ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing

Virtual Conference
7-13 May 2022

Singapore
22-27 May 2022

Pages 1-535

IEEE Catalog Number: CFP22ICA-POD
ISBN: 978-1-6654-0541-6

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP22ICA-POD
ISBN (Print-On-Demand):	978-1-6654-0541-6
ISBN (Online):	978-1-6654-0540-9
ISSN:	1520-6149

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings .com

TABLE OF CONTENTS

ASPS-1: SIGNAL PROCESSING FOR HEALTH AND WELLNESS MONITORING

ASPS-1.1: COUGHTRIGGER: EARBUDS IMU BASED COUGH DETECTION ACTIVATOR USING AN ENERGY-EFFICIENT SENSITIVITY-PRIORITIZED TIME SERIES CLASSIFIER 1
Shibo Zhang, Nabil Alshurafa, Northwestern University, United States of America; Ebrahim Nemati, Tousif Ahmed, Mahbubur Rahman, Jilong Kuang, Alex Gao, Samsung Research America, United States of America; Minh Dinh, Nathan Folkman, Samsung Design Innovation Center, United States of America

ASPS-1.2: NON-INVASIVE BLOOD PRESSURE MONITORING WITH MULTI-MODAL IN-EAR SENSING 6
Hoang Truong, Alessandro Montanari, Fahim Kawsar, Nokia Bell Labs, United Kingdom of Great Britain and Northern Ireland

ASPS-1.3: INTELLIGENT WI-FI BASED CHILD PRESENCE DETECTION SYSTEM 11
Xiaolu Zeng, K. J. Ray Liu, University of Maryland College Park, United States of America; Beibei Wang, Sai Deepika Regani, Origin Wireless Inc, United States of America; Chenshu Wu, The University of Hong Kong, Hong Kong

ASPS-1.4: REAL-TIME FALL DETECTION USING MMWAVE RADAR 16
Wenxuan Li, Dongheng Zhang, Yadong Li, Zhi Wu, Jinbo Chen, Dong Zhang, Yang Hu, Qibin Sun, Yan Chen, University of Science and Technology of China, China

ASPS-1.5: HIERARCHICAL DEEP LEARNING MODEL WITH INERTIAL AND PHYSIOLOGICAL SENSORS FUSION FOR WEARABLE-BASED HUMAN ACTIVITY RECOGNITION 21
Dae Yon Hwang, Pai Chet Ng, Dimitrios Hatzinakos, Konstantinos N. Plataniotis, University of Toronto, Canada; Yuanhao Yu, Yang Wang, Huawei Technologies, Canada; Petros Spachos, University of Guelph, Canada

ASPS-2: AUDIO AND ACOUSTIC SIGNAL PROCESSING

ASPS-2.1: SPEECH RECOVERY FOR REAL-WORLD SELF-POWERED INTERMITTENT DEVICES 26
Yu-Chen Lin, Tsun-An Hsieh, Kuo-Hsuan Hung, Cheng Yu, Yu Tsao, Academia Sinica, Taiwan; Harinath Garudadri, University of California, San Diego, United States of America; Tei-Wei Kuo, City University of Hong Kong, Hong Kong

ASPS-2.2: PHASE CONTROL OF PARAMETRIC ARRAY LOUDSPEAKER BY OPTIMIZING SIDEBAND WEIGHTS 31
Ai Okano, Yoshinobu Kajikawa, Kansai University, Japan

ASPS-2.3: LOW-LATENCY HUMAN-COMPUTER AUDITORY INTERFACE BASED ON REAL-TIME VISION ANALYSIS 36
Florian Scalvini, Cyrille Migniot, Julien Dubois, ImViA, France; Camille Bordeau, Maxime Ambard, LEAD, France

ASPS-2.4: ROBUST ADAPTIVE NOISE CANCELLER ALGORITHM WITH SNR-BASED STEPSIZE CONTROL AND NOISE-PATH GAIN COMPENSATION 41
Akihiko Sugiyama, Yahoo Japan Corporation, Japan

ASPS-2.5: NEARTRACKER: ACOUSTIC 2-D TARGET TRACKING WITH NEARBY REFLECTOR IN SISO SYSTEM 46
Chao Liu, Linlin Gao, Ruobing Jiang, Ocean University of China, China

ASPS-3: SIGNAL PROCESSING, MACHINE LEARNING, AND ARCHITECTURE DESIGN OPTIMIZATION

ASPS-3.1: AN EFFICIENT METHOD FOR GENERIC DSP IMPLEMENTATION OF DILATED CONVOLUTION 51
Harinarayanan EV, Sachin Ghanekar, Cadence Design Systems, India

ASPS-3.2: COMPRESSION-AWARE PROJECTION WITH GREEDY DIMENSION REDUCTION FOR CONVOLUTIONAL NEURAL NETWORK ACTIVATIONS 56
Yu-Shan Tai, Chieh-Fang Teng, Cheng-Yang Chang, An-Yeu Wu, National Taiwan University, Taiwan

ASPS-3.3: OPTIMIZING THE CONSUMPTION OF SPIKING NEURAL NETWORKS WITH ACTIVITY REGULARIZATION 61
Simon Narduzzi, Siavash A. Bigdeli, L. Andrea Dunbar, Centre suisse d'électronique et de microtechnique, Switzerland; Shih-Chii Liu, University of Zürich, Switzerland

ASPS-3.4: IMPQ: REDUCED COMPLEXITY NEURAL NETWORKS VIA GRANULAR PRECISION ASSIGNMENT 66
Sujan Kumar Gonugondla, Amazon, United States of America; Naresh Shanbhag, University of Illinois at Urbana-Champaign, United States of America

ASPS-3.5: RATE CODING OR DIRECT CODING: WHICH ONE IS BETTER FOR ACCURATE, ROBUST, AND ENERGY-EFFICIENT SPIKING NEURAL NETWORKS? 71
Youngeun Kim, Hyoungseob Park, Abhishek Moitra, Abhiroop Bhattacharjee, Yeshwanth Venkatesha, Priyadarshini Panda, Yale University, United States of America

ASPS-3.6: PYXIS: AN OPEN-SOURCE PERFORMANCE DATASET OF SPARSE ACCELERATORS 76
Linghao Song, Yuze Chi, Jason Cong, University of California, Los Angeles, United States of America

ASPS-4: INTEGRATIVE SIGNAL PROCESSING AND MACHINE LEARNING FOR MULTIMODAL SENSING AND PROCESSING

ASPS-4.1: FAST FAULT DIAGNOSIS METHOD OF ROLLING BEARINGS IN MULTI-SENSOR MEASUREMENT ENVIROMENT 81
Zuozhou Pan, Zong Meng, Yanshan University, China; Zhiping Lin, Yuanjin Zheng, Nanyang Technological University, Singapore

ASPS-4.2: DETECTING ANOMALY IN CHEMICAL SENSORS VIA REGULARIZED CONTRASTIVE LEARNING 86
Diaa Badawi, Ahmet Enis Cetin, University of Illinois at Chicago, United States of America; Ishaan Bassi, Sule Ozev, Arizona State University, United States of America

ASPS-4.3: EVOLUTIONARY NEURAL ARCHITECTURE DESIGN OF LIQUID STATE MACHINE FOR IMAGE CLASSIFICATION 91
Cheng Tang, University of Toyama, Japan; Junkai Ji, Qiuzhen Lin, Shenzhen University, China; Yan Zhou, Northeastern University, China

ASPS-4.4: INVISIBLE AND EFFICIENT BACKDOOR ATTACKS FOR COMPRESSED DEEP NEURAL NETWORKS 96
Huy Phan, Yi Xie, Yingying Chen, Bo Yuan, Rutgers University, United States of America; Jian Liu, University of Tennessee, United States of America

ASPS-4.5: TENSOR-BASED ORTHOGONAL MATCHING PURSUIT WITH PHASE ROTATION FOR CHANNEL ESTIMATION IN HYBRID BEAMFORMING MIMO-OFDM SYSTEMS 101
Cheng-Hung Lo, Pei-Yun Tsai, National Central University, Taiwan, Province of China

AUD-1: MUSIC SOURCE SEPARATION I: INTERACTIVE, CONDITIONED, AND INFORMED MODELS

AUD-1.1: SPAIN-NET: SPATIALLY-INFORMED STEREOPHONIC MUSIC SOURCE SEPARATION 106
Darius Petermann, Minje Kim, Indiana University, United States of America

AUD-1.3: IMPROVED SINGING VOICE SEPARATION WITH CHROMAGRAM-BASED PITCH-AWARE REMIXING 111
Siyuan Yuan, Stanford University, United States of America; Zhepei Wang, UIUC, United States of America; Umut Isik, Ritwik Giri, Jean-Marc Valin, Michael M. Goodwin, Arvindh Krishnaswamy, Amazon Web Services, United States of America

AUD-1.4: DON'T SEPARATE, LEARN TO REMIX: END-TO-END NEURAL REMIXING WITH JOINT OPTIMIZATION 116
Haici Yang, Shivani Firodiya, Minje Kim, Indiana University, United States of America; Nicholas Bryan, Adobe Research, United States of America

AUD-1.5: FEW-SHOT MUSICAL SOURCE SEPARATION 121
Yu Wang, Juan Pablo Bello, New York University, United States of America; Daniel Stoller, Rachel M. Bittner, Spotify, United States of America

AUD-1.6: SOURCE SEPARATION BY STEERING PRETRAINED MUSIC MODELS 126
Ethan Manilow, Patrick O'Reilly, Bryan Pardo, Northwestern University, United States of America; Prem Seetharaman, Descript, Inc., United States of America

AUD-2: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS I: DATASETS

AUD-2.1: INFANT CRYING DETECTION IN REAL-WORLD ENVIRONMENTS 131
Xuewen Yao, Megan Micheletti, Mckensey Johnson, Edison Thomaz, Kaya de Barbaro, University of Texas at Austin, United States of America

AUD-2.2: WIKITAG: WIKIPEDIA-BASED KNOWLEDGE EMBEDDINGS TOWARDS IMPROVED ACOUSTIC EVENT CLASSIFICATION 136
Qin Zhang, Qingming Tang, Chieh-Chi Kao, Ming Sun, Yang Liu, Chao Wang, Amazon Inc, United States of America

AUD-2.3: URBAN SOUND & SIGHT: DATASET AND BENCHMARK FOR AUDIO-VISUAL URBAN SCENE UNDERSTANDING 141
Magdalena Fuentes, Bea Steers, Julia Wilkins, Qianyi Shi, Yao Hou, Juan Pablo Bello, New York University, United States of America; Pablo Zinemanas, Xavier Serra, Universitat Pompeu Fabra, Spain; Martin Rocamora, Universidad de la Republica, Uruguay; Luca Bondi, Samarjit Das, Bosch Research, United States of America

AUD-2.4: REAL-WORLD ON-BOARD UAV AUDIO DATA SET FOR PROPELLER ANOMALIES 146
Sai Srinadhu Katta, Kide Vuojärvi, Sivaprasad Nandyala, Ulla-Maria Kovalainen, Lauren Baddeley, Secure Systems Research Centre (SSRC) - Technology Innovation Institute (TII), United Arab Emirates

AUD-2.5: VOCALSOUND: A DATASET FOR IMPROVING HUMAN VOCAL SOUNDS RECOGNITION 151
Yuan Gong, James Glass, Massachusetts Institute of Technology, United States of America; Jin Yu, Signify Research, United States of America

AUD-2.6: WEARABLE SELD DATASET: DATASET FOR SOUND EVENT LOCALIZATION AND DETECTION USING WEARABLE DEVICES AROUND HEAD 156
Kento Nagatomo, Kohei Yatabe, Yasuhiro Oikawa, Waseda University, Japan; Masahiro Yasuda, Shoichiro Saito, NTT Corporation, Japan

AUD-3: SYSTEM IDENTIFICATION AND REVERBERATION REDUCTION

AUD-3.1: TUNET: A BLOCK-ONLINE BANDWIDTH EXTENSION MODEL BASED ON TRANSFORMERS AND SELF-SUPERVISED PRETRAINING 161
Viet-Anh Nguyen, Anh H. T. Nguyen, FPT Software, Viet Nam; Andy W. H. Khong, Nanyang Technological University, Singapore

AUD-3.2: DRC-NET: DENSELY CONNECTED RECURRENT CONVOLUTIONAL NEURAL NETWORK FOR SPEECH DEREVERBERATION 166
Jinjiang Liu, Xueliang Zhang, College of Computer Science, Inner Mongolia University, China

AUD-3.3: CUSTOMIZABLE END-TO-END OPTIMIZATION OF ONLINE NEURAL NETWORK-SUPPORTED DEREVERBERATION FOR HEARING DEVICES 171
Jean-Marie Lemercier, Timo Gerkmann, Universität Hamburg, Germany; Joachim Thiemann, Raphael Koning, Advanced Bionics GmbH, Germany

AUD-3.5: IMPORTANCE OF SWITCH OPTIMIZATION CRITERION IN SWITCHING WPE DEREVERBERATION 176
Naoyuki Kamo, Rintaro Ikeshita, Keisuke Kinoshita, Tomohiro Nakatani, NTT, Japan

AUD-4: SYMBOLIC ANALYSIS AND SYNTHESIS OF MUSIC

AUD-4.1: AUDIO-TO-SYMBOLIC ARRANGEMENT VIA CROSS-MODAL MUSIC REPRESENTATION LEARNING 181
Ziyu Wang, Gus Xia, NYU Shanghai, China; Dejing Xu, Tencent ARC Lab, China; Ying Shan, Tencnet ARC Lab, China

AUD-4.2: MUSIC PHRASE INPAINTING USING LONG-TERM REPRESENTATION AND CONTRASTIVE LOSS 186
Shiqi Wei, Weiguo Gao, Fudan University, China; Gus Xia, Liwei Lin, New York University Shanghai, China; Yixiao Zhang, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland

AUD-4.3: MELONS: GENERATING MELODY WITH LONG-TERM STRUCTURE USING TRANSFORMERS AND STRUCTURE GRAPH 191
Yi Zou, Communication University of China, China; Pei Zou, Peking University, China; Yi Zhao, Ran Zhang, Xiaorui Wang, Kuaishou Technology Co., Beijing, China; Kaixiang Zhang, Kuaishou Technology Co, China

AUD-4.4: DIFFICULTY-AWARE NEURAL BAND-TO-PIANO SCORE ARRANGEMENT BASED ON NOTE- AND STATISTIC-LEVEL CRITERIA 196
Moyu Terao, Yuki Hiramatsu, Ryoto Ishizuka, Yiming Wu, Kazuyoshi Yoshii, Kyoto University, Japan

AUD-4.5: SCORE DIFFICULTY ANALYSIS FOR PIANO PERFORMANCE EDUCATION BASED ON FINGERING 201
Pedro Ramoneda, Music Technology Group (MTG), Spain; Nazif Can Tamer, Vsevolod Eremenko, Xavier Serra, Marius Mirón, MTG, Spain

AUD-5: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS II: TASKS

AUD-5.1: A NEURAL NETWORK-BASED HOWLING DETECTION METHOD FOR REAL-TIME COMMUNICATION APPLICATIONS 206
Zhipeng Chen, Yiya Hao, Yaobin Chen, Gong Chen, Liang Ruan, Netease GrowthEase, China

AUD-5.2: ALARM SOUND DETECTION USING TOPOLOGICAL SIGNAL PROCESSING 211
Tomer Fireaizen, Saar Ron, Omer Bobrowski, Technion - Israel Institute of Technology, Israel

AUD-5.3: A METHOD FOR ESTIMATING THE GROUPING OF PARTICIPANTS IN CLASSROOM GROUP WORK USING ONLY AUDIO INFORMATION 216
Osamu Ichikawa, Yuuto Shima, Shiga University, Japan; Takahiro Nakayama, The University of Tokyo, Japan; Hajime Shirouzu, National Institute for Educational Policy Research, Japan

AUD-5.4: ENVIRONMENTAL SOUND EXTRACTION USING ONOMATOPOEIC WORDS 221
Yuki Okamoto, Ritsumeikan University, Japan; Shota Horiguchi, Masaaki Yamamoto, Yohei Kawaguchi, Hitachi Ltd., Japan; Keisuke Imoto, Doshisha University, Japan

AUD-5.6: ECHO-AWARE ADAPTATION OF SOUND EVENT LOCALIZATION AND DETECTION IN UNKNOWN ENVIRONMENTS 226
Masahiro Yasuda, Yasunori Ohishi, Shoichiro Saito, NTT Corporation, Japan

AUD-6: AUDIO SECURITY AND MULTI-MODAL SYSTEMS

AUD-6.1: ON ADVERSARIAL ROBUSTNESS OF LARGE-SCALE AUDIO VISUAL LEARNING 231
Juncheng B Li, Xinjian Li, Po-Yao (Bernie) Huang, Florian Metze, Carnegie Mellon University, United States of America; Shuhui Qu, Stanford University, United States of America

AUD-6.2: ADVERSARIAL SAMPLE DETECTION FOR SPEAKER VERIFICATION BY NEURAL VOCODERS 236
Haibin Wu, Po-chun Hsu, Hung-yi Lee, National Taiwan University, China; Ji Gao, Shanshan Zhang, Shen Huang, Jian Kang, Tencent, China; Zhiyong Wu, Shenzhen International Graduate School, Tsinghua University, China; Helen Meng, The Chinese University of Hong Kong, Hong Kong

AUD-6.3: SOURCE MIXING AND SEPARATION ROBUST AUDIO STEGANOGRAPHY 241
Naoya Takahashi, Mayank Kumar Singh, Yuki Mitsufuji, Sony Group Corporation, Japan

AUD-6.4: MULTI-MODAL PRE-TRAINING FOR AUTOMATED SPEECH RECOGNITION 246
David Chan, University of California, Berkeley, United States of America; Shalini Ghosh, Debmalya Chakrabarty, Björn Hoffmeister, Amazon, United States of America

AUD-6.5: SPEAKER-TARGETED AUDIO-VISUAL SPEECH RECOGNITION USING A HYBRID CTC/ATTENTION MODEL WITH INTERFERENCE LOSS 251
Ryota Tsunoda, Ryoichi Takashima, Tetsuya Takiguchi, Kobe University, Japan; Ryo Aihara, Yoshie Imai, Mitsubishi Electric Corporation, Japan

AUD-6.6: TIME-DOMAIN AUDIO-VISUAL SPEECH SEPARATION ON LOW QUALITY VIDEOS 256
Yifei Wu, Chenda Li, Yanmin Qian, Shanghai Jiao Tong University, China; Jinfeng Bai, Zhongqin Wu, TAL Education Group, China, China

AUD-7: SIGNAL ENHANCEMENT AND RESTORATION I: MULTICHANNEL

AUD-7.1: COMPLEX-VALUED SPATIAL AUTOENCODERS FOR MULTICHANNEL SPEECH ENHANCEMENT 261
Mhd Modar Halimeh, Walter Kellermann, Friedrich–Alexander University Erlangen–Nürnberg, Germany

AUD-7.2: MULTICHANNEL NOISE REDUCTION USING DILATED MULTICHANNEL U-NET AND PRE-TRAINED SINGLE-CHANNEL NETWORK 266
Zhi-Wei Tan, Yuan Liu, Andy W. H. Khong, Nanyang Technological University, Singapore; Anh H. T. Nguyen, FPT Software, Viet Nam

AUD-7.3: ONE MODEL TO ENHANCE THEM ALL: ARRAY GEOMETRY AGNOSTIC MULTI-CHANNEL PERSONALIZED SPEECH ENHANCEMENT 271
Hassan Taherian, The Ohio State University, United States of America; Sefik Emre Eskimez, Takuya Yoshioka, Huaming Wang, Zhuo Chen, Xuedong Huang, Microsoft, United States of America

AUD-7.4: MULTI-CHANNEL SPEECH DENOISING FOR MACHINE EARS 276
Cong Han, Columbia University, United States of America; Merve Kaya, Kyle Hoefer, Simon Carlile, X Development, LLC, United States of America; Malcolm Slaney, Google, LLC, United States of America

AUD-7.5: LOCALIZATION BASED SEQUENTIAL GROUPING FOR CONTINUOUS SPEECH SEPARATION 281
Zhong-Qiu Wang, DeLiang Wang, The Ohio State University, United States of America

AUD-7.6: CONVOLUTIONAL WEIGHTED MINIMUM MEAN SQUARE ERROR FILTER FOR JOINT SOURCE SEPARATION AND DEREVERBERATION 286
Mieszko Fraś, Marcin Witkowski, Konrad Kowalczyk, AGH University of Science and Technology, Poland

AUD-8: MUSIC SOURCE SEPARATION II: METHODS AND ARCHITECTURES

AUD-8.1: IMPROVING SOURCE SEPARATION BY EXPLICITLY MODELING DEPENDENCIES BETWEEN SOURCES 291
Ethan Manilow, Bryan Pardo, Northwestern University, United States of America; Curtis Hawthorne, Cheng-Zhi Anna Huang, Jesse Engel, Google, United States of America

AUD-8.2: MUSIC SOURCE SEPARATION WITH DEEP EQUILIBRIUM MODELS 296
Yuichiro Koyama, Naoki Murata, Shusuke Takahashi, Yuki Mitsufuji, Sony Group Corporation, Japan; Stefan Uhlich, Giorgio Fabbro, Sony Europe B.V., Germany

AUD-8.5: HARMONIC AND PERCUSSIVE SOUND SEPARATION BASED ON MIXED PARTIAL DERIVATIVE OF PHASE SPECTROGRAM 301
Natsuki Akaishi, Kohei Yatabe, Yasuhiro Oikawa, Waseda University, Japan

AUD-8.6: ON LOSS FUNCTIONS AND EVALUATION METRICS FOR MUSIC SOURCE SEPARATION 306
Enric Gusó, Universitat Pompeu Fabra, Spain; Jordi Pons, Santiago Pascual, Joan Serrà, Dolby Laboratories, Spain

AUD-9: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS III: LOSSES AND TRAINING

AUD-9.1: TIME-BALANCED FOCAL LOSS FOR AUDIO EVENT DETECTION 311
Sangwook Park, Mounya Elhilali, Johns Hopkins University, United States of America

AUD-9.2: MULTI-ACCDOA: LOCALIZING AND DETECTING OVERLAPPING SOUNDS FROM THE SAME CLASS WITH AUXILIARY DUPLICATING PERMUTATION INVARIANT TRAINING 316
Kazuki Shimada, Yuichiro Koyama, Shusuke Takahashi, Naoya Takahashi, Emiru Tsunoo, Yuki Mitsufuji, Sony Group Corporation, Japan

AUD-9.3: IMPROVED REPRESENTATION LEARNING FOR ACOUSTIC EVENT CLASSIFICATION USING TREE-STRUCTURED ONTOLOGY 321
Arman Zharmagambetov, University of California, Merced, United States of America; Qingming Tang, Chieh-Chi Kao, Qin Zhang, Ming Sun, Viktor Rozgic, Jasha Droppo, Chao Wang, Amazon, Alexa, United States of America

AUD-9.4: TEMPORAL CONTRASTIVE-LOSS FOR AUDIO EVENT DETECTION 326
Sandeep Kothinti, Mounya Elhilali, Johns Hopkins University, United States of America

AUD-9.5: A FRAME LOSS OF MULTIPLE INSTANCE LEARNING FOR WEAKLY SUPERVISED SOUND EVENT DETECTION 331
Xu Wang, Xiangjinzi Zhang, Shengwu Xiong, School of Computer and Artificial Intelligence, Wuhan University of Technology, Wuhan, China, China; Yunfei Zi, ,

AUD-9.6: PSEUDO STRONG LABELS FOR LARGE SCALE WEAKLY SUPERVISED AUDIO TAGGING 336
Heinrich Dinkel, Zhiyong Yan, Yongqing Wang, Junbo Zhang, Yujun Wang, Xiaomi Corporation, China

AUD-10: HEARING INSTRUMENTS AND ON-DEVICE PROCESSING

AUD-10.1: INDIVIDUALIZED HEAR-THROUGH FOR ACOUSTIC TRANSPARENCY USING PCA-BASED SOUND PRESSURE ESTIMATION AT THE EARDRUM 341
Wenyu Jin, Tim Schoof, Henning Schepker, Starkey Hearing Technologies, United States of America

AUD-10.2: ON SPECTRAL AND TEMPORAL SPARSIFICATION OF SPEECH SIGNALS FOR THE IMPROVEMENT OF SPEECH PERCEPTION IN CI LISTENERS 346
Benjamin Lentz, Rainer Martin, Ruhr-University Bochum, Germany; Kirsten Oberländer, Christiane Völter, St. Elisabeth-Hospital, Ruhr-Universität Bochum, Germany

AUD-10.3: A DIFFERENTIABLE OPTIMISATION FRAMEWORK FOR THE DESIGN OF INDIVIDUALISED DNN-BASED HEARING-AID STRATEGIES 351
Fotios Drakopoulos, Sarah Verhulst, Ghent University, Belgium

AUD-10.4: PERSONALIZED SPEECH ENHANCEMENT: NEW MODELS AND COMPREHENSIVE EVALUATION 356
Sefik Emre Eskimez, Takuya Yoshioka, Huaming Wang, Xiaofei Wang, Zhuo Chen, Xuedong Huang, Microsoft, United States of America

AUD-10.5: DYNAMIC SLIDING WINDOW FOR REALTIME DENOISING NETWORKS 361
Jinxu Xiang, Yuyang Zhu, Rundi Wu, Ruilin Xu, Changxi Zheng, Columbia University, United States of America; Yuko Ishiwaka, SoftBank Group, Japan

AUD-11: DEEP LEARNING-BASED SINGLE-CHANNEL SPEECH ENHANCEMENT

AUD-11.1: BLOOM-NET: BLOCKWISE OPTIMIZATION FOR MASKING NETWORKS TOWARD SCALABLE AND EFFICIENT SPEECH ENHANCEMENT 366
Sunwoo Kim, Minje Kim, Indiana University, United States of America

AUD-11.2: HGCN: HARMONIC GATED COMPENSATION NETWORK FOR SPEECH ENHANCEMENT 371
Tianrui Wang, Weibin Zhu, Beijing Jiaotong University, China; Yingying Gao, Junlan Feng, Shilei Zhang, China Mobile Research Institute, China

AUD-11.3: SPEECH ENHANCEMENT WITH NEURAL HOMOMORPHIC SYNTHESIS 376
Wenbin Jiang, Zhijun Liu, Kai Yu, Fei Wen, Shanghai Jiao Tong University, China

AUD-11.4: A BAYESIAN PERMUTATION TRAINING DEEP REPRESENTATION LEARNING METHOD FOR SPEECH ENHANCEMENT WITH VARIATIONAL AUTOENCODER 381
Yang Xiang, Aalborg University & Capturi A/S, Denmark; Jesper Lisby Højvang, Morten Højfeldt Rasmussen, Capturi A/S, Denmark; Mads Græsbøll Christensen, Aalborg University, Denmark

AUD-11.5: INTEGRATING STATISTICAL UNCERTAINTY INTO NEURAL NETWORK-BASED SPEECH ENHANCEMENT 386
Huajian Fang, Tal Peer, Stefan Wermter, Timo Gerkmann, Universität Hamburg, Germany

AUD-11.6: UNSUPERVISED SPEECH ENHANCEMENT WITH SPEECH RECOGNITION EMBEDDING AND DISENTANGLEMENT LOSSES 391
Viet Anh Trinh, The City University of New York, United States of America; Sebastian Braun, Microsoft, United States of America

AUD-12: MUSICAL EVENT DETECTION AND STRUCTURE ANALYSIS

AUD-12.1: MUSICYOLO: A SIGHT-SINGING ONSET/OFFSET DETECTION FRAMEWORK BASED ON OBJECT DETECTION INSTEAD OF SPECTRUM FRAMES 396
Xianke Wang, Wei Xu, Weiming Yang, Wenqing Cheng, Huazhong University of Science and Technology, China

AUD-12.2: MODELING BEATS AND DOWNBEATS WITH A TIME-FREQUENCY TRANSFORMER 401
Yun-Ning Hung, Georgia Institute of Technology, United States of America; Ju-Chiang Wang, Xuchen Song, Wei-Tsung Lu, Minz Won, ByteDance, United States of America

AUD-12.3: HIERARCHICAL CLASSIFICATION OF SINGING ACTIVITY, GENDER, AND TYPE IN COMPLEX MUSIC RECORDINGS 406
Michael Krause, Meinard Müller, International Audio Laboratories Erlangen, Germany

AUD-12.4: DEEPCHORUS: A HYBRID MODEL OF MULTI-SCALE CONVOLUTION AND SELF-ATTENTION FOR CHORUS DETECTION 411
Qiqi He, Xiaoheng Sun, Wei Li, Fudan University, School of Computer Science and Technology, China, China; Yi Yu, National Institute of Informatics (NII), Tokyo, Japan, Japan

AUD-12.5: TO CATCH A CHORUS, VERSE, INTRO, OR ANYTHING ELSE: ANALYZING A SONG WITH STRUCTURAL FUNCTIONS 416
Ju-Chiang Wang, Jordan B. L. Smith, ByteDance, United States of America; Yun-Ning Hung, Georgia Institute of Technology, United States of America

AUD-12.6: A NOVEL 1D STATE SPACE FOR EFFICIENT MUSIC RHYTHMIC ANALYSIS 421
Mojtaba Heydari, Zhiyao Duan, University of Rochester, United States of America; Matthew McCallum, Andreas Ehmann, Pandora Media, Inc., United States of America

AUD-13: SPATIAL AUDIO I

AUD-13.1: UPMIXING VIA STYLE TRANSFER: A VARIATIONAL AUTOENCODER FOR DISENTANGLING SPATIAL IMAGES AND MUSICAL CONTENT 426
Haici Yang, Minje Kim, Indiana University, United States of America; Sanna Wager, Spencer Russell, Mike Luo, Wontak Kim, Amazon Lab 126, United States of America

AUD-13.2: SPATIAL MIXUP: DIRECTIONAL LOUDNESS MODIFICATION AS DATA AUGMENTATION FOR SOUND EVENT LOCALIZATION AND DETECTION 431
Ricardo Falcon Perez, Aalto University, Finland; Kazuki Shimada, Yuichiro Koyama, Shusuke Takahashi, Yuki Mitsufuji, Sony Group Corporation, Japan

AUD-13.3: TOWARDS FASTER CONTINUOUS MULTI-CHANNEL HRTF MEASUREMENTS BASED ON LEARNING SYSTEM MODELS 436
Tobias Kabzinski, Peter Jax, RWTH Aachen University, Germany

AUD-13.4: TOWARDS FAST AND CONVENIENT END-TO-END HRTF PERSONALIZATION 441
Bowen Zhi, Dmitry Zotkin, Ramani Duraiswami, VisiSonics Corporation, United States of America

AUD-13.6: WISHART LOCALIZATION PRIOR ON SPATIAL COVARIANCE MATRIX IN AMBISONIC SOURCE SEPARATION USING NON-NEGATIVE TENSOR FACTORIZATION 446
Mateusz Guzik, Konrad Kowalczyk, AGH University of Science and Technology, Poland

AUD-14: MUSIC APPLICATIONS

AUD-14.1: IMPROVING LYRICS ALIGNMENT THROUGH JOINT PITCH DETECTION 451
Jiawen Huang, Emmanouil Benetos, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland; Sebastian Ewert, Spotify, Germany

AUD-14.2: LEARNING MUSIC AUDIO REPRESENTATIONS VIA WEAK LANGUAGE SUPERVISION 456
Ilaria Manco, Emmanouil Benetos, Gyorgy Fazekas, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland; Elio Quinton, Universal Music Group, United Kingdom of Great Britain and Northern Ireland

AUD-14.3: ON THE PREDICTION OF THE FREQUENCY RESPONSE OF A WOODEN PLATE FROM ITS MECHANICAL PARAMETERS 461
David Giuseppe Badiane, Raffaele Malvermi, Sebastian Gonzalez, Fabio Antonacci, Augusto Sarti, Politecnico di Milano, Italy

AUD-14.5: AUTOMATIC DJ TRANSITIONS WITH DIFFERENTIABLE AUDIO EFFECTS AND GENERATIVE ADVERSARIAL NETWORKS 466
Bo-Yu Chen, Wei-Han Hsu, Yi-Hsuan Yang, Academia Sinica, Taiwan; Wei-Hsiang Liao, Marco Antonio Martinez Ramirez, Yuki Mitsufuji, Sony Group Corporation, Japan

AUD-15: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS IV: ON-DEVICE CONSIDERATIONS

AUD-15.1: SELF-SUPERVISED REPRESENTATION LEARNING FOR UNSUPERVISED ANOMALOUS SOUND DETECTION UNDER DOMAIN SHIFT 471
Han Chen, Yan Song, Li-Rong Dai, Ian McLoughlin, University of Science and Technology of China, China; Lin Liu, iFLYTEK Research, iFLYTEK CO., LTD, China

AUD-15.3: FEDERATED SELF-TRAINING FOR DATA-EFFICIENT AUDIO RECOGNITION 476
Vasileios Tsouvalas, Tanir Ozcelebi, Eindhoven University of Technology, Netherlands; Aaqib Saeed, Philips Research, Netherlands

AUD-15.4: FEDERATED SELF-SUPERVISED LEARNING FOR ACOUSTIC EVENT CLASSIFICATION 481
Meng Feng, Massachusetts Institute of Technology, United States of America; Chieh-Chi Kao, Qingming Tang, Ming Sun, Viktor Rozgic, Spyros Matsoukas, Chao Wang, Amazon, United States of America

AUD-15.5: TEMPORAL KNOWLEDGE DISTILLATION FOR ON-DEVICE AUDIO CLASSIFICATION 486
Kwanghee Choi, Martin Kersner, Jacob Morton, Buru Chang, Hyperconnect, Korea, Republic of

AUD-15.6: STREAMING ON-DEVICE DETECTION OF DEVICE DIRECTED SPEECH FROM VOICE AND TOUCH-BASED INVOCATION 491
Ognjen (Oggi) Rudovic, Akanksha Bindal, Vineet Garg, Pramod Simha, Pranay Dighe, Sachin Kajarekar, Apple Inc., United States of America

AUD-16: MULTICHANNEL BLIND SOURCE SEPARATION: LATENT VARIABLE ANALYSIS

AUD-16.1: MULTI-FRAME FULL-RANK SPATIAL COVARIANCE ANALYSIS FOR UNDERDETERMINED BSS IN REVERBERANT ENVIRONMENTS 496
Hiroshi Sawada, Rintaro Ikeshita, Keisuke Kinoshita, Tomohiro Nakatani, NTT Corporation, Japan

AUD-16.2: FLOW-BASED FAST MULTICHANNEL NONNEGATIVE MATRIX FACTORIZATION FOR BLIND SOURCE SEPARATION 501
Aditya Arie Nugraha, Kouhei Sekiguchi, RIKEN, Japan; Mathieu Fontaine, Télécom Paris, France; Yoshiaki Bando, National Institute of Advanced Industrial Science and Technology (AIST), Japan; Kazuyoshi Yoshii, Kyoto University, Japan

AUD-16.3: HARVESTING PARTIALLY-DISJOINT TIME-FREQUENCY INFORMATION FOR IMPROVING DEGENERATE UNMIXING ESTIMATION TECHNIQUE 506
Yudong He, Qifeng Chen, Richard H.Y. So, The Hong Kong University of Science and Technology, Hong Kong; He Wang, HKUST-Shenzhen Research Institute, China

AUD-16.4: INVESTIGATION AND COMPARISON OF OPTIMIZATION METHODS FOR VARIATIONAL AUTOENCODER-BASED UNDERDETERMINED MULTICHANNEL SOURCE SEPARATION 511
Shogo Seki, Hirokazu Kameoka, NTT Corporation, Japan; Li Li, Nagoya University, Japan

AUD-16.6: HBP: AN EFFICIENT BLOCK PERMUTATION SOLVER USING HUNGARIAN ALGORITHM AND SPECTROGRAM INPAINTING FOR MULTICHANNEL AUDIO SOURCE SEPARATION 516
Li Li, Nagoya University, Japan; Hirokazu Kameoka, Shogo Seki, NTT Communication Science Laboratories, NTT Corporation, Japan

AUD-17: SPEECH SEPARATION I

AUD-17.1: EAD-CONFORMER: A CONFORMER-BASED ENCODER-ATTENTION-DECODER-NETWORK FOR MULTI-TASK AUDIO SOURCE SEPARATION 521
Chenxing Li, Yang Wang, Feng Deng, Zhuo Zhang, Xiaorui Wang, Zhongyuan Wang, Kuai Shou, China

AUD-17.2: THE COCKTAIL FORK PROBLEM: THREE-STEM AUDIO SEPARATION FOR REAL-WORLD SOUNDTRACKS 526
Darius Petermann, Indiana University, United States of America; Gordon Wichern, Jonathan Le Roux, Mitsubishi Electric Research Laboratories (MERL), United States of America; Zhong-Qiu Wang, Carnegie Mellon University, United States of America

AUD-17.3: PHASE SHIFTED BEDROSIAN FILTERBANK: AN INTERPRETABLE AUDIO FRONT-END FOR TIME-DOMAIN AUDIO SOURCE SEPARATION 531
Félix Mathieu, Gael Richard, Geoffroy Peeters, Telecom Paris, France; Thomas Courtat, Thales, France

AUD-17.4: HARMONICITY PLAYS A CRITICAL ROLE IN DNN BASED VERSUS IN BIOLOGICALLY-INSPIRED MONAURAL SPEECH SEGREGATION SYSTEMS 536
Rahil Parikh, Carol Espy-Wilson, Shihab Shamma, University of Maryland College Park, United States of America; Ilya Kavalerov, Google Inc., United States of America

AUD-17.5: MULTI-CHANNEL NARROW-BAND DEEP SPEECH SEPARATION WITH FULL-BAND PERMUTATION INVARIANT TRAINING 541
Changsheng Quan, Zhejiang University, China; Xiaofei Li, Westlake University, China

AUD-18: MEDICAL ACOUSTICS

AUD-18.1: CSENET: COMPLEX SQUEEZE-AND-EXCITATION NETWORK FOR SPEECH DEPRESSION LEVEL PREDICTION 546
Cunhang Fan, Zhao Lv, Shengbing Pei, Anhui Province Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University, Hefei 230601, China, China; Mingyue Niu, School of Computer and Information Engineering, Tianjin Normal University, Tianjin 300387, China, China

AUD-18.2: UBILUNG: MULTI-MODAL PASSIVE-BASED LUNG HEALTH ASSESSMENT 551
Ebrahim Nemati, Viswam Nathan, Korosh Vatanparvar, Tousif Ahmed, Md. Mahbubur Rahman, Dan McCaffrey, Jilong Kuang, Alex Gao, Samsung Research America, United States of America; Xuhai Xu, University of Washington, United States of America

AUD-18.3: THE SECOND DICOVA CHALLENGE: DATASET AND PERFORMANCE ANALYSIS FOR DIAGNOSIS OF COVID-19 USING ACOUSTICS 556
Neeraj Kumar Sharma, Srikanth Raj Chetupalli, Debarpan Bhattacharya, Debottam Dutta, Pravin Mote, Sriram Ganapathy, Indian Institute of Science, India

AUD-18.4: SUPERVISED AND SELF-SUPERVISED PRETRAINING BASED COVID-19 DETECTION USING ACOUSTIC BREATHING/COUGH/SPEECH SIGNALS 561
Xing-Yu Chen, Qiu-Shi Zhu, Jie Zhang, Li-Rong Dai, University of Science and Technology of China, China

AUD-18.5: EXPLORING AUDITORY ACOUSTIC FEATURES FOR THE DIAGNOSIS OF COVID-19 566
Madhu Kamble, Jose Patino, Maria A. Zuluaga, Massimiliano Todisco, EURECOM, France, France

AUD-19: MODELING, ANALYSIS, AND SYNTHESIS OF ACOUSTIC ENVIRONMENTS

AUD-19.1: FAST-RIR: FAST NEURAL DIFFUSE ROOM IMPULSE RESPONSE GENERATOR 571
Anton Jeran Ratnarajah, Zhenyu Tang, Dinesh Manocha, University of Maryland, College Park, United States of America; Shi-Xiong Zhang, Meng Yu, Dong Yu, Tencent AI Lab, United States of America

AUD-19.2: REGION-TO-REGION KERNEL INTERPOLATION OF ACOUSTIC TRANSFER FUNCTION WITH DIRECTIONAL WEIGHTING 576
Juliano Garcia do Carmo Ribeiro, Shoichi Koyama, Hiroshi Saruwatari, The University of Tokyo, Japan

AUD-19.3: BLIND REVERBERATION TIME ESTIMATION IN DYNAMIC ACOUSTIC CONDITIONS 581
Philipp Götz, Emanuël A. P. Habets, International Audio Laboratories Erlangen, Germany; Cagdas Tuna, Andreas Walther, Fraunhofer Institute for Integrated Circuits IIS, Germany

AUD-19.4: SPARSE MODELING OF THE EARLY PART OF NOISY ROOM IMPULSE RESPONSES WITH SPARSE BAYESIAN LEARNING 586
Maozhong Fu, Yuhan Li, Xiamen University, China; Jesper Rindom Jensen, Mads Græsbøll Christensen, Aalborg University, Denmark

AUD-19.5: IMPROVED SIMULATION OF REALISTICALLY-SPATIALISED SIMULTANEOUS SPEECH USING MULTI-CAMERA ANALYSIS IN THE CHIME-5 DATASET 591
Jack Deadman, Jon Barker, University of Sheffield, United Kingdom of Great Britain and Northern Ireland

AUD-19.6: A DATA-DRIVEN APPROACH FOR ACOUSTIC PARAMETER SIMILARITY ESTIMATION OF SPEECH RECORDING 596
Mattia Papa, Clara Borrelli, Paolo Bestagini, Fabio Antonacci, Augusto Sarti, Stefano Tubaro, Politecnico di Milano, Italy

AUD-20: MUSIC CLASSIFICATION, IDENTIFICATION, AND MELODY EXTRACTION

AUD-20.1: VIOLINIST IDENTIFICATION USING NOTE-LEVEL TIMBRE FEATURE DISTRIBUTIONS 601
Yudong Zhao, György Fazekas, Mark Sandler, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland

AUD-20.2: S3T: SELF-SUPERVISED PRE-TRAINING WITH SWIN TRANSFORMER FOR MUSIC CLASSIFICATION 606
Hang Zhao, Bilei Zhu, Zejun Ma, ByteDance, China; Chen Zhang, Kejun Zhang, Zhejiang University, China

AUD-20.3: AMBIGUITY MODELLING WITH LABEL DISTRIBUTION LEARNING FOR MUSIC CLASSIFICATION 611
Morgan Buisson, Pablo Alonso Jimenez, Dmitry Bogdanov, Universitat Pompeu Fabra, Spain

AUD-20.4: BYTECOVER2: TOWARDS DIMENSIONALITY REDUCTION OF LATENT EMBEDDING FOR EFFICIENT COVER SONG IDENTIFICATION 616
Xingjian Du, Zijie Wang, Bilei Zhu, Zejun Ma, Bytedance AI Lab, China; Ke Chen, University of California San Diego, China

AUD-20.5: TONET: TONE-OCTAVE NETWORK FOR SINGING MELODY EXTRACTION FROM POLYPHONIC MUSIC 621
Ke Chen, Taylor Berg-Kirkpatrick, Shlomo Dubnov, University of California San Diego, United States of America; Shuai Yu, Wei Li, Fudan University, China; Cheng-i Wang, Smule Inc., United States of America

AUD-20.6: HIERARCHICAL GRAPH-BASED NEURAL NETWORK FOR SINGING MELODY EXTRACTION 626
Shuai Yu, Xi Chen, Wei Li, Fudan University, China

AUD-21: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS V: CLASSIFICATION

AUD-21.1: ON THE IMPACT OF NORMALIZATION STRATEGIES IN UNSUPERVISED ADVERSARIAL DOMAIN ADAPTATION FOR ACOUSTIC SCENE CLASSIFICATION 631
Michel Olvera, Emmanuel Vincent, Université de Lorraine, CNRS, Inria, Loria, France; Gilles Gasso, LITIS, Université & INSA Rouen Normandie, France

AUD-21.2: IMPROVING BIRD CLASSIFICATION WITH UNSUPERVISED SOUND SEPARATION 636
Tom Denton, Google, United States of America; Scott Wisdom, John R. Hershey, Google Research, United States of America

AUD-21.3: SCALABLE NEURAL ARCHITECTURES FOR END-TO-END ENVIRONMENTAL SOUND CLASSIFICATION 641
Francesco Paissan, Alberto Ancilotto, Alessio Brutti, Elisabetta Farella, Fondazione Bruno Kessler, Italy

AUD-21.4: HTS-AT: A HIERARCHICAL TOKEN-SEMANTIC AUDIO TRANSFORMER FOR SOUND CLASSIFICATION AND DETECTION 646
Ke Chen, Taylor Berg-Kirkpatrick, Shlomo Dubnov, University of California San Diego, United States of America; Xingjian Du, Bilei Zhu, Zejun Ma, Bytedance Inc., China

AUD-21.5: HYBRID ATTENTION-BASED PROTOTYPICAL NETWORKS FOR FEW-SHOT SOUND CLASSIFICATION 651
You Wang, David Anderson, Georgia Institute of Technology, United States of America

AUD-22: ACTIVE NOISE CONTROL, ECHO REDUCTION, AND FEEDBACK REDUCTION I: NOISE ROBUSTNESS

AUD-22.1: END-TO-END COMPLEX-VALUED MULTIDILATED CONVOLUTIONAL NEURAL NETWORK FOR JOINT ACOUSTIC ECHO CANCELLATION AND NOISE SUPPRESSION 656
Karn N. Watcharasupat, Thi Ngoc Tho Nguyen, Woon-Seng Gan, Nanyang Technological University, Singapore; Shengkui Zhao, Bin Ma, Alibaba Group, Singapore

AUD-22.2: NN3A: NEURAL NETWORK SUPPORTED ACOUSTIC ECHO CANCELLATION, NOISE SUPPRESSION AND AUTOMATIC GAIN CONTROL FOR REAL-TIME COMMUNICATIONS 661
Ziteng Wang, Yueyue Na, Biao Tian, Qiang Fu, Alibaba Group, China

AUD-22.3: DEEP RESIDUAL ECHO SUPPRESSION AND NOISE REDUCTION: A MULTI-INPUT FCRN APPROACH IN A HYBRID SPEECH ENHANCEMENT SYSTEM 666
Jan Franzen, Tim Fingscheidt, Technische Universität Braunschweig, Germany

AUD-22.4: NEURAL CASCADE ARCHITECTURE FOR JOINT ACOUSTIC ECHO AND NOISE SUPPRESSION 671
Hao Zhang, DeLiang Wang, The Ohio State University, United States of America

AUD-22.5: CASCADE MULTI-CHANNEL NOISE REDUCTION AND ACOUSTIC FEEDBACK CANCELLATION 676
Santiago Ruiz, Toon van Waterschoot, Marc Moonen, KU Leuven, Belgium

AUD-23: SPEECH SEPARATION II

AUD-23.1: SKIM: SKIPPING MEMORY LSTM FOR LOW-LATENCY REAL-TIME CONTINUOUS SPEECH SEPARATION 681
Chenda Li, Yanmin Qian, Shanghai Jiao Tong University, China; Lei Yang, Weiqin Wang, Samsung Research China - Beijing, China

AUD-23.2: ADAPTING SPEECH SEPARATION TO REAL-WORLD MEETINGS USING MIXTURE INVARIANT TRAINING 686
Aswin Sivaraman, Indiana University, United States of America; Scott Wisdom, Hakan Erdogan, John R. Hershey, Google Research, United States of America

AUD-23.3: QUANTIFYING DISCRIMINABILITY BETWEEN NMF BASES 691
Eisuke Konno, Daisuke Saito, Nobuaki Minematsu, The University of Tokyo, Japan

AUD-23.4: LOCATION-BASED TRAINING FOR MULTI-CHANNEL TALKER-INDEPENDENT SPEAKER SEPARATION 696
Hassan Taherian, Ke Tan, DeLiang Wang, The Ohio State University, United States of America

AUD-23.5: SDR — MEDIUM RARE WITH FAST COMPUTATIONS 701
Robin Scheibler, LINE Corporation, Japan

AUD-23.6: ATTENTIONPIT: SOFT PERMUTATION INVARIANT TRAINING FOR AUDIO SOURCE SEPARATION WITH ATTENTION MECHANISM 706
HIrokazu Kameoka, Shogo Seki, Li Li, Chihiro Watanabe, Nippon Telegraph and Telephone Corporation, Japan

AUD-24: ACOUSTIC SENSOR ARRAY PROCESSING I: LOCALIZATION

AUD-24.1: LOCATE THIS, NOT THAT: CLASS-CONDITIONED SOUND EVENT DOA ESTIMATION 711
Olga Slizovskaia, Gordon Wichern, Zhong-Qiu Wang, Jonathan Le Roux, Mitsubishi Electric Research Laboratories (MERL), United States of America

AUD-24.2: SALSA-LITE: A FAST AND EFFECTIVE FEATURE FOR POLYPHONIC SOUND EVENT LOCALIZATION AND DETECTION WITH MICROPHONE ARRAYS 716
Thi Ngoc Tho Nguyen, Karn N. Watcharasupat, Woon-Seng Gan, Nanyang Technological University, Singapore; Douglas L. Jones, University of Illinois Urbana-Champaign, United States of America; Huy Phan, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland

AUD-24.3: SRP-DNN: LEARNING DIRECT-PATH PHASE DIFFERENCE FOR MULTIPLE MOVING SOUND SOURCE LOCALIZATION 721
Bing Yang, Hong Liu, Peking University, China; Xiaofei Li, Westlake University & Westlake Institute for Advanced Study, China

AUD-24.4: CLOSED-FORM SINGLE SOURCE DIRECTION-OF-ARRIVAL ESTIMATOR USING FIRST-ORDER RELATIVE HARMONIC COEFFICIENTS 726
Yonggang Hu, Australian National University, Australia; Sharon Gannot, Bar-Ilan University, Israel

AUD-24.5: A SLIDE-SAVE BASED FRAMEWORK FOR MULTI-SOURCE DOA EXTRACTION WITH SPATIAL CLOSELY SEPARATED SOURCES 731
Jianhua Geng, Sifan Wang, Xin Lou, ShanghaiTech University, China

AUD-24.6: AN END-TO-END DEEP LEARNING FRAMEWORK FOR MULTIPLE AUDIO SOURCE SEPARATION AND LOCALIZATION 736
Yu Chen, Bowen Liu, Zijian Zhang, Hun-Seok Kim, University of Michigan, United States of America

AUD-25: ACTIVE NOISE CONTROL, ECHO REDUCTION, AND FEEDBACK REDUCTION II

AUD-25.1: DEEP ADAPTATION CONTROL FOR ACOUSTIC ECHO CANCELLATION 741
Amir Ivry, Israel Cohen, Baruch Berdugo, Technion - Israel Institute of Technology, Israel

AUD-25.2: OFF-THE-SHELF DEEP INTEGRATION FOR RESIDUAL-ECHO SUPPRESSION 746
Amir Ivry, Israel Cohen, Baruch Berdugo, Technion - Israel Institute of Technology, Israel

AUD-25.3: A COMPLEX SPECTRAL MAPPING WITH INPLACE CONVOLUTION RECURRENT NEURAL NETWORKS FOR ACOUSTIC ECHO CANCELLATION 751
Chenggang Zhang, Jinjiang Liu, Xueliang Zhang, Inner Mongolia University, China

AUD-25.4: DEEP ADAPTIVE AEC: HYBRID OF DEEP LEARNING AND ADAPTIVE ACOUSTIC ECHO CANCELLATION 756
Hao Zhang, The Ohio State University, United States of America; Srivatsan Kandadai, Harsha Rao, Tarun Pruthi, Trausti Kristjansson, Amazon Inc, United States of America; Minje Kim, Indiana University, United States of America

AUD-25.5: COMPUTATIONALLY EFFICIENT FIXED-FILTER ANC FOR SPEECH BASED ON LONG-TERM PREDICTION FOR HEADPHONE APPLICATIONS 761
Yurii Iotov, Mads Græsbøll Christensen, Aalborg University, Denmark; Sidsel Marie Nørholm, Valiantsin Belyi, Mads Dyrholm, GN Audio A/S, Denmark

AUD-25.6: END-TO-END DEEP LEARNING-BASED ADAPTATION CONTROL FOR FREQUENCY-DOMAIN ADAPTIVE SYSTEM IDENTIFICATION 766
Thomas Haubner, Andreas Brendel, Walter Kellermann, Friedrich-Alexander-University Erlangen-Nuremberg, Germany

AUD-26: MUSIC AND LYRICS TRANSCRIPTION

AUD-26.1: A FEW-SAMPLE STRATEGY FOR GUITAR TABLATURE TRANSCRIPTION BASED ON INHARMONICITY ANALYSIS AND PLAYABILITY CONSTRAINTS 771
Grigoris Bastas, Maximos Kaliakatsos-Papakostas, Vassilis Katsouros, Athena Research Center, Greece; Stefanos Koutoupis, Petros Maragos, National Technical University of Athens, Greece

AUD-26.2: EXPLORING TRANSFORMER'S POTENTIAL ON AUTOMATIC PIANO TRANSCRIPTION 776
Longshen Ou, Ziyi Guo, Ye Wang, National University of Singapore, Singapore; Emmanouil Benetos, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland; Jiqing Han, Harbin Institute of Technology, China

AUD-26.3: A LIGHTWEIGHT INSTRUMENT-AGNOSTIC MODEL FOR POLYPHONIC NOTE TRANSCRIPTION AND MULTIPITCH ESTIMATION 781
Rachel Bittner, Juan José Bosch, David Rubinstein, Sebastian Ewert, Spotify, France; Gabriel Meseguer Brocal, Institut de recherche et coordination acoustique/musique (IRCAM), France

AUD-26.4: TOWARDS AUTOMATIC TRANSCRIPTION OF POLYPHONIC ELECTRIC GUITAR MUSIC: A NEW DATASET AND A MULTI-LOSS TRANSFORMER MODEL 786
Yu-Hua Chen, Jyh-Shing Roger Jang, National Taiwan University, Taiwan; Wen-Yi Hsiao, Tsu-Kuang Hsieh, Yi-Hsuan Yang, Taiwan AI Labs, Taiwan

AUD-26.5: GENRE-CONDITIONED ACOUSTIC MODELS FOR AUTOMATIC LYRICS TRANSCRIPTION OF POLYPHONIC MUSIC 791
Xiaoxue Gao, Chitralekha Gupta, Haizhou Li, National University of Singapore, Singapore

AUD-26.6: PSEUDO-LABEL TRANSFER FROM FRAME-LEVEL TO NOTE-LEVEL IN A TEACHER-STUDENT FRAMEWORK FOR SINGING TRANSCRIPTION FROM POLYPHONIC MUSIC 796
Sangeun Kum, Jongpil Lee, Keunhyoung Luke Kim, Taehyoung Kim, Neutune Research, Korea, Republic of; Juhan Nam, KAIST, Korea, Republic of

AUD-27: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS VI: EVENTS

AUD-27.1: SOUND EVENT DETECTION GUIDED BY SEMANTIC CONTEXTS OF SCENES 801
Noriyuki Tonami, Ryotaro Nagase, Yuki Okamoto, Takahiro Fukumori, Yoichi Yamashita, Ritsumeikan University, Japan; Keisuke Imoto, Doshisha University, Japan

AUD-27.2: CNN-TRANSFORMER WITH SELF-ATTENTION NETWORK FOR SOUND EVENT DETECTION 806
Keigo Wakayama, Shoichiro Saito, NTT Corporation, Japan

AUD-27.3: A MUTUAL LEARNING FRAMEWORK FOR FEW-SHOT SOUND EVENT DETECTION 811
Dongchao Yang, Helin Wang, Yuexian Zou, Zhongjie Ye, Peking University, China; Wenwu Wang, University of Surrey, United Kingdom of Great Britain and Northern Ireland

AUD-27.4: ANOMALOUS SOUND DETECTION USING SPECTRAL-TEMPORAL INFORMATION FUSION 816
Youde Liu, Jian Guan, Harbin Engineering University, China; Qiaoxi Zhu, University of Technology Sydney, Australia; Wenwu Wang, University of Surrey, United Kingdom of Great Britain and Northern Ireland

AUD-27.5: SPARSE SELF-ATTENTION FOR SEMI-SUPERVISED SOUND EVENT DETECTION 821
Yadong Guan, Jiabin Xue, Guibin Zheng, Jiqing Han, Harbin Institute of Technology, China

AUD-27.6: PEER COLLABORATIVE LEARNING FOR POLYPHONIC SOUND EVENT DETECTION 826
Hayato Endo, Hiromitsu Nishizaki, University of Yamanashi, Japan

AUD-28: SIGNAL ENHANCEMENT AND RESTORATION II

AUD-28.1: POSTGAN: A GAN-BASED POST-PROCESSOR TO ENHANCE THE QUALITY OF CODED SPEECH 831
Srikanth Korse, Nicola Pia, Kishan Gupta, Guillaume Fuchs, Fraunhofer IIS, Erlangen, Germany, Germany

AUD-28.2: A DNN BASED POST-FILTER TO ENHANCE THE QUALITY OF CODED SPEECH IN MDCT DOMAIN 836
Kishan Gupta, Srikanth Korse, Bernd Edler, Guillaume Fuchs, Fraunhofer IIS, Germany

AUD-28.3: A TWO-STAGE U-NET FOR HIGH-FIDELITY DENOISING OF HISTORICAL RECORDINGS 841
Eloi Moliner, Vesa Välimäki, Aalto University, Finland

AUD-28.4: EXPERTS VERSUS ALL-ROUNDERS: TARGET LANGUAGE EXTRACTION FOR MULTIPLE TARGET LANGUAGES 846
Marvin Borsdorf, Kevin Scheck, Tanja Schultz, University of Bremen, Germany; Haizhou Li, National University of Singapore, Singapore

AUD-28.5: CATEGORY-ADAPTED SOUND EVENT ENHANCEMENT WITH WEAKLY LABELED DATA 851
Guangwei Li, Xuenan Xu, Mengyue Wu, Kai Yu, Shanghai Jiao Tong University, China; Heinrich Dinkel, Xiaomi Corporation, China

AUD-28.6: SEQUENTIAL MCMC METHODS FOR AUDIO SIGNAL ENHANCEMENT 856
Ruben Claveria, Simon Godsill, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

AUD-29: AUDIO AND SPEECH CODING

AUD-29.3: ARCHITECTURE FOR VARIABLE BITRATE NEURAL SPEECH CODEC 861
WITH CONFIGURABLE COMPUTATION COMPLEXITY
Tejas Jayashankar, Massachusetts Institute of Technology, United States of America; Thilo Koehler, Kaustubh Kalgaonkar, Zhiping Xiu, Jilong Wu, Ju Lin, Prabhav Agrawal, Qing He, Facebook AI, United States of America

AUD-29.4: END-TO-END NEURAL SPEECH CODING FOR REAL-TIME 866
COMMUNICATIONS
Xue Jiang, Chengyu Zheng, Yuan Zhang, Communication University of China, China; Xiulian Peng, Huaying Xue, Yan Lu, Microsoft Research Asia, China

AUD-29.5: DEEP NEURAL NETWORK (DNN) AUDIO CODER USING A 871
PERCEPTUALLY IMPROVED TRAINING METHOD
Seungmin Shin, Joon Byun, Youngcheol Park, Intelligent Signal Processing Lab, Yonsei University, Wonju, Korea, Republic of; Jongmo Sung, Seungkwon Beack, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea, Republic of

AUD-29.6: PROGRESSIVE MULTI-STAGE NEURAL AUDIO CODING WITH GUIDED 876
REFERENCES
Chanwoo Lee, Hyungseob Lim, Jihyun Lee, Hong-Goo Kang, Yonsei university, Korea, Republic of; Inseon Jang, Electronics and Telecommunications Research Institution, Korea, Republic of

AUD-30: AUDIO QUALITY AND SPEECH INTELLIGIBILITY MEASURES

AUD-30.1: VOCBENCH: A NEURAL VOCODER BENCHMARK FOR SPEECH 881
SYNTHESIS
Ehab A. AlBadawy, Ming-Ching Chang, University at Albany, State University of New York, United States of America; Andrew Gibiansky, Qing He, Jilong Wu, Facebook AI, United States of America; Siwei Lyu, University at Buffalo, State University of New York, USA, United States of America

AUD-30.2: DNSMOS P.835: A NON-INTRUSIVE PERCEPTUAL OBJECTIVE SPEECH 886
QUALITY METRIC TO EVALUATE NOISE SUPPRESSORS
Chandan Reddy, Google, United States of America; Vishak Gopal, Ross Cutler, Microsoft, United States of America

AUD-30.3: SQAPP: NO-REFERENCE SPEECH QUALITY ASSESSMENT VIA PAIRWISE 891
PREFERENCE
Pranay Manocha, Adam Finkelstein, Princeton University, United States of America; Zeyu Jin, Adobe Research, United States of America

AUD-30.4: LDNET: UNIFIED LISTENER DEPENDENT MODELING IN MOS 896
PREDICTION FOR SYNTHETIC SPEECH
Wen-Chin Huang, Tomoki Toda, Nagoya University, Japan; Erica Cooper, Junichi Yamagishi, National Institute of Informatics, Japan

AUD-30.5: AECMOS: A SPEECH QUALITY ASSESSMENT METRIC FOR ECHO 901
IMPAIRMENT
Marju Purin, Sten Sootla, Mateja Sponza, Ando Saabas, Ross Cutler, Microsoft Corporation, Estonia

AUD-30.6: MOS PREDICTOR FOR SYNTHETIC SPEECH WITH I-VECTOR INPUTS 906
Miao Liu, Jing Wang, Beijing Institute of Technology, China; Shicong Li, Fei Xiang, Xiaomi Inc., China; Yue Yao, Lidong Yang, Inner Mongolia University of Science and Technology, China

AUD-31: MULTICHANNEL PROCESSING

AUD-31.1: WAVE-DOMAIN APPROACH FOR CANCELLING NOISE ENTERING OPEN WINDOWS 911
Daan Ratering, Technical University Delft, Netherlands; W. Bastiaan Kleijn, Victoria University of Wellington, New Zealand; Jean Gonzalez Silva, Riccardo M. G. Ferrari, Delft University of Technology, Netherlands

AUD-31.2: ON SYNCHRONIZATION OF WIRELESS ACOUSTIC SENSOR NETWORKS IN THE PRESENCE OF TIME-VARYING SAMPLING RATE OFFSETS AND SPEAKER CHANGES 916
Tobias Gburrek, Joerg Schmalenstroeer, Reinhold Haeb-Umbach, Paderborn University, Germany

AUD-31.3: PICKNET: REAL-TIME CHANNEL SELECTION FOR AD HOC MICROPHONE ARRAYS 921
Takuya Yoshioka, Xiaofei Wang, Dongmei Wang, Microsoft, United States of America

AUD-31.4: END-TO-END ALEXA DEVICE ARBITRATION 926
Jarred Barber, Tao Zhang, Amazon, United States of America; Yifeng Fan, University of Illinois, Urbana-Champaign, United States of America

AUD-31.5: INSTANTANEOUS LINEAR DIMENSIONALITY REDUCTION OF MULTICHANNEL TIME-SERIES SIGNAL FOR ARRAY SIGNAL PROCESSING 931
Natsuki Ueno, Nobutaka Ono, Tokyo Metropolitan University, Japan

AUD-31.6: GENERALIZED TIME DOMAIN VELOCITY VECTOR 936
Srđan Kitić, Jérôme Daniel, Orange, France

AUD-32: AUDIO MUSIC SYNTHESIS

AUD-32.1: DIFFERENTIABLE DIGITAL SIGNAL PROCESSING MIXTURE MODEL FOR SYNTHESIS PARAMETER EXTRACTION FROM MIXTURE OF HARMONIC SOUNDS 941
Masaya Kawamura, Tomohiko Nakamura, Hiroshi Saruwatari, The University of Tokyo, Japan; Daichi Kitamura, National Institute of Technology, Kagawa College, Japan; Yu Takahashi, Kazunobu Kondo, Yamaha Corporation, Japan

AUD-32.2: THE MIRRORNET : LEARNING AUDIO SYNTHESIZER CONTROLS INSPIRED BY SENSORIMOTOR INTERACTION 946
Yashish M. Siriwardena, Shihab Shamma, Institute for Systems Research, University of Maryland College Park, United States of America; Guilhem Marion, Laboratoire des Systèmes Perceptifs, École Normale Supérieure, PSL University, France

AUD-32.3: DEEP PERFORMER: SCORE-TO-AUDIO MUSIC PERFORMANCE SYNTHESIS 951
Hao-Wen Dong, Taylor Berg-Kirkpatrick, Julian McAuley, University of California San Diego, United States of America; Cong Zhou, Dolby Laboratories, United States of America

AUD-32.4: KARASINGER: SCORE-FREE SINGING VOICE SYNTHESIS WITH VQ-VAE USING MEL-SPECTROGRAMS 956
Chien-Feng Liao, Jen-Yu Liu, Yi-Hsuan Yang, Taiwan AI Labs, Taiwan

AUD-32.5: ADVERSARIAL AUDIO SYNTHESIS USING A HARMONIC-PERCUSSIVE DISCRIMINATOR 961
Jihyun Lee, Hyungseob Lim, Chanwoo Lee, Hong-Goo Kang, Yonsei University, Korea, Republic of; Inseon Jang, Electronics and Telecommunications Research Institution, Korea, Republic of

AUD-32.6: SLEEPGAN: TOWARDS PERSONALIZED SLEEP THERAPY MUSIC 966
Jing Yang, Swiss Federal Institute of Technology in Zurich, Switzerland; Chulhong Min, Akhil Mathur, Fahim Kawsar, Nokia Bell Labs, United Kingdom of Great Britain and Northern Ireland

AUD-33: EXTENDED EVALUATION AND CAPTIONING

AUD-33.1: DIVERSITY-CONTROLLABLE AND ACCURATE AUDIO CAPTIONING BASED ON NEURAL CONDITION 971
Xuenan Xu, Mengyue Wu, Kai Yu, Shanghai Jiao Tong University, China

AUD-33.2: AUDIOCLIP: EXTENDING CLIP TO IMAGE, TEXT AND AUDIO 976
Andrey Guzhov, Federico Raue, Jörn Hees, Andreas Dengel, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Germany

AUD-33.3: CAN AUDIO CAPTIONS BE EVALUATED WITH IMAGE CAPTION METRICS? 981
Zelin Zhou, Zhiling Zhang, Xuenan Xu, Zeyu Xie, Mengyue Wu, Kenny Zhu, Shanghai Jiao Tong University, China

AUD-33.4: A DATA-DRIVEN COGNITIVE SALIENCE MODEL FOR OBJECTIVE PERCEPTUAL AUDIO QUALITY ASSESSMENT 986
Pablo M. Delgado, Jürgen Herre, International Audio Laboratories Erlangen, Germany

AUD-33.5: IMPROVING CHARACTER ERROR RATE IS NOT EQUAL TO HAVING CLEAN SPEECH: SPEECH ENHANCEMENT FOR ASR SYSTEMS WITH BLACK-BOX ACOUSTIC MODELS 991
Ryosuke Sawata, Yosuke Kashiwagi, Shusuke Takahashi, Sony Group Corporation, Japan

AUD-33.6: EFFECT OF NOISE SUPPRESSION LOSSES ON SPEECH DISTORTION AND ASR PERFORMANCE 996
Sebastian Braun, Hannes Gamper, Microsoft, United States of America

AUD-34: DECLIPPING, LOUDNESS, AND TIME DOMAIN MODIFICATION

AUD-34.1: INCREASING LOUDNESS IN AUDIO SIGNALS: A PERCEPTUALLY MOTIVATED APPROACH TO PRESERVE AUDIO QUALITY 1001
Alix Jeannerot, Paolo Prandoni, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland; Niels de Koeijer, Delft University of Technology (TU Delft), Netherlands; Pablo Martínez-Nuevo, Martin Bo Møller, Jakob Dyreby, Bang & Olufsen, Denmark

AUD-34.2: AUDIO PEAK REDUCTION USING A SYNCED ALLPASS FILTER 1006
Sebastian J. Schlecht, Leonardo Fierro, Vesa Välimäki, Aalto University, Finland; Juha Backman, AAC Technologies Solutions, Finland

AUD-34.3: APPLADE: ADJUSTABLE PLUG-AND-PLAY AUDIO DECLIPPER COMBINING DNN WITH SPARSE OPTIMIZATION 1011
Tomoro Tanaka, Kohei Yatabe, Yasuhiro Oikawa, Waseda University, Japan; Masahiro Yasuda, NTT Corporation, Japan

: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS VII: EVALUATION

.1: MAXIMIZING AUDIO EVENT DETECTION MODEL PERFORMANCE ON SMALL DATASETS THROUGH KNOWLEDGE TRANSFER, DATA AUGMENTATION, AND PRETRAINING: AN ABLATION STUDY 1016
Daniel Tompkins, Kshitiz Kumar, Jian Wu, Microsoft, United States of America

.2: THRESHOLD INDEPENDENT EVALUATION OF SOUND EVENT DETECTION SCORES 1021
Janek Ebbers, Reinhold Haeb-Umbach, Paderborn University, Germany; Romain Serizel, Université de Lorraine, France

.3: MULTIMODAL EVALUATION METHOD FOR SOUND EVENT DETECTION..... 1026
Seyed Mohammad Reza Modaresi, Aomar Osmani, Sorbonne Paris Nord University, France; Mohammadreza Razzazi, Amirkabir University of Technology, Iran (Islamic Republic of); Abdelghani Chibani, Université Paris-Est Créteil, France

.4: A BENCHMARK OF STATE-OF-THE-ART SOUND EVENT DETECTION SYSTEMS 1031
EVALUATED ON SYNTHETIC SOUNDSCAPES
Francesca Ronchini, Romain Serizel, INRIA Grand Est, France

AUD-35: TOPICS IN CONVOLUTIONAL AND LATENT VARIABLE MODELING

AUD-35.1: ATTENTIVE MAX FEATURE MAP AND JOINT TRAINING FOR ACOUSTIC 1036
SCENE CLASSIFICATION
Hye-jin Shim, Ju-ho Kim, Ha-Jin Yu, University of Seoul, Korea, Republic of; Jee-weon Jung, Naver corporation, Korea, Republic of

AUD-35.2: A VARIATIONAL BAYESIAN APPROACH TO LEARNING LATENT VARIABLES 1041
FOR ACOUSTIC KNOWLEDGE TRANSFER
Hu Hu, Chao-Han Huck Yang, Chin-Hui Lee, Georgia Institute of Technology, United States of America; Sabato Marco Siniscalchi, University of Enna Kore, Italy

AUD-35.5: ORCA-PARTY: AN AUTOMATIC KILLER WHALE SOUND TYPE SEPARATION 1046
TOOLKIT USING DEEP LEARNING
Christian Bergler, Manuel Schmitt, Andreas Maier, Elmar Nöth, Friedrich-Alexander-University Erlangen-Nuremberg, Germany; Rachael Xi Cheng, Leibniz Institute for Zoo and Wildlife Research, Germany; Volker Barth, Anthro-Media, Germany

AUD-36: SPATIAL AUDIO II

AUD-36.1: SPARSITY-BASED SOUND FIELD SEPARATION IN THE SPHERICAL 1051
HARMONICS DOMAIN
Mirco Pezzoli, Fabio Antonacci, Augusto Sarti, Politecnico di Milano, Italy; Maximo Cobos, Universitat de Valencia, Spain

AUD-36.2: SPATIAL ACTIVE NOISE CONTROL BASED ON INDIVIDUAL KERNEL 1056
INTERPOLATION OF PRIMARY AND SECONDARY SOUND FIELDS
Kazuyuki Arikawa, Shoichi Koyama, Hiroshi Saruwatari, The University of Tokyo, Japan

AUD-36.3: TIME-DOMAIN ACOUSTIC CONTRAST CONTROL WITH A SPATIAL 1061
UNIFORMITY CONSTRAINT FOR PERSONAL AUDIO SYSTEMS
Sipei Zhao, Ian Burnett, University of Technology Sydney, Australia

AUD-36.4: GENERATION OF PERSONAL SOUND FIELDS IN REVERBERANT 1066
ENVIRONMENTS USING INTERFRAME CORRELATION
Liming Shi, Mads Graesboll Christensen, Aalborg University, Denmark; Guoli Ping, Xiaoxiang Shen, Huawei Technologies Co., Ltd, Denmark

AUD-36.5: VARIABLE SPAN TRADE-OFF FILTER FOR SOUND ZONE CONTROL 1071
WITH KERNEL INTERPOLATION WEIGHTING
Jesper Brunnström, Marc Moonen, KU Leuven, Belgium; Shoichi Koyama, The University of Tokyo, Japan

AUD-36.6: TIME DOMAIN RADIAL FILTER DESIGN FOR SPHERICAL WAVES..... 1076
Nara Hahn, Frank Schultz, Sascha Spors, University of Rostock, Germany

BIO-1: MEDICAL IMAGE SEGMENTATION I

BIO-1.1: FEATURE SPACE MESSAGE PASSING NETWORK FOR MEDICAL IMAGE SEMANTIC SEGMENTATION **1081**
Junxiao Sun, Ke Zhang, Shuyi Niu, Yan Zhang, Youyong Kong, Southeast University, China

BIO-1.2: CROSS-DOMAIN FEW-SHOT LEARNING FOR RARE-DISEASE SKIN LESION SEGMENTATION **1086**
Yixin Wang, Institute of Computing Technology, Chinese Academy of Sciences, China; Zhe Xu, Department of Biomedical Engineering, The Chinese University of Hong Kong, China; Jiang Tian, Zhongchao Shi, AI Lab, Lenovo Research, China; Jie Luo, Brigham and Women's Hospital, Harvard Medical School, China; Yang Zhang, Zhiqiang He, Lenovo Ltd, China; Jianping Fan, UNCC; AI Lab, Lenovo Research, China

BIO-1.3: ADAPTIVE PSEUDO LABELING FOR SOURCE-FREE DOMAIN ADAPTATION IN MEDICAL IMAGE SEGMENTATION **1091**
Chen Li, Wei Chen, Xin Luo, Yulin He, Yusong Tan, National University of Defense Technology, China

BIO-1.4: OBJECT DETECTION AND TRACKING IN ULTRASOUND SCANS USING AN OPTICAL FLOW AND SEMANTIC SEGMENTATION FRAMEWORK BASED ON CONVOLUTIONAL NEURAL NETWORKS **1096**
Abdullah Al-Battal, Imanuel Lerman, Truong Nguyen, University of California - San Diego, United States of America

BIO-1.5: HEURISTIC DROPOUT: AN EFFICIENT REGULARIZATION METHOD FOR MEDICAL IMAGE SEGMENTATION MODELS **1101**
Dachuan Shi, Ruiyang Liu, Linmi Tao, Chun Yuan, Tsinghua University, China

BIO-1.6: SUPERRESOLUTION AND SEGMENTATION OF OCT SCANS USING MULTI-STAGE ADVERSARIAL GUIDED ATTENTION TRAINING **1106**
Paria Jeihouni, Omid Dehzangi, Nasser M. Nasrabadi, West Virginia University, United States of America; Annahita Amireskandari, Ali Dabouei, Ali Rezai, West Virginia Univeristy, United States of America

BIO-2: MACHINE LEARNING FOR PHYSIOLOGICAL SIGNALS I

BIO-2.1: HEART RATE AND OXYGEN SATURATION ESTIMATION FROM FACIAL VIDEO WITH MULTIMODAL PHYSIOLOGICAL DATA GENERATION **1111**
Yusuke Akamatsu, Yoshifumi Onishi, Hitoshi Imaoka, NEC Corporation, Japan

BIO-2.2: EMGSE: ACOUSTIC/EMG FUSION FOR MULTIMODAL SPEECH ENHANCEMENT **1116**
Kuan-Chen Wang, National Taiwan University, Taiwan; Kai-Chun Liu, Hsin-Min Wang, Yu Tsao, Academia Sinica, Taiwan

BIO-2.3: A DILATED RESIDUAL VISION TRANSFORMER FOR ATRIAL FIBRILLATION DETECTION FROM STACKED TIME-FREQUENCY ECG REPRESENTATIONS **1121**
Sawon Pratiher, Apoorva Srivastava, Yedla Bindu Priyatha, Nirmalya Ghosh, Amit Patra, Indian Institute of Technology Kharagpur, India

BIO-2.4: CONTRASTIVE HEARTBEATS: CONTRASTIVE LEARNING FOR SELF-SUPERVISED ECG REPRESENTATION AND PHENOTYPING **1126**
Crystal T. Wei, Ming-En Hsieh, Chien-Liang Liu, Vincent S. Tseng, National Yang Ming Chiao Tung University, Taiwan, Province of China

BIO-2.5: UBIQUITOUS PHYSIOLOGICAL PREDICTION OF SUD PATIENTS' WELLNESS STATE USING MEMORY-BASED CONVOLUTIONAL MODELS **1131**
Omid Dehzangi, Paria Jeihouni, Victor Finomore, Nasser M. Nasrabadi, Ali Rezai, West Virginia University, United States of America; Jad Ramadan, West Virginia Univeristy, United States of America

BIO-2.6: JOINT HYPOGLYCEMIA PREDICTION AND GLUCOSE FORECASTING VIA DEEP MULTI-TASK LEARNING 1136
Mu Yang, University of Texas at Dallas, United States of America; Darpit Dave, Madhav Erraguntla, Gerard Cote, Ricardo Gutierrez-Osuna, Texas A&M University, United States of America

BIO-3: DEEP LEARNING FOR PHYSIOLOGICAL SIGNAL ANALYSIS I

BIO-3.1: SEGNET-BASED DEEP REPRESENTATION LEARNING FOR DYSPHAGIA CLASSIFICATION 1141
Siddharth Subramani, Achuth Rao M V, Anwesha Roy, Prasanta Ghosh, Indian Institute of Science, India; Prasanna Suresh Hegde, Health Care Global Enterprises Ltd, India

BIO-3.2: ROBUST COLLABORATIVE LEARNING FOR SEQUENCE MODELLING 1146
Francois Buet-Golfouse, University College London, United Kingdom of Great Britain and Northern Ireland; Hans Roggeman, Islam Utyagulov, Independent, United Kingdom of Great Britain and Northern Ireland

BIO-3.3: A SELF-SUPERVISED PRE-TRAINING FRAMEWORK FOR VISION-BASED SEIZURE CLASSIFICATION 1151
Jen-Cheng Hou, Monique Thonnat, University of Côte d'Azur, INRIA, France; Aileen McGonigal, Fabrice Bartolomei, Aix-Marseille University, France

BIO-3.4: DESIGN OF REAL-TIME SYSTEM BASED ON MACHINE LEARNING FOR SNORING AND OSA DETECTION 1156
Huaiwen Luo, Lu Zhang, Lianyu Zhou, Xu Lin, Zehuai Zhang, Mingjiang Wang, Harbin Institute of Technology, Shenzhen, China

BIO-3.5: PARAMETRIC MODELING OF HUMAN WRIST FOR BIOIMPEDANCE-BASED PHYSIOLOGICAL SENSING 1161
Kaan Sel, Noah Huerta, Roozbeh Jafari, Texas A&M University, United States of America; Michael S. Sacks, The University of Texas at Austin, United States of America

BIO-3.6: PRELIMINARY RESULTS ON THE GENERATION OF ARTIFICIAL HANDWRITING DATA USING A DECOMPOSITION-RECOMBINATION STRATEGY 1166
Jose Fernando Adrán Otero, Oscar Solans Caballer, Open University of Catalonia, Spain; Pere Marti-Puig, Jordi Solé-Casals, University of Vic - Central University of Catalonia, Spain; Zhe Sun, RIKEN, Japan; Toshihisa Tanaka, Tokyo University of Agriculture and Technology, Japan

BIO-4: MACHINE LEARNING FOR PHYSIOLOGICAL SIGNALS II

BIO-4.1: A STYLE TRANSFER MAPPING AND FINE-TUNING SUBJECT TRANSFER FRAMEWORK USING CONVOLUTIONAL NEURAL NETWORKS FOR SURFACE ELECTROMYOGRAM PATTERN RECOGNITION 1171
Suguru Kanoga, Takayuki Hoshino, Mitsunori Tada, National Institute of Advanced Industrial Science and Technology (AIST), Japan

BIO-4.2: FEATURE-BASED SENSING MATRIX DESIGN FOR ANALOG TO INFORMATION CONVERTERS 1176
Chencheng Guo, Hui Qian, Baoling Hong, Fuzhou University, China

BIO-4.3: ALSNET: A DILATED 1-D CNN FOR IDENTIFYING ALS FROM RAW EMG SIGNAL 1181
K. M. Naimul Hassan, Md. Shamiul Alam Hridoy, Naima Tasnim, Atia Faria Chowdhury, Tanvir Alam Roni, Sheikh Tabrez, Arik Subhana, Celia Shahnaz, Bangladesh University of Engineering and Technology, Bangladesh

BIO-4.4: JOINT MODEL ORDER ESTIMATION FOR MULTIPLE TENSORS WITH A COUPLED MODE AND APPLICATIONS TO THE JOINT DECOMPOSITION OF EEG, MEG MAGNETOMETER, AND GRADIOMETER TENSORS 1186
Bilal Ahmad, Liana Khamidullina, Alla Manina, Jens Haueisen, Martin Haardt, Ilmenau University of Technology, Germany; Alexey Korobkov, Kazan National Research Technical University n.a. A.N. Tupolev-KAI, Russian Federation

BIO-4.5: AN EXPERIMENTAL STUDY ON TRANSFERRING DATA-DRIVEN IMAGE COMPRESSIVE SENSING TO BIOELECTRIC SIGNAL 1191
Zhikang Zhang, Jonathan Zhao, Fengbo Ren, Arizona State University, United States of America

BIO-4.6: HAND GESTURE RECOGNITION USING TEMPORAL CONVOLUTIONS AND ATTENTION MECHANISM 1196
Elahe Rahimian, Soheil Zabihi, Amir Asif, Arash Mohammadi, Concordia University, Canada; Dario Farina, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Farokh Atashzar, New York University, United States of America

BIO-5: DEEP LEARNING FOR MEDICAL IMAGE ANALYSIS

BIO-5.1: COMBINING MULTIPLE STYLE TRANSFER NETWORKS AND TRANSFER LEARNING FOR LGE-CMR SEGMENTATION 1201
Bo Fang, Junxin Chen, Northeastern University, China; Wei Wang, Sun Yat-sen University, China; Yicong Zhou, University of Macau, China

BIO-5.2: MULTI-DOMAIN UNPAIRED ULTRASOUND IMAGE ARTIFACT REMOVAL USING A SINGLE CONVOLUTIONAL NEURAL NETWORK 1206
Jaeyoung Huh, Shujaat Khan, Jong Chul Ye, Korea Advanced Institute of Science and Technology, Korea, Republic of

BIO-5.3: IMPROVING ULTRASOUND IMAGE CLASSIFICATION WITH LOCAL TEXTURE QUANTISATION 1211
Xiao Li, Huizhi Liang, University of Reading, United Kingdom of Great Britain and Northern Ireland; Sidhartha Nagala, Jane Chen, Royal Berkshire Hospital, United Kingdom of Great Britain and Northern Ireland

BIO-5.4: ACCELERATED INTRAVASCULAR ULTRASOUND IMAGING USING DEEP REINFORCEMENT LEARNING 1216
Tristan Stevens, Nishith Chennakeshava, Ruud van Sloun, Eindhoven University of Technology, Netherlands; Frederik de Bruijn, Martin Pekař, Philips Research Eindhoven, Netherlands

BIO-5.5: DEEP PROXIMAL UNFOLDING FOR IMAGE RECOVERY FROM UNDER-SAMPLED CHANNEL DATA IN INTRAVASCULAR ULTRASOUND 1221
Nishith Chennakeshava, Tristan S.W. Stevens, Massimo Mischi, Eindhoven University of Technology, Netherlands; Frederik J. de Bruijn, Andrew Hancock, Martin Pekarˇ, Philips Research, Netherlands; Yonina C. Eldar, The Weizmann Institute of Science, Netherlands; Ruud J.G. van Sloun, Eindhoven University of Technology, Philips Research, Netherlands

BIO-5.6: MULTIVIEW LONG-SHORT SPATIAL CONTRASTIVE LEARNING FOR 3D MEDICAL IMAGE ANALYSIS 1226
Gongpeng Cao, Yiping Wang, Manli Zhang, Jing Zhang, Guixia Kang, Beijing University of Posts and Telecommunications, China; Xin Xu, General Hospital of People's Liberation Army, China

BIO-6: MACHINE LEARNING AND SIGNAL PROCESSING FOR EEG

BIO-6.1: COMPOSING GRAPHICAL MODELS WITH GENERATIVE ADVERSARIAL NETWORKS FOR EEG SIGNAL MODELING 1231
Khuong Vo, Manoj Vishwanath, Ramesh Srinivasan, Nikil Dutt, Hung Cao, University of California, Irvine, United States of America

BIO-6.2: DOMAIN-INVARIANT REPRESENTATION LEARNING FROM EEG WITH PRIVATE ENCODERS 1236
David Bethge, Porsche, LMU, Germany; Philipp Hallgarten, Porsche, KIT, Germany; Tobias Grosse-Puppendahl, Mohamed Kari, Porsche, Germany; Ralf Mikut, KIT, Germany; Albrecht Schmidt, LMU, Germany; Ozan Özdenizci, TU Graz, Austria

BIO-6.3: HOLISTIC SEMI-SUPERVISED APPROACHES FOR EEG REPRESENTATION LEARNING 1241
Guangyi Zhang, Ali Etemad, Queen's University, Canada

BIO-6.4: MUSIC IDENTIFICATION USING BRAIN RESPONSES TO INITIAL SNIPPETS 1246
Pankaj Pandey, Krishna. P. Miyapuram, IIT Gandhinagar, India; Gulshan Sharma, IIT Ropar, India; Ramanathan Subramanian, University of Canberra, Australia; Derek Lomas, TU Delft, Netherlands

BIO-6.5: MULTI-LEVEL SPATIAL-TEMPORAL ADAPTATION NETWORK FOR MOTOR IMAGERY CLASSIFICATION 1251
Wei Xu, Jing Wang, Ziyu Jia, Zhiqing Hong, Yunze Li, Youfang Lin, Beijing Jiaotong University, China

BIO-6.6: LEARNING SUBJECT-INVARIANT REPRESENTATIONS FROM SPEECH-EVOKED EEG USING VARIATIONAL AUTOENCODERS 1256
Lies Bollens, Tom Francart, Hugo Van hamme, Katholieke Universiteit Leuven, Belgium

BIO-7: MACHINE LEARNING FOR MEDICAL IMAGING

BIO-7.1: UNSUPERVISED HIERARCHICAL TRANSLATION-BASED MODEL FOR MULTI-MODAL MEDICAL IMAGE REGISTRATION 1261
Xinru Dai, Tai Ma, Haibin Cai, Ying Wen, East China Normal University, China

BIO-7.2: FAZ-BV: A DIABETIC MACULAR ISCHEMIA GRADING FRAMEWORK COMBINING FAZ ATTENTION NETWORK AND BLOOD VESSEL ENHANCEMENT FILTERS 1266
Zailiang Chen, Hailei Lan, Yuchen Xiong, Hailan Shen, School of Computer Science and Engineering, Central South University, Changsha, China, China; Yongan Meng, Jing Luo, The Second Xiangya Hospital, Central South University, Changsha, China, China

BIO-7.3: FRACTURE DETECTION AND LOCALIZATION IN CHEST X-RAYS USING SEMI-SUPERVISED LEARNING WITH DYNAMIC SHARPENING 1271
Lijuan Lu, Tsinghua University, China; Shun Miao, PAII Inc., United States of America; Ling Ye, Ping An Technology, China

BIO-7.4: HISTOKT: CROSS KNOWLEDGE TRANSFER IN COMPUTATIONAL PATHOLOGY 1276
Ryan Zhang, Jiadai Zhu, Stephen Yang, Konstantinos Plataniotis, University of Toronto, Canada; Mahdi Hosseini, University of New Brunswick, Canada; Angelo Genovese, Università degli Studi di Milano, Italy; Lina Chen, Sunnybrook Health Sciences Centre, Canada; Corwyn Rowsell, St. Michaels Hospital, Canada; Savvas Damaskinos, Huron Digital Pathology, Canada; Sonal Varma, Kingston Health Sciences Center, Canada

BIO-7.5: UNSUPERVISED DEEP LEARNING NETWORK FOR DEFORMABLE FUNDUS IMAGE REGISTRATION 1281
Giovana Benvenuto, Marilaine Colnago, Wallace Casaca, São Paulo State University, Brazil

BIO-7.6: A MINIMALLY SUPERVISED APPROACH FOR MEDICAL IMAGE QUALITY ASSESSMENT IN DOMAIN SHIFT SETTINGS 1286
Huijuan Yang, Feri Guretno, Ivan Ho Mien, Chuan Sheng Foo, Pavitra Krishnaswamy, Institute for Infocomm Research, Singapore; Aaron S. Coyner, J. Peter Campbell, Susan Ostmo, Oregon Health & Science University, Singapore; Michael F. Chiang, National Institutes of Health, United States of America

BIO-8: BRAIN-COMPUTER INTERFACES

BIO-8.1: A CHANNEL ATTENTION BASED MLP-MIXER NETWORK FOR MOTOR IMAGERY DECODING WITH EEG 1291
Yanbin He, Zhiyang Lu, Jun Wang, Jun Shi, Shanghai University, China

BIO-8.2: TOWARDS CLOSED-LOOP SPEECH SYNTHESIS FROM STEREOTACTIC EEG: A UNIT SELECTION APPROACH 1296
Miguel Angrick, Lorenz Diener, Darius Ivucic, Gabriel Ivucic, Tanja Schultz, University of Bremen, Germany; Maarten Ottenhoff, Sophocles Goulis, Pieter L. Kubben, Christian Herff, Maastricht University, Netherlands; Albert J. Colon, Louis Wagner, Epilepsy Center Kempenhaeghe, Netherlands; Dean J. Krusienski, Virginia Commonwealth University, United States of America

BIO-8.3: ENHANCING CONTEXTUAL ENCODING WITH STAGE-CONFUSION AND STAGE-TRANSITION ESTIMATION FOR EEG-BASED SLEEP STAGING 1301
Jaeun Phyo, Wonjun Ko, Eunjin Jeon, Heung-Il Suk, Korea University, Korea, Republic of

BIO-8.4: IMPROVING BCI-BASED COLOR VISION ASSESSMENT USING GAUSSIAN PROCESS REGRESSION 1306
Hadi Habibzadeh, Daphney-Stavroula Zois, University at Albany, State University of New York, United States of America; Kevin J. Long, James J. S. Norton, US Department of Veterans Affairs, United States of America; Ally E. Atkins, Union College, United States of America

BIO-8.5: TRANSFORMER-BASED ESTIMATION OF SPOKEN SENTENCES USING ELECTROCORTICOGRAPHY 1311
Shuji Komeiji, Kai Shigemi, Toshihisa Tanaka, Tokyo University of Agriculture and Technology, Japan; Takumi Mitsuhashi, Yasushi Iimura, Hiroharu Suzuki, Hidenori Sugano, Juntendo University, Japan; Koichi Shinoda, Tokyo Institute of Technology, Japan

BIO-9: DEEP LEARNING FOR PHYSIOLOGICAL SIGNAL ANALYSIS II

BIO-9.1: BOOST ENSEMBLE LEARNING FOR CLASSIFICATION OF CTG SIGNALS 1316
Marzieh Ajirak, Petar Djuric, Stony Brook University, United States of America; Cassandra Heiselman, Gerald Quirk, Stony Brook Medicine, United States of America

BIO-9.2: MULTI-VIEW LEARNING BASED ON NON-REDUNDANT FUSION FOR ICU PATIENT MORTALITY PREDICTION 1321
Yifan Wang, Ying Lan, Peking University, China

BIO-9.3: IMPROVING PHASE-RECTIFIED SIGNAL AVERAGING FOR FETAL HEART RATE ANALYSIS 1326
Tong Chen, Guanchao Feng, Cassandra Heiselman, J. Gerald Quirk, Petar M. Djuric, Stony Brook University, United States of America

BIO-9.4: UNSUPERVISED CLUSTERING AND ANALYSIS OF CONTRACTION-DEPENDENT FETAL HEART RATE SEGMENTS 1331
Liu Yang, Petar M. Djurić, Stony Brook University, United States of America; Cassandra Heiselman, J. Gerald Quirk, Stony Brook University Hospital, United States of America

BIO-9.5: A METHOD FOR DETECTING CORONARY ARTERY DISEASE USING NOISY ULTRASHORT ELECTROCARDIOGRAM RECORDINGS 1336
Orestis Apostolou, Vasileios Charisis, Georgios Apostolidis, Aristotle University of Thessaloniki, Greece; Leontios J. Hadjileontiadis, Aristotle University of Thessaloniki, Khalifa University, Greece

BIO-9.6: MULTI-TASK GAUSSIAN PROCESS REGRESSION FOR THE DETECTION OF SLEEP CYCLES IN PREMATURE INFANTS 1341
Nele Sophie Brügge, University of Lübeck, Germany; Jan Grasshoff, Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering, Germany; Arne Weigenand, Drägerwerk AG & Co. KGaA, Germany; Philipp Rostalski, University of Lübeck and Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering, Germany

BIO-10: DETECTION, ESTIMATION AND CLASSIFICATION FOR MEDICAL IMAGES

BIO-10.1: FAST LOW RANK COLUMN-WISE COMPRESSIVE SENSING FOR ACCELERATED DYNAMIC MRI 1346
Silpa Babu, Seyedehsara Nayer, Namrata Vaswani, Iowa State University, United States of America; Sajan Lingala, University of Iowa, United States of America

BIO-10.2: MRI RECOVERY WITH A SELF-CALIBRATED DENOISER 1351
Sizhuo Liu, Philip Schniter, Rizwan Ahmad, Ohio State University, United States of America

BIO-10.3: 3D CROSS-SCALE FEATURE TRANSFORMER NETWORK FOR BRAIN MR IMAGE SUPER-RESOLUTION 1356
Wanqi Zhang, Lulu Wang, Wei Chen, Zhongshi He, Chongqing University, China; Yuanyuan Jia, Jinglong Du, Chongqing Medical University, China

BIO-10.4: DATA EFFICIENT SUPPORT VECTOR MACHINE TRAINING USING THE MINIMUM DESCRIPTION LENGTH PRINCIPLE 1361
Harsh Singh, International Institute of Information Technology, India; Ognjen Arandjelovic, University of St Andrews, United Kingdom of Great Britain and Northern Ireland

BIO-10.5: MULTIPLE INSTANCE LEARNING WITH TASK-SPECIFIC MULTI-LEVEL FEATURES FOR WEAKLY ANNOTATED HISTOPATHOLOGICAL IMAGE CLASSIFICATION 1366
Yuanpin Zhou, Yao Lu, Sun Yat-sen University, China

BIO-11: MEDICAL IMAGE ANALYSIS FOR DIAGNOSIS

BIO-11.1: SELF-KNOWLEDGE DISTILLATION BASED SELF-SUPERVISED LEARNING FOR COVID-19 DETECTION FROM CHEST X-RAY IMAGES 1371
Guang Li, Ren Togo, Takahiro Ogawa, Miki Haseyama, Hokkaido University, Japan

BIO-11.2: PIXEL-LEVEL AND AFFINITY-LEVEL KNOWLEDGE DISTILLATION FOR UNSUPERVISED SEGMENTATION OF COVID-19 LESIONS 1376
Rui Xu, Yufeng Wang, Xinchen Ye, Pengcheng Wu, Dalian University of Technology, China; Yen-Wei Chen, Ritsumeikan University, Japan; Fangyi Xu, Wenchao Zhu, Chao Chen, Yong Zhou, Hongjie Hu, Zhejiang University School of Medicine, China; Xiaofeng Qu, the Second Hospital of Dalian Medical University, China; Shoji Kido, Noriyuki Tomiyama, Osaka University, Japan

BIO-11.3: DATA SHAPLEY VALUE FOR HANDLING NOISY LABELS: AN APPLICATION IN SCREENING COVID-19 PNEUMONIA FROM CHEST CT SCANS 1381
Nastaran Enshaei, Arash Mohammadi, Farnoosh Naderkhani, Concordia University, Canada; Moezedin Javad Rafiee, McGill University, Canada

BIO-11.4: ACCURATE MULTISCALE SELECTIVE FUSION OF CT AND VIDEO IMAGES FOR REAL-TIME ENDOSCOPIC CAMERA 3D TRACKING IN ROBOTIC SURGERY 1386
Xiongbiao Luo, Xiamen University, China

BIO-11.5: LEARNING DEEP PATHOLOGICAL FEATURES FOR WSI-LEVEL CERVICAL CANCER GRADING 1391
Ruixiang Geng, Qing Liu, Shuo Feng, Yixiong Liang, Central South University, China

BIO-11.6: SELECTIVE SCALE CASCADE ATTENTION NETWORK FOR BREAST CANCER HISTOPATHOLOGY IMAGE CLASSIFICATION 1396
Bowen Xu, Wenqiang Zhang, Fudan University, China

BIO-12: NEURAL SIGNAL PROCESSING

BIO-12.1: FREQUENCY-SPECIFIC NON-LINEAR GRANGER CAUSALITY IN A NETWORK OF BRAIN SIGNALS 1401
Archishman Biswas, Hernando Ombao, King Abdullah University of Science and Technology, Saudi Arabia

BIO-12.2: EPILEPTIC SPIKE DETECTION BY RECURRENT NEURAL NETWORKS WITH SELF-ATTENTION MECHANISM 1406
Kosuke Fukumori, Toshihisa Tanaka, Tokyo University of Agriculture and Technology, Japan; Noboru Yoshida, Juntendo University Nerima Hospital, Japan; Hidenori Sugano, Madoka Nakajima, Juntendo University Hospital, Japan

BIO-12.3: TOPOLOGICAL CORRELATION OF BRAIN SIGNALS 1411
Jian Yin, Nanjing Medical University, China; Yuan Wang, University of South Carolina, United States of America

BIO-12.4: ONLINE DETECTION OF SCALP-INVISIBLE MESIAL-TEMPORAL BRAIN INTERICTAL EPILEPTIFORM DISCHARGES FROM EEG 1416
Bahman Abdi-Sargezeh, Saeid Sanei, Nottingham Trent University, United Kingdom of Great Britain and Northern Ireland; Antonio Valentin, King's College London, United Kingdom of Great Britain and Northern Ireland; Gonzalo Alarcon, Hamad General Hospital, Qatar

BIO-12.5: LEVERAGING SPARSE CODING FOR EEG BASED EMOTION RECOGNITION IN SHOOTING 1421
Yulu Wang, Yiwen Sun, Changshui Zhang, Tsinghua University, China; Fang Lei, DataCanvas Technology Co., Ltd., China

BIO-12.6: A NOVEL UNSUPERVISED AUTOENCODER-BASED HFOS DETECTOR IN INTRACRANIAL EEG SIGNALS 1426
Weilai Li, Lanfeng Zhong, Weixi Xiang, Tongzhou Kang, Dakun Lai, University of Electronic Science and Technology of China, China

BIO-13: MEDICAL IMAGE SEGMENTATION II

BIO-13.1: A NOVEL CONVOLUTIONAL NEURAL NETWORK BASED ON ADAPTIVE MULTI-SCALE AGGREGATION AND BOUNDARY-AWARE FOR LATERAL VENTRICLE SEGMENTATION ON MR IMAGES 1431
Fei Ye, Kai Hu, Xiangtan University, China; Zhiqiang Wang, Sheng Zhu, Affiliated Hospital of Xiangnan University, China; Xuanya Li, Baidu Inc., China

BIO-13.2: MULTISCALE ATTENTION AGGREGATION NETWORK FOR 2D VESSEL SEGMENTATION 1436
Wentao Liu, Huihua Yang, Weijin Xu, School of Artificial Intelligence, Beijing University of Posts and Telecommunications, China; Tong Tian, the State Key Laboratory of Structural Analysis for Industrial Equipment, School of Aeronautics and Astronautics, Dalian University of Technology, China; Xipeng Pan, the School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, China, China

BIO-13.3: TCRNET: MAKE TRANSFORMER, CNN AND RNN COMPLEMENT EACH OTHER 1441
Xinxin Shan, Tai Ma, Anqi Gu, Haibin Cai, Ying Wen, East China Normal University, China

BIO-13.4: DOUBLE NOISE MEAN TEACHER SELF-ENSEMBLING MODEL FOR SEMI-SUPERVISED TUMOR SEGMENTATION 1446
Ke Zheng, Junhai Xu, Jianguo Wei, Tianjin University, China

BIO-13.5: RETHINKING COMPUTER-AIDED PELVIS SEGMENTATION 1451
Siming Yuan, Qing Liu, Shenghui Liao, Fuchang Han, Central South University, China; Haitao Wei, Central South E3D Digital Medical and Virtual Reality Research Center, China; Yingqi Zhang, Tongji University, China

BIO-13.6: VISION TRANSFORMER-BASED RETINA VESSEL SEGMENTATION WITH DEEP ADAPTIVE GAMMA CORRECTION 1456
Hyunwoo Yu, Jae-hun Shim, Jaeho Kwak, Jou won Song, Suk-Ju Kang, Sogang University, Korea, Republic of

BIO-14: NEUROIMAGING DATA ANALYSIS

BIO-14.1: SPECTRAL PERMUTATION TEST ON PERSISTENCE DIAGRAMS 1461
Yuan Wang, Julius Fridriksson, University of South Carolina, United States of America; Moo Chung, University of Wisconsin - Madison, United States of America

BIO-14.2: MULTI-TASK FMRI DATA FUSION USING IVA AND PARAFAC2 1466
Isabell Lehmann, Tanuj Hasija, Peter J. Schreier, Paderborn University, Germany; Evrim Acar, Simula Research Laboratory, Norway; M.A.B.S. Akhonda, University of Maryland, Baltimore County, United States of America; Vince D. Calhoun, Tri-Institutional Georgia State University/Georgia Institute of Technology/Emory University Center, United States of America; Tülay Adali, University of University of Maryland, Baltimore County, United States of America

BIO-14.3: INDEPENDENT VECTOR ANALYSIS BASED SUBGROUP IDENTIFICATION FROM MULTISUBJECT FMRI DATA 1471
Hanlu Yang, Mohammad Abu Baker Siddique Akhonda, Fatemeh Ghayyem, Qunfang Long, Tülay Adali, University of Maryland Baltimore County, United States of America; Vince Calhoun, Georgia State University, Georgia Institute of Technology, and Emory University, United States of America

BIO-14.4: IMPROVING BRAIN DECODING METHODS AND EVALUATION 1476
Damián Pascual, Béni Egressy, Nicolas Affolter, Yiming Cai, Oliver Richter, Roger Wattenhofer, ETH Zurich, Switzerland

BIO-14.5: CMRI2SPEC: CINE MRI SEQUENCE TO SPECTROGRAM SYNTHESIS VIA A PAIRWISE HETEROGENEOUS TRANSLATOR 1481
Xiaofeng Liu, Jangwon Kim, Georges El Fakhri, Jonghye Woo, Harvard Medical School, United States of America; Fangxu Xing, MGH and Harvard Medical School, United States of America; Maureen Stone, UMD, United States of America; Jerry Prince, JHU, United States of America

BIO-14.6: SPATIO-TEMPORAL ATTENTION GRAPH CONVOLUTION NETWORK FOR FUNCTIONAL CONNECTOME CLASSIFICATION 1486
Wenhan Wang, Youyong Kong, Zhenghua Hou, Chunfeng Yang, Yonggui Yuan, Southeast University, China

CI-1: COMPUTATIONAL IMAGING METHODS AND MODELS

CI-1.1: BILEVEL LEARNING OF L1 REGULARIZERS WITH CLOSED-FORM GRADIENTS (BLORC) 1491
Avrajit Ghosh, Saiprasad Ravishankar, Michigan State University, United States of America; Michael Mccann, Los Alamos National Laboratory, United States of America

CI-1.2: MULTIBAND IMAGE FUSION WITH CONTROLLABLE ERROR GUARANTEES 1496
Unni V. S., Ruturaj Gavaskar, Kunal Chaudhury, Indian Institute of Science, India

CI-1.3: WEIGHTED GRAPH EMBEDDED LOW-RANK PROJECTION LEARNING FOR FEATURE EXTRACTION 1501
Zhuojie Huang, Shuping Zhao, Lunke Fei, Jigang Wu, Guangdong University of Technology, China

CI-1.4: ADMM-DAD NET: A DEEP UNFOLDING NETWORK FOR ANALYSIS COMPRESSED SENSING 1506
Vasiliki Kouni, George C. Alexandropoulos, National and Kapodistrian University of Athens, Greece; Georgios Paraskevopoulos, National Technical University of Athens, Greece; Holger Rauhut, Rheinisch-Westfälische Technische Hochschule Aachen University, Germany

CI-1.5: HIGH-DIMENSIONAL SPARSE BAYESIAN LEARNING WITHOUT COVARIANCE MATRICES 1511
Alexander Lin, Demba Ba, Harvard University, United States of America; Andrew Song, Massachusetts Institute of Technology, United States of America; Berkin Bilgic, Athinoula A. Martinos Center for Biomedical Imaging, United States of America

CI-1.6: A TRAINABLE BOUNDED DENOISER USING DOUBLE TIGHT FRAME NETWORK FOR SNAPSHOT COMPRESSIVE IMAGING 1516
Baoshun Shi, Yuxin Wang, Qiusheng Lian, Yanshan University, China

CI-1.7: PROGRESSIVE IMAGE SUPER-RESOLUTION VIA NEURAL DIFFERENTIAL EQUATION 1521
Seobin Park, Tae Hyun Kim, Hanyang University, Korea, Republic of

CI-2: COMPUTATIONAL IMAGING FORMATION

CI-2.1: HIGH-QUALITY SELF-SUPERVISED SNAPSHOT HYPERSPECTRAL IMAGING 1526
Yuhui Quan, Xinran Qin, Mingqin Chen, Yan Huang, South China University of Technology, China

CI-2.2: ROBUST BAYESIAN RECONSTRUCTION OF MULTISPECTRAL SINGLE-PHOTON 3D LIDAR DATA WITH NON-UNIFORM BACKGROUND 1531
Abderrahim Halimi, Jakeoung Koo, Gerald S. Buller, Stephen McLaughlin, Heriot-Watt University, United Kingdom of Great Britain and Northern Ireland; Robert Lamb, Leonardo MW Ltd,, United Kingdom of Great Britain and Northern Ireland

CI-2.3: JOINT CALIBRATION AND MAPPING OF SATELLITE ALTIMETRY DATA USING TRAINABLE VARIATIONAL MODELS 1536
Quentin Febvre, Ronan Fablet, IMT Atlantique, France; Julien Le Sommer, Université Grenoble Alpes, France; Clément Ubelmann, OceanNext, France

CI-2.4: 4D CONVOLUTIONAL NEURAL NETWORKS FOR MULTI-SPECTRAL AND MULTI-TEMPORAL REMOTE SENSING DATA CLASSIFICATION 1541
Michalis Giannopoulos, Panagiotis Tsakalides, University of Crete, Greece; Grigorios Tsagkatakis, Foundation for Research and Technology Hellas, Greece

CI-2.5: A NEW DEEP LEARNING METHOD FOR MULTISPECTRAL IMAGE TIME SERIES COMPLETION USING HYPERSPECTRAL DATA 1546
Cheick T. Cissé, Ahed Alboody, Matthieu Puigt, Gilles Roussel, Univ. Littoral Côte d'Opale, France; Vincent Vantrepotte, Cédric Jamet, Trung-Kien Tran, Univ. Littoral Côte d'Opale, CNRS, France

CI-2.6: IMAGE DENOISING WITH DEEP UNFOLDING AND NORMALIZING FLOWS 1551
Xinyi Wei, Hans van Gorp, Lizeth Gonzalez Carabarin, Ruud van Sloun, Eindhoven University of Technology, Netherlands; Daniel Freedman, Verily Research, Israel; Yonina Eldar, Weizmann Institute of Science, Israel

CI-2.7: 3D TEXTURE SUPER RESOLUTION VIA THE RENDERING LOSS 1556
Rohit Ranade, Yangwen Liang, Shuangquan Wang, Dongwoon Bai, Jungwon Lee, Samsung Semiconductor Inc., United States of America

CI-3: COMPUTATIONAL PHOTOGRAPHY

CI-3.1: BUNDLE ICP WITH VIRTUAL DEPTH FOR HAND-HELD 3D SCANNER 1561
Changhun Sung, Byungdeok Kim, SAMSUNG ELECTRONICS, Korea, Republic of

CI-3.2: SKETCHED RT3D: HOW TO RECONSTRUCT BILLIONS OF PHOTONS PER SECOND 1566
Julian Tachella, Michael P. Sheehan, Mike E. Davies, University of Edinburgh, United Kingdom of Great Britain and Northern Ireland

CI-3.3: A GENERIC METHOD TO ESTIMATE CAMERA EXTRINSIC PARAMETERS 1571
Naveen Kuruba, Neel Badadare, Vikram Narayan, Satish Putta, Visteon, India

CI-3.4: PHOTON-LIMITED DEBLURRING USING ALGORITHM UNROLLING 1576
Yash Sanghvi, Abhiram Gnanasambandan, Stanley Chan, Purdue University, United States of America

CI-3.5: NEX+: NOVEL VIEW SYNTHESIS WITH NEURAL REGULARISATION OVER MULTI-PLANE IMAGES 1581
Wenpeng Xing, Jie Chen, Hong Kong Baptist University, Hong Kong

CI-4: COMPUTATIONAL IMAGING SYSTEMS

CI-4.1: COMPRESSIVE SCANNING TRANSMISSION ELECTRON MICROSCOPY 1586
Daniel Nicholls, Alex Robinson, Jack Wells, Mounib Bahri, Nigel Browning, University of Liverpool, United Kingdom of Great Britain and Northern Ireland; Amirafshar Moshtaghpour, Rosalind Franklin Institute, United Kingdom of Great Britain and Northern Ireland; Angus Kirkland, University of Oxford, United Kingdom of Great Britain and Northern Ireland

CI-4.2: DEEP ITERATIVE PHASE RETRIEVAL FOR PTYCHOGRAPHY 1591
Simon Welker, Tal Peer, Timo Gerkmann, Universität Hamburg, Germany; Henry Chapman, Deutsches Elektronen-Synchrotron, Germany

CI-4.3: COMPRESSIVE PHASE RETRIEVAL BASED ON SPARSE LATENT GENERATIVE PRIORS 1596
Vinayak Killedar, Chandra Sekhar Seelamantula, Indian Institute of Science, India

CI-4.4: MODEL-BASED RECONSTRUCTION FOR COLLIMATED BEAM ULTRASOUND SYSTEMS 1601
Abdulrahman Alanazi, Gregery Buzzard, Charles Bouman, Purdue University-Main Campus, United States of America; Singanallur Venkatakrishnan, Hector Santos-Villalobos, Oak Ridge National Laboratory, United States of America

CI-4.5: LEARNED ACOUSTIC RECONSTRUCTION USING SYNTHETIC APERTURE FOCUSING 1606
Tim Straubinger, Robert Xiao, Helge Rhodin, University of British Columbia, Canada

CI-4.6: SDETR: ATTENTION-GUIDED SALIENT OBJECT DETECTION WITH TRANSFORMER 1611
Guanze Liu, Bo Xu, Han Huang, Yandong Guo, OPPO Research Institute, China; Cheng Lu, Xmotors, China

IVMSP-1: IMAGE AND VIDEO CODING I

IVMSP-1.1: EVALUATION OF VIDEO CODING FOR MACHINES WITHOUT GROUND TRUTH 1616
Kristian Fischer, André Kaup, Friedrich-Alexander-University Erlangen-Nürnberg, Germany; Markus Hofbauer, Christopher Kuhn, Eckehard Steinbach, Technical University of Munich, Germany

IVMSP-1.2: RAW PLENOPTIC VIDEO CODING UNDER HEXAGONAL LATTICE RESOLUTION OF MOTION VECTORS 1621
Thuc Nguyen Huu, Vinh Duong Van, Jonghoon Yim, Byeungwoo Jeon, Sungkyunkwan University, Korea, Republic of

IVMSP-1.3: COMPARISON OF BOUNDARY ARTIFACT REMOVAL METHODS IN CODING OF GENERALIZED CUBEMAP PROJECTION USING VVC 1625
Kianoush Jafari, Alireza Aminlou, Miska Hannuksela, Nokia Technologies, Finland

IVMSP-1.4: LOW-COMPLEXITY MULTI-MODEL CNN IN-LOOP FILTER FOR AVS3 1630
Shen Wang, Yibing Fu, Chen Zhu, Li Song, Wenjun Zhang, Shanghai Jiao Tong University, China

IVMSP-1.5: UNIFIED MATRIX CODING FOR NN ORIGINATED MIP IN H.266/VVC 1635
Junyan Huo, Yu Sun, Haixin Wang, Fuzheng Yang, Xidian University, China; Shuai Wan, Northwestern Polytechnical University, China; Ming Li, OPPO Mobile Telecommunications Corp., Ltd, China

IVMSP-1.6: FOV-BASED CODING OPTIMIZATION FOR 360-DEGREE VIRTUAL REALITY VIDEOS 1640
Yuanyuan Xu, Taoyu Yang, Zengjie Tan, Haolun Lan, Hohai University, China

IVMSP-2: MACHINE LEARNING FOR IMAGE PROCESSING I

IVMSP-2.1: MULTI-HIERARCHY PROXY STRUCTURE FOR DEEP METRIC LEARNING 1645
Jian Wang, Xinyue Li, Wei Song, Shanghai Ocean University, China; Zhichao Zhang, Shanghai Jiao Tong University, China; Weiqi Guo, East Sea Oceanographic Engineering Investigation, Design and Research Institute, China

IVMSP-2.2: EXPLOITING CAPTION DIVERSITY FOR UNSUPERVISED VIDEO SUMMARIZATION 1650
Michail Kaseris, Ioannis Mademlis, Ioannis Pitas, Aristotle University of Thessaloniki, Greece

IVMSP-2.3: CLUSTERING AND SEPARATING SIMILARITIES FOR DEEP UNSUPERVISED HASHING 1655
Wanqian Zhang, Dayan Wu, Bo Li, Weiping Wang, Institute of Information Engineering, Chinese Academy of Sciences, China; Chule Yang, National Innovation Institute of Defense Technology, Academy of Military Science, China

IVMSP-2.4: ENHANCING PROTOTYPICAL FEW-SHOT LEARNING BY LEVERAGING THE LOCAL-LEVEL STRATEGY 1660
Junying Huang, Fan Chen, Keze Wang, Liang Lin, Dongyu Zhang, Sun Yat-sen University, China

IVMSP-2.5: BLIND UNMIXING USING A DOUBLE DEEP IMAGE PRIOR 1665
Chao Zhou, Miguel Rodrigues, University College London, United Kingdom of Great Britain and Northern Ireland

IVMSP-2.6: A NEW FRAMEWORK FOR MULTIPLE DEEP CORRELATION FILTERS BASED OBJECT TRACKING 1670
Yi Liu, Qiangqiang Wu, Hanzi Wang, Xiamen University, China; Yanjie Liang, Xiamen University, Peng Cheng Laboratory, China; Liming Zhang, University of Macau, China

IVMSP-3: IMAGE PROCESSING

IVMSP-3.1: ADAPTIVE ACTOR-CRITIC BILATERAL FILTER 1675
Bo-Hao Chen, Hsiang-Yin Cheng, Yuan Ze University, Taiwan, Province of China; Jia-Li Yin, Fuzhou University, China

IVMSP-3.2: DOMAIN DECOMPOSITION ALGORITHMS FOR REAL-TIME HOMOGENEOUS DIFFUSION INPAINTING IN 4K 1680
Niklas Kämper, Joachim Weickert, Saarland University, Germany

IVMSP-3.3: DEEP TEMPORAL INTERPOLATION OF RADAR-BASED PRECIPITATION 1685
Michiaki Tatsubori, Takao Moriyama, Tatsuya Ishikawa, Paolo Fraccaro, Anne Jones, Blair Edwards, Julian Kuehnert, Sekou L. Remy, IBM Research, Japan

IVMSP-3.4: A NONLINEAR STEERABLE COMPLEX WAVELET DECOMPOSITION OF IMAGES 1690
Zikai Sun, Thierry Blu, The Chinese University of Hong Kong, Hong Kong

IVMSP-3.5: KERNEL ESTIMATION NETWORK FOR BLIND SUPER-RESOLUTION 1695
Xiang Cao, Haibo Shen, Liangqi Zhang, Yihao Luo, Tianjiang Wang, Huazhong University of Science and Technology, China

IVMSP-3.6: TERAHERTZ IMAGE RESTORATION BENCHMARKING DATASET..... 1700
Yixiong Zhang, Zhipeng Su, School of Informatics Xiamen University (National Demonstrative Software School), Xiamen 361005, China, China; Feng Qi, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China, China; Jianyang Zhou, School of Electronic Science and Engineering, Xiamen University, Xiamen, China, China; Xiao-Ping Zhang, Department of Electrical, Computer and Biomedical Engineering, Ryerson University, Canada, Canada

IVMSP-4: IMAGE AND VIDEO UNDERSTANDING

IVMSP-4.1: BINARY DENSE PREDICTORS FOR HUMAN POSE ESTIMATION BASED 1705
ON DYNAMIC THRESHOLDS AND FILTERING
Xingrun Xing, Yalong Jiang, Baochang Zhang, Wenrui Ding, Hongguang Li, Beihang University, China; Yangguang Li, SenseTime Group, China; Huan Peng, Huazhong University of Science and Technology, China

IVMSP-4.2: SELF-SUPERVISED LEARNING FOR SENTIMENT ANALYSIS VIA 1710
IMAGE-TEXT MATCHING
Haidong Zhu, Zhaoheng Zheng, Ram Nevatia, University of Southern California, United States of America; Mohammad Soleymani, USC-ICT, United States of America

IVMSP-4.3: DOMAIN-AGNOSTIC META-LEARNING FOR CROSS-DOMAIN FEW-SHOT 1715
CLASSIFICATION
Wei-Yu Lee, Jheng-Yu Wang, MOXA Inc., Taiwan; Yu-Chiang Wang, National Taiwan University, Taiwan

IVMSP-4.4: SEMANTIC ASSOCIATION NETWORK FOR VIDEO CORPUS MOMENT 1720
RETRIEVAL
Dahyun Kim, Sunjae Yoon, Ji Woo Hong, Chang D Yoo, KAIST, Korea, Republic of

IVMSP-4.5: STATISTICAL, SPECTRAL AND GRAPH REPRESENTATIONS FOR 1725
VIDEO-BASED FACIAL EXPRESSION RECOGNITION IN CHILDREN
Nida Itrat Abbasi, Siyang Song, Hatice Gunes, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

IVMSP-4.6: DERIVING EXPLAINABLE DISCRIMINATIVE ATTRIBUTES USING 1730
CONFUSION ABOUT COUNTERFACTUAL CLASS
Nakyeong Yang, Taegwan Kang, Kyomin Jung, Seoul National University, Korea, Republic of

IVMSP-5: IMAGE AND VIDEO SYNTHESIS

IVMSP-5.1: REALISTIC MONOCULAR-TO-3D VIRTUAL TRY-ON VIA MULTI-SCALE 1735
CHARACTERISTICS CAPTURE
Chenghu Du, Feng Yu, Minghua Jiang, Yaxin Zhao, Xiong Wei, Tao Peng, Xinrong Hu, Wuhan Textile University, China

IVMSP-5.2: OPTIMIZING LATENT SPACE DIRECTIONS FOR GAN-BASED LOCAL 1740
IMAGE EDITING
Ehsan Pajouheshgar, Tong Zhang, Sabine Süsstrunk, EPFL, Switzerland

IVMSP-5.3: TOWARDS USING CLOTHES STYLE TRANSFER FOR SCENARIO-AWARE 1745
PERSON VIDEO GENERATION
Jingning Xu, School of Software Engineering, Tongji University, Shanghai, China, China; Benlai Tang, Wenyi Guo, Xiang Yin, Zejun Ma, ByteDance AI Lab, China; Mingjie Wang, School of Computer Science, University of Guelph, ON, Canada, Canada; Siyuan Bian, Department of Computer Science and Engineering, Shanghai Jiao Tong University, China, China

IVMSP-5.4: MULTI-DOMAIN UNSUPERVISED IMAGE-TO-IMAGE TRANSLATION WITH APPEARANCE ADAPTIVE CONVOLUTION 1750
Somi Jeong, NAVER LABS, Korea, Republic of; Jiyoung Lee, NAVER AI Lab, Korea, Republic of; Kwanghoon Sohn, Yonsei University, Korea, Republic of

IVMSP-5.5: VR-FAM: VARIANCE-REDUCED ENCODER WITH NONLINEAR TRANSFORMATION FOR FACIAL ATTRIBUTE MANIPULATION 1755
Yifan Yuan, Siteng Ma, Junping Zhang, Fudan University, China

IVMSP-5.6: WAVELET-BASED UNSUPERVISED LABEL-TO-IMAGE TRANSLATION 1760
George Eskandar, Mohamed Abdelsamad, Karim Armanious, Shuai Zhang, Bin Yang, University of Stuttgart, Germany

IVMSP-6: VIDEO SUMMARIZATION AND ANOMALY DETECTION

IVMSP-6.1: FAST GRAPH SAMPLING FOR SHORT VIDEO SUMMARIZATION USING GERSHGORIN DISC ALIGNMENT 1765
Sadid Sahami, Chia-Wen Lin, National Tsing Hua University, Taiwan; Gene Cheung, York University, Canada

IVMSP-6.2: TOWARDS PRACTICAL AND EFFICIENT LONG VIDEO SUMMARY 1770
Xiaopeng Ke, Boyu Chang, Hao Wu, Fengyuan Xu, Sheng Zhong, Nanjing University, China

IVMSP-6.3: CUT AND CONTINUOUS PASTE TOWARDS REAL-TIME DEEP FALL DETECTION 1775
Sunhee Hwang, Minsong Ki, Seung-Hyun Lee, Sanghoon Park, Byoung-Ki Jeon, LG Uplus, Korea, Republic of

IVMSP-6.5: MANNET: A LARGE-SCALE MANIPULATED IMAGE DETECTION DATASET AND BASELINE EVALUATIONS 1780
Aditya Singh, Saheb Chhabra, Puspita Majumdar, IIIT-Delhi, India; Richa Singh, Mayank Vatsa, IIT Jodhpur, India

IVMSP-6.6: APPROACHES TOWARD PHYSICAL AND GENERAL VIDEO ANOMALY DETECTION 1785
Laura Kart, Niv Cohen, The Hebrew University of Jerusalem, Israel

IVMSP-7: PERCEPTION AND QUALITY MODELS

IVMSP-7.1: CONSIDERING USER AGREEMENT IN LEARNING TO PREDICT THE AESTHETIC QUALITY 1790
Suiyi Ling, Andreas Pastor, University of Nantes, France; Junle Wang, Tencent, China; Patrick Le Callet, Nantes Universite, Ecole Centrale Nantes, France

IVMSP-7.2: NO-REFERENCE QUALITY ASSESSMENT OF VARIABLE FRAME-RATE VIDEOS USING TEMPORAL BANDPASS STATISTICS 1795
Qi Zheng, Yibo Fan, Xiaoyang Zeng, Fudan University, China; Zhengzhong Tu, Alan Bovik, University of Texas at Austin, United States of America

IVMSP-7.3: TOWARDS JOINT FRAME-LEVEL AND MOS QUALITY PREDICTIONS WITH LOW-COMPLEXITY OBJECTIVE MODELS 1800
Joel Jung, Alexandre Giraud, Meijia Song, Songnan Li, Xiang Li, Shan Liu, Tencent Media Lab, United States of America

IVMSP-7.4: TEACHING CNNS TO MIMIC HUMAN VISUAL COGNITIVE PROCESS & REGULARISE TEXTURE-SHAPE BIAS 1805
Satyam Mohla, Anshul Nasery, Biplab Banerjee, IIT Bombay, India

IVMSP-7.5: SUBJECTIVE AND OBJECTIVE QUALITY ASSESSMENT OF MOBILE GAMING VIDEO 1810
Shaoguo Wen, Junle Wang, Ximing Chen, Yanqing Jing, Tencent, China; Suiyi Ling, University of Nantes, France; Patrick Le Callet, Nantes Université, Ecole Centrale Nantes, CNRS, LS2N, Nantes, France, France

IVMSP-7.6: ER-PIQA: A TASK-GUIDED PEDESTRIAN IMAGE QUALITY ASSESSMENT VIA EMBEDDING RECONSTRUCTION 1815
Yanzhe Zhong, Bangjie Tang, Zhonggeng Liu, Yiming Zhu, Jun Yin, Advanced Research Institute of Zhejiang Dahua Technology Co.,Ltd., China; Huadong Pan, Advanced Research Institute of Zhejiang Dahua Technology Co.,Ltd., Zhejiang Provincial Key Laboratory of Harmonized Application of Vision & Transmission, China

IVMSP-8: APPLICATIONS I

IVMSP-8.1: MULTISCALE CROWD COUNTING AND LOCALIZATION BY MULTITASK POINT SUPERVISION 1820
Mohsen Zand, Haleh Damirchi, Andrew Farley, Mahdiyar Molahasani Majdabadi, Michael Greenspan, Ali Etemad, Queeen's University, Canada

IVMSP-8.2: SUPER-RESOLUTION OF SATELLITE IMAGES BY TWO-DIMENSIONAL RRDB AND EDGE-ENHANCEMENT GENERATIVE ADVERSARIAL NETWORK 1825
Yu-Zhang Chen, Tsung-Jung Liu, National Chung Hsing University, Taiwan; Kuan-Hsien Liu, National Taichung University of Science and Technology, Taiwan

IVMSP-8.3: LEVERAGING LOCAL TEMPORAL INFORMATION FOR MULTIMODAL SCENE CLASSIFICATION 1830
Saurabh Sahu, Palash Goyal, Samsung Research America, United States of America

IVMSP-8.4: PREDICTING HUMAN MOTION USING KEY SUBSEQUENCES 1835
Menghao Li, Mingtao Pei, Wei Liang, Beijing Institute of Technology, China

IVMSP-8.5: DYNAMIC TEXTURE RECOGNITION USING PDV HASHING AND DICTIONARY LEARNING ON MULTI-SCALE VOLUME LOCAL BINARY PATTERN 1840
Ruxin Ding, Jianfeng Ren, Heng Yu, Jiawei Li, University of Nottingham Ningbo China, China

IVMSP-8.6: DO YOU LIVE A HEALTHY LIFE? ANALYZING LIFESTYLE BY VISUAL LIFE LOGGING 1845
Qing Gao, Mingtao Pei, Hongyu Shen, Beijing Institute of technology, China

IVMSP-9: IMAGE AND VIDEO CODING II

IVMSP-9.1: WEIGHTED WAVELET-BASED SPECTRAL-SPATIAL TRANSFORMS FOR CFA-SAMPLED RAW CAMERA IMAGE COMPRESSION CONSIDERING IMAGE FEATURES 1850
Liping Huang, Taizo Suzuki, University of Tsukuba, Japan

IVMSP-9.2: JMPNET: JOINT MOTION PREDICTION FOR LEARNING-BASED VIDEO COMPRESSION 1855
Dongyang Li, Zhenhong Sun, Zhiyu Tan, Xiuyu Sun, Fangyi Zhang, Yichen Qian, Hao Li, Alibaba Group, China, China

IVMSP-9.3: A LOW-PARAMETRIC MODEL FOR BIT-RATE ESTIMATION OF VVC RESIDUAL CODING 1860
Fabian Brand, Christian Herglotz, André Kaup, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

IVMSP-9.4: OPTE: ONLINE PER-TITLE ENCODING FOR LIVE VIDEO STREAMING 1865
Vignesh V Menon, Hadi Amirpour, Christian Timmerer, University of Klagenfurt, Austria; Mohammad Ghanbari, University of Essex, United Kingdom of Great Britain and Northern Ireland

IVMSP-9.5: SADN: LEARNED LIGHT FIELD IMAGE COMPRESSION WITH SPATIAL-ANGULAR DECORRELATION 1870
Kedeng Tong, Xin Jin, Chen Wang, Fan Jiang, Tsinghua University, China

IVMSP-9.6: HIERARCHICAL FEATURE AGGREGATION NETWORK FOR DEEP IMAGE COMPRESSION 1875
Wenfeng Li, Zongcai Du, Hao He, Jie Tang, Gangshan Wu, Nanjing University, China

IVMSP-10: MACHINE LEARNING FOR IMAGE PROCESSING II

IVMSP-10.1: ACCURATE INSTANCE SEGMENTATION VIA COLLABORATIVE LEARNING 1880
Tianyou Chen, Xiaoguang Hu, Jin Xiao, Guofeng Zhang, Shaojie Wang, Beihang University, China

IVMSP-10.2: DYNAMIC BINARY NEURAL NETWORK BY LEARNING CHANNEL-WISE THRESHOLDS 1885
Jiehua Zhang, Zhuo Su, Matti Pietikäinen, University of Oulu, China; Yanghe Feng, Xin Lu, National University of Defense Technology, China; Li Liu, 1. National University of Defense Technology; 2. University of Oulu, China

IVMSP-10.3: SELF-SUPERVISED LEARNING ON A LIGHTWEIGHT LOW-LIGHT IMAGE ENHANCEMENT MODEL WITH CURVE REFINEMENT 1890
Wanyu Wu, Wei Wang, Xin Xu, Wuhan University of Science and Technology, China; Kui Jiang, Ruimin Hu, Wuhan University, China

IVMSP-10.4: SEMANTICALLY PROPORTIONAL PATCHMIX FOR FEW-SHOT LEARNING 1895
Jingquan Wang, Jing Xu, Yu Pan, Harbin Institute of Technology (Shenzhen), China; Zenglin Xu, Harbin Institute of Technology (Shenzhen), PengCheng Laboratory, China

IVMSP-10.5: NOISE SUPPRESSION FOR IMPROVED FEW-SHOT LEARNING 1900
Zhikui Chen, Tiandong Ji, Suhua Zhang, Fangming Zhong, Dalian University of Technology, China

IVMSP-10.6: ONLINE CONTINUAL LEARNING USING ENHANCED RANDOM VECTOR FUNCTIONAL LINK NETWORKS 1905
Cheryl Sze Yin Wong, Yang Guo, ArulMurugan Ambikapathi, Savitha Ramasamy, Institute for Infocomm Research, Singapore

IVMSP-11: VIDEO PROCESSING

IVMSP-11.1: A GENERALIZED KERNEL RISK SENSITIVE LOSS FOR ROBUST TWO-DIMENSIONAL SINGULAR VALUE DECOMPOSITION 1910
Miaohua Zhang, Yongsheng Gao, Jun Zhou, Griffith University, Australia

IVMSP-11.2: VIDEO FRAME INTERPOLATION VIA LOCAL LIGHTWEIGHT BIDIRECTIONAL ENCODING WITH CHANNEL ATTENTION CASCADE 1915
Xiangling Ding, Hunan University of Science and Technology, China; Pu Huang, Dengyong Zhang, Changsha University of Science and Technology, China; Xianfeng Zhao, Chinese Academy of Sciences, China

IVMSP-11.3: SAIN: SIMILARITY-AWARE VIDEO FRAME INTERPOLATION 1920
Yue Lv, Wenming Yang, Qingmin Liao, Tsinghua University, China; Wangmeng Zuo, Harbin Institute of Technology, China; Rui Zhu, City, University of London, United Kingdom of Great Britain and Northern Ireland

IVMSP-11.4: SELF-LEARNED VIDEO SUPER-RESOLUTION WITH AUGMENTED SPATIAL AND TEMPORAL CONTEXT 1925
Zejia Fan, Jiaying Liu, Wenhan Yang, Zongming Guo, Wangxuan Institute of Computer Technology, Peking University, China; Wei Xiang, Bigo Technology PTE. LTD, China

IVMSP-11.5: DEFORMABLE CONVOLUTION DENSE NETWORK FOR COMPRESSED VIDEO QUALITY ENHANCEMENT 1930
Jiahui Liu, Mingcai Zhou, Meng Xiao, Alibaba Group, China

IVMSP-11.6: CONVOLUTIONAL ISTA NETWORK WITH TEMPORAL CONSISTENCY CONSTRAINTS FOR VIDEO RECONSTRUCTION FROM EVENT CAMERAS 1935
Siying Liu, Roxana Alexandru, Pier Luigi Dragotti, Imperial College London, United Kingdom of Great Britain and Northern Ireland

IVMSP-12: IMAGE AND VIDEO UNDERSTANDING II

IVMSP-12.1: PMP-NET: RETHINKING VISUAL CONTEXT FOR SCENE GRAPH GENERATION 1940
Xuezhi Tong, College of Intelligence and Computing, Tianjin University, China; Rui Wang, Zhejiang Lab and State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, China; Chuan Wang, Sanyi Zhang, Xiaochun Cao, State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, China

IVMSP-12.2: IMPROVE IMAGE CAPTIONING VIA RELATION MODELING 1945
Feicheng Huang, Zhixin Li, Guangxi Normal University, China

IVMSP-12.3: EQUAL LOSS: A SIMPLE LOSS FUNCTION FOR NOISE ROBUST LEARNING 1950
Lei Cui, Tsinghua University, China; Huan Peng, Huazhong University of Science and Technology, China; Yangguang Li, Chuming Li, Sensetime Group, China; Xingrun Xing, Beihang University, China

IVMSP-12.4: INFORMATIVE ATTENTION SUPERVISION FOR GROUNDED VIDEO DESCRIPTION 1955
Boyang Wan, Wenhui Jiang, Yuming Fang, Jiangxi University of Finance and Economics, China

IVMSP-12.5: SPATIAL-CONTEXT-AWARE DEEP NEURAL NETWORK FOR MULTI-CLASS IMAGE CLASSIFICATION 1960
Jialu Zhang, Qian Zhang, Jianfeng Ren, University of Nottingham Ningbo China, China; Yitian Zhao, Chinese Academy of Sciences, China; Jiang Liu, Southern University of Science and Technology, China

IVMSP-12.6: TRANSTL: SPATIAL-TEMPORAL LOCALIZATION TRANSFORMER FOR MULTI-LABEL VIDEO CLASSIFICATION 1965
Hongjun Wu, Mengzhu Li, Hongzhe Liu, Cheng Xu, Xuewei Li, Beijing Union University, China; Yongcheng Liu, Chinese Academy of Sciences, China

IVMSP-13: MULTI-MODAL ANALYSIS AND SYNTHESIS

IVMSP-13.1: DEEP VIDEO INPAINTING GUIDED BY AUDIO-VISUAL SELF-SUPERVISION 1970
Kyuyeon Kim, Junsik Jung, Woo Jae Kim, Sung-Eui Yoon, Korea Advanced Institute of Science and Technology, Korea, Republic of

IVMSP-13.2: NAVIGATING AUDIO-VISUAL EVENT DETECTION ACROSS MISMATCHED MODALITIES 1975
Guangwei Li, Xuenan Xu, Mengyue Wu, Kai Yu, Shanghai Jiao Tong University, China

IVMSP-13.3: LOOK, LISTEN AND PAY MORE ATTENTION: FUSING MULTI-MODAL INFORMATION FOR VIDEO VIOLENCE DETECTION 1980
Dong-Lai Wei, Yang Liu, Jing Liu, Xin-Hua Zeng, Fudan University, China; Chen-Geng Liu, The University of Melbourne, China; Xiao-Guang Zhu, Shanghai Jiao Tong University, China

IVMSP-13.4: MULTI-MODAL LEARNING WITH TEXT MERGING FOR TEXTVQA 1985
Changsheng Xu, Zhenlong Xu, Yifan He, Shuigeng Zhou, Fudan Univeristy, China; Jihong Guan, Tongji University, China

IVMSP-13.6: A NOVEL PART FEATURE INTEGRATION AND FUSION METHOD FOR FINE-GRAINED VEHICLE RECOGNITION 1990
Ping Wang, Yijie Cao, Lei Lu, Xi'an Jiaotong University, China

IVMSP-14: OBJECT DETECTION II

IVMSP-14.1: MONOCULAR VEHICLE 3D BOUNDING BOX ESTIMATION USING HOMOGRAHY AND GEOMETRY IN TRAFFIC SCENE 1995
Yiqiang Chen, Feng Liu, Ke Pei, Huawei Technology, China

IVMSP-14.2: FSM: FEATURE SAMPLING MODULE FOR OBJECT DETECTION 2000
Xin Yi, Bo Ma, JiaHao Wu, Beijing Institute of Technology, China

IVMSP-14.3: RETHINKING TWO-B-REAL NET FOR REAL-TIME SALIENT OBJECT DETECTION 2005
Senyun Kuang, Shijin Meng, Bo Xiao, Southwest Jiaotong University, China; Lv Tang, Nanjing University, China; Bo Li, Youtu Lab, Tencent, China; , ,

IVMSP-14.4: BALANCED RANKING AND SORTING FOR CLASS INCREMENTAL OBJECT DETECTION 2010
Bo Cui, Shan Yu, CASIA, China; Hui Qu, Xuhui Huang, X-Lab, the Second Academy of CASIC, China

IVMSP-14.5: MULTI-SCALE REINFORCEMENT LEARNING STRATEGY FOR OBJECT DETECTION 2015
Yihao Luo, Xiang Cao, Juntao Zhang, Leixilan Pan, Tianjiang Wang, Qi Feng, Huazhong University of Science and Technology, China

IVMSP-14.6: DEEP OBJECT DETECTION WITH EXAMPLE ATTRIBUTE BASED PREDICTION MODULATION 2020
Zhihao Wu, Chengliang Liu, Chao Huang, Jie Wen, Yong Xu, Harbin Institute of Technology, Shenzhen, China

IVMSP-15: IMAGE COMMUNICATION AND QUALITY MODELS

IVMSP-15.1: UNIVERSAL EFFICIENT VARIABLE-RATE NEURAL IMAGE COMPRESSION 2025
Shanzhi Yin, Chao Li, Youneng Bao, Yongsheng Liang, Harbin Institute of Technology, Shenzhen, China; Fanyang Meng, Wei Liu, Peng Cheng Laboratory, China

IVMSP-15.2: ADDERIC: TOWARDS LOW COMPUTATION COST IMAGE COMPRESSION 2030
Bowen Li, Chao Li, Youneng Bao, Yongsheng Liang, Harbin Institute of Technology, Shenzhen, China, China; Yao Xin, Fanyang Meng, Peng Cheng Laboratory, Shenzhen, China, China

IVMSP-15.3: DCNGAN: A DEFORMABLE CONVOLUTION-BASED GAN WITH QP ADAPTATION FOR PERCEPTUAL QUALITY ENHANCEMENT OF COMPRESSED VIDEO 2035
Saiping Zhang, Fuzheng Yang, Xidian University, China; Luis Herranz, Universitat Autònoma de Barcelona, Spain; Marta Mrak, Marc Gorriz Blanch, British Broadcasting Corporation, United Kingdom of Great Britain and Northern Ireland; Shuai Wan, Northwestern Polytechnical University, China

IVMSP-15.4: SPECIALISED VIDEO QUALITY MODEL FOR ENHANCED USER GENERATED CONTENT (UGC) WITH SPECIAL EFFECTS 2040
Anne-Flore Perrin, Yiejing Xie, Nantes Université, Ecole Centrale Nantes, CAPACITES SAS, CNRS, LS2N, UMR 6004, F-44000 Nantes, France, France; Tao Zhang, Yiting Liao, Junlin Li, ByteDance Inc., United States of America; Patrick Le Callet, Nantes Université, Ecole Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France

IVMSP-15.5: IMPROVING MAXIMUM LIKELIHOOD DIFFERENCE SCALING METHOD TO MEASURE INTER CONTENT SCALE 2045
Andréas Pastor, Patrick Le Callet, University of Nantes, France; Lukáš Krasula, Xiaoqing Zhu, Zhi Li, Netflix Inc., United States of America

IVMSP-15.6: TEXTURE INFORMATION BOOSTS VIDEO QUALITY ASSESSMENT 2050
Ao-Xiang Zhang, Yuan-Gen Wang, Guangzhou University, China

IVMSP-16: APPLICATIONS II

IVMSP-16.1: PLUG-AND-PLAY AND RELAY REGULARIZATIONS ON NOISY LOW RANK TENSOR COMPLETION FOR SNAPSHOT MULTISPECTRAL IMAGE RESTORATION 2055
Keisuke Ozawa, DENSO IT Laboratory, Japan

IVMSP-16.2: LERPS: LIGHTING ESTIMATION AND RELIGHTING FOR PHOTOMETRIC STEREO 2060
Ashish Tiwari, Shanmuganathan Raman, IIT Gandhinagar, India

IVMSP-16.3: A UNIFIED TWO-STAGE MODEL FOR SEPARATING SUPERIMPOSED IMAGES 2065
Huiyu Duan, Xiongkuo Min, Wei Shen, Guangtao Zhai, Shanghai Jiao Tong University, China

IVMSP-16.4: PARAMETER-FREE STYLE PROJECTION FOR ARBITRARY IMAGE STYLE TRANSFER 2070
Siyu Huang, Harvard University, China; Haoyi Xiong, Qingzhong Wang, Zeyu Chen, Dejing Dou, Baidu Research, China; Tianyang Wang, Austin Peay State University, United States of America; Bihan Wen, Nanyang Technological University, Singapore; Jun Huan, Amazon, China

IVMSP-16.5: OPTIMIZATION OF COMPRESSIVE LIGHT FIELD DISPLAY IN DUAL-GUIDED LEARNING 2075
Yangfan Sun, Zhu Li, UMKC, United States of America; Li Li, University of Science and Technology of China, United States of America; Shizheng Wang, Chinese Academy of Sciences, United States of America; Wei Gao, Peking University, United States of America

IVMSP-16.6: ARM 4-BIT PQ: SIMD-BASED ACCELERATION FOR APPROXIMATE NEAREST NEIGHBOR SEARCH ON ARM 2080
Yusuke Matsui, The University of Tokyo, Japan; Yoshiki Imaizumi, Naoya Miyamoto, Naoki Yoshifuji, Fixstars Corporation, Japan

IVMSP-17: ENHANCEMENT AND RESTORATION I

IVMSP-17.1: ITERATIVE LEARNING FOR DISTORTED IMAGE RESTORATION 2085
Chao Wang, Yi Gu, Jie Li, Xinlei He, Zirui Zhang, Chentao Wu, Shanghai Jiao Tong University, China; Yuting Gao, Texas A&M University, United States of America

IVMSP-17.2: JE2NET: JOINT EXPLOITATION AND EXPLORATION IN REINFORCEMENT LEARNING BASED IMAGE RESTORATION 2090
Xiaoyu Zhang, Wei Gao, Ge Li, School of Electronic and Computer Engineering, Peking University, Shenzhen, China. Peng Cheng Laboratory, Shenzhen, China., China; Hui Yuan, School of Control Science and Engineering, Shandong University, Jinan, China, China

IVMSP-17.3: MULTIPLE PATCH-AWARE NETWORK FOR FASTER REAL-WORLD IMAGE DEHAZING 2095
Kun Yang, Juan Zhang, Xiaoqi Lang, Shanghai University of Engineering Science, China

IVMSP-17.4: LEARNING TO FUSE HETEROGENEOUS FEATURES FOR LOW-LIGHT IMAGE ENHANCEMENT 2100
Zhenyu Tang, Long Ma, Xiaoke Shang, Xin Fan, Dalian University of Technology, China

IVMSP-17.5: DEEP SCALE-AWARE IMAGE SMOOTHING 2105
Jiachun Li, Kunkun Qin, Ruotao Xu, South China University of Technology, China; Hui Ji, National University of Singapore, Singapore

IVMSP-17.6: A MULTISCALE GRADIENT-BACKPROPAGATION OPTIMIZATION FRAMEWORK FOR DEFORMABLE CONVOLUTION BASED COMPRESSED VIDEO ENHANCEMENT 2110
Yanbo Gao, Shuai Li, Xun Cai, Shandong University, China; Menghu Jia, Mao Ye, University of Electronic Science and Technology of China, China; Frédéric Dufaux, Université Paris-Saclay, France

IVMSP-18: MACHINE LEARNING FOR IMAGE PROCESSING III

IVMSP-18.1: DOWNSTREAM AUGMENTATION GENERATION FOR CONTRASTIVE LEARNING 2115
Tomohiro Hayase, Suguru Yasutomi, Fujitsu, Ltd., Japan; Nakamasa Inoue, Tokyo Institute of Technology, Japan

IVMSP-18.2: FEW-SHOT LEARNING WITH IMPROVED LOCAL REPRESENTATIONS VIA BIAS RECTIFY MODULE 2120
Chao Dong, Qi Ye, Wenchao Meng, Kaixiang Yang, Zhejiang University, China

IVMSP-18.3: IMAGE-TO-VIDEO RE-IDENTIFICATION VIA MUTUAL DISCRIMINATIVE KNOWLEDGE TRANSFER 2125
Pichao Wang, Fan Wang, Hao Li, Alibaba Group, United States of America

IVMSP-18.4: DYNSNN: A DYNAMIC APPROACH TO REDUCE REDUNDANCY IN SPIKING NEURAL NETWORKS 2130
Fangxin Liu, Wenbo Zhao, Yongbiao Chen, Zongwu Wang, Shanghai Jiaotong University, China; Dai Fei, Academy of System Engineering, Academy of Military Sciences, China

IVMSP-18.5: MEJIGCLU: MORE EFFECTIVE JIGSAW CLUSTERING FOR UNSUPERVISED VISUAL REPRESENTATION LEARNING 2135
Yongsheng Zhang, Qing Liu, Yang Zhao, Yixiong Liang, Central South University, China

IVMSP-18.6: GANET: UNARY ATTENTION REACHES PAIRWISE ATTENTION VIA IMPLICIT GROUP CLUSTERING IN LIGHT-WEIGHT CNNS 2140
Cheng Zhuang, Yunlian Sun, Nanjing University of Science and Technology, China

IVMSP-19: SUPER-RESOLUTION

IVMSP-19.1: FIND THE WAY BACK: INVERTIBLE KERNEL ESTIMATOR FOR BLIND IMAGE SUPER-RESOLUTION 2145
Ting-Wei Chang, Wei-Chen Chiu, Ching-Chun Huang, National Yang Ming Chiao Tung University, Taiwan, Province of China

IVMSP-19.3: FINE-GRAINED DYNAMIC LOSS FOR ACCURATE SINGLE-IMAGE SUPER-RESOLUTION 2150
Haoquan Wang, Gang Zhang, Tianjin university, China; Zhichun Lei, University of Applied Science Ruhr West (HRW) in Germany, Germany

IVMSP-19.4: MULTI-FRAME SUPER-RESOLUTION WITH RAW IMAGES VIA MODIFIED DEFORMABLE CONVOLUTION 2155
Gongzhe Li, Linwei Qiu, Haopeng Zhang, Fengying Xie, Zhiguo Jiang, Beihang University, China

IVMSP-19.5: LOCAL-GLOBAL FEATURE AGGREGATION FOR LIGHT FIELD IMAGE SUPER-RESOLUTION 2160
Yan Wang, Yao Lu, Shunzhou Wang, Wenyao Zhang, Zijian Wang, Beijing Institute of Technology, China

IVMSP-19.6: PYRAMID FUSION ATTENTION NETWORK FOR SINGLE IMAGE SUPER-RESOLUTION 2165
Hao He, Zongcai Du, Wenfeng Li, Jie Tang, Gangshan Wu, Nanjing University, China

IVMSP-20: ACTION RECOGNITION

IVMSP-20.1: VCD: VIEW-CONSTRAINT DISENTANGLEMENT FOR ACTION RECOGNITION 2170
Xian Zhong, Zhuo Zhou, Wenxuan Liu, Xuemei Jia, Wuhan University of Technology, China; Kui Jiang, Zheng Wang, Wuhan University, China; Wenxin Huang, Hubei University, China

IVMSP-20.2: PRIVACY-PRESERVING ACTION RECOGNITION 2175
Chengming Zou, Hubei Key Laboratory of Transportation Internet of Things, Wuhan University of Technology, Wuhan, 430070, China, China; Ducheng Yuan, School of Computer Science and Technology, Wuhan University of Technology, Wuhan, 430070, China, China; Long Lan, Haoang Chi, Institute for Quantum Information & State Key Laboratory of High Performance Computing,College of Computer, National University of Defense Technology, Changsha 410073, China, China

IVMSP-20.3: SPATIO-TEMPORAL MOTION AGGREGATION NETWORK FOR VIDEO ACTION DETECTION 2180
Hongcheng Zhang, Xu Zhao, Shanghai Jiao Tong University, China

IVMSP-20.4: TP-VIT: A TWO-PATHWAY VISION TRANSFORMER FOR VIDEO ACTION RECOGNITION 2185
Yanhao Jing, Feng Wang, East China Normal University, China

IVMSP-20.5: LEARNING TASK-SPECIFIC REPRESENTATION FOR VIDEO ANOMALY DETECTION WITH SPATIAL-TEMPORAL ATTENTION 2190
Yang Liu, Jing Liu, Donglai Wei, Xiaohong Huang, Liang Song, Fudan University, China; Xiaoguang Zhu, Shanghai Jiao Tong University, China

IVMSP-20.6: W-ART: ACTION RELATION TRANSFORMER FOR WEAKLY-SUPERVISED TEMPORAL ACTION LOCALIZATION 2195
Mengzhu Li, Hongjun Wu, Hongzhe Liu, Cheng Xu, Xuewei Li, Beijing Union University, China; Yongcheng Liu, Chinese Academy of Sciences, China

IVMSP-21: ATTENTION FOR IMAGE AND VIDEO PROCESSING

IVMSP-21.1: MS-ROCANET: MULTI-SCALE RESIDUAL ORTHOGONAL-CHANNEL ATTENTION NETWORK FOR SCENE TEXT DETECTION 2200
Jinpeng Liu, Song Wu, DeHong He, Guoqiang Xiao, SouthWest University, China

IVMSP-21.2: BI-DIRECTIONAL NORMALIZATION AND COLOR ATTENTION-GUIDED GENERATIVE ADVERSARIAL NETWORK FOR IMAGE ENHANCEMENT 2205
Shan Liu, Guoqiang Xiao, Xiaohui Xu, Song Wu, Southwest University, China

IVMSP-21.3: DUAL-ATTENTION NETWORK FOR FEW-SHOT SEGMENTATION 2210
Zhikui Chen, Han Wang, Suhua Zhang, Fangming Zhong, Dalian University of Technology, China

IVMSP-21.4: ATTENTION GUIDED INVARIANCE SELECTION FOR LOCAL FEATURE DESCRIPTORS 2215
Jiapeng Li, Ge Li, Peking University, China; Thomas H Li, Advanced Institute of Information Technology, Peking University, China

IVMSP-21.5: ATTENTION PROBE: VISION TRANSFORMER DISTILLATION IN THE WILD 2220
Jiahao Wang, Mingdeng Cao, Shuwei Shi, Yujiu Yang, Tsinghua Shenzhen International Graduate School, China; Baoyuan Wu, The Chinese University of Hong Kong, Shenzhen, China

IVMSP-21.6: STACKED MULTI-SCALE ATTENTION NETWORK FOR IMAGE COLORIZATION 2225
Bin Jiang, Fangqiang Xu, Jun Xia, Chao Yang, Wei Huang, Yun Huang, Hunan university, China

IVMSP-22: OBJECT DETECTION I

IVMSP-22.1: CRPN: DISTINGUISH NOVEL CATEGORIES VIA CLASS-RELEVANT REGION PROPOSAL NETWORK FOR FEW-SHOT OBJECT DETECTION 2230
Han Wang, Yali Li, Shengjin Wang, Tsinghua University, China

IVMSP-22.2: AN EFFICIENT FRAMEWORK FOR DETECTION AND RECOGNITION OF NUMERICAL TRAFFIC SIGNS 2235
Zhishan Li, Lei Xie, Hongye Su, Zhejiang University, China; Mingmu Chen, Yifan He, Reconova Technologies Co., Ltd, China

IVMSP-22.3: DIVERGENCE-GUIDED FEATURE ALIGNMENT FOR CROSS-DOMAIN OBJECT DETECTION 2240
Zongyao Li, Ren Togo, Takahiro Ogawa, Miki Haseyama, Hokkaido University, Japan

IVMSP-22.4: PGTRNET: TWO-PHASE WEAKLY SUPERVISED OBJECT DETECTION WITH PSEUDO GROUND TRUTH REFINEMENT 2245
Jun Wang, University of Warwick, United Kingdom of Great Britain and Northern Ireland; Hefeng Zhou, Shanghai Jiao Tong University, China; Xiaohan Yu, School of Engineering and Built Environment, Australia

IVMSP-22.5: NOVEL INSTANCE MINING WITH PSEUDO-MARGIN EVALUATION FOR FEW-SHOT OBJECT DETECTION 2250
Weijie Liu, Chong Wang, Shenghao Yu, Chenchen Tao, Ningbo University, China; Jun Wang, China University of Mining and Technology, China; Jiafei Wu, SenseTime Research, China

IVMSP-22.6: BIP-NET: BIDIRECTIONAL PERSPECTIVE STRATEGY BASED ARBITRARY-SHAPED TEXT DETECTION NETWORK 2255
Chuang Yang, Mulin Chen, Yuan Yuan, Qi Wang, School of Computer Science, and School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an 710072, Shaanxi, P. R. China, China

IVMSP-23: DEPTH ESTIMATION & POINT CLOUD PROCESSING

IVMSP-23.1: A NOVEL LIGHTWEIGHT NETWORK FOR FAST MONOCULAR DEPTH ESTIMATION 2260
Tim Heydrich, Yimin Yang, Lakehead University, Canada; Xiangyu Ma, Yu Liu, Tianjin University, China; Shan Du, The University of British Columbia, Canada

IVMSP-23.2: A LIGHTWEIGHT SELF-SUPERVISED TRAINING FRAMEWORK FOR MONOCULAR DEPTH ESTIMATION 2265
Tim Heydrich, Yimin Yang, Lakehead University, Canada; Shan Du, The University of British Columbia, Canada

IVMSP-23.3: PU-REFINER: A GEOMETRY REFINER WITH ADVERSARIAL LEARNING FOR POINT CLOUD UPSAMPLING 2270
Hao Liu, Hui Yuan, Shuai Li, Shandong University, China; Hamzaoui Raouf, De Montfort University, United Kingdom of Great Britain and Northern Ireland; Wei Gao, Peking University, China

IVMSP-23.4: CF-NET: COMPLEMENTARY FUSION NETWORK FOR ROTATION INVARIANT POINT CLOUD COMPLETION 2275
Bo-Fan Chen, Yang-Ming Yeh, Yi-Chang Lu, National Taiwan University, Taiwan

IVMSP-23.5: TH-NET: A METHOD OF SINGLE 3D OBJECT TRACKING BASED ON TRANSFORMERS AND HAUSDORFF DISTANCE 2280
Zihao Zhang, Nan Sang, Xupeng Wang, University of Electronic Science and Technology of China, China

IVMSP-23.6: ENRICH FEATURES FOR FEW-SHOT POINT CLOUD CLASSIFICATION 2285
Hengxin Feng, Weifeng Liu, Yanjiang Wang, Baodi Liu, China University of Petroleum (East China), China

IVMSP-24: APPLICATIONS III

IVMSP-24.1: SEMI-SUPERVISED 360° DEPTH ESTIMATION FROM MULTIPLE FISHEYE CAMERAS WITH PIXEL-LEVEL SELECTIVE LOSS 2290
Jaewoo Lee, Daeul Park, Dongwook Lee, Daehyun Ji, Samsung Advanced Institute of Technology, Korea, Republic of

IVMSP-24.2: UNDERWATER STEREO MATCHING VIA UNSUPERVISED APPEARANCE AND FEATURE ADAPTATION NETWORKS 2295
Wei Zhong, Yazhi Yuan, Xinchen Ye, Dian Zheng, Rui Xu, Dalian University of Technology, China

IVMSP-24.3: DOMAIN ADAPTATION VIA MUTUAL INFORMATION MAXIMIZATION FOR HANDWRITING RECOGNITION 2300
Pei Tang, Liangrui Peng, Ruijie Yan, Haodong Shi, Gang Yao, Changsong Liu, Tsinghua University, China; Jie Li, Yuqi Zhang, Shanghai Pudong Development Bank, China

IVMSP-24.4: ATTRIBUTE-CONDITIONED FACE SWAPPING NETWORK FOR LOW-RESOLUTION IMAGES 2305
Ang Li, Chilin Fu, Xiaolu Zhang, Jun Zhou, Ant Group, China; Jian Hu, Queen mary university of London, China

IVMSP-24.5: LEARNING MULTIPLE EXPLAINABLE AND GENERALIZABLE CUES FOR FACE ANTI-SPOOFING 2310
Ying Bian, Peng Zhang, Jingjing Wang, Chunmao Wang, Shiliang Pu, Hikvision Research Institute, China, China

IVMSP-24.6: OFF-THE-GRID COVARIANCE-BASED SUPER-RESOLUTION FLUCTUATION MICROSCOPY 2315
Bastien Laville, Laure Blanc-Féraud, Gilles Aubert, Université Côte d'Azur, France

IVMSP-25: ENHANCEMENT AND RESTORATION II

IVMSP-25.1: SIMULTANEOUS NONLOCAL LOW-RANK AND DEEP PRIORS FOR POISSON DENOISING 2320
Zhiyuan Zha, Bihan Wen, Nanyang Technological University, Singapore; Xin Yuan, Nokia Bell Labs, United States of America; Jiantao Zhou, University of Macau, China; Ce Zhu, University of Electronic Science and Technology of China, China

IVMSP-25.2: DOUBLE CLOSED-LOOP NETWORK FOR IMAGE DEBLURRING 2325
Yiming Liu, Hunan Normal University, China; Yanni Zhang, Qiang Li, Jun Kong, Jianzhong Wang, Northeast Normal University, China; Miao Qi, Changchun Humanities and Sciences College, China

IVMSP-25.3: SINGLE IMAGE DE-RAINING WITH HIGH-LOW FREQUENCY GUIDANCE 2330
Ying Zhang, Youjun Xiang, Lei Cai, Yuli Fu, Wanliang Huo, Junjun Xia, South China University of Technology, China

IVMSP-25.4: DETAIL GENERATION AND FUSION NETWORKS FOR IMAGE INPAINTING 2335
Wu Yang, Wuzhen Shi, Shenzhen University, China

IVMSP-25.6: ADAPTIVE WEIGHTED NETWORK WITH EDGE ENHANCEMENT MODULE FOR MONOCULAR SELF-SUPERVISED DEPTH ESTIMATION 2340
Hong Liu, Ying Zhu, Guoliang Hua, Weibo Huang, Runwei Ding, Key Laboratory of Machine Perception, Shenzhen Graduate School, Peking University, China

IVMSP-25.7: PAS-MEF: MULTI-EXPOSURE IMAGE FUSION BASED ON PRINCIPAL COMPONENT ANALYSIS, ADAPTIVE WELL-EXPOSEDNESS AND SALIENCY MAP 2345
Diclehan Karakaya, Oguzhan Ulucan, Mehmet Turkan, Izmir University of Economics, Turkey

IVMSP-26: MACHINE LEARNING FOR IMAGE PROCESSING IV

IVMSP-26.1: PDD-NET: A PRECISE DEFECT DETECTION NETWORK BASED ON POINT SET REPRESENTATION 2350
Miaoju Ban, Runwei Ding, Jian Zhang, Tianyu Guo, Tao Wang, Key Laboratory of Machine Perception, Peking University, Shenzhen Graduate School, China

IVMSP-26.2: SOLVING THE LONG-TAILED PROBLEM VIA INTRA- AND INTER-CATEGORY BALANCE 2355
Renhui Zhang, Tiancheng Lin, Rui Zhang, Yi Xu, Shanghai Jiao Tong University, China

IVMSP-26.3: EXTRACTING AND DISTILLING DIRECTION-ADAPTIVE KNOWLEDGE FOR LIGHTWEIGHT OBJECT DETECTION IN REMOTE SENSING IMAGES 2360
Zhanchao Huang, Wei Li, Ran Tao, Beijing Institute of Technology, China

IVMSP-26.4: PSEUDO-INTERACTING GUIDED NETWORK FOR FEW-SHOT SEGMENTATION 2365
Xiaoliu Luo, Zhao Duan, Jin Tan, Taiping Zhang, Chongqing University, China; Jing Luo, Zhengzhou University, China

IVMSP-26.5: FEW-SHOT GENERATION BY MODELING STEREOSCOPIC PRIORS 2370
Yuehui Wang, Qing Wang, Dongyu Zhang, Sun-yat sen university, China

IVMSP-26.6: RELATIVE VIEWPOINT ESTIMATION BASED ON STRUCTURED 3D REPRESENTATION ALIGNMENT 2375
Kohei Matsuzaki, Kei Kawamura, KDDI Research, Inc., Japan

IVMSP-27: SEGMENTATION

IVMSP-27.1: DEEP MARKOV CLUSTERING FOR PANOPTIC SEGMENTATION 2380
Minxiang Ye, YiFei Zhang, Shiqiang Zhu, Anhuan Xie, Dan Zhang, Zhejiang Lab, China

IVMSP-27.2: MULTI-TASK LEARNING IMPROVES THE BRAIN STOKE LESION SEGMENTATION 2385
Libo Liu, Qingmao Hu, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China; School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China, China; Chengjian Huang, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China; University of Chinese Academy of Sciences, Beijing, China, China; Chunsheng Cai, Department of Cerebrovascular Disease, Huizhou First People's Hospital, Huizhou, China, China; Xiaodong Zhang, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, China

IVMSP-27.3: MIXED TRANSFORMER U-NET FOR MEDICAL IMAGE SEGMENTATION 2390
Hongyi Wang, Shiao Xie, Lanfen Lin, Ruofeng Tong, Zhejiang University, China; Yutaro Iwamoto, Yen-Wei Chen, Ritsumeikan University, Japan; Xian-Hua Han, Yamaguchi University, Japan

IVMSP-27.4: CONTRASTIVE TRANSLATION LEARNING FOR MEDICAL IMAGE SEGMENTATION 2395
Wankang Zeng, Wenkang Fan, Dongfang Shen, Yinran Chen, Xiongbiao Luo, Xiamen University, China

IVMSP-27.5: FAST VIDEO OBJECT SEGMENTATION VIA DYNAMIC YOLACT 2400
Tianfang Meng, Wenqiang Zhang, Fudan University, China

IVMSP-27.6: DEPTH REMOVAL DISTILLATION FOR RGB-D SEMANTIC SEGMENTATION 2405
Tiyu Fang, Zhen Liang, Zihao Dong, Jinping Li, University of Jinan, China; Xiuli Shao, Nankai University, China

IVMSP-28: FACES

IVMSP-28.1: MASK-BASED ATTENTION PARALLEL NETWORK FOR IN-THE-WILD FACIAL EXPRESSION RECOGNITION **2410**
Lingzhao Ju, Xu Zhao, Shanghai Jiao Tong University, Department of Automation, Shanghai, China, China

IVMSP-28.2: SDNET: LIGHTWEIGHT FACIAL EXPRESSION RECOGNITION FOR SAMPLE DISEQUILIBRIUM **2415**
Lifang Zhou, Siqin Li, Yi Wang, Junlin Liu, Chongqing University of Posts and Telecommunications, China

IVMSP-28.3: A NOVEL MICRO-EXPRESSION RECOGNITION APPROACH USING ATTENTION-BASED MAGNIFICATION-ADAPTIVE NETWORKS **2420**
Wei Mengting, Zheng Wenming, Zong Yuan, Jiang Xingxun, Lu Cheng, Liu Jiateng, Southeast University, China

IVMSP-28.4: LIPREADING MODEL BASED ON WHOLE-PART COLLABORATIVE LEARNING **2425**
Weidong Tian, Housen Zhang, Chen Peng, Zhong-Qiu Zhao, Hefei University of Technology, China

IVMSP-28.5: WHAT IS THE PATIENT LOOKING AT? ROBUST GAZE-SCENE INTERSECTION UNDER FREE-VIEWING CONDITIONS **2430**
Ahmed Al-Hindawi, Marcela Vizcaychipi, Yiannis Demiris, Imperial College London, United Kingdom of Great Britain and Northern Ireland

IVMSP-28.6: GAZEATTENTIONNET: GAZE ESTIMATION WITH ATTENTIONS **2435**
Haoxian Huang, Luqian Ren, Zhuo Yang, Yinwei Zhan, Guangdong University of Technology, China; Qieshi Zhang, Chinese Academy of Sciences, China; Jujian Lv, Guangdong Polytechnic Normal University, China

IVMSP-29: INVERSE PROBLEMS

IVMSP-29.1: LOW-LIGHT IMAGE ENHANCEMENT VIA FEATURE RESTORATION **2440**
Yang Yang, Xiaojie Guo, Tianjin university, China; Yonghua Zhang, Henan University, China

IVMSP-29.2: HIRL: HYBRID IMAGE RESTORATION BASED ON HIERARCHICAL DEEP REINFORCEMENT LEARNING VIA TWO-STEP ANALYSIS **2445**
Xiaoyu Zhang, Wei Gao, School of Electronic and Computer Engineering, Peking University, Shenzhen, China. Peng Cheng Laboratory, Shenzhen, China., China; , ,

IVMSP-29.3: HIGH-FIDELITY PORTRAIT EDITING VIA EXPLORING DIFFERENTIABLE GUIDED SKETCHES FROM THE LATENT SPACE **2450**
Chengrong Wang, Chenjie Cao, Yanwei Fu, Xiangyang Xue, Fudan University, China

IVMSP-29.4: LEARNING ADJUSTABLE IMAGE RESCALING WITH JOINT OPTIMIZATION OF PERCEPTION AND DISTORTION **2455**
Zhihong Pan, Baidu Research (USA), United States of America

IVMSP-29.5: FSOINET: FEATURE-SPACE OPTIMIZATION-INSPIRED NETWORK FOR IMAGE COMPRESSIVE SENSING **2460**
Wenjun Chen, Chunling Yang, Xin Yang, South China University of Technology, China

IVMSP-29.6: DISENTANGLED FEATURE-GUIDED MULTI-EXPOSURE HIGH DYNAMIC RANGE IMAGING **2465**
Keuntek Lee, Yeong Il Jang, Nam Ik Cho, Seoul National University, Korea, Republic of

IVMSP-30: SECURITY AND IMAGE INTERPRETATION

IVMSP-30.1: DEFENDING AGAINST UNIVERSAL ATTACK VIA CURVATURE-AWARE CATEGORY ADVERSARIAL TRAINING **2470**
Peilun Du, Xiaolong Zheng, Liang Liu, Huadong Ma, Beijing University of Posts and Telecommunications, China

IVMSP-30.2: SP ATTACK: SINGLE-PERSPECTIVE ATTACK FOR GENERATING ADVERSARIAL OMNIDIRECTIONAL IMAGES 2475
Yunjian Zhang, Yanwei Liu, Pengwei Zhan, Liming Wang, Zhen Xu, Institute of Information Engineering, Chinese Academy of Sciences, China; Jinxia Liu, Zhejiang Wanli University, China

IVMSP-30.3: FEW-SHOT ONE-CLASS DOMAIN ADAPTATION BASED ON FREQUENCY FOR IRIS PRESENTATION ATTACK DETECTION 2480
Yachun Li, Ying Lian, Jingjing Wang, Yuhui Chen, Chunmao Wang, Shiliang Pu, Hikvision Research Institute, China

IVMSP-30.4: PIXINWAV: RESIDUAL STEGANOGRAPHY FOR HIDING PIXELS IN AUDIO 2485
Margarita Geleta, University of California, Irvine, United States of America; Cristina Punti, Kevin McGuinness, Dublin City University, Ireland; Jordi Pons, Dolby Laboratories, Spain; Cristian Canton, Xavier Giro-i-Nieto, Universitat Politecnica de Catalunya, Spain

IVMSP-30.5: A SEMI-HANDCRAFTED KEYPOINT DETECTOR WITH DISCRIMINATIVE FEATURE ENCODING 2490
Yurui Xie, Ling Guan, Ryerson University, Canada

IVMSP-30.6: SAFARI FROM VISUAL SIGNALS: RECOVERING VOLUMETRIC 3D SHAPES 2495
Antonio Agudo, Institut de Robotica i Informatica Industrial, CSIC-UPC, Spain

IVMSP-31: IMAGE AND VIDEO REPRESENTATION

IVMSP-31.1: COUPLED FEATURE LEARNING VIA STRUCTURED CONVOLUTIONAL SPARSE CODING FOR MULTIMODAL IMAGE FUSION 2500
Farshad G. Veshki, Sergiy A. Vorobyov, Aalto university, Finland

IVMSP-31.2: DOMAINDESC: LEARNING LOCAL DESCRIPTORS WITH DOMAIN ADAPTATION 2505
Rongtao Xu, Weiliang Meng, Xiaopeng Zhang, Institute of Automation, Chinese Academy of Sciences, Beijing, China, China; Changwei Wang, Institute of Automation□Chinese Academy of Sciences, China; Bin Fan, University of Science and Technology, China; Yuyang Zhang, Institute of Automation, Chinese Academy of Sciences, China; Shibiao Xu, Beijing University of Posts and Telecommunications, China

IVMSP-31.3: MULTI-HEAD RELU IMPLICIT NEURAL REPRESENTATION NETWORKS 2510
Arya Aftab, Sharif University of Technology, Iran (Islamic Republic of); Alireza Morsali, McGill University, Canada; Shahrokh Ghaemmaghami, Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran, Iran (Islamic Republic of)

IVMSP-31.4: AN EFFICIENT METHOD FOR MODEL PRUNING USING KNOWLEDGE DISTILLATION WITH FEW SAMPLES 2515
Zhaojing Zhou, Zhuqing Jiang, Aidong Men, Haiying Wang, Beijing University of Posts and Telecommunications, China; Yun Zhou, National Radio and Television Administration, China

IVMSP-31.5: ADAPTIVE INTRA-GROUP AGGREGATION FOR CO-SALIENCY DETECTION 2520
Guangyu Ren, Tianhong Dai, Tania Stathaki, Imperial College London, United Kingdom of Great Britain and Northern Ireland

IVMSP-31.6: NOVEL CLASS DISCOVERY: A DEPENDENCY APPROACH 2525
Tanmoy Mukherjee, Nikos Deligiannis, VUB/IMEC, Belgium

IVMSP-32: APPLICATIONS IV

IVMSP-32.1: SINGLE-SHOT BALANCED DETECTOR FOR GEOSPATIAL OBJECT DETECTION 2529
Yanfeng Liu, Qiang Li, Yuan Yuan, Qi Wang, Northwestern Polytechnical University,, China

IVMSP-32.2: REGULARIZED LATENT SPACE EXPLORATION FOR DISCRIMINATIVE FACE SUPER-RESOLUTION 2534
Ruixin Shi, Junzheng Zhang, Shiming Ge, Institute of Information Engineering, Chinese Academy of Sciences; School of Cyber Security, University of Chinese Academy of Sciences, China; Yong Li, Institute of Information Engineering, Chinese Academy of Sciences, China

IVMSP-32.3: ENHANCING AND DISSECTING CROWD COUNTING BY SYNTHETIC DATA 2539
Yi Hou, Yuheng Lu, Yuan Li, Huizhu Jia, Xiaodong Xie, National Engineering Laboratory for Video Technology, Peking University, China; Chengyang Li, School of Electronics Engineering and Computer Science, Peking University, Beijing, China, China; Liping Zhu, Key Lab of Petroleum Data Mining, China University of Petroleum (Beijing), China

IVMSP-32.4: MULTI-POSE VIRTUAL TRY-ON VIA SELF-ADAPTIVE FEATURE FILTERING 2544
Chenghu Du, Feng Yu, Minghua Jiang, Xiong Wei, Tao Peng, Xinrong Hu, Wuhan Textile University, China

IVMSP-32.5: HISTOGRAM-GUIDED SEMANTIC-AWARE COLORIZATION 2549
Jie Zhang, Yi Xiao, Guo Chen, Qingping Sun, Fangqiang Xu, Hunan University, China; Chi-Sing Leung, City University of Hong Kong, Hong Kong

IVMSP-32.6: CONTENT PRESERVING SCALE SPACE NETWORK FOR FAST IMAGE RESTORATION FROM NOISY-BLURRY PAIRS 2554
Green Rosh K S, Nikhil Krishnan, B H Pawan Prasad, Sachin Lomte, Samsung R&D Institute India, Bangalore, India

IVMSP-33: POINT CLOUD PROCESSING

IVMSP-33.1: FLOW-BASED POINT CLOUD COMPLETION NETWORK WITH ADVERSARIAL REFINEMENT 2559
Rong Bao, Yurui Ren, Ge Li, Wei Gao, Peking University Shenzhen Graduate School, China; Shan Liu, Tencent, China

IVMSP-33.2: WEAKLY SUPERVISED POINT CLOUD UPSAMPLING VIA OPTIMAL TRANSPORT 2564
Zezeng Li, Weimin Wang, Na Lei, Rui Wang, Dalian University of Technology, China

IVMSP-33.3: POINT CLOUD DENOISING USING NORMAL VECTOR-BASED GRAPH WAVELET SHRINKAGE 2569
Ryosuke Watanabe, Keisuke Nonaka, Haruhisa Kato, Tatsuya Kobayashi, KDDI Research, Inc., Japan; Eduardo Pavez, Antonio Ortega, University of Southern California, United States of America

IVMSP-33.4: DYNAMIC POINT CLOUD INTERPOLATION 2574
Anique Akhtar, Zhu Li, University of Missouri Kansas City, United States of America; Geert Van der Auwera, Jianle Chen, Qualcomm Technologies Inc., United States of America

IVMSP-33.5: POINT CLOUD ATTRIBUTE COMPRESSION VIA CHROMA SUBSAMPLING 2579
Shashank Nelamangala Sridhara, Eduardo Pavez, Antonio Ortega, University of Southern California, United States of America; Ryosuke Watanabe, Keisuke Nonaka, KDDI Research, Inc, Japan

IVMSP-33.6: RANGEINET: FAST LIDAR POINT CLOUD TEMPORAL INTERPOLATION 2584
Lili Zhao, Xuhu Lin, Wenyi Wang, Jianwen Chen, University of Electronic Science and Technology of China, China; Kai-Kuang Ma, Nanyang Technological University, Singapore

IVMSP-34: DEEP NEURAL NETWORKS FOR IMAGE AND VIDEO PROCESSING

IVMSP-34.1: MBNET: A MULTI-RESOLUTION BRANCH NETWORK FOR SEMANTIC SEGMENTATION OF ULTRA-HIGH RESOLUTION IMAGES 2589
Lianlei Shan, University of Chinese Academy of Sciences, China; Weiqiang Wang, University of Chinese Academy of Science, China

IVMSP-34.2: BSOLO: BOUNDARY-AWARE ONE-STAGE INSTANCE SEGMENTATION SOLO 2594
Yuxuan Zhang, Wei Yang, University of Science and Technology of China, China

IVMSP-34.3: CS-GRESNET: A SIMPLE AND HIGHLY EFFICIENT NETWORK FOR FACIAL EXPRESSION RECOGNITION 2599
Shaoping Jiang, Xiangmin Xu, Xiaofen Xing, Lin Wang, South China University of Technology, China; Fang Liu, Guangdong University of Finance, China

IVMSP-34.4: RCANET: ROW-COLUMN ATTENTION NETWORK FOR SEMANTIC SEGMENTATION 2604
Bingxu Lu, Qinghua Hu, Yu Wang, Tianjin University, China; Guosheng Hu, AnyVision, China

IVMSP-34.5: EXPLORING CATEGORY CONSISTENCY FOR WEAKLY SUPERVISED SEMANTIC SEGMENTATION 2609
Zhaozhi Xie, Lu Hongtao, Shanghai Jiao Tong University, China

IVMSP-34.6: VISION TRANSFORMER EQUIPPED WITH NEURAL RESIZER ON FACIAL EXPRESSION RECOGNITION TASK 2614
Hyeonbin Hwang, Soyeon Kim, Hyeon Yeo, KAIST, Korea, Republic of; Wei-Jin Park, Jiho Seo, Kyungtae Ko, ACRYL, Korea, Republic of

IVMSP-35: IMAGE SEGMENTATION

IVMSP-35.1: ISDA: POSITION-AWARE INSTANCE SEGMENTATION WITH DEFORMABLE ATTENTION 2619
Kaining Ying, Zhenhua Wang, Cong Bai, Pengfei Zhou, Zhejiang University of Technology, China

IVMSP-35.2: IMPROVING CLASS ACTIVATION MAP FOR WEAKLY SUPERVISED OBJECT LOCALIZATION 2624
Zhenfei Zhang, Ming-Ching Chang, University at Albany, United States of America; D.Bui Tien, Concordia University, Canada

IVMSP-35.3: A ROBUST OBJECT SEGMENTATION NETWORK FOR UNDERWATER SCENES 2629
Ruizhe Chen, Zhenqi Fu, Yue Huang, En Cheng, Xinghao Ding, Xiamen University, China

IVMSP-35.4: A FAST AND EFFICIENT NETWORK FOR SINGLE IMAGE SHADOW DETECTION 2634
Leiping Jie, Hong Kong Baptist University, Hong Kong; Hui Zhang, United International College, BNU-HKBU, China

IVMSP-35.5: IMPORTANCE SAMPLING CAMS FOR WEAKLY-SUPERVISED SEGMENTATION 2639
Arvi Jonnarth, Michael Felsberg, Linköping University, Sweden

IVMSP-35.6: DEEPGBASS: DEEP GUIDED BOUNDARY-AWARE SEMANTIC SEGMENTATION 2644
Qingfeng Liu, Hai Su, Mostafa El-Khamy, Kee-Bong Song, Samsung USA, United States of America

IVMSP-36: CAMERA CALIBRATION AND HUMAN POSE

IVMSP-36.1: CAMERA CALIBRATION THROUGH CAMERA PROJECTION LOSS 2649
Talha Butt, Murtaza Taj, Lahore University of Management Sciences, Pakistan

IVMSP-36.2: INFERRING CAMERA INTRINSICS BASED ON SURFACES OF REVOLUTION: A SINGLE IMAGE GEOMETRIC NETWORK APPROACH FOR CAMERA CALIBRATION 2654
Christopher Walker, Yuxing Wang, Yawen Lu, Guoyu Lu, Rochester Institute of Technology, United States of America

IVMSP-36.3: TEXT2VIDEO: TEXT-DRIVEN TALKING-HEAD VIDEO SYNTHESIS WITH PERSONALIZED PHONEME - POSE DICTIONARY 2659
Sibo Zhang, Jiahong Yuan, Miao Liao, Liangjun Zhang, Baidu Research, United States of America

IVMSP-36.4: TOWARDS ACCURATE CROSS-DOMAIN IN-BED HUMAN POSE ESTIMATION 2664
Mohamed Afham, Udith Haputhanthri, Jathurshan Pradeepkumar, Mithunjha Anandakumar, Chamira Edussooriya, University of Moratuwa, Sri Lanka; Ashwin De Silva, Johns Hopkins University, United States of America

IVMSP-36.5: LEARNING MONOCULAR MESH RECOVERY OF MULTIPLE BODY PARTS VIA SYNTHETICS 2669
Yu Sun, Wenpeng Gao, Yili Fu, Harbin Institute of Technology, China; Tianyu Huang, University of Maryland, China; Qian Bao, Wu Liu, AI Research of JD.com, China

IVMSP-36.6: LIGHTPOSE: A LIGHTWEIGHT AND EFFICIENT MODEL WITH TRANSFORMER FOR HUMAN POSE ESTIMATION 2674
Xiyang Liu, Peng Li, Ding Ni, Yan Wang, Hui Xue, Alibaba Group, China

IVMSP-37: STATISTICAL AND STRUCTURAL MODELS

IVMSP-37.1: ON THE OBSERVABILITY IN VISUAL SLAM NETWORKS 2679
Qier An, Yuan Shen, Tsinghua University, China

IVMSP-37.2: VARIATIONAL BAYESIAN FRAMEWORK FOR ADVANCED IMAGE GENERATION WITH DOMAIN-RELATED VARIABLES 2684
Yuxiao Li, Yuan Shen, Tsinghua University, China; Santiago Mazuelas, BCAM-Basque Center for Applied Mathematics, Spain

IVMSP-37.3: THE IMPACT OF JPEG COMPRESSION ON PRIOR IMAGE NOISE 2689
Marina Gardella, Tina Nikoukhah, Yanhao Li, Quentin Bammey, École Normale Supérieure Paris-Saclay, France

IVMSP-37.5: ON THE USE OF COMPONENT STRUCTURAL CHARACTERISTICS FOR VOXEL SEGMENTATION IN SEMICON 3D IMAGES 2694
Tin Lay Nwe, Singh Pahwa Ramanpreet, Chang Richard, Zaw Min Oo, Jie Wang, Yiqun Li, Dongyun Lin, Prasad Shitala, Sheng Dong, Institute for Infocomm Research (I2R), Singapore

IVMSP-37.6: BLIND SOURCE SEPARATION VIA A WEAK EXCLUSION PRINCIPLE 2699
Zihan Zhang, Thierry Blu, The Chinese University of Hong Kong, Hong Kong

IVMSP-38: PERSON RE-IDENTIFICATION

IVMSP-38.1: GRAPH CONVOLUTION FOR RE-RANKING IN PERSON RE-IDENTIFICATION 2704
Yuqi Zhang, Qi Qian, Chong Liu, Weihua Chen, Fan Wang, Hao Li, Rong Jin, Alibaba Group, China

IVMSP-38.2: MULTI-LEVEL RELATION AWARE NETWORK FOR PERSON RE-IDENTIFICATION 2709
Jing Yang, Canlong Zhang, Zhixin Li, Guangxi Normal University, China; Yanping Tang, Guilin University of Electronic Technology, China

IVMSP-38.3: PROGRESSIVE-GRANULARITY RETRIEVAL VIA HIERARCHICAL FEATURE ALIGNMENT FOR PERSON RE-IDENTIFICATION 2714
Zhaopeng Dou, Zhongdao Wang, Yali Li, Shengjin Wang, Tsinghua University, China

IVMSP-38.4: OCCLUDED PERSON RE-IDENTIFICATION VIA RELATIONAL ADAPTIVE FEATURE CORRECTION LEARNING 2719
Minjung Kim, MyeongAh Cho, Heansung Lee, Suhwan Cho, Sangyoun Lee, Yonsei University, Korea, Republic of

IVMSP-38.5: LEARNING SEMANTIC-ALIGNED FEATURE REPRESENTATION FOR TEXT-BASED PERSON SEARCH 2724
Shiping Li, Min Cao, Min Zhang, Soochow University, China

IVMSP-38.6: TRANSFORMER-BASED PERSON SEARCH MODEL WITH SYMMETRIC ONLINE INSTANCE MATCHING 2729
Xuezhi Xiang, Ning Lv, Yulong Qiao, Harbin Engineering University, China

IVMSP-39: SECURITY

IVMSP-39.1: WASSERTRAIN: AN ADVERSARIAL TRAINING FRAMEWORK AGAINST WASSERSTEIN ADVERSARIAL ATTACKS 2734
Qingye Zhao, Xin Chen, Enyi Tang, Xuandong Li, Nanjing University, China; Zhuoyu Zhao, Henan University of Economics and Law, China

IVMSP-39.2: EFFICIENT UNIVERSAL SHUFFLE ATTACK FOR VISUAL OBJECT TRACKING 2739
Siao Liu, Zhaoyu Chen, Wei Li, Jiwei Zhu, Jiafeng Wang, Wenqiang Zhang, Zhongxue Gan, Fudan University, China

IVMSP-39.3: NON-RIGID TRANSFORMATION BASED ADVERSARIAL ATTACK AGAINST 3D OBJECT TRACKING 2744
Riran Cheng, Nan Sang, Yinyuan Zhou, Xupeng Wang, University of Electronic Science and Technology of China, China

IVMSP-39.4: ADVERSARY DISTILLATION FOR ONE-SHOT ATTACKS ON 3D TARGET TRACKING 2749
Zhengyi Wang, Xupeng Wang, Yong Liao, Jiali Yu, University of Electronic Science and Technology of China, China; Ferdous Sohel, Murdoch University, Australia; Mohammed Bennamoun, The University of Western Australia, Australia

IVMSP-39.5: ADVERFACIAL: PRIVACY-PRESERVING UNIVERSAL ADVERSARIAL PERTURBATION AGAINST FACIAL MICRO-EXPRESSION LEAKAGES 2754
Yin Yin Low, Angeline Tanvy, Raphael C.W. Phan, Monash University, Malaysia; Xiaojun Chang, RMIT University, Australia

IVMSP-39.6: INTERPRETABLE IMAGE CLASSIFICATION USING SPARSE OBLIQUE DECISION TREES 2759
Suryabhan Singh Hada, Miguel A. Carreira-Perpinan, University of California, Merced, United States of America

IVMSP-40: UNDERWATER IMAGING

IVMSP-40.1: UNDERWATER IMAGE ENHANCEMENT VIA LEARNING WATER TYPE DESENSITIZED REPRESENTATIONS 2764
Zhenqi Fu, Xiaopeng Lin, Wu Wang, Yue Huang, Xinghao Ding, Xiamen University, China

IVMSP-40.2: A WAVELET-BASED DUAL-STREAM NETWORK FOR UNDERWATER IMAGE ENHANCEMENT 2769
Ziyin Ma, Changjae Oh, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland

IVMSP-40.3: UNSUPERVISED AND UNTRAINED UNDERWATER IMAGE RESTORATION BASED ON PHYSICAL IMAGE FORMATION MODEL 2774
Shu Chai, Zhenqi Fu, Yue Huang, Xiaotong Tu, Xinghao Ding, Xiamen University, China

IVMSP-40.4: AGCYCLEGAN: ATTENTION-GUIDED CYCLEGAN FOR SINGLE UNDERWATER IMAGE RESTORATION 2779
Zhenlong Wang, Weifeng Liu, Yanjiang Wang, Baodi Liu, China University of Petroleum (East China), China

IVMSP-40.5: UNDERWATER SMALL TARGET DETECTION BASED ON DEFORMABLE CONVOLUTIONAL PYRAMID 2784
Shuhan Qi, Jianjun Du, Mingyan Wu, Linlin Tang, Tao Qian, Xuan Wang, School of Computer Science and Technology, Harbin Institute of Technology Shenzhen, China; Hong Yi, Ricoh Software Research Center (Beijing), China

IVMSP-40.6: TOWARDS CONTROLLABLE AND PHYSICAL INTERPRETABLE UNDERWATER SCENE SIMULATION 2789
Kaixin Chen, Lin Zhang, Ying Shen, Tongji University, China; Yicong Zhou, University of Macau, China

IVMSP-41: HYPESPECTRAL IMAGING

IVMSP-41.1: GRAPH LEARNING BASED AUTOENCODER FOR HYPERSPECTRAL BAND SELECTION 2794
Yongshan Zhang, Zhenyu Wang, Xinwei Jiang, China University of Geosciences, China; Xinxin Wang, Yicong Zhou, University of Macau, China

IVMSP-41.2: MULTITASK SPARSE NEURAL NETWORK FOR HYPERSPECTRAL IMAGE DENOISING 2799
Fengchao Xiong, Jianfeng Lu, Nanjing University of Science and Technology, China; Minchao Ye, China Jiliang University, China; Jun Zhou, Griffith University, Australia; Yuntao Qian, College of Computer Science, China

IVMSP-41.3: HYPERSPECTRAL IMAGE CLASSIFICATION BASED ON CO-LEARNING THROUGH DUAL-ARCHITECTURE ENSEMBLE 2804
Xiaoyue Chen, Xianghai Cao, Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education School of Artificial Intelligence, Xidian University, China

IVMSP-41.4: MATERIAL-GUIDED SIAMESE FUSION NETWORK FOR HYPERSPECTRAL OBJECT TRACKING 2809
Zhuanfeng Li, Fengchao Xiong, Jianfeng Lu, Nanjing University of Science and Technology, China; Jun Zhou, Griffith University, Australia; Yuntao Qian, College of Computer Science, China

IVMSP-41.5: HYPERSPECTRAL IMAGE SUPER-RESOLUTION WITH DEEP PRIORS AND DEGRADATION MODEL INVERSION 2814
Xiuheng Wang, Cédric Richard, Université Côte d'Azur, CNRS, OCA., France; Jie Chen, School of Marine Science and Technology, Northwestern Polytechnical University., China

IVMSP-41.6: GEOMETRIC LOW-RANK TENSOR APPROXIMATION FOR REMOTELY SENSED HYPERSPECTRAL AND MULTISPECTRAL IMAGERY FUSION 2819
Na Liu, Wei Li, Ran Tao, Beijing Institute of Technology, China

IVMSP-42: VIDEO COMMUNICATION

IVMSP-42.1: DILATED CONVOLUTIONAL NEURAL NETWORK-BASED DEEP REFERENCE PICTURE GENERATION FOR VIDEO COMPRESSION **2824**
Haoyue Tian, Pan Gao, Nanjing University of Aeronautics and Astronautics, China; Ran Wei, Science and Technology on Electro-optic Control Laboratory, China; Manoranjan Paul, Charles Sturt University, Australia

IVMSP-42.2: RATE CONTROL FOR LEARNED VIDEO COMPRESSION **2829**
Yanghao Li, Xinyao Chen, Jisheng Li, Jiangtao Wen, Tsinghua University, China; Yuxing Han, Research Institute of Tsinghua University in Shenzhen, China; Shan Liu, Xiaozhong Xu, Tencent Media Lab, China

IVMSP-42.3: GLOBAL OPTIMIZATION SOLUTION FOR DYNAMIC ADAPTIVE 360-DEGREE STREAMING **2834**
Xuekai Wei, Weijia Jia, Beijing Normal University, China; Mingliang Zhou, Chongqing University, China

IVMSP-42.4: COLLABORATIVE OBJECT DETECTORS ADAPTIVE TO BANDWIDTH AND COMPUTATION **2839**
Juliano S. Assine, José C. S. Santos Filho, Eduardo Valle, Unicamp, Brazil

IVMSP-42.5: MA-NET: MULTI-SCALE ATTENTION-AWARE NETWORK FOR OPTICAL FLOW ESTIMATION **2844**
Mu Li, Baojiang Zhong, Soochow University, China; Kai-Kuang Ma, Nanyang Technological University, Singapore

IVMSP-42.6: MODELING HUMAN MEMORY IN MULTI-OBJECT TRACKING WITH TRANSFORMERS **2849**
Yizhuo Li, Cewu Lu, Shanghai Jiao Tong University, China

IFS-1: MULTIMEDIA FORENSICS I

IFS-1.1: REAL-WORLD ADVERSARIAL EXAMPLES VIA MAKEUP **2854**
Chang-Sheng Lin, National Chung Hsing University, Taiwan, Province of China; Chia-Yi Hsu, Chia-Mu Yu, National Yang Ming Chiao Tung University, Taiwan, Province of China; Pin-Yu Chen, IBM Research, Taiwan, Province of China

IFS-1.2: IN PURSUIT OF PRESERVING THE FIDELITY OF ADVERSARIAL IMAGES **2859**
Joseph Clements, Yingjie Lao, Clemson University, United States of America

IFS-1.3: OBJECT-ORIENTED BACKDOOR ATTACK AGAINST IMAGE CAPTIONING **2864**
Meiling Li, Nan Zhong, Xinpeng Zhang, Zhenxing Qian, Sheng Li, Fudan University, China

IFS-1.4: TOWARDS ROBUST SPEECH-TO-TEXT ADVERSARIAL ATTACK **2869**
Mohammad Esmaeilpour, Patrick Cardinal, Alessandro Lameiras Koerich, École de Technologie Supérieure, Canada

IFS-1.5: SPARSE ADVERSARIAL ATTACK FOR VIDEO VIA GRADIENT-BASED KEYFRAME SELECTION **2874**
Yixiao Xu, Xiaolei Liu, Mingyong Yin, Teng Hu, Kangyi Ding, China Academy of Engineering Physics, China

IFS-1.6: HOW SECURE ARE THE ADVERSARIAL EXAMPLES THEMSELVES? **2879**
Hui Zeng, Southwest university of science and technology, China; Kang Deng, Anjie Peng, Southwest University of Science and Technology, China; Biwei Chen, Houghton College, United States of America

IFS-2: MULTIMEDIA FORENSICS II

IFS-2.1: EXPLORING COMPLEMENTARITY OF GLOBAL AND LOCAL SPATIOTEMPORAL INFORMATION FOR FAKE FACE VIDEO DETECTION **2884**
Xiaohui Zhao, Yang Yu, Rongrong Ni, Yao Zhao, Beijing Jiaotong University, China

IFS-2.2: PANCHROMATIC IMAGERY COPY-PASTE LOCALIZATION THROUGH DATA-DRIVEN SENSOR ATTRIBUTION 2889
Edoardo Daniele Cannas, Paolo Bestagini, Stefano Tubaro, Politecnico di Milano, Italy; János Horváth, Sriram Baireddy, Edward J. Delp, Purdue University, Italy

IFS-2.3: ROBUST VIDEO HASHING BASED ON LOCAL FLUCTUATION PRESERVING FOR TRACKING DEEP FAKE VIDEOS 2894
Lv Chen, Dengpan Ye, Jiaqing Huang, wuhan university, China; Yueyun Shang, South Central University for Nationalitics, China

IFS-2.4: ADT: ANTI-DEEPFAKE TRANSFORMER 2899
Ping Wang, Kunlin Liu, Wenbo Zhou, Honggu Liu, Weiming Zhang, Nenghai Yu, University of Science and Technology of China, China; Hang Zhou, Simon Fraser University, Canada

IFS-2.5: EYES TELL ALL: IRREGULAR PUPIL SHAPES REVEAL GAN-GENERATED FACES 2904
Hui Guo, Ming-Ching Chang, SUNY at Albany, United States of America; Shu Hu, Siwei Lyu, SUNY at Buffalo, United States of America; Xin Wang, keya medical, United States of America

IFS-2.6: EXPLAINABLE ARTIFICIAL INTELLIGENCE FOR AUTHORSHIP ATTRIBUTION ON SOCIAL MEDIA 2909
Antonio Theophilo, Rafael Padilha, Fernanda A. Andaló, Anderson Rocha, University of Campinas, Brazil

IFS-3: FORENSICS AND BIOMETRICS

IFS-3.1: DUAL-DOMAIN LOW-RANK FUSION DEEP METRIC LEARNING FOR OFF-THE-PERSON ECG BIOMETRICS 2914
Guiping Zhu, Mingzhu Ma, Kuikui Wang, Gongping Yang, Shandong University, China; Yuwen Huang, Heze University, China

IFS-3.2: A ROBUST DEEP AUDIO SPLICING DETECTION METHOD VIA SINGULARITY DETECTION FEATURE 2919
Kanghao Zhang, Shulin He, Jiahui Pan, Xueliang Zhang, College of Computer Science, Inner Mongolia University, China, China; Shan Liang, Shuai Nie, Haoxin Ma, Jiangyan Yi, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China

IFS-3.3: ONLINE ECG BIOMETRICS VIA HADAMARD CODE 2924
Kuikui Wang, Gongping Yang, Yilong Yin, Shandong University, China; Yuwen Huang, Heze University, China; Lu Yang, Shandong Jianzhu University, China

IFS-3.4: FORENSIC ANALYSIS AND LOCALIZATION OF MULTIPLY COMPRESSED MP3 AUDIO USING TRANSFORMERS 2929
Ziyue Xiang, Edward Delp, Purdue University, United States of America; Paolo Bestagini, Stefano Tubaro, Politecnico di Milano, Italy

IFS-4: SURVEILLANCE, BIOMETRICS AND SECURITY

IFS-4.1: ADAPTIVE MATCHING STRATEGY FOR MULTI-TARGET MULTI-CAMERA TRACKING 2934
Chong Liu, University of Chinese Academy of Sciences, China; Yuqi Zhang, Weihua Chen, Fan Wang, Hao Li, Alibaba Group, China; Yi-Dong Shen, Institute of Software Chinese Academy of Sciences, China

IFS-4.2: GENERALIZED FACE ANTI-SPOOFING VIA CROSS-ADVERSARIAL DISENTANGLEMENT WITH MIXING AUGMENTATION 2939
Hanye Huang, Youjun Xiang, Guodong Yang, Lingling Lv, Xianfeng Li, Zichun Weng, Yuli Fu, South China University of Technology, China

IFS-4.3: FREE LUNCH FOR CROSS-DOMAIN OCCLUDED FACE RECOGNITION WITHOUT SOURCE DATA 2944
Taoshan Zhang, Youjun Xiang, Xianfeng Li, Zichun Weng, Zhen Chen, Yuli Fu, South China University of Technology, China

IFS-4.4: CONEFACE: APPROXIMATE PAIRWISE LOSS FOR FACE RECOGNITION 2949
Zijun Zhuang, Hongtao Lu, Shanghai Jiao Tong University, China

IFS-4.5: DEPTH-BASED ENSEMBLE LEARNING NETWORK FOR FACE ANTI-SPOOFING 2954
Jie Jiang, Yunlian Sun, Nanjing University of Science and Technology, China

IFS-4.6: ARE GAN-BASED MORPHS THREATENING FACE RECOGNITION? 2959
Eklavya Sarkar, Pavel Korshunov, Laurent Colbois, Sébastien Marcel, Idiap Research Institute, Switzerland

IFS-5: PRIVACY, INFORMATION, AND NETWORK SECURITY

IFS-5.1: PRIVACY PROTECTION IN LEARNING FAIR REPRESENTATIONS 2964
Yulu Jin, Lifeng Lai, University of California, Davis, United States of America

IFS-5.2: STEALTHY BACKDOOR ATTACK WITH ADVERSARIAL TRAINING 2969
Le Feng, Sheng Li, Zhenxing Qian, Xinpeng Zhang, Fudan university, China

IFS-5.3: FLDP: FLEXIBLE STRATEGY FOR LOCAL DIFFERENTIAL PRIVACY 2974
Dan Zhao, Hong Chen, Suyun Zhao, Ruixuan Liu, Cuiping Li, Xiaoying Zhang, Renmin University of China, China

IFS-5.4: ENHANCING UTILITY IN THE WATCHDOG PRIVACY MECHANISM 2979
Mohammad Amin Zarrabian, Australian National University, Australia; Ni Ding, University of Melbourne, Australia; Parastoo Sadeghi, University of New South Wales, Australia; Thierry Rakotoarivelo, The Commonwealth Scientific and Industrial Research Organisation, Australia

IFS-5.5: CYBER-THREAT PROPAGATION OVER NETWORK-SLICING ARCHITECTURES 2984
Michele Cirillo, Mario Di Mauro, Vincenzo Matta, Giuseppe Basileo, University of Salerno, Italy

IFS-5.6: PRIVACY-AWARE COMMUNICATION OVER A WIRETAP CHANNEL WITH GENERATIVE NETWORKS 2989
Ecenaz Erdemir, Pier Luigi Dragotti, Deniz Gunduz, Imperial College London, United Kingdom of Great Britain and Northern Ireland

IFS-6: ANONYMIZATION, SECURITY, AND PRIVACY

IFS-6.1: ENCRYPTED IMAGE VISUAL SECURITY INDEX VIA NON-LOCAL RECOGNIZABLE DEGREE EVALUATION 2994
Ran Shi, Nanjing University of Science and Technology, China; Jian Xiong, Nanjing University of Posts and Telecommunications, China; Tong Qiao, Hangzhou Dianzi University, China

IFS-6.2: DEFENDING AGAINST BACKDOOR ATTACKS IN FEDERATED LEARNING WITH DIFFERENTIAL PRIVACY 2999
Lu Miao, Wei Yang, Rong Hu, Liusheng Huang, University of Science and Technology of China, China; Lu Li, Yancheng Teachers University, China

IFS-6.3: SECMPNN: 3-PARTY PRIVACY-PRESERVING MOLECULAR STRUCTURE PROPERTIES INFERENCE 3004
Xinying Liao, Jiaye Xue, Ximeng Liu, Fuzhou University, China; Shengxing Yu, Peking University, China; Jiangang Shu, Peng Cheng Laboratory, China

IFS-6.4: COMPRESSED DATA SHARING BASED ON INFORMATION BOTTLENECK MODEL 3009
Behrooz Razeghi, Shideh Rezaeifar, Taras Holotyak, Slava Voloshynovskiy, University of Geneva, Switzerland; Sohrab Ferdowsi, HES-SO Geneva, Switzerland

IFS-6.5: RANDOMIZED SMOOTHING UNDER ATTACK: HOW GOOD IS IT IN PRACTICE? 3014
Thibault Maho, Teddy Furon, Erwan Le Merrer, INRIA Rennes, France

IFS-6.6: TRAINING PRIVACY-PRESERVING VIDEO ANALYTICS PIPELINES BY SUPPRESSING FEATURES THAT REVEAL INFORMATION ABOUT PRIVATE ATTRIBUTES 3019
Chau Yi Li, Andrea Cavallaro, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland

IFS-7: APPLIED CRYPTOGRAPHY AND CYBERSECURITY

IFS-7.1: UNSUPERVISED ANOMALY DETECTION FOR CONTAINER CLOUD VIA BILSTM-BASED VARIATIONAL AUTO-ENCODER 3024
Yulong Wang, Xingshu Chen, Qixu Wang, Sichuan University, China; Run Yang, Bangzhou Xin, China Academy of Engineering Physics, China

IFS-7.2: APPLYING DEEP LEARNING TO KNOWN-PLAINTEXT ATTACK ON CHAOTIC IMAGE ENCRYPTION SCHEMES 3029
Fusen Wang, Jun Sang, Chunlin Huang, Bin Cai, Hong Xiang, Chongqing University, China; Nong Sang, Huazhong University of Science and Technology, China

IFS-7.3: WORDMARKOV: A NEW PASSWORD PROBABILITY MODEL OF SEMANTICS 3034
Jiahong Xie, Haibo Cheng, Rong Zhu, Ping Wang, Peking University, China; Kaitai Liang, Delft University of Technology, Netherlands

IFS-7.4: EFFICIENT IDENTITY-BASED CHAMELEON HASH FOR MOBILE DEVICES 3039
Cong Li, Qingni Shen, Zhikang Xie, Jisheng Dong, Yuejian Fang, Zhonghai Wu, Peking University, China

IFS-7.5: PASSTRANS: AN IMPROVED PASSWORD REUSE MODEL BASED ON TRANSFORMER 3044
Xiaoxi He, Haibo Cheng, Jiahong Xie, Ping Wang, Peking University, China; Kaitai Liang, Delft University of Technology, Netherlands

IFS-8: WATERMARKING AND DATA HIDING

IFS-8.1: FOSTERING THE ROBUSTNESS OF WHITE-BOX DEEP NEURAL NETWORK WATERMARKS BY NEURON ALIGNMENT 3049
Fangqi Li, Shilin Wang, Yun Zhu, Shanghai Jiao Tong University, China

IFS-8.2: WATERMARKING IMAGES IN SELF-SUPERVISED LATENT SPACES 3054
Pierre Fernandez, Alexandre Sablayrolles, Hervé Jégou, Matthijs Douze, Meta AI, France; Teddy Furon, Univ. Rennes, Inria, CNRS, IRISA, France

IFS-8.3: SPEECH PATTERN BASED BLACK-BOX MODEL WATERMARKING FOR AUTOMATIC SPEECH RECOGNITION 3059
Haozhe Chen, Weiming Zhang, Kunlin Liu, Kejiang Chen, Han Fang, Nenghai Yu, University of Science and Technology of China, China

IFS-8.4: ENCRYPTION RESISTANT DEEP NEURAL NETWORK WATERMARKING 3064
Guobiao Li, Sheng Li, Zhenxing Qian, Xinpeng Zhang, Fudan university, China

IFS-8.5: ATTRIBUTABLE WATERMARKING OF SPEECH GENERATIVE MODELS 3069
Yongbaek Cho, Changhoon Kim, Yezhou Yang, Yi Ren, Arizona State University, United States of America

IFS-9: STEGANOGRAPHY, STEGANALYSIS, AND DATA PRIVACY

IFS-9.1: EXPLOITING LANGUAGE MODEL FOR EFFICIENT LINGUISTIC STEGANALYSIS 3074
Biao Yi, Hanzhou Wu, Guorui Feng, Xinpeng Zhang, Shanghai University, China

IFS-9.2: PATCH STEGANALYSIS: A SAMPLING BASED DEFENSE AGAINST ADVERSARIAL STEGANOGRAPHY 3079
Chuan Qin, Na Zhao, Weiming Zhang, Nenghai Yu, University of Science and Technology of China, China

IFS-9.3: AN EFFECTIVE STEGANALYSIS FOR ROBUST STEGANOGRAPHY WITH REPETITIVE JPEG COMPRESSION 3084
Jinliu Feng, Yaofei Wang, Kejiang Chen, Weiming Zhang, Nenghai Yu, University of Science and Technology of China, China

IFS-9.4: IMAGE STEGANALYSIS WITH CONVOLUTIONAL VISION TRANSFORMER 3089
Ge Luo, Ping Wei, Shuwen Zhu, Xinpeng Zhang, Zhenxing Qian, Sheng Li, Fudan University, China

IFS-9.5: A BRIDGE BETWEEN FEATURES AND EVIDENCE FOR BINARY ATTRIBUTE-DRIVEN PERFECT PRIVACY 3094
Paul-Gauthier Noé, Driss Matrouf, Pierre-Michel Bousquet, Jean-François Bonastre, Avignon University, France; Andreas Nautsch, vitas.ai, Germany

IFS-9.6: PRESERVING TRAJECTORY PRIVACY IN DRIVING DATA RELEASE 3099
Yi Xu, Chong Xiao Wang, Yang Song, Wee Peng Tay, Nanyang Technological University, Singapore

MLSP-1: DEEP LEARNING FOR AUDIO AND MUSIC APPLICATIONS

MLSP-1.1: DIRECT DESIGN OF BIQUAD FILTER CASCADES WITH DEEP LEARNING BY SAMPLING RANDOM POLYNOMIALS 3104
Joseph Colonel, Christian Steinmetz, Joshua Reiss, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland; Marcus Michelen, University of Illinois at Chicago, United States of America

MLSP-1.2: AN END-TO-END DEEP LEARNING SPEECH CODING AND DENOISING STRATEGY FOR COCHLEAR IMPLANTS 3109
Tom Gajecki, Waldo Nogueira, Hannover Medical School, Germany

MLSP-1.3: EXPLOITING HYBRID MODELS OF TENSOR-TRAIN NETWORKS FOR SPOKEN COMMAND RECOGNITION 3114
Jun Qi, Georgia Institute of Technology, United States of America; Javier Tejedor, Universidad San Pablo-CEU, CEU Universities, Spain

MLSP-1.4: LEARNABLE WAVELET PACKET TRANSFORM FOR DATA-ADAPTED SPECTROGRAMS 3119
Gaëtan Frusque, Olga Fink, ETH Zürich, Switzerland

MLSP-1.5: MUSIC ENHANCEMENT VIA IMAGE TRANSLATION AND VOCODING 3124
Nikhil Kandpal, University of North Carolina at Chapel Hill, United States of America; Oriol Nieto, Zeyu Jin, Adobe, United States of America

MLSP-1.6: PROGRESSIVE TEACHER-STUDENT TRAINING FRAMEWORK FOR MUSIC TAGGING 3129
Rui Lu, Baigong Zheng, Jiarui Hai, Fei Tao, Zhiyao Duan, Ji Liu, Kuaishou Technology, China

MLSP-2: PATTERN RECOGNITION AND CLASSIFICATION I

MLSP-2.1: JOINT DUAL-DOMAIN MATRIX FACTORIZATION FOR ECG BIOMETRIC RECOGNITION 3134
Kuikui Wang, Gongping Yang, Yilong Yin, Shandong University, China; Yuwen Huang, Heze University, China; Lu Yang, Shandong Jianzhu University, China

MLSP-2.2: ITERATIVE SELF KNOWLEDGE DISTILLATION --- FROM POTHOLE CLASSIFICATION TO FINE-GRAINED AND COVID RECOGNITION 3139
Kuan-Chuan Peng, Mitsubishi Electric Research Laboratories (MERL), United States of America

MLSP-2.3: ATTENTION-BASED ADVERSARIAL PARTIAL DOMAIN ADAPTATION 3144
Mengzhu Wang, Xiong Peng, Wei Yu, Zhigang Luo, National University of Defense Technology, China; Shan An, Tech & Data Center JD.COM Inc, China; Xiao Luo, Peking University, China; Junyang Chen, Shenzhen University, China

MLSP-2.4: GROUP-WISE FEATURE SELECTION FOR SUPERVISED LEARNING 3149
Qi Xiao, Hebi Li, Jin Tian, Zhengdao Wang, Iowa State University, United States of America

MLSP-2.5: A LIGHT WEIGHT MODEL FOR VIDEO SHOT OCCLUSION DETECTION 3154
Junhua Liao, Wanbin Zhao, Yanbing Yang, Liangyin Chen, Sichuan University, China; Haihan Duan, The Chinese University of Hong Kong, China

MLSP-2.6: DETECTING BACKDOOR ATTACKS AGAINST POINT CLOUD CLASSIFIERS 3159
Zhen Xiang, David J. Miller, Xi Li, George Kesidis, Pennsylvania State University, United States of America; Siheng Chen, Shanghai Jiao Tong University, China

MLSP-3: SELF-SUPERVISED LEARNING FOR SPEECH AND AUDIO PROCESSING I

MLSP-3.1: CHARACTERIZING THE ADVERSARIAL VULNERABILITY OF SPEECH SELF-SUPERVISED LEARNING 3164
Haibin Wu, Hung-yi Lee, National Taiwan University, China; Bo Zheng, Xu Li, Xixin Wu, Helen Meng, The Chinese University of Hong Kong, Hong Kong

MLSP-3.2: UNIVERSAL PARALINGUISTIC SPEECH REPRESENTATIONS USING SELF-SUPERVISED CONFORMERS 3169
Joel Shor, Verily Life Sciences, United States of America; Aren Jansen, Wei Han, Daniel Park, Yu Zhang, Google, United States of America

MLSP-3.3: A NOISE-ROBUST SELF-SUPERVISED PRE-TRAINING MODEL BASED SPEECH REPRESENTATION LEARNING FOR AUTOMATIC SPEECH RECOGNITION 3174
Qiu-Shi Zhu, Jie Zhang, Zi-Qiang Zhang, Ming-Hui Wu, Xin Fang, Li-Rong Dai, University of Science and Technology of China, China

MLSP-3.4: AN ADAPTER BASED PRE-TRAINING FOR EFFICIENT AND SCALABLE SELF-SUPERVISED SPEECH REPRESENTATION LEARNING 3179
Samuel Kessler, University of Oxford, United Kingdom of Great Britain and Northern Ireland; Bethan Thomas, Salah Karout, Huawei R&D UK, United Kingdom of Great Britain and Northern Ireland

MLSP-3.5: DRVC: A FRAMEWORK OF ANY-TO-ANY VOICE CONVERSION WITH SELF-SUPERVISED LEARNING 3184
Qiqi Wang, The University of Auckland, New Zealand; Xulong Zhang, Jianzong Wang, Ning Cheng, Jing Xiao, Ping An Technology (Shenzhen) Co., Ltd., China

MLSP-3.6: CONTRASTIVE PREDICTION STRATEGIES FOR UNSUPERVISED SEGMENTATION AND CATEGORIZATION OF PHONEMES AND WORDS 3189
Santiago Cuervo, Maciej Grabias, Grzegorz Ciesielski, Paweł Rychlikowski, University of Wroclaw, Poland; Jan Chorowski, University of Wroclaw / NavAlgo, Poland; Adrian Łańcucki, NVIDIA, Poland; Ricard Marxer, Université de Toulon, Aix Marseille Univ, CNRS, LIS, France

MLSP-4: DEEP LEARNING I

MLSP-4.1: UNCERTAINTY IN DATA-DRIVEN KALMAN FILTERING FOR PARTIALLY KNOWN STATE-SPACE MODELS 3194
Itzik Klein, University of Haifa, Israel; Guy Revach, Jonas E. Mehr, ETH Zürich, Switzerland; Nir Shlezinger, Ben-Gurion University of the Negev, Israel; Ruud J. G. van Sloun, Eindhoven University of Technology, and with Phillips Research, Netherlands; Yonina. C. Eldar, Weizmann Institute of Science, Israel

MLSP-4.2: DEEP PIECEWISE HASHING FOR EFFICIENT HAMMING SPACE RETRIEVAL 3199
Jingzi Gu, Dayan Wu, Peng Fu, Bo Li, Weiping Wang, Institute of Information Engineering, Chinese Academy of Sciences, China

MLSP-4.3: SODA: SELF-ORGANIZING DATA AUGMENTATION IN DEEP NEURAL NETWORKS - APPLICATION TO BIOMEDICAL IMAGE SEGMENTATION TASKS 3204
Arnaud Deleruyelle, John Klein, Cristian Versari, Université de Lille, France

MLSP-4.4: DEEP IMPULSE RESPONSES: ESTIMATING AND PARAMETERIZING FILTERS WITH DEEP NETWORKS 3209
Alexander Richard, Peter Dodds, Vamsi Krishna Ithapu, Facebook Inc, United States of America

MLSP-4.5: JOINT TEMPORAL CONVOLUTIONAL NETWORKS AND ADVERSARIAL DISCRIMINATIVE DOMAIN ADAPTATION FOR EEG-BASED CROSS-SUBJECT EMOTION RECOGNITION 3214
Zhipeng He, Yongshi Zhong, Jiahui Pan, South China Normal University, China

MLSP-4.6: GRADIENT VARIANCE LOSS FOR STRUCTURE-ENHANCED IMAGE SUPER-RESOLUTION 3219
Lusine Abrahamyan, Nikos Deligiannis, Vrije Universiteit Brussel, Belgium; Anh Minh Truong, Wilfried Philips, Ghent University, Belgium

MLSP-5: PATTERN RECOGNITION AND CLASSIFICATION II

MLSP-5.1: LABEL-OCCURRENCE-BALANCED MIXUP FOR LONG-TAILED RECOGNITION 3224
Shaoyu Zhang, Chen Chen, Silong Peng, Institute of Automation, Chinese Academy of Sciences, China; Xiujuan Zhang, Inner Mongolia Key Laboratory of Molecular Biology on Featured Plants, China

MLSP-5.2: TNTC: TWO-STREAM NETWORK WITH TRANSFORMER-BASED COMPLEMENTARITY FOR GAIT-BASED EMOTION RECOGNITION 3229
Chuanfei Hu, Weijie Sheng, Xinde Li, Southeast University, China; Bo Dong, Zhejiang University, China

MLSP-5.3: A FREE LUNCH FROM VIT: ADAPTIVE ATTENTION MULTI-SCALE FUSION TRANSFORMER FOR FINE-GRAINED VISUAL RECOGNITION 3234
Yuan Zhang, Jian Cao, Ling Zhang, Xiangcheng Liu, Zhiyi Wang, Feng Ling, Peking University, China; Weiqian Chen, Peking University, Alibaba Group, Inc., China

MLSP-5.4: SELF-SUPERVISED CONTRASTIVE LEARNING FOR CROSS-DOMAIN HYPERSPECTRAL IMAGE REPRESENTATION 3239
Hyungtae Lee, Heesung Kwon, DEVCOM Army Research Laboratory, United States of America

MLSP-5.5: GOS: A LARGE-SCALE ANNOTATED OUTDOOR SCENE SYNTHETIC DATASET 3244
Mingye Xie, Ting Liu, Yuzhuo Fu, Shanghai Jiao Tong University, China

MLSP-5.6: OUT-OF-DISTRIBUTION AS A TARGET CLASS IN SEMI-SUPERVISED LEARNING 3249
Antoine Tadros, Sébastien Drouyer, Rafael Grompone von Gioi, Ecole Normale Supérieur Paris-Saclay, France

MLSP-6: SELF-SUPERVISED LEARNING FOR SPEECH AND AUDIO PROCESSING II

MLSP-6.1: SELF-SUPERVISED ACOUSTIC ANOMALY DETECTION VIA CONTRASTIVE LEARNING 3253
Hadi Hojjati, Narges Armanfard, McGill University, Canada

MLSP-6.2: DON'T SPEAK TOO FAST: THE IMPACT OF DATA BIAS ON SELF-SUPERVISED SPEECH MODELS 3258
Yen Meng, Yi-Hui Chou, Andy T. Liu, Hung-yi Lee, National Taiwan University, Taiwan

MLSP-6.3: SELF-SUPERVISED LEARNING METHOD USING MULTIPLE SAMPLING STRATEGIES FOR GENERAL-PURPOSE AUDIO REPRESENTATION 3263
Ibuki Kuroyanagi, Nagoya University, Japan; Tatsuya Komatsu, LINE Corporation, Japan

MLSP-6.4: SELF SUPERVISED REPRESENTATION LEARNING WITH DEEP CLUSTERING FOR ACOUSTIC UNIT DISCOVERY FROM RAW SPEECH 3268
Varun Krishna, Sriram Ganapathy, LEAP lab, Indian Institute of Science, Bangalore, India., India

MLSP-6.5: T-NGA: TEMPORAL NETWORK GRAFTING ALGORITHM FOR LEARNING TO PROCESS SPIKING AUDIO SENSOR EVENTS 3273
Shu Wang, Yuhuang Hu, Shih-Chii Liu, Institute of Neuroinformatics, University of Zürich and ETH Zürich, Switzerland

MLSP-6.6: CONTRASTIVE KNOWLEDGE GRAPH ATTENTION NETWORK FOR REQUEST-BASED RECIPE RECOMMENDATION 3278
Xiyao Ma, Zheng Gao, Qian Hu, Mohamed AbdelHady, Amazon Alexa AI, United States of America

MLSP-7: DEEP LEARNING FOR IMAGE AND VIDEO PROCESSING I

MLSP-7.1: TARGETDROP: A TARGETED REGULARIZATION METHOD FOR CONVOLUTIONAL NEURAL NETWORKS 3283
Hui Zhu, Xiaofang Zhao, Institute of Computing Technology, Chinese Academy of Sciences, China

MLSP-7.2: COARSE-TO-FINE UNSUPERVISED CHANGE DETECTION FOR REMOTE SENSING IMAGES VIA OBJECT-BASED MRF AND INCEPTION UNET 3288
Xuan Hou, Haonan Shi, Ying Li, Northwestern Polytechnical University, China; Yunpeng Bai, Aberystwyth University, China

MLSP-7.3: COMBATING FALSE SENSE OF SECURITY: BREAKING THE DEFENSE OF ADVERSARIAL TRAINING VIA NON-GRADIENT ADVERSARIAL ATTACK 3293
Mingyuan Fan, Wenzhong Guo, Ximeng Liu, Fuzhou University, China; Yang Liu, Xidian University, China; Cen Chen, East China Normal University, China; Shengxing Yu, Peking University, China

MLSP-7.4: DYNAMICALLY PRUNING SEGFORMER FOR EFFICIENT SEMANTIC SEGMENTATION 3298
Haoli Bai, The Chinese University of Hong Kong, China; Hongda Mao, Dinesh Nair, Amazon.com, United States of America

MLSP-7.5: DEFORMABLE VISTR: SPATIO TEMPORAL DEFORMABLE ATTENTION FOR VIDEO INSTANCE SEGMENTATION 3303
Sudhir Yarram, Jialian Wu, Junsong Yuan, University at Buffalo, United States of America; Pan Ji, Yi Xu, OPPO US Research Center, InnoPeak Technology Inc., United States of America

MLSP-8: PATTERN RECOGNITION AND CLASSIFICATION III

MLSP-8.1: ATTENTIONAL GATED RES2NET FOR MULTIVARIATE TIME SERIES CLASSIFICATION 3308
Chao Yang, Xianzhi Wang, University of Technology, Sydney, Australia; Lina Yao, University of New South Wales, Australia; Guodong Long, Jing Jiang, Guandong Xu, University of Technology Sydney, Australia

MLSP-8.2: CONVEX CLUSTERING FOR AUTOCORRELATED TIME SERIES 3313
Max Revay, Victor Solo, University of New South Wales, Australia

MLSP-8.3: INVESTIGATING THE POTENTIAL OF AUXILIARY-CLASSIFIER GANS FOR IMAGE CLASSIFICATION IN LOW DATA REGIMES 3318
Amil Dravid, Florian Schiffers, Yunan Wu, Oliver Cossairt, Aggelos Katsaggelos, Northwestern University, United States of America

MLSP-8.4: FEATURE AUGMENTATION LEARNING FOR FEW-SHOT PALMPRINT IMAGE RECOGNITION WITH UNCONSTRAINED ACQUISITION 3323
Kunlei Jing, Xinman Zhang, Xi'an Jiaotong University, China; Zhiyuan Yang, Bihan Wen, Nanyang Technological University, China

MLSP-8.5: PRIME KNOWLEDGE WITH LOCAL PATTERN CONSISTENCY FOR KNOWLEDGE DISTILLATION 3328
Qiankun Tang, Jun Wang, Zhejiang Lab, China; Xiaogang Xu, Zhejiang Gongshang University, China

MLSP-8.6: TEST-TIME DETECTION OF BACKDOOR TRIGGERS FOR POISONED DEEP NEURAL NETWORKS 3333
Xi Li, Zhen Xiang, David Miller, George Kesidis, Pennsylvania State University, United States of America

MLSP-9: MATRIX FACTORIZATION AND COMPLETION: THEORY AND APPLICATIONS

MLSP-9.1: MULTI-VIEW DATA REPRESENTATION VIA DEEP AUTOENCODER-LIKE NONNEGATIVE MATRIX FACTORIZATION 3338
Haonan Huang, Yihao Luo, Guoxu Zhou, Qibin Zhao, Guangdong University of Technology, China

MLSP-9.2: ON IDENTIFIABLE POLYTOPE CHARACTERIZATION FOR POLYTOPIC MATRIX FACTORIZATION 3343
Bariscan Bozkurt, Alper Erdogan, Koc University, Turkey

MLSP-9.3: FAST LEARNING OF FAST TRANSFORMS, WITH GUARANTEES 3348
Quoc-Tung Le, Elisa Riccietti, Ecole normale supérieure de Lyon, France; Léon Zheng, valeo.ai, France; Rémi Gribonval, Inria, France

MLSP-9.4: REGRESSION ASSISTED MATRIX COMPLETION FOR RECONSTRUCTING A PROPAGATION FIELD WITH APPLICATION TO SOURCE LOCALIZATION 3353
Hao Sun, Junting Chen, The Chinese University of Hong Kong(Shenzhen), China

MLSP-9.5: MATRIX DECOMPOSITION ON GRAPHS: A SIMPLIFIED FUNCTIONAL VIEW 3358
Abhishek Sharma, Maks Ovsjanikov, Ecole Polytechnique, IP Paris, France

MLSP-10: SPARSITY AWARE PROCESSING

MLSP-10.1: LEARNING TO SAMPLE FOR SPARSE SIGNALS 3363
Satish Mulleti, Haiyang Zhang, Weizmann Institute of Science, Rehovot, Israel, Israel; Yonina C. Eldar, Weizmann Institute of Science, Israel

MLSP-10.2: MIXTURE MODEL AUTO-ENCODERS: DEEP CLUSTERING THROUGH DICTIONARY LEARNING 3368
Alexander Lin, Demba Ba, Harvard University, United States of America; Andrew Song, Massachusetts Institute of Technology, United States of America

MLSP-10.3: EXPLORING THE EFFECT OF L0/L2 REGULARIZATION IN NEURAL NETWORK PRUNING USING THE LC TOOLKIT 3373
Yerlan Idelbayev, Miguel Á. Carreira-Perpiñán, UC Merced, United States of America

MLSP-10.4: DICTIONARY LEARNING WITH UNIFORM SPARSE REPRESENTATIONS FOR ANOMALY DETECTION 3378
Paul Irofti, Cristian Rusu, Andrei Patrascu, University of Bucharest, Faculty of Mathematics and Computer Science, Research Center for Logic, Optimization and Security (LOS), Romania

MLSP-10.5: DATA-DRIVEN SPATIALLY DEPENDENT PDE IDENTIFICATION 3383
Ruixian Liu, Michael Bianco, Peter Gerstoft, Bhaskar Rao, University of California, San Diego, United States of America

MLSP-10.6: SPARSITY IMPROVES UNSUPERVISED ATTRIBUTE DISCOVERY IN STYLEGAN 3388
Shusen Liu, Rushil Anirudh, Jayaraman J. Thiagarajan, Peer-Timo Bremer, Lawrence Livermore National Laboratory, United States of America

MLSP-11: DEEP LEARNING FOR IMAGE AND VIDEO PROCESSING II

MLSP-11.1: IMAGE-TO-GRAPH TRANSFORMERS FOR CHEMICAL STRUCTURE RECOGNITION 3393
Sanghyun Yoo, Ohyun Kwon, Hoshik Lee, Samsung Advanced Institute of Technology, Korea, Republic of

MLSP-11.2: A SIMPLE HYBRID FILTER PRUNING FOR EFFICIENT EDGE INFERENCE 3398
Shabbeer Basha Shaik Hussain, Sheethal N Gowda, Jayachandra Dakala, DryvAmigo, Bangalore, India., India

MLSP-11.3: AN ENHANCED DEEP LEARNING APPROACH FOR TECTONIC FAULT AND FRACTURE EXTRACTION IN VERY HIGH RESOLUTION OPTICAL IMAGES 3403
Bilel Kanoun, Géoazur, INRIA, France; Mohamed Abderrazak Cherif, Géoazur, LuxCarta Technology, France; Isabelle Manighetti, Géoazur, France; Yuliya Tarabalka, LuxCarta Technology, France; Josiane Zerubia, INRIA, France

MLSP-11.4: JOINT LEARNING OF FEATURE EXTRACTION AND COST AGGREGATION FOR SEMANTIC CORRESPONDENCE 3408
Jiwon Kim, Youngjo Min, Mira Kim, Seungryong Kim, Korea University, Korea, Republic of

MLSP-11.5: GENERALIZED ZERO-SHOT LEARNING USING CONDITIONAL WASSERSTEIN AUTOENCODER 3413
Junhan Kim, Byonghyo Shim, Seoul National University, Korea, Republic of

MLSP-11.6: MBA-RAINGAN: A MULTI-BRANCH ATTENTION GENERATIVE ADVERSARIAL NETWORK FOR MIXTURE OF RAIN REMOVAL 3418
Yiyang Shen, Yidan Feng, Dong Liang, Mingqiang Wei, Nanjing University of Aeronautics and Astronautics, China; Weiming Wang, Hong Kong Metropolitan University, China; Jing Qin, The Hong Kong Polytechnic University, China; Haoran Xie, Lingnan University, China

MLSP-12: PATTERN RECOGNITION AND CLASSIFICATION IV

MLSP-12.1: END-TO-END KEYWORD SPOTTING USING NEURAL ARCHITECTURE SEARCH AND QUANTIZATION 3423
David Peter, Wolfgang Roth, Franz Pernkopf, Graz University of Technology, Austria

MLSP-12.2: SYNPOSE: A LARGE-SCALE AND DENSELY ANNOTATED SYNTHETIC DATASET FOR HUMAN POSE ESTIMATION IN CLASSROOM 3428
Zefang Yu, Yangcheng Li, Yicheng Liu, Ting Liu, Yuzhuo Fu, Shanghai Jiao Tong University, China

MLSP-12.3: STPOINTGCN: SPATIAL TEMPORAL GRAPH CONVOLUTIONAL NETWORK FOR MULTIPLE PEOPLE RECOGNITION USING MILLIMETER-WAVE RADAR 3433
Chunyu Wang, Peixian Gong, Lihua Zhang, Fudan University, China

MLSP-12.4: MULTIPLE TEMPORAL CONTEXT EMBEDDING NETWORKS FOR UNSUPERVISED TIME SERIES ANOMALY DETECTION 3438
Hanhui Li, Xinggan Peng, Huiping Zhuang, Zhiping Lin, Nanyang Technological University, Singapore

MLSP-12.5: INTERMIX: AN INTERFERENCE-BASED DATA AUGMENTATION AND REGULARIZATION TECHNIQUE FOR AUTOMATIC DEEP SOUND CLASSIFICATION 3443
Ramit Sawhney, Scaler, India; Atula Tejaswi Neerkaje, Manipal Institute of Technology, India

MLSP-12.6: CROSS-LAYER AGGREGATION WITH TRANSFORMERS FOR MULTI-LABEL IMAGE CLASSIFICATION 3448
Weibo Zhang, Fuqing Zhu, Songlin Hu, 1, Institute of Information Engineering, Chinese Academy of Sciences;2,School of Cyber Security, University of Chinese Academy of Sciences, China; Jizhong Han, Tao Guo, Institute of Information Engineering, Chinese Academy of Sciences, China

MLSP-13: SELF-SUPERVISED LEARNING METHODS I

MLSP-13.1: VISUAL REPRESENTATION LEARNING WITH SELF-SUPERVISED ATTENTION FOR LOW-LABEL HIGH-DATA REGIME ..... N/A
Prarthana Bhattacharyya, University of Waterloo, Canada; Chenge Li, Xiaonan Zhao, István Fehérvári, Jason Sun, Amazon Inc., Canada

MLSP-13.2: TRIBYOL: TRIPLET BYOL FOR SELF-SUPERVISED REPRESENTATION LEARNING 3458
Guang Li, Ren Togo, Takahiro Ogawa, Miki Haseyama, Hokkaido University, Japan

MLSP-13.3: SAGA: SELF-AUGMENTATION WITH GUIDED ATTENTION FOR REPRESENTATION LEARNING 3463
Chun-Hsiao Yeh, Academia Sinica / UC Berkeley, Taiwan; Cheng-Yao Hong, Yen-Chi Hsu, Tyng-Luh Liu, Academia Sinica, Taiwan

MLSP-13.4: AN ANOMALY DETECTION METHOD BASED ON SELF-SUPERVISED LEARNING WITH SOFT LABEL ASSIGNMENT FOR DEFECT VISUAL INSPECTION 3468
Chuanfei Hu, Yongxiong Wang, University of Shanghai for Science and Technology, China

MLSP-13.5: CONTRASTIVE PREDICTIVE CODING FOR ANOMALY DETECTION OF FETAL HEALTH FROM THE CARDIOTOCOGRAM 3473
Ivar R. de Vries, Iris A.M. Huijben, Ruud J.G. van Sloun, Rik Vullings, Eindhoven University of Technology, Netherlands; René D. Kok, Nemo Healthcare BV, Netherlands

MLSP-13.6: GRAPH FINE-GRAINED CONTRASTIVE REPRESENTATION LEARNING 3478
Hui Tang, Xun Liang, Yuhui Guo, Xiangping Zheng, Bo Wu, Renmin University of China, China

MLSP-14: DEEP LEARNING II

MLSP-14.1: POSITION-INVARIANT ADVERSARIAL ATTACKS ON NEURAL MODULATION RECOGNITION 3483
Zhen Yu, Yifeng Xiong, Kun He, Huazhong University of Science and Technology, China; Shao Huang, Yaodong Zhao, Jie Gu, Science and Technology on Electronic Information Control Laboratory, China

MLSP-14.2: USING A SINGLE INPUT TO FORECAST HUMAN ACTION KEYSTATES IN EVERYDAY PICK AND PLACE ACTIONS **3488**
Haziq Razali, Yiannis Demiris, Imperial College London, United Kingdom of Great Britain and Northern Ireland

MLSP-14.3: ADVERSARIAL ROBUSTNESS BY DESIGN THROUGH ANALOG COMPUTING AND SYNTHETIC GRADIENTS **3493**
Alessandro Cappelli, Iacopo Poli, LightOn, France; Ruben Ohana, Julien Launay, LightOn/ Ecole Normale Superieure, France; Laurent Meunier, Université Dauphine/ Facebook AI Research, France; Florent Krzakala, LightOn/ EPFL, France

MLSP-14.4: DIFFERENTIABLE PROGRAMMING A LA MOREAU **3498**
Vincent Roulet, Zaid Harchaoui, University of Washington, United States of America

MLSP-14.5: DATA AGNOSTIC FILTER GATING FOR EFFICIENT DEEP NETWORKS **3503**
Hongyan Xu, Arcot Sowmya, the University of New South Wales, Australia; Xiu Su, Chang Xu, the University of Sydney, Australia; Shan You, Tao Huang, Fei Wang, Chen Qian, Sensetime, China; Changshui Zhang, Tsinghua University, China; Dadong Wang, Data61, The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

MLSP-15: PATTERN RECOGNITION AND CLASSIFICATION V

MLSP-15.1: NEAREST SUBSPACE SEARCH IN THE SIGNED CUMULATIVE DISTRIBUTION TRANSFORM SPACE FOR 1D SIGNAL CLASSIFICATION **3508**
Abu Hasnat Mohammad Rubaiyat, Mohammad Shifat-E-Rabbi, Yan Zhuang, Shiying Li, Gustavo Rohde, University of Virginia, United States of America

MLSP-15.2: ENERGY ALIGNMENT FOR BIAS RECTIFICATION IN CLASS INCREMENTAL LEARNING **3513**
Bowen Zhao, Xi Xiao, Shutao Xia, Tsinghua University, China; Chen Chen, Qi Ju, Tencent, China

MLSP-15.3: A TWO-STAGE CONTRASTIVE LEARNING FRAMEWORK FOR IMBALANCED AERIAL SCENE RECOGNITION **3518**
Lexing Huang, Senlin Cai, Yihong Zhuang, Changxing Jing, Yue Huang, Xiaotong Tu, Xinghao Ding, Xiamen University, China

MLSP-15.4: A MAXIMAL CORRELATION APPROACH TO IMPOSING FAIRNESS IN MACHINE LEARNING **3523**
Joshua Lee, Yuheng Bu, Gregory Wornell, MIT, United States of America; Prasanna Sattigeri, Rameswar Panda, Leonid Karlinsky, Rogerio Feris, IBM Research, United States of America

MLSP-15.5: BOUNDARY-AWARE BIAS LOSS FOR TRANSFORMER-BASED AERIAL IMAGE SEGMENTATION MODEL **3528**
Yan Zhang, Xue Jiang, Bo Hu, Xinbo Gao, Chongqing University of Posts and Telecommunications, China; Siqi Liu, Peking University, China

MLSP-15.6: INVESTIGATING ROBUSTNESS OF BIOLOGICAL VS. BACKPROP BASED LEARNING **3533**
*Yanpeng Zhou, Maosen Wang, Ponnuthurai Nagaratnam Suganthan, Nanyang Technological University, Singapore; Manas Gupta, Arulmurugan Ambikapathi, Ramasamy Savitha, A*STAR, Singapore*

MLSP-16: SELF-SUPERVISED LEARNING METHODS II

MLSP-16.1: SEMI-SUPERVISED GAUSSIAN MIXTURE VARIATIONAL AUTOENCODER FOR PULSE SHAPE DISCRIMINATION **3538**
Abdullah Abdulaziz, Yoann Altmann, Stephen McLaughlin, Heriot-Watt University, United Kingdom of Great Britain and Northern Ireland; Jianxin Zhou, Angela Di Fulvio, University of Illinois, United States of America

MLSP-16.2: HOW NEURAL PROCESSES IMPROVE GRAPH LINK PREDICTION 3543
Huidong Liang, Junbin Gao, The University of Sydney, Australia

MLSP-16.3: UNCERTAINTY ESTIMATION WITH A VAE-CLASSIFIER HYBRID MODEL 3548
Shuyu Lin, Niki Trigoni, Stephen Roberts, University of Oxford, United Kingdom of Great Britain and Northern Ireland; Ronald Clark, Imperial College London, United Kingdom of Great Britain and Northern Ireland

MLSP-16.4: CONTEXT-AWARE GRAPH-BASED SELF-SUPERVISED LEARNING OF 3553
WHOLE SLIDE IMAGES
Milan Aryal, Nasim Yahyasoltani, Marquette University, United States of America

MLSP-16.5: CONTRASTIVE SENSOR TRANSFORMER FOR PREDICTIVE 3558
MAINTENANCE OF INDUSTRIAL ASSETS
Zaharah Bukhsh, Eindhoven University of Technology, Netherlands

MLSP-16.6: IMPROVING ANOMALY DETECTION WITH A SELF-SUPERVISED TASK 3563
BASED ON GENERATIVE ADVERSARIAL NETWORK
Heyan Chai, Weijun Su, Siyu Tang, Harbin Institute of Technology (Shenzhen), China; Ye Ding, Dongguan University of Technology, China; Binxing Fang, Qing Liao, Harbin Institute of Technology (Shenzhen), China

MLSP-17: GRAPH NEURAL NETWORKS

MLSP-17.1: STGAT-MAD : SPATIAL-TEMPORAL GRAPH ATTENTION NETWORK FOR 3568
MULTIVARIATE TIME SERIES ANOMALY DETECTION
Jun Zhan, Siqi Wang, Chengkun Wu, Detian Zeng, National University of Defense Technology, China; Xiandong Ma, Lancaster University, United Kingdom of Great Britain and Northern Ireland; Canqun Yang, National Supercomputing Center of Tianjin, China; Shilin Wang, Beijing Goldwind Smart Energy Technology Co., Ltd., China

MLSP-17.2: DUAL GRAPH CROSS-DOMAIN FEW-SHOT LEARNING FOR 3573
HYPERSPECTRAL IMAGE CLASSIFICATION
Yuxiang Zhang, Wei Li, Mengmeng Zhang, Ran Tao, Beijing Institute of Technology, China

MLSP-17.3: PERSONALIZED PAGERANK GRAPH ATTENTION NETWORKS 3578
Julie Choi, Amazon, United States of America

MLSP-17.4: MULTI-RELATION MESSAGE PASSING FOR MULTI-LABEL TEXT 3583
CLASSIFICATION
Muberra Ozmen, Mark Coates, McGill University, Canada; Hao Zhang, Hong Kong University of Science and Technology, Hong Kong; Pengyun Wang, Huawei Noah's Ark Lab, China

MLSP-17.5: ADAPTIVE ATTENTION GRAPH CAPSULE NETWORK 3588
Xiangping Zheng, Xun Liang, Bo Wu, Yuhui Guo, Hui Tang, Renmin University of China, China

MLSP-17.6: GRAPH CONVOLUTIONAL NETWORKS WITH AUTOENCODER-BASED 3593
COMPRESSION AND MULTI-LAYER GRAPH LEARNING
Lorenzo Giusti, Claudio Battiloro, Paolo Di Lorenzo, Sergio Barbarossa, Sapienza University of Rome, Italy

MLSP-18: MACHINE LEARNING FOR AUDIO, ACOUSTICS, AND SPEECH PROCESSING

MLSP-18.1: DEEP AUGMENTED MUSIC ALGORITHM FOR DATA-DRIVEN DOA 3598
ESTIMATION
Julian P. Merkofer, Guy Revach, ETH Zurich, Switzerland; Nir Shlezinger, Ben-Gurion University of the Negev, Israel; Ruud J. G. van Sloun, Eindhoven University of Technology, Netherlands

MLSP-18.2: CONVMIXER: FEATURE INTERACTIVE CONVOLUTION WITH CURRICULUM LEARNING FOR SMALL FOOTPRINT AND NOISY FAR-FIELD KEYWORD SPOTTING 3603
Dianwen Ng, Yunqi Chen, Eng Siong Chng, Nanyang Technological University, Singapore; Biao Tian, Qiang Fu, Alibaba Group, China

MLSP-18.3: SEMI-SUPERVISED SOURCE LOCALIZATION WITH RESIDUAL PHYSICAL LEARNING 3608
Michael Bianco, Peter Gerstoft, University of California San Diego, United States of America

MLSP-18.4: AUTOMATED PROSODY CLASSIFICATION FOR ORAL READING FLUENCY WITH QUADRATIC KAPPA LOSS AND ATTENTIVE X-VECTORS 3613
George Sammit, Zhongjie Wu, Yihao Wang, Zhongdi Wu, Akihito Kamata, Eric C. Larson, Southern Methodist University, United States of America; Joseph Nese, University of Oregon, United States of America

MLSP-18.5: SEED: SOUND EVENT EARLY DETECTION VIA EVIDENTIAL UNCERTAINTY 3618
Xujiang Zhao, Feng Chen, University of Texas at Dallas, United States of America; Xuchao Zhang, Wei Cheng, Wenchao Yu, Yuncong Chen, Haifeng Chen, NEC Laboratories America, United States of America

MLSP-18.6: RANK-BASED LOSS FOR LEARNING HIERARCHICAL REPRESENTATIONS 3623
Ines Nolasco, Queen Mary university of London, United Kingdom of Great Britain and Northern Ireland; Dan Stowell, Tilburg University, Netherlands

MLSP-19: TENSOR-BASED SIGNAL PROCESSING

MLSP-19.1: ON THE RELAXATION OF ORTHOGONAL TENSOR RANK AND ITS NONCONVEX RIEMANNIAN OPTIMIZATION FOR TENSOR COMPLETION 3628
Keisuke Ozawa, DENSO IT Laboratory, Japan

MLSP-19.2: ROBUST HIGH-ORDER TENSOR RECOVERY VIA NONCONVEX LOW-RANK APPROXIMATION 3633
Wenjin Qin, Jianjun Wang, Southwest University, School of Mathematic and Statistics, China; Hailin Wang, Xi'an Jiaotong University, School of Mathematics and Statistics, China; Weijun Ma, Ningxia University, School of Information Engineering, China; , ,

MLSP-19.3: VARIATIONAL BAYESIAN TENSOR NETWORKS WITH STRUCTURED POSTERIORS 3638
Kriton Konstantinidis, Yao Xu, Danilo Mandic, IMPERIAL COLLEGE LONDON, United Kingdom of Great Britain and Northern Ireland; Qibin Zhao, RIKEN CENTER FOR ADVANCED INTELLIGENCE PROJECT, Japan

MLSP-19.4: LOW-RANK PHASE RETRIEVAL WITH STRUCTURED TENSOR MODELS 3643
Soo Min Kwon, Xin Li, Anand Sarwate, Rutgers, The State University of New Jersey, United States of America

MLSP-19.5: HOQRI: HIGHER-ORDER QR ITERATION FOR SCALABLE TUCKER DECOMPOSITION 3648
Yuchen Sun, Kejun Huang, University of Florida, United States of America

MLSP-19.6: A MULTI-RESOLUTION LOW-RANK TENSOR DECOMPOSITION 3653
Sergio Rozada Doval, Antonio G. Marques, King Juan Carlos University, Spain

MLSP-20: DATA SCIENCE INITIATIVE I

MLSP-20.1: EXPLORING HETEROGENEOUS CHARACTERISTICS OF LAYERS IN ASR MODELS FOR MORE EFFICIENT TRAINING 3658
Lillian Zhou, Dhruv Guliani, Andreas Kabel, Giovanni Motta, Françoise Beaufays, Google LLC, United States of America

MLSP-20.2: PROBABILISTIC FINE-GRAINED URBAN FLOW INFERENCE WITH NORMALIZING FLOWS 3663
Ting Zhong, Haoyang Yu, Rongfan Li, Xovee Xu, Xucheng Luo, Fan Zhou, University of Electronic Science and Technology of China, China

MLSP-20.3: ATTENTION-BASED DUAL-STREAM VISION TRANSFORMER FOR RADAR GAIT RECOGNITION 3668
Shiliang Chen, Wentao He, Jianfeng Ren, University of Nottingham Ningbo China, China; Xudong Jiang, Nanyang Technological University, Singapore

MLSP-20.4: DEEP-MLE: FUSION BETWEEN A NEURAL NETWORK AND MLE FOR A SINGLE SNAPSHOT DOA ESTIMATION 3673
Marcio L. Lima de Oliveira, University of Twente, Netherlands; Marco Bekooij, NXP Semiconductors, Netherlands

MLSP-20.5: SELECTIVE MUTUAL LEARNING: AN EFFICIENT APPROACH FOR SINGLE CHANNEL SPEECH SEPARATION 3678
MinhTan Ha, Duc-Quang Vu, Chung-Ting Lee, Yung-Hui Li, Jia-Ching Wang, National Central University, Taiwan

MLSP-20.6: DETECTION OF COVID-19 FROM JOINT TIME AND FREQUENCY ANALYSIS OF SPEECH, BREATHING AND COUGH AUDIO 3683
John Harvill, Moitreya Chatterjee, Mark Hasegawa-Johnson, Narendra Ahuja, University of Illinois at Urbana-Champaign, United States of America; Yash Wani, Mustafa Alam, David Beiser, University of Chicago, United States of America; David Chestek, University of Illinois at Chicago, United States of America

MLSP-21: DEEP LEARNING FOR SPEECH AND AUDIO PROCESSING

MLSP-21.1: SELF-CRITICAL SEQUENCE TRAINING FOR AUTOMATIC SPEECH RECOGNITION 3688
Chen Chen, Yuchen Hu, Nana Hou, Heqing Zou, Eng Siong Chng, Nanyang Technological University, Singapore; Xiaofeng Qi, Hachibot, China

MLSP-21.2: FASTAUDIO: A LEARNABLE AUDIO FRONT-END FOR SPOOF SPEECH DETECTION 3693
Quchen Fu, Zhongwei Teng, Jules White, Douglas Schmidt, Vanderbilt University, United States of America; Maria Powell, Vanderbilt University Medical Center, United States of America

MLSP-21.3: COMPLEX IRM-AWARE TRAINING FOR VOICE ACTIVITY DETECTION USING ATTENTION MODEL 3698
Yifei Zhao, Yazid Attabi, Benoit Champagne, McGill University, Canada; Wei-Ping Zhu, Concordia University, Canada

MLSP-21.4: LEARNING CONTINUOUS REPRESENTATION OF AUDIO FOR ARBITRARY SCALE SUPER RESOLUTION 3703
Jaechang Kim, Yunjoo Lee, Jungseul Ok, Pohang University of Science and Technology, Korea, Republic of; Seunghoon Hong, Korea Advanced Institute of Science and Technology, Korea, Republic of

MLSP-21.5: AN INVESTIGATION OF THE EFFECTIVENESS OF PHASE FOR AUDIO CLASSIFICATION 3708
Shunsuke Hidaka, Kohei Wakamiya, Tokihiko Kaburagi, Kyushu University, Japan

MLSP-21.6: STUDY OF POSITIONAL ENCODING APPROACHES FOR AUDIO SPECTROGRAM TRANSFORMERS 3713
Leonardo Pepino, Pablo Riera, Luciana Ferrer, University of Buenos Aires, Argentina

MLSP-22: MACHINE LEARNING FOR IMAGE AND VIDEO PROCESSING

MLSP-22.1: FEW-SHOT OBJECT DETECTION WITH LOCAL CORRESPONDENCE RPN AND ATTENTIVE HEAD 3718
Jian Han, Yali Li, Shengjin Wang, tsinghua.edu.cn, China

MLSP-22.2: NATURAL-LOOKING ADVERSARIAL EXAMPLES FROM FREEHAND SKETCHES 3723
Hak Gu Kim, Chung-Ang University, Korea, Republic of; Davide Nanni, Sabine Süsstrunk, École Polytechnique Fédérale de Lausanne, Switzerland

MLSP-22.3: VIDEO ANOMALY DETECTION VIA PREDICTION NETWORK WITH ENHANCED SPATIO-TEMPORAL MEMORY EXCHANGE 3728
Guodong Shen, Yuqi Ouyang, Victor Sanchez, University of Warwick, United Kingdom of Great Britain and Northern Ireland

MLSP-22.4: SIGNAL COMPRESSION VIA NEURAL IMPLICIT REPRESENTATIONS 3733
Francesca Pistilli, Diego Valsesia, Giulia Fracastoro, Enrico Magli, Politecnico di Torino, Italy

MLSP-22.5: HYBRID WEIGHTING LOSS FOR PRECIPITATION NOWCASTING FROM RADAR IMAGES 3738
Yuan Cao, Hongming Shan, Fudan University, China; Lei Chen, Leiming Ma, Shanghai Central Meteorological Observatory, China; Danchen Zhang, University of Pittsburgh, China

MLSP-22.6: ADVERSARIAL LEARNING ENHANCEMENT FOR 3D HUMAN POSE AND SHAPE ESTIMATION 3743
Yidian Sun, Jiwei Zhang, Wendong Wang, Beijing University of Posts and Telecommunications, China

MLSP-23: TRANSFER LEARNING I

MLSP-23.1: DOMAIN GENERALIZED FEW-SHOT IMAGE CLASSIFICATION VIA META REGULARIZATION NETWORK 3748
Min Zhang, Zhejiang University, China; Siteng Huang, Donglin Wang, Westlake University, China

MLSP-23.2: GENERATION FOR UNSUPERVISED DOMAIN ADAPTATION: A GAN-BASED APPROACH FOR OBJECT CLASSIFICATION WITH 3D POINT CLOUD DATA 3753
Junxuan Huang, Junsong Yuan, Chunming Qiao, University at buffalo, United States of America

MLSP-23.3: EXPLORING TRANSFERABILITY MEASURES AND DOMAIN SELECTION IN CROSS-DOMAIN SLOT FILLING 3758
Xin-Chun Li, Yan-Jia Wang, Le Gan, De-Chuan Zhan, Nanjing University, China

MLSP-23.4: MAXIMUM BATCH FROBENIUS NORM FOR MULTI-DOMAIN TEXT CLASSIFICATION 3763
Yuan Wu, Ahmed El-Roby, Carleton University, Canada; Diana Inkpen, University of Ottawa, Canada

MLSP-23.5: JOINT GLOBAL-LOCAL ALIGNMENT FOR DOMAIN ADAPTIVE SEMANTIC SEGMENTATION 3768
Sudhir Yarram, Yuan Junsong, Chunming Qiao, University at Buffalo, United States of America; Ming Yang, Horizon Robotics, United States of America

MLSP-23.6: CATEGORY-ADAPTIVE DOMAIN ADAPTATION FOR SEMANTIC SEGMENTATION 3773
Zhiming Wang, Yantian Luo, Danlan Huang, Ning Ge, Jianhua Lu, Tsinghua University, China

MLSP-24: DEEP GENERATIVE MODEL

MLSP-24.1: SIMPLER IS BETTER: SPECTRAL REGULARIZATION AND UP-SAMPLING TECHNIQUES FOR VARIATIONAL AUTOENCODERS 3778
Sara Björk, UiT The Arctic University of Norway, Norway; Jonas Nordhaug Myhre, Thomas Haugland Johansen, NORCE Norwegian Research Centre, Norway

MLSP-24.2: AUGMENTING MOLECULAR DEEP GENERATIVE MODELS WITH TOPOLOGICAL DATA ANALYSIS REPRESENTATIONS 3783
Yair Schiff, Vijil Chenthamarakshan, Samuel Hoffman, Karthikeyan Natesan Ramamurthy, Payel Das, IBM, United States of America

MLSP-24.3: STYLEGAN-INDUCED DATA-DRIVEN REGULARIZATION FOR INVERSE PROBLEMS 3788
Arthur Conmy, Subhadip Mukherjee, Carola-Bibiane Schönlieb, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

MLSP-24.4: A CLOSER LOOK AT AUTOENCODERS FOR UNSUPERVISED ANOMALY DETECTION 3793
Oyebade Oyedotun, Djamila Aouada, University of Luxembourg, Luxembourg

MLSP-24.5: NFT-K: NON-FUNGIBLE TANGENT KERNELS 3798
Sina Alemohammad, Hossein Babaei, CJ Barberan, Naiming Liu, Lorenzo Luzi, Blake Mason, Richard Baraniuk, Rice University, United States of America

MLSP-25: MACHINE LEARNING FOR EARTH SCIENCES AND MATERIALS

MLSP-25.1: FDSNET: AN ACCURATE REAL-TIME SURFACE DEFECT SEGMENTATION NETWORK 3803
Jian Zhang, Runwei Ding, Miaoju Ban, Tianyu Guo, Key Laboratory of Machine Perception, Peking University, Shenzhen Graduate School, China

MLSP-25.2: PATH SIGNATURES FOR NON-INTRUSIVE LOAD MONITORING 3808
Paul Moore, Theodor-Mihai Iliant, Filip-Alexandru Ion, Terry Lyons, University of Oxford, United Kingdom of Great Britain and Northern Ireland; Yue Wu, University of Strathclyde, United Kingdom of Great Britain and Northern Ireland

MLSP-25.3: DATA-DRIVEN APPROACH FOR THE FLOQUET PROPAGATOR INVERSE PROBLEM SOLUTION 3813
Alexander Hvatov, ITMO University, Russian Federation

MLSP-25.4: CHUNKFUSION: A LEARNING-BASED RGB-D 3D RECONSTRUCTION FRAMEWORK VIA CHUNK-WISE INTEGRATION 3818
Chaozheng Guo, Lin Zhang, Ying Shen, Tongji University, China; Yicong Zhou, University of Macau, China

MLSP-25.5: CLOSING THE SIM-TO-REAL GAP IN GUIDED WAVE DAMAGE DETECTION WITH ADVERSARIAL TRAINING OF VARIATIONAL AUTO-ENCODERS 3823
Ishan Khurjekar, Joel Harley, University of Florida, United States of America

MLSP-25.6: DEEP LEARNING ON THE SPHERE FOR MULTI-MODEL ENSEMBLING OF SIGNIFICANT WAVE HEIGHT 3828
Andrea Littardi, Offshore Monitoring Ltd., Cyprus. The Cyprus Institute, Cyprus, Italy; Anders Hildeman, Chalmers University of Technology, Sweden., Sweden; Mihalis A. Nicolaou, The Cyprus Institute, Cyprus., Cyprus

MLSP-26: TRANSFER LEARNING II

MLSP-26.1: LOCAL AND GLOBAL ALIGNMENTS FOR GENERALIZABLE SENSOR-BASED HUMAN ACTIVITY RECOGNITION 3833
Wang Lu, Jindong Wang, Yiqiang Chen, Institute of Computing Technology, Chinese Academy of Sciences, China

MLSP-26.2: STUDY ON TIME-OF-FLIGHT ESTIMATION IN ULTRASONIC WELL LOGGING TOOL: MODEL-DRIVEN TRANSFER LEARNING 3838
Wei Zhang, Zhipeng Li, Yiduo Guo, Ao Qiu, Yanjun Li, Yibing Shi, School of Automation Engineering, University of Electronic Science and Technology of China, China

MLSP-26.3: SIMULATION-AND-MINING: TOWARDS ACCURATE SOURCE-FREE UNSUPERVISED DOMAIN ADAPTIVE OBJECT DETECTION 3843
Peng Yuan, Weijie Chen, Shicai Yang, Yunyi Xuan, Di Xie, Shiliang Pu, Hikvision Research Institute, China; Yueting Zhuang, Zhejiang University, China; , ,

MLSP-26.4: TARGET-AWARE AUTO-AUGMENTATION FOR UNSUPERVISED DOMAIN ADAPTIVE OBJECT DETECTION 3848
Zhaoyang Li, Long Zhao, Weijie Chen, Shicai Yang, Di Xie, Shiliang Pu, Hikvision Research Institute, China

MLSP-26.5: SELF-ENSEMBLE VARIANCE REGULARIZATION FOR DOMAIN ADAPTATION 3853
Xinyi Liu, Tao Dai, Shu-Tao Xia, Yong Jiang, Tsinghua University, China

MLSP-26.6: TRANSDUCTIVE CLIP WITH CLASS-CONDITIONAL CONTRASTIVE LEARNING 3858
Junchu Huang, South China University of Technology, China; Weijie Chen, Shicai Yang, Di Xie, Shiliang Pu, Hikvision Research Institute, China; Yueting Zhuang, Zhejiang University, China

MLSP-27: SUBSPACE AND MANIFOLD LEARNING

MLSP-27.1: CONTROLLING THE FRÉCHET VARIANCE IMPROVES BATCH NORMALIZATION ON THE SYMMETRIC POSITIVE DEFINITE MANIFOLD 3863
Reinmar Kobler, Jun-ichiro Hirayama, Motoaki Kawanabe, RIKEN, Japan

MLSP-27.2: SUBSPACE CLUSTERING USING UNSUPERVISED DATA AUGMENTATION 3868
Maryam Abdolali, Nicolas Gillis, University of Mons, Belgium

MLSP-27.3: PRIVATE LEARNING VIA KNOWLEDGE TRANSFER WITH HIGH-DIMENSIONAL TARGETS 3873
Dominik Fay, Tobias Oechtering, KTH Royal Institute of Technology, Sweden; Jens Sjölund, Uppsala University, Sweden

MLSP-27.4: DEEP DETERMINISTIC INDEPENDENT COMPONENT ANALYSIS FOR HYPERSPECTRAL UNMIXING 3878
Hongming Li, Jose Principe, University of Florida, United States of America; Shujian Yu, The Arctic University of Norway, Norway

MLSP-27.5: LABEL-AWARE RANKED LOSS FOR ROBUST PEOPLE COUNTING USING AUTOMOTIVE IN-CABIN RADAR 3883
Lorenzo Servadei, Huawei Sun, Julius Ott, Daniela Sanchéz Lopera, Infineon Technologies AG / Technical University of Munich, Germany; Michael Stephan, Thomas Stadelmayer, Infineon Technologies AG / Friedrich-Alexander-University of Erlangen Nuremberg, Germany; Souvik Hazra, Avik Santra, Infineon Technologies AG, Germany; Robert Wille, Johannes Kepler University Linz, Germany

MLSP-27.6: DEEPHULL: FAST CONVEX HULL APPROXIMATION IN HIGH DIMENSIONS 3888
Randall Balestriero, Facebook AI Research, United States of America; Zichao Wang, Richard Baraniuk, Rice University, United States of America

MLSP-28: MACHINE LEARNING FOR SOCIAL SCIENCES

MLSP-28.1: NEIGHBOR-AUGMENTED TRANSFORMER-BASED EMBEDDING FOR RETRIEVAL 3893
Jihai Zhang, Fangquan Lin, Wei Jiang, Cheng Yang, Alibaba Group, China; Gaoge Liu, Columbia University, United States of America

MLSP-28.2: SENTIMENT-AWARE DISTILLATION FOR BITCOIN TREND FORECASTING UNDER PARTIAL OBSERVABILITY 3898
Georgios Panagiotatos, Nikolaos Passalis, Avraam Tsantekidis, Anastasios Tefas, Aristotle University of Thessaloniki, Greece

MLSP-28.3: ROBUST NONPARAMETRIC DISTRIBUTION FORECAST WITH BACKTEST-BASED BOOTSTRAP AND ADAPTIVE RESIDUAL SELECTION 3903
Longshaokan Wang, Mina Georgieva, Paulo Machado, Abinaya Ulagappa, Safwan Ahmed, Yan Lu, Arjun Bakshi, Farhad Ghassemi, Amazon, United States of America; Lingda Wang, University of Illinois at Urbana-Champaign, United States of America

MLSP-28.4: VARIATIONAL BAYESIAN GRAPH CONVOLUTIONAL NETWORK FOR ROBUST COLLABORATIVE FILTERING 3908
Nozomu Onodera, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama, Hokkaido University, Japan

MLSP-28.5: FINT: FIELD-AWARE INTERACTION NEURAL NETWORK FOR CLICK-THROUGH RATE PREDICTION 3913
Zhishan Zhao, Guohui Liu, iQIYI Inc., China; Sen Yang, Dawei Feng, Kele Xu, National University of Defense Technology, China

MLSP-28.6: MAKING THE UNKNOWN MORE CERTAIN: A STACKED ENSEMBLE CLASSIFIER FOR OPEN GESTURE RECOGNITION WITH A SOCIAL ROBOT 3918
Heike Brock, Randy Gomez, Honda Research Institute Japan, Co Ltd., Japan

MLSP-29: SIGNAL PROCESSING THEORY AND METHODS

MLSP-29.1: APPLYING DIFFERENTIAL PRIVACY TO TENSOR COMPLETION 3923
Zheng Wei, Zhengpin Li, Jian Wang, Fudan University, China; Xiaojun Mao, Shanghai Jiao Tong University, China

MLSP-29.2: LOW-COMPLEXITY ATTENTION MODELLING VIA GRAPH TENSOR NETWORKS 3928
Yao Lei Xu, Kriton Konstantinidis, Shengxi Li, Danilo P. Mandic, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Ljubiša Stanković, University of Montenegro, Montenegro

MLSP-29.3: AN ACCELERATED RANK-(L,L,1,1) BLOCK TERM DECOMPOSITION OF MULTI-SUBJECT FMRI DATA UNDER SPATIAL ORTHONORMALITY CONSTRAINT 3933
Li-Dan Kuang, Biao Wang, Hao-Peng Zhang, Jianming Zhang, Wenjun Li, Feng Li, Changsha University of Science and Technology, China; Qiu-Hua Lin, Dalian University of Technology, China; Vince D. Calhoun, Georgia State University, China

MLSP-29.5: IMPROVING DYNAMIC GRAPH CONVOLUTIONAL NETWORK WITH FINE-GRAINED ATTENTION MECHANISM 3938
Bo Wu, Xun Liang, Xiangping Zheng, Yuhui Guo, Hui Tang, Renmin University of China, China

MLSP-29.6: ADAPID: AN ADAPTIVE PID OPTIMIZER FOR TRAINING DEEP NEURAL NETWORKS 3943
Boxi Weng, Jian Sun, Gang Wang, School of Automation, Beijing Institute of Technology, China; Alireza Sadeghi, Dept. of Electrical and Computer Engineering, University of Minnesota, United States of America

MLSP-30: DEEP LEARNING III

MLSP-30.1: MEMORY IN ECHO STATE NETWORKS AND THE CONTROLLABILITY MATRIX RANK 3948
Brian Whiteaker, Peter Gerstoft, UC San Diego and Scripps Inst. of Oceanography, United States of America

MLSP-30.2: OT CLEANER: LABEL CORRECTION AS OPTIMAL TRANSPORT 3953
Jun Xia, Zhejiang University, Westlake University, China; Cheng Tan, Lirong Wu, Yongjie Xu, Stan Z. Li, Westlake University, China

MLSP-30.3: DEMON: IMPROVED NEURAL NETWORK TRAINING WITH MOMENTUM DECAY 3958
John Chen, Cameron Wolfe, Anastasios Kyrillidis, Rice University, United States of America; Zhao Li, UT Health, United States of America

MLSP-30.4: DEPTH PRUNING WITH AUXILIARY NETWORKS FOR TINYML 3963
Josen Daniel De Leon, Rowel Atienza, University of the Philippines, Philippines

MLSP-30.5: GLASSOFORMER: A QUERY-SPARSE TRANSFORMER FOR POST-FAULT POWER GRID VOLTAGE PREDICTION 3968
Yunling Zheng, Carson Hu, Jack Xin, University of California, Irvine, United States of America; Guang Lin, Purdue University, United States of America; Meng Yue, Brookhaven National Laboratory, United States of America; Bao Wang, University of Utah, United States of America

MLSP-30.6: SAR-SHIPNET: SAR-SHIP DETECTION NEURAL NETWORK VIA BIDIRECTIONAL COORDINATE ATTENTION AND MULTI-RESOLUTION FEATURE FUSION 3973
Yuwen Deng, Donghai Guan, Weiwei Yuan, Jiemin Ji, Mingqiang Wei, Nanjing University of Aeronautics and Astronautics, China; Yanyu Chen, Fudan University, China

MLSP-31: MACHINE LEARNING FOR HEALTHCARE AND LIFE SCIENCES

MLSP-31.1: SPATIO-TEMPORAL PRRS EPIDEMIC FORECASTING VIA FACTORIZED DEEP GENERATIVE MODELING 3978
Mohammadsadegh Shamsabardeh, Beatriz Martínez-López, University of Califronia, Davis, United States of America; Bahar Azari, Northeastern University, United States of America

MLSP-31.2: FUSION-ID: A PHOTOPLETHYSMOGRAPHY AND MOTION SENSOR FUSION BIOMETRIC AUTHENTICATOR WITH FEW-SHOT ON-BOARDING 3983
Harshat Kumar, University of Pennsylvania, United States of America; Hojjat Seyed Mousavi, Behrooz Shahsavari, Apple Inc., United States of America

MLSP-31.3: DYNIMP: DYNAMIC IMPUTATION FOR WEARABLE SENSING DATA THROUGH SENSORY AND TEMPORAL RELATEDNESS 3988
Zepeng Huo, Taowei Ji, Yifei Liang, Xiaoning Qian, Bobak Mortazavi, Texas A&M University, United States of America; Shuai Huang, University of Washington, United States of America; Zhangyang Wang, University of Texas at Austin, United States of America

MLSP-31.4: INCREMENTAL CONTEXT AWARE ATTENTIVE KNOWLEDGE TRACING 3993
Cheryl Sze Yin Wong, Yang Guo, Nancy F. Chen, Savitha Ramasamy, Institute for Infocomm Research, Singapore

MLSP-31.5: ROBUST AND EFFICIENT UNCERTAINTY AWARE BIOSIGNAL CLASSIFICATION VIA EARLY EXIT ENSEMBLES 3998
Alexander Campbell, Lorena Qendro, Pietro Liò, Cecilia Mascolo, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

MLSP-31.6: TEMPORAL CROSS-GRAPH NETWORK FOR BRAIN FUNCTIONAL ACTIVITY PREDICTION 4003
Xinyu Yuan, Wenhan Wang, Youyong Kong, Jiasong Wu, Guanyu Yang, Huazhong Shu, School of Computer Science and Engineering, Southeast University, China

MLSP-32: REINFORCEMENT LEARNING I

MLSP-32.1: POPO: PESSIMISTIC OFFLINE POLICY OPTIMIZATION 4008
Qiang He, Xinwen Hou, Yu Liu, Institute of Automation, Chinese Academy of Sciences, China

MLSP-32.2: BYZANTINE-ROBUST FEDERATED DEEP DETERMINISTIC POLICY GRADIENT 4013
Qifeng Lin, Qing Ling, School of Computer Science and Engineering, Sun Yat-Sen University, China

MLSP-32.3: IMPROVING ACTOR-CRITIC REINFORCEMENT LEARNING VIA HAMILTONIAN MONTE CARLO METHOD 4018
Duo Xu, Faramarz Fekri, Georgia Institute of Technology, United States of America

MLSP-32.4: EFFICIENT AND STABLE INFORMATION DIRECTED EXPLORATION FOR CONTINUOUS REINFORCEMENT LEARNING 4023
Mingzhe Chen, Xi Xiao, Wanpeng Zhang, Tsinghua University, China; Xiaotian Gao, Microsoft Research Asia, China

MLSP-32.5: HYPERGRAPH-BASED REINFORCEMENT LEARNING FOR STOCK PORTFOLIO SELECTION 4028
Xiaojie Li, Chaoran Cui, Donglin Cao, Juan Du, Chunyun Zhang, Shandong University of Finance and Economics, China

MLSP-33: DEEP LEARNING IV

MLSP-33.1: MEMORY-BASED MESSAGE PASSING: DECOUPLING THE MESSAGE FOR PROPAGATION FROM DISCRIMINATION 4033
Jie Chen, Weiqi Liu, Jian Pu, Fudan University, China

MLSP-33.2: PEAR: PHOTOGRAPHIC EMBEDDING FOR AESTHETIC RATING 4038
Hao Wu, Jiangchao Yao, Alibaba Group, China; , ,

MLSP-33.3: GRADIENT-WEIGHTED CLASS ACTIVATION MAPPING FOR SPATIO TEMPORAL GRAPH CONVOLUTIONAL NETWORK 4043
Pratyusha Das, Antonio Ortega, University of Southern California, United States of America

MLSP-33.4: DEEP LEARNING BASED OFF-ANGLE IRIS RECOGNITION 4048
Ehsaneddin Jalilian, Georg Wimmer, Andreas Uhl, University of Salzburg, Austria; Mahmut Karakaya, Kennesaw State University, United States of America

MLSP-33.5: TOWARDS ROBUST VISUAL TRANSFORMER NETWORKS VIA K-SPARSE ATTENTION 4053
Sajjad Amini, Shahrokh Ghaemmaghami, Sharif University of Technology, Iran (Islamic Republic of)

MLSP-33.6: A GLOBAL TO LOCAL GUIDING NETWORK FOR MISSING DATA IMPUTATION 4058
Wei Wang, Yimeng Chai, Yue Li, nankai university, China

MLSP-34: MACHINE LEARNING FOR TELECOMMUNICATIONS

MLSP-34.1: LOCUNET: FAST URBAN POSITIONING USING RADIO MAPS AND DEEP LEARNING 4063
Çağkan Yapar, Giuseppe Caire, Technische Universität Berlin, Germany; Ron Levie, Gitta Kutyniok, Ludwig-Maximilians-Universität München, Germany

MLSP-34.2: LITEHAR: LIGHTWEIGHT HUMAN ACTIVITY RECOGNITION FROM WIFI SIGNALS WITH RANDOM CONVOLUTION KERNELS 4068
Hojjat Salehinejad, Shahrokh Valaee, University of Toronto, Canada

MLSP-34.3: CDX-NET: CROSS-DOMAIN MULTI-FEATURE FUSION MODELING VIA DEEP NEURAL NETWORKS FOR MULTIVARIATE TIME SERIES FORECASTING IN AIOPS 4073
Jiajia Li, Ling Dai, Bin Sheng, Shanghai Jiao Tong University, China; Feng Tan, Zikai Wang, Shanghai Artificial Intelligence Research Institute, China; Hui Shen, Shanghai Dingmao information technology Inc., China; Pengwei Hu, Merck China Innovation Hub, China

MLSP-34.4: A CLUSTERING-BASED ML SCHEME FOR CAPACITY APPROACHING SOFT LEVEL SENSING IN 3D TLC NAND 4078
Li-Wei Liu, Yen-Ching Liao, Hsie-Chia Chang, National Yang Ming Chiao Tung University, Taiwan

MLSP-34.5: DYNAMIC RESOURCE OPTIMIZATION FOR ADAPTIVE FEDERATED LEARNING EMPOWERED BY RECONFIGURABLE INTELLIGENT SURFACES 4083
Claudio Battiloro, Paolo Di Lorenzo, Sergio Barbarossa, Sapienza University of Rome, Italy; Mattia Merluzzi, CEA-Leti, Université Grenoble Alpes, France

MLSP-34.6: LEARNING-BASED RESOURCE ALLOCATION WITH DYNAMIC DATA RATE CONSTRAINTS 4088
Pourya Behmandpoor, Panagiotis Patrinos, Marc Moonen, KU Leuven, Belgium

MLSP-35: REINFORCEMENT LEARNING II

MLSP-35.1: DENOISING-ORIENTED DEEP HIERARCHICAL REINFORCEMENT LEARNING FOR NEXT-BASKET RECOMMENDATION 4093
Qihan Du, Li Yu, Huiyuan Li, Youfang Leng, Ningrui Ou, Renmin University of China, China

MLSP-35.2: COMPETITIVE MULTI-AGENT REINFORCEMENT LEARNING WITH SELF-SUPERVISED REPRESENTATION 4098
DiJia Su, Jason D. Lee, John M. Mulvey, H. Vincent Poor, Princeton University, United States of America

MLSP-35.3: MODEL-BASED ONLINE LEARNING FOR RESOURCE SHARING IN JOINT RADAR-COMMUNICATION SYSTEMS 4103
Petteri Pulkkinen, Aalto University, Saab Finland Oy, Finland; Visa Koivunen, Aalto University, Finland

MLSP-35.4: QRELATION: AN AGENT RELATION-BASED APPROACH FOR MULTI-AGENT REINFORCEMENT LEARNING VALUE FUNCTION FACTORIZATION 4108
Siqi Shen, Jun Liu, Mengwei Qiu, Weiquan Liu, Cheng Wang, Xiamen University, China; Yongquan Fu, Qinglin Wang, Peng Qiao, National University of Defense Technology, China, China

MLSP-35.5: DENOISING-GUIDED DEEP REINFORCEMENT LEARNING FOR SOCIAL RECOMMENDATION 4113
Qihan Du, Li Yu, Huiyuan Li, Youfang Leng, Ningrui Ou, Junyao Xiang, Renmin University of China, China

MLSP-36: APPLICATIONS OF MACHINE LEARNING

MLSP-36.1: AN EFFICIENT DP-SGD MECHANISM FOR LARGE SCALE NLU MODELS 4118
Christophe Dupuy, Radhika Arava, Rahul Gupta, Amazon Alexa AI, United States of America; Anna Rumshisky, Amazon Alexa AI; University of Massachusetts, Lowell, United States of America

MLSP-36.2: MAKD:MULTIPLE AUXILIARY KNOWLEDGE DISTILLATION 4123
Zehan Chen, Xuan Jin, Yuan He, Hui Xue, Alibaba Group, China

MLSP-36.3: FEATURE IMITATING NETWORKS 4128
Sari Saba-Sadiya, Mohammad Ghassemi, Michigan State University, United States of America; Tuka Alhanai, New York University Abu Dhabi, United Arab Emirates

MLSP-36.5: OVER-PARAMETERIZED NETWORK SOLVES PHASE RETRIEVAL EFFECTIVELY 4133
Ji Li, Chao Wang, National University of Singapore, Singapore

MLSP-36.6: DEEP SPATIO-TEMPORAL WIND POWER FORECASTING 4138
Jiangyuan Li, Mohammadreza Armandpour, Texas A&M University, United States of America

MLSP-37: KERNEL METHODS

MLSP-37.1: MULTIPLE KERNEL K-MEANS CLUSTERING WITH SIMULTANEOUS SPECTRAL ROTATION 4143
Jitao Lu, Yihang Lu, Rong Wang, Feiping Nie, Xuelong Li, Northwestern Polytechnical University, China

MLSP-37.2: MULTITASK GAUSSIAN PROCESS WITH HIERARCHICAL LATENT INTERACTIONS 4148
Kai Chen, Feng Yin, Shuguang Cui, The Chinese University of Hong Kong, Shenzhen, China; Twan van Laarhoven, Elena Marchiori, Radboud University, Nijmegen, Netherlands

MLSP-37.3: DISCRETE MULTI-KERNEL K-MEANS WITH DIVERSE AND OPTIMAL KERNEL LEARNING 4153
Yihang Lu, Jitao Lu, Rong Wang, Feiping Nie, Northwestern Polytechnical University, China

MLSP-37.4: ACCESS CONTROL FOR PRIVACY-PRESERVING GAUSSIAN PROCESS REGRESSION 4158
Takayuki Nakachi, University of the Ryukyus, Japan; Yitu Wang, Nippon Telegraph and Telephone Corporation, Japan

MLSP-37.5: SCALABLE RIDGE LEVERAGE SCORE SAMPLING FOR THE NYSTRÖM METHOD 4163
Farah Cherfaoui, Hachem Kadri, LIS, France; Liva Ralaivola, Criteo IA lab, Paris, France, France

MLSP-37.6: ON SUBMODULAR SET COVER PROBLEMS FOR NEAR-OPTIMAL ONLINE KERNEL BASIS SELECTION 4168
Hrusikesha Pradhan, IIT Kanpur, India; Alec Koppel, Amazon, United States of America; Ketan Rajawat, Indian Institute of Technology Kanpur, India

MLSP-38: SEQUENTIAL LEARNING

MLSP-38.1: IMPROVING FEATURE GENERALIZABILITY WITH MULTITASK LEARNING IN CLASS INCREMENTAL LEARNING 4173
Dong Ma, Chi Ian Tang, Cecilia Mascolo, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

MLSP-38.2: ADVERSARIAL MASK TRANSFORMER FOR SEQUENTIAL LEARNING 4178
Hou Lio, Shang-En Li, Jen-Tzung Chien, National Yang Ming Chiao Tung University, Taiwan

MLSP-38.3: ONLINE LEARNING WITH PROBABILISTIC FEEDBACK 4183
Pouya M. Ghari, Yanning Shen, University of California Irvine, United States of America

MLSP-38.4: DATA INCUBATION — SYNTHESIZING MISSING DATA FOR HANDWRITING RECOGNITION 4188
Jen-Hao Rick Chang, Martin Bresler, Youssouf Chherawala, Adrien Delaye, Thomas Deselaers, Ryan Dixon, Oncel Tuzel, Apple, United States of America

MLSP-38.5: TRACKING THE DIMENSIONS OF LATENT SPACES OF GAUSSIAN PROCESS LATENT VARIABLE MODELS 4193
Yuhao Liu, Petar M. Djuric, Stony Brook University, United States of America

MLSP-38.6: CONTROLLED SENSING AND ANOMALY DETECTION VIA SOFT ACTOR-CRITIC REINFORCEMENT LEARNING 4198
Chen Zhong, M. Cenk Gursoy, Senem Velipasalar, Syracuse University, United States of America

MLSP-39: DATA SCIENCE INITIATIVE II

MLSP-40: DEEP LEARNING V

MLSP-40.1: WIN THE LOTTERY TICKET VIA FOURIER ANALYSIS: FREQUENCIES GUIDED NETWORK PRUNING 4203
Yuzhang Shang, Bin Duan, Yan Yan, Illinois Institute of Technology, United States of America; Ziliang Zong, Texas State University, United States of America; Liqiang Nie, Shandong University, China; , ,

MLSP-40.2: SPARSEBFA: ATTACKING SPARSE DEEP NEURAL NETWORKS WITH THE WORST-CASE BIT FLIPS ON COORDINATES 4208
Kyungmi Lee, Anantha P. Chandrakasan, Massachusetts Institute of Technology, United States of America

MLSP-40.4: ADVERSARIAL EXAMPLES DETECTION BASED ON ERROR LEVEL ANALYSIS AND SPACE MAPPING 4213
Sizhao Huang, Shuai Wang, Jian Chen, Guozhi Li, School of Information and Communication Engineering University of Electronic Science and Technology of China;Yangtze Delta Region Institute of University of Electronic Science and Technology of China, China; Wenyi Wang, School of Information and Communication Engineering University of Electronic Science and Technology of China, China

MLSP-40.5: LEARNING MONOCULAR 3D HUMAN POSE ESTIMATION WITH SKELETAL INTERPOLATION 4218
Ziyi Chen, Shang-Hong Lai, National Tsing Hua University, Taiwan; Akihiro Sugimoto, National Institute of Informatics, Japan

MLSP-40.6: TRAINING STABLE GRAPH NEURAL NETWORKS THROUGH CONSTRAINED LEARNING 4223
Juan Cervino, Luana Ruiz, Alejandro Ribeiro, University of Pennsylvania, United States of America

MLSP-41: SUPERVISED LEARNING

MLSP-41.1: MISMATCHED SUPERVISED LEARNING 4228
Xun Xian, Mingyi Hong, Jie Ding, University of Minnesota, United States of America

MLSP-41.2: SUPERVISED TRAINING OF SIAMESE SPIKING NEURAL NETWORKS WITH EARTH MOVER'S DISTANCE 4233
Mateusz Pabian, Dominik Rzepka, Mirosław Pawlak, AGH University of Science and Technology, Poland

MLSP-41.3: ON THE EFFECTIVENESS OF ACTIVE LEARNING BY UNCERTAINTY SAMPLING IN CLASSIFICATION OF HIGH-DIMENSIONAL GAUSSIAN MIXTURE DATA 4238
Xiaoyi Mai, French National Centre for Scientific Research (CNRS), France; Salman Avestimehr, Antonio Ortega, Mahdi Soltanolkotabi, University of Southern California, United States of America

MLSP-41.4: NEURAL COLLAPSE IN DEEP HOMOGENEOUS CLASSIFIERS AND THE ROLE OF WEIGHT DECAY 4243
Akshay Rangamani, Andrzej Banburski-Fahey, Massachusetts Institute of Technology, United States of America

MLSP-41.5: SYNTHESIS OF ADVERSARIAL SAMPLES IN TWO-STAGE CLASSIFIERS 4248
Ismail Alkhouri, George Atia, University of Central Florida, United States of America; Alvaro Velasquez, Air Force Research Laboratory, United States of America

MLSP-41.6: SYNERGISTIC NETWORK LEARNING AND LABEL CORRECTION FOR NOISE-ROBUST IMAGE CLASSIFICATION 4253
Chen Gong, University of Washington, United States of America; Kong Bin, Xin Wang, Youbing Yin, Qi Song, Keya medical, United States of America; Eric Seibel, Univerisity of Washington, United States of America

MLSP-42: DISTRIBUTED AND FEDERATED LEARNING I

MLSP-42.1: SOCIAL WELFARE MAXIMIZATION IN CROSS-SILO FEDERATED LEARNING 4258
Jianan Chen, Qin Hu, Indiana University Purdue University Indianapolis, United States of America; Honglu Jiang, The University of Texas Rio Grande Valley, United States of America

MLSP-42.2: PRIVACY-PRESERVING DISTRIBUTED EXPECTATION MAXIMIZATION FOR GAUSSIAN MIXTURE MODEL USING SUBSPACE PERTURBATION 4263
Qiongxiu Li, Jaron Skovsted Gundersen, Katrine Tjell, Rafal Wisniewski, Mads Græsbøll Christensen, Aalborg university, Denmark

MLSP-42.3: A COMMUNICATION EFFICIENT QUASI-NEWTON METHOD FOR LARGE-SCALE DISTRIBUTED MULTI-AGENT OPTIMIZATION 4268
Yichuan Li, University of Illinois Urbana Champaign, United States of America; Petros Voulgaris, University of Nevada, Reno, United States of America; Nikolaos Freris, University of Science and Technology of China, China

MLSP-42.4: A BYZANTINE-RESILIENT DUAL SUBGRADIENT METHOD FOR VERTICAL FEDERATED LEARNING 4273
Kun Yuan, Alibaba Group (US), United States of America; Zhaoxian Wu, Qing Ling, Sun Yat-Sen University, China

MLSP-42.5: BYZANTINE-ROBUST AGGREGATION WITH GRADIENT DIFFERENCE COMPRESSION AND STOCHASTIC VARIANCE REDUCTION FOR FEDERATED LEARNING 4278
Heng Zhu, University of California, San Diego, United States of America; Qing Ling, Sun Yat-Sen University, China

MLSP-42.6: VARIANCE REDUCTION-BOOSTED BYZANTINE ROBUSTNESS IN DECENTRALIZED STOCHASTIC OPTIMIZATION 4283
Jie Peng, Qing Ling, Sun Yat-Sen University, China; Weiyu Li, Harvard Univeristy, China

MLSP-43: DEEP LEARNING FOR SPEECH AND LANGUAGE PROCESSING

MLSP-43.1: INTEGER-ONLY ZERO-SHOT QUANTIZATION FOR EFFICIENT SPEECH RECOGNITION 4288
Sehoon Kim, Amir Gholami, Zhewei Yao, Nicholas Lee, Patrick Wang, Aniruddha Nrusimha, Bohan Zhai, Tianren Gao, Michael Mahoney, Kurt Keutzer, University of California, Berkeley, United States of America

MLSP-43.3: NNSPEECH: SPEAKER-GUIDED CONDITIONAL VARIATIONAL AUTOENCODER FOR ZERO-SHOT MULTI-SPEAKER TEXT-TO-SPEECH 4293
Botao Zhao, Fudan University, China; Xulong Zhang, Jianzong Wang, Ning Cheng, Jing Xiao, Ping An Technology (Shenzhen) Co., Ltd., China

MLSP-43.4: NOISE-ROBUST SPEECH RECOGNITION WITH 10 MINUTES UNPARALLELED IN-DOMAIN DATA 4298
Chen Chen, Nana Hou, Yuchen Hu, Eng Siong Chng, Nanyang Technological University, Singapore; Shashank Shirol, Manipal Institute of Technology, India

MLSP-43.5: ENHANCING CLASS UNDERSTANDING VIA PROMPT-TUNING FOR ZERO-SHOT TEXT CLASSIFICATION 4303
Yuhao Dan, Qin Chen, Liang He, East China Normal University, China; Jie Zhou, Fudan Univerisity, China; Qingchun Bai, Shanghai Open University, China

MLSP-43.6: FILTERAUGMENT: AN ACOUSTIC ENVIRONMENTAL DATA AUGMENTATION METHOD 4308
Hyeonuk Nam, Seong-Hu Kim, Yong-Hwa Park, KAIST, Korea, Republic of

MLSP-44: LEARNING THEORY AND ALGORITHMS I

MLSP-44.1: THE REPRESENTATION JENSEN-RÉNYI DIVERGENCE 4313
Jhoan Keider Hoyos Osorio, Oscar Skean, Luis Gonzalo Sanchez Giraldo, University of Kentucky, United States of America; Austin J. Brockmeier, University of Delaware, United States of America

MLSP-44.3: MULTI-VIEW INFORMATION BOTTLENECK WITHOUT VARIATIONAL APPROXIMATION 4318
Qi Zhang, Jingmin Xin, Badong Chen, Xi'an Jiaotong University, China; Shujian Yu, UiT - The Arctic University of Norway, Norway

MLSP-44.4: TIME-FREQUENCY AND GEOMETRIC ANALYSIS OF TASK-DEPENDENT LEARNING IN RAW WAVEFORM BASED ACOUSTIC MODELS 4323
Devansh Gupta, Vinayak Abrol, Indraprastha Institute of Information Technology, Delhi, India

MLSP-44.6: CHANNEL REDUNDANCY AND OVERLAP IN CONVOLUTIONAL NEURAL NETWORKS WITH CHANNEL-WISE NNK GRAPHS 4328
David Bonet, Javier Ruiz-Hidalgo, Universitat Politècnica de Catalunya, Spain; Antonio Ortega, Sarath Shekkizhar, University of Southern California, United States of America

MLSP-45: DISTRIBUTED AND FEDERATED LEARNING II

MLSP-45.1: FEDCLEAN: A DEFENSE MECHANISM AGAINST PARAMETER POISONING ATTACKS IN FEDERATED LEARNING 4333
Abhishek Kumar, University of Oulu & University of Helsinki, Finland; Vivek Khimani, Drexel University, United States of America; Dimitris Chatzopoulos, University College Dublin, Ireland; Pan Hui, Hong Kong University of Science and Technology & University of Helsinki, Hong Kong

MLSP-45.2: A METHOD TO REVEAL SPEAKER IDENTITY IN DISTRIBUTED ASR TRAINING, AND HOW TO COUNTER IT 4338
Trung Dang, Peter Chin, Boston University, United States of America; Om Thakkar, Swaroop Ramaswamy, Rajiv Mathews, Françoise Beaufays, Google LLC, United States of America

MLSP-45.3: ON THE CONVERGENCE OF ADAM-TYPE ALGORITHMS FOR SOLVING STRUCTURED SINGLE NODE AND DECENTRALIZED MIN-MAX SADDLE POINT GAMES 4343
Babak Barazandeh, Kristal Curtis, Chandrima Sarkar, Splunk, United States of America; Ram Sriharsha, Pinecone, United States of America; George Michailidis, University of Florida, United States of America

MLSP-45.4: PARTIAL VARIABLE TRAINING FOR EFFICIENT ON-DEVICE FEDERATED LEARNING 4348
Tien-Ju Yang, Dhruv Guliani, Françoise Beaufays, Giovanni Motta, Google LLC, United States of America

MLSP-45.5: GRADIENT STALENESS IN ASYNCHRONOUS OPTIMIZATION UNDER RANDOM COMMUNICATION DELAYS 4353
Haider Al-Lawati, Stark Draper, University of Toronto, Canada

MLSP-45.6: TEMPO: IMPROVING TRAINING PERFORMANCE IN CROSS-SILO FEDERATED LEARNING 4358
Chen Ying, Baochun Li, University of Toronto, Canada; Bo Li, Hong Kong University of Science and Technology, Hong Kong

MLSP-46: SOURCE SEPARATION AND ICA

MLSP-46.1: DMANET: DEEP LEARNING-BASED DIFFERENTIAL MICROPHONE ARRAYS FOR MULTI-CHANNEL SPEECH SEPARATION 4363
Xiaokang Yang, Jianguo Wei, Tianjin University, China

MLSP-46.2: AMICABLE EXAMPLES FOR INFORMED SOURCE SEPARATION 4368
Naoya Takahashi, Yuki Mitsufuji, Sony Group Corporation, Japan

MLSP-46.3: REMIX-CYCLE-CONSISTENT LEARNING ON ADVERSARIALLY LEARNED SEPARATOR FOR ACCURATE AND STABLE UNSUPERVISED SPEECH SEPARATION 4373
Kohei Saijo, Tetsuji Ogawa, Waseda University, Japan

MLSP-46.4: AN INFORMATION MAXIMIZATION BASED BLIND SOURCE SEPARATION APPROACH FOR DEPENDENT AND INDEPENDENT SOURCES 4378
Alper Erdogan, Koc University, Turkey

MLSP-46.5: BLIND SEPARATION OF LINEAR-QUADRATIC MIXTURES OF MUTUALLY INDEPENDENT AND AUTOCORRELATED SOURCES 4383
Shahram Hosseini, Yannick Deville, Toulouse university, France

MLSP-46.6: LARGE-SCALE INDEPENDENT COMPONENT ANALYSIS BY SPEEDING UP LIE GROUP TECHNIQUES 4388
Matthias Hermann, Georg Umlauf, Matthias O. Franz, HTWG Konstanz, Germany

MLSP-47: DEEP LEARNING VI

MLSP-47.1: PREDICTING THE GENERALIZATION GAP IN DEEP MODELS USING ANCHORING 4393
Vivek Narayanaswamy, Andreas Spanias, Arizona State University, United States of America; Rushil Anirudh, Irene Kim, Yamen Mubarka, Jayaraman J Thiagarajan, Lawrence Livermore National Laboratory, United States of America

MLSP-47.2: WHEN DOES BACKDOOR ATTACK SUCCEED IN IMAGE RECONSTRUCTION? A STUDY OF HEURISTICS VS. BI-LEVEL SOLUTION 4398
Vardaan Taneja, Indian Institute of Technology, Delhi, India; Pin-Yu Chen, IBM Research, United States of America; Yuguang Yao, Sijia Liu, Michigan State University, United States of America

MLSP-47.3: MIXED KNOWLEDGE RELATION TRANSFORMER FOR IMAGE CAPTIONING 4403
Tianyu Chen, Zhixin Li, Jiahui Wei, Tiantao Xian, Guangxi Normal University, China

MLSP-47.4: BALANCED STRIPE-WISE PRUNING IN THE FILTER 4408
Zheng Huo, Chong Wang, Weiwei Chen, Yuqi Li, Ningbo University, China; Jun Wang, China University of Mining and Technology, China; Jiafei Wu, SenseTime Research, China

MLSP-47.5: GAN-BASED JOINT ACTIVITY DETECTION AND CHANNEL ESTIMATION FOR GRANT-FREE RANDOM ACCESS 4413
Shuang Liang, Yinan Zou, Shanghaitech University, Shanghai Institute of Microsystem and Information Technology and University of Chinese Academy of Sciences, China; Yong Zhou, Shanghaitech University, China

MLSP-48: LEARNING THEORY AND ALGORITHMS II

MLSP-48.1: CASCADING BANDIT UNDER DIFFERENTIAL PRIVACY 4418
Kun Wang, Shuai Li, Shanghai Jiao Tong University, China; Jing Dong, University of Michigan, China; Baoxiang Wang, The Chinese University of Hong Kong, Shenzhen, China

MLSP-48.2: ITERATIVE RE-WEIGHTED LEAST SQUARES ALGORITHMS FOR NON-NEGATIVE SPARSE AND GROUP-SPARSE RECOVERY 4423
Angshul Majumdar, Indraprastha Institute of Information Technology, India

MLSP-48.3: EXACT PARTITIONING OF HIGH-ORDER PLANTED MODELS WITH A TENSOR NUCLEAR NORM CONSTRAINT 4428
Chuyang Ke, Jean Honorio, Purdue University, United States of America

MLSP-48.4: NO MORE THAN 6FT APART: ROBUST K-MEANS VIA RADIUS UPPER BOUNDS 4433
Ahmed Imtiaz Humayun, Anastasios Kyrillidis, Richard Baraniuk, Rice University, United States of America; Randall Balestriero, Meta AI Research, United States of America

MLSP-48.5: DEEP KERNEL LEARNING NETWORKS WITH MULTIPLE LEARNING PATHS 4438
Ping Xu, Yue Wang, Xiang Chen, Zhi Tian, George Mason University, United States of America

MLSP-48.6: PROVABLE SAMPLE COMPLEXITY GUARANTEES FOR LEARNING OF CONTINUOUS-ACTION GRAPHICAL GAMES WITH NONPARAMETRIC UTILITIES 4443
Adarsh Barik, Jean Honorio, Purdue University, United States of America

MLSP-49: LEARNING FROM MULTIMODAL DATA

MLSP-49.1: CROSS-MODAL KNOWLEDGE DISTILLATION FOR VISION-TO-SENSOR ACTION RECOGNITION 4448
Jianyuan Ni, Anne H. H. Ngu, Texas State University, United States of America; Raunak Sarbajna, University of Houston, United States of America; Yang Liu, Sun Yat-sen University, China; Yan Yan, Illinois Institute of Technology, United States of America

MLSP-49.2: CLIPCAM: A SIMPLE BASELINE FOR ZERO-SHOT TEXT-GUIDED OBJECT AND ACTION LOCALIZATION 4453
Hsuan-An Hsia, Che-Hsien Lin, Bo-Han Kung, Jhao-Ting Chen, Daniel Stanley Tan, Jun-Cheng Chen, Academia Sinica, Taiwan; Kai-Lung Hua, National Taiwan University of Science and Technology, Taiwan

MLSP-49.3: EXPLORING DUAL STREAM GLOBAL INFORMATION FOR IMAGE CAPTIONING 4458
Tiantao Xian, Zhixin Li, Tianyu Chen, Guangxi Normal University, China; Huifang Ma, Northwest Normal University, China

MLSP-49.4: UNSUPERVISED CONTRASTIVE HASHING FOR CROSS-MODAL RETRIEVAL IN REMOTE SENSING 4463
Georgii Mikriukov, Mahdyar Ravanbakhsh, Begüm Demir, Technische Universität Berlin, Germany

MLSP-49.5: ROBUST THERMAL INFRARED PEDESTRIAN DETECTION BY ASSOCIATING VISIBLE PEDESTRIAN KNOWLEDGE 4468
Sungjune Park, Dae Hwi Choi, Jung Uk Kim, Yong Man Ro, Korea Advanced Institute of Science and Technology (KAIST), Korea, Republic of

MLSP-49.6: A GENERALIZED HIERARCHICAL NONNEGATIVE TENSOR DECOMPOSITION 4473
Joshua Vendrow, Deanna Needell, UCLA, United States of America; Jamie Haddock, Harvey Mudd College, United States of America

MLSP-50: DEEP LEARNING IV

MLSP-50.1: TWO STRATEGIES TOWARD LIGHTWEIGHT IMAGE SUPER-RESOLUTION 4478
Zongcai Du, Jie Liu, Jie Tang, Gangshan Wu, Nanjing University, China

MLSP-50.2: ON MINI-BATCH TRAINING WITH VARYING LENGTH TIME SERIES 4483
Brian Kenji Iwana, Kyushu University, Japan

MLSP-50.3: ACP: ADAPTIVE CHANNEL PRUNING FOR EFFICIENT NEURAL NETWORKS 4488
Yuan Zhang, Yuan Yuan, Qi Wang, Northwestern Polytechnical University, China

MLSP-50.4: BAYESIAN CONTINUAL IMPUTATION AND PREDICTION FOR IRREGULARLY SAMPLED TIME SERIES DATA 4493
*Yang Guo, Cheryl Sze Yin Wong, Savitha Ramasamy, Institute for Infocomm Research, A*STAR, Singapore; Jeanette Wen Jun Poh, Nanyang Technological University Singapore, Singapore*

MLSP-50.5: CONFIDENCE-AWARE MULTI-TEACHER KNOWLEDGE DISTILLATION 4498
Hailin Zhang, Defang Chen, Can Wang, Zhejiang University, China

MLSP-50.6: LEARNABLE HYPERGRAPH LAPLACIAN FOR HYPERGRAPH LEARNING 4503
Jiying Zhang, Yuzhao Chen, Xi Xiao, Runiu Lu, Shu-Tao Xia, Tsinghua University, China

MLSP-51: LEARNING THEORY AND ALGORITHMS III

MLSP-51.1: GRAPH LEARNING FROM MULTIVARIATE DEPENDENT TIME SERIES VIA A MULTI-ATTRIBUTE FORMULATION 4508
Jitendra Tugnait, Auburn University, United States of America

MLSP-51.2: GENERALIZED SLICED PROBABILITY METRICS 4513
Soheil Kolouri, Vanderbilt University, United States of America; Kimia Nadjahi, Telecom Paris, France; Shahin Shahrampour, Northeastern University, United States of America; Umut Simsekli, INRIA / ENS, France

MLSP-51.3: RECOVERY OF NOISY POOLED TESTS VIA LEARNED FACTOR GRAPHS WITH APPLICATION TO COVID-19 TESTING 4518
Eyal Fishel Ben-Knaan, Nir Shlezinger, Ben Gurion University, Israel; Yonina Eldar, Weizmann Institute of Science, Israel

MLSP-51.4: A REMEDY FOR DISTRIBUTIONAL SHIFTS THROUGH EXPECTED DOMAIN TRANSLATION 4523
Jean-Christophe Gagnon-Audet, Irina Rish, University of Montréal and Mila - Québec AI Institute, Canada; Soroosh Shahtalebi, Vector Institute for Artificial intelligence, Canada; Frank Rudzicz, University of Toronto and Vector Institute for Artificial intelligence, Canada

MLSP-51.5: DETERMINISTIC TRANSFORM BASED WEIGHT MATRICES FOR NEURAL NETWORKS 4528
Pol Grau Jurado, Saikat Chatterjee, KTH Royal Institute of Technology, Sweden; Xinyue Liang, Ocean University of China, China

MLSP-51.6: ADAPTIVE GROUP TESTING WITH MISMATCHED MODELS 4533
Mingzhou Fan, Byung-Jun Yoon, Edward R. Dougherty, Xiaoning Qian, Texas A&M University, United States of America; Francis Alexander, Brookhaven National Laboratory, United States of America

MLSP-52: MULTIMODAL DATA FUSION AND PROCESSING

MLSP-52.1: MIXED IN TIME AND MODALITY: CURSE OR BLESSING? CROSS-INSTANCE DATA AUGMENTATION FOR WEAKLY SUPERVISED MULTIMODAL TEMPORAL FUSION 4538
Yonggang Zhu, Beijing University of Posts and Telecommunications, Peking University, China; Chao Tian, China Xi'an Satellite Control Center, China; Zhuqing Jiang, Aidong Men, Haiying Wang, Beijing University of Posts and Telecommunications, China; Qingchao Chen, Peking University, China

MLSP-52.2: MTAF: SHOPPING GUIDE MICRO-VIDEOS POPULARITY PREDICTION USING MULTIMODAL AND TEMPORAL ATTENTION FUSION APPROACH 4543
Ningrui Ou, Li Yu, Huiyuan Li, Qihan Du, Junyao Xiang, Wei Gong, Renmin University of China, China

MLSP-52.3: LEARNING TO INTEGRATE VISION DATA INTO ROAD NETWORK DATA 4548
Oliver Stromann, Linköping University, Scania CV AB, Sweden; Alireza Razavi, Scania CV AB, Sweden; Michael Felsberg, Linköping University, Sweden

MLSP-52.4: HIERARCHICAL SIGNAL FUSION NETWORK FOR PULSAR DETECTION WITH PHASE-CORRELATION AND SIGNAL ATTENTIONS 4553
Huajian Wu, Mingmin Chi, Fudan University, China

MLSP-52.5: RECOGNITION OF SILENTLY SPOKEN WORD FROM EEG SIGNALS USING DENSE ATTENTION NETWORK (DAN). 4558
Sahil Datta, Akuha Aondoakaa, Jorunn Jo Holmberg, Elena Antonova, Brunel University London, United Kingdom of Great Britain and Northern Ireland

MLSP-53: MULTIMODAL ANALYSIS IN AUDIO APPLICATIONS

MLSP-53.1: WAV2CLIP: LEARNING ROBUST AUDIO REPRESENTATIONS FROM CLIP 4563
Ho-Hsiang Wu, Juan Bello, New York University, United States of America; Prem Seetharaman, Kundan Kumar, Descript, United States of America

MLSP-53.2: ASD-TRANSFORMER: EFFICIENT ACTIVE SPEAKER DETECTION USING SELF AND MULTIMODAL TRANSFORMERS 4568
Gourav Datta, University of Southern California, United States of America; Tyler Etchart, Vivek Yadav, Varsha Hedau, Pradeep Natarajan, Shih-Fu Chang, Amazon, United States of America

MLSP-53.3: MMLATCH: BOTTOM-UP TOP-DOWN FUSION FOR MULTIMODAL SENTIMENT ANALYSIS 4573
Georgios Paraskevopoulos, Efthymios Georgiou, Alexandros Potamianos, National Technical University of Athens, Greece

MLSP-53.4: MULTI-CHANNEL ATTENTIVE GRAPH CONVOLUTIONAL NETWORK WITH SENTIMENT FUSION FOR MULTIMODAL SENTIMENT ANALYSIS 4578
Luwei Xiao, Xingjiao Wu, Wen Wu, Jing Yang, Liang He, East China Normal University, China

MLSP-53.5: LEARNING MUSIC SEQUENCE REPRESENTATION FROM TEXT SUPERVISION 4583
Tianyu Chen, Shuai Zhang, Haoyi Zhou, Jianxin Li, Beihang University, China; Yuan Xie, The Institute of Acoustics of the Chinese Academy of Sciences, China; Shaohan Huang, Microsoft, China

MLSP-53.6: ENHANCING AFFECTIVE REPRESENTATIONS OF MUSIC-INDUCED EEG THROUGH MULTIMODAL SUPERVISION AND LATENT DOMAIN ADAPTATION 4588
Kleanthis Avramidis, University of Southern California, United States of America; Christos Garoufis, Athanasia Zlatintsi, Petros Maragos, National Technical University of Athens, Greece

MLSP-54: NEURAL AUDIO REPRESENTATION AND SYNTHESIS

MLSP-54.1: TOWARDS LEARNING UNIVERSAL AUDIO REPRESENTATIONS 4593
Luyu Wang, Pauline Luc, Yan Wu, Adria Recasens, Lucas Smaira, Andrew Brock, Andrew Jaegle, Jean-Baptiste Alayrac, Sander Dieleman, Joao Carreira, Aaron van den Oord, DeepMind, United Kingdom of Great Britain and Northern Ireland

MLSP-54.2: DIFFERENTIABLE WAVETABLE SYNTHESIS 4598
Siyuan Shan, University of North Carolina at Chapel Hill, United States of America; Lamtharn Hantrakul, Jitong Chen, Matt Avent, David Trevelyan, ByteDance, Thailand

MLSP-54.3: NEURAL AUDIO-TO-SCORE MUSIC TRANSCRIPTION FOR UNCONSTRAINED POLYPHONY USING COMPACT OUTPUT REPRESENTATIONS 4603
Víctor Arroyo, Jose J. Valero-Mas, Jorge Calvo-Zaragoza, Antonio Pertusa, University of Alicante, Spain

MLSP-54.4: END-TO-END MUSIC REMASTERING SYSTEM USING SELF-SUPERVISED AND ADVERSARIAL TRAINING 4608
Junghyun Koo, Seungryeol Paik, Kyogu Lee, Seoul National University, Korea, Republic of

MLSP-54.5: AVQVC: ONE-SHOT VOICE CONVERSION BY VECTOR QUANTIZATION WITH APPLYING CONTRASTIVE LEARNING 4613
Huaizhen Tang, University of Science and Technology of China, China; Xulong Zhang, Jianzong Wang, Ning Cheng, Jing Xiao, Ping An Technology (Shenzhen) Co., Ltd., China

MLSP-54.6: TOWARDS SPEAKER AGE ESTIMATION WITH LABEL DISTRIBUTION LEARNING 4618
Shijing Si, Jianzong Wang, Junqing Peng, Jing Xiao, Ping An Technology (Shenzhen) Co., Ltd., China

MMSP-1: MULTIMODAL SIGNAL PROCESSING, ANALYSIS, AND SYNTHESIS I

MMSP-1.1: DISTRIBUTED AUDIO-VISUAL PARSING BASED ON MULTIMODAL TRANSFORMER AND DEEP JOINT SOURCE CHANNEL CODING 4623
Penghong Wang, Xiaopeng Fan, Harbin Institute of Technology, China; Jiahui Li, Mengyao Ma, Huawei, China

MMSP-1.2: TALKINGFLOW: TALKING FACIAL LANDMARK GENERATION WITH MULTI-SCALE NORMALIZING FLOW NETWORK 4628
Sen Liang, Zhize Zhou, Hujun Bao, Zhejiang University, China; Rong Li, Zhejiang Lab, China; Juyong Zhang, University of Science and Technology of China, China

MMSP-1.3: INCORPORATING GAZE BEHAVIOR USING JOINT EMBEDDING WITH SCENE CONTEXT FOR DRIVER TAKEOVER DETECTION 4633
Yuning Qiu, Carlos Busso, University of Texas at Dallas, United States of America; Teruhisa Misu, Kumar Akash, Honda Research Institute USA, Inc., United States of America

MMSP-1.4: MULTI-VIEW AND MULTI-MODAL EVENT DETECTION UTILIZING TRANSFORMER-BASED MULTI-SENSOR FUSION 4638
Masahiro Yasuda, Yasunori Ohishi, Shoichiro Saito, Noboru Harada, NTT Corporation, Japan

MMSP-1.5: DISTRIBUTED LABEL DEQUANTIZED GAUSSIAN PROCESS LATENT VARIABLE MODEL FOR MULTI-VIEW DATA INTEGRATION 4643
Koshi Watanabe, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama, Hokkaido University, Japan

MMSP-1.6: CO-ATTENTION-GUIDED BILINEAR MODEL FOR ECHO-BASED DEPTH ESTIMATION 4648
Go Irie, Takashi Shibata, Akisato Kimura, NTT Corporation, Japan

MMSP-2: VIDEO PROCESSING, ANALYSIS, AND SYNTHESIS

MMSP-2.1: MODELING THE DETECTION CAPABILITY OF HIGH-SPEED SPIKING CAMERAS 4653
Junwei Zhao, Zhaofei Yu, Lei Ma, Ziluo Ding, Shiliang Zhang, Yonghong Tian, Tiejun Huang, Peking University, China

MMSP-2.2: MODERNN: TOWARDS FINE-GRAINED MOTION DETAILS FOR SPATIOTEMPORAL PREDICTIVE LEARNING 4658
Zenghao Chai, Zhengzhuo Xu, Chun Yuan, Tsinghua University, China

MMSP-2.3: GRAPH-BASED POINT CLOUD DENOISING USING SHAPE-AWARE CONSISTENCY FOR FREE-VIEWPOINT VIDEO 4663
Keisuke Nonaka, Ryosuke Watanabe, Haruhisa Kato, KDDI Research, Inc., Japan; Tatsuya Kobayashi, KDDI Research Inc., Japan; Eduardo Pavez, Antonio Ortega, University of Southern California, United States of America

MMSP-2.4: DCSN: DEFORMABLE CONVOLUTIONAL SEMANTIC SEGMENTATION NEURAL NETWORK FOR NON-RIGID SCENES 4668
Bor-Sheng Huang, Chih-Chung Hsu, Han-Yi Kao, Xian-Yun Wang, National Cheng Kung University, Taiwan; Wo-Ting Liao, National Pingtung University of science and technology, Taiwan

MMSP-2.5: TRANSFORMER-BASED DOMAIN ADAPTATION FOR EVENT DATA CLASSIFICATION 4673
Junwei Zhao, Shiliang Zhang, Tiejun Huang, Peking University, China

MMSP-3: EMOTION RECOGNITION

MMSP-3.1: MULTIMODAL EMOTION RECOGNITION WITH SURGICAL AND FABRIC MASKS 4678
Ziqing Yang, Houwei Cao, New York Institute of Technology, United States of America; Katherine Nayan, Rochester Institute of Technology, United States of America; Zehao Fan, New York University Tandon School of Engineering, United States of America

MMSP-3.2: HUMAN EMOTION RECOGNITION USING MULTI-MODAL BIOLOGICAL SIGNALS BASED ON TIME LAG-CONSIDERED CORRELATION MAXIMIZATION 4683
Yuya Moroto, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama, Hokkaido University, Japan

MMSP-3.3: MULTI-MODAL EMOTION RECOGNITION WITH SELF-GUIDED MODALITY CALIBRATION 4688
Mixiao Hou, Zheng Zhang, Guangming Lu, Harbin Institute of Technology, Shenzhen, China

MMSP-3.4: IS CROSS-ATTENTION PREFERABLE TO SELF-ATTENTION FOR MULTI-MODAL EMOTION RECOGNITION? 4693
Vandana Rajan, Andrea Cavallaro, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland; Alessio Brutti, Fondazione Bruno Kessler, Italy

MMSP-3.5: A PRE-TRAINED AUDIO-VISUAL TRANSFORMER FOR EMOTION RECOGNITION 4698
Minh Tran, Mohammad Soleymani, University of Southern California, United States of America

MMSP-3.6: MEMOBERT: PRE-TRAINING MODEL WITH PROMPT-BASED LEARNING FOR MULTIMODAL EMOTION RECOGNITION 4703
Jinming Zhao, Ruichen Li, Qin Jin, Renmin University of China, China; Xinchao Wang, Haizhou Li, National University of Singapore, Singapore

MMSP-4: OBJECT DETECTION AND TRACKING

MMSP-4.1: GLOBAL-LOCAL FEATURE ENHANCEMENT NETWORK FOR ROBUST OBJECT DETECTION USING MMWAVE RADAR AND CAMERA 4708
Kaikai Deng, Dong Zhao, Qiaoyue Han, Zihan Zhang, Shuyue Wang, Huadong Ma, Beijing University of Posts and Telecommunications, China

MMSP-4.2: LEARNING CORRELATION FOR ONLINE MULTIPLE OBJECT TRACKING 4713
Ying Wang, Chihui Zhuang, Haihui Ye, Yan Yan, Hanzi Wang, Xiamen University, China

MMSP-4.3: BOUNDING BOX DISTRIBUTION LEARNING AND CENTER POINT CALIBRATION FOR ROBUST VISUAL TRACKING 4718
Chihui Zhuang, Yan Yan, Yang Lu, Hanzi Wang, Xiamen University, China; Yanjie Liang, Xiamen University, Peng Cheng Laboratory, China

MMSP-4.4: MULTI-FOCUS GUIDED SEMANTIC AGGREGATION FOR VIDEO OBJECT DETECTION 4723
Haihui Ye, Guangge Wang, Yang Lu, Yan Yan, Hanzi Wang, Xiamen University, China

MMSP-5: MULTIMODAL SIGNAL PROCESSING, ANALYSIS, AND SYNTHESIS II

MMSP-5.1: ENHANCING CONTRASTIVE LEARNING WITH TEMPORAL COGNIZANCE FOR AUDIO-VISUAL REPRESENTATION GENERATION 4728
Chandrashekhar Lavania, Shiva Sundaram, Sundararajan Srinivasan, Katrin Kirchhoff, Amazon, United States of America

MMSP-5.2: CROSS-MODAL KNOWLEDGE DISTILLATION IN MULTI-MODAL FAKE NEWS DETECTION 4733
Zimian Wei, Hengyue Pan, Linbo Qiao, Xin Niu, Peijie Dong, Dongsheng Li, College of Computer, National University of Defense Technology, China

MMSP-5.3: TRAINING STRATEGIES FOR AUTOMATIC SONG WRITING: A UNIFIED FRAMEWORK PERSPECTIVE 4738
Tao Qian, Shuai Guo, Qin Jin, Renmin University of China, China; Jiatong Shi, Peter Wu, Carnegie Mellon University, United States of America

MMSP-5.4: RESIDUAL-GUIDED PERSONALIZED SPEECH SYNTHESIS BASED ON FACE IMAGE 4743
Jianrong Wang, Zixuan Wang, Xiaosheng Hu, Xuewei Li, Tianjin University, China; Qiang Fang, Chinese Academy of Social Sciences, China; Li Liu, the Chinese University of Hong Kong, China

MMSP-5.5: SKETCH STORYTELLING 4748
Yucheng Zhou, Fudan University, China

MMSP-5.6: MAG+: AN EXTENDED MULTIMODAL ADAPTATION GATE FOR MULTIMODAL SENTIMENT ANALYSIS 4753
Xianbing Zhao, Yixin Chen, Wanting Li, Buzhou Tang, Harbin Institute of Technology (Shenzhen), China; Lei Gao, University College London, United Kingdom of Great Britain and Northern Ireland

MMSP-6: MULTIMEDIA DATABASES

MMSP-6.1: IMAGE-TEXT ALIGNMENT AND RETRIEVAL USING LIGHT-WEIGHT TRANSFORMER 4758
Wenrui Li, Xiaopeng Fan, Harbin Institute of Technology, School of Computer Science & Technology, China

MMSP-6.2: A GENERAL FRAMEWORK FOR INCOMPLETE CROSS-MODAL RETRIEVAL WITH MISSING LABELS AND MISSING MODALITIES 4763
Mingyang Li, Shao-Lun Huang, Lin Zhang, Tsinghua University, China

MMSP-6.3: SUBGRAPH REPRESENTATION LEARNING WITH HARD NEGATIVE SAMPLES FOR INDUCTIVE LINK PREDICTION 4768
Heeyoung Kwak, Hyunkyung Bae, Kyomin Jung, Seoul National University, Korea, Republic of

MMSP-6.4: DEEP HASHING WITH HASH CENTER UPDATE FOR EFFICIENT IMAGE RETRIEVAL 4773
Abin Jose, Daniel Filbert, Christian Rohlfing, Jens-Rainer Ohm, RWTH Aachen University, Germany

MMSP-6.5: PROTOTYPE-BASED INTER-CAMERA LEARNING FOR PERSON RE-IDENTIFICATION 4778
Lin Wang, Wanqian Zhang, Dayan Wu, Pingting Hong, Bo Li, Institute of Information Engineering, Chinese Academy of Sciences, China

MMSP-6.6: DHWP: LEARNING HIGH-QUALITY SHORT HASH CODES VIA WEIGHT PRUNING 4783
Zeyu Ma, Guangming Lu, Harbin Institute of Technology, Shenzhen, China; Yuhang Guo, Xiao Luo, Minghua Deng, Peking University, China; Chong Chen, Alibaba Group, China; Wei Cheng, University of Chinese Academy of Sciences, China

MMSP-7: AUDIO/SPEECH/LANGUAGE ANALYSIS AND SYNTHESIS

MMSP-7.1: NODE SLICING BROAD LEARNING SYSTEM FOR TEXT CLASSIFICATION 4788
Fagui Liu, Xinjie Wu, Chao Li, South China University of Technology, China

MMSP-7.2: AUDIO-TEXT RETRIEVAL IN CONTEXT 4793
Siyu Lou, Xuenan Xu, Mengyue Wu, Kai Yu, Shanghai Jiao Tong University, China

MMSP-7.3: IMPROVED META LEARNING FOR LOW RESOURCE SPEECH 4798
RECOGNITION
Satwinder Singh, Ruili Wang, Feng Hou, Massey University, New Zealand

MMSP-7.4: QUANTIZED WINOGRAD ACCELERATION FOR CONV1D EQUIPPED ASR 4803
MODELS ON MOBILE DEVICES
Yiwu Yao, Chengyu Wang, Jun Huang, Alibaba Group, China

MMSP-7.5: ACOUSTIC-TO-ARTICULATORY INVERSION BASED ON SPEECH 4808
DECOMPOSITION AND AUXILIARY FEATURE
Jianrong Wang, Jinyu Liu, Longxuan Zhao, Shanyu Wang, Ruiguo Yu, Tianjin University, China; Li Liu, the Chinese University of Hong Kong, China

MMSP-7.6: AN AUDIO-SALIENCY MASKING TRANSFORMER FOR AUDIO EMOTION 4813
CLASSIFICATION IN MOVIES
Ya-Tse Wu, Jeng-Lin Li, Chi-Chun Lee, National Tsing Hua University, Taiwan

MMSP-8: JOINT IMAGE AND TEXT PROCESSING

MMSP-8.1: GENERATIVE ADVERSARIAL NETWORK INCLUDING REFERRING 4818
IMAGE SEGMENTATION FOR TEXT-GUIDED IMAGE MANIPULATION
Yuto Watanabe, Ren Togo, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama, Hokkaido University, Japan

MMSP-8.2: TEXT2POSTER: LAYING OUT STYLIZED TEXTS ON RETRIEVED 4823
IMAGES
Chuhao Jin, Hongteng Xu, Ruihua Song, Zhiwu Lu, Renmin University of China, China

MMSP-8.3: DEEP RANK CROSS-MODAL HASHING WITH SEMANTIC CONSISTENT 4828
FOR IMAGE-TEXT RETRIEVAL
Xiaoqing Liu, Huanqiang Zeng, Yifan Shi, Jianqing Zhu, Huaqiao University, China; Kai-Kuang Ma, Nanyang Technological University, China

MMSP-8.4: VQA-BC: ROBUST VISUAL QUESTION ANSWERING VIA 4833
BIDIRECTIONAL CHAINING
Mingrui Lao, Wei Chen, Nan Pu, Michael Lew, Leiden University, Netherlands; Yanming Guo, National University of Defense Technology, China

MMSP-8.5: TYPE-AWARE MEDICAL VISUAL QUESTION ANSWERING 4838
Anda Zhang, Wei Tao, Ziyan Li, Wenqiang Zhang, Academy for Engineering and Technology, Fudan University, China; Haofen Wang, College of Design and Innovation, Tongji University, China

MMSP-9: HUMAN CENTRIC MULTIMEDIA

MMSP-9.1: FROM BOTTOM-UP TO TOP-DOWN: CHARACTERIZATION OF 4843
TRAINING PROCESS IN GAZE MODELING
Ron Hecht, Ke Liu, Noa Garnett, Ariel Telpaz, Omer Tsimhoni, General Motors, Israel

MMSP-9.2: META TALK: LEARNING TO DATA-EFFICIENTLY GENERATE 4848
AUDIO-DRIVEN LIP-SYNCHRONIZED TALKING FACE WITH HIGH DEFINITION
Yuhan Zhang, Weihua He, Minglei Li, Kun Tian, Ziyang Zhang, Jie Cheng, Yaoyuan Wang, Jianxing Liao, Huawei Technologies Co., Ltd., China

MMSP-9.3: MAP: MULTISPECTRAL ADVERSARIAL PATCH TO ATTACK PERSON 4853
DETECTION
Taeheon Kim, Hong Joo Lee, Yong Man Ro, KAIST, Korea, Republic of

MMSP-9.4: GENRE-CONDITIONED LONG-TERM 3D DANCE GENERATION DRIVEN BY MUSIC 4858
Yuhang Huang, Junjie Zhang, Dan Zeng, Shanghai University, China; Shuyan Liu, University of Chinese Academy of Sciences, China; Qian Bao, Wu Liu, JD AI Research, China; Zhineng Chen, Fudan University, China

MMSP-9.5: LEARNING SOUND LOCALIZATION BETTER FROM SEMANTICALLY SIMILAR SAMPLES 4863
Arda Senocak, Hyeonggon Ryu, In So Kweon, KAIST, Korea, Republic of; Junsik Kim, Harvard University, Korea, Republic of

MMSP-9.6: BI-DIRECTIONAL MODALITY FUSION NETWORK FOR AUDIO-VISUAL EVENT LOCALIZATION 4868
Shuo Liu, Weize Quan, University of Chinese Academy of Sciences; NLPR, Institute of Automation, Chinese Academy of Sciences, China; Yuan Liu, Speech Lab, Alibaba Group, China; Dong-Ming Yan, NLPR, Institute of Automation, Chinese Academy of Sciences, China

MMSP-10: IMAGE/GRAPHICS PROCESSING, ANALYSIS, AND CODING

MMSP-10.1: DYNAMIC MULTI-SCALE LOSS BALANCE FOR OBJECT DETECTION 4873
Yihao Luo, Xiang Cao, Juntao Zhang, Tianjiang Wang, Qi Feng, Huazhong University of Science and Technology, China; Peng Cheng, Coolanyp Limited Liability Company, Kirkland, State of Washington, USA, United States of America

MMSP-10.2: LATENT SPACE SLICING FOR ENHANCED ENTROPY MODELING IN LEARNING-BASED POINT CLOUD GEOMETRY COMPRESSION 4878
Nicolas Frank, Davi Lazzarotto, Touradj Ebrahimi, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

MMSP-10.3: DAM-GAN : IMAGE INPAINTING USING DYNAMIC ATTENTION MAP BASED ON FAKE TEXTURE DETECTION 4883
Dongmin Cha, Daijin Kim, Pohang University of Science and Technology, Korea, Republic of

MMSP-10.4: IMPROVING REFERENCE-BASED IMAGE COLORIZATION FOR LINE ARTS VIA FEATURE AGGREGATION AND CONTRASTIVE LEARNING 4888
Shukai Wu, Qingqin Wang, Sanyuan Zhang, Zhejiang University, China; Shuchang Xu, Zhejiang University & Hangzhou Normal University & Wenzhou University, China

MMSP-10.5: FEW-SHOT GAZE ESTIMATION WITH MODEL OFFSET PREDICTORS 4893
Jiawei Ma, Shih-Fu Chang, Columbia University, United States of America; Xu Zhang, Yue Wu, Varsha Hedau, Amazon Alexa AI, United States of America

MMSP-10.6: ADVERSARIAL EXAMPLES FOR IMAGE CROPPING IN SOCIAL MEDIA 4898
Masatomo Yoshida, Masahiro Okuda, Doshisha University, Japan

SAM-1: BEAMFORMING

SAM-1.1: ROBUST ADAPTIVE BEAMFORMING BASED ON POWER METHOD PROCESSING AND SPATIAL SPECTRUM MATCHING 4903
Saeed Mohammadzadeh, Vitor H. Nascimento, University of São Paulo, Brazil; Rodrigo C. de Lamare, Pontifícia Universidade Católica do Rio de Janeiro, Brazil; Osman Kukrer, Eastern Mediterranean University, Cyprus

SAM-1.2: AN ADAPTIVE ORIENTATIONAL BEAMFORMING TECHNIQUE FOR NARROWBAND INTERFERENCE REJECTION 4908
Jiangyan Han, Boon Poh Ng, Meng Hwa Er, Nanyang Technological University, Singapore

SAM-1.3: PHASE-ONLY RECONFIGURABLE SPARSE ARRAY BEAMFORMING USING DEEP LEARNING 4913
Syed Ali Hamza, Widener University, United States of America; Moeness Amin, Villanova University, United States of America; Batu Chalise, New York Institute of Technology, United States of America

SAM-1.4: ROBUST ADAPTIVE BEAMFORMING MAXIMIZING THE WORST-CASE SINR OVER DISTRIBUTIONAL UNCERTAINTY SETS FOR RANDOM INC MATRIX AND SIGNAL STEERING VECTOR 4918
Yongwei Huang, Wenzheng Yang, Guangdong University of Technology, China; Sergiy A. Vorobyov, Aalto University, Finland

SAM-1.5: IMPROVED BEAMFORMING ENCODING FOR JOINT RADAR AND COMMUNICATION 4923
Tuomas Aittomaki, Visa Koivunen, Aalto University, Finland

SAM-1.6: STUDY OF THE NULL DIRECTIONS ON THE PERFORMANCE OF DIFFERENTIAL BEAMFORMERS 4928
Xuehan Wang, Jingdong Chen, Northwestern Polytechnical University, China; Israel Cohen, Technion - Israel Institute of Technology, Israel; Jacob Benesty, University of Quebec, Canada

SAM-2: DOA ESTIMATION AND LEARNING BASED METHODS

SAM-2.3: DOA M-ESTIMATION USING SPARSE BAYESIAN LEARNING 4933
Christoph Mecklenbraeuker, TU Wien, Austria; Peter Gerstoft, University of California, San Diego, United States of America; Esa Ollila, Aalto University, Finland

SAM-2.4: LEARNING-AIDED INITIALIZATION FOR VARIATIONAL BAYESIAN DOA ESTIMATION 4938
Yongsung Park, Florian Meyer, Peter Gerstoft, University of California, San Diego, United States of America

SAM-2.5: NEURAL NETWORK-BASED COMPRESSION FRAMEWORK FOR DOA ESTIMATION EXPLOITING DISTRIBUTED ARRAY 4943
Saidur Pavel, Yimin Zhang, Temple University, United States of America

SAM-3: WIDEBAND AND MICROPHONE ARRAY PROCESSING

SAM-3.1: T-SVD BASED BROADBAND NON-SYNCHRONOUS MEASUREMENTS 4948
Long Chen, Weize Sun, Lei Huang, Guitong Chen, Shenzhen University, China

SAM-3.2: LOW COMPLEX ACCURATE MULTI-SOURCE RTF ESTIMATION 4953
Changheng Li, Jorge Martinez, Richard Christian Hendriks, Delft University of Technology, Netherlands

SAM-3.3: MULTIPLE OFFSETS MULTILATERATION: A NEW PARADIGM FOR SENSOR NETWORK CALIBRATION WITH UNSYNCHRONIZED REFERENCE NODES 4958
Luca Ferranti, Jani Boutellier, University of Vaasa, Finland; Kalle Åström, Magnus Oskarsson, Lund University, Sweden; Juho Kannala, Aalto University, Finland

SAM-3.4: REFERENCE MICROPHONE SELECTION AND LOW-RANK APPROXIMATION BASED MULTICHANNEL WIENER FILTER WITH APPLICATION TO SPEECH RECOGNITION 4963
Xing-yu Chen, Jie Zhang, Li-rong Dai, University of Science and Technology of China, China

SAM-3.5: INCOHERENT SYNTHESIS OF SPARSE BROADBAND ARRAYS BASED ON A PARAMETER-FREE SUBSPACE CLUSTERING 4968
Guy Gubnitsky, University of Haifa, Israel; Yaakov Buchris, Israel Cohen, Technion–Israel Institute of Technology, Israel

SAM-3.6: INITIALIZATION-FREE IMPLICIT-FOCUSING (IF2) FOR WIDEBAND DIRECTION-OF-ARRIVAL ESTIMATION 4973
Jake Millhiser, Pulak Sarangi, Piya Pal, UC San Diego, United States of America

SAM-4: MIMO RADAR AND WAVEFORM DESIGN

SAM-4.1: RECURRENT DESIGN OF PROBING WAVEFORM FOR SPARSE BAYESIAN LEARNING BASED DOA ESTIMATION 4978
Linlong Wu, Bhavani Shankar, Ruizhi Hu, Björn Ottersten, University of Luxembourg, Luxembourg; Jisheng Dai, Jiangsu University, China

SAM-4.2: UNIMODULAR WAVEFORM DESIGN WITH LOW CORRELATION LEVELS: A FAST ALGORITHM DEVELOPMENT TO SUPPORT LARGE-SCALE CODE LENGTHS 4983
Yongzhe Li, Chunxuan Shi, Ran Tao, Beijing Institute of Technology, China

SAM-4.3: AIRBORNE MIMO RADAR TRANSMIT-RECEIVE DESIGN UNDER SPECTRAL CONSTRAINT IN SIGNAL-DEPENDENT CLUTTER 4988
Zhihui Li, Junpeng Shi, Dongming Wu, Shujie Shi, Qingsong Zhou, National University of Defense and Technology, China

SAM-4.5: WEAK TARGET DETECTION IN MASSIVE MIMO RADAR VIA AN IMPROVED REINFORCEMENT LEARNING APPROACH 4993
Weitong Zhai, Xiangrong Wang, School of Electronic and Information Engineering, Beihang University, Beijing, 100083, China, China; Maria S. Greco, Fulvio Gini, Department of Information Engineering, University of Pisa, Pisa, Italy, Italy

SAM-4.6: RIS-AIDED MONOSTATIC MIMO RADAR WITH CO-LOCATED ANTENNAS 4998
Stefano Buzzi, Emanuele Grossi, Luca Venturino, University of Cassino and Southern Lazio, Italy; Marco Lops, University of Naples Federico II, Italy

SAM-5: DIRECTION OF ARRIVAL ESTIMATION

SAM-5.1: CONVOLUTIONAL BEAMSPACE USING IIR FILTERS 5003
Po-Chih Chen, P. P. Vaidyanathan, California Institute of Technology, United States of America

SAM-5.2: RATIONAL ARRAYS FOR DOA ESTIMATION 5008
Pranav Kulkarni, P. P. Vaidyanathan, California Institute of Technology, United States of America

SAM-5.3: LOCALIZING MORE SOURCES THAN SENSORS IN PRESENCE OF COHERENT SOURCES 5013
Xinyao Chen, Zai Yang, Xi'an Jiaotong University, China

SAM-5.4: TWO-SNAPSHOT DOA ESTIMATION VIA HANKEL-STRUCTURED MATRIX COMPLETION 5018
Mohammad Bokaei, Saeed Razavikia, Arash Amini, Sharif University of Technology, Iran (Islamic Republic of); Stefano Rini, National Chao Tung University, Taiwan

SAM-5.5: A NOVEL ANGULAR ESTIMATION METHOD IN THE PRESENCE OF NONUNIFORM NOISE 5023
Majdoddin Esfandiari, Sergiy A. Vorobyov, Aalto University, Finland

SAM-5.6: PARTIALLY RELAXED ORTHOGONAL LEAST SQUARES WEIGHTED SUBSPACE FITTING DIRECTION-OF-ARRIVAL ESTIMATION 5028
David Schenck, Katja Lübbe, Minh Trinh-Hoang, Marius Pesavento, Technische Universität Darmstadt, Germany

SAM-6: CO-ARRAY BASED ARRAY PROCESSING

SAM-6.1: A NEW COPRIME-ARRAY-BASED CONFIGURATION WITH AUGMENTED DEGREES OF FREEDOM AND REDUCED MUTUAL COUPLING 5033
Nabil Mohsen, Ammar Hawbani, University of Science and Technology of China, China; Monika Agrawal, Indian Institute of Technology Delhi, India; Saeed Alsamhi, AthloneInstitute of Technology, Ireland; Liang Zhao, Shenyang Aerospace University, China

SAM-6.2: COARRAY MANIFOLD SEPARATION IN THE SPHERICAL HARMONICS DOMAIN FOR ENHANCED SOURCE LOCALIZATION 5038
Shekhar Kumar Yadav, Nithin V George, Indian Institute of Technology Gandhinagar, India

SAM-6.3: SPARSE ARRAY SOURCE ENUMERATION VIA COARRAY SUBSPACE OPTIMIZATION 5043
Chun-Lin Liu, National Taiwan University, Taiwan

SAM-6.4: THE PROTOTYPE CO-PRIME ARRAY WITH A ROBUST DIFFERENCE CO-ARRAY 5048
Ahmed M. A. Shaalan, Jun Du, University of Science and Technology of China, China

SAM-6.5: DOA ESTIMATION VIA COARRAY TENSOR COMPLETION WITH MISSING SLICES 5053
Hang Zheng, Chengwei Zhou, Zhejiang University, China; André L. F. de Almeida, Federal University of Ceará, Brazil; Yujie Gu, Aptiv, United States of America; Zhiguo Shi, Zhejiang University, International Joint Innovation Center, China

SAM-6.6: HALF INVERTED NESTED ARRAYS WITH LARGE HOLE-FREE FOURTH-ORDER DIFFERENCE CO-ARRAYS 5058
Yuan-Pon Chen, Chun-Lin Liu, National Taiwan University, Taiwan

SAM-7: TRACKING AND LOCALIZATION

SAM-7.1: SPHERICAL CONVOLUTIONAL RECURRENT NEURAL NETWORK FOR REAL-TIME SOUND SOURCE TRACKING 5063
Tianle Zhong, University of Electronic Science and Technology of China, China; Israel Mendoza Velazquez, Yi Ren, Yoichi Haneda, The University of Electro-Communications, Japan; Hector Manuel Perez Meana, National Polytechnic Institute of Mexico, Mexico

SAM-7.2: AUDIO-VISUAL TRACKING OF MULTIPLE SPEAKERS VIA A PMBM FILTER 5068
Jinzheng Zhao, Peipei Wu, Xubo Liu, Wenwu Wang, University of Surrey, United Kingdom of Great Britain and Northern Ireland; Yong Xu, Tencent, United States of America; Lyudmila Mihaylova, University of Sheffield, United Kingdom of Great Britain and Northern Ireland; Simon Godsill, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

SAM-7.3: FLOOR PLAN RECONSTRUCTION WITH HIGH-PRECISION RF-BASED TRACKING 5073
Guozhen Zhu, K. J. Ray Liu, University of Maryland, College Park, United States of America; Chenshu Wu, University of Hong Kong, China; Beibei Wang, Origin Wireless, Inc., United States of America

SAM-7.4: PARTIAL ARITHMETIC CONSENSUS BASED DISTRIBUTED INTENSITY PARTICLE FLOW SMC-PHD FILTER FOR MULTI-TARGET TRACKING 5078
Peipei Wu, Jinzheng Zhao, Shidrokh Goudarzi, Wenwu Wang, University of Surrey, United Kingdom of Great Britain and Northern Ireland

SAM-7.6: MULTI-MODAL RECURRENT FUSION FOR INDOOR LOCALIZATION 5083
Jianyuan Yu, Virginia Polytechnic Institute and State University, United States of America; Pu Wang, Toshiaki Koike-Akino, Philip Orlik, MITSUBISHI ELECTRIC RESEARCH LABORATORIES, United States of America

SAM-8: IMAGING AND ENVIRONMENT MONITORING

SAM-8.1: IMPROVING JOINT SPARSE HYPERSPECTRAL UNMIXING BY SIMULTANEOUSLY CLUSTERING PIXELS ACCORDING TO THEIR MIXTURES 5088
Seyyede Fatemeh Seyyedsalehi, Hamid Rabiee, Sharif University of Technology, Iran (Islamic Republic of)

SAM-8.2: GLOBAL EVOLUTION NEURAL NETWORK FOR SEGMENTATION OF REMOTE SENSING IMAGES 5093
Xinzhe Geng, Tao Lei, Qi Chen, Jian Su, Xi He, School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, China; Qi Wang, School of Computer Science and the Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, China; Asoke K. Nandi, Department of Electronic and Computer Engineering, Brunel University London, United Kingdom of Great Britain and Northern Ireland

SAM-8.3: SPECTRAL-SPATIAL SYMMETRICAL AGGREGATION CROSS-LINKING MULTI-MODAL DATA FUSION NETWORK 5098
Jinping Wang, Jun Li, Xiaojun Tan, Sun Yat-sen University, China

SAM-8.4: RELATION DISCOVERY IN NONLINEARLY RELATED LARGE-SCALE SETTINGS 5103
Ali Vosoughi, Adora DSouza, Anas Abidin, Axel Wismuller, University of Rochester, United States of America

SAM-8.5: ACOUSTIC IMAGING ABOARD THE INTERNATIONAL SPACE STATION (ISS): CHALLENGES AND PRELIMINARY RESULTS 5108
Luca Bondi, Charles Shelton, Samarjit Das, Bosch Research, United States of America; Gabriel Chuang, Adarsh Dave, Carnegie Mellon University, United States of America; Christopher Ick, New York University, United States of America; Brian Coltin, KBR, United States of America; Trey Smith, NASA Ames Research Center, United States of America

SAM-9: SOURCE LOCALIZATION

SAM-9.1: CONJUGATE AUGMENTED SPATIAL-TEMPORAL NEAR-FIELD SOURCES LOCALIZATION WITH CROSS ARRAY 5113
Zhiwei Jiang, Hua Chen, Ye Tian, Gang Wang, Ningbo University, China; Wei Liu, University of Sheffield, United Kingdom of Great Britain and Northern Ireland

SAM-9.2: PARAMETRIC MODELS FOR DOA TRAJECTORY LOCALIZATION 5118
Ruchi Pandey, Santosh Nannuru, IIIT Hyderabad, India

SAM-9.3: JOINT SOURCE LOCALIZATION AND ASSOCIATION THROUGH OVERCOMPLETE REPRESENTATION UNDER MULTIPATH PROPAGATION ENVIRONMENT 5123
Yuan Liu, Zhi-Wei Tan, Andy W. H. Khong, Nanyang Technological University, Singapore; Hongwei Liu, Xidian University, China

SAM-9.4: SEMIDEFINITE RELAXATION METHOD FOR MOVING OBJECT LOCALIZATION USING A STATIONARY TRANSMITTER AT UNKNOWN POSITION 5128
Ruichao Zheng, Gang Wang, Ningbo University, China; K. C. Ho, University of Missouri, United States of America; Lei Huang, Shenzhen University, China

SAM-9.5: UNDERDETERMINED TWO-DIMENSIONAL LOCALIZATION FOR WIDEBAND SOURCES BASED ON DISTRIBUTED SENSOR ARRAY NETWORKS 5133
Hantian Wu, Qing Shen, Yibao Liang, Beijing Institute of Technology, China; Wei Liu, University of Sheffield, United Kingdom of Great Britain and Northern Ireland

SAM-9.6: DIRECT LOCALIZATION: AN ISING MODEL APPROACH 5138
Shiva Akbari, Shahrokh Valaee, University of Toronto, Canada

SAM-10: DETECTION AND ESTIMATION

SAM-10.4: TRANSIENT DETECTION WITH UNKNOWN STATISTICS VIA SOURCE CODING 5143
Andrew Finelli, Peter Willett, Yaakov Bar-Shalom, Stefano Marano, University of Connecticut, United States of America

SAM-10.5: IDENTIFICATION OF PULSE STREAMS OF UNKNOWN SHAPE FROM TIME ENCODING MACHINE SAMPLES 5148
Meghna Kalra, Kiryung Lee, Ohio State University, United States of America; Yoram Bresler, University of Illinois at Urbana-Champaign, United States of America

SAM-10.6: EXACT SPARSE SUPER-RESOLUTION VIA MODEL AGGREGATION 5153
Hongqing Yu, Heng Qiao, Shanghai Jiao Tong University, China

SAM-11: ARRAY CALIBRATION AND PERFORMANCE ANALYSIS

SAM-11.2: A CRLB ANALYSIS OF AOA ESTIMATION USING BLUETOOTH 5 5158
Wentao Shi, Baoqi Huang, Kai Sun, Inner Mongolia University, China

SAM-11.3: CRAMER-RAO BOUND ANALYSIS OF DISTRIBUTED DOA ESTIMATION EXPLOITING MIXED-PRECISION COVARIANCE MATRIX 5163
Md. Waqeeb T. S. Chowdhury, Yimin Zhang, Temple University, United States of America

SAM-11.4: CRAMÉR-RAO BOUND AND ANTENNA SELECTION OPTIMIZATION FOR DUAL RADAR-COMMUNICATION DESIGN 5168
Zhaoyi Xu, Athina Petropulu, Rutgers, the State University of New Jersey, United States of America; Fan Liu, Southern University of Science and Technology, China

SAM-11.5: INFORMATION THEORETIC LIMITS FOR STANDARD AND ONE-BIT COMPRESSED SENSING WITH GRAPH-STRUCTURED SPARSITY 5173
Adarsh Barik, Jean Honorio, Purdue University, United States of America

SAM-12: MULTICHANNEL PROCESSING AND COMMUNICATIONS

SAM-12.3: THE DATA/IDENTITY TRADEOFF WITH CENSORED SENSORS 5178
Zachariah Sutton, Peter Willett, University of Connecticut, United States of America; Stefano Marano, University of Salerno, Italy

SAM-12.4: DOUBLE-RIS VERSUS SINGLE-RIS AIDED SYSTEMS: TENSOR-BASED MIMO CHANNEL ESTIMATION AND DESIGN PERSPECTIVES 5183
Khaled Ardah, Sepideh Gherekhloo, Martin Haardt, Communications Research Laboratory (CRL)/TU Ilmenau, Germany; André L. F. de Almeida, GTEL/Universidade Federal do Ceará (UFC), Brazil

SAM-12.5: EFFICIENT TWO-STAGE BEAM TRAINING AND CHANNEL ESTIMATION FOR RIS-AIDED MMWAVE SYSTEMS VIA FAST ALTERNATING LEAST SQUARES 5188
Hyeonjin Chung, Sunwoo Kim, Hanyang University, Korea, Republic of

SPCOM-1: MACHINE LEARNING FOR COMMUNICATION SYSTEMS

SPCOM-1.1: DEEP JOINT SOURCE-CHANNEL CODING FOR WIRELESS IMAGE TRANSMISSION WITH ADAPTIVE RATE CONTROL 5193
Mingyu Yang, Hun-Seok Kim, University of Michigan, United States of America

SPCOM-1.2: DEEP SEQUENTIAL BEAMFORMER LEARNING FOR MULTIPATH CHANNELS IN MMWAVE COMMUNICATION SYSTEMS 5198
Aditya Sant, University of California San Diego, United States of America; Afshin Abdi, Joseph Soriaga, Qualcomm Technologies Inc., United States of America

SPCOM-1.3: DATA-DRIVEN OPTIMIZATION FOR ZERO-DELAY LOSSY SOURCE CODING WITH SIDE INFORMATION 5203
Elad Domanovitz, Daniel Severo, Ashish Khisti, Wei Yu, University of Toronto, Canada

SPCOM-1.4: DISTRIBUTED IMAGE TRANSMISSION USING DEEP JOINT SOURCE-CHANNEL CODING 5208
Sixian Wang, Ke Yang, Jincheng Dai, Kai Niu, Beijing University of Posts and Telecommunications, China

SPCOM-1.5: ADAPTIVE WIRELESS POWER ALLOCATION WITH GRAPH NEURAL NETWORKS 5213
Navid NaderiAlizadeh, Alejandro Ribeiro, University of Pennsylvania, United States of America; Mark Eisen, Intel Corporation, United States of America

SPCOM-1.6: RESTLESS MULTI-ARMED BANDITS UNDER EXOGENOUS GLOBAL MARKOV PROCESS 5218
Tomer Gafni, Ben-Gurion university of the Negev, Israel; Michal Yemini, Princeton University, United States of America; Kobi Cohen, Ben-Gurion University of the Negev, Israel

SPCOM-2: DECENTRALIZED LEARNING IN COMMUNICATION SYSTEMS

SPCOM-2.1: BYZANTINE-ROBUST AND COMMUNICATION-EFFICIENT DISTRIBUTED NON-CONVEX LEARNING OVER NON-IID DATA 5223
Xuechao He, Sun Yat-Sen University, China; Heng Zhu, University of Science and Technology of China, China; Qing Ling, Sun Yat-sen University, China

SPCOM-2.2: COMMUNICATION-EFFICIENT ONLINE FEDERATED LEARNING FRAMEWORK FOR NONLINEAR REGRESSION 5228
Vinay Chakravarthi Gogineni, Stefan Werner, Norwegian University of Science and Technology-NTNU, Norway; Yih-Fang Huang, University of Notre Dame, United States of America; Anthony Kuh, University of Hawaii, United States of America

SPCOM-2.3: S2 REDUCER: HIGH-PERFORMANCE SPARSE COMMUNICATION TO ACCELERATE DISTRIBUTED DEEP LEARNING 5233
Keshi Ge, Yongquan Fu, Yiming Zhang, Zhiquan Lai, Xiaoge Deng, Dongsheng Li, National University of Defense Technology, China

SPCOM-2.4: A DATA-DRIVEN QUANTIZATION DESIGN FOR DISTRIBUTED TESTING AGAINST INDEPENDENCE WITH COMMUNICATION CONSTRAINTS 5238
Sebastian Espinosa, Jorge Silva, Universidad de Chile, Chile; Pablo Piantanida, CentraleSupélec CNRS Université Paris-Saclay, France; , ,

SPCOM-2.5: POWER ALLOCATION FOR WIRELESS FEDERATED LEARNING USING GRAPH NEURAL NETWORKS 5243
Boning Li, Santiago Segarra, Rice University, United States of America; Ananthram Swami, U.S. Army Combat Capabilities Development Command (DEVCOM) Army Research Laboratory, United States of America

SPCOM-2.6: FEDERATED MULTI-ARMED BANDIT VIA UNCOORDINATED EXPLORATION 5248
Zirui Yan, Quan Xiao, Tianyi Chen, Ali Tajer, Rensselaer Polytechnic institute, United States of America

SPCOM-2.7: BYZANTINE-RESILIENT DECENTRALIZED COLLABORATIVE LEARNING 5253
Jian Xu, Shao-Lun Huang, Tsinghua University, China

SPCOM-3: DETECTION AND ESTIMATION FOR COMMUNICATIONS

SPCOM-3.1: ADAPTIVE IDENTIFICATION OF UNDERWATER ACOUSTIC CHANNEL WITH A MIX OF STATIC AND TIME-VARYING PARAMETERS 5258
Maciej Niedzwiecki, Artur Gancza, Gdansk University of Technology, Poland; Lu Shen, Yuriy Zakharov, University of York, United Kingdom of Great Britain and Northern Ireland

SPCOM-3.2: ITERATIVE CHANNEL ESTIMATION AND DATA DETECTION ALGORITHM FOR OTFS MODULATION 5263
Rabah Ouchikh, Mustapha Djeddou, École Militaire Polytechnique, Algeria; Abdeldjalil Aissa-El-Bey, Thierry Chonavel, École nationale supérieure Mines-Télécom Atlantique Bretagne-Pays de la Loire, France

SPCOM-3.3: AN ASYMPTOTICALLY OPTIMAL APPROXIMATION OF THE CONDITIONAL MEAN CHANNEL ESTIMATOR BASED ON GAUSSIAN MIXTURE MODELS 5268
Michael Koller, Benedikt Fesl, Nurettin Turan, Wolfgang Utschick, Technical University Munich, Germany

SPCOM-3.4: LOW COMPLEXITY EQUALIZATION FOR AFDM IN DOUBLY DISPERSIVE CHANNELS 5273
Ali Bemani, Marios Kountouris, Eurecom, France; Nassar Ksairi, Huawei France R&D, France

SPCOM-3.5: CSI CLUSTERING WITH VARIATIONAL AUTOENCODING 5278
Michael Baur, Michael Würth, Michael Koller, Vlad-Costin Andrei, Wolfgang Utschick, Technical University of Munich, Germany

SPCOM-3.6: MASSIVE UNSOURCED RANDOM ACCESS BASED ON BILINEAR VECTOR APPROXIMATE MESSAGE PASSING 5283
Ramzi Ayachi, Mohamed Akrout, Volodymyr Shyianov, Faouzi Bellili, Amine Mezghani, University of Manitoba, Canada

SPCOM-4: RESOURCE ALLOCATION IN COMMUNICATION SYSTEMS

SPCOM-4.1: OPTIMAL QOS-AWARE NETWORK SLICING FOR SERVICE-ORIENTED NETWORKS WITH FLEXIBLE ROUTING 5288
Wei-Kun Chen, Beijing Institute of Technology, China; Ya-Feng Liu, Yu-Hong Dai, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China; Zhi-Quan Luo, Shenzhen Research Institute of Big Data and The Chinese University of Hong Kong, China

SPCOM-4.2: BYZANTINE-RESILIENT DECENTRALIZED RESOURCE ALLOCATION 5293
Runhua Wang, Qing Ling, Sun Yat-Sen University, China; Yaohua Liu, Nanjing University of Information Science and Technology, China

SPCOM-4.3: STABILITY ANALYSIS OF UNFOLDED WMMSE FOR POWER ALLOCATION 5298
Arindam Chowdhury, Fernando Gama, Santiago Segarra, Rice University, United States of America

SPCOM-4.4: MONOTONIC GENERALIZED NASH GAMES WITH APPLICATION TO THE MANAGEMENT OF ENERGY-AWARE ALOHA NETWORKS 5303
Wenbo Wang, Amir Leshem, Bar-Ilan University, Israel

SPCOM-4.5: DISTRIBUTED LINK SPARSIFICATION FOR SCALABLE SCHEDULING USING GRAPH NEURAL NETWORKS 5308
Zhongyuan Zhao, Santiago Segarra, Rice University, United States of America; Ananthram Swami, US Army's DEVCOM Army Research Laboratory, United States of America

SPCOM-4.6: A PERFORMANCE ANALYSIS FOR MULTI-RIS-ASSISTED FULL DUPLEX WIRELESS COMMUNICATION SYSTEM 5313
Farjam Karim, Bishmita Hazarika, Sandeep Kumar Singh, Keshav Singh, National Sun Yat-sen University, Taiwan

SPCOM-5: MIMO SIGNAL PROCESSING FOR WIRELESS SYSTEMS

SPCOM-5.1: JOINT BEAM SELECTION AND PRECODING BASED ON DIFFERENTIAL EVOLUTION FOR MILLIMETER-WAVE MASSIVE MIMO SYSTEMS 5318
Yang Liu, Yancheng Hou, Jiaxuan Wei, Yinghui Zhang, Junxing Zhang, Inner Mongolia University, China; Tiankui Zhang, Beijing University of Posts and Telecommunications, China

SPCOM-5.2: A NOVEL NEGATIVE L1 PENALTY APPROACH FOR MULTIUSER ONE-BIT MASSIVE MIMO DOWNLINK WITH PSK SIGNALING 5323
Zheyu Wu, Ya-Feng Liu, Yu-Hong Dai, Academy of Mathematics and Systems Science/Chinese Academy of Sciences, China; Bo Jiang, Nanjing Normal University, China

SPCOM-5.3: A SET-THEORETIC APPROACH TO MIMO DETECTION 5328
Jochen Fink, Renato Cavalcante, Zoran Utkovski, Slawomir Stanczak, Fraunhofer Heinrich-Hertz-Institute, Germany

SPCOM-5.4: DESIGNING A QAM SIGNAL DETECTOR FOR MASSIVE MIMO SYSTEMS VIA PS-ADMM APPROACH 5333
Quan Zhang, Xuyang Zhao, Jiangtao Wang, Yongchao Wang, Xidian University, China

SPCOM-5.5: POWER-EFFICIENT HYBRID MIMO RECEIVER WITH TASK-SPECIFIC BEAMFORMING USING LOW-RESOLUTION ADCS 5338
Timur Zirtiloglu, Rabia Tugce Yazicigil, Boston University, United States of America; Nir Shlezinger, Ben-Gurion University, Israel; Yonina Eldar, Weizmann Institute of Science, Israel

SPCOM-5.6: MIMO DETECTION BY VARIATIONAL POSTERIOR INFERENCE 5343
Junbin Liu, Mingjie Shao, Wing-Kin Ma, The Chinese University of Hong Kong, China

SPCOM-6: RECONFIGURABLE INTELLIGENT SURFACES, RELAYS, AND DRONES

SPCOM-6.1: CONTROLLING SMART PROPAGATION ENVIRONMENTS: LONG-TERM VERSUS SHORT-TERM PHASE SHIFT OPTIMIZATION 5348
Trinh Van Chien, Dinh-Hieu Tran, Hieu Van Nguyen, Symeon Chatzinotas, Björn Ottersten, University of Luxembourg, Luxembourg; Lam Thanh Tu, University of Poitiers, France; Marco Di Renzo, Université Paris-Saclay, CNRS, CentraleSupélec, France

SPCOM-6.2: DEEP ACTOR-CRITIC FOR CONTINUOUS 3D MOTION CONTROL IN MOBILE RELAY BEAMFORMING NETWORKS 5353
Spilios Evmorfos, Athina Petropulu, RUTGERS UNIVERSITY, United States of America

SPCOM-6.3: AERIAL BASE STATION PLACEMENT LEVERAGING RADIO TOMOGRAPHIC MAPS 5358
Daniel Romero, Viet Pham, University of Agder, Norway; Geert Leus, TU Delft, Netherlands

SPCOM-6.5: ATOMIC NORM BASED LOCALIZATION AND ORIENTATION ESTIMATION FOR MILLIMETER-WAVE MIMO OFDM SYSTEMS 5363
Jianxiu Li, Maxime Ferreira Da Costa, Urbashi Mitra, University of Southern California, United States of America

SPCOM-6.6: ESTIMATION OF CHANNELS IN SYSTEMS WITH INTELLIGENT REFLECTING SURFACES 5368
Michael Joham, Hangze Gao, Wolfgang Utschick, Technical University of Munich, Germany

SPCOM-7: CELL-FREE AND DISTRIBUTED ANTENNA SYSTEMS

SPCOM-7.1: DISTRIBUTED HYBRID BEAMFORMING FOR MMWAVE CELL-FREE MASSIVE MIMO 5373
Nuan Song, Tao Yang, Nokia Bell Labs China, China

SPCOM-7.2: QUANTIZATION-AWARE PRECODING FOR MU-MIMO WITH LIMITED-CAPACITY FRONTHAUL 5378
Yasaman Khorsandmanesh, Emil Björnson, Joakim Jaldén, KTH Royal Institute of Technology, Sweden

SPCOM-7.3: AN ONLINE THROUGHPUT MAXIMIZATION ALGORITHM FOR GREEN COORDINATED MULTI-POINT SYSTEMS 5383
Yanjie Dong, Jianqiang Li, Shenzhen University, China; Haijun Zhang, University of Science and Technology Beijing, China; F. Richard Yu, Carleton University, Canada; Song Guo, The Hong Kong Polytechnic University, Hong Kong; Victor C. M. Leung, Shenzhen University/The University of British Columbia, China

SPCOM-7.4: EFFICIENTLY AND GLOBALLY SOLVING JOINT BEAMFORMING AND COMPRESSION PROBLEM IN THE COOPERATIVE CELLULAR NETWORK VIA LAGRANGIAN DUALITY 5388
Xilai Fan, Ya-Feng Liu, Academy of Mathematics and Systems Science/Chinese Academy of Sciences, China; Liang Liu, The Hong Kong Polytechnic University, China

SPCOM-7.5: CELL-FREE MASSIVE MIMO: EXPLOITING THE WAX DECOMPOSITION 5393
Juan Vidal Alegría, Fredrik Rusek, Lund University, Sweden; Jinliang Huang, Huawei Technologies AB, Sweden

SPTM-1: SAMPLING, MULTIRATE, AND DIGITAL SIGNAL PROCESSING I

SPTM-1.1: LEARNING STRUCTURED SPARSITY FOR TIME-FREQUENCY RECONSTRUCTION 5398
Lei Jiang, Haijian Zhang, Lei Yu, Wuhan University, China

SPTM-1.2: UNLIMITED SAMPLING WITH SPARSE OUTLIERS: EXPERIMENTS WITH IMPULSIVE AND JUMP OR RESET NOISE 5403
Ayush Bhandari, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SPTM-1.3: LEARNING APPROACH FOR FAST APPROXIMATE MATRIX FACTORIZATIONS 5408
Haiyan Yu, Zhen Qin, Zhihui Zhu, University of Denver, United States of America

SPTM-1.4: PARAMETER ESTIMATION IN SPARSE INVERSE PROBLEMS USING BERNOULLI-GAUSSIAN PRIOR 5413
Pierre Barbault, Matthieu Kowalski, Universite Paris Saclay, France; Charles Soussen, CentraleSupelec, France

SPTM-1.5: SPARSE RECOVERY OF ACOUSTIC WAVES 5418
Mohamed Mansour, Amazon, United States of America

SPTM-1.6: NONLINEAR SIGNAL DECOMPOSITION BASED ON BLOCK SPARSE APPROXIMATION 5423
El Hadji Diop, University Iba Der Thiam of Thies, Senegal; Karl Skretting, University of Stavanger, Senegal

SPTM-2: OPTIMIZATION METHODS FOR SIGNAL PROCESSING I

SPTM-2.1: BLOCK-ACTIVATED ALGORITHMS FOR MULTICOMPONENT FULLY NONSMOOTH MINIMIZATION 5428
Minh N. Bùi, Patrick L. Combettes, Zev C. Woodstock, North Carolina State University, United States of America

SPTM-2.2: BLOCK-COORDINATE FRANK-WOLFE ALGORITHM AND CONVERGENCE ANALYSIS FOR SEMI-RELAXED OPTIMAL TRANSPORT PROBLEM 5433
Takumi Fukunaga, Hiroyuki Kasai, Waseda University, Japan

SPTM-2.3: AN IMPLICIT GRADIENT-TYPE METHOD FOR LINEARLY CONSTRAINED BILEVEL PROBLEMS 5438
Ioannis Tsaknakis, Prashant Khanduri, Mingyi Hong, University of Minnesota Twin Cities, United States of America

SPTM-2.4: SCREEN & RELAX: ACCELERATING THE RESOLUTION OF ELASTIC-NET BY SAFE IDENTIFICATION OF THE SOLUTION SUPPORT 5443
Théo Guyard, Univ Rennes, INSA Rennes, CNRS, IRMAR-UMR 6625, F-35000 Rennes, France, France; Cédric Herzet, INRIA Rennes-Bretagne Atlantique, Campus de Beaulieu, 35000 Rennes, France, France; Clément Elvira, SCEE/IETR UMR CNRS 6164, CentraleSupélec, 35510 Cesson Sévigné, France, France

SPTM-2.5: NODE-SCREENING TESTS FOR THE L0-PENALIZED LEAST-SQUARES PROBLEM 5448
Théo Guyard, Univ Rennes, INSA Rennes, CNRS, IRMAR-UMR 6625, F-35000 Rennes, France, France; Cédric Herzet, INRIA Rennes-Bretagne Atlantique, Campus de Beaulieu, 35000 Rennes, France, France; Clément Elvira, SCEE/IETR UMR CNRS 6164, CentraleSupélec, 35510 Cesson Sévigné, France, France

SPTM-2.6: PROXIMAL-BASED ADAPTIVE SIMULATED ANNEALING FOR GLOBAL OPTIMIZATION 5453
Thomas Guilmeau, Emilie Chouzenoux, Université Paris-Saclay, CentraleSupélec, INRIA, France; Víctor Elvira, University of Edinburgh, United Kingdom of Great Britain and Northern Ireland

SPTM-3: GRAPH REPRESENTATION AND ANALYSIS I

SPTM-3.1: GRAPHON-AIDED JOINT ESTIMATION OF MULTIPLE GRAPHS 5458
Madeline Navarro, Santiago Segarra, William Marsh Rice University, United States of America

SPTM-3.2: EXPLORING DEEPER GRAPH CONVOLUTIONS FOR SEMI-SUPERVISED NODE CLASSIFICATION 5463
Ashish Tiwari, Shanmuganathan Raman, IIT Gandhinagar, India; Richeek Das, IIT Bombay, India

SPTM-3.3: DYNAMIC PORTFOLIO CUTS: A SPECTRAL APPROACH TO GRAPH-THEORETIC DIVERSIFICATION 5468
Alvaro Arroyo, Bruno Scalzo, Danilo Mandic, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Ljubisa Stankovic, University of Montenegro, United Kingdom of Great Britain and Northern Ireland

SPTM-3.4: STABILITY OF NEURAL NETWORKS ON MANIFOLDS TO RELATIVE PERTURBATIONS 5473
Zhiyang Wang, Luana Ruiz, Alejandro Ribeiro, University of Pennsylvania, United States of America

SPTM-3.5: ADA-STNET: A DYNAMIC ADABOOST SPATIO-TEMPORAL NETWORK FOR TRAFFIC FLOW PREDICTION 5478
Jiawei Sun, Jie Li, Chentao Wu, Zili Tang, Shanghai Jiao Tong University, China; Celimuge Wu, The University of Electro-Communications, Japan

SPTM-3.6: LABEL PROPAGATION ACROSS GRAPHS: NODE CLASSIFICATION USING GRAPH NEURAL TANGENT KERNELS 5483
Artun Bayer, Arindam Chowdhury, Santiago Segarra, Rice University, United States of America

SPTM-4: BAYESIAN SIGNAL PROCESSING

SPTM-4.1: DISTRIBUTED PARTICLE FILTERS FOR STATE TRACKING ON THE STIEFEL MANIFOLD USING TANGENT SPACE STATISTICS 5488
Claudio Bordin, Universidade Federal do ABC, Brazil; Caio de Figueredo, Escola Naval, Brazil; Marcelo Bruno, Instituto Tecnologico de Aeronautica, Brazil

SPTM-4.2: HUMAN DECISION MAKING WITH BOUNDED RATIONALITY 5493
Baocheng Geng, University of Alabama at Birmingham, United States of America; Qunwei Li, Ant Group, China; Pramod Varshney, Syracuse University, United States of America

SPTM-4.3: UNROLLING PARTICLES: UNSUPERVISED LEARNING OF SAMPLING DISTRIBUTIONS 5498
Fernando Gama, Nicolas Zilberstein, Richard Baraniuk, Santiago Segarra, Rice University, United States of America

SPTM-4.4: SCALABLE DATA ASSOCIATION AND MULTI-TARGET TRACKING UNDER A POISSON MIXTURE MEASUREMENT PROCESS 5503
Qing Li, Jiaming Liang, Simon Godsill, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

SPTM-4.5: ONLINE LEARNING FOR LATENT YULE-SIMON PROCESSES 5508
Asher Hensley, Petar Djuric, Stony Brook University, United States of America

SPTM-5: SAMPLING, MULTIRATE, AND DIGITAL SIGNAL PROCESSING II

SPTM-5.1: COUNTING THE NUMBER OF DIFFERENT SCALING EXPONENTS IN MULTIVARIATE SCALE-FREE DYNAMICS: CLUSTERING BY BOOTSTRAP IN THE WAVELET DOMAIN 5513
Charles-Gérard Lucas, ENS de Lyon, France; Patrice Abry, CNRS, ENS de Lyon, France; Herwig Wendt, CNRS, Université de Toulouse, France; Gustavo Didier, Tulane University, United States of America

SPTM-5.2: ON THE ACQUISITION OF STATIONARY SIGNALS USING UNIFORM ADCS 5518
Peter Neuhaus, Meik Dörpinghaus, Gerhard Fettweis, Technische Universität Dresden, Germany; Nir Shlezinger, Ben-Gurion University of the Negev, Israel; Yonina Eldar, Weizmann Institute of Science, Israel

SPTM-5.3: DATA-DRIVEN ALGORITHMS FOR GAUSSIAN MEASUREMENT MATRIX DESIGN IN COMPRESSIVE SENSING 5523
Yang Sun, Jonathan Scarlett, National University of Singapore, Singapore

SPTM-5.4: SCATTERING STATISTICS OF GENERALIZED SPATIAL POISSON POINT PROCESSES 5528
Michael Perlmutter, UCLA, United States of America; Jieqian He, Matthew Hirn, Michigan State University, United States of America

SPTM-5.5: REGULARIZATION USING DENOISING: EXACT AND ROBUST SIGNAL RECOVERY 5533
Ruturaj Gavaskar, Kunal Chaudhury, Indian Institute of Science, India

SPTM-5.6: GRAPH-STRUCTURED SPARSE REGULARIZATION VIA CONVEX OPTIMIZATION 5538
Hiroki Kuroda, Daichi Kitahara, Ritsumeikan University, Japan

SPTM-6: OPTIMIZATION METHODS FOR SIGNAL PROCESSING II

SPTM-6.1: DECENTRALIZED BILEVEL OPTIMIZATION FOR PERSONALIZED CLIENT LEARNING 5543
Songtao Lu, Xiaodong Cui, Mark Squillante, Brian Kingsbury, Lior Horesh, IBM Thomas J. Watson Research Center, United States of America

SPTM-6.3: EXTREME-POINT PURSUIT FOR UNIT-MODULUS OPTIMIZATION 5548
Mingjie Shao, Qi Dai, Wing-Kin Ma, The Chinese University of Hong Kong, China

SPTM-6.4: GENERALIZED MATCHING PURSUITS FOR THE SPARSE OPTIMIZATION OF SEPARABLE OBJECTIVES 5553
Sebastian Ament, Carla Gomes, Cornell University, United States of America

SPTM-6.6: DELTA DISTANCING: A LIFTING APPROACH TO LOCALIZING ITEMS FROM USER COMPARISONS 5558
Andrew McRae, Austin Xu, Jihui Jin, Namrata Nadagouda, Nauman Ahad, Peimeng Guan, Mark Davenport, Georgia Institute of Technology, United States of America; Santhosh Karnik, Michigan State University, United States of America

SPTM-7: GRAPH REPRESENTATION AND ANALYSIS II

SPTM-7.1: DUAL PATH GRAPH CONVOLUTIONAL NETWORKS 5563
Yunhe Li, University of Montreal, Canada; Yaochen Hu, Yingxue Zhang, Huawei Technologies Canada, Canada

SPTM-7.2: ON THE STABILITY OF LOW PASS GRAPH FILTER WITH A LARGE NUMBER OF EDGE REWIRES 5568
Hoang-Son Nguyen, Yiran He, Hoi-To Wai, The Chinese University of Hong Kong, Hong Kong

SPTM-7.3: SPATIO-TEMPORAL GRAPH COMPLEMENTARY SCATTERING NETWORKS 5573
Zida Cheng, Siheng Chen, Ya Zhang, Shanghai Jiao Tong University, China

SPTM-7.4: CONVOLUTIONAL FILTERING IN SIMPLICIAL COMPLEXES 5578
Elvin Isufi, Maosheng Yang, Delft University of Technology, Netherlands

SPTM-7.5: ANNIHILATION FILTER APPROACH FOR ESTIMATING GRAPH DYNAMICS FROM DIFFUSION PROCESSES 5583
Arun Venkitaraman, Pascal Frossard, EPFL Switzerland, Switzerland

SPTM-8: SPARSITY-AWARE PROCESSING

SPTM-8.1: LEARNING GAUSSIAN GRAPHICAL MODELS WITH DIFFERING PAIRWISE SAMPLE SIZES 5588
Lili Zheng, Genevera Allen, Rice University, United States of America

SPTM-8.2: R-LOCAL UNLABELED SENSING: IMPROVED ALGORITHM AND APPLICATIONS 5593
Ahmed Ali Abbasi, Shuchin Aeron, Tufts ECE, United States of America; Abiy Tasissa, Tufts Department of Mathematics, United States of America

SPTM-8.3: FEDERATED OVER-AIR ROBUST SUBSPACE TRACKING FROM MISSING DATA 5598
Praneeth Narayanamurthy, Namrata Vaswani, Aditya Ramamoorthy, Iowa State University, United States of America

SPTM-8.4: ON CONTINUOUS-DOMAIN INVERSE PROBLEMS WITH SPARSE SUPERPOSITIONS OF DECAYING SINUSOIDS AS SOLUTIONS 5603
Rahul Parhi, Robert Nowak, University of Wisconsin-Madison, United States of America

SPTM-8.6: MULTIPLICATION-AVOIDING VARIANT OF POWER ITERATION WITH APPLICATIONS 5608
Hongyi Pan, Diaa Badawi, Runxuan Miao, Erdem Koyuncu, Ahmet Enis Cetin, Univerity of Illiniois at Chicago, United States of America

SPTM-9: SAMPLING, MULTIRATE, AND DIGITAL SIGNAL PROCESSING III

SPTM-9.1: BONA FIDE RIESZ PROJECTIONS FOR DENSITY ESTIMATION 5613
Pol del Aguila Pla, Michael Unser, École polytechnique fédérale de Lausanne, Switzerland

SPTM-9.2: BLIND MODULO ANALOG-TO-DIGITAL CONVERSION OF VECTOR PROCESSES **5617**
Amir Weiss, Gregory W. Wornell, Massachusetts Institute of Technology, United States of America; Everest Huang, MIT Lincoln Laboratory, United States of America; Or Ordentlich, Hebrew University of Jerusalem, Israel

SPTM-9.3: JOINT RADAR-COMMUNICATIONS PROCESSING FROM A DUAL-BLIND DECONVOLUTION PERSPECTIVE **5622**
Edwin Vargas, Roman Jacome, Henry Arguello, UNIVERSIDAD INDUSTRIAL DE SANTANDER, Colombia; Kumar Vijay Mishra, Brian Sadler, United States CCDC Army Research Laboratory, United States of America

SPTM-9.4: FAST MULTISCALE DIFFUSION ON GRAPHS **5627**
Sibylle Marcotte, ENS Rennes, France; Amélie Barbe, Rémi Gribonval, Titouan Vayer, Paulo Gonçalves, Université de Lyon, Inria, CNRS, ENSL, LIP, France; Marc Sebban, Université de Lyon, UJM-St-Etienne, CNRS, Lab. Hubert Curien, France; Pierre Borgnat, Université de Lyon, CNRS, ENSL, Lab. de Physique, France

SPTM-9.5: ADAPTIVE VARIATIONAL NONLINEAR CHIRP MODE DECOMPOSITION **5632**
Hao Liang, Xinghao Ding, Xiaotong Tu, Yue Huang, Xiamen University, China; Andreas Jakobsson, Lund University, Sweden

SPTM-9.6: DIFFERENTIATE-AND-FIRE TIME-ENCODING OF FINITE-RATE-OF-INNOVATION SIGNALS **5637**
Abijith Jagannath Kamath, Chandra Sekhar Seelamantula, Indian Institute of Science, India

SPTM-10: SIGNALS ON GRAPHS AND NETWORKS

SPTM-10.1: GRAPH LEARNING INFORMATION CRITERION **5642**
Koki Yamada, Yuichi Tanaka, Tokyo University of Agriculture and Technology, Japan

SPTM-10.3: EMBEDDING SIGNALS ON GRAPHS WITH UNBALANCED DIFFUSION EARTH MOVER'S DISTANCE **5647**
Alexander Tong, Dennis Shung, Manik Kuchroo, Smita Krishnaswamy, Yale University, United States of America; Guillaume Huguet, Amine Natik, Guillaume Lajoie, Guy Wolf, Université de Montréal, Canada

SPTM-10.5: MESSAGE PASSING-BASED COOPERATIVE LOCALIZATION WITH EMBEDDED PARTICLE FLOW **5652**
Lukas Wielandner, Erik Leitinger, Klaus Witrisal, Graz University of Technology, Austria; Florian Meyer, University of California San Diego, United States of America; Bryan Teague, MIT Lincoln Laboratory, United States of America

SPTM-10.6: A FRAMEWORK FOR PRIVATE COMMUNICATION WITH SECRET BLOCK STRUCTURE **5657**
Maxime Ferreira Da Costa, Urbashi Mitra, University of Southern California, United States of America

SPTM-11: ADAPTIVE SIGNAL PROCESSING

SPTM-11.1: LMS AND NLMS ALGORITHMS FOR THE IDENTIFICATION OF IMPULSE RESPONSES WITH INTRINSIC SYMMETRIC OR ANTISYMMETRIC PROPERTIES **5662**
Jacob Benesty, INRS-EMT, University of Quebec, Canada; Constantin Paleologu, Silviu Ciochina, University Politehnica of Bucharest, Romania; Eduardo Vinicius Kuhn, Khaled Jamal Bakri, Rui Seara, Federal University of Santa Catarina, Brazil

SPTM-11.2: DECENTRALIZED LEARNING IN THE PRESENCE OF LOW-RANK NOISE **5667**
Roula Nassif, Université Côte d'Azur, France; Virginia Bordignon, Ali Sayed, Ecole Polytechnique Fédérale de Lausanne, Switzerland; Stefan Vlaski, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SPTM-11.3: ADAPTIVE DIFFUSION WITH COMPRESSED COMMUNICATION..... 5672
Marco Carpentiero, Vincenzo Matta, University of Salerno, Italy; Ali H. Sayed, École Polytechnique Fédérale de Lausanne, Switzerland

SPTM-11.4: JOINT CENTRALITY ESTIMATION AND GRAPH IDENTIFICATION FROM 5677
MIXTURE OF LOW PASS GRAPH SIGNALS
Yiran He, Hoi-To Wai, The Chinese University of Hong Kong, Hong Kong

SPTM-11.5: FAIRNESS-AWARE SELECTIVE SAMPLING ON ATTRIBUTED GRAPHS 5682
Oyku Kose, Yanning Shen, University of California, Irvine, United States of America

SPTM-11.6: A SIMPLE GRAPH NEURAL NETWORK VIA LAYER SNIFFER..... 5687
Dingyi Zeng, Li Zhou, Wanlong Liu, Hong Qu, Wenyu Chen, University of Electronic Science and Technology of China, China

SPTM-12: CLASSIFICATION AND LEARNING

SPTM-12.1: NEW IMPROVED CRITERION FOR MODEL SELECTION IN SPARSE 5692
HIGH-DIMENSIONAL LINEAR REGRESSION MODELS
Prakash Borpatra Gohain, Magnus Jansson, KTH Royal Institute of Technology, Sweden, Sweden

SPTM-12.2: ON THE USE OF GEODESIC TRIANGLES BETWEEN GAUSSIAN 5697
DISTRIBUTIONS FOR CLASSIFICATION PROBLEMS
Antoine Collas, Chengfang Ren, CentraleSupélec, Université Paris-Saclay, France; Florent Bouchard, CNRS, CentraleSupélec, Université Paris-Saclay, France; Guillaume Ginolhac, Université Savoie Mont Blanc, France; Arnaud Breloy, Université Paris Nanterre, France; Jean-Philippe Ovarlez, Onera, Université Paris-Saclay, France

SPTM-12.3: A NON-CONVEX PROXIMAL APPROACH FOR CENTROID-BASED 5702
CLASSIFICATION
Mewe-Hezoudah Kahanam, Laurent Le Brusquet, Ségolène Martin, Jean-Christophe Pesquet, CentraleSupelec, Université Paris-Saclay, France

SPTM-12.4: EXTENDING THE USE OF MDL FOR HIGH-DIMENSIONAL 5707
PROBLEMS: VARIABLE SELECTION, ROBUST FITTING, AND ADDITIVE MODELING
Zhenyu Wei, Thomas Lee, University of California, Davis, United States of America; Raymond Wong, Texas A&M University, United States of America

SPTM-12.5: CLUSTERING COMPLEX SUBSPACES IN LARGE DIMENSIONS..... 5712
Roberto Pereira, Xavier Mestre, Centre Tecnològic de Telecomunicacions de Catalunya, Spain; David Gregoratti, Software Radio Systems, Spain

SPTM-12.6: ROBUST CLASSIFICATION WITH FLEXIBLE DISCRIMINANT ANALYSIS 5717
IN HETEROGENEOUS DATA
Pierre Houdouin, Frédéric Pascal, University Paris-Saclay, CentraleSupélec, France; Andrew Wang, University of Cambridge, United Kingdom of Great Britain and Northern Ireland; Matthieu Jonckheere, CNRS, France

SPTM-13: SAMPLING, MULTIRATE, AND DIGITAL SIGNAL PROCESSING IV

SPTM-13.1: RESIDUAL RECOVERY ALGORITHM FOR MODULO SAMPLING 5722
Eyar Azar, Satish Mulleti, Yonina Eldar, Weizmann Institute of Science, Rehovot, Israel, Israel

SPTM-13.2: OPERATOR FORMULATION FOR LINEAR TRANSFORMATIONS AND 5727
SIGNAL ESTIMATION IN THE JOINT SPATIAL-SLEPIAN DOMAIN
Adeem Aslam, University of Engineering and Technology, Pakistan; Zubair Khalid, Lahore University of Management Sciences, Pakistan

SPTM-13.3: SAMPLING SET SELECTION FOR GRAPH SIGNALS UNDER ARBITRARY SIGNAL PRIORS 5732
Junya Hara, Yuichi Tanaka, Tokyo University of Agriculture and Technology, Japan

SPTM-13.4: DETERMINING JOINT PERIODICITIES IN MULTI-TIME DATA WITH SAMPLING UNCERTAINTIES 5737
David Svedberg, Uppsala University, Sweden; Filip Elvander, KU Leuven, Belgium; Andreas Jakobsson, Lund University, Sweden

SPTM-13.5: UNLIMITED SAMPLING WITH LOCAL AVERAGES 5742
Dorian Florescu, Ayush Bhandari, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SPTM-13.6: MODULO EVENT-DRIVEN SAMPLING: SYSTEM IDENTIFICATION AND HARDWARE EXPERIMENTS 5747
Dorian Florescu, Ayush Bhandari, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SPTM-14: SIGNAL PROCESSING ON GRAPHS AND NETWORKS

SPTM-15: ESTIMATION THEORY AND METHODS I

SPTM-15.1: POINT-MASS FILTER WITH DECOMPOSITION OF TRANSIENT DENSITY 5752
Petr Tichavský, Czech Academy of Sciences, Czechia; Ondřej Straka, Jindřich Duník, University of West Bohemia, Czechia

SPTM-15.2: CRAMER-RAO BOUND FOR THE TIME-VARYING POISSON 5757
Xinhui Rong, Victor Solo, UNSW, Australia

SPTM-15.3: MODEL SELECTION VIA MISSPECIFIED CRAMER-RAO BOUND MINIMIZATION 5762
Nadav Eliran Rosenthal, Joseph Tabrikian, Ben-Gurion University of the Negev, Israel

SPTM-15.4: ROBUST PARAMETER ESTIMATION BASED ON THE K-DIVERGENCE 5767
Yair Sorek, Koby Todros, Ben-Gurion University of the Negev, Israel

SPTM-15.5: A CONVEX FORMULATION FOR THE ROBUST ESTIMATION OF MULTIVARIATE EXPONENTIAL POWER MODELS 5772
Nora Ouzir, Jean-Christophe Pesquet, CentraleSupélec Inria Paris Saclay, France; Frédéric Pascal, CentraleSupélec Paris Saclay, France

SPTM-15.6: CONDITIONALLY FACTORIZED VARIATIONAL BAYES WITH IMPORTANCE SAMPLING 5777
Runze Gan, Simon Godsill, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

SPTM-16: DETECTION THEORY AND METHODS

SPTM-16.1: ON THE FALSE ALARM PROBABILITY OF THE NORMALIZED MATCHED FILTER FOR OFF-GRID TARGET DETECTION 5782
Pierre Develter, Jonathan Bosse, Olivier Rabaste, Jean-Philippe Ovarlez, ONERA, France; Philippe Forster, Université Paris-Saclay, France

SPTM-16.2: A TWO-STREAM INFORMATION FUSION APPROACH TO ABNORMAL EVENT DETECTION IN VIDEO 5787
Yuxing Yang, Syed Mohsen Naqvi, Newcastle University, United Kingdom of Great Britain and Northern Ireland; Zeyu Fu, University of Oxford, United Kingdom of Great Britain and Northern Ireland

SPTM-16.3: A TEST FOR CONDITIONAL CORRELATION BETWEEN RANDOM VECTORS BASED ON WEIGHTED U-STATISTICS 5792
Marc Vilà, Jaume Riba, Technical University of Catalonia (UPC), Spain

SPTM-16.4: SEMI-SUPERVISED STANDARDIZED DETECTION OF PERIODIC SIGNALS WITH APPLICATION TO EXOPLANET DETECTION 5797
Sophia Sulis, Univ. Aix-Marseille, Laboratoire d'Astrophysique de Marseille, CNRS, France; David Mary, Lionel Bigot, Univ. Cote d'Azur, Observatoire de la Cote d'Azur, CNRS, France

SPTM-16.5: JOINT NORMALITY TEST VIA TWO-DIMENSIONAL PROJECTION 5802
Sara ElBouch, GIPSA-Lab, France; Olivier Michel, Pierre Comon, GIPSA-lab, France

SPTM-16.6: QUICKEST DETECTION OF COMPOSITE AND NON-STATIONARY CHANGES WITH APPLICATION TO PANDEMIC MONITORING 5807
Yuchen Liang, Venugopal Veeravalli, University of Illinois at Urbana-Champaign, United States of America

SPTM-17: SIGNAL AND INFORMATION PROCESSING OVER GRAPHS I: STATISTICAL APPROACHES

SPTM-17.1: A STIMULI-RELEVANT DIRECTED DEPENDENCY INDEX FOR TIME SERIES 5812
Payam Shahsavari Baboukani, Sergios Theodoridis, Jan Østergaard, Aalborg University, Denmark

SPTM-17.2: JOINT INFERENCE OF MULTIPLE GRAPHS WITH HIDDEN VARIABLES FROM STATIONARY GRAPH SIGNALS 5817
Samuel Rey, Andrei Buciulea, Antonio Garcia Marques, King Juan Carlos University, Spain; Madeline Navarro, Santiago Segarra, Rice University, United States of America

SPTM-17.3: SPARSE-GROUP LOG-SUM PENALIZED GRAPHICAL MODEL LEARNING FOR TIME SERIES 5822
Jitendra Tugnait, Auburn University, United States of America

SPTM-17.4: WIDE-SENSE STATIONARITY AND SPECTRAL ESTIMATION FOR GENERALIZED GRAPH SIGNAL 5827
Xingchao Jian, Wee Peng Tay, Nanyang Technological University, Singapore

SPTM-17.5: BLIND EXTRACTION OF EQUITABLE PARTITIONS FROM GRAPH SIGNALS 5832
Michael Scholkemper, Michael Schaub, RWTH Aachen, Germany

SPTM-17.6: LEARNING SPARSE GRAPHS WITH A CORE-PERIPHERY STRUCTURE 5837
Sravanthi Gurugubelli, Sundeep Prabhakar Chepuri, Indian Institute of Science, India

SPTM-18: ADAPTATION AND LEARNING OVER GRAPHS

SPTM-18.1: OPTIMAL COMBINATION POLICIES FOR ADAPTIVE SOCIAL LEARNING 5842
Ping Hu, Virginia Bordignon, Ali H. Sayed, EPFL, Switzerland; Stefan Vlaski, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SPTM-18.2: SEISMIC FAULT IDENTIFICATION USING GRAPH HIGH-FREQUENCY COMPONENTS AS INPUT TO GRAPH CONVOLUTIONAL NETWORK 5847
Patitapaban Palo, Aurobinda Routray, IIT Kharagpur, India

SPTM-18.3: DISTRIBUTED GRAPH LEARNING WITH SMOOTH DATA PRIORS 5852
Isabela Cunha Maia Nobre, Pascal Frossard, EPFL, Switzerland; Mireille El Gheche, Sony AI, Switzerland

SPTM-18.4: ADVERSPARSE: AN ADVERSARIAL ATTACK FRAMEWORK FOR DEEP SPATIAL-TEMPORAL GRAPH NEURAL NETWORKS **5857**
Jiayu Li, Shengmin Jin, Makan Fardad, Reza Zafarani, Syracuse University, United States of America; Tianyun Zhang, Cleveland State University, United States of America

SPTM-18.5: MULTIMODAL GRAPH SIGNAL DENOISING VIA TWOFOLD GRAPH SMOOTHNESS REGULARIZATION WITH DEEP ALGORITHM UNROLLING **5862**
Masatoshi Nagahama, Yuichi Tanaka, Tokyo University of Agriculture and Technology, Japan

SPTM-18.6: HETEROGENEOUS GRAPH NODE CLASSIFICATION WITH MULTI-HOPS RELATION FEATURES **5867**
Xiaolong Xu, Lingjuan Lyu, Hong Jin, Weiqiang Wang, Shuo Jia, ANT GROUP, China

SPTM-19: ESTIMATION, RECONSTRUCTION, AND DECONVOLUTION

SPTM-19.1: SIGNAL RECOVERY FROM INCONSISTENT NONLINEAR OBSERVATIONS **5872**
Patrick L. Combettes, Zev C. Woodstock, North Carolina State University, United States of America

SPTM-19.2: PERFECT RECONSTRUCTION OF CLASSES OF NON-BANDLIMITED SIGNALS FROM PROJECTIONS WITH UNKNOWN ANGLES **5877**
Renke Wang, Roxana Alexandru, Pier Luigi Dragotti, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SPTM-19.3: SHORT-AND-SPARSE DECONVOLUTION VIA RANK-ONE CONSTRAINED OPTIMIZATION (ROCO) **5882**
Cheng Cheng, Wei Dai, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SPTM-19.6: BLIND EQUALIZATION OF MOVING AVERAGE CHANNELS OVER GALOIS FIELDS **5887**
Arie Yeredor, Tel Aviv University, Israel

SPTM-20: DETECTION AND TRACKING

SPTM-20.1: SPARSE SUBSPACE TRACKING IN HIGH DIMENSIONS **5892**
Trung Thanh Le, Karim Abed-Meraim, University of Orleans, France; Adel Hafiane, INSA Centre Val de Loire, France; Linh Trung Nguyen, VNU University of Engineering and Technology, Viet Nam

SPTM-20.2: HOW CAN A COGNITIVE RADAR MASK ITS COGNITION? **5897**
Kunal Pattanayak, Vikram Krishnamurthy, Cornell University, United States of America; Christopher Berry, Lockheed Martin Advanced Technologies Laboratory, United States of America

SPTM-20.3: RTSNET: DEEP LEARNING AIDED KALMAN SMOOTHING **5902**
Xiaoyong Ni, Guy Revach, Eidgenössische Technische Hochschule Zürich, Switzerland; Nir Shlezinger, Ben-Gurion University of the Negev, Israel; Ruud J. G. van Sloun, Eindhoven University of Technology, Netherlands; Yonina C. Eldar, Weizmann Institute of Science, Israel

SPTM-20.5: GENERALIZED AUTOCORRELATION ANALYSIS FOR MULTI-TARGET DETECTION **5907**
Ye'Ela Shalit, Ran Weber, Asaf Abas, Shay Kreymer, Tamir Bendory, Tel Aviv University, Israel

SPTM-20.6: APPROXIMATING THE LIKELIHOOD RATIO IN LINEAR-GAUSSIAN STATE-SPACE MODELS FOR CHANGE DETECTION **5912**
Kostas Tsampourakis, Víctor Elvira, University of Edinburgh, United Kingdom of Great Britain and Northern Ireland

SPTM-21: SIGNAL AND INFORMATION PROCESSING OVER GRAPHS II: FILTERING AND TRANSFORMS

SPTM-21.1: LEARNING EXPANDING GRAPHS FOR SIGNAL INTERPOLATION 5917
Bishwadeep Das, Elvin Isufi, TU Delft, Netherlands

SPTM-21.2: HODGELETS: LOCALIZED SPECTRAL REPRESENTATIONS OF FLOWS ON SIMPLICIAL COMPLEXES 5922
T. Mitchell Roddenberry, Santiago Segarra, Rice University, United States of America; Florian Frantzen, Michael T. Schaub, RWTH Aachen University, Germany

SPTM-21.3: RECOVERY OF GRAPH SIGNALS FROM SIGN MEASUREMENTS 5927
Wenwei Liu, Hui Feng, Kaixuan Wang, Bo Hu, Fudan University, China; Feng Ji, Nanyang Technological University, Singapore

SPTM-21.4: EDGE SAMPLING OF GRAPHS BASED ON EDGE SMOOTHNESS 5932
Kenta Yanagiya, Koki Yamada, Yuichi Tanaka, Tokyo University of Agriculture and Technology, Japan; Yasuo Katsuhara, Tomoya Takatani, Frontier Research Center, Toyota Motor Corporation, Japan

SPTM-21.5: WLS DESIGN OF ARMA GRAPH FILTERS USING ITERATIVE SECOND-ORDER CONE PROGRAMMING 5937
Darukeesan Pakiyarajah, University of Jaffna, Sri Lanka; Chamira Edussooriya, University of Moratuwa, Sri Lanka

SPTM-21.6: LINEAR-TIME SAMPLING ON SIGNED GRAPHS VIA GERSHGORIN DISC PERFECT ALIGNMENT 5942
Chinthaka Dinesh, Ivan Bajic, Simon Fraser University, Canada; Saghar Bagheri, Gene Cheung, York University, Canada

SPTM-22: SIGNAL PROCESSING OVER NETWORKS

SPTM-22.1: PRIVACY-PRESERVING FEDERATED MULTI-TASK LINEAR REGRESSION: A ONE-SHOT LINEAR MIXING APPROACH INSPIRED BY GRAPH REGULARIZATION 5947
Harlin Lee, Andrea Bertozzi, University of California, Los Angeles, United States of America; Jelena Kovačević, New York University, United States of America; Yuejie Chi, Carnegie Mellon University, United States of America

SPTM-22.2: ECO-FEDSPLIT: FEDERATED LEARNING WITH ERROR-COMPENSATED COMPRESSION 5952
Sarit Khirirat, Division of Decision and Control Systems, KTH Royal Institute of Technology, Sweden, Sweden; Sindri Magnússon, Stockholm University, Sweden; Mikael Johansson, KTH Royal Institute of Technology, Sweden

SPTM-22.3: A TIME ENCODING APPROACH TO TRAINING SPIKING NEURAL NETWORKS 5957
Karen Adam, Ecole Polytechnique Fédérale de Lausanne, Switzerland

SPTM-22.4: TRANSIENT ANALYSIS OF CLUSTERED MULTITASK DIFFUSION RLS ALGORITHM 5962
Wei Gao, Jiangsu University, China; Jie Chen, Wentao Shi, Qunfei Zhang, Northwestern Polytechnical University, China; Cedric Richard, Université Côte d'Azur, France

SPTM-22.5: IMPROVING INFERENCE FOR SPATIAL SIGNALS BY CONTEXTUAL FALSE DISCOVERY RATES 5967
Martin Gölz, Abdelhak M. Zoubir, Technische Universität Darmstadt, Germany; Visa Koivunen, Aalto University, Finland

SPTM-22.6: ESTIMATION OF THE ADMITTANCE MATRIX IN POWER SYSTEMS UNDER LAPLACIAN AND PHYSICAL CONSTRAINTS 5972
Morad Halihal, Tirza Routtenberg, Ben Gurion University of the Negev, Israel

SPTM-23: ESTIMATION THEORY AND APPLICATIONS

SPTM-23.2: INCIPIENT FAULT SEVERITY ESTIMATION USING LOCAL MAHALANOBIS DISTANCE 5977
Junjie Yang, Claude Delpha, Universite Paris Saclay, France

SPTM-23.3: GRIDLESS DOA ESTIMATION UNDER THE MULTI-FREQUENCY MODEL 5982
Yifan Wu, Peter Gerstoft, University of California, San Diego, United States of America; Michael Wakin, Colorado School of Mines, United States of America

SPTM-24: NMF, DIVERGENCES AND OPTIMIZATION

SPTM-24.1: ORTHOGONAL NONNEGATIVE MATRIX TRI-FACTORIZATION FOR COMMUNITY DETECTION IN MULTIPLEX NETWORKS 5987
Meiby Ortiz-Bouza, Selin Aviyente, Michigan State University, United States of America

SPTM-24.2: STUDYING THREE FAMILIES OF DIVERGENCES TO COMPARE WIDE-SENSE STATIONARY GAUSSIAN ARMA PROCESSES 5992
Eric Grivel, Bordeaux INP, France

SPTM-24.3: MULTIVARIATE MULTISCALE COSINE SIMILARITY ENTROPY 5997
Hongjian Xiao, Danilo Mandic, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Theerasak Chanwimalueang, Srinakharinwirot University, Thailand

SPTM-24.4: ZEROTH-ORDER RANDOMIZED SUBSPACE NEWTON METHODS 6002
Erik Berglund, Sarit Khirirat, Xiaoyu Wang, KTH Royal Institute of Technology, Sweden

SPTM-24.5: FAST AND STABLE CONVERGENCE OF ONLINE SGD FOR CV@R-BASED RISK-AWARE LEARNING 6007
Dionysios Kalogerias, Yale University, United States of America

SPTM-24.6: DEEP INITIALIZATION FOR GUARANTEED UNIMODULAR QUADRATIC PROGRAMMING 6012
Amrutha Varshini Ramesh, University of British Columbia, Canada; Mojtaba Soltanalian, University of Illinois at Chicago, United States of America

SPE-1: SPEECH SEPARATION

SPE-1.1: CONTINUOUS SPEECH SEPARATION WITH RECURRENT SELECTIVE ATTENTION NETWORK 6017
Yixuan Zhang, The Ohio State University, United States of America; Zhuo Chen, Jian Wu, Takuya Yoshioka, Peidong Wang, Zhong Meng, Jinyu Li, Microsoft, United States of America

SPE-1.2: SA-SDR: A NOVEL LOSS FUNCTION FOR SEPARATION OF MEETING STYLE DATA 6022
Thilo von Neumann, Christoph Boeddeker, Reinhold Haeb-Umbach, Paderborn University, Germany; Keisuke Kinoshita, Marc Delcroix, NTT Corporation, Japan

SPE-1.3: VARARRAY: ARRAY-GEOMETRY-AGNOSTIC CONTINUOUS SPEECH SEPARATION 6027
Takuya Yoshioka, Xiaofei Wang, Dongmei Wang, Min Tang, Zirun Zhu, Zhuo Chen, Naoyuki Kanda, Microsoft, United States of America

SPE-1.4: ALL-NEURAL BEAMFORMER FOR CONTINUOUS SPEECH SEPARATION 6032
Zhuohuang Zhang, Indiana University Bloomington, United States of America; Takuya Yoshioka, Naoyuki Kanda, Zhuo Chen, Xiaofei Wang, Dongmei Wang, Sefik Emre Eskimez, Microsoft, United States of America

SPE-1.6: MINING HARD SAMPLES LOCALLY AND GLOBALLY FOR IMPROVED SPEECH SEPARATION **6037**
Kai Wang, Yizhou Peng, Hao Huang, Ying Hu, Xinjiang University, China; Sheng Li, National Institute of Information and Communications Technology, China

SPE-2: SPEECH RECOGNITION: ROBUST SPEECH RECOGNITION I

SPE-2.1: AUDIO-VISUAL MULTI-CHANNEL SPEECH SEPARATION, DEREVERBERATION AND RECOGNITION **6042**
Guinan Li, Jiajun Deng, Xunying Liu, Helen Meng, The Chinese University of Hong Kong, China; Jianwei Yu, Tencent AI lab, China

SPE-2.2: BEST OF BOTH WORLDS: MULTI-TASK AUDIO-VISUAL AUTOMATIC SPEECH RECOGNITION AND ACTIVE SPEAKER DETECTION **6047**
Otavio Braga, Olivier Siohan, Google, United States of America

SPE-2.3: END-TO-END MULTI-MODAL SPEECH RECOGNITION WITH AIR AND BONE CONDUCTED SPEECH **6052**
Junqi Chen, Mou Wang, Northwestern Polytechnical University, China; Xiao-Lei Zhang, Northwestern Polytechnic University, China; Zhiyong Huang, Susanto Rahardja, National University of Singapore, Singapore

SPE-2.4: END-TO-END SPEECH RECOGNITION WITH JOINT DEREVERBERATION OF SUB-BAND AUTOREGRESSIVE ENVELOPES **6057**
Rohit Kumar, Anurenjan Purushothaman, Sriram Ganapathy, Indian Institute of Science, Bangalore, India; Anirudh Sreeram, University of Southern California, United States of America

SPE-2.5: IMPROVING NOISE ROBUSTNESS OF CONTRASTIVE SPEECH REPRESENTATION LEARNING WITH SPEECH RECONSTRUCTION **6062**
Heming Wang, DeLiang Wang, The Ohio State University, United States of America; Yao Qian, Xiaofei Wang, Yiming Wang, Chengyi Wang, Shujie Liu, Takuya Yoshioka, Jinyu Li, Microsoft Corporation, United States of America

SPE-2.6: MULTI-CHANNEL MULTI-SPEAKER ASR USING 3D SPATIAL FEATURE **6067**
Yiwen Shao, Johns Hopkins University, United States of America; Shi-Xiong Zhang, Dong Yu, Tencent AI Lab, United States of America

SPE-3: SPEECH SYNTHESIS: GENERAL TOPICS I

SPE-3.1: IMPROVING CROSS-LINGUAL SPEECH SYNTHESIS WITH TRIPLET TRAINING SCHEME **6072**
Jianhao Ye, Hongbin Zhou, Zhiba Su, Wendi He, Kaimeng Ren, Lin Li, Heng Lu, Ximalaya Inc., China

SPE-3.2: IMPROVING PHONETIC REALIZATIONS IN TTS BY USING PHONEME-ALIGNED GRAPHEMES **6077**
Manish Sharma, Yizhi Hong, Emily Kaplan, Siamak Tazari, Rob Clark, Google, United Kingdom of Great Britain and Northern Ireland

SPE-3.3: CONTEXT-AWARE MASK PREDICTION NETWORK FOR END-TO-END TEXT-BASED SPEECH EDITING **6082**
Tao Wang, Jiangyan Yi, Ruibo Fu, Jianhua Tao, Zhengqi Wen, Institute of Automation, Chinese Academy of Sciences, China; Liqun Deng, Huawei Noah's Ark Lab, Shenzhen, China, China

SPE-3.4: A STUDY ON THE EFFICACY OF MODEL PRE-TRAINING IN DEVELOPING NEURAL TEXT-TO-SPEECH SYSTEM **6087**
Guangyan Zhang, Daxin Tan, Tan Lee, Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong; Yichong Leng, University of Science and Technology of China, China; Ying Qin, Institute of Information Science, Beijing Jiaotong University, China; Kaitao Song, Xu Tan, Microsoft Research Asia, China; Sheng Zhao, Microsoft Azure Speech, China

SPE-3.6: ONE TTS ALIGNMENT TO RULE THEM ALL 6092
Rohan Badlani, Adrian Lancucki, Kevin J. Shih, Rafael Valle, Wei Ping, Bryan Catanzaro, NVIDIA, United States of America

SPE-4: LANGUAGE MODELING

SPE-4.1: CAPITALIZATION NORMALIZATION FOR LANGUAGE MODELING WITH AN ACCURATE AND EFFICIENT HIERARCHICAL RNN MODEL 6097
Hao Zhang, You-Chi Cheng, Shankar Kumar, W. Ronny Huang, Mingqing Chen, Rajiv Mathews, Google, United States of America

SPE-4.2: ENHANCE RNNLMS WITH HIERARCHICAL MULTI-TASK LEARNING FOR ASR 6102
Minguang Song, Yunxin Zhao, University of Missouri-Columbia, United States of America

SPE-4.3: NEURAL-FST CLASS LANGUAGE MODEL FOR END-TO-END SPEECH RECOGNITION 6107
Antoine Bruguier, Duc Le, Dangna Li, Zhe Liu, Bo Wang, Eun Chang, Fuchun Peng, Ozlem Kalinli, Mike Seltzer, Facebook, United States of America; Rohit Prabhavalkar, Google, United States of America

SPE-4.4: LATTICEBART: LATTICE-TO-LATTICE PRE-TRAINING FOR SPEECH RECOGNITION 6112
Lingfeng Dai, Lu Chen, Zhikai Zhou, Kai Yu, Shanghai Jiao Tong University, China

SPE-4.5: RESCOREBERT: DISCRIMINATIVE SPEECH RECOGNITION RESCORING WITH BERT 6117
Liyan Xu, Yile Gu, Jari Kolehmainen, Haidar Khan, Ankur Gandhe, Ariya Rastrow, Andreas Stolcke, Ivan Bulyko, Amazon Alexa AI, United States of America

SPE-4.6: HYBRID SUB-WORD SEGMENTATION FOR HANDLING LONG TAIL IN MORPHOLOGICALLY RICH LOW RESOURCE LANGUAGES 6122
Sreeja Manghat, Sreeram Manghat, IEEE Graduate Member, India; Tanja Schultz, Cognitive Systems Lab, University Bremen, Germany, Germany

SPE-5: SPEAKER RECOGNITION I: SELF SUPERVISION

SPE-5.1: SELF-SUPERVISED SPEAKER VERIFICATION WITH SIMPLE SIAMESE NETWORK AND SELF-SUPERVISED REGULARIZATION 6127
Mufan Sang, The University of Texas at Dallas, United States of America; Haoqi Li, Fang Liu, Andrew O. Arnold, Li Wan, Amazon, United States of America

SPE-5.2: SELF-SUPERVISED SPEAKER RECOGNITION TRAINING USING HUMAN-MACHINE DIALOGUES 6132
Metehan Cekic, Upamanyu Madhow, UCSB, United States of America; Ruirui Li, Zeya Chen, Yuguang Yang, Andreas Stolcke, Amazon, United States of America

SPE-5.3: MULTI-TASK VOICE ACTIVATED FRAMEWORK USING SELF-SUPERVISED LEARNING 6137
Shehzeen Hussain, University of California, San Diego, United States of America; Van Nguyen, Shuhua Zhang, Erik Visser, Qualcomm Technologies, Inc., United States of America

SPE-5.4: SELF-SUPERVISED SPEAKER RECOGNITION WITH LOSS-GATED LEARNING 6142
*Ruijie Tao, Ville Hautamäki, Haizhou Li, National University of Singapore, Singapore; Kong Aik Lee, A*STAR, Singapore; Rohan Kumar Das, Fortemedia Singapore, Singapore*

SPE-5.5: LARGE-SCALE SELF-SUPERVISED SPEECH REPRESENTATION LEARNING FOR AUTOMATIC SPEAKER VERIFICATION 6147
Zhengyang Chen, Yanmin Qian, Shanghai Jiao Tong University, China; Sanyuan Chen, Yu Wu, Yao Qian, Chengyi Wang, Shujie Liu, Michael Zeng, Microsoft Corporation, China

SPE-5.6: UNISPEECH-SAT: UNIVERSAL SPEECH REPRESENTATION LEARNING WITH SPEAKER AWARE PRE-TRAINING 6152
Sanyuan Chen, Xiangzhan Yu, Harbin Institute of Technology, China; Yu Wu, Chengyi Wang, Zhengyang Chen, Zhuo Chen, Shujie Liu, Jian Wu, Yao Qian, Furu Wei, Jinyu Li, Microsoft Corporation, China

SPE-6: SPEECH AND SPOKEN LANGUAGE CORPORA

SPE-6.1: CARINA – A CORPUS OF ALIGNED GERMAN READ SPEECH INCLUDING ANNOTATIONS 6157
Hannes Kath, Simon Stone, Stefan Rapp, Peter Birkholz, Technische Universität Dresden, Germany

SPE-6.2: TOWARDS MEASURING FAIRNESS IN SPEECH RECOGNITION: CASUAL CONVERSATIONS DATASET TRANSCRIPTIONS 6162
Chunxi Liu, Michael Picheny, Leda Sari, Pooja Chitkara, Alex Xiao, Xiaohui Zhang, Mark Chou, Andrés Alvarado, Caner Hazirbas, Yatharth Saraf, Meta AI, United States of America

SPE-6.3: M2MET: THE ICASSP 2022 MULTI-CHANNEL MULTI-PARTY MEETING TRANSCRIPTION CHALLENGE 6167
Fan Yu, Shiliang Zhang, Siqi Zheng, Zhihao Du, Weilong Huang, Zhijie Yan, Bin Ma, Alibaba, China; Yihui Fu, Lei Xie, Pengcheng Guo, AISHELL Foundation, China; Xin Xu, Hui Bu, Beijing Shell Shell Technology Co., Ltd., China

SPE-6.4: ADIMA: ABUSE DETECTION IN MULTILINGUAL AUDIO 6172
Vikram Gupta, Rini Sharon, Ramit Sawhney, Debdoot Mukherjee, ShareChat, India, India

SPE-6.5: ANNO-MI: A DATASET OF EXPERT-ANNOTATED COUNSELLING DIALOGUES 6177
Zixiu Wu, Philips Research and University of Cagliari, Netherlands; Simone Balloccu, Ehud Reiter, University of Aberdeen, United Kingdom of Great Britain and Northern Ireland; Vivek Kumar, Diego Reforgiato Recupero, Daniele Riboni, University of Cagliari, Italy; Rim Helaoui, Philips Research, Netherlands

SPE-6.6: WENETSPEECH: A 10000+ HOURS MULTI-DOMAIN MANDARIN CORPUS FOR SPEECH RECOGNITION 6182
Binbin Zhang, Northwestern Polytechnical University; Mobvoi Inc, China; Hang Lv, Pengcheng Guo, Qijie Shao, Lei Xie, Northwestern Polytechnical University, China; Chao Yang, Xiaoyu Chen, Chenchen Zeng, Di Wu, Zhendong Peng, Mobvoi Inc., China; Xin Xu, Hui Bu, Beijing Shell Shell Technology Co., Ltd., China

SPE-7: SPEECH SYNTHESIS: GENERAL TOPICS II

SPE-7.1: WAVEBENDER GAN: AN ARCHITECTURE FOR PHONETICALLY MEANINGFUL SPEECH MANIPULATION 6187
Gustavo Teodoro Döhler Beck, Ulme Wennberg, Zofia Malisz, Gustav Eje Henter, KTH Royal Institute of Technology, Sweden

SPE-7.2: FRE-GAN 2: FAST AND EFFICIENT FREQUENCY-CONSISTENT AUDIO SYNTHESIS 6192
Sang-Hoon Lee, Ji-Hoon Kim, Kang-Eun Lee, Seong-Whan Lee, Korea University, Korea, Republic of

SPE-7.3: R-G2P: EVALUATING AND ENHANCING ROBUSTNESS OF GRAPHEME TO PHONEME CONVERSION BY CONTROLLED NOISE INTRODUCING AND CONTEXTUAL INFORMATION INCORPORATION 6197
Chendong Zhao, Haoqian Wang, The Shenzhen International Graduate School, Tsinghua University, China, China; Jianzong Wang, Xiaoyang Qu, Jing Xiao, Ping An Technology (Shenzhen) Co., Ltd., China

SPE-7.4: NEURAL GRAPHEME-TO-PHONEME CONVERSION WITH PRE-TRAINED GRAPHEME MODELS 6202
Lu Dong, Zhi-Qiang Guo, Chao-Hong Tan, Zhen-Hua Ling, National Engineering Laboratory for Speech and Language Information Processing, University of Science and Technology of China, Hefei, P. R. China, China; Ya-Jun Hu, Yuan Jiang, iFLYTEK Research, iFLYTEK Co., Ltd., Hefei, P. R. China, China

SPE-7.5: ISTFTNET: FAST AND LIGHTWEIGHT MEL-SPECTROGRAM VOCODER INCORPORATING INVERSE SHORT-TIME FOURIER TRANSFORM 6207
Takuhiro Kaneko, Kou Tanaka, Hirokazu Kameoka, Shogo Seki, NTT Corporation, Japan

SPE-7.6: ACOUSTIC APPLICATION OF PHASE RECONSTRUCTION ALGORITHMS IN OPTICS 6212
Tomoki Kobayashi, Tomoro Tanaka, Kohei Yatabe, Yasuhiro Oikawa, Waseda University, Japan

SPE-8: MACHINE TRANSLATION FOR SPOKEN AND WRITTEN LANGUAGE

SPE-8.1: CPT: CROSS-MODAL PREFIX-TUNING FOR SPEECH-TO-TEXT TRANSLATION 6217
Yukun Ma, Trung Hieu Nguyen, Bin Ma, Alibaba Group, Singapore

SPE-8.2: TACKLING DATA SCARCITY IN SPEECH TRANSLATION USING ZERO-SHOT MULTILINGUAL MACHINE TRANSLATION TECHNIQUES 6222
Tu Anh Dinh, Danni Liu, Jan Niehues, Maastricht University, Netherlands

SPE-8.3: IMPROVING END-TO-END SPEECH TRANSLATION MODEL WITH BERT-BASED CONTEXTUAL INFORMATION 6227
Jeong-Uk Bang, Min-Kyu Lee, Seung Yun, Sang-Hun Kim, Electronics and Telecommunications Research Institute (ETRI), Korea, Republic of

SPE-8.4: CONTEXT-ADAPTIVE DOCUMENT-LEVEL NEURAL MACHINE TRANSLATION 6232
Linlin Zhang, Zhejiang University, China; Zhirui Zhang, Boxing Chen, Weihua Luo, Luo Si, Alibaba DAMO Academy, China

SPE-8.5: INTEGRATING MULTIPLE ASR SYSTEMS INTO NLP BACKEND WITH ATTENTION FUSION 6237
Takatomo Kano, Atsunori Ogawa, Marc Delcroix, NTT Corporation, Japan; Shinji Watanabe, Language Technologies Institute, Carnegie Mellon University, United States of America

SPE-8.6: ISOMETRIC MT: NEURAL MACHINE TRANSLATION FOR AUTOMATIC DUBBING 6242
Surafel Lakew, Yogesh Virkar, Prashant Mathur, Marcello Federico, Amazon AI, United States of America

SPE-9: DEPRESSION DETECTION

SPE-9.1: AUTOMATIC DEPRESSION DETECTION: AN EMOTIONAL AUDIO-TEXTUAL CORPUS AND A GRU/BILSTM-BASED MODEL 6247
Ying Shen, Huiyu Yang, Lin Lin, Tongji University, China

SPE-9.2: MULTIMODAL DEPRESSION CLASSIFICATION USING ARTICULATORY COORDINATION FEATURES AND HIERARCHICAL ATTENTION BASED TEXT EMBEDDINGS 6252
Nadee Seneviratne, Carol Espy-Wilson, University of Maryland - College Park, United States of America

SPE-9.3: THIN SLICES OF DEPRESSION: IMPROVING DEPRESSION DETECTION PERFORMANCE THROUGH DATA SEGMENTATION 6257
Rawan Alsarrani, Alessandro Vinciarelli, University of Glasgow, United Kingdom of Great Britain and Northern Ireland; Anna Esposito, Universita' degli Studi della Campania, Italy

SPE-9.4: CLIMATE AND WEATHER: INSPECTING DEPRESSION DETECTION VIA EMOTION RECOGNITION 6262
Wen Wu, University of Cambridge, United Kingdom of Great Britain and Northern Ireland; Mengyue Wu, Kai Yu, Shanghai Jiao Tong University, China

SPE-9.5: FRAUG: A FRAME RATE BASED DATA AUGMENTATION METHOD FOR DEPRESSION DETECTION FROM SPEECH SIGNALS 6267
Vijay Ravi, Jinhan Wang, Jonathan Flint, Abeer Alwan, University of California Los Angeles, United States of America

SPE-9.6: PRIVACY SENSITIVE SPEECH ANALYSIS USING FEDERATED LEARNING TO ASSESS DEPRESSION 6272
Suhas Bn, Saeed Abdullah, Pennsylvania State University, United States of America

SPE-10: SPEECH RECOGNITION: ROBUST SPEECH RECOGNITION II

SPE-10.1: A TIME DOMAIN PROGRESSIVE LEARNING APPROACH WITH SNR CONSTRICTION FOR SINGLE-CHANNEL SPEECH ENHANCEMENT AND RECOGNITION 6277
Zhaoxu Nian, Jun Du, University of Science and Technology of China, China; Yu Ting Yeung, Renyu Wang, Huawei Noah's Ark Lab, China

SPE-10.2: A TWO-STEP APPROACH TO LEVERAGE CONTEXTUAL DATA: SPEECH RECOGNITION IN AIR-TRAFFIC COMMUNICATION 6282
Iuliia Nigmatulina, Juan Zuluaga-Gomez, Amrutha Prasad, Seyyed Saeed Sarfjoo, Petr Motlicek, Idiap Research Institute, Switzerland

SPE-10.3: LEARNING TO ENHANCE OR NOT: NEURAL NETWORK-BASED SWITCHING OF ENHANCED AND OBSERVED SIGNALS FOR OVERLAPPING SPEECH RECOGNITION 6287
Hiroshi Sato, Tsubasa Ochiai, Marc Delcroix, Keisuke Kinoshita, Naoyuki Kamo, Takafumi Moriya, NTT Corporation, Japan

SPE-10.4: INTERACTIVE FEATURE FUSION FOR END-TO-END NOISE-ROBUST SPEECH RECOGNITION 6292
Yuchen Hu, Nana Hou, Chen Chen, Eng Siong Chng, Nanyang Technological University, Singapore

SPE-10.5: SPEAKER REINFORCEMENT USING TARGET SOURCE EXTRACTION FOR ROBUST AUTOMATIC SPEECH RECOGNITION 6297
Catalin Zorila, Rama Doddipatla, Toshiba Cambridge Research Laboratory, United Kingdom of Great Britain and Northern Ireland

SPE-10.6: MITIGATING CLOSED-MODEL ADVERSARIAL EXAMPLES WITH BAYESIAN NEURAL MODELING FOR ENHANCED END-TO-END SPEECH RECOGNITION 6302
Chao-Han Huck Yang, Georgia Institute of Technology, United States of America; Zeeshan Ahmed, Yile Gu, Joseph Szurley, Roger Ren, Linda Liu, Andreas Stolcke, Ivan Bulyko, Amazon, United States of America

SPE-11: SPEECH SYNTHESIS: STYLE & EXPRESSIVENESS

SPE-11.1: REFEREE: TOWARDS REFERENCE-FREE CROSS-SPEAKER STYLE TRANSFER WITH LOW-QUALITY DATA FOR EXPRESSIVE SPEECH SYNTHESIS 6307
Songxiang Liu, Shan Yang, Dan Su, Dong Yu, Tencent, China

SPE-11.2: PVAE-TTS: ADAPTIVE TEXT-TO-SPEECH VIA PROGRESSIVE STYLE ADAPTATION 6312
Ji-Hyun Lee, Sang-Hoon Lee, Ji-Hoon Kim, Seong-Whan Lee, Korea University, Korea, Republic of

SPE-11.3: EMOQ-TTS: EMOTION INTENSITY QUANTIZATION FOR FINE-GRAINED CONTROLLABLE EMOTIONAL TEXT-TO-SPEECH 6317
Chae-Bin Im, Sang-Hoon Lee, Seung-Bin Kim, Seong-Whan Lee, Korea University, Korea, Republic of

SPE-11.4: JOINT AND ADVERSARIAL TRAINING WITH ASR FOR EXPRESSIVE SPEECH SYNTHESIS 6322
Kaili Zhang, Cheng Gong, Wenhuan Lu, Longbiao Wang, Jianguo Wei, Dawei Liu, Tianjin University, China

SPE-11.5: MSDTRON: A HIGH-CAPABILITY MULTI-SPEAKER SPEECH SYNTHESIS SYSTEM FOR DIVERSE DATA USING CHARACTERISTIC INFORMATION 6327
Qinghua Wu, Quanbo Shen, Jian Luan, Yujun Wang, Xiaomi Company, China

SPE-11.6: SPEECHSPLIT2.0: UNSUPERVISED SPEECH DISENTANGLEMENT FOR VOICE CONVERSION WITHOUT TUNING AUTOENCODER BOTTLENECKS 6332
Chak Ho Chan, Mark Hasegawa-Johnson, University of Illinois at Urbana-Champaign, United States of America; Kaizhi Qian, Yang Zhang, MIT-IBM Watson AI Lab, United States of America

SPE-12: EVENT AND RELATION EXTRACTION

SPE-12.1: DOCUMENT-LEVEL EVENT EXTRACTION VIA HUMAN-LIKE READING PROCESS 6337
Shiyao Cui, Xin Cong, Bowen Yu, Tingwen Liu, Yucheng Wang, Institute of Information Engineering, Chinese Academy of Sciences., China; Jinqiao Shi, Beijing University of Posts and Telecommunications, China

SPE-12.2: GENERATING DISENTANGLED ARGUMENTS WITH PROMPTS: A SIMPLE EVENT EXTRACTION FRAMEWORK THAT WORKS 6342
Jinghui Si, Beihang University; Alibaba Group, China; Xutan Peng, The University of Sheffield, United Kingdom of Great Britain and Northern Ireland; Chen Li, Jianxin Li, Beihang University, China; Haotian Xu, Alibaba Group, China

SPE-12.3: MULTI-ROLE EVENT ARGUMENT EXTRACTION AS MACHINE READING COMPREHENSION WITH ARGUMENT MATCH OPTIMIZATION 6347
Jingcong Tao, Youcheng Pan, Baotian Hu, Xiaolong Wang, Harbin Institute of Technology, Shenzhen, China; Xinyu Li, Weihua Peng, Cuiyun Han, Baidu International Technology (Shenzhen), China

SPE-12.4: BNU: A BALANCE-NORMALIZATION-UNCERTAINTY MODEL FOR INCREMENTAL EVENT DETECTION 6352
Jia Li, Peking University, China; Yunyan Zhang, Yifan Yang, Yefeng Zheng, Tencent, China; Zhicheng An, Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, China

SPE-12.5: WLINKER: MODELING RELATIONAL TRIPLET EXTRACTION AS WORD LINKING 6357
Yongxiu Xu, Chuan Zhou, Jing Yu, Yue Hu, University of Chinese Academy of Sciences, China; Heyan Huang, Beijing Institute of Technology, China

SPE-12.6: A KNOWLEDGE/DATA ENHANCED METHOD FOR JOINT EVENT AND TEMPORAL RELATION EXTRACTION 6362
Xiaobin Zhang, Liangjun Zang, Songlin Hu, Institute of Information Engineering, Chinese Academy of Sciences, China; Peng Cheng, Yuqi Wang, State Key Laboratory of Media Convergence Production Technology and Systems, China

SPE-13: SPEAKER RECOGNITION II: ANTI-SPOOFING

SPE-13.1: AASIST: AUDIO ANTI-SPOOFING USING INTEGRATED SPECTRO-TEMPORAL GRAPH ATTENTION NETWORKS 6367
Jee-weon Jung, Hee-Soo Heo, Joon Son Chung, Bong-Jin Lee, Naver Corporation, Korea, Republic of; Hemlata Tak, Nicholas Evans, EURECOM, France; Hye-jin Shim, Ha-Jin Yu, University of Seoul, France

SPE-13.2: ESTIMATING THE CONFIDENCE OF SPEECH SPOOFING COUNTERMEASURE 6372
Xin Wang, Junichi Yamagishi, National Institute of Informatics, Japan

SPE-13.3: TWO-PATH GMM-RESNET AND GMM-SENET FOR ASV SPOOFING DETECTION 6377
Zhenchun Lei, Hui Yan, Changhong Liu, Minglei Ma, Yingen Yang, Jiangxi Normal University, China

SPE-13.4: RAWBOOST: A RAW DATA BOOSTING AND AUGMENTATION METHOD APPLIED TO AUTOMATIC SPEAKER VERIFICATION ANTI-SPOOFING 6382
Hemlata Tak, Madhu Kamble, Jose Patino, Massimiliano Todisco, Nicholas Evans, EURECOM, France

SPE-13.5: EXPLAINING DEEP LEARNING MODELS FOR SPOOFING AND DEEPFAKE DETECTION WITH SHAPLEY ADDITIVE EXPLANATIONS 6387
Wanying Ge, Jose Patino, Massimiliano Todisco, Nicholas Evans, EURECOM, France

SPE-13.6: MULTI-TASK LEARNING IMPROVES SYNTHETIC SPEECH DETECTION 6392
Yichuan Mo, Shilin Wang, Shanghai Jiao Tong University, China

SPE-14: MULTI-LINGUAL ASR

SPE-14.1: MASSIVELY MULTILINGUAL ASR: A LIFELONG LEARNING SOLUTION 6397
Bo Li, Ruoming Pang, Yu Zhang, Tara Sainath, Trevor Strohman, Parisa Haghani, Yun Zhu, Brian Farris, Neeraj Gaur, Manasa Prasad, Google, United States of America

SPE-14.2: JOINT UNSUPERVISED AND SUPERVISED TRAINING FOR MULTILINGUAL ASR 6402
Junwen Bai, Cornell University, United States of America; Bo Li, Yu Zhang, Ankur Bapna, Nikhil Siddhartha, Khe Chai Sim, Tara Sainath, Google, United States of America

SPE-14.3: MULTILINGUAL SECOND-PASS RESCORING FOR AUTOMATIC SPEECH RECOGNITION SYSTEMS 6407
Neeraj Gaur, Tongzhou Chen, Ehsan Variani, Parisa Haghani, Bhuvana Ramabhadran, Pedro Moreno, Google, United States of America

SPE-14.4: JOINT MODELING OF CODE-SWITCHED AND MONOLINGUAL ASR VIA CONDITIONAL FACTORIZATION 6412
Brian Yan, Siddharth Dalmia, Dan Berrebbi, Shinji Watanabe, Carnegie Mellon University, United States of America; Chunlei Zhang, Meng Yu, Shi-Xiong Zhang, Chao Weng, Dong Yu, Tencent AI Lab, United States of America

SPE-14.5: BILINGUAL END-TO-END ASR WITH BYTE-LEVEL SUBWORDS 6417
Liuhui Deng, Roger Hsiao, Arnab Ghoshal, Apple, United States of America

SPE-14.6: A CONFIGURABLE MULTILINGUAL MODEL IS ALL YOU NEED TO RECOGNIZE ALL LANGUAGES 6422
Long Zhou, Shujie Liu, Microsoft Research Asia, China; Jinyu Li, Eric Sun, Microsoft Speech and Language Group, United States of America

SPE-15: EMOTION RECOGNITION: REPRESENTATION LEARNING

SPE-15.1: DOMAIN-INVARIANT FEATURE LEARNING FOR CROSS CORPUS SPEECH EMOTION RECOGNITION 6427
Yuan Gao, Longbiao Wang, Jiaxing Liu, Jianwu Dang, Tianjin University, China; Shogo Okada, Japan Advanced Institute of Science and Technology, Japan

SPE-15.2: MULTI-STAGE GRAPH REPRESENTATION LEARNING FOR DIALOGUE-LEVEL SPEECH EMOTION RECOGNITION 6432
Yaodong Song, Jiaxing Liu, Longbiao Wang, Ruiguo Yu, Tianjin University, China; Jianwu Dang, Japan Advanced Institute of Science and Technology, Japan

SPE-15.3: SPEECH EMOTION RECOGNITION WITH GLOBAL-AWARE FUSION ON MULTI-SCALE FEATURE REPRESENTATION 6437
Wenjing Zhu, Xiang Li, DXM Financial, China

SPE-15.4: REPRESENTATION LEARNING THROUGH CROSS-MODAL CONDITIONAL TEACHER-STUDENT TRAINING FOR SPEECH EMOTION RECOGNITION 6442
Sundararajan Srinivasan, Zhaocheng Huang, Katrin Kirchhoff, Amazon.com, United States of America

SPE-15.5: NOT ALL FEATURES ARE EQUAL: SELECTION OF ROBUST FEATURES FOR SPEECH EMOTION RECOGNITION IN NOISY ENVIRONMENTS 6447
Seong-Gyun Leem, Carlos Busso, The University of Texas at Dallas, United States of America; Daniel Fulford, Boston University, United States of America; Jukka-Pekka Onnela, Harvard University, United States of America; David Gard, San Francisco State University, United States of America

SPE-15.6: TOWARDS TRANSFERABLE SPEECH EMOTION REPRESENTATION: ON LOSS FUNCTIONS FOR CROSS-LINGUAL LATENT REPRESENTATIONS 6452
Sneha Das, Line H. Clemmensen, Technical University of Denmark, Denmark; Nicole Nadine Lønfeldt, Copenhagen University Hospital, Capital Region, Denmark; Anne Katrine Pagsberg, Copenhagen University Hospital and Copenhagen University, Denmark

SPE-16: LANGUAGE DISORDERS: DETECTION I

SPE-16.1: DEMENTIA DETECTION BY FUSING SPEECH AND EYE-TRACKING REPRESENTATION 6457
Zhengyan Sheng, Zhiqiang Guo, University of Science and Technology of China, China; Xin Li, iFlytek, China; Yunxia Li, Shanghai Tongji Hospital, China; Zhenhua Ling, National Engineering Laboratory for Speech and Language Information Processing,University of Science and Technology of China, Hefei, P.R.China, China

SPE-16.2: TOWARDS INTERPRETABILITY OF SPEECH PAUSE IN DEMENTIA DETECTION USING ADVERSARIAL LEARNING 6462
Youxiang Zhu, Bang Tran, Xiaohui Liang, University of Massachusetts Boston, United States of America; John Batsis, University of North Carolina at Chapel Hill, United States of America; Robert Roth, Geisel School of Medicine at Dartmouth, United States of America

SPE-16.3: USING SPECTRAL SEQUENCE-TO-SEQUENCE AUTOENCODERS TO ASSESS MILD COGNITIVE IMPAIRMENT 6467
Mercedes Vetráb, José Vicente Egas-López, Réka Balogh, Nóra Imre, László Tóth, Magdolna Pákáski, János Kálmán, University of Szeged, Hungary; Ildikó Hoffmann, Research Institute for Linguistics, Hungary; Gábor Gosztolya, ELRN-SZTE Research Group on Artificial Intelligence, Hungary

SPE-16.4: EXPLORING DEMENTIA DETECTION FROM SPEECH: CROSS CORPUS ANALYSIS 6472
Ayimnisagul Ablimit, Tanja Schultz, University of Bremen, Germany; Catarina Botelho, Alberto Abad, Isabel Trancoso, INESC-ID/Instituto Superior Técnico, Portugal

SPE-16.5: EXPERIMENTAL INVESTIGATION ON STFT PHASE REPRESENTATIONS FOR DEEP LEARNING-BASED DYSARTHRIC SPEECH DETECTION 6477
Parvaneh Janbakhshi, Ina Kodrasi, Idiap Research Institute, Switzerland

SPE-16.6: DYSFLUENCY CLASSIFICATION IN STUTTERED SPEECH USING DEEP LEARNING FOR REAL-TIME APPLICATIONS 6482
Melanie Jouaiti, Kerstin Dautenhahn, University of Waterloo, Canada

SPE-17: SPEECH ENHANCEMENT: MULTI-CHANNEL

SPE-17.1: EMBEDDING AND BEAMFORMING: ALL-NEURAL CAUSAL BEAMFORMER FOR MULTICHANNEL SPEECH ENHANCEMENT 6487
Andong Li, Wenzhe Liu, Chengshi Zheng, Xiaodong Li, Institute of Acoustics, Chinese Academy of Sciences, China

SPE-17.2: IMPROVING DUAL-MICROPHONE SPEECH ENHANCEMENT BY LEARNING CROSS-CHANNEL FEATURES WITH MULTI-HEAD ATTENTION 6492
Xinmeng Xu, Rongzhi Gu, Yuexian Zou, Peking University, China

SPE-17.3: TPARN: TRIPLE-PATH ATTENTIVE RECURRENT NETWORK FOR TIME-DOMAIN MULTICHANNEL SPEECH ENHANCEMENT 6497
Ashutosh Pandey, DeLiang Wang, The Ohio State University, United States of America; Buye Xu, Anurag Kumar, Jacob Donley, Paul Calamia, Facebook Reality Labs Research, United States of America

SPE-17.4: MULTICHANNEL SPEECH ENHANCEMENT WITHOUT BEAMFORMING 6502
Ashutosh Pandey, DeLiang Wang, The Ohio State University, United States of America; Buye Xu, Anurag Kumar, Jacob Donley, Paul Calamia, Facebook Reality Labs Research, United States of America

SPE-17.5: LEARNING FILTERBANKS FOR END-TO-END ACOUSTIC BEAMFORMING 6507
Samuele Cornell, Stefano Squartini, Università Politecnica delle Marche, Italy; Manuel Pariente, Universite de Lorraine, CNRS, Inria, LORIA,, France; François Grondin, Université de Sherbrooke, Canada

SPE-17.6: SPATIAL-TEMPORAL GRAPH CONVOLUTION NETWORK FOR MULTICHANNEL SPEECH ENHANCEMENT 6512
Minghui Hao, Jingjing Yu, Luyao Zhang, Beijing Jiaotong University, China

SPE-18: SPEECH RECOGNITION: GENERAL TOPICS I

SPE-18.1: LATTICE RESCORING BASED ON LARGE ENSEMBLE OF COMPLEMENTARY NEURAL LANGUAGE MODELS 6517
Atsunori Ogawa, Naohiro Tawara, Marc Delcroix, Shoko Araki, NTT Corporation, Japan

SPE-18.2: CONTINUAL LEARNING USING LATTICE-FREE MMI FOR SPEECH RECOGNITION 6522
Hossein Hadian, Arseniy Gorin, Behavox, Canada

SPE-18.3: NON-AUTOREGRESSIVE TRANSFORMER WITH UNIFIED BIDIRECTIONAL DECODER FOR AUTOMATIC SPEECH RECOGNITION 6527
Chuan-Fei Zhang, Yan Liu, Tian-Hao Zhang, Song-Lu Chen, Xu-Cheng Yin, University of Science and Technology Beijing, China; Feng Chen, EEasy Technology Company Ltd., China

SPE-18.4: MODEL-BASED APPROACH FOR MEASURING THE FAIRNESS IN ASR 6532
Zhe Liu, Irina-Elena Veliche, Fuchun Peng, Meta Platforms, Inc, United States of America

SPE-18.5: IMPROVING CONFIDENCE ESTIMATION ON OUT-OF-DOMAIN DATA FOR END-TO-END SPEECH RECOGNITION 6537
Qiujia Li, Phil Woodland, University of Cambridge, United Kingdom of Great Britain and Northern Ireland; Yu Zhang, David Qiu, Yanzhang He, Liangliang Cao, Google LLC, United States of America

SPE-18.6: PARALLEL COMPOSITION OF WEIGHTED FINITE-STATE TRANSDUCERS 6542
Shubho Sengupta, Vineel Pratap, Facebook, United States of America; Awni Hannun, Zoom, United States of America

SPE-19: VOICE CONVERSION: REPRESENTATION

SPE-19.1: DGC-VECTOR: A NEW SPEAKER EMBEDDING FOR ZERO-SHOT VOICE CONVERSION 6547
Ruitong Xiao, South China University of Technology, China; Haitong Zhang, Yue Lin, Netease Games, China

SPE-19.2: S3PRL-VC: OPEN-SOURCE VOICE CONVERSION FRAMEWORK WITH SELF-SUPERVISED SPEECH REPRESENTATIONS 6552
Wen-Chin Huang, Tomoki Hayashi, Tomoki Toda, Nagoya University, Japan; Shu-wen Yang, Hung-yi Lee, National Taiwan University, Taiwan; Shinji Watanabe, Carnegie Mellon University, United States of America

SPE-19.3: TRAINING ROBUST ZERO-SHOT VOICE CONVERSION MODELS WITH SELF-SUPERVISED FEATURES 6557
Trung Dang, Peter Chin, Boston University, United States of America; Dung Tran, Kazuhito Koishida, Microsoft Corp., United States of America

SPE-19.4: A COMPARISON OF DISCRETE AND SOFT SPEECH UNITS FOR IMPROVED VOICE CONVERSION 6562
Benjamin van Niekerk, Matthew Baas, Herman Kamper, Stellenbosch University, South Africa; Marc-André Carbonneau, Julian Zaïdi, Hugo Seuté, Ubisoft, Canada

SPE-19.5: SIG-VC: A SPEAKER INFORMATION GUIDED ZERO-SHOT VOICE CONVERSION SYSTEM FOR BOTH HUMAN BEINGS AND MACHINES 6567
Haozhe Zhang, Zexin Cai, Xiaoyi Qin, Ming Li, Duke Kunshan University, China

SPE-19.6: ROBUST DISENTANGLED VARIATIONAL SPEECH REPRESENTATION LEARNING FOR ZERO-SHOT VOICE CONVERSION 6572
Jiachen Lian, UC Berkeley, United States of America; Chunlei Zhang, Dong Yu, Tencent AI Lab, United States of America

SPE-20: DIALOG SYSTEM I: RESPONSE RETRIEVAL/GENERATION

SPE-20.1: IMPROVING DIALOGUE GENERATION VIA PROACTIVELY QUERYING GROUNDED KNOWLEDGE 6577
Xiangyu Zhao, Longbiao Wang, Tianjin University, China; Jianwu Dang, Japan Advanced Institute of Science and Technology, China; , ,

SPE-20.2: A NON-HIERARCHICAL ATTENTION NETWORK WITH MODALITY DROPOUT FOR TEXTUAL RESPONSE GENERATION IN MULTIMODAL DIALOGUE SYSTEMS 6582
Rongyi Sun, Borun Chen, Yinghui Li, Hai-Tao Zheng, Shenzhen International Graduate School, Tsinghua University, China; Qingyu Zhou, Yunbo Cao, Tencent Cloud Xiaowei, China; , ,

SPE-20.3: JOINT LEARNING FOR ADDRESSEE SELECTION AND RESPONSE GENERATION IN MULTI-PARTY CONVERSATION 6587
Qi Song, Sheng Li, Ping Wei, Ge Luo, Xinpeng Zhang, Zhenxing Qian, Fudan University, China

SPE-20.4: RETRIEVAL ENHANCED SEGMENT GENERATION NEURAL NETWORK FOR TASK-ORIENTED DIALOGUE SYSTEMS 6592
Miaoxin Chen, Rongyi Sun, Kai Ouyang, Hai-Tao Zheng, Shenzhen International Graduate School, Tsinghua University, China, China; Zibo Lin, WeChat Search Application Department, Tencent, China, China; Rui Xie, Wei Wu, Meituan, Shanghai, China, China

SPE-20.5: A MULTI DOMAIN KNOWLEDGE ENHANCED MATCHING NETWORK FOR RESPONSE SELECTION IN RETRIEVAL-BASED DIALOGUE SYSTEMS 6597
Xiuyi Chen, Feilong Chen, Shuang Xu, Bo Xu, Institute of Automation, Chinese Academy of Science, China

SPE-20.6: RETRIEVAL BIAS AWARE ENSEMBLE MODEL FOR CONDITIONAL SENTENCE GENERATION 6602
Yiping Song, Zheng Xie, Jianping Li, National University of Defense Technology, China; Luchen Liu, Ming Zhang, Peking University, China; Zhiliang Tian, The Hong Kong University of Science and Technology, China

SPE-21: SPEAKER RECOGNITION III: ADVERSARIAL, SPOOFING AND ROBUSTNESS

SPE-21.2: EFFECTIVE AND INCONSPICUOUS OVER-THE-AIR ADVERSARIAL EXAMPLES WITH ADAPTIVE FILTERING 6607
Patrick O'Reilly, Aravindan Vijayaraghavan, Bryan Pardo, Northwestern University, United States of America; Pranjal Awasthi, Google Research, United States of America

SPE-21.3: LRPD: LARGE REPLAY PARALLEL DATASET 6612
Ivan Yakovlev, Mikhail Melnikov, Nikita Bukhal, Rostislav Makarov, Alexander Alenin, Nikita Torgashov, Anton Okhotnikov, ID R&D Inc., Russian Federation

SPE-21.4: ROBUST SELF-SUPERVISED SPEAKER REPRESENTATION LEARNING VIA INSTANCE MIX REGULARIZATION 6617
Woohyun Kang, Jahangir Alam, Abderrahim Fathan, Computer Research Institute of Montreal, Canada

SPE-21.5: GRAPH CONVOLUTIONAL NETWORK BASED SEMI-SUPERVISED LEARNING ON MULTI-SPEAKER MEETING DATA 6622
Fuchuan Tong, Lin Li, Qingyang Hong, Xiamen University, China; Siqi Zheng, Hongbin Suo, Alibaba Group, China; Min Zhang, Zhejiang University, China; Yafeng Chen, University of Science and Technology of China, China

SPE-22: ADAPTATION AND PERSONALIZATION FOR SPEECH MODELS

SPE-22.1: LARGE-SCALE ASR DOMAIN ADAPTATION USING SELF- AND SEMI-SUPERVISED LEARNING 6627
Dongseong Hwang, Ananya Misra, Zhouyuan Huo, Nikhil Siddhartha, Shefali Garg, David Qiu, Khe Chai Sim, Trevor Strohman, Françoise Beaufays, Yanzhang He, Google, United States of America

SPE-22.2: FAST CONTEXTUAL ADAPTATION WITH NEURAL ASSOCIATIVE MEMORY FOR ON-DEVICE PERSONALIZED SPEECH RECOGNITION 6632
Tsendsuren Munkhdalai, Khe Chai Sim, Angad Chandorkar, Fan Gao, Mason Chua, Trevor Strohman, Françoise Beaufays, Google, United States of America

SPE-22.3: PERSONALIZED AUTOMATIC SPEECH RECOGNITION TRAINED ON SMALL DISORDERED SPEECH DATASETS 6637
Jimmy Tobin, Katrin Tomanek, Google, United States of America

SPE-22.4: SPELL MY NAME: KEYWORD BOOSTED SPEECH RECOGNITION 6642
Namkyu Jung, Geonmin Kim, NAVER Corporation, Korea, Republic of; Joon Son Chung, Korea Advanced Institute of Science and Technology, Korea, Republic of

SPE-22.5: MAGIC DUST FOR CROSS-LINGUAL ADAPTATION OF MONOLINGUAL WAV2VEC-2.0 6647
Sameer Khurana, James Glass, Massachusetts Institute of Technology, United States of America; Antoine Laurent, Le Mans University, France

SPE-22.6: TEXT ADAPTIVE DETECTION FOR CUSTOMIZABLE KEYWORD SPOTTING 6652
Yu Xi, Wangyou Zhang, Baochen Yang, Kai Yu, Shanghai Jiao Tong University, China; Tian Tan, AISpeech Ltd, China

SPE-23: VOICE CONVERSION: SINGING VOICE AND OTHERS

SPE-23.1: IMPROVING ADVERSARIAL WAVEFORM GENERATION BASED SINGING VOICE CONVERSION WITH HARMONIC SIGNALS 6657
Haohan Guo, Chinese University of Hong Kong, Hong Kong; Zhiping Zhou, Fanbo Meng, Kai Liu, Sogou, China

SPE-23.2: K-CONVERTER: AN UNSUPERVISED SINGING VOICE CONVERSION SYSTEM 6662
Ying Zhang, Peng Yang, Jinba Xiao, Ye Bai, Hao Che, Xiaorui Wang, kwai, China

SPE-23.3: HIFI-SVC: FAST HIGH FIDELITY CROSS-DOMAIN SINGING VOICE CONVERSION 6667
Yong Zhou, Xiangju Lu, iQIYI Inc., China

SPE-23.4: TOWARDS IDENTITY PRESERVING NORMAL TO DYSARTHRIC VOICE CONVERSION 6672
Wen-Chin Huang, Lester Phillip Violeta, Tomoki Toda, Nagoya University, Japan; Bence Mark Halpern, Odette Scharenborg, Delft University of Technology, Netherlands

SPE-23.5: SPEAKER IDENTITY PRESERVATION IN DYSARTHRIC SPEECH RECONSTRUCTION BY ADVERSARIAL SPEAKER ADAPTATION 6677
Disong Wang, Songxiang Liu, Xixin Wu, Hui Lu, Xunying Liu, Helen Meng, The Chinese University of Hong Kong, Hong Kong; Lifa Sun, SpeechX Limited, China

SPE-23.6: CONTROLLABLE SPEECH REPRESENTATION LEARNING VIA VOICE CONVERSION AND AIC LOSS 6682
Yunyun Wang, Jiaqi Su, Adam Finkelstein, Princeton University, United States of America; Zeyu Jin, Adobe Research, United States of America

SPE-24: DIALOG SYSTEM II: GENERAL TOPICS

SPE-24.1: VU-BERT: A UNIFIED FRAMEWORK FOR VISUAL DIALOG 6687
Tong Ye, Ping An Technology (Shenzhen) Co., Ltd. ;University of Science and Technology of China, China; Shijing Si, Jianzong Wang, Ning Cheng, Jing Xiao, Ping An Technology (Shenzhen) Co., Ltd., China; Rui Wang, Duke University, China

SPE-24.2: INTEGRATING PRETRAINED LANGUAGE MODEL FOR DIALOGUE POLICY EVALUATION 6692
Hongru Wang, Huimin Wang, Zezhong Wang, Kam-Fai Wong, The Chinese University of Hong Kong, China

SPE-24.3: CACHE: MODELING CONTRIBUTION-AWARE CONTEXT HIERARCHICALLY FOR LONG-RANGE DIALOGUE STATE TRACKING 6697
Jianshu Qi, Yuke Si, Longbiao Wang, Tianjin University, China; Jianwu Dang, Japan Advanced Institute of Science and Technology, Japan

SPE-24.4: ROBUST UNSTRUCTURED KNOWLEDGE ACCESS IN CONVERSATIONAL DIALOGUE WITH ASR ERRORS 6702
Yik-Cheung Tam, Jiacheng Xu, Zecheng Wang, Tinglong Liao, Shuhan Yuan, NYU Shanghai, China; Jiakai Zou, NYU, China

SPE-24.5: AN EMBARRASSINGLY SIMPLE MODEL FOR DIALOGUE RELATION EXTRACTION 6707
Fuzhao Xue, Aixin Sun, Hao Zhang, Jinjie Ni, Eng-Siong Chng, Nanyang Technological University, Singapore

SPE-24.6: A GAUSSIAN MIXTURE MODEL FOR DIALOGUE GENERATION WITH DYNAMIC PARAMETER SHARING STRATEGY 6712
Qingqing Zhu, Pengfei Wu, Zhouxing Tan, Jiaxin Duan, Fengyu Lu, Junfei Liu, Peking University, China

SPE-25: SPEAKER RECOGNITION IV: ATTENTION MECHANISM

SPE-25.1: ATTENTION BACK-END FOR AUTOMATIC SPEAKER VERIFICATION WITH MULTIPLE ENROLLMENT UTTERANCES 6717
Chang Zeng, Junichi Yamagishi, National Institute of Informatics, SOKENDAI, Japan; Xin Wang, Erica Cooper, Xiaoxiao Miao, National Institute of Informatics, Japan

SPE-25.2: SIMPLE ATTENTION MODULE BASED SPEAKER VERIFICATION WITH ITERATIVE NOISY LABEL DETECTION 6722
Xiaoyi Qin, Wuhan University, China; Na Li, Chao Weng, Dan Su, Tencent AI Lab, China; Ming Li, Duke Kunshan University, China

SPE-25.3: LOCAL INFORMATION MODELING WITH SELF-ATTENTION FOR SPEAKER VERIFICATION 6727
Bing Han, Zhengyang Chen, Yanmin Qian, Shanghai Jiao Tong University, China

SPE-25.4: MULTI-VIEW SELF-ATTENTION BASED TRANSFORMER FOR SPEAKER RECOGNITION 6732
Rui Wang, Zhihua Wei, Tongji University, China; Junyi Ao, Tom Ko, Yu Zhang, Southern University of Science and Technology, China; Long Zhou, Shujie Liu, Microsoft Research Asia, China; Qing Li, The Hong Kong Polytechnic University, Hong Kong

SPE-25.5: MULTI-QUERY MULTI-HEAD ATTENTION POOLING AND INTER-TOPK PENALTY FOR SPEAKER VERIFICATION 6737
Miao Zhao, Yufeng Ma, Yu Zheng, Min Liu, Minqiang Xu, SpeakIn Technologies Co. Ltd., China; Yiwei Ding, Fudan University, China

SPE-25.6: TEMPORAL DYNAMIC CONVOLUTIONAL NEURAL NETWORK FOR TEXT-INDEPENDENT SPEAKER VERIFICATION AND PHONEMIC ANALYSIS 6742
Seong-Hu Kim, Hyeonuk Nam, Yong-Hwa Park, Korea Advanced Institute of Science and Technology, Korea, Republic of

SPE-26: SPEECH RECOGNITION: GENERAL TOPICS II

SPE-26.1: EXPLOITING CROSS DOMAIN ACOUSTIC-TO-ARTICULATORY INVERTED FEATURES FOR DISORDERED SPEECH RECOGNITION 6747
Shujie Hu, Shansong Liu, Xurong Xie, Mengzhe Geng, Tianzi Wang, Shoukang Hu, Mingyu Cui, Xunying Liu, Helen Meng, The Chinese University of Hong Kong, Hong Kong

SPE-26.2: CONVERSATIONAL SPEECH RECOGNITION BY LEARNING CONVERSATION-LEVEL CHARACTERISTICS 6752
Kun Wei, Lei Xie, Audio, Speech and Language Processing Group, School of Computer Science, Northwestern Polytechnical University, China; Yike Zhang, Sining Sun, Long Ma, Tencent, xiaowei, China

SPE-26.3: EXPLORING MACHINE SPEECH CHAIN FOR DOMAIN ADAPTATION 6757
Fengpeng Yue, Tom Ko, Yu Zhang, Southern University of Science and Technology, China; Yan Deng, Lei He, Microsoft China, China

SPE-26.4: A LIKELIHOOD RATIO BASED DOMAIN ADAPTATION METHOD FOR E2E MODELS 6762
Chhavi Choudhury, Ankur Gandhe, Xiaohan Ding, Ivan Bulyko, Amazon, United States of America

SPE-26.5: RETRIEVING SPEAKER INFORMATION FROM PERSONALIZED ACOUSTIC MODELS FOR SPEECH RECOGNITION 6767
Salima Mdhaffar, Jean-François Bonastre, Natalia Tomashenko, Yannick Estève, LIA - University of Avignon, France; Marc Tommasi, Inria - University of Lille, France

SPE-26.6: NON-AUTOREGRESSIVE END-TO-END AUTOMATIC SPEECH RECOGNITION INCORPORATING DOWNSTREAM NATURAL LANGUAGE PROCESSING 6772
Motoi Omachi, Yuya Fujita, Yahoo Japan Corporation, Japan; Shinji Watanabe, Carnegie Mellon University, Japan; Tianzi Wang, Johns Hopkins University, Japan

SPE-27: VOICE CONVERSION I

SPE-27.1: TOWARD DEGRADATION-ROBUST VOICE CONVERSION 6777
Chien-yu Huang, Kai-Wei Chang, Hung-yi Lee, National Taiwan University, Taiwan

SPE-27.2: TEXT-FREE NON-PARALLEL MANY-TO-MANY VOICE CONVERSION USING NORMALISING FLOWS 6782
Thomas Merritt, Abdelhamid Ezzerg, Piotr Bilinski, Kamil Pokora, Roberto Barra-Chicote, Daniel Korzekwa, Amazon, United Kingdom of Great Britain and Northern Ireland; Magdalena Proszewska, Jagiellonian University, Poland

SPE-27.3: DIRECT NOISY SPEECH MODELING FOR NOISY-TO-NOISY VOICE CONVERSION 6787
Chao Xie, Yi-Chiao Wu, Patrick Lumban Tobing, Wen-Chin Huang, Tomoki Toda, Nagoya University, Japan

SPE-27.4: ONE-SHOT VOICE CONVERSION FOR STYLE TRANSFER BASED ON SPEAKER ADAPTATION 6792
Zhichao Wang, Qicong Xie, Tao Li, Hongqiang Du, Lei Xie, Northwestern Polytechnical University, China; Pengcheng Zhu, Mengxiao Bi, Fuxi AI Lab, NetEase Inc., China

SPE-27.5: CROSS-SPEAKER STYLE TRANSFER FOR TEXT-TO-SPEECH USING DATA AUGMENTATION 6797
Manuel Sam Ribeiro, Julian Roth, Giulia Comini, Goeric Huybrechts, Adam Gabryś, Jaime Lorenzo-Trueba, Amazon, Poland

SPE-27.6: AN INVESTIGATION OF STREAMING NON-AUTOREGRESSIVE SEQUENCE-TO-SEQUENCE VOICE CONVERSION 6802
Tomoki Hayashi, Kazuhiro Kobayashi, Tomoki Toda, Nagoya University, Japan

SPE-28: PRONUNCIATION ASSESSMENT

SPE-28.1: A UNIVERSAL ORDINAL REGRESSION FOR ASSESSING PHONEME-LEVEL PRONUNCIATION 6807
Shaoguang Mao, Frank Soong, Yan Xia, Jonathan Tien, Microsoft Research Asia, China

SPE-28.2: A TRANSFER LEARNING APPROACH FOR PRONUNCIATION SCORING 6812
Marcelo Sancinetti, Jazmín Vidal, Cyntia Bonomi, Universidad de Buenos Aires, Argentina; Luciana Ferrer, Consejo Nacional de Investigaciones y Técnicas, Argentina

SPE-28.3: EXPLORING NON-AUTOREGRESSIVE END-TO-END NEURAL MODELING FOR ENGLISH MISPRONUNCIATION DETECTION AND DIAGNOSIS 6817
Hsin-Wei Wang, Bi-Cheng Yan, Berlin Chen, National Taiwan Normal University, Taiwan; Hsuan-Sheng Chiu, Chunghwa Telecom Laboratories, Taiwan; Yung-Chang Hsu, EZAI, Taiwan

SPE-28.4: PHONEME MISPRONUNCIATION DETECTION BY JOINTLY LEARNING TO ALIGN 6822
Binghuai Lin, Liyuan Wang, Tencent Technology Co., Ltd, China

SPE-28.5: AN APPROACH TO MISPRONUNCIATION DETECTION AND DIAGNOSIS WITH ACOUSTIC, PHONETIC AND LINGUISTIC (APL) EMBEDDINGS 6827
Wenxuan Ye, Zhiyong Wu, Tsinghua University, China; Shaoguang Mao, Frank Soong, Wenshan Wu, Yan Xia, Jonathan Tien, Microsoft Research Asia, China

SPE-28.6: MASKED ACOUSTIC UNIT FOR MISPRONUNCIATION DETECTION AND CORRECTION 6832
Zhan Zhang, Yuehai Wang, Jianyi Yang, Zhejiang University, China

SPE-29: SPEECH ENHANCEMENT AND SEPARATION

SPE-29.1: INVESTIGATING SELF-SUPERVISED LEARNING FOR SPEECH ENHANCEMENT AND SEPARATION 6837
Zili Huang, Paola Garcia, Sanjeev Khudanpur, Johns Hopkins University, United States of America; Shinji Watanabe, Carnegie Mellon University, United States of America; Shu-wen Yang, National Taiwan University, United States of America

SPE-29.2: TFPSNET: TIME-FREQUENCY DOMAIN PATH SCANNING NETWORK FOR SPEECH SEPARATION 6842
Lei Yang, Wei Liu, Weiqin Wang, Samsung Research China – Beijing (SRC-B), China

SPE-29.3: EFFICIENT MONAURAL SPEECH SEPARATION WITH MULTISCALE TIME-DELAY SAMPLING 6847
Shuangqing Qian, Lijian Gao, Hongjie Jia, Qirong Mao, Jiangsu University, China

SPE-29.4: TOWARD MMWAVE-BASED SOUND ENHANCEMENT AND SEPARATION 6852
Muhammed Zahid Ozturk, K. J. Ray Liu, University of Maryland, College Park, United States of America; Chenshu Wu, The University of Hong Kong, Hong Kong; Beibei Wang, Origin Wireless Inc., United States of America

SPE-29.5: DPT-FSNET: DUAL-PATH TRANSFORMER BASED FULL-BAND AND SUB-BAND FUSION NETWORK FOR SPEECH ENHANCEMENT 6857
Feng Dang, Hangting Chen, Pengyuan Zhang, University of Chinese Academy of Sciences, China

SPE-29.6: REAL-M: TOWARDS SPEECH SEPARATION ON REAL MIXTURES 6862
Cem Subakan, Universite de Sherbrooke / Mila, Canada; Mirco Ravanelli, Concordia University / Mila, Canada; Samuele Cornell, Universita Politecnica delle Marche, Italy; Francois Grondin, Universite de Sherbrooke, Canada

SPE-30: LANGUAGE IDENTIFICATION AND LOW RESOURCE SPEECH RECOGNITION

SPE-30.1: SPOKEN LANGUAGE RECOGNITION WITH CLUSTER-BASED MODELING 6867
Stanisław Kacprzak, Magdalena Rybicka, Konrad Kowalczyk, Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, Poland

SPE-30.2: PHONOTACTIC LANGUAGE RECOGNITION USING A UNIVERSAL PHONEME RECOGNIZER AND A TRANSFORMER ARCHITECTURE 6872
David Romero, Christian Salamea, Universidad Politecnica Salesiana, Ecuador; Luis Fernando D'Haro, Marcos Estecha-Garitagoitia, Universidad Politécnica de Madrid, Spain

SPE-30.3: IMPROVED LANGUAGE IDENTIFICATION THROUGH CROSS-LINGUAL SELF-SUPERVISED LEARNING 6877
Andros Tjandra, Diptanu Gon Choudhury, Frank Zhang, Kritika Singh, Alexis Conneau, Alexei Baevski, Assaf Sela, Yatharth Saraf, Michael Auli, Facebook AI, United States of America

SPE-30.4: LANGUAGE ADAPTIVE CROSS-LINGUAL SPEECH REPRESENTATION LEARNING WITH SPARSE SHARING SUB-NETWORKS 6882
Yizhou Lu, Mingkun Huang, Xinghua Qu, Pengfei Wei, Zejun Ma, ByteDance AI Lab, China

SPE-30.5: INVESTIGATION OF ROBUSTNESS OF HUBERT FEATURES FROM DIFFERENT LAYERS TO DOMAIN, ACCENT AND LANGUAGE VARIATIONS 6887
Pratik Kumar, Vrunda N. Sukhadia, Srinivasan Umesh, Indian Institute of Technology Madras, India

SPE-30.6: COMBINING UNSUPERVISED AND TEXT AUGMENTED SEMI-SUPERVISED LEARNING FOR LOW RESOURCED AUTOREGRESSIVE SPEECH RECOGNITION 6892
Chak-Fai Li, Francis Keith, William Hartmann, Matthew Snover, Raytheon BBN Technologies, United States of America

SPE-31: EMOTION RECOGNITION: NEURAL ARCHITECTURE

SPE-31.1: KEY-SPARSE TRANSFORMER FOR MULTIMODAL SPEECH EMOTION RECOGNITION 6897
Weidong Chen, Xiaofeng Xing, Xiangmin Xu, Jichen Yang, South China University of Technology, China; Jianxin Pang, UBTECH Robotics Corp, China

SPE-31.2: NEURAL ARCHITECTURE SEARCH FOR SPEECH EMOTION RECOGNITION 6902
Xixin Wu, Shoukang Hu, Zhiyong Wu, Xunying Liu, Helen Meng, The Chinese University of Hong Kong, Hong Kong

SPE-31.3: MULTI-LINGUAL MULTI-TASK SPEECH EMOTION RECOGNITION USING WAV2VEC 2.0 6907
Mayank Sharma, Amazon, India

SPE-31.4: LIGHT-SERNET: A LIGHTWEIGHT FULLY CONVOLUTIONAL NEURAL NETWORK FOR SPEECH EMOTION RECOGNITION 6912
Arya Aftab, Shahrokh Ghaemmaghami, Sharif University of Technology, Iran (Islamic Republic of); Alireza Morsali, Benoit Champagne, McGill University, Canada

SPE-31.5: MULTIMODAL TRANSFORMER WITH LEARNABLE FRONTEND AND SELF ATTENTION FOR EMOTION RECOGNITION 6917
Soumya Dutta, Sriram Ganapathy, LEAP lab, Indian Institute of Science, Bangalore, India., India

SPE-31.6: SPEECH EMOTION RECOGNITION USING SELF-SUPERVISED FEATURES 6922
Edmilson Morais, Ron Hoory, Weizhong Zhu, Itai Gat, Matheus Damasceno, Hagai Aronowitz, IBM Research, Brazil

SPE-32: LANGUAGE DISORDERS: DETECTION II

SPE-32.1: USING ACOUSTIC DEEP NEURAL NETWORK EMBEDDINGS TO DETECT MULTIPLE SCLEROSIS FROM SPEECH 6927
Gábor Gosztolya, ELKH-SZTE Research Group on Artificial Intelligence, Hungary; László Tóth, University of Szeged, Hungary; Veronika Svindt, Ildikó Hoffmann, Research Center for Linguistics, Hungary; Judit Bóna, Eötvös Loránd University, Hungary

SPE-32.2: REPETITION ASSESSMENT FOR SPEECH AND LANGUAGE DISORDERS: A STUDY OF THE LOGOPENIC VARIANT OF PRIMARY PROGRESSIVE APHASIA 6932
R'mani Haulcy, James Glass, Massachusetts Institute of Technology, United States of America; Katerina Placek, Brian Tracey, Takeda Pharmaceutical Company, United States of America; Adam Vogel, University of Melbourne/Redenlab Inc, Australia

SPE-32.3: SPEECH TASKS RELEVANT TO SLEEPINESS DETERMINED WITH DEEP TRANSFER LEARNING 6937
Bang Tran, Youxiang Zhu, Xiaohui Liang, University of Massachusetts Boston, United States of America; James Schwoebel, Lindsay Warrenburg, Sonde Health Inc., United States of America

SPE-33: SPEECH ENHANCEMENT: TRAINING SCHEMES AND LOSSES

SPE-33.1: PHASE CONTINUITY: LEARNING DERIVATIVES OF PHASE SPECTRUM FOR SPEECH ENHANCEMENT 6942
Doyeon Kim, Hyewon Han, Hong-Goo Kang, Yonsei University, Korea, Republic of; Hyeon-Kyeong Shin, Soo-Whan Chung, Naver Corporation, Korea, Republic of

SPE-33.2: CONTINUAL SELF-TRAINING WITH BOOTSTRAPPED REMIXING FOR SPEECH ENHANCEMENT 6947
Efthymios Tzinis, University of Illinois at Urbana-Champaign, United States of America; Yossi Adi, Meta AI Research, Israel; Vamsi Ithapu, Buye Xu, Anurag Kumar, Meta Reality Labs Research, United States of America

SPE-33.3: ALLEVIATING THE LOSS-METRIC MISMATCH IN SUPERVISED SINGLE-CHANNEL SPEECH ENHANCEMENT 6952
Yang Yang, Hui Zhang, Xueliang Zhang, Huaiwen Zhang, Inner Mongolia University, China

SPE-33.4: A PRIORI SNR ESTIMATION FOR SPEECH ENHANCEMENT BASED ON PESQ-INDUCED REINFORCEMENT LEARNING 6957
Tong Lei, Haoxin Ruan, Kai Chen, Jing Lu, Key Laboratory of Modern Acoustics, Nanjing University; NJU-Horizon Intelligent Audio Lab, Horizon Robotics; Nanjing Institute of Advanced Artificial Intelligence., China

SPE-33.5: A TRAINING FRAMEWORK FOR STEREO-AWARE SPEECH ENHANCEMENT USING DEEP NEURAL NETWORKS 6962
Bahareh Tolooshams, Harvard University, United States of America; Kazuhito Koishida, Microsoft Corporation, United States of America

SPE-33.6: JOINT MAGNITUDE ESTIMATION AND PHASE RECOVERY USING CYCLE-IN-CYCLE GAN FOR NON-PARALLEL SPEECH ENHANCEMENT 6967
Guochen Yu, Communication University of China/Institute of Acoustics, Chinese Academy of Sciences, China; Andong Li, Chengshi Zheng, Institute of Acoustics, Chinese Academy of Sciences, China; Yutian Wang, Hui Wang, Communication University of China, China; Yinuo Guo, Bytedance, China

SPE-34: SPEECH RECOGNITION: GENERAL TOPICS III

SPE-34.1: PRIVACY ATTACKS FOR AUTOMATIC SPEECH RECOGNITION ACOUSTIC MODELS IN A FEDERATED LEARNING FRAMEWORK 6972
Natalia Tomashenko, Salima Mdhaffar, Yannick Estève, Jean-Francois Bonastre, University of Avignon, France; Marc Tommasi, University of Lille, France

SPE-34.2: VADOI: VOICE-ACTIVITY-DETECTION OVERLAPPING INFERENCE FOR END-TO-END LONG-FORM SPEECH RECOGNITION 6977
Jinhan Wang, University of California, Los Angeles, United States of America; Xiaosu Tong, Jinxi Guo, Di He, Roland Maas, Amazon, United States of America

SPE-34.3: TORCHAUDIO: BUILDING BLOCKS FOR AUDIO AND SPEECH PROCESSING 6982
Yao-Yuan Yang, University of California, San Diego, United States of America; Moto Hira, Zhaoheng Ni, Artyom Astafurov, Caroline Chen, Christian Puhrsch, Dmitriy Genzel, Donny Greenberg, Edward Yang, Jason Lian, Jeff Hwang, Ji Chen, Soumith Chintala, Vincent Quenneville-Bélair, Facebook AI, United States of America; David Pollack, Solvemate Gmbh, Germany; Peter Goldsborough, Anduril Industries, United States of America; Sean Narenthiran, Grid AI Labs, United Kingdom of Great Britain and Northern Ireland; Shinji Watanabe, Carnegie Mellon University, United States of America

SPE-34.4: UNSUPERVISED MODEL ADAPTATION FOR END-TO-END ASR 6987
Ganesh Sivaraman, Ricardo Casal, Matthew Garland, Elie Khoury, Pindrop, United States of America

SPE-34.5: SPEECH RECOGNITION USING BIOLOGICALLY-INSPIRED NEURAL NETWORKS 6992
Thomas Bohnstingl, Ayush Garg, Stanisław Woźniak, Evangelos Eleftheriou, Angeliki Pantazi, IBM Research Zurich, Switzerland; George Saon, IBM Research AI, United States of America

SPE-35: VOICE CONVERSION II

SPE-35.1: ASSEM-VC: REALISTIC VOICE CONVERSION BY ASSEMBLING MODERN SPEECH SYNTHESIS TECHNIQUES 6997
Kang-wook Kim, Junhyeok Lee, Myun-chul Joe, MINDsLab Inc., Korea, Republic of; Seung-won Park, Seoul National University, Korea, Republic of

SPE-35.2: MINIMIZING RESIDUALS FOR NATIVE-NONNATIVE VOICE CONVERSION IN A SPARSE, ANCHOR-BASED REPRESENTATION OF SPEECH 7002
Christopher Liberatore, Ricardo Gutierrez-Osuna, Texas A&M University, United States of America

SPE-35.3: IMPROVING RECOGNITION-SYNTHESIS BASED ANY-TO-ONE VOICE CONVERSION WITH CYCLIC TRAINING 7007
Yannian Chen, Zhenhua Ling, National Engineering Laboratory for Speech and Language Information Processing, University of Science and Technology of China, Hefei, P.R.China, China; Lijuan Liu, Yajun Hu, Yuan Jiang, iFLYTEK Research, iFLYTEK Co., Ltd., Hefei, P.R.China, China

SPE-35.4: NVC-NET: END-TO-END ADVERSARIAL VOICE CONVERSION 7012
Bac Nguyen, Fabien Cardinaux, Sony Europe B.V., Germany

SPE-35.5: U-GAT-VC: UNSUPERVISED GENERATIVE ATTENTIONAL NETWORKS FOR NON-PARALLEL VOICE CONVERSION 7017
Sheng Shi, Yangzhou Du, Jianping Fan, Lenovo Research, China; Jiahao Shao, Tsinghua University, China; Yifei Hao, University of Southern California, China

SPE-35.6: DISENTANGLING CONTENT AND FINE-GRAINED PROSODY INFORMATION VIA HYBRID ASR BOTTLENECK FEATURES FOR VOICE CONVERSION 7022
Xintao Zhao, Changhe Song, Zhiyong Wu, Tsinghua University, China; Feng Liu, Shiyin Kang, Deyi Tuo, Huya Inc, China; Helen Meng, The Chinese University of Hong Kong, China

SPE-36: DIALOG SYSTEM III: EMOTION RECOGNITION AND PARALINGUISTIC MODELING

SPE-36.1: A COMMONSENSE KNOWLEDGE ENHANCED NETWORK WITH RETROSPECTIVE LOSS FOR EMOTION RECOGNITION IN SPOKEN DIALOG 7027
Yunhe Xie, Chengjie Sun, Zhenzhou Ji, Harbin Institute of Technology, China

SPE-36.2: HIERARCHICAL AND MULTI-VIEW DEPENDENCY MODELLING NETWORK FOR CONVERSATIONAL EMOTION RECOGNITION 7032
Yu-Ping Ruan, Shu-Kai Zheng, Taihao Li, Fen Wang, Guanxiong Pei, Zhejiang Lab, China

SPE-36.3: MM-DFN: MULTIMODAL DYNAMIC FUSION NETWORK FOR EMOTION RECOGNITION IN CONVERSATIONS 7037
Dou Hu, Xiaolong Hou, Lianxin Jiang, Ping An Life Insurance Company of China, Ltd., China; Lingwei Wei, Institute of Information Engineering, Chinese Academy of Sciences, China; Yang Mo, Ping An Life Insurance Company of China, Ltd, China

SPE-36.4: MODELING INTENTION, EMOTION AND EXTERNAL WORLD IN DIALOGUE SYSTEMS 7042
Wei Peng, Yue Hu, Luxi Xing, Yuqiang Xie, Xingsheng Zhang, Yajing Sun, Institute of Information Engineering, Chinese Academy of Sciences, China

SPE-36.5: A NEURAL PROSODY ENCODER FOR END-TO-END DIALOGUE ACT CLASSIFICATION 7047
Kai Wei, Martin Radfar, Thanh Tran, Markus Mueller, Grant Strimel, Nathan Susanj, Athanasios Mouchtaris, Maurizio Omologo, Alexa Speech, Amazon, United States of America; Dillon Knox, University of Southern California, United States of America

SPE-36.6: IMPROVING CONTEXTUAL COHERENCE IN VARIATIONAL PERSONALIZED AND EMPATHETIC DIALOGUE AGENTS 7052
*Jing Yang Lee, Woon Seng Gan, Nanyang Technological University, Singapore; Kong Aik Lee, A*STAR, Singapore*

SPE-37: SPEAKER RECOGNITION V: APPLICATIONS

SPE-37.1: FUSION AND ORTHOGONAL PROJECTION FOR IMPROVED FACE-VOICE ASSOCIATION 7057
Muhammad Saad Saeed, Muhammad Haroon Yousaf, University of Engineering and Technology Taxila, Pakistan; Muhammad Haris Khan, Mohamed Bin Zayed University of Artificial Intelligence, United Arab Emirates; Shah Nawaz, Alessio Del Bue, Istituto Italiano di Tecnologia, Italy

SPE-37.2: OPENFEAT: IMPROVING SPEAKER IDENTIFICATION BY OPEN-SET FEW-SHOT EMBEDDING ADAPTATION WITH TRANSFORMER 7062
Kishan K C, Rochester Institute of Technology, United States of America; Zhenning Tan, Long Chen, Minho Jin, Eunjung Han, Andreas Stolcke, Chul Lee, Amazon Alexa AI, United States of America

SPE-37.3: TOWARDS LIGHTWEIGHT APPLICATIONS: ASYMMETRIC ENROLL-VERIFY STRUCTURE FOR SPEAKER VERIFICATION 7067
Qingjian Lin, Lin Yang, Xuyang Wang, Junjie Wang, Lenovo, China; Xiaoyi Qin, Ming Li, Duke Kunshan University, China

SPE-37.4: SPEAKER EMBEDDING CONVERSION FOR BACKWARD AND CROSS-CHANNEL COMPATIBILITY 7072
Tianxiang Chen, Elie Khoury, Pindrop, United States of America

SPE-37.5: IMPROVING FAIRNESS IN SPEAKER VERIFICATION VIA GROUP-ADAPTED FUSION NETWORK 7077
Hua Shen, The Pennsylvania State University, United States of America; Yuguang Yang, Guoli Sun, Ryan Langman, Eunjung Han, Jasha Droppo, Andreas Stolcke, Amazon, United States of America

SPE-37.6: CS-REP: MAKING SPEAKER VERIFICATION NETWORKS EMBRACING RE-PARAMETERIZATION 7082
Ruiteng Zhang, Wenhuan Lu, Junhai Xu, Tianjin University, China; Jianguo Wei, Tianjin University; Qinghai Nationalities University, China; Lin Zhang, National Institute of Informatics, Japan; Yantao Ji, Xi'an Jiaotong University, China; Xugang Lu, National Institute of Information and Communications Technology, Japan

SPE-38: SELF SUPERVISED LEARNING FOR SPEECH RECOGNITION

SPE-38.1: DISTILHUBERT: SPEECH REPRESENTATION LEARNING BY LAYER-WISE DISTILLATION OF HIDDEN-UNIT BERT 7087
Heng-Jui Chang, Shu-wen Yang, Hung-yi Lee, National Taiwan University, Taiwan

SPE-38.2: IMPROVING SELF-SUPERVISED LEARNING FOR SPEECH RECOGNITION WITH INTERMEDIATE LAYER SUPERVISION 7092
Chengyi Wang, Zhenglu Yang, Nankai University, China; Yu Wu, Sanyuan Chen, Shujie Liu, Jinyu Li, Yao Qian, Microsoft, China

SPE-38.3: WAV2VEC-SWITCH: CONTRASTIVE LEARNING FROM ORIGINAL-NOISY SPEECH PAIRS FOR ROBUST SPEECH RECOGNITION 7097
Yiming Wang, Jinyu Li, Yao Qian, Chengyi Wang, Yu Wu, Microsoft Corporation, United States of America; Heming Wang, The Ohio State University, United States of America

SPE-38.4: EFFICIENT ADAPTER TRANSFER OF SELF-SUPERVISED SPEECH MODELS FOR AUTOMATIC SPEECH RECOGNITION 7102
Bethan Thomas, Salah Karout, Huawei R&D UK, United Kingdom of Great Britain and Northern Ireland; Samuel Kessler, University of Oxford, United Kingdom of Great Britain and Northern Ireland

SPE-38.5: AN EXPLORATION OF HUBERT WITH LARGE NUMBER OF CLUSTER UNITS AND MODEL ASSESSMENT USING BAYESIAN INFORMATION CRITERION 7107
Takashi Maekaku, Yuya Fujita, Yahoo Japan Corporation, Japan; Xuankai Chang, Shinji Watanabe, Carnegie Mellon University, United States of America

SPE-38.6: OPTIMIZE WAV2VEC2S ARCHITECTURE FOR SMALL TRAINING SET THROUGH ANALYZING ITS PRE-TRAINED MODELS ATTENTION PATTERN 7112
Liu Chen, Meysam Asgari, Hiroko Dodge, Oregon Health & Science University, United States of America

SPE-39: SPEECH SYNTHESIS: FRONT-END

SPE-39.1: PART-OF-SPEECH MODELS COMPRESSION METHODS FOR ON-DEVICE GRAPHEME-TO-PHONEME CONVERSION 7117
Marek Kubis, Paweł Skórzewski, Adam Mickiewicz University in Poznań, Poland; Maxime Méloux, Marcin Lewandowski, Samsung R&D Institute Poland, Poland; Gunu Jho, Hyoungmin Park, Samsung Electronics, Korea, Republic of

SPE-39.2: AN END-TO-END CHINESE TEXT NORMALIZATION MODEL BASED ON RULE-GUIDED FLAT-LATTICE TRANSFORMER 7122
Wenlin Dai, Changhe Song, Xiang Li, Zhiyong Wu, Tsinghua University, China; Huashan Pan, Xiulin Li, Databaker Technology Co., Ltd, China; Helen Meng, The Chinese University of Hong Kong, China

SPE-39.3: CHINESE SPELLING TEXT GENERATION OF MATHEMATICAL FORMULAS 7127
Su Dong, Sicen Liu, Buzhou Tang, Harbin Institute of Technology, China; Shan Liu, Tencent Technology Company, China

SPE-39.4: POLYPHONE DISAMBIGUATION AND ACCENT PREDICTION USING PRE-TRAINED LANGUAGE MODELS IN JAPANESE TTS FRONT-END 7132
Rem Hida, Masaki Hamada, Emiru Tsunoo, Toshiyuki Sekiya, Sony Group Corporation, Japan; Chie Kamada, Toshiyuki Kumakura, Sony Corporation of America, United States of America

SPE-39.6: DATA AUGMENTATION FOR LONG-TAILED AND IMBALANCED POLYPHONE DISAMBIGUATION IN MANDARIN 7137
Yang Zhang, Haitong Zhang, Yue Lin, NetEase Games AI Lab, China

SPE-40: LANGUAGE UNDERSTANDING I: END-TO-END FRAMEWORK

SPE-40.1: LEVERAGING BILINEAR ATTENTION TO IMPROVE SPOKEN LANGUAGE UNDERSTANDING 7142
Dongsheng Chen, Zhiqi Huang, Yuexian Zou, Peking University, China

SPE-40.2: BUILDING ROBUST SPOKEN LANGUAGE UNDERSTANDING BY CROSS ATTENTION BETWEEN PHONEME SEQUENCE AND ASR HYPOTHESIS 7147
Zexun Wang, Yuming Zhao, Mingchao Feng, Meng Chen, Xiaodong He, JD AI, China; Yuquan Le, Hunan University, China; Yi Zhu, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

SPE-40.3: INTEGRATION OF PRE-TRAINED NETWORKS WITH CONTINUOUS TOKEN INTERFACE FOR END-TO-END SPOKEN LANGUAGE UNDERSTANDING 7152
Seunghyun Seo, Donghyun Kwak, Naver Corporation, Korea, Republic of; Bowon Lee, Inha University, Korea, Republic of

SPE-40.4: TIE YOUR EMBEDDINGS DOWN: CROSS-MODAL LATENT SPACES FOR END-TO-END SPOKEN LANGUAGE UNDERSTANDING 7157
Bhuvan Agrawal, Markus Müller, Samridhi Choudhary, Martin Radfar, Athanasios Mouchtaris, Ross McGowan, Nathan Susanj, Siegfried Kunzmann, Amazon, United States of America

SPE-40.5: IMPROVING END-TO-END MODELS FOR SET PREDICTION IN SPOKEN LANGUAGE UNDERSTANDING 7162
Hong-Kwang Kuo, Zoltan Tuske, Samuel Thomas, Brian Kingsbury, George Saon, IBM Research AI, United States of America

SPE-40.6: ESPNET-SLU: ADVANCING SPOKEN LANGUAGE UNDERSTANDING THROUGH ESPNET 7167
Siddhant Arora, Siddharth Dalmia, Xuankai Chang, Yushi Ueda, Yifan Peng, Sujay Kumar, Karthik Ganesan, Brian Yan, Alan W Black, Shinji Watanabe, Carnegie Mellon University, United States of America; Pavel Denisov, Ngoc Thang Vu, University of Stuttgart, Germany; Yuekai Zhang, Zoom Video Communications, China

SPE-41: SPEAKER RECOGNITION VI: DOMAIN ADAPTATION

SPE-41.1: THE CORAL+ + ALGORITHM FOR UNSUPERVISED DOMAIN ADAPTATION OF SPEAKER RECOGNITION 7172
Rongjin Li, Weibin Zhang, Dongpeng Chen, VoiceAI Technologies, Co. Ltd. Shenzhen, China, China

SPE-41.2: LEARNING DOMAIN-INVARIANT TRANSFORMATION FOR SPEAKER VERIFICATION 7177
*Hanyi Zhang, Longbiao Wang, Meng Liu, Jianwu Dang, Hui Chen, Tianjin University, China; Kong Aik Lee, A*STAR, Singapore*

SPE-41.3: DOMAIN ROBUST DEEP EMBEDDING LEARNING FOR SPEAKER RECOGNITION 7182
Hang-Rui Hu, Yan Song, Ying Liu, Li-Rong Dai, University of Science and Technology of China, China; Ian McLoughlin, Singapore Institute of Technology, Singapore; Lin Liu, iFLYTEK CO. LTD., China

SPE-41.4: TACKLING THE SCORE SHIFT IN CROSS-LINGUAL SPEAKER VERIFICATION BY EXPLOITING LANGUAGE INFORMATION 7187
Jenthe Thienpondt, Brecht Desplanques, Kris Demuynck, Ghent University, Belgium

SPE-41.5: DOMAIN ADAPTATION FOR SPEAKER RECOGNITION IN SINGING AND SPOKEN VOICE 7192
Anurag Chowdhury, Austin Cozzo, Arun Ross, Michigan State University, United States of America

SPE-41.6: CDMA: CROSS-DOMAIN DISTANCE METRIC ADAPTATION FOR SPEAKER VERIFICATION 7197
Jianchen Li, Jiqing Han, Hongwei Song, Harbin Institute of Technology, China

SPE-42: SPEECH RECOGNITION: ACOUSTIC MODELING I

SPE-42.1: WORD ORDER DOES NOT MATTER FOR SPEECH RECOGNITION 7202
Vineel Pratap Konduru, Qiantong Xu, Tatiana Likhomanenko, Gabriel Synnaeve, Ronan Collobert, Facebook AI, United States of America

SPE-42.2: CONTRASTIVE SIAMESE NETWORK FOR SEMI-SUPERVISED SPEECH RECOGNITION 7207
Soheil Khorram, Jaeyoung Kim, Anshuman Tripathi, Han Lu, Qian Zhang, Hasim Sak, Google Inc. USA, United States of America

SPE-42.3: SEQUENCE TRANSDUCTION WITH GRAPH-BASED SUPERVISION 7212
Niko Moritz, Facebook AI, United Kingdom of Great Britain and Northern Ireland; Takaaki Hori, Jonathan Le Roux, MERL, United States of America; Shinji Watanabe, Carnegie Mellon University, United States of America

SPE-42.4: SPEECHMOE2: MIXTURE-OF-EXPERTS MODEL WITH IMPROVED ROUTING 7217
Zhao You, Shulin Feng, Dan Su, Dong Yu, Tencent, China

SPE-42.5: SUPERVISED ATTENTION IN SEQUENCE-TO-SEQUENCE MODELS FOR SPEECH RECOGNITION 7222
Gene-Ping Yang, Hao Tang, University of Edinburgh, United Kingdom of Great Britain and Northern Ireland

SPE-42.6: END-TO-END SPEECH RECOGNITION FROM FEDERATED ACOUSTIC MODELS 7227
Yan Gao, Pedro P. B. de Gusmao, Nicholas D. Lane, University of Cambridge, United Kingdom of Great Britain and Northern Ireland; Titouan Parcollet, Avignon University, France; Salah Zaiem, Telecom Paris, France; Javier Fernandez-Marques, University of Oxford, United Kingdom of Great Britain and Northern Ireland; Daniel J. Beutel, Adap GmbH; University of Cambridge, Germany

SPE-43: SPEECH SYNTHESIS: SINGING VOICE AND OTHERS

SPE-43.1: HIFIDENOISE: HIGH-FIDELITY DENOISING TEXT TO SPEECH WITH ADVERSARIAL NETWORKS 7232
Lichao Zhang, Yi Ren, Zhou Zhao, Zhejiang University, China; Liqun Deng, Huawei Noah's Ark Lab, China

SPE-43.2: VISINGER: VARIATIONAL INFERENCE WITH ADVERSARIAL LEARNING FOR END-TO-END SINGING VOICE SYNTHESIS 7237
Yongmao Zhang, Jian Cong, Heyang Xue, Lei Xie, Northwestern Polytechnical University, China; Pengcheng Zhu, Mengxiao Bi, Fuxi AI Lab, NetEase Inc., China

SPE-43.3: A MELODY-UNSUPERVISION MODEL FOR SINGING VOICE SYNTHESIS 7242
Soonbeom Choi, Juhan Nam, Korea Advanced Institute of Science and Technology, Korea, Republic of

SPE-43.4: TRANSFORMER-S2A: ROBUST AND EFFICIENT SPEECH-TO-ANIMATION 7247
Liyang Chen, Zhiyong Wu, Shenzhen International Graduate School, China; Jun Ling, Shanghai Jiao Tong University, China; Runnan Li, Xu Tan, Sheng Zhao, Microsoft, China, China

SPE-43.5: VCVTS: MULTI-SPEAKER VIDEO-TO-SPEECH SYNTHESIS VIA CROSS-MODAL KNOWLEDGE TRANSFER FROM VOICE CONVERSION 7252
Disong Wang, Xunying Liu, Helen Meng, The Chinese University of Hong Kong, Hong Kong; Shan Yang, Dan Su, Dong Yu, Tencent AI Lab, China

SPE-44: LANGUAGE ASSESSMENT / TRANSCRIPTION AUGMENTATION

SPE-44.1: FAST TASK-SPECIFIC ADAPTATION IN SPOKEN LANGUAGE ASSESSMENT WITH META-LEARNING 7257
Binghuai Lin, Liyuan Wang, Tencent Technology Co., Ltd, China

SPE-44.2: TRANSFORMER-BASED MULTI-ASPECT MULTI-GRANULARITY NON-NATIVE ENGLISH SPEAKER PRONUNCIATION ASSESSMENT 7262
Yuan Gong, James Glass, Massachusetts Institute of Technology, United States of America; Ziyi Chen, Iek-Heng Chu, Peng Chang, PAII Inc., United States of America

SPE-44.3: A MODEL FOR ASSESSOR BIAS IN AUTOMATIC PRONUNCIATION ASSESSMENT 7267
Jose Antonio Lopez Saenz, Thomas Hain, University of Sheffield, United Kingdom of Great Britain and Northern Ireland

SPE-44.4: UNIFIED MULTIMODAL PUNCTUATION RESTORATION FRAMEWORK FOR MIXED-MODALITY CORPUS 7272
Yaoming Zhu, Liwei Wu, Shanbo Cheng, Mingxuan Wang, Bytedance, China

SPE-44.5: PUNCTUATION PREDICTION FOR STREAMING ON-DEVICE SPEECH RECOGNITION 7277
Zhikai Zhou, Yanmin Qian, Shanghai Jiao Tong University, China; Tian Tan, AISpeech Ltd., China

SPE-44.6: ASR ERROR CORRECTION WITH DUAL-CHANNEL SELF-SUPERVISED LEARNING 7282
Fan Zhang, Jinyao Yan, Communication Unversity of China, China; Mei Tu, Song Liu, Samsung Research China - Beijing (SRCB), China

SPE-45: SPEECH ENHANCEMENT: SPEECH EXTRACTION AND AUDIO-VISUAL ENHANCEMENT

SPE-45.1: L-SPEX: LOCALIZED TARGET SPEAKER EXTRACTION 7287
Meng Ge, Longbiao Wang, Jianwu Dang, Tianjin University, China; Chenglin Xu, Kuaishou Technology, China; Eng Siong Chng, Nanyang Technological University, Singapore; Haizhou Li, National University of Singapore, Singapore

SPE-45.3: DPCCN: DENSELY-CONNECTED PYRAMID COMPLEX CONVOLUTIONAL NETWORK FOR ROBUST SPEECH SEPARATION AND EXTRACTION 7292
Jiangyu Han, Yanhua Long, Shanghai Normal University, China; Lukas Burget, Jan Cernocky, Brno University of Technology, Czechia

SPE-45.4: MIXED PRECISION DNN QUANTIZATION FOR OVERLAPPED SPEECH SEPARATION AND RECOGNITION 7297
Junhao Xu, Xunying Liu, Helen Mei-Ling Meng, The Chinese University of Hong Kong, China; Jianwei Yu, Tencent AI Lab, China

SPE-45.5: THE IMPACT OF REMOVING HEAD MOVEMENTS ON AUDIO-VISUAL SPEECH ENHANCEMENT 7302
Zhiqi Kang, Radu Horaud, Xavier Alameda-Pineda, Inria Grenoble Rhône-Alpes & Univ. Grenoble Alpes, France, France; Mostafa Sadeghi, Inria Nancy Grand-Est, France, France; Jacob Donley, Anurag Kumar, Facebook Reality Labs Research, Redmond WA, USA, United States of America

SPE-45.6: VSEGAN: VISUAL SPEECH ENHANCEMENT GENERATIVE ADVERSARIAL NETWORK 7307
Xinmeng Xu, Trinity College Dublin, China; Yang Wang, Dongxiang Xu, Yiyuan Peng, Cong Zhang, Jie Jia, Binbin Chen, Vivo Communication Technology Co. Ltd., China

SPE-46: MULTI SPEAKER AND MULTI-CHANNEL SPEECH RECOGNITION AND PROCESSING

SPE-46.1: ENDPOINT DETECTION FOR STREAMING END-TO-END MULTI-TALKER ASR 7312
Liang Lu, Otter.ai, United States of America; Jinyu Li, Yifan Gong, Microsoft, USA, United States of America

SPE-46.2: CONTINUOUS STREAMING MULTI-TALKER ASR WITH DUAL-PATH TRANSDUCERS 7317
Desh Raj, Johns Hopkins University, United States of America; Liang Lu, Zhuo Chen, Yashesh Gaur, Jinyu Li, Microsoft Corp., United States of America

SPE-46.3: EXTENDED GRAPH TEMPORAL CLASSIFICATION FOR MULTI-SPEAKER END-TO-END ASR 7322
Xuankai Chang, Shinji Watanabe, Carnegie Mellon University, United States of America; Niko Moritz, Facebook AI, United Kingdom of Great Britain and Northern Ireland; Takaaki Hori, Jonathan Le Roux, Mitsubishi Electric Research Laboratories (MERL), United States of America

SPE-46.4: ADA-VAD: UNPAIRED ADVERSARIAL DOMAIN ADAPTATION FOR NOISE-ROBUST VOICE ACTIVITY DETECTION 7327
Taesoo Kim, Jong Hwan Ko, Sungkyunkwan university, Korea, Republic of; Jiho Chang, Korea Research Institute of Standards and Science, Korea, Republic of

SPE-46.5: MULTI-CHANNEL END-TO-END NEURAL DIARIZATION WITH DISTRIBUTED MICROPHONES 7332
Shota Horiguchi, Yuki Takashima, Yohei Kawaguchi, Hitachi, Ltd., Japan; Paola Garcia, Johns Hopkins University, United States of America; Shinji Watanabe, Carnegie Mellon University, United States of America

SPE-46.6: MULTI-CHANNEL SPEAKER DIARIZATION USING SPATIAL FEATURES FOR MEETINGS 7337
Naijun Zheng, XunYing Liu, Helen Meng, The Chinese University of Hong Kong, Hong Kong; Na Li, Jianwei Yu, Chao Weng, Dan Su, Tencent AI Lab, Hong Kong

SPE-47: EMOTION RECOGNITION: GENERAL TOPICS I

SPE-47.1: SPEAKER NORMALIZATION FOR SELF-SUPERVISED SPEECH EMOTION RECOGNITION 7342
Itai Gat, Hagai Aronowitz, Weizhong Zhu, Edmilson Morais, Ron Hoory, IBM Research, Israel

SPE-47.2: SENTIMENT-AWARE AUTOMATIC SPEECH RECOGNITION PRE-TRAINING FOR ENHANCED SPEECH EMOTION RECOGNITION 7347
Ayoub Ghriss, University of Colorado, Boulder, United States of America; Bo Yang, Viktor Rozgic, Wang Chao, Amazon LLC, United States of America; Elizabeth Shriberg, Ellipsis Health, United States of America

SPE-47.3: CONFIDENCE ESTIMATION FOR SPEECH EMOTION RECOGNITION BASED ON THE RELATIONSHIP BETWEEN EMOTION CATEGORIES AND PRIMITIVES 7352
Yang Li, University of Rochester, United States of America; Constantinos Papayiannis, Viktor Rozgic, Elizabeth Shriberg, Chao Wang, Amazon, United States of America

SPE-47.4: AUXFORMER: ROBUST APPROACH TO AUDIOVISUAL EMOTION RECOGNITION 7357
Lucas Goncalves, Carlos Busso, The University of Texas at Dallas, United States of America

SPE-47.5: FUSING ASR OUTPUTS IN JOINT TRAINING FOR SPEECH EMOTION RECOGNITION 7362
Yuanchao Li, Peter Bell, Catherine Lai, University of Edinburgh, United Kingdom of Great Britain and Northern Ireland

SPE-47.6: SPEECH EMOTION RECOGNITION WITH CO-ATTENTION BASED MULTI-LEVEL ACOUSTIC INFORMATION 7367
Heqing Zou, Chen Chen, Deepu Rajan, Eng Siong Chng, Nanyang Technological University, Singapore; Yuke Si, Tianjin University, China

SPE-48: LANGUAGE DISORDERS: RECOGNITION

SPE-48.1: MULTI-MODAL ACOUSTIC-ARTICULATORY FEATURE FUSION FOR DYSARTHRIC SPEECH RECOGNITION 7372
Zhengjun Yue, Heidi Christensen, Jon Barker, The University of Sheffield, United Kingdom of Great Britain and Northern Ireland; Erfan Loweimi, Zoran Cvetkovic, King's College London, United Kingdom of Great Britain and Northern Ireland

SPE-48.2: RAW SOURCE AND FILTER MODELLING FOR DYSARTHRIC SPEECH RECOGNITION 7377
Zhengjun Yue, The University of Sheffield, United Kingdom of Great Britain and Northern Ireland; Erfan Loweimi, Zoran Cvetkovic, King's College London, United Kingdom of Great Britain and Northern Ireland

SPE-48.3: SYNTHESIZING DYSARTHRIC SPEECH USING MULTI-SPEAKER TTS FOR DYSARTHRIC SPEECH RECOGNITION 7382
Mohammad Soleymanpour, Michael T. Johnson, University of Kentucky, United States of America; Rahim Soleymanpour, University of Connecticut, United States of America; Jeffrey Berry, Marquette University, United States of America

SPE-48.4: TOWARDS INTERPRETING DEEP LEARNING MODELS TO UNDERSTAND LOSS OF SPEECH INTELLIGIBILITY IN SPEECH DISORDERS STEP 2: CONTRIBUTION OF THE EMERGENCE OF PHONETIC TRAITS 7387
Sondes Abderrazek, Corinne Fredouille, LIA - Avignon Université, France; Alain Ghio, Muriel Lalain, Christine Meunier, LPL, Aix-Marseille Université, France; Virginie Woisard, UT2J, Octogone-Lordat, Toulouse University, Toulouse Hospital, France

SPE-48.5: CONSTANT Q CEPSTRAL COEFFICIENTS FOR CLASSIFICATION OF NORMAL VS. PATHOLOGICAL INFANT CRY 7392
Hemant A. Patil, Ankur T. Patil, Aastha Kachhi, Dhirubhai Ambani Institute of Information and Communication Technology, India

SPE-48.6: NONVERBAL SOUND DETECTION FOR DISORDERED SPEECH 7397
Colin Lea, Zifang Huang, Lauren Tooley, Zeinab Liaghat, Shri Thelapurath, Leah Findlater, Jeffrey P. Bigham, Apple, United States of America; Dhruv Jain, University of Washington, United States of America

SPE-49: SPEECH ENHANCEMENT AND DEREVERBERATION

SPE-49.1: CONDITIONAL DIFFUSION PROBABILISTIC MODEL FOR SPEECH ENHANCEMENT 7402
Yen-Ju Lu, Zhong-Qiu Wang, Shinji Watanabe, Carnegie Mellon University, United States of America; Alexander Richard, Facebook, United States of America; Cheng Yu, Yu Tsao, Academia Sinica, Taiwan

SPE-49.2: DEEPFILTERNET: A LOW COMPLEXITY SPEECH ENHANCEMENT FRAMEWORK FOR FULL-BAND AUDIO BASED ON DEEP FILTERING 7407
Hendrik Schröter, Andreas Maier, Friedrich-Alexander University Erlangen-Nuremberg (FAU), Germany; Alberto N. Escalante-B., Tobias Rosenkranz, WS Audiology, Germany

SPE-49.3: METRICGAN-U: UNSUPERVISED SPEECH ENHANCEMENT/ DEREVERBERATION BASED ONLY ON NOISY/ REVERBERATED SPEECH 7412
Szu-Wei Fu, Microsoft, United States of America; Cheng Yu, Kuo-Hsuan Hung, Yu Tsao, Academia Sinica, United States of America; Mirco Ravanelli, Mila-Quebec AI Institute, United States of America

SPE-49.4: UFORMER: A UNET BASED DILATED COMPLEX & REAL DUAL-PATH CONFORMER NETWORK FOR SIMULTANEOUS SPEECH ENHANCEMENT AND DEREVERBERATION 7417
Yihui Fu, Shubo Lv, Yukai Jv, Lei Xie, Audio, Speech and Language Processing Group (ASLP@NPU), Northwestern Polytechnical University, Xi'an, China, China; Yun Liu, Jingdong Li, Dawei Luo, AI Interaction Division, Sogou Inc., Beijing, China, China

SPE-49.5: ATTENUATION OF ACOUSTIC EARLY REFLECTIONS IN TELEVISION STUDIOS USING PRETRAINED SPEECH SYNTHESIS NEURAL NETWORK 7422
Tomer Rosenbaum, Israel Cohen, Technion – Israel Institute of Technology, Israel; Emil Winebrand, Insoundz Ltd., Israel

SPE-50: SPEECH RECOGNITION: ACOUSTIC MODELING II

SPE-50.1: NON-AUTOREGRESSIVE ASR WITH SELF-CONDITIONED FOLDED ENCODERS 7427
Tatsuya Komatsu, LINE Corporation, Japan

SPE-50.2: HAVE BEST OF BOTH WORLDS: TWO-PASS HYBRID AND E2E CASCADING FRAMEWORK FOR SPEECH RECOGNITION 7432
Guoli Ye, Vadim Mazalov, Jinyu Li, Yifan Gong, Microsoft, United States of America

SPE-50.3: CONFORMER-BASED HYBRID ASR SYSTEM FOR SWITCHBOARD DATASET 7437
Mohammad Zeineldeen, Christoph Lüscher, Wilfried Michel, Ralf Schlüter, Hermann Ney, RWTH Aachen University / AppTek, Germany; Jingjing Xu, Alexander Gerstenberger, RWTH Aachen University, Germany

SPE-50.4: IMPROVING FACTORED HYBRID HMM ACOUSTIC MODELING WITHOUT STATE TYING 7442
Tina Raissi, Ralf Schlüter, RWTH Aachen University, Germany; Eugen Beck, Apptek GmbH, Germany; Hermann Ney, RWTH Aachen Univeristy, Germany

SPE-50.5: AUDITORY-BASED DATA AUGMENTATION FOR END-TO-END AUTOMATIC SPEECH RECOGNITION 7447
Zehai Tu, Jack Deadman, Ning Ma, Jon Barker, University of Sheffield, United Kingdom of Great Britain and Northern Ireland

SPE-50.6: DELIBERATION OF STREAMING RNN-TRANSDUCER BY NON-AUTOREGRESSIVE DECODING 7452
Weiran Wang, Ke Hu, Tara Sainath, Google, United States of America

SPE-51: SPEECH SYNTHESIS: NOVEL ACOUSTIC MODELS

SPE-51.1: NEURAL HMMS ARE ALL YOU NEED (FOR HIGH-QUALITY ATTENTION-FREE TTS) 7457
Shivam Mehta, Éva Székely, Jonas Beskow, Gustav Eje Henter, KTH Royal Institute of Technology, Sweden

SPE-51.2: AUTOREGRESSIVE VARIATIONAL AUTOENCODER WITH A HIDDEN SEMI-MARKOV MODEL-BASED STRUCTURED ATTENTION FOR SPEECH SYNTHESIS 7462
Takato Fujimoto, Kei Hashimoto, Yoshihiko Nankaku, Keiichi Tokuda, Nagoya Institute of Technology, Japan

SPE-51.3: PAMA-TTS: PROGRESSION-AWARE MONOTONIC ATTENTION FOR STABLE SEQ2SEQ TTS WITH ACCURATE PHONEME DURATION CONTROL 7467
Yunchao He, Jian Luan, Yujun Wang, Xiaomi Corporation, China

SPE-51.4: IMPROVING FASTSPEECH TTS WITH EFFICIENT SELF-ATTENTION AND COMPACT FEED-FORWARD NETWORK 7472
Yujia Xiao, Xi Wang, Lei He, Frank K. Soong, Microsoft, China

SPE-51.5: VARIANCEFLOW: HIGH-QUALITY AND CONTROLLABLE TEXT-TO-SPEECH USING VARIANCE INFORMATION VIA NORMALIZING FLOW 7477
Yoonhyung Lee, Kyomin Jung, Seoul National University, Korea, Republic of; Jinhyeok Yang, NCSOFT, Korea, Republic of

SPE-51.6: MIXER-TTS: NON-AUTOREGRESSIVE, FAST AND COMPACT TEXT-TO-SPEECH MODEL CONDITIONED ON LANGUAGE MODEL EMBEDDINGS 7482
Oktai Tatanov, Stanislav Beliaev, Boris Ginsburg, NVIDIA, Russian Federation

SPE-52: LANGUAGE UNDERSTANDING II: CONTEXT/KNOWLEDGE INTEGRATION

SPE-52.1: KNOWLEDGE AUGMENTED BERT MUTUAL NETWORK IN MULTI-TURN SPOKEN DIALOGUES 7487
Ting-Wei Wu, Biing-Hwang Juang, Georgia Institute of Technology, United States of America

SPE-52.2: TINYS2I: A SMALL-FOOTPRINT UTTERANCE CLASSIFICATION MODEL WITH CONTEXTUAL SUPPORT FOR ON-DEVICE SLU 7492
Anastasios Alexandridis, Kanthashree Mysore Sathyendra, Grant Strimel, Pavel Kveton, Jon Webb, Athanasios Mouchtaris, Amazon.com, United States of America

SPE-52.3: TOWARDS END-TO-END INTEGRATION OF DIALOG HISTORY FOR IMPROVED SPOKEN LANGUAGE UNDERSTANDING 7497
Vishal Sunder, Eric Fosler-Lussier, The Ohio State University, United States of America; Samuel Thomas, Hong-Kwang Kuo, Jatin Ganhotra, Brian Kingsbury, IBM, United States of America

SPE-52.4: IMPROVING SPOKEN LANGUAGE UNDERSTANDING BY ENHANCING TEXT REPRESENTATION 7502
Thai Binh Nguyen, Karlsruhe Institute of Technology, Germany; , ,

SPE-52.5: MULTI-TASK RNN-T WITH SEMANTIC DECODER FOR STREAMABLE SPOKEN LANGUAGE UNDERSTANDING 7507
Xuandi Fu, Carnegie Mellon University, United States of America; Feng-Ju Chang, Martin Radfar, Kai Wei, Jing Liu, Grant P. Strimel, Kanthashree Mysore Sathyendra, Amazon, United States of America

SPE-52.6: A BERT BASED JOINT LEARNING MODEL WITH FEATURE GATED MECHANISM FOR SPOKEN LANGUAGE UNDERSTANDING 7512
Wang Zhang, Lei Jiang, Shaokang Zhang, Shuo Wang, Jianlong Tan, Chinese Academic of Science, China

SPE-53: SPEAKER RECOGNITION VII: NEURAL NETWORK ARCHITECTURE

SPE-53.1: MFA: TDNN WITH MULTI-SCALE FREQUENCY-CHANNEL ATTENTION FOR TEXT-INDEPENDENT SPEAKER VERIFICATION WITH SHORT UTTERANCES 7517
Tianchi Liu, Kong Aik Lee, Agency for Science, Technology and Research, Singapore; Rohan Kumar Das, Fortemedia Singapore, Singapore; Haizhou Li, National University of Singapore, Singapore

SPE-53.2: MLP-SVNET : A MULTI-LAYER PERCEPTRONS BASED NETWORK FOR SPEAKER VERIFICATION 7522
Bing Han, Zhengyang Chen, Bei Liu, Yanmin Qian, Shanghai Jiao Tong University, China

SPE-53.3: REAL ADDITIVE MARGIN SOFTMAX FOR SPEAKER VERIFICATION 7527
Lantian Li, Ruiqian Nai, Dong Wang, Tsinghua University, China

SPE-53.4: STATISTICAL PYRAMID DENSE TIME DELAY NEURAL NETWORK FOR SPEAKER VERIFICATION 7532
Zi-Kai Wan, Qing-Hua Ren, You-Cai Qin, Qi-Rong Mao, Jiangsu University, China

SPE-53.5: ON THE IMPORTANCE OF DIFFERENT FREQUENCY BINS FOR SPEAKER VERIFICATION 7537
Aiwen Deng, Wenxiong Kang, Feiqi Deng, South China University of Technology, China; Shuai Wang, Shanghai Jiao Tong University, China

SPE-53.6: SELF-KNOWLEDGE DISTILLATION VIA FEATURE ENHANCEMENT FOR SPEAKER VERIFICATION 7542
Bei Liu, Haoyu Wang, Zhengyang Chen, Shuai Wang, Yanmin Qian, Shanghai Jiao Tong University, China

SPE-54: KEYWORD SPOTTING

SPE-54.1: JOINT EGO-NOISE SUPPRESSION AND KEYWORD SPOTTING ON SWEEPING ROBOTS 7547
Yueyue Na, Ziteng Wang, Qiang Fu, Alibaba Group, China; Liang Wang, Sun Yat-sen University (SYSU), China

SPE-54.2: PROGRESSIVE CONTINUAL LEARNING FOR SPOKEN KEYWORD SPOTTING 7552
*Yizheng Huang, Nancy Chen, Institute for Infocomm Research, A*STAR, Singapore, Singapore; Nana Hou, Nanyang Technological University, Singapore, Singapore*

SPE-54.3: UNIFIED SPECULATION, DETECTION, AND VERIFICATION KEYWORD SPOTTING 7557
Geng-shen Fu, Thibaud Senechal, Aaron Challenner, Tao Zhang, Amazon, United States of America

SPE-54.4: LEARNING DECOUPLING FEATURES THROUGH ORTHOGONALITY REGULARIZATION 7562
Li Wang, Rongzhi Gu, Yuexian Zou, Peking University, China; Weiji Zhuang, Peng Gao, Yujun Wang, Xiaomi Inc., China

SPE-54.5: TEMPORAL EARLY EXITING FOR STREAMING SPEECH COMMANDS RECOGNITION 7567
Raphael Tang, Karun Kumar, Piyush Vyas, Gefei Yang, Yajie Mao, Craig Murray, Comcast, United States of America; Ji Xin, Jimmy Lin, University of Waterloo, Canada; Wenyan Li, SenseTime, China

SPE-54.6: A STUDY OF DESIGNING COMPACT AUDIO-VISUAL WAKE WORD SPOTTING SYSTEM BASED ON ITERATIVE FINE-TUNING IN NEURAL NETWORK PRUNING 7572
Hengshun Zhou, Jun Du, University of Science and Technology of China, China; Chao-Han Huck Yang, Chin-Hui Lee, Georgia Institute of Technology, United States of America; Shifu Xiong, iFlytek Research, China

SPE-55: SPEECH SYNTHESIS: PROSODY

SPE-55.1: PROSOSPEECH: ENHANCING PROSODY WITH QUANTIZED VECTOR PRE-TRAINING IN TEXT-TO-SPEECH 7577
Yi Ren, Zhou Zhao, Zhejiang University, China; Ming Lei, Zhiying Huang, Shiliang Zhang, Qian Chen, Zhijie Yan, Alibaba Group, China

SPE-55.2: PROSODYSPEECH: TOWARDS ADVANCED PROSODY MODEL FOR NEURAL TEXT-TO-SPEECH 7582
Yuanhao Yi, Lei He, Shifeng Pan, Xi Wang, Yujia Xiao, Microsoft, China

SPE-55.3: HIERARCHICAL PROSODY MODELING AND CONTROL IN NON-AUTOREGRESSIVE PARALLEL NEURAL TTS 7587
Tuomo Raitio, Jiangchuan Li, Shreyas Seshadri, Apple, United States of America

SPE-55.4: DISCOURSE-LEVEL PROSODY MODELING WITH A VARIATIONAL AUTOENCODER FOR NON-AUTOREGRESSIVE EXPRESSIVE SPEECH SYNTHESIS 7592
Ning-Qian Wu, Zhao-Ci Liu, Zhen-Hua Ling, University of Science and Technology of China, China

SPE-55.5: UNSUPERVISED WORD-LEVEL PROSODY TAGGING FOR CONTROLLABLE SPEECH SYNTHESIS 7597
Yiwei Guo, Chenpeng Du, Kai Yu, Shanghai Jiaotong University, China

SPE-55.6: A CHARACTER-LEVEL SPAN-BASED MODEL FOR MANDARIN PROSODIC STRUCTURE PREDICTION 7602
Xueyuan Chen, Changhe Song, Yixuan Zhou, Zhiyong Wu, Tsinghua University, China; Changbin Chen, Zhongqin Wu, TAL Education Group, China; Helen Meng, The Chinese University of Hong Kong, China

SPE-56: LANGUAGE UNDERSTANDING III: INTENT DETECTION AND SLOT FILLING

SPE-56.1: SLIM: EXPLICIT SLOT-INTENT MAPPING WITH BERT FOR JOINT MULTI-INTENT DETECTION AND SLOT FILLING 7607
Fengyu Cai, Wanhao Zhou, Boi Faltings, EPFL, Switzerland; Fei Mi, Huawei Noah's Ark Lab, China

SPE-56.2: JOINT MULTIPLE INTENT DETECTION AND SLOT FILLING VIA SELF-DISTILLATION 7612
Lisong Chen, Peilin Zhou, Yuexian Zou, Peking University, China

SPE-56.3: A GRAPH ATTENTION INTERACTIVE REFINE FRAMEWORK WITH CONTEXTUAL REGULARIZATION FOR JOINTING INTENT DETECTION AND SLOT FILLING 7617
Zhanbiao Zhu, Peijie Huang, Haojing Huang, Shudong Liu, Leyi Lao, South China Agricultural University, China

SPE-56.4: ADJACENCY PAIRS-AWARE HIERARCHICAL ATTENTION NETWORKS FOR DIALOGUE INTENT CLASSIFICATION 7622
Jiabao Xu, Peijie Huang, Youming Peng, Jiande Ding, Boxi Huang, Simin Huang, South China Agricultural University, China

SPE-56.5: ADVIN: AUTOMATICALLY DISCOVERING NOVEL DOMAINS AND INTENTS FROM USER TEXT UTTERANCES 7627
Nikhita Vedula, Rahul Gupta, Aman Alok, Mukund Sridhar, Shankar Ananthakrishnan, Amazon, United States of America

SPE-56.6: A NEW DATA AUGMENTATION METHOD FOR INTENT CLASSIFICATION ENHANCEMENT AND ITS APPLICATION ON SPOKEN CONVERSATION DATASETS 7632
Zvi Kons, Aharon Satt, Hong-Kwang Kuo, Samuel Thomas, Boaz Carmeli, Ron Hoory, Brian Kingsbury, IBM, Israel

SPE-57: SPEAKER RECOGNITION VIII: ROBUSTNESS

SPE-57.1: ROBUST SPEAKER VERIFICATION WITH JOINT SELF-SUPERVISED AND SUPERVISED LEARNING 7637
Kai Wang, Xiaolei Zhang, Miao Zhang, Yuguang Li, Samsung R&D Institute China Xian, China; Jaeyun Lee, Kiho Cho, Sung-Un Park, Samsung Advanced Institute of Technology, Korea, Republic of

SPE-57.2: ROBUST SPEAKER VERIFICATION USING POPULATION-BASED DATA AUGMENTATION 7642
Weiwei Lin, Man-Wai Mak, The Hong Kong Polytechnic University, Hong Kong

SPE-57.3: RAWNEXT: SPEAKER VERIFICATION SYSTEM FOR VARIABLE-DURATION UTTERANCES WITH DEEP LAYER AGGREGATION AND EXTENDED DYNAMIC SCALING POLICIES 7647
Ju-ho Kim, Hye-jin Shim, Jungwoo Heo, Ha-Jin Yu, University of Seoul, Korea, Republic of

SPE-57.4: CONTRASTIVE-MIXUP LEARNING FOR IMPROVED SPEAKER VERIFICATION 7652
Xin Zhang, Texas A&M University, United States of America; Minho Jin, Roger Cheng, Ruirui Li, Eunjung Han, Andreas Stolcke, Amazon Inc., United States of America

SPE-57.5: A STUDY OF THE ROBUSTNESS OF RAW WAVEFORM BASED SPEAKER EMBEDDINGS UNDER MISMATCHED CONDITIONS 7657
Ge Zhu, Frank Cwitkowitz, Zhiyao Duan, University of Rochester, United States of America

SPE-57.6: DISENTANGLED SPEAKER EMBEDDING FOR ROBUST SPEAKER VERIFICATION 7662
Lu Yi, Man-Wai Mak, The Hong Kong Polytechnic University, Hong Kong

SPE-58: SPEECH RECOGNITION: ACOUSTIC MODELING III

SPE-58.1: PERFORMANCE-EFFICIENCY TRADE-OFFS IN UNSUPERVISED PRE-TRAINING FOR SPEECH RECOGNITION 7667
Felix Wu, Kwangyoun Kim, Jing Pan, Kyu Han, ASAPP, United States of America; Kilian Weinberger, Yoav Artzi, ASAPP and Cornell University, United States of America

SPE-58.2: ADVANCING MOMENTUM PSEUDO-LABELING WITH CONFORMER AND INITIALIZATION STRATEGY 7672
Yosuke Higuchi, Waseda University, Mitsubishi Electric Research Laboratories, Japan; Niko Moritz, Facebook AI, United Kingdom of Great Britain and Northern Ireland; Jonathan Le Roux, Takaaki Hori, Mitsubishi Electric Research Laboratories, United States of America

SPE-58.3: TTS4PRETRAIN 2.0: ADVANCING THE USE OF TEXT AND SPEECH IN ASR PRETRAINING WITH CONSISTENCY AND CONTRASTIVE LOSSES 7677
Zhehuai Chen, Yu Zhang, Andrew Rosenberg, Bhuvana Ramabhadran, Pedro Moreno, Gary Wang, Google, United States of America

SPE-58.4: SYNT++: UTILIZING IMPERFECT SYNTHETIC DATA TO IMPROVE SPEECH RECOGNITION 7682
Ting-Yao Hu, Ashish Shrivastava, Jen-Hao Rick Chang, Hema Koppula, Oncel Tuzel, Apple, United States of America; Mohammadreza Armandpour, Texas A&M University, United States of America

SPE-58.5: PSEUDO-LABELING FOR MASSIVELY MULTILINGUAL SPEECH RECOGNITION 7687
Loren Lugosch, McGill University/Mila, Canada; Tatiana Likhomanenko, Gabriel Synnaeve, Ronan Collobert, Facebook AI Research, United States of America

SPE-58.6: DP-DWA: DUAL-PATH DYNAMIC WEIGHT ATTENTION NETWORK WITH STREAMING DFSMN-SAN FOR AUTOMATIC SPEECH RECOGNITION 7692
Dongpeng Ma, Yiwen Wang, Liqiang He, Mingjie Jin, Dan Su, Dong Yu, Tencent, China

SPE-59: EMOTION RECOGNITION: GENERAL TOPICS II

SPE-59.2: SERAB: A MULTI-LINGUAL BENCHMARK FOR SPEECH EMOTION RECOGNITION 7697
Neil Scheidwasser-Clow, Pierre Beckmann, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland; Mikolaj Kegler, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Milos Cernak, Logitech Europe S.A., Switzerland

SPE-59.3: ENHANCING PRIVACY THROUGH DOMAIN ADAPTIVE NOISE INJECTION FOR SPEECH EMOTION RECOGNITION 7702
Tiantian Feng, Hanieh Hashemi, Murali Annavaram, Shrikanth Narayanan, University of Southern California, United States of America

SPE-59.4: SELECTIVE MULTI-TASK LEARNING FOR SPEECH EMOTION RECOGNITION USING CORPORA OF DIFFERENT STYLES 7707
Heran Zhang, Masato Mimura, Tatsuya Kawahara, Kyoto University, Japan; Kenkichi Ishizuka, Revcomm, Japan

SPE-59.5: FRONTEND ATTRIBUTES DISENTANGLEMENT FOR SPEECH EMOTION RECOGNITION 7712
Yu-Xuan Xi, Yan Song, Li-Rong Dai, National Engineering Laboratory for Speech and Language Information Processing,University of Science and Technology of China, China; Ian McLoughlin, ICT Cluster, Singapore Institute of Technology, Singapore; Lin Liu, iFLYTEK Research, iFLYTEK CO., LTD, China

SPE-59.6: EXPLOITING ANNOTATORS' TYPED DESCRIPTION OF EMOTION PERCEPTION TO MAXIMIZE UTILIZATION OF RATINGS FOR SPEECH EMOTION RECOGNITION 7717
Huang-Cheng Chou, Wei-Cheng Lin, Carlos Busso, The University of Texas at Dallas, United States of America; Chi-Chun Lee, National Tsing Hua University,, Taiwan

SPE-60: MULTIMODAL LANGUAGE PROCESSING

SPE-60.1: AUTOMATED AUDIO CAPTIONING USING TRANSFER LEARNING AND RECONSTRUCTION LATENT SPACE SIMILARITY REGULARIZATION 7722
Andrew Koh, Fuzhao Xue, Eng Siong Chng, Nanyang Technological University, Singapore

SPE-60.2: FAST-SLOW TRANSFORMER FOR VISUALLY GROUNDING SPEECH 7727
Puyuan Peng, David Harwath, The University of Texas at Austin, United States of America

SPE-60.3: AUDIO-VISUAL SCENE-AWARE DIALOG AND REASONING USING 7732
AUDIO-VISUAL TRANSFORMERS WITH JOINT STUDENT-TEACHER LEARNING
Ankit Parag Shah, Carnegie Mellon University, United States of America; Shijie Geng, Rutgers University, United States of America; Gao Peng, Chinese University of Hong Kong, United States of America; Anoop Cherian, Takaaki Hori, Tim K. Marks, Jonathan Le Roux, Chiori Hori, Mitsubishi Electric Research Laboratories (MERL), United States of America

SPE-60.4: AIMNET: ADAPTIVE IMAGE-TAG MERGING NETWORK FOR AUTOMATIC 7737
MEDICAL REPORT GENERATION
Jijun Shi, Shanshe Wang, Ronggang Wang, Siwei Ma, Peking University, China

SPE-60.5: ADVERSARIAL INPUT ABLATION FOR AUDIO-VISUAL LEARNING..... 7742
David Xu, David Harwath, The University of Texas at Austin, United States of America

SPE-60.6: GATED MULTIMODAL FUSION WITH CONTRASTIVE LEARNING FOR 7747
TURN-TAKING PREDICTION IN HUMAN-ROBOT DIALOGUE
Jiudong Yang, Peiying Wang, Mingchao Feng, Meng Chen, Xiaodong He, JD AI, China; Yi Zhu, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

SPE-61: SPEECH ENHANCEMENT AND CODING

SPE-61.1: JOINT FAR- AND NEAR-END SPEECH INTELLIGIBILITY ENHANCEMENT 7752
BASED ON THE APPROXIMATED SPEECH INTELLIGIBILITY INDEX
Andreas Jonas Fuglsig, Lars Søndergaard Bertelsen, Peter Mariager, RTX A/S, Denmark; Jan Østergaard, Jesper Jensen, Zheng-Hua Tan, Aalborg University, Denmark

SPE-61.2: ATTENTION-BASED FUSION FOR BONE-CONDUCTED AND 7757
AIR-CONDUCTED SPEECH ENHANCEMENT IN THE COMPLEX DOMAIN
Heming Wang, DeLiang Wang, The Ohio State University, United States of America; Xueliang Zhang, Inner Mongolia University, China

SPE-61.3: A TWO-STEP BACKWARD COMPATIBLE FULLBAND SPEECH 7762
ENHANCEMENT SYSTEM
Xu Zhang, Lianwu Chen, Xiguang Zheng, Xinlei Ren, Chen Zhang, Liang Guo, Bing Yu, Kuai Shou Technology Co. Beijing, China, China

SPE-61.4: S-DCCRN: SUPER WIDE BAND DCCRN WITH LEARNABLE COMPLEX 7767
FEATURE FOR SPEECH ENHANCEMENT
Shubo Lv, Yihui Fu, Mengtao Xing, Jiayao Sun, Lei Xie, Northwestern Polytechnical University, Xi'an, China, China; Jun Huang, Yannan Wang, Tao Yu, Tencent Ethereal Audio Lab, Tencent Corporation, Shenzhen, China, China

SPE-61.5: COGNITIVE CODING OF SPEECH 7772
Reza Lotfidereshgi, Philippe Gournay, Universite de Sherbrooke, Canada

SPE-61.6: SPEECH ENHANCEMENT FOR LOW BIT RATE SPEECH CODEC 7777
Ju Lin, Clemson University, United States of America; Kaustubh Kalgaonkar, Qing He, Xin Lei, Facebook, United States of America

SPE-62: SPEECH RECOGNITION: TRAINING METHODS FOR E2E ASR

SPE-62.1: CONSISTENT TRAINING AND DECODING FOR END-TO-END SPEECH 7782
RECOGNITION USING LATTICE-FREE MMI
Jinchuan Tian, Yuexian Zou, Peking University, China; Jianwei Yu, Chao Weng, Shi-Xiong Zhang, Dan Su, Dong Yu, Tencent, China

SPE-62.2: BEING GREEDY DOES NOT HURT: SAMPLING STRATEGIES FOR END-TO-END SPEECH RECOGNITION 7787
Jahn Heymann, Egor Lakomkin, Leif Raedel, Amazon Inc., Germany

SPE-62.3: INVESTIGATING SEQUENCE-LEVEL NORMALISATION FOR CTC-LIKE END-TO-END ASR 7792
Zeyu Zhao, Peter Bell, Centre for Speech Technology Research, University of Edinburgh, UK, United Kingdom of Great Britain and Northern Ireland

SPE-62.4: HIERARCHICAL CONDITIONAL END-TO-END ASR WITH CTC AND MULTI-GRANULAR SUBWORD UNITS 7797
Yosuke Higuchi, Keita Karube, Tetsuji Ogawa, Tetsunori Kobayashi, Waseda University, Japan

SPE-62.5: OPTIMIZING ALIGNMENT OF SPEECH AND LANGUAGE LATENT SPACES FOR END-TO-END SPEECH RECOGNITION AND UNDERSTANDING 7802
Wei Wang, Yanmin Qian, Shanghai Jiao Tong University, China; Shuo Ren, Yao Qian, Shujie Liu, Yu Shi, Michael Zeng, Microsoft Corporation, China

SPE-62.6: MINIMUM WORD ERROR TRAINING FOR NON-AUTOREGRESSIVE TRANSFORMER-BASED CODE-SWITCHING ASR 7807
Yizhou Peng, Jicheng Zhang, Hao Huang, Xinjiang University, China; Haihua Xu, Bytedance, Singapore; Eng Siong Chng, Nanyang Technological University, Singapore

SPE-63: TEXT CLASSIFICATION

SPE-63.1: CONNECTING TARGETS VIA LATENT TOPICS AND CONTRASTIVE LEARNING: A UNIFIED FRAMEWORK FOR ROBUST ZERO-SHOT AND FEW-SHOT STANCE DETECTION 7812
Rui Liu, Zheng Lin, Peng Fu, Yuanxin Liu, Weiping Wang, Institute of Information Engineering, Chinese Academy of Sciences, China

SPE-63.2: PRIOR-BERT AND MULTI-TASK LEARNING FOR TARGET-ASPECT-SENTIMENT JOINT DETECTION 7817
Cai Ke, Qingyu Xiong, Chao Wu, Hualing Yi, Chongqing University, China; Zikai Liao, Xidian University, China

SPE-63.3: CROSS-TARGET STANCE DETECTION VIA REFINED META-LEARNING 7822
Huishan Ji, Zheng Lin, Peng Fu, Weiping Wang, Institute of Information Engineering, Chinese Academy of Sciences, China

SPE-63.4: A ROBUST CONTRASTIVE ALIGNMENT METHOD FOR MULTI-DOMAIN TEXT CLASSIFICATION 7827
Xuefeng Li, Hao Lei, Liwen Wang, Guanting Dong, Jiachi Liu, Weiran Xu, Beijing University of Posts and Telecommunications, China; Jinzheng Zhao, University of Surrey, United Kingdom of Great Britain and Northern Ireland; Chunyun Zhang, Shandong University of Finance and Economics, China

SPE-63.5: INCREMENTAL USER EMBEDDING MODELING FOR PERSONALIZED TEXT CLASSIFICATION 7832
Ruixue Lian, University of Wisconsin-Madison, United States of America; Che-Wei Huang, Yuqing Tang, Qilong Gu, Chengyuan Ma, Chenlei Guo, Amazon Alexa, United States of America

SPE-63.6: BLOCK-SPARSE ADVERSARIAL ATTACK TO FOOL TRANSFORMER-BASED TEXT CLASSIFIERS 7837
Sahar Sadrizadeh, Pascal Frossard, École polytechnique fédérale de Lausanne (EPFL), Switzerland; Ljiljana Dolamic, Armasuisse S+T, Switzerland

SPE-64: SPEECH ENHANCEMENT: DNN ARCHITECTURES

SPE-64.1: MANNER: MULTI-VIEW ATTENTION NETWORK FOR NOISE ERASURE 7842
Hyun Joon Park, Byung Ha Kang, Wooseok Shin, Jin Sob Kim, Sung Won Han, Korea University, Korea, Republic of

SPE-64.2: DUAL-BRANCH ATTENTION-IN-ATTENTION TRANSFORMER FOR 7847
SINGLE-CHANNEL SPEECH ENHANCEMENT
Guochen Yu, Communication University of China\ Institute of Acoustics, Chinese Academy of Sciences, China; Andong Li, Chengshi Zheng, Institute of Acoustics, Chinese Academy of Sciences, China; Yinuo Guo, Bytedance, China; Yutian Wang, Hui Wang, Communication University of China, China

SPE-64.3: TIME-FREQUENCY ATTENTION FOR MONAURAL SPEECH 7852
ENHANCEMENT
Qiquan Zhang, Haizhou Li, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore; Qi Song, Alibaba Group, China, China; Zhaoheng Ni, Meta AI, United States of America; Aaron Nicolson, Australian eHealth Research Centre, CSIRO, Herston, QLD, 4006, Australia, Australia

SPE-64.4: FULLSUBNET+: CHANNEL ATTENTION FULLSUBNET WITH COMPLEX 7857
SPECTROGRAMS FOR SPEECH ENHANCEMENT
Jun Chen, Zilin Wang, Zhiyong Wu, Tsinghua University, China; Deyi Tuo, Shiyin Kang, Huya Inc., China; Helen Meng, The Chinese University of Hong Kong, China

SPE-64.5: CROSS-DOMAIN SPEECH ENHANCEMENT WITH A NEURAL CASCADE 7862
ARCHITECTURE
Heming Wang, DeLiang Wang, The Ohio State University, United States of America

SPE-64.6: SPEECH DENOISING IN THE WAVEFORM DOMAIN WITH 7867
SELF-ATTENTION
Zhifeng Kong, University of California San Diego, United States of America; Wei Ping, Ambrish Dantrey, Bryan Catanzaro, Nvidia, United States of America

SPE-65: SPEECH RECOGNITION: GENERAL TOPICS IV

SPE-65.1: SRU++: PIONEERING FAST RECURRENCE WITH ATTENTION FOR 7872
SPEECH RECOGNITION
Jing Pan, Tao Lei, Kwangyoun Kim, Kyu Han, ASAPP Inc, United States of America; Shinji Watanabe, Carnegie Mellon University, United States of America

SPE-65.2: LATTENTION: LATTICE-ATTENTION IN ASR RESCORING 7877
Prabhat Pandey, Sergio Duarte Torres, Ali Orkan Bayer, Ankur Gandhe, Volker Leutnant, Amazon, Germany

SPE-65.4: LEARNING ACOUSTIC FRAME LABELING FOR PHONEME 7882
SEGMENTATION WITH REGULARIZED ATTENTION MECHANISM
Binghuai Lin, Liyuan Wang, Tencent Technology Co., Ltd, China

SPE-65.5: LISTEN, KNOW AND SPELL: KNOWLEDGE-INFUSED SUBWORD 7887
MODELING FOR IMPROVING ASR PERFORMANCE OF OOV NAMED ENTITIES
Nilaksh Das, Duen Horng Chau, Georgia Institute of Technology, United States of America; Monica Sunkara, Dhanush Bekal, Sravan Bodapati, Katrin Kirchhoff, Amazon AWS AI, United States of America

SPE-65.6: JOINT SPEECH RECOGNITION AND AUDIO CAPTIONING 7892
Chaitanya Narisetty, Xuankai Chang, Shinji Watanabe, Carnegie Mellon University, United States of America; Emiru Tsunoo, Yosuke Kashiwagi, Michael Hentschel, Sony Group Corporation, Japan

SPE-66: SPEECH SYNTHESIS: STYLE CONTROL, TRANSFER, AND ADAPTATION

SPE-66.1: SPEAKER GENERATION 7897
Daisy Stanton, Matt Shannon, Soroosh Mariooryad, RJ Skerry-Ryan, Eric Battenberg, Tom Bagby, David Kao, Google, United States of America

SPE-66.2: VOICE FILTER: FEW-SHOT TEXT-TO-SPEECH SPEAKER ADAPTATION 7902
USING VOICE CONVERSION AS A POST-PROCESSING MODULE
Adam Gabryś, Goeric Huybrechts, Manuel Sam Ribeiro, Julian Roth, Giulia Comini, Roberto Barra-Chicote, Bartek Perz, Jaime Lorenzo-Trueba, Amazon, Poland; Chung-Ming Chien, National Taiwan University (NTU), United Kingdom of Great Britain and Northern Ireland

SPE-66.3: FINE-GRAINED STYLE CONTROL IN TRANSFORMER-BASED 7907
TEXT-TO-SPEECH SYNTHESIS
Li-Wei Chen, Alexander Rudnicky, Carnegie Mellon University, United States of America

SPE-66.4: USING MULTIPLE REFERENCE AUDIOS AND STYLE EMBEDDING 7912
CONSTRAINTS FOR SPEECH SYNTHESIS
Cheng Gong, Longbiao Wang, Tianjin University, China; Zhenhua Ling, University of Science and Technology of China, China; Ju Zhang, Huiyan Technology (Tianjin) Co., Ltd, China; Jianwu Dang, Japan Advanced Institute of Science and Technology, Japan

SPE-66.5: ENHANCING SPEAKING STYLES IN CONVERSATIONAL 7917
TEXT-TO-SPEECH SYNTHESIS WITH GRAPH-BASED MULTI-MODAL CONTEXT MODELING
Jingbei Li, Yi Meng, Chenyi Li, Zhiyong Wu, Tsinghua University, China; Helen Meng, The Chinese University of Hong Kong, China; Chao Weng, Dan Su, Tencent, China

SPE-66.6: TOWARDS EXPRESSIVE SPEAKING STYLE MODELLING WITH 7922
HIERARCHICAL CONTEXT INFORMATION FOR MANDARIN SPEECH SYNTHESIS
Shun Lei, Yixuan Zhou, Liyang Chen, Zhiyong Wu, Tsinghua University, China; Shiyin Kang, Huya Inc., China; Helen Meng, The Chinese University of Hong Kong, Hong Kong

SPE-67: LANGUAGE UNDERSTANDING IV: GENERAL TOPICS

SPE-67.1: SLUE: NEW BENCHMARK TASKS FOR SPOKEN LANGUAGE 7927
UNDERSTANDING EVALUATION ON NATURAL SPEECH
Suwon Shon, Felix Wu, Pablo Brusco, Kyu J. Han, ASAPP, United States of America; Ankita Pasad, Karen Livescu, Toyota Technological Institute at Chicago, United States of America; Yoav Artzi, Cornell University, ASAPP, United States of America

SPE-67.2: TOWARDS REDUCING THE NEED FOR SPEECH TRAINING DATA TO 7932
BUILD SPOKEN LANGUAGE UNDERSTANDING SYSTEMS
Samuel Thomas, Hong-Kwang Kuo, Brian Kingsbury, George Saon, IBM Research AI, United States of America

SPE-67.3: IMPROVING CROSS-MODAL UNDERSTANDING IN VISUAL DIALOG VIA 7937
CONTRASTIVE LEARNING
Feilong Chen, Xiuyi Chen, Shuang Xu, Bo Xu, Institute of Automation, Chinese Academy of Sciences (CASIA), China

SPE-67.4: NEWS RECOMMENDATION VIA MULTI-INTEREST NEWS SEQUENCE 7942
MODELLING
Rongyao Wang, Wenpeng Lu, Qilu University of Technology (Shandong Academy of Sciences), China; Shoujin Wang, RMIT University, Australia; Xueping Peng, University of Technology Sydney, Australia

SPE-67.5: MULTI-LEVEL CONTRASTIVE LEARNING FOR CROSS-LINGUAL 7947
ALIGNMENT
Beiduo Chen, Wu Guo, Bin Gu, University of Science and Technology of China, China; Quan Liu, Yongchao Wang, iFLYTEK Research, China

SPE-67.6: AUGMENTATION STRATEGY OPTIMIZATION FOR LANGUAGE UNDERSTANDING 7952
Chang-Ting Chu, Mahdin Rohmatillah, Jen-Tzung Chien, National Yang Ming Chiao Tung University, Taiwan; Ching-hsien Lee, Industrial Technology Research Institute, Taiwan

SPE-68: SPEAKER RECOGNITION IX: SINGLE AND MULTI CHANNEL

SPE-68.1: MULTI-FEATURE INTEGRATION FOR SPEAKER EMBEDDING EXTRACTION 7957
Sreekanth Sankala, Shaik Mohammad Rafi B, Sri Rama Murty K, Indian Institute of Technology Hyderabad, India

SPE-68.2: LEARNABLE NONLINEAR COMPRESSION FOR ROBUST SPEAKER VERIFICATION 7962
Xuechen Liu, University of Eastern Finland & Inria, Finland; Md Sahidullah, Inria, France; Tomi Kinnunen, University of Eastern Finland, Finland

SPE-68.3: FINE-TUNING WAV2VEC2 FOR SPEAKER RECOGNITION 7967
Nik Vaessen, David A. van Leeuwen, Radboud University, Netherlands

SPE-68.4: GRAPH ATTENTIVE FEATURE AGGREGATION FOR TEXT-INDEPENDENT SPEAKER VERIFICATION 7972
Hye-jin Shim, Jungwoo Heo, Ha-Jin Yu, University of Seoul, Korea, Republic of; Jae-han Park, Ga-Hui Lee, KT Corporation, Korea, Republic of

SPE-68.5: MULTISV: DATASET FOR FAR-FIELD MULTI-CHANNEL SPEAKER VERIFICATION 7977
Ladislav Mošner, Oldřich Plchot, Lukáš Burget, Jan Černocký, Faculty of Information Technology, Brno University of Technology, Czechia

SPE-68.6: MULTI-CHANNEL SPEAKER VERIFICATION WITH CONV-TASNET BASED BEAMFORMER 7982
Ladislav Mošner, Oldřich Plchot, Lukáš Burget, Jan Černocký, Faculty of Information Technology, Brno University of Technology, Czechia

SPE-69: ATTENTION MECHANISM FOR SPEECH MODELS

SPE-69.1: LETR: A LIGHTWEIGHT AND EFFICIENT TRANSFORMER FOR KEYWORD SPOTTING 7987
Kevin Ding, Martin Zong, Jiakui Li, Baoxiang Li, SenseTime, China

SPE-69.2: COMPRESSING TRANSFORMER-BASED ASR MODEL BY TASK-DRIVEN LOSS AND ATTENTION-BASED MULTI-LEVEL FEATURE DISTILLATION 7992
Yongjie Lv, Longbiao Wang, Meng Ge, Kiyoshi Honda, Tianjin University, China; Sheng Li, Chenchen Ding, National Institute of Information & Communications Technology (NICT), Japan; Lixin Pan, Yuguang Wang, Huiyan Technology (TianJin) Co., LTD., China; Jianwu Dang, Japan Advanced Institute of Science and Technology, Japan

SPE-69.3: SPATIAL PROCESSING FRONT-END FOR DISTANT ASR EXPLOITING SELF-ATTENTION CHANNEL COMBINATOR 7997
Dushyant Sharma, Rong Gong, James Fosburgh, Stanislav Kruchinin, Patrick Naylor, Ljubomir Milanovic, Nuance Communications Inc, United States of America

SPE-69.4: EFFICIENT SEQUENCE TRAINING OF ATTENTION MODELS USING APPROXIMATIVE RECOMBINATION 8002
Nils-Philipp Wynands, Wilfried Michel, Jan Rosendahl, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany

SPE-69.5: NEUFA: NEURAL NETWORK BASED END-TO-END FORCED ALIGNMENT WITH BIDIRECTIONAL ATTENTION MECHANISM 8007
Jingbei Li, Yi Meng, Zhiyong Wu, Tsinghua University, China; Helen Meng, The Chinese University of Hong Kong, China; Qiao Tian, Yuping Wang, Yuxuan Wang, ByteDance, China

SPE-69.6: CONFORMER-BASED SPEECH RECOGNITION WITH LINEAR NYSTRÖM ATTENTION AND ROTARY POSITION EMBEDDING 8012
Lahiru Samarakoon, Tsun-Yat Leung, Fano Labs, Hong Kong

SPE-70: SPEECH SYNTHESIS: MULTI-LINGUAL AND MULTIMODAL SYNTHESIS

SPE-70.1: MULTILINGUAL TEXT-TO-SPEECH TRAINING USING CROSS LANGUAGE VOICE CONVERSION AND SELF-SUPERVISED LEARNING OF SPEECH REPRESENTATIONS 8017
Jilong Wu, Adam Polyak, Yaniv Taigman, Prabhav Agrawal, Qing He, Facebook, United States of America; Jason Fong, The University of Edinburgh, United Kingdom of Great Britain and Northern Ireland

SPE-70.2: TOWARDS LIFELONG LEARNING OF MULTILINGUAL TEXT-TO-SPEECH SYNTHESIS 8022
Mu Yang, University of Texas at Dallas, United States of America; Shaojin Ding, Texas A&M University, United States of America; Tianlong Chen, Zhangyang Wang, University of Texas at Austin, United States of America; Tong Wang, University of Science and Technology of China, China

SPE-70.3: ZERO-SHOT CROSS-LINGUAL TRANSFER USING MULTI-STREAM ENCODER AND EFFICIENT SPEAKER REPRESENTATION 8027
Yibin Zheng, Zewang Zhang, Xinhui Li, Wenchao Su, Li Lu, Tencent Inc, China, China

SPE-70.4: VISUALTTS: TTS WITH ACCURATE LIP-SPEECH SYNCHRONIZATION FOR AUTOMATIC VOICE OVER 8032
Junchen Lu, Mingyang Zhang, Haizhou Li, National University of Singapore, Singapore; Berrak Sisman, Singapore University of Technology and Design, Singapore; Rui Liu, National University of Singapore and Singapore University of Technology and Design, Singapore

SPE-70.5: DURATION MODELING OF NEURAL TTS FOR AUTOMATIC DUBBING 8037
Johanes Effendi, Yogesh Virkar, Roberto Barra-Chicote, Marcello Federico, Amazon AI, Japan

SPE-70.6: LEARNING TO PREDICT SPEECH IN SILENT VIDEOS VIA AUDIOVISUAL ANALOGY 8042
Ravindra Yadav, Vinay Namboodiri, Rajesh Hegde, Indian Institute of Technology Kanpur, India; Ashish Sardana, NVIDIA, India

SPE-71: TEXT COMPLETION AND SUMMARIZATION

SPE-71.1: SELF-ATTENTION FOR INCOMPLETE UTTERANCE REWRITING 8047
Yong Zhang, Zhitao Li, Jianzong Wang, Ning Cheng, Jing Xiao, Ping An Technology (Shenzhen) Co., Ltd., China

SPE-71.2: MULTI-TURN INCOMPLETE UTTERANCE RESTORATION AS OBJECT DETECTION 8052
Wangjie Jiang, Siheng Li, Jiayi Li, Yujiu Yang, Tsinghua University, China

SPE-71.3: CLSEG: CONTRASTIVE LEARNING OF STORY ENDING GENERATION 8057
Yuqiang Xie, Yue Hu, Luxi Xing, Yunpeng Li, Wei Peng, Ping Guo, Institute of Information Engineering, Chinese Academy of Sciences, China

SPE-71.4: EXPLICITLY MODELING IMPORTANCE AND COHERENCE FOR TIMELINE SUMMARIZATION 8062
Qianren Mao, Jianxin Li, Jiazheng Wang, Xi Li, Peng Hao, Beihang University (BUAA), China; Lihong Wang, CNCERT/CC, China; Zheng Wang, University of Leeds, United Kingdom of Great Britain and Northern Ireland

SPE-71.5: TED TALK TEASER GENERATION WITH PRE-TRAINED MODELS 8067
Gianluca Vico, Jan Niehues, Maastricht University, Netherlands

SPE-71.6: END-TO-END SPEECH SUMMARIZATION USING RESTRICTED SELF-ATTENTION 8072
Roshan Sharma, Shruti Palaskar, Alan W Black, Florian Metze, Carnegie Mellon University, United States of America

SPE-72: SPEAKER DIARIZATION I

SPE-72.1: TURN-TO-DIARIZE: ONLINE SPEAKER DIARIZATION CONSTRAINED BY TRANSFORMER TRANSDUCER SPEAKER TURN DETECTION 8077
Wei Xia, University of Texas at Dallas, United States of America; Han Lu, Quan Wang, Anshuman Tripathi, Yiling Huang, Ignacio Lopez Moreno, Hasim Sak, Google, United States of America

SPE-72.2: TRANSCRIBE-TO-DIARIZE: NEURAL SPEAKER DIARIZATION FOR UNLIMITED NUMBER OF SPEAKERS USING END-TO-END SPEAKER-ATTRIBUTED ASR 8082
Naoyuki Kanda, Xiong Xiao, Yashesh Gaur, Xiaofei Wang, Zhong Meng, Zhuo Chen, Takuya Yoshioka, Microsoft, United States of America

SPE-72.3: A MULTITASK LEARNING FRAMEWORK FOR SPEAKER CHANGE DETECTION WITH CONTENT INFORMATION FROM UNSUPERVISED SPEECH DECOMPOSITION 8087
Hang Su, Danyang Zhao, Long Dang, Xixin Wu, Xunying Liu, Helen Meng, The Chinese University of Hong Kong, Hong Kong; Minglei Li, Huawei Cloud, China

SPE-72.4: ASR-AWARE END-TO-END NEURAL DIARIZATION 8092
Aparna Khare, Eunjung Han, Yuguang Yang, Andreas Stolcke, Amazon, United States of America

SPE-72.5: REFORMULATING SPEAKER DIARIZATION AS COMMUNITY DETECTION WITH EMPHASIS ON TOPOLOGICAL STRUCTURE 8097
Siqi Zheng, Hongbin Suo, Alibaba Group, China

SPE-72.6: TITANET: NEURAL MODEL FOR SPEAKER REPRESENTATION WITH 1D DEPTH-WISE SEPARABLE CONVOLUTIONS AND GLOBAL CONTEXT 8102
Nithin Rao Koluguri, Taejin Park, Boris Ginsburg, NVIDIA, United States of America

SPE-73: SPEECH RECOGNITION: NEURAL TRANSDUCER MODELS

SPE-73.1: TRANSDUCER-BASED STREAMING DELIBERATION FOR CASCADED ENCODERS 8107
Ke Hu, Tara Sainath, Arun Narayanan, Ruoming Pang, Trevor Strohman, Google, United States of America

SPE-73.2: IMPROVING THE LATENCY AND QUALITY OF CASCADED ENCODERS 8112
Tara Sainath, Yanzhang He, Arun Narayanan, Rami Botros, Weiran Wang, David Qiu, Chung-cheng Chiu, Rohit Prabhavalkar, Alexander Gruenstein, Anmol Gulati, Bo Li, David Rybach, Emmanuel Guzman, Ian McGraw, James Qin, Krzysztof Choromanski, Qiao Liang, Robert David, Ruoming Pang, Shuoyiin Chang, Trevor Strohman, W. Ronny Huang, Wei Han, Yonghui Wu, Yu Zhang, Google, Inc., United States of America

SPE-73.3: IMPROVING THE FUSION OF ACOUSTIC AND TEXT REPRESENTATIONS IN RNN-T 8117
Chao Zhang, Bo Li, Zhiyun Lu, Tara Sainath, Shuo-Yiin Chang, Google LLC, United Kingdom of Great Britain and Northern Ireland

SPE-73.4: ADAPTIVE DISCOUNTING OF IMPLICIT LANGUAGE MODELS IN RNN-TRANSDUCERS 8122
Vinit Unni, Preethi Jyothi, Sunita Sarawagi, Indian Institute of Technology Bombay, India; Shreya Khare, Ashish Mittal, Samarth Bharadwaj, IBM Research, India

SPE-73.5: INTEGRATING TEXT INPUTS FOR TRAINING AND ADAPTING RNN TRANSDUCER ASR MODELS 8127
Samuel Thomas, Brian Kingsbury, George Saon, Hong-Kwang Kuo, IBM Research AI, United States of America

SPE-73.6: FACTORIZED NEURAL TRANSDUCER FOR EFFICIENT LANGUAGE MODEL ADAPTATION 8132
Xie Chen, Zhong Meng, Sarangarajan Parthasarathy, Jinyu Li, Microsoft Corporation, United States of America

SPE-74: REPRESENTATIONS AND RELATIONSHIPS IN LANGUAGE

SPE-74.1: INTEGRATING DEPENDENCY TREE INTO SELF-ATTENTION FOR SENTENCE REPRESENTATION 8137
Junhua Ma, Yuxuan Liu, Shangbo Zhou, Chongqing University, China; Jiajun Li, Xue Li, The University of Queensland, Australia

SPE-74.2: METRICBERT: TEXT REPRESENTATION LEARNING VIA SELF-SUPERVISED TRIPLET TRAINING 8142
Itzik Malkiel, Dvir Ginzburg, Oren Barkan, Avi Caciularu, Yoni Weill, Noam Koenigstein, microsoft, Israel

SPE-74.4: END-TO-END NEURAL COREFERENCE RESOLUTION REVISITED: A SIMPLE YET EFFECTIVE BASELINE 8147
Tuan Lai, University of Illinois at Urbana-Champaign, United States of America; Trung Bui, Adobe, United States of America; Doo-Soon Kim, Roku, Inc., United States of America

SPE-74.5: LOCAL CONTEXT INTERACTION-AWARE GLYPH-VECTORS FOR CHINESE SEQUENCE TAGGING 8152
JunYu Lu, PingJian Zhang, South China University of Technology, China

SPE-75: SPEECH ANALYSIS

SPE-75.2: DEEP LEARNING FOR PROMINENCE DETECTION IN CHILDREN'S READ SPEECH 8157
Mithilesh Vaidya, Kamini Sabu, Preeti Rao, Indian Institute of Technology, Bombay, India

SPE-75.3: TOWARDS A COMMON SPEECH ANALYSIS ENGINE 8162
Hagai Aronowitz, Itai Gat, Edmilson Morais, Weizhong Zhu, Ron Hoory, IBM Research AI, Israel

SPE-75.4: PHONE-TO-AUDIO ALIGNMENT WITHOUT TEXT: A SEMI-SUPERVISED APPROACH 8167
Jian Zhu, David Jurgens, University of Michigan, United States of America; Cong Zhang, Radboud University, Netherlands

SPE-75.5: ATTACHMENT RECOGNITION IN SCHOOL-AGE CHILDREN: A MULTIMODAL APPROACH BASED ON LANGUAGE AND PARALANGUAGE ANALYSIS 8172
Huda Alsofyani, Alessandro Vinciarelli, University of Glasgow, United Kingdom of Great Britain and Northern Ireland

SPE-75.6: DETERMINING THE BEST ACOUSTIC FEATURES FOR SMOKER IDENTIFICATION 8177
Zhizhong Ma, Yuanhang Qiu, Feng Hou, Ruili Wang, Massey University, New Zealand; Joanna Ting Wai Chu, Christopher Bullen, University of Auckland, New Zealand

SPE-76: RESOURCE CONSTRAINED SPEECH MODELING

SPE-76.1: END-TO-END LOW RESOURCE KEYWORD SPOTTING THROUGH CHARACTER RECOGNITION AND BEAM-SEARCH RE-SCORING 8182
Ephrem Tibebe Mekonnen, University of Trento, Italy; Alessio Brutti, Daniele Falavigna, Fondazione Bruno Kessler, Italy

SPE-76.2: CURRICULUM OPTIMIZATION FOR LOW-RESOURCE SPEECH RECOGNITION 8187
Anastasia Kuznetsova, Francis Tyers, Indiana University Bloomington, United States of America; Anurag Kumar, Jennifer Drexler Fox, Rev.com, United States of America

SPE-76.3: EXPLORING EFFECTIVE DATA UTILIZATION FOR LOW-RESOURCE SPEECH RECOGNITION 8192
Zhikai Zhou, Wei Wang, Wangyou Zhang, Yanmin Qian, Shanghai Jiao Tong University, China

SPE-76.4: OMNI-SPARSITY DNN: FAST SPARSITY OPTIMIZATION FOR ON-DEVICE STREAMING E2E ASR VIA SUPERNET 8197
Haichuan Yang, Yuan Shangguan, Dilin Wang, Meng Li, Pierce Chuang, Xiaohui Zhang, Ganesh Venkatesh, Ozlem Kalinli, Vikas Chandra, Facebook, United States of America

SPE-76.5: ANALYZING THE ROBUSTNESS OF UNSUPERVISED SPEECH RECOGNITION 8202
Guan-Ting Lin, Chan-Jan Hsu, Da-Rong Liu, Hung-Yi Lee, National Taiwan University, Taiwan; Yu Tsao, Academia Sinica, Taiwan

SPE-76.6: INTERPRETING INTERMEDIATE CONVOLUTIONAL LAYERS IN UNSUPERVISED ACOUSTIC WORD CLASSIFICATION 8207
Gašper Beguš, Alan Zhou, University of California, Berkeley, United States of America

SPE-77: QUESTION ANSWERING AND INFERENCE

SPE-77.1: CONTEXT MODELING WITH EVIDENCE FILTER FOR MULTIPLE CHOICE QUESTION ANSWERING 8212
Sicheng Yu, Jing Jiang, Singapore Management University, Singapore; Hao Zhang, Nanyang Technological University, Singapore; Wei Jing, Alibaba Damo Academy, China

SPE-77.2: FROM SHALLOW TO DEEP: COMPOSITIONAL REASONING OVER GRAPHS FOR VISUAL QUESTION ANSWERING 8217
Zihao Zhu, University of Chinese Academy of Sciences, China

SPE-77.3: A QUESTION-ORIENTED PROPAGATION NETWORK FOR NEWS READING COMPREHENSION 8222
Liang Wen, Houfeng Wang, Yingwei Luo, Xiaolin Wang, Peking University, China; Dehong Ma, Jun Fan, Daiting Shi, Zhicong Cheng, Dawei Yin, Baidu Inc., China

SPE-77.4: SYNTAX-BASED GRAPH MATCHING FOR KNOWLEDGE BASE QUESTION ANSWERING 8227
Lu Ma, Dan Luo, Xi Zhu, Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China; Peng Zhang, Meilin Zhou, Qi Liang, Institute of Information Engineering, Chinese Academy of Sciences, China; Bin Wang, Xiaomi Inc., China

SPE-77.5: QA4QG: USING QUESTION ANSWERING TO CONSTRAIN MULTI-HOP QUESTION GENERATION 8232
Dan Su, Peng Xu, Pascale Fung, Hong Kong University of Science and Technology, Hong Kong

SPE-77.6: PAIR-LEVEL SUPERVISED CONTRASTIVE LEARNING FOR NATURAL LANGUAGE INFERENCE 8237
Shu'ang Li, Xuming Hu, Li Lin, Lijie Wen, Tsinghua University, China

SPE-78: SPEECH PRODUCTION AND PERCEPTION

SPE-78.1: ACOUSTIC COMPARISON OF PHYSICAL VOCAL TRACT MODELS WITH HARD AND SOFT WALLS 8242
Peter Birkholz, Patrick Häsner, Steffen Kürbis, TU Dresden, Germany

SPE-78.2: AN ERROR CORRECTION SCHEME FOR IMPROVED AIR-TISSUE BOUNDARY IN REAL-TIME MRI VIDEO FOR SPEECH PRODUCTION 8247
Anwesha Roy, Varun Belagali, Prasanta Kumar Ghosh, Indian Institute of Science (IISc), Bangalore, India

SPE-78.3: REPEAT AFTER ME: SELF-SUPERVISED LEARNING OF ACOUSTIC-TO-ARTICULATORY MAPPING BY VOCAL IMITATION 8252
Marc-Antoine Georges, Laurent Girin, Jean-Luc Schwartz, Thomas Hueber, GIPSA-lab, France; Julien Diard, LPNC, France

SPE-78.4: MULTI-SPEAKER PITCH TRACKING VIA EMBODIED SELF-SUPERVISED LEARNING 8257
Xiang Li, Yifan Sun, Xihong Wu, Jing Chen, Peking University, China

SPE-78.5: IMPROVING THE CLASSIFICATION OF PHONETIC SEGMENTS FROM RAW ULTRASOUND USING SELF-SUPERVISED LEARNING AND HARD EXAMPLE MINING 8262
Yunsheng Xiong, Kele Xu, Yong Dou, National University of Defense Technology, China; Meng Jiang, Changsha Medical University, China; Liang Cheng, Xiangya Hospital, Central South University, China; Jinjia Wang, Yanshan University, China

SPE-78.6: THE IMPACT OF CROSS LANGUAGE ON ACOUSTIC-TO-ARTICULATORY INVERSION AND ITS INFLUENCE ON ARTICULATORY SPEECH SYNTHESIS 8267
Aravind Illa, Prasanta Kumar Ghosh, Indian Institute of Science, Bangalore, India; Aanish Nair, Georgia Institute of Technology, United States of America

SPE-79: SPEECH RECOGNITION: TRANSFORMERS

SPE-79.1: TRANSFORMER-BASED STREAMING ASR WITH CUMULATIVE ATTENTION 8272
Mohan Li, Shucong Zhang, Cătălin Zorilă, Rama Doddipatla, Toshiba Cambridge Research Laboratory, Toshiba Europe Ltd, United Kingdom of Great Britain and Northern Ireland

SPE-79.2: STREAMING TRANSFORMER TRANSDUCER BASED SPEECH RECOGNITION USING NON-CAUSAL CONVOLUTION 8277
Yangyang Shi, Chunyang Wu, Dilin Wang, Alex Xiao, Jay Mahadeokar, Xiaohui Zhang, Chunxi Liu, Ke Li, Yuan Shangguan, Varun Nagaraja, Ozlem Kalinli, Mike Seltzer, Facebook AI, United States of America

SPE-79.3: HYBRID RNN-T/ATTENTION-BASED STREAMING ASR WITH TRIGGERED CHUNKWISE ATTENTION AND DUAL INTERNAL LANGUAGE MODEL INTEGRATION 8282
Takafumi Moriya, Takanori Ashihara, Atsushi Ando, Hiroshi Sato, Tomohiro Tanaka, Kohei Matsuura, Ryo Masumura, Marc Delcroix, NTT Corporation, Japan; Takahiro Shinozaki, Tokyo Institute of Technology, Japan

SPE-79.4: RUN-AND-BACK STITCH SEARCH: NOVEL BLOCK SYNCHRONOUS DECODING FOR STREAMING ENCODER-DECODER ASR 8287
Emiru Tsunoo, Michael Hentschel, Yosuke Kashiwagi, Sony Group Corporation, Japan; Chaitanya Narisetty, Shinji Watanabe, Carnegie Mellon University, United States of America

SPE-79.5: ALIGNMENT-LEARNING BASED SINGLE-STEP DECODING FOR ACCURATE AND FAST NON-AUTOREGRESSIVE SPEECH RECOGNITION 8292
Yonghe Wang, Rui Liu, Feilong Bao, Hui Zhang, Guanglai Gao, Inner Mongolia University, China

SPE-79.6: USTED: IMPROVING ASR WITH A UNIFIED SPEECH AND TEXT ENCODER-DECODER 8297
Bolaji Yusuf, Bogazici University, Turkey; Ankur Gandhe, Alex Sokolov, Amazon, United States of America

SPE-80: SPEECH SYNTHESIS: EXPRESSIVENESS AND VOICE CLONING

SPE-80.1: IMPROVING EMOTIONAL SPEECH SYNTHESIS BY USING SUS-CONSTRAINED VAE AND TEXT ENCODER AGGREGATION 8302
Fengyu Yang, Jian Luan, Yujun Wang, Xiaomi Corporation, China

SPE-80.2: DISTRIBUTION AUGMENTATION FOR LOW-RESOURCE EXPRESSIVE TEXT-TO-SPEECH 8307
Mateusz Lajszczak, Animesh Prasad, Arent van Korlaar, Bajibabu Bollepalli, Antonio Bonafonte, Arnaud Joly, Marco Nicolis, Alexis Moinet, Thomas Drugman, Trevor Wood, Elena Sokolova, Amazon, United Kingdom of Great Britain and Northern Ireland

SPE-80.3: INTERACTIVE MULTI-LEVEL PROSODY CONTROL FOR EXPRESSIVE SPEECH SYNTHESIS 8312
Tobias Cornille, Jessa Bekker, KU Leuven, Belgium; Fengna Wang, Acapela-Group, Belgium

SPE-80.4: IMPROVE FEW-SHOT VOICE CLONING USING MULTI-MODAL LEARNING 8317
Haitong Zhang, Yue Lin, Netease Games AI Lab, China

SPE-80.5: CLONING ONE'S VOICE USING VERY LIMITED DATA IN THE WILD 8322
Dongyang Dai, Yuanzhe Chen, Li Chen, Ming Tu, Lu Liu, Rui Xia, Qiao Tian, Yuping Wang, Yuxuan Wang, bytedance, China

SPE-80.6: UNET-TTS: IMPROVING UNSEEN SPEAKER AND STYLE TRANSFER IN ONE-SHOT VOICE CLONING 8327
Rui Li, Dong Pu, Minnie Huang, Bill Huang, CloudMinds, China

SPE-81: ENTITY RECOGNITION

SPE-81.1: IMPROVING BIOMEDICAL NAMED ENTITY RECOGNITION WITH A UNIFIED MULTI-TASK MRC FRAMEWORK 8332
Yiqi Tong, Fuzhen Zhuang, Deqing Wang, Beihang University, China; Haochao Ying, Zhejiang University, China; Binling Wang, Xiamen University, China

SPE-81.2: A MULTI-TASK LEARNING FRAMEWORK FOR CHINESE MEDICAL PROCEDURE ENTITY NORMALIZATION 8337
Xuhui Sui, Kehui Song, Baohang Zhou, Ying Zhang, Xiaojie Yuan, Nankai University, China

SPE-81.3: WASSERSTEIN CROSS-LINGUAL ALIGNMENT FOR NAMED ENTITY RECOGNITION 8342
Rui Wang, Ricardo Henao, Duke University, United States of America

SPE-81.4: LEARNING COMMON DEPENDENCY STRUCTURE FOR UNSUPERVISED CROSS-DOMAIN NER 8347
Luchen Liu, Bin Wang, Xiaomi Inc., China; Xixun Lin, Peng Zhang, Lei Zhang, Institute of Information Engineering, Chinese Academy of Sciences, China

SPE-81.5: AISHELL-NER: NAMED ENTITY RECOGNITION FROM CHINESE SPEECH 8352
Boli Chen, Guangwei Xu, Xiaobin Wang, Pengjun Xie, Meishan Zhang, Fei Huang, Alibaba Group, China

SPE-81.6: CALL-SIGN RECOGNITION AND UNDERSTANDING FOR NOISY AIR-TRAFFIC TRANSCRIPTS USING SURVEILLANCE INFORMATION 8357
Alexander Blatt, Dietrich Klakow, Saarland University, Germany; Martin Kocour, Karel Veselý, Igor Szöke, Brno University of Technology, Germany

SPE-82: SPEAKER DIARIZATION II

SPE-82.1: INCORPORATING END-TO-END FRAMEWORK INTO TARGET-SPEAKER VOICE ACTIVITY DETECTION 8362
Weiqing Wang, Duke University, United States of America; Ming Li, Duke Kunshan University, China

SPE-82.2: MULTI-SCALE SPEAKER EMBEDDING-BASED GRAPH ATTENTION NETWORKS FOR SPEAKER DIARISATION 8367
Youngki Kwon, Hee-Soo Heo, Jee-weon Jung, You Jin Kim, Bong-Jin Lee, Naver Corporation, Korea, Republic of; Joon Son Chung, Korea Advanced Institute of Science and Technology, Korea, Republic of

SPE-82.3: TOWARDS END-TO-END SPEAKER DIARIZATION WITH GENERALIZED NEURAL SPEAKER CLUSTERING 8372
Chunlei Zhang, Chao Weng, Meng Yu, Dong Yu, Tencent AI Lab, United States of America; Jiatong Shi, Carnegie Mellon University, United States of America

SPE-82.4: AUXILIARY LOSS OF TRANSFORMER WITH RESIDUAL CONNECTION FOR END-TO-END SPEAKER DIARIZATION 8377
Yechan Yu, Dongkeon Park, Hong Kook Kim, Gwangju Institute of Science and Technology, Korea, Republic of

SPE-82.5: TIGHT INTEGRATION OF NEURAL- AND CLUSTERING-BASED DIARIZATION THROUGH DEEP UNFOLDING OF INFINITE GAUSSIAN MIXTURE MODEL 8382
Keisuke Kinoshita, Marc Delcroix, Tomoharu Iwata, NTT, Japan

SPE-82.6: IMPROVING SEPARATION-BASED SPEAKER DIARIZATION VIA ITERATIVE MODEL REFINEMENT AND SPEAKER EMBEDDING BASED POST-PROCESSING 8387
Shu-Tong Niu, Jun Du, University of Science and Technology of China, China; Lei Sun, iFlytek Research, China; Chin-Hui Lee, Georgia Institute of Technology, United States of America

SPE-83: NEW ALGORITHM FOR SPEECH RECOGNITION

SPE-83.1: KNOWLEDGE DISTILLATION FROM LANGUAGE MODEL TO ACOUSTIC MODEL: A HIERARCHICAL MULTI-TASK LEARNING APPROACH 8392
Munhak Lee, Joonhyuk Chang, Hanyang University, Korea, Republic of

SPE-83.2: IMPROVING PSEUDO-LABEL TRAINING FOR END-TO-END SPEECH RECOGNITION USING GRADIENT MASK 8397
Shaoshi Ling, Chen Shen, Meng Cai, Zejun Ma, bytedance, China

SPE-83.3: MULTI-TURN RNN-T FOR STREAMING RECOGNITION OF MULTI-PARTY SPEECH 8402
Ilya Sklyar, Anna Piunova, Amazon, Germany; Xianrui Zheng, University of Cambridge, United Kingdom of Great Britain and Northern Ireland; Yulan Liu, DeepMind, United Kingdom of Great Britain and Northern Ireland

SPE-83.4: ON LANGUAGE MODEL INTEGRATION FOR RNN TRANSDUCER BASED SPEECH RECOGNITION 8407
Wei Zhou, Zuoyun Zheng, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany

SPE-83.5: CACHING NETWORKS: CAPITALIZING ON COMMON SPEECH FOR ASR 8412
Anastasios Alexandridis, Grant Strimel, Ariya Rastrow, Pavel Kveton, Jon Webb, Maurizio Omologo, Siegfried Kunzmann, Athanasios Mouchtaris, Amazon.com, United States of America

SPE-83.6: GPU-ACCELERATED FORWARD-BACKWARD ALGORITHM WITH APPLICATION TO LATTICE-FREE MMI 8417
Lucas Ondel, Léa-Marie Lam-Yee-Mui, Caio Corro, Laboratoire Interdisciplinaire des Sciences du Numérique, France; Martin Kocour, Lukas Burget, Brno University of Technology, France

SPE-84: SPEECH SYNTHESIS: VOCODER AND EVALUATION

SPE-84.1: ITOWAVE: ITO STOCHASTIC DIFFERENTIAL EQUATION IS ALL YOU NEED FOR WAVE GENERATION 8422
Shoule Wu, Yangzhou University, China; Ziqiang Shi, Fujitsu R & D Center, China

SPE-84.2: MULTI-SAMPLE SUBBAND WAVERNN VIA MULTIVARIATE GAUSSIAN 8427
Hiroki Kanagawa, Yusuke Ijima, NTT Corporation, Japan

SPE-84.3: INFERGRAD: IMPROVING DIFFUSION MODELS FOR VOCODER BY 8432
CONSIDERING INFERENCE IN TRAINING
Zehua Chen, Danilo Mandic, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Xu Tan, Ke Wang, Shifeng Pan, Lei He, Sheng Zhao, Microsoft, China; , ,

SPE-84.4: NEURAL SPEECH SYNTHESIS ON A SHOESTRING: IMPROVING THE 8437
EFFICIENCY OF LPCNET
Jean-Marc Valin, Umut Isik, Paris Smaragdis, Arvindh Krishnaswamy, Amazon, Canada

SPE-84.5: GENERALIZATION ABILITY OF MOS PREDICTION NETWORKS..... 8442
Erica Cooper, Junichi Yamagishi, National Institute of Informatics, Japan; Wen-Chin Huang, Tomoki Toda, Nagoya University, Japan

SPE-84.6: ON THE INTERPLAY BETWEEN SPARSITY, NATURALNESS, ... 8447
INTELLIGIBILITY, AND PROSODY IN SPEECH SYNTHESIS
Cheng-I Lai, Yi-Lun Liao, Yung-Sung Chuang, Alexander Liu, James Glass, MIT CSAIL, United States of America; Erica Cooper, Junichi Yamagishi, National Institute of Informatics, Japan; Yang Zhang, Shiyu Chang, Kaizhi Qian, David Cox, MIT-IBM Watson AI Lab, United States of America

SPE-85: SIGN LANGUAGE AND LIP READING

SPE-85.1: PHONOLOGY RECOGNITION IN AMERICAN SIGN LANGUAGE 8452
Federico Tavella, Aphrodite Galata, Angelo Cangelosi, The University of Manchester, United Kingdom of Great Britain and Northern Ireland

SPE-85.2: SPATIO-TEMPORAL GRAPH CONVOLUTIONAL NETWORKS FOR 8457
CONTINUOUS SIGN LANGUAGE RECOGNITION
Maria Parelli, Petros Maragos, National Technical University of Athens, Greece; Katerina Papadimitriou, Gerasimos Potamianos, University of Thessaly, Greece; Georgios Pavlakos, University of California, United States of America

SPE-85.3: SENSORS TO SIGN LANGUAGE: A NATURAL APPROACH TO EQUITABLE 8462
COMMUNICATION
Thomas Fouts, University of Michigan, United States of America; Ali Hindy, Stanford University, United States of America; Chris Tanner, Harvard University, United States of America

SPE-85.4: ACCURATE AND RESOURCE-EFFICIENT LIPREADING WITH 8467
EFFICIENTNETV2 AND TRANSFORMERS
Alexandros Koumparoulis, Gerasimos Potamianos, University of Thessaly, Greece

SPE-85.5: TRAINING STRATEGIES FOR IMPROVED LIP-READING 8472
Pingchuan Ma, Yujiang Wang, Stavros Petridis, Jie Shen, Maja Pantic, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SPE-85.6: MULTISTREAM NEURAL ARCHITECTURES FOR CUED SPEECH 8477
RECOGNITION USING A PRE-TRAINED VISUAL FEATURE EXTRACTOR AND CONSTRAINED
CTC DECODING
Sanjana Sankar, Denis Beautemps, Thomas Hueber, Centre National de la Recherche Scientifique, France

SPE-86: SPEECH PARALINGUISTICS: REPRESENTATION LEARNING

SPE-86.1: MODELING OF PRE-TRAINED NEURAL NETWORK EMBEDDINGS 8482
LEARNED FROM RAW WAVEFORM FOR COVID-19 INFECTION DETECTION
Zohreh Mostaani, RaviShankar Prasad, Bogdan Vlasenko, Mathew Magimai Doss, Idiap Research Institute, Switzerland

SPE-86.2: DUAL ATTENTION POOLING NETWORK FOR RECORDING DEVICE CLASSIFICATION USING NEUTRAL AND WHISPERED SPEECH 8487
Abinay Reddy Naini, Prasanta Kumar Ghosh, Indian Institute of Science, Bangalore, India; Bhavuk Singhal, Information Technology, Bundelkhand Institute of Engineering and Technology, Jhansi, India, India

SPE-86.3: ENTRAINMENT ANALYSIS FOR ASSESSMENT OF AUTISTIC SPEECH PROSODY USING BOTTLENECK FEATURES OF DEEP NEURAL NETWORK 8492
Keiko Ochi, Kyoto University, Japan; Nobutaka Ono, Tokyo Metropolitan University, Japan; Keiho Owada, Miho Kuroda, Shigeki Sagayama, University of Tokyo, Japan; Hidenori Yamasue, Hamamatsu University School of Medicine, Japan

SPE-86.4: CUSTOMER SATISFACTION ESTIMATION USING UNSUPERVISED REPRESENTATION LEARNING WITH MULTI-FORMAT PREDICTION LOSS 8497
Atsushi Ando, Yumiko Murata, Ryo Masumura, Satoshi Suzuki, Naoki Makishima, Takafumi Moriya, Takanori Ashihara, Hiroshi Sato, NTT Corporation, Japan

SPE-86.5: AUTOMATIC ASSESSMENT OF THE DEGREE OF CLINICAL DEPRESSION FROM SPEECH USING X-VECTORS 8502
José Vicente Egas-López, University of Szeged, Hungary; Gábor Kiss, Sztahó David, Budapest University of Technology and Economics, Hungary; Gábor Gosztolya, MTA-SZTE Research Group on Artificial Intelligence, Hungary

SPE-86.6: AUTOMATIC DEPRESSION LEVEL ASSESSMENT FROM SPEECH BY LONG-TERM GLOBAL INFORMATION EMBEDDING 8507
Ya Li, Beijing University of Posts and Telecommunications, China; Mingyue Niu, Ziping Zhao, Tianjin Normal University, China; Jianhua Tao, Institute of Automation, Chinese Academy of Sciences, China

SPE-87: SPEECH RECOGNITION: KNOWLEDGE TRANSFER AND CONTEXTUALIZATION

SPE-87.1: KNOWLEDGE TRANSFER FROM LARGE-SCALE PRETRAINED LANGUAGE MODELS TO END-TO-END SPEECH RECOGNIZERS 8512
Yotaro Kubo, Shigeki Karita, Michiel Bacchiani, Google, Japan

SPE-87.2: IMPROVING CTC-BASED SPEECH RECOGNITION VIA KNOWLEDGE TRANSFERRING FROM PRE-TRAINED LANGUAGE MODELS 8517
Keqi Deng, Gaofeng Cheng, Ji Xu, Pengyuan Zhang, Institute of Acoustics, Chinese Academy of Sciences, China; Songjun Cao, Yike Zhang, Long Ma, Tencent Cloud Xiaowei, China

SPE-87.3: IMPROVING NON-AUTOREGRESSIVE END-TO-END SPEECH RECOGNITION WITH PRE-TRAINED ACOUSTIC AND LANGUAGE MODELS 8522
Keqi Deng, Zehui Yang, Gaofeng Cheng, Pengyuan Zhang, Institute of Acoustics, Chinese Academy of Sciences, China; Shinji Watanabe, Carnegie Mellon University, United States of America; Yosuke Higuchi, Waseda University, Japan

SPE-87.4: KNOWLEDGE DISTILLATION FOR NEURAL TRANSDUCERS FROM LARGE SELF-SUPERVISED PRE-TRAINED MODELS 8527
Xiaoyu Yang, Qiujia Li, Phil Woodland, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

SPE-87.5: IMPROVING END-TO-END CONTEXTUAL SPEECH RECOGNITION WITH FINE-GRAINED CONTEXTUAL KNOWLEDGE SELECTION 8532
Minglun Han, Shiyu Zhou, Bo Xu, Institute of Automation, Chinese Academy of Sciences, China; Linhao Dong, Zhenlin Liang, Meng Cai, Zejun Ma, Bytedance AI Lab, China

SPE-87.6: CONTEXTUAL ADAPTERS FOR PERSONALIZED SPEECH RECOGNITION IN NEURAL TRANSDUCERS 8537
Kanthashree Mysore Sathyendra, Thejaswi Muniyappa, Feng-Ju Chang, Jing Liu, Jinru Su, Grant P. Strimel, Athanasios Mouchtaris, Siegfried Kunzmann, Amazon, United States of America

SPE-88: SPEECH ANALYSIS: PARALINGUISTICS

SPE-88.1: EMOTIONFLOW: CAPTURE THE DIALOGUE LEVEL EMOTION TRANSITIONS 8542
Xiaohui Song, 1. Institute of Information Engineering, Chinese Academy of Sciences; 2. School of Cyber Security, University of Chinese Academy of Sciences; 3. Alibaba Group, China; Liangjun Zang, Songlin Hu, 1. Institute of Information Engineering, Chinese Academy of Sciences;2. School of Cyber Security, University of Chinese Academy of Sciences, China; Rong Zhang, Longtao Huang, Alibaba Group, China

SPE-88.2: MULTIMODAL SENTIMENT ANALYSIS ON UNALIGNED SEQUENCES VIA HOLOGRAPHIC EMBEDDING 8547
Yukun Ma, Bin Ma, Alibaba Group, Singapore

SPE-88.3: DISTRIBUTION LEARNING FOR AGE ESTIMATION FROM SPEECH 8552
Amruta Saraf, Elie Khoury, Pindrop, United States of America

SPE-88.4: DISPEECH: A SYNTHETIC TOY DATASET FOR SPEECH DISENTANGLING 8557
Olivier Zhang, Nicolas Gengembre, Olivier Le Blouch, Orange, France; Damien Lolive, IRISA, France

SPE-88.5: END-TO-END ASR-ENHANCED NEURAL NETWORK FOR ALZHEIMER'S DISEASE DIAGNOSIS 8562
Jiancheng Gui, Yikai Li, Kai Chen, Joanna Siebert, Qingcai Chen, Harbin Institute of Technology (Shenzhen), China, China

SPE-88.6: A NOVEL SEQUENTIAL MONTE CARLO FRAMEWORK FOR PREDICTING AMBIGUOUS EMOTION STATES 8567
Jingyao Wu, Vidhyasaharan Sethu, Eliathamby Ambikairajah, University of New South Wales, Australia; Ting Dang, University of Cambidge, United Kingdom of Great Britain and Northern Ireland

SPE-89: AUGMENTATION FOR SPEECH RECOGNITION

SPE-89.1: PHONE-INFORMED REFINEMENT OF SYNTHESIZED MEL SPECTROGRAM FOR DATA AUGMENTATION IN SPEECH RECOGNITION 8572
Sei Ueno, Tatsuya Kawahara, Graduate School of Informatics, Kyoto University, Japan

SPE-89.2: LPC AUGMENT: AN LPC-BASED ASR DATA AUGMENTATION ALGORITHM FOR LOW AND ZERO-RESOURCE CHILDREN'S DIALECTS 8577
Alexander Johnson, Ruchao Fan, Abeer Alwan, University of California - Los Angeles, United States of America; Robin Morris, Georgia State University, United States of America

SPE-89.3: TOWARDS BETTER META-INITIALIZATION WITH TASK AUGMENTATION FOR KINDERGARTEN-AGED SPEECH RECOGNITION 8582
Yunzheng Zhu, Ruchao Fan, Abeer Alwan, University of California, Los Angeles, United States of America

SPE-89.4: UNSUPERVISED DATA SELECTION FOR SPEECH RECOGNITION WITH CONTRASTIVE LOSS RATIOS 8587
Chanho Park, Rehan Ahmad, Thomas Hain, The University of Sheffield, United Kingdom of Great Britain and Northern Ireland

SPE-89.5: IMPORTANTAUG: A DATA AUGMENTATION AGENT FOR SPEECH 8592
Viet Anh Trinh, Hassan Kavaki, The City University of New York, United States of America; Michael Mandel, Brooklyn College, United States of America

SPE-89.6: INJECTING TEXT AND CROSS-LINGUAL SUPERVISION IN FEW-SHOT LEARNING FROM SELF-SUPERVISED MODELS 8597
Matthew Wiesner, Human Language Technology Center of Excellence, United States of America; Desh Raj, Sanjeev Khudanpur, Johns Hopkins University, United States of America

SS-1: QUANTUM MACHINE LEARNING FOR SPEECH AND LANGUAGE PROCESSING

SS-1.1: WHEN BERT MEETS QUANTUM TEMPORAL CONVOLUTION LEARNING 8602
FOR TEXT CLASSIFICATION IN HETEROGENEOUS COMPUTING
Chao-Han Huck Yang, Jun Qi, Georgia Institute of Technology, United States of America; Samuel Yen-Chi Chen, Brookhaven National Laboratory, United States of America; Yu Tsao, Academia Sinica, Taiwan, Province of China; Pin-Yu Chen, IBM Research, United States of America

SS-1.2: MATCHING POINT SETS WITH QUANTUM CIRCUIT LEARNING..... 8607
Mohammadreza Noormandipour, Hanchen Wang, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

SS-1.3: THE DAWN OF QUANTUM NATURAL LANGUAGE PROCESSING 8612
Riccardo Di Sipio, Ceridian HCM Inc., Italy; Jia-Hong Huang, Marcel Worring, University of Amsterdam, Netherlands; Samuel Yen-Chi Chen, Brookhaven National Laboratory, United States of America; Stefano Mangini, University of Pavia, Italy

SS-1.4: QUANTUM FEDERATED LEARNING WITH QUANTUM DATA 8617
Mahdi Chehimi, Walid Saad, Virginia Tech, United States of America

SS-1.5: QUANTUM LONG SHORT-TERM MEMORY 8622
Samuel Yen-Chi Chen, Shinjae Yoo, Yao-Lung L. Fang, Brookhaven National Laboratory, United States of America

SS-1.6: CLASSICAL-TO-QUANTUM TRANSFER LEARNING FOR SPOKEN COMMAND 8627
RECOGNITION BASED ON QUANTUM NEURAL NETWORKS
Jun Qi, Georgia Institute of Technology, United States of America; Javier Tejedor, Universidad San Pablo-CEU, CEU Universities, Spain

SS-2: SIGNAL PROCESSING AND MACHINE LEARNING FOR ENERGY HARVESTING COMMUNICATIONS

SS-2.1: WAVEFORM OPTIMIZATION FOR WIRELESS POWER TRANSFER WITH 8632
POWER AMPLIFIER AND ENERGY HARVESTER NON-LINEARITIES
Yumeng Zhang, Bruno Clerckx, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SS-2.2: ECONOMICS OF SEMANTIC COMMUNICATION SYSTEM IN WIRELESS 8637
POWERED INTERNET OF THINGS
Zi Qin Liew, Yanyu Cheng, Wei Yang Bryan Lim, Alibaba-NTU Singapore Joint Research Institute, Nanyang Technological Univerity, Singapore, Singapore; Dusit Niyato, School of Computer Science and Engineering, Nanyang Technological University, Singapore, Singapore; Chunyan Miao, Alibaba-NTU Singapore Joint Research Institute, Nanyang Technological Univerity, Singapore and School of Computer Science and Engineering, Nanyang Technological University, Singapore, Singapore; Sumei Sun, Institute of Infocomm Research, Agency for Science, Technology and Research, Singapore, Singapore

SS-2.3: OPTIMAL RESOURCE ALLOCATION AND BEAMFORMING FOR TWO-USER 8642
MISO WPCNS FOR A NON-LINEAR CIRCUIT-BASED EH MODEL
Nikita Shanin, Moritz Garkisch, Robert Schober, Laura Cottatellucci, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany; Amelie Hagelauer, Fraunhofer EMFT Einrichtung für Mikrosysteme und Festkörper-Technologien; Technische Universität München, Germany

SS-2.4: PERFORMANCE OPTIMIZATION FOR WIRELESS SEMANTIC 8647
COMMUNICATIONS OVER ENERGY HARVESTING NETWORKS
Mingzhe Chen, H. Vincent Poor, Princeton University, United States of America; Yining Wang, Beijing University of Posts and Telecommunications, China

SS-2.5: DEEP LEARNING BASED PASSIVE BEAMFORMING FOR IRS-ASSISTED 8652
MONOSTATIC BACKSCATTER SYSTEMS
Sahar Idrees, Xiaolun Jia, Saud Khan, Salman Durrani, Xiangyun Zhou, Australian National University, Australia

SS-2.6: ON FEDERATED LEARNING WITH ENERGY HARVESTING CLIENTS 8657
Cong Shen, University of Virginia, United States of America; Jing Yang, The Pennsylvania State University, United States of America; Jie Xu, University of Miami, United States of America

SS-3: MODEL-BASED DEEP LEARNING IN SIGNAL PROCESSING

SS-3.1: STRUCTURAL PRIOR MODELS FOR 3-D DEEP VESSEL SEGMENTATION 8662
Xuelu Li, Vishal Monga, Pennsylvania State University, University Park, United States of America; Raja Bala, Amazon Research, United States of America

SS-3.2: EXPECTATION CONSISTENT PLUG-AND-PLAY FOR MRI 8667
Saurav Shastri, Rizwan Ahmad, Philip Schniter, Ohio State, United States of America; Christopher Metzler, Univ. Maryland, United States of America

SS-3.3: INVERSE IMAGING WITH GENERATIVE PRIORS VIA LANGEVIN DYNAMICS 8672
Thanh Nguyen, Unaffiliated, United States of America; Gauri Jagatap, Chinmay Hegde, New York University, United States of America

SS-3.4: CNN-AIDED FACTOR GRAPHS WITH ESTIMATED MUTUAL INFORMATION 8677
FEATURES FOR SEIZURE DETECTION
Bahareh Salafian, Nariman Farsad, Ryerson University, Canada; Eyal Fishel Ben-Knaan, Nir Shlezinger, Ben-Gurion University of the Negev, Israel; Sandrine de Ribaupierre, University of Western Ontario, Canada

SS-3.5: UNFOLDING MODEL-BASED BEAMFORMING FOR HIGH QUALITY 8682
ULTRASOUND IMAGING
Christopher Khan, Brett Byram, Vanderbilt University, United States of America; Ruud van Sloun, Eindhoven University of Technology, Netherlands

SS-3.6: OPTIMIZATION GUARANTEES FOR ISTA AND ADMM BASED UNFOLDED 8687
NETWORKS
Wei Pu, Miguel Rodrigues, UCL, United Kingdom of Great Britain and Northern Ireland; Yonina Eldar, Weizmann institute of science, Israel

SS-4: HUMAN CENTRIC ACTIVE NOISE CONTROL APPLICATIONS

SS-4.1: INTEGRATION OF ANOMALY MACHINE SOUND DETECTION INTO ACTIVE 8692
NOISE CONTROL TO SHAPE THE RESIDUAL SOUND
Chuang Shi, Mengjie Huang, Huitian Jiang, Huiyong Li, University of Electronic Science and Technology of China, China

SS-4.2: DUAL ACTIVE NOISE CONTROL WITH COMMON SENSORS 8697
Ryosuke Okajima, Yoshinobu Kajikawa, Kansai University, Japan; Kohei Oto, Nitto Denko Corporation, Japan

SS-4.3: A HYBRID APPROACH TO COMBINE WIRELESS AND EARCUP 8702
MICROPHONES FOR ANC HEADPHONES WITH ERROR SEPARATION MODULE
Xiaoyi Shen, Dongyuan Shi, Woon-Seng Gan, Nanyang Technological University, Singapore

SS-4.4: SPATIAL ACTIVE NOISE CONTROL WITH THE REMOTE MICROPHONE 8707
TECHNIQUE: AN APPROACH WITH A MOVING HIGHER ORDER MICROPHONE
Huiyuan Sun, Jihui Zhang, Thushara Abhayapala, Prasanga Samarasinghe, The Australian National University, Australia

SS-4.5: ROBUST PRESSURE MATCHING WITH ATF PERTURBATION CONSTRAINTS 8712
FOR SOUND FIELD CONTROL
Junqing Zhang, Wen Zhang, Lijun Zhang, Jingdong Chen, Center of Intelligent Acoustics and Immersive Communications, School of Marine Science and Technology, Northwestern Polytechnical University, China; Liming Shi, Mads Græsbøll Christensen, Audio Analysis Lab, CREATE, Aalborg University, China

SS-4.6: OPTIMIZATION OF A FIXED VIRTUAL SENSING FEEDBACK ANC CONTROLLER FOR IN-EAR HEADPHONES WITH MULTIPLE LOUDSPEAKERS 8717
Piero Rivera Benois, Simon Doclo, Signal Processing Group, University of Oldenburg, Oldenburg, Germany, Germany; Reinhild Roden, Matthias Blau, Institut fuer Hoertechnik und Audiologie, Jade Hochschule, Oldenburg, Germany, Germany

SS-5: INTEGRATED SENSING AND COMMUNICATIONS FOR FUTURE CELLULAR AND VEHICULAR NETWORKS

SS-5.1: ON THE POTENTIAL OF SPATIALLY-SPREAD ORTHOGONAL TIME FREQUENCY SPACE MODULATION FOR ISAC TRANSMISSIONS 8722
Shuangyang Li, Jinhong Yuan, University of New South Wales, Australia; Weijie Yuan, Southern University of Science and Technology, China; Giuseppe Caire, Technische Universität Berlin, China

SS-5.2: SENSING-ASSISTED BEAM TRACKING IN V2I NETWORKS: EXTENDED TARGET CASE 8727
Zhen Du, Zenghui Zhang, Shanghai Jiao Tong University, China; Fan Liu, Southern University of Science and Technology, China

SS-5.3: TRANSMIT BEAMFORMING WITH FIXED COVARIANCE FOR INTEGRATED MIMO RADAR AND MULTIUSER COMMUNICATIONS 8732
Xiang Liu, Tianyao Huang, Yimin Liu, Tsinghua University, China; Yonina Eldar, Weizmann institute of Science, Israel

SS-5.4: SAFEGUARDING UAV NETWORKS THROUGH INTEGRATED SENSING, JAMMING, AND COMMUNICATIONS 8737
Zhiqiang Wei, Robert Schober, Friedrich-Alexander University Erlangen-Nuremberg, Germany; Fan Liu, Southern University of Science and Technology, China; Derrick Wing Kwan Ng, University of New South Wales, Australia

SS-5.5: EVALUATION OF ORTHOGONAL CHIRP DIVISION MULTIPLEXING FOR AUTOMOTIVE INTEGRATED SENSING AND COMMUNICATIONS 8742
Sangeeta Bhattacharjee, Chandra Murthy, Indian Institute of Science, Bangalore, India; Kumar Vijay Mishra, United States CCDC Army Research Laboratory, United States of America; Ramesh Annavajjala, University of Massachusetts, United States of America

SS-5.6: INTEGRATED SENSING AND COMMUNICATIONS VIA 5G NR WAVEFORM: PERFORMANCE ANALYSIS 8747
Yuanhao Cui, Xiaojun Jing, Junsheng Mu, Beijing University of Posts and Telecommunications, China

SS-6: FRONTIERS OF FEDERATED LEARNING: APPLICATIONS, CHALLENGES, AND OPPORTUNITIES

SS-6.1: FEDERATED LEARNING CHALLENGES AND OPPORTUNITIES: AN OUTLOOK 8752
Jie Ding, Eric Tramel, Anit Kumar Sahu, Shuang Wu, Salman Avestimehr, Tao Zhang, Amazon, United States of America

SS-6.2: ENABLING ON-DEVICE TRAINING OF SPEECH RECOGNITION MODELS WITH FEDERATED DROPOUT 8757
Dhruv Guliani, Lillian Zhou, Changwan Ryu, Tien-Ju Yang, Harry Zhang, Yonghui Xiao, Francoise Beaufays, Giovanni Motta, Google Inc, United States of America

SS-6.3: ADAPTIVE NODE PARTICIPATION FOR STRAGGLER-RESILIENT FEDERATED LEARNING 8762
Amirhossein Reisizadeh, MIT, United States of America; Isidoros Tziotis, Aryan Mokhtari, UT Austin, United States of America; Hamed Hassani, UPenn, United States of America; Ramtin Pedarsani, UC Santa Barbara, United States of America

SS-6.4: LEARNINGS FROM FEDERATED LEARNING IN THE REAL WORLD 8767
Christophe Dupuy, Tanya Roosta, Leo Long, Clement Chung, Rahul Gupta, Amazon Alexa AI, United States of America; Salman Avestimehr, Amazon Alexa AI ; University of Southern California (USC), United States of America

SS-6.5: A DYNAMIC REWEIGHTING STRATEGY FOR FAIR FEDERATED LEARNING..... 8772
Zhiyuan Zhao, Gauri Joshi, Carnegie Mellon University, United States of America

SS-6.6: OVER-THE-AIR PERSONALIZED FEDERATED LEARNING..... 8777
Hasin Us Sami, Basak Guler, University of California, Riverside, United States of America

SS-7: LEARNING-BASED ACOUSTIC SIGNAL PROCESSING FOR STAY-AT-HOME APPLICATIONS IN PANDEMIC TIMES

SS-7.1: DNN BASED MULTIFRAME SINGLE-CHANNEL NOISE REDUCTION 8782
FILTERS
Ningning Pan, Jingdong Chen, Northwestern Polytechnical University, China; Jacob Benesty, University of Quebec, Canada

SS-7.2: LEARNING-BASED PERSONAL SPEECH ENHANCEMENT FOR 8787
TELECONFERENCING BY EXPLOITING SPATIAL-SPECTRAL FEATURES
Yicheng Hsu, Yonghan Lee, Mingsian Bai, National Tsing Hua University, Taiwan

SS-7.3: MANIFOLD LEARNING-SUPPORTED ESTIMATION OF RELATIVE TRANSFER 8792
FUNCTIONS FOR SPATIAL FILTERING
Andreas Brendel, Johannes Zeitler, Walter Kellermann, Friedrich-Alexander University Erlangen-Nuremberg, Germany

SS-7.4: AUDIO SIGNAL PROCESSING FOR TELEPRESENCE BASED ON WEARABLE 8797
ARRAY IN NOISY AND DYNAMIC SCENES
Hanan Beit-On, Moti Lugasi, Lior Madmoni, Boaz Rafaely, Ben-Gurion University, Israel; Anjali Menon, Anurag Kumar, Jacob Donley, Vladimir Tourbabin, Facebook, United States of America

SS-7.5: A MULTI-TASK LEARNING METHOD FOR WEAKLY SUPERVISED SOUND 8802
EVENT DETECTION
Sichen Liu, Feiran Yang, Fang Kang, Jun Yang, Institute of Acoustics, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China

SS-7.6: LOW RESOURCES ONLINE SINGLE-MICROPHONE SPEECH 8807
ENHANCEMENT WITH HARMONIC EMPHASIS
Nir Raviv, Ofer Schwartz, CEVA-DSP, Israel; Sharon Gannot, Bar-Ilan University, Israel

SS-8: MACHINE LEARNING IN BEYOND 5G WIRELESS NETWORKS

SS-8.1: DEEP LEARNING FOR LOCATION BASED BEAMFORMING WITH NLOS 8812
CHANNELS
Luc Le Magoarou, Stéphane Paquelet, bcom, France; Taha Yassine, Matthieu Crussière, bcom and INSA Rennes, France

SS-8.2: PREDICTING FLAT-FADING CHANNELS VIA META-LEARNED CLOSED-FORM 8817
LINEAR FILTERS AND EQUILIBRIUM PROPAGATION
Sangwoo Park, Osvaldo Simeone, King's College London, United Kingdom of Great Britain and Northern Ireland

SS-8.3: DEEP-LEARNING-ASSISTED CONFIGURATION OF RECONFIGURABLE 8822
INTELLIGENT SURFACES IN DYNAMIC RICH-SCATTERING ENVIRONMENTS
Kyriakos Stylianopoulos, George Alexandropoulos, National and Kapodistrian University of Athens, Greece; Nir Shlezinger, Ben-Gurion University of the Negev, Israel; Philipp del Hougne, Univ Rennes, France

SS-8.4: SUPERVISED LEARNING BASED SPARSE CHANNEL ESTIMATION FOR RIS AIDED COMMUNICATIONS **8827**
Dilin Dampahalage, K. B. Shashika Manosha, Nandana Rajatheva, Matti Latva-aho, University of Oulu, Finland

SS-8.5: GOAL-ORIENTED COMMUNICATION FOR EDGE LEARNING BASED ON THE INFORMATION BOTTLENECK **8832**
Francesco Pezone, Sergio Barbarossa, Paolo Di Lorenzo, Sapienza University of Rome, Italy

SS-9: PROCESSING HIGHER ORDER NETWORK DATA

SS-9.1: HYPERGRAPHS WITH EDGE-DEPENDENT VERTEX WEIGHTS: SPECTRAL CLUSTERING BASED ON THE 1-LAPLACIAN **8837**
Yu Zhu, Boning Li, Santiago Segarra, Rice University, United States of America

SS-9.2: CAUSAL LINEAR TOPOLOGICAL FILTERS OVER A 2-SIMPLEX **8842**
Georg Essl, University of Wisconsin - Milwaukee, United States of America

SS-9.3: SIMPLICIAL CONVOLUTIONAL NEURAL NETWORKS **8847**
Maosheng Yang, Elvin Isufi, Geert Leus, Delft University of Technology, Netherlands

SS-9.4: SIGNAL PROCESSING ON CELL COMPLEXES **8852**
T. Mitchell Roddenberry, Rice University, United States of America; Michael T. Schaub, RWTH Aachen University, Germany; Mustafa Hajij, Santa Clara University, United States of America

SS-9.5: ROBUST SIGNAL PROCESSING OVER SIMPLICIAL COMPLEXES **8857**
Stefania Sardellitti, Sergio Barbarossa, Sapienza University of Rome, Italy

SS-10: SIGNAL PROCESSING AND NEURAL APPROACHES FOR SOUNDSCAPES (SINAPS)

SS-10.1: CONFORMER-BASED SELF-SUPERVISED LEARNING FOR NON-SPEECH AUDIO TASKS **8862**
Sangeeta Srivastava, The Ohio State University, United States of America; Yun Wang, Andros Tjandra, Anurag Kumar, Chunxi Liu, Kritika Singh, Yatharth Saraf, Meta, United States of America

SS-10.2: UNSUPERVISED AUDIO-CAPTION ALIGNING LEARNS CORRESPONDENCES BETWEEN INDIVIDUAL SOUND EVENTS AND TEXTUAL PHRASES **8867**
Huang Xie, Okko Räsänen, Konstantinos Drossos, Tuomas Virtanen, Tampere University, Finland

SS-10.3: SPATIAL DATA AUGMENTATION WITH SIMULATED ROOM IMPULSE RESPONSES FOR SOUND EVENT LOCALIZATION AND DETECTION **8872**
Yuichiro Koyama, Masafumi Takahashi, Kazuki Shimada, Naoya Takahashi, Emiru Tsunoo, Shusuke Takahashi, Yuki Mitsufuji, Sony Group Corporation, Japan; Kazuhide Shigemi, The University of Tokyo, Japan

SS-10.4: POLYPHONIC AUDIO EVENT DETECTION: MULTI-LABEL OR MULTI-CLASS MULTI-TASK CLASSIFICATION PROBLEM? **8877**
Huy Phan, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland; Thi Ngoc Tho Nguyen, Nanyang Technological University, Singapore; Philipp Koch, Alfred Mertins, University of Lübeck, Germany

SS-10.5: DIVERSE AUDIO CAPTIONING VIA ADVERSARIAL TRAINING **8882**
Xinhao Mei, Xubo Liu, Jianyuan Sun, Mark Plumbley, Wenwu Wang, University of Surrey, United Kingdom of Great Britain and Northern Ireland

SS-10.6: PROBABLY PLEASANT? A NEURAL-PROBABILISTIC APPROACH TO AUTOMATIC MASKER SELECTION FOR URBAN SOUNDSCAPE AUGMENTATION **8887**
Kenneth Ooi, Karn N. Watcharasupat, Bhan Lam, Zhen-Ting Ong, Woon-Seng Gan, Nanyang Technological University, Singapore

SS-11: GRAPH-BASED DEEP LEARNING FOR COMMUNICATIONS

SS-11.1: USER SCHEDULING USING GRAPH NEURAL NETWORKS FOR RECONFIGURABLE INTELLIGENT SURFACE ASSISTED MULTIUSER DOWNLINK COMMUNICATIONS 8892
Zhongze Zhang, Tao Jiang, Wei Yu, University of Toronto (St. George), Canada

SS-11.2: SYMBOL-LEVEL ONLINE CHANNEL TRACKING FOR DEEP RECEIVERS 8897
Ron Aharon Finish, Yoav Cohen, Tomer Raviv, Nir Shlezinger, Ben-Gurion University of the Negev, Israel

SS-11.3: DELAY-ORIENTED DISTRIBUTED SCHEDULING USING GRAPH NEURAL NETWORKS 8902
Zhongyuan Zhao, Santiago Segarra, Rice University, United States of America; Gunjan Verma, Ananthram Swami, US Army's DEVCOM Army Research Laboratory, United States of America

SS-11.4: FLOWDT: A FLOW-AWARE DIGITAL TWIN FOR COMPUTER NETWORKS 8907
Miquel Ferriol-Galmés, Pere Barlet-Ros, Albert Cabellos-Aparicio, Universitat Politècnica de Catalunya, Spain; Xiangle Cheng, University of Exeter, United Kingdom of Great Britain and Northern Ireland; Xiang Shi, Shihan Xiao, Huawei Technologies, China

SS-11.5: STABLE AND TRANSFERABLE WIRELESS RESOURCE ALLOCATION POLICIES VIA MANIFOLD NEURAL NETWORKS 8912
Zhiyang Wang, Luana Ruiz, Alejandro Ribeiro, University of Pennsylvania, United States of America; Mark Eisen, Intel Labs, United States of America

SS-12: TOWARDS NEUROMORPHIC MACHINE INTELLIGENCE: SPIKE-BASED REPRESENTATION, LEARNING, AND SIGNAL PROCESSING

SS-12.1: MOTIF-TOPOLOGY AND REWARD-LEARNING IMPROVED SPIKING NEURAL NETWORK FOR EFFICIENT MULTI-SENSORY INTEGRATION 8917
Shuncheng Jia, Tielin Zhang, Hongxing Liu, Bo Xu, Institute of Automation, Chinese Academy of Sciences, China; Ruichen Zuo, School of Information and Electronics, Beijing Institute of Technology, China

SS-12.2: EVENT-BASED MULTIMODAL SPIKING NEURAL NETWORK WITH ATTENTION MECHANISM 8922
Qianhui Liu, Dong Xing, Lang Feng, Huajin Tang, Gang Pan, Zhejiang University, China

SS-12.3: GRADUAL SURROGATE GRADIENT LEARNING IN DEEP SPIKING NEURAL NETWORKS 8927
Yi Chen, Silin Zhang, Shiyu Ren, Hong Qu, University of Electronic Science and Technology of China, China

SS-12.4: AXONAL DELAY AS A SHORT-TERM MEMORY FOR FEED FORWARD DEEP SPIKING NEURAL NETWORKS 8932
*Pengfei Sun, Dick Botteldooren, Ghent University, Belgium; Longwei Zhu, Institute for Infocomm Research, A*STAR, Singapore*

SS-12.5: LOW PRECISION LOCAL LEARNING FOR HARDWARE-FRIENDLY NEUROMORPHIC VISUAL RECOGNITION 8937
*Jyotibdha Acharya, Laxmi R Iyer, Wenyu Jiang, Institute of Infocomm Research, A*STAR, Singapore*

SS-12.6: A HYBRID LEARNING FRAMEWORK FOR DEEP SPIKING NEURAL NETWORKS WITH ONE-SPIKE TEMPORAL CODING 8942
Jiadong Wang, Jibin Wu, Malu Zhang, Qi Liu, Haizhou Li, National University of Singapore, Singapore

SS-13: SEMANTIC AND INTERPRETABLE ANALYSIS IN MULTIMEDIA FORENSICS

SS-13.1: A-PIXELHOP: A GREEN, ROBUST AND EXPLAINABLE FAKE-IMAGE DETECTOR 8947
Yao Zhu, Xinyu Wang, Hong-Shuo Chen, C.-C. Jay Kuo, University of Southern California, United States of America; Ronald Salloum, California State University, San Bernardino, United States of America

SS-13.2: EXPLAINABLE FACT-CHECKING THROUGH QUESTION ANSWERING 8952
Jing Yang, Didier Vega-Oliveros, Anderson Rocha, Recod.ai, University of Campinas, Brazil; Taís Seibt, Universidade do Vale do Rio dos Sinos, Brazil

SS-13.3: DEEP VIDEO INPAINTING LOCALIZATION USING SPATIAL AND TEMPORAL TRACES 8957
Shujin Wei, Haodong Li, Jiwu Huang, Shenzhen University, China

SS-13.4: DEEPFAKE SPEECH DETECTION THROUGH EMOTION RECOGNITION: A SEMANTIC APPROACH 8962
Emanuele Conti, Davide Salvi, Clara Borrelli, Paolo Bestagini, Fabio Antonacci, Augusto Sarti, Stefano Tubaro, Politecnico di Milano, Italy; Brian Hosler, Matthew C. Stamm, Drexel University, United States of America

SS-13.5: TEXT-IMAGE DE-CONTEXTUALIZATION DETECTION USING VISION-LANGUAGE MODELS 8967
Mingzhen Huang, Shan Jia, Siwei Lyu, University at Buffalo, United States of America; Ming-Ching Chang, University at Albany, United States of America

SS-13.6: CUSTOM ATTRIBUTION LOSS FOR IMPROVING GENERALIZATION AND INTERPRETABILITY OF DEEPFAKE DETECTION 8972
Pavel Korshunov, Anubhav Jain, Sebastien Marcel, Idiap Research Institute, Switzerland

SS-14: GROUP THEORY FOR SENSING, IMAGING, AND LEARNING APPLICATIONS

SS-14.1: SPARSE MULTI-REFERENCE ALIGNMENT: SAMPLE COMPLEXITY AND COMPUTATIONAL HARDNESS 8977
Tamir Bendory, Tel Aviv University, Israel; Oscar Mickelin, Amit Singer, Princeton University, United States of America

SS-14.2: GRASSMANNIAN DIMENSIONALITY REDUCTION USING TRIPLET MARGIN LOSS FOR UME CLASSIFICATION OF 3D POINT CLOUDS 8982
Yuval Haitman, Joseph Francos, Ben-Gurion University, Israel; Louis Scharf, Colorado State University, United States of America

SS-14.3: A NOTE ON TOTALLY SYMMETRIC EQUI-ISOCLINIC TIGHT FUSION FRAMES 8987
Matthew Fickus, Air Force Institute of Technology, United States of America; Joseph Iverson, Iowa State University, United States of America; John Jasper, South Dakota State University, United States of America; Dustin Mixon, The Ohio State University, United States of America

SS-14.4: A SIMPLE FORMULA FOR THE MOMENTS OF UNITARILY INVARIANT MATRIX DISTRIBUTIONS 8992
Stephen D. Howard, Defence Science and Technology Group, Australia; Ali Pezeshki, Colorado State University, United States of America

SS-15: COMPUTER AUDITION FOR HEALTHCARE (CA4H)

SS-15.1: FUSION OF MODULATION SPECTRAL AND SPECTRAL FEATURES WITH SYMPTOM METADATA FOR IMPROVED SPEECH-BASED COVID-19 DETECTION 8997
Yi Zhu, Tiago Falk, Institut national de la recherche scientifique, Canada

SS-15.2: AN OVERVIEW OF THE FIRST ICASSP SPECIAL SESSION ON COMPUTER AUDITION FOR HEALTHCARE 9002
Kun Qian, Beijing Institute of Technology, China; Tanja Schultz, University of Bremen, Germany; Björn Schuller, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SS-15.3: A GLANCE-AND-GAZE NETWORK FOR RESPIRATORY SOUND CLASSIFICATION 9007
Shuai Yu, Yiwei Ding, Wei Li, Fudan University, China; Kun Qian, Bin Hu, Beijing Institute of Technology, China; Bjoern Schuller, Imperial College London, United Kingdom of Great Britain and Northern Ireland

SS-15.4: INTERNET STREAMING AUDIO BASED SPEECH RECEPTION THRESHOLD MEASUREMENT IN COCHLEAR IMPLANT USERS 9012
Xi Chen, Kang Ouyang, Mingyue Shi, Huali Zhou, Nengheng Zheng, Shenzhen University, China; Yefei Mo, Qinglin Meng, South China University of Technology, China; Yupeng Shi, Wei Xiao, Shidong Shang, Tencent, China

SS-15.5: A DOMAIN TRANSFER BASED DATA AUGMENTATION METHOD FOR AUTOMATED RESPIRATORY CLASSIFICATION 9017
Zijie Wang, Zhao Wang, Zhejiang University, China

SS-16: SECURITY- AND PRIVACY-PRESERVING SIGNAL PROCESSING TECHNIQUES FOR EMERGING APPLICATIONS

SS-16.1: PHYSICAL LAYER ANONYMOUS COMMUNICATIONS: AN ANONYMITY ENTROPY ORIENTED PRECODING DESIGN 9022
Zhongxiang Wei, Tongji University, China; Christos Masouros, University College London, United Kingdom of Great Britain and Northern Ireland; Sumei Sun, Institute for Infocomm Research (I2R), Singapore

SS-16.2: FEDERATED STOCHASTIC GRADIENT DESCENT BEGETS SELF-INDUCED MOMENTUM 9027
Howard Yang, Zuozhu Liu, Zhejiang University, China; Yaru Fu, The Open University of Hong Kong, China; Tony Quek, Singapore University of Technology and Design, Singapore; Vincent Poor, Princeton University, United States of America

SS-16.3: ADVERSARIAL LEARNING IN TRANSFORMER BASED NEURAL NETWORK IN RADIO SIGNAL CLASSIFICATION 9032
Lu Zhang, Sangarapillai Lambotharan, Gan Zheng, Loughborough University, United Kingdom of Great Britain and Northern Ireland

SS-16.4: OPTM3SEC: OPTIMIZING MULTICAST IRS-AIDED MULTIANTENNA DFRC SECRECY CHANNEL WITH MULTIPLE EAVESDROPPERS 9037
Kumar Vijay Mishra, United States Army Research Laboratory, United States of America; Arpan Chattopadhyay, Siddharth Sankar Acharjee, Indian Institute of Technology, India; Athina Petropulu, Rutgers - The State University of New Jersey, United States of America

SS-16.5: PRIVACY-ENHANCING APPLIANCE FILTERING FOR SMART METERS 9042
Ramana Reddy Avula, Tobias Oechtering, KTH Royal Institute of Technology, Sweden

SS-16.6: ADVERSARIAL LINEAR QUADRATIC REGULATOR UNDER FALSIFIED ACTIONS 9047
Chenglong Sun, Zuxing Li, Chao Wang, Tongji University, China

SS-17: PROVABLE TENSOR AND MATRIX METHODS FOR SENSING AND LEARNING

SS-17.1: COMMUNICATION-EFFICIENT DISTRIBUTED MAX-VAR GENERALIZED CCA VIA ERROR FEEDBACK-ASSISTED QUANTIZATION 9052
Sagar Shrestha, Xiao Fu, Oregon State University, United States of America

SS-17.2: PROVABLE SECOND-ORDER RIEMANNIAN GAUSS-NEWTON METHOD FOR LOW-RANK TENSOR ESTIMATION 9057
Yuetian Luo, University of Wisconsin-Madison, United States of America; Qin Ma, Ohio State University, United States of America; Chi Zhang, Indiana University, United States of America; Anru Zhang, Duke University, United States of America

SS-17.3: BOUNDED SIMPLEX-STRUCTURED MATRIX FACTORIZATION 9062
Olivier Vu Thanh, Nicolas Gillis, Fabian Lecron, Université de Mons, Belgium

SS-17.4: CPD COMPUTATION VIA RECURSIVE EIGENSPACE DECOMPOSITIONS 9067
Eric Evert, Michiel Vandecappelle, Lieven De Lathauwer, KU Leuven, Belgium

SS-17.5: ACCELERATING ILL-CONDITIONED ROBUST LOW-RANK TENSOR REGRESSION 9072
Tian Tong, Yuejie Chi, Carnegie Mellon University, United States of America; Cong Ma, University of Chicago, United States of America

SS-17.6: ADA-JSR: SAMPLE EFFICIENT ADAPTIVE JOINT SUPPORT RECOVERY FROM EXTREMELY COMPRESSED MEASUREMENT VECTORS 9077
Sina Shahsavari, Pulak Sarangi, Mehmet Hucumenoglu, Piya Pal, University of California San Diego, United States of America

SS-18: SIGNAL PROCESSING AND MACHINE LEARNING FOR RESPIRATORY DISEASE ANALYSIS

SS-18.1: END-TO-END NETWORK BASED ON TRANSFORMER FOR AUTOMATIC DETECTION OF COVID-19 9082
Cong Cai, Bin Liu, Jianhua Tao, Zhengkun Tian, Kexin Wang, Institute of Automation, Chinese Academy of Sciences, China; Jiahao Lu, Tianjin Normal University, China

SS-18.2: PROTOTYPE LEARNING FOR INTERPRETABLE RESPIRATORY SOUND ANALYSIS 9087
Zhao Ren, Thanh Tam Nguyen, Wolfgang Nejdl, Leibniz Universität Hannover, Germany

SS-18.3: CONVOLUATIONAL TRANSFORMER WITH ADAPTIVE POSITION EMBEDDING FOR COVID-19 DETECTION FROM COUGH SOUNDS 9092
Tianhao Yan, Harbin Engineering University and University of Augsburg, China; Hao Meng, Harbin Engineering University, China; Shuo Liu, University of Augsburg, Germany; Emilia Parada-Cabaleiro, Johannes Kepler University Linz, Austria; Zhao Ren, Leibniz Universit¨at Hannover, Germany; Björn W. Schuller, University of Augsburg and Imperial College London, Germany

SS-18.4: DETECTION OF COPD EXACERBATION FROM SPEECH: COMPARISON OF ACOUSTIC FEATURES AND DEEP LEARNING BASED SPEECH BREATHING MODELS 9097
Venkata Srikanth Nallanthighal, Philips Research Eindhoven and Radboud University, Nijmegen, Netherlands; Aki Härmä, Philips Research Eindhoven, Netherlands; Helmer Strik, Radboud University, Nijmegen, Netherlands

SS-18.5: AUTOMATIC RESPIRATORY SOUND CLASSIFICATION VIA MULTI-BRANCH TEMPORAL CONVOLUTIONAL NETWORK 9102
Ziping Zhao, Zhen Gong, Mingyue Niu, Jiali Ma, Tianjin Normal University, China; Haishuai Wang, Fairfield University, United States of America; Zixing Zhang, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Ya Li, Beijing University of Posts and Telecommunications, China

CHAL-1: ACOUSTIC ECHO CANCELLATION

CHAL-1.1: ICASSP 2022 ACOUSTIC ECHO CANCELLATION CHALLENGE 9107
Ross Cutler, Ando Saabas, Tanel Parnamaa, Marju Purin, Hannes Gamper, Sebastian Braun, Karsten Sørensen, Robert Aichner, Microsoft, United States of America

CHAL-1.2: A DEEP HIERARCHICAL FUSION NETWORK FOR FULLBAND ACOUSTIC ECHO CANCELLATION 9112
Haoran Zhao, Nan Li, Runqiang Han, Lianwu Chen, Xiguang Zheng, Chen Zhang, Liang Guo, Bing Yu, Kuaishou Technology, China

CHAL-1.3: EXPLORE RELATIVE AND CONTEXT INFORMATION WITH TRANSFORMER FOR JOINT ACOUSTIC ECHO CANCELLATION AND SPEECH ENHANCEMENT 9117
Xingwei Sun, Chenbin Cao, Qinglong Li, Linzhang Wang, Fei Xiang, Xiaomi Corporation, China

CHAL-1.4: MULTI-SCALE TEMPORAL FREQUENCY CONVOLUTIONAL NETWORK WITH AXIAL ATTENTION FOR SPEECH ENHANCEMENT 9122
Guochang Zhang, Libiao Yu, Chunliang Wang, Jianqiang Wei, Baidu, China

CHAL-1.5: MULTI-TASK DEEP RESIDUAL ECHO SUPPRESSION WITH ECHO-AWARE LOSS 9127
Shimin Zhang, Jiayao Sun, Yihui Fu, Lei Xie, Northwestern Polytechnical University, China; Ziteng Wang, Biao Tian, Qiang Fu, Personal, China

CHAL-1.6: MULTI-SCALE REFINEMENT NETWORK BASED ACOUSTIC ECHO CANCELLATION 9132
Fan Cui, Liyong Guo, Wenfeng Li, Peng Gao, Yujun Wang, Xiaomi Inc, China

CHAL-2: AUDIO-VISUAL OBJECT CLASSIFICATION FOR HUMAN-ROBOT COLLABORATION

CHAL-2.1: AUDIO-VISUAL OBJECT CLASSIFICATION FOR HUMAN-ROBOT COLLABORATION 9137
Alessio Xompero, Yik Lung Pang, Andrea Cavallaro, Queen Mary University of London, United Kingdom of Great Britain and Northern Ireland; Timothy Patten, University of Technology Sydney, Australia; Ahalya Prabhakar, École Polytechnique Fédérale de Lausanne, Switzerland; Berk Calli, Worcester Polytechnic Institute, United States of America

CHAL-2.2: SHARED TRANSFORMER ENCODER WITH MASK-BASED 3D MODEL ESTIMATION FOR CONTAINER MASS ESTIMATION 9142
Tomoya Matsubara, Seitaro Otsuki, Yuiga Wada, Haruka Matsuo, Takumi Komatsu, Yui Iioka, Komei Sugiura, Hideo Saito, Keio University, Japan

CHAL-2.3: IMPROVING GENERALIZATION OF DEEP NETWORKS FOR ESTIMATING PHYSICAL PROPERTIES OF CONTAINERS AND FILLINGS 9147
Hengyi Wang, University College London, United Kingdom of Great Britain and Northern Ireland; Chaoran Zhu, Changjae Oh, Queen Mary University of London, China; Ziyin Ma, University of Birmingham, United Kingdom of Great Britain and Northern Ireland

CHAL-2.4: CONTAINER LOCALISATION AND MASS ESTIMATION WITH AN RGB-D CAMERA 9152
Tommaso Apicella, Giulia Slavic, Edoardo Ragusa, Paolo Gastaldo, Lucio Marcenaro, University of Genoa, Italy

CHAL-3: MULTI-CHANNEL MULTI-PARTY MEETING TRANSCRIPTION

CHAL-3.1: SUMMARY ON THE ICASSP 2022 MULTI-CHANNEL MULTI-PARTY MEETING TRANSCRIPTION GRAND CHALLENGE 9156
*Fan Yu, Shiliang Zhang, Zhihao Du, Siqi Zheng, Weilong Huang, Zhijie Yan, Bin Ma, Alibaba, China; Pengcheng Guo, Yihui Fu, Lei Xie, AISHELL Foundation, China; Zheng-Hua Tan, Aalborg University, Denmark; DeLiang Wang, The Ohio State University, United States of America; Yanmin Qian, Shanghai Jiao Tong University,, China; Kong Aik Lee, Institute for Infocomm Research, A*STAR, Singapore; Xin Xu, Hui Bu, Beijing Shell Shell Technology Co., Ltd., China*

CHAL-3.2: THE CUHK-TENCENT SPEAKER DIARIZATION SYSTEM FOR THE ICASSP 2022 MULTI-CHANNEL MULTI-PARTY MEETING TRANSCRIPTION CHALLENGE 9161
Naijun Zheng, Xixin Wu, Lingwei Meng, Jiawen Kang, Helen Meng, The Chinese University of Hong Kong, China; Na Li, Chao Weng, Dan Su, Tencent AI Lab, China; Haibin Wu, National Taiwan University, China

CHAL-3.3: THE USTC-XIMALAYA SYSTEM FOR THE ICASSP 2022 MULTI-CHANNEL MULTI-PARTY MEETING TRANSCRIPTION (M2MET) CHALLENGE 9166
Maokui He, Xiaoqi Zhang, Yuxuan Wang, Shutong Niu, Jun Du, University of Science and Technology of China, China; Xiang Lv, Weilin Zhou, Jingjing Yin, Yuhang Cao, Heng Lu, Ximalaya Inc., China; Chin-Hui Lee, Georgia Institute of Technology, United States of America

CHAL-3.4: CROSS-CHANNEL ATTENTION-BASED TARGET SPEAKER VOICE ACTIVITY DETECTION: EXPERIMENTAL RESULTS FOR M2MET CHALLENGE 9171
Weiqing Wang, Duke University, United States of America; Xiaoyi Qin, Ming Li, Duke Kunshan University, China

CHAL-3.5: THE VOLCSPEECH SYSTEM FOR THE ICASSP 2022 MULTI-CHANNEL MULTI-PARTY MEETING TRANSCRIPTION CHALLENGE 9176
Chen Shen, Yi Liu, Wenzhi Fan, Bin Wang, Shixue Wen, Yao Tian, Jun Zhang, Jingsheng Yang, Zejun Ma, bytedance, China

CHAL-3.6: THE ROYALFLUSH SYSTEM OF SPEECH RECOGNITION FOR M2MET CHALLENGE 9181
Shuaishuai Ye, Peiyao Wang, Shunfei Chen, Xinhui Hu, Xinkang Xu, Hithink RoyalFlush Information Network Co.,Ltd., China

CHAL-4: L3DAS22 MACHINE LEARNING FOR 3D AUDIO SIGNAL PROCESSING

CHAL-4.1: L3DAS22 CHALLENGE: LEARNING 3D AUDIO SOURCES IN A REAL OFFICE ENVIRONMENT 9186
Eric Guizzo, Christian Marinoni, Marco Pennese, Aurelio Uncini, Danilo Comminiello, Sapienza University of Rome, Italy; Xinlei Ren, Xiguang Zheng, Chen Zhang, Kuaishou Technology, China; Bruno Masiero, University of Campinas, Brazil

CHAL-4.2: ICASSP 2022 L3DAS22 CHALLENGE: ENSEMBLE OF RESNET-CONFORMERS WITH AMBISONICS DATA AUGMENTATION FOR SOUND EVENT LOCALIZATION AND DETECTION 9191
Yongjian Mao, Ying Zeng, Hongqing Liu, Wenbin Zhu, Yi Zhou, Chongqing University of Posts and Telecommunications, China

CHAL-4.3: A TRACK-WISE ENSEMBLE EVENT INDEPENDENT NETWORK FOR POLYPHONIC SOUND EVENT LOCALIZATION AND DETECTION 9196
Jinbo Hu, Ming Wu, Feiran Yang, Jun Yang, Institute of Acoustics, Chinese Academy of Sciences, China; Yin Cao, Mark Plumbley, University of Surrey, United Kingdom of Great Britain and Northern Ireland; Qiuqiang Kong, ByteDance, China

CHAL-4.4: TOWARDS LOW-DISTORTION MULTI-CHANNEL SPEECH ENHANCEMENT: THE ESPNET-SE SUBMISSION TO THE L3DAS22 CHALLENGE 9201
Yen-Ju Lu, Xuankai Chang, Zhong-Qiu Wang, Shinji Watanabe, Carnegie Mellon University, United States of America; Samuele Cornell, Università Politecnica delle Marche, Italy; Wangyou Zhang, Chenda Li, Shanghai Jiao Tong University, China; Zhaoheng Ni, Meta, United States of America

CHAL-4.5: MULTI-SCALE TEMPORAL FREQUENCY CONVOLUTIONAL NETWORK WITH AXIAL ATTENTION FOR MULTI-CHANNEL SPEECH ENHANCEMENT 9206
Guochang Zhang, Chunliang Wang, Libiao Yu, Jianqiang Wei, Baidu, China

CHAL-4.6: THE PCG-AIID SYSTEM FOR L3DAS22 CHALLENGE: MIMO AND MISO CONVOLUTIONAL RECURRENT NETWORK FOR MULTI CHANNEL SPEECH ENHANCEMENT AND SPEECH RECOGNITION 9211
Jingdong Li, Yuanyuan Zhu, Dawei Luo, Yun Liu, Guohui Cui, Zhaoxia Li, Tencent, China

CHAL-5: AUDIO DEEPFAKE DETECTION

CHAL-5.1: ADD 2022: THE FIRST AUDIO DEEP SYNTHESIS DETECTION CHALLENGE 9216
Jiangyan Yi, Ruibo Fu, Jianhua Tao, Shuai Nie, Haoxin Ma, Chenglong Wang, Tao Wang, Zhengkun Tian, Ye Bai, Cunhang Fan, Shan Liang, Shiming Wang, Shuai Zhang, Zhengqi Wen, Institute of Automation, Chinese Academy of Sciences, China; Xinrui Yan, Le Xu, University of Chinese Academy of Sciences, China; Haizhou Li, National University of Singapore, Singapore

CHAL-5.2: TIME DOMAIN ADVERSARIAL VOICE CONVERSION FOR ADD 2022 9221
Cheng Wen, Tingwei Guo, Xingjun Tan, Rui Yan, Shuran Zhou, Chuandong Xie, Wei Zou, Xiangang Li, Beike, China

CHAL-5.3: AUDIO DEEPFAKE DETECTION SYSTEM WITH NEURAL STITCHING FOR ADD 2022 9226
Rui Yan, Cheng Wen, Shuran Zhou, Tingwei Guo, Wei Zou, Xiangang Li, Beike, China

CHAL-5.4: FAKE AUDIO DETECTION BASED ON UNSUPERVISED PRETRAINING MODELS 9231
Zhiqiang Lv, Shanshan Zhang, Kai Tang, Pengfei Hu, Tencent, China

CHAL-5.5: PARTIALLY FAKE AUDIO DETECTION BY SELF-ATTENTION-BASED FAKE SPAN DISCOVERY 9236
Haibin Wu, Hung-yi Lee, National Taiwan University, Taiwan; Heng-Cheng Kuo, Kuo-Hsuan Hung, Yu Tsao, Research Center for Information Technology Innovation, Academia Sinica, Taiwan, Taiwan; Naijun Zheng, Helen Meng, The Chinese University of Hong Kong, Hong Kong; Hsin-Min Wang, Institute of Information Science, Academia Sinica, Taiwan, Taiwan

CHAL-5.6: THE VICOMTECH AUDIO DEEPFAKE DETECTION SYSTEM BASED ON WAV2VEC2 FOR THE 2022 ADD CHALLENGE 9241
Juan M. Martín-Doñas, Aitor Álvarez, Vicomtech Foundation, Spain

CHAL-6: MULTIMODAL INFORMATION BASED SPEECH PROCESSING

CHAL-6.1: AUDIO-VISUAL WAKE WORD SPOTTING SYSTEM FOR MISP CHALLENGE 2021 9246
Yanguang Xu, Jianwei Sun, Yang Han, Shuaijiang Zhao, Chaoyang Mei, Tingwei Guo, Shuran Zhou, Chuandong Xie, Wei Zou, Xiangang Li, Beike, China

CHAL-6.2: CHANNEL-WISE AV-FUSION ATTENTION FOR MULTI-CHANNEL AUDIO-VISUAL SPEECH RECOGNITION 9251
Gaopeng Xu, Song Yang, Wei Li, Sang Wang, Wei Guo, Junfeng Yuan, Jie Gao, NIO Co., Ltd., China

CHAL-6.3: THE DKU AUDIO-VISUAL WAKE WORD SPOTTING SYSTEM FOR THE 2021 MISP CHALLENGE 9256
Ming Cheng, Haoxu Wang, Ming Li, Wuhan University, China; Yechen Wang, Duke Kunshan University, China

CHAL-6.4: THE SJTU SYSTEM FOR MULTIMODAL INFORMATION BASED SPEECH PROCESSING CHALLENGE 2021 9261
Wei Wang, Xun Gong, Yifei Wu, Zhikai Zhou, Chenda Li, Wangyou Zhang, Bing Han, Yanmin Qian, Shanghai Jiao Tong University, China

CHAL-6.5: THE FIRST MULTIMODAL INFORMATION BASED SPEECH PROCESSING (MISP) CHALLENGE: DATA, TASKS, BASELINES AND RESULTS 9266
Hang Chen, Hengshun Zhou, Jun Du, University of Science and Technology of China, China; Chin-Hui Lee, Georgia Institute of Technology, United States of America; Jingdong Chen, Northwestern Polytechnical University, China; Shinji Watanabe, Carnegie Mellon University, United States of America; Sabato Marco Siniscalchi, Kore University of Enna, Italy; Odette Scharenborg, Delft University of Technology, Netherlands; Di-Yuan Liu, Bao-Cai Yin, Jia Pan, Jian-Qing Gao, Cong Liu, iFlytek, China

CHAL-7: DEEP NOISE SUPPRESSION

CHAL-7.1: ICASSP 2022 DEEP NOISE SUPPRESSION CHALLENGE 9271
Harishchandra Dubey, Vishak Gopal, Ross Cutler, Ashkan Aazami, Sergiy Matusevych, Sebastian Braun, Sefik Emre Eskimez, Manthan Thakker, Takuya Yoshioka, Hannes Gamper, Robert Aichner, Microsoft Corporation, Redmond, WA, USA, United States of America

CHAL-7.2: FB-MSTCN: A FULL-BAND SINGLE-CHANNEL SPEECH ENHANCEMENT 9276
METHOD BASED ON MULTI-SCALE TEMPORAL CONVOLUTIONAL NETWORK
Zehua Zhang, Lu Zhang, Xuyi Zhuang, Yukun Qian, Heng Li, Mingjiang Wang, Harbin Institute of Technology(Shenzhen), China

CHAL-7.3: FRCRN: BOOSTING FEATURE REPRESENTATION USING FREQUENCY 9281
RECURRENCE FOR MONAURAL SPEECH ENHANCEMENT
Shengkui Zhao, Bin Ma, Alibaba Group, Singapore; Karn N. Watcharasupat, Woon-Seng Gan, Nanyang Technological University (NTU), Singapore

CHAL-7.4: HARMONIC GATED COMPENSATION NETWORK PLUS FOR ICASSP 2022 9286
DNS CHALLENGE
Tianrui Wang, Weibin Zhu, Beijing Jiaotong University, China; Yingying Gao, Yanan Chen, Junlan Feng, Shilei Zhang, China Mobile Research Institute, China

CHAL-7.5: TEA-PSE: TENCENT-ETHEREAL-AUDIO-LAB PERSONALIZED SPEECH 9291
ENHANCEMENT SYSTEM FOR ICASSP 2022 DNS CHALLENGE
Yukai Ju, Xiaopeng Yan, Yihui Fu, Shubo Lv, Luyao Cheng, Lei Xie, Northwestern Polytechnical University, China; Wei Rao, Yannan Wang, Shidong Shang, Tencent, China

CHAL-7.6: MULTI-STAGE AND MULTI-LOSS TRAINING FOR FULLBAND 9296
NON-PERSONALIZED AND PERSONALIZED SPEECH ENHANCEMENT
Lianwu Chen, Chenglin Xu, Xu Zhang, Xinlei Ren, Xiguang Zheng, Chen Zhang, Liang Guo, Bing Yu, Kuaishou Technology, China

CHAL-8: ROOT CAUSE ANALYSIS FOR WIRELESS NETWORK FAULTS LOCALIZATION

CHAL-8.1: ICASSP-SPGC 2022: ROOT CAUSE ANALYSIS FOR WIRELESS NETWORK 9301
FAULT LOCALIZATION
Tianjian Zhang, Qian Chen, Feng Yin, Zhi-Quan Luo, The Chinese University of Hong Kong, Shenzhen; Shenzhen Research Institute of Big Data, China; Yi Jiang, The Chinese University of Hong Kong, Shenzhen, China; Dandan Miao, Tao Quan, Huawei Technologies Co., Ltd, China; Qingjiang Shi, Shenzhen Research Institute of Big Data, China

CHAL-8.2: ACCURATE INFERENCE OF UNSEEN COMBINATIONS OF MULTIPLE 9306
ROOTCAUSES WITH CLASSIFIER ENSEMBLE
Xuan Zhang, Longxiang Xiong, Ningyuan Sun, Mingxia Wang, Hao Tang, Network Intelligent Laboratory, School of Electronics and Information Engineering, Beijing Jiaotong Unviersity, China; Yanxing Zhao, Beijing Baolande Software Corporation, China

CHAL-8.3: CAUSAL ALIGNMENT BASED FAULT ROOT CAUSES LOCALIZATION FOR 9311
WIRELESS NETWORK
Yuequn Liu, Wenhui Zhu, Jie Qiao, Zhiyi Huang, Yu Xiang, Xuanzhi Chen, Wei Chen, Ruichu Cai, Guangdong University of Technology, China

CHAL-8.4: NETRCA: AN EFFECTIVE NETWORK FAULT CAUSE LOCALIZATION 9316
ALGORITHM
Chaoli Zhang, Zhiqiang Zhou, Yingying Zhang, Linxiao Yang, Kai He, Qingsong Wen, Liang Sun, Alibaba Group, China