

2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW 2024)

**Seoul, South Korea
14-19 April 2024**



**IEEE Catalog Number: CFP24JZ8-POD
ISBN: 979-8-3503-7452-0**

**Copyright © 2024 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** 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:	CFP24JZ8-POD
ISBN (Print-On-Demand):	979-8-3503-7452-0
ISBN (Online):	979-8-3503-7451-3

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

GC-L1: AUDITORY EEG DECODING CHALLENGE 2024

GC-L1.1: ICASSP 2024 AUDITORY EEG DECODING CHALLENGE 127
Lies Bollens, KU Leuven, Belgium; Corentin Puffay, Bernd Accou, Jonas Vanthornhout, Tom Francart, Hugo Van hamme, KU Leuven, Belgium

GC-L1.2: DETECTING GAMMA-BAND RESPONSES TO THE SPEECH ENVELOPE 55
FOR THE ICASSP 2024 AUDITORY EEG DECODING SIGNAL PROCESSING GRAND CHALLENGE
Michael Thornton, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Jonas Auernheimer, Constantin Jehn, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany; Danilo Mandic, Imperial College London, United Kingdom of Great Britain and Northern Ireland; Tobias Reichenbach, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

GC-L1.3: SELF-SUPERVISED SPEECH REPRESENTATION AND CONTEXTUAL 111
TEXT EMBEDDING FOR MATCH-MISMATCH CLASSIFICATION WITH EEG RECORDING
Bo Wang, Xiran Xu, Zechen Zhang, Haolin Zhu, Yujie Yan, Xihong Wu, Jing Chen, Peking University, China

GC-L1.4: BMMSNET: BIDIRECTIONAL MAPPING AND MULTILEVEL SIMILARITY 117
COMPARISON FOR EEG-SPEECH MATCH-MISMATCH PROBLEM
Zelin Qiu, Jianjun Gu, Dingding Yao, Junfeng Li, Yonghong Yan, Institute of Acoustics, Chinese Academy of Sciences, China

GC-L1.5: CROSS-ATTENTION-GUIDED WAVENET FOR MEL SPECTROGRAM 7
RECONSTRUCTION IN THE ICASSP 2024 AUDITORY EEG CHALLENGE
Yuan Fang, Inner Mongolia University, China; Hao Li, Southern University of Science and Technology, China; Xueliang Zhang, Inner Mongolia University, China; Fei Chen, Southern University of Science and Technology, China; Guanglai Gao, Inner Mongolia University, China

GC-L1.6: CONVCONCATNET: A DEEP CONVOLUTIONAL NEURAL NETWORK TO 113
RECONSTRUCT MEL SPECTROGRAM FROM THE EEG
Xiran Xu, Bo Wang, Yujie Yan, Haolin Zhu, Zechen Zhang, Xihong Wu, Jing Chen, Peking University, China

GC-L2: THE 2ND E-PREVENTION CHALLENGE: PSYCHOTIC AND NON-PSYCHOTIC RELAPSE DETECTION USING WEARABLE-BASED DIGITAL PHENOTYPING

GC-L2.1: THE 2ND E-PREVENTION CHALLENGE: PSYCHOTIC AND 125
NON-PSYCHOTIC RELAPSE DETECTION USING WEARABLE-BASED DIGITAL PHENOTYPING
Panagiotis Filintisis, Niki Efthymiou, George Retsinas, Athanasia Zlatintsi, Christos Garoufis, Institute of Robotics, AthenaRC, Greece; T. Sounapoglou, Blockachain, Greece; P. Tsanakas, NTUA, Greece; N. Smyrnis, University General Hospital "ATTIKON", National & Kapodistrian University of Athens, Medical School, Greece; Petros Maragos, NTUA, Institute of Robotics, Athena RC, Greece

GC-L2.2: PATIENT-SPECIFIC MODELING OF DAILY ACTIVITY PATTERNS FOR 43
UNSUPERVISED DETECTION OF PSYCHOTIC AND NON-PSYCHOTIC RELAPSES
Alice Hein, Sven Gronauer, Klaus Diepold, Technical University of Munich, Germany

GC-L2.3: UNSUPERVISED RELAPSE DETECTION USING WEARABLE-BASED 47
DIGITAL PHENOTYPING FOR THE 2ND E-PREVENTION CHALLENGE
Jinting Wu, Mei Tu, Samsung Research China-Beijing, China

GC-L2.4: PERSONALISED ANOMALY DETECTORS AND PROTOTYPICAL REPRESENTATIONS FOR RELAPSE DETECTION FROM WEARABLE-BASED DIGITAL PHENOTYPING	103
<i>Adria Mallol-Ragolta, Anika Spiesberger, Andreas Triantafyllopoulos, Björn Schuller, Technical University of Munich, Germany</i>	
GC-L2.5: A SELF-SUPERVISED LEARNING APPROACH FOR DETECTING NON-PSYCHOTIC RELAPSES USING WEARABLE-BASED DIGITAL PHENOTYPING	97
<i>Panagiotis Kaliosis, Sofia Eleftheriou, Christos Nikou, Theodoros Giannakopoulos, National Centre for Scientific Research Demokritos, Greece</i>	
GC-L2.6: BUMBLEBEE YOUR WAY TO RECOVERY: TRANSFORMING THE APPROACH TO DETECTION OF MENTAL HEALTH RELAPSES	89
<i>Kamil Górzyski, Anna Płes, Ivan Ryzhankow, Bartłomiej Zych, Samsung R&D Institute Poland, Poland</i>	
GC-L3: IN-CAR MULTI-CHANNEL AUTOMATIC SPEECH RECOGNITION CHALLENGE (ICMC-ASR)	
GC-L3.1: ICMC-ASR: THE ICASSP 2024 IN-CAR MULTI-CHANNEL AUTOMATIC SPEECH RECOGNITION CHALLENGE	63
<i>He Wang, Pengcheng Guo, Yue Li, Ao Zhang, Jiayao Sun, Lei Xie, Northwestern Polytechnical University, China; Wei Chen, Pan Zhou, Li Auto, China; Hui Bu, Xin Xu, Beijing AI Shell Technology Co. Ltd, China; Binbin Zhang, WeNet Open Source Community, China; Zhuo Chen, ByteDance, China; Jian Wu, Microsoft Corporation, China; Longbiao Wang, Tianjin University, China; Eng Siong Chng, Nanyang Technological University, China; Sun Li, China Academy of Information and Communication Technology, China</i>	
GC-L3.2: THE USTC-NERC SLIP SYSTEMS FOR THE ICMC-ASR CHALLENGE	3
<i>Minghui Wu, Luzhen Xu, Jie Zhang, University of Science and Technology of China, China; Haitao Tang, Yanyan Yue, Ruizhi Liao, Jintao Zhao, Anhui iFLYTEK Universal Language Technology, China; Zhengzhe Zhang, Yichi Wang, Haoyin Yan, University of Science and Technology of China, China; Hongliang Yu, Tongle Ma, Jiachen Liu, Chongliang Wu, Yongchao Li, Anhui iFLYTEK Universal Language Technology, China; Yanyong Zhang, University of Science and Technology of China, China; Xin Fang, Anhui iFLYTEK Universal Language Technology, China; Yue Zhang, Huawei Technology, China</i>	
GC-L3.3: THE FOSA FER SYSTEM FOR THE ICASSP2024 IN-CAR MULTI-CHANNEL AUTOMATIC SPEECH RECOGNITION CHALLENGE	5
<i>Shangkun Huang, Yuxuan Du, Yankai Wang, Jing Deng, Rong Zheng, Beijing Fosafer Information Technology Co., Ltd., China</i>	
GC-L3.4: THE ROYALFLUSH AUTOMATIC SPEECH DIARIZATION AND RECOGNITION SYSTEM FOR IN-CAR MULTI-CHANNEL AUTOMATIC SPEECH RECOGNITION CHALLENGE	1
<i>Jingguang Tian, Shuaishuai Ye, Shunfei Chen, Yang Xiang, Zhaohui Yin, Xinhui Hu, Xinkang Xu, Hithink RoyalFlush AI Research Institute, Zhejiang, China, China</i>	
GC-L3.5: THE FAWAISPEECH SYSTEM FOR MULTI-CHANNEL SPEECH RECOGNITION IN ICMC-ASR CHALLENGE	19
<i>yujia sun, jinxin he, yi zhang, xiaoming liang, ziyang wang, zhen fu, bo chen, China FAW Group Co. Ltd., China</i>	
GC-L3.6: XIMALAYA ASDR SYSTEM FOR ICASSP 2024 IN-CAR MULTI-CHANNEL (ICMC) ASR CHALLENGE	29
<i>Xiang Lyu, Yuhang Cao, Pengpeng Zou, Weilin Zhou, Ximalaya Inc., Shanghai, China, China</i>	

GC-L4: LIMMITS'24: MULTI-SPEAKER, MULTI-LINGUAL INDIC TTS WITH VOICE CLONING

GC-L4.1: LIMMITS'24: MULTI-SPEAKER, MULTI-LINGUAL INDIC TTS WITH VOICE CLONING 61

Abhayjeet Singh, Amala Nagireddi, Deekshitha G, Jesuraja Bandekar, Roopa R, Sandhya Badiger, Sathvik Udupa, Prasanta Kumar Ghosh, Indian Institute of Science, India; Hema A Murthy, Indian Institute of Technology, Madras, India; Pranaw Kumar, Centre for Development of Advanced Computing, India; Keiichi Tokuda, Nagoya Institute of Technology, Japan, India; Mark Hasegawa-Johnson, University of Illinois, India; Philipp Olbrich, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), India

GC-L4.2: LEVERAGING EFFECTIVE LANGUAGE AND SPEAKER CONDITIONING IN INDIC TTS FOR LIMMITS 2024 CHALLENGE 67

Yejin Jeon, Youngjae Kim, Gary Geunbae Lee, POSTECH, Korea, Republic of

GC-L4.3: SINGLE-STAGE TTS WITH ADAPTED VOCODER AND CROSS-ATTENTION: TALTECH SYSTEMS FOR THE LIMMITS'24 CHALLENGE 121

Daniil Rõbnikov, Tanel Alumäe, Tallinn University of Technology, Estonia

GC-L4.4: SCALING NVIDIA'S MULTI-SPEAKER MULTI-LINGUAL TTS SYSTEMS WITH ZERO-SHOT TTS TO INDIC LANGUAGES 115

Akshit Arora, Rohan Badlani, Sungwon Kim, Rafael Valle, Bryan Catanzaro, NVIDIA, United States of America

GC-L4.5: THE THU-HCSI MULTI-SPEAKER MULTI-LINGUAL FEW-SHOT VOICE CLONING SYSTEM FOR LIMMITS'24 CHALLENGE 71

Yixuan Zhou, Shuoyi Zhou, Shun Lei, Zhiyong Wu, Tsinghua University, China; Menglin Wu, ByteDance, China

GC-L4.6: CROSS-LINGUAL TEXT-TO-SPEECH VIA HIERARCHICAL STYLE TRANSFER 25

Sang-Hoon Lee, Ha-Yeong Choi, Seong-Whan Lee, Korea University, Korea, Republic of

GC-L5: GRAND CHALLENGE ON HYPERSPECTRAL SKIN VISION

GC-L5.1: HYPERSPECTRAL SKIN VISION CHALLENGE: CAN YOUR CAMERA SEE BEYOND YOUR SKIN? 59

Pai Chet Ng, Singapore Institute of Technology, Singapore; Zhixiang Chi, University of Toronto, Canada; Malcolm Low, Singapore Institute of Technology, Singapore; Juwei Lu, Huawei Noah's Ark Lab, Singapore; Konstantinos Plataniotis, University of Toronto, Canada; Nikolaos Boulgouris, Brunel University, United Kingdom of Great Britain and Northern Ireland; Thirimachos Bourlai, University of Georgia, United States of America; Yong Man Ro, Korea Advanced Institute of Science and Technology, Korea, Republic of

GC-L5.2: HYSAT++: HYBRID SPECTRAL-WISE ATTENTION TRANSFORMER FOR SKIN SPECTRAL RECONSTRUCTION 87

Yuchen Wang, Hongyuan Wang, Jiang Xu, Beijing Institute of Technology, China; Chang Chen, Xue Hu, Fenglong Song, Huawei Noah's Ark Lab, China; Lizhi Wang, Beijing Institute of Technology, China

GC-L5.3: MST--: A MODIFICATION OF MST++ FOR NARROW DOMAIN HYPERSPECTRAL RECONSTRUCTION 73

Daniil Vladimirov, Daniil Reutsky, Pavel Pischev, Vsevolod Plokhotnyuk, Egor Ershov, Institute of Information Transmission Problems named after A.A. Kharkevich, RAS, Russian Federation

GC-L5.4: VISION TRANSFORMER MST++: EFFICIENT HYPERSPECTRAL SKIN RECONSTRUCTION 51

Ana Clara Caznok Silveira, Diedre S. do Carmo, Lucas H. Ueda, Denis G. Fantinato, Paula D. P. Costa, Leticia Rittner, Universidade Estadual de Campinas (UNICAMP), Brazil

GC-L5.5: RGBT2HS-NET: RECONSTRUCTING A HYPER-SPECTRAL VOLUME FROM AN RGB-T STACK VIA AN ATTENTION-POWERED MULTIREOLUTION FRAMEWORK	83
<i>Hao-Chiang Shao, Szu-Chi Wu, Yen-Liang Chuo, Jyun-Hao Lin, National Chung Hsing University, Taiwan; Yuan-Rong Liao, Tse-Yu Tseng, Fu Jen Catholic University, Taiwan</i>	
GC-L5.6: HYPERSPECTRAL RECONSTRUCTION OF SKIN THROUGH FUSION OF SCATTERING TRANSFORM FEATURES	105
<i>Wojciech Czaja, Jeremiah Emidih, Brandon Kolstoe, University of Maryland, United States of America; Richard Spencer, National Institutes of Health, National Institute on Aging, United States of America</i>	
GC-L6: THE RF SIGNAL SEPARATION CHALLENGE	
GC-L6.1: THE DATA-DRIVEN RADIO FREQUENCY SIGNAL SEPARATION CHALLENGE	53
<i>Tejas Jayashankar, Massachusetts Institute of Technology, United States of America; Benoy Kurien, MIT Lincoln Laboratories, United States of America; Alejandro Lancho, Universidad Carlos III de Madrid, Spain; Gary Lee, Institute for Infocomm Research, Singapore; Yury Polyanskiy, Amir Weiss, Gregory Wornell, Massachusetts Institute of Technology, United States of America</i>	
GC-L6.2: A NOVEL APPROACH TO WAVENET ARCHITECTURE FOR RF SIGNAL SEPARATION WITH LEARNABLE DILATION AND DATA AUGMENTATION	79
<i>Yu Tian, Ahmed Alhammedi, Technology Innovation Institute, United Arab Emirates; Abdullah Quran, Abubakar Sani Ali, Khalifa University, United Arab Emirates</i>	
GC-L6.3: SIGNAL SEPARATION IN RADIO SPECTRUM USING SELF-ATTENTION MECHANISM	99
<i>Fadli Damara, Technical University of Berlin, Germany; Zoran Utkovski, Slawomir Stanczak, Fraunhofer Heinrich Hertz Institute, Germany</i>	
GC-L6.4: IMPROVING DATA-DRIVEN RF SIGNAL SEPARATION WITH SOI-MATCHED AUTOENCODERS	45
<i>Lukas Henneke, Fraunhofer FKIE, Germany</i>	
GC-L6.5: DEMUCS FOR DATA-DRIVEN RF SIGNAL DENOISING	95
<i>Çağkan Yapar, Fabian Jaensch, Jan Christian Hauffen, TU Berlin, Germany; Francesco Pezone, Sapienza University of Rome, Italy; Peter Jung, TU Berlin, DLR, Germany; Saeid Khalili Dehkordi, Giuseppe Caire, TU Berlin, Germany</i>	
GC-L6.6: A U-NET ARCHITECTURE FOR TIME-FREQUENCY INTERFERENCE SIGNAL SEPARATION OF RF WAVEFORMS	91
<i>Mostafa Naseri, Jaron Fontaine, Ingrid Moerman, Eli De Poorter,, Adnan Shahid, IDLab, Department of Information Technology at Ghent University - imec, Belgium</i>	
GC-L7: MULTIMODAL INFORMATION BASED SPEECH PROCESSING (MISP) 2023 CHALLENGE	
GC-L7.1: SUMMARY ON THE MULTIMODAL INFORMATION-BASED SPEECH PROCESSING (MISP) 2023 CHALLENGE	123
<i>Hang Chen, Shilong Wu, Chenxi Wang, Jun Du, University of Science and Technology of China, China; Chin-hui Lee, Georgia Institute of Technology, United States of America; Sabato Marco Siniscalchi, Kore University of Enna, Italy; Shinji Watanabe, Carnegie Mellon University, United States of America; Jingdong Chen, Northwestern Polytechnical University, China; Odette Scharenborg, Delft University of Technology, Netherlands; Zhong-Qiu Wang, Carnegie Mellon University, United States of America; Baocai Yin, Jianqing Gao, iFlytek Research, China</i>	
GC-L7.2: SIR-PROGRESSIVE AUDIO-VISUAL TF-GRIDNET WITH ASR-AWARE SELECTOR FOR TARGET SPEAKER EXTRACTION IN MISP 2023 CHALLENGE	11
<i>Zhongshu Hou, Tianchi Sun, Nanjing University, China; Yuxiang Hu, Changbao Zhu, Horizon Robotics, China; Kai Chen, Jing Lu, Nanjing University, China</i>	

GC-L7.3: AN AUDIO-QUALITY-BASED MULTI-STRATEGY APPROACH FOR TARGET SPEAKER EXTRACTION IN THE MISP 2023 CHALLENGE	27
<i>Runuduo Han, Xiaopeng Yan, Weiming Xu, Pengcheng Guo, Jiayao Sun, He Wang, Northwestern Polytechnical University, China; Quan Lu, Ning Jiang, Mashang Consumer Finance Co., Ltd., China; Lei Xie, Northwestern Polytechnical University, China</i>	
GC-L7.4: THE XMUSPEECH SYSTEM FOR AUDIO-VISUAL TARGET SPEAKER EXTRACTION IN MISP 2023 CHALLENGE	39
<i>Longjie Luo, Tao Li, Lin Li, Qingyang Hong, Xiamen University, China</i>	
 GC-L8: ADVANCING THE FRONTIERS OF DEEP LEARNING FOR LOW-DOSE 3D CONE-BEAM CT RECONSTRUCTION	
GC-L8.1: ADVANCING THE FRONTIERS OF DEEP LEARNING FOR LOW-DOSE 3D CONE-BEAM COMPUTED TOMOGRAPHY (CT) RECONSTRUCTION	81
<i>Ander Biguri, University of Cambridge, United Kingdom of Great Britain and Northern Ireland; Subhadip Mukherjee, IIT Kharagpur, India</i>	
GC-L8.2: DUAL-DOMAIN NEURAL NETWORKS FOR CLINICAL AND LOW-DOSE CBCT RECONSTRUCTION	17
<i>Xuzhi Zhao, Xi Liu, Xinyi Wang, Rui Yang, Beijing Jiaotong University, China; Yi Du, Peking University Cancer Hospital & Institute, China; Yahui Peng, Beijing Jiaotong University, China</i>	
GC-L8.3: MONAI FOR DEEP-LEARNING BASED CBCT RECONSTRUCTION	75
<i>Mikael Brudfors, NVIDIA, United Kingdom of Great Britain and Northern Ireland; Mark Graham, King's College London, United Kingdom of Great Britain and Northern Ireland; Hyungon Ryu, Oliver Kutter, NVIDIA, Korea, Republic of</i>	
GC-L8.4: A MULTI-FILTER AND MULTI-SCALE U-NET FOR CONE-BEAM COMPUTED TOMOGRAPHY WITH HARDWARE CONSTRAINTS	69
<i>Andreas Hauptmann, University of Oulu; University College London, Finland; Mustafa Al-Rubaye, University of Oulu, Finland; Miika T. Nieminen, Mikael A.K. Brix, University of Oulu; Oulu University Hospital, Finland</i>	
GC-L8.5: LOW DOSE CBCT DENOISING USING A 3D U-NET	85
<i>Austin Yunker, Argonne National Laboratory, United States of America; John Roeske, Loyola University Chicago, United States of America; Rajkumar Kettimuthu, Argonne National Laboratory, United States of America</i>	
GC-L8.6: 3D CBCT CHALLENGE 2024: IMPROVED CONE BEAM CT RECONSTRUCTION USING SWINIR-BASED SINOGRAM AND IMAGE ENHANCEMENT	101
<i>Sasidhar Alavala, Subrahmanyam Gorthi, Indian Institute of Technology Tirupati, India</i>	
 GC-L9: ICASSP SP CADENZA CHALLENGE: MUSIC DEMIXING/REMIXING FOR HEARING AIDS	
GC-L9.1: THE ICASSP SP CADENZA CHALLENGE: MUSIC DEMIXING/REMIXING FOR HEARING AIDS	93
<i>Gerardo Roa Dabike, University of Salford, United Kingdom of Great Britain and Northern Ireland; Michael Akeroyd, University of Nottingham, United Kingdom of Great Britain and Northern Ireland; Scott Bannister, University of Leeds, United Kingdom of Great Britain and Northern Ireland; Jon Barker, University of Sheffield, United Kingdom of Great Britain and Northern Ireland; Trevor Cox, Bruno Fazenda, University of Salford, United Kingdom of Great Britain and Northern Ireland; Jennifer Firth, University of Nottingham, United Kingdom of Great Britain and Northern Ireland; Simone Graetzer, University of Salford, United Kingdom of Great Britain and Northern Ireland; Alinka Greasley, University of Leeds, United Kingdom of Great Britain and Northern Ireland; Rebecca Vos, University of Salford, United Kingdom of Great Britain and Northern Ireland; William Whitmer, University of Nottingham, United Kingdom of Great Britain and Northern Ireland</i>	

GC-L9.2: THE USTC SYSTEM FOR CADENZA 2024 CHALLENGE	57
<i>Hongbo Lan, Tianyou Cheng, Maokui He, Hang Chen, Jun Du, University of Science and Technology of China, China</i>	
GC-L9.3: OPTIMIZING MUSIC SOURCE SEPARATION IN COMPLEX AUDIO	13
ENVIRONMENTS THROUGH PROGRESSIVE SELF-KNOWLEDGE DISTILLATION	
<i>ChangHeon Han, SuHyun Lee, Hanyang University, Korea, Republic of</i>	
GC-L9.4: SUB-BAND AND FULL-BAND INTERACTIVE U-NET WITH DPRNN FOR	21
DEMIXING CROSS-TALK STEREO MUSIC	
<i>Han Yin, School of Marine Science and Technology, Northwestern Polytechnical University, China; Mou Wang, Institute of Acoustics, Chinese Academy of Sciences, China; Jisheng Bai, School of Marine Science and Technology, Northwestern Polytechnical University, China; Dongyuan Shi, Woon-Seng Gan, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore; Jianfeng Chen, School of Marine Science and Technology, Northwestern Polytechnical University, China</i>	
GC-L9.5: REMIXING MUSIC FOR HEARING AIDS USING ENSEMBLE OF	109
FINE-TUNED SOURCE SEPARATORS	
<i>Matthew Daly, Johns Hopkins University, United States of America</i>	
GC-L9.6: MUSIC ENHANCEMENT WITH DEEP FILTERS: A TECHNICAL REPORT	119
FOR THE ICASSP 2024 CADENZA CHALLENGE	
<i>Keren Shao, Ke Chen, Shlomo Dubnov, UC San Diego, United States of America</i>	
GC-L11: THE ICASSP 2024 AUDIO DEEP PACKET LOSS CONCEALMENT GRAND CHALLENGE	
GC-L11.1: THE ICASSP 2024 AUDIO DEEP PACKET LOSS CONCEALMENT GRAND	41
CHALLENGE	
<i>Lorenz Diener, Solomiya Branets, Ando Saabas, Ross Cutler, Microsoft, Estonia</i>	
GC-L11.2: MULTI-STAGE TRAINING FOR CROSS-DOMAIN FULL-BAND AUDIO	35
PACKET LOSS CONCEALMENT	
<i>Nan Li, Guochen Yu, Chen Zhang, Chao Zhou, Qi Huang, Bing Yu, 中国科学院声学研究所, China</i>	
GC-L11.3: BS-PLCNET: BAND-SPLIT PACKET LOSS CONCEALMENT NETWORK	23
WITH MULTI-TASK LEARNING FRAMEWORK AND MULTI-DISCRIMINATORS	
<i>Zihan Zhang, Jiayao Sun, Northwestern Polytechnical University, China; Xianjun Xia, Chuanzeng Huang, Yijian Xiao, Bytedance, China; Lei Xie, Northwestern Polytechnical University, China</i>	
GC-L11.4: A TIME-FREQUENCY BAND-SPLIT NEURAL NETWORK FOR REAL-TIME	77
FULL-BAND PACKET LOSS CONCEALMENT	
<i>Lingling Dai, Yuxuan Ke, Huiyong Zhang, Fengyuan Hao, Xiaoxue Luo, Xiaodong Li, Chengshi Zheng, Institute of Acoustics, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China</i>	
GC-L11.5: A FULLBAND NEURAL NETWORK FOR AUDIO PACKET LOSS	107
CONCEALMENT	
<i>Bryce Irvin, Sile Yin, Shuo Zhang, Marko Stamenovic, Bose Corporation, United States of America</i>	
GC-L11.6: TWO-STAGE NEURAL NETWORK MODEL WITH PACKET LOSS	9
DETECTION FOR ICASSP 2024 PLC CHALLENGE	
<i>Xingwei Sun, Qinglong Li, Kaichi Ma, Linzhang Wang, Yujun Wang, Xiaomi Corporation, China</i>	
GC-L10: ICASSP 2024 SPEECH SIGNAL IMPROVEMENT CHALLENGE	
GC-L10.1: ICASSP 2024 SPEECH SIGNAL IMPROVEMENT CHALLENGE	15
<i>Nicolae Catalin Ristea, Ando Saabas, Ross Cutler, Babak Naderi, Sebastian Braun, Solomiya Branets, Microsoft, Romania</i>	

GC-L10.2: GENERAL SPEECH RESTORATION USING TWO-STAGE GENERATIVE ADVERSARIAL NETWORKS	31
<i>Qinwen Hu, Tianyi Tan, Nanjing University, China; Ming Tang, State Grid Jiangsu Electric Power Company Limited, China; Yuxiang Hu, Changbao Zhu, Horizon Robotics, China; Jing Lu, Nanjing University, China</i>	
GC-L10.3: RENET: A TIME-FREQUENCY DOMAIN GENERAL SPEECH RESTORATION NETWORK FOR ICASSP 2024 SPEECH SIGNAL IMPROVEMENT CHALLENGE	37
<i>Fengyuan Hao, Huiyong Zhang, Lingling Dai, Xiaoxue Luo, Xiaodong Li, Chengshi Zheng, Key Laboratory of Noise and Vibration Research, Institute of Acoustics, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China</i>	
GC-L10.4: RAD-NET: A REPAIRING AND DENOISING NETWORK FOR SPEECH SIGNAL IMPROVEMENT	49
<i>Mingshuai Liu, NWPU, China; Zhuangqi Chen, ByteDance, China; Xiaopeng Yan, Yuanju Lv, NWPU, China; Xianjun Xia, ByteDance, China; Chuanzeng Huang, Speech, Audio and Music Intelligence (SAMI) group, ByteDance, China; Yijian Xiao, ByteDance, China; Lei Xie, NWPU, China</i>	
GC-L10.5: KS-NET: MULTI-BAND JOINT SPEECH RESTORATION AND ENHANCEMENT NETWORK FOR 2024 ICASSP SSI CHALLENGE	33
<i>Guochen Yu, Runqiang Han, Chenglin Xu, Haoran Zhao, Nan Li, Chen Zhang, Xiguang Zheng, Chao Zhou, Qi Huang, Bing Yu, Kuaishou Technology, Beijing, China, China</i>	
GC-L10.6: REBUILD, REGENERATE: A GATED TEMPORAL CONVOLUTION BASED GAN FOR SPEECH SIGNAL IMPROVEMENT	65
<i>Nikhil Das, Rakesh Pogula, Mohammad Imtiaz Ali, Sasank Kottapalli, Sanniboyina Venkata Kiran, Independent Researcher, India</i>	
 WS-2: TRUSTWORTHY SPEECH PROCESSING (TSP)	
WS-2.1: LKANETS + + : ROBUST AUDIO CLASSIFICATION USING MULTI-LAYER NEURAL NETWORKS WITH LATERAL COMPETITION	129
<i>Sayanton V. Dibbo, Dartmouth College and Los Alamos National Laboratory, United States of America; Juston S. Moore, Garrett T. Kenyon, Michael A. Teti, Los Alamos National Laboratory, United States of America</i>	
WS-2.2: MULTI-MODAL APPROACHES FOR IMPROVING THE ROBUSTNESS OF AUDIO-BASED COVID-19 DETECTION SYSTEMS	144
<i>Drew Grant, Helena Hahn, Adebayo Eisape, Valerie Rennoll, James West, Johns Hopkins University, United States of America</i>	
WS-2.3: IMPROVING MEMBERSHIP INFERENCE IN ASR MODEL AUDITING WITH PERTURBED LOSS FEATURES	N/A
<i>Francisco Teixeira, INESC-ID/Técnico, University of Lisbon, Portugal; Karla Pizzi, Technical University Munich / Fraunhofer AISEC, Germany; Raphael Olivier, Vatic Labs, United States of America; Alberto Abad, INESC-ID/Técnico, University of Lisbon, Portugal; Bhiksha Raj, LTI, Carnegie Mellon University, United States of America; Isabel Trancoso, INESC-ID/Técnico, University of Lisbon, Portugal</i>	
WS-2.4: CHARACTERIZING THE TEMPORAL DYNAMICS OF UNIVERSAL SPEECH REPRESENTATIONS FOR GENERALIZABLE DEEPFAKE DETECTION	139
<i>Yi Zhu, Institut national de la recherche scientifique, Canada; Saurabh Powar, Veermata Jijabai Technological Institute, India; Tiago Falk, Institut national de la recherche scientifique, Canada</i>	
WS-2.5: LEVERAGING CONFIDENCE MODELS FOR IDENTIFYING CHALLENGING DATA SUBGROUPS IN SPEECH MODELS	134
<i>Alkis Koudounas, Eliana Pastor, Politecnico di Torino, Italy; Vittorio Mazzia, Manuel Giollo, Thomas Gueudre, Elisa Reale, Amazon Alexa AI, Italy; Giuseppe Attanasio, Bocconi University, Italy; Luca Cagliero, Sandro Cumani, Politecnico di Torino, Italy; Luca de Alfaro, University of California, Santa Cruz, United States of America; Elena Baralis, Politecnico di Torino, Italy; Daniele Amberti, Amazon Alexa AI, Italy</i>	

WS-1: DEEP NEURAL NETWORK MODEL COMPRESSION

WS-1.1: DEEP NEURAL NETWORK MODEL COMPRESSION AND SIGNAL PROCESSING 179

Arijit Ukil, TCS Research, India; Angshul Majumdar, Indraprastha Institute of Information Technology, Delhi, India, India; Antonio J. Jara, Libelium, Spain; Joao Gamma, University of Porto, Portugal

WS-1.2: A PRE-TRAINING PRUNING STRATEGY FOR ENABLING LIGHTWEIGHT NON-INTRUSIVE LOAD MONITORING ON EDGE DEVICES 249

Sotirios Athanasoulis, Stavros Sykiotis, Nikos Temenos, Anastasios Doulamis, Nikolaos Doulamis, National Technical University of Athens, Greece

WS-1.3: EDGE-OPTIMIZED MODEL FOR MULTIMEDIA CLASSIFICATION USING LINGUISTIC METADATA 269

Sunil Bharitkar, Samsung Research America, United States of America; Thaddeus Paez, Samsung Mexico, Mexico

WS-1.4: A VARIATIONAL BAYESIAN INFERENCE-INSPIRED UNROLLED DEEP NETWORK FOR COMPRESSED SENSING 304

Menghong Cai, Jun Fang, Huiping Duan, Xiaoyu Li, University of Electronic Science and Technology of China, China; Hongbin Li, Stevens Institute of Technology, United States of America

WS-1.5: STRUCTURED LOTTERY TICKET HYPOTHESIS FOR EFFECTIVE DEEP NEURAL NETWORK MODEL SIZE REDUCTION 330

Arijit Ukil, Ishan Sahu, Mridul Biswas, Arpan Pal, Tata Consultancy Services, India; Angshul Majumdar, Indraprastha Institute of Information Technology, India

WS-1.6: EMBEDDING COMPRESSION FOR TEACHER-TO-STUDENT KNOWLEDGE TRANSFER 403

Yiwei Ding, Alexander Lerch, Georgia Institute of Technology, United States of America

WS-1.7: DUAL KNOWLEDGE DISTILLATION FOR EFFICIENT SOUND EVENT DETECTION 690

Yang Xiao, Rohan Kumar Das, Fortemedia Singapore, Singapore

WS-5: COMPUTATIONAL IMAGING USING SYNTHETIC APERTURES

WS-5.1: PREDICTING THE SPECTRUM: DEEP ADAPTIVE SENSING FOR HADAMARD SINGLE PIXEL SPECTRAL IMAGING 154

Brayan Monroy, Jorge Bacca, Henry Arguello, Universidad Industrial de Santander, Colombia

WS-5.2: WIDEBAND ADAPTIVE BEAMFORMING FOR A PARTIALLY-CALIBRATED DISTRIBUTED ARRAY 164

Anil Ganti, Michael Martinez, Granger Hickman, Jeffrey Krolik, Duke University, United States of America

WS-5.3: NON-INVASIVE OCCUPANCY MONITORING USING SUSTAINABLE BACKSCATTER TAGS 259

Edmond Chan, Yirui Deng, Amus Goay, Deepak Mishra, Aruna Seneviratne, University of New South Wales, Australia

WS-5.4: IRS ASSISTED SECURE NOMA FOR UNTRUSTED USERS..... 356

Kaneez Khatoon, Indian Institute of Technology Jammu, India; Deepak Mishra, University of New South Wales, Australia, Australia; Rvaikant Saini, Indian Institute of Technology Jammu, India

WS-5.5: AN EFFECTIVE COMPOSITE JAMMING METHOD FOR SYNTHETIC APERTURE RADAR IN PRACTICAL ELECTRONIC COUNTERMEASURES 370

Xinrui Li, Baixiao Chen, Xidian University, China

WS-5.6: K-SPACE BEAMFORMING FOR AN ARRAY OF QUANTUM SENSORS 439

Peter Vouras, U.S. Department of Defense, United States of America

WS-6: TIMELY AND PRIVATE MACHINE LEARNING OVER NETWORKS

WS-6.1: CONVERGENCE ANALYSIS OF SEMI-FEDERATED LEARNING WITH NON-IID DATA 214

Wanli Ni, Tsinghua University, China; Jiachen Han, Beijing University of Posts and Telecommunications, China; Zhijin Qin, Tsinghua University, China

WS-6.2: RECOMMENDATION ALGORITHM BASED ON FEDERATED MULTI-MODAL LEARNING 229

Chenyuan Feng, Guangdong-Hong Kong Joint Laboratory for Big Data Imaging and Communication, China; Zhenyu Feng, Guangdong-Hong Kong Joint Laboratory for Big Data Imaging and Communication, Shenzhen University, China, China; Qing Wang, National Key Laboratory for Mobile Network and Mobile Multimedia Technology, ZTE Corporation, China, China

WS-6.3: A COMPRESSED DECENTRALIZED FEDERATED LEARNING FRAMEWORK FOR ENHANCED ENVIRONMENTAL AWARENESS IN V2V NETWORKS 374

Luca Barbieri, Mattia Brambilla, Monica Nicoli, Politecnico di Milano, Italy

WS-6.4: OVER THE AIR FEDERATED LEARNING IN THE PRESENCE OF IMPULSIVE NOISE 429

Zubair Shaban, Ranjitha Prasad, Pooja Kumari, IIIT Delhi, India

WS-6.5: FEDRF-ADAPT: ROBUST AND COMMUNICATION-EFFICIENT FEDERATED DOMAIN ADAPTATION VIA RANDOM FEATURES 615

Yuanjie Wang, Huazhong University of Science and Technology, China; Zhanbo Feng, Shanghai Jiao Tong University, China; Zhenyu Liao, Huazhong University of Science and Technology, China

WS-6.6: PERSOTA FL: A ROBUST-TO-NOISE PERSONALIZED OVER THE AIR FEDERATED LEARNING FOR HUMAN ACTIVITY RECOGNITION 620

Arash Rasti-Meymandi, Pai Chet Ng, University of Toronto, Canada; Huan Liu, Yuanhao Yu, Huawei Noah's Ark Lab, Canada; Konstantinos N. Plataniotis, University of Toronto, Canada

WS-6.7: TOWARDS COLLABORATIVE MULTI-MODAL FEDERATED LEARNING FOR HUMAN ACTIVITY RECOGNITION IN SMART WORKPLACE ENVIRONMENTS 635

Seyed Mohammad Sheikholeslami, Pai Chet Ng, University of Toronto, Canada; Huan Liu, Yuanhao Yu, Huawei Noah's Ark Lab, Canada, Canada; Konstantinos N. Plataniotis, University of Toronto, Canada

WS-6.8: VERSION AGE-BASED CLIENT SCHEDULING POLICY FOR FEDERATED LEARNING 695

Xinyi Hu, Zhejiang University, China; Nikolaos Pappas, Linköping University, Sweden; Howard H. Yang, Zhejiang University, China

WS-6.9: BYZANTINE-RESILIENT HIERARCHICAL FEDERATED LEARNING WITH CLUSTERED OVER-THE-AIR AGGREGATION 715

David Nordlund, Jialing Liao, Zheng Chen, Linköping University, Sweden

WS-9: SPID-CPS: SIGNAL PROCESSING FOR RESILIENT INTRUSION DETECTION IN CYBER-PHYSICAL SYSTEMS

WS-9.1: ROBUST FILTERING OF DISTRIBUTED CYBER-PHYSICAL SYSTEMS WITH CYBER-ATTACK DETECTION 294

Hongwei Wang, Xi Zheng, University of Electronic Science and Technology of China, China

WS-9.2: INTERPRETABILITY AND COMPLEXITY REDUCTION IN IOT NETWORK ANOMALY DETECTION VIA XAI 325

Alfredo Nascita, Raffaele Carillo, Federica Giampetraglia, Antonio Iacono, Valerio Persico, Antonio Pescapé, University of Napoli Federico II, Italy

WS-9.3: EMPOWERING NETWORK SECURITY WITH AUTOENCODERS.....	346
<i>Antonino Ferraro, Valerio La Gatta, Marco Postiglione, University of Naples Federico II, Italy</i>	
WS-9.4: ANOMALY DETECTION IN CYBER-PHYSICAL SYSTEMS: A CASE STUDY ON PUMP HEALTH MONITORING	361
<i>Giancarlo Sperli, Andrea Vignali, Università degli Studi di Napoli Federico II, Italy</i>	
WS-9.5: DEEP SUB-IMAGE SAMPLING BASED DEFENSE AGAINST SPATIAL-DOMAIN ADVERSARIAL STEGANOGRAPHY	383
<i>Xinyu Huang, Northwestern Polytechnical University, China; Yuwen Cao, Northwestern Polytechnical University, and Donghua University, China; Tomoaki Ohtsuki, Keio University, Japan</i>	
 WS-11: SIGNAL PROCESSING AND MACHINE LEARNING ADVANCES IN AUTOMOTIVE RADARS	
WS-11.1: A METHOD FOR ESTIMATING STATIONARY REFLECTIVE SURFACES AND EXTRACTING ROAD GUARDRAILS	149
<i>Yu Chen, Minglei Yang, Baixiao Chen, Meng Liu, Teng Ma, Xidian University, China</i>	
WS-11.2: JAMMING DEFENSE SCHEME FOR IRS-ASSISTED JRC SYSTEM IN VEHICULAR NETWORKS	254
<i>Yu Yao, Feng Shu, Hainan University, China; Linlong Wu, University of Luxembourg, Luxembourg; Xu Cheng, Sun Yatsen University, China; Xuan Li, East China Jiaotong University, China; Jiangzhou Wang, University of Kent, United Kingdom of Great Britain and Northern Ireland</i>	
WS-11.3: MODEL-BASED DIFFUSION FOR MITIGATING AUTOMOTIVE RADAR INTERFERENCE	284
<i>Jeroen Overvest, NXP Semiconductors, Netherlands; Xinyi Wei, Hans van Gorp, Ruud J.G. van Sloun, Eindhoven University of Technology, Netherlands</i>	
 WS-12: RADIO MAPS AND THEIR APPLICATIONS (RMA)	
WS-12.1: A TRANSFORMER-BASED NETWORK FOR UNIFYING RADIO MAP ESTIMATION AND OPTIMIZED SITE SELECTION	610
<i>Yi Zheng, Jiangnan University, China; CunYi Liao, Ji Wang, Shouyin Liu, Central China Normal University, China</i>	
WS-12.2: A NEW APPROACH TO PREDICT RADIO MAP VIA LEARNING-BASED SPATIAL LOSS FIELD	770
<i>Zhiqiang Tan, Zhiwei Yao, Tsinghua University, China; Limin Xiao, Ming Zhao, Yunzhou Li, Beijing National Research Center for Information Science and Technology, China</i>	
WS-12.3: MULTIAN TENNA CHANNEL MAP ESTIMATION USING DEEP SPATIAL INTERPOLATION	790
<i>Gyungmin Kim, Gwangju Institute of Science & Technology, Korea, Republic of; Rui Jin, Wilson Funk, Seung-Jun Kim, University of Maryland, Baltimore County, United States of America; Hyuk Lim, KENTECH, Korea, Republic of</i>	
 WS-13: SUPER-RESOLUTION INTEGRATED COMMUNICATIONS, LOCALIZATION, VISION AND RADIO MAPPING (SUPER-CLAM)	
WS-13.1: SMART CSI PROCESSING FOR ACCURATE COMMODITY WIFI-BASED HUMIDITY SENSING	279
<i>Yirui Deng, Deepak Mishra, Shaghik Atakaramians, Aruna Seneviratne, University of New South Wales, Australia</i>	
WS-13.2: COSMOS: COHERENT SCENE WITH MULTIPLE OBJECTS RECONSTRUCTION	660
<i>Byoungsung Lim, Korea University, Korea, Republic of; David Han, Drexel University, United States of America</i>	

WS-13.3: A RAY LAUNCHING APPROACH FOR COMPUTING EXACT PATHS WITH POINT CLOUDS	800
<i>Niklas Vaara, Pekka Sangi, Miguel Bordallo López, Janne Heikkilä, University of Oulu, Finland</i>	
WS-13.4: SUPER-RESOLUTION SENSING OF USER EQUIPMENT USING DELAY-DOPPLER PILOT-DATA STRUCTURE IN RIS-AIDED OTFS SYSTEMS	805
<i>OLIVIA ZACHARIA, Vani Devi M, Indian Institute of Space Science and Technology, India</i>	
WS-14: FEARLESS STEPS APOLLO: A NATURALISTIC TEAM BASED SPEECH COMMUNICATIONS COMMUNITY RESOURCE (FS-APOLLO)	
WS-14.1: FEARLESS STEPS APOLLO: OPERATIONAL DISCONNECT DETECTION IN MISSION CONTROL	810
<i>Douglas Oard, University of Maryland, United States of America; Christopher Bearman, CQ University, Australia; David Baker, Lobethal Apollo Association, Australia; Susannah Paletz, University of Maryland, United States of America; Johanne Trippas, RMIT University, Australia</i>	
WS-14.2: HOUSTON WE HAVE A DIVERGENCE: A SUBGROUP PERFORMANCE ANALYSIS OF ASR MODELS	812
<i>Alkis Koudounas, Flavio Giobergia, Politecnico di Torino, Italy</i>	
WS-3a: SELF-SUPERVISION IN AUDIO, SPEECH AND BEYOND I	
WS-3a.1: EXPLORING FEDERATED SELF-SUPERVISED LEARNING FOR GENERAL-PURPOSE AUDIO UNDERSTANDING	335
<i>Yasar Abbas UR Rehman, Kin Wai Lau, Yuyang Xie, Lan Ma, Jiajun Shen, TCL, Hong Kong</i>	
WS-3a.2: INVESTIGATING SELF-SUPERVISED FEATURES FOR EXPRESSIVE, MULTILINGUAL VOICE CONVERSION	341
<i>Álvaro Martín-Cortinas, Daniel Sáez-Trigueros, Grzegorz Beringer, Iván Vallés-Pérez, Roberto Barra-Chicote, Biel Tura-Vecino, Adam Gabryś, Thomas Merritt, Piotr Biliński, Jaime Lorenzo-Trueba, Alexa AI, Spain</i>	
WS-3a.3: ACOUSTIC-TO-ARTICULATORY INVERSION FOR DYSARTHIC SPEECH: ARE PRE-TRAINED SELF-SUPERVISED REPRESENTATIONS FAVORABLE?	408
<i>Sarthak Kumar Maharana, Krishna Kamal Adidam, Shoumik Nandi, Ajitesh Srivastava, University of Southern California, United States of America</i>	
WS-3a.4: OPEN IMPLEMENTATION AND STUDY OF BEST-RQ FOR SPEECH PROCESSING	460
<i>Ryan Whetten, Avignon Université, France; Titouan Parcollet, Samsung AI, United Kingdom of Great Britain and Northern Ireland; Marco Dinarelli, Univervisté Grenoble Alpes, France; Yannick Estève, Avignon Université, France</i>	
WS-3a.5: SPEECHCLIP + : SELF-SUPERVISED MULTI-TASK REPRESENTATION LEARNING FOR SPEECH VIA CLIP AND SPEECH-IMAGE DATA	465
<i>Hsuan-Fu Wang, National Taiwan University, Taiwan; Yi-Jen Shih, The University of Texas at Austin, USA, Taiwan; Heng-Jui Chang, Massachusetts Institute of Technology, Taiwan; Layne Berry, Puyuan Peng, The University of Texas at Austin, USA, Taiwan; Hung-yi Lee, National Taiwan University, Taiwan, Taiwan; Hsin-Min Wang, Institute of Information Science, Academia Sinica, Taiwan, Taiwan; David Harwath, The University of Texas at Austin, USA, Taiwan</i>	
WS-3a.6: VICMUS: VARIANCE-INVARIANCE-COVARIANCE REGULARIZATION FOR MUSIC REPRESENTATION LEARNING	475
<i>Sebastian Löf, Epidemic Sound, Sweden; Cody Hesse, Chalmers University of Technology, Sweden; Carl Thomé, Carlos Lordelo, Epidemic Sound, Sweden; Jens Ahrens, Chalmers University of Technology, Sweden</i>	

WS-3a.7: ENHANCING GAN-BASED VOCODERS WITH CONTRASTIVE LEARNING UNDER DATA-LIMITED CONDITION	480
<i>Haoming Guo, Seth Zhao, Jiachen Lian, Gerald Friedland, Gopala Anumanchipalli, UC Berkeley, United States of America</i>	
WS-3a.8: NOISE ROBUST DISTILLATION OF SELF-SUPERVISED SPEECH MODELS VIA CORRELATION METRICS	495
<i>Fabian Ritter-Gutierrez, Nanyang Technological University, Singapore; Kuan-Po Huang, National Taiwan University, Taiwan; Dianwen Ng, Alibaba Group / Nanyang Technological University, Singapore; Jeremy Wong, Institute for Infocomm Research, Singapore; Hung-yi Lee, National Taiwan University, Taiwan; Eng Siong Chng, Nanyang Technological University, Singapore; Nancy Chen, Institute for Infocomm Research, Singapore</i>	
WS-3a.9: BENCHMARKING REPRESENTATIONS FOR SPEECH, MUSIC, AND ACOUSTIC EVENTS	505
<i>Moreno La Quatra, Kore University of Enna, Italy; Alkis Koudounas, Lorenzo Vaiani, Elena Baralis, Luca Cagliero, Paolo Garza, Politecnico di Torino, Italy; Sabato Marco Siniscalchi, Università degli Studi di Palermo, Italy</i>	
WS-3a.10: ON THE TRANSFERABILITY OF LARGE-SCALE SELF-SUPERVISION TO FEW-SHOT AUDIO CLASSIFICATION	515
<i>Calum Heggan, University Of Edinburgh, United Kingdom of Great Britain and Northern Ireland; Sam Budgett, Thales UK, United Kingdom of Great Britain and Northern Ireland; Timothy Hosepedales, Mehrdad Yeghoobi, University Of Edinburgh, United Kingdom of Great Britain and Northern Ireland</i>	
WS-3a.11: PROBING SELF-SUPERVISED LEARNING MODELS WITH TARGET SPEECH EXTRACTION	535
<i>Junyi Peng, Brno University of Technology, Czechia; Marc Delcroix, Tsubasa Ochiai, NTT Corporation, Japan; Oldrich Plchot, Brno University of Technology, Czechia; Takanori Ashihara, Shoko Araki, NTT Corporation, Japan; Honza Černocký, Brno University of Technology, Czechia</i>	
WS-3a.12: INVESTIGATING ZERO-SHOT GENERALIZABILITY ON MANDARIN-ENGLISH CODE-SWITCHED ASR AND SPEECH-TO-TEXT TRANSLATION OF RECENT FOUNDATION MODELS WITH SELF-SUPERVISION AND WEAK SUPERVISION	540
<i>Chih-Kai Yang, Kuan Po Huang, Ke-Han Lu, Chun-Yi Kuan, Chi-Yuan Hsiao, Hung-yi Lee, National Taiwan University, Taiwan</i>	
WS-3b: SELF-SUPERVISION IN AUDIO, SPEECH AND BEYOND II	
WS-3b.1: POSITIVE AND NEGATIVE SAMPLING STRATEGIES FOR SELF-SUPERVISED LEARNING ON AUDIO-VIDEO DATA	545
<i>Shanshan Wang, Soumya Tripathy, Toni Heittola, Annamaria Mesaros, Tampere University, Finland</i>	
WS-3b.2: ANALYSIS OF SELF-SUPERVISED SPEECH MODELS ON CHILDREN'S SPEECH AND INFANT VOCALIZATIONS	550
<i>Jialu Li, Mark Hasegawa-Johnson, Nancy McElwain, University of Illinois Urbana-Champaign, United States of America</i>	
WS-3b.3: ATTENTION OR CONVOLUTION: TRANSFORMER ENCODERS IN AUDIO LANGUAGE MODELS FOR INFERENCE EFFICIENCY	555
<i>Sungho Jeon, Heidelberg Institute of Theoretical Studies, Germany; Ching-Feng Yeh, Hakan Inan, Wei-Ning Hsu, Rashi Rungta, Yashar Mehdad, Daniel Bikel, Meta, United States of America</i>	
WS-3b.4: SOA: REDUCING DOMAIN MISMATCH IN SSL PIPELINE BY SPEECH ONLY ADAPTATION FOR LOW RESOURCE ASR	560
<i>Natarajan Balaji Shankar, Ruchao Fan, Abeer Alwan, University of California, Los Angeles, United States of America</i>	
WS-3b.5: UNCERTAINTY AS A PREDICTOR: LEVERAGING SELF-SUPERVISED LEARNING FOR ZERO-SHOT MOS PREDICTION	580
<i>Aditya Ravuri, University of Cambridge, United Kingdom of Great Britain and Northern Ireland; Erica Cooper, Junichi Yamagishi, National Institute of Informatics, Japan</i>	

WS-3b.6: SYNTHETIC SPEECH DETECTION WITH WAV2VEC 2.0 IN VARIOUS LANGUAGE SETTINGS	585
<i>Branimir Droupuljić, Miljenko Šuflaj, Andrej Jertec, Leo Obadić, RealNetworks, Croatia</i>	
WS-3b.7: LAYER-WISE ANALYSIS OF SELF-SUPERVISED ACOUSTIC WORD EMBEDDINGS: A STUDY ON SPEECH EMOTION RECOGNITION	590
<i>Alexandra Saliba, Yuanchao Li, Ramon Sanabria, Catherine Lai, University of Edinburgh, United Kingdom of Great Britain and Northern Ireland</i>	
WS-3b.8: SELF-SUPERVISED LEARNING FOR FEW-SHOT BIRD SOUND CLASSIFICATION	600
<i>Ilyass Moummad, IMT Atlantique, France; Romain Serizel, University of Lorraine, INRIA, LORIA, France; Nicolas Farrugia, IMT Atlantique, France</i>	
WS-3b.9: LOW-RESOURCE CROSS-DOMAIN SINGING VOICE SYNTHESIS VIA REDUCED SELF-SUPERVISED SPEECH REPRESENTATIONS	605
<i>Panagiotis Kakoulidis, Nikolaos Ellinas, Georgios Vamvoukakis, Myrsini Christidou, Alexandra Vioni, Georgia Maniati, Junkwang Oh, Gunu Jho, Inchul Hwang, Pirros Tsiakoulis, Aimilios Chalamandaris, Samsung Electronics, Greece</i>	
WS-3b.10: A STUDY ON THE IMPACT OF SELF-SUPERVISED LEARNING ON AUTOMATIC DYSARTHIC SPEECH ASSESSMENT	630
<i>Xavier Cadet, Ranya Aloufi, Sara Ahmadi-Abhari, Hamed Haddadi, Imperial College London, United Kingdom of Great Britain and Northern Ireland</i>	
WS-3b.11: LOW-RESOURCED PHONETIC AND PROSODIC FEATURE ESTIMATION WITH SELF-SUPERVISED-LEARNING-BASED ACOUSTIC MODELING	640
<i>Kiyoshi Kurihara, Masanori Sano, NHK (Japan Broadcasting Corporation), Japan</i>	
WS-3b.12: INTEGRATING SELF-SUPERVISED SPEECH MODEL WITH PSEUDO WORD-LEVEL TARGETS FROM VISUALLY-GROUNDED SPEECH MODEL	645
<i>Hung-Chieh Fang, Nai-Xuan Ye, National Taiwan University, Taiwan; Yi-Jen Shih, Puyuan Peng, The University of Texas at Austin, Taiwan; Hsuan-Fu Wang, National Taiwan University, Taiwan; Layne Berry, The University of Texas at Austin, United States of America; Hung-yi Lee, National Taiwan University, Taiwan; David Harwath, The University of Texas at Austin, United States of America</i>	
WS-3c: SELF-SUPERVISION IN AUDIO, SPEECH AND BEYOND III	
WS-3c.1: MULTIMODAL ATTENTION MERGING FOR IMPROVED SPEECH RECOGNITION AND AUDIO EVENT CLASSIFICATION	655
<i>Anirudh S. Sundar, Georgia Institute of Technology, United States of America; Chao-Han Huck Yang, Nvidia Research, United States of America; David M. Chan, University of California, Berkeley, United States of America; Shalini Ghosh, Venkatesh Ravichandran, Phani Sankar Nidadavolu, Amazon Alexa AI, United States of America</i>	
WS-3c.2: CROSS-LINGUAL TRANSFER LEARNING FOR LOW-RESOURCE SPEECH TRANSLATION	670
<i>Sameer Khurana, Mitsubishi Electric Research Lab, United States of America; Nauman Dawalatabad, MIT Computer Science and Artificial Intelligence Laboratory, United States of America; Antoine Laurent, LIUM - Le Mans University, Spain; Luis Vicente, Pablo Gimeno, Victoria Mingote, University of Zaragoza, Spain; James Glass, MIT Computer Science and Artificial Intelligence Laboratory, United States of America</i>	
WS-3c.3: SKILL: SIMILARITY-AWARE KNOWLEDGE DISTILLATION FOR SPEECH SELF-SUPERVISED LEARNING	675
<i>Luca Zampierin, Ecole Polytechnique Fédérale de Lausanne, Switzerland; Ghouthi Boukli Hacene, Bac Nguyen, Sony Europe B.V., Stuttgart Laboratory 1, Canada; Mirco Ravanelli, Concordia University, Canada</i>	

WS-3c.4: INVESTIGATING DESIGN CHOICES IN JOINT-EMBEDDING PREDICTIVE ARCHITECTURES FOR GENERAL AUDIO REPRESENTATION LEARNING	680
<i>Alain Riou, Stefan Lattner, Sony Computer Science Laboratories, France; Gaëtan Hadjeres, Sony AI, Switzerland; Geoffroy Peeters, LTCI, Télécom-Paris, Institut Polytechnique de Paris, France</i>	
WS-3c.5: TRAINING EARLY-EXIT ARCHITECTURES FOR AUTOMATIC SPEECH RECOGNITION: FINE-TUNING PRE-TRAINED MODELS OR TRAINING FROM SCRATCH	685
<i>George August Wright, Umberto Cappellazzo, University of Trento, Italy; Salah Zaiem, Télécom Paris, Institut Polytechnique de Paris, France; Desh Raj, Johns Hopkins University, United States of America; Lucas Ondel Yang, Université Paris-Saclay, French National Centre for Scientific Research, France; Daniele Falavigna, Mohamed Nabih Ali, Alessio Brutti, Fondazione Bruno Kessler, Italy</i>	
WS-3c.6: COCHCEPS-AUGMENT: A NOVEL SELF-SUPERVISED CONTRASTIVE LEARNING USING COCHLEAR CEPSTRUM-BASED MASKING FOR SPEECH EMOTION RECOGNITION	700
<i>Ioannis Ziogas, Hessa Alfalahi, Ahsan Khandoker, Leontios Hadjileontiadis, Khalifa University of Science and Technology, United Arab Emirates</i>	
WS-3c.7: PEFT FOR SPEECH: UNVEILING OPTIMAL PLACEMENT, MERGING STRATEGIES, AND ENSEMBLE TECHNIQUES	705
<i>Tzu-Han Lin, How-Shing Wang, Hao-Yung Weng, Kaung-Chen Peng, Zih-Ching Chen, Hung-yi Lee, National Taiwan University, Taiwan</i>	
 WS-7: SECOND WORKSHOP ON SIGNAL PROCESSING FOR AUTONOMOUS SYSTEMS (SPAS)	
WS-7.1: NAVIGATION FOR AUTONOMOUS ROBOTS WITH ADAPTIVE INFORMATION-SEEKING	159
<i>Kyriakos Lite, Bernhard Rinner, University of Klagenfurt, Austria</i>	
WS-7.2: SIMULTANEOUS GAS EXPLORATION AND NETWORK LOCALIZATION WITH ROBOTIC SWARMS	169
<i>Fabio Broghammer, Thomas Wiedemann, Siwei Zhang, German Aerospace Center, Germany; Benjamin Noack, Otto von Guericke University Magdeburg, Germany</i>	
WS-7.3: DRONAR: OBSTACLE ECHOLOCAION USING DRONE EGO-NOISE	184
<i>Henrik Nilsson, Joakim Rydell, Swedish Defence Research Agency, Sweden; Anton Kullberg, Gustaf Hendeby, Linköping University, Sweden</i>	
WS-7.4: PERFORMANCE EVALUATION AND ANALYSIS OF THRESHOLDING-BASED INTERFERENCE MITIGATION FOR AUTOMOTIVE RADAR SYSTEMS	204
<i>Jun Li, Jihwan Youn, Ryan Wu, Jeroen Overdevest, NXP Semiconductors, United States of America; Shunqiao Sun, The University of Alabama, United States of America</i>	
WS-7.5: LASDNET: A LIGHTWEIGHT ADAPTIVE SURFACE DEFECT DETECTION NETWORK	209
<i>Qiancheng Wei, Yi Yuan, Ying Liu, MOHAMMED ALI MOHAMMED AL-HABABI, University of Chinese Academy of Sciences, China; Qiya Su, Muyao Yu, Beijing Institute of Remote Sensing Equipment, China</i>	
WS-7.6: SPARSE ARRAY DESIGN FOR MIMO RADAR IN MULTIPATH SCENARIOS	239
<i>Xuchen Li, Ronghao Lin, University of Science and Technology of China, China; Hing Cheung So, City University of Hong Kong, Hong Kong</i>	
WS-7.7: JOINT DATA-DRIVEN ANALYSIS OF VISUAL-ODOMETRIC ANOMALY SIGNALS IN GENERATIVE AI-BASED AGENTS	264
<i>Giulia Slavic, Mattia Bracco, Lucio Marcenaro, University of Genoa, Italy; David Martin Gomez, University Carlos III of Madrid, Spain; Carlo Regazzoni, Pamela Zontone, University of Genoa, Italy</i>	
WS-7.8: SOFT ROI POOLING DISTILLATION FOR DETECTION TRANSFORMERS	314
<i>Nikolaos Marios Militsis, Vasileios Mygdalis, Ioannis Pitas, Aristotle University of Thessaloniki, Greece</i>	

WS-7.9: QMARL: A QUANTUM MULTI-AGENT REINFORCEMENT LEARNING FRAMEWORK FOR SWARM ROBOTS NAVIGATION	388
<i>Weizhao Chen, Jiawang Wan, Fangwen Ye, Ran Wang, Cheng Xu, University of Science and Technology Beijing, China</i>	
WS-7.10: AN OBJECT-ORIENTED DEEP LEARNING METHOD FOR VIDEO FRAME PREDICTION	398
<i>Saad Mokssit, International University of Rabat, Morocco; Daniel Bonilla Licea, Czech Technical University in Prague, Czechia; Bassma Guermah, Mounir Ghogho, International University of Rabat, Morocco</i>	
WS-7.11: PARAMETER BLENDING FOR MULTI-CAMERA HARMONIZATION FOR AUTOMOTIVE SURROUND VIEW SYSTEMS	424
<i>Yuzhuo Ren, Yining Deng, David Pajak, Robin Jenkin, Niranjan Avadhanam, Varsha Hedau, NVIDIA, United States of America</i>	
WS-7.12: DRIVESAFE: POSE-BASED DRIVER MONITORING IN SOFTWARE DEFINED VEHICLE USING STGCN	434
<i>Patitapaban Palo, Satyajit Nayak, Durga Nagendra Raghava Kumar Modhugu, Satarupa Uttarkabat, Kwanit Gupta, Valeo India Private Limited, India</i>	
WS-7.13: ADVANCING INDUSTRIAL INSPECTION: A DATASET FOR AUTOMATED DAMAGE DETECTION IN INSULATED PIPES	720
<i>Pantelis Mentesisidis, Christos Papaioannidis, Ioannis Pitas, Aristotle University of Thessaloniki, Greece</i>	
 WS-8: REVOLUTIONIZING INTERACTION: EMBODIED INTELLIGENCE AND THE NEW ERA OF HUMAN-ROBOT COLLABORATION	
WS-8.1: NEURO-MECHANICAL SYNERGY: A SMART IDEA OF INTELLIGENT EXOSKELETONS FOR STROKE REHABILITATION	194
<i>Wei qi Huang, Chenxu Zhang, Wen Qi, South China University of Technology, China</i>	
WS-8.2: HARNESSING VIDEO INTELLIGENCE: INTELLIGENT SYSTEM FOR ADHD DETECTION	274
<i>Yichun Li, Newcastle University, United Kingdom of Great Britain and Northern Ireland; Rajesh Nair, Cumbria, Northumberland, Tyne and Wear (CNTW), NHS Foundation Trust, United Kingdom of Great Britain and Northern Ireland; Mohsen Naqvi, Newcastle University, United Kingdom of Great Britain and Northern Ireland</i>	
WS-8.3: A CROSS-ATTENTION EMOTION RECOGNITION ALGORITHM BASED ON AUDIO AND VIDEO MODALITIES	309
<i>Xiao Wu, xuan MU, Qingdao University, China; Wen qi, South China University of Technology, China; xiaorui liu, Qingdao University, China</i>	
WS-8.4: EVOLVING COMPLEX ENVIRONMENTS IN EVOLUTION GYM USING LARGE LANGUAGE MODELS	320
<i>Fuma Aki, Riku Ikeda, Takumi Saito, Ciaran Regan, Mizuki Oka, University of Tsukuba, Japan</i>	
WS-8.5: DOMAIN-SPECIFIC IMPROVEMENT ON PSYCHOTHERAPY CHATBOT USING ASSISTANT	351
<i>Cheng Kang, Daniel Novak, Czech Technical University in Prague, Czechia; Katerina Urbanova, National Institute of Mental Health, Czechia; Yuqing Cheng, Shenzhen Mental Health Centre, China; Yong Hu, The University of Hong Kong, Hong Kong</i>	
WS-8.6: INTEGRATING AI IN HUMAN-ROBOT INTERACTION: EMERGING CHALLENGES AND FUTURE DIRECTIONS	379
<i>Chenwei Lin, Wen Qi, South China University of Technology, China</i>	

WS-8.7: COLLABORATIVE LEARNING OF COMMON LATENT REPRESENTATIONS IN ROUTINELY COLLECTED MULTIVARIATE ICU PHYSIOLOGICAL SIGNALS	393
<i>Hollan Haule, Ian Piper, Patricia Jones, Tsz-Yan Milly Lo, Javier Escudero, University of Edinburgh, United Kingdom of Great Britain and Northern Ireland</i>	
WS-8.8: DEEP LEARNING EEG TECHNOLOGY DEVELOPMENT AND BRAIN COMPUTER INTERFACE CHIPS PROGRESS	413
<i>Wenjie Li, Beihang University, China; Lei Chen, China Academy of Aerospace Electronics Technology, China; Bing Lv, Beijing Microelectronics Technology Institute, China; Weijun Wang, Robotics Engineering Center, The 21ST Research Institute of China Electronic Technology Group Corporation, China; Kun Xu, Xilun Ding, Beihang University, China</i>	
 WS-10: 1ST WORKSHOP ON INTEGRATION OF SENSING, COMMUNICATION, AND COMPUTATION (ISCC)	
WS-10.1: RECONFIGURABLE INTELLIGENT SURFACE ASSISTED INTEGRATED SENSING, COMMUNICATION AND COMPUTATION SYSTEMS	189
<i>Jiahua Wan, Hong Ren, Cunhua Pan, Zhiyuan Yu, Zhenkun Zhang, Yang Zhang, National Mobile Communications Research Laboratory, Southeast University, China</i>	
WS-10.2: PARETO-OPTIMAL WAVEFORM DESIGN FOR MULTI-USER AND MULTI-TARGET MIMO-ISAC SYSTEMS	219
<i>Peng Wang, North China Electric Power University, China; Yashuai Cao, Wanli Ni, Tsinghua University, China; Dongsheng Han, North China Electric Power University, China</i>	
WS-10.3: LOW-OVERHEAD OFDM SENSING WAVEFORM DESIGN FOR 5G NR: A COPRIME-BASED SUBCARRIER ALLOCATION APPROACH	224
<i>Xiaoye Wang, Huizhou University, China; Fan Liu, Southern University of Science and Technology, China; Le Zheng, Beijing Institute of Technology, China</i>	
WS-10.4: SPECTRUM DATA GRAPH STRUCTURE LEARNING BASED ON DUAL-VIEW CONTRASTIVE LEARNING FOR SPECTRUM PREDICTION OF ISCC	234
<i>Shuang Li, Yaxiu Sun, Zherui Zhang, Han Zhang, Yun Lin, Harbin Engineering University, China</i>	
WS-10.5: FULL-DUPLEX DECODE-AND-FORWARD RELAY FOR JOINT ENVIRONMENTAL SENSING AND SELF-INTERFERENCE CANCELLATION	289
<i>Pu Yang, Xidian University, China; Xiang-Gen Xia, University of Delaware, United States of America; Xichang Zhang, Jiaming Zhang, Yi Liu, Xidian University, China</i>	
WS-10.6: ITERATIVE DETECTION ALGORITHM WITH LDPC DECODER IN CYCLIC-PREFIX OTFS SYSTEM	299
<i>Mingfeng Zhang, Jinhua Sun, Xidian University, China; Xiaojun Wu, Chang' an University, China</i>	
WS-10.7: CRB OPTIMIZATION FOR INTEGRATED SENSING AND COMMUNICATION SYSTEMS USING HYBRID LINEAR-NONLINEAR PRECODING	365
<i>Yating Chen, Southeast University, China; Cai Wen, Northwest University, China; Yan Huang, Southeast University, China; Timothy N. Davidson, McMaster University, Canada</i>	
WS-10.8: INTEGRATED RNNs FOR RAINFALL SENSING WITH WIRELESS COMMUNICATION NETWORKS	419
<i>Dror Jacoby, Jonathan Ostrometzky, Hagit Messer, Tel Aviv University, Israel</i>	
WS-10.9: JOINT TRANSMIT AND RECEIVE CODEBOOK DESIGN FOR SELF-INTERFERENCE SUPPRESSION IN A MMWAVE ISAC TERMINAL DEVICE	575
<i>Guang Chai, Technical University of Berlin, Germany; Zhibin Yu, Xiaofeng Wu, Muhammad Nabeel, Huawei Technologies Duesseldorf GmbH, Germany; Giuseppe Caire, Technical University of Berlin, Germany</i>	

WS-10.10: UAV POSITION DEPLOYMENT AND POWER OPTIMIZATION BASED ON USER CLUSTERING IN IOT NETWORK	625
<i>Xiaoye Shi, Xiangcheng Su, Shu Cai, Zhaowei Zhang, Haibo Dai, Nanjing University of Posts and Telecommunications, China</i>	
WS-10.11: INTELLIGENT HEAVY TRUCK PLATOONING WITH ISCC FOR ENCLOSED INTERMODAL RAILWAY-ROAD TRANSPORT PARKS	650
<i>Weibin Li, Trolley Technology Group Co., Ltd., China; Haifeng Zheng, Fuzhou University, China; Jun Fang, Trolley Technology Group Co., Ltd., China; Xinxin Feng, Fuzhou University, China; Chunyan Cheng, Trolley Technology Group Co., Ltd., China</i>	
WS-10.12: RELIABILITY STUDY OF LEO SATELLITE NETWORKS BASED ON RANDOM LINEAR NETWORK CODING	665
<i>Jiahui Wang, Zan Xu, Ruxin Zhi, Liming Wang, Beijing Information Science and Technology University, China</i>	
WS-10.13: DIRECTIONAL TERAHERTZ RADIO MAP-ASSISTED FEDERATED OBSTACLE SENSING	710
<i>Tianyu Hu, Yunhang Xie, Shuai Wang, Lingxiang Li, Zhi Chen, University of Electronic Science and Technology of China, China</i>	
 WS-15a: HANDS-FREE SPEECH COMMUNICATION AND MICROPHONE ARRAYS (HSCMA 2024): EFFICIENT AND PERSONALIZED SPEECH PROCESSING THROUGH DATA SCIENCE I	
WS-15a.1: LATE AUDIO-VISUAL FUSION FOR IN-THE-WILD SPEAKER DIARIZATION	174
<i>Zexu Pan, Gordon Wichern, Francois Germain, Aswin Subramanian, Jonathan Le Roux, Mitsubishi Electric Research Labs, United States of America</i>	
WS-15a.2: MICROPHONE ALIGNED CONTINUOUS WEARABLE DEVICE-RELATED TRANSFER FUNCTION: EFFICIENT MODELING AND MEASUREMENTS	199
<i>Wageesha Manamperi, Thushara Abhayapala, Australian National University, Australia; Paul Holmberg, Dolby Laboratories, Australia</i>	
WS-15a.3: AMBISONICS ENCODING FOR ARBITRARY MICROPHONE ARRAYS INCORPORATING RESIDUAL CHANNELS FOR BINAURAL REPRODUCTION	244
<i>Yhonatan Gayer, Ben Gurion University of the Negev., Israel; Vladimir Tourbabin, Zamir Ben-Hur, Jacob Donley, Meta, United States of America; Boaz Rafaely, Ben Gurion University of the Negev., Israel</i>	
WS-15a.4: A CROSS-TALK ROBUST MULTICHANNEL VAD MODEL FOR MULTIPARTY AGENT INTERACTIONS TRAINED USING SYNTHETIC RE-RECORDINGS	444
<i>Hyewon Han, Yonsei University, Korea, Republic of; Naveen Kumar, Disney Research Imagineering, United States of America</i>	
WS-15a.5: FAST RANDOM APPROXIMATION OF MULTI-CHANNEL ROOM IMPULSE RESPONSE	449
<i>Yi Luo, Rongzhi Gu, Tencent AI Lab, China</i>	
WS-15a.6: DIFFUSION MODEL-BASED MIMO SPEECH DENOISING AND DEREVERBERATION	455
<i>Rino Kimura, Waseda University, NTT Corporation, Japan; Tomohiro Nakatani, Naoyuki Kamo, Marc Delcroix, Shoko Araki, NTT Corporation, Japan; Tetsuya Ueda, Shoji Makino, Waseda University, Japan</i>	
WS-15a.7: BLIND ESTIMATION OF SPATIAL ROOM IMPULSE RESPONSES USING A PSEUDO REFERENCE SIGNAL	470
<i>Thomas Deppisch, Jens Ahrens, Chalmers University of Technology, Sweden; Sebastià V. Amengual Garí, Paul Calamia, Meta, United States of America</i>	

WS-15a.8: JOINT MINIMUM PROCESSING BEAMFORMING AND NEAR-END LISTENING ENHANCEMENT	485
<i>Andreas Jonas Fuglsig, Jesper Jensen, Zheng-Hua Tan, Aalborg University, Denmark; Lars Søndergaard Bertelsen, Jens Christian Lindof, RTX A/S, Denmark; Jan Østergaard, Aalborg University, Denmark</i>	
WS-15a.9: A TWO-STEP APPROACH FOR NARROWBAND SOURCE LOCALIZATION IN REVERBERANT ROOMS	490
<i>Wei-Ting Lai, Lachlan Birnie, Thushara Abhayapala, Amy Bastine, Shaoheng Xu, Prasanga Samarasinghe, Australian National University, Australia</i>	
WS-15a.10: INSIGHTS INTO MAGNITUDE AND PHASE ESTIMATION BY MASKING AND MAPPING IN DNN-BASED MULTICHANNEL SPEAKER SEPARATION	500
<i>Alexander Bohlender, Ghent University - imec, Belgium; Ann Spriet, Wouter Tirry, Goodix Technology (Belgium) B.V., Belgium; Nilesh Madhu, Ghent University - imec, Belgium</i>	
WS-15a.11: MULTICHANNEL VOICE TRIGGER DETECTION BASED ON TRANSFORM-AVERAGE-CONCATENATE	510
<i>Takuya Higuchi, Apple, United States of America; Avamarie Brueggeman, The University of Texas at Dallas, United States of America; Masood Delfarah, Stephen Shum, Apple, United States of America</i>	
WS-15a.12: EXTERNAL KNOWLEDGE AUGMENTED POLYPHONE DISAMBIGUATION USING LARGE LANGUAGE MODEL	520
<i>Chen Li, Ant Group, China</i>	
 WS-15b: HANDS-FREE SPEECH COMMUNICATION AND MICROPHONE ARRAYS (HSCMA 2024): EFFICIENT AND PERSONALIZED SPEECH PROCESSING THROUGH DATA SCIENCE II	
WS-15b.1: DEEP LOW-LATENCY JOINT SPEECH TRANSMISSION AND ENHANCEMENT OVER A GAUSSIAN CHANNEL	525
<i>Mohammad Bokaei, Jesper Jensen, Aalborg university, Denmark; Simon Doclo, University of Oldenburg, Germany; Jan Østergaard, Aalborg university, Denmark</i>	
WS-15b.2: GEOMETRICALLY CONSTRAINED JOINT MOVING SOURCE EXTRACTION AND DEREVERBERATION BASED ON CONSTANT SEPARATING VECTOR MIXING MODEL	530
<i>Mingxue Song, Tetsuya Ueda, Ruifeng Zhang, Jiahui Hu, Shoji Makino, Waseda University, Japan</i>	
WS-15b.3: CONVOIFILTER: A CASE STUDY OF DOING COCKTAIL PARTY SPEECH RECOGNITION	565
<i>Thai Binh Nguyen, Karlsruhe Institute of Technology, Germany; Alexander Waibel, Carnegie Mellon University, United States of America</i>	
WS-15b.4: JOINT OPTIMIZATION OF STREAMING AND NON-STREAMING AUTOMATIC SPEECH RECOGNITION WITH MULTI-DECODER AND KNOWLEDGE DISTILLATION	570
<i>Muhammad Shakeel, Yui Sudo, Honda Research Institute Japan Co. Ltd., Japan; Yifan Peng, Shinji Watanabe, Carnegie Mellon University, United States of America</i>	
WS-15b.5: TRAINING STRATEGIES FOR MODALITY DROPOUT RESILIENT MULTI-MODAL TARGET SPEAKER EXTRACTION	595
<i>Srikanth Korse, Mohamed Elminshawi, Emanuël Habets, Srikanth Raj Chetupalli, International Audio Laboratories Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany</i>	
WS-15b.6: A UNIFIED GEOMETRY-AWARE SOURCE LOCALIZATION AND SEPARATION FRAMEWORK FOR AD-HOC MICROPHONE ARRAY	725
<i>Jingjie Fan, Southeast University, China; Rongzhi Gu, Yi Luo, Tencent, China; Cong Pang, Southeast University, China</i>	

WS-15b.7: REAL-TIME SPEECH EXTRACTION USING SPATIALLY REGULARIZED INDEPENDENT LOW-RANK MATRIX ANALYSIS AND RANK-CONSTRAINED SPATIAL COVARIANCE MATRIX ESTIMATION	730
<i>Yuto Ishikawa, Kohei Konaka, The University of Tokyo, Japan; Tomohiko Nakamura, The National Institute of Advanced Industrial Science and Technology (AIST), Japan; Norihiro Takamune, Hiroshi Saruwatari, The University of Tokyo, Japan</i>	
WS-15b.8: ENSEMBLE INFERENCE FOR DIFFUSION MODEL-BASED SPEECH ENHANCEMENT	735
<i>Hao Shi, Kyoto University, Japan; Naoyuki Kamo, Marc Delcroix, Tomohiro Nakatani, Shoko Araki, NTT Corporation, Japan</i>	
WS-15b.9: NEURAL STEERER: NOVEL STEERING VECTOR SYNTHESIS WITH A CAUSAL NEURAL FIELD OVER FREQUENCY AND DIRECTION	740
<i>Diego Di Carlo, Aditya Arie Nugraha, Center for Advanced Intelligence Project (AIP), RIKEN, Japan; Mathieu Fontaine, LTCL, Telecom Paris, Institut Polytechnique de Paris, France, France; Yoshiaki Bando, National Institute of Advanced Industrial Science and Technology (AIST), Center for Advanced Intelligence Project (AIP), RIKEN, Japan; Kazuyoshi Yoshii, Graduate School of Informatics, Kyoto University, Center for Advanced Intelligence Project (AIP), RIKEN, Japan</i>	
WS-15b.10: DATA-DRIVEN JOINT DETECTION AND LOCALIZATION OF ACOUSTIC REFLECTORS	745
<i>H. Nazim Bicer, Cagdas Tuna, Andreas Walther, Fraunhofer Institute for Integrated Circuits IIS, Germany; Emanuël A. P. Habets, International Audio Laboratories Erlangen, Germany</i>	
WS-15b.11: SOUND SOURCE SEPARATION USING LATENT VARIATIONAL BLOCK-WISE DISENTANGLEMENT	750
<i>Karim Helwani, Masahito Togami, Paris Smaragdis, Michael M. Goodwin, Amazon, United States of America</i>	
WS-15b.12: MULTI-CHANNEL SPEECH ENHANCEMENT USING BEAMFORMING AND NULLFORMING FOR SEVERELY ADVERSE DRONE ENVIRONMENT	755
<i>Seokhyun Kim, Won Jeong, Hyung-Min Park, Sogang Univ., Korea, Republic of</i>	
 WS-15c: HANDS-FREE SPEECH COMMUNICATION AND MICROPHONE ARRAYS (HSCMA 2024): EFFICIENT AND PERSONALIZED SPEECH PROCESSING THROUGH DATA SCIENCE III	
WS-15c.1: GATED LOW-RANK ADAPTATION FOR PERSONALIZED CODE-SWITCHING AUTOMATIC SPEECH RECOGNITION ON THE LOW-SPEC DEVICES	760
<i>Gwantae Kim, Bokyeung Lee, Donghyeon Kim, Hanseok Ko, Korea University, Korea, Republic of</i>	
WS-15c.2: LIGHT GATED MULTI MINI-PATCH EXTRACTOR FOR AUDIO CLASSIFICATION	765
<i>Bo He, Shiqi Zhang, Xianrui Wang, Zheng Qiu, Waseda University, Japan; Daiki Takeuchi, Daisuke Niizumi, Noboru Harada, NTT Corporation, Japan; Shoji Makino, Waseda University, Japan</i>	
WS-15c.3: MEETING RECOGNITION WITH CONTINUOUS SPEECH SEPARATION AND TRANSCRIPTION-SUPPORTED DIARIZATION	775
<i>Thilo von Neumann, Christoph Boeddeker, Tobias Cord-Landwehr, Paderborn University, Germany; Marc Delcroix, NTT Corporation, Japan; Reinhold Haeb-Umbach, Paderborn University, Germany</i>	
WS-15c.4: A LIGHTWEIGHT DUAL-STAGE FRAMEWORK FOR PERSONALIZED SPEECH ENHANCEMENT BASED ON DEEPPILTERNET2	780
<i>Thomas Serre, Orosound, France; Mathieu Fontaine, Télécom Paris, Institut Polytechnique de Paris, France; Éric Benhaim, Geoffroy Dutour, Orosound, France; Slim Essid, Télécom Paris, Institut Polytechnique de Paris, France</i>	
WS-15c.5: VOICEVECTOR: MULTIMODAL ENROLMENT VECTORS FOR SPEAKER SEPARATION	785
<i>Akam Rahimi, Andrew Zisserman, Triantafyllos Afouras, University of Oxford, United Kingdom of Great Britain and Northern Ireland</i>	

WS-15c.6: HOMULA-RIR: A ROOM IMPULSE RESPONSE DATASET FOR 795
TELECONFERENCING AND SPATIAL AUDIO APPLICATIONS ACQUIRED THROUGH HIGHER-
ORDER MICROPHONES AND UNIFORM LINEAR MICROPHONE ARRAYS
Federico Miotello, Paolo Ostan, Mirco Pezzoli, Luca Comanducci, Alberto Bernardini, Fabio Antonacci, Augusto Sarti,
Politecnico di Milano, Italy