^2 process

Vasileios Kalantzis (IBM Research); Georgios Kollias (IBM Research); Shashanka Ubaru (IBM Research); Theodoros Salonidis (IBM T.J. Watson Research Center)

1152 (MLSP-P28.6): Input-dependent Dynamical Channel Association for Knowledge Distillation

Qiankun Tang (Zhejiang Lab); Yuan Zhang (China Telecom); Xiaogang Xu (Zhejiang Gongshang University); Jun Wang (Zhejiang Lab); Yimin Guo (China Telecom Research Institute)

Wednesday, June 7

1266 (MLSP-P28.7): T5-SR: A Unified Seq-to-Seq Decoding Strategy for Semantic Parsing

Yuntao Li (Peking University); Zhenpeng Su (University of Chinese Academy of Sciences); yutian li (Meituan Group); Zhang Hanchu (meituan); Sirui Wang (Meituan); Wei Wu (Meituan-Dianping Group); Yan Zhang (Peking University)

794 (MLSP-P28.8): SEMI-SUPERVISED LOCAL STRUCTURED FEATURE LEARNING WITH DYNAMIC MAXIMUM ENTROPY GRAPH

Rui Xu (Renmin University of China); Xun Liang (Renmin University of China)

3645 (MLSP-P28.9): Hint-dynamic Knowledge Distillation

Yiyang Liu (Xiamen University); Chenxin Li (Xiamen University); Xiaotong Tu (Xiamen University); Xinghao Ding (Xiamen University); Yue Huang (Xiamen University)

3817 (MLSP-P28.10): Learning from Label Proportion with Online Pseudo-Label Decision by Regret Minimization

Shinnosuke Matsuo (Kyushu University); Seiichi Uchida (Kyushu University); Ryoma Bise (Kyushu University); Daiki Suehiro (Kyushu University)

4099 (MLSP-P28.11): INTER-SCALE SURE-LET DENOISE WITH STRUCTURED DEEP IMAGE PRIOR: INTERPRETABLE SELF-SUPERVISED LEARNING

JIKAI LI (Niigata University); Shogo Muramatsu (Niigata Univ.)

4469 (MLSP-P28.12): ASYMMETRIC POLYNOMIAL LOSS FOR MULTI-LABEL CLASSIFICATION

Yusheng Huang (Shanghai Jiao Tong University); Jiexing Qi (Shanghai Jiao Tong University); Xinbing Wang (Shanghai Jiao Tong University); Zhouhan Lin (Shanghai Jiao Tong University)

SAM-P3: Sparsity, Compressed Sensing, and Tensor Decomposition

Room: Dome M

Type: Poster

03:40 PM to 05:10 PM

Chair(s): TBA

1089 (SAM-P3.1): Super-resolution harmonic retrieval of non-circular signals

Yu Zhang (Nanjing University of Aeronautics and Astronautics); Yue Wang (George Mason University); Zhi Tian (George Mason University); Geert Leus (TU Delft); Gong Zhang (Nanjing University of Aeronautics and Astronautics)

1816 (SAM-P3.2): Radio-astronomy imaging and interference excision using tensor decomposition and canonical correlation analysis

Mikael Sorensen (University of Virginia); Nicholas D Sidiropoulos (University of Virginia)

3226 (SAM-P3.3): TO REGULARIZE OR NOT TO REGULARIZE: THE ROLE OF POSITIVITY IN SPARSE ARRAY INTERPOLATION WITH A SINGLE SNAPSHOT

Mehmet Hucumenoglu (University of California San Diego); Pulak Sarangi (UCSD); Robin Rajamaki (UCSD); Piya Pal (Nil)

3472 (SAM-P3.4): LIGHT-WEIGHT SEQUENTIAL SBL ALGORITHM: AN ALTERNATIVE TO OMP

Rohan Ramchandra Pote (University of California San Diego); Bhaskar Rao (UC San Diego)

3948 (SAM-P3.5): Sparse Bayesian Learning Based Three-Dimensional Imaging for Antenna Array Radar

Yuhan Li (Xiamen University); Jesper Rindom Jensen (Aalborg University); Maozhong Fu (Xiamen University of Technology); Zhenmiao Deng (Sun Yat-sen University); Mads G. Christensen (Audio Analysis Lab., AD:MT, Aalborg University, Denmark)

4132 (SAM-P3.6): Sparse Non-Contact Multiple People Localization and Vital Signs Monitoring Via FMCW Radar

Yonathan Eder (Weizmann Institute of Science); Zhuoyang Liu (Weizmann Institute of Science); Yonina Eldar ()

4392 (SAM-P3.7): Deep learning-based compressive sampling optimization in massive MIMO systems

Saidur Pavel (Temple University); Yimin D Zhang (Temple University); Maria S. Greco (University of Pisa); Fulvio Gini (University of Pisa)

4491 (SAM-P3.8): Sample-Efficient Robust MMV Recovery Algorithm

Yuvraj Singh (IIT Bombay); Jahnavi S Rohela (Indian Institute of Technology Bombay); Satish Mulleti (Indian Institute of Technology Bombay, India)

4658 (SAM-P3.9): Deep Learning Sparse Array Design Using Binary Switching Configurations

Syed Ali Hamza (Widener University); Kyle Juretus (Villanova University); Moeness Amin (Villanova University); Fauzia Ahmad (Temple University)

4824 (SAM-P3.10): Low-rank tensor decompositions for quaternion multiway arrays

Osimone Imhogiemhe (Université de Lorraine, CNRS, CRAN); Julien Flamant (CNRS); Xavier Luciani (Université de Toulon, Aix Marseille Université, CNRS, Seatech, LIS); Yassine ZNIYED (LIS/SeaTech); Sebastian Miron (University of Lorraine)

Wednesday, June 7

4838 (SAM-P3.11): SPARSITY CONSTRAINT IMPLEMENTATION FOR THE JOINT EIGENVALUE DECOMPOSITION OF MATRICES

Rémi ANDRÉ (Institut Fresnel); Xavier Luciani (LIS)

5896 (SAM-P3.12): SPARSE ERROR CORRECTION FOR POWER NETWORK PARAMETERS

Dilan S Senaratne (Oregon State University); Jinsub Kim (Oregon State)

SLT-P24: Natural Language Processing IV

Room: Dome C

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Pegah Kharazmi,

5463 (SLT-P24.1): KG-ECO: Knowledge Graph Enhanced Entity Correction for Query Rewriting

Jinglun Cai (Amazon.com, Inc); Mingda Li (Amazon); Ziyang Jiang (Amazon); Eunah Cho (Amazon); Zheng Chen (Amazon Alexa AI); Yang Liu (Amazon, Alexa AI); Xing Fan (Amazon); Chenlei Guo (Amazon)

400 (SLT-P24.2): From Easy to Hard: Two-stage Selector and Reader for Multi-hop Question Answering

Xin-Yi Li (State Key Laboratory for Novel Software Technology, Nanjing University); Wei-Jun Lei (State Key Laboratory for Novel Software Technology, Nanjing University); Yu-Bin Yang (State Key Laboratory for Novel Software Technology, Nanjing University)

443 (SLT-P24.3): Tell Model Where to Attend: Improving Interpretability of Aspect-Based Sentiment Classification via Small Explanation Annotations

Zhenxiao Cheng (East China Normal University); Jie Zhou (Fudan University); Wen Wu (East China Normal University); Qin Chen (East China Normal University); Liang He (ECNU)

640 (SLT-P24.4): LED: Label Correlation Enhanced Decoder for Multi-Label Text Classification

Kefan Ma (Shanghai Jiao Tong University); Zheng Huang (Shanghai Jiao Tong University); Xinrui Deng (Shanghai Jiao Tong University); Jie Guo (Shanghai Jiao Tong University); Weidong Qiu (Shanghai Jiaotong University)

823 (SLT-P24.5): SADE: A Self-adaptive Expert for Multi-dataset Question Answering

Yixing Peng (State Key Laboratory of Communication Content Cognition, University of Science and Technology of China); Quan Wang (Beijing University of Posts and Telecommunications); Zhendong Mao (University of Science and Technology of China); Yongdong Zhang (University of Science and Technology of China)

1792 (SLT-P24.6): Time-Aware Multiway Adaptive Fusion Network for Temporal Knowledge Graph Question Answering

Yonghao Liu (Centre for Natural Language Processing, Meituan Inc., Beijing, China); Di Liang (Centre for Natural Language Processing, Meituan Inc., Beijing, China); Fang Fang (Department of Automation, Tsinghua University, Beijing, China); Sirui Wang (Centre for Natural Language Processing, Meituan Inc., Beijing, China); Wei Wu (Centre for Natural Language Processing, Meituan Inc., Beijing, China); Rui Jiang (Department of Automation, Tsinghua University, Beijing, China)

3083 (SLT-P24.7): Narrow Down Before Selection: A Dynamic Exclusion Model For Multiple-Choice QA

Xiyun Liu (Beijing University of Posts and Telecommunications); Yidong Shi (Beijing University of Posts and Telecommunications); Ruifang Liu (Beijing University of Posts and Telecommunications); Ge Bai (Beijing University of Posts and Telecommunications); Yanyi Chen (Beijing University of Posts and Telecommunications)

3507 (SLT-P24.8): SELF-ADAPTIVE REASONING ON SUB-QUESTIONS FOR MULTI-HOP QUESTION ANSWERING

ZeKai Li (National University of Singapore); Wei Peng (Institute of Information Engineering, Chinese Academy of Sciences)

3785 (SLT-P24.9): Disentangled and Robust Representation Learning for Bragging Classification in Social Media

Xiang Li (Tianjin university); Yucheng Zhou (University of Technology Sydney)

4608 (SLT-P24.10): SynGen: A Syntactic Plug-and-play Module for Generative Aspect-based Sentiment Analysis

Chengze Yu (Tsinghua University); Taiqiang Wu (Tsinghua University); Jiayi Li (Tsinghua University); Xingyu Bai (Tsinghua University); Yujiu Yang (Tsinghua University)

6318 (SLT-P24.11): A Sentiment and Syntactic-Aware Graph Convolutional Network for Aspect-level Sentiment Classification

Yuxin Yang (Northwest University); Xia Sun (Northwest University); Qiang Lu (Northwest university); Richard F E Sutcliffe (Northwest University); Jun Feng (Northwest University)

6376 (SLT-P24.12): SELF SUPERVISED BERT FOR LEGAL TEXT CLASSIFICATION

Arghya Pal (Monash University); Sailaja Rajanala (Monash University Malaysia); Raphael CW Phan (Monash University); KokSheik Wong (Monash University Malaysia)

Wednesday, June 7

SLT-P25: Resource Constrained ASR

Room: Dome D

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Samuel Thomas, Jen-Tzung Chien

2456 (SLT-P25.1): Adversarial Data Augmentation Using VAE-GAN for Disordered Speech Recognition

Zengrui Jin (The Chinese University of Hong Kong); Xurong Xie (Institute of Software, Chinese Academy of Sciences); Mengzhe GENG (The Chinese University of Hong Kong); Tianzi Wang (The Chinese University of Hong Kong); Shujie HU (The Chinese University of Hong Kong); Jiajun Deng (The Chinese University of Hong Kong); Guinan Li (Chinese University of Hong Kong); Xunying Liu (The Chinese University of Hong Kong)

2842 (SLT-P25.2): BEBERT: Efficient and Robust Binary Ensemble BERT

Jiayi Tian (Nanjing University); Chao Fang (Nanjing University); Haonan Wang (University of Southern California); Zhongfeng Wang (Nanjing University)

3153 (SLT-P25.3): Sharing Low Rank Conformer Weights for Tiny Always-On Ambient Speech Recognition Models

Steven M. Hernandez (Virginia Commonwealth University); Ding Zhao (Google); Shaojin Ding (Google); Antoine Bruguier (Google); Rohit Prabhavalkar (Google); Tara Sainath (Google); Yanzhang He (Google); Ian McGraw ()

3488 (SLT-P25.4): INTERMEDIATE FINE-TUNING USING IMPERFECT SYNTHETIC SPEECH FOR IMPROVING ELECTROLARYNGEAL SPEECH RECOGNITION

Lester Phillip G Violeta (Nagoya University); Ding Ma (Nagoya University); Wen-Chin Huang (Nagoya University); Tomoki Toda (Nagoya University)

3930 (SLT-P25.5): META LEARNING WITH ADAPTIVE LOSS WEIGHT FOR LOW-RESOURCE SPEECH RECOGNITION

Qiulin Wang (Xiamen University); Wenxuan Hu (Xiamen University); Lin Li (Xiamen University); Qingyang Hong (Xiamen University)

4546 (SLT-P25.6): USING MODIFIED ADULT SPEECH AS DATA AUGMENTATION FOR CHILD SPEECH RECOGNITION

Zijian Fan (Norwegian University of Science and Technology); Xinwei Cao (NTNU); Giampiero Salvi (NTNU); Torbjørn Svendsen (NTNU)

4747 (SLT-P25.7): Make More of Your Data: Minimal Effort Data Augmentation for Automatic Speech Recognition and Translation

Tsz Kin Lam (Heidelberg University); Shigehiko Schamoni (Heidelberg University); Stefan Riezler (Heidelberg University)

4842 (SLT-P25.8): Multilingual end-to-end spoken language understanding for ultra-low footprint applications

Markus Mueller (Amazon Alexa); Anastasios Alexandridis (Amazon.com); Zach Trozenski (Amazon Alexa); Joel Whiteman (Amazon Alexa); Grant Strimel (Amazon Alexa); Nathan Susanj (Amazon Alexa); Athanasios Mouchtaris (Amazon Alexa); Siegfried Kunzmann (Amazon Alexa)

5084 (SLT-P25.9): AN ISOTROPY ANALYSIS FOR SELF-SUPERVISED ACOUSTIC UNIT EMBEDDINGS ON THE ZERO RESOURCE SPEECH CHALLENGE 2021 FRAMEWORK

Jianan Chen (Japan Advanced Institute of Science and Technology); Sakriani Sakti (Japan Advanced Institute of Science and Technology)

5687 (SLT-P25.10): SELF-SUPERVISED ACCENT LEARNING FOR UNDER-RESOURCED ACCENTS USING NATIVE LANGUAGE DATA

Mehul Kumar (Samsung Research); Jiyeon Kim (Samsung Research); Dhananjaya Gowda (Samsung Electronics); Abhinav Garg (Stanford); Chanwoo Kim (Samsung Electronics)

5935 (SLT-P25.11): Enhancing Unsupervised Speech Recognition with Diffusion GANs

Xianchao Wu (NVIDIA Japan)

5976 (SLT-P25.12): Effectiveness of Mining Audio and Text Pairs from Public Data for Improving ASR Systems for Low-Resource Languages

Kaushal Bhogale (Indian Institute of Technology, Madras); Abhigyan Raman (AI4Bharat); Tahir Javed (Indian Institute of Technology Madras); Sumanth Doddapaneni (Robert Bosch Centre for Data Science and AI); Anoop Kunchukuttan (Microsoft); Pratyush Kumar (Indian Institute of Technology Madras); Mitesh M. Khapra (Indian Institute of Technology Madras)

SLT-P26: Singing Voice Synthesis/Conversion and Pretrained TTS

Room: Dome E

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Erica Cooper,

404 (SLT-P26.1): UNSUPERVISED PRE-TRAINING FOR DATA-EFFICIENT TEXT-TO-SPEECH ON LOW RESOURCE LANGUAGES

Seongyeon Park (Seoul National University); Myungseo Song (CNAI); bohuyng kim (CNAI); Tae-Hyun Oh (POSTECH)

Wednesday, June 7

510 (SLT-P26.2): WeSinger 2: Fully Parallel Singing Voice Synthesis via Multi-Singer Conditional Adversarial Training
Zewang Zhang (Tencent Inc.); Yibin Zheng (Tencent Inc, China); Xinhui Li (Tencent Inc); Li Lu (Tencent Inc)

1240 (SLT-P26.3): Hierarchical Diffusion Models for Singing Voice Neural Vocoder
Naoya Takahashi (Sony Group); Mayank Kumar Singh (Sony Research India); Yuki Mitsufuji (Sony Group Corporation)

2314 (SLT-P26.4): A FEW SHOT LEARNING OF SINGING TECHNIQUE CONVERSION BASED ON CYCLE CONSISTENCY GENERATIVE ADVERSARIAL NETWORKS
Po-Wei Chen (National Tsing Hua University); Von-Wun Soo (nthu)

2624 (SLT-P26.5): Self-supervised Representations for Singing Voice Conversion
Tejas Jayashankar (MIT); Jilong Wu (Meta Platforms Inc.); Leda Sari (Meta Platforms Inc.); David Kant (Meta Platforms Inc.); Vimal Manohar (Meta Platforms Inc.); Qing He (Meta)

3191 (SLT-P26.6): NNSVS: A Neural Network-Based Singing Voice Synthesis Toolkit
Ryuichi Yamamoto (LINE Corp.); Reo Yoneyama (Nagoya University); Tomoki Toda (Nagoya University)

4738 (SLT-P26.7): Analysis and transformation of voice level in singing voice
Frederik Bous (STMS - IRCAM, Sorbonne Université, CNRS); Axel Roebel (Ircam)

6177 (SLT-P26.8): Singing Voice Synthesis Based on a Musical Note Position-aware Attention Mechanism
Yukiya Hono (Nagoya Institute of Technology); Kei Hashimoto (Nagoya Institute of Technology); Yoshihiko Nankaku (Nagoya Institute of Technology); Keiichi Tokuda (Department of Computer Science and Engineering, Nagoya Institute of Technology)

6200 (SLT-P26.9): Enhancing the Vocal Range of Single-Speaker Singing Voice Synthesis with Melody-Unsupervised Pre-training
Shaohuan Zhou (Tsinghua University); Xu Li (ARC Lab, Tencent); Zhiyong Wu (Tsinghua University); Ying Shan (Tencent); Helen Meng (The Chinese University of Hong Kong)

6458 (SLT-P26.10): PHONEix: Acoustic Feature Processing Strategy for Enhanced Singing Pronunciation with Phoneme Distribution Predictor
Yuning Wu (Renmin University of China); Jiatong Shi (Carnegie Mellon University); Tao Qian (RUC); Dongji Gao (Johns Hopkins University); Qin Jin (Renmin University of China)

6778 (SLT-P26.11): Exploring the Role of Language Families for Building Indic Speech Synthesisers (SPS Journal Paper)*
Anusha Prakash (Indian Institute of Technology Madras); Hema A Murthy (IIT Madras)

Thursday, June 8

BISP-L3: Medical Image Reconstruction

Room: Nafsika A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Ulugbek Kamilov,

08:15 AM

6097 (BISP-L3.1): Hankel Structured Low Rank and Sparse Representation via L0-Norm Optimization for Compressed Ultrasound Plane Wave Signal Reconstruction

Miaomiao Zhang (Capital Normal University); Ji Chen (Capital Normal University); Xiaoyan Fu (Capital Normal University); Xin Ge (Beijing Jiaotong University); Jingzhi Zhang (Capital Normal University); Na Jiang (Information Engineering College, Capital Normal University); Jan D'Hooge (KU Leuven)

08:30 AM

280 (BISP-L3.2): OCT image blind despeckling based on gradient guided filter with speckle statistical prior

sanqian Li (Southern University of Science and Technology); Muxing Xiong (Southern University of Science and Technology); Bing Yang (Southern University of Science and Technology); Xiaoqing Zhang (Southern University of Science and Technology); Risa Higashita (tomey corporation); Jiang Liu (Southern University of Science and Technology)

08:45 AM

3547 (BISP-L3.3): Simultaneous Reconstruction and Uncertainty Quantification for Tomography

Agnimitra Dasgupta (University of Southern California); Carlo Graziani (Argonne National Laboratory); Zichao Di (Argonne National Laboratory)

09:00 AM

1852 (BISP-L3.4): Perspective Projection-Based 3D CT Reconstruction from Biplanar X-rays

Daeun Kyung (KAIST); Kyungmin Jo (Korea Advanced Institute of Science and Technology); Jaegul Choo (Korea Advanced Institute of Science and Technology); Joonseok Lee (Google Research & Seoul National University); Edward Choi (KAIST)

09:15 AM

5626 (BISP-L3.5): DGN: DESCRIPTOR GENERATION NETWORK FOR FEATURE MATCHING IN MONOCULAR ENDOSCOPY 3D RECONSTRUCTION

KaiYun Zhang (Xiamen University); Wenkang Fan (Xiamen University); Yinran Chen (Xiamen University); Xiongbiao Luo (Xiamen University)

09:30 AM

3469 (BISP-L3.6): UNeXt: a Low-Dose CT denoising UNet model with the modified ConvNeXt block

Farzan Niknejad Mazandarani (Toronto Metropolitan university); Paul Babyn (Physician Executive, Saskatchewan Health Authority, Saskatoon, S7K 0M7, Canada,); Javad Alirezaie (Toronto Metropolitan University, Dept of Electrical Eng.)

GC-8: ICASSP2023 General Meeting Understanding and Generation (MUG) Challenge

Room: Nefeli B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): TBA

08:15 AM

6655 (GC-L8.1): Introduction

Qinglin Zhang (Alibaba Group); Chong Deng (Alibaba inc); Jiaqing Liu (Speech Lab, Alibaba Group); Hai Yu (Alibaba Group); Qian Chen (Speech Lab, DAMO Academy, Alibaba Group); Wen Wang (Alibaba Group); Zhijie Yan (Alibaba Inc.); Jinglin Liu (Zhejiang University); Yi Ren (Zhejiang University); Zhou Zhao (Zhejiang University)

08:35 AM

6853 (GC-L8.2): Dialogue Context Modelling for Action Item Detection: Solution for ICASSP 2023 MUG Challenge Track 5

Jie Huang (Harbin Institute of Technology); Xiachong Feng (Harbin Institute of Technology); Ye Yangfan (HIT); Liang Zhao (HIT); Xiaocheng Feng (Harbin Institute of Technology); Bing Qin (Harbin Institute of Technology); Ting Liu (哈尔滨工业大学)

08:47 AM

6865 (GC-L8.3): W2KPE: Keyphrase Extraction with Word-Word Relation

Wen Cheng (Nanjing University); Shichen Dong (Nanjing University); Wei Wang (Nanjing University)

08:59 AM

6902 (GC-L8.4): HITsz TMG at ICASSP 2023 SPGC: Leveraging pre-training and distillation method for title generation with limited resource

Tianxiao Xu (Harbin Institute of Technology, shenzhen); Zihao Zheng (Harbin Institute of Technology, shenzhen); Xinchuo Hu (Harbin Institute of Technology, Shenzhen); Zetian Sun (Harbin Institute of Technology, shenzhen); Yu Zhao (Harbin Institute of Technology, Shenzhen); Baotian Hu (Harbin Institute of Technology, Shenzhen)

Thursday, June 8

09:11 AM

6912 (GC-L8.5): POST-TRAINED LANGUAGE MODEL ADAPTIVE TO EXTRACTIVE SUMMARIZATION OF LONG SPOKEN DOCUMENTS

Hyunjong Ok (Kyung Hee University); Seong-Bae Park (Kyung Hee University)

09:23 AM

6915 (GC-L8.6): THE AJMIDE TOPIC SEGMENTATION SYSTEM FOR THE ICASSP 2023 GENERAL MEETING UNDERSTANDING AND GENERATION CHALLENGE

Beibei Hu (Ajmide Media); Qiang Li (Ajmide Media); Xianjun Xia (Ajmide Media)

IFS-L2: Multimedia Forensics

Room: Salon de Roses B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): TBA

08:15 AM

1149 (IFS-L2.1): Image Sharing Chain Detection via Sequence-to-Sequence Model

Jiaxiang You (Shenzhen University); Yuanman Li (Shenzhen University); Rongqin Liang (Shenzhen University); Yuxuan Tan (Shenzhen University); Jiatao Zhou (University of Macau); Xia Li (Shenzhen University)

08:30 AM

1299 (IFS-L2.2): CONTENT-INSENSITIVE DYNAMIC LIP FEATURE EXTRACION FOR VISUAL SPEAKER AUTHENTICATION AGAINST DEEPFAKE ATTACKS

Zihao Guo (Shanghai Jiao Tong University); shilin wang (SEIEE, Shanghai Jiaotong University)

08:45 AM

1917 (IFS-L2.3): EXPLOITING PRNU AND LINEAR PATTERNS IN FORENSIC CAMERA ATTRIBUTION UNDER COMPLEX LENS DISTORTION CORRECTION

Andrea AM Montibeller (University of Trento); Fernando Perez-Gonzalez (Universidad de Vigo)

09:00 AM

1950 (IFS-L2.4): WHICH COUNTRY IS THIS PICTURE FROM? NEW DATA AND METHODS FOR DNN-BASED COUNTRY RECOGNITION

Omran Alamayreh (University of Siena); Giovanna Dimitri (University of Siena); Jun Wang (University of Siena); Benedetta Tondi (University of Siena); Mauro Barni (University of Siena)

09:15 AM

2345 (IFS-L2.5): UNTAG: Learning Generic Features for Unsupervised Type-Agnostic Deepfake Detection

Nesryne Mejri (Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg); Enjie Ghorbel (SnT, University of Luxembourg); Djamila Aouada (SnT, University of Luxembourg)

09:30 AM

2479 (IFS-L2.6): A 3D-ASSISTED FRAMEWORK TO EVALUATE THE QUALITY OF HEAD MOTION REPLICATION BY REENACTMENT DEEPFAKE GENERATORS

Sahar Hussein (Eurecom); Jean-Luc DUGELAY (Eurecom); Fabien Aili (Docaposte); Emmanuel Nars (Docaposte)

MLSP-L8: Distributed and Federated Learning I

Room: Jupiter

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Qing Liu,

08:15 AM

2526 (MLSP-L8.1): Local Graph-homomorphic Processing for Privatized Distributed Systems

Elsa Rizk (EPFL); Stefan Vlaski (Imperial College London); Ali H. Sayed (Ecole Polytechnique Fédérale de Lausanne)

08:30 AM

5542 (MLSP-L8.2): CODED MATRIX COMPUTATIONS FOR D2D-ENABLED LINEARIZED FEDERATED LEARNING

Anindya Bijoy Das (Purdue University); Aditya RAMAMOORTHY (Iowa State University); David Love (Purdue University); Christopher Brinton (Purdue University)

08:45 AM

2943 (MLSP-L8.3): PERSONALIZED FEDERATED LEARNING ON LONG-TAILED DATA VIA ADVERSARIAL FEATURE AUGMENTATION

Yang Lu (Xiamen University); Pinxin Qian (Xiamen University); Gang Huang (Zhejiang Lab); Hanzi Wang (Xiamen University)

Thursday, June 8

09:00 AM

4809 (MLSP-L8.4): Speech Privacy Leakage from Shared Gradients in Distributed Learning

Zhuohang Li (University of Tennessee, Knoxville); Jiaxin Zhang (Intuit AI Research); Jian Liu (The University of Tennessee, Knoxville)

09:15 AM

2672 (MLSP-L8.5): Fed-3DA: A Dynamic and Personalized Federated Learning Framework

Hui Wang (SKLSDE, School of Computer Science and Engineering, Beihang University, Beijing, China); Jie Sun (Beihang University); Tianyu Wo (Beihang University); Xudong Liu (Beihang University)

09:30 AM

4123 (MLSP-L8.6): Subgradient Descent Learning with Over-the-Air Computation

Tamir L.S. Gez (Ben-Gurion University of the Negev); Kobi Cohen (Ben-Gurion University of the Negev)

SAM-L2: MIMO Radars and Waveform Design

Room: Salon de Roses A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): TBA

08:15 AM

3098 (SAM-L2.1): Target Velocity Estimation for Quantization-Based Cooperative MIMO Radar and Communications System

Zhen Wang (Southwest Petroleum University); xue xue yan (yanxuedan); Qian He (University of Electronic Science and Technology of China); Rick S Blum (Lehigh University)

08:30 AM

3121 (SAM-L2.2): MIMO RADAR TRANSMIT BEAMPATTERN MATCHING VIA MANIFOLD OPTIMIZATION

Weijie Xiong (University of Electronic Science and Technology of China); Jinfeng Hu (University of Electronic Science and Technology of China); Kai Zhong (University of Electronic Science and Technology of China)

08:45 AM

3298 (SAM-L2.3): Optimal Carrier Frequency Design for Frequency Diverse Array MIMO Radar

Jie Cheng (University of electronic science and technology of China); Maria Juhlin (Lund University); Wen-Qin Wang (University of electronic science and technology of China); Andreas Jakobsson (Lund University)

09:00 AM

4027 (SAM-L2.4): SUBSPACE-BASED DETECTOR FOR DISTRIBUTED MMWAVE MIMO RADAR SENSORS

Moein Ahmadi (University of Luxembourg, SNT); Mohammad Alaei (University of Luxembourg); Bhavani Shankar Mysore Ramarao (University of Luxembourg); Bjorn Ottersten (SnT)

09:15 AM

4108 (SAM-L2.5): Transmit Energy Focusing for Parameter Estimation in Transmit Beamspace Slow-time MIMO Radar

Tingting Zhang (Nanjing University of Science and Technology); Feng Xu (Aalto University); Sergiy Vorobyov ()

09:30 AM

4640 (SAM-L2.6): RANGE-ISL MINIMIZATION AND SPECTRAL SHAPING IN MIMO RADAR SYSTEMS VIA WAVEFORM DESIGN

Ehsan Raei (SnT, University of Luxembourg); Mohammad Alaei (University of Luxembourg); Bhavani Shankar Mysore Ramarao (University of Luxembourg); Bjorn Ottersten (SnT)

SLT-L15: Speech Dysarthria

Room: Delphi

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Stefan Goetze, Sudarsana Kadiri

08:15 AM

1699 (SLT-L15.1): On Using the UA-Speech and TORGO Databases to Validate Automatic Dysarthric Speech Classification Approaches

Guilherme Schu (Idiap); Parvaneh janbakhshi (Bayer AG); Ina Kodrasi (Idiap Research Institute)

08:30 AM

2179 (SLT-L15.2): Automatic Severity Classification of Dysarthric speech by using Self-supervised Model with Multi-task Learning

Eun Jung Yeo (Seoul National University); Kwanghee Choi (Sogang University); Sunhee Kim (Seoul National University); Minhwa Chung (Seoul National University)

Thursday, June 8

08:45 AM

2575 (SLT-L15.3): AN ANALYSIS OF DEGENERATING SPEECH DUE TO PROGRESSIVE DYSARTHRIA ON ASR PERFORMANCE

Katrin Tomanek (Google); Katie Seaver (Google); Pan-Pan Jiang (Google); Richard Cave (Google); Lauren Harrell (Google); Jordan Green (MGH Institute of Health Professions)

09:00 AM

3635 (SLT-L15.4): STATISTICAL ANALYSIS OF SPEECH DISORDER SPECIFIC FEATURES TO CHARACTERISE DYSARTHRIA SEVERITY LEVEL

AMLU ANNA JOSHY (COLLEGE OF ENGINEERING TRIVANDRUM); P. N. PARAMESWARAN (COLLEGE OF ENGINEERING TRIVANDRUM); Siddharth R. Nair (College of Engineering Trivandrum); Rajeev Rajan (Government Engineering College, Barton Hill, Trivandrum)

09:15 AM

6225 (SLT-L15.5): Wav2vec-based Detection and Severity Level Classification of Dysarthria from Speech

Farhad Javanmardi (Aalto University); Saska Tirronen (Aalto University); Manila Kodali (Aalto University); Sudarsana Reddy Kadiri (Aalto University); Paavo Alku (Aalto University)

09:30 AM

6762 (SLT-L15.6): Acoustic Modelling from Raw Source and Filter Components for Dysarthric Speech Recognition (SPS Journal Paper)*

Zhengjun Yue (King's College London); Erfan Loweimi (University of Cambridge); Zoran Cvetkovic (King's College London); Heidi Christensen (University of Sheffield); Jon Barker (Professor)

SLT-L16: Speech Emotion Recognition: General Topics I

Room: Athina

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Chi-Chun Lee, Jaebok Kim

08:15 AM

2490 (SLT-L16.1): Multi-Scale Receptive Field Graph Model for Emotion Recognition in Conversations

JIE WEI (Xi'an Jiaotong University); Guanyu Hu (Xi'an Jiaotong University); Anh Tuan Luu (Nanyang Technological University); Xinyu Yang (Xi'an Jiaotong University); WenJing Zhu (DXM)

08:30 AM

3918 (SLT-L16.2): MGAT: Multi-granularity Attention based Transformers for Multi-modal Emotion Recognition

WeiQuan Fan (South China University of Technology); Xiaofen Xing (South China University of Technology); Bolun Cai (Shopee); Xiangmin Xu (South China University of Technology)

08:45 AM

4523 (SLT-L16.3): ACHIEVING FAIR SPEECH EMOTION RECOGNITION VIA PERCEPTUAL FAIRNESS

Woan-Shiuan Chien (Department of Electrical Engineering, National Tsing Hua University); Chi-Chun Lee (National Tsing Hua University)

09:00 AM

5023 (SLT-L16.4): Personalized Task Load Prediction in Speech Communication

Robert P Spang (TU Berlin); Karl El Hajal (EPFL); Sebastian Möller (TU Berlin); Milos Cernak (Logitech Europe)

09:15 AM

5075 (SLT-L16.5): DWFORMER: DYNAMIC WINDOW TRANSFORMER FOR SPEECH EMOTION RECOGNITION

Shuaiqi Chen (School of Electronic and Information Engineering, South China University of Technology); Xiaofen Xing (South China University of Technology); Weibin Zhang (VoiceAI Technologies); Weidong Chen (South China University of Technology); Xiangmin Xu (South China University of Technology)

09:30 AM

5730 (SLT-L16.6): Multi-View Learning for Speech Emotion Recognition With Categorical Emotion, Categorical Sentiment, and Dimensional Scores

Daniel Tompkins (Microsoft); Dimitra Emmanouilidou (Microsoft Research); Soham Deshmukh (Microsoft); Benjamin Elizalde (Microsoft)

Thursday, June 8

SS-L10: Learning on graphs for biology and medicine

Room: Nafsika B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Dorina Thanou,

08:15 AM

2914 (SS-L10.1): Deep spatio-temporal multiplex graph learning for cardiac imaging classification

Jaume Banus (Lausanne University Hospital (CHUV)); Augustin Ogier (Lausanne University Hospital (CHUV)); Roger Hullin (Lausanne University Hospital (CHUV)); Philippe Meyer (Geneva University Hospital (HUG)); Ruud Van Heeswijk (Lausanne University Hospital (CHUV)); Jonas Richiardi (Lausanne University Hospital (CHUV))

08:30 AM

4165 (SS-L10.2): GRAPH SIGNAL PROCESSING FOR NEUROIMAGING TO REVEAL DYNAMICS OF BRAIN STRUCTURE-FUNCTION COUPLING

Maria Giulia Preti (EPFL); Thomas A.W. Bolton (Centre Hospitalier Universitaire Vaudois); Alessandra Griffa (EPFL/UNIGE/CHUV); Dimitri Van De Ville (Ecole Polytechnique Fédérale de Lausanne - LIB)

08:45 AM

4375 (SS-L10.3): Multiple Signed Graph Learning for Gene Regulatory Network Inference

Abdullah Karaaslanli (Michigan State University); Satabdi Saha (Michigan State University); Taps Maiti (Michigan State University); Selin Aviyente (Michigan State University)

09:00 AM

4599 (SS-L10.4): Predicting Brain Age using Transferable CoVariance Neural Networks

Saurabh Sihag (University of Pennsylvania); Gonzalo Mateos (University of Rochester); Corey McMillan (University of Pennsylvania); Alejandro Ribeiro (University of Pennsylvania)

09:15 AM

6456 (SS-L10.5): Spatial Graph Signal Interpolation with an Application for Merging BCI Datasets with Various Dimensionalities

Yassine El Ouahidi (IMT Atlantique); Lucas Drumetz (IMT Atlantique); Giulia Lioi (IMT Atlantique); Nicolas Farrugia (IMT Atlantique); Bastien Pasdeloup (IMT Atlantique, Lab-STICC); Vincent Gripon (IMT Atlantique)

SS-L9: Intelligent and Semantic Communications for 5G Mobile Networks and Beyond

Room: Nefeli A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Zhaohui Yang, Mingzhe Chen

08:15 AM

2242 (SS-L9.1): Rate Region Characterization for Semantics and Bits Based Multiuser Communications

Xidong Mu (Queen Mary University of London); Yuanwei Liu (Queen Mary University of London)

08:30 AM

3359 (SS-L9.2): HARQ Delay Minimization of 5G Wireless Network with Imperfect Feedback

Weihang Ding (King's College London); Mohammad Shikh-Bahaei (King's College London)

08:45 AM

3484 (SS-L9.3): Multi-Agent Reinforcement Learning for Covert Semantic Communications over Wireless Networks

Yining Wang (Beijing University of Posts and Telecommunications); Ye Hu (Columbia University); HONGYANG DU (Nanyang Technological University); Tao Luo (Beijing University of Posts and Communications); Dusit Niyato ()

09:00 AM

5623 (SS-L9.4): Asynchronous Federated Learning for Real-time Multiple Licence Plate Recognition through Semantic Communication

renyou xie (Central South University); Chaojie Li (The University of New South Wales); Xiaojun Zhou (Central South University); Zhao Yang Dong (The University of New South Wales)

09:15 AM

6234 (SS-L9.5): An Efficient Relay Selection Scheme for Relay-assisted HARQ

Weihang Ding (King's College London); Mohammad Shikh-Bahaei (King's College London)

09:30 AM

6293 (SS-L9.6): Adaptive CSI Feedback with Hidden Semantic Information Transfer

Jiaqi Cao (ShanghaiTech University); Lixiang Lian (ShanghaiTech University); Yijie Mao (ShanghaiTech University); Bruno Clerckx (Imperial College London)

Thursday, June 8

AASP-P5: Audio and Speech Quality Measurements

Room: Dome A

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Ante Jukić,

6220 (AASP-P5.1): AUDIO QUALITY ASSESSMENT OF VINYL MUSIC COLLECTIONS USING SELF-SUPERVISED LEARNING

Alessandro Ragano (University College Dublin); Emmanouil Benetos (Queen Mary University of London); Andrew Hines (University College Dublin)

4685 (AASP-P5.2): Speech Intelligibility Classifiers from 550k Disordered Speech Samples

Subhashini Venugopalan (Google); Jimmy Tobin (Google); Samuel J. Yang (Google); Katie Seaver (Google); Richard Cave (Google); Pan-Pan Jiang (Google); Neil Zeghidour (Google); Rus Heywood (Google); Jordan Green (MGH Institute of Health Professions); Michael Brenner (Google/Harvard)

1493 (AASP-P5.3): NORD: Non-Matching Reference Based Relative Depth Estimation From Binaural Audio

Pranay Manocha (Princeton University); Israel D Gebru (Facebook); Anurag Kumar (Facebook Research); Dejan Markovic (Facebook Reality Labs); Alexander Richard (Facebook Reality Labs)

2988 (AASP-P5.4): Ensemble of Deep Neural Network Models for MOS Prediction

Marie Kunešová (University of West Bohemia); Jindrich Matousek (University of West Bohemia, Pilsen, Czech Republic); Jan Lehečka (University of West Bohemia); Jan Svec (University of West Bohemia); Josef Michalek (University of West Bohemia); Daniel Tihelka (University of West Bohemia); Martin Bulin (University of West Bohemia); Zdenek Hanzlicek (University of West Bohemia); Marketa Rezaczkova (University of West Bohemia)

4448 (AASP-P5.5): On Crowdsourcing-Design with Comparison Category Rating for Evaluating Speech Enhancement algorithms

Angélica Stephania Zambrano Suárez (DTU); Clement Laroche (GN Audio); Line Clemmensen (DTU); Sneha Das (Technical University of Denmark)

4660 (AASP-P5.6): EFFICIENT INTELLIGIBILITY EVALUATION USING KEYWORD SPOTTING: A STUDY ON AUDIO-VISUAL SPEECH ENHANCEMENT

Cassia Valentini (University of Edinburgh); Andrea L Aldana (Edinburgh University); Ondrej Klejch (University of Edinburgh); Peter Bell (University of Edinburgh)

4898 (AASP-P5.7): TORCHAUDIO-SQUIM: REFERENCE-LESS SPEECH QUALITY AND INTELLIGIBILITY MEASURES IN TORCHAUDIO

Anurag Kumar (Facebook Reality Labs); Ke Tan (Meta Platforms, Inc.); Zhaoheng Ni (Meta); Pranay Manocha (Princeton University); Xiaohui Zhang (Meta); Ethan Henderson (Meta Reality Labs Research); Buye Xu (Meta Reality Labs Research)

4980 (AASP-P5.8): Speech MOS multi-task learning and rater bias correction

Haleh Akrami (Signal and Image Processing Institute at University of Southern California); Hannes Gamper (Microsoft)

5008 (AASP-P5.9): SpeechLMScore: Evaluating speech generation using speech language model

Soumi Maiti (CMU); Yifan Peng (Carnegie Mellon University); Takaaki Saeki (The University of Tokyo); Shinji Watanabe (Carnegie Mellon University)

5519 (AASP-P5.10): On the robustness of non-intrusive speech quality model by adversarial examples

Hsin-Yi Lin (Academia Sinica); Huan-Hsin Tseng (Academia Sinica); Yu Tsao (Academia Sinica)

3676 (AASP-P5.11): TG-Critic: A Timbre-Guided Model for Reference-Independent Singing Evaluation

Xiaoheng Sun (NetEase Cloud Music); Yuejie Gao (Hangzhou Netease cloud Music Technology Co., Ltd); Hanyao Lin (Fudan University); Huaping Liu (Hangzhou Netease cloud Music Technology Co., Ltd)

AASP-P6: Acoustic Modeling; Auditory Modeling for Hearing Instruments

Room: Dome B

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Antoine Deleforge,

5004 (AASP-P6.1): On the Reduction of Large-Scale Room Acoustic Models

Pavlos Stoikos (University of Thessaly); Olympia Axelou (University of Thessaly); George Floros (University of Thessaly); Nestor Evmorfopoulos (University of Thessaly); George Stamoulis (University of Thessaly)

5278 (AASP-P6.2): Spherical sector harmonics based soundfield radial extrapolation and robustness analysis

Hanwen Bi (ANU); thushara abhayapala (The Australian National University); fei ma (Australian National University); Prasanga Samarasinghe (Australian National University)

Thursday, June 8

2620 (AASP-P6.3): SPARSE AND STRUCTURED MODELLING OF UNDERWATER ACOUSTIC CHANNEL IMPULSE RESPONSES

Chaoran Yang (Harbin Engineering University); Qing Ling (Harbin Engineering University); Xueli Sheng (Harbin Engineering University); Mengfei Mu (Harbin Engineering University); Andreas Jakobsson (Lund University)

6487 (AASP-P6.4): Image source method based on the directional impulse responses

Jiarui Wang (The Australian National University); Prasanga Samarasinghe (Australian National University); thushara abhayapala (The Australian National University); Jihui (Aimee) Zhang (University of Southampton)

2454 (AASP-P6.5): Fast Low-latency Convolution by Low-rank Tensor Approximation

Martin Jälmbly (KU Leuven); Filip Elvander (Aalto University); Toon van Waterschoot (Department of Electrical Engineering (ESAT-STADIUS/ETC))

4976 (AASP-P6.6): LEARNING ENVIRONMENTAL STRUCTURE USING ACOUSTIC PROBES WITH A DEEP NEURAL NETWORK

Toros ARIKAN (MIT); Amir Weiss (Massachusetts Institute of Technology); Hari Vishnu (NUS); Grant Deane (UCSD); Andrew C Singer (University of Illinois); Gregory W Wornell (MIT)

3074 (AASP-P6.7): SPECTRO-TEMPORAL POST-FILTERING VIA SHORT-TIME TARGET CANCELLATION FOR DIRECTIONAL SPEECH ENHANCEMENT IN A DUAL-MICROPHONE HEARING AID

Marcos A Cantu (Carl von Ossietzky University of Oldenburg); Volker Hohmann (Carl von Ossietzky University of Oldenburg)

5924 (AASP-P6.8): Immersive enhancement and removal of loudspeaker sound using wireless assistive listening systems and binaural hearing devices

Ryan M Corey (University of Illinois Chicago); Andrew C Singer (University of Illinois)

6169 (AASP-P6.9): A MODEL-BASED HEARING COMPENSATION METHOD USING A SELF-SUPERVISED FRAMEWORK

Yadong Niu (Peking University); Nan Li (peking university); Xihong Wu (Peking University); Jing Chen (Peking University)

5195 (AASP-P6.10): RAPID AUDIOMETRIC EVALUATION FOR PERSONALIZED HEADPHONE LISTENING

Matthew J. Goupell (University of Maryland - College Park); Marjan Davoodian (University of Maryland - College Park); Sarah Weinstein (University of Maryland - College Park); David Gadzinski (Visisonics Corporation); Dmitry Zotkin (Visisonics); Kaushik Sethunath (Visisonics Corporation); Ramani ramani.d@visisonics.com (Visisonics Corporation)

2393 (AASP-P6.11): A DNN-based hearing-aid strategy for real-time processing: One size fits all

Fotios Drakopoulos (Ghent University); Arthur Van Den Broucke (Ghent University); Sarah Verhulst (Ghent University)

3898 (AASP-P6.12): GRAD-CAM-INSPIRED INTERPRETATION OF NEARFIELD ACOUSTIC HOLOGRAPHY USING PHYSICS-INFORMED EXPLAINABLE NEURAL NETWORK

Hagar Kafri (Bar Ilan University); Marco Olivieri (Politecnico di Milano); Fabio Antonacci (Politecnico di Milano); Mordehay Moradi (Bar illan University); Augusto Sarti (Politecnico di Milano); Sharon Gannot (Bar-Ilan University)

IFS-P1: Anonymization, Data Privacy, and Biometrics

Room: Dome H

Type: Poster

08:15 AM to 09:45 AM

Chair(s): TBA

121 (IFS-P1.1): FedPrompt: Communication-Efficient and Privacy-Preserving Prompt Tuning in Federated Learning

Haodong Zhao (Shanghai Jiao Tong University); Wei Du (Shanghai Jiao Tong University); Fangqi Li (SEIEE, Shanghai Jiao Tong University); Peixuan Li (Shanghai Jiao Tong University); Gongshen Liu (Shanghai Jiao Tong University)

4062 (IFS-P1.2): FINE-GRAINED PRIVATE KNOWLEDGE DISTILLATION

Yuntong Li (Guangzhou University); Shaowei Wang (Guangzhou University); Yingying Wang (Guangzhou University); Jin Li (Guangzhou University); Yuqiu Qian (Tencent Inc.); Bangzhou Xin (University of Science and Technology of China); Wei Yang (University of Science and Technology of China)

2183 (IFS-P1.3): CPA: Compressed Private Aggregation for Scalable Federated Learning over Massive Networks

Natalie Lang (Ben-Gurion University of the Negev); Elad Sofer (Ben-Gurion University of the Negev); Nir Shlezinger (Ben-Gurion University); Rafael D'Oliveira (Clemson University); Salim El Rouayheb (Rutgers University)

2570 (IFS-P1.4): HE-GAN: Differentially Private GAN using Hamiltonian Monte Carlo based Exponential Mechanism

Usman Hassan (University of Kentucky); Dongjie Chen (University of California, Davis); Sen-ching S Cheung (University of Kentucky); Chen-Nee Chuah (University of California Davis)

3110 (IFS-P1.5): Backdoor Attack Against Automatic Speaker Verification Models in Federated Learning

Dan Meng (OPPO Research Institute); Xue Wang (Wuhan University); Jun Wang (OPPO Research Institute)

Thursday, June 8

3529 (IFS-P1.6): Sparse Black-box Inversion Attack With Limited Information

Yixiao Xu (Institute of Computer Application, China Academy of Engineering Physics); Xiaolei Liu (Institute of Computer Application, China Academy of Engineering Physics); Teng Hu (Institute of Computer Application, China Academy of Engineering Physics); Bangzhou Xin (Institute of Computer Application, China Academy of Engineering Physics); Run Yang (Institute of computer application, Chinese Academy of Engineering Physics)

4888 (IFS-P1.7): ROW CONDITIONAL-TGAN FOR GENERATING SYNTHETIC RELATIONAL DATABASES

Mohamed Gueye (CROESUS); Yazid Attabi (CROESUS); Maxime Dumas (CROESUS)

5386 (IFS-P1.8): A Speech Representation Anonymization Framework via Selective Noise Perturbation

Minh Tran (University of Southern California); Mohammad Soleymani (University of Southern California)

1560 (IFS-P1.9): Liveness Score-Based Regression Neural Networks for Face Anti-Spoofing

Youngjun Kwak (Kakaobank); Minyoung Jung (KETI); Hunjae Yoo (Kakaobank); Jinho Shin (Kakaobank); Changick Kim (KAIST)

1665 (IFS-P1.10): Benchmarking Cross-Domain Face Recognition with Avatars, Caricatures and Sketches

Ahmad Foroughi (Hochschule Darmstadt); Christian Rathgeb (Hochschule Darmstadt); Mathias Ibsen (Hochschule Darmstadt); Christoph Busch (Hochschule Darmstadt)

2449 (IFS-P1.11): Effect of Lossy Compression Algorithms on Face Image Quality and Recognition

Torsten Schlett (Hochschule Darmstadt); Sebastian Schachner (Hochschule Darmstadt); Christian Rathgeb (Hochschule Darmstadt); Juan Tapia (hda); Christoph Busch (Hochschule Darmstadt)

3662 (IFS-P1.12): Single Domain Dynamic Generalization for Iris Presentation Attack Detection

Yachun Li (Hikvision Research Institute); Jingjing Wang (Hikvision Research Institute); yuhui chen (HIKVISION); Di Xie (Hikvision Research Institute); Shiliang Pu (Hikvision Research Institute)

IFS-P2: Multimedia Forensics

Room: Dome I

Type: Poster

08:15 AM to 09:45 AM

Chair(s): TBA

290 (IFS-P2.1): Audio Cross Verification Using Dual Alignment Likelihood Ratio Test

Heidi Lei (MIT); Arm Wonghirundacha (Pomona College); Irmak Bukey (Pomona College); Timothy Tsai (Harvey Mudd College)

606 (IFS-P2.2): Classification of Synthetic Facial Attributes by Means of Hybrid Classification/Localization Patch-based Analysis

Jun Wang (University of Siena); Benedetta Tondi (University of Siena); Mauro Barni (University of Siena)

3684 (IFS-P2.3): Learning Expressive and Generalizable Motion Features for Face Forgery Detection

Jingyi Zhang (Hikvision Research Institute); Peng Zhang (Hikvision Research Institute); Jingjing Wang (Hikvision Research Institute); Di Xie (Hikvision Research Institute); Shiliang Pu (Hikvision Research Institute)

3899 (IFS-P2.4): DOUBLE COMPRESSION DETECTION BASED ON THE DE-BLOCKING FILTERING OF HEVC VIDEOS

Xiangui Kang (Sun Yat-Sen University); pengcheng su (Sun Yat-sen University); Zisheng Huang (Sun Yat-sen University); Yifang Chen (Guangdong Polytechnic Normal University); Jie Wang (Sun Yat-sen University)

4031 (IFS-P2.5): Electric Network Frequency Detection Using Least Absolute Deviations

Christos Korgialas (Aristotle University of Thessaloniki); Constantine Kotropoulos (Aristotle University of Thessaloniki)

4490 (IFS-P2.6): Hearing and Seeing Abnormality: Self-supervised Audio-Visual Mutual Learning for Deepfake Detection

ChangSung Sung (National Taiwan University); Jun-Cheng Chen (Academia Sinica); Chu-Song Chen (National Taiwan University)

4514 (IFS-P2.7): Two-branch multi-scale deep neural network for generalized document recapture attack detection

Li Jiaying (City University of Hong Kong); Chenqi KONG (City University of Hong Kong); Shiqi Wang (City University of Hong Kong); Haoliang Li (CityU)

5689 (IFS-P2.8): LEARNING TO LOCATE THE TEXT FORGERY IN SMARTPHONE SCREENSHOTS

Zeqin Yu (shenzhen university); Bin Li (Shenzhen University); Yuzhen Lin (Shenzhen University); Jinhua Zeng (Academy of Forensic Science); Jishen Zeng (Alibaba Group)

5944 (IFS-P2.9): MAKE YOUR ENEMY YOUR FRIEND: IMPROVING IMAGE ROTATION ANGLE ESTIMATION WITH HARMONICS

yu kun (School of Computer Science & Technology Southwest University of Science and Technology Mianyang, China); Morteza Darvish Morshedi Hosseini (State University of New York at Binghamton); Anjie Peng (Southwest University of Science and Technology); Hui Zeng (Southwest University of Science and Technology); Miroslav Goljan (State University of New York at Binghamton)

Thursday, June 8

6310 (IFS-P2.10): On the detection of synthetic images generated by diffusion models

Riccardo Corvi (University Federico II of Naples); Davide Cozzolino (University Federico II of Naples); Giada Zingarini (University Federico II of Naples); Giovanni Poggi (University Federico II of Naples); Koki Nagano (NVIDIA); Luisa Verdoliva (University Federico II of Naples)

2597 (IFS-P2.11): TRUSTERA: A LIVE CONVERSATION REDACTION SYSTEM

Evandro Gouvea (Interaction); Ali Dadgar (Interactions LLC); Shahab Jalalvand (Interactions LLC); Rathi Chengalvarayan (Interactions LLC); Badrinath Jayakumar (Interactions LLC); Ryan Price (Interactions); Nicholas Ruiz (Interactions); Jennifer McGovern (Interactions LLC); Srinivas Bangalore (Interactions LLC); Ben Stern (Interactions LLC)

6096 (IFS-P2.12): Prototype-Based Layered Federated Cross-Modal Hashing

Jiale Liu (Shandong University); Yu-Wei Zhan (Shandong University); Xin Luo (Shandong University); Zhen-Duo Chen (Shandong University); Yongxin Wang (Shandong Jianzhu University); Xin-Shun Xu (Shandong University)

IVMSP-P23: Object Recognition

Room: Dome J

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Wei Feng, Weiming Hu

2770 (IVMSP-P23.1): SQA: STRONG GUIDANCE QUERY WITH SELF-SELECTED ATTENTION FOR HUMAN-OBJECT INTERACTION DETECTION

Feng Zhang (Zhejiang University of Technology); Sheng Liu (Zhejiang University of Technology); Bingnan Guo (Zhejiang University of Technology); ruixiang chen (Zhejiang University of Technology); Junhao Chen (Zhejiang University of Technology)

1033 (IVMSP-P23.2): COMBINING THE SILHOUETTE AND SKELETON DATA FOR GAIT RECOGNITION

Likai Wang (Tianjin University); Ruize Han (College of Intelligence and Computing, Tianjin University); Wei Feng (College of Intelligence and Computing, Tianjin University, China)

1139 (IVMSP-P23.3): Learning from the raw domain: cross modality distillation for compressed video action recognition

Yufan Liu (Institute of Automation, Chinese Academy Sciences); Jiajiong Cao (Ant Financial Service Group); Weiming Bai (Chinese Academy of Sciences); Bing Li (National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences); Weiming Hu (Institute of Automation, Chinese Academy of Sciences)

1646 (IVMSP-P23.4): GAITCOTR: improved spatial-temporal representation for gait recognition with a hybrid convolution-transformer framework

Jingqi Li (Fudan University); Yuzhen Zhang (Fudan University); Hongming Shan (Fudan University); Junping Zhang (Fudan University)

2455 (IVMSP-P23.5): TrOMR:Transformer-based Polyphonic Optical Music Recognition

Yixuan Li (Hangzhou Netease cloud Music Technology Co., Ltd); Huaping Liu (Hangzhou Netease cloud Music Technology Co., Ltd); Qiang Jin (Hangzhou Netease cloud Music Technology Co., Ltd); Miaomiao Cai (Hangzhou Netease cloud Music Technology Co., Ltd); Peng Li (NetEase Cloud Music)

3021 (IVMSP-P23.6): RETHINK LONG-TAILED RECOGNITION WITH VISION TRANSFORMS

Zhenghuo Xu (Tsinghua University); Shuo Yang (Tsinghua university); Xingjun Wang (Tsinghua University); Chun Yuan (Graduate school at ShenZhen, Tsinghua university)

3602 (IVMSP-P23.7): Mask Guided Selective Context Decoding for Handwritten Chinese Text Recognition

tao li (University of Science and Technology of China); shilian wu (University of Science and Technology of China); Zengfu Wang (Institute of Intelligent Machines, Chinese Academy of Sciences)

3838 (IVMSP-P23.8): Long-tailed Image Recognition with Dynamic Re-weighting

Xinyuan LI (Ritsumeikan University); Yu Wang (Hitotsubashi University); Jien Kato (Ritsumeikan University)

4836 (IVMSP-P23.9): Motion Matters: A Novel Motion Modeling For Cross-View Gait Feature Learning

Jingqi Li (Fudan University); Jiaqi Gao (Fudan University); Yuzhen Zhang (Fudan University); Hongming Shan (Fudan University); Junping Zhang (Fudan University)

5194 (IVMSP-P23.10): CROSS-MODAL MATCHING AND ADAPTIVE GRAPH ATTENTION NETWORK FOR RGB-D SCENE RECOGNITION

Yuhui Guo (Renmin University of China); Xun Liang (Renmin University of China); james kwok (The Hong Kong University of Science and Technology); Xiangping Zheng (Renmin University of China); Bo Wu (Renmin University of China); Yuefeng Ma (Qufu Normal University)

5794 (IVMSP-P23.11): EXPLOITING 3D HUMAN RECOVERY FOR ACTION RECOGNITION WITH SPATIO-TEMPORAL BIFURCATION FUSION

Na Jiang (Information Engineering College, Capital Normal University); Wei Quan (Capital Normal University); Qichuan Geng (Capital Normal University); Zhiping Shi (Capital Normal University); Peng Xu (Capital Normal University)

Thursday, June 8

6387 (IVMSP-P23.12): In-Sensor & Neuromorphic Computing are all you need for Efficient Computer Vision

Gourav Datta (University of Southern California); Zeyu Liu (University of Southern California); Md Abdullah-AI Kaiser (University of Southern California); Souvik Kundu (Intel Labs); Joe Mathai (Information Sciences Institute); Zihan Yin (USC); Ajey Jacob (USC); Akhilesh Jaiswal (USC); Peter A. Bearel (University of Southern California)

IVMSP-P24: Identification Detection

Room: Dome K

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Fanman Meng, Nenghai Yu

6015 (IVMSP-P24.1): Neighborhood Information-Based Label Refinement for Person Re-Identification with Label Noise

Xian Zhong (Wuhan University of Technology); Shuaipeng Su (Wuhan University of Technology); Wenxuan Liu (Wuhan University of Technology); Xuemei Jia (Wuhan University); Wenxin Huang (Hubei University); Mengdie Wang (Wuhan University Of Technology)

3588 (IVMSP-P24.2): MFAT: A Multi-level Feature Aggregated Transformer for person re-identification

Bowen Tan (University of Electronic Science and Technology of China); Linfeng Xu (University of Electronic Science and Technology of China); Zihuan Qiu (University of Electronic Science and Technology of China); Qingbo Wu (University of Electronic Science and Technology of China); Fanman Meng (University of Electronic Science and Technology of China)

3730 (IVMSP-P24.3): DUAL-UNCERTAINTY GUIDED CURRICULUM LEARNING AND PART-AWARE FEATURE REFINEMENT FOR DOMAIN ADAPTIVE PERSON RE-IDENTIFICATION

Zhangping Liu (University of Science and Technology of China); Bin Liu (University of Science and Technology of China); Zhiwei Zhao (University of Science and Technology of China); Qi Chu (University of Science and Technology of China); Nenghai Yu (University of Science and Technology of China)

5841 (IVMSP-P24.4): Joint Multi-Level Feature Network for Lightweight Person Re-Identification

Yunzuo Zhang (Shijiazhuang Tiedao University); Weili Kang (Shijiazhuang Tiedao University); Yameng Liu (Shijiazhuang Tiedao University); Pengfei Zhu (Shijiazhuang Tiedao University)

1967 (IVMSP-P24.5): Cross-head supervision for crowd counting with noisy annotations

Mingliang Dai (Fudan University); Zhizhong Huang (Fudan University); Jiaqi Gao (Fudan University); Hongming Shan (Fudan University); Junping Zhang (Fudan University)

4037 (IVMSP-P24.6): Recurrent Fine-Grained Self-Attention Network for Video Crowd Counting

Jifan Zhang (School of Electronic and Computer Engineering, Peking University); Zhe Wu (Peng Cheng Laboratory); xinfeng zhang (University of Chinese Academy of Sciences); Guoli Song (Peng Cheng Laboratory); Yaowei Wang (PengCheng Laboratory); Jie Chen (Peking University)

4414 (IVMSP-P24.7): DENSITYTOKEN: WEAKLY-SUPERVISED CROWD COUNTING WITH DENSITY CLASSIFICATION

Zaiyi Hu (Northwestern Polytechnical University); Binglu Wang (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

4549 (IVMSP-P24.8): Weakly-Supervised Scene-Specific Crowd Counting Using Real-Synthetic Hybrid Data

Yaowu Fan (Northwestern Polytechnical University); Jia Wan (University of California, San Diego); Yuan Yuan (Northwestern Polytechnical University); Qi Wang (Northwestern Polytechnical University)

MLSP-P29: Pattern Recognition and Classification II

Room: Dome F

Type: Poster

08:15 AM to 09:45 AM

Chair(s): TBA

3740 (MLSP-P29.1): A MEMORY-FREE EVOLVING BIPOLAR NEURAL NETWORK FOR EFFICIENT MULTI-LABEL STREAM LEARNING

Sourav Mishra (Indian Institute of Science, Bangalore); Suresh Sundaram (Indian Institute of Science)

5829 (MLSP-P29.2): Multi-view K-means with Laplacian Embedding

zhezhen hao (Northwestern Polytechnical University); Zhoumin Lu (Northwestern Polytechnical University); Feiping Nie (Northwestern Polytechnical University); Rong Wang (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

5682 (MLSP-P29.3): M-CTRL: A CONTINUAL REPRESENTATION LEARNING FRAMEWORK WITH SLOWLY IMPROVING PAST PRE-TRAINED MODEL

Jin-Seong Choi (Hanyang university); Jae-Hong Lee (Hanyang University); Chae-Won Lee (Hanyang University, Seoul); Joon-Hyuk Chang (Hanyang University)

Thursday, June 8

1411 (MLSP-P29.4): PSEUDO-QUERY GENERATION FOR SEMI-SUPERVISED VISUAL GROUNDING WITH KNOWLEDGE DISTILLATION

Jianglin Jin (East China Normal University); Jiabo Ye (East China Normal University); Xin Lin (ECNU); Liang He (ECNU)

1262 (MLSP-P29.5): CLASSIFICATION VIA SUBSPACE LEARNING MACHINE (SLM): METHODOLOGY AND PERFORMANCE EVALUATION

Hongyu Fu (University of Southern California); Yijing Yang (University of Southern California); Vinod Mishra (Army Research Lab); C.-C. Jay Kuo (USC)

1271 (MLSP-P29.6): Bag of Tricks with Quantized Convolutional Neural Networks for image classification

Jie Hu (Institute of Software Chinese Academy of Sciences); Mengze Zeng (Momenta); Enhua Wu (SKLCS, Institute of Software, Chinese Academy of Sciences, Beijing, China); Faculty of Science and Technology, University of Macau, Macao, China)

3557 (MLSP-P29.7): REDUCING THE COMPUTATIONAL COMPLEXITY OF LEARNING WITH RANDOM CONVOLUTIONAL FEATURES

Mohammad Amin Omid (Shahed University); Babak Seyfe (Shahed University); Shahrokh Valaee (University of Toronto)

427 (MLSP-P29.8): Pseudo-Inverted Bottleneck Convolution for DARTS Search Space

Arash Ahmadian (University of Toronto); Yue Fei (University of Toronto); Louis S.P. Liu (University of Toronto); Konstantinos N Plataniotis (UofT); Mahdi S Hosseini (Concordia University)

596 (MLSP-P29.9): Adaptive Scale and Spatial Aggregation for Real-time Object Detection

Wei Chen (College of Computer, National University of Defense Technology); Yulin He (National University of Defense Technology); Zhengfa Liang (Defense Innovation Institute); Yulan Guo (National University of Defense Technology)

5034 (MLSP-P29.10): Variable Rate Allocation for Vector-Quantized Autoencoders

Federico Baldassarre (KTH - Royal Institute of Technology); Alaaeldin M El-Nouby (Facebook AI Research); Herve Jegou (Facebook AI Research)

1083 (MLSP-P29.11): Mixed Sample Augmentation for Online Distillation

Yiqing Shen (Johns Hopkins University)

1126 (MLSP-P29.12): FlowReg: Latent Space Regularization using Normalizing Flow for Limited Samples Learning

Chi Wang (Queen's University Belfast); Jian Gao (Queen's University Belfast); Yang Hua (Queen's University Belfast); Hui Wang (Queen's University Belfast)

MLSP-P30: Pattern Recognition and Classification III

Room: Dome G

Type: Poster

08:15 AM to 09:45 AM

Chair(s): TBA

1400 (MLSP-P30.1): Bias Identification with RankPix Saliency

salamata konate (QUT); Leo Lebrat (CSIRO); Rodrigo Santa Cruz (CSIRO); Clinton Fookes (Queensland University of Technology); Andrew Bradley (Queensland University of Technology); Olivier Salvado (CSIRO)

371 (MLSP-P30.2): ModalDrop: Modality-aware Regularization for Temporal-Spectral Fusion in Human Activity Recognition

Xin Zeng (Institute of Computing Technology, Chinese Academy of Sciences); Yiqiang Chen (Institute of Computing Technology, Chinese Academy of Sciences); Benfeng Xu (University of Science and Technology of China); Tengxiang Zhang (institute of computing technology, Chinese academy of sciences)

5237 (MLSP-P30.3): Modeling the Wave Equation Using Physics-Informed Neural Networks Enhanced with Attention to Loss Weights

Shaikhah Alkhadhr (Pennsylvania State University); Mohamed Almekkawy (Pennsylvania State University)

5327 (MLSP-P30.4): Fast and Exact Enumeration of Deep Networks Partitions Regions

Randall Balestriero (Facebook AI Research); Yann Lecun (Facebook)

1738 (MLSP-P30.5): Training Robust Spiking Neural Networks on Neuromorphic Data with Spatiotemporal Fragments

Haibo Shen (Huazhong University of Science and Technology); Yihao Luo (Yichang Testing Technique R&D Institute); Xiang Cao (School of Computer Science and Technology, Huazhong University of Science and Technology); Liangqi Zhang (Huazhong University of Science and Technology); Juyu Xiao (Huazhong University of Science and Technology); Tianjiang Wang (School of Computer Science and Technology, Huazhong University of Science and Technology)

1802 (MLSP-P30.6): Introducing topography in convolutional neural networks

Maxime Poli (École Normale Supérieure); Emmanuel Dupoux (EHESS, ENS, PSL University, CNRS, INRIA, META); Rachid Riad (CoML/NPI/ENS/CNRS/EHESS/INRIA/PSL/INSERM/UPEC)

Thursday, June 8

1832 (MLSP-P30.7): HiFi++: a Unified Framework for Bandwidth Extension and Speech Enhancement

Pavel Andreev (Samsung AI Center Moscow); Aibek Alanov (Artificial Intelligence Research Institute); Oleg Ivanov (Samsung AI Center Moscow); Dmitry P Vetrov (Higher School of Economics)

5459 (MLSP-P30.8): A Closer Look at Scoring Functions and Generalization Prediction

Puja Trivedi (University of Michigan); Danai Koutra (U Michigan); Jayaraman J. Thiagarajan (Lawrence Livermore National Laboratory)

2119 (MLSP-P30.9): Is Multi-Task Learning an Upper Bound for Continual Learning?

Zihao Wu (Vanderbilt University); Huy Tran (Vanderbilt University); Hamed Pirsiavash (University of California Davis); Soheil Kolouri (Vanderbilt University)

2131 (MLSP-P30.10): SUPER-RESOLUTION INFORMATION ENHANCEMENT FOR CROWD COUNTING

Jiahao Xie (Beijing University of Posts and Telecommunications); Wei Xu (Beijing University of Posts and Telecommunications); Dingkan Liang (Huazhong University of Science and Technology); Zhanyu Ma (Beijing University of Posts and Telecommunications); Kongming Liang (Beijing University of Posts and Telecommunications); Weidong Liu (China Mobile Research Institute); Rui Wang (China Mobile Research Institute); Ling Jin (China Mobile Research Institute)

5490 (MLSP-P30.11): A METHOD OF CONSTRUCTING AND AUTOMATICALLY LABELING RADIO FREQUENCY SIGNAL TRAINING DATASET FOR UAV

Chao Liu (Fudan University); Ruipeng Ma (ZhengZhou University); Zheng Si (ZhengZhou University); mingmin Chi (Fudan university)

5703 (MLSP-P30.12): A hybrid deep neural network for nonlinear causality analysis in complex industrial control system

Tian Feng (Zhejiang University); Qiming Chen (DAMOAcademy,AlibabaGroup); Yao Shi (Zhejiang University); Xun Lang (Yunnan University); Lei Xie (Zhejiang University); Hongye Su (Zhejiang University)

SAM-P4: Tracking, Data Fusion, and Sensor Networks

Room: Dome M

Type: Poster

08:15 AM to 09:45 AM

Chair(s): TBA

268 (SAM-P4.1): Deep fusion of multi-object densities using transformer

Lechi Li (Chalmers University of Technology); Chen Dai (Chalmers University of Technology); Yuxuan Xia (Chalmers University of Technology); Lennart Svensson (Chalmers University of Technology)

6240 (SAM-P4.10): Nonnegative block-term decomposition with the β -divergence: joint data fusion and blind spectral unmixing

Clémence Prévost (University of Lille); Valentin Leplat (Skoltech)

2238 (SAM-P4.2): ROBUST SUBSPACE TRACKING WITH CONTAMINATION MITIGATION VIA α -DIVERGENCE

LE Trung Thanh (University of Orleans); Aref Miri Rekavandi (University of Melbourne, Melbourne, Australia); Abd-Krim Seghouane (University of Mebourne); KARIM ABED-MERAIM (PRISME laboratory, university of Orleans, France)

2321 (SAM-P4.3): Wireless location tracking via complex-domain Super MDS with time series self-localization information

Yuya Nishi (Osaka University); Takumi Takahashi (Osaka University); Hiroki Iimori (Ericsson Research); Giuseppe Abreu (Jacobs University Bremen); Shinsuke Ibi (Doshisha University); Seichi Sampei (Osaka University)

2463 (SAM-P4.4): Angle-of-arrival Target Tracking Using a Mobile UAV In External Signal-denied Environment

Bing Zhu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences); Sheng Xu (Shenzhen Institute of Advanced technology, Chinese Academy of Sciences); Feng Rice (QinetiQ Australia); Kutluyil Dogancay (University of South Australia)

2821 (SAM-P4.5): A Distributed Adaptive Algorithm for Non-Smooth Spatial Filtering Problems

Charles Hovine (KULeuven); Alexander Bertrand (KU Leuven)

2937 (SAM-P4.6): A Computationally Efficient Algorithm for Distributed Adaptive Signal Fusion based on Fractional Programs

Cem A. Musluoglu (KU Leuven); Alexander Bertrand (KU Leuven)

3217 (SAM-P4.7): Data Driven Joint Sensor Fusion and Regression based on Geometric Mean Squared Error

Carlos A Lopez Molina (Polytechnic University of Catalonia); Jaume Riba (UPC)

4043 (SAM-P4.8): Sensor Selection for Angle of Arrival Estimation Based on the Two-Target Cramér-Rao Bound

Costas A Kokke (Delft University of Technology); Mario Coutino (TNO); Laura Anitori (TNO); Richard Heusdens (Netherlands Defence Academy); Geert Leus (TU Delft)

4149 (SAM-P4.9): Clustered Greedy Algorithm for Large-Scale Sensor Selection

Kaushani Majumder (Indian Institute of Technology, Bombay); Sibi Raj B. Pillai (Indian Institute of Technology, Bombay); Satish Mulleti (Indian Institute of Technology Bombay, India)

Thursday, June 8

SLT-P27: Speaker Recognition III: Neural Network Architecture

Room: Dome C

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Xin Wang, Themis Stafylakis

2485 (SLT-P27.1): Cross-Modal Audio-Visual Co-learning for Text-independent Speaker Verification

Meng Liu (Tianjin University); Kong Aik Lee (Institute for Infocomm Research, ASTAR); Longbiao Wang (Tianjin University); Hanyi Zhang (Tianjin University); Chang Zeng (National Institute of Informatics); Jianwu Dang (School of Computer Science and Technology, Tianjin University, Tianjin, China; School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa, Japan)

2681 (SLT-P27.2): DASA: Difficulty-Aware Semantic Augmentation for Speaker Verification

Yuanyuan Wang (Tsinghua University); Yang Zhang (Tsinghua University); Zhiyong Wu (Tsinghua University); Zhihan Yang (tsinghua); Tao Wei (Ping An Technology); Kun Zou (Ping An Technology); Helen Meng (The Chinese University of Hong Kong)

2803 (SLT-P27.3): CONVOLUTION-BASED CHANNEL-FREQUENCY ATTENTION FOR TEXT-INDEPENDENT SPEAKER VERIFICATION

Jingyu Li (The Chinese University of Hong Kong); Yusheng Tian (The Chinese University of Hong Kong); Tan Lee (The Chinese University of Hong Kong)

2889 (SLT-P27.4): PCF: ECAPA-TDNN with Progressive Channel Fusion for Speaker Verification

Zhenduo Zhao (Institute of Acoustics, Chinese Academy of Sciences); Zhuo Li (Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics, Chinese Academy of Sciences); Wenchao Wang (Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics, Chinese Academy of Sciences, Beijing, China); pengyuan zhang (Institute of Acoustics, Chinese Academy of Sciences)

2892 (SLT-P27.5): Short-segment speaker verification using ECAPA-TDNN with multi-resolution encoder

Sangwook Han (GIST); Youngdo Ahn (GIST); Kyeognmuk Kang (GIST); Jong Won Shin (Gwangju Institute of Science and Technology)

3796 (SLT-P27.6): Hybrid Neural Network With Cross- and Self-Module Attention Pooling for Text-Independent Speaker Verification

Jahangir Alam (Computer Research Institute of Montreal (CRIM), Montreal (Quebec) Canada); Woohyun Kang (Amazon Web Services); Abderrahim Fathan (Computer Research Institute of Montreal (CRIM), Montreal, Quebec, Canada)

4441 (SLT-P27.7): Wespeaker: A Research and Production oriented Speaker Embedding Learning Toolkit

Hongji Wang (None); Chengdong Liang (Northwestern Polytechnical University); Shuai Wang (Shanghai Jiao Tong University); Binbin Zhang (Horizon Robotics); Zhengyang Chen (Shanghai Jiao Tong University); Xu Xiang (AISpeech Ltd); Slyne Deng (NVIDIA); Yanmin Qian (Shanghai Jiao Tong University)

4623 (SLT-P27.8): Distance-based Weight Transfer for Fine-tuning from Near-field to Far-field Speaker Verification

Li Zhang (Northwestern Polytechnical University); Qing Wang (Northwestern Polytechnical University); Hongji Wang (None); Yue Li (Northwestern Polytechnical University); Wei Rao (Tencent); Yannan Wang (Tencent); Lei Xie (NWPU)

5259 (SLT-P27.9): towards a unified Conformer structure: from ASR to ASV task

Dexin Liao (Xiamen University); Tao Jiang (Xiamen Talentedsoft Co., Ltd.); Feng Wang (Xiamen University); Lin Li (Xiamen University); Qingyang Hong (Xiamen University)

5646 (SLT-P27.10): Pretraining Conformer with ASR for Speaker Verification

Danwei Cai (Duke university); Weiqing Wang (Duke University); Ming Li (Duke Kunshan University); Rui Xia (ByteDance AI Lab); Chuanzeng Huang (Speech, Audio and Music Intelligence (SAMI) group, ByteDance)

5742 (SLT-P27.11): StarGAN-VC based Cross-Domain Data Augmentation for Speaker Verification

Hang-Rui Hu (University of Science and Technology of China); Yan Song (USTC); Jian-Tao Zhang (University of Science and Technology of China); Lirong Dai (University of Science and Technology of China); Ian v McLoughlin (The University of Science and Technology of China); ZHU ZHUO (alibaba); Yu Zhou (alibaba); Yuhong Li (Alibaba); hui xue (Alibaba)

6353 (SLT-P27.12): Improving Transformer-Based Networks with Locality for Automatic Speaker Verification

Mufan Sang (University of Texas at Dallas); Yong Zhao (Microsoft Corporation); GANG Liu (Microsoft); John H Hansen (Univ. of Texas at Dallas); Jian WU (Microsoft Corp)

SLT-P28: Speaker Recognition IV: Verification and Diarization

Room: Dome D

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Ming Li, Alicia Lozano Diez

Thursday, June 8

595 (SLT-P28.1): Jeffreys divergence-based regularization of neural network output distribution applied to speaker recognition

Pierre-Michel Bousquet (Avignon University); Mickael Rouvier (LIA - Avignon University)

1688 (SLT-P28.2): Probabilistic back-ends for online speaker recognition and clustering

Alexey Sholokhov (Huawei Technologies Co., Ltd.); Nikita Kuzmin (NTU); Kong Aik Lee (Institute for Infocomm Research, ASTAR); Eng Siong Chng (Nanyang Technological University)

1815 (SLT-P28.3): Adaptive Large Margin Fine-tuning for Robust Speaker Verification

Leying Zhang (Shanghai Jiao Tong University); Zhengyang Chen (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)

2423 (SLT-P28.4): Parameter-efficient Transfer Learning of Pre-trained Transformer Models for Speaker Verification using Adapters

Junyi Peng (Brno University of Technology); Themis Stafylakis (Omilia - Conversational Intelligence); rongzhi gu (Tencent); Oldrich Pichot (Brno University of Technology); Ladislav Mosner (Brno University of Technology); Lukas Burget (Brno University of Technology); Jan Cernocky (Brno University of Technology)

2687 (SLT-P28.5): Augmenting Transformer-Transducer Based Speaker Change Detection With Token-Level Training Loss

Guanlong Zhao (Google); Quan Wang (Google); Han Lu (Google); Yiling Huang (Google); Ignacio Lopez Moreno (Google)

2982 (SLT-P28.6): Multitask Detection of Speaker Changes, Overlapping Speech and Voice Activity Using wav2vec 2.0

Marie Kunešová (University of West Bohemia); Zbyněk Zajíc (University of West Bohemia)

3179 (SLT-P28.7): Speaker recognition with two-step multi-modal deep cleansing

Ruijie Tao (National University of Singapore); Kong Aik Lee (Institute for Infocomm Research, ASTAR); Zhan Shi (Chinese University of Hong Kong, Shenzhen); Haizhou Li (The Chinese University of Hong Kong, Shenzhen)

3379 (SLT-P28.8): STUDY ON THE FAIRNESS OF SPEAKER VERIFICATION SYSTEMS ACROSS ACCENT AND GENDER GROUPS

Mariel Estevez (CONICET / Universidad de Buenos Aires); Luciana Ferrer (CONICET / Universidad de Buenos Aires)

3825 (SLT-P28.9): Discriminative Speaker Representation via Contrastive Learning with Class-Aware Attention in Angular Space

Zhe Li (Hong Kong Polytechnic University); Man-Wai MAK (The Hong Kong Polytechnic University); Helen Meng (The Chinese University of Hong Kong)

4092 (SLT-P28.10): Self-Supervised Audio-Visual Speaker Representation with Co-Meta Learning

Hui Chen (Tianjin university); Hanyi Zhang (Tianjin university); Longbiao Wang (Tianjin University); Kong Aik Lee (Institute for Infocomm Research, ASTAR); Meng Liu (Tianjin University); Jianwu Dang (Tianjin University)

4364 (SLT-P28.11): Exploring Binary Classification Loss for Speaker Verification

Bing Han (Shanghai Jiao Tong University); Zhengyang Chen (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)

4537 (SLT-P28.12): Target-Speaker Voice Activity Detection via Sequence-to-Sequence Prediction

Ming Cheng (Duke Kunshan University); Weiqing Wang (Duke University); Yucong Zhang (Duke Kunshan University); Xiaoyi Qin (Duke Kunshan University); Ming Li (Duke Kunshan University)

SLT-P29: Speaker Recognition V: Anti-spoofing and Verification

Room: Dome E

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Yanmin Qian, Pierre-Michel Bousquet

303 (SLT-P29.1): Parameter Efficient Transfer Learning for Various Speech Processing Tasks

Shinta Otake (Tokyo Institute of Technology); Rei Kawakami (Tokyo Institute of Technology); Nakamasa Inoue (Tokyo Institute of Technology)

1638 (SLT-P29.2): Spoofed training data for speech spoofing countermeasure can be efficiently created using neural vocoders

Xin Wang (National Institute of Informatics); Junichi Yamagishi (National Institute of Informatics)

2280 (SLT-P29.3): SHIFT TO YOUR DEVICE: DATA AUGMENTATION FOR DEVICE-INDEPENDENT SPEAKER VERIFICATION ANTI-SPOOFING

Junhao Wang (Zhejiang University); Li Lu (Zhejiang University); Zhongjie Ba (Zhejiang University); Feng Lin (Zhejiang University); Kui Ren (Zhejiang University)

Thursday, June 8

4087 (SLT-P29.4): Noise-Disentanglement Metric Learning for Robust Speaker Verification

Yao Sun (Tianjin University); Hanyi zhang (tianjin university); Longbiao Wang (Tianjin University); Kong Aik Lee (Institute for Infocomm Research, ASTAR); Meng Liu (Tianjin University); Jianwu Dang (Tianjin University)

4340 (SLT-P29.5): Phase-Aware Spoof Speech Detection Based on Res2Net with Phase Network

Juntae Kim (SK Telecom); Sung Min Ban (SK Telecom)

4502 (SLT-P29.6): GRAPH-BASED SPECTRO-TEMPORAL DEPENDENCY MODELING FOR ANTI-SPOOFING

Feng Chen (Harbin Institute of Technology); Shiwen Deng (Harbin Normal University); 铁然 郑 (哈尔滨工业大学); 勇军 何 (50+); jiqing Han (Harbin Institute of Technology)

4507 (SLT-P29.7): Leveraging Positional-Related Local-Global Dependency for Synthetic Speech Detection

Xiaohui Liu (Tianjin University, Tianjin, China); Meng Liu (Tianjin University); Longbiao Wang (Tianjin University); Kong Aik Lee (Institute for Infocomm Research, ASTAR); Hanyi Zhang (Tianjin University); Jianwu Dang (Tianjin University)

4651 (SLT-P29.8): UNIVERSAL SPEAKER RECOGNITION ENCODERS FOR DIFFERENT SPEECH SEGMENTS DURATION

Sergey Novoselov (ITMO University); Vladimir Volokhov (STC-innovations Ltd., ITMO University); Galina Lavrentyeva (ITMO University)

4882 (SLT-P29.9): Waveform Boundary Detection for Partially Spoofed Audio

Zexin Cai (Duke University); Weiqing Wang (Duke University); Ming Li (Duke Kunshan University)

5361 (SLT-P29.10): Identifying Source Speakers for Voice Conversion based Spoofing Attacks on Speaker Verification

Danwei Cai (Duke university); Zexin Cai (Duke University); Ming Li (Duke Kunshan University)

5447 (SLT-P29.11): SAMO: Speaker Attractor Multi-Center One-Class Learning for Voice Anti-Spoofing

Siwen Ding (Columbia University); You Zhang (University of Rochester); Zhiyao Duan (University of Rochester)

5693 (SLT-P29.12): Learning From Yourself: A Self-Distillation Method for Fake Speech Detection

Jun Xue (Anhui Province Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University); Cunhang Fan (Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University); Jiangyan Yi (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences); chenglong wang (CASIA); Zhengqi Wen (Qiyuan Laboratory); Dan Zhang (Department of Psychology, Tsinghua University); zhao lv (anhui university)

SPTM-P1: Bayesian Signal Processing

Room: Dome L

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Victor Elvira, Clausas Pau

6813 (SPTM-P1.1): Coherent long-time integration and Bayesian detection with Bernoulli track-before-detect (SPS Journal Paper)*

Murat Uney (University of Liverpool); Paul Horridge (University of Liverpool); Bernie Mulgrew (University of Edinburgh); Simon Maskell (University of Liverpool)

1324 (SPTM-P1.2): A Robust Kalman Filter Based Approach for Indoor Robot Positioning with Multi-Path Contaminated UWB Data

Justin Cano (ISAE-Supaéro); Yi Ding (ISAE-Supaéro); Gaël Pagès (ISAE-Supaéro); Eric Chaumette (ISAE-Supaéro); Jerome Le Ny (Polytechnique Montreal)

4266 (SPTM-P1.3): On Parametric Misspecified Bayesian Cramér–Rao bound: An application to linear/Gaussian systems

Shuo Tang (Northeastern University); Gerald LaMountain (Northeastern University); Tales Imbiriba (Northeastern University); Pau Closas (Northeastern University)

4705 (SPTM-P1.4): Bayesian Cramér–Rao Bound Estimation with Score-Based Models

Evan Scope Crafts (The University of Texas at Austin); Bo Zhao (University of Texas at Austin)

5176 (SPTM-P1.5): Generalized Two-Stage Particle Filter for High Dimensions

Marija Iloska (Stony Brook University); Monica Bugallo (Stony Brook University)

581 (SPTM-P1.6): Distributed Bayesian Tracking on the Special Euclidean Group using Lie Algebra Parametric Approximations

CLAUDIO JOSE BORDIN JUNIOR (Universidade Federal do ABC); CAIO DE FIGUEREDO (INSTITUTO FEDERAL DO CEARA); Marcelo G S Bruno (ITA)

3924 (SPTM-P1.7): Adaptive Gaussian nested filter for parameter estimation and state tracking in dynamical systems

Sara Pérez-Vieites (IMT Nord Europe); Victor Elvira (University of Edinburgh)

Thursday, June 8

4751 (SPTM-P1.8): A Bayesian Perspective on Noise2Noise: Theory and Extensions

Sarah Miller (University of Dayton); Christina M Karam (Huddly); Achour Idoughi (University of Dayton); Kodai Kikuchi (Japan broadcasting corporation); Keigo Hiraoka (University of Dayton)

4859 (SPTM-P1.09): An Augmented Gaussian Sum Filter Through a Mixture Decomposition

Kostas Tsampourakis (University of Edinburgh); Victor Elvira (University of Edinburgh)

6751 (SPTM-P1.10): Bayesian Estimation of Graph Signals (SPS Journal Paper)*

Ariel Kroizer (Ben Gurion University); Tirza S Routtenberg (Ben Gurion University of the Negev); Yonina Eldar (Weizmann Institute of Science)

BISP-L4: Learning from Neuroimaging Data

Room: Nafsika A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Javier Escudero,

10:50 AM

4990 (BISP-L4.1): COUPLED CP TENSOR DECOMPOSITION WITH SHARED AND DISTINCT COMPONENTS FOR MULTI-TASK FMRI DATA FUSION

Ricardo Borsari (CNRS); Isabell Lehmann (University of Paderborn); Mohammad Akhond (UMBC); Vince Calhoun (TReNDS); Konstantin Usevich (CNRS); David BRIE (Université de Lorraine); Tulay Adali (University of Maryland, Baltimore County)

11:05 AM

2213 (BISP-L4.2): BrainNetFormer: Decoding Brain Cognitive States With Spatial-Temporal Cross Attention

Leheng Sheng (Tsinghua University); Wenhan Wang (Southeast University); Zhiyi Shi (Carnegie Mellon University); Jichao Zhan (Southeast University); Youyong Kong (Southeast University)

11:20 AM

2926 (BISP-L4.3): Decoding musical pitch from human brain activity with automatic voxel-wise whole-brain fMRI feature selection

Vincent K.M. Cheung (Sony Computer Science Laboratories, Inc.); Yueh-Po Peng (Institute of Information Science, Academia Sinica); Jing-Hua Lin (Academia Sinica); Li Su (Academia Sinica)

11:35 AM

3124 (BISP-L4.4): Interpretable Nonnegative Incoherent Deep Dictionary Learning for fMRI data analysis

Manuel Morante (AAU); Jan Ostergaard (Aalborg University); Sergios Theodoridis (Aalborg University)

11:50 AM

3219 (BISP-L4.5): Time-Resolved fMRI Shared Response Model Using Gaussian Process Factor Analysis

MohammadReza Ebrahimi (University of Toronto); Navona Calarco (University of Toronto); Colin Hawco (Centre for Addiction and Mental Health); Aristotle Voineskos (CAMH); Ashish Khisti (University of Toronto)

12:05 PM

133 (BISP-L4.6): Tensor-Based Complex-valued Graph Neural Network for Dynamic Coupling Multimodal Brain Networks

Yanwu Yang (HIT at shenzhen); Guoqing Cai (Harbin Institute of Technology, Shenzhen); Chenfei Ye (Harbin Institute of Technology at Shenzhen); Yang Xiang (Peng Cheng Laboratory); Ting Ma (Harbin Institute of Technology, Shenzhen)

GC-9: Lightweight, Multi-Speaker, Multi-Lingual Indic Text-to-Speech

Room: Nefeli B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): TBA

10:50 AM

6628 (GC-L9.1): Introduction

Deekshitha G (IISc); Prasanta Ghosh (); Hema A Murthy (IIT Madras); Philipp Olbrich (Gesellschaft für Internationale Zusammenarbeit (GIZ)); Pranaw Kumar (CDAC Mumbai); Keiichi Tokuda (Department of Computer Science and Engineering, Nagoya Institute of Technology); Mark Hasegawa-Johnson (University of Illinois); Heiga Zen (Google); Sathvik Udapa (Indian Institute of Science); Abhayjeet Singh (IISc); Jesuraj Bandekar (IISc); Sandhya Badiger (IISc)

11:10 AM

6860 (GC-L9.2): Multi-speaker Multi-lingual VQTTTS System for LIMMITS 2023 Challenge

Chenpeng Du (Shanghai Jiao Tong University); Yiwei Guo (Shanghai Jiao Tong University); Feiyu Shen (Shanghai Jiao Tong University); Kai Yu (Shanghai Jiao Tong University)

Thursday, June 8

11:22 AM

6897 (GC-L9.3): VANI: Very-lightweight Accent-controllable TTS for Native and Non-native speakers with Identity Preservation

Rohan Badlani (NVIDIA); Akshit Arora (NVIDIA); Subhankar Ghosh (NVIDIA); Rafael Valle (NVIDIA); Kevin Shih (NVIDIA); João Felipe Santos (NVIDIA); Boris Ginsburg (NVIDIA); Bryan Catanzaro (NVIDIA)

11:34 AM

6903 (GC-L9.4): LeanSpeech: The Microsoft Lightweight Speech Synthesis System for LIMMITS Challenge 2023

Chen Zhang (Microsoft); SHUBHAM BANSAL (Microsoft); Aakash Lakhera (Microsoft); Jinzhu Li (Microsoft); Gag Wang (Microsoft); Sandeep kumar Satpal (Microsoft,India); sheng zhao (microsoft); Lei He (Microsoft Cloud and AI)

11:46 AM

6918 (GC-L9.5): Lightweight Prosody-TTS for multi-lingual multi-speaker scenario

Girdhar Pamisetty (IIT Hyderabad); Chaitanya Varun Sahukari (IIT Hyderabad); Sri Rama Murty Kodukula (IIT Hyderabad)

IVMSP-L6: Quality Assessment and Anomaly Detection

Room: Athina

Type: Oral

10:50 AM to 12:10 PM

Chair(s): Aykut Erdem, Ross Cutler

10:50 AM

890 (IVMSP-L6.1): ST360IQ: No-Reference Omnidirectional Image Quality Assessment with Spherical Vision Transformers

Nafiseh Jabbari Tofighi (Koc University); Mohamed Hedi elfkir (hacettepe university); Nevrez Imamoglu (AIST); Cagri Ozcinar (Samsung); Erkut Erdem (Hacettepe University); Aykut Erdem (Koc University)

11:05 AM

2917 (IVMSP-L6.2): JNDMix: JND-Based Data Augmentation for No-reference Image Quality Assessment

Jiamu Sheng (Fudan University); Jiayuan Fan (Fudan University); peng ye (fudan university); Jianjian Cao (Fudan University)

11:20 AM

665 (IVMSP-L6.3): Continuous Learning for Blind Image Quality Assessment with Contrastive Transformer

Jifan Yang (National Engineering Research Center for Multimedia Software, School of Computer Science, Wuhan University); Zhongyuan Wang (Wuhan University); Baojin Huang (National Engineering Research Center for Multimedia Software, School of Computer Science, Wuhan University); Lianbing Deng (Guangdong-Macau Joint Laboratory for Advanced and Intelligent Computing)

11:35 AM

1553 (IVMSP-L6.4): Learning Hybrid Representations of Semantics and Distortion for Blind Image Quality Assessment

Xiaoqi Wang (Nanjing University of Posts and Telecommunications); Jian Xiong (Nanjing Univeristy of Posts and Telecommunications); Bo Li (Xihua University); Jinli Suo (Tsinghua University); Hao Gao (Nanjing University of Posts and Telecommunications)

11:50 AM

5185 (IVMSP-L6.5): Multi-modal Approach to Food Classification Diet Tracking System with spoken and visual inputs

Shivani Gowda Kallappanahalli (Loyola Marymount University); Yifan Hu (Loyola Marymount University); Mandy B Korpusik (Loyola Marymount University)

12:05 PM

1041 (IVMSP-L6.6): Robust Video Anomaly Detection Framework via Prior Knowledge and Multi-Path Frame Prediction

Menghao Zhang (Beijing University of Posts and Telecommunications); Jingyu Wang (Beijing University of Posts and Telecommunications); Jing Wang (Beijing University of Posts and Telecommunications); Qi Qi (Beijing University of Posts and Telecommunications); Zirui Zhuang (Beijing University of Posts and Telecommunications); Haifeng Sun (Beijing university of posts and telecommunications); Ning Xiao (Didi Chuxing)

MLSP-L9: Deep Learning I

Room: Jupiter

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Robert Jenssen,

10:50 AM

4764 (MLSP-L9.1): FULLY COMPLEX-VALUED DEEP LEARNING MODEL FOR VISUAL PERCEPTION

Aniruddh Sanjoy Sikdar (Indian Institute of Science); Sumanth V Udupa (Indian Institute of Science); Suresh Sundaram (Indian Institute of Science)

Thursday, June 8

11:05 AM

4795 (MLSP-L9.2): InfoShape: Task-Based Neural Data Shaping via Mutual Information

Homa Esfahanizadeh (Massachusetts Institute of Technology); William Wu (MIT); Manya Ghobadi (Massachusetts Institute of Technology (MIT)); Dr.Regina Barzilay (Massachusetts institute of technology); Muriel Medard (MIT)

11:20 AM

4911 (MLSP-L9.3): Improved Projection Learning for Lower Dimensional Feature Maps

Ilan Price (University of Oxford); Jared Tanner (Oxford University)

11:35 AM

1312 (MLSP-L9.4): MULTI-HEAD UNCERTAINTY INFERENCE FOR ADVERSARIAL ATTACK DETECTION

Yuqi Yang (Beijing University of Posts and Telecommunications); Songyun Yang (Beijing University of Posts and Telecommunications); Jiyang Xie (Beijing University of Posts and Telecommunications); Zhongwei Si (Beijing University of Posts and Telecommunications); Kai Guo (BUPT); Ke Zhang (North China Electric Power University); Kongming Liang (Beijing University of Posts and Telecommunications)

11:50 AM

4810 (MLSP-L9.5): Neural networks with quantization constraints.

Ignacio Hounie (University of Pennsylvania); Juan Elenter (University of Pennsylvania); Alejandro Ribeiro (University of Pennsylvania)

12:05 PM

5679 (MLSP-L9.6): Multi-task Bias-Variance Trade-off Through Functional Constraints

Juan Cervino (University of Pennsylvania); Juan Andres Bazerque (University of Pittsburgh); Miguel Calvo-Fullana (Universitat Pompeu Fabra); Alejandro Ribeiro (University of Pennsylvania)

MSP-L1: Human-Centric Multimedia and Human-Machine Interaction

Room: Salon de Roses

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Ron Hecht,

10:50 AM

2015 (MSP-L1.01): NATURALISTIC HEAD MOTION GENERATION FROM SPEECH

Trisha Mittal (University of Maryland); Zakaria Aldeneh (Apple); Masha Fedzechkina (Apple); Anurag Ranjan (Apple); Barry Theobald (Apple)

11:05 AM

3853 (MSP-L1.02): Visual Answer Localization with Cross-modal Mutual Knowledge Transfer

Yixuan Weng (CASIA); Bin Li (Hunan University)

11:20 AM

3186 (MSP-L1.03): CC-POSENET: TOWARDS HUMAN POSE ESTIMATION IN CROWDED CLASSROOMS

Zefang Yu (Shanghai Jiao Tong University); Yanping Hu (Shanghai Jiao Tong University); Suncheng Xiang (Shanghai Jiao Tong University); Ting Liu (Shanghai Jiao Tong University); Yuzhuo Fu (sjtu)

11:35 AM

4260 (MSP-L1.04): GAZE PRE-TRAIN FOR IMPROVING DISPARITY ESTIMATION NETWORKS

Ron M Hecht (General Motors); Ohad Rahamim (General Motors); Shaul Oron (GM); Andrea Forgacs (General Motors); Gershon Celniker (General Motors); Dan Levi (General Motors); Omer Tsimhoni (General Motors)

11:50 AM

5289 (MSP-L1.05): Contextually-rich human affect perception using multimodal scene information

Digbalay Bose (University of Southern California); Rajat Hebbar (University of Southern California); Krishna Somandepalli (University of Southern California); Shrikanth Narayanan (USC)

12:05 PM

6090 (MSP-L1.06): TWO-STREAM JOINT-TRAINING FOR SPEAKER INDEPENDENT ACOUSTIC-TO-ARTICULATORY INVERSION

Jianrong Wang (School of Computer Science and Technology, Tianjin University, Tianjin, China); Jinyu Liu (Tianjin University); Xuewei Li (Tianjin University); Mei Yu (Tianjin University); Jie Gao (Tianjin University); Qiang Fang (Chinese Academy of Social Sciences); Li Liu (Shenzhen Research Institute of Big Data, the chinese university of hong kong shenzhen)

SLT-L17: Speech Emotion Recognition: Transfer Learning

Room: Delphi

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Carlos Busso, Douglas O'Shaughnessy

Thursday, June 8

10:50 AM

457 (SLT-L17.1): A Generalized Subspace Distribution Adaptation Framework for Cross-Corpus Speech Emotion Recognition

Shaokai Li (Yantai University); Peng Song (Yantai University); Liang Ji (Yantai University); Yun Jin (Jiangsu Normal University); Wenming Zheng (Southeast University)

11:05 AM

3755 (SLT-L17.2): Fast Yet Effective Speech Emotion Recognition with Self-Distillation

Zhao Ren (L3S Research Center); Thanh Tam Nguyen (Griffith University); Yi Chang (Imperial College London); Bjoern W. Schuller (Imperial College London)

11:20 AM

3954 (SLT-L17.3): DOMAIN ADAPTATION WITHOUT CATASTROPHIC FORGETTING ON A SMALL-SCALE PARTIALLY-LABELED CORPUS FOR SPEECH EMOTION RECOGNITION

Zhi Zhu (Fairy Devices Inc.); Yoshinao Sato (Fairy Devices Inc.)

11:35 AM

4547 (SLT-L17.4): PHONETIC ANCHOR-BASED TRANSFER LEARNING TO FACILITATE UNSUPERVISED CROSS-LINGUAL SPEECH EMOTION RECOGNITION

Shreya G Upadhyay (National Tsing Hua University); Luz Martinez-Lucas (Department of Electrical and Computer Engineering, University of Texas at Dallas); Bo-Hao Su (Department of Electrical Engineering, National Tsing Hua University); Wei-Cheng Lin (The University of Texas at Dallas); Woan-Shiuan Chien (Department of Electrical Engineering, National Tsing Hua University); Ya-Tse Wu (Department of Electrical Engineering, National Tsing Hua University); William F Katz (UT Dallas); Carlos Busso (University of Texas at Dallas); Chi-Chun Lee (National Tsing Hua University)

11:50 AM

4559 (SLT-L17.5): Zero-Shot Speech Emotion Recognition Using Generative Learning with Reconstructed Prototypes

Xinzhou Xu (Nanjing University of Posts and Telecommunications); Jun Deng (Agile Robots AG); Xizong Zhang (Imperial College London); Zhen Yang (Nanjing University of Posts and Telecommunication); Bjorn W. Schuller (Imperial College London)

12:05 PM

4858 (SLT-L17.6): Unsupervised domain adaptation for preference learning based speech emotion recognition

Abinay Reddy Naini (The University of Texas at Dallas); Mary Kohler (Laboratory for Analytic Sciences, North Carolina State University); Carlos Busso (University of Texas at Dallas)

SPCN-L1: Multi-Antenna Communications and Sensing

Room: Salon de Roses B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Osvaldo Simeone,

10:50 AM

569 (SPCN-L1.01): Noncoherent multiuser Grassmannian Constellations for the MIMO Multiple Access Channel

Javier Álvarez Vizoso (Universidad de Cantabria); Diego Cuevas (Universidad de Cantabria); Carlos Beltrán (Universidad de Cantabria); Ignacio Santamaria (University of Cantabria); Vit Tucek (Huawei Technologies); Gunnar Peters (Huawei Sweden)

11:05 AM

4297 (SPCN-L1.02): Distributed Signal Processing for Out-of-System Interference Suppression in Cell-Free Massive MIMO

Zakir Hussain Shaik (Linköping University); Erik G. Larsson (Nil)

11:20 AM

4374 (SPCN-L1.03): Information and Sensing Beamforming Optimization for Multi-User Multi-Target MIMO ISAC Systems

Minghe Zhu (The Chinese University of Hong Kong, Shenzhen); Lei Li (CUHK-Shenzhen); Shuqiang Xia (ZTE Corporation); Tsung-Hui Chang ("The Chinese University of Hong Kong,")

11:35 AM

5341 (SPCN-L1.04): WHEN IS MIMO MASSIVE IN RADAR?

Jaimin Shah (University of Minnesota, Twin-cities); Martina Cardone (University of Minnesota, Twin Cities); Alex R Dytso (New Jersey Institute of Technology); Cynthia Rush (Columbia University)

11:50 AM

6018 (SPCN-L1.05): Joint Millimeter-Wave AoD and AoA Estimation Using One OFDM Symbol and Frequency-Dependent Beams

Veljko Boljanovic (University of California, Los Angeles); Danijela Cabric (University of California, Los Angeles)

12:05 PM

6327 (SPCN-L1.06): Promoting the Efficiency of WMMSE and FP for Beamforming by Minorization-Maximization

Zepeng Zhang (ShanghaiTech University); Ziping Zhao (ShanghaiTech University); Kaiming Shen (The Chinese University of Hong Kong, Shenzhen)

Thursday, June 8

SS-L11: Quantum Machine Learning Algorithms and Applications on NISQ Devices

Room: Nefeli A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Jun Qi,

10:50 AM

1857 (SS-L11.01): A Quantum Kernel Learning Approach to Low-Resource Spoken Command Recognition

Chao-Han Huck Yang (Georgia Institute of Technology); Bo Li (Google); Yu Zhang (Google); Nanxin Chen (John Hopkins University); Tara Sainath (Google); Sabato M Siniscalchi (Kore University of Enna); Chin-hui Lee (Georgia Institute of Technology)

11:05 AM

2107 (SS-L11.02): PQLM - Multilingual Decentralized Portable Quantum Language Model

Shuyue Stella Li (Johns Hopkins University); Xiangyu Zhang (Johns Hopkins University); Shu Zhou (HKUST); Hongchao Shu (Johns Hopkins University); Ruixing Liang (Johns Hopkins University); Hexin Liu (Nanyang Technological University); Paola Garcia (Johns Hopkins University)

11:20 AM

2691 (SS-L11.03): OPTIMIZING QUANTUM FEDERATED LEARNING BASED ON FEDERATED QUANTUM NATURAL GRADIENT DESCENT

Jun Qi (Georgia Institute of Technology); Zhang XiaoLei (Northwestern Polytechnical University); Javier Tejedor (Institute of Technology, Universidad San Pablo-CEU, CEU Universities)

11:35 AM

3265 (SS-L11.04): Quantum deep recurrent reinforcement learning

Samuel Yen-Chi Chen (Wells Fargo)

11:50 AM

3634 (SS-L11.05): Certified Robustness of Quantum Classifiers against Adversarial Examples through Quantum Noise

Jih-Cing Huang (National Taiwan University); Yu-Lin Tsai (National Yang Ming Chiao Tung University); Chao-Han Huck Yang (Georgia Institute of Technology); Cheng-Fang Su (National Yang Ming Chiao Tung University); Chia-Mu Yu (National Yang Ming Chiao Tung University); Pin-Yu Chen (IBM Research); Sy-Yen Kuo (National Taiwan University)

12:05 PM

5392 (SS-L11.06): Quantum transfer learning using the large-scale unsupervised pre-trained model WavLM-Large for synthetic speech detection

Ruoyu Wang (University of Science and Technology of China); Jun Du (University of Science and Technology of China); Tian Gao (iFlytek Research)

AASP-P7: Medical and Environmental Acoustics; Audio Security

Room: Dome A

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Dimitra Emmanouilidou,

2008 (AASP-P7.1): On the importance of different cough phases for COVID-19 detection

Yi Zhu (Institut national de la recherche scientifique (INRS)); Mahil Shaik (Indian Institute of Technology Kharagpur); Tiago Falk (Institut national de la recherche scientifique (INRS))

1231 (AASP-P7.2): ANALYSIS AND RE-SYNTHESIS OF NATURAL CRICKET SOUNDS ASSESSING THE PERCEPTUAL RELEVANCE OF IDIOSYNCRATIC PARAMETERS

Anibal JS Ferreira (University of Porto - Faculty of Engineering); Marco Oliveira (University of Porto - Faculty of Engineering); João Silva (University of Porto - Faculty of Engineering); Vítor Almeida (University of Porto - Faculty of Engineering)

3527 (AASP-P7.3): COVID-19 Detection from Speech in Noisy Conditions

Shuo Liu (University of Augsburg); Adria Mallol-Ragolta (University of Augsburg); Björn Schuller (University of Augsburg)

6117 (AASP-P7.4): POSITIVE-PAIR REDUNDANCY REDUCTION REGULARISATION FOR SPEECH-BASED ASTHMA DIAGNOSIS PREDICTION

Georgios Rizos (Imperial College London); Rafael Calvo (Imperial College London); Bjoern W. Schuller (Imperial College London)

6346 (AASP-P7.5): Piecewise position encoding in convolutional neural network for cough-based COVID-19 detection

Jiakun Shen (Institute of Acoustics, Chinese Academy of Sciences); XueShuai Zhang (University of Chinese Academy of Sciences); pengyuan zhang (Institute of Acoustics, Chinese Academy of Sciences); Yonghong Yan (Institute of Acoustics, Chinese Academy of Sciences); Shaoxing Zhang (Peking University Third Hospital); Zhihua Huang (Xinjiang University); Yanfen Tang (Beijing Ditan Hospital Capital Medical University); Yu Wang (Beijing Ditan Hospital Capital Medical University); Fujie Zhang (Beijing Ditan Hospital Capital Medical University); Aijun Sun (Dalian Public Health Clinical Center)

Thursday, June 8

6537 (AASP-P7.6): A Contrastive Embedding-based Domain Adaptation method for Lung Sound Recognition in Children Community-Acquired Pneumonia

Dongmin Huang (Southern University of Science and Technology); Lingwei Wang (Shenzhen People's Hospital); Hongzhou Lu (Department of Infectious Diseases, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China); Wenjin Wang (Southern University of Science and Technology)

834 (AASP-P7.7): BTS-E: Audio Deepfake Detection using Breathing-Talking-Silence Encoder

Thien-Phuc Doan (Soongsil university); Long Nguyen-Vu (Soongsil university); Souhwan Jung (Soongsil university); Kihun Hong (Soongsil university)

1468 (AASP-P7.8): TransAudio: Towards the Transferable Adversarial Audio Attack via Learning Contextualized Perturbations

Gege Qi (Alibaba); Yuefeng Chen (Alibaba Group); Yao Zhu (Zhejiang University); Binyuan Hui (Alibaba Group); Xiaodan Li (Alibaba Group); Xiaofeng Mao (Alibaba Group); rong zhang (Alibaba); hui xue (Alibaba)

3908 (AASP-P7.9): GENERAL OR SPECIFIC? INVESTIGATING EFFECTIVE PRIVACY PROTECTION IN FEDERATED LEARNING FOR SPEECH EMOTION RECOGNITION

Chao Tan (Kyoto University); Yang Cao (Hokkaido University); Sheng Li (National Institute of Information & Communications Technology (NICT)); Masatoshi Yoshikawa (Kyoto University)

6685 (AASP-P7.10): The PartialSpoof Database and Countermeasures for the Detection of Short Fake Speech Segments Embedded in an Utterance (SPS Journal Paper)*

Lin Zhang (National Institute of Informatics); Xin Wang (National Institute of Informatics); Erica Cooper (); Nicholas Evans (EURECOM); Junichi Yamagishi (National Institute of Informatics)

4446 (AASP-P7.11): Adversarial Guitar Amplifier Modelling With Unpaired Data

Alec P Wright (Aalto University); Vesa Valimaki (Aalto University); Lauri Juvela (Aalto University)

1035 (AASP-P7.12): End-to-End Amp Modelling: From Data to Controllable Guitar Amplifier Models

Lauri Juvela (Aalto University); Eero-Pekka Damsk agg (Neural DSP); Aleks Peussa (Neural DSP); Jaakko M akinen (Neural DSP); Thomas Sherson (Neural DSP); Stylianos I Mimilakis (Neural DSP); Kimmo Rauhanen (Neural DSP); Athanasios Gotsopoulos (Neural DSP)

AASP-P8: Classification of Acoustic Scenes and Events

Room: Dome B

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Annamaria Mesaros,

1509 (AASP-P8.1): Weight-sharing Supernet for Searching Specialized Acoustic Event Classification Networks Across Device Constraints

Guan-Ting Lin (National Taiwan University); Qingming Tang (Amazon, Alexa); Chieh-Chi Kao (Amazon); Viktor Rozgic (Amazon Alexa); Chao Wang (Amazon)

5071 (AASP-P8.2): SEMANTICAC: SEMANTICS-ASSISTED FRAMEWORK FOR AUDIO CLASSIFICATION

Yicheng Xiao (Tsinghua Shenzhen International Graduate School, Tsinghua University); Yue Ma (Tsinghua University); SHUYAN LI (University of Cambridge); Hantao Zhou (Tsinghua Shenzhen International Graduate School, Tsinghua University); Ran Liao (Tsinghua Shenzhen International Graduate School, Tsinghua University); Xiu Li (Tsinghua University)

6135 (AASP-P8.3): Lightweight Annotation and Class Weight Training for Automatic Estimation of Alarm Audibility in Noise

Fran ois Effa (INRS); romain serizel (Universit  de Lorraine); Jean-Pierre Arz (INRS); Nicolas Grimault (Universit  Lyon 1)

6341 (AASP-P8.4): Effectiveness of Inter- and Intra-Subarray Spatial Features for Acoustic Scene Classification

Takao Kawamura (Tokyo Metropolitan University); Yuma Kinoshita (Tokai University); Nobutaka Ono (Tokyo Metropolitan University); Robin Scheibler (LINE Corporation)

1222 (AASP-P8.5): Simple Pooling Front-ends for Efficient Audio Classification

Xubo Liu (University of Surrey); Haohe Liu (University of Surrey); Qiuqiang Kong (Byte Dance); Xinhao Mei (University of Surrey); Mark D. Plumbley (University of Surrey); Wenwu Wang (University of Surrey)

5112 (AASP-P8.6): Mouth Breathing Detection Using Audio Captured Through Earbuds

Tousif Ahmed (Samsung Research America, Inc.); Md Mahbubur Rahman (Samsung Research America); Ebrahim Nemati (Samsung Research America); Jilong Kuang (Samsung Research America); Jun Alex Gao (Samsung Research America)

5275 (AASP-P8.7): FEDRPO: FEDERATED RELAXED PARETO OPTIMIZATION FOR ACOUSTIC EVENT CLASSIFICATION

Meng Feng (MIT); Chieh-Chi Kao (Amazon); Qingming Tang (Amazon, Alexa); Amit Solomon (Amazon); Viktor Rozgic (Amazon Alexa); Chao Wang (Amazon)

Thursday, June 8

1578 (AASP-P8.8): RANDMASKING AUGMENT: A SIMPLE AND RANDOMIZED DATA AUGMENTATION FOR ACOUSTIC SCENE CLASSIFICATION

JuBum Han (Samsung Research); Mateusz Matuszewski (Samsung R&D Institute Poland); Olaf Sikorski (Samsung R&D Poland); Hosang Sung (Samsung Research); Hoonyoung Cho (Samsung Research)

2866 (AASP-P8.9): Zero-shot Sound Event Classification Using a Sound Attribute Vector with Global and Local Feature Learning

Yi-Han Lin (Kobe University); Xunquan Chen (Kobe University); Ryoichi Takashima (Kobe University); Tetsuya Takiguchi (Kobe University)

3421 (AASP-P8.10): Efficient similarity-based passive filter pruning for compressing CNNs

Arshdeep Singh (University of Surrey); Mark D. Plumbley (University of Surrey)

4106 (AASP-P8.11): AN EXPERIMENTAL STUDY ON SOUND EVENT LOCALIZATION AND DETECTION UNDER REALISTIC TESTING CONDITIONS

Shutong Niu (University of Science and Technology of China); Jun Du (University of Science and Technology of China); Qing Wang (University of Science and Technology of China); Li Chai (University of Science and Technology of China); Huaxin Wu (iFlytek Research); Zhaoxu Nian (University of Science and Technology of China); Lei Sun (University of Science and Technology of China); Yi Fang (iFlytek Research); Jia Pan (University of Science and Technology of China); Chin-Hui Lee (Georgia Institute of Technology)

5721 (AASP-P8.12): COSMOPOLITE SOUND MONITORING (COSMO) : A STUDY OF URBAN SOUND EVENT DETECTION SYSTEMS GENERALIZING TO MULTIPLE CITIES

Florian Angulo (LTCI - Télécom Paris, IP Paris); Slim Essid (Telecom Paristech); Geoffroy Peeters (LTCI - Télécom Paris, IP Paris); Christophe Mietlicki (Bruitparif)

BISP-P1: Learning from EEG Data

Room: Dome F

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Justin Dauwels,

888 (BISP-P1.1): Wavelet2Vec: A Filter Bank Masked Autoencoder for EEG-based Seizure Subtype Classification

Ruimin Peng (Huazhong University of Science and Technology); changming zhao (Huazhong University of Science and Technology); Yifan Xu (Huazhong University of Science and Technology); Jun Jiang (Wuhan Children's Hospital); Guangtao Kuang (Wuhan Children's Hospital); Jianbo Shao (Wuhan Children's Hospital); Dongrui Wu (Huazhong University of Science and Technology)

3910 (BISP-P1.2): Brain network features differentiate intentions from different emotional expressions of the same text

Zhongjie Li (Tianjin University); Bin Zhao (Japan Advanced Institute of Science and Technology); Gaoyan Zhang (Tianjin University); Jianwu Dang (Tianjin University)

2441 (BISP-P1.3): CROSS-SITE GENERALIZATION FOR IMBALANCED EPILEPTIC CLASSIFICATION

Tala Raif Abdallah (Université d'Angers); Nisrine Jrad (Université d'Angers/UCO); Fahed Abdallah (Lebanese University); Anne heurtier (Université d'Angers); Patrick Van Bogaert (CHU)

1342 (BISP-P1.4): FedEEG: Federated EEG Decoding via Inter-subject Structure Matching

Wenlong Hang (Nanjing TECH University); Jiaxing Li (School of Computer Science and Technology, Nanjing Tech University); Shuang Liang (Nanjing University of Posts and Telecommunications); yuan wu (Nanjing Tech University); Baiying Lei (Shenzhen University); Jing Qin (The Hong Kong Polytechnic University); Yu Zhang (Lehigh University, BIOE); Kup-Sze Choi (The Hong Kong Polytechnic University)

6227 (BISP-P1.5): Graph based semantic ensemble of Riemannian Neural Structured Learning for BCI-EEG signal classification

KURUSETTI VINAY GUPTA (IIT KANPUR); Prof Laxmidhar Behera (IIT Kanpur); Tushar Sandhan (Indian Institute of Technology Kanpur)

4560 (BISP-P1.6): DISAMBIGUATION OF COGNITIVE IMPAIRMENT DIAGNOSIS WITH EEG-BASED DUAL-CONTRASTIVE LEARNING

Zhenxi Song (Harbin Institute of Technology (Shenzhen)); Zian Pei (Shenzhen Bay Laboratory); Huixia Ren (Shenzhen People's Hospital); Lin Zhu (Shenzhen People's Hospital); Yi Guo (Shenzhen People's Hospital; Shenzhen Bay Laboratory); Zhiguo Zhang (Harbin Institute of Technology (Shenzhen))

3205 (BISP-P1.7): SS-ADMM: STATIONARY AND SPARSE GRANGER CAUSAL DISCOVERY FOR CORTICO-MUSCULAR COUPLING

Farwa Abbas (Imperial College London); Verity McClelland (King's College London); Zoran Cvetkovic (King's College London); Wei Dai (Imperial College London)

Thursday, June 8

5926 (BISP-P1.8): CROSS-SUBJECT MENTAL FATIGUE DETECTION BASED ON SEPARABLE SPATIO-TEMPORAL FEATURE AGGREGATION

Yalan Ye (*University of Electronic Science and Technology of China*); Yutuo He (*University of Electronic Science and Technology of China*); Wanjing Huang (*University of Electronic Science and Technology of China*); Qiaosen Dong (*Sichuan University*); Chong Wang (*University of Electronic Science and Technology of China*); Guoqing Wang (*University of Electronic Science and Technology of China*)

5501 (BISP-P1.9): Improving EEG-based Emotion Recognition by Fusing Time-frequency And Spatial Representations

Kexin Zhu (*Fudan University*); Xulong Zhang (*Ping An Technology (Shenzhen) Co., Ltd.*); Jianzong Wang (*Ping An Technology (Shenzhen) Co., Ltd.*); Ning Cheng (*Ping An Technology (Shenzhen) Co., Ltd.*); Jing Xiao (*Ping An Insurance (Group) Company of China*)

925 (BISP-P1.10): Subject-specific Adaptation for a Causally-Trained Auditory-Attention Decoding System

Christine Beauchene (*MIT Lincoln Laboratory*); Mike Brandstein (*MIT Lincoln Laboratory*); Stephanie Haro (*Harvard University*); Thomas Quatieri (*Massachusetts Institute of Technology Lincoln Laboratory*); Christopher Smalt (*Massachusetts Institute of Technology Lincoln Laboratory*)

6506 (BISP-P1.11): MOTOR ACTIVITY RECOGNITION USING EEG DATA AND ENSEMBLE OF STACKED BLSTM-LSTM NETWORK AND TRANSFORMER MODEL

Pallavi Kaushik (*Indian Institute of Technology Roorkee*); Ilina Tripathi (*Thapar Institute of Engineering*); Dr. Partha Pratim Roy (*IIT Roorkee*)

6528 (BISP-P1.12): NODE-WISE DOMAIN ADAPTATION BASED ON TRANSFERABLE ATTENTION FOR RECOGNIZING ROAD RAGE VIA EEG

Xueqi Gao (*College of Intelligence and Computing, Tianjin University*); Chao Xu (*College of Intelligence and Computing, Tianjin University*); Yihang Song (*College of Intelligence and Computing, Tianjin University*); Jing Hu (*College of Intelligence and Computing, Tianjin University*); Jian Xiao (*College of Intelligence and Computing, Tianjin University*); Zhaopeng Meng (*College of Intelligence and Computing, Tianjin University*)

BISP-P2: Physiological Signal Processing I

Room: Dome G

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

269 (BISP-P2.1): Parasympathetic-Sympathetic Causal Interactions and Perceived Workload for Varying Difficulty Affective Computing Tasks

Pravallika Lavanuru (*Human Space Flight Centre, Indian Space Research Organization, Bangalore, India.*); Sawon Pratiher (*IIT Kharagpur*); Karuna P Sahoo (*IIT Kharagpur*); Mrinal Acharya (*Dr. B C Roy Multi-speciality Medical Research Centre, Indian Institute of Technology Kharagpur, India.*); Sreejith S (*Human Space Flight Centre, Indian Space Research Organization, Bangalore, India.*); Nirmalya Ghosh (*Indian Institute of Technology Kharagpur*); Amit Patra (*IIT Kharagpur*)

645 (BISP-P2.2): Exploiting Interactivity and Heterogeneity for Sleep Stage Classification via Heterogeneous Graph Neural Network

Ziyu Jia (*Beijing Jiaotong University*); Youfang Lin (*Beijing Jiaotong University*); Yuhao Zhou (*Beijing Jiaotong University*); Xiyang Cai (*University of California, Los Angeles*); Peng Zheng (*Beijing Jiaotong University*); Qiang Li (*RWTH Aachen University*); Jing Wang (*Beijing Jiaotong University*)

1539 (BISP-P2.3): A Mathematical Model for Neuronal Activity and Brain Information Processing Capacity

Yu Zheng (*Michigan State University*); David Zhu (*Michigan State University*); Jian Ren (*Michigan State University*); Taosheng Liu (*Michigan State University*); Karl Friston (*University College London*); Tongtong Li (*Michigan State University*)

4782 (BISP-P2.4): Improving Automatic Sleep Staging via Temporal Smoothness Regularization

Huy Phan (*Amazon Alexa*); Elisabeth Heremans (*KU Leuven*); Oliver Y. Chén (*University of Bristol*); Philipp Koch (*University of Luebeck*); Alfred Mertins (*University of Luebeck*); Maarten De Vos (*KU Leuven*)

2984 (BISP-P2.5): BLOOD OXYGEN SATURATION ESTIMATION FROM FACIAL VIDEO VIA DC AND AC COMPONENTS OF SPATIO-TEMPORAL MAP

Yusuke Akamatsu (*NEC Corporation*); Yoshifumi Onishi (*NEC Corporation*); Hitoshi Imaoka (*NEC Corporation*)

3248 (BISP-P2.6): Structured Errors-in-variables Modelling for Cortico-muscular Coherence Enhancement

Zhenghao Guo (*King's College London*); Verity McClelland (*King's College London*); Wei Dai (*Imperial College London*); Zoran Cvetkovic (*King's College London*)

3619 (BISP-P2.7): A SPATIAL-TEMPORAL ECG EMOTION RECOGNITION MODEL BASED ON DYNAMIC FEATURE FUSION

Shuo Xiao (*China University of Mining and Technology*); Xiaojing Qiu (*China University of Mining and Technology*); Chaogang Tang (*China University of Mining and Technology*); Zhenzhen Huang (*China University of Mining and Technology*)

Thursday, June 8

3757 (BISP-P2.8): Heart Rate Estimation and Performance Analysis using MIMO Radar with Dispersed Antennas
PeiChao Wang (University of Electronic Science and Technology of China); Qian He (University of Electronic Science and Technology of China)

4195 (BISP-P2.9): Constrained non-negative PARAFAC2 for electromyogram separation
MAGBONDE Abilé Serge (GIPSA LAB); QUAINÉ Franck (GIPSA LAB); Bertrand Rivet (Grenoble-INP)

3237 (BISP-P2.10): A PATIENT INVARIANT MODEL TOWARDS THE PREDICTION OF FREEZING OF GAIT
Nasimuddin Ahmed (TCS Research); Shivam Singhal (TCS Research); Aniruddha Sinha (TCS Research); Avik Ghose (TCS)

6444 (BISP-P2.11): Multi-Observation Hidden Semi-Markov Model for Photoplethysmogram Signal Semantic Segmentation
Navid Hasanzadeh (University of Toronto); Shahrokh Valaee (University of Toronto); Hojjat Salehinejad (Mayo Clinic)

2116 (BISP-P2.12): HIERARCHICAL FILTERING WITH ONLINE LEARNED PRIORS FOR ECG DENOISING
Timur Locher (ETH Zurich); Guy Revach (ETH Zürich); Nir Shlezinger (Ben-Gurion University); Ruud J. G. van Sloun (Technical university of Eindhoven); Rik Vullings (Technical university of Eindhoven)

IFS-P3: Adversarial Machine Learning and Information Theoretic Security

Room: Dome H

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

276 (IFS-P3.1): SC-NET: SALIENT POINT AND CURVATURE BASED ADVERSARIAL POINT CLOUD GENERATION NETWORK

Zihao Zhang (The University of Electronic Science and Technology of China); Nan Sang (UESTC); Xupeng Wang (University of Electronic Science and Technology of China); Mumuxin Cai (University of Electronic Science and Technology of China)

300 (IFS-P3.2): NCL: Textual Backdoor Defense Using Noise-augmented Contrastive Learning

Shengfang Zhai (Peking University); Qingni Shen (Peking University); Xiaoyi Chen (Peking University); Weilong Wang (Peking University); Cong Li (Peking University); Yuejian Fang (Peking University); Zhonghai Wu (Peking University)

494 (IFS-P3.3): Enhance Transferability of Adversarial Examples with Model Architecture

Mingyuan Fan (Fuzhou University); Wenzhong Guo (Fuzhou University); Zuobin Ying (Anhui University); Ximeng Liu (Fuzhou University)

1428 (IFS-P3.4): BadRes: Reveal the Backdoors through Residual Connection

Mingrui He (Beihang University); Tianyu Chen (Beihang University); Haoyi Zhou (Beihang University); Shanghang Zhang (Peking University); Jianxin Li (Beihang University)

1285 (IFS-P3.5): Adversarial Network Pruning By Filter Robustness Estimation

Xinlu Zhuang (Wuhan University); Yunjie Ge (Wuhan University); Baolin Zheng (Alibaba Group); Qian Wang (Wuhan University)

2535 (IFS-P3.6): SORA: Scalable Black-box Reachability Analyser on Neural Networks

Peipei Xu (University of Liverpool); Fu Wang (University of Exeter); Wenjie Ruan (University of Exeter); Chi Zhang (University of Exeter); Xiaowei Huang (Liverpool University)

2780 (IFS-P3.7): MULTI-LAYER FEATURE DIVISION TRANSFERABLE ADVERSARIAL ATTACK

Zikang Jin (Nanjing University of Aeronautics and Astronautics); Changchun Yin (Nanjing University of Aeronautics and Astronautics); Piji Li (Nanjing University of Aeronautics and Astronautics); Lu Zhou (Nanjing University of aeronautics and astronautics); Liming Fang (Nanjing University of Aeronautics and Astronautics); Xiangmao Chang (Nanjing University of Aeronautics and Astronautics); Zhe Liu (Nanjing University of Aeronautics and Astronautics)

3221 (IFS-P3.8): Model Fingerprinting with Benign Inputs

Thibault Maho (Inria); Teddy Furon (Inria); Erwan Le Merrer (Inria)

4630 (IFS-P3.9): Reliability Estimation for Synthetic Speech Detection

Daive Salvi (Politecnico di Milano); Paolo Bestagini (Politecnico di Milano); Stefano Tubaro (Politecnico di Milano, Italy)

5189 (IFS-P3.10): A Role Engineering Approach based on Spectral Clustering Analysis for RESTful Permissions in Cloud

Yutang Xia (Peking University); Yang Luo (Peking University); Wu Luo (Peking University); Qingni Shen (Peking University); Yahui Yang (Peking University); Zhonghai Wu (Peking University)

IFS-P4: Watermarking, Data Hiding and Human Factors in Security

Room: Dome E

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

Thursday, June 8

297 (IFS-P4.1): ENHANCING ROBUSTNESS AND IMPERCEPTIBILITY OF BLIND WATERMARKING WITH IMPROVED MESSAGE PROCESSOR

Yufeng Wu (Nanjing University of Information Science and Technology); Baowei Wang (Nanjing University of Information Science and Technology); Changyu Dai (Nanjing University of Information Science and Technology); Yi Yuan (Nanjing University of Information Science and Technology); Bin Li (Nanjing University of Information Science and Technology); Weiqian Zheng (Nanjing University of Information Science and Technology); Hao Wu (Nanjing University of Information Science and Technology)

376 (IFS-P4.2): Measure and Countermeasure of the Capsulation Attack against Backdoor-based Deep Neural Network Watermarks

Fangqi Li (SEIEE, Shanghai Jiao Tong University); shilin wang (SEIEE, Shanghai Jiaotong University); Yun Zhu (Shanghai Jiaotong University)

906 (IFS-P4.3): A study on the invariance in security whatever the dimension of images for the steganalysis by deep-learning

Kévin Planolles (LIRMM (Montpellier)); Marc Chaumont (LIRMM (Montpellier), UNimes); Frédéric Comby (LIRMM)

3540 (IFS-P4.4): Image Adversarial Steganography Based on Joint Distortion

Zexin Fan (University of Science and Technology of China); Kejiang Chen (University of Science and Technology of China); Chuan Qin (University of Science and Technology of China); Kai Zeng (University of Science and Technology of China); Weiming Zhang (University of Science and Technology of China); Nenghai Yu (University of Science and Technology of China)

4324 (IFS-P4.5): ICStega: Image Captioning-based Semantically Controllable Linguistic Steganography

Xilong Wang (University of Science and Technology of China); Yaofei Wang (Hefei University of Technology); Kejiang Chen (University of Science and Technology of China); Jinyang Ding (University of Science and Technology of China); Weiming Zhang (University of Science and Technology of China); Nenghai Yu (University of Science and Technology of China)

4399 (IFS-P4.6): ROBUST WATERMARKING SCHEME IN ENCRYPTED DOMAIN BASED ON INTEGER LIFTING WAVELET TRANSFORM AND COMPRESSED SENSING

Di Xiao (Chongqing University); Qin Tang (Chongqing University); Aozhu Zhao (Chongqing University); Min Li (Chongqing University)

4506 (IFS-P4.7): Mixer: DNN Watermarking using Image Mixup

Kassem Kallas (National Institute for Research in Digital Science and Technology (INRIA)); Teddy Furon (Inria)

4632 (IFS-P4.8): LINK: Linguistic Steganalysis Framework with External Knowledge

Jinshuai Yang (Tsinghua University); zhongliang yang (tsinghua university); Xinrui Ge (Beijing University of Posts and Telecommunications); Jiajun Zou (Tsinghua University); yue gao (tsinghua); Yongfeng Huang (Tsinghua University)

4877 (IFS-P4.9): CSM in Motion Vector Steganalysis: The Effect of Coders on Motion Vectors in H.264 Video Encoding

Verena Lachner (ZITIS); Katharina Schaar (ZITIS); Ralf Zimmermann (ZITIS)

6296 (IFS-P4.10): Improved WordPCFG for Passwords with Maximum Probability Segmentation

Wenting Li (Peking University); Jiahong Yang (Peking University); Haibo Cheng (Peking University); Ping Wang (Peking University); Kaitai Liang (Delft University of Technology)

IVMSP-P25: 3D Point Cloud/Stereo Video

Room: Dome J

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Zhihong Zhang, Patrick Le Callet

5286 (IVMSP-P25.1): 3D Point Cloud Completion based on Multi-Scale Degradation

long jianing (Institute of Software, Chinese Academy of Sciences); Hao He (Institute of Software, Chinese Academy of Sciences); Qingmeng Zhu (Institute of Software, Chinese Academy of Sciences); Zhipeng Yu (Institute of Software, Chinese Academy of Sciences); Qilin Zhang (Institute of Software, Chinese Academy of Sciences); Zhihong Zhang (Kunming University of Science and Technology)

4957 (IVMSP-P25.2): Soft 2D-to-3D Delivery Using Deep Graph Neural Networks for Holographic-Type Communication

Takuya Fujihashi (Osaka University); Toshiaki Koike-Akino (Mitsubishi Electric Research Laboratories); Takashi Watanabe (Osaka University)

2071 (IVMSP-P25.3): Semantic Preserving Learning for Task-oriented Point Cloud Downsampling

Jianyu Xiong (Tsinghua University); Tao Dai (Shenzhen University); Yaohua Zha (Tsinghua University); Xin Wang (Tsinghua University); Shu-Tao Xia (Tsinghua University)

2757 (IVMSP-P25.4): BILATERAL COARSE-TO-FINE NETWORK FOR POINT CLOUD COMPLETION

Tran Thanh Phong Nguyen (University of Wollongong); Son Lam Phung (University of Wollongong, VinAI); Vinod Gopaldasani (University of Wollongong); Jane L Whitelaw (University of Wollongong)

Thursday, June 8

2680 (IVMSP-P25.5): Deep3DSketch: 3D modeling from Free-hand Sketches with View- and Structural-Aware Adversarial Training

Tianrun Chen (Zhejiang University); Chenglong Fu (Huzhou University); Lanyun Zhu (Singapore University of Technology and Design); Mao Papa (Moxin (Huzhou) Technology Co., LTD); Ying Zang (Huzhou University); Jia Zhang (Yangzhou Polytechnic College); Lingyun Sun (Zhejiang University)

2553 (IVMSP-P25.6): RATE-DISTORTION OPTIMIZED VARIABLE-NODE-SIZE TRISOUP FOR POINT CLOUD CODING

Kyohei Unno (KDDI Research); Kohei Matsuzaki (KDDI Research); Satoshi Komorita (KDDI Research, Inc.); Kei Kawamura (KDDI Research)

2467 (IVMSP-P25.7): VOLUMETRIC ATTRIBUTE COMPRESSION FOR 3D POINT CLOUDS USING FEEDFORWARD NETWORK WITH GEOMETRIC ATTENTION

Tam Thuc V.H Do (York University); Philip A Chou (Google); Gene Cheung (York University)

2160 (IVMSP-P25.8): SFR: Semantic-aware Feature Rendering of Point Cloud

Yaohua Zha (Tsinghua University); Rongsheng Li (Tsinghua University); Tao Dai (Shenzhen University); Jianyu Xiong (Tsinghua University); Xin Wang (Tsinghua University); Shu-Tao Xia (Tsinghua University)

1974 (IVMSP-P25.9): N2MVSNet: Non-local Neighbors Aware Multi-View Stereo Network

Zhe Zhang (Peking University); Huachen Gao (Peking University); Yuxi Hu (The Chinese University of Hong Kong, Shenzhen); Ronggang Wang (Peking University)

1675 (IVMSP-P25.10): Deep learning-based stereo camera multi-video synchronization

Nicolas Boizard (University Of Mons); Kevin El Haddad (University of Mons/The Big Projects); Thierry Ravet (UMONS); Francois Cresson (UMONS); Thierry Dutoit (University of Mons)

6698 (IVMSP-P25.12): Coded Illumination for 3D Lensless Imaging (OJSP Paper)*

Yucheng Zheng (University of California, Riverside); M. Salman Asif (University of California, Riverside)

IVMSP-P26: Face Processing

Room: Dome K

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Bo Peng, Zhiyong Wu

2126 (IVMSP-P26.1): Designing a 3D-Aware StyleNeRF Encoder for Face Editing

Songlin Yang (Institute of Automation, Chinese Academy of Sciences); Wei Wang (Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences); Bo Peng (Institute of Automation, Chinese Academy of Sciences); Jing Dong (Chinese Academy of Sciences)

6496 (IVMSP-P26.2): DisCoHead: Audio-and-Video-Driven Talking Head Generation by Disentangled Control of Head Pose and Facial Expressions

Geumbyeol Hwang (DeepBrain AI Inc.); Sunwon Hong (DeepBrain AI Inc.); Seunghyun Lee (DeepBrain AI Inc.); Sungwoo Park (DeepBrain AI Inc.); Gyeongsu Chae (DeepBrain AI Inc.)

6284 (IVMSP-P26.3): MEMORY-AUGMENTED CONTRASTIVE LEARNING FOR TALKING HEAD GENERATION

Jianrong Wang (School of Computer Science and Technology, Tianjin University, Tianjin, China); Yaxin Zhao (Tianjin International Engineering Institute, Tianjin University, Tianjin, China); Hongkai Fan (School of Computer Science and Technology, Tianjin University, Tianjin, China); Tianyi Xu (Tianjin University); Qi Li (School of Electrical and Information Engineering, Tianjin University, Tianjin, China); Sen Li (School of Computer Science and Technology, Tianjin University, Tianjin, China); Li Liu (Shenzhen Research Institute of Big Data, the chinese university of hong kong shenzhen)

4057 (IVMSP-P26.4): WavSyncSwap: End-to-End Portrait-Customized Audio-Driven Talking Face Generation

WeiHong Bao (Tsinghua University); Liyang Chen (Tsinghua University); Chaoyong Zhou (Ping An Technology); Sicheng Yang (Tsinghua University); Zhiyong Wu (Tsinghua University)

2636 (IVMSP-P26.5): Customized Automatic Face Beautification

Wang Chen (Fuzhou University); Peizhen Chen (Fuzhou University); Weijie Chen (Zhejiang University); LuoJun Lin (Fuzhou University)

1279 (IVMSP-P26.6): MODIFY: Model-driven Face Stylization without Style Images

Yuhe Ding (Institute of Automation, Chinese Academy of Sciences); Jian Liang (CASIA); Jie Cao (Institute of Automation, Chinese Academy of Sciences); Aihua Zheng (Anhui University); Ran He (Institute of Automation, Chinese Academy of Sciences)

2824 (IVMSP-P26.7): Structure-Aware Multi-Feature Co-Learning for Dual Branch Face Super Resolution

Kangli Zeng (School of Computer Science, Wuhan University); Zhongyuan Wang (Wuhan University); Tao Lu (Wuhan Institute of Technology); Jianyu Chen (Wuhan University)

Thursday, June 8

2134 (IVMSP-P26.8): Fine-grained Blind Face Inpainting with 3D Face Component Disentanglement

Yu Bai (Fudan University); Ruian He (Fudan University); Weimin Tan (Fudan University); Bo Yan (Fudan University); Yangle Lin (Fudan University)

737 (IVMSP-P26.9): Semantics-Guided Object Removal for Facial Images: with Broad Applicability and Robust Style Preservation

Jookyung Song (Seoul National University); Yeonjin Chang (Seoul National University); SeongUk Park (Seoul National University); Nojun Kwak (Seoul National University)

SAM-P5: MIMO Radars and MIMO Communications

Room: Dome M

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

1762 (SAM-P5.1): LiQuiD-MIMO Radar: Distributed MIMO Radar with Low-Bit Quantization

Yikun Xiang (Nanjing University of Science and Technology); Feng Xi (Nanjing University of Science and Technology); Shengyao Chen (Nanjing University of Science and Technology)

2049 (SAM-P5.2): Resolving Doppler Ambiguity via Spread Phase Alignment in FDA-MIMO Radar

Yanxing Wang (National Laboratory of Radar Signal Processing, Xidian University); Shengqi Zhu (National Laboratory of Radar Signal Processing, Xidian University); Guisheng Liao (National Laboratory of Radar Signal Processing, Xidian University); Lan Lan (National Laboratory of Radar Signal Processing, Xidian University); Zhuochen Chen (National Laboratory of Radar Signal Processing, Xidian University); Feilong Liu (National Laboratory of Radar Signal Processing, Xidian University)

2207 (SAM-P5.3): Gridless Target Localization for FDA-MIMO Radar with Sparse Arrays

Xiaohuan Wu (Nanjing University of Posts and Telecommunications); yaxin liu (Nanjing University of Posts and Telecommunications); Xiaoyuan Jia (Nanjing University of Posts and Telecommunications)

3637 (SAM-P5.4): Binary sequence set optimization for CDMA applications via mixed-integer quadratic programming

Alan Yang (Stanford University); Tara Mina (Stanford University); Grace Gao (Stanford University)

4332 (SAM-P5.5): Multi-User Data Detection in Massive MIMO with 1-Bit ADCs

Amin Radbord (Centre for Wireless Communications (CWC) at University of Oulu); Italo Atzeni (University of Oulu); Antti Tölli (University of Oulu)

4633 (SAM-P5.6): RATE SPLITTING AND PRECODING STRATEGIES FOR MULTI-USER MIMO BROADCAST CHANNELS WITH COMMON AND PRIVATE STREAMS

Liana Khamidullina (Ilmenau University of Technology); André Almeida (Federal University of Ceará); Martin Haardt (Ilmenau University of Technology)

5233 (SAM-P5.7): Multi-User Methods for Vibrational Radar Backscatter Communications

Jessica M Centers (Duke University); Jeffrey Krolik (Duke University)

5410 (SAM-P5.8): Active IRS-Assisted MIMO Channel Estimation and Prediction

Mirza Asif Haider (Temple University); Saidur Pavel (Temple University); Yimin D Zhang (Temple University); Elias Aboutanios (University of New South Wales)

5483 (SAM-P5.9): Waveform design to improve the estimation of target parameters using the Fourier Transform method in a MIMO OFDM DFRC system

Satwika Bhogavalli (Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore); Eric Grivel (Bordeaux INP, IMS laboratory); KVS Hari (Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore); Vincent Corretja (THALES)

6019 (SAM-P5.10): Quantized Precoding and RIS-Assisted Modulation for Integrated Sensing and Communications Systems

R.S. Prasobh Sankar (Indian Institute of Science Bangalore); Sundeep Prabhakar Chepuri (Indian Institute of Science)

6466 (SAM-P5.11): EFFICIENT LARGE-SCALE MULTI-UNIMODULAR WAVEFORM DESIGN WITH GOOD CORRELATION PROPERTIES VIA DIRECT PHASE OPTIMIZATIONS

xiaohan zhao (Beijing Institute of Technology); Yongzhe Li (Beijing Institute of Technology); Ran Tao (Beijing Institute of Technology)

6817 (SAM-P5.12): Joint radar and communications for frequency-hopped MIMO systems (SPS Journal Paper)*

William Baxter (UNSW); Elias Aboutanios (University of New South Wales); Aboulnasr Hassani (Wright State University)

Thursday, June 8

SLT-P30: Speaker Recognition VI: Diarization

Room: Dome C

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Leibny Garcia Perera, Kong Aik Lee

963 (SLT-P30.1): Target Speaker Voice Activity Detection with Transformers and Its Integration with End-to-End Neural Diarization

Dongmei Wang (Microsoft); Xiong Xiao (Microsoft); Naoyuki Kanda (Microsoft); Takuya Yoshioka (Microsoft); Jian Wu (Microsoft)

1290 (SLT-P30.2): In search of strong embedding extractors for speaker diarisation

Jee-weon Jung (Naver Corp.); Heesoo Heo (Naver Corp.); Bong-Jin Lee (Naver Corporation); Jaesung Huh (University of Oxford); Andrew Brown (University of Oxford); Youngki Kwon (Naver Corporation); Shinji Watanabe (Carnegie Mellon University); Joon Son Chung (KAIST)

1883 (SLT-P30.3): Supervised Hierarchical Clustering using Graph Neural Networks for Speaker Diarization

Prachi Singh (Indian Institute of Science, Bangalore); Amrit Kaul (Indian Institute of Science, Bangalore); Sriram Ganapathy (Indian Institute of Science, Bangalore, India, 560012)

2608 (SLT-P30.4): Spectral Clustering-aware Learning of Embeddings for Speaker Diarisation

Evonne Lee (University of Cambridge); Guangzhi Sun (University of Cambridge Department of Engineering); Chao Zhang (Tsinghua University); Phil Woodland (Machine Intelligence Laboratory, Cambridge University Department of Engineering)

3210 (SLT-P30.5): Frame-wise and overlap-robust speaker embeddings for meeting diarization

Tobias Cord-Landwehr (Paderborn University); Christoph B Boeddeker (Paderborn University); Catalin Zorila (Toshiba Cambridge Research Laboratory); Rama S Doddipatla (Toshiba Europe LTD); Reinhold Haeb-Umbach (University of Paderborn)

3840 (SLT-P30.6): Neural Diarization with Non-autoregressive Intermediate Attractors

Yusuke Fujita (LINE Corporation); Tatsuya Komatsu (LINE Corporation); Robin Scheibler (LINE Corporation); Yusuke Kida (LINE Corp); Tetsuji Ogawa (Waseda University)

3953 (SLT-P30.7): TOLD: A Novel Two-stage Overlap-aware Framework for Speaker Diarization

Jiaming Wang (Alibaba Group); Zhihao Du (Speech Lab, Alibaba Group); Shiliang Zhang (Alibaba Group)

4655 (SLT-P30.8): High-resolution embedding extractor for speaker diarisation

Heesoo Heo (Naver Corp.); Youngki Kwon (Naver Corporation); Bong-Jin Lee (Naver Corporation); You Jin Kim (Naver Corporation); Jee-weon Jung (Naver Corp.)

4662 (SLT-P30.9): COMMUNITY DETECTION GRAPH CONVOLUTIONAL NETWORK FOR OVERLAP-AWARE SPEAKER DIARIZATION

Jie Wang (Xiamen University); Zhicong Chen (Xiamen University); Haodong Zhou (Xiamen University); Lin Li (Xiamen University); Qingyang Hong (Xiamen University)

5423 (SLT-P30.10): IMPROVING TRANSFORMER-BASED END-TO-END SPEAKER DIARIZATION BY ASSIGNING AUXILIARY LOSSES TO ATTENTION HEADS

Ye-Rin Jeoung (Hanyang University); Joon-Young Yang (Hanyang University); Jeong-Hwan Choi (Hanyang University); Joon-Hyuk Chang (Hanyang University)

6079 (SLT-P30.11): Absolute decision corrupts absolutely: conservative online speaker diarisation

Youngki Kwon (Naver Corporation); Heesoo Heo (Naver Corp.); Bong-Jin Lee (Naver Corporation); You Jin Kim (Naver Corporation); Jee-weon Jung (Naver Corp.)

6683 (SLT-P30.12): Online Neural Diarization of Unlimited Numbers of Speakers Using Global and Local Attractors (SPS Journal Paper)*

Shota Horiguchi (Hitachi, Ltd.); Shinji Watanabe (Carnegie Mellon University); Paola Garcia (Johns Hopkins University); Yuki Takashima (Hitachi, Ltd.); Yohei Kawaguchi (Hitachi, Ltd.)

SLT-P31: Speech Analysis

Room: Dome D

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Stefan Goetze,

1298 (SLT-P31.1): Raw Ultrasound-based Phonetic Segments Classification Via Mask Modeling

kang you (Shanghai Jiao Tong University); Bo Liu (National University of Defense Technology); Kele Xu (National Key Laboratory of Parallel and Distributed Processing (PDL)); Yunsheng Xiong (National University of Defense Technology); Qisheng Xu (National University of Defense Technology); Ming Feng (Tongji University); Tamás G Csapó (Budapest University of Technology and Economics); Boqing Zhu (National University of Defense Technology)

Thursday, June 8

1981 (SLT-P31.2): Pitch Mark Detection from Noisy Speech Waveform using Wave-U-Net

Hyun-Joon Nam (Pohang University of Science and Technology); Hong-June Park (Pohang University of Science and Technology)

3423 (SLT-P31.3): Leveraging Multiple Sources in Automatic African American English Dialect Detection for Adults and Children

Alexander Johnson (UCLA); Vishwas Shetty (UCLA); Mari Ostendorf (University of Washington); Abeer Alwan (UCLA)

3539 (SLT-P31.4): Does human speech follow Benford's Law?

Leo Hsu (Arizona State University); Visar Berisha (Arizona State University)

3984 (SLT-P31.5): Real-Time MRI Video synthesis from time aligned phonemes with sequence-to-sequence networks

Sathvik Udupa (Indian Institute of Science); Prasanta Dr Ghosh (Indian Institute of Science (IISc), Bangalore)

4959 (SLT-P31.6): Exploring Subgroup Performance in End-to-End Speech Models

Alkis Koudounas (Politecnico di Torino); Eliana Pastor (Politecnico di Torino); Giuseppe Attanasio (Bocconi University); Vittorio Mazzia (Amazon Alexa AI); Manuel Giollo (Amazon); Thomas Gueudre (Amazon Alexa AI); Luca Cagliero (Dipartimento di Automatica e Informatica Politecnico di Torino); Luca de Alfaro (University of California, Santa Cruz); Elena Baralis (Politecnico di Torino); Daniele Amberti (Amazon Alexa AI)

5418 (SLT-P31.7): EFFICIENT STUTTERING EVENT DETECTION USING SIAMESE NETWORKS

Payal Mohapatra (Northwestern University); Bashima Islam (Worcester Polytechnic Institute); MD Tamzeed Islam (Amazon); Ruo Chen Jiao (Northwestern University); Zhu Qi (Northwestern University)

5473 (SLT-P31.8): IMPORTANCE OF DIFFERENT TEMPORAL MODULATIONS OF SPEECH: A TALE OF TWO PERSPECTIVES

Samik Sadhu (Johns Hopkins University); Hynek Hermansky (The Johns Hopkins University, USA)

5943 (SLT-P31.9): REPRESENTATION OF VOCAL TRACT LENGTH TRANSFORMATION BASED ON GROUP THEORY

Atsushi Miyashita (Nagoya University); Tomoki Toda (Nagoya University)

6035 (SLT-P31.10): Moving Towards Non-Binary Gender Identification Via Analysis of System Errors in Binary Gender Classification

Sebastian CG Ellis (University of Sheffield); Stefan Goetze (University of Sheffield); Heidi Christensen (University of Sheffield)

6066 (SLT-P31.11): UNSUPERVISED WORD SEGMENTATION BASED ON WORD INFLUENCE

ruohao yan (Beijing Institute of Technology & xinjiang university); Hua-Ping Zhang (Beijing Institute of Technology); Wushour Slamur (xinjiang university); Askar Hamdulla (Xinjiang University)

1695 (SLT-P31.12): Toward a Multimodal Approach for Disfluency Detection and Categorization

Amrit Romana (University of Michigan); Kazuhito Koishida (Microsoft)

SPTM-P2: Estimation, Detection, and Classification

Room: Dome L

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Koby Todros, Vikram Krishnamurthy

1079 (SPTM-P2.1): ENTROPY BASED FEATURE REGULARIZATION TO IMPROVE TRANSFERABILITY OF DEEP LEARNING MODELS

Raphaël Baena (IMT Atlantique); Lucas Drumetz (IMT Atlantique); Vincent Gripon (IMT Atlantique)

3051 (SPTM-P2.2): A NOVEL APPROACH BASED ON VORONOÏ CELLS TO CLASSIFY SPECTROGRAM ZEROS OF MULTICOMPONENT SIGNALS

Nils Laurent (University Grenoble Alpes); Sylvain Meignen (University Grenoble Alpes); Marcelo A Colominas (CONICET); Juan M Miramont Taurel (Instituto de Investigación y Desarrollo en Bioingeniería y Bioinformática (UNER-CONICET)); Francois Auger (Université de Nantes - Laboratoire IREENA)

3738 (SPTM-P2.3): ACHIEVABLE ERROR EXPONENTS FOR ALMOST FIXED-LENGTH M-ARY HYPOTHESIS TESTING

jun diao (北京航空航天大学); lin zhou (Beihang University); Lin Bai (BUAA)

5773 (SPTM-P2.4): Second-Order Statistic Deviation to Model Anomalies in the Design of Unsupervised Detectors

Andriy Enttsel (University of Bologna); Filippo Martinini (University of Bologna); Alex Marchioni (University of Bologna); Mauro Mangia (University of Bologna); Riccardo Rovatti (University of Bologna); Gianluca Setti (Politecnico di Torino)

1619 (SPTM-P2.5): Estimating Inharmonic Signals with Optimal Transport Priors

Filip Elvander (Aalto University)

Thursday, June 8

1806 (SPTM-P2.6): Adaptive Filtering Algorithms for Set-Valued Observations--Symmetric Measurement Approach to Unlabeled and Anonymized Data

Vikram Krishnamurthy (Cornell University)

2211 (SPTM-P2.7): Robust GMM parameter estimation via the K-BM algorithm

Ori Kenig (Ben Gurion University of The Negev); Koby Todros (Ben-Gurion University of the Negev); Tulay Adali (University of Maryland, Baltimore County)

2592 (SPTM-P2.8): A New Probabilistic Distance Metric with Application in Gaussian Mixture Reduction

Ahmad Sajedi (University of Toronto); Yuri Lawryshyn (University of Toronto); Konstantinos N Plataniotis (UoT)

3330 (SPTM-P2.9): Learned Generative Misspecified Lower Bound

Hai Victor Habi (Tel Aviv University); Hagit Messer (Tel Aviv University); Yoram Bresler (UIUC)

4750 (SPTM-P2.10): Phase Unwrapping in Correlated Noise for FMCW Lidar Depth Estimation

Alfred Ulvog (Mitsubishi Electric Research Laboratories); Joshua Rapp (Mitsubishi Electric Research Laboratories); Toshiaki Koike-Akino (Mitsubishi Electric Research Laboratories); Hassan Mansour (Mitsubishi Electric Research Laboratories (MERL)); Petros Boufounos (Mitsubishi Electric Research Laboratories); Kieran Parsons (Mitsubishi Electric Research Laboratories)

4880 (SPTM-P2.11): Regularized EM algorithm

Pierre HOUDOUIN (CentraleSupélec); Esa Ollila (Aalto University); Frédéric Pascal (CentraleSupélec)

GC-10: L3DAS23: Learning 3D Audio Sources for Audio-Visual Extended Reality

Room: Nefeli B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): TBA

02:00 PM

6642 (GC-L10.1): Introduction

Christian Marinoni (Sapienza University of Rome); Riccardo Fosco Gramaccioni (La Sapienza); Changan Chen (UT Austin); Aurelio Uncini (); Danilo Comminiello (Sapienza University of Rome, Italy)

02:20 PM

6917 (GC-L10.2): Dual-Path Dilated Convolutional Recurrent Network with Group Attention for Multi-Channel Speech Enhancement

Jiaming Cheng (Southeast University); Cong Pang (Southeast University); Ruiyu Liang (Southeast University); Jingjie Fan (Southeast University); Li Zhao (Southeast University)

02:32 PM

6925 (GC-L10.3): STREAM ATTENTION BASED U-NET FOR L3DAS23 CHALLENGE

Honglong Wang (Tianjin University); Yanjie Fu (Tianjin University); Junjie Li (Tianjin University); Meng Ge (Tianjin University); Longbiao Wang (Tianjin University); xinyuan qian (National University of Singapore)

02:44 PM

6926 (GC-L10.4): 3D audio signal processing systems for speech enhancement and sound localization and detection

Jisheng Bai (School of Marine Science and Technology, Northwestern Polytechnical University); Siwei Huang (JLESS); Han Yin (JLESS); Mou Wang (Northwestern Polytechnical University); Yafei Jia (School of Marine Science and Technology, Northwestern Polytechnical University); Jianfeng Chen (School of Marine Science and Technology, Northwestern Polytechnical University)

02:56 PM

6929 (GC-L10.5): THE NERCSLIP-USTC SYSTEM FOR THE L3DAS23 CHALLENGE TASK2: 3D SOUND EVENT LOCALIZATION AND DETECTION (SELD)

Haoyin Yan (University of Science and Technology of China); Haitao Xu (University of Science and Technology of China); Jie Zhang (University of Science and Technology of China); Qing Wang (University of Science and Technology of China)

IVMSP-L7: Model Lightweight and Video Compression

Room: Athina

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Dongsheng Li, Adrian Bors

02:00 PM

413 (IVMSP-L7.1): SR-init: An Interpretable Layer Pruning Method

Hui Tang (Zhejiang University of Technology); Yao Lu (Zhejiang University of Technology); Qi Xuan (Zhejiang University of Technology)

Thursday, June 8

02:15 PM

1437 (IVMSP-L7.2): DUAL META CALIBRATION MIX FOR IMPROVING GENERALIZATION IN META-LEARNING

Ze-Yu Mi (Nanjing university); Yu-Bin Yang (State Key Laboratory for Novel Software Technology, Nanjing University)

02:30 PM

4748 (IVMSP-L7.3): Compressing Cross-Domain Representation via Lifelong Knowledge Distillation

Fei Ye (University of York); Adrian Bors (University of York)

02:45 PM

3318 (IVMSP-L7.4): Multi-rate adaptive transform coding for video compression

Lyndon Duong (New York University); Bohan Li (Google LLC); Cheng Chen (Google Inc.); Jingning Han (Google Inc.)

03:00 PM

4692 (IVMSP-L7.5): REPETITION COUNTING FROM COMPRESSED VIDEOS USING SPARSE RESIDUAL SIMILARITY

Rishabh Khurana (Samsung Research, Bangalore); Jayesh Rajkumar Vachhani (Samsung R&D Institute Bengaluru); Sourabh Vasant Gothe (SAMSUNG R&D INSTITUTE BANGALORE, KARNATAKA, INDIA); Pranay Kashyap (Samsung Research Institute Bangalore)

03:15 PM

3503 (IVMSP-L7.6): Lightweight Fisher Vector Transfer Learning for Video Deduplication

Chris Henry (University of Missouri-Kansas City); Rijun Liao (University of Missouri-Kansas City); Ruiyuan Lin (InnoPeak Technology (Oppo US Research Center)); Zhebin Zhang (OPPO); Hongyu Sun (Oppo); Zhu Li (university of missouri-kansas city)

MLSP-L10: Subspace and Manifold Learning

Room: Jupiter

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Pin-Yu Chen,

02:00 PM

2651 (MLSP-L10.1): Generative Modeling Based Manifold Learning for Adaptive Filtering Guidance

Karim Helwani (Amazon); Paris Smaragdīs (University of Illinois at Urbana-Champaign); Michael M Goodwin (AWS)

02:15 PM

684 (MLSP-L10.2): TENSOR COMPLETION FOR EFFICIENT AND ACCURATE HYPERPARAMETER OPTIMISATION IN LARGE-SCALE STATISTICAL LEARNING

Aaman Rebello (Imperial College London); Kriton Konstantinidis (Imperial College London); Yao Lei Xu (Imperial College London); Danilo P. Mandic ((Imperial College of London, UK))

02:30 PM

903 (MLSP-L10.3): CO-Net: Classification-oriented Point Cloud Sampling via Informative Feature Learning and Non-overlapped Local Adjustment

Yanan Lin (Xiamen University); Keyu Chen (East China Normal University); Shihao Zhou (Xiamen University); Yunan Huang (Xiamen University); Yunqi Lei (Xiamen university)

02:45 PM

2091 (MLSP-L10.4): Deep Survival Analysis and Counterfactual Inference Using Balanced Representations

Muskan Gupta (Tata Consultancy Services - Research); Gokul Kannan (NITT); Ranjitha Prasad (IIIT Delhi); Garima Gupta (TCS Innovation Labs, Delhi)

03:00 PM

3045 (MLSP-L10.5): Feature Space Recovery for Incomplete Multi-view Clustering

Zhen Long (University of Electronic Science and Technology of China); Ce Zhu (University of Electronic Science & Technology of China); Pierre Comon (Univ. Grenoble Alpes); Yipeng Liu (University of Electronic Science and Technology of China)

03:15 PM

4602 (MLSP-L10.6): Study of Manifold Geometry using Multiscale Non-Negative Kernel Graphs

Carlos Hurtado (Universitat Politècnica de Catalunya); Sarath Shekkizhar (University of Southern California); Javier Ruiz-Hidalgo (Universitat Politècnica de Catalunya); Antonio Ortega (University of Southern California)

SLT-L18: Speech Enhancement - Diffusion and Other Generative Models

Room: Delphi

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Timo Gerkmann, Yu Tsao

02:00 PM

2594 (SLT-L18.1): Cross-domain Diffusion based Speech Enhancement for Very Noisy Speech

Heming Wang (The Ohio State University); DeLiang Wang (Ohio State University)

Thursday, June 8

02:15 PM

3643 (SLT-L18.2): SRTNet: Time Domain Speech Enhancement Via Stochastic Refinement

Zhibin Qiu (XinJiang University); Mengfan Fu (XinJiang University); Yinfeng Yu (Department of Computer Science and Technology, State Key Lab on Intelligent Technology and Systems, Tsinghua University, Beijing, China;Xinjiang University); Lili Yin (Xinjiang University); Fuchun Sun (Tsinghua University); Hao Huang (Xinjiang University)

02:30 PM

4671 (SLT-L18.3): DIFFUSION-BASED GENERATIVE SPEECH SOURCE SEPARATION

Robin Scheibler (LINE Corporation); Youna Ji (NAVER Corporation); Soo-Whan Chung (Naver Corporation); Jaeuk Byun (Naver Corporation); Soyeon Choe (NAVER Corporation); Min-Seok Choi (NAVER)

02:45 PM

4716 (SLT-L18.4): SEPDIF: SPEECH SEPARATION BASED ON DENOISING DIFFUSION MODEL

Bo Chen (Huawei Technologies); Chao Wu (Huawei Technologies); Wenbin Zhao (Huawei Technologies)

03:00 PM

5798 (SLT-L18.5): Fast and Efficient Speech Enhancement with Variational Autoencoders

Mostafa Sadeghi (INRIA); romain serizel (Université de Lorraine)

03:15 PM

6105 (SLT-L18.6): Metric-oriented Speech Enhancement using Diffusion Probabilistic Model

Chen Chen (Nanyang Technological University); Yuchen Hu (Nanyang Technological University); Weiwei Weng (Nanyang Technological University); Eng Siong Chng (Nanyang Technological University)

SS-L19: Signal Processing for Smart City Applications and the Internet of Things

Room: Salon de Roses B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Petros Spachos, Konstantinos N. Plataniotis

02:00 PM

1180 (SS-L19.1): MSN-net: Multi-Scale Normality Network for Video Anomaly Detection

Yang Liu (Fudan University); Di Li (Shanghai East-bund Research Institute on NSAI); Wei Zhu (Fudan University); Dingkan Yang (Fudan University); Jing Liu (Fudan University); Liang Song (Fudan University)

02:15 PM

2262 (SS-L19.2): EEG Emotion Recognition via Ensemble Learning Representations

Bilal Taha (University of Toronto); Dae Yon Hwang (University of Toronto); Dimitrios Hatzinakos (University of Toronto)

02:30 PM

3043 (SS-L19.3): Hybrid Indoor Localization via Reinforcement Learning-based Information Fusion

Mohammad Salimibeni (Concordia University); Arash Mohammadi (Concordia University)

02:45 PM

3111 (SS-L19.4): Adapting exploratory behaviour in Active Inference for Autonomous Driving

Sheida Nozari (University of Genoa); Ali Krayani (University of Genoa); Pablo Marín (University Carlos III de Madrid); LUCIO MARCENARO (Università degli Studi di Genoa, Genoa); David Martín (University Carlos III de Madrid); Carlo Regazzoni (Università degli Studi di Genoa, Genoa)

03:00 PM

5446 (SS-L19.5): Federated Semi-Supervised Learning for Object Detection in Autonomous Driving

Fangyuan Chi (The University of British Columbia); Yixiao Wang (University of British Columbia); Panos Nasiopoulos (University of British Columbia); Victor C. M. Leung (Shenzhen University); Mahsa Pourazad (TELUS Communications Inc.)

03:15 PM

6114 (SS-L19.6): SINGLE-SAMPLE DIRECTION-OF-ARRIVAL ESTIMATION FOR FAST AND ROBUST 3D LOCALIZATION WITH REAL MEASUREMENTS FROM A MASSIVE MIMO SYSTEM

Stephan Mazokha (Florida Atlantic University); Sanaz Naderi (Florida Atlantic University); Georgios Orfanidis (Florida Atlantic University); George Sklivanitis (Florida Atlantic University); Dimitris Pados (Florida Atlantic University); Jason Hallstrom (Florida Atlantic University)

Thursday, June 8

SS-L20: Symbol-Level Precoding: Recent Advance and New Applications in 6G and Beyond

Room: Salon de Roses A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Hei Victor Cheng,

02:00 PM

924 (SS-L20.1): OVERLAY COGNITIVE RADIO USING SYMBOL LEVEL PRECODING WITH QUANTIZED CSI

Lu Liu (University Of California, Irvine); Lee Swindlehurst (University of California at Irvine)

02:15 PM

2299 (SS-L20.2): Efficient Quantized Constant Envelope Precoding for Multiuser Downlink Massive MIMO Systems

Zheyu Wu (Academy of Mathematics and Systems Science); Ya-Feng Liu (Chinese Academy of Sciences); Bo Jiang (Nanjing Normal University); Yu-Hong Dai (Academy of Mathematics and Systems Science)

02:30 PM

3177 (SS-L20.3): Joint Symbol-Level Precoding and Sub-Block-Level RIS Design for Dual-Function Radar-Communications

Linlong Wu (University of Luxembourg); Bowen Wang (University of Electronic Science and Technology of China); Ziyang Cheng (University of Electronic Science and Technology of China); Bhavani Shankar Mysore Ramarao (University of Luxembourg); Bjorn Ottersten (SnT)

02:45 PM

3644 (SS-L20.4): SYMBOL LEVEL PRECODING IN THE RF DOMAIN FOR LOW HARDWARE COMPLEXITY RIS-ASSISTED MU-MISO SYSTEMS

Christos Tsinos (University of Athens); Theodoros Tsiftsis (Jinan University); Robert Schober (Friedrich-Alexander University Erlangen-Nurnberg)

03:00 PM

6212 (SS-L20.5): SYMBOL-LEVEL PRECODING IS RELATED TO PARAMETER ESTIMATION FROM QUANTIZED DATA

Mingjie Shao (The Chinese University of Hong Kong, Shandong University); Wing-Kin Ma (The Chinese University of Hong Kong); Yatao Liu (The Chinese University of Hong Kong)

SS-L8: Graphical Inference and Modeling in Dynamical Systems

Room: Nefeli A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Emilie Chouzenoux, Petar Djuric

02:00 PM

649 (SS-L8.1): GraphIT: Iterative reweighted l1 algorithm for sparse graph inference in state-space models

Emilie Chouzenoux (Inria Saclay); Victor Elvira (University of Edinburgh)

02:15 PM

1742 (SS-L8.2): MATRIX RESOLVENT EIGENEMBEDDINGS FOR DYNAMIC GRAPHS

Vasileios Kalantzis (IBM Research); Panagiotis Traganitis (Michigan State University)

02:30 PM

3263 (SS-L8.3): Extended Kalman Filter for Graph Signals in Nonlinear Dynamic Systems

Guy Sagi (Ben Gurion University of the Negev); Nir Shlezinger (Ben-Gurion University); Tirza S Rauttenberg (Ben Gurion University of the Negev)

02:45 PM

4316 (SS-L8.4): Estimating Normalized Graph Laplacians in Financial Markets

José Vinícius de Miranda Cardoso (HKUST); Jiayi Ying (The Hong Kong University of Science and Technology); Sandeep Prof. Kumar (IIT Delhi); Daniel Palomar (The Hong Kong University of Science and Technology)

03:00 PM

4779 (SS-L8.5): Dual-based Online Learning of Dynamic Network Topologies

Seyed Saman Saboksayr (University of Rochester); Gonzalo Mateos (University of Rochester)

03:15 PM

5427 (SS-L8.6): ESTIMATION OF TIME-VARYING GRAPH TOPOLOGIES FROM GRAPH SIGNALS

Yuhao Liu (Stony Brook University); Cui Chen (Stony Brook University); Marzieh Ajirak (Stony Brook University); Petar Djuric ()

AASP-P10: Deep Learning-Based Source Separation II

Room: Dome B

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Emmanuël Habets,

Thursday, June 8

3007 (AASP-P10.1): Ripple Sparse Self-Attention For Monaural Speech Enhancement

Qiquan Zhang (The University of New South Wales); Hongxu Zhu (Department of Electrical and Computer Engineering, National University of Singapore); Qi Song (Alibaba); Xinyua Qian (Department of Electrical and Computer Engineering, National University of Singapore); Zhaoheng Ni (Meta AI); Haizhou Li (The Chinese University of Hong Kong, Shenzhen)

2305 (AASP-P10.2): ICCRN: INPLACE CEPSTRAL CONVOLUTIONAL RECURRENT NEURAL NETWORK FOR MONAURAL SPEECH ENHANCEMENT

Jinjiang Liu (College of Computer Science, Inner Mongolia University); Xueliang zhang (Inner Mongolia University)

6388 (AASP-P10.3): Aiding speech harmonic recovery in DNN-based single channel noise reduction using cepstral excitation manipulation (CEM) components

Yanjue Song (Ghent University - imec); Nilesh Madhu (IDLab, Ghent University - imec)

4152 (AASP-P10.4): Frequency bin-wise single channel speech presence probability estimation using multiple DNNs

Shuai Tao (Aalborg University); Himavanth Reddy Pundla (Aalborg University); Jesper Rindom Jensen (Aalborg University); Mads G. Christensen (Audio Analysis Lab., AD:MT, Aalborg University, Denmark)

5588 (AASP-P10.5): JOINT NOISE REDUCTION AND LISTENING ENHANCEMENT FOR FULL-END SPEECH ENHANCEMENT

Haoyu Li (National Institute of Informatics); Yun Liu (National Institute of Informatics); Junichi Yamagishi (National Institute of Informatics)

5610 (AASP-P10.6): Partially Adaptive Multichannel Joint Reduction of Ego-noise and Environmental Noise

Huajian Fang (Universität Hamburg); Niklas Wittmer (Universität Hamburg); Johannes Twiefel (Universität Hamburg); Stefan Wermter (University of Hamburg); Timo Gerkmann (Universität Hamburg)

5724 (AASP-P10.7): Uncertainty Estimation in Deep Speech Enhancement Using Complex Gaussian Mixture Models

Huajian Fang (Universität Hamburg); Timo Gerkmann (Universität Hamburg)

6687 (AASP-P10.8): Online Phase Reconstruction via DNN-Based Phase Differences Estimation (SPS Journal Paper)*

Yoshiki Masuyama (Tokyo Metropolitan University); Kohei Yatabe (Tokyo University of Agriculture and Technology); Kento Nagatomo (Waseda University); Yasuhiro Oikawa (Waseda University)

6709 (AASP-P10.9): Factorized MVDR Deep Beamforming for Multi-Channel Speech Enhancement (SPS Journal Paper)*

Hansol Kim (GIST); Kyeognmuk Kang (GIST); Jong Won Shin (Gwangju Institute of Science and Technology)

6723 (AASP-P10.10): Insights Into Deep Non-Linear Filters for Improved Multi-Channel Speech Enhancement (SPS Journal Paper)*

Kristina Tesch (Universität Hamburg); Timo Gerkmann (Universität Hamburg)

6785 (AASP-P10.11): Boosted Locality Sensitive Hashing: Discriminative, Efficient, and Scalable Binary Codes for Source Separation (SPS Journal Paper)*

Sunwoo Kim (Indiana University); Minje Kim (Indiana University)

6745 (AASP-P10.12): MixCycle: Unsupervised Speech Separation via Cyclic Mixture Permutation Invariant Training (SPS Journal Paper)*

Serap Kirbız (MEF Üniversitesi); Ertuğ Karamatlı (Boğaziçi University)

AASP-P9: Deep Learning-Based Source Separation I

Room: Dome A

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Lukas Drude,

1501 (AASP-P9.1): Multi-resolution Location-based training for multi-channel continuous speech separation

Hassan Taherian (The Ohio State University); DeLiang Wang (Ohio State University)

5089 (AASP-P9.2): TF-GridNet: Making Time-Frequency Domain Models Great Again for Monaural Speaker Separation

Zhong-Qiu Wang (Carnegie Mellon University); Samuele Cornell (Università Politecnica delle Marche); Shukjae Choi (Hyundai Motor Company); Youngjo Lee (42dot); Byeong-Yeol Kim (42dot); Shinji Watanabe (Carnegie Mellon University)

5736 (AASP-P9.3): SELF-REMIXING: UNSUPERVISED SPEECH SEPARATION VIA SEPARATION AND REMIXING

Kohei Saijo (Waseda University); Tetsuji Ogawa (Waseda University)

2881 (AASP-P9.4): DEEPSPACE: DYNAMIC SPATIAL AND SOURCE CUE BASED SOURCE SEPARATION FOR DIALOG ENHANCEMENT

Aaron S Master (Dolby Laboratories, Inc); Lie Lu (Dolby Laboratories); Jonas Samuelsson (Dolby Laboratories, Inc); Heidi-Maria Lehtonen (Dolby Sweden AB); Scott Norcross (Dolby Laboratories, Inc); Nathan Swedlow (Dolby Laboratories, Inc); Audrey Howard (Dolby Laboratories, Inc)

Thursday, June 8

3426 (AASP-P9.5): Blind source counting and separation with relative harmonic coefficients

Huiyuan Sun (The Australian National University); Prasanga Samarasinghe (Australian National University); thushara abhayapala (The Australian National University)

408 (AASP-P9.6): Multi-Dimensional and Multi-Scale Modeling for Speech Separation Optimized by Discriminative Learning

Zhaoxi Mu (Xi'an Jiaotong University); Xinyu Yang (Xi'an Jiaotong University); WenJing Zhu (DXM)

1367 (AASP-P9.7): UX-Net: Filter-and-Process-based Improved U-Net for Real-time Time-domain Audio Separation

Kashyap Patel (The University of Texas at Dallas); Anton Kovalyov (Electrical and Computer Engineering, University of Texas at Dallas, Richardson, TX, USA); Issa Panahi (UTD)

2572 (AASP-P9.8): Hyperbolic Audio Source Separation

Darius Petermann (Indiana University - Bloomington); Gordon Wichern (Mitsubishi Electric Research Laboratories (MERL)); Aswin Shanmugam Subramanian (Mitsubishi Electric Research Laboratories (MERL)); Jonathan LeRoux (Mitsubishi Electric Research Laboratories (MERL))

5457 (AASP-P9.9): Reverberation as supervision for speech separation

Rohith Aralikatti (University of Maryland at College Park); Christoph B Boeddeker (Paderborn University); Gordon Wichern (Mitsubishi Electric Research Laboratories (MERL)); Aswin Shanmugam Subramanian (Mitsubishi Electric Research Laboratories (MERL)); Jonathan LeRoux (Mitsubishi Electric Research Laboratories (MERL))

5400 (AASP-P9.10): JACAPPELLA CORPUS: A JAPANESE A CAPPELLA VOCAL ENSEMBLE CORPUS

Tomohiko Nakamura (The University of Tokyo); Shinnosuke Takamichi (The University of Tokyo); Naoko Tanji (The University of Tokyo); Satoru Fukayama (National Institute of Advanced Industrial Science and Technology (AIST)); Hiroshi Saruwatari (The University of Tokyo)

1958 (AASP-P9.11): Better Together: Dialogue Separation and Voice Activity Detection for Audio Personalization in TV

Matteo Torcoli (International Audio Laboratories Erlangen); Emanuel Habets (AudioLabs Erlangen)

6735 (AASP-P9.12): Sampling-Frequency-Independent Convolutional Layer and its Application to Audio Source Separation (SPS Journal Paper)*

Koichi Saito (The University of Tokyo); Tomohiko Nakamura (The University of Tokyo); Kohei Yatabe (Tokyo University of Agriculture and Technology); Hiroshi Saruwatari (The University of Tokyo)

BISP-P3: Medical Image Segmentation

Room: Dome F

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

1147 (BISP-P3.1): MLP-GAN for Brain Vessel Image Segmentation

Bin Xie (Illinois Institute of Technology); Hao Tang (ETH Zurich); Bin Duan (Illinois Institute of Technology); Dawen Cai (University of Michigan); Yan Yan (Illinois Institute of Technology)

1164 (BISP-P3.2): SCSGNet: Spatial-Correlated and Shape-Guided Network for Breast Mass Segmentation

Qingqiu Li (Fudan University); Jilan Xu (Fudan University); Runtian Yuan (Fudan University); Yuejie Zhang (Fudan University); Rui Feng (Fudan University)

1865 (BISP-P3.3): Domain Generalized Fundus Image Segmentation via Dual-Level Mixing

Xin Luo (College of Computer, National University of Defense Technology); Wei Chen (College of Computer, National University of Defense Technology); Chen Li (National University of Defense Technology); Bin Zhou (National University of Defense Technology); yusong tan (College of Computer, National University of Defense Technology)

2297 (BISP-P3.4): Multi-stage Aggregation Transformer for Medical Image Segmentation

Xiaoyan Wang (Zhejiang University of Technology); Minghan Shao (Zhejiang University of Technology); Dongyan Guo (Zhejiang University of Technology); Ying Cui (Zhejiang University of Technology); Xiaojie Huang (Zhejiang University); Ming Xia (Zhejiang University of Technology); Cong Bai (Zhejiang University of Technology)

2854 (BISP-P3.5): LSSD: A robust segmentation network for inflamed appendix from CT images

Wing W.Y. Ng (South China University of Technology); Peixin Zheng (South China University of Technology); Ting Wang (South China University of Technology); Jianjun Zhang (South China University of Technology); Hui Zhou (The Sixth Affiliated Hospital of Guangzhou Medical University, Qingyuan People's Hospital); GuangMing Li (The Sixth Affiliated Hospital of Guangzhou Medical University, Qingyuan People's Hospital); Dan Liang (Guangzhou First People's Hospital/The Second Affiliated Hospital, South China University of Technology); Yin hao Liang (South China University of Technology); Xinhua Wei (Department of Radiology, Guangzhou First People's Hospital, South China University of Technology)

Thursday, June 8

2964 (BISP-P3.6): LightVessel: Exploring Lightweight Coronary Artery Vessel Segmentation via Similarity Knowledge Distillation

Hao Dang (Henan University of Chinese Medicine); Yuekai Zhang (Beijing University of Posts and Telecommunications); Xingqun Qi (University of Technology Sydney); Wanting Zhou (Beijing University of Posts and Telecommunications); Muyi Sun (CRIPAC, Institute of Automation, Chinese Academy of Sciences)

3763 (BISP-P3.7): Automatic segmentation of nasopharyngeal carcinoma in CT images using dual attention and edge detection

Qizhi Wang (Xiangtan University); Wei Huang (The First Hospital of ChangSha); Yuan Zhang (Xiangtan University); Xuanya Li (Baidu); Xiongjun Ye (Chinese Academy of Medical Sciences and Peking Union Medical College); Kai Hu (Xiangtan University)

4009 (BISP-P3.8): Pseudo Multi-Source Domain Extension and Selective Pseudo-labeling for Unsupervised Domain Adaptive Medical Image Segmentation

Xiaokang Liu (Xiangtan University); Zhiqiang Wang (Xiangnan University); Kai Hu (Xiangtan University); Xieping Gao (Hunan Normal University)

4357 (BISP-P3.9): U-Shifformer: brain tumor segmentation using a shifted attention mechanism

Chih-Wei Lin (Fujian Agriculture and Forestry University); Zhongsheng Chen (Fujian Agriculture and Forestry University)

4393 (BISP-P3.10): MPS-AMS: Masked Patches Selection and Adaptive Masking Strategy Based Self-Supervised Medical Image Segmentation

Xiangtao Wang (Hebei University of Technology); Ruizhi Wang (Hebei University of Technology); Tian Biao (Hebei University of Technology); Jiaojiao Zhang (Hebei University of Technology); Shuo Zhang (Hebei University of Technology); Junyang Chen (Shenzhen University); Thomas Lukasiewicz (University of Oxford); Zhenghua Xu (Hebei University of Technology)

4430 (BISP-P3.11): Exploiting Multi-Decision and Deep Refinement for Ultrasound Image Segmentation

Wenjing Liu (Xiangtan University); Xuanya Li (Baidu); Kai Hu (Xiangtan University); Xieping Gao (Hunan Normal University)

4647 (BISP-P3.12): Optimizing Vision Transformers for Medical Image Segmentation

qianying liu (university of glasgow); Chaitanya Kaul (University of Glasgow); Jun Wang (University of Warwick); Christos Anagnostopoulos (University of Glasgow); Roderick Murray-Smith (University of Glasgow); Fani Deligianni (University of Glasgow)

BISP-P4: Bioinformatics

Room: Dome G

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

139 (BISP-P4.1): A new Semi-supervised classification method using a supervised autoencoder for biomedical applications
Cyprien Gille (UMONS); Frederic Guyard (Orange Labs); Michel Barlaud (University of Nice)

415 (BISP-P4.2): END-TO-END CLASSIFICATION OF CELL-CYCLE STAGES WITH CENTER-CELL FOCUS TRACKER USING RECURRENT NEURAL NETWORKS

Abin Jose (RWTH); Rijo Roy (RWTH Aachen); Dennis Eschweiler (RWTH Aachen University); Ina Laube (Lehrstuhl für Bildverarbeitung, RWTH Aachen); reza azad (rwth); Daniel Moreno-Andreas (RWTH Aachen University); Johannes Stegmaier (RWTH Aachen University)

5228 (BISP-P4.3): Rethinking Learning-based Method for Lossless Genome Compression

Han Yang (Alibaba Group); Fei Gu (Alibaba Group); Jieping Ye (Alibaba Group)

5582 (BISP-P4.4): Adversarial Attacks on Genotype Sequences

Daniel Mas Montserrat (Stanford University); Alexander Ioannidis (Stanford University)

197 (BISP-P4.5): DIGITAL PHENOTYPE REPRESENTATION BY STATISTICAL, INFORMATION THEORY, DATA-DRIVEN APPROACH WITH DIGITAL HEALTH DATA

Binh Nguyen (TMU); Michael Nigro (Toronto Metropolitan University); Alice Rueda (Ryerson University); Venkat Bhat (University of Toronto); Sri Krishnan (Ryerson University)

3419 (BISP-P4.6): EFFICIENT PROTEIN STRUCTURAL CLASS PREDICTION VIA CHAOS GAME REPRESENTATION AND RECURRENT NEURAL NETWORKS

Michaela Areti Zervou (University of Crete, ICS-FORTH); Effrosyni Doutsis (Foundation for Research and Technology - Hellas (FORTH)); Panagiotis Tsakalides (University of Crete, Foundation for Research and Technology - Hellas (FORTH))

1270 (BISP-P4.7): Generative De Novo Protein Design with Global Context

Cheng Tan (Zhejiang University & Westlake University); Zhangyang Gao (westlake university); Jun Xia (Westlake University); Bozhen Hu (Zhejiang University & Westlake University); Stan Z. Li (Westlake University)

601 (BISP-P4.8): Efficient implementation of robust CUSUM algorithm to characterize nanogaps measurements with heavy-tailed noise

Javier Kipen (KTH); Joakim Jalden (KTH); Shyamprasad Raja (KTH); Saumey Jain (KTH Royal Institute of Technology)

Thursday, June 8

5360 (BISP-P4.9): Representation Learning of Clinical Multivariate Time Series with Random Filter Banks

Alireza Keshavarzian (University of Toronto); Hojjat Salehinejad (Mayo Clinic); Shahrokh Valaee (University of Toronto)

6504 (BISP-P4.10): MvCo-DoT: Multi-View Contrastive Domain Transfer Network for Medical Report Generation

Ruizhi Wang (Hebei University of Technology); Xiangtao Wang (Hebei University of Technology); Zhenghua Xu (Hebei University of Technology); Wenting Xu (Hebei University of Technology); Junyang Chen (Shenzhen University); Thomas Lukasiewicz (University of Oxford)

5691 (BISP-P4.11): Attention-Guided Deep Learning Framework for Movement Quality Assessment

Aditya S Kanade (Indian Institute of Technology Madras); Mansi Sharma (Department of Computer Science and Engineering, Amrita School of Computing, Coimbatore, Amrita Vishwa Vidyapeetham, India and Department of Electrical Engineering, IIT Madras); M Manivannan ("Indian Institute of Technology Madras, India")

4279 (BISP-P4.12): Learning from single-expert annotated labels for automatic sleep staging

Zhiheng Luan (School of Cyber Science and Engineering, Wuhan University); Yanzhen Ren (Computer School of Wuhan University); Li Peng (Wuhan University); Xiong Chen (Sleep Medicine Centre, Zhongnan Hospital of Wuhan University); Xiuping Yang (Sleep Medicine Centre, Zhongnan Hospital of Wuhan University); Weiping Tu (Wuhan University); Yuhong Yang (Wuhan University)

IFS-P5: Cybersecurity, Hardware and Network Security

Room: Dome H

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

2322 (IFS-P5.1): A Graph Neural Network Multi-task Learning-Based Approach for Detection and Localization of Cyberattacks in Smart Grids

Abdulrahman Takiddin (Texas A&M University); Rachad Atat (Texas A&M University at Qatar); Muhammad Ismail (Tennessee Tech University); Katherine Davis (Texas A&M University); Erchin Serpedin ()

2761 (IFS-P5.2): Light Projection-based Physical-world Vanishing Attack against Car Detection

Huixiang Wen (Donghua University); Shan Chang (Donghua University); Luo Zhou (Donghua University)

3940 (IFS-P5.3): AN EMPIRICAL STUDY OF BACKDOOR ATTACKS ON MASKED AUTOENCODERS

Shuli Zhuang (University of Science and Technology of China); Pengfei Xia (University of Science and Technology of China); Bin Li (University of Science and Technology of China)

4473 (IFS-P5.4): RUMOR DETECTION VIA ASSESSING THE SPREADING PROPENSITY OF USERS

Peng Zheng (National University of Defense Technology); Zhen Huang (National Laboratory for Parallel and Distributed Processing, National University of Defense Technology, Changsha, Hunan); Yong Dou (National University of Defense Technology); Yeqing Yan (National University of Defense Technology)

5769 (IFS-P5.5): Detecting Malicious Migration on Edge to Prevent Running Data Leakage

Yuchen Wong (软件工程中心); Qingni Shen (Peking University); Cong Li (Peking University); Cunzhan Liu (Peking University); Tianxiang Ai (Peking University)

6121 (IFS-P5.6): A LARGE-SCALE PRETRAINED DEEP MODEL FOR PHISHING URL DETECTION

Yanbin Wang (Zhejiang university); wei fan zhu (zhejiang university); Haitao Xu (Zhejiang University); Zhan Qin (Zhejiang University); Kui Ren (Zhejiang University); Wenrui Ma (Zhejiang Gongshang University)

6761 (IFS-P5.7): Detecting and Interpreting Changes in Scanning Behavior in Large Network Telescopes (SPS Journal Paper)*

Michael Kallitsis (Merit/UM); Rupesh Prajapati (Penn State University); Dinghao Wu (Penn State University); Vasant Honavar (The Pennsylvania State University); John Yen (Penn State University)

6716 (IFS-P5.8): Synthetic speech detection based on local autoregression and variance statistics (SPS Journal Paper)*

Sanshuai Cui (Sun Yat-sen University); Bingyuan Huang (Sun Yat-Sen University); Jiwu Huang (Shenzhen University); Xiangui Kang (Sun Yat-Sen University)

223 (IFS-P5.9): QTrojan: A Circuit Backdoor Against Quantum Neural Networks

Cheng Chu (Indiana University Bloomington); Lei Jiang (Indiana University); Martin Swamy (Indiana University); Fan Chen (Indiana University Bloomington)

874 (IFS-P5.10): A Multi-modal Approach for Context-aware Network Traffic Classification

Bo Pang (哈尔滨工业大学); Yongquan Fu (National University of Defense Technology); Siyuan Ren (Department of Computer Science and Technology, Harbin Institute of Technology(Shenzhen)); Siqi Shen (Xiamen University); Ye Wang (National University of Defense Technology); Qing Liao (Harbin Institute of Technology (Shenzhen)); Yan Jia (National University of Defense Technology)

Thursday, June 8

4121 (IFS-P5.11): Efficient Privacy Preserving Graph Neural Network for Node Classification

Xinjun Pei (Central South University); Xiaoheng Deng (Central South University); Shengwei Tian (Xinjiang University); Kaiping Xue (University of Science and Technology of China)

5936 (IFS-P5.12): LEARNING SPARSE ALIGNMENTS VIA OPTIMAL TRANSPORT FOR CROSS-DOMAIN FAKE NEWS DETECTION

Wei Tang (Beijing University of Posts and Telecommunications); zuyao ma (Beijing University of Posts and Telecommunications)

MSP-P1: Multimedia Compression and Quality

Room: Dome L

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

265 (MSP-P1.1): Surface-Sampling based Objective Quality Assessment Metrics for Meshes

Chunyang Fu (SECE, Shenzhen Graduate School, Peking University); Xiang Zhang (Tencent America); Thuong Nguyen Canh (Tencent America); Xiaozhong Xu (Tencent America); Ge Li (SECE, Shenzhen Graduate School, Peking University); Shan Liu (Tencent America)

345 (MSP-P1.2): Perceptual Quality Assessment for Digital Human Heads

Zicheng Zhang (Shanghai Jiaotong University); Yingjie Zhou (Shanghai Jiao Tong University); Wei Sun (Shanghai Jiao Tong University); Xiongkuo Min (Shanghai Jiao Tong University); Yuzhe Wu (Donghua University); Guangtao Zhai (Shanghai Jiao Tong University)

1051 (MSP-P1.3): NF-PCAC: Normalizing Flow based Point Cloud Attribute Compression

Rodrigo Borba Pinheiro (InterDigital); Jean-Eudes Marvie (InterDigital); Giuseppe Valenzise (CNRS); Frederic Dufaux (CNRS)

2460 (MSP-P1.4): Your Camera Improves Your Point Cloud Compression

Lin Yuhuan (Tsinghua University); Tongda Xu (Tsinghua University); ziyu zhu (Tsinghua University); Yanghao Li (Tsinghua University); Zhe Wang (Tsinghua University); Yan Wang (Tsinghua University)

3117 (MSP-P1.5): CNN Filter for Super-Resolution with RPR functionality in VVC

Shimin Huang (Xidian University); Cheolkon Jung (Xidian University); Yang Liu (OPPO Mobile); Ming Li (OPPO)

3337 (MSP-P1.6): SPARSE CONVOLUTION BASED OCTREE FEATURE PROPAGATION FOR LIDAR POINT CLOUD COMPRESSION

Muhammad Asad Lodhi (InterDigital); Jiahao Pang (InterDigital); Dong Tian (InterDigital)

3997 (MSP-P1.7): EFFICIENT SUPER-RESOLUTION FOR COMPRESSION OF GAMING VIDEOS

Yifan Wang (Xidian University); Luka Murn (British Broadcasting Corporation); Luis Herranz (Computer Vision Center); Fei Yang (Universitat Autònoma de Barcelona); Marta Mrak (Queen Mary University of London); Wei Zhang (Xidian University); Shuai Wan (Northwestern Polytechnical University); Marc Gorris Blanch (BBC)

4241 (MSP-P1.8): Estimating Uncertainty on Video Quality Metrics

Pierre David (Capacités); Patrick Le Callet ("Université de Nantes, France"); Suiyi Ling (University of Nantes); Haixiong Wang (Meta); Ioannis Katsavounidis (Facebook); Zafar Shahid (Facebook); Cosmin Stejerean (Meta)

5005 (MSP-P1.9): JPEG Pleno Call for Proposals responses quality assessment

João P.C. Prazeres (Universidade da Beira Interior); Zhe Luo (University of Technology Sydney); Antonio Pinheiro (U.B.I. & I.T.); Luis A da Silva Cruz (Dep. Electrical and Computer Engineering - Univ. of Coimbra); Stuart Pery (University of Technology Sydney)

3489 (MSP-P1.10): SEMANTIC PREPROCESSOR FOR IMAGE COMPRESSION FOR MACHINES

Mingyi Yang (State Key Laboratory of ISN, Xidian University, Xi'an, China); Luis Herranz (Computer Vision Center); Fei Yang (Universitat Autònoma de Barcelona); Luka Murn (British Broadcasting Corporation); Marc Gorris Blanch (BBC); Shuai Wan (Northwestern Polytechnical University); FuZheng Yang (Xidian University); Marta Mrak (Queen Mary University of London)

5893 (MSP-P1.11): BIRD-PCC: Bi-directional Range Image-based Deep LiDAR Point Cloud Compression

Chia-Sheng Liu (National Taiwan University); Jia-Fong Yeh (National Taiwan University); Hao Hsu (National Taiwan University); Hung-Ting Su (National Taiwan University); Ming-Sui Lee (National Taiwan University); Winston H. Hsu (National Taiwan University)

6294 (MSP-P1.12): Deep probabilistic model for lossless scalable point cloud attribute compression

Dat Thanh Nguyen (University of Erlangen-Nuremberg); Kamal Nambiar (Friedrich-Alexander-Universität Erlangen-Nürnberg); Andre Kaup (Friedrich-Alexander-Universität Erlangen-Nürnberg)

Thursday, June 8

MSP-P2: Multimedia Analysis, Synthesis, and Learning

Room: Dome J

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

332 (MSP-P2.1): TOWARDS EXPLAINABLE RECOMMENDATION VIA BERT-GUIDED EXPLANATION GENERATOR

Huijing Zhan (I2R, Astar); LING LI (Nanyang Technological University); Shaohua Li (IHPC, ASTAR); Weide Liu (Institute for Infocomm Research); Manas Gupta (Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (ASTAR), Singapore); Alex Kot (Nanyang Technological University)

481 (MSP-P2.2): VarietySound: Timbre-Controllable Video to Sound Generation via Unsupervised Information Disentanglement

Chenyue Cui (Zhejiang University); Zhou Zhao (Zhejiang University); Yi Ren (Bytedance); Jinglin Liu (Zhejiang University); Rongjie Huang (Zhejiang University); chen feiyang (huawei); Zhefeng Wang (Huawei Cloud); Baoxing Huai (Huawei Cloud); Fei Wu (Zhejiang University, China)

1700 (MSP-P2.3): Detection of Real-time DeepFakes in Video Conferencing with Active Probing and Corneal Reflection

Hui Guo (University at Buffalo, SUNY); Xin Wang (University at Buffalo, SUNY); Siwei Lyu (University at Buffalo)

2017 (MSP-P2.4): ON THE ROLE OF LIP ARTICULATION IN VISUAL SPEECH PERCEPTION

Zakaria Aldeneh (Apple); Masha Fedzechkina (Apple); Skyler Seto (Apple); Katherine Metcalf (Apple, Inc.); Miguel Sarabia (Apple); Nicholas Apostoloff (Apple Inc.); Barry Theobald (Apple)

3133 (MSP-P2.5): Audio-driven Talking Head Video Generation with Diffusion Model

Yizhe Zhu (Shanghai Jiao Tong University); Chunhui Zhang (Shanghai Jiaotong University, CloudWalk Technology Co., Ltd); Qiong Liu (CloudWalk); Xi Zhou (CloudWalk Technology)

3476 (MSP-P2.6): On the Role of Visual Context in Enriching Music Representations

Kleanthis Avramidis (University of Southern California); Shanti Stewart (University of Southern California); Shrikanth Narayanan (USC)

3854 (MSP-P2.7): Visual Graph Reasoning Network

Dingbang Li (ECNU); Xin Lin (ECNU); Haibin Cai (East China Normal University); Wenzhou Chen (Zhejiang University)

4268 (MSP-P2.8): A novel efficient multi-view traffic-related object detection framework

Kun Yang (Fudan University); Jing Liu (Fudan University); Dingkang Yang (Fudan University); Hanqi Wang (Fudan University); Peng Sun (Duke Kunshan University); Liang Song (Fudan University)

5749 (MSP-P2.9): BAT: Bi-Alignment Based on Transformation in Multi-Target Domain Adaptation for Semantic Segmentation

Xian Zhong (Wuhan University of Technology); Wei Li (Wuhan University of Technology); Liang Liao (Nanyang Technological University); Jing Xiao (Wuhan University); Wenxuan Liu (Wuhan University of Technology); Wenxin Huang (Hubei University); Zheng Wang (Wuhan University)

6037 (MSP-P2.10): Improving Few-Shot Learning for Talking Face System with TTS Data Augmentation

Qi Chen (Shanghai Jiao Tong University); Ziyang Ma (Shanghai Jiao Tong University); Tao Liu (Shanghai Jiao Tong University); Xu Tan (Microsoft Research Asia); Qu Lu (Shanghai Media Tech); Kai Yu (Shanghai Jiao Tong University); Xie Chen (Shanghai Jiaotong University)

6122 (MSP-P2.11): Code-Switching Speech Synthesis Based on Self-Supervised Learning and Domain Adaptive Speaker Encoder

YiXing Lin (National Central University); Cheng-Hsun Pai (National Central University); Le Phuong (National Central University); Bima Prihasto (National Central University); CHIEN-LIN HUANG (NCKU); Jia-Ching Wang (National Central University)

SAM-P6: DoA Estimation and Beamforming

Room: Dome M

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

327 (SAM-P6.1): ENHANCED COPRIME ARRAY CONFIGURATION FOR DOA ESTIMATION OF NON-CIRCULAR SIGNALS

Nabil Mohsen (University of Science and Technology of China (USTC)); Ammar Hawbani (University of Science and Technology of China); Xing-Fu Wang (USTC); Benjamin Bairrington (USTC); Liang Zhao (Shenyang Aerospace University); Saeed Alsamhi (Software Research Institute, AthlonInstitute of Technology)

981 (SAM-P6.2): Optimal Mixed-ADC arrangement for DOA estimation via CRB under ULA

Xinnan Zhang (University of Science and Technology of China); Yuanbo Cheng (University of Science and Technology of China); Xiaolei Shang (University of Science and Technology of China); Jun Liu (University of Science and Technology of China)

Thursday, June 8

1361 (SAM-P6.3): ERROR ANALYSIS OF CONVOLUTIONAL BEAMSPACE ALGORITHMS

Po-Chih Chen (California Institute of Technology); Dr.P P Vaidyanathan (California Institute of Technology)

2229 (SAM-P6.4): Exploiting Sparse Recovery Algorithms for Semi-Supervised Training of Deep Neural Networks for Direction-of-Arrival Estimation

Murtiza Ali (Indian Institute of Technology, Jammu); Aditya Arie Nugraha (RIKEN); Karan Nathwani (Indian Institute of Technology, Jammu)

2604 (SAM-P6.5): DIFFERENCE COARRAYS OF RATIONAL ARRAYS

Pranav D Kulkarni (California Institute of Technology); Dr.P P Vaidyanathan (California Institute of Technology)

3324 (SAM-P6.6): Equivalence of aperture reduction in element space and constrained combination of DFT beams in beamspace

Damir Rakhimov (TU-Ilmenau); Martin Haardt (Ilmenau University of Technology)

3824 (SAM-P6.7): Graph Signal Processing for Narrowband Direction of Arrival Estimation

Disheng Li (University of Sheffield); Wei Liu (University of Sheffield); Yuriy Zakharov (University of York); Paul Mitchell (University of York)

4638 (SAM-P6.8): Robust Adaptive Beamforming with Proximal Method

Ruifu Li (UCLA); Danijela Cabric (University of California, Los Angeles)

4688 (SAM-P6.9): Neural Maximum-A-Posteriori Beamforming For Ultrasound Imaging

Ben Luijten (Eindhoven University of Technology); Boudewine Ossenkopppele (Delft University of Technology); Nico de Jong (Delft University of Technology); Martin Verweij (Delft University of Technology); Yonina Eldar (); Massimo Mischi (Eindhoven University of Technology); Ruud J. G. van Sloun (Technical university of Eindhoven)

4947 (SAM-P6.10): SUPER DILATED NESTED ARRAYS WITH IDEAL CRITICAL WEIGHTS AND INCREASED DEGREES OF FREEDOM

Ahmed Mohammed Shaalan (University of Science and Technology of China); Jun Du (University of Science and Technology of China)

6802 (SAM-P6.11): Difference-Frequency MUSIC for DOAs (SPS Journal Paper)*

YONGSUNG PARK (Scripps Institution of Oceanography, University of California San Diego); Peter Gerstoft (UCSD)

SLT-P33: Speech Emotion Recognition: General Topics II

Room: Dome C

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Laurence Devillers, Theodora Chaspari

540 (SLT-P33.1): Mingling or Misalignment? Temporal Shift for Speech Emotion Recognition with Pre-trained Representations

Siyuan Shen (East China Normal University); Feng Liu (East China Normal University); Aimin Zhou (East China Normal University)

563 (SLT-P33.2): Emotion Recognition in Conversation from Variable-Length Context

Mian Zhang (Soochow University); Xiabing Zhou (soochow university); Wenliang Chen (Soochow University); Min Zhang (Soochow University)

1423 (SLT-P33.3): Knowledge-Aware Graph Convolutional Network with Utterance-Specific Window Search for Emotion Recognition in Conversations

Xiaotong Zhang (School of Software, Dalian University of Technology); Peng He (School of Software, Dalian University of Technology); Han Liu (Dalian University of Technology); Zhengxi Yin (Huawei Technologies Co. Ltd); Xinyue Liu (School of Software, Dalian University of Technology); Xianchao Zhang (School of Software, Dalian University of Technology)

1611 (SLT-P33.4): Masking speech contents by random splicing: Is emotional expression preserved?

Felix Burkhardt (audEERING GmbH); Anna Derington (audEERING GmbH); Matthias Kahlau (audEERING GmbH); Klaus Scherer (University of Geneva); Florian Eyben (audEERING); Bjoern Schuller (audEERING)

3129 (SLT-P33.5): Multi-Local Attention for Speech-based Depression Detection

Fuxiang Tao (University of Glasgow); Xuri Ge (University of Glasgow); Wei Ma (University of Glasgow); Anna Esposito (Università di Napoli (Italy)); Alessandro Vinciarelli (UNiversity of Glasgow)

3130 (SLT-P33.6): DAILY MENTAL HEALTH MONITORING FROM SPEECH: A REAL-WORLD JAPANESE DATASET AND MULTITASK LEARNING ANALYSIS

Meishu Song (University of Augsburg); Andreas Triantafyllopoulos (University of Augsburg); Zijiang Yang (University of Augsburg); Hiroki Takeuchi (University of Tokyo); Toru Nakamura (Osaka University); Akifumi Kishi (University of Tokyo); Tetsuro Ishizawa (University of Tokyo); Kazuhiro Yoshiuchi (University of Tokyo); Xin Jing (Universität Augsburg); Zhonghao Zhao (Beijing Institute of Technology); Vincent Karas (University of Augsburg); Kun Qian (Beijing Institute of Technology); Bin Hu (Beijing Institute of Technology); Bjorn W. Schuller (Imperial College London); Yamamoto Yoshiharu (University of Tokyo)

Thursday, June 8

3830 (SLT-P33.7): SDTN: SPEAKER DYNAMICS TRACKING NETWORK FOR EMOTION RECOGNITION IN CONVERSATION
Jiawei Chen (South China Agricultural University); Peijie Huang (South China Agricultural University); Guotai Huang (South China Agricultural University); Qianer Li (South China Agricultural University); Yuhong Xu (South China Agricultural University)

4065 (SLT-P33.8): Temporal Modeling Matters: A Novel Temporal Emotional Modeling Approach for Speech Emotion Recognition
JiaXin Ye (Fudan University); Xin-Cheng Wen (Harbin Institute of Technology (Shenzhen)); Yujie Wei (Fudan University); Yong Xu (Fujian University of Technology); KunHong Liu (Xiamen University); Hongming Shan (Fudan University)

5683 (SLT-P33.9): Designing and Evaluating Speech Emotion Recognition Systems: A reality check case study with IEMOCAP
Nikolaos Antoniou (National Technical University of Athens); Athanasios Katsamanis ("ATHENA R.C., Behavioral Signal Technologies"); Theodoros Giannakopoulos (NCSR Demokritos); Shrikanth Narayanan (University of Southern California)

5711 (SLT-P33.10): EMix: A Data Augmentation Method for Speech Emotion Recognition
An Dang (National Central University); Toan H Vu (National Central University); Nguyen Dinh Le (National Central University); Jia-Ching Wang (National Central University)

6131 (SLT-P33.11): A Hierarchical Regression Chain Framework for Affective Vocal Burst Recognition
Jinchao Li (The Chinese University of Hong Kong); Xixin Wu (The Chinese University of Hong Kong); Kaitao Song (Microsoft Research Asia); Dongsheng Li (Microsoft Research Asia); Xunying Liu (The Chinese University of Hong Kong); Helen Meng (The Chinese University of Hong Kong)

6316 (SLT-P33.12): Automatic classification of vocal intensity category from speech
Manila Kodali (Aalto University); Sudarsana Reddy Kadiri (Aalto University); laura laaksonen (Huawei); Paavo Alku (Aalto University)

SLT-P34: Speech Emotion Recognition: Multimodality

Room: Dome D

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Engin Erzin, Prasanta Ghosh

108 (SLT-P34.1): Exploring complementary features in multi-modal speech emotion recognition
Suzhen Wang (Netease Fuxi AI Lab); Yifeng Ma (Tsinghua University); Yu Ding (Netease Fuxi AI Lab)

930 (SLT-P34.2): Learning Cross-modal Audiovisual Representations with Ladder Networks for Emotion Recognition
Lucas Goncalves (The University of Texas at Dallas); Carlos Busso (University of Texas at Dallas)

2151 (SLT-P34.3): PRE-TRAINED MODEL REPRESENTATIONS AND THEIR ROBUSTNESS AGAINST NOISE FOR SPEECH EMOTION ANALYSIS
Vikramjit Mitra (Apple Inc.); Vasudha Kowtha (Apple); Hsiang-Yun Sherry Chien (Apple); Erdrin Azemi (Apple); Carlos Avendano (Apple)

2963 (SLT-P34.4): CROSS-MODAL FUSION TECHNIQUES FOR UTTERANCE-LEVEL EMOTION RECOGNITION FROM TEXT AND SPEECH
JiACHEN LUO (Queen Mary University of London); Huy Phan (Amazon Alexa); Joshua D. Reiss (Queen Mary University of London)

2990 (SLT-P34.5): Using Auxiliary Tasks In Multimodal Fusion Of Wav2vec 2.0 And BERT For Multimodal Emotion Recognition
Dekai Sun (Harbin Institute of Technology); yancheng He (Harbin Institute of Technology); jiqing Han (Harbin Institute of Technology)

3072 (SLT-P34.6): Robust multi-modal speech emotion recognition with ASR error adaptation
Binghui Lin (MIG, Tencent Science and Technology Ltd.); Liyuan wang (Tencent Technology Co., Ltd)

3859 (SLT-P34.7): MULTILEVEL TRANSFORMER FOR MULTIMODAL EMOTION RECOGNITION
Junyi He (360 DigiTech); Meimei Wu (360DigiTech); Meng Li (360 DigitalTech); Xiaobo Zhu (360DigiTech); Feng Ye (360DigiTech, Inc.)

4354 (SLT-P34.8): Role of Lexical Boundary Information in Chunk-Level Segmentation for Speech Emotion Recognition
Wei-Cheng Lin (The University of Texas at Dallas); Carlos Busso (University of Texas at Dallas)

4412 (SLT-P34.9): Knowledge-aware Bayesian Co-attention for Multimodal Emotion Recognition
Zihan Zhao (Shanghai Jiao Tong University); Yu Wang (Shanghai Jiao Tong University); Yan-Feng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University)

4629 (SLT-P34.10): MULTIMODAL EMOTION RECOGNITION BASED ON DEEP TEMPORAL FEATURES USING CROSS-MODAL TRANSFORMER AND SELF-ATTENTION
Bubai Maji (Indian Institute of Technology Kharagpur); Monorama Swain (Silicon Institute of Technology, Bhubaneswar); Rajlakshmi Guha (IIT Kharagpur); Aurobinda Routray (IIT Kharagpur)

Thursday, June 8

5065 (SLT-P34.11): Exploring Attention Mechanisms for Multimodal Emotion Recognition in an Emergency Call Center Corpus

Theo Deschamps-Berger (Paris-Saclay University, CNRS); Lori Lamel (CNRS LIMSI); Laurence Y. Devillers (LISN-CNRS)

5966 (SLT-P34.12): AN EMPIRICAL STUDY AND IMPROVEMENT FOR SPEECH EMOTION RECOGNITION

Zhen Wu (Nanjing University); Yizhe Lu (Nanjing University); Xin-yu Dai (Nanjing University)

SLT-P35: Speech Emotion Recognition: Neural Architectures

Room: Dome E

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Bjorn Schuller,

111 (SLT-P35.1): DST: DEFORMABLE SPEECH TRANSFORMER FOR EMOTION RECOGNITION

Weidong Chen (South China University of Technology); Xiaofen Xing (South China University of Technology); Xiangmin Xu (South China University of Technology); Jianxin Pang (Ubtech Robotics Corp.); Lan Du (iFLYTEK Research)

421 (SLT-P35.2): SPEECH EMOTION RECOGNITION BASED ON LOW-LEVEL AUTO-EXTRACTED TIME-FREQUENCY FEATURES

Ke Liu (Northwest University); Jingzhao Hu (Northwest University); Jun Feng (Northwest University)

1195 (SLT-P35.3): MULTIPLE ACOUSTIC FEATURES SPEECH EMOTION RECOGNITION USING CROSS-ATTENTION TRANSFORMER

Yurun He (The University of Tokyo); Nobuaki Minematsu (The University of Tokyo); Daisuke Saito (The University of Tokyo)

1858 (SLT-P35.4): SPEECH EMOTION RECOGNITION VIA TWO-STREAM POOLING ATTENTION WITH DISCRIMINATIVE CHANNEL WEIGHTING

Ke Liu (Northwest University); Dekui Wang (Northwest University); Dongya Wu (Northwest University); Jun Feng (Northwest University)

2573 (SLT-P35.5): LEARNING ROBUST SELF-ATTENTION FEATURES FOR SPEECH EMOTION RECOGNITION WITH LABEL-ADAPTIVE MIXUP

Lei Kang (Shantou University); Lichao Zhang (Air Force Engineering University); Dazhi Jiang (Shantou University)

2712 (SLT-P35.6): HIERARCHICAL NETWORK WITH DECOUPLED KNOWLEDGE DISTILLATION FOR SPEECH EMOTION RECOGNITION

Ziping Zhao (Tianjin Normal University); Huan Wang (Tianjin Normal University); Haishuai Wang (Zhejiang University); Prof. Dr. Bjoern Schuller (Imperial College London)

3398 (SLT-P35.7): Adapting a self-supervised speech representation for noisy speech emotion recognition by using contrastive teacher-student learning

Seong-Gyun Leem (University of Texas at Dallas); Daniel Fulford (Boston University); JP Onnela (T.H. Chan School of Public Health Harvard University); David Gard (San Francisco State University); Carlos Busso (University of Texas at Dallas)

3963 (SLT-P35.8): SPEECH-BASED EMOTION RECOGNITION WITH SELF-SUPERVISED MODELS USING ATTENTIVE CHANNEL-WISE CORRELATIONS AND LABEL SMOOTHING

Sofoklis Kakouros (University of Helsinki); Themis Stafylakis (Omilia - Conversational Intelligence); Ladislav Mošner (Brno University of Technology); Lukáš Burget (Brno University of Technology)

4135 (SLT-P35.9): Knowledge Transfer For On-Device Speech Emotion Recognition with Neural Structured Learning

Yi Chang (Imperial College London); Zhao Ren (L3S Research Center); Thanh Tam Nguyen (Griffith University); Kun Qian (Beijing Institute of Technology); Bjoern W. Schuller (Imperial College London)

4968 (SLT-P35.10): EXPLORING WAV2VEC 2.0 FINE TUNING FOR IMPROVED SPEECH EMOTION RECOGNITION

Li-Wei Chen (Carnegie Mellon University); Alexander I. Rudnicky (Carnegie Mellon University)

5113 (SLT-P35.11): Deep Implicit Distribution Alignment Networks for Cross-Corpus Speech Emotion Recognition

Yan Zhao (Southeast University); Jincen Wang (Southeast University); Yuan Zong (Southeast University); Wenming Zheng (Southeast University); Hailun Iian (Southeast University); Li Zhao (Southeast University)

6064 (SLT-P35.12): TOWARDS LEARNING EMOTION INFORMATION FROM SHORT SEGMENTS OF SPEECH

Tilak Purohit (Idiap Research Institute); Sarthak Yadav (Aalborg University); Bogdan Vlasenko (Idiap Research Institute); S. Pavankumar Dubagunta (Uniphore Software Systems); Mathew Magimai.-Doss (Idiap Research Institute)

SPTM-P3: Optimization methods for signal processing

Room: Dome K

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Geert Leus, Emilie Chouzenoux

Thursday, June 8

1466 (SPTM-P3.1): COMBINING DUAL-TREE WAVELET ANALYSIS AND PROXIMAL OPTIMIZATION FOR ANISOTROPIC SCALEFREE TEXTURE SEGMENTATION

Leo Davy (ENS Lyon); Nelly Pustelnik (); Patrice Abry (CNRS, Physics Department, Ecole Normale Supérieure de Lyon)

1511 (SPTM-P3.2): Fast convolution algorithm for Real valued finite length sequences

Weiwei Wang (FSU); Victor DeBrunner (FSU); Linda DeBrunner (FSU)

1660 (SPTM-P3.3): Multilevel FISTA for Image Restoration

Guillaume Lauga (Inria/ENS Lyon); Elisa Riccietti (ENS Lyon); Nelly Pustelnik (); Paulo Goncalves (ENS de Lyon)

1869 (SPTM-P3.4): Distributed Online Learning with Adversarial Participants in An Adversarial Environment

XingRong Dong (Sun Yat-Sen University); Zhaoxian Wu (Sun Yat-Sen University); Qing Ling (Sun Yat-Sen University); Zhi Tian (George Mason University)

1940 (SPTM-P3.5): Blind Polynomial Regression

Alberto Natali (Delft University of Technology); Geert Leus (TU Delft)

1956 (SPTM-P3.6): Byzantine-Robust and Communication-Efficient Personalized Federated Learning

Xuechao He (Sun Yat-sen University); Jiaojiao Zhang (The Chinese University of Hong Kong); Qing Ling (Sun Yat-Sen University)

2401 (SPTM-P3.7): Optimized Dithering for Quantization Index Modulation

Shanxiang Lyu ()

2492 (SPTM-P3.8): Meta-DAG: Meta Causal Discovery via Bilevel Optimization

Songtao Lu (IBM Thomas J. Watson Research Center); Tian Gao (IBM Research)

3839 (SPTM-P3.9): Adaptive Simulated Annealing through Alternating Rényi Divergence Minimization

Thomas Guilmeau (Université Paris-Saclay, CentraleSupélec, Inria, CVN); Emilie Chouzenoux (Inria Saclay); Victor Elvira (University of Edinburgh)

4964 (SPTM-P3.10): AN IMPLICIT GRADIENT METHOD FOR CONSTRAINED BILEVEL PROBLEMS USING BARRIER APPROXIMATION

Ioannis Tsaknakis (University of Minnesota); Prashant Khanduri (Wayne State University); Mingyi Hong (University of Minnesota)

5077 (SPTM-P3.11): Unique Bispectrum Inversion for Signals with Finite Spectral/Temporal Support

Samuel Pinilla (STFC); Kumar Vijay Mishra (United States DEVCOM Army Research Laboratory); Brian M Sadler (Army Research Laboratory, USA)

GC-11: 5TH DNS CHALLENGE AT IEEE ICASSP 2023

Room: Nefeli B

Type: Oral

03:40 PM to 05:10 PM

Chair(s): TBA

03:40 PM

6654 (GC-L11.01): Introduction

Harishchandra Dubey (Microsoft Corporation); Ross Cutler (Microsoft Corporation); Vishak Gopal (Microsoft Corporation); Sebastian Braun (Microsoft)

04:00 PM

6850 (GC-L11.02): Personalized speech enhancement combining band-split RNN and speaker attentive module

Xiaohuai Le (Nanjing University ; ByteDance); Li Chen (ByteDance); Yiqing Guo (ByteDance); Chao He (ByteDance); Cheng Chen (ByteDance); Xianjun Xia (NA); Jing Lu (Nanjing University)

04:12 PM

6904 (GC-L11.03): The NPU-Elevoc Personalized Speech Enhancement System for ICASSP2023 DNS Challenge

Xiaopeng Yan (Northwestern Polytechnical University); Yindi Yang (Elevoc); Zhihao Guo (Elevoc); Liangliang Peng (Elevoc); Lei Xie (NWPU)

04:24 PM

6909 (GC-L11.04): TEA-PSE 3.0: TENCENT-ETHEREAL-AUDIO-LAB PERSONALIZED SPEECH ENHANCEMENT SYSTEM FOR ICASSP 2023 DNS-CHALLENGE

Yukai Jv (Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, School of Computer Science, Northwestern Polytechnical University); Jun Chen (Tencent); Shimin Zhang (Northwestern Polytechnical University); Shulin He (College of Computer Science, Inner Mongolia University); Wei Rao (Tencent); weixin zhu (tencent); Yannan Wang (Tencent); Tao Yu (Tencent); Shi-dong Shang (tencent)

Thursday, June 8

04:36 PM

6910 (GC-L11.05): Tspeech-AI System Description to the 5th Deep Noise Suppression (DNS) Challenge

Jianwei Yu (Tencent AI lab); Hangting Chen (Tencent ASSP OTeam); Yi Luo (Tencent AI Lab); Rongzhi Gu (Tencent); Chao Weng (Tencent AI Lab)

04:48 PM

6914 (GC-L11.06): CONVOLUTIONAL RECURRENT METRICGAN WITH SPECTRAL DIMENSION COMPRESSION FOR FULL-BAND SPEECH ENHANCEMENT

Zhongshu Hou (Nanjing University); Qinwen Hu (Nanjing University); Tianchi Sun (Nanjing University); Yuxiang Hu (Horizon Robotics); Changbao Zhu (Horizon Robotics); Kai Chen (Nanjing University)

IVMSP-L8: Human Action Recognition

Room: Athina

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Haoqian Wang, Joon Son Chung

03:40 PM

2972 (IVMSP-L8.1): Prior-Enhanced Temporal Action Localization using Subject-aware Spatial Attention

Yifan Liu (Tsinghua University); Yubao Tang (PAII Inc.); Ning Zhang (PAII Inc.); Rwei-Sung Lin (PAII Inc.); Haoqian Wang (Tsinghua Shenzhen International Graduate School, Tsinghua University)

03:55 PM

6198 (IVMSP-L8.2): AV-TAD: AUDIO-VISUAL TEMPORAL ACTION DETECTION WITH TRANSFORMER

Yangcheng Li (Shanghai Jiao Tong University); Zefang Yu (Shanghai Jiao Tong University); Suncheng Xiang (Shanghai Jiao Tong University); Ting Liu (Shanghai Jiao Tong University); Yuzhuo Fu (sjtu)

04:10 PM

388 (IVMSP-L8.3): SELF-SUFFICIENT FRAMEWORK FOR CONTINUOUS SIGN LANGUAGE RECOGNITION

Youngjoon Jang (KAIST); Youngtaek Oh (KAIST); Jae Won Cho (KAIST); Myungchul Kim (KAIST); Dong-Jin Kim (Hanyang University); In So Kweon (KAIST); Joon Son Chung (KAIST)

04:25 PM

3983 (IVMSP-L8.4): BODY PRIOR GUIDED GRAPH CONVOLUTIONAL NEURAL NETWORK FOR SKELETON-BASED ACTION RECOGNITION

Qianshuo Hu (Chongqing university of technology); Hong Liu (Peking University Shenzhen Graduate School); Hua-qiu Wang (Chongqing University of Technology); Mengyuan Liu (Peking University, Shenzhen Graduate School)

04:40 PM

1901 (IVMSP-L8.5): TAMFormer: Multi-Modal Transformer with Learned Attention Mask for Early Intent Prediction

Nada Osman (University of Padova); Guglielmo Camporese (University of Padova); Lamberto Ballan (University of Padova)

04:55 PM

437 (IVMSP-L8.6): Dual-Feature Enhancement for Weakly Supervised Temporal Action Localization

Siyang Liu (University of Science and Technology of China); Qiankun Liu (Beijing Institute of Technology); Qi Chu (University of Science and Technology of China); Bin Liu (University of Science and Technology of China); Nenghai Yu (University of Science and Technology of China)

MLSP-L11: Deep Generative Model

Room: Jupiter

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Danilo Comminiello,

03:40 PM

1565 (MLSP-L11.01): STRING-BASED MOLECULE GENERATION VIA MULTI-DECODER VAE

Kisoo Kwon (Samsung Advanced Institute of Technology, Samsung Electronics); Kuhwan Jeong (Samsung Advanced Institute of Technology); Junghyun Park (samsung electronics); HWIDONG NA (Samsung Electronics.); Jinwoo Shin (KAIST)

03:55 PM

4161 (MLSP-L11.02): Graph Contrastive Learning with Learnable Graph Augmentation

Xinyan Pu (Southeast University); Ke Zhang (Southeast University); Huazhong Shu (Southeast University); Jean-Louis Coatrieux ("LTSI, Rennes, France"); Youyong Kong (Southeast University)

04:10 PM

3180 (MLSP-L11.03): CONDITIONING AND SAMPLING IN VARIATIONAL DIFFUSION MODELS FOR SPEECH SUPER-RESOLUTION

Chin-Yun Yu (Queen Mary University of London); Sung-Lin Yeh (University of Edinburgh); George Fazekas (QMUL); Hao Tang (The University of Edinburgh)

Thursday, June 8

04:25 PM

5068 (MLSP-L11.04): EVALUATION OF CATEGORICAL GENERATIVE MODELS - BRIDGING THE GAP BETWEEN REAL AND SYNTHETIC DATA

Florence Regol (McGill University); Anja M Kroon (McGill University); Mark Coates (McGill University)

04:40 PM

6053 (MLSP-L11.05): Diffusion Probabilistic Modeling for Fine-Grained Urban Traffic Flow Inference With Relaxed Structural Constraint

Xovee Xu (University of Electronic Science and Technology of China); Yutao Wei (University of Electronic Science and Technology of China); Pengyu Wang (School of Information and Software Engineering, University of Electronic Science and Technology of China); Xucheng Luo (University of Electronic Science and Technology of China); Fan Zhou (School of Information and Software Engineering, University of Electronic Science and Technology of China); Goce Trajcevski (Iowa State University)

04:55 PM

4977 (MLSP-L11.06): Single-Shot Domain Adaptation via Target-Aware Generative Augmentations

Rakshith Subramanyam (Arizona State University); Kowshik Thopalli (Arizona State University); Spring Berman (Arizona State University, USA); Pavan Turaga (Arizona State University); Jayaraman J. Thiagarajan (Lawrence Livermore National Laboratory)

MSP-L3: Multimodal Signal Processing and Analysis I

Room: Salon de Roses A

Type: Oral

03:40 PM to 05:10 PM

Chair(s): TBA

03:40 PM

493 (MSP-L3.1): A Point Is A Wave: Point-wave Network for Place Recognition

Ge Li (SECE, Shenzhen Graduate School, Peking University); Ruonan Zhang (Peking University, shenzhen graduate school)

03:55 PM

726 (MSP-L3.2): C2BN: Cross-modality and Cross-scale Balance Network for multi-modal 3D Object Detection

BoNan Ding (Chongqing University); Jin Xie (Chongqing University); Jing Nie (Chongqing University)

04:10 PM

1106 (MSP-L3.3): Adaptive Mask Co-optimization for Modal Dependence in Multimodal Learning

Ying Zhou (Xidian University); Xuefeng Liang (Xidian University); ShiQuan Zheng (Xidian University); Huijun Xuan (Xidian University); Takatsune Kumada (Kyoto University)

04:25 PM

3605 (MSP-L3.4): A Mutli-stage Hierarchical Relational Graph Neural Network for Multimodal Sentiment Analysis

Peizhu Gong (Shanghai Maritime University); Jin Liu (Shanghai Maritime University); Xiliang Zhang (Shanghai Maritime University); XingYe Li (Shanghai Maritime University)

04:40 PM

1467 (MSP-L3.5): Interaction-Assisted Multi-Modal Representation Learning for Recommendation

Hao Wu (Alibaba Group); Jiajie Wang (Alibaba Group); Zhonglin Zu (Alibaba Group)

04:55 PM

1519 (MSP-L3.6): MULTI-SCALE COMPOSITIONAL CONSTRAINTS FOR REPRESENTATION LEARNING ON VIDEOS

Georgios Paraskevopoulos (National Technical University of Athens); Chandrashekhar Lavania (AWS AI Labs); Lovish Chum (Amazon Inc.); Shiva Sundaram (Amazon)

SLT-L19: Speech Enhancement - Self-Supervised Learning

Room: Delphi

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Takuya Yoshioka, Shinji Watanabe

03:40 PM

915 (SLT-L19.1): PERCEIVE AND PREDICT: SELF SUPERVISED SPEECH REPRESENTATION BASED LOSS FUNCTIONS FOR SPEECH ENHANCEMENT

George L Close (University of Sheffield); William Ravenscroft (The University of Sheffield); Thomas Hain (University of Sheffield); Stefan Goetze (University of Sheffield)

03:55 PM

2006 (SLT-L19.2): Data2vec-SG: Improving Self-supervised Learning Representations for Speech Generation Tasks

Heming Wang (The Ohio State University); Yao Qian (Microsoft); Hemin Yang (Microsoft); Naoyuki Kanda (Microsoft); Peidong Wang (Microsoft); Takuya Yoshioka (Microsoft); Xiaofei Wang (Microsoft); Yiming Wang (Microsoft Corporation); Shujie Liu (Microsoft Research Asia); Zhuo Chen (Microsoft); DeLiang Wang (Ohio State University); Michael Zeng (Microsoft)

Thursday, June 8

04:10 PM

3343 (SLT-L19.3): Speech separation with large-scale self-supervised learning

Zhuo Chen (Microsoft); Naoyuki Kanda (Microsoft); Jian Wu (Microsoft); Yu Wu (Microsoft Research Asia); Xiaofei Wang (Microsoft); Takuya Yoshioka (Microsoft); Jinyu Li (Microsoft); Sunit Sivasankaran (Microsoft); Sefik Emre Eskimez (Microsoft)

04:25 PM

3511 (SLT-L19.4): Self-Supervised Learning-Based Source Separation for Meeting Data

Yuang Li (University of Cambridge); Xianrui Zheng (University of Cambridge); Phil Woodland (Machine Intelligence Laboratory, Cambridge University Department of Engineering)

04:40 PM

4456 (SLT-L19.5): AN ADAPTER BASED MULTI-LABEL PRE-TRAINING FOR SPEECH SEPARATION AND ENHANCEMENT

Tianrui Wang (Beijing Jiaotong University); Xie Chen (Shanghai Jiaotong University); Zhuo Chen (Microsoft); Shu Yu (SJTU); Weibin Zhu (Beijing Jiaotong University(China))

04:55 PM

5785 (SLT-L19.6): Self-Supervised Learning for Speech Enhancement Through Synthesis

Bryce Irvin (Bose Corporation); Marko Stamenovic (Bose Corp.); Mikolaj Kegler (Bose Corp.); Li-Chia Yang (Bose Corp.)

SPCN-L2: Distributed and Reliable Signal Processing and Communications

Room: Salon de Roses B

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Waheed Bajwa,

03:40 PM

3282 (SPCN-L2.1): EXPECTATION PROPAGATION ON FACTOR GRAPHS BASED ON MATRIX DECOMPOSITION

Adam Mekhiche (Thales); Antonio Maria Cipriano (Thales); Charly Poulliat (IRIT/INPT)

03:55 PM

3955 (SPCN-L2.2): Codes Correcting Burst and Arbitrary Erasures for Reliable and Low-Latency Communication

Serge Kas Hanna (Aalto); Zhiyuan Tan (Huawei); Wen Xu (Huawei); Antonia Wachter-Zeh (TUM)

04:10 PM

4320 (SPCN-L2.3): Joint Estimation of Clustered User Activity and Correlated Channels with Unknown Covariance in mMTC

Hamza Djelouat (University of Oulu); Markus Leinonen (University of Oulu); Markku Juntti (OULU, Finland)

04:25 PM

4734 (SPCN-L2.4): Reducing the communication and computational cost of random Fourier features Kernel LMS in diffusion networks

Daniel G Tiglea (Universidade de Sao Paulo); Renato Candido (Universidade de São Paulo); Luis Antonio Azpicueta-Ruiz (Universidad Carlos III de Madrid); Magno T.M. Silva (University of Sao Paulo)

04:40 PM

6757 (SPCN-L2.5): A Sequence Repetition Node-Based Successive Cancellation List Decoder for 5G Polar Codes: Algorithm and Implementation (SPS Journal Paper)*

Yuqing Ren (EPFL); Andreas Toftegaard Kristensen (EPFL); Yifei Shen (EPFL); Alexios Balatsoukas-Stimming (Eindhoven University of Technology); Chuan Zhang (Southeast University); Andreas Burg (EPFL)

04:55 PM

6730 (SPCN-L2.6): Message Passing-Based 9-D Cooperative Localization and Navigation with Embedded Particle Flow (SPS Journal Paper)*

Lukas Wielandner (Graz University of Technology)

SS-L13: Resource-efficient Real-time Neural Speech Separation

Room: Nefeli A

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Zhong-Qiu Wang,

03:40 PM

444 (SS-L13.1): On the Design and Training Strategies for RNN-based Online Neural Speech Separation Systems

Kai Li (Tsinghua University); Yi Luo (Tencent AI Lab)

03:55 PM

2493 (SS-L13.2): Computational Efficient Monaural Speech Enhancement with Universal Sample rate Band-split RNN

Jianwei Yu (Tencent AI lab); Yi Luo (Tencent AI Lab)

Thursday, June 8

04:10 PM

3416 (SS-L13.3): Towards Real-Time Single-Channel Speech Separation in Noisy and Reverberant Environments

Julian Neri (McGill University); Sebastian Braun (Microsoft)

04:25 PM

4244 (SS-L13.4): Predictive SkIM: Contrastive Predictive Coding for Low-Latency Online Speech Separation

Chenda Li (Shanghai Jiao Tong University); Yifei Wu (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)

04:40 PM

5485 (SS-L13.5): Neural Speech Enhancement with Very Low Algorithmic Latency and Complexity via Integrated Full- and Sub-Band Modeling

Zhong-Qiu Wang (Carnegie Mellon University); Samuele Cornell (Università Politecnica delle Marche); Shukjae Choi (Hyundai Motor Company); Younglo Lee (42dot); Byeong-Yeol Kim (42dot); Shinji Watanabe (Carnegie Mellon University)

04:55 PM

6492 (SS-L13.6): Latent Iterative Refinement for Modular Source Separation

Dimitrios Bralios (University of Illinois at Urbana-Champaign); Efthymios Tzinis (University of Illinois at Urbana-Champaign); Gordon Wichern (Mitsubishi Electric Research Laboratories (MERL)); Paris Smaragdis (University of Illinois at Urbana-Champaign); Jonathan LeRoux (Mitsubishi Electric Research Laboratories (MERL))

AASP-P11: Multichannel Speech Enhancement, Dereverberation, and System Identification

Room: Dome A

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Simon Doclo,

2110 (AASP-P11.1): SWITCHING KRONECKER PRODUCT LINEAR FILTERING FOR MULTISPEAKER ADAPTIVE SPEECH DEREVERBERATION

Gongping Huang (University of Erlangen-Nuremberg); Jacob Benesty (INRS); Israel Cohen (Technion); Emil Winebrand (Insoundz Ltd.); Jingdong Chen (Northwestern Polytechnical University); Walter Kellermann (Friedrich-Alexander-University Erlangen-Nürnberg)

4930 (AASP-P11.2): Exploiting speaker embeddings for improved microphone clustering and speech separation in ad-hoc microphone arrays

Stijn Kindt (UGent); Jenthe Thienpondt (IDLab, Ghent University); Nilesh Madhu (IDLab, Ghent University - imec)

1720 (AASP-P11.3): Linear Microphone Array Parallel to the Driving Direction for In-Car Speech Enhancement

Masanori Tsujikawa (NEC Corporation); Akihiko K. Sugiyama (Yahoo Japan Corporation); Ken Hanazawa (NEC Laboratories America, Inc.); Yoshinobu Kajikawa (Kansai University)

2350 (AASP-P11.4): A LIGHTWEIGHT FOURIER CONVOLUTIONAL ATTENTION ENCODER FOR MULTI-CHANNEL SPEECH ENHANCEMENT

Siyu Sun (Wuhan University); Jian Jin (RTC Lab, ByteDance); Zhe Han (RTC Lab, ByteDance); Xianjun Xia (RTC Lab, ByteDance); Li Chen (ByteDance); Yijian Xiao (RTC Lab, ByteDance); Piao Ding (RTC Lab, ByteDance); Shenyi Song (RTC Engineering, ByteDance); Roberto Togneri (The University of Western Australian); Haijian Zhang (Wuhan University)

922 (AASP-P11.5): Subspace Hybrid Beamforming for Head-worn Microphone Arrays

Sina Hafezi (Imperial College London); Alastair H Moore (Imperial College London); Pierre Guiraud (Imperial College London); Patrick A. Naylor (Imperial College London); Jacob Donley (Facebook); Vladimir Tourbabin (Meta); Thomas Lunner (Meta)

4205 (AASP-P11.6): Dereverberation in Acoustic Sensor Networks using Weighted Prediction Error with Microphone-dependent Prediction Delays

Anselm Lohmann (University of Oldenburg); Toon van Waterschoot (Department of Electrical Engineering (ESAT-STADIUS/ETC)); Joerg Bitzer (Institute of Hearing Technology and Audiology, Jade University of Applied Sciences, Oldenburg); Simon Doclo (University of Oldenburg)

1754 (AASP-P11.7): SPEECH DEREVERBERATION WITH A REVERBERATION TIME SHORTENING TARGET

Rui Zhou (Westlake University); Wenye Zhu (Zhejiang University); Xiaofei Li (Westlake University)

1393 (AASP-P11.8): A FREQUENCY-DOMAIN RECURSIVE LEAST-SQUARES ADAPTIVE FILTERING ALGORITHM BASED ON A KRONECKER PRODUCT DECOMPOSITION

Hongsen He (Southwest University of Science and Technology); Jingdong Chen (Northwestern Polytechnical University); Jacob Benesty (INRS); Yi Yu (Southwest University of Science and Technology)

6684 (AASP-P11.9): Meta-AF: Meta-Learning for Adaptive Filters (SPS Journal Paper)*

Jonah Casebeer (University of Illinois at Urbana-Champaign); Nicholas J. Bryan (Adobe Research); Paris Smaragdis (University of Illinois at Urbana-Champaign)

Thursday, June 8

6722 (AASP-P11.10): Mask-based Neural Beamforming for Moving Speakers with Self-Attention-based Tracking (SPS Journal Paper)*

Tsubasa Ochiai (NTT); Marc Delcroix (NTT); Tomohiro Nakatani (NTT Communication Science Laboratories); Shoko Araki (NTT Corporation)

6749 (AASP-P11.11): Towards Unified All-Neural Beamforming for Time and Frequency Domain Speech Separation (SPS Journal Paper)*

rongzhi gu (Tencent); Shixiong Zhang (Tencent); Yuexian Zou (Peking University); Dong Yu (Tencent)

6792 (AASP-P11.12): Spatial Audio Signal Enhancement by a Two-Stage Source—System Estimation With Frequency Smoothing for Improved Perception (SPS Journal Paper)*

Moti Lugasi (Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, 84105)

AASP-P12: Multilabel Acoustic Event Classification

Room: Dome B

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Shoko Araki,

1368 (AASP-P12.1): SARdBSScene: Dataset and ResNet Baseline for Audio Scene Source Counting and Analysis

Michael Nigro (Toronto Metropolitan University); Sri Krishnan (Ryerson University)

2027 (AASP-P12.2): GCT: GATED CONTEXTUAL TRANSFORMER FOR SEQUENTIAL AUDIO TAGGING

Yuanbo Hou (Ghent University); Yun Wang (Meta); Wenwu Wang (University of Surrey); Dick Botteldooren (Ghent University)

185 (AASP-P12.3): Play It Back: Iterative Attention for Audio Recognition

Alexandros Stergiou (Vrije Universiteit Brussel); Dima Damen (University of Bristol)

325 (AASP-P12.4): Multiscale Audio Spectrogram Transformer for Efficient Audio Classification

Wentao Zhu (Amazon); Mohamed Omar (Amazon)

1610 (AASP-P12.5): Efficient Large-scale Audio Tagging via Transformer-to-CNN Knowledge Distillation

Florian Schmid (Johannes Kepler University); Khaled Koutini (Johannes Kepler University); Gerhard Widmer (Johannes Kepler University)

2247 (AASP-P12.6): Heterogeneous Graph Learning for Acoustic Event Classification

Amir Shirian (University of Warwick); Mona Ahmadian (University of Surrey); Krishna Somandepalli (University of Southern California); Tanaya Guha (University of Glasgow)

4119 (AASP-P12.7): Attention Mixup: An Accurate Mixup Scheme based on Interpretable Attention Mechanism for Multi-label Audio Classification

Wuyang Liu (School of Cyber Science and Engineering, Wuhan University); Yanzhen Ren (Computer School of Wuhan University); Jingru Wang (School of Cyber Science and Engineering, Wuhan University)

3307 (AASP-P12.8): DATASET BALANCING CAN HURT MODEL PERFORMANCE

Richard C Moore (Google Research); Dan Ellis (Google, Inc); Eduardo Fonseca (Google Research); Shawn Hershey (Google); Aren Jansen (Google); Manoj Plakal (Google, Inc.)

5098 (AASP-P12.9): Does a quieter city mean fewer complaints? The Sounds of New York City During COVID-19 Lockdown

Mark Cartwright (New Jersey Institute of Technology); Magdalena Fuentes (New York University); Charlie Mydlarz (New York University); Fabio Miranda (University of Illinois, USA); Juan P Bello (New York University)

5442 (AASP-P12.10): LEARNING TO DETECT NOVEL AND FINE-GRAINED ACOUSTIC SEQUENCES USING PRETRAINED AUDIO REPRESENTATIONS

Vasudha Kowtha (Apple); Miquel Espi (Apple); Jonathan J Huang (Apple); Yichi Zhang (Apple); Carlos Avendano (Apple)

BISP-P6: Deep Learning for Medical Imaging

Room: Dome G

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Vishal Monga,

634 (BISP-P6.1): LOW-DOSE CT RECONSTRUCTION VIA OPTIMIZATION-INSPIRED GAN

jiawei jiang (zhejiang university of technology); Yuchao Feng (Zhejiang University of Technology); Honghui Xu (Zhejiang University of Technology); Jianwei Zheng (Zhejiang University of Technology)

Thursday, June 8

5648 (BISP-P6.2): Deep Triple-Supervision Learning Unannotated Surgical Endoscopic Video Data for Monocular Dense Depth Estimation

Wenkang Fan (Xiamen University); KaiYun Zhang (Xiamen University); Hong Shi (Fujian Cancer Hospital & Fujian Medical University Cancer Hospital); Jianhua Chen (Fujian Cancer Hospital & Fujian Medical University Cancer Hospital); Yinran Chen (Xiamen University); Xiongbiao Luo (Xiamen University)

4226 (BISP-P6.3): ADAPTIVE NON-LOCAL GENERATIVE ADVERSARIAL NETWORKS FOR LOW-DOSE CT IMAGE DENOISING

Linlin Yang (Xidian University); Hongying Liu (Key Lab. of Intelligent Perception and Image Understanding of Ministry of Education, School of Artificial Intelligence, Xidian University, China); Fanhua Shang (Tianjin University); Yuanyuan Liu (Xidian University)

1771 (BISP-P6.4): This changes to that : Combining causal and non-causal explanations to generate disease progression in capsule endoscop

Anuja Vats (NTNU); Ahmed Mohammed (NTNU); Marius Pedersen (NTNU); Nirmalie Wiratunga (Robert Gordon University)

4973 (BISP-P6.5): CO-OPERATIVE CNN FOR VISUAL SALIENCY PREDICTION ON WCE IMAGES

George Dimas (Department of Computer Science and Biomedical Informatics, University of Thessaly, Greece); Anastasios Koulaouzidis (The Royal Infirmary of Edinburgh); Dimitris K Iakovidis (Department of Computer Science and Biomedical Informatics, University of Thessaly, Greece)

2159 (BISP-P6.6): Assessing the Robustness of Deep Learning-Assisted Pathological Image Analysis under Practical Variables of Imaging System

YUXUAN SUN (Westlake University); Chenglu Zhu (Westlake University); Yunlong Zhang (Westlake University); Honglin Li (Westlake University); Pingyi Chen (Westlake University); Lin Yang (Westlake University)

2114 (BISP-P6.7): DIFFUSIONNET: AN EFFICIENT FRAMEWORK TO CLASSIFY SINGLE-MOLECULE IMAGES WITH LATENT ENTROPY MINIMIZATION

Soumee Guha (University of Virginia); Olivia de Cuba (University of Virginia); Andreas Gahlmann (University of Virginia); Scott Acton (University of Virginia)

2228 (BISP-P6.8): Benchmarking White Blood Cell Classification Under Domain Shift

Satoshi Tsutsui (Nanyang Technological University, Singapore); Zhengyang Su (NTU); Bihan Wen (Nanyang Technological University)

2836 (BISP-P6.9): Spatio-Temporal Structure Consistency for Semi-supervised Medical Image Classification

Lei Wentao (The Chinese University of Hongkong, Shenzhen); Lei Liu (The Chinese University of Hong Kong, Shenzhen); Li Liu (Shenzhen Research Institute of Big Data, the chinese university of hong kong shenzhen)

5086 (BISP-P6.10): AN ADAPTIVE ENHANCEMENT METHOD FOR GASTROINTESTINAL LOW-LIGHT IMAGES OF CAPSULE ENDOSCOPE

Peixuan Liu (Jiangnan University); Yinghui Wang (Jiangnan University); Jinlong Yang (Jiangnan University); Wei Li (Jiangnan University)

3346 (BISP-P6.11): High-dimensional confidence regions in sparse MRI

Frederik Hoppe (RWTH Aachen University); Felix Kraemer (Technical University of Munich); Claudio Mayrink Verdun (Technical University of Munich); Marion Menzel (GE Global Research); Holger Rauhut (RWTH Aachen University)

4600 (BISP-P6.12): RETINAL BIOMARKERS FOR DETECTING DIABETIC RETINOPATY USING SMARTPHONE-BASED DEEP LEARNING FRAMEWORKS

Mahmut Karakaya (Kennesaw State University); Ramazan Aygun (Kennesaw State University)

MSP-P3: Machine/Deep Learning Methodologies for Multimedia

Room: Dome F

Type: Poster

03:40 PM to 05:10 PM

Chair(s): TBA

793 (MSP-P3.1): Step Restriction for Improving Adversarial Attacks

Keita Goto (Tokyo Institute of Technology); Shinta Otake (Tokyo Institute of Technology); Rei Kawakami (Tokyo Institute of Technology); Nakamasa Inoue (Tokyo Institute of Technology)

768 (MSP-P3.2): FedVMR: A New Federated Learning method for Video Moment Retrieval

Yan Wang (Shandong University); Xin Luo (Shandong University); Zhen-Duo Chen (Shandong University); Peng-Fei Zhang (University of Queensland); Meng Liu (Shandong Jianzhu University); Xin-Shun Xu (Shandong University)

976 (MSP-P3.3): Multi-source Templates Learning for Real-time Aerial Tracking

Yiming Sun (East China Normal University); Yang Li (East China Normal University); Changbo Wang (East China Normal University)

1264 (MSP-P3.4): Embrace Smaller Attention: Efficient Cross-Modal Matching with Dual Gated Attention Fusion.

Weikuo Guo (Dalian University of Technology); Xiangwei Kong (Zhejiang University)

Thursday, June 8

1923 (MSP-P3.5): Region-awared transformer with asymmetric loss in multi-label classification

Long Feng (Northwest University); Guohua Geng (Northwest University); Chen Guo (Shaanxi Normal University); longquan yan (Northwest University); Xingrui Ma (Northwest University); Zhan Li (Northwest University); Kang Li (Northwest University)

1289 (MSP-P3.6): CF-VTON: Multi-Pose Virtual Try-On with Cross-domain Fusion

Chenghu Du (Wuhan university of technology); Shengwu Xiong (Wuhan University of Technology)

919 (MSP-P3.7): LEARNING TO LOCATE VISUAL ANSWER IN VIDEO CORPUS USING QUESTION

Bin Li (Hunan University); Yixuan Weng (CASIA); Bin Sun (Hunan University); Shutao Li (Hunan University)

3694 (MSP-P3.8): An End-to-End Framework for Partial View-aligned Clustering with Graph Structure

Liang Zhao (Dalian University of Technology); Qiongie Xie (大连理工大学); Songtao Wu (大连理工大学); shubin ma (Dalian University of Technology)

6261 (MSP-P3.9): Detecting Out-of-distribution Examples via Class-conditional Impressions Reappearing

Jinggang Chen (Huazhong University of Science and Technology); Xiaoyang Qu (Ping An Technology (Shenzhen) Co., Ltd); Junjie Li (Huazhong University of Science and Technology); Jianzong Wang (Ping An Technology (Shenzhen) Co., Ltd); Jiguang Wan (Huazhong University of Science and Technology); Jing Xiao (Ping An Insurance (Group) Company of China)

6290 (MSP-P3.10): Guide and Select: A Transformer-based Multimodal Fusion Method for Points of Interest Description Generation

Hanqing Liu (Tsinghua Shenzhen International Graduate School); Wei Wang (Tsinghua University); Niu Hu (Tsinghua University); Hai-Tao Zheng (Tsinghua University); Rui Xie (Meituan); Wei Wu (Meituan); Yang Bai (Tsinghua University)

954 (MSP-P3.11): Boosting Fine-grained Sketch-based Image Retrieval with Self-supervised Learning

Zhaolong Zhang (Fudan University); Yangdong Chen (Fudan University); Yuejie Zhang (Fudan University); Rui Feng (Fudan University); Tao Zhang (Shanghai University of Finance and Economics)

MSP-P4: Human-Centric Multimedia

Room: Dome J

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Theodora Chaspari,

5707 (MSP-P4.1): Binauralization Robust to Camera Rotation Using 360° Videos

Masaki Yoshida (Hokkaido University); Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

1238 (MSP-P4.2): Salient Co-Speech Gesture Synthesizing with Discrete Motion Representation

Zijie Ye (Tsinghua University); Jia Jia (Tsinghua University); Haozhe Wu (Tsinghua University); Shuo Huang (Tsinghua University); Shikun Sun (Tsinghua University); Junliang Xing (Tsinghua University)

1382 (MSP-P4.3): Shuffled Autoregression For Motion Interpolation

Shuo Huang (Tsinghua University); Jia Jia (Tsinghua University); Zongxin Yang (Zhejiang University); Wei Wang (University of Oxford); Haozhe Wu (Tsinghua University); Yi Yang (Zhejiang University); Junliang Xing (Tsinghua University)

3113 (MSP-P4.4): GTN-Bailando: Genre Consistent Long-Term 3D Dance Generation based on Pre-trained Genre Token Network

Haolin Zhuang (Tsinghua University); Shun Lei (Tsinghua University); Long Xiao (Tsinghua University); Weiqin Li (Tsinghua University); Liyang Chen (Tsinghua University); Sicheng Yang (Tsinghua University); Zhiyong Wu (Tsinghua University); Shiyin Kang (XVerse Inc.); Helen Meng (The Chinese University of Hong Kong)

4007 (MSP-P4.5): DyLiteRADHAR: DYNAMIC LIGHTWEIGHT SLOWFAST NETWORK FOR HUMAN ACTIVITY RECOGNITION USING MMWAVE RADAR

Biyun Sheng (Nanjing University of Posts and Telecommunications); Yan Bao (Nanjing University of Posts and Telecommunications); Fu Xiao (Nanjing University of Posts and Telecommunications); Linqing Gui (Nanjing University of Posts and Telecommunications)

3786 (MSP-P4.6): Whether Contribution of Features Differ Between Video-mediated and In-person Meetings in Important Utterance Estimation

Fumio Nihei (NTT); Ryo Ishii (NTT); Yukiko Nakano (Seikei Univeristy); Atsushi Fukayama (NTT); Takao Nakamura (NTT)

4025 (MSP-P4.7): JOINT ROBUST REPRESENTATION AND GENERALIZATION ENHANCEMENT FOR CROSS-MODALITY PERSON RE-IDENTIFICATION

Heqing Cheng (Chongqing University); Yong Feng (Chongqing University); Mingliang ZHOU (Chongqing University); Xian-cai Xiong (Key Laboratory of Monitoring, Evaluation and Early Warning of Territorial Spatial Planning Implementation, Ministry of Natural Resources); Yongheng Wang (Zhejiang Lab); Qiang Baohua (Guilin University of Electronic Technology)

Thursday, June 8

4561 (MSP-P4.8): CN-CVS: A Mandarin Audio-Visual Dataset for Large Vocabulary Continuous Visual to Speech Synthesis
Chen Chen (Tsinghua University); Dong Wang (Tsinghua University); Thomas Fang Zheng ("CSLT, Tsinghua University")

4791 (MSP-P4.9): Toward privacy-enhancing ambulatory-based well-being monitoring: Investigating user re-identification risk in multimodal data
Ravi Pranjali (Texas A&M University); Ranjana Seshadri (Texas A&M University); Rakesh Kumar Sanath Kumar Kadaba (Texas A&M University); Tiantian Feng (University of Southern California); Shrikanth Narayanan (University of Southern California); Theodora Chaspari (Texas A&M University)

4860 (MSP-P4.10): Contrastive Self-Supervised Learning for Automated Multi-Modal Dance Performance Assessment
Yun Zhong (Imperial College London); Fan Zhang (Imperial College London); Yiannis Demiris (Imperial College London)

5528 (MSP-P4.11): Unsupervised Video Anomaly Detection for Stereotypical Behaviours in Autism
Jiaqi Gao (Fudan University); Xinyang Jiang (Microsoft Research Asia); Yuqing Yang (Microsoft Research); Dongsheng Li (Microsoft Research Asia); Lili Qiu (The University of Texas at Austin)

6773 (MSP-P4.12): Cross modal video representations for weakly supervised active speaker localization (SPS Journal Paper)*
Rahul Sharma (University of Southern California); Krishna Somandepalli (University of Southern California); Shrikanth Narayanan (University of Southern California)

SAM-P7: Source Localization and Separation

Room: Dome M

Type: Poster

03:40 PM to 05:10 PM

Chair(s): TBA

986 (SAM-P7.1): Robust Iterative Solution for Linear Array-Based 3-D Localization By Message Passing
Yimao Sun (Sichuan University); Dominic Ho (Nil); Yanbing Yang (Sichuan University); Lei Zhang (Sichuan University); Liangyin Chen (Sichuan University)

2130 (SAM-P7.2): Bias Reduced Semidefinite Relaxation Method for Multistatic Localization in the Absence of Transmitter Position and Its Synchronization
Pei Jian (Ningbo University); Gang Wang (Ningbo University); Dominic Ho (Nil); Lei Huang (Shenzhen University)

2973 (SAM-P7.3): Dynamic Independent Component Extraction with Blending Mixing Vector: Lower Bound on Mean Interference-to-Signal Ratio
Jaroslav Čmejla (Technical University of Liberec); Zbynek Koldovsky (Technical University of Liberec); Václav Kautský (Technical University of Liberec); Tulay Adali (University of Maryland, Baltimore County)

3235 (SAM-P7.4): ON SUPER-RESOLUTION WITH SEPARATION PRIOR
Xingyun Mao (Shanghai Jiao Tong university); HENG QIAO (SHANGHAI JIAO TONG UNIVERSITY)

4228 (SAM-P7.5): Near-field Localization with Dynamic Metasurface Antennas
Qianyu Yang (Nanjing University of Posts and Telecommunications); Anna Guerra (the National Research Council of Italy, Institute of Electronics, Computer and Telecommunication Engineering); Francesco Guidi (the National Research Council of Italy, Institute of Electronics, Computer and Telecommunication Engineering); Nir Shlezinger (Ben-Gurion University); Haiyang Zhang (Nanjing University of Posts and Telecommunications); Davide Dardari (DEIS-University of Bologna, Italy); Baoyun Wang (Nanjing University of Posts and Telecommunications); Yonina Eldar ()

4528 (SAM-P7.6): Soft label coding for end-to-end sound source localization with ad-hoc microphone arrays
Linfeng Feng (Northwestern Polytechnical University); Yijun Gong (Northwestern Polytechnical University); Zhang XiaoLei (Northwestern Polytechnical University)

5973 (SAM-P7.7): SOURCE LOCALIZATION FOR EXTREMELY LARGE-SCALE ANTENNA ARRAYS WITH SPATIAL NON-STATIONARITY
Xiaohuan Wu (Nanjing University of Posts and Telecommunications); Ji Sun (Nanjing University of Posts and Telecommunications); Xiaoyuan Jia (Nanjing University of Posts and Telecommunications); Shuxin Wang (Nanjing University of Posts and Telecommunications)

6338 (SAM-P7.8): A PROXIMAL APPROACH TO IVA-G WITH CONVERGENCE GUARANTEES
Clément Cosserat (CVN); Ben Gabrielson (University of Maryland, Baltimore County); Emilie Chouzenoux (Inria Saclay); Jean-Christophe Pesquet (CentraleSupélec); Tulay Adali (University of Maryland, Baltimore County)

6422 (SAM-P7.9): Mixed Far-field and Near-field Source Localization Based on Low-Rank Matrix Reconstruction
Yunchang Liu (Jilin University); Hong Jiang (Jilin University); Qi Zhang (Jilin University)

6694 (SAM-P7.10): Location Estimates from Channel State Information Via Binary Programming (SPS Journal Paper)*
Muhammed Tahsin Rahman (University of Toronto); Shahrokh Valaee (University of Toronto)

Thursday, June 8

6820 (SAM-P7.11): Localizing Unsynchronized Sensors with Unknown Sources (SPS Journal Paper)*

Dalia El Badawy (-); Ivan Dokmanic (University of Basel); Viktor Larsson (Lund University); Marc Pollefeys (ETH Zurich / Microsoft)

SLT-P36: Speech Enhancement /Audio-Visual, Multi-Channel, and Other

Room: Dome C

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Li Li, sabato marco siniscalchi

1336 (SLT-P36.1): Extreme bandwidth extension network applied to speech signals captured with noise-resilient body-conduction microphones

Julien Hauret (Conservatoire national des arts et métiers); Thomas Joubaud (ISL); Véronique Zimpfer (Department of Acoustics and Soldier Protection, French-German Research Institute of Saint-Louis (ISL)); Éric BAVU (Conservatoire National des Arts et Métiers)

2499 (SLT-P36.2): LA-VocE: Low-SNR Audio-visual Speech Enhancement using Neural Vocoders

Rodrigo Mira (Imperial College London); Buye Xu (Meta Reality Labs Research); Jacob Donley (Facebook); Anurag Kumar (Meta Reality Labs Research); Stavros Petridis (Imperial College London / Meta); Vamsi Krishna Ithapu (Meta Reality Labs Research); Maja Pantic (Imperial College London / Meta)

3063 (SLT-P36.3): A MULTI-SCALE FEATURE AGGREGATION BASED LIGHTWEIGHT NETWORK FOR AUDIO-VISUAL SPEECH ENHANCEMENT

Haitao Xu (University of Science and Technology of China); Liangfa Wei (Tencent); Jie Zhang (University of Science and Technology of China); Jianming Yang (Tsinghua University); Yannan Wang (Tencent); Tian Gao (University of Science and Technology of China); Xin Fang (iFlytek Research); Lirong Dai (University of Science and Technology of China)

3241 (SLT-P36.4): Deep Subband Network for Joint Suppression of Echo, Noise and Reverberation in Real-Time Fullband Speech Communication

Feifei Xiong (Alibaba Group); Minya Dong (Alibaba Group); Kechenyong Zhou (Alibaba Group); Houwei Zhu (Alibaba Group); Jinwei Feng (Alibaba Group)

3522 (SLT-P36.5): Nonparallel High-Quality Audio Super Resolution with Domain Adaptation and Resampling CycleGANs

Reo Yoneyama (Nagoya University); Ryuichi Yamamoto (LINE Corp.); Kentaro Tachibana (LINE Corp.)

4063 (SLT-P36.6): Two-stage UNet with multi-axis gated multilayer perceptron for monaural noisy-reverberant speech enhancement

Zehua Zhang (Harbin Institute of Technology(Shenzhen)); Shiyun Xu (Harbin Institute of Technology(Shenzhen)); Xuyi Zhuang (Harbin Institute of Technology(Shenzhen)); Liyanu Zhou (Harbin Institute of Technology(Shenzhen)); Heng Li (Harbin Institute of Technology(Shenzhen)); Mingjiang Wang (Harbin Institute of Technology Shenzhen)

4421 (SLT-P36.7): Audio-visual Speech Enhancement with a Deep Kalman Filter Generative Model

Ali Golmakani (Inria Nancy Grand); Mostafa Sadeghi (INRIA); romain serizel (Université de Lorraine)

4789 (SLT-P36.8): Egocentric Audio-Visual Noise Suppression

Roshan S Sharma (Carnegie Mellon University); Weipeng He (Idiap Research Institute); Egor Lakomkin (Meta); Ju Lin (Meta); Yang Liu (Meta); Kaustubh Kalgaonkar (Meta)

4829 (SLT-P36.9): The 2nd Clarity Enhancement Challenge for hearing aid speech intelligibility enhancement: Overview and Outcomes

Michael Akeroyd (University of Nottingham); Will Bailey (University of Sheffield); Jon Barker (Professor); Trevor Cox (University of Salford); John F Culling (Cardiff University); Simone Graetzer (University of Salford); Graham Naylor (University of Nottingham); Zuzanna Podwinska (University of Salford); Zehai Tu (University of Sheffield)

6477 (SLT-P36.10): Incorporating Visual Information Reconstruction into Progressive Learning for Optimizing Audio-Visual Speech Enhancement

Chen-Yue Zhang (USTC); Hang Chen (USTC); Jun Du (University of Science and Technology of China); Baocai Yin (USTC,iFLYTEK); Jia Pan (iFlytek Research); Chin-hui Lee (Georgia Institute of Technology)

6708 (SLT-P36.11): Alias-and-Separate: Wideband Speech Coding Using Sub-Nyquist Sampling and Speech Separation (SPS Journal Paper)*

Soojoong Hwang (Hyundai Motor Company); Eunkyun Lee (GIST); Inseon Jang (Electronics and Telecommunications Research Institution); Jong Won Shin (Gwangju Institute of Science and Technology)

6806 (SLT-P36.12): DeFT-AN: Dense Frequency-Time Attentive Network for Multichannel Speech Enhancement (SPS Journal Paper)*

Dongheon Lee (KAIST); Jung-Woo Choi (KAIST)

Thursday, June 8

SLT-P37: Speech Enhancement - Separation and Target Speech Extraction

Room: Dome D

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Shoko Araki, Katerina Zmolikova

466 (SLT-P37.1): Papez: Resource-efficient Speech Separation with Auditory Working Memory

Hyunseok Oh (Seoul National University); Juheon Yi (Seoul National University); Youngki Lee (Seoul National University)

959 (SLT-P37.2): MossFormer: Pushing the Performance Limit of Monaural Speech Separation using Gated Single-head Transformer with Convolution-augmented Joint Self-Attentions

Shengkui Zhao (Alibaba Group); Bin Ma ("Alibaba, Singapore R&D Center")

1982 (SLT-P37.3): A Framework for Unified Real-time Personalized and Non-Personalized Speech Enhancement

Zhepei Wang (University of Illinois at Urbana-Champaign); Ritwik Giri (Amazon); Devansh Shah (Amazon Web Services); Jean-Marc Valin (Amazon); Michael M Goodwin (AWS); Paris Smaragdis (University of Illinois at Urbana-Champaign)

2003 (SLT-P37.4): ImagineNET: Target Speaker Extraction with Intermittent Visual Cue through Embedding Inpainting

Zexu Pan (National University of Singapore); Wupeng Wang (NUS); Marvin Borsdorf (University of Bremen); Haizhou Li (The Chinese University of Hong Kong (Shenzhen))

2661 (SLT-P37.5): QUANTITATIVE EVIDENCE ON OVERLOOKED ASPECTS OF ENROLLMENT SPEAKER EMBEDDINGS FOR TARGET SPEAKER SEPARATION

Xiaoyu Liu (Dolby Laboratories); Xu Li (Dolby Laboratories); Joan Serra (Dolby Laboratories)

3175 (SLT-P37.6): Unifying Speech Enhancement and Separation with Gradient Modulation for End-to-End Noise-Robust Speech Separation

Yuchen Hu (Nanyang Technological University); Chen Chen (Nanyang Technological University); Heqing Zou (Nanyang Technological University); Xionghu Zhong (Hunan University); Eng Siong Chng (Nanyang Technological University)

4212 (SLT-P37.7): X-SepFormer: End-to-end Speaker Extraction Network with Explicit Optimization on Speaker Confusion

KAI LIU (Huawei Technologies Co., Ltd.); Ziqing Du (Huawei Technologies Co., Ltd.); Xucheng Wan (Huawei Technologies Co., Ltd.); Zhou Huan (AARC, Huawei Technologies Co., Ltd.)

4305 (SLT-P37.8): Target Speaker Extraction with Ultra-Short Reference Speech by VE-VE Framework

Lei Yang (Samsung); Wei Liu (Samsung); Lufen Tan (Samsung); Jaemo Yang (Samsung); Han-gil Moon (Samsung)

4805 (SLT-P37.9): Deformable Temporal Convolutional Networks for Monaural Noisy Reverberant Speech Separation

William Ravenscroft (The University of Sheffield); Stefan Goetze (University of Sheffield); Thomas Hain (University of Sheffield)

5532 (SLT-P37.10): DasFormer: deep alternating spectrogram transformer for multi/single-channel speech separation

Shuo Wang (MSFT); Xiangyu Kong (Microsoft Research Asia); Xiulian Peng (Microsoft Research Asia); Mahmood Movassagh (Microsoft); Vinod Prakash (Microsoft); Yan Lu (Microsoft Research Asia)

5557 (SLT-P37.11): TFCNET: TIME-FREQUENCY DOMAIN CORRECTOR FOR SPEECH SEPARATION

Weinan Tong (Tsinghua University); Jiayu Zhu (Tsinghua University); Jun Chen (Tsinghua University); Zhiyong Wu (Tsinghua University); Shiyin Kang (XVerse Inc.); Helen Meng (The Chinese University of Hong Kong)

6737 (SLT-P37.12): USEV: Universal Speaker Extraction With Visual Cue (SPS Journal Paper)*

Zexu Pan (National University of Singapore)

SLT-P38: Speech Enhancement - Single Channel

Room: Dome E

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Jun Du, Emanuel Habets

766 (SLT-P38.1): Time-domain Speech Enhancement Assisted by Multi-resolution Frequency Encoder and Decoder

Hao Shi (Kyoto University); Masato Mimura (Kyoto University); Longbiao Wang (Tianjin University); Jianwu Dang (Tianjin University); Tatsuya Kawahara (Kyoto University)

935 (SLT-P38.2): PAAPLoss: A Phonetic-Aligned Acoustic Parameter Loss for Speech Enhancement

Muqiao Yang (Carnegie Mellon University); Joseph Konan (Carnegie Mellon University); David Bick (Carnegie Mellon University); Yunyang Zeng (Carnegie Mellon University); Shuo Han (Carnegie Mellon University); Anurag Kumar (Facebook Research); Shinji Watanabe (Carnegie Mellon University); Bhiksha Raj (Carnegie Mellon University)

958 (SLT-P38.3): D2Former: A Fully Complex Dual-Path Dual-Decoder Conformer Network using Joint Complex Masking and Complex Spectral Mapping for Monaural Speech Enhancement

Shengkui Zhao (Alibaba Group); Bin Ma ("Alibaba, Singapore R&D Center")

Thursday, June 8

1273 (SLT-P38.4): Semi-supervised speech enhancement based on speech purity

Zihao Cui (China Mobile Research Institute); Shilei Zhang (China Mobile Research Institute); Yanan Chen (China Mobile Research Institute); Yingying Gao (China Mobile Research Institute); Chao Deng (China Mobile Research Institute); Junlan Feng (China Mobile Research)

2693 (SLT-P38.5): TAPLoss: A Temporal Acoustic Parameter Loss for Speech Enhancement

Yunyang Zeng (Carnegie Mellon University); Joseph Konan (Carnegie Mellon University); Shuo Han (Carnegie Mellon University); Muqiao Yang (Carnegie Mellon University); David Bick (Carnegie Mellon University); Anurag Kumar (Facebook Research); Shinji Watanabe (Carnegie Mellon University); Bhiksha Raj (Carnegie Mellon University)

2778 (SLT-P38.6): SELECTIVE FILM CONDITIONING WITH CTC-BASED ASR PROBABILITY FOR SPEECH ENHANCEMENT

Da-Hee Yang (Hanyang University); Joon-Hyuk Chang (Hanyang University)

2899 (SLT-P38.7): Inter-SubNet: Speech Enhancement with Subband Interaction

Jun Chen (Tsinghua University); Wei Rao (Tencent); Zilin Wang (Tsinghua University); Jiuxin Lin (Tsinghua University); Zhiyong Wu (Tsinghua University); Yannan Wang (Tencent); Shi-dong Shang (tencent); Helen Meng (The Chinese University of Hong Kong)

3613 (SLT-P38.8): LEVERAGING HETEROSCEDASTIC UNCERTAINTY IN LEARNING COMPLEX SPECTRAL MAPPING FOR SINGLE-CHANNEL SPEECH ENHANCEMENT

Kuan-Lin Chen (University of California San Diego); Daniel D.E. Wong (Meta Platforms Inc.); Ke Tan (Meta Platforms, Inc.); Buye Xu (Work); Anurag Kumar (Facebook Research); Vamsi Krishna K Ithapu (Facebook Reality Labs)

3692 (SLT-P38.9): Mitigating Domain Dependency for Improved Speech Enhancement via SNR Loss Boosting

Lili Yin (Xinjiang University); Di Wu (Xinjiangdaxue); Zhibin Qiu (XinJiang University); Hao Huang (Xinjiang University)

4756 (SLT-P38.10): On batching variable size inputs for training end-to-end speech enhancement systems

Philippe Gonzalez (Technical University of Denmark); Tommy Sonne Alstrøm (Technical University of Denmark); Tobias May (Technical University of Denmark)

5826 (SLT-P38.11): Learning interpretable filters in Wav-Unet for speech enhancement

Félix MATHIEU (Telecom Paris); Thomas Courtat (Thales); Gaël Richard (Telecom Paris, Institut polytechnique de Paris); Geoffroy Peeters (LTCL - Télécom Paris, IP Paris)

5950 (SLT-P38.12): SINGLE-CHANNEL SPEECH ENHANCEMENT WITH DEEP COMPLEX U-NETWORKS AND PROBABILISTIC LATENT SPACE MODELS

Eike J Nustede (Carl von Ossietzky University Oldenburg); Jörn Anemüller (Carl von Ossietzky University Oldenburg)

SPCN-P2: Machine Learning Applications to Communications

Room: Dome H

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Zhongyuan Zhao,

3944 (SPCN-P2.1): Delay-aware Backpressure Routing Using Graph Neural Networks

Zhongyuan Zhao (Rice University); Bojan Radojicic (University of Novi Sad); Gunjan Verma (US Army's DEVCOM Army Research Laboratory); Ananthram Swami (ARL); Santiago Segarra (Rice University)

584 (SPCN-P2.2): Model-based vs. Data-driven Approaches for Predicting Rain-induced Attenuation in Commercial Microwave Links: A Comparative Empirical Study

Dror Jacoby (Tel Aviv University); Jonatan Ostrometzky (Tel Aviv University); Hagit Messer (Tel Aviv University)

1020 (SPCN-P2.3): EMC²-Net: JOINT EQUALIZATION AND MODULATION CLASSIFICATION BASED ON CONSTELLATION NETWORK

Hyun Ryu (KAIST); Junil Choi (KAIST)

1960 (SPCN-P2.4): A Causal Convolutional Approach for Packet Loss Concealment in Low Powered Devices

Steven Davy (Huawei); Niamh Belton (Science Foundation Ireland Centre for Research Training in Machine Learning, University College Dublin); Joshua Tobin (Huawei); Owais Bin Zuber (Huawei); Liu Dong (Huawei); Yuan Xuewen (Huawei)

3612 (SPCN-P2.5): Boosting Signal Modulation Few-Shot Learning with Pre-transformation

peng sun (Zhejiang University of Technology); Jie Su (Newcastle University); Zhenyu Wen (Zhejiang University of Technology); Yejian Zhou (Zhejiang University of Technology); Zhen Hong (Zhejiang University of Technology); Shanqing Yu (Zhejiang University of Technology); Huaji Zhou (Xidian University)

3775 (SPCN-P2.6): Bit Error and Block Error Rate Training for ML-Assisted Communication

Reinhard Wiesmayr (ETH Zurich); Gian Marti (ETH Zurich); Chris Dick (NVIDIA); Haochuan Song (Southeast University); Christoph Studer (ETH Zurich)

Thursday, June 8

3952 (SPCN-P2.7): Structural Optimization of Factor Graphs for Symbol Detection via Continuous Clustering and Machine Learning

Lukas Rapp (Communications Engineering Lab, Karlsruhe Institute of Technology (KIT)); Luca Schmid (Communications Engineering Lab, Karlsruhe Institute of Technology (KIT)); Andrej Rode (Communications Engineering Lab, Karlsruhe Institute of Technology (KIT)); Laurent Schmalen (Communications Engineering Lab, Karlsruhe Institute of Technology (KIT))

4424 (SPCN-P2.8): DEEP UNFOLDING-ENABLED HYBRID BEAMFORMING DESIGN FOR MMWAVE MASSIVE MIMO SYSTEMS

Nhan Nguyen (University of Oulu (UOULU)); Mengyuan Ma (University of Oulu); Nir Shlezinger (Ben-Gurion University); Yonina Eldar (); Lee Swindlehurst (University of California at Irvine); Markku Juntti (OULU, Finland)

4691 (SPCN-P2.9): Accelerated massive MIMO detector based on annealed underdamped Langevin dynamics

Nicolas M Zilberstein (Rice University); Chris Dick (Nvidia); Rahman Doost-Mohammady (Rice University); Ashutosh Sabharwal (Rice University); Santiago Segarra (Rice University)

4817 (SPCN-P2.10): Machine Learning-Aided Piece-wise Modeling Technique of Power Amplifier for Digital Predistortion

Sri Satish Krishna Chaitanya Bulusu (University of Oulu); Nuutti Tervo (University of Oulu); Praneeth Susarla (University of Oulu); Mikko Sillanpää (University of Oulu); Olli Silven (University of Oulu); Markku Juntti (OULU, Finland); Aarno Pärssinen (University of Oulu)

4926 (SPCN-P2.11): ViT-CAT: Parallel Vision Transformers with Cross Attention Fusion for Popularity Prediction in MEC Networks

Zohreh HajiAkhondi-Meybodi (Concordia University); Arash Mohammadi (Concordia University); Ming Hou (Defence Research and Development Canada (DRDC)); Jamshid Abouei (Yazd University); Konstantinos N Plataniotis (UofT)

5166 (SPCN-P2.12): Deep Spectrum Cartography Using Quantized Measurements

Subash Timilsina (Oregon State University); Sagar Shrestha (Oregon State University); Xiao Fu (Oregon State University)

SPCN-P6: Multi-Antenna and Multi-Carrier Communications

Room: Dome L

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Tareq Al-Naffouri,

5577 (SPCN-P6.1): Sparse Delay Doppler Channel Estimation for OTFS Modulation using 2D MUSIC

Akshay S Bondre (Arizona State University); Christ Richmond (Duke University); Ahmed Alkhateeb (Arizona State University); Nicolo Michelusi (Arizona State University)

6712 (SPCN-P6.2): Symbol-Level Precoding Through the Lens of Zero Forcing and Vector Perturbation (SPS Journal Paper)*

Yatao Liu (The Chinese University of Hong Kong); Mingjie Shao (The Chinese University of Hong Kong, Shandong University); Wing-Kin Ma (The Chinese University of Hong Kong); Qiang Li (University of Electronic Science and Technology of China)

6720 (SPCN-P6.3): AoA Estimation Based on Fourth-Order Coarrays and Cumulants for Mmwave Systems (SPS Journal Paper)*

Yuan-Pei Lin (National Yang Ming Chiao Tung University)

6790 (SPCN-P6.4): Asymptotic Analysis of RLS-Based Digital Precoder With Limited PAPR in Massive MIMO (SPS Journal Paper)*

Xiuxiu Ma (KAUST); Abba Kammoun (KAUST); Ayed M. Alrashdi (University of Ha'il); Tarig Ballal (KAUST); Mohamed-Slim Alouini (King Abdullah University of Science and Technology); Tareq Al-Naffouri (CEMSE, KAUST)

6770 (SPCN-P6.5): DCT-based Air Interface Design for Function Computation (OJSP Paper)*

Marc Martínez Gost (Centre Tecnològic de Telecomunicacions de Catalunya); Ana Pérez-Neira (Centre Tecnològic de Telecomunicacions de Catalunya); Miguel Ángel Lagunas (Centre Tecnològic de Telecomunicacions de Catalunya)

7168 (SPCN-P6.6): Per-Wavelet Equalization for Discrete Wavelet Transform Based Multi-Carrier Modulation Systems

Milan Wils (KU Leuven, ESAT-STADIUS); Mohit Sharma (KU Leuven, ESAT-STADIUS); Marc Moonen (KU Leuven)

7170 (SPCN-P6.7): An Explanation of Deep MIMO Detection from a Perspective of Homotopy Optimization

Mingjie Shao (The Chinese University of Hong Kong, Shandong University); Wing-Kin Ma (The Chinese University of Hong Kong); Junbin Liu (CUHK)

7172 (SPCN-P6.8): Diversity Order Analysis for Quantized Constant Envelope Transmission

Zheyu Wu (Academy of Mathematics and Systems Science); Jiageng Wu (Jilin University); Wei-Kun Chen (Beijing Institute of Technology); Ya-Feng Liu (Chinese Academy of Sciences)

Thursday, June 8

7186 (SPCN-P6.9): Robust Tensor Decomposition for Heterogeneous Beamforming Under Imperfect Channel State Information

Kengo Ando (Constructor University); Koji Ishibashi (The University of Electro-Communications); Giuseppe Abreu (Jacobs University Bremen)

3025 (SPCN-P6.10): Regularized Neural Detection for Millimeter Wave Massive MIMO Communication Systems with One-bit ADCs

Aditya Sant ("University of California, San Diego"); Bhaskar Rao (UC San Diego)

5029 (SPCN-P6.11): Model-Free Learning of Optimal Beamformers for Passive IRS-Assisted Sumrate Maximization

Hassaan Hashmi (Yale University); Spyridon Pougkakiotis (Yale University); Dionysios Kalogerias (Yale University)

SPTM-P5: Signal Filtering, Restoration, Enhancement, and Reconstruction

Room: Dome K

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Mojtaba Soltanalian, Shunsuke Ono

1396 (SPTM-P5.1): ASYMPTOTIC BIAS AND VARIANCE OF KERNEL RIDGE REGRESSION

Victor Solo (University of New South Wales)

1502 (SPTM-P5.2): CyPMLI: WISL-Minimized Unimodular Sequence Design via Power Method-Like Iterations

Arian Eamaz (University of Illinois - Chicago, IL); Farhang Yeganegi (University of Illinois Chicago); Mojtaba Soltanalian (University of Illinois)

1836 (SPTM-P5.3): MAKING SYNCHROSQUEEZING LOCALLY ADAPTIVE IN THE TIME-FREQUENCY PLANE

Marcelo A Colominas (CONICET); Sylvain Meignen (University Grenoble Alpes)

2667 (SPTM-P5.4): WIENER FILTERING WITHOUT COVARIANCE MATRIX INVERSION

Pranav U Damale (Colorado State University); Edwin Chong (Colorado State University); Louis Scharf (Colorado State University)

2908 (SPTM-P5.5): REGRESSION TO CLASSIFICATION: WAVEFORM ENCODING FOR NEURAL FIELD-BASED AUDIO SIGNAL REPRESENTATION

TaeSoo Kim (KT Corporation); Daniel Rho (KT Corporation); GaHui Lee (KT Corporation); JaeHan Park (KT Corporation); Jong Hwan Ko (Sungkyunkwan University)

4875 (SPTM-P5.6): On the primal and dual formulations of the Discrete Mumford-Shah functional

Nelly Pustelnik ()

5115 (SPTM-P5.7): Discriminative Vector Learning with Application To Single Channel Speech Separation

Ha Minh Tan (National Central University); Kai-Wen Liang (Department of Communication Engineering, National Central University); Jia-Ching Wang (National Central University)

5330 (SPTM-P5.8): ITER-SIS: ROBUST UNLIMITED SAMPLING VIA ITERATIVE SIGNAL SIEVING

Ruiming Guo (Imperial College London); Ayush Bhandari (Imperial College London)

5364 (SPTM-P5.9): ENHANCING SPATIO-SPECTRAL REGULARIZATION BY STRUCTURE TENSOR MODELING FOR HYPERSPECTRAL IMAGE DENOISING

Shingo Takemoto (Tokyo Institute of Technology); Shunsuke Ono (Tokyo Institute of Technology)

2663 (SPTM-P5.10): INTERPOLATION FILTER MODEL FOR RAMANUJAN SUBSPACE SIGNALS

Pranav D Kulkarni (California Institute of Technology); Dr.P.P Vaidyanathan (California Institute of Technology)

5589 (SPTM-P5.11): Restoration of Time-varying Graph Signals Using Deep Algorithm Unrolling

Hayate KOJIMA (Department of Electrical Engineering and Computer Science, Tokyo University of Agriculture and Technology); Hikari Noguchi (Tokyo University of Agriculture and Technology); Koki Yamada (Tokyo University of Science); Yuichi Tanaka (Osaka University)

Friday, June 9

CI-L1: Computational Imaging I

Room: Nafsika A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Hassan Mansour,

08:15 AM

532 (CI-L1.01): Minimising Distortion for GAN-based Facial Attribute Manipulation

Mingyu Shao (Dongguan University of Technology); Li Lu (Dongguan University of Technology); Ye Ding (Dongguan University of Technology); Qing Liao (Harbin Institute of Technology (Shenzhen))

08:30 AM

1341 (CI-L1.02): Long Range Imaging Using Multispectral Fusion of RGB and NIR Images

Hao Zhang (Xidian University); Lin Mei (Xidian University); Cheolkon Jung (Xidian University)

08:45 AM

1954 (CI-L1.03): Capturing Cross-Scale Disparity for Stereo Image Super-Resolution

Kun He (University of Electronic Science and Technology of China); Changyu Li (University of Electronic Science and Technology of China); Dongyang Zhang (University of Electronic Science and Technology of China); Jie Shao (University of Electronic Science and Technology of China)

09:00 AM

2519 (CI-L1.04): Single-photon Image Super-resolution via Self-supervised Learning

Yiwei Chen (Zhejiang University); Chen Jiang (Zhejiang University); Yu Pan (Zhejiang University)

09:15 AM

3323 (CI-L1.05): Deep Adaptive Superpixels for Hadamard Single Pixel Imaging in Near-Infrared Spectrum

Brayan Monroy (Universidad Industrial de Santander); Jorge Bacca (Universidad Industrial de Santander); Henry Arguello (Universidad Industrial Santander)

09:30 AM

3393 (CI-L1.06): SINCO: A NOVEL STRUCTURAL REGULARIZER FOR IMAGE COMPRESSION USING IMPLICIT NEURAL REPRESENTATIONS

Harry Gao (Washington University in St. Louis); Weijie Gan (Washington University in St. Louis); Zhixin Sun (Washington University in St. Louis); Ulugbek S. Kamilov (Washington University in St. Louis)

GC-12: ICASSP SP Clarity Challenge: Speech Enhancement for Hearing Aids

Room: Nefeli B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): TBA

08:15 AM

6621 (GC-L12.1): Introduction

Trevor Cox (University of Salford)

08:35 AM

6854 (GC-L12.2): A Multi-stage Low-latency Enhancement System for Hearing Aids

Chengwei Ouyang (Orka Inc.); Kexin Fei (Orka Inc.); Haoshuai Zhou (Orka Inc.); Congxi Lu (Orka Inc.); Linkai Li (Orka Inc.)

08:47 AM

6857 (GC-L12.3): INPLACE CEPSTRAL SPEECH ENHANCEMENT SYSTEM FOR THE ICASSP 2023 CLARITY CHALLENGE

Jinjiang Liu (College of Computer Science, Inner Mongolia University); Xueliang zhang (Inner Mongolia University)

08:59 AM

6887 (GC-L12.4): Multi-Channel Speaker Extraction with Adversarial Training: the Wavlab submission to the Clarity ICASSP 2023 Grand Challenge

Samuele Cornell (Università Politecnica delle Marche); Zhong-Qiu Wang (Carnegie Mellon University); Yoshiki Masuyama (Tokyo Metropolitan University); Shinji Watanabe (Carnegie Mellon University); Manuel Pariente (Pulse Audition); Nobutaka Ono (Tokyo Metropolitan University); Stefano Squartini (Università Politecnica delle Marche)

09:11 AM

6908 (GC-L12.5): A LOW-LATENCY HYBRID MULTI-CHANNEL SPEECH ENHANCEMENT SYSTEM FOR HEARING AIDS

Tong Lei (Nanjing University); Zhongshu Hou (Nanjing University); Yuxiang Hu (Horizon Robotics); Wanyu Yang (Horizon Robotics); Tianchi Sun (Nanjing University); Xiaobin Rong (Nanjing University); Dahan Wang (Nanjing University); Kai Chen (Nanjing University); Jing Lu (Nanjing University)

Friday, June 9

IVMSP-L9: Image and Video Enhancement

Room: Athina

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Akisato Kimura, Anastasios Drosou

08:15 AM

1584 (IVMSP-L9.1): Two-Stage Video De-raining with Spatio-Temporal Fusion and Illumination-Invariant Detail Preservation
Yufeng Tan (South China University of Technology); Youjun Xiang (South China University of Technology); Lei Cai (South China University of Technology); Pengcheng Wang (South China University of Technology); Ying Zhang (South China University of Technology); Yuli Fu (South China University of Technology)

08:30 AM

3206 (IVMSP-L9.2): LP-IOANET: EFFICIENT HIGH RESOLUTION DOCUMENT SHADOW REMOVAL

Kostas Georgiadis (CERTH/ITI); Mehmet Kerim Yücel (Samsung R&D UK); Evangelos Skartados (Centre for Research and Technology, Hellas, Information Technologies Institute); Valia Dimaridou (CERTH-ITI); Anastasios Drosou (CERTH-ITI); Albert Saà-Garriga (Samsung R&D UK); Bruno Manganelli (Samsung Research UK)

08:45 AM

653 (IVMSP-L9.3): RETIFORMER: RETINEX-BASED ENHANCEMENT IN TRANSFORMER FOR LOW-LIGHT IMAGE

Junxiang Ruan (Tsinghua University); Xiangtao Kong (SIAT); Wenqi Huang (China southern power grid); Wenming Yang (Tsinghua University)

09:00 AM

1585 (IVMSP-L9.4): AugTarget Data Augmentation for Infrared Small Target Detection

Shengjia Chen (University of Electronic Science and Technology of China); Jiewen Zhu (UESTC); Luping Ji (UESTC); Hongjun Pan (Sichuan University); Yuhao Xu (Sichuan University)

09:15 AM

5200 (IVMSP-L9.5): Exploring Progressive Hybrid-degraded Image Processing for Homography Estimation

Yijun Lin (University of Chinese Academy of Sciences); Xingzhe Su (Institute of Software Chinese Academy of Sciences); Fengge Wu (Institute of Software Chinese Academy of Sciences); Junsuo Zhao (Science and Technology on Integrated Information System Laboratory Institute of Software Chinese Academy of Sciences)

09:30 AM

940 (IVMSP-L9.6): Deep Quantigraphic Image Enhancement via Comparametric Equations

Xiaomeng Wu (NTT Corporation); Yongqing Sun (NTT, Japan); Akisato Kimura (NTT Communication Science Laboratories)

MLSP-L12: Graph based Learning I

Room: Jupiter

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Luana Ruiz,

08:15 AM

5959 (MLSP-L12.1): Towards a More Stable and General Subgraph Information Bottleneck

Hongzhi Liu (Xi'an Jiaotong University); Kaizhong Zheng (Xi'an Jiaotong University); Shujian Yu (Vrije Universiteit Amsterdam); Badong Chen ("Xi'an Jiaotong University, China")

08:30 AM

1523 (MLSP-L12.2): A Magnetic Framelet-Based Convolutional Neural Network for Directed Graphs

Lequan Lin (The University of Sydney); Junbin Gao (University of Sydney, Australia)

08:45 AM

418 (MLSP-L12.3): Hierarchical Hypergraph Recurrent Attention Network for Temporal Knowledge Graph Reasoning

Jiayan Guo (Peking University); Meiqi Chen (Peking University); Yan Zhang (Peking University); Jianqiang Huang (Meituan); zhiwei liu (meituan)

09:00 AM

5277 (MLSP-L12.4): Global and Nodal Mutual Information Maximization in Heterogeneous Graphs

Costas Mavromatis (University of Minnesota); George Karypis (University of Minnesota, Twin Cities)

09:15 AM

5201 (MLSP-L12.5): Training Graph Neural Networks on Growing Stochastic Graphs

Juan Cervino (University of Pennsylvania); Luana Ruiz (University of Pennsylvania); Alejandro Ribeiro (University of Pennsylvania)

09:30 AM

5510 (MLSP-L12.6): AutoGCF: Personalized Aggregation on Neural Graph Collaborative Filtering

Xiaoyu You (Fudan University); Chi Li (Fudan University); Jianwei Xu (Fudan University); Mi Zhang (Fudan University)

Friday, June 9

MSP-L4: Multimodal Signal Processing and Analysis II

Room: Salon de Roses A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Theodoros Giannakopoulos,

08:15 AM

1586 (MSP-L4.1): MMATR: A lightweight approach for Multimodal Sentiment Analysis based on tensor methods

Panagiotis Koromilas (University of Athens); Mihalios A Nicolaou (The Cyprus Institute); Theodoros Giannakopoulos (NCSR Demokritos); Yannis Panagakis (University of Athens)

08:30 AM

1751 (MSP-L4.2): Class-aware Shared Gaussian Process Dynamic Model

Ryosuke Sawata (Sony Group Corporation / Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

08:45 AM

2464 (MSP-L4.3): INDUCTIVE RELATION PREDICTION FROM RELATIONAL PATHS AND CONTEXT WITH HIERARCHICAL TRANSFORMERS

Jiaang Li (University of Science and Technology of China); Quan Wang (Beijing University of Posts and Telecommunications); Zhendong Mao (University of Science and Technology of China)

09:00 AM

2507 (MSP-L4.4): A Multi-signal Perception Network For Textile Composition Identification

Bo Peng (Fudan University); Liren He (Fudan University); Dong Wu (Fudan University); mingmin Chi (Fudan university); Jintao Chen (Shanghai Fabric Eyes Artificial Intelligence Technology Co., Ltd)

09:15 AM

2823 (MSP-L4.5): Single-branch Network for Multimodal Training

Muhammad Saad Saeed (University of Engineering and Technology); Shah Nawaz (German Electron Synchrotron); Muhammad Haris Khan (Muhammad Bin Zayed University of Artificial Intelligence); Muhammad Zaigham Zaheer (Mohamed bin Zayed University of Artificial Intelligence); Karthik Nandakumar (Mohamed Bin Zayed University of Artificial Intelligence); Mohammad Haroon Yousaf (UET Taxila, Pakistan); Arif Mahmood (Information Technology University)

09:30 AM

2980 (MSP-L4.6): Multimodal Propaganda Detection via Anti-persuasion Prompt Enhanced Contrastive Learning

Jian Cui (Wuhan University of Technology); Lin Li (Wuhan University of Technology); Xin Zhang (Wuhan University of Technology); Jingling Yuan (Wuhan University of Technology)

SLT-L20: Speech recognition- training/adaptation

Room: Delphi

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Bo Li, George Saon

08:15 AM

785 (SLT-L20.1): Stabilising and accelerating light gated recurrent units for automatic speech recognition

Adel Moumen (Avignon University); Titouan Parcollet (Samsung AI Research)

08:30 AM

1616 (SLT-L20.2): Improving Scheduled Sampling for Neural Transducer-based ASR

Takafumi Moriya (NTT Corporation); Takanori Ashihara (NTT Corp.); Hiroshi Sato (NTT Corporation); Kohei Matsuura (NTT); Tomohiro Tanaka (NTT Corporation); Ryo Masumura (NTT Corporation)

08:45 AM

3453 (SLT-L20.3): Cross-Training: A Semi-Supervised Training Scheme for Speech Recognition

Soheil Khorram (Google Inc. USA); Anshuman Tripathi (Google); Jaeyoung Kim (Google); Han Lu (Google Inc. USA); Qian Zhang (Google Inc. USA); Rohit Prabhavalkar (Google); Hasim Sak (Google)

09:00 AM

4313 (SLT-L20.4): Self-Convolution for Automatic Speech Recognition

Tian-Hao Zhang (University of Science and Technology Beijing); Qi Liu (University of Science and Technology Beijing); Xinyuan Qian (USTB); Song-Lu Chen (University of Science and Technology); Feng Chen (EEasy Technology Co. LTD); Xu-Cheng Yin (University of Science and Technology Beijing)

Friday, June 9

09:15 AM

5051 (SLT-L20.5): Efficient Domain Adaptation for Speech Foundation Models

Bo Li (Google); Dongseong Hwang (Google); Zhouyuan Huo (Google); Junwen Bai (Google); Guru Prakash Arumugam (Google LLC); Tara Sainath (Google); Khe C Sim (Google Inc.); Yu Zhang (Google); Wei Han (Google); Trevor Strohman (Google); Françoise Beaufays (Google)

09:30 AM

6379 (SLT-L20.6): I3D: Transformer architectures with input-dependent dynamic depth for speech recognition

Yifan Peng (Carnegie Mellon University); Jaesong Lee (NAVER); Shinji Watanabe (Carnegie Mellon University)

SPCN-L3: Decentralized Wireless Systems and Energy Harvesting

Room: Salon de Roses B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Namrata Vaswani,

08:15 AM

3148 (SPCN-L3.1): COMPARING DECENTRALIZED GRADIENT DESCENT APPROACHES AND GUARANTEES

Shana Moothedath (Iowa State University); Namrata Vaswani (Iowa State University)

08:30 AM

4878 (SPCN-L3.2): DISTRIBUTED GAUSSIAN PROCESS HYPERPARAMETER OPTIMIZATION FOR MULTI-AGENT SYSTEMS

Peiyuan Zhai (Delft University of Technology); Raj Thilak Rajan (Delft university of technology)

08:45 AM

4812 (SPCN-L3.3): Wireless Power Transfer using Chirp Waveforms

Arijit Roy (University of Cyprus); Constantinos Psomas (University of Cyprus); Ioannis Krikidis (University of Cyprus)

09:00 AM

6759 (SPCN-L3.4): State-Augmented Learnable Algorithms for Resource Management in Wireless Networks (SPS Journal Paper)*

Navid Naderializadeh (University of Pennsylvania); Mark Eisen (Intel Corporation); Alejandro Ribeiro (University of Pennsylvania)

09:15 AM

6767 (SPCN-L3.5): Non-Convex Generalized Nash Games for Energy Efficient Power Allocation and Beamforming in mmWave Networks (SPS Journal Paper)*

Wenbo Wang (Kunming University); Amir Leshem (Bar-Ilan University)

SS-L14: Robust Learning and Inference

Room: Nefeli A

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Shaofeng Zou,

08:15 AM

783 (SS-L14.1): Adversarially Robust Fairness-aware Regression

Yulu Jin (University of California, Davis); Lifeng Lai (UC Davis)

08:30 AM

1100 (SS-L14.2): Distributionally Robust Multiclass Classification and Applications in Deep Image Classifiers

Ruidi Chen (Amazon); Boran Hao (Boston University); Ioannis C Paschalidis (Boston University)

08:45 AM

3165 (SS-L14.3): Training Neural networks for sequential change-point detection

Junghwan Lee (Georgia Institute of Technology); Yao Xie (Georgia Tech); Xiuyuan Cheng (Duke University)

09:00 AM

3334 (SS-L14.4): Robust and Parallelizable Tensor Completion based on Tensor Factorization and Maximum Correntropy Criterion

Yicong He (University of Central Florida); George Atia (University of Central Florida)

09:15 AM

3373 (SS-L14.5): ROBUST HYPOTHESIS TESTING WITH MOMENT CONSTRAINED UNCERTAINTY SETS

Akshayaa Magesh (University of Illinois at Urbana-Champaign); Zhongchang Sun (University at Buffalo, the State University of New York); Venugopal V. Veeravalli (University of Illinois at Urbana Champaign); Shaofeng Zou (University at Buffalo, the State University of New York)

Friday, June 9

09:30 AM

3905 (SS-L14.6): LABEL-EFFICIENT AND ROBUST LEARNING FROM MULTIPLE EXPERTS

Bojan Kolosnjaji (Technical University of Munich); Apostolis Zarras (Delft University of Technology)

SS-L15: Signal Processing and Learning over Dynamic Graphs

Room: Nafsika B

Type: Oral

08:15 AM to 09:45 AM

Chair(s): Elvin Isufi, Geert Leus

08:15 AM

2557 (SS-L15.1): LEARNING DYNAMIC GRAPHS UNDER PARTIAL OBSERVABILITY

Michele Cirillo (University of Salerno); Vincenzo Matta (DIEM, University of Salerno); Ali H. Sayed (Ecole Polytechnique Fédérale de Lausanne)

08:30 AM

3321 (SS-L15.2): Dynamic Signed Graph Learning

Abdullah Karaaslanli (Michigan State University); Selin Aviyente (Michigan State University)

08:45 AM

3388 (SS-L15.3): Gaussian process dynamical modeling for adaptive inference over graphs

Qin Lu (University of Minnesota); Konstantinos D. Polyzos (University of Minnesota)

09:00 AM

3987 (SS-L15.4): Online Vector Autoregressive Models over Expanding Graphs

Bishwadeep Das (TU Delft); Elvin Isufi (Tu Delft)

09:15 AM

4681 (SS-L15.5): Dynamic Fair Node Representation Learning

Oyku D Kose (University of California Irvine); Yanning Shen (University of California, Irvine)

AASP-P13: Music Classification and Transcription

Room: Dome A

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Geoffroy Peeters,

455 (AASP-P13.1): Phonation Mode Detection in Singing: a Singer Adapted Model

Yixin Wang (Xi'an Jiaotong University; National University of Singapore); Wei Wei (National University of Singapore); Ye Wang (National University of Singapore)

977 (AASP-P13.2): DiffRoll: Diffusion-based Generative Music Transcription with Unsupervised Pretraining Capability

Kin Wai Cheuk (Singapore University of Technology and Design); ryosuke sawata (Sony); Toshimitsu Uesaka (Sony Group Corporation); Naoki Murata (Sony Group Corporation); Naoya Takahashi (Sony Group); Shusuke Takahashi (Sony Group Corporation); Dorien Herremans (Singapore University of Technology and Design); Yuki Mitsufuji (Sony Group Corporation)

2200 (AASP-P13.3): TAPE: An End-to-End Timbre-Aware Pitch Estimator

Nazif Can Tamer (Universitat Pompeu Fabra); Yigitcan Özer (International Audio Laboratories Erlangen); Meinard Müller (International Audio Laboratories Erlangen); Xavier Serra (Universitat Pompeu Fabra)

2232 (AASP-P13.4): RAT: Radial Attention Transformer for Singing Technique Recognition

Guan-Yuan Chen (National Tsing Hua University); Ya-Fen Yeh (National Tsing Hua University); Von-Wun Soo (nthu)

3765 (AASP-P13.5): A study of audio mixing methods for piano transcription in violin-piano ensembles

Hyemi Kim (KAIST / ETRI); Jiyun Park (KAIST); Taegyun Kwon (KAIST); Dasaem Jeong (Sogang University); Juhan Nam (KAIST)

4936 (AASP-P13.6): Explainable Audio Classification of Playing Techniques with Layer-wise Relevance Propagation

Changhong Wang (LS2N); Vincent Lostanlen (Cornell Lab of Ornithology); Mathieu Lagrange (LS2N)

5316 (AASP-P13.7): FretNet: Continuous-Valued Pitch Contour Streaming for Polyphonic Guitar Tablature Transcription

Frank Cwitkowitz (University of Rochester); Toni Hirvonen (Yousician); Anssi Klapuri (Yousician)

348 (AASP-P13.8): Soft Dynamic Time Warping for Multi-Pitch Estimation and Beyond

Michael Krause (International Audio Laboratories Erlangen); Christof Weiß (University of Würzburg); Meinard Müller (International Audio Laboratories Erlangen)

729 (AASP-P13.9): An Attention-based Approach to Hierarchical Multi-label Music Instrument Classification

Zhi Zhong (Sony Group Corporation); Masato Hirano (Sony Group Corporation); Kazuki Shimada (SONY); Kazuya Tateishi (Sony Group Corporation); Shusuke Takahashi (Sony Group Corporation); Yuki Mitsufuji (Sony Group Corporation)

Friday, June 9

1207 (AASP-P13.10): Frame-Level Multi-Label Playing Technique Detection Using Multi-Scale Network and Self-Attention Mechanism

Dichucheng Li (Fudan University); Mingjin Che (Sichuan Conservatory of Music); Wen wu Meng (Sichuan Conservatory of Music); Yulun Wu (Fudan University); Yi Yu (NII); Fan Xia (Sichuan Conservatory of Music); Wei Li (Fudan University)

4540 (AASP-P13.11): Pre-training strategies using contrastive learning and playlist information for music classification and similarity

Pablo Alonso-Jiménez (Universitat Pompeu Fabra); Xavier Favory (Utopia Music); Hadrien Foroughmand (Utopia Music); Grigoris Bourdalas (Utopia Music); Xavier Serra (Universitat Pompeu Fabra); Thomas Lidy (Utopia Music); Dmitry Bogdanov (Universitat Pompeu Fabra)

5466 (AASP-P13.12): A Phoneme-Informed Neural Network Model for Note-Level Singing Transcription

Sangeon Yong (KAIST); Li Su (Academia Sinica); Juhan Nam (KAIST)

AASP-P14: Music Information Retrieval

Room: Dome B

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Zafar Rafii, Minje Kim

4908 (AASP-P14.1): Contrastive Learning-based Audio to Lyrics Alignment for Multiple Languages

Simon Durand (Spotify); Daniel Stoller (Spotify); Sebastian Ewert (Spotify)

572 (AASP-P14.2): Deep Self-Supervised Hierarchical Metrical Structure Modeling

Junyan Jiang (NYU Shanghai); Gus Xia (New York University Shanghai)

2362 (AASP-P14.3): Self-Transcriber: Few-shot Lyrics Transcription with Self-training

Xiaoxue Gao (National University of Singapore); Xianghu Yue (National University of Singapore); Haizhou Li (The Chinese University of Hong Kong, Shenzhen)

3549 (AASP-P14.4): SingNet: A Real-time Singing Voice Beat and Downbeat Tracking System

Mojtaba Heydari (University of Rochester); Ju-Chiang Wang (TikTok); Zhiyao Duan (University of Rochester)

3819 (AASP-P14.5): Show Me the Instruments: Musical Instrument Retrieval from Mixture Audio

Kyungsu Kim (Seoul National University); Minju Park (Seoul National University); Haesun Joung (Seoul National University); Yunkee Chae (Seoul National University); Yeongbeom Hong (Seoul National University); Seonghyeon Go (Seoul National University); Kyogu Lee (Seoul National University)

6130 (AASP-P14.6): Toward Universal Text-to-Music Retrieval

Seunghoon Doh (KAIST); Minz Won (ByteDance); Keunwoo Choi (Gaudio Lab); Juhan Nam (KAIST)

6141 (AASP-P14.7): Textless Speech-to-Music Retrieval Using Emotion Similarity

Seunghoon Doh (KAIST); Minz Won (ByteDance); Keunwoo Choi (Gaudio Lab); Juhan Nam (KAIST)

3434 (AASP-P14.8): U-BEAT: A MULTI-SCALE BEAT TRACKING MODEL BASED ON WAVE-U-NET

Tian Cheng (National Institute of Advanced Industrial Science and Technology (AIST)); Masataka Goto (National Institute of Advanced Industrial Science and Technology (AIST))

4886 (AASP-P14.9): SIMULTANEOUSLY LEARNING ROBUST AUDIO EMBEDDINGS AND BALANCED HASH CODES FOR QUERY-BY-EXAMPLE

Anup Singh (Ghent University); Kris Demuyne (Ghent University); Vipul Arora (IIT Kanpur)

6381 (AASP-P14.10): ByteCover3: Accurate Cover Song Identification on Short Queries

Xingjian Du (ByteDance); Xia Liang (Bytedance); Zijie Wang (Zhejiang University); Huidong Liang (Oxford University); Bilei Zhu (ByteDance AI Lab); Zejun Ma (Bytedance)

6689 (AASP-P14.11): Grid-Based Decimation for Wavelet Transforms with Stably Invertible Implementation (SPS Journal Paper)*

Nicki Holighaus (Acoustics Research Institute, Austrian Academy of Sciences); Günther Koliander (Acoustics Research Institute, Austrian Academy of Sciences); Clara Hollomey (Acoustics Research Institute, Austrian Academy of Sciences); Friedrich Pillichshammer (Institute of Financial Mathematics and Applied Number Theory, Johannes Kepler University Linz)

BISP-P7: Deep Learning for Medical Image Segmentation

Room: Dome F

Type: Poster

08:15 AM to 09:45 AM

Chair(s): TBA

Friday, June 9

201 (BISP-P7.1): Towards simultaneous segmentation of liver tumors and intrahepatic vessels via cross-attention mechanism

Haopeng Kuang (Fudan University); Dingkang Yang (Fudan University); Shunli Wang (Fudan University); Xiaoying Wang (Zhongshan Hospital, Fudan University); Lihua Zhang (Fudan University)

387 (BISP-P7.2): LDTSF: A LABEL-DECOUPLING TEACHER-STUDENT FRAMEWORK FOR SEMI-SUPERVISED ECHOCARDIOGRAPHY SEGMENTATION

Jiapeng Zhang (University Of Shanghai For Science And Technology); Yongxiong Wang (University of Shanghai for Science and Technology); Zhiqun Pan (University of Shanghai for Science and Technology); Zhenhui Tang (Shanghai Jiao Tong University); Lijun Chen (Shanghai Children's Medical Center); Jinlong Liu (Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University)

538 (BISP-P7.3): IDEAL: Improved DENSE LOCAL Contrastive Learning for Semi-Supervised Medical Image Segmentation

Hritam Basak (Stony Brook University); Soumitri Chattopadhyay (Jadavpur University); Rohit Kundu (University of California, Riverside); Sayan Nag (University of Toronto); Rammohan Mallipeddi (Kyungpook national University)

3055 (BISP-P7.4): Estimation of cardiac fibre direction based on activation maps

Johannes W. de Vries (TU Delft); Miao Sun (TU Delft); Natasja de Groot (Erasmus MC); Richard Hendriks (TU Delft)

3246 (BISP-P7.5): FAN-Net: Fourier-based Adaptive Normalization for Cross-Domain Stroke Lesion Segmentation

Weiyi Yu (Fudan University); Yiming Lei (Fudan University); Hongming Shan (Fudan University)

4663 (BISP-P7.6): Smart Split-Federated Learning Over Noisy Channels for Embryo Image Segmentation

Zahra Hafezi Kafshgari (Simon Fraser University); Ivan Bajic (Simon Fraser University); Parvaneh Saeedi (Simon Fraser University)

5301 (BISP-P7.7): Multi-object Localization and Irrelevant-semantic Separation for Nuclei Segmentation in Histopathology Images

Ya Tang (Xiangtan University); Xiongjun Ye (Department of Urology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021); Xuanya Li (Baidu); Zhineng Chen (School of Computer Science, Fudan University)

5652 (BISP-P7.8): Local-Global Progressive U-Transformers for Accurate Hepatic and Portal Veins Segmentation in Abdominal MR Images

Yu Wu (Xiamen University); Dongfang Shen (Xiamen University); Jiabao Jin (Xiamen University); Guanping Xu (Xiamen University); Yinran Chen (Xiamen University); Xiongbiao Luo (Xiamen University)

5655 (BISP-P7.9): DB-UNet: MLP Based Dual Branch UNet for Accurate Vessel Segmentation in OCTA Images

Chengliang Wang (Chongqing University); Haojian Ning (Chongqing University); Xinrun Chen (Chongqing University); Shiyang Li (Xiamen University)

3019 (BISP-P7.10): Semantic Memory Guided Image Representation for Polyp Segmentation

Zijin Yin (Beijing University of Posts and Telecommunications); Runpu Wei (Beijing University of Posts and Telecommunications); Kongming Liang (Beijing University of Posts and Telecommunications); Yiyang Lin (Beijing University of Posts and Telecommunications); Wei Liu (Beijing University of Posts and Telecommunications); Zhanyu Ma (Beijing University of Posts and Telecommunications); Min Min (The Fifth Medical Center of Chinese PLA General Hospital); Jun Guo (Beijing University of Posts and Telecommunications)

1384 (BISP-P7.11): Coarse-to-Fine Covid-19 Segmentation via Vision-Language Alignment

dandan shan (Xiamen University); Zihan Li (University of Illinois at Urbana-Champaign); Wentao Chen (Beijing University of Posts and Telecommunications); Qingde Li (University of Hull); Jie Tian (); Qingqi Hong (Xiamen University)

2206 (BISP-P7.12): Diabetic Retinopathy Grading with Weakly-supervised Lesion Priors

Junlin Hou (Fudan University); Fan Xiao (Fudan University); Jilan Xu (Fudan University); Rui Feng (Fudan University); Yuejie Zhang (Fudan University); Haidong Zou (Shanghai Eye Diseases Prevention and Treatment Center); Lina Lu (Shanghai Eye Diseases Prevention and Treatment Center); Wenwen Xue (Shanghai Eye Diseases Prevention and Treatment Center)

BISP-P8: Detection and Classification in Medical Imaging

Room: Dome G

Type: Poster

08:15 AM to 09:45 AM

Chair(s): TBA

6083 (BISP-P8.1): Heart Rate Extraction from Abdominal Audio Signals

Jake Stuchbury-Wass (University of Cambridge); Erika Bondareva (University of Cambridge); Kayla-Jade Butkow (University of Cambridge); Sanja sceanovic (NOKIA BELL LABS); Zoran Radivojevic (NOKIA BELL LABS); Cecilia Mascolo (University of Cambridge)

6068 (BISP-P8.2): VISION TRANSFORMER WITH PROGRESSIVE TOKENIZATION FOR CT METAL ARTIFACT REDUCTION

Songwei Zheng (Fuzhou University); Dong Zhang (Fuzhou University); ChunYan Yu (Fuzhou University); Danhong Zhu (Fuzhou University); Longlong Zhu (Fuzhou University); Hao Liu (Fuzhou University); Zhongzheng Huang (Fuzhou University)

Friday, June 9

6034 (BISP-P8.3): Spatio-Temporal Hybrid Fusion of CAE and SWIn Transformers for Lung Cancer Malignancy Prediction
Sadaf Khademi (Concordia University); Shahin Heidarian (Concordia University); Parnian Afshar (Concordia University); Farnoosh Naderkhani (Concordia University); Anastasia Oikonomou (University of Toronto); Konstantinos N Plataniotis (UofT); Arash Mohammadi (Concordia University)

5234 (BISP-P8.4): A New Personalized Efficacy Atlas for Pallidal Deep Brain Stimulation
Xiongbiao Luo (Xiamen University)

5105 (BISP-P8.5): Active selection of source patients in transfer learning for epileptic seizure detection using Riemannian Manifold
Toshiki Orihara (Tokyo University of Agriculture and Technology); Kazi Mahmudul Hassan (Tokyo University of Agriculture and Technology); Toshihisa Tanaka (Tokyo University of Agriculture and Technology)

424 (BISP-P8.6): Cardiac Disease Diagnosis on Imbalanced Electrocardiography Data Through Optimal Transport Augmentation
Jielin Qiu (Carnegie Mellon University); Jiacheng Zhu (Carnegie Mellon University); Mengdi Xu (Carnegie Mellon University); Peide Huang (Carnegie Mellon University); Michael Rosenberg (University of Colorado Denver - Anschutz Medical Campus); Douglas J Weber (Carnegie Mellon University); Emerson Liu (Allegheny General Hospital); DING ZHAO (Carnegie Mellon University)

370 (BISP-P8.7): ViTASD: Robust Vision Transformer Baselines for Autism Spectrum Disorder Facial Diagnosis
Xu Cao (NYU); Wenqian Ye (NYU); Elena Sizikova (FDA); Xue Bai (Shenzhen children's hospital); Megan Coffee (NYU); Hongwu Zeng (Shenzhen Children's Hospital); Jianguo Cao (Shenzhen Children's Hospital)

4439 (BISP-P8.8): TRANSFORMER-BASED MULTI-PROTOTYPE APPROACH FOR DIABETIC MACULAR EDEMA ANALYSIS IN OCT IMAGES
Plácido L. Vidal (University of A Coruña); José Joaquim de Moura Ramos (University of A Coruña); Jorge Novo (University of A Coruña); Marcos Ortega (University of A Coruña); Jaime S Cardoso (INESC Porto, Universidade do Porto)

2298 (BISP-P8.9): ULTRASOUND IMAGE QUALITY CONTROL USING SPEECH-ASSISTED SWITCHABLE CYCLEGAN
Jaeyoung Huh (KAIST); Shujaat Khan (Korea Advanced Institute of Science and Technology (KAIST)); Eun Sun Lee (Chung-Ang University Hospital); Jong Chul Ye (Kim Jaechul Graduate School of AI, KAIST, Korea)

4270 (BISP-P8.10): Multi-Head Feature Pyramid Networks for Breast Mass Detection
Hexiang Zhang (Hebei University of Technology); Zhenghua Xu (Hebei University of Technology); Dan Yao (Hebei University of Technology); Shuo Zhang (Hebei University of Technology); Junyang Chen (Shenzhen Univeristy); Thomas Lukasiewicz (University of Oxford)

4661 (BISP-P8.11): A Meta-GNN approach to personalized seizure detection and classification
Abdellah RAHMANI (École polytechnique fédérale de Lausanne); Arun VENKITARAMAN (École polytechnique fédérale de Lausanne); Pascal Frossard (EPFL)

6028 (BISP-P8.12): A NOVEL TRANSFORMER-BASED PIPELINE FOR LUNG CYTOPATHOLOGICAL WHOLE SLIDE IMAGE CLASSIFICATION
Gaojie Li (Central South University); Qing Liu (Central South University); Haotian Liu (Central South University); Yixiong Liang (Central South University)

IVMSP-P27: Image Coding/Compression

Room: Dome M

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Andre Kaup, Yuan Li

372 (IVMSP-P27.1): Saliency-Driven Hierarchical Learned Image Coding for Machines
Kristian Fischer (Friedrich-Alexander-University Erlangen-Nürnberg); Fabian Brand (Friedrich-Alexander University Erlangen-Nürnberg (FAU)); Christian Blum (Friedrich-Alexander University Erlangen-Nürnberg (FAU)); Andre Kaup (Friedrich-Alexander-Universität Erlangen-Nürnberg)

403 (IVMSP-P27.2): Multistage Spatial Context Models for Learned Image Compression
Fangzheng Lin (Waseda University); Heming Sun (Waseda University, Japan); Jinming Liu (Shanghai Jiao Tong University); Jiro Katto (Waseda University)

707 (IVMSP-P27.3): Learnt Mutual Feature Compression for Machine Vision
Tie Liu (BUAA); Mai Xu (BUAA); Shengxi Li (Beihang University); Chaoran Chen (Beihang University); Li Yang (Beihang university); Zhuoyi Lv (vivo)

1310 (IVMSP-P27.4): A novel Cross-Component Context Model for End-to-End Wavelet Image Coding
Anna Meyer (Friedrich-Alexander-Universität Erlangen-Nürnberg); Andre Kaup (Friedrich-Alexander-Universität Erlangen-Nürnberg)

Friday, June 9

1496 (IVMSP-P27.5): ROI-BASED DEEP IMAGE COMPRESSION WITH SWIN TRANSFORMERS

Binglin Li (Simon Fraser University); Jie Liang (Simon Fraser University); Haisheng Fu (Xi'an Jiaotong University); Jingning Han (Google Inc.)

2929 (IVMSP-P27.6): Residual Hybrid Attention Network for Compression Artifact Reduction

bingchun luo (Harbin Institute of Technology); Wei Yu (Harbin Institute of Technology)

5157 (IVMSP-P27.7): TENSOR LOWRANK COLUMN-WISE COMPRESSIVE SENSING FOR DYNAMIC IMAGING

Silpa Babu (IOWA STATE UNIVERSITY); Selin Aviyente (Michigan State University); Namrata Vaswani (Iowa State University)

6158 (IVMSP-P27.8): SigVIC: Spatial Importance Guided Variable-Rate Image Compression

Jiaming Liang (Beijing Jiaotong University); Meiqin Liu (Beijing Jiaotong University); Chao Yao (University of Science and Technology, Beijing); Chunyu Lin (Beijing Jiaotong University); Yao Zhao (Beijing Jiaotong University)

3470 (IVMSP-P27.9): A NOVEL STATE CONNECTION STRATEGY FOR QUANTUM COMPUTING TO REPRESENT AND COMPRESS DIGITAL IMAGES

MD ERSHADUL HAQUE (Charles Sturt University); Manoranjan Paul (Charles Sturt University, Australia); Anwaar Ulhaq (Charles Sturt University); Tanmoy Debnath (Charles Sturt University, Australia)

MSP-P5: Audio-Visual Signal Processing and Analysis

Room: Dome I

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Emre Eskimez,

822 (MSP-P5.1): AUDIO-DRIVEN HIGH DEFINITION AND LIP-SYNCHRONIZED TALKING FACE GENERATION BASED ON FACE REENACTMENT

Xianyu Wang (Huawei Technologies Co., Ltd.); Yuhan Zhang (Peking University); Weihua He (Tsinghua University); Yaoyuan Wang (Huawei Technologies Co., Ltd.); Minglei Li (Huawei Technologies Co., Ltd.); Yuchen Wang (Huawei Technologies Co., Ltd.); Jingyi Zhang (Huawei Technologies Co., Ltd.); Shunbo Zhou (Huawei Cloud); Ziyang Zhang (HUAWEI TECHNOLOGIES CO.LTD)

1185 (MSP-P5.2): Lip-to-speech Synthesis in the Wild with Multi-task Learning

Minsu Kim (KAIST); Joanna Hong (KAIST); Yong Man Ro (KAIST)

2103 (MSP-P5.3): Multi-Temporal Lip-Audio Memory for Visual Speech Recognition

Jeong Hun Yeo (Korea Advanced Institute of Science and Technology); Minsu Kim (KAIST); Yong Man Ro (KAIST)

2730 (MSP-P5.4): Dual-Path Cross-Modal Attention for better Audio-Visual Speech Extraction

Zhongweiyang Xu (University of Illinois Urbana-Champaign); Xulin Fan (University of Illinois at Urbana-Champaign); Mark Hasegawa-Johnson (University of Illinois)

4345 (MSP-P5.5): AV-SepFormer: Cross-attention SepFormer for Audio-Visual Target Speaker Extraction

Jiuxin Lin (Tsinghua University); Xinyu Cai (Tsinghua University); Heinrich Dinkel (Xiaomi); Jun Chen (Tsinghua University); Zhiyong Yan (xiaomi); Yongqing Wang (xiaomi); Junbo Zhang (Xiaomi); Zhiyong Wu (Tsinghua University); Yujun Wang (xiaomi); Helen Meng (The Chinese University of Hong Kong)

4422 (MSP-P5.6): WL-MSR: Watch and Listen for Multimodal Subtitle Recognition

Jiawei Liu (Institute of Automation, Chinese Academy of Sciences and School of Artificial Intelligence, University of Chinese Academy of Sciences); Hao Wang (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences and School of Artificial Intelligence, University of Chinese Academy of Sciences); Weining Wang (The Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences); Xingjian He (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences and School of Artificial Intelligence, University of Chinese Academy of Sciences); Jing Liu (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences)

5104 (MSP-P5.7): CM-CS : CROSS-MODAL COMMON-SPECIFIC FEATURE LEARNING FOR AUDIO-VISUAL VIDEO PARSING

Hongbo Chen (ShanghaiTech University); Dongchen Zhu (SIMIT); Guanghui Zhang (SIMIT); Wenjun Shi (SIMIT); Xiaolin Zhang (SIMIT); Jiamao Li (SIMIT)

5444 (MSP-P5.8): ModEFormer: Modality-Preserving Embedding for Audio-Video Synchronization using Transformers

Akash Gupta (New York University); Rohun Tripathi (Amazon); Won-Dong Jang (Amazon Studios)

5559 (MSP-P5.9): Audio-driven facial landmark generation in violin performance using 3DCNN network with self attention model

Ting-Wei Lin (Academia Sinica); Chao-Lin Liu (National Chengchi University); Li Su (Academia Sinica)

5564 (MSP-P5.10): REAL-TIME AUDIO-VISUAL END-TO-END SPEECH ENHANCEMENT

Zirun Zhu (Microsoft); Hemin Yang (Microsoft); Min Tang (Microsoft); Ziyi Yang (Microsoft); Sefik Emre Eskimez (Microsoft); Huaming Wang (Microsoft)

Friday, June 9

5772 (MSP-P5.11): Next-speaker Prediction Based on Non-Verbal Information in Multi-party Video Conversation
Saki Mizuno (NTT Computer & Data Science Laboratories); Nobukatsu Hojo (NTT); Satoshi Kobashikawa (NTT Corporation); Ryo Masumura (NTT Corporation)

6163 (MSP-P5.12): The Multimodal Information Based Speech Processing (MISP) 2022 Challenge: Audio-Visual Diarization and Recognition

Zhe Wang (University of Science and Technology of China); Shilong Wu (University of Science and Technology of China); Hang Chen (USTC); Mao-Kui He (University of Science and Technology of China); Jun Du (University of Science and Technology of China); Chin-hui Lee (Georgia Institute of Technology); Shinji Watanabe (Carnegie Mellon University); Sabato M Siniscalchi (Korea University of Enna); Odette Scharenborg (Multimedia Computing Group, Delft University of Technology); Baocai Yin (USTC,iFLYTEK); Jia Pan (iFlytek Research); Cong Liu (iFLYTEK Research)

MSP-P6: Multimodal Signal Processing and Analysis III

Room: Dome J

Type: Poster

08:15 AM to 09:45 AM

Chair(s): TBA

3515 (MSP-P6.1): Unrestricted Anchor Graph based GCN for Incomplete Multi-view Clustering

Liang Zhao (Dalian University of Technology); Zihao Wang (Dalian University of Technology); Yukun Yuan (Dalian University of Technology); Feng Ding (Dalian University of Technology)

1258 (MSP-P6.2): Rethink pair-wise self-supervised cross-modal retrieval from a contrastive learning perspective

Tiantian Gong (Nanjing University of Aeronautics and Astronautics); Junsheng Wang (Nanjing University of Science And Technology); Liyan Zhang (Nanjing University of Aeronautics and Astronautics)

3616 (MSP-P6.3): Multi-view Graph Regularized Deep Autoencoder-like NMF Framework

Liang Zhao (Dalian University of Technology); Zihao Wang (Dalian University of Technology); Ziyue Wang (Dalian University of Technology); Zhikui Chen (Dalian University of Technology)

4532 (MSP-P6.4): Exploiting modality-invariant feature for robust multimodal emotion recognition with missing modalities

Haolin Zuo (Inner Mongolia University); Rui Liu (Inner Mongolia University); Jinming Zhao (Qiyuan Lab); Guanglai Gao (Inner Mongolia University); Haizhou Li (The Chinese University of Hong Kong (Shenzhen))

4567 (MSP-P6.5): FlowGrad: Using Motion for Visual Sound Source Localization

Rajsuryan Singh (Universitat Pompeu Fabra); Pablo Zinemanas (Universitat Pompeu Fabra); Xavier Serra (Universitat Pompeu Fabra); Juan P Bello (New York University); Magdalena Fuentes (New York University)

4648 (MSP-P6.6): Multimodal Dyadic Impression Recognition via Listener Adaptive Cross-Domain Fusion

Yuanchao Li (University of Edinburgh); Peter Bell (University of Edinburgh); Catherine Lai (University of Edinburgh)

5856 (MSP-P6.7): Confidence-based Event-centric Online Video Question Answering on a Newly Constructed ATBS Dataset

Weikai Kong (University of Nottingham Ningbo, China); Shuhong Ye (University of Nottingham Ningbo China); Chenglin Yao (UNNC); Jianfeng Ren (University of Nottingham Ningbo China)

5879 (MSP-P6.8): Imaginary Voice: Face-styled Diffusion Model for Text-to-Speech

Jiyoung Lee (NAVER AI Lab); Joon Son Chung (KAIST); Soo-Whan Chung (Naver Corporation)

6206 (MSP-P6.9): SEMGEO: SEMANTIC KEYWORDS FOR CROSS-VIEW IMAGE GEO-LOCALIZATION

Royston Rodrigues (NEC); Masahiro Tani (NEC)

6332 (MSP-P6.10): Abusive activity detection with multi-modality based on convolutional neural network

Jisoo Kim (Korea Institute of Science and Technology (KIST)); Hyebin Ahn (Korea Institute of Science and Technology (KIST)); Byounghyun Yoo (Korea Institute of Science and Technology (KIST))

6404 (MSP-P6.11): MRML: Multimodal Rumor Detection by Deep Metric Learning

Liwen Peng (National University of Defence Technology); Songlei Jian (National University of Defense Technology); Dongsheng Li (School of Computer Science, National University of Defense Technology); Siqi Shen (Xiamen University)

6478 (MSP-P6.12): IMPROVING THE MODALITY REPRESENTATION WITH MULTI-VIEW CONTRASTIVE LEARNING FOR MULTIMODAL SENTIMENT ANALYSIS

Peipei Liu (School of Cyber Security, University of Chinese Academy of Sciences); Xin Zheng (Henan University); Hong Li (Institute of Information Engineering, Chinese Academy of Sciences); Liu Jie (Institute of Information Engineering, Chinese Academy of Sciences); Yimo Ren (Beijing Haidian); Hongsong Zhu (Institute of information Engineering, CAS); Limin Sun (Institute of Information Engineering, Chinese Academy of Sciences)

Friday, June 9

SLT-P39: Speech Production, Perception, and Psychoacoustics

Room: Dome C

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Yossi Kshet,

486 (SLT-P39.1): EfficientSpeech: An On-Device Text to Speech Model

Rowel O Atienza (University of the Philippines)

5640 (SLT-P39.10): Articulatory Representation Learning Via Joint Factor Analysis and Neural Matrix Factorization

2619 (SLT-P39.2): nVOC-22: A low cost Mel Spectrogram vocoder for mobile devices

Rakesh Iyer (Google Inc)

3227 (SLT-P39.3): STYLE MODELING FOR MULTI-SPEAKER ARTICULATION-TO-SPEECH

Miseul Kim (Yonsei University); Zhenyu Piao (Yonsei University); Jihyun Lee (yonsei university); Hong-Goo Kang (Yonsei University)

3399 (SLT-P39.4): Evidence of Vocal Tract Articulation in Self-Supervised Learning of Speech

Cheol Jun Cho (UC Berkeley); Peter Wu (UC Berkeley); Abdelrahman Mohamed (Meta); Gopala Krishna Anumanchipalli (UC Berkeley)

3992 (SLT-P39.5): Improved acoustic-to-articulatory inversion using representations from pretrained self-supervised learning models

Sathvik Udupa (Indian Institute of Science); Siddarth C (Robert Bosch Centre for Data Science and AI, Indian Institute of Technology Madras); Prasanta Dr Ghosh (Indian Institute of Science (IISc), Bangalore)

4823 (SLT-P39.6): Performance comparison of TTS models for Brazilian Portuguese to establish a baseline

Wilmer Johan Lobato (Alana AI); Felipe Farias (Alana AI); William Cruz (Alana AI); Marcellus Amadeus (Alana AI)

5060 (SLT-P39.7): Speaker-Independent Acoustic-to-Articulatory Speech Inversion

Peter Wu (UC Berkeley); Li-Wei Chen (Carnegie Mellon University); Cheol Jun Cho (UC Berkeley); Shinji Watanabe (Carnegie Mellon University); Louis Goldstein (University of Southern California); Alan Black (CMU); Gopala Krishna Anumanchipalli (UC Berkeley)

5096 (SLT-P39.8): THE SECRET SOURCE : INCORPORATING SOURCE FEATURES TO IMPROVE ACOUSTIC-TO-ARTICULATORY SPEECH INVERSION

Yashish M. Siriwardena (University of Maryland College Park); Carol Y Espy-Wilson (University of Maryland)

5406 (SLT-P39.9): Articulation GAN: Unsupervised modeling of articulatory learning

Gasper Begus (UC Berkeley); Alan Zhou (Johns Hopkins University); Peter Wu (UC Berkeley); Gopala Krishna Anumanchipalli (UC Berkeley)

Jiachen Lian (University of California Berkeley); Alan Black (CMU); Yijing Lu (University of Southern California); Louis Goldstein (USC); Shinji Watanabe (Carnegie Mellon University); Gopala Krishna Anumanchipalli (UC Berkeley)

6012 (SLT-P39.11): PMMSD: DEVELOPMENT OF THE MATRIX SENTENCE INTELLIGIBILITY DATASET FOR MANDARIN WITH LOMBARD EFFECT

Hanchen Pei (Wuhan University); Yuhong Yang (Wuhan University); Xufeng Chen (School of Computer Science, Wuhan University); Qingmu Liu (Wuhan University); Hongyang Chen (Wuhan University); Weiping Tu (Wuhan University); Song Lin (Oppo)

6216 (SLT-P39.12): Efficient Speech Quality Assessment using Self-supervised Framewise Embeddings

Karl El Hajal (EPFL); Zihan Wu (EPFL); Neil Scheidwasser-Clow (University of Copenhagen); Gasser Elbanna (MIT); Milos Cernak (Logitech Europe)

SLT-P40: Speech Recognition: Modeling and Context

Room: Dome D

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Wei Zhou,

1524 (SLT-P40.1): Multi-output RNN-T Joint Networks for Multi-task Learning of {ASR} and Auxiliary Tasks

Weiran Wang (Google); Ding Zhao (Google); Shaojin Ding (Google); Hao Zhang (Google); Shuo-yiin Chang (Google); David Rybach (Google); Tara Sainath (Google); Yanzhang He (Google); Ian McGraw (); Shankar Kumar (Google)

1702 (SLT-P40.2): CONTEXT-AWARE END-TO-END ASR USING SELF-ATTENTIVE EMBEDDING AND TENSOR FUSION

Shuo-yiin Chang (Google); Chao Zhang (Cambridge University); Tara Sainath (Google); Bo Li (Google); Trevor Strohman (Google)

2478 (SLT-P40.3): CUMULATIVE ATTENTION BASED STREAMING TRANSFORMER ASR WITH INTERNAL LANGUAGE MODEL JOINT TRAINING AND RESCORING

Mohan LI (Toshiba Europe Ltd); Cong-Thanh Do (Toshiba Research Europe Ltd.); Rama S Doddipatla (Toshiba Europe LTD)

Friday, June 9

2609 (SLT-P40.4): IMPROVEMENTS TO EMBEDDING-MATCHING ACOUSTIC-TO-WORD ASR USING MULTIPLE-HYPOTHESIS PRONUNCIATION-BASED EMBEDDINGS

Hao Yen (Georgia Institute of Technology); Woojay Jeon (Apple)

2643 (SLT-P40.5): Conformer-based Target-Speaker Automatic Speech Recognition for Single-Channel Audio

Yang Zhang (NVIDIA); Krishna C Puvvada (NVIDIA); Vitaly Lavrukhin (NVIDIA); Boris Ginsburg (NVIDIA)

2700 (SLT-P40.6): Improving CTC-based ASR Models with Gated Interlayer Collaboration

Yuting Yang (NetEase Yidun AI Lab); Yuke Li (NetEase Yidun AI Lab); Binbin Du (NetEase Yidun AI Lab)

3042 (SLT-P40.7): LongFNT: Long-form Speech Recognition with Factorized Neural Transducer

Xun Gong (Shanghai Jiaotong University); Yu Wu (Microsoft Research Asia); Jinyu Li (Microsoft); Shujie Liu (Microsoft Research Asia); Rui Zhao (Microsoft); Xie Chen (Shanghai Jiaotong University); Yanmin Qian (Shanghai Jiao Tong University)

3867 (SLT-P40.8): Enhancing and Adversarial: Improve ASR with Speaker Labels

Wei Zhou (RWTH Aachen University); Haotian Wu (RWTH Aachen University); Jingjing Xu (RWTH i6); Mohammad Zeineldeen (RWTH Aachen University / AppTek); Christoph M. Lüscher (Informatik 6, RWTH Aachen University); Ralf Schlüter (RWTH Aachen University); Hermann Ney (RWTH Aachen University)

4592 (SLT-P40.9): Dynamic Chunk Convolution for Unified Streaming and Non-Streaming Conformer ASR

Xilai Li (Amazon); Goeric Huybrechts (Amazon); Srikanth Ronanki (Amazon); Jeff Farris (Amazon); Sravan Babu Bodapati (Amazon)

5356 (SLT-P40.10): Cross-utterance ASR Rescoring with Graph-based Label Propagation

Srinath Tankasala (The University of Texas at Austin); Long Chen (Amazon); Andreas Stolcke (Amazon); Anirudh Raju (Amazon Alexa); Qianli Deng (Amazon); Chander Chandak (Amazon); Aparna Khare (Amazon); Roland Maas (Amazon Inc.); Venkatesh Ravichandran (Amazon)

6093 (SLT-P40.11): Lattice-free Sequence Discriminative Training for Phoneme-based Neural Transducers

Zijian Yang (Lehrstuhl fuer Informatik 6, RWTH Aachen); Wei Zhou (RWTH Aachen University); Ralf Schlüter (RWTH Aachen University); Hermann Ney (RWTH Aachen University)

6729 (SLT-P40.12): Spelling-Aware Word-Based End-to-End ASR (SPS Journal Paper)*

Ekaterina Egorova (Brno University of Technology); Hari Krishna Vydana (CERENCE INC.); Lukáš Burget (Brno University of Technology); Jan Honza Cernocky (Brno University of Technology)

SLT-P41: Speech Recognition: Self-Supervised Models

Room: Dome E

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Yifan Gong,

797 (SLT-P41.1): Cocktail HuBERT: Generalized Self-Supervised Pre-training for Mixture and Single-Source Speech

Maryam Fazel-Zarandi (Meta); Wei-Ning Hsu (Massachusetts Institute of Technology)

2240 (SLT-P41.2): SELF-SUPERVISED LEARNING WITH BI-LABEL MASKED SPEECH PREDICTION FOR STREAMING MULTI-TALKER SPEECH RECOGNITION

Zili Huang (Johns Hopkins University); Zhuo Chen (Microsoft); Naoyuki Kanda (Microsoft); Jian Wu (Microsoft); Yiming Wang (Microsoft Corporation); Jinyu Li (Microsoft); Takuya Yoshioka (Microsoft); Xiaofei Wang (Microsoft Corp.); Peidong Wang (Microsoft)

1230 (SLT-P41.3): CTCBERT: ADVANCING HIDDEN-UNIT BERT WITH CTC OBJECTIVES

Ruchao Fan (University of California, Los Angeles); Yiming Wang (Microsoft Corporation); Yashesh Gaur (Microsoft); Jinyu Li (Microsoft)

2123 (SLT-P41.4): Exploring Self-supervised Pre-trained ASR Models For Dysarthric and Elderly Speech Recognition

Shujie HU (The Chinese University of Hong Kong); Xurong Xie (Institute of Software, Chinese Academy of Sciences); Zengrui Jin (The Chinese University of Hong Kong); Mengzhe GENG (The Chinese University of Hong Kong); Yi Wang (The Chinese University of Hong Kong); Mingyu Cui (The Chinese University of Hong Kong); Jiajun Deng (The Chinese University of Hong Kong); Xunying Liu (The Chinese University of Hong Kong); Helen Meng (The Chinese University of Hong Kong)

2210 (SLT-P41.5): BECTRA: Transducer-based End-to-End ASR with BERT-Enhanced Encoder

Yosuke Higuchi (Waseda University); Tetsuji Ogawa (Waseda University); Tetsunori Kobayashi (Waseda University); Shinji Watanabe (Carnegie Mellon University)

3033 (SLT-P41.6): UFO2: A unified pre-training framework for online and offline speech recognition

Li Fu (JD); Siqi Li (JD Technology); Qingtao Li (JD Technology); Liping Deng (JD Technology); Fangzhu Li (JD Technology); fan lu (JD); Meng Chen (JD AI); Xiaodong He (JDT)

Friday, June 9

5123 (SLT-P41.7): EURO: ESPnet Unsupervised ASR Open-source Toolkit

Dongji Gao (Johns Hopkins University); Jiatong Shi (Carnegie Mellon University); Shun-Po Chuang (National Taiwan University); Paola Garcia (Johns Hopkins University); Hung-yi Lee (National Taiwan University); Shinji Watanabe (Carnegie Mellon University); Sanjeev Khudanpur (Johns Hopkins University)

5207 (SLT-P41.8): Federated Self-Learning with Weak Supervision for Speech Recognition

Milind M Rao (Amazon); Gopinath Chennupati (Amazon Alexa); Gautam Tiwari (Amazon); Anit Kumar Sahu (Amazon Alexa AI); Anirudh Raju (Amazon Alexa); Ariya Rastrow (Amazon); Jasha Droppo (Amazon)

5342 (SLT-P41.9): Adapting self-supervised models to multi-talker speech recognition using speaker embeddings

Zili Huang (Johns Hopkins University); Desh Raj (Johns Hopkins University); Paola Garcia (Johns Hopkins University); Sanjeev Khudanpur (Johns Hopkins University)

5502 (SLT-P41.10): Avoid Overthinking in Self-Supervised Models for Speech Recognition

Dan Berrebbi (Carnegie Mellon University); Brian Yan (Carnegie Mellon University); Shinji Watanabe (Carnegie Mellon University)

5660 (SLT-P41.11): Structured Pruning of Self-Supervised Pre-trained Models for Speech Recognition and Understanding

Yifan Peng (Carnegie Mellon University); Kwangyoung Kim (ASAPP); Felix Wu (ASAPP); Prashant Sridhar (ASAPP); Shinji Watanabe (Carnegie Mellon University)

6344 (SLT-P41.12): PRACTICE OF THE CONFORMER ENHANCED AUDIO-VISUAL HUBERT ON MANDARIN AND ENGLISH

Xiaoming Ren (OPPO); Chao Li (OPPO); Shenjian Wang (OPPO); Li Biao (oppo)

SPCN-P1: Channel State Estimation

Room: Dome H

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Kfir Cohen,

174 (SPCN-P1.1): ON BIDIRECTIONAL PREESTIMATES AND THEIR APPLICATION TO IDENTIFICATION OF FAST TIME-VARYING SYSTEMS

Maciej Niedzwiecki (); Artur Gancza (Gdansk University of Technology); Lu Shen (School of Physics, Engineering and Technology, University of York); Yuriy Zakharov (School of Physics, Engineering and Technology, University of York)

219 (SPCN-P1.2): A CRITICAL LOOK AT RECENT TRENDS IN COMPRESSION OF CHANNEL STATE INFORMATION

Marcus Valtonen Örnhog (Ericsson Research); Stefan Adalbjörnsson (Ericsson Research); Püren Güler (Ericsson); Mojtaba Mahdavi (Ericsson)

478 (SPCN-P1.3): Downlink Covariance Estimation in URA FDD Massive MIMO systems

Salime Bameri (Carleton University); Khalid Almahrog (Carleton university); Ramy Gohary (Carleton University); Amr El-Keyi (Ericsson); Yahia Ahmed (Ericsson)

918 (SPCN-P1.4): CHANNEL ESTIMATION IN MASSIVE MIMO WITH HEAVY-TAILED NOISE: GAUSSIAN-MIXTURE VERSUS CAUCHY MODELS

Ziya Gülgün (Linköping University); Erik G. Larsson (Nil)

2949 (SPCN-P1.5): COMPRESSIVE CHANNEL ESTIMATION FOR IRS-AIDED MILLIMETER-WAVE SYSTEMS VIA TWO-STAGE LAMP NETWORK

Wen-Chiao Tsai (National Taiwan University); Chi-Wei Chen (National Taiwan University); An-Yeu (Andy) Wu (National Taiwan University)

3048 (SPCN-P1.6): ON THE JOINT ESTIMATION OF PHASE NOISE AND TIME-VARYING CHANNELS FOR OFDM UNDER HIGH-MOBILITY CONDITIONS

Francesco Linsalata (Politecnico di Milano); Nassar Ksairi (Huawei Technologies France)

4497 (SPCN-P1.7): Joint Channel and Direction Estimation for Ground-to-UAV Communications Enabled by A Simultaneous Reflecting and Sensing RIS

Jiguang He (Technology Innovation Institute, 9639 Masdar City, Abu Dhabi); Aymen Fakhreddine (Technology Innovation Institute); George Alexandropoulos (National and Kapodistrian University of Athens)

4664 (SPCN-P1.8): Variational Inference Aided Estimation of Time Varying Channels

Benedikt Böck (Technische Universität München); Michael Baur (Technische Universität München); Valentina Rizzello (Technische Universität München); Wolfgang Utschick (Technische Universität München)

4697 (SPCN-P1.9): Robust Angle Estimation for Hybrid mmWave Systems

Yuan-Pei Lin (National Yang Ming Chiao Tung University); Ting-Ming Yang (MediaTek inc.)

5505 (SPCN-P1.10): Channel Estimation with Tightly-Coupled Antenna Arrays

Bamelak H Tadele (University of Manitoba); Volodymyr Shyianov (University of Manitoba); Faouzi Bellili (University of Manitoba); Amine Mezghani (University of Manitoba)

Friday, June 9

2115 (SPCN-P1.11): ANTENNA IMPEDANCE ESTIMATION IN CORRELATED RAYLEIGH FADING CHANNELS
Shaohan Wu (MediaTek USA); Brian Hughes (North Carolina State University)

SPTM-P7: Signal processing over graphs and networks

Room: Dome K

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Selin Aviyente, Hoi-To Wai

7165 (SPTM-P7.1): Multi-channel Sampling on Graphs and Its Relationship to Graph Filter Banks
Junya Hara (Tokyo University of Agriculture and Technology)

3808 (SPTM-P7.2): Product Graph Learning from Multi-attribute Graph Signals with Inter-layer Coupling
Chenyue Zhang (The Chinese University of Hong Kong); Yiran HE (The Chinese University of Hong Kong); Hoi-To Wai (Chinese University of Hong Kong)

6726 (SPTM-P7.3): Explainability in Graph Data Science: Interpretability, replicability, and reproducibility of community detection (SPS Journal Paper)*
Selin Aviyente (Michigan State University); Abdullah Karaaslanli (Michigan State University)

6728 (SPTM-P7.4): Permutation Entropy for Graph Signals (SPS Journal Paper)*
John Stewart Fabila-Carrasco (The University of Edinburgh); Chao TAN (Tianjin University); Javier Escudero ("University of Edinburgh, UK")

6731 (SPTM-P7.5): Simplicial Convolution Filters (SPS Journal Paper)*
Maosheng Yang (Delft University of Technology)

6742 (SPTM-P7.6): Graph Signal Processing: Dualizing GSP Sampling in the Vertex and Spectral Domains (SPS Journal Paper)*
John Shi (Carnegie Mellon University); José M. F. Moura (Carnegie Mellon University)

6796 (SPTM-P7.7): Learning Sparse Graphs via Majorization-Minimization for Smooth Node Signals (SPS Journal Paper)*
Ghania Fatima (Indian Institute of Technology, Delhi); Aakash Arora (University of Luxembourg); Prabhu Babu (IIT Delhi); Petre Stoica (Uppsala University)

6805 (SPTM-P7.8): Online Inference for Mixture Model of Streaming Graph Signals With Sparse Excitation (SPS Journal Paper)*
Yiran HE (The Chinese University of Hong Kong); Hoi-To Wai (Chinese University of Hong Kong)

6779 (SPTM-P7.9): Decentralized Non-Convex Learning With Linearly Coupled Constraints: Algorithm Designs and Application to Vertical Learning Problem (SPS Journal Paper)*
Jiawei Zhang (The Chinese University of Hong Kong, Shenzhen); Songyang Ge (The Chinese University of Hong Kong, Shenzhen); Tsung-Hui Chang ("The Chinese University of Hong Kong,"); Zhiqian Luo (The Chinese University of Hong Kong, Shenzhen and Shenzhen Research Institute of Big Data)

7164 (SPTM-P7.10): Sampling Rate Offset Estimation and Compensation for Distributed Adaptive Node-Specific Signal Estimation in Wireless Acoustic Sensor Networks
Paul HM Didier (KU Leuven University, ESAT, STADIUS); Toon van Waterschoot (Department of Electrical Engineering (ESAT-STADIUS/ETC)); Simon Doclo (University of Oldenburg); Marc Moonen (KU Leuven)

SPTM-P8: Signal Processing Over Networks II

Room: Dome L

Type: Poster

08:15 AM to 09:45 AM

Chair(s): Sergio Barbarossa, Stefan Vlaski

1028 (SPTM-P8.1): Learning graph Laplacian from intrinsic patterns via Gaussian process
Koshi Watanabe (Hokkaido University); Keisuke Maeda (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

1143 (SPTM-P8.2): Identifying Opinion Influencers over Social Networks
Valentina Shumovskaia (Ecole Polytechnique Fédérale de Lausanne); Mert Kayaalp (Ecole Polytechnique Fédérale de Lausanne); Ali H. Sayed (Ecole Polytechnique Fédérale de Lausanne)

1333 (SPTM-P8.3): Sparse Graph Learning with Spectrum Prior for Deep Graph Convolutional Networks
Jin Zeng (Tongji University); Yang Liu (Peking University); Gene Cheung (York University); Wei Hu (Peking University)

Friday, June 9

1352 (SPTM-P8.4): Asynchronous Social Learning

Mert Cemri (Bilkent University); Virginia Bordignon (EPFL); Mert Kayaalp (Ecole Polytechnique Fédérale de Lausanne); Valentina Shumovskaia (Ecole Polytechnique Fédérale de Lausanne); Ali H. Sayed (Ecole Polytechnique Fédérale de Lausanne)

1631 (SPTM-P8.5): MÖBIUS TOTAL VARIATION FOR DIRECTED ACYCLIC GRAPHS

Vedran Mihal (ETH Zurich); Markus Püschel (ETH Zurich)

2556 (SPTM-P8.6): Windowed Fourier Analysis for Signal Processing on Graph Bundles

T. Mitchell Roddenberry (Rice University); Santiago Segarra (Rice University)

4054 (SPTM-P8.7): Learning Hypergraphs From Signals With Dual Smoothness Prior

Bohan Tang (University of Oxford); Siheng Chen (Shanghai Jiao Tong University, Shanghai AI Laboratory); Xiaowen Dong (University of Oxford)

4249 (SPTM-P8.8): Topological Slepians: Maximally Localized Representations of Signals over Simplicial Complexes

Claudio Battiloro (Sapienza University of Rome); Paolo Di Lorenzo (Sapienza University of Rome); Sergio Barbarossa (Sapienza University of Rome)

5172 (SPTM-P8.9): Robust Network Topologies for Distributed Learning

Chutian Wang (Imperial College London); Stefan Vlaski (Imperial College London)

1009 (SPTM-P8.10): THE ROLE OF MEMORY IN SOCIAL LEARNING WHEN SHARING PARTIAL OPINIONS

Michele Cirillo (University of Salerno); Virginia Bordignon (EPFL); Vincenzo Matta (DIEM, University of Salerno); Ali H. Sayed (Ecole Polytechnique Fédérale de Lausanne)

1045 (SPTM-P8.11): Adaptive Axonal Delays in feedforward spiking neural networks for accurate spoken word recognition

Pengfei SUN (Ghent University); Ehsan Eqlimi (Ghent University); Yansong Chua (China Nanhu Academy of Electronics and Information Technology); Paul Devos (Ghent University); Dick Botteldooren (Ghent University)

ASPS-L1: Applications to Vision, Speech, and Robotics

Room: Salon de Roses B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Brian Telfer,

10:50 AM

6443 (ASPS-L1.01): LMBAO: A Landmark Map for Bundle Adjustment Odometry in LiDAR SLAM

Letian Zhang (Sun Yat-sen University); Jinping Wang (Sun Yat-sen University); Jie Lu (Sun Yat-sen University); Nanjie Chen (Sun Yat-sen University); Xiaojun Tan (Sun Yat-sen University); Duan Zhifei (XPeng Inc)

11:05 AM

1069 (ASPS-L1.02): Residual Squeeze-and-Excitation U-shaped Network for Minutia Extraction in Contactless Fingerprint Images

Anderson Cotrim (Institute of Computing - UNICAMP); Helio Pedrini (Institute of Computing - UNICAMP)

11:20 AM

1603 (ASPS-L1.03): TSPTQ-viT: TWO-SCALED POST-TRAINING QUANTIZATION FOR VISION TRANSFORMER

Yu Shan Tai (National Taiwan University GIEE); Ming Guang Lin (National Taiwan University GIEE); An-Yeu (Andy) Wu (National Taiwan University)

11:35 AM

3925 (ASPS-L1.04): Low-Complexity Low-Rank Approximation SVD for Massive Matrix in Tensor Train Format

Jung-Chun Chi (National Tsing Hua University); Chiao-En Chen (National Chung Hsing University); Yuan-Hao Huang (National Tsing Hua University)

11:50 AM

2043 (ASPS-L1.05): DailyTalk: Spoken Dialogue Dataset for Conversational Text-to-Speech

Keon Lee (KRAFTON, Inc.); Kyumin Park (KAIST); Daeyoung Kim (KAIST)

12:05 PM

3040 (ASPS-L1.06): COOPERATIVE FIVE DEGREES OF FREEDOM MOTION ESTIMATION FOR A SWARM OF AUTONOMOUS VEHICLES

Nikos Piperigkos (University of Patras/ATHENA Research Center); Aris Lalos (Industrial Systems Institute, Athena Research Center); Kostas Berberidis (University of Patras); Christos Anagnostopoulos (Industrial Systems Institute, Athena Research and Innovation Center)

Friday, June 9

GC-13: Person Identification and Relapse Detection from Continuous Recordings of Biosignals

Room: Nefeli B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): TBA

10:50 AM

6637 (GC-L13.1): Introduction

Athanasia Zlatintsi (National Technical Univ. of Athens, Greece); Panagiotis P Filntisis (National Technical University of Athens); Niki Efthymiou (NTUA); Christos Garoufis (National Technical University of Athens); George Retsinas (National Technical University of Athens); Thomas Sounapoglou (Blockchain PC); Ilias Maglogiannis (University of Piraeus); Panayiotis Tsanakas (National Technical University of Athens); Nikolaos Smyrnis (Laboratory of Cognitive Neuroscience and Sensorimotor Control); Petros Maragos (National Technical University of Athens)

11:10 AM

6877 (GC-L13.2): Ensemble and personalized Transformer models for subject identification and relapse detection in e-Prevention Challenge

Salvatore Calcagno (University of Catania); Raffaele Mineo (University of Catania); Daniela Giordano (University of Catania); Concetto Spampinato (University of Catania)

11:22 AM

6900 (GC-L13.3): A person identification system for the ICASSP 2023 e-Prevention challenge

Jinting Wu (Samsung Research China-Beijing (SRC-B)); Mei Tu (Samsung)

11:34 AM

6916 (GC-L13.4): Signal Processing Grand Challenge 2023 - e-Prevention: Sleep Behavior as an Indicator of Relapses in Psychotic Patients

Kleanthis Avramidis (University of Southern California); Kranti Adsul (University of Southern California); Digbalay Bose (University of Southern California); Shrikanth Narayanan (USC)

11:46 AM

6921 (GC-L13.5): REPLAPSE DETECTION IN PATIENTS WITH PSYCHOTIC DISORDERS USING UNSUPERVISED LEARNING ON SMARTWATCH SIGNALS

Salam Hamieh (CEA); Christelle Godin (CEA); vincent heiries (CEA); Hussein Al Osman (University of Ottawa)

11:58 AM

6965 (GC-L13.6): PERSON IDENTIFICATION WITH WEARABLE SENSING USING MISSING FEATURE ENCODING AND MULTI-STAGE MODALITY FUSION

Payal Mohapatra (Northwestern University); Akash Pandey (Northwestern University); Sinan Keten (Northwestern University); Wei Chen (" Northwestern University, UK"); Zhu Qi (Northwestern University)

IFS-L3: Adversarial Machine Learning

Room: Nafsika A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): TBA

10:50 AM

6822 (IFS-L3.1): Comments on "Privacy-Enhanced Federated Learning Against Poisoning Adversaries" (SPS Journal Paper)*

Thomas Schneider (TU Darmstadt); Hossein Yalame (TU Darmstadt); Ajith Suresh (Technical University of Darmstadt)

11:05 AM

6809 (IFS-L3.2): Joint Privacy Enhancement and Quantization in Federated Learning (SPS Journal Paper)*

Natalie Lang (Ben-Gurion University of the Negev); Elad Sofer (Ben-Gurion University of the Negev); Tomer Shaked (Ben-Gurion University of the Negev); Nir Shlezinger (Ben-Gurion University)

11:20 AM

5839 (IFS-L3.3): Going in Style: Audio Backdoors through Stylistic Transformations

Stefanos Koffas (Technical University of Delft); Luca pajola (University of Padova); Stjepan Picek (Delft University of Technology); Mauro Conti (University of Padua)

11:35 AM

2407 (IFS-L3.4): Defense against black-box adversarial attacks via heterogeneous fusion features

Jiahuan Zhang (Hokkaido University); Keisuke Maeda (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

Friday, June 9

11:50 AM

4206 (IFS-L3.5): Towards Adversarially Robust Continual Learning

Tao Bai (Nanyang Technological University); Chen Chen (Sony AI); Lingjuan Lyu (Sony AI); Jun Zhao (Nanyang Technological University); Bihan Wen (Nanyang Technological University)

12:05 PM

5628 (IFS-L3.6): Styx: Adaptive Poisoning Attacks against Byzantine-Robust defenses in Federated Learning

Yuxin Wen (University of Maryland); Jonas A. Geiping (University of Maryland, College Park); Micah Goldblum (University of Maryland); Tom Goldstein (University of Maryland, College Park)

IVMSP-L10: Vision and Language Model

Room: Athina

Type: Oral

10:50 AM to 12:10 PM

Chair(s): Jie Wu, Tao Mei

10:50 AM

872 (IVMSP-L10.1): Nested Attention Network with Graph Filtering for Visual Question and Answering

Jing Lu (China University of Petroleum (East China)); Chunlei Wu (China University Of Petroleum(East China)); Leiquan Wang (UPC); Shaozu Yuan (UPC); Jie Wu (China University Of Petroleum)

11:05 AM

1528 (IVMSP-L10.2): Visual-Aware Text-to-Speech

Mohan Zhou (Harbin Institute of Technology); Yalong Bai (JD AI Research); Wei Zhang (JD AI Research); Ting Yao (JD AI Research); Tiejun Zhao (Harbin Institute of Technology); Tao Mei (AI Research of JD.com)

11:20 AM

1769 (IVMSP-L10.3): Streaming Stroke Classification of Online Handwriting

Jingyu Liu (Institute of Automation of Chinese Academy of Sciences); Yanming Zhang (Institute of Automation of Chinese Academy of Sciences); Fei yin (Institute of Automation of Chinese Academy of Sciences); Cheng-Lin Liu (Institute of Automation of Chinese Academy of Sciences)

11:35 AM

1335 (IVMSP-L10.4): DivCon: Learning Concept Sequences for Semantically Diverse Image Captioning

Yue Zheng (Tsinghua University); Ya-Li Li (Tsinghua University); Shengjin Wang (Tsinghua University)

11:50 AM

1680 (IVMSP-L10.5): Instance-Aware Hierarchical Structured Policy for Prompt learning in Vision-Language Models

Xun Wu (school of software, tsinghua university); Guolong Wang (University of International Business and Economics); Zhaoyuan Liu (Qilu University of Technology (Shandong Academy of Sciences)); Xuan Dang (Tsinghua University); Zheng Qin (Tsinghua University)

12:05 PM

1303 (IVMSP-L10.6): A3S: ADVERSARIAL LEARNING OF SEMANTIC REPRESENTATIONS FOR SCENE-TEXT SPOTTING

Masato Fujitake (Fast accounting co., ltd.)

MLSP-L13: Transfer Learning I

Room: Jupiter

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Thuan Nguyen,

10:50 AM

2612 (MLSP-L13.1): Implicit Bayes Adaptation: A Collaborative Transport Approach

Bo Jiang (North Carolina State University); Hamid Krim (North Carolin. State Univ.); Tianfu Wu (NC State University); Derya Cansever (US Army Research Office)

11:05 AM

238 (MLSP-L13.2): Centroid Distance Distillation for Effective Rehearsal in Continual Learning

Liu Daofeng (Suzhou University of Science and Technology); Fan Lyu (College of Intelligence and Computing, Tianjin University); Linyan Li (Suzhou Institute of Trade & Commerce); Zhenping Xia (Suzhou University of Science and Technology); Fuyuan Hu (Suzhou University of Science and Technology)

11:20 AM

3565 (MLSP-L13.3): Scalable Weight Reparametrization for Efficient Transfer Learning

Byeonggeun Kim (Amazon Alexa AI); Jun-Tae Lee (Qualcomm AI Research); Seunghan Yang (Qualcomm AI Research); Simyung Chang (Qualcomm AI Research)

Friday, June 9

11:35 AM

743 (MLSP-L13.4): Bipartite Graph Convolutional Networks with Adversarial Domain Transfer

Dong Wu (Fudan University); Bin Liang (Fudan University); Xiangjun Liu (Fudan University); Xuan Zang (Fudan University); mingmin Chi (Fudan university)

11:50 AM

2154 (MLSP-L13.5): Cross-Domain Learning with Normalizing Flow

Chi Wang (Queen's University Belfast); Jian Gao (Queen's University Belfast); Yang Hua (Queen's University Belfast); Hui Wang (Queen's University Belfast)

12:05 PM

2391 (MLSP-L13.6): MCKD: Mutually Collaborative Knowledge Distillation for Federated Domain Adaptation and Generalization

Ziwei Niu (Zhejiang University); Hongyi Wang (Zhejiang University); Hao Sun (Zhejiang University); Shuyi Ouyang (Zhejiang University); Yen-Wei Chen (Ritsumeikan University); Lanfen Lin (Zhejiang University)

SLT-L21: TTS: AM and Vocoder I

Room: Delphi

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Tomoki Toda, Nicolas Obin

10:50 AM

2679 (SLT-L21.01): Virtuoso: Massive Multilingual Speech-Text Joint Semi-Supervised Learning for Text-to-Speech

Takaaki Saeki (The University of Tokyo); Heiga Zen (Google); Zhehuai Chen (Google); Nobuyuki Morioka (Google); Yuan Wang (Google); Yu Zhang (Google); Ankur Bapna (Google Research); Andrew Rosenberg (Google LLC); Bhuvana Ramabhadran (Google)

11:05 AM

3440 (SLT-L21.02): Text-to-speech synthesis based on latent variable conversion using diffusion probabilistic model and variational autoencoder

Yusuke Yasuda (Nagoya university); Tomoki Toda (Nagoya University)

11:20 AM

3935 (SLT-L21.03): WAVE-U-NET DISCRIMINATOR: FAST AND LIGHTWEIGHT DISCRIMINATOR FOR GENERATIVE ADVERSARIAL NETWORK-BASED SPEECH SYNTHESIS

Takuhiro Kaneko (NTT Corporation); Hirokazu Kameoka (NTT Communication Science Laboratories, NTT Corporation); Kou Tanaka (NTT corporation); Shogo Seki (NTT Corporation)

11:35 AM

4104 (SLT-L21.04): Autovocoder: Fast Waveform Generation from a Learned Speech Representation using Differentiable Digital Signal Processing

Jacob J Webber (The Centre for Speech Technology Research, University of Edinburgh); Cassia Valentini (University of Edinburgh); Evelyn Williams (University of Edinburgh); Gustav Eje Henter (KTH Royal Institute of Technology); Simon King (University of Edinburgh)

11:50 AM

6085 (SLT-L21.05): PUFFIN: PITCH-SYNCHRONOUS NEURAL WAVEFORM GENERATION FOR FULLBAND SPEECH ON MODEST DEVICES

Oliver Watts (SpeakUnique); Lovisa Wihlborg (SpeakUnique); Cassia Valentini (SpeakUnique)

12:05 PM

6411 (SLT-L21.06): DiffVoice: Text-to-Speech with Latent Diffusion

Zhijun Liu (Shanghai Jiao Tong University); Yiwei Guo (Shanghai Jiao Tong University); Kai Yu (Shanghai Jiao Tong University)

SPED-L1: Signal Processing Education

Room: Salon de Roses A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): TBA

10:50 AM

3212 (SPED-L1.01): EFFECTIVE GRAPH-BASED MODELING OF ARTICULATION TRAITS FOR MISPRONUNCIATION DETECTION AND DIAGNOSIS

Bi-Cheng Yan (National Taiwan Normal University); Hsin-Wei Wang (National Taiwan Normal University); Yi-Cheng Wang (National Taiwan Normal University); Berlin Chen (National Taiwan Normal University)

Friday, June 9

11:05 AM

4731 (SPED-L1.02): StuArt: Individualized Classroom Observation of Students with Automatic Behavior Recognition and Tracking

Huayi Zhou (Shanghai Jiao Tong University); Fei Jiang (East China Normal University); Jiaxin Si (Shanghai Jiao Tong University); Lili Xiong (Chongqing Academy of Science and Technology); Hongtao Lu (Shanghai Jiao Tong University)

11:20 AM

4966 (SPED-L1.03): SIGNAL ANALYSIS-SYNTHESIS USING THE QUANTUM FOURIER TRANSFORM

Aradhita Sharma (Arizona State University); Glen Uehara (Arizona State University); Vivek Narayanaswamy (Arizona State University); Leslie Miller (Arizona State University); Andreas Spanias (ASU)

11:35 AM

5324 (SPED-L1.04): On Designing A 3D Imaging Summer Project for Ontario's High School Students during Covid-19 Pandemic

Fengbo Lan (York University); Gene Cheung (York University); Prabhkirat Arora (York University); Deinabo Richard-Koko (York University); Lisa Cole (York University)

11:50 AM

3480 (SPED-L1.05): Classification of the Cervical Vertebrae Maturation (CVM) stages Using the Tripod Network

Salih Furkan Atici (University of Illinois Chicago); Hongyi Pan (University of Illinois Chicago); Mohammed Elnagar (University of Illinois Chicago); Veerasathpurush Allareddy (University of Illinois Chicago); Omar Suhaym (University of Illinois Chicago); Rashid Ansari (); Ahmet E Cetin (University of Illinois at Chicago)

12:05 PM

250 (SPED-L1.06): OUTLIER-INSENSITIVE KALMAN FILTERING USING NUV PRIORS

Shunit Truzman (University of Haifa); Guy Revach (ETH Zürich); Nir Shlezinger (Ben-Gurion University); Itzik Klein (University of Haifa)

SS-L17: Signal Processing and Systems for Remote Biometrics

Room: Nefeli A

Type: Oral

10:50 AM to 12:20 PM

Chair(s): Yu Rong,

10:50 AM

2669 (SS-L17.1): COUGH DETECTION USING MILLIMETER-WAVE FMCW RADAR

Kawon Han (KAIST); Songcheol Hong (KAIST)

11:05 AM

3921 (SS-L17.2): Wireless sensing for simultaneous human vocal sound and heart sound recognition

yu rong (Arizona State University); Kumar Vijay Mishra (United States DEVCOM Army Research Laboratory); Daniel Bliss (Arizona State University)

11:35 AM

6245 (SS-L17.4): Exploiting CCTV Cameras for Hand Hygiene Recognition in ICU

Weijun Huang (Institute of Basic Medicine and Cancer, China); Jia Huang (The Third People's Hospital of Shenzhen); Guowei Wang (The Third People's Hospital of Shenzhen, China); Hongzhou Lu (Department of Infectious Diseases, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China); Min He (College of Electrical and Information Engineering, Hunan University; Institute of Basic Medicine and Cancer, Chinese Academy of Sciences); Wenjin Wang (Southern University of Science and Technology)

11:50 AM

6452 (SS-L17.5): BENCHMARK OF PHYSIOLOGICAL MODEL BASED AND DEEP LEARNING BASED REMOTE PHOTOPLETHYSMOGRAPHY IN AUTOMOTIVE

Zhiyu Wang (Shandong University of Science and Technology); Xuezhi Yang (Hefei University of Technology); Hongzhou Lu (Department of Infectious Diseases, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China); Caifeng Shan (Shandong Univ. Science & Technology); Wenjin Wang (Southern University of Science and Technology)

SS-L18: Signal Processing for RIS-Enabled Smart Wireless Environments

Room: Nafsika B

Type: Oral

10:50 AM to 12:20 PM

Chair(s): George Alexandropoulos,

10:50 AM

646 (SS-L18.1): Codebook-Based User Tracking in IRS-Assisted mmWave Communication Networks

Moritz Garkisch (Friedrich-Alexander University of Erlangen-Nuremberg); Vahid Jamali (Technical University of Darmstadt); Robert Schober (Friedrich-Alexander University Erlangen-Nuremberg)

Friday, June 9

11:05 AM

1144 (SS-L18.2): Beamforming Optimization in RIS-Aided MIMO Systems Under Multiple-Reflection Effects

Dilki Wijekoon (University of Manitoba); Amine Mezghani (University of Manitoba); Ekram Hossain (University of Manitoba)

11:20 AM

1846 (SS-L18.3): Hybrid RIS-Assisted Interference Mitigation for Spectrum Sharing

Fangzhou Wang (University of California Irvine); Lee Swindlehurst (University of California at Irvine)

11:35 AM

2477 (SS-L18.4): An Efficient Beam-Sharing Algorithm for RIS-aided Simultaneous Wireless Information and Power Transfer Applications

Tran Minh Nguyen (Sungkyunkwan University); Muhammad Miftahul Amri (Sungkyunkwan University); Je Hyeon Park (Sungkyunkwan University); Dong In Kim (Sungkyunkwan University); Kae Won Choi (Sungkyunkwan University)

11:50 AM

3625 (SS-L18.5): Compressed-Sensing-Based 3D Localization with Distributed Passive Reconfigurable Intelligent Surfaces

Jiguang He (Technology Innovation Institute, 9639 Masdar City, Abu Dhabi); Aymen Fakhreddine (Technology Innovation Institute, 9639 Masdar City, Abu Dhabi); Henk Wymeersch (Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden); George Alexandropoulos (National and Kapodistrian University of Athens)

12:05 PM

4022 (SS-L18.6): ENERGY EFFICIENCY MAXIMIZATION IN RIS-AIDED NETWORKS WITH GLOBAL REFLECTION CONSTRAINTS

Robert Fotock (University of Cassino and Southern Lazio); Alessio Zappone (University of Cassino and Southern Lazio); Marco Di Renzo (Université Paris Saclay)

BISP-P9: Multimodal Learning

Room: Dome G

Type: Poster

10:50 AM to 12:20 PM

Chair(s): TBA

3342 (BISP-P9.1): A non-contact SpO2 estimation using video magnification and infrared data

Thomas Stogiannopoulos (DUTH Dept. of Electrical Engineering); Grigorios-Aris Cheimariotis (DUTH Dept. of Electrical Engineering); Nikolaos Mitianoudis (DUTH Dept. of Electrical Engineering)

6780 (BISP-P9.2): EPG2S: Speech Generation and Speech Enhancement based on Electropalatography and Audio Signals using Multimodal Learning (SPS Journal Paper)*

Li-Chin Chen (Academia Sinica); Po-Hsun Chen (National Central University); Richard Tzong-Han Tsai (National Central University); Yu Tsao (Academia Sinica)

5024 (BISP-P9.3): TEXT-TO-ECG: 12-LEAD ELECTROCARDIOGRAM SYNTHESIS CONDITIONED ON CLINICAL TEXT REPORTS

Hyunseung Chung (KAIST); Jiho Kim (KAIST); Joon-myung Kwon (Medical AI); Ki-Hyun Jeon (Seoul National University Bundang Hospital); Min Sung Lee (Medical AI); Edward Choi (KAIST)

5223 (BISP-P9.4): Multimodal microscopy image alignment using spatial and shape information and a branch-and-bound algorithm

Shuonan Chen (Columbia University); Bovey Y Rao (Columbia University); Stephanie Herrlinger (Columbia University); Attila Losonczy (Columbia University); Liam Paninski (Department of Statistics, Columbia University); Erdem Varol (Columbia University)

3058 (BISP-P9.5): BIMODAL FUSION NETWORK FOR BASIC TASTE SENSATION RECOGNITION FROM ELECTROENCEPHALOGRAPHY AND ELECTROMYOGRAPHY

Han Gao (Zhejiang University); Shuo Zhao (Zhejiang university); Huiyan Li (Zhejiang University); Li Liu (Zhejiang University); You Wang (Zhejiang University); Ruifen Hu (Zhejiang University); Jin Zhang (Hunan Normal University); Guang Li (Zhejiang University)

4743 (BISP-P9.6): Robust online multiband drift estimation in electrophysiology data

Charles Windolf (Columbia University); Angelique Paulk (Massachusetts General Hospital); Yoav Kfir (Massachusetts General Hospital); Eric Eric Trautmann (Columbia University); Domokos Meszéna (MGH / Harvard Medical School); William Muñoz (Massachusetts General Hospital); Irene Caprara (Massachusetts General Hospital); Mohsen Jamali (Massachusetts General Hospital); Julien Boussard (Columbia University); Ziv Williams (Massachusetts General Hospital); Sydney Cash (Harvard Medical School); Liam Paninski (Department of Statistics, Columbia University); Erdem Varol (Columbia University)

1295 (BISP-P9.7): Elastic Graph Transformer Networks for EEG-based Emotion Recognition

Wei-Bang Jiang (Shanghai Jiao Tong University); Xu Yan (University of Washington); Wei-Long Zheng (Shanghai Jiao Tong University); Bao-Liang Lu (Shanghai Jiao Tong University)

2097 (BISP-P9.8): Prototype Knowledge Distillation for Medical Segmentation with Missing Modality

Shuai Wang (Tsinghua University); Zipei Yan (The Hong Kong Polytechnic University); Daoan Zhang (Southern University of Science and Technology); Haining Wei (Tsinghua University); Zhongsen Li (Tsinghua University); Rui Li (Tsinghua University)

Friday, June 9

3224 (BISP-P9.9): AUTOMATIC CAMERA POSE ESTIMATION BY KEY-POINT MATCHING OF REFERENCE OBJECTS

Jinchen Zeng (TU Delft); Rick Butler (TU Delft); Benno Hendriks (Philips); John.J van den Dobbelsteen (Delft university of technology); Maarten Van der Elst (Reinier de Graaf Groep); Justin Dauwels (TU Delft)

4674 (BISP-P9.10): RELAPSE PREDICTION FROM LONG-TERM WEARABLE DATA USING SELF-SUPERVISED LEARNING AND SURVIVAL ANALYSIS

Evangelos Fekas (National Technical University of Athens); Athanasia Zlatintsi (National Technical Univ. of Athens, Greece); Panagiotis P Filintisis (National Technical University of Athens); Christos Garoufis (National Technical University of Athens); Niki Efthymiou (NTUA); Petros Maragos (National Technical University of Athens)

IVMSP-P28: Video Coding/Compression

Room: Dome J

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Antonio Ortega, Feng Jiang

4890 (IVMSP-P28.1): Comprehensive Complexity Assessment of Emerging Learned Image Compression on CPU And GPU

Farhad Pakdaman (Tampere University); Moncef Gabbouj (Tampere University)

5694 (IVMSP-P28.2): RATE-DISTORTION OPTIMIZATION WITH ALTERNATIVE REFERENCES FOR UGC VIDEO COMPRESSION

Xin Xiong (University of Southern California); Eduardo Pavez (University of Southern California); Antonio Ortega (University of Southern California); Balu Adsumilli (YouTube/Google)

362 (IVMSP-P28.3): A Flow-Guided Non-Local Alignment Network for Video Compressive Sensing Reconstruction

Chao Zhou (Nanjing University of Posts and Telecommunications); Can Chen (Nanjing University of Posts and Telecommunications); Dengyin Zhang (School of Internet of Things Nanjing University of Posts and Telecommunications Nanjing, China)

4235 (IVMSP-P28.4): HIERARCHICAL INTERACTIVE RECONSTRUCTION NETWORK FOR VIDEO COMPRESSIVE SENSING

Tong Zhang (Harbin Institute of Technology); Wenxue Cui (Harbin Institute of Technology); Chen Hui (Harbin Institute of Technology); Feng Jiang (Harbin Institute of Technology, Harbin)

1514 (IVMSP-P28.5): LEARNED VIDEO CODING WITH MOTION COMPENSATION MIXTURE MODEL

Khanh Quoc Dinh (Samsung Research); Kwang Pyo Choi (Samsung Electronics)

1627 (IVMSP-P28.6): GOP-based Latent Refinement for Learned Video Coding

Mohsen Abdoli (IRT b-com); Gordon Clare (IRT b-com); Felix E Henry (Orange)

2088 (IVMSP-P28.7): A NOVEL MODE SELECTION-BASED FAST INTRA PREDICTION ALGORITHM FOR SPATIAL SHVC

Dayong Wang (Institute of Bioinformatics, Chongqing University of Posts & Telecommunications, Chongqing, China); Yu Sun (University of Central Arkansas); Weisheng Li (Chongqing University of Posts and Telecommunications); Lele Xie (Chongqing University of Posts & Telecommunications); Xin Lu (De Montfort University); Frederic Dufaux (CNRS); Ce Zhu (University of Electronic Science & Technology of China)

2927 (IVMSP-P28.8): RDO CANDIDATE SELECTION FOR MAXIMIZING CODING EFFICIENCY IN A PRACTICAL HEVC ENCODER

Joose Sainio (Tampere University); Alexandre MERCAT (Tampere University); Jarno Vanne (Tampere University)

3031 (IVMSP-P28.9): DISTORTION-AWARE CONVOLUTIONAL NEURAL NETWORK-BASED INTERPOLATION FILTER FOR AVS3

Ying Zhang (Samsung Electronics); liang wen (Samsung Research China-Beijing (SRC-B)); Lihong Wang (Samsung); Yinji Piao (Samsung Electronics); Weijing Shi (Samsung Electronics); Kwang Pyo Choi (Samsung Electronics)

IVMSP-P29: Object Tracking

Room: Dome K

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Shunli Zhang, Wenjing Yang

3296 (IVMSP-P29.1): Efficient Siamese Network for UAV Tracking

Xiaohan Zhang (Dalian University of Technology); Dong Wang (Dalian University of Technology); Xiaohong Ma (Dalian University of Technology)

3589 (IVMSP-P29.2): Single-Particle Tracking by Graph Transformer

Satoshi Kamiya (Meijo university); Kazuhiro Hotta (Meijo University); Taka Aki Tsunoyama (OIST); Akihiro Kusumi (Okinawa Institute of Science and Technology Graduate University)

Friday, June 9

5432 (IVMSP-P29.3): ON DESIGNING LIGHT-WEIGHT OBJECT TRACKERS THROUGH NETWORK PRUNING: USE CNNs OR TRANSFORMERS?

Saksham Aggarwal (IIT (ISM) Dhanbad); Taneesh Gupta (Indian Institute of Technology, Dhanbad); Pawan Kumar Sahu (IIT Dhanbad); Arnab Santosh Chavan (Indian Institute of Technology - Dhanbad); Rishabh Tiwari (Google Research, India); Dilip K Prasad (UiT The Arctic University of Norway); Deepak K Gupta (UiT The Arctic University of Norway)

5574 (IVMSP-P29.4): Enhanced DCF Tracker Regularized by Reliable Sample Construction

Kun Hu (National University of Defense Technology); Mingyu Cao (NUDT); Mengzhu Wang (NUDT); long lan (NUDT); Wenjing Yang (National University of Defense Technology); Huibin Tan (NUDT)

6106 (IVMSP-P29.5): Decomposition, Interaction, Reconstruction Meets Global Context Learning in Visual Tracking

Huibin Tan (NUDT); Kun Hu (National University of Defense Technology); Mingyu Cao (NUDT); Mengzhu Wang (NUDT); liyang xu (National University of Defense Technology); Wenjing Yang (National University of Defense Technology)

217 (IVMSP-P29.6): ENHANCED GM-PHD FILTER FOR REAL TIME SATELLITE MULTI-TARGET TRACKING

Camilo G Aguilar (Inria); Mathias Ortner (Airbus); Josiane Zerubia ()

476 (IVMSP-P29.7): Learning to Reconnect Interrupted Trajectories for Weakly Supervised Multi-Object Tracking

Yu-Lei Li (Xiamen University); Yang Lu (Xiamen University); Jie Li (Xidian University); Hanzi Wang (Xiamen University)

1014 (IVMSP-P29.8): TransLink: Transformer-based Embedding for Tracklets' Global Link

Yanting Zhang (Donghua University); Shunghong Wang (Donghua University); Yuxuan Fan (Donghua University); Gaoang Wang (Zhejiang University); Cairong Yan (Donghua University)

3972 (IVMSP-P29.9): Hierarchical Spatiotemporal Feature Fusion Network for Video Saliency Prediction

Yunzuo Zhang (Shijiazhuang Tiedao University); Tian Zhang (Shijiazhuang Tiedao University); Cunyu Wu (Shijiazhuang Tiedao University); Yuxin Zheng (Shijiazhuang Tiedao University)

IVMSP-P30: Image Generation

Room: Dome M

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Junliang Xing, Nicu Sebe

105 (IVMSP-P30.1): Multispectral image fusion based on super pixel segmentation

Nati Ofir (Kingston University London)

1446 (IVMSP-P30.2): Infrared and visible image fusion by using multi-scale transformation and fractional-order gradient information

Shiwei Wu (Nanjing University of Science and Technology); Kang Zhang (Nanjing University of Science and Technology); Xia Yuan (Nanjing University of Science and Technology); ChunXia Zhao (Nanjing university of science and technology)

2198 (IVMSP-P30.3): TRANSFORMER-BASED DEEP HASHING METHOD FOR MULTI-SCALE FEATURE FUSION

Chao He (Inner Mongolia University); Hongxi Wei (Inner Mongolia University)

2303 (IVMSP-P30.4): IMAGE FUSION VIA SLICE_BASED CONVOLUTIONAL SPARSE REPRESENTATION

Jingchen Xu (Yanshan University); Yali Zhang (Yanshan University); Ze Li (YanShan University); Jinjia Wang (Yanshan University)

2476 (IVMSP-P30.5): HIGH-FREQUENCY TRANSFORMER NETWORK BASED ON WINDOW CROSS-ATTENTION FOR PANSHARPENING

Chengjie Ke (Wuhan University); Hao Liang (Wuhan University); Duidui Li (China Centre for Resources Satellite Data and Application); Xin Tian (Wuhan University)

3551 (IVMSP-P30.6): RCDPT: Radar-Camera fusion Dense Prediction Transformer

Lo Chen-Chou (KU Leuven); Vandewalle Patrick (KU Leuven)

5684 (IVMSP-P30.7): A Deep Fusion Rule for Infrared and Visible Image Fusion: Feature Communication for Importance Assessment

Xuran Lv (Qilu University of Technology (Shandong Academy of Sciences)); Jinyong Cheng (Qilu University of Technology (Shandong Academy of Sciences)); Guohua Lv (Qilu University of Technology (Shandong Academy of Sciences)); Zhonghe Wei (Qilu University of Technology (Shandong Academy of Sciences))

4113 (IVMSP-P30.8): Vehicle View Synthesis by Generative Adversarial Network

Chan-Shuo Hu (National Chung-Cheng University); Sung-Wei Tseng (National Chung Cheng University); Xin-Yun Fan (National Chung Cheng University); Chen-Kuo Chiang (National Chung Cheng University)

5586 (IVMSP-P30.9): Pyramid Spatial Feature Transform And Shared-Offsets Deformable Alignment Based Convolutional Network for HDR Imaging

Junda Liao (Nanjing University); Waseda University); Qin Liu (Nanjing University); Takeshi Ikenaga (Waseda University)

Friday, June 9

2569 (IVMSP-P30.10): ERBNet: An Effective Representation Based Network for Unbiased Scene Graph Generation
Wenxi Ma (Xiamen University); Tianxiang Hou (Xiamen University); Qianji Di (Xiamen University); Zhongang Qi (Tencent); Ying Shan (Tencent); Hanzi Wang (Xiamen University)

110 (IVMSP-P30.11): Learning Generalizable Light Field Networks from Few Images
QIAN LI (INRIA); Franck Multon (INRIA); Adnane Boukhayma (INRIA)

2361 (IVMSP-P30.12): PI-Trans: Parallel-ConvMLP and Implicit-Transformation Based GAN for Cross-View Image Translation
Bin Ren (University of Trento); Hao Tang (ETH Zurich); Yiming Wang (Fondazione Bruno Kessler); Xia Li (ETH Zurich); Wei Wang (EPFL); Nicu Sebe (University of Trento)

SLT-P42: Spoken Language Understanding

Room: Dome C

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Patrick Nguyen,

4868 (SLT-P42.1): End-to-end spoken language understanding using joint CTC loss and self-supervised, pretrained acoustic encoders
Jixuan Wang (Amazon); Martin Radfar (Amazon); Kai Wei (Amazon); Clement Chung (Amazon)

4889 (SLT-P42.2): DISTILL-QUANTIZE-TUNE - LEVERAGING LARGE TEACHERS FOR LOW-FOOTPRINT EFFICIENT MULTILINGUAL NLU ON EDGE
Pegah Kharazmi (Amazon); Zhewei Zhao (Amazon); Clement Chung (Amazon); Samridhi Choudhary (Amazon)

636 (SLT-P42.3): HAG: Hierarchical Attention with Graph Network for Dialogue Act Classification in Conversation
Changzeng Fu (Osaka University); Zhenghan Chen (Peking University); Jiaqi Shi (Osaka University; RIKEN); Bowen Wu (Osaka University); Chaoran Liu (Advanced Telecommunications Research Institute International); Carlos Toshinori Ishi (Advanced Telecommunications Research Institute International); Hiroshi Ishiguro (Osaka University)

2671 (SLT-P42.4): Bridging Speech and Text Pre-trained Models with Unsupervised ASR
Jiatong Shi (Carnegie Mellon University); Chan-Jan Hsu (National Taiwan University); ho lam Chung (National Taiwan University); Dongji Gao (Johns Hopkins University); Paola Garcia (Johns Hopkins University); Shinji Watanabe (Carnegie Mellon University); Ann Lee (Meta, Inc.); Hung-yi Lee (National Taiwan University)

2756 (SLT-P42.5): Auxiliary Pooling Layer For Spoken Language Understanding
Yukun Ma (Alibaba Group); Trung Hieu Nguyen (Alibaba Group); Jinjie Ni (Nanyang Technological University); Wen Wang (Alibaba Group); Qian Chen (Speech Lab, DAMO Academy, Alibaba Group); Chong Zhang (Alibaba Group); Bin Ma ("Alibaba, Singapore R&D Center")

3093 (SLT-P42.6): Joint Modeling for ASR Correction and Dialog State Tracking
Deyuan Wang (Beijing University of Posts and Telecommunications); Tiantian Zhang (Beijing University of Posts and Telecommunications); Caixia Yuan (Beijing University of Posts and Telecommunications); Xiaojie Wang (Beijing University of Posts and Telecommunications)

3462 (SLT-P42.7): T5lephone: Bridging Speech and Text Self-supervised Models for Spoken Language Understanding via Phoneme level T5
Chan-Jan Hsu (National Taiwan University); ho lam Chung (National Taiwan University); Hung-yi Lee (National Taiwan University); Yu Tsao (Academia Sinica)

3815 (SLT-P42.8): A BIDIRECTIONAL JOINT MODEL FOR SPOKEN LANGUAGE UNDERSTANDING
Nguyen Anh Tu (Posts and Telecommunications Institute of Technology); Duong Xuan Hieu (Posts and Telecommunications Institute of Technology); Tu Minh Phuong (Posts and Telecommunications Institute of Technology, Ha Noi, Vietnam); Ngo Xuan Bach (Posts and Telecommunications Institute of Technology, Vietnam)

3864 (SLT-P42.9): Speaker-aware Hierarchical Transformer for Personality Recognition in Multiparty Dialogues
Wenjing Han (South China University of Technology); Yirong Chen (South China University of Technology); Xiaofen Xing (South China University of Technology); Guohua Zhou (iFlytek South China AI Institute(Guangzhou) Co.,Ltd); Xiangmin Xu (South China University of Technology)

3881 (SLT-P42.10): Leveraging Large Text Corpora for End-to-End Speech Summarization
Kohei Matsuura (NTT); Takanori Ashihara (NTT Corp.); Takafumi Moriya (NTT); Tomohiro Tanaka (NTT); Marc Delcroix (NTT); Atsunori Ogawa (NTT Corporation); Ryo Masumura (NTT Corporation)

4819 (SLT-P42.11): End-to-end Spoken Language Understanding with Tree-constrained Pointer Generator
Guangzhi Sun (University of Cambridge Department of Engineering); Chao Zhang (Tsinghua University); Phil Woodland (Machine Intelligence Laboratory, Cambridge University Department of Engineering)

Friday, June 9

4876 (SLT-P42.12): Abstract Representation for Multi-Intent Spoken Language Understanding

Rim Abrougui (Orange Innovation Lannion); Geraldine Damnati (Orange Innovation); Johannes Heinecke (Orange Innovation); FREDERIC BECHET (Aix Marseille University)

SPCN-P3: Optimization and Machine Learning for Communications

Room: Dome H

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Nir Shlezinger,

952 (SPCN-P3.1): INVERSE QUADRATIC TRANSFORM FOR MINIMIZING A SUM OF RATIOS

Yannan CHEN (The Chinese University of Hong Kong, Shenzhen); Licheng Zhao (Shenzhen Research Institute of Big Data); Yaowen Zhang (CUHKSZ); Kaiming Shen (The Chinese University of Hong Kong, Shenzhen)

1256 (SPCN-P3.2): Optimizing distributed multi-sensor multi-target tracking algorithm based on labeled multi-bernoulli filter

Honggang Liu (Fudan University); Jinlong Yang (Jiangnan university); Yue Xu (Jiangnan University); Le Yang (University of Canterbury)

2413 (SPCN-P3.3): Model-Free Online Learning for Waveform Optimization in Integrated Sensing and Communications

Petteri Pulkkinen (Aalto University, Saab Finland Oy); Visa Koivunen (Aalto university)

2438 (SPCN-P3.4): INVERSE REINFORCEMENT LEARNING WITH GRAPH NEURAL NETWORKS FOR IOT RESOURCE ALLOCATION

Guangchen Wang (The University of Sydney); Peng Cheng (La Trobe University); Zhuo Chen (CSIRO); Wei Xiang (La Trobe University); Branka Vucetic (University of Sydney); Yonghui Li (THE UNIVERSITY OF SYDNEY)

3158 (SPCN-P3.5): Anomaly Detection in Optical Spectra via Joint Optimization

Antonino M Rizzo (Politecnico di Milano); Luca Magri (Politecnico di Milano); Pietro Invernizzi (Cisco Photonics); Enrico Sozio (Cisco Photonics); Stefano Piciaccia (Cisco Photonics); Alberto Tanzi (Cisco Photonics); Stefano Binetti (Cisco Photonics); Cesare Alippi (Università della Svizzera Italiana); Giacomo Boracchi (Politecnico di Milano)

3506 (SPCN-P3.6): ITERATIVE WATER-FILLING POWER AND SUBCARRIER ALLOCATION FOR MULTICARRIER NOMA DOWNLINK

Chin Choy Chai (Toronto Metropolitan University (formerly Ryerson University)); Xiao-Ping Steven Zhang (Ryerson University)

3534 (SPCN-P3.7): A Novel Extrapolation Technique to Accelerate WMMSE

Kaiwen Zhou (The Chinese University of Hong Kong); Zhilin Chen (Huawei Noah's Ark Lab); Guochen Liu (Huawei); Zhitang Chen (Huawei Noah's Ark Lab)

4596 (SPCN-P3.8): DEEP-UNFOLDED ADAPTIVE PROJECTED SUBGRADIENT METHOD FOR MIMO DETECTION

Jochen Fink (Fraunhofer HHI); Renato Luis Garrido Cavalcante (Fraunhofer Heinrich Hertz Inst); Zoran Utkovski (Fraunhofer Heinrich hertz Institute); Slawomir Stanczak (Fraunhofer HHI)

5634 (SPCN-P3.9): Learning-based Resource Allocation with Regularization

Shaoke Fang (Peking University); Qingsong Liu (Tsinghua University); Lei Xu (University of Southern California); Wenfei Wu (Peking University)

5199 (SPCN-P3.10): Towards Efficient and Optimal Joint Beamforming and Antenna Selection: A Machine Learning Approach

Sagar Shrestha (Oregon State University); Xiao Fu (Oregon State University); Mingyi Hong (University of Minnesota)

4899 (SPCN-P3.11): Multicast Beamformer Design for MIMO Coded Caching Systems

MohammadJavad Salehi (University of Oulu); Mohammad Naseri Tehrani (University of Oulu); Antti Tölli (University of Oulu)

4825 (SPCN-P3.12): Received Power Maximization with Practical Phase-dependent Amplitude Response in RIS-Aided OFDM Wireless Communications

Dimitris Kompostiotis (University Of Patras); Dimitris Vordonis (University of Patras); Vassilis Paliouras (University of Patras)

SPCN-P4: Multi-Antenna Communications and Intelligent Reflecting Surfaces

Room: Dome I

Type: Poster

10:50 AM to 12:20 PM

Chair(s): George Alexandropoulos,

1417 (SPCN-P4.1): Interference Leakage Minimization in RIS-assisted MIMO Interference Channels

Ignacio Santamaria (University of Cantabria); Mohammad Soleymani (Universität Paderborn); Eduard A Jorswieck (TU Braunschweig); Jesús Gutiérrez (IHP)

Friday, June 9

1704 (SPCN-P4.2): EH-Enabled Distributed Detection Over Temporally Correlated Markovian MIMO Channels

Ghazaleh Ardeshiri (University of central Florida); Azadeh Vosoughi (University of Central Florida)

1808 (SPCN-P4.3): Structure-aware Sparse Bayesian Learning-based Channel Estimation for Intelligent Reflecting Surface-aided MIMO

Yanbin He (Delft University of Technology); Geethu Joseph (TU Delft)

1879 (SPCN-P4.4): Multi-Functional Reconfigurable Intelligent Surface

Wen Wang (BUPT); Wanli Ni (Beijing University of Posts and Telecommunications); Hui Tian (Beijing university of posts and telecommunications); Yonina Eldar ()

1937 (SPCN-P4.5): Scaling Law Analysis for Covariance Based Activity Detection in Cooperative Multi-Cell Massive MIMO

Ziyue Wang (Chinese Academy of Sciences); Ya-Feng Liu (Chinese Academy of Sciences); Zhaorui Wang (The Chinese University of Hong Kong); Wei Yu (University of Toronto)

2036 (SPCN-P4.6): Transceiver Design for MIMO-DFRC Systems

Cai Wen (Northwest University); Timothy N. Davidson (McMaster University)

2312 (SPCN-P4.7): Sparse Aggregation-Based Channel Estimation for Massive MIMO Systems With Decentralized Baseband Processing

Yanqing Xu (The Chinese University of HongKong, Shenzhen); Enbin Song (Sichuan University); Qingjiang Shi (Tongji University); Tsung-Hui Chang ("The Chinese University of Hong Kong,")

3878 (SPCN-P4.8): Frequency-Selective Hybrid Beamforming for mmWave Full-Duplex

Andrea Guamo-Morocho (atlanTTic Research Center, Universidade de Vigo); Roberto Lopez-Valcarce (atlanTTic Research Center, Universidade de Vigo)

3960 (SPCN-P4.9): Capacity Maximization for Active RIS Assisted Outdoor-to-Indoor Communication System

Chen He (Northwest University); GONG WEISHENG (Northwest University); Yangrui Dong (Northwest University); Xie Xie (Northwest University); Z. Jane Wang (University of British Columbia)

3979 (SPCN-P4.10): Joint Microstrip Selection and Beamforming Design for MmWave Systems with Dynamic Metasurface Antennas

Wei Huang (Hefei University of Technology); Haiyang Zhang (Nanjing University of Posts and Telecommunications); Nir Shlezinger (Ben-Gurion University); Yonina Eldar ()

4084 (SPCN-P4.11): Misspecified Cramér-Rao Bound of RIS-aided Localization under Geometry Mismatch

Pinjun Zheng (King Abdullah University of Science and Technology); Hui Chen (Chalmers University of Technology); Tarig Ballal (KAUST); Henk Wymeersch (Chalmers University of Technology); Tareq Al-NAffouri (CEMSE, KAUST)

4343 (SPCN-P4.12): RADIO SENSING WITH LARGE INTELLIGENT SURFACE FOR 6G

Cristian J Vaca Rubio (Aalborg University); Pablo Ramirez Espinosa (University of Granada); Kimmo Kansanen (Norwegian University of Science and Technology); Zheng-Hua Tan (Aalborg University); Elisabeth de Carvalho (Aalborg University)

SPTM-P10: Sparse/Low-Dimensional Signal Processing

Room: Dome L

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Masahiro Yukawa, Julien Flamant

6700 (SPTM-P10.1): Measurement Matrix Design for Sample-Efficient Binary Compressed Sensing (SPS Journal Paper)*

Pulak Sarangi (UCSD); Piya Pal (Nil)

6715 (SPTM-P10.2): On Landscape of Nonconvex Regularized Least Squares for Sparse Support Recovery (SPS Journal Paper)*

HENG QIAO (SHANGHAI JIAO TONG UNIVERSITY); Hongqing Yu (Shanghai Jiao Tong university)

980 (SPTM-P10.3): Sparsity-Smoothness-Aware Power Spectral Density Estimation with Application to Phased Array Weather Radar

Hiroki Kuroda (Nagaoka University of Technology); Daichi Kitahara (Osaka University); Eiichi Yoshikawa (Japan Aerospace Exploration Agency); Hiroshi Kikuchi (The University of Electro-Communications); Tomoo Ushio (Osaka University)

995 (SPTM-P10.4): Sparse representations with cone atoms

Denis C Ilie-Ablachim (University Politehnica of Bucharest); Andra Băltioiu (University Politehnica of Bucharest); Bogdan Dumitrescu (University Politehnica of Bucharest)

3008 (SPTM-P10.5): SPECTRAL SUPER-RESOLUTION ON THE UNIT CIRCLE VIA GRADIENT DESCENT

Xunmeng Wu (Xi'an Jiaotong University); Zai Yang (Xi'an Jiaotong University); Jian-Feng Cai (The Hong Kong University of Science and Technology); Zongben Xu (XJTU)

Friday, June 9

4405 (SPTM-P10.6): Polarized signal singular spectrum analysis with complex SSA

Sébastien Journé (Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble France); Nicolas Le Bihan (Gipsa-lab/CNRS); Florent Chatelain (Gipsa-lab); Julien Flamant (Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France)

4581 (SPTM-P10.7): Deep Unfolded Tensor Robust PCA with Self-supervised Learning

Harry Dong (Carnegie Mellon University); Megna Shah (Air Force Research Laboratory); Sean Donegan (Air Force Research Laboratory); Yuejie Chi (Carnegie Mellon University)

4659 (SPTM-P10.8): SPARSITY-DRIVEN JOINT BLIND DECONVOLUTION-DEMODULATION WITH APPLICATION TO MOTOR FAULT DETECTION

Varun A Kelkar (University of Illinois at Urbana-Champaign); Dehong Liu (Mitsubishi Electric Research Laboratories (MERL)); Hiroshi Inoue (Mitsubishi Electric Corporation); Makoto Kanemaru (Mitsubishi Electric Corporation)

5127 (SPTM-P10.9): HeMPPCAT: Mixtures of Probabilistic Principal Component Analysers for Data with Heteroscedastic Noise

Alec Xu (University of Michigan); Laura Balzano (University of Michigan); Jeffrey A Fessler (University of Michigan)

6007 (SPTM-P10.10): Element Selection with Wide Class of Optimization Criteria Using Non-convex Sparse Optimization

Taiga Kawamura (Tokyo Metropolitan University); Natsuki Ueno (Tokyo Metropolitan University); Nobutaka Ono (Tokyo Metropolitan University)

6378 (SPTM-P10.11): Global Localisation in Continuous Magnetic Vector Fields Using Gaussian Processes

William T McDonald (University of Technology, Sydney); Cedric Le Gentil (University of Technology Sydney); Teresa A. Vidal-Calleja (University of Technology Sydney)

SPTM-P9: Signal Processing Theory and Methods I

Room: Dome F

Type: Poster

10:50 AM to 12:20 PM

Chair(s): David Tay, Tareq Al-Naffouri

6691 (SPTM-P9.1): Spectral Mappings for Graph Wavelets (SPS Journal Paper)*

David Tay (Deakin University)

6823 (SPTM-P9.2): Sliding Short-Time Fractional Fourier Transform (SPS Journal Paper)*

Gaowa Huang (Beijing Institute of Technology); Feng Zhang (Beijing Institute of Technology); Ran Tao (Beijing Institute of Technology)

315 (SPTM-P9.3): IQGAN: Robust Quantum Generative Adversarial Network for Image Synthesis On NISQ Devices

Cheng Chu (Indiana University Bloomington); Grant Skipper (Indiana University Bloomington); Martin Swamy (Indiana University Bloomington); Fan Chen (Indiana University Bloomington)

2792 (SPTM-P9.4): Efficient Data Loading with Quantum Autoencoder

Siang-Ruei Wu (National Taiwan University); Chun-Tse Li (National Taiwan University); Hao-Chung Cheng (National Taiwan University)

5548 (SPTM-P9.5): SIGNAL PROCESSING AND QUANTUM STATE TOMOGRAPHY ON NOISY DEVICES

Wenbo Shi (The University of New South Wales); Robert Malaney (University of New South Wales)

6714 (SPTM-P9.6): NLMS is More Robust to Input-Correlation Than LMS: A Proof (SPS Journal Paper)*

Anum Ali (Samsung Research America); Muhammad Moinuddin (King Abdulaziz University); Tareq Al-Naffouri (CEMSE, KAUST)

5851 (SPTM-P9.7): ROBUST HYPERSPECTRAL ANOMALY DETECTION WITH SIMULTANEOUS MIXED NOISE REMOVAL VIA CONSTRAINED CONVEX OPTIMIZATION

Koyo Sato (Tokyo Institute of Technology); Shunsuke Ono (Tokyo Institute of Technology)

6784 (SPTM-P9.8): Anomaly Search over Discrete Composite Hypotheses in Hierarchical Statistical Models (SPS Journal Paper)*

Tomer Gafni (Ben Gurion University); Benjamin Wolff (Institute for Signal and Information Processing (ISI), D-ITET, ETH Zürich); Guy Revach (ETH Zürich); Nir Shlezinger (Ben-Gurion University); Kobi Cohen (Ben-Gurion University of the Negev)

6739 (SPTM-P9.9): Non-Bayesian Parametric Missing-Mass Estimation (SPS Journal Paper)*

Shir Cohen (Ben Gurion University of the Negev); Tirza S Rountenberg (Ben Gurion University of the Negev); Lang Tong ()

6753 (SPTM-P9.10): Covariance Matrix Estimation Under Positivity Constraints With Application to Portfolio Selection (SPS Journal Paper)*

Ghania Fatima (Indian Institute of Technology, Delhi); Prabhu Babu (IIT Delhi); Petre Stoica (Uppsala University)

6797 (SPTM-P9.11): Gridless DOA Estimation with Multiple Frequencies (SPS Journal Paper)*

Yifan Wu (University of California, San Diego); Peter Gerstoft (University of California San Diego)

Friday, June 9

SS-P2: Radar/Array Signal Processing. Networks and Communications

Room: Dome D

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Shunqiao Sun, Peter Vouras, Brian M Sadler

2374 (SS-P2.1): DOPPLER-CODED JOINT DIVISION MULTIPLE ACCESS WAVEFORM FOR AUTOMOTIVE MIMO RADAR

Yanhua Wang (School of Information and Electronics, Beijing Institute of Technology; Electromagnetic Sensing Research Center of CEMEE State Key Laboratory, Beijing Institute of Technology, Beijing, China); Qiubo Pei (School of Information and Electronics, Beijing Institute of Technology; Chongqing Innovation Center, Beijing Institute of Technology, Chongqing, China); Xueyao Hu (School of Information and Electronics, Beijing Institute of Technology; Chongqing Innovation Center, Beijing Institute of Technology, Chongqing, China); Jiamin Long (School of Information and Electronics, Beijing Institute of Technology; Chongqing Innovation Center, Beijing Institute of Technology, Chongqing, China); Hao Yu (School of Information and Electronics, Beijing Institute of Technology; Chongqing Innovation Center, Beijing Institute of Technology, Chongqing, China); Le Zheng (School of Information and Electronics, Beijing Institute of Technology; Chongqing Innovation Center, Beijing Institute of Technology, Chongqing, China)

3982 (SS-P2.2): Multi-Carrier Wideband OCDM-Based THz Automotive Radar

Sangeeta Bhattacharjee (Indian Institute of Science, Bangalore); Kumar Vijay Mishra (United States DEVCOM Army Research Laboratory); Ramesh Annavaajala (University of Massachusetts Boston); Chandra Murthy (Indian Institute of Science)

5043 (SS-P2.3): Machine learning based early debris detection using automotive low level radar data

Kanishka Tyagi (Aptiv Advance Research Center); Shan Zhang (Aptiv Advance Research Center); Yihang Zhang (Aptiv Advance Research Center); John Kirkwood (Aptiv Advance Research Center); Sanling Song (Aptiv); Narbik Manukian (Aptiv Advance Research Center)

5064 (SS-P2.4): Joint Antenna Selection and Beamforming in Integrated Automotive Radar Sensing-Communications with Quantized Double Phase Shifters

lifan xu (University of Alabama); Shunqiao Sun (The University of Alabama); Yimin D Zhang (Temple University); Athina Petropulu (Rutgers)

2309 (SS-P2.5): Channel State Information-Free Artificial Noise-Aided Location-Privacy Enhancement

Jianxiu Li (University of Southern California); Urbashi Mitra (USC)

6678 (SS-P2.6): Utilization of Bessel Beams in Wideband Sub Terahertz Communication Systems to Mitigate Beamsplit Effects in the Near-field

Arjun Singh (SUNY Polytechnic Institute); Vitaly Petrov (Northeastern University); Josep Jornet (Northeastern University)

6681 (SS-P2.7): NBA-OMP: NEAR-FIELD BEAM-SPLIT-AWARE ORTHOGONAL MATCHING PURSUIT FOR WIDEBAND THZ CHANNEL ESTIMATION

Ahmet M Elbir (University of Luxembourg); Kumar Vijay Mishra (United States DEVCOM Army Research Laboratory); Symeon Chatzinotas (University of Luxembourg)

3446 (SS-P2.8): Phase Retrieval for Rydberg Quantum Arrays

Peter Vouras (U.S Department of Defense); Kumar Vijay Mishra (United States DEVCOM Army Research Laboratory); Alexandra Artusio-Glimpse (National Institute of Standards and Technology)

4418 (SS-P2.9): Compressive estimation of near field channels for ultra massive-MIMO wideband THz systems

Simon Tarboush (Independent Researcher); Anum Ali (Samsung Research America); Tareq Al-NAffouri (CEMSE, KAUST)

6483 (SS-P2.10): Generative Model Based Highly Efficient Semantic Communication Approach for Image Transmission

TIANXIAO HAN (Zhejiang University); Jiancheng Tang (Zhejiang University); Qianqian Yang (Zhejiang University); Yiping Duan (Tsinghua University); Zhaoyang Zhang (Zhejiang University); Zhiguo Shi (Zhejiang University)

312 (SS-P2.11): Semi-Federated Learning for Edge Intelligence with Imperfect SIC

Wanli Ni (Beijing University of Posts and Telecommunications); Jingheng Zheng (Beijing University of Posts and Telecommunications); Yonina Eldar (); Changsheng You (Southern University of Science and Technology); Kaibin Huang (University of Hong Kong)

2250 (SS-P2.12): ENSEMBLE GRAPH Q-LEARNING FOR LARGE SCALE NETWORKS

Talha Bozkus (University of Southern California); Urbashi Mitra (USC)

SS-P3: Neural speech and audio coding: emerging challenges and opportunities

Room: Dome E

Type: Poster

10:50 AM to 12:20 PM

Chair(s): Minje Kim, Jan Skoglund

929 (SS-P3.1): AudioDec: An Open-source Streaming High-fidelity Neural Audio Codec

Yi-Chiao Wu (META); Israel Dejene Gebru (Reality Labs Research); Dejan Markovic (META); Alexander Richard (META)

Friday, June 9

3436 (SS-P3.2): Low-Bitrate Redundancy Coding of Speech Using a Rate-Distortion-Optimized Variational Autoencoder
Jean-Marc Valin (Amazon); Jan Bütthe (Amazon); Ahmed Mustafa (Amazon)

3491 (SS-P3.3): High Quality Audio Coding with MDCTNet

Grant Davidson (Dolby Laboratories); Mark Vinton (Dolby Laboratories); Per Ekstrand (Dolby Sweden AB); Cong Zhou (Dolby Laboratories); Lars F Villemoes (Dolby Sweden AB); Lie Lu (Dolby Laboratories)

3543 (SS-P3.4): END-TO-END NEURAL AUDIO CODING IN THE MDCT DOMAIN

Hyungseob Lim (Yonsei University); Jihyun Lee (yonsei university); Byeong Hyeon Kim (Yonsei University); Inseon Jang (Electronics and Telecommunications Research Institution); Hong-Goo Kang (Yonsei University)

3657 (SS-P3.5): A Perceptual Neural Audio Coder with A Mean-Scale Hyperprior

Joon Byun (Yonsei University); Seungmin Shin (Yonsei University); Young-Cheol Park (Yonsei University); Jongmo Sung (ETRI); Seung-Kwon Beack (IEEE Broadcast Technology Society (BTS))

4687 (SS-P3.6): DISENTANGLED FEATURE LEARNING FOR REAL-TIME NEURAL SPEECH CODING

Xue Jiang (Communication University of China); Xiulian Peng (Microsoft Research Asia); Yuan Zhang (Communication University of China); Yan Lu (Microsoft Research Asia)

4715 (SS-P3.7): LMCCodec: A Low Bitrate Speech Codec With Causal Transformer Models

Teerapat Jenrungrot (University of Washington); Michael Chinen (Google); W. Bastiaan Kleijn (Google); Jan Skoglund (Google); Zalán Borsos (Google); Neil Zeghidour (Google); Marco Tagliasacchi (Google)

4906 (SS-P3.8): Native Multi-Band Audio Coding within Hyper-Autoencoded Reconstruction Propagation Networks

Darius Petermann (Indiana University - Bloomington); Inseon Jang (Electronics and Telecommunications Research Institution); Minje Kim (Indiana University)

5088 (SS-P3.9): MULTI-CHANNEL AUDIO SIGNAL GENERATION

W. Bastiaan Kleijn (Google); Michael Chinen (Google); Felicia S. C. Lim (Google); Jan Skoglund (Google)

5161 (SS-P3.10): Neural Feature Predictor and Discriminative Residual Coding for Low-Bitrate Speech Coding

Haici Yang (Indiana University); Wootae Lim (ETRI); Minje Kim (Indiana University)

5911 (SS-P3.11): Disentangling speech from surroundings with neural embeddings

Ahmed Omran (Google); Neil Zeghidour (Google); Zalán Borsos (Google); Félix de Chaumont Quitry (Google); Malcolm Slaney (Google); Marco Tagliasacchi (Google)

5425 (SS-P3.12): Progressive Multi-stage Neural Audio Codec with Psychoacoustic Loss and Discriminator

Byeong Hyeon Kim (Yonsei University); Hyungseob Lim (Yonsei University); Jihyun Lee (yonsei university); Inseon Jang (Electronics and Telecommunications Research Institution); Hong-Goo Kang (Yonsei University)

ASPS-L2.1: Applications to Communications

Room: Salon de Roses B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Akihiko Sugiyama,

02:00 PM

328 (ASPS-L2.1): Hardware-limited Non-uniform Task-based Quantizers

Neil Irwin M Bernardo (University of Melbourne); Jingge Zhu (University of Melbourne); Yonina Eldar (); Jamie S Evans (University of Melbourne)

02:15 PM

3293 (ASPS-L2.2): Single-anchor UWB Localization using Channel Impulse Response Distributions

Sitian Li (EPFL); Alexios Balatsoukas-Stimming (Eindhoven University of Technology); Andreas Burg (EPFL)

02:30 PM

1724 (ASPS-L2.3): Adaptive Noise Canceller Algorithm with SNR-Based Step Size and Data-Dependent Averaging

Akihiko K. Sugiyama (Yahoo Japan Corporation)

02:45 PM

4081 (ASPS-L2.4): RIS REFLECTION AND PLACEMENT OPTIMISATION FOR UNDERLAY D2D COMMUNICATIONS IN COGNITIVE CELLULAR NETWORKS

Sarbani Ghose (DSZ Innovation Labs Private Limited); Deepak Mishra (University of New South Wales); Santi P. Maity (Indian Institution of Engineering Science and Technology); George Alexandropoulos (National and Kapodistrian University of Athens)

Friday, June 9

03:00 PM

4985 (ASPS-L2.5): Boosting the Accuracy of SRAM-Based In-Memory Architectures via Maximum Likelihood-based Error Compensation Methods

Hyungyo Kim (University of Illinois at Urbana-Champaign); Naresh Shanbhag (University of Illinois at Urbana-Champaign)

03:15 PM

5968 (ASPS-L2.6): CANCELLING INTERMODULATION DISTORTIONS FOR OTOACOUSTIC EMISSION MEASUREMENTS WITH EARBUDS

Berken Utku Demirel (Nokia Bell Labs); Khaldoon T Al-Naimi (Nokia Bell Labs); Fahim Kawsar (Nokia Bell Labs); Alessandro Montanari (Nokia Bell Labs)

GC-14: The First Pathloss Radio Map Prediction Challenge

Room: Nefeli B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): TBA

02:00 PM

6658 (GC-L14.1): Introduction

Cagkan Yapar (TU Berlin); Fabian Jaensch (TU Berlin); Ron Levie (Technion); Gitta Kutyniok (Ludwig Maximilian University of Munich); Giuseppe Caire (TU Berlin)

02:20 PM

6875 (GC-L14.2): DEEP LEARNING-BASED PATH LOSS PREDICTION FOR OUTDOOR WIRELESS COMMUNICATION SYSTEMS

Kehai Qiu (University of Cambridge); Stefanos Bakirtzis (University of Cambridge); Hui Song (Ranplan Wireless Network Design Ltd); Ian J Wassell (University of Cambridge); Jie Zhang (University of Sheffield)

02:32 PM

6883 (GC-L14.3): Agile Radio Map Prediction Using Deep Learning

Enes Krijestorac (University of California, Los Angeles); Hazem Sallouha (KU Leuven); Shamik Sarkar (University of California, Los Angeles); Danijela Cabric (University of California, Los Angeles)

02:44 PM

6891 (GC-L14.4): PMNet: Large-Scale Channel Prediction System for ICASSP 2023 First Pathloss Radio Map Prediction Challenge

Ju-Hyung Lee (University of Southern California); Joochan Lee (University of Southern California); Seon-Ho Lee (MCL, Korea University); Andreas Molisch (University of Southern California)

IVMSP-L11: Human Video Generation and Editing

Room: Jupiter

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Shiliang Pu, Junping Zhang

02:00 PM

490 (IVMSP-L11.01): PRIME: 3D Human Pose and Body Shape Recovery with Perspective Projection

Baobei Xu (Hikvision Research Institute); Shukai Fang (Hikvision Research Institute); Zhaoyang Li (Hikvision Research Institute); Shicai Yang (Hikvision Research Institute); Di Xie (Hikvision Research Institute); Shiliang Pu (Hikvision Research Institute)

02:15 PM

309 (IVMSP-L11.02): OPT: One-shot Pose-Controllable Talking Head Generation

Jin Liu (1. Institute of Information Engineering, Chinese Academy of Sciences. 2. School of Cyber Security, University of Chinese Academy of Sciences); Xi Wang (Institute of Information Engineering, Chinese Academy of Sciences); Xiaomeng Fu (1. Institute of Information Engineering, Chinese Academy of Sciences. 2. School of Cyber Security, University of Chinese Academy of Sciences); chai yesheng (Institute of Information Engineering, Chinese Academy of Sciences); Cai Yu (1. Institute of Information Engineering, Chinese Academy of Sciences. 2. School of Cyber Security, University of Chinese Academy of Sciences); Jiao Dai (Institute of Information Engineering, Chinese Academy of Sciences); Jizhong Han (Institute of Information Engineering, Chinese Academy of Sciences)

02:30 PM

1800 (IVMSP-L11.03): CRFAST: CLIP-BASED REFERENCE-GUIDED FACIAL IMAGE SEMANTIC TRANSFER

Ailin Li (College of Computer Science and Technology, Zhejiang University); Lei Zhao (Zhejiang University); Zhizhong Wang (Zhejiang University); Zhiwen Zuo (Zhejiang University); Wei Xing (Zhejiang University); Dongming Lu (Zhejiang University)

02:45 PM

2847 (IVMSP-L11.04): End-to-End Unsupervised sketch to image generation

Xingming Lv (Shandong University); Lei Wu (Shandong University); Zhenwei Cheng (Shandong University); Xiangxu Meng (Shandong University)

Friday, June 9

03:00 PM

4168 (IVMSP-L11.05): DO-FAM: Disentangled Non-Linear Latent Navigation for Facial Attribute Manipulation
Yifan Yuan (Fudan University); Siteng Ma (Fudan University); Hongming Shan (Fudan University); Junping Zhang (Fudan University)

03:15 PM

6223 (IVMSP-L11.06): Free-view Expressive Talking Head Video Editing
Yuantian Huang (University of Tsukuba); Satoshi Iizuka (University of Tsukuba); Kazuhiro Fukui (University of Tsukuba)

IVMSP-L12: Point Cloud Processing

Room: Athina

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Fu Lee Wang, Xudong Jiang

02:00 PM

2318 (IVMSP-L12.1): GEOGCN: GEOMETRIC DUAL-DOMAIN GRAPH CONVOLUTION NETWORK FOR POINT CLOUD DENOISING

ZhaoWei Chen (Nanjing University of Aeronautics and Astronautics); Peng Li (Nanjing University of Aeronautics and Astronautics); Zeyong Wei (Nanjing University of Aeronautics and Astronautics); Honghua Chen (Nanyang Technological University); Haoran Xie (Lingnan University); Mingqiang Wei (Nanjing University of Aeronautics and Astronautics); Fu Lee Wang (Hong Kong Metropolitan University)

02:15 PM

2765 (IVMSP-L12.2): Face Recognition on Point Cloud with cGAN-TOP for Denoising

Junyu Liu (University of Nottingham Ningbo China); Jianfeng Ren (University of Nottingham Ningbo China); Hong-liang Sun (UNNC); Xudong Jiang (Nanyang Technological University)

02:30 PM

5095 (IVMSP-L12.3): GRAPH WAVELET-BASED POINT CLOUD GEOMETRIC DENOISING WITH SURFACE-CONSISTENT NON-NEGATIVE KERNEL REGRESSION

Ryosuke Watanabe (KDDI Research, Inc.); Keisuke Nonaka (KDDI Research Inc.); Eduardo Pavez (University of Southern California); Tatsuya Kobayashi (KDDI Research Inc.); Antonio Ortega (University of Southern California)

02:45 PM

4310 (IVMSP-L12.4): Could the BubbleView metaphor be used to infer visual attention on 3D graphical content ?

Alexandre Bruckert (Nantes Université); Mona Abid (Nantes université); Matthieu Perreira Da Silva (Université de Nantes); Patrick Le Callet ("Universite de Nantes, France")

03:00 PM

1077 (IVMSP-L12.5): PCSalMix: Gradient Saliency-based Mix Augmentation for Point Cloud Classification

Tao Hong (Peking University); Zeren Zhang (Peking University); Jinwen Ma (Peking University)

03:15 PM

5696 (IVMSP-L12.6): MDR-MFI: Multi-Branch Decoupled Regression and Multi-Scale Feature Interaction for Partial-to-Partial Cloud Registration

Weidong Dai (Hikvision Research Institute); Xuejun Yan (Hikvision Research Institute); Jingjing Wang (Hikvision Research Institute); Di Xie (Hikvision Research Institute); Shiliang Pu (Hikvision Research Institute)

MSP-L2: Multimedia Databases and Information Retrieval

Room: Salon de Roses A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Chi-Man Pun,

02:00 PM

6513 (MSP-L2.1): SSVMR: SALIENCY-BASED SELF-TRAINING FOR VIDEO-MUSIC RETRIEVAL

Xuxin Cheng (Peking University); Zhihong Zhu (Peking University); Hongxiang Li (Peking University); Yaowei Li (Peking University); Yuexian Zou (Peking University)

02:15 PM

1030 (MSP-L2.2): IMPROVING DROPOUT IN GRAPH CONVOLUTIONAL NETWORKS FOR RECOMMENDATION VIA CONTRASTIVE LOSS

Hiroki Okamura (Hokkaido University); Keisuke Maeda (Hokkaido University); Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

Friday, June 9

02:30 PM

2337 (MSP-L2.3): Locality Preserving Multiview Graph Hashing for Large Scale Remote Sensing Image Search

Wenyun Li (University of Macau); Guo Zhong (University of Macau); XINGYU LU (University of Macau); Chi-Man Pun (University of Macau)

02:45 PM

3883 (MSP-L2.4): A dataset for Audio-Visual Sound Event Detection in Movies

Rajat Hebbar (University of Southern California); Digbalay Bose (University of Southern California); Krishna Somandepalli (University of Southern California); Veena Vijai (University of Southern California); Shrikanth Narayanan (USC)

03:00 PM

4725 (MSP-L2.5): Vision, Deduction and Alignment: An Empirical Study on Multi-modal Knowledge Graph Alignment

Li Yangning (Tsinghua Shenzhen International Graduate School); Jiaoyan Chen (The University of Manchester); Yinghui Li (Tsinghua University); Yuejia Xiang (Tencent); Xi Chen (Tencent); Hai-Tao Zheng (Tsinghua University)

03:15 PM

6251 (MSP-L2.6): A DATABASE FOR MULTI-MODAL SHORT VIDEO QUALITY ASSESSMENT

Yukun Zhang (Institute of Information Engineering, Chinese Academy of Sciences); Chuan Wang (Chinese Academy of Sciences); Sanyi Zhang (Institute of Information Engineering, Chinese Academy of Sciences); Xiaochun Cao (Sun Yat-sen University)

SLT-L22: Voice and Style Conversion I

Room: Delphi

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Berrak Sisman, Satoshi Nakamura

02:00 PM

2508 (SLT-L22.1): Cross-speaker Emotion Transfer by Manipulating Speech Style Latents

Suhee Jo (Neosapience, Inc.); Younggun Lee (Neosapience); Yookyung Shin (Neosapience, Inc.); Yeongtae Hwang (Neosapience, Inc.); Taesu Kim (Neosapience, Inc.)

02:15 PM

2640 (SLT-L22.2): A UNIFIED ONE-SHOT PROSODY AND SPEAKER CONVERSION SYSTEM WITH SELF-SUPERVISED DISCRETE SPEECH UNITS

Li-Wei Chen (Carnegie Mellon University); Shinji Watanabe (Carnegie Mellon University); Alexander I. Rudnicky (Carnegie Mellon University)

02:30 PM

3966 (SLT-L22.3): JSV-VC: JOINTLY TRAINED SPEAKER VERIFICATION AND VOICE CONVERSION MODELS

Shogo Seki (NTT Corporation); Hirokazu Kameoka (NTT Communication Science Laboratories, NTT Corporation); Kou Tanaka (NTT Corporation); Takuhiro Kaneko (NTT Corporation)

02:45 PM

5134 (SLT-L22.4): Voice-preserving Zero-shot Multiple Accent Conversion

Mumin Jin (MIT); Prashant Serai (Meta AI); Jilong Wu (Meta AI); Andros Tjandra (Meta Platforms, Inc); Vimal Manohar (Meta Platforms Inc.); Qing He (Meta)

03:00 PM

5209 (SLT-L22.5): FREEVC: TOWARDS HIGH-QUALITY TEXT-FREE ONE-SHOT VOICE CONVERSION

Jingyi Li (Wuhan University); Weiping Tu (Wuhan University); Li Xiao (School of Computer Science, Wuhan University)

03:15 PM

5930 (SLT-L22.6): Preserving background sound in noise-robust voice conversion via multi-task learning

Jixun Yao (Northwestern Polytechnical University); Yi Lei (Northwestern Polytechnical University); Qing Wang (Northwestern Polytechnical University); Pengcheng Guo (Northwestern Polytechnical University); Ziqian Ning (Northwestern Polytechnical University); Lei Xie (NWPU); Hai Li (iQIYI Inc); Junhui Liu (iQIYI Inc); Danming Xie (iQIYI)

SS-L21: Synergy between human and machine approaches to sound/scene recognition and processing

Room: Nafsika A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Benjamin Elizalde,

02:00 PM

2581 (SS-L21.01): Perceptual analysis of speaker embeddings for voice discrimination between machine and human listening

Iordanis Thoidis (Aristotle University of Thessaloniki); Clément Gaultier (University of Cambridge); Tobias Goehring (University of Cambridge)

Friday, June 9

02:15 PM

4176 (SS-L21.02): Semantically-informed Deep Neural Networks for sound recognition

Michele Esposito (Maastricht University); Giancarlo Valente (Maastricht University); Yenisel Plasencia-Calaña (Maastricht University); Michel Dumontier (Maastricht University); Bruno L. Giordano (CNRS); Elia Formisano (Maastricht University)

02:30 PM

4472 (SS-L21.03): An Approach to Ontological Learning from Weak Labels

Ankit Parag Shah (Carnegie Mellon University); Larry Tang (Carnegie Mellon University); Po Hao Chou (Carnegie Mellon University); Yi Yu Zheng (Carnegie Mellon University); Ziqiang Ge (Carnegie Mellon University); Bhiksha Raj (Carnegie Mellon University)

02:45 PM

4813 (SS-L21.04): CLASSIFYING NON-INDIVIDUAL HEAD-RELATED TRANSFER FUNCTIONS WITH A COMPUTATIONAL AUDITORY MODEL: CALIBRATION AND METRICS

Rapolas Daugintis (Imperial College London); Roberto Barumerli (Austrian Academy of Sciences); Lorenzo Picinali (Imperial College London); Michele Geronazzo (Imperial College London)

03:00 PM

4960 (SS-L21.05): Perceptual-Neural-Physical Sound Matching

Han Han (Ecole Centrale Nantes); Vincent LOSTANLEN (Cornell Lab of Ornithology); Mathieu Lagrange (LS2N)

03:15 PM

4988 (SS-L21.06): USING MACHINE LEARNING TO UNDERSTAND THE RELATIONSHIPS BETWEEN AUDIOMETRIC DATA, SPEECH PERCEPTION, TEMPORAL PROCESSING, AND COGNITION

Rana Khalil (University of Maryland - College Park); Alexandra Papanicolaou (University of Maryland - College Park); Renee Chou (University of Maryland - College Park); Bobby Gibbs (University of Maryland - College Park); Samira B Anderson (University of Maryland); Sandra Gordon-Salant (University of Maryland - College Park); Michael Cummings (University of Maryland - College Park); Matthew J. Goupell (University of Maryland - College Park)

SS-L22: Topological and Simplicial Data Processing

Room: Nefeli A

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Santiago Segarra, Elvin Isufi

02:00 PM

2602 (SS-L22.1): Signal Processing on Product Spaces

T. Mitchell Roddenberry (Rice University); Vincent P Grande (RWTH Aachen University); Florian Frantzen (RWTH Aachen University); Michael Schaub (RWTH Aachen University); Santiago Segarra (Rice University)

02:15 PM

2607 (SS-L22.2): Online Edge Flow Prediction over Expanding Simplicial Complexes

Maosheng Yang (Delft University of Technology); Bishwadeep Das (TU Delft); Elvin Isufi (TU Delft)

02:30 PM

3653 (SS-L22.3): TOPOLOGICAL SIGNAL PROCESSING OVER WEIGHTED SIMPLICIAL COMPLEXES

Claudio Battiloro (Sapienza University of Rome); Stefania Sardellitti (Sapienza University of Rome); Sergio Barbarossa (Sapienza University of Rome); Paolo Di Lorenzo (Sapienza University of Rome)

02:45 PM

5420 (SS-L22.4): Topo-MLP : A Simplicial Network Without Message Passing

Karthikeyan Natesan Ramamurthy (IBM Research); Aldo Guzmán-Sáenz (IBM); Mustafa Hajj (USFCA)

03:00 PM

5898 (SS-L22.5): Higher-order Spatio-temporal Neural Networks for COVID-19 Forecasting

Yuzhou Chen (Temple University); Sotirios P Batsakis (University of Huddersfield); H. Vincent Poor (Princeton University)

SS-L23: Unsupervised Deep Learning of Image Priors for Inverse Problems

Room: Nafsika B

Type: Oral

02:00 PM to 03:30 PM

Chair(s): Bihan Wen,

02:00 PM

357 (SS-L23.1): Stay in the Middle: A Semi-Supervised Model for CT Metal Artifact Reduction

Tao Wang (Sichuan University); Hui Yu (Sichuan University); Zexin Lu (Sichuan University); Zhongzhou Zhang (Sichuan university); Jiliu Zhou (Chengdu University of Information Technology); Yi Zhang (Sichuan University)

Friday, June 9

02:15 PM

873 (SS-L23.2): Unsupervised Deep Virtual Staining for Microscopic Cell Images via Knowledge Distillation
Ziwan Xu (School of Electrical and Electronic Engineering, Nanyang Technological University); Lanqing Guo (Nanyang Technological University); Shuyan Zhang (Agency for Science, Technology and Research); Alex Kot (Nanyang Technological University); Bihan Wen (Nanyang Technological University)

02:30 PM

4737 (SS-L23.3): DEEP PROXIMAL GRADIENT METHOD FOR LEARNED CONVEX REGULARIZERS
Aaron Berk (McGill University); Yanting Ma (Mitsubishi Electric Research Laboratories, USA); Petros Boufounos (Mitsubishi Electric Research Laboratories); Pu Wang (MERL); Hassan Mansour (Mitsubishi Electric Research Laboratories (MERL))

02:45 PM

4993 (SS-L23.4): ROBUST SELF-GUIDED DEEP IMAGE PRIOR
Evan Bell (Michigan State University); Shijun Liang (michigan state university); Qing Qu (University of Michigan); Saiprasad Ravishankar (Michigan State University)

03:00 PM

5248 (SS-L23.5): CryoSVD: Sliced Wasserstein Distance Minimization for 3D Reconstruction in Cryo-Electron Microscopy
Mona Zehni (University of Illinois at Urbana-Champaign); Zhizhen Zhao (University of Illinois at Urbana-Champaign)

AASP-P15: Self-Supervised Learning and Data-Efficiency for Speech and Audio

Room: Dome A

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Minje Kim,

1513 (AASP-P15.1): EFFECT OF ACOUSTIC UNIT GRANULARITY ON SEQ2SEQ REPRESENTATION LEARNING

Ali Elkahky (Meta, Inc); Wei-Ning Hsu (Meta, Inc); Paden P Tomasello (Meta); Tu Anh Nguyen (Meta, Inc); Robin Algayres (Inria, Paris, France); Yossi Adi (Facebook AI Research); Jade Copet (Meta, Inc); Emmanuel Dupoux (Meta, Inc); Abdelrahman Mohamed (Meta, Inc)

5676 (AASP-P15.2): The Potential of Neural Speech Synthesis-Based Data Augmentation for Personalized Speech Enhancement

Anastasia Kuznetsova (Indiana University); Aswin Sivaraman (Indiana University Bloomington); Minje Kim (Indiana University)

4214 (AASP-P15.3): SPEAKERAUGMENT: DATA AUGMENTATION FOR GENERALIZABLE SOURCE SEPARATION VIA SPEAKER PARAMETER MANIPULATION

Kai Wang (Xinjiang University); Yuhang Yang (School of Information Science and Engineering, Xinjiang University, China); Hao Huang (Xinjiang University); Ying Hu (Xinjiang University); Sheng Li (National Institute of Information & Communications Technology (NICT))

5842 (AASP-P15.4): AUDIO SIGNAL ENHANCEMENT WITH LEARNING FROM POSITIVE AND UNLABELLED DATA

Nobutaka Ito (UTokyo); Masashi Sugiyama (RIKEN/The University of Tokyo)

4192 (AASP-P15.5): Analysis of Noisy-target Training for DNN-based speech enhancement

Takuya Fujimura (Nagoya University); Tomoki Toda (Nagoya University)

5992 (AASP-P15.6): An empirical study on speech restoration guided by self-supervised speech representation

Jaeuk Byun (Naver Corporation); Youna Ji (NAVER Corporation); Soo-Whan Chung (Naver Corporation); Soyeon Choe (NAVER Corporation); Min-Seok Choi (NAVER)

6686 (AASP-P15.7): Self-Supervised Contrastive Learning for Singing Voices (SPS Journal Paper)*

Hiromu Yakura (University of Tsukuba); Kento Watanabe (National Institute of Advanced Industrial Science and Technology (AIST)); Masataka Goto (National Institute of Advanced Industrial Science and Technology (AIST))

6703 (AASP-P15.8): Pretext Tasks Selection for Multitask Self-Supervised Audio Representation Learning (SPS Journal Paper)*

Salah Zaiem (Telecom Paris); Titouan Parcollet (Samsung AI Research); Slim Essid (Telecom Paris - Institut Polytechnique de Paris); Abdelwaheb Heba (Microsoft)

6743 (AASP-P15.9): BYOL for Audio: Exploring Pre-Trained General-Purpose Audio Representations (SPS Journal Paper)*

Daisuke Niizumi (NTT Corporation); Daiki Takeuchi (NTT Corporation); Yasunori Ohishi (NTT Corporation); Noboru Harada (NTT); Kunio Kashino (NTT Communication Science Laboratories)

6786 (AASP-P15.10): Efficient Personalized Speech Enhancement Through Self-Supervised Learning (SPS Journal Paper)*

Aswin Sivaraman (Indiana University Bloomington); Minje Kim (Indiana University)

6801 (AASP-P15.11): Self-Supervised Graphs for Audio Representation Learning With Limited Labeled Data (SPS Journal Paper)*

Amir Shirian (University of Warwick); Krishna Somandepalli (Google Research); Tanaya Guha (University of Glasgow)

Friday, June 9

AASP-P16: Sound Event Detection and Localization; Bioacoustic Event Detection

Room: Dome B

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Michael Mandel, Annamaria Mesaros

2148 (AASP-P16.1): INTER-PULSE ESTIMATION FOR SPERM WHALE CLICK DETECTION

Guy Gubnitsky (University of Haifa); Roe Diamant (University of Haifa)

3383 (AASP-P16.2): BEANS: The Benchmark of Animal Sounds

Masato Hagiwara (Earth Species Project); Benjamin Hoffman (Earth Species Project); Jen-Yu Liu (Earth Species Project); Maddie Cusimano (Earth Species Project); Felix Effenberger (Earth Species Project); Katie Zacarian (Earth Species Project)

3610 (AASP-P16.3): AVES: Animal Vocalization Encoder based on Self-Supervision

Masato Hagiwara (Earth Species Project)

984 (AASP-P16.4): Improving Weakly Supervised Sound Event Detection with Causal Intervention

Yifei Xin (Peking University); Dongchao Yang (Peking university); fan cui (xiaomi); Yujun Wang (xiaomi); Yuexian Zou (Peking University)

1325 (AASP-P16.5): SEMI-SUPERVISED SOUND EVENT DETECTION WITH PRE-TRAINED MODEL

Liang Xu (Beijing Institute of Technology); Lizhong Wang (Samsung); Sijun Bi (Beijing Institute of Technology); Hanyue Liu (Beijing Institute of Technology); Jing Wang (Beijing Institute of Technology)

5672 (AASP-P16.6): Multi-dimensional frequency dynamic convolution with confident mean teacher for sound event detection

shengchang xiao (UCAS); xueshuai zhang (UCAS); pengyuan zhang (Institute of Acoustics, Chinese Academy of Sciences)

5297 (AASP-P16.7): LOSS FUNCTION DESIGN FOR DNN-BASED SOUND EVENT LOCALIZATION AND DETECTION ON LOW-RESOURCE REALISTIC DATA

Qing Wang (University of Science and Technology of China); Jun Du (University of Science and Technology of China); Zhaoxu Nian (University of Science and Technology of China); Shutong Niu (University of Science and Technology of China); Li Chai (University of Science and Technology of China); Huaxin Wu (iFlytek Research); Jia Pan (University of Science and Technology of China); Chin-Hui Lee (Georgia Institute of Technology)

317 (AASP-P16.8): Few-shot continual learning with weight alignment and positive enhancement for bioacoustic event detection

Xiaoxiao Wu (Shanghai Normal University); Dongxing Xu (Unisound AI Technology Co., Ltd., Beijing); Haoran Wei (University of Texas at Dallas); yanhua long (Shanghai Normal University)

2743 (AASP-P16.9): EVALUATING VARIANTS OF WAV2VEC 2.0 ON AFFECTIVE VOCAL BURST TASKS

Bagus Tris Atmaja (Sepuluh Nopember Institute of Technology); Akira Sasou (AIST)

3356 (AASP-P16.10): TRANSFORMER-BASED BIOACOUSTIC SOUND EVENT DETECTION ON FEW-SHOT LEARNING TASKS

Liwen You (Amazon); Erika Pelaez Coyotl (Amazon); Suren Gunturu (Amazon); Maarten Van Segbroeck (Amazon)

4137 (AASP-P16.11): CONVOLUTIONAL RECURRENT NEURAL NETWORKS FOR THE CLASSIFICATION OF CETACEAN BIOACOUSTIC PATTERNS

Dimitris Makropoulos (National Technical University of Athens); Antigonis Tsiami (National Technical University of Athens); Aristides M Prospathopoulos (HCMR); DIMITRIS KASSIS (HCMR); Alexandros Frantzis (Pelagos Cetacean Research Institute); Emmanuel Skarsoulis (Foundation of Research and Technology - HELLAS); George Piperakis (Foundation of Research and Technology - HELLAS); Petros Maragos (National Technical University of Athens)

4094 (AASP-P16.12): Training sound event detection with soft labels from crowdsourced annotations

Irene Martin (Tampere University); Manu Harju (Tampere University); Paul Ahokas (Tampere University); Annamaria Mesaros (Tampere University)

IVMSP-P31: Classification/ Anomaly Detection/ Object Recognition/ Image Generation/ Object Detection/ Domain-Specific Detection

Room: Dome J

Type: Poster

02:00 PM to 03:30 PM

Chair(s): C.-C. Jay Kuo, Ran Tao, Zhenhua Guo, Wei Feng, Weiming Hu, Zhiqiang Wu, Hanzi Wang, F Richard Yu, Junliang Xing, Nicu Sebe, Kaihua Zhang, Haoran Xie

Friday, June 9

5876 (IVMSP-P31.1): Classifying Pathological Images Based on Multi-Instance Learning and End-to-End Attention Pooling
Yuqi Chen (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Juan Liu (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Zhiqun Zuo (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Peng Jiang (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Yu Jin (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Guangsheng Wu (School of Mathematics and Computer Science, Xinyu University)

3711 (IVMSP-P31.2): A Simulation-Based Framework for Urban Road Accident Detection
Haohan Luo (East China Normal University); Feng Wang (East China Normal University)

4690 (IVMSP-P31.3): Collaborative Audio-Visual Event Localization based on Sequential Decision and Cross-modal Consistency
Yuqian Kuang (Harbin Institute of Technology); Xiaopeng Fan (Harbin Institute of Technology)

1834 (IVMSP-P31.4): Continuous interaction with a smart speaker via low-dimensional embeddings of dynamic hand pose
songpei xu (University of Glasgow); Chaitanya Kaul (University of Glasgow); Xuri Ge (University of Glasgow); Roderick Murray-Smith (University of Glasgow)

3993 (IVMSP-P31.5): MPE4G : Multimodal Pretrained Encoder for Co-Speech Gesture Generation
Gwantae Kim (Korea University); Seonghyeok Noh (Korea University); Insung Ham (Korea University); Hanseok Ko (Korea University)

3115 (IVMSP-P31.6): MSNet: A Deep Architecture using Multi-Sentiment Semantics for Sentiment-Aware Image Style Transfer
Shikun Sun (Tsinghua University); Jia Jia (Tsinghua University); Haozhe Wu (Tsinghua University); Zijie Ye (Tsinghua University); Junliang Xing (Tsinghua University)

6215 (IVMSP-P31.7): Pondering about Task Spatial Misalignment: Classification-Localization Equilibrated Object Detection
Yudong Zhang (University of Science and Technology of China); Wei Lu (University of Science and Technology of China); Xu Wang (University of Science and Technology of China); Pengkun Wang (University of Science and Technology of China); Yang Wang (University of Science and Technology of China)

4803 (IVMSP-P31.8): Efficient Feature Extraction for Non-Maximum Suppression in Visual Person Detection
Charalampos Symeonidis (AUTH); Ioannis Mademlis (Department of Informatics, Aristotle University of Thessaloniki); Ioannis Pitas (Aristotle University of Thessaloniki); Nikolaos Nikolaidis (Aristotle University of Thessaloniki)

5280 (IVMSP-P31.9): Boundary Cue Guidance and Contextual Feature Mining for Glass Segmentation
Qiquan Xiao (Xiangtan University); Yuan Zhang (Xiangtan University); Xuanya Li (Baidu); Kai Hu (Xiangtan University)

5685 (IVMSP-P31.10): Optimal Kernel for Real-Time Arbitrary-Shaped Text Detection
Haozhao Ma (Northwestern Polytechnical University); Chuang Yang (Northwestern Polytechnical University); Yuan Yuan (Northwestern Polytechnical University); Qi Wang (Northwestern Polytechnical University)

5708 (IVMSP-P31.11): KEPS-NET: Robust Parking Slot Detection based Keypoint Estimation for High Localization Accuracy
Jaewoo Lee (Samsung Electronics); Kapje Sung (Samsung Electronics); Daeul Park (Samsung Electronics); Younghwan Jeon (Seoul National University)

IVMSP-P32: Super Resolution, Denoising, Image Restoration, Image/Video Enhancement

Room: Dome K

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Chun Yuan, Xinbo Gao, Qiang Shen, Yijin Liu, Xuelong Li, Ju Sun, Yu-Chiang Frank, Yao Zhao

3945 (IVMSP-P32.1): Deep Feature Aggregation for Lightweight Single Image Super-Resolution
Yanchun Li (Xiangtan university); Xinan He (Xiangtan University); Shujuan Tian (Xiangtan University); Zhetao Li (湘潭大学); Saiqin Long (Jinan University)

4175 (IVMSP-P32.2): Stochastic super-resolution for Gaussian textures
Emile Pierret (Institut Denis Poisson); Bruno Galerne (University of Orléans)

4263 (IVMSP-P32.3): COLOR GUIDED DEPTH MAP SUPER-RESOLUTION WITH NONLOCLA AUTOREGRESSIVE MODELING
Wei Xu (Faculty of Information Technology, Beijing University of Technology); Na Qi (Beijing University of Technology); Qing Zhu (Beijing University of Technology); Jingzhong Qi (Beijing University of Technology); Longlu Huang (Beijing University of Technology); Kun Cao (Beijing University of Technology); Yuxin Bao (Beijing University of Technology); Qianwen Wang (Beijing university of technology)

6067 (IVMSP-P32.4): Classification-based Dynamic Network for Efficient Super-Resolution

Friday, June 9

Qi Wang (Beijing Jiaotong University); Weiwei Fang (Beijing Jiaotong University); Meng Wang (Beijing Jiaotong University); Yusong Cheng (Beijing Jiaotong University)

212 (IVMSP-P32.5): M2TSR: Multi-range and Mix-grained Transformer for Single Image Super-Resolution

Zhong-Han Niu (State Key Laboratory for Novel Software Technology, Nanjing University); Qinglong Zhang (State Key Laboratory for Novel Software Technology, Nanjing University); Yi Fan (State Key Laboratory for Novel Software Technology, Nanjing University); Yu-Bin Yang (State Key Laboratory for Novel Software Technology, Nanjing University)

3986 (IVMSP-P32.6): A highly Interpretable Deep equilibrium network for hyperspectral image deconvolution

Alexandros Gkillas (University of Patras); Dimitris Ampeliotis (Digital Media and Communication Department, Ionian University, Greece); Kostas Berberidis (University of Patras)

6815 (IVMSP-P32.7): PoGalN: Poisson-Gaussian Image Noise Modeling From Paired Samples (SPS Journal Paper)*

Nicolas L Bähler (EPFL); Majed El Helou (ETHZ); Kaan Okumuş (EPFL); Étienne Objois (EPFL); Sabine Süsstrunk (EPFL)

242 (IVMSP-P32.8): IMAGE COMPLETION VIA DUAL-PATH COOPERATIVE FILTERING

Pourya Shamsolmoali (East China Normal University); Masoumeh Zareapoor (Shanghai Jiao Tong University); Eric Granger (ETS Montreal)

6808 (IVMSP-P32.9): Trainable Subspaces for Low Rank Tensor Completion: Model and Analysis (SPS Journal Paper)*

Zhen Long (University of Electronic Science and Technology of China); Ce Zhu (University of Electronic Science & Technology of China); Jiani Liu (University of Electronic Science and Technology of China); Pierre Comon (gipsa-lab); Yipeng Liu (University of Electronic Science and Technology of China)

1854 (IVMSP-P32.10): SHADOW REMOVAL OF TEXT DOCUMENT IMAGES USING BACKGROUND ESTIMATION and ADAPTIVE TEXT ENHANCEMENT

Wenjie Liu (Northwestern Polytechnical University); Bingshu Wang (Northwestern Polytechnical University); Jiangbin Zheng (Northwestern Polytechnical University); Wenmin Wang (Macau University of Science and Technology)

2133 (IVMSP-P32.11): ShaDocNet: Learning Spatial-Aware Tokens in Transformer for Document Shadow Removal

Xuhang Chen (University of Macau); Xiaodong Cun (Tencent AI Lab); Chi-Man Pun (University of Macau); Shuqiang Wang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences)

1759 (IVMSP-P32.12): DEWARPING DOCUMENTS USING C2 CONTINUOUS BOUNDARY ESTIMATION

Prasenjit Mondal (Adobe); Ayush Pant (Adobe); Sachin Soni (Adobe)

IVMSP-P33: Representation Learning, Anomaly Detection, Image Generation, Temporal Video Analysis and Detection, Image/Video Retrieval

Room: Dome M

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Ju Sun, Ye Peng, Zhenhua Guo, Zhiqiang Wu, Junliang Xing, Nicu Sebe, Tyng-Luh Liu, Lijun Yin, Bingbing Ni, Pan Zhou

6816 (IVMSP-P33.1): Efficient ADMM-based Algorithms for Convolutional Sparse Coding (SPS Journal Paper)*

Farshad G Veshki (Aalto university); Sergiy Vorobyov ()

6768 (IVMSP-P33.2): ISP Distillation (OJSP Paper)*

Eli Schwartz (Tel Aviv University); Alex Bronstein (Technion); Raja Giryes (Tel Aviv University)

7176 (IVMSP-P33.3): Self-Supervised Learning Based Anomaly Detection in Synthetic Aperture Radar Imaging

Max Muzeau (ONERA); Chengfang Ren (Sondra - CentraleSupélec); Sebastien Angelliaume (Onera); Mihai Dactu (Politehnica University of Bucharest); Jean-Philippe Ovarlez (ONERA, CentraleSupélec, SONDRRA, Université Paris-Saclay)

295 (IVMSP-P33.4): A discriminative multi-channel noise feature representation method for image manipulation localization

yang zhou (sichuan university); Hongxia Wang (Sichuan University); Qiang Zeng (Sichuan University); Rui Zhang (Sichuan University); Sijiang Meng (Sichuan University)

1945 (IVMSP-P33.5): Semantic-Aware Gated Fusion Network for Interactive Colorization

Jie Zhang (Hunan University); yi xiao (Hunan University); yan zheng (Hunan University); Zhenni Wang (City University of Hong Kong); Chi Sing Leung (City University of Hong Kong)

1670 (IVMSP-P33.6): Line segment matching based on intersection-enhanced point correspondences

Zhiyu Liu (School of Computer Science and Technology, Soochow University); Baojiang Zhong (School of Computer Science and Technology, Soochow University)

6819 (IVMSP-P33.7): BL-JUNIPER: A CNN-Assisted Framework for Perceptual Video Coding Leveraging Block-Level JND (SPS Journal Paper)*

Sanaz Nami (Tampere University); Farhad Pakdaman (Tampere University); Mahmoud R. Hashemi (University of Tehran, Iran); Shervin Shirmohammadi (University of Ottawa, Canada)

Friday, June 9

1031 (IVMSP-P33.8): ESTIMATION OF VISUAL CONTENTS FROM HUMAN BRAIN SIGNALS VIA VQA BASED ON BRAIN-SPECIFIC ATTENTION

Ryo Shichida (Hokkaido University); Ren Togo (Hokkaido University); Keisuke Maeda (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

7174 (IVMSP-P33.9): Recallable Question Answering-based Re-ranking Considering Semantic Region for Cross-modal Retrieval

Rintaro Yanagi (Hokkaido University); Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

2226 (IVMSP-P33.10): SELF-DISTILLATION HASHING FOR EFFICIENT HAMMING SPACE RETRIEVAL

Hongjia HJ Zhai (Zhejiang University); Hai Li (Zhejiang University); hanzhi zhang (Zhejiang University); Hujun Bao (Zhejiang University); Guofeng Zhang (Zhejiang University)

MLSP-P31: Learning Algorithms and Applications

Room: Dome C

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

893 (MLSP-P31.1): Quantifying Catastrophic Forgetting in Continual Federated Learning

Christophe Dupuy (Amazon); Jimit Majmudar (Amazon); Jixuan Wang (Amazon); Tanya G Roosta (Amazon); Rahul Gupta (Amazon); Clement Chung (Amazon); Jie Ding (Amazon); Salman Avestimehr (University of Southern California)

4840 (MLSP-P31.2): Client Selection for Generalization in Accelerated Federated Learning: A Bandit Approach

Dan Ben Ami (Ben-Gurion University of the Negev); Kobi Cohen (Ben-Gurion University of the Negev); Qing Zhao (Cornell University)

5461 (MLSP-P31.3): Accelerated Distributed Stochastic Non-Convex Optimization over Time-Varying Directed Networks

Yiyue Chen (University of Texas at Austin); Abolfazl Hashemi (Purdue University); Haris Vikalo (University of Texas at Austin)

2659 (MLSP-P31.4): QuantPipe: Applying Adaptive Post-Training Quantization for Distributed Transformer Pipelines in Dynamic Edge Environments

Haonan Wang (University of Southern California); Connor Imes (Information Sciences Institute, USC); Souvik Kundu (University of Southern California); Peter A. Beerel (University of Southern California); Stephen Crago (Information Sciences Institute, USC); John Paul Walters (Information Sciences Institute, USC)

2626 (MLSP-P31.5): Utility pole localization by learning from ambient traces on distributed acoustic sensing

Zhuocheng Jiang (NEC laboratories America, Inc.); Yue Tian (NEC laboratories America, Inc.); Yangmin Ding (NEC Labs America); Sarper Ozharar (NEC laboratories America, Inc.); Ting Wang (NEC laboratories America, Inc.)

1872 (MLSP-P31.6): A geometric surrogate for simulation calibration

Lincon Souza (National Institute of Advanced Industrial Science and Technology (AIST)); Bojan Batalo (University of Tsukuba); Keisuke Yamazaki (National Institute of Advanced Industrial Science and Technology)

4431 (MLSP-P31.7): LEARNING FROM POSITIVE AND UNLABELED DATA USING OBSERVER-GAN

Omar Zamzam (University of Southern California); Haleh Akrami (Signal and Image Processing Institute at University of Southern California); Richard Leahy (Signal and Image Processing Institute at University of Southern California)

2518 (MLSP-P31.8): PRIV-AUG-SHAP-EGRESNET: PRIVACY PRESERVING SHAPLEY-VALUE ATTRIBUTED AUGMENTED RESNET FOR PRACTICAL SINGLE-LEAD ELECTROCARDIOGRAM CLASSIFICATION

Arijit Ukil (Tata Consultancy Services); Leandro Marin (University of Murcia); Antonio J. Jara (Libelium)

4628 (MLSP-P31.9): WATER LEAK DETECTION AND LOCALIZATION USING CONVOLUTIONAL AUTOENCODERS

Daniele Ugo Leonzio (Politecnico di Milano); Paolo Bestagini (Politecnico di Milano); Marco Marcon (Politecnico di Milano); Gian Paolo Quarta (Onyax); Stefano Tubaro (Politecnico di Milano, Italy)

5372 (MLSP-P31.10): Towards Scale Adaptive Underwater Detection through Refined Pyramid Grid

Xiaoheng Deng (Central South University); Lirong Liao (Xinjiang University); Ping Jiang (Central South University); Yurong Qian (Xinjiang University)

4498 (MLSP-P31.11): Towards low-power heart rate estimation based on user's demographics and activity level for wearables

Andre GC Pacheco (Samsung); Frank Cabello (Samsung); Paula Rodrigues (Samsung); Paula Pinto (Samsung); Adriana Fonoff (Samsung); Otávio Penatti (SAMSUNG)

Friday, June 9

MLSP-P32: Optimization Methods in Machine Learning

Room: Dome D

Type: Poster

02:00 PM to 03:30 PM

Chair(s): TBA

5487 (MLSP-P32.1): A Gaussian Latent Variable Model for Incomplete Mixed Type Data

Marzieh Ajirak (Stony Brook University); Petar Djuric ()

3392 (MLSP-P32.2): Bayesian Optimization with Ensemble Learning Models and Adaptive Expected Improvement

Konstantinos D. Polyzos (University of Minnesota); Qin Lu (University of Minnesota); Georgios B. Giannakis (University of Minnesota)

681 (MLSP-P32.3): Stochastic Optimization of Vector Quantization Methods in Application to Speech and Image Processing

Mohammad Hassan Vali (Aalto University); Tom Bäckström (Aalto University)

3202 (MLSP-P32.4): WordReg: Mitigating the Gap between Training and Inference with Worst-case Drop Regularization

Jun Xia (Westlake University); Ge Wang (Westlake University); Bozhen Hu (Zhejiang University & Westlake University); Cheng Tan (Zhejiang University & Westlake University); Jiangbin Zheng (Westlake University); Yongjie Xu (Westlake University); Stan Z. Li (Westlake University)

4551 (MLSP-P32.5): Improving the Stochastic Gradient Descent's test accuracy by manipulating the l_{∞} norm of its gradient approximation

Paul Rodriguez (PUCP)

3670 (MLSP-P32.6): Convex Optimization of Deep Polynomial and ReLU Activation Neural Networks

Burak Bartan (Stanford University); Mert Pilanci (Stanford University)

3748 (MLSP-P32.7): UAV Local Path Planning Based on Improved Proximal Policy Optimization Algorithm

Jiahao xu (Nanjing University of Aeronautics and Astronautics); Xuefeng Yan (Nanjing University of Aeronautics and Astronautics); Peng Cui (Dalian Naval Academy); Xinquan Wu (Nanjing University of Aeronautics and Astronautics); Lipeng Gu (Nanjing University of Aeronautics and Astronautics); Yan biao Niu (Nanjing University of Aeronautics and Astronautics)

4172 (MLSP-P32.8): Newton-based Trainable Learning Rate

George Retsinas (National Technical University of Athens); Giorgos Sfikas (University of West Attica); Panagiotis P Filintisis (National Technical University of Athens); Petros Maragos (National Technical University of Athens)

6453 (MLSP-P32.9): Learning Gradients of Convex Functions with Monotone Gradient Networks

Shreyas Chaudhari (Carnegie Mellon University); Srinivasa Pranav (Carnegie Mellon University); José M. F. Moura (Carnegie Mellon University)

2246 (MLSP-P32.10): POLICE: Provably Optimal Linear Constraint Enforcement for Deep Neural Networks

Randall Balestriero (Facebook AI Research); Yann Lecun (Facebook)

1190 (MLSP-P32.11): Toward Asymptotic Optimality: Sequential Unsupervised Regression of Density Ratio for Early Classification

Akinori F Ebihara (NEC Corporation); Taiki Miyagawa (NEC Corporation); Kazuyuki Sakurai (NEC Biometrics Research Laboratories); Hitoshi Imaoka (NEC Corporation)

5346 (MLSP-P32.12): Algebraic Convolutional Filters on Lie Group Algebras

Harshat Kumar (University of Pennsylvania); Alejandro Parada-Mayorga (University of Pennsylvania); Alejandro Ribeiro (University of Pennsylvania)

MLSP-P33: Applications of Machine Learning

Room: Dome E

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Chang D. Yoo,

4256 (MLSP-P33.1): HyperSteg: Hyperbolic Learning for Deep Steganography

Shivam Agarwal (University of Illinois Urbana-Champaign); Ritesh Singh Soun (Sri Venkateswara College); Rahul Shivani (M.B.M. Engineering College, Jodhpur); Vishnuvardhan Varanasi V (IIT Kanpur); Navroop Gill (Scaler); Ramit Sawhney (IIIT Delhi)

156 (MLSP-P33.2): HDNet: Hierarchical Dynamic Network for Gait Recognition using Millimeter-Wave Radar

Yanyan Huang (Zhejiang University); Yong Wang (Zhejiang University); Kun Shi (Zhejiang University); Chaojie Gu (Zhejiang University); Yu Fu (Zhejiang University); Cheng Zhuo (Zhejiang University); Zhiguo Shi (Zhejiang University)

Friday, June 9

162 (MLSP-P33.3): Gluformer: Transformer-Based Personalized Glucose Forecasting with Uncertainty Quantification
Renat Sergazinov (Texas A&M University); Mohammadreza Armandpour (Texas A&M University); Irina Gaynanova (Texas A&M University)

2257 (MLSP-P33.4): Prefix-Level Detection and Autocorrection of Keyboard Input Errors
Jerome R Bellegarda (Apple)

3308 (MLSP-P33.5): MCROOD: MULTI-CLASS RADAR OUT-OF-DISTRIBUTION DETECTION
Sabri Mustafa Kahya (Technical University of Munich); Muhammet Sami Yavuz (Technical University of Munich); Eckehard Steinbach (TUM)

4293 (MLSP-P33.6): QUANTILE ONLINE LEARNING FOR SEMICONDUCTOR FAILURE ANALYSIS
bangjian Zhou (ASTAR,I2R,MI); Pan Jierning (Electrical and Computer Engineering, National University of Singapore); Maheswari Sivan (Electrical and Computer Engineering, National University of Singapore); Aaron Voon-Yew Thean (Department of Electrical and Computer Engineering, NUS, Singapore); Senthilnath Jayavelu (Institute for Infocomm Research , ASTAR, Singapore)

4644 (MLSP-P33.7): Self-attention for Enhanced OAMP Detection in MIMO Systems
Alexander Fuchs (University of Technology Graz); Christian Knoll (Graz, University of Technology); Nima Najari Moghadam (Huawei Technologies Sweden AB); Alexey Pak (Huawei Technologies Sweden AB); Jinliang Huang (Huawei Technologies Sweden AB); Erik Leitinger (Graz University of Technology); Franz Pernkopf (Graz University of Technology)

5222 (MLSP-P33.8): RECURSIVE ESTIMATION OF USER INTENT FROM NONINVASIVE ELECTROENCEPHALOGRAPHY USING DISCRIMINATIVE MODELS
Niklas Smedemark-Margulies (Northeastern University); Basak Celik (Northeastern University); Tales Imbiriba (Northeastern University); Aziz Kocanaogullari (Northeastern University); Deniz Erdogmus (Northeastern University)

2193 (MLSP-P33.9): AMC-Net: An Effective Network for Automatic Modulation Classification
JiaWei Zhang (Xidian University); Tiantian Wang (Xidian University); Zhixi Feng (Xidian university); Shuyuan Yang (Xidian University)

5518 (MLSP-P33.10): Communication-Constrained Exchange of Zeroth-Order Information with Application to Collaborative Target Tracking
Ege Can Kaya (Purdue University); Mehmet Berk Şahin (Purdue University); Abolfazl Hashemi (Purdue University)

6385 (MLSP-P33.11): Joint Cryo-ET Alignment and Reconstruction with Neural Deformation Fields
Valentin Debarnot (University of Basel); Sidharth Gupta (University of Illinois at Urbana-Champaign); Konik Kothari (University of Illinois at Urbana-Champaign); Ivan Dokmanic (University of Basel)

SLT-P44: Various Aspects In Speech and Language Processing

Room: Dome D

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Constantine Kotropoulos,

4626 (SLT-P44.1): Utilizing Wav2vec in Database-independent Voice Disorder Detection
Saska Tirronen (Aalto University); Farhad Javanmardi (Aalto University); Manila Kodali (Aalto University); Sudarsana Reddy Kadiri (Aalto University); Paavo Alku (Aalto University)

7169 (SLT-P44.2): Hierarchical Multi-class Classification of Voice Disorders Using Self-supervised Models and Glottal Features
Saska Tirronen (Aalto University); Sudarsana Reddy Kadiri (Aalto University); Paavo Alku (Aalto University)

6340 (SLT-P44.3): SPASHT: Semantic and PrAgmatic SpeeCh Features for automatic assessment of autism
B Ashwini (Indraprastha Institute of Information Technology, New Delhi, India); Vrinda Narayan (Indraprastha Institute of Information Technology, New Delhi, India); Jainendra Shukla (IIIT-Delhi)

6791 (SLT-P44.4): Time-domain speech separation networks with graph encoding auxiliary (SPS Journal Paper)*
Wang Tingting (Nanjing University of Posts and Tel); Zexu Pan (National University of Singapore); Meng Ge (Tianjin University); Zhen Yang (Nanjing University of Posts and Telecommunication); Haizhou Li (The Chinese University of Hong Kong, Shenzhen)

6307 (SLT-P44.5): Continuous Action Space-based Spoken Language Acquisition Agent Using Residual Sentence Embedding and Transformer Decoder
Ryota Komatsu (Tokyo Institute of Technology); Yusuke Kimura (Tokyo Institute of Technology); Takuma Okamoto (National Institute of Information and Communications Technology); Takahiro Shinozaki (Tokyo Institute of Technology)

6560 (SLT-P44.6): MFCCGAN: A Novel MFCC-Based Speech Synthesizer Using Adversarial Learning
Mohammad Reza Hasanabadi (Shahid Beheshti University); Majid - Behdad (Shahid Beheshti University); Davood Gharavian (Shahid Beheshti University)

Friday, June 9

5665 (SLT-P44.7): Numerical Semantic Modeling for Implicit Discourse Relation Recognition

Chenxu Wang (Department of Computer Science and Technology, Beijing Institute of Technology); Ping Jian (Beijing Engineering Research Center of High Volume Language Information Processing and Cloud Computing Applications, Department of Computer Science and Technology, Beijing Institute of Technology); Hai Wang (Beijing Institute of Technology)

6219 (SLT-P44.8): History, Present and Future: Enhancing Dialogue Generation with Few-shot History-Future Prompt

Yihe Wang (Wuhan University); Yitong Li (Huawei Technologies Co., Ltd.); Yasheng Wang (NoahArk Lab, Huawei); Fei Mi (Huawei); pingyi zhou (Noah's Ark Lab, Huawei); Jin Liu (School of Computer Science, Wuhan University); Xin Jiang (Huawei Noah's Ark Lab); Qun Liu (Huawei Noah's Ark Lab)

5456 (SLT-P44.9): Text Classification in the Wild: A Large-Scale Long-Tailed Name Normalization Dataset

Jiexing Qi (Shanghai Jiao Tong University); Shuhao Li (Shanghai Jiao Tong University); Zhixin Guo (Shanghai Jiao Tong University); Yusheng Huang (Shanghai Jiao Tong University); Chenghu Zhou (Shanghai Jiao Tong University); Weinan Zhang (Shanghai Jiao Tong University); Xinbing Wang (Shanghai Jiao Tong University); Zhouhan Lin (Shanghai Jiao Tong University)

6420 (SLT-P44.10): Good Neighbors Are All You Need for Chinese Grapheme-to-Phoneme Conversion

Jungjun Kim (DeepBrain AI Inc.); Changjin Han (DeepBrain AI Inc.); Gyuhyeon Nam (DeepBrain AI Inc.); Gyeongsu Chae (DeepBrain AI Inc.)

5325 (SLT-P44.11): MUG: A General Meeting Understanding and Generation Benchmark

Qinglin Zhang (Alibaba); Chong Deng (Alibaba inc); Jiaqing Liu (Speech Lab, Alibaba Group); Hai Yu (Alibaba); Qian Chen (Speech Lab, DAMO Academy, Alibaba Group); Wen Wang (Alibaba Group); Zhijie Yan (Alibaba Inc.); Jinglin Liu (Zhejiang University); Yi Ren (Bytedance); Zhou Zhao (Zhejiang University)

6320 (SLT-P44.12): Effectiveness of Text, Acoustic, and Lattice-based representations in Spoken Language Understanding tasks

Esaú Villatoro-Tello (Idiap Research Institute); Srikanth Madikeri (Idiap); Juan Pablo Zuluaga Gomez (Idiap Research Institute); Bidisha Sharma (Uniphore); Seyyed Saeed Sarfjoo (Idiap Research Institute); Luliia Nigmatulina (Idiap Research Institute); Petr Motlicek (Idiap); Aliaksei V. IVANOU (Uniphore Inc.); Aravind Ganapathiraju (Uniphore Software Systems Inc.)

SLT-P45: TTS: AM and Vocoder II

Room: Dome F

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Jan Skoglund

700 (SLT-P45.1): PhaseAug: A Differentiable Augmentation for Speech Synthesis to Simulate One-to-Many Mapping

Junhyeok Lee (MINDsLab Inc.); Seungu Han (MINDsLab Inc.); Hyunjae Cho (MINDsLab Inc.); Wonbin Jung (MINDsLab Inc.)

1027 (SLT-P45.2): CYFI-TTS: CYCLIC NORMALIZING FLOW WITH FINE-GRAINED REPRESENTATION FOR END-TO-END TEXT-TO-SPEECH

Insun Hwang (LG Uplus); Youngsub Han (LG Uplus); Byoung-Ki Jeon (LG Uplus)

1826 (SLT-P45.3): EGAN: A Neural Excitation Generation Model based on Generative Adversarial Networks with Harmonics and Noise Input

Yen-Ting Lin (National Taipei University); Chen Yu CHIANG (National Taipei University)

2945 (SLT-P45.4): AutoTTS: End-to-End Text-to-Speech Synthesis through Differentiable Duration Modeling

Bac Nguyen (Sony Europe B.V.); Fabien Cardinaux (Sony European Technology Center); Stefan Uhlich (Sony European Technology Center)

3410 (SLT-P45.5): A Synthetic Corpus Generation Method for Neural Vocoder Training

Zilin Wang (Tsinghua University); peng liu (transsion); Jun Chen (Tsinghua University); Sipan Li (Tsinghua University); Baijin Feng (TAL Education Group); He Gang (TAL Education Group); Zhiyong Wu (Tsinghua University); Helen Meng (The Chinese University of Hong Kong)

3486 (SLT-P45.6): Framewise WaveGAN: High Speed Adversarial Vocoder in Time Domain with Very Low Computational Complexity

Ahmed Mustafa (Amazon); Jean-Marc Valin (Amazon); Jan Bütthe (Amazon); Paris Smaragdis (Amazon); Michael M Goodwin (AWS)

4154 (SLT-P45.7): SYNTACC : Synthesizing multi-accent speech by weight factorization

Tuan-Nam Nguyen (Karlsruhe Institute of Technology); Quan Pham (Karlsruhe Institute of Technology); Alexander Waibel (Karlsruhe Institute of Technology (KIT))

4166 (SLT-P45.8): Modelling low-resource accents without accent-specific TTS frontend

Georgi Tinchev (Amazon); Marta Czarnowska (Amazon); Kamil Deja (Warsaw University of Technology); Kayoko Yanagisawa (Amazon); Marius Cotescu (Amazon)

Friday, June 9

4480 (SLT-P45.9): TEXT-TO-SPEECH SYNTHESIS FROM DARK DATA WITH EVALUATION-IN-THE-LOOP DATA SELECTION
Kentaro Seki (The University of Tokyo); Shinnosuke Takamichi (The University of Tokyo); Takaaki Saeki (The University of Tokyo); Hiroshi Saruwatari (The University of Tokyo)

5990 (SLT-P45.10): AE-Flow: AutoEncoder Normalizing Flow

Jakub Mosiński (Amazon); Piotr Bilinski (Amazon); Thomas Merritt (Amazon); Abdelhamid Ezzer (Amazon); Daniel Korzekwa (Nvidia)

6418 (SLT-P45.11): DSPGAN: a GAN-based universal vocoder for high-fidelity TTS by time-frequency domain supervision from DSP

Kun Song (Northwestern Polytechnical University); yongmao zhang (Audio, Speech and Language Processing Group (ASLP@NPU), School of Computer Science, Northwestern Polytechnical University, Xi'an, China); Yi Lei (Northwestern Polytechnical University); Jian Cong (Northwestern Polytechnical University); Hanzhao Li (Northwestern Polytechnical University); Lei Xie (NWPUP); Gang He (TAL Education Group); Jinfeng Bai (TAL Education Group)

6776 (SLT-P45.12): Investigation of Japanese PnG BERT Language Model in Text-to-Speech Synthesis for Pitch Accent Language (SPS Journal Paper)*

Yusuke Yasuda (Nagoya university); Tomoki Toda (Nagoya University)

SLT-P46: Voice and Style Conversion II

Room: Dome G

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Rui Liu,

253 (SLT-P46.1): ACE-VC: Adaptive and Controllable Voice Conversion using Explicitly Disentangled Self-supervised Speech Representations

Shehzeen S Hussain (UCSD); Paarth Neekhara (UCSD); Jocelyn Huang (NVIDIA); Jason Li (NVIDIA); Boris Ginsburg (NVIDIA)

672 (SLT-P46.2): DELIVERING SPEAKING STYLE IN LOW-RESOURCE VOICE CONVERSION WITH MULTI-FACTOR CONSTRAINTS

Zhichao Wang (Northwestern Polytechnical University); Xinsheng Wang (Northwestern Polytechnical University); Lei Xie (NWPUP); yuanzhe chen (Bytedance); qiao tian (ByteDance); wang yuping (bytedance)

1025 (SLT-P46.3): Improving Prosody for Cross-Speaker Style Transfer by Semi-Supervised Style Extractor and Hierarchical Modeling in Speech Synthesis

Chunyu Qiang (Kwai); Peng Yang (Kwai); Hao Che (Kwai); Ying Zhang (Kwai); Xiaorui Wang (Kwai); Zhongyuan Wang (Kwai)

1398 (SLT-P46.4): Streaming Voice Conversion Via Intermediate Bottleneck Features And Non-streaming Teacher Guidance

yuanzhe chen (Bytedance); Ming Tu (ByteDance AI Lab); Tang Li (ByteDance Ltd); Xin Li (ByteDance); Qiuqiang Kong (ByteDance); Jiixin Li (ByteDance); Zhichao Wang (ByteDance); qiao tian (ByteDance); wang yuping (bytedance); Yuxuan Wang (ByteDance AI Lab)

1512 (SLT-P46.5): Do Prosody Transfer Models Transfer Prosody?

Atli Thor Sigurgeirsson (University of Edinburgh); Simon King (University of Edinburgh)

2775 (SLT-P46.6): LIM-VC: A LIGHT WEIGHT VOICE CONVERSION MODEL WITH MUTUAL INFORMATION DISENTANGLEMENT

Liangjie Huang (Beijing Language and Culture University); Tian Yuan (Baidu (China) Co., Ltd); Yunming Liang (Baidu (China) Co., Ltd); Zeyu Chen (Baidu, Inc.); Can Wen (Baidu (China) Co., Ltd); Yanlu Xie (Beijing Language and Culture University); Jinsong Zhang (Beijing Language and Culture University); dengfeng ke (blcu.edu.cn)

3500 (SLT-P46.7): NONPARALLEL EMOTIONAL VOICE CONVERSION FOR UNSEEN SPEAKER-EMOTION PAIRS USING DUAL DOMAIN ADVERSARIAL NETWORK & VIRTUAL DOMAIN PAIRING

Nirmesh J Shah (Sony Research India); Mayank Kumar Singh (Sony Research India); Naoya Takahashi (Sony Group); Naoyuki Onoe (Sony)

4242 (SLT-P46.8): Hiding speaker's sex in speech using zero-evidence speaker representation in an analysis/synthesis pipeline

Paul-Gauthier Noé (Avignon University); Xiaoxiao Miao (national institute of informatics); Xin Wang (National Institute of Informatics); Junichi Yamagishi (National Institute of Informatics); Jean-Francois Bonastre (Université d'Avignon); Driss Matrouf (Avignon University)

5439 (SLT-P46.9): DVQVC: AN UNSUPERVISED ZERO-SHOT VOICE CONVERSION FRAMEWORK

Dayong Li (Westlake University); xian li (westlake university); Xiaofei Li (Westlake University)

5568 (SLT-P46.10): VQ-CL: Learning disentangled speech representations with contrastive learning and vector quantization

Huaizhen Tang (University of Science and Technology of China); Xulong Zhang (Ping An Technology (Shenzhen) Co., Ltd.); Jianzong Wang (Ping An Technology (Shenzhen) Co., Ltd); Ning Cheng (Ping An Technology (Shenzhen) Co., Ltd); Jing Xiao (Ping An Insurance (Group) Company of China)

Friday, June 9

5932 (SLT-P46.11): ANY-TO-ANY VOICE CONVERSION WITH F0 AND TIMBRE DISENTANGLEMENT AND NOVEL TIMBRE CONDITIONING

Sudheer Kumar Kovela (Nvidia); Rafael Valle (NVIDIA); Ambrish Dantrey (Nvidia); Bryan Catanzaro (NVIDIA)

6774 (SLT-P46.12): A Comparative Study of Self-Supervised Speech Representation Based Voice Conversion (SPS Journal Paper)*

Wen-Chin Huang (Nagoya University); Shu-wen Yang (National Taiwan University); Tomoki Hayashi (Nagoya University); Tomoki Toda (Nagoya University)

SPCN-P5: Sensing, Computing, and Semantic Communications

Room: Dome H

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Apoorva Chawla

4317 (SPCN-P5.1): Radio Map based UAV Target Localization

Chen He (Northwest University); GONG WEISHENG (Northwest University); Yangrui Dong (Northwest University); Xie Xie (Northwest University); Z. Jane Wang (University of British Columbia)

3004 (SPCN-P5.2): Integrated Sensing and Full-Duplex Communication: Joint Transceiver Beamforming and Power Allocation

Zhenyao He (Southeast University); Wei Xu (Southeast University); Hong Shen (Southeast University); Derrick Wing Kwan Ng (University of New South Wales); Yonina C. Eldar (Weizmann Institute of Science); Xiaohu You (Southeast University)

2457 (SPCN-P5.3): SCALABLE MULTI-TASK SEMANTIC COMMUNICATION SYSTEM WITH FEATURE IMPORTANCE RANKING

Jiangjing Hu (Beijing University of Posts and Telecommunications); Fengyu WANG (Beijing University of Posts and Telecommunications); Wenjun Xu (Beijing University of Posts and Telecommunications); Hui Gao (Beijing University of Posts and Telecommunications); Ping Zhang (Beijing University of Posts and Telecommunications)

6150 (SPCN-P5.4): WITT: A Wireless Image Transmission Transformer For Semantic Communications

Ke Yang (Beijing University of Posts and Telecommunications); Sixian Wang (Beijing University of Posts and Telecommunications); Jincheng Dai (Beijing University of Posts and Telecommunications); Kailin Tan (Beijing University of Posts and Telecommunications); kai niu (Beijing University of Posts and Telecommunications); Ping Zhang (Beijing University of Posts and Telecommunications)

4635 (SPCN-P5.5): Sparse Bayesian Learning Assisted Decision Fusion in Millimeter Wave Massive MIMO Sensor Networks

Apoorva Chawla (Norwegian University of Science and Technology); Domenico Ciuonzo (University of Naples Federico II); Pierluigi Salvo Rossi (NTNU)

5174 (SPCN-P5.6): TDMA-Based Multi-User Binary Computation Offloading in the Finite-Block-Length Regime

Mohammad Amin Manouchehrpour (McMaster University); Harvinder Lehal (McMaster University); Mahsa Salmani (McMaster University); Timothy N Davidson (McMaster University)

5212 (SPCN-P5.7): Managing Information Updating with Edge Computing: A Distributed and Learning Approach

Junyi He (Beijing Jiaotong University); Di Zhang (Beijing Jiaotong University); Shumeng Liu (Beijing Jiaotong University); Yuezhi Zhou (Tsinghua University); Yaoxue Zhang (Tsinghua University)

5345 (SPCN-P5.8): Multiple Access Computation Offloading for the K-user Case

Xiaomeng Liu (McMaster University); Christian Schaible (McMaster University); Timothy N Davidson (McMaster University)

5155 (SPCN-P5.9): Comparative Study of IRS Assisted Opportunistic Communications Over i.i.d. and LoS Channels

L Yashvanth (Indian Institute of Science, Bangalore); Chandra Murthy (Indian Institute of Science)

SPTM-P11: Sparsity and Low-Rank Models

Room: Dome L

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Georgios Giannakis, Arnaud Breloy

5332 (SPTM-P11.1): Higher-order Link Prediction via Learnable Maximum Mean Discrepancy

Georgios V. Karanikolas (University of Minnesota); Alba Pagès Zamora (Universitat Politècnica de Catalunya); Georgios B. Giannakis (University of Minnesota)

6690 (SPTM-P11.2): Diffusion Particle Filtering on the Special Orthogonal Group Using Lie Algebra Statistics (SPS Journal Paper)*

CLAUDIO JOSE BORDIN JUNIOR (Universidade Federal do ABC); Caio Gomes de Figueiredo (Instituto Federal do Ceará); Marcelo G S Bruno (ITA)

Friday, June 9

7171 (SPTM-P11.3): Scalable and Privacy-aware Online Learning of Nonlinear Structural Equation Models

Rohan Money (UiA); Joshin P. Krishnan (Simula Metropolitan Center for Digital Engineering); Baltasar Beferull-Lozano (University of Agder); Elvin Isufi (Tu Delft)

6711 (SPTM-P11.4): The Mirror Transform (SPS Journal Paper)*

Fabrizio Guerrini (University of Brescia); Alessandro Gnutti (University of Brescia); Riccardo Leonardi (UNIBS)

6744 (SPTM-P11.5): Signal Useful Information Recovery by Overlapping Supports of Time-Frequency Representations (SPS Journal Paper)*

Nicoletta Saulig (Faculty of Engineering, University of Pula); Miloš Milovanović (Mathematical Institute of the Serbian Academy of Sciences and Arts National Institute of the Republic of Serbia)

6789 (SPTM-P11.6): Binary Signal Perfect Recovery from partial DFT coefficients (SPS Journal Paper)*

Soo-Chang Pei (National Taiwan University); Kuo-Wei Chang (Chunghwa Telecom)

2564 (SPTM-P11.7): Column-based matrix approximation with quasi-polynomial structure

Jeongmin Chae (University of Southern California); Praneeth Narayanamurthy (University of Southern California); Selin Bac (University of Southern California); Shaama Mallikarjun Sharada (University of Southern California); Urbashi Mitra (USC)

5373 (SPTM-P11.8): Fast robust principle component analysis using Gauss-Newton iterations

William Chettleburgh (Michigan State University); Zhishen Huang (Amazon Inc.); Ming Yan (The Chinese University of Hong Kong, Shenzhen)

6159 (SPTM-P11.9): A SIMPLE SCHEME FOR COUPLED FACTORIZATION FOR HYPERSPECTRAL SUPER-RESOLUTION: EXPLOITING SPARSITY IN AN EASY WAY

Yuening Li (The Chinese University of Hong Kong); Wing-Kin Ma (The Chinese University of Hong Kong); Ruiyuan Wu (Meituan); Huikang Liu (Shanghai University of Finance and Economics)

3107 (SPTM-P11.10): ROBUST AND GLOBALLY SPARSE PCA VIA MAJORIZATION-MINIMIZATION AND VARIABLE SPLITTING

Hugo Brehier (SONDRRA, CentraleSupélec); Arnaud Breloy (Université Paris Nanterre); Mohammed Nabil EL KORSO (Paris Nanterre University); Sandeep Prof. Kumar (IIT Delhi)

5300 (SPTM-P11.11): An Online Algorithm for Contrastive Principal Component Analysis

Siavash Golkar (Flatiron Institute); David Lipshutz (Flatiron Institute); Tiberiu Tesileanu (Flatiron Institute); Dmitri Chklovskii (Flatiron Institute)

SPTM-P6: Signal Processing Over Graphs

Room: Dome I

Type: Poster

02:00 PM to 03:30 PM

Chair(s): Antonio Ortega, Antonio Marques

592 (SPTM-P6.1): Pooling Strategies for Simplicial Convolutional Networks

Domenico Mattia Cinque (Sapienza University of Rome); Claudio Battiloro (Sapienza University of Rome); Paolo Di Lorenzo (Sapienza University of Rome)

965 (SPTM-P6.2): SPACE-TIME VARIABLE DENSITY SAMPLINGS FOR SPARSE BANDLIMITED GRAPH SIGNALS DRIVEN BY DIFFUSION OPERATORS

Qing Yao (UCSB); Longxiu Huang (Michigan State University); Sui Tang (UCSB)

2415 (SPTM-P6.3): GRAPH LEARNING FROM GAUSSIAN AND STATIONARY GRAPH SIGNALS

Andrei Buciualea Vlas (Universidad Rey Juan Carlos); Antonio G. Marques (King Juan Carlos University)

2822 (SPTM-P6.4): Performance of Social Machine Learning under Limited Data

Ping Hu (EPFL); Virginia Bordignon (EPFL); Mert Kayaalp (Ecole Polytechnique Fédérale de Lausanne); Ali H. Sayed (Ecole Polytechnique Fédérale de Lausanne)

2947 (SPTM-P6.5): Central nodes detection from partially observed graph signals

Yiran HE (The Chinese University of Hong Kong); Hoi-To Wai (Chinese University of Hong Kong)

3027 (SPTM-P6.6): Smoothing complex-valued signals on Graphs with Monte-Carlo

Hugo Jaquard (GIPSA-lab); Michaël Fanuel (CRISTAL); Pierre-Olivier Amblard ("CNRS, Grenoble"); Rémi Bardenet (CRISTAL); Simon Barthelmé (CNRS); Nicolas Tremblay (CNRS)

3569 (SPTM-P6.7): Eigen-Decomposition-Free Directed Graph Sampling via Gershgorin Disc Alignment

Yuejiang Li (Tsinghua University); H. Vicky Zhao (Tsinghua University); Gene Cheung (York University)

4892 (SPTM-P6.8): Convolutional Filtering on Sampled Manifolds

Zhiyang Wang (University of Pennsylvania); Luana Ruiz (MIT CSAIL); Alejandro Ribeiro (University of Pennsylvania)

Friday, June 9

4901 (SPTM-P6.9): Tangent Bundle Filters and Neural Networks: from Manifolds to Cellular Sheaves and Back
Claudio Battiloro (Sapienza University of Rome); Zhiyang Wang (University of Pennsylvania); Hans Riess (Duke University); Paolo Di Lorenzo (Sapienza University of Rome); Alejandro Ribeiro (University of Pennsylvania)

4953 (SPTM-P6.10): Estimating and Analyzing Neural Information Flow Using Signal Processing on Graphs
Felix Schwock (University of Washington); Julien Bloch (University of Washington); Les Atlas (University of Washington); Shima Abadi (University of Washington); Azadeh Yazdan-Shahmorad (University of Washington)

5165 (SPTM-P6.11): Towards bandwidth estimation for graph signal reconstruction
Ajinkya Jayawant (University of Southern California); Antonio Ortega (University of Southern California)

AASP-L7: Target Source Extraction
Room: Salon de Roses A
Type: Oral
03:40 PM to 05:10 PM
Chair(s): Nobutaka Ito, Gordon Wichern

03:40 PM

3894 (AASP-L7.1): REAL-TIME MULTICHANNEL SPEECH SEPARATION AND ENHANCEMENT USING A BEAMSPACE-DOMAIN-BASED LIGHTWEIGHT CNN

Marco Olivieri (Politecnico di Milano); Luca Comanducci (Politecnico di Milano); Mirco Pezzoli (Politecnico di Milano); Davide Balsarri (BdSound); Luca Menescardi (BdSound); Michele Buccoli (BdSound S.r.l.); Simone Pecorino (BdSound); Antonio Grosso (BdSound); Fabio Antonacci (Politecnico di Milano); Augusto Sarti (Politecnico di Milano)

03:55 PM

3700 (AASP-L7.2): REAL-TIME TARGET SOUND EXTRACTION

Bandhav Veluri (University of Washington); Justin Chan (University of Washington); Malek Itani (University of Washington); Tuochoa Chen (University of Washington); Takuya Yoshioka (Microsoft); Shyamnath Gollakota (University of Washington)

04:10 PM

4894 (AASP-L7.3): Optimal Condition Training for Target Source Separation

Efthymios Tzinis (University of Illinois at Urbana-Champaign); Gordon Wichern (Mitsubishi Electric Research Laboratories (MERL)); Paris Smaragdis (University of Illinois at Urbana-Champaign); Jonathan LeRoux (Mitsubishi Electric Research Laboratories (MERL))

04:25 PM

4591 (AASP-L7.4): Target Sound Extraction with Variable Cross-modality Clues

Chenda Li (Shanghai Jiao Tong University); Yao Qian (Microsoft); Zhuo Chen (Microsoft); Dongmei Wang (Microsoft); Takuya Yoshioka (Microsoft); Shujie Liu (Microsoft Research Asia); Yanmin Qian (Shanghai Jiao Tong University); Michael Zeng (Microsoft)

04:40 PM

897 (AASP-L7.5): Breaking the trade-off in personalized speech enhancement with cross-task knowledge distillation

Hassan Taherian (The Ohio State University); Sefik Emre Eskimez (Microsoft); Takuya Yoshioka (Microsoft)

04:55 PM

6704 (AASP-L7.6): SoundBeam: Target Sound Extraction Conditioned on Sound-Class Labels and Enrollment Clues for Increased Performance and Continuous Learning (SPS Journal Paper)*

Marc Delcroix (NTT); Jorge Bennisar Vázquez (NTT); Tsubasa Ochiai (NTT); Keisuke Kinoshita (Google); Yasunori Ohishi (NTT); Shoko Araki (NTT Corporation)

AASP-L8: Music Generation and Arrangement

Room: Salon de Roses B
Type: Oral
03:40 PM to 05:10 PM
Chair(s): Juan Pablo Bello, Kyogu Lee

03:40 PM

339 (AASP-L8.1): Pop2Piano : Pop Audio-based Piano Cover Generation

Jongho Choi (rebellions Inc.); Kyogu Lee (Seoul National University)

03:55 PM

1415 (AASP-L8.2): Neural Band-to-Piano Score Arrangement with Stepless Difficulty Control

Moyu Terao (Kyoto University); Eita Nakamura (Kyoto University); Kazuyoshi Yoshii (Kyoto University)

04:10 PM

407 (AASP-L8.3): Multitrack Music Transformer

Hao-Wen Dong (University of California San Diego); Ke Chen (University of California San Diego); Shlomo Dubnov (UC San Diego); Julian McAuley (University of California, San Diego); Taylor Berg-Kirkpatrick (UCSD)

Friday, June 9

04:25 PM

6488 (AASP-L8.4): MUSIC REARRANGEMENT USING HIERARCHICAL SEGMENTATION

Christos Plachouras (Universitat Pompeu Fabra); Marius Miron (Music Technology Group, Universitat Pompeu Fabra)

04:40 PM

6740 (AASP-L8.5): Use of Speaker Recognition Approaches for Learning and Evaluating Embedding Representations of Musical Instrument Sounds (SPS Journal Paper)*

Xuan Shi (University of Southern California); Erica Cooper (); Junichi Yamagishi (National Institute of Informatics)

GC-15: Multimodal Information Based Speech Processing (MISP) 2022 Challenge Proposal

Room: Nefeli B

Type: Oral

03:40 PM to 05:10 PM

Chair(s): TBA

03:40 PM

6619 (GC-L15.1): Introduction

Hang Chen (USTC); Shilong Wu (University of Science and Technology of China); Yusheng Dai (University of Science and Technology of China); Zhe Wang (University of Science and Technology of China); Jun Du (University of Science and Technology of China); Chin-hui Lee (Georgia Institute of Technology); Jingdong Chen (Northwestern Polytechnical University); Shinji Watanabe (Carnegie Mellon University); Sabato M Siniscalchi (Kore University of Enna); Odette Scharenborg (Multimedia Computing Group, Delft University of Technology); Diyuang Liu (USTC,iFLYTEK); Baocai Yin (USTC,iFLYTEK); Jia Pan (iFlytek Research); Jianqing Gao (iFLYTEK); Cong Liu (iFLYTEK Research)

04:00 PM

6829 (GC-L15.2): Multi-Speaker End-to-end Multi-modal Speaker Diarization System for the MISP 2022 CHALLENGE

Tao Liu (Shanghai Jiao Tong University); Zhengyang Chen (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University); Kai Yu (Shanghai Jiao Tong University)

04:12 PM

6848 (GC-L15.3): The NPU-ASLP System for Audio-Visual Speech Recognition in MISP 2022 Challenge

Pengcheng Guo (Northwestern Polytechnical University); He Wang (NWPU); Bingshen Mu (Northwestern Polytechnical University); Ao Zhang (Northwestern Polytechnical University); Peikun Chen (Northwestern Polytechnical University)

04:24 PM

6867 (GC-L15.4): The WHU-Alibaba Audio-Visual Speaker Diarization System for the MISP 2022 Challenge

Ming Cheng (Duke Kunshan University); Haoxu Wang (Wuhan University); Ziteng Wang (Alibaba Group); Qiang Fu (Alibaba Group); Ming Li (Duke Kunshan University)

04:36 PM

6905 (GC-L15.5): The XMU system for audio-visual diarization and recognition in MISP challenge 2022

Tao Li (Xiamen University); Haodong Zhou (Xiamen University); Jie Wang (Xiamen University); Qingyang Hong (Xiamen University); Lin Li (Xiamen University)

04:48 PM

6923 (GC-L15.6): THE NIO SYSTEM FOR AUDIO-VISUAL DIARIZATION AND RECOGNITION IN MISP CHALLENGE 2022

Gaopeng Xu (nio); Xianliang Wang (nio); Sang Wang (nio); junfeng yuan (nio); Wei Guo (nio); Wei Li (nio); Jie Gao (nio)

IVMSP-L13: Image Retrieval and Classification

Room: Athina

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Hujun Bao, Juan Liu

03:40 PM

100 (IVMSP-L13.1): Counterfactual Two-stage Debiasing for Video Corpus Moment Retrieval

Sunjae Yoon (KAIST); Ji Woo Hong (KAIST); SooHwan Eom (KAIST); Hee Suk Yoon (KAIST); Eunseop Yoon (KAIST); Daehyeok Kim (KAIST); Junyeong Kim (Chung-Ang University); Chanwoo Kim (Samsung Electronics); Chang D. Yoo (KAIST)

03:55 PM

528 (IVMSP-L13.2): MABNet: Master Assistant Buddy Network with Hybrid Learning for Image Retrieval

Rohit Agarwal (UiT The Arctic University of Norway, Tromsø); Gyanendra Das (Indian Institute of Technology, Dhanbad); Saksham Aggarwal (IIT (ISM) Dhanbad); Alexander Horsch (UiT The Arctic University of Norway); Dilip K Prasad (UiT The Arctic University of Norway)

Friday, June 9

04:10 PM

2993 (IVMSP-L13.3): A2SConv: Asymmetric Spectral-Spatial Neural Architecture Search for Hyperspectral Image

Classification

Zhan Lin (School of Information Science and Technology, Fudan); Jiayuan Fan (Fudan University); peng ye (fudan university); Cao Jianjian (Fudan University)

04:25 PM

3874 (IVMSP-L13.4): Masked-AP: Attention Pyramid Convolutional Neural Network with mask for Cervical Cell

Classification

yu jin (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Juan Liu (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Hua Chen (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Wensi Duan (Institute of Artificial Intelligence, School of Computer Science, Wuhan University); Dehua Cao (Landing Artificial Intelligence Center for Pathological Diagnosis); Baochuan Pang (Landing Artificial Intelligence Center for Pathological Diagnosis)

04:40 PM

4300 (IVMSP-L13.5): An Auto-Encoder Based Method for Camera Fingerprint Compression

Kaixuan Zhang (Shanghai Jiao Tong University); Zihan Liu (Shanghai Jiao Tong University); Jiashang Hu (Shanghai Jiao Tong University); shilin wang (SEIEE, Shanghai Jiaotong University)

04:55 PM

760 (IVMSP-L13.6): CRGAN: Learning image-to-image translation of two unpaired images by cross-attention mechanism

Chengyuan Xu (Northwestern Polytechnical University); Kang Liu (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University)

MLSP-L5: Adversarial Machine Learning I

Room: Delphi

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Sijia Liu

03:40 PM

829 (MLSP-L5.1): IMPROVING ADVERSARIAL ROBUSTNESS WITH HYPERSPHERE EMBEDDING AND ANGULAR-BASED REGULARIZATIONS

Olukorede J Fakorede (Iowa State University); Ashutosh Nirala (Iowa State University); Modeste Atsague (Iowa State University); Jin Tian (Iowa State University)

03:55 PM

5666 (MLSP-L5.2): Forensics for Adversarial Machine Learning through Attack Mapping Identification

Allen H Yan (Oregon State University); Jinsub Kim (""); Raviv Raich (Oregon State University)

04:10 PM

623 (MLSP-L5.3): Measuring the Transferability of L-infty Attacks by the L-2 Norm

Sizhe Chen (Shanghai Jiao Tong University); Qinghua Tao (KU Leuven); Zhixing Ye (Shanghai Jiao Tong University); Xiaolin Huang (Shanghai Jiao Tong University)

04:25 PM

6016 (MLSP-L5.4): A Game of Snakes and GANs

Siddarth Asokan (Indian Institute of Science); Fatwir Sheikh Mohammed (University of Washington); Chandra Sekhar Seelamantula (IISc Bangalore)

04:40 PM

6181 (MLSP-L5.5): BOOSTING TRANSFERABILITY OF ADVERSARIAL EXAMPLE VIA AN ENHANCED EULER'S METHOD

Anjie Peng (Southwest University of Science and Technology); Zhi Lin (Southwest University of Science and Technology); Hui Zeng (Southwest University of Science and Technology); Wenxin Yu (Southwest University of Science and Technology); Xiangui Kang (Sun Yat-Sen University)

04:55 PM

5101 (MLSP-L5.6): Robustness-preserving Lifelong Learning via Dataset Condensation

jinghan jia (Michigan state university); Yihua Zhang (Michigan State University); Dogyoon Song (University of Michigan); Sijia Liu (Michigan State University); Alfred Hero (University of Michigan)

Friday, June 9

SPTM-L4: Signal Processing Over Networks I

Room: Jupiter

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Abdelhak Zoubir, Visa Koivunen

03:40 PM

1008 (SPTM-L4.1): COMPRESSED DISTRIBUTED REGRESSION OVER ADAPTIVE NETWORKS

Marco Carpentiero (University of Salerno); Vincenzo Matta (DIEM, University of Salerno); Ali H. Sayed (Ecole Polytechnique Fédérale de Lausanne)

03:55 PM

2218 (SPTM-L4.2): Convergence of Stochastic PDMM

Sebastian Jordan (TU Delft); Thomas Sherson (Delft University of Technology); Richard Heusdens (Netherlands Defence Academy)

04:10 PM

4570 (SPTM-L4.3): Simplicial Vector Autoregressive Model for Streaming Edge Flows

Joshin P. Krishnan (Simula Metropolitan Center for Digital Engineering); Rohan Money (UiA); Baltasar Beferull-Lozano (University of Agder); Elvin Isufi (Tu Delft)

04:25 PM

4585 (SPTM-L4.4): Dynamic Distributed Convex Optimization "Over-the-Air" in Decentralized Wireless Networks

Navneet Agrawal (TU Berlin); Renato Luis Garrido Cavalcante (Fraunhofer Heinrich Hertz Inst); Slawomir Stanczak (TU Berlin)

04:40 PM

5580 (SPTM-L4.5): Kernel Ridge Regression for Generalized Graph Signal Processing

Xingchao Jian (School of Electrical and Electronic Engineering, Nanyang Technological University); Wee Peng Tay (Nanyang Technological University)

04:55 PM

5951 (SPTM-L4.6): SPATIAL INFERENCE USING CENSORED MULTIPLE TESTING WITH FDR CONTROL

Martin Götz (Technische Universität Darmstadt); Abdelhak M Zoubir (Technische Universität Darmstadt); Visa Koivunen (Aalto university)

SS-L24: Variational Inference and Approximate Bayesian Techniques

Room: Nafsika A

Type: Oral

03:40 PM to 05:10 PM

Chair(s): Dirk Stöck

03:40 PM

838 (SS-L24.1): Long-Memory Message-Passing for Spatially Coupled Systems

Keigo Takeuchi (Toyohashi University of Technology)

03:55 PM

1444 (SS-L24.2): A unitary transform based generalized approximate message passing

Jiang Zhu (Zhejiang University); Xiangming Meng (The University of Tokyo); Xupeng Lei (Zhejiang University); Qinghua Guo (UNIVERSITY OF WOLLONGONG)

04:10 PM

1782 (SS-L24.3): QUANTUM VARIATIONAL BAYES ON MANIFOLDS

Anna Lopatnikova (U of Sydney); Minh-Ngoc Tran (U of Sydney)

04:25 PM

2545 (SS-L24.4): Overcoming Posterior Collapse in Variational Autoencoders via EM-type Training

Ying Li (The University of Hong Kong); Lei Cheng (Zhejiang University); Feng Yin (The Chinese University of Hong Kong, Shenzhen); Michael Zhang (University of Hong Kong); Sergios Theodoridis (National and Kapodistrian University of Athens)

04:40 PM

3413 (SS-L24.5): Alternating Constrained Minimization based Approximate Message Passing

Christo Kurisummoottil Thomas (Virginia Tech); Dirk Stöck (EURECOM, France)

04:55 PM

6098 (SS-L24.6): Variational Bayesian Channel Estimation in Wideband Multi-Scale Multi-Lag Channels

Niladri Halder (Indian Institute of Science); Arunkumar K. P. (Indian Institute of Science); Chandra Murthy (Indian Institute of Science)

Friday, June 9

AASP-P17: Spatial Audio Recording and Reproduction

Room: Dome A

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Mark Thomas

3513 (AASP-P17.1): The R3VIVAL dataset: Repository of room responses and 360 videos of a variable acoustics room
Florian Klein (TU Ilmenau); Sebastia V. Amengual Garí (Reality Labs Research, Meta)

1785 (AASP-P17.2): Robust FIR Filters for Wireless Low-frequency Sound Zones
Mo Zhou (Aalborg University); Martin Bo Møller (Bang & Olufsen); Christian Sejer Pedersen (Aalborg University); Jan Ostergaard (Aalborg University)

2387 (AASP-P17.3): TT-NET: DUAL-PATH TRANSFORMER BASED SOUND FIELD TRANSLATION IN THE SPHERICAL HARMONIC DOMAIN
Yiwen Wang (Peking University); Zijian Lan (Peking University); Xihong Wu (Peking University); Tianshu Qu (Peking University)

1884 (AASP-P17.4): Kernel interpolation of acoustic transfer functions with adaptive kernel for directed and residual reverberations
Juliano G. C. Ribeiro (The University of Tokyo); Shoichi Koyama (The University of Tokyo); Hiroshi Saruwatari (The University of Tokyo)

1860 (AASP-P17.5): Spatial active noise control method based on sound field interpolation from reference microphone signals
Kazuyuki Arikawa (The University of Tokyo); Shoichi Koyama (The University of Tokyo); Hiroshi Saruwatari (The University of Tokyo)

4586 (AASP-P17.6): STUDY AND DESIGN OF ROBUST PERSONAL SOUND ZONES WITH VAST USING LOW RANK RIRs
Sankha Subhra Bhattacharjee (Audio Analysis Lab, CREATE, Aalborg University); Liming Shi (CIE, Chongqing University of Posts and Telecommunications); Guoli Ping (Acoustic Engineering Lab, Huawei Technologies Co., Ltd); Xiaoxiang Shen (Acoustic Engineering Lab, Huawei Technologies Co., Ltd); Mads G. Christensen (Audio Analysis Lab., AD:MT, Aalborg University, Denmark)

7161 (AASP-P17.7): Signal-to-interference-plus-noise ratio based optimization for sound zone control
Jesper Brunström (KU Leuven); Toon van Waterschoot (Department of Electrical Engineering (ESAT-STADIUS/ETC)); Marc Moonen (KU Leuven)

6682 (AASP-P17.8): Autonomous In-Situ Soundscape Augmentation via Joint Selection of Masker and Gain (SPS Journal Paper)*
Karn N Watcharasupat (Georgia Institute of Technology); Kenneth Ooi (Nanyang Technological University); Bhan Lam (NTU); Trevor Wong (Nanyang Technological University); Zhen-Ting Ong (Nanyang Technological University); Woon Seng Gan (NTU)

6732 (AASP-P17.9): Gridless 3D Recovery of Image Sources From Room Impulse Responses (SPS Journal Paper)*
Tom Sprunck (Inria); Antoine Deleforge (INRIA); Yannick Privat (IRMA); Cédric Foy (UMRAE)

6752 (AASP-P17.10): Amplitude Matching for Multizone Sound Field Control (SPS Journal Paper)*
Takumi Abe (The University of Tokyo); Shoichi Koyama (National Institute of Informatics); Natsuki Ueno (Tokyo Metropolitan University); Hiroshi Saruwatari (The University of Tokyo)

6800 (AASP-P17.11): Parametric Ambisonic Encoding of Arbitrary Microphone Arrays (SPS Journal Paper)*
Leo McCormack (Aalto University); Archontis Politis (Tampere University); Raimundo Gonzalez Diaz (Aalto University); Tapio Lokki (Aalto University); Ville T Pulkki (Aalto University)

AASP-P18: Speech Modeling and Audio Coding

Room: Dome B

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Jan Skoglund,

2877 (AASP-P18.1): Speech Enhancement with Intelligent Neural Homomorphic Synthesis
Shulin He (College of Computer Science, Inner Mongolia University); Wei Rao (Tencent); Jinjiang Liu (College of Computer Science, Inner Mongolia University); Jun Chen (Tencent); Yukai Ju (Tencent); Xueliang zhang (Inner Mongolia University); Yannan Wang (Tencent); Shi-dong Shang (tencent)

4051 (AASP-P18.2): F0 ESTIMATION FROM TELEPHONE SPEECH USING DEEP FEATURE LOSS
Supriya M Shetty (Indian Institute of Information Technology, Dharwad); Shraddha Revankar (K L E Technological University); Nalini Iyer ("KLETech, Hubballi"); Deepak T (IIIT-Dharwad)

3718 (AASP-P18.3): LEARNING TO AUTO-CORRECT FOR HIGH-QUALITY SPECTROGRAMS
Zhiyang Zhou (Beijing Bombax Xiaolce Technology Co., Ltd); Shihui Liu (Beijing Bombax Xiaolce Technology Co., Ltd)

Friday, June 9

1516 (AASP-P18.4): Spherical vector quantization for spatial direction coding

Stéphane Ragot (Orange); Adriana Vasilache (Nokia Technologies)

1117 (AASP-P18.5): Audio Coding With Unified Noise Shaping And Phase Contrast Control

Byeongho Jo (Electronics and Telecommunications Research Institute); Seung-Kwon Beack (IEEE Broadcast Technology Society (BTS)); Taejin Lee (ETRI)

1825 (AASP-P18.6): HybridFormer: Improving SqueezeFormer with Hybrid Attention and NSR Mechanism

Yuguang Yang (Ximalaya Inc., ShangHai, China); Yu Pan (University of Alberta); Jingjing Yin (Ximalaya); jiangyu Han (Ximalaya); Lei Ma (University of Alberta); heng lu (Ximalaya Inc., ShangHai, China)

4577 (AASP-P18.7): SPEECH MODELING WITH A HIERARCHICAL TRANSFORMER DYNAMICAL VAE

Xiaoyu Lin (Inria Grenoble-Rhône-Alpes); Xiaoyu Bie (INRIA); Simon Leglaive (CentraleSupélec); Laurent Girin (); Xavier Alameda-Pineda (INRIA)

4946 (AASP-P18.8): MASKED AUTOENCODERS ARE ARTICULATORY LEARNERS

Ahmed A Attia (University Of Maryland College Park); Carol Y Espy-Wilson (University of Maryland)

6101 (AASP-P18.9): Wireless Deep Speech Semantic Transmission

Zixuan Xiao (Beijing University of Posts and Telecommunications); Shengshi Yao (Beijing University of Posts and Telecommunications); Jincheng Dai (Beijing University of Posts and Telecommunications); Sixian Wang (Beijing University of Posts and Telecommunications); kai niu (Beijing University of Posts and Telecommunications); Ping Zhang (Beijing University of Posts and Telecommunications)

6433 (AASP-P18.10): Building Keyword Search System from End-to-End ASR Systems

Ruizhe Huang (Johns Hopkins University); Matthew S Wiesner (Johns Hopkins University); Paola Garcia (Johns Hopkins University); Daniel Povey (Johns Hopkins University); Jan Trmal (Johns Hopkins University); Sanjeev Khudanpur (Johns Hopkins University)

6733 (AASP-P18.11): Multitaper-Mel Spectrograms for Keyword Spotting (SPS Journal Paper)*

Khaled Jamal Bakri (SiDi); Juliana Camilo (SiDi); Fernanda Ferreira (SiDi); Douglas Souza (-)

AASP-P19: Audio Processing and Analysis

Room: Dome G

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Constantine Kotropoulos,

5468 (AASP-P19.1): Improving Music Genre Classification from Multi-Modal Properties of Music and Genre Correlations Perspective

Ganghui Ru (Fudan University); Xulong Zhang (Ping An Technology (Shenzhen) Co., Ltd.); Jianzong Wang (Ping An Technology (Shenzhen) Co., Ltd); Ning Cheng (Ping An Technology (Shenzhen) Co., Ltd); Jing Xiao (Ping An Insurance (Group) Company of China)

5884 (AASP-P19.2): NOTE AND PLAYING TECHNIQUE TRANSCRIPTION OF ELECTRIC GUITAR SOLOS IN REAL-WORLD MUSIC PERFORMANCE

TungSheng Huang (Georgia Institute of Technology); Ping-Chung Yu (National Tsing Hua University); Li Su (Academia Sinica)

6176 (AASP-P19.3): Multitrack Music Transcription with a Time-Frequency Perceiver

Wei-Tsung Lu (TikTok); Ju-Chiang Wang (TikTok); Yun-Ning Hung (TikTok)

659 (AASP-P19.4): Towards Controllable Audio Texture Morphing

Chitralekha Gupta (National University of Singapore); Purnima Kamath (National University of Singapore); Yize Wei (National University of Singapore); Zhuoyao Li (National University of Singapore); Suranga Nanayakkara (National University of Singapore); Lonce Wyse (National University of Singapore)

3088 (AASP-P19.5): Improving Text-Audio Retrieval by Text-aware Attention Pooling and Prior Matrix Revised Loss

Yifei Xin (Peking University); Dongchao Yang (Peking university); Yuexian Zou (Peking University)

2089 (AASP-P19.6): SPEECH EMOTION RECOGNITION VIA HETEROGENEOUS FEATURE LEARNING

Ke Liu (Northwest University); Dongya Wu (Northwest University); Dekui Wang (Northwest University); Jun Feng (Northwest University)

4461 (AASP-P19.7): CONTRASTIVE SPEECH MIXUP FOR LOW-RESOURCE KEYWORD SPOTTING

Dianwen Ng (Alibaba Group/Nanyang Technological University); Ruixi Zhang (National University of Singapore); Jia Qi Yip (Alibaba Group); Chong Zhang (Alibaba Group); Yukun Ma (Alibaba Group); Trung Hieu Nguyen (Alibaba Group); Chongjia Ni (Alibaba); Eng Siong Chng (Nanyang Technological University); Bin Ma ("Alibaba, Singapore R&D Center")

Friday, June 9

1587 (AASP-P19.8): Real-time speech enhancement with dynamic attention span

Chengyu Zheng (Communication University of China); Yuan Zhou (Microsoft Research Asia); Xiulian Peng (Microsoft Research Asia); Yuan Zhang (Communication University of China); Yan Lu (Microsoft Research Asia)

1953 (AASP-P19.9): Improving Speech Enhancement via Event-based Query

Yifei Xin (Peking University); Xiulian Peng (Microsoft Research Asia); Yan Lu (Microsoft Research Asia)

1606 (AASP-P19.10): CONTRAST-PLC: CONTRASTIVE LEARNING FOR PACKET LOSS CONCEALMENT

Huaying Xue (Microsoft); Xiulian Peng (Microsoft Research Asia); Yan Lu (Microsoft Research Asia)

447 (AASP-P19.11): MAID: A Conditional Diffusion Model For Long Music Audio Inpainting

Kaiyang Liu (Sichuan university); Wendong Gan (Wiz Holdings Pte Ltd); Chenchen Yuan (Sichuan university)

6688 (AASP-P19.12): A Hybrid SFANC-FxNLMS Algorithm for Active Noise Control based on Deep Learning (SPS Journal Paper)*

Zhengding Luo (Nanyang Technological University); Dongyuan Shi (NTU DSP lab); Woon Seng Gan (NTU)

BISP-P5: Physiological Signal Processing II

Room: Dome H

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Richard Hendricks

468 (BISP-P5.1): IR-ECG: Invertible Reconstruction of ECG

Peng Wang (Institute of Computing Technology); Xi Huang (Institute of computing technology of the Chinese Academy of Sciences); Li Cui (Institute of computing technology of the Chinese Academy of Sciences)

2029 (BISP-P5.2): Real-time Wireless ECG-derived Respiration Rate Estimation Using an Autoencoder with a DCT Layer

Hongyi Pan (University of Illinois Chicago); Xin Zhu (UIC); Zhilu Ye (University of Illinois Chicago); Pai-Yen Chen (University of Illinois Chicago); Ahmet E Cetin (University of Illinois at Chicago)

2466 (BISP-P5.3): ECG Artifact Removal from Single-Channel Surface EMG Using Fully Convolutional Networks

Kuan-Chen Wang (National Taiwan University); Kai-Chun Liu (Academia Sinica); Sheng-Yu Peng (National Taiwan University of Science and Technology); Yu Tsao (Academia Sinica)

2673 (BISP-P5.4): A New Approach to Extract Fetal Electrocardiogram Using Affine Combination of Adaptive Filters

Yu Xuan (University of California San Diego); Xiangyu Zhang (Johns Hopkins University); Shuyue Stella Li (Johns Hopkins University); zihan shen (University of Chinese Academy of Sciences); XIN XIE (University of California, San Diego); Paola Garcia (Johns Hopkins University); Roberto Togneri (The University of Western Australian)

2754 (BISP-P5.5): ECGT2T: Towards Synthesizing Twelve-Lead Electrocardiograms from Two Asynchronous Leads

Yong-Yeon Jo (Medical AI Inc.); Young Sang Choi (National Cancer Center); Jong-Hwan Jang (Medical AI); Joon-myung Kwon (Medical AI Co. Ltd.)

3424 (BISP-P5.6): MTDL-Net: Morphological and Temporal Discriminative Learning for Heartbeat Classification

Can Han (Shanghai Jiao Tong University); Suncheng Xiang (Shanghai Jiao Tong University); Dahong Qian (Shanghai Jiao Tong University)

6265 (BISP-P5.7): MLCGAN: MULTI-LEAD ECG SYNTHESIS WITH MULTI LABEL CONDITIONAL GENERATIVE ADVERSARIAL NETWORK

Jian Wu (East China Normal University); Liping Wang (ECNU); Hailin Pan (East China Normal University); Binyu Wang (East China Normal University)

5722 (BISP-P5.8): IMPROVING HEART RATE AND HEART RATE VARIABILITY ESTIMATION FROM VIDEO THROUGH A HR-RR-TUNED FILTER

Michael Chan (Georgia Institute of Technology); Li Zhu (Samsung Research America); Korosh Vatanparvar (Samsung Research America); Hewon Jung (Georgia Institute of Technology); Jilong Kuang (Samsung Research America); Alex Gao (Samsung Research America)

4999 (BISP-P5.9): HYDRA-HGR: A Hybrid Transformer-based Architecture for Fusion of Macroscopic and Microscopic Neural Drive Information

Mansoorh Montazerin (Concordia University); Elahe Rahimian (Concordia University); Farnoosh Naderkhani (Concordia University); S. Farokh Atashzar (NYU); Hamid Alinejad-Rokny (UNSW); Arash Mohammadi (Concordia University)

Friday, June 9

5007 (BISP-P5.10): Light-weighted CNN-Attention based architecture for Hand Gesture Recognition via ElectroMyography
Soheil Zabihi (Concordia University); Elahe Rahimian (Concordia University); Amir Asif (York University); Arash Mohammadi (Concordia University)

2378 (BISP-P5.11): A Novel Heart Rate Estimation Method Exploiting Heartbeat Second Harmonic Reconstruction via Millimeter Wave Radar
Tao Li (China University of Mining and Technology); Huayu Shou (China University of Mining and Technology); Yuchen Deng (China University of Mining and Technology); Yu Zhou (China University of Mining and Technology); Chenqi Shi (China University of Mining and Technology); Pengpeng Chen (China University of Mining and Technology)

4927 (BISP-P5.12): BreathIE: Estimating Breathing Inhale Exhale Ratio Using Motion Sensor Data from Consumer Earbuds
Nafiul Rashid (Samsung Research America); Md Mahbubur Rahman (Samsung Research America); Tousif Ahmed (Samsung Research America, Inc.); Jilong Kuang (Samsung Research America); Jun Alex Gao (Samsung Research America)

CI-P3: Computational Imaging IV

Room: Dome F

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Salman Asif

127 (CI-P3.1): Dual-Cycle: Self-Supervised Dual-View Fluorescence Microscopy Image Reconstruction using CycleGAN
Tomas Kerepecky (Czech Academy of Sciences); Jiaming Liu (Washington University in St. Louis); Xue Wen Ng (Washington University School of Medicine); David Piston (Washington University School of Medicine); Ulugbek S. Kamilov (Washington University in St. Louis)

731 (CI-P3.2): LIGHT FIELD COMPRESSION VIA COMPACT NEURAL SCENE REPRESENTATION
Jinglei Shi (Nankai University); Christine Guillemot (INRIA)

1068 (CI-P3.3): A DEEP DISENTANGLED APPROACH FOR INTERPRETABLE HYPERSPECTRAL UNMIXING
Ricardo Borsoi (CNRS); Tales C O Imbiriba (Northeastern University); Deniz Erdogmus (Northeastern University)

1287 (CI-P3.4): CTSR: A Hybrid CNN-Transformer Network for Scene Text Image Super-Resolution
Kaiwei Dai (Central South University); Nan Kang (Central South University); Li Kuang (Central South University)

2603 (CI-P3.5): BLOCK-BASED COLOR CONSTANCY: THE DEVIATION OF SALIENT PIXELS
Oguzhan Ulucan (Universität Greifswald); Diclehan Ulucan (Universität Greifswald); Marc Ebner (Universität Greifswald)

4131 (CI-P3.6): Zone Plate Virtual Lenses for Memory-Constrained NLOS Imaging
Pablo Luesia-Lahoz (Universidad de Zaragoza); Diego Gutierrez (University of Zaragoza); Adolfo Muñoz (U. Zaragoza)

5264 (CI-P3.7): ROBUST SPATIOTEMPORAL FUSION OF SATELLITE IMAGES VIA CONVEX OPTIMIZATION
Ryosuke Isono (Tokyo Institute of Technology); Kazuki Naganuma (Tokyo Institute of Technology); Shunsuke Ono (Tokyo Institute of Technology)

1097 (CI-P3.8): Semi-SwinDerain: Semi-supervised Image Deraining Network using Swin Transformer
Chun Ren (Beijing University of Posts and Telecommunications); Danfeng Yan (State Key Laboratory of Networking and Switching Technology Beijing University of Posts and Telecommunications); Yuanqiang Cai (Beijing University of Posts and Telecommunications); Li Yang-chun (Chinese Academy of Cyberspace Studies)

1541 (CI-P3.9): Attention Based Relation Network for Facial Action Units Recognition
Yao Wei (South China University of Technology); Haoxiang Wang (South China University of Technology); Mingze Sun (South China University of Technology); Liu Jiawang (SCUT)

1898 (CI-P3.10): DEEP LOW LIGHT IMAGE ENHANCEMENT VIA MULTI-SCALE RECURSIVE FEATURE ENHANCEMENT AND CURVE ADJUSTMENT
Haiyan Jin (Xi'an University of Technology); Dawei Wei (Xi'an University of Technology); Haonan Su (Xi'an University of Technology)

6775 (CI-P3.11): D2UF: Deep Coded Aperture Design and Unrolling Algorithm for Compressive Spectral Image Fusion (SPS Journal Paper)*
Roman A Jacome (Universidad Industrial de Santander); Jorge Bacca (Universidad Industrial de Santander); Henry Arguello (Universidad Industrial Santander)

Friday, June 9

IVMSP-P12: Image/Video Enhancement

Room: Dome M

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Yu-Chiang Frank, Yao Zhao

1892 (IVMSP-P12.1): Thermal Infrared Image Inpainting via Edge-Aware Guidance

Zeyu Wang (Zhejiang University); Haibin Shen (Zhejiang University); Changyou Men (Hangzhou Vango Technologies, Inc.); Quan Sun (Hangzhou Vango Technologies, Inc.); Kejie Huang (Zhejiang University)

889 (IVMSP-P12.2): SEMANTICS-AWARE GAMMA CORRECTION FOR UNSUPERVISED LOW-LIGHT IMAGE ENHANCEMENT

Yu-Hsuan Chen (National Taiwan University); Fu-Cheng Pan (National Taiwan University); Yu-Chien Liao (National Taiwan University); Jao-Hong Kao (novatek inc.); Yu-Chiang Frank Wang (National Taiwan University)

2155 (IVMSP-P12.3): A fusion-based and multi-layer method for low light image enhancement

Xueyan Zhou (Nankai University); Jiacen Guo (Nankai University); Hao Liu (Nankai University); Chao Wang (Nankai University)

3287 (IVMSP-P12.4): Lit the Darkness: Three-stage zero-shot learning for low-light enhancement with multi-neighbor enhancement factors

Mariam Saeed (Alexandria University); Marwan Torki (Alexandria University)

997 (IVMSP-P12.5): A Spatio-Temporal Decomposition Network for Compressed Video Quality Enhancement

Kai Wang (Hikvision Research Institute); Fangdong Chen (Hikvision Research Institute); Zongmiao Ye (Hikvision Research Institute); Li Wang (Hikvision Research Institute); xiaoyang wu (Hikvision Research Institute); Shiliang Pu (Hikvision Research Institute)

2268 (IVMSP-P12.6): DUAL-HEAD FUSION NETWORK FOR IMAGE ENHANCEMENT

Yuhong Zhang (Shanghai Jiao Tong University); Hengsheng Zhang (Shanghai Jiao Tong University); Li Song (Shanghai Jiao Tong University); Rong Xie (Shanghai Jiao Tong University); Wenjun Zhang (Shanghai Jiao Tong University)

6396 (IVMSP-P12.7): SVMV: SPATIOTEMPORAL VARIANCE-SUPERVISED MOTION VOLUME FOR VIDEO FRAME INTERPOLATION

Yao Luo (Nanjing University of Science and Technology); Jinshan Pan (Nanjing University of Science and Technology); Jinhui Tang (Nanjing University of Science and Technology)

3311 (IVMSP-P12.8): Data Augmentation based on Invariant Shape Blending for Deep Learning Classification

Emna Ghorbel (National School of Computer Science (ENSI)); Mahmoud Ghorbel (National School of Computer Science (ENSI)); Slim Mhiri (ENSI)

2713 (IVMSP-P12.9): UAV REMOTE SENSING IMAGE DEHAZING BASED ON MULTI-DIMENSIONAL SALIENCY AWARENESS UNEQUAL NETWORK

Ruohui Zheng (Beijing Normal University); Libao Zhang (Beijing Normal University)

453 (IVMSP-P12.10): Towards Reliable Image Outpainting: Learning Structure-Aware Multimodal Fusion with Depth Guidance

Lei Zhang (Beijing Jiaotong University); Chunyu Lin (Beijing Jiaotong University); Kang Liao (Beijing Jiaotong University); Yao Zhao (Beijing Jiaotong University)

347 (IVMSP-P12.11): Flow-Guided Deformable Alignment Network with Self-Supervision for Video Inpainting

Zhiliang Wu (Nanjing University of Science and Technology); Kang Zhang (Nanjing University of Science and Technology); Changchang Sun (Illinois Institute of Technology); Hanyu Xuan (Anhui University); Yan Yan (Illinois Institute of Technology)

916 (IVMSP-P12.12): A Template Matching Approach for Reference Picture Padding in Video Coding

Nicolas Horst (Institute of Imaging & Computer Vision, RWTH Aachen University); Priyanka Das (RWTH Aachen University, Germany); Mathias Wien (RWTH Aachen University, Germany)

MLSP-P24: Zero or Few Shot Learning

Room: Dome G

Type: Poster

03:40 PM to 05:10 PM

Chair(s): TBA

3438 (MLSP-P24.1): Zero-shot domain adaptation of anomalous samples for semi-supervised anomaly detection

Tomoya Nishida (Hitachi, Ltd.); Takashi Endo (Hitachi, Ltd.); Yohei Kawaguchi (Hitachi, Ltd.)

5514 (MLSP-P24.2): Dual Collaborative Visual-Semantic Mapping for Multi-Label Zero-Shot Image Recognition

Yunqing Hu (Zhejiang University); Xuan Jin (Alibaba Turing Lab, Alibaba Group); Xi Chen (Zhejiang University); Yin Zhang (Zhejiang University)

Friday, June 9

3663 (MLSP-P24.3): EXPLOITING SEMANTIC ATTRIBUTES FOR TRANSDUCTIVE ZERO-SHOT LEARNING

Zhengbo Wang (University of Science and Technology of China); Jian Liang (CASIA); Zilei Wang (University of Science and Technology of China); Tieniu Tan (NLPR, China)

2259 (MLSP-P24.4): Enhanced Embeddings in Zero-Shot Learning for Environmental Audio

Ysobel Sims (The University of Newcastle); Alexandre Mendes (The University of Newcastle); Stephan K Chalup (The University of Newcastle)

3638 (MLSP-P24.5): Zero-Shot Anomalous Sound Detection in Domestic Environments Using Large-Scale Pretrained Audio Pattern Recognition Models

Alessandro I Mezza (Politecnico di Milano); Giulio Zanetti (Politecnico di Milano); Maximo Cobos (Universitat de Valencia); Fabio Antonacci (Politecnico di Milano)

361 (MLSP-P24.6): HalluAudio: Hallucinate frequency as concepts for few-shot audio classification

Zhongjie Yu (Wyze Labs, Inc.); Shuyang Wang (Shiseido Americas); Lin Chen (Wyze Labs Inc.); Zhongwei Cheng (Wyze Labs, Inc.)

3811 (MLSP-P24.7): TeAw: Text-Aware Few-Shot Remote Sensing Image Scene Classification

Kaihui Cheng (National Innovation Institute of Defense Technology, Academy of Military Science); Chule Yang (Defense Innovation Institute (DII)); Zunlin Fan (National Innovation Institute of Defense Technology, China); Dayan Wu (Institute of Information Engineering, Chinese Academy of Sciences); Naiyang Guan (National Innovation Institute of Defense Technology; Tianjin Artificial Intelligence Innovation Center)

1125 (MLSP-P24.8): Uncertainty-Aware Few-Shot Class-Incremental Learning

Zhu Jiancai (East China Normal University); Jiabao Zhao (East China Normal University); Jiayi Zhou (East China Normal University); Liang He (ECNU); Jing Yang (ECNU); Zhi Zhang (Shanghai Educational Technology Center)

4631 (MLSP-P24.9): VPPT: Visual Pre-trained Prompt Tuning Framework for Few-Shot Image Classification

Zhao Song (National Innovation Institute of Defense Technology); Ke YANG (NIIDT); Naiyang Guan (National Innovation Institute of Defense Technology; Tianjin Artificial Intelligence Innovation Center); Junjie Zhu (NIIDT); Peng Qiao (NUDT); Qingyong Hu (University of Oxford)

1155 (MLSP-P24.10): An Adaptive Plug-and-Play Network for Few-Shot Learning

Hao Li (Beihang University); Li Li (Beihang university); Yunmeng Huang (Beihang University); Ning Li (Beihang University); Yongtao Zhang (Nanjing University of Aeronautics and Astronautics)

526 (MLSP-P24.11): ACTIVE LEARNING FOR EFFICIENT FEW-SHOT CLASSIFICATION

Aymane Abdali (IMT Atlantique); Vincent Gripon (IMT Atlantique); Lucas Drumetz (IMT Atlantique); Bartosz Boguslawski (Schneider Electric)

SAM-P1: Acoustic and Microphone Array Processing

Room: Dome E

Type: Poster

03:40 PM to 05:10 PM

Chair(s): TBA

278 (SAM-P1.1): Improved Deep Speaker Localization and Tracking: Revised Training Paradigm and Controlled Latency

Alexander Bohlender (IDLab, Ghent University - imec); Liesbeth Roelens (IDLab, Ghent University - imec); Nilesh Madhu (IDLab, Ghent University - imec)

1482 (SAM-P1.2): GCC-speaker: Target Speaker Localization with Optimal Speaker-dependent Weighting in Multi-speaker Scenarios

Guanjun Li (Institute of Automation, Chinese Academy of Sciences); Wei Xue (Department of Computer Science, Hong Kong Baptist University, Hong Kong SAR, China); Wenju Liu (National Laboratory of Pattern Recognition, Institute of Automation, University of Chinese Academy of Sciences, Beijing, China); Jiangyan Yi (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences); Jianhua Tao ("National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences")

1669 (SAM-P1.3): Using Received Power in Microphone Arrays to Estimate Direction of Arrival

Gustav Zetterqvist (Linköping University); Fredrik Gustafsson (Linköping University); Gustaf Hendeby (Linköping University)

2026 (SAM-P1.4): Improved Mask-Based Neural Beamforming for Multichannel Speech Enhancement by Snapshot Matching Masking

Ching-Hua Lee (Samsung Research America); Chouchang Yang (Samsung Research America); Yilin Shen (Samsung Research America); Hongxia Jin (Samsung Research America)

2571 (SAM-P1.5): Order Reduction of Multi-Channel FIR Filters by Balanced Truncation

Florian Hilgemann (Institute of Communication Systems, RWTH Aachen University); Peter Jax (RWTH Aachen University, Institute of Communication Systems (IKS))

Friday, June 9

2632 (SAM-P1.6): Passive Acoustic Tracking of Whales in 3-D

Junsu Jang (UC San Diego); Florian Meyer (UC San Diego); Eric Snyder (UC San Diego); Sean Wiggins (UC San Diego); Simone Baumann-Pickering (UC San Diego); John Hildebrand (Univ of California San Diego)

2662 (SAM-P1.7): Fast Cross-Correlation for TDoA Estimation on Small Aperture Microphone Arrays

François Grondin (Université de Sherbrooke); Marc-Antoine Maheux (Université de Sherbrooke); Jean-Samuel Lauzon (Université de Sherbrooke); Jonathan Vincent (Université de Sherbrooke); Francois Michaud (Universite de Sherbrooke)

3913 (SAM-P1.8): DIFFUSION-BASED SOUND SOURCE LOCALIZATION USING NETWORKS OF PLANAR MICROPHONE ARRAYS

Davide Albertini (Politecnico di Milano); Gioele Greco (Politecnico di Milano); Alberto Bernardini (Politecnico di Milano); Augusto Sarti (Politecnico di Milano)

4481 (SAM-P1.9): Towards improved sonar performance using environment-informed sparse sub-array processing

Alexandre L'Her (Thales DMS); Angélique Drémeau (ENSTA Bretagne); Florent Le Courtois (DGA TN); Gaultier Real (DGA TN); Xavier Cristol (Thales DMS); Yann Stéphan (SHOM)

5035 (SAM-P1.10): SIMULTANEOUS ESTIMATION OF DIRECTION OF ARRIVAL AND SOUND SPEED USING A NON-UNIFORM SENSOR ARRAY

Ryouichi Nishimura (National Institute of Information and Communications Technology); Kenichi Takizawa (National Institute of Information and Communications Technology)

5517 (SAM-P1.11): JOINT ESTIMATION OF DOA AND DISTANCE IN NOISY REVERBERANT CONDITIONS

Suliang Bu (University of Missouri); Tuo Zhao (University of Missouri); Yunxin Zhao (University of Missouri)

5677 (SAM-P1.12): DEEP NEURAL MEL-SUBBAND BEAMFORMER FOR IN-CAR SPEECH SEPARATION

Vinay Kothapally (Tencent AI); Yong Xu (Tencent); Meng Yu (Tencent); Shixiong Zhang (Tencent); Dong Yu (tencent)

SLT-P32: Speech and Language Disorders

Room: Dome J

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Milos Cernak,

1905 (SLT-P32.1): "Prediction of Sleepiness Ratings from Voice by Man and Machine": a perceptual experiment replication study

Vincent P. Martin (Université de Bordeaux); Aymeric Ferron (INRIA Bordeaux); Jean-Luc Rouas (CNRS); Pierre Philip (Université de Bordeaux)

2074 (SLT-P32.2): Self-supervised representations in speech-based depression detection

Wen Wu (University of Cambridge); Chao Zhang (University of Cambridge); Phil Woodland (Machine Intelligence Laboratory, Cambridge University Department of Engineering)

3001 (SLT-P32.3): Towards Reducing Patient Effort for the Automatic Prediction of Speech Intelligibility in Head and Neck Cancers

Sebastião Quintas (IRIT, Université de Toulouse, CNRS, Toulouse, France); Alberto Abad (INESC-ID); Julie Maclair (IRIT); Virginie Woisard (Hospitals of Toulouse); Julien Pinquier (IRIT)

3760 (SLT-P32.4): A Context-Aware Computational Approach for Measuring Vocal Entrainment in Dyadic Conversations

Rimita Lahiri (University of Southern California); Md Nasir (Microsoft); Catherine Lord (UCLA); So Hyun Kim (Korea University); Shrikanth Narayanan (USC)

4107 (SLT-P32.5): Feature Selection and Text Embedding For Detecting Dementia from Spontaneous Cantonese

Xiaoquan Ke (The Hong Kong Polytechnic University); Man-Wai MAK (The Hong Kong Polytechnic University); Mei Ling MENG (The Chinese University of Hong Kong)

4792 (SLT-P32.6): A knowledge-driven vowel-based approach of depression classification from speech using data augmentation

Kexin Feng (Texas A&M University); Theodora Chaspari (Texas A&M University)

4931 (SLT-P32.7): NEURAL ARCHITECTURE SEARCH WITH MULTIMODAL FUSION METHODS FOR DIAGNOSING DEMENTIA

Michail Chatzianastasis (École Polytechnique); Loukas Ilias (National Technical University of Athens); Dimitris Askounis (National Technical University of Athens); Michalis Vazirgiannis (École Polytechnique)

4984 (SLT-P32.8): EARLY DETECTION OF COGNITIVE DECLINE USING VOICE ASSISTANT COMMANDS

Eli Kurtz (UMass Boston); Youxiang Zhu (UMass Boston); Tiffany Driesse (University of North Carolina); Bang Tran (UMass Boston); John Batsis (University of North Carolina); Robert Roth (Geisel School of Medicine); Xiaohui Liang (University of Massachusetts Boston)

Friday, June 9

5344 (SLT-P32.9): EXPLORING THE ROLE OF FRICATIVES IN CLASSIFYING HEALTHY SUBJECTS AND PATIENTS WITH AMYOTROPHIC LATERAL SCLEROSIS AND PARKINSON'S DISEASE

Tanuka Bhattacharjee (Indian Institute of Science); Yamini BK (NIMHANS); Nalini Atchayaram (NIMHANS); Ravi Yadav (NIMHANS); Prasanta Dr Ghosh (Indian Institute of Science (IISc), Bangalore)

5363 (SLT-P32.10): STATIC AND DYNAMIC SOURCE AND FILTER CUES FOR CLASSIFICATION OF AMYOTROPHIC LATERAL SCLEROSIS PATIENTS AND HEALTHY SUBJECTS

Tanuka Bhattacharjee (Indian Institute of Science); Chowdam Venkata Thirumala Kumar (Indian Institute of Science, Bengaluru); Yamini BK (NIMHANS); Nalini Atchayaram (NIMHANS); Ravi Yadav (NIMHANS); Prasanta Dr Ghosh (Indian Institute of Science (IISc), Bangalore)

5998 (SLT-P32.11): Transferring Quantified Emotion Knowledge for the Detection of Depression in Alzheimer's Disease Using ForestNets

Paula Andrea Pérez-Toro (Friedrich-Alexander-Universität Erlangen-Nürnberg); Dalia Rodríguez-Salas (Friedrich-Alexander-Universität Erlangen-Nürnberg); Tomas Arias-Vergara (Friedrich-Alexander-Universität Erlangen-Nürnberg); Sebastian P Bayerl (Technische Hochschule Nürnberg Georg Simon Ohm); Philipp Klumpp (Pattern Recognition Lab, FAU Erlangen-Nuremberg); Korbinian Riedhammer (Technische Hochschule Nürnberg Georg Simon Ohm); Maria Schuster (Ludwig Maximilian University of Munich); Elmar Noeth (Friedrich Alexander Universität, Erlangen-Nuremberg); Andreas K Maier (Pattern Recognition Lab, FAU Erlangen-Nuremberg); Juan Rafael Orozco-Aroyave (University of Antioquia)

6063 (SLT-P32.12): Leveraging Pretrained Representations with Task-related Keywords for Alzheimer's Disease Detection

Jinchao Li (The Chinese University of Hong Kong); Kaitao Song (Microsoft Research Asia); Junan Li (The Chinese University of Hong Kong); Bo ZHENG (The Chinese University of Hong Kong); Dongsheng Li (Microsoft Research Asia); Xixin Wu (The Chinese University of Hong Kong); Xunying Liu (The Chinese University of Hong Kong); Helen Meng (The Chinese University of Hong Kong)

SLT-P43: Various Aspects in Speech and Speaker Recognition

Room: Dome C

Type: Poster

03:40 PM to 05:10 PM

Chair(s): Shri Narayanan,

5862 (SLT-P43.1): Auto-AVSR: Audio-Visual Speech Recognition with Automatic Labels

Pingchuan Ma (Meta); Alexandros Haliassos (Imperial College London); Adriana Fernandez-Lopez (Meta); Honglie Chen (Meta); Stavros Petridis (Imperial College London); Maja Pantic (Facebook / Imperial College London)

6156 (SLT-P43.2): DELAY-PENALIZED TRANSDUCER FOR LOW-LATENCY STREAMING ASR

Wei Kang (Xiaomi Corp., Beijing, China); Zengwei Yao (Xiaomi Corp.); Fangjun Kuang (Xiaomi Corp.); Liyong Guo (Xiaomi Corp.); Xiaoyu Yang (Xiaomi Corp.); Long Lin (Xiaomi Corp.); Piotr Żelasko (Johns Hopkins University); Daniel Povey (Johns Hopkins University)

6305 (SLT-P43.3): FINDADAPNET: FIND AND INSERT ADAPTERS BY LEARNED LAYER IMPORTANCE

Junwei Huang (Carnegie Mellon University); Karthik Ganesan (CARNEGIE MELLON UNIVERSITY); Soumi Maiti (CMU); Young Min Kim (Carnegie Mellon University); Xuankai Chang (Carnegie Mellon University); Paul Pu Liang (Carnegie Mellon University); Shinji Watanabe (Carnegie Mellon University)

6281 (SLT-P43.4): ROBUST ACOUSTIC AND SEMANTIC CONTEXTUAL BIASING IN NEURAL TRANSDUCERS FOR SPEECH RECOGNITION

Xuandi FU (Amazon Alexa); Kanthashree Mysore Sathyendra (Amazon); Ankur Gandhe (Amazon Alexa); Jing Liu (Amazon.com); Grant P. Strimel (Amazon Alexa); Ross McGowan (Amazon Alexa); Athanasios Mouchtaris (Amazon Alexa)

2212 (SLT-P43.5): InterMPL: Momentum Pseudo-Labeling with Intermediate CTC Loss

Yosuke Higuchi (Waseda University); Tetsuji Ogawa (Waseda University); Tetsunori Kobayashi (Waseda University); Shinji Watanabe (Carnegie Mellon University)

3562 (SLT-P43.6): On Unsupervised Uncertainty-Driven Speech Pseudo-Label Filtering and Model Calibration

Nauman Dawalatabad (Massachusetts Institute of Technology); Sameer Khurana (Massachusetts Institute of Technology); Antoine Laurent (Le Mans University); James Glass (Massachusetts Institute of Technology)

6721 (SLT-P43.7): Towards Better Domain Adaptation for Self-supervised Models: A Case Study of Child ASR (SPS Journal Paper)*

Ruchao Fan (University of California, Los Angeles); Yuzheng Zhu (University of California, Los Angeles); Jinhan Wang (UCLA); Abeer Alwan (UCLA)

4652 (SLT-P43.8): UNSUPERVISED SPEAKER VERIFICATION USING PRE-TRAINED MODEL AND LABEL CORRECTION

Zhicong Chen (Xiamen University); Jie Wang (Xiamen University); Wenxuan Hu (Xiamen University); Lin Li (Xiamen University); Qingyang Hong (Xiamen University)

6738 (SLT-P43.9): Selective Listening by Synchronizing Speech With Lips (SPS Journal Paper)*

Zexu Pan (National University of Singapore)

Friday, June 9

6769 (SLT-P43.10): Multi-Classifer Interactive Learning for Ambiguous Speech Emotion Recognition (SPS Journal Paper)*
Ying Zhou (Xidian University); Xuefeng Liang (Xidian University); Yu Gu (School of Artificial Intelligence, Xi'dian University); Yin Yifei (Xidian University); longshan yao (xidian university)

6004 (SLT-P43.11): Unsupervised Noise Adaptation using Data Simulation
Chen Chen (Nanyang Technological University); Yuchen Hu (Nanyang Technological University); Heqing Zou (Nanyang Technological University); Linhui Sun (Nanjing University of Posts and Telecommunications); Eng Siong Chng (Nanyang Technological University)

SPTM-P12: Estimation, detection, classification, Optimization methods for signal processing, Sampling theory, compressed and non-uniform sampling, Signal filtering, restoration, enhancement, and reconstruction, Signal processing over graphs, Signal Processing over networks II
Room: Dome L
Type: Poster
03:40 PM to 05:10 PM
Chair(s): Athina Petropulu

5070 (SPTM-P12.1): Signal processing with optical quadratic random sketches
Remi AAV Delogne (Université catholique de Louvain); Vincent Schellekens (CEA); Laurent Daudet (LightOn); Laurent Jacques (Université catholique de Louvain)

5398 (SPTM-P12.2): Low Precision Representations for High Dimensional Models
Rajarshi Saha (Stanford University); Mert Pilanci (Stanford University); Andrea Goldsmith (Princeton University)

6707 (SPTM-P12.3): The Fixed Point Iteration of Positive Concave Mappings Converges Geometrically if a Fixed Point Exists: Implications to Wireless Systems (SPS Journal Paper)*
Tomasz Piotrowski (NCU); Renato Luis Garrido Cavalcante (Fraunhofer Heinrich Hertz Inst)

6696 (SPTM-P12.4): Convex quantization preserves logconcavity (SPS Journal Paper)*
Pol del Aguila Pla (EPFL - CIBM); Aleix Boquet-Pujadas (EPFL); Joakim Jalden (KTH)

6781 (SPTM-P12.5): Sparse Stable Outlier-Robust Signal Recovery Under Gaussian Noise (SPS Journal Paper)*
Kyohei Suzuki (Keio University); Masahiro Yukawa (Keio University)

5037 (SPTM-P12.6): A Statistical Interpretation of the Maximum Subarray Problem
Dennis Wei (IBM Research); Dmitry M Malioutov (Scarsdale)

5175 (SPTM-P12.7): Robustness and Convergence of Mirror Descent for Blind Deconvolution
Ronak Mehta (University of Wisconsin-Madison); Sathya Ravi (University of Illinois at Chicago); Vikas Singh (University of Wisconsin Madison)

6804 (SPTM-P12.8): On the Benefits of Progressively Increasing Sampling Sizes in Stochastic Greedy Weak Submodular Maximization (SPS Journal Paper)*
Abolfazl Hashemi (Purdue University)

6057 (SPTM-P12.9): Low-rank plus sparse trajectory decomposition for direct exoplanet imaging
Simon Vary (ICTEAM/INMA, UCLouvain); Hazan Daglayan (UCLouvain); Laurent Jacques (Université catholique de Louvain); P.-A. Absil (UCLouvain)

6329 (SPTM-P12.10): Efficient Learning of Balanced Signature Graphs
Gerald Matz (Technische Universität Wien); Claudio Verardo (University of Udine); Thomas Dittrich (Technische Universität Wien)

6351 (SPTM-P12.11): Robust M-Estimation based Distributed Expectation Maximization Algorithm with Robust Aggregation
Christian A. Schroth (Technische Universität Darmstadt); Stefan Vlaski (Imperial College London); Abdelhak M Zoubir (Technische Universität Darmstadt)

SPTM-P4: Sampling Theory, Compressed and Non-uniform Sampling
Room: Dome K
Type: Poster
03:40 PM to 05:10 PM
Chair(s): Yonina Eldar, Ayush Bhandari

1182 (SPTM-P4.1): High-Dynamic Range ADC for Finite-Rate-of-Innovation Signals
Satish Mulleti (Indian Institute of Technology Bombay, India); Yonina Eldar ()

Friday, June 9

1914 (SPTM-P4.2): Online Residual-Based Key Frame Sampling with Self-Coach Mechanism and Adaptive Multi-Level Feature Fusion

Rui Zhang (Shanghai Jiao Tong University); Yang Hua (Queen's University Belfast); Tao Song (Shanghai Jiao Tong University); Zhengui Xue (Shanghai Jiao Tong University); Ruhui Ma (Shanghai Jiao Tong University); Haibing Guan (Shanghai Jiao Tong University)

4000 (SPTM-P4.3): Sparse asynchronous samples from networks of TEMS for reconstruction of classes of non-bandlimited signals

Marek Hilton (Imperial College London); Pier Luigi Dragotti (Imperial College London)

4029 (SPTM-P4.4): Revisit Sampling Theory of Bandlimited Graph Signals: One Bridge Between GSP and DSP

Fen Wang (Zhejiang Lab); Taihao Li (zhejianglab); Xue Zhang (Shandong University of Science and Technology)

4101 (SPTM-P4.5): Multichannel Time-Encoding of Finite-Rate-of-Innovation Signals

Abijith Jagannath Kamath (Indian Institute of Science); Chandra Sekhar Seelamantula (IISc Bangalore)

4774 (SPTM-P4.6): LASSO-BASED FAST RESIDUAL RECOVERY FOR MODULO SAMPLING

Shaik Basheeruddin Shah (Weizmann Institute of Science); Satish Mulleti (Indian Institute of Technology Bombay, India); Yonina Eldar ()

5198 (SPTM-P4.7): Unlimited Sampling in Phase Space

Peiyu Zhang (Imperial College London); Ayush Bhandari (Imperial College London)

5765 (SPTM-P4.8): BEYOND RATE CODING: SIGNAL CODING AND RECONSTRUCTION USING LEAN SPIKE TRAINS

Anik Chattopadhyay (University of Florida); Arunava Banerjee (University of Florida)

5938 (SPTM-P4.9): SAMPLING ORDER-LIMITED SIGNALS ON THE SPHERE

Salaar Khan (LUMS); Salman Nadeem (Lahore University of Management Sciences); Zubair Khalid (Lahore University of Management Sciences)

6401 (SPTM-P4.10): UNLIMITED SAMPLING OF FRI SIGNALS INDEPENDENT OF SAMPLING RATE

Ruiming Guo (Imperial College London); Ayush Bhandari (Imperial College London)

6788 (SPTM-P4.11): Compressive Sensing With Wigner D-Functions on Subsets of the Sphere (SPS Journal Paper)*

Marc A Valdez (Colorado School of Mines); Michael Wakin (CO school of Mines)