



第二十届

全国人机语音通讯学术会议

2025 年 10 月 16 - 19 日 江苏·镇江

Thursday, October 16

[14:30-20:00] Thu-R - 会议注册 1

Friday, October 17

[9:00-10:30] Fri-M-T-1 - Deep Speaker Modeling: Theories, Applications and Practice 1

[11:00-12:30] Fri-M-T-2 - 分布式麦克风阵列拾音理论与方法 1

[13:30-15:00] Fri-A-T - 音频深度伪造检测 1

[14:00-16:00] Fri-A-ES - 科普教育走进中小学活动 1

[15:30-17:00] Fri-A-IND - Industry Forum 1 2

Saturday, October 18

[9:15-10:00] Sat-M-PLEN-1 - The Zapformer: New work in Model architectures 3

[10:00-10:45] Sat-M-PLEN-2 - 面向类人智能的语音机理研究初探 3

[11:00-12:30] Sat-M-O-1 - Best Paper Presentation 3

[11:00-12:30] Sat-M-YF - Young Forum 1 4

[11:00-12:30] Sat-M-SS-1 - 言语障碍人群汉语普通话的语音感知与产出 4

[11:00-12:30] Sat-M-SS-2 - 语音，情感与精神健康 5

[11:00-12:30] Sat-M-P-1 - Phonetics, Phonology and Prosody; Speech Production, Perception and Acquisition 5

[14:00-14:45] Sat-A-PLEN - 迈向可靠的任务型大语言模型智能体 8

[15:00-16:30] Sat-A-O-1 - Speech Synthesis and Conversion	9
[15:00-16:30] Sat-A-YF - Young Forum 2	9
[15:00-16:30] Sat-A-SS-1 - AI 语音赋能生命健康：无障碍沟通与人机协同创新	10
[15:00-16:30] Sat-A-SS-2 - 复杂场景下的通用声音增强与分离	10
[15:00-16:30] Sat-A-SS-3 - 音色属性强度检测竞赛（NCMMSC2025-VtaD）	10
[16:45-18:15] Sat-A-O-2 - Speech Coding and Speech Enhancement	10
[16:45-18:15] Sat-A-SS-4 - 中文连续视觉语音识别挑战赛（CNVSRC 2025）	11
[16:45-18:15] Sat-A-SS-5 - 第二届半监督智能语音与语言技术研讨会	11
[16:45-18:15] Sat-A-SS-6 - 开源论坛 OASIS · 开源音频语音处理分论坛	12
[16:45-18:15] Sat-A-P-1 - Detection and Classification of Acoustic Scenes and Events	12
[20:30-21:30] Sat-E-FS - 基金分享沙龙活动	15
Sunday, October 19	
[9:00-9:45] Sun-M-PLEN-1 - VoxCPM：面向高拟真语音生成的高效大模型架构探索与应用	16
[9:45-10:30] Sun-M-PLEN-2 - 《机器合成普通话水平测评等级标准及测评大纲》研制与实践	16
[10:45-12:15] Sun-M-O-1 - Automatic Speech Recognition	16
[10:45-12:15] Sun-M-SF - Student Forum 1	17
[10:45-12:15] Sun-M-YF - Young Forum 3	17
[10:45-12:15] Sun-M-P-1 - Speech Coding and Speech Enhancement	17
[13:15-14:45] Sun-A-O-1 - Speech Technology	20
[13:15-14:45] Sun-A-SF - Student Forum 2	21
[13:15-14:45] Sun-A-SS-1 - 人机语音交互中的多模态智能建模与个性化感知	22
[13:15-14:45] Sun-A-SS-1 - 2025 CCF 先进音频技术竞赛	22
[15:00-16:30] Sun-A-O-2 - Spoof Speech Detection & Others	22
[15:00-16:30] Sun-A-O-3 - Speech, Paralinguistics and Language Analysis	23

[15:00-16:30] Sun-A-O-4 - Audio Processing	24
[15:00-16:30] Sun-A-P-1 - Automatic Speech Recognition	24
Organizer List	28

Thursday, October 16

Thu-R: 会议注册
Room: 商务楼一楼大厅
Type: Registration
October 16, 2:30 PM to 8: 00 PM

Friday, October 17

Fri-M-T-1: Deep Speaker Modeling: Theories, Applications and Practice
Room: 研发楼 401
Type: Tutorial
October 17, 9:00 AM to 10:30 AM
Presenter: 王帅老师, 钱彦旻老师
Chair: 毛启容老师

Fri-M-T-2: 分布式麦克风阵列拾音理论与方法
Room: 研发楼 401
Type: Tutorial
October 17, 11:00 AM to 12:30 PM
Presenter: 张结老师
Chair: 毕凡老师

Fri-A-T: 音频深度伪造检测
Room: 研发楼 401
Type: Tutorial
October 17, 1:30 PM to 3:00 PM
Presenter: 程皓楠老师
Chair: 高利剑老师

Fri-A-ES: 科普教育走进中小学活动
Room: 江苏大学附属学校
Type: Education Session
October 17, 2:00 PM to 4:00 PM
Presenter: 马少平老师, 王东老师
Chair: 毛启容老师

14: 00-15: 00

Fri-A-ES-1: 人工智能的奥秘
Presenter: 马少平 (清华大学)

15: 00-16: 00
Fri-A-ES-1: 机器如何能听会说？
Presenter: 王东 (清华大学)

Fri-A-IND: Industry Forum 1
Room: 研发楼 401
Type: Industry Forum
October 17, 3:30 PM to 5:30 PM
Chair: 陈谐老师

15:30-16:00
Fri-A-IND-1: OPPO 小布助手语音交互中 TTS 的技术演进
Presenter: 梁敏 (OPPO)

16:00-16:15
Fri-A-IND-2: 赋能语音大模型：高质量数据实践路径
Presenter: 肖永红 (数据堂)

16:15-16:30
Fri-A-IND: Break time

16:30-16:45
Fri-A-IND-3: 智能计算: 期刊简介与展望
Presenter: 金熠 (之江实验室)

16:45-17:00
Fri-A-IND-4: 高质量语音数据：驱动人工智能语音技术创新的基石
Presenter: 胡婕 (慧听科技)

17:00-17:15
Fri-A-IND-5: 构建高质量端到端语音大模型数据集的关键技术与实践
Presenter: 穆向禹 (标贝科技)

17:15-17:30
Fri-A-IND-6: Dolphin 语音大模型技术框架及底层数据构建
Presenter: 王冠博 (海天瑞声)

Saturday, October 18

Sat-M-PLEN-1: The Zapformer: New work in Model architectures

Room: 商务楼二楼-华怡厅

Type: Plenary

October 18, 9:15 AM to 10:00 AM

Presenter: Daniel Povey 老师

Chair: 郑方老师

Sat-M-PLEN-2: 面向类人智能的语音机理研究初探

Room: 商务楼二楼-华怡厅

Type: Plenary

October 18, 10:00 AM to 10:45 AM

Presenter: 魏建国老师

Chair: 郑方老师

Sat-M-O-1: Best Paper Presentation

Room: 商务楼三楼-明都厅

Type: Oral

October 18, 11:00 AM to 12:30 PM

Chair: 凌震华老师

Sat-M-O-1-1: MPO: Multidimensional Preference Optimization for Language Model-based Text-to-Speech

夏康翔 (西北工业大学)*; Xinfu Zhu (Northwestern Polytechnical University); Jixun Yao (Northwestern Polytechnical University); Xie Lei (Northwestern Polytechnical University)

Sat-M-O-1-2: De-Codec: Decoupling Speech and Background-sound in Codec to Allow Explicit Feature Selections for Downstream Audio Tasks

Xiaoxue Luo (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University)*; Jinwei Huang (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University); Zihao Cui (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University); Tao Li (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University); Yanan Chen (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University); Yingying Gao (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University); Chao Deng (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University); Junlan Feng (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University); Shilei Zhang (China Mobile Research Institute; State Key Laboratory of Multimedia Information Processing, Peking University);

Sat-M-O-1-3: ZO-ASR: Zeroth-Order Fine-Tuning of Speech Foundation Models without Back-propagation

Yuezhang PENG (Shanghai Jiao Tong University)*; Yuxin Liu (Shanghai Jiao Tong University); Yao Li (AVIC); Sheng Wang (Shanghai Jiao Tong University); Fei Wen (Shanghai Jiao Tong University); Xie Chen (Shanghai Jiao Tong University)

Sat-M-O-1-4: 面向构音障碍者的音素识别与评估方法

陈晓玥 (深圳大学)*; 江志健 (深圳大学); 郑能恒 (深圳大学); 朝魏刚 (深圳市龙岗区耳鼻咽喉医院, 深圳市耳鼻咽喉研究所, 深圳市龙岗区耳鼻咽喉医院-华南理工大学数字听力健康联合实验室); 郑慧丽 (深圳市龙岗区耳鼻咽喉医院, 深圳市耳鼻咽喉研究所, 深圳市龙岗区耳鼻咽喉医院-华南理工大学数字听力健康联合实验室); 李娟娟 (深圳市龙岗区耳鼻咽喉医院, 深圳市耳鼻咽喉研究所, 深圳市龙岗区耳鼻咽喉医院-华南理工大学数字听力健康联合实验室)

Sat-M-O-1-5: Advancing Speech Data Collection and Annotation: General Methods, Practical Experience, and Future Perspectives

Yingming Gao (Beijing University of Posts and Telecommunications)*; Xiaoli Feng (Ministry of Education Institute of Applied Linguistics); Boxue Li (Yunfan Massive Data Technology (Beijing) Co.); Yanlu Xie (Beijing Language and Culture University); Jinsong Zhang (Beijing Language and Culture University); Ya Li (Beijing University of Posts and Telecommunications)

Sat-M-O-1-6: Exploring Audio-Visual Fusion for Sound Event Localization and Detection with BEATs

Chengnuo Sun (Jiangsu University); Lijian Gao (Jiangsu University); Wenlong Dong (Jiangsu University); Shiyu Yu (Jiangsu University); Qirong Mao (Jiangsu University)*

Sat-M-YF: Young Forum 1

Room: 商务楼二楼-华怡厅

Type: Young Forum

October 18, 11:00 AM to 12:30 PM

Chair: 欧智坚老师

Sat-M-YF-1: 面向声音信号理解的多模态持续学习机理研究

Presenter: 许可乐老师, 国防科技大学计算机学院副研究员

Sat-M-YF-2: 研究情智一体的音视频内容生成研究进展

Presenter: 刘李老师, 香港科技大学 (广州) 助理教授、博导

Sat-M-YF-3: 轻量级神经网络语音编码方法研究

Presenter: 艾杨老师, 中国科学技术大学语音及语言信息处理国家工程研究中心副研究员

Sat-M-SS-1: 言语障碍人群汉语普通话的语音感知与产出

Room: 研发楼 401

Type: Special Session

October 18, 11:00 AM to 12:30 PM

Organizer: 陈莹老师, 龚箭老师

Chair: 陈莹老师

Sat-M-SS-1-1: 语前聋成人普通话单字调生成的时域特征研究

Presenter: 陈彧老师 (东南大学)

Sat-M-SS-1-2: 韵律和表情线索促进汉语人工耳蜗儿童的反语理解

Presenter: 汤平老师 (南京理工大学)

Sat-M-SS-1-3: 腭裂儿童汉语普通话口鼻塞音的感知和产出研究
 Presenter: 陈莹老师（南京理工大学）

Sat-M-SS-1-4: 语义对汉语孤独症儿童情绪韵律感知的调节机制研究
 Presenter: 王婷老师（同济大学）

Sat-M-SS-1-5: 老年人汉语普通话的辅音感知
 Presenter: 冯燕老师（南京理工大学）

Sat-M-SS-1-6: 隆巴德效应对老年人噪声中言语接收阈值的影响
 Presenter: 龚箭老师（江苏科技大学）

Sat-M-SS-2: 语音，情感与精神健康
Room: 研发楼 402
Type: Special Session
October 18, 11:00 AM to 12:30 PM
Organizer: 罗兆杰老师，宗源老师
Chair: 罗兆杰老师

Sat-M-SS-2-1: 音频--数字健康的新启发
 Presenter: 钱昆老师（北京理工大学）

Sat-M-SS-2-2: 基于多模态大模型的情感理解
 Presenter: 连政老师（中国科学院自动化研究所）

Sat-M-SS-2-3: 情智兼备的对话语音生成研究
 Presenter: 刘瑞老师（内蒙古大学计算机学院、软件学院、人工智能学院）

Sat-M-SS-2-4: 多模态协同增强的情感分析
 Presenter: 李勇老师（东南大学）

Sat-M-SS-2-5: 基于可穿戴生理信号的细粒度情感识别研究
 Presenter: 张天翼老师（东南大学）

Sat-M-P-1: Phonetics, Phonology and Prosody; Speech Production, Perception and Acquisition
Room: 研发楼四楼-海报区
Type: Poster
October 18, 11:00 AM to 12:30 PM

Sat-M-P-1-1: Experimental Study of Tonal Type and Tone Sandhi Pattern in Zhangzhou: A Jointly Tone Circulatory Theory Analysis
 WANG Chun (Deparment of Linguistcis and Modern Languages,The Chinese University of Hongkong)*

Sat-M-P-1-2: 基于 Z 值法的五度调值划分研究
 Cao Hao yang (BLCU)*; Lu Xi yao (BLCU); Zhang Jin song (BLCU)

Sat-M-P-1-3: EMA 发音生理特征与说话人个性化声学表征的相关性初探

Xinyu CHAN (Beijing Language and Culture University); Hui JIA (Beijing Language and Culture University); Yanlu XIE (Beijing Language and Culture University)*

Sat-M-P-1-4: 普通话声调音高的加速度特征考察

Cao Hao yang (BLCU); Zhang Jinsong (北京语言大学)*

Sat-M-P-1-5: The determinants of Chinese lexical stress

Zhigang Yin (Institute of Linguistics, Chinese Academy of Social Sciences)*

Sat-M-P-1-6: 越南高级阶段汉语学习者句子停顿习得研究

范氏诗 (琴趋汉语培训中心)*

Sat-M-P-1-7: Research on the Generation Mechanism of Prosodic Boundaries at Different Speech Rates – Analysis of pause patterns based on natural speech

刘璐 (辽宁师范大学)*; 许嘉文 (辽宁师范大学); 王蓓 (北京理工大学)

Sat-M-P-1-8: 武汉方言两字组连读变调声学实现线索分析

徐琬萍 (Shandong University); 李爱军 (Institute of Linguistics, Chinese Academy of Social Sciences)*

Sat-M-P-1-9: An acoustic comparison of tones and vowels in Standard Mandarin and Dongbei Mandarin: Evidence from Changchun

Wei Zhang (Nanjing University)*; Xuanda Chen (McGill); Aini Li (City University of Hong Kong)

Sat-M-P-1-10: The Phonetic Properties of Sibilant Place Contrasts in Three Chinese Dialects

Mingxing Li (Hong Kong Baptist University)*; Pauline Bolin Liu (Hong Kong Baptist University)

Sat-M-P-1-11: The Phonetics of Tonal Contrast in Apical Vowel: A Case Study of Meihua Dialect

Pauline Bolin Liu (Hong Kong Baptist University); Mingxing Li (Hong Kong Baptist University)*

Sat-M-P-1-12: The influence of prosodic structure on tone variation and change in Chengdu Mandarin

Liang Zhao (Beijing Foreign Studies University)*; Aini Li (City University of Hong Kong)

Sat-M-P-1-13: 浅谈越南语声调音位功能负载及其标记性研究

Phat Bach (Chinese Language and Chinese Studies Department, School of Information Science, Beijing Language and Culture University); 吴晓琳 (北京语言大学信息科学学院); 张劲松 (北京语言大学信息科学学院)*

Sat-M-P-1-14: A Comparison of the Syllable-initial /n/ and /l/ in Standard Mandarin: Articulatory, Acoustical, and Aerodynamic Characteristics

Yijing He (City University of Hong Kong)*; Wai-Sum LEE (City University of Hong Kong)

Sat-M-P-1-15: Lingual configuration in whispered Mandarin across phonation types: an ultrasound study

Chutian Zhao (City university of Hong Kong)*; Ying Lu (City University of Hong Kong); Bin Li (City University of Hong Kong)

Sat-M-P-1-16: 普通话语速的影响因素探究

王毅慧 (北京师范大学); 汪高武 (北京师范大学)*

Sat-M-P-1-17: 北京话朗读与口语的节奏特征研究

Zhiwei Wang (University of Chinese Academy of Social Sciences); Aijun LI ("Institute of Linguistics, CASS")*

Sat-M-P-1-18: Accent Familiarity and Phonological Weighting in Spoken-Word Recognition

Jia Ying (The Chinese University of Hong Kong, Shenzhen)*

Sat-M-P-1-19: 基于声学生理特征的泰国留学生汉语元音“ie”偏误研究

Jinshuan Zhang (Zhang Jinshuang)*

Sat-M-P-1-20: Early Production of Successive Addition Boundary Tones in Mandarin: Preliminary Evidence from Monolingual and Bilingual Children's Corpora

Wangqian Zhou (The Chinese University of Hong Kong)*

Sat-M-P-1-21: Cross-Linguistic Control of L2 Consonantal F0: Interaction Between L1 Phonology and Prosodic Planning

Junjie Zhang (University of Cambridge)*; Brechtje Post (University of Cambridge)

Sat-M-P-1-22: Investigating the Mechanisms of Suprasegmental Adaptation: The Role of Input Variance in Mandarin Tone Perception

Muxuan Liu (The University of Edinburgh)*

Sat-M-P-1-23: SAIT-EMA: 多语背景汉语发音者三维电磁发音数据库

HUI JIA (北京语言大学); FANG QIANG (中国社会科学院语言研究所); SUN Yifeng (北京语言大学); XIE YANLU (北京语言大学)*

Sat-M-P-1-24: 基于 EMA 发声生理数据的越南汉语学习者舌尖前音“z[ts]”偏误初探

HUI JIA (北京语言大学); QIANG FANG (中国社会科学院语言研究所); YANLU XIE (北京语言大学)*

Sat-M-P-1-25: 越南学习者普通话持续短语边界调的韵律特征分析

Yurong Wu (Beijing Language and Culture University); Jinsong Zhang (Beijing Language and Culture University)*

Sat-M-P-1-26: An Acoustic Study on Intonation Production of English Learners from Guanzhong Region in Shaanxi Province

Chenyang Zhao (College of Language and Culture, Northwest A&F University); Yuxuan Wu (College of Language and Culture, Northwest A&F University)*

Sat-M-P-1-27: 孟加拉国学习者汉语声调、语调偏误研究——基于单音节字、单音节句的考察

景佳 (北京语言大学)*

Sat-M-P-1-28: Effects of Explicit Phonetic Instruction in Shaping L2 Perceptual Cue Weighting: Evidence from English /i:/-//

Ziyue Zheng (University of Oxford)*

Sat-M-P-1-29: 反语的声学线索：基于合成语音的感知研究

Shanpeng Li (Nanjing University of Science and Technology); Wentao Gu (Nanjing Normal University)*

Sat-M-P-1-30: Speech Perception in Blind Individuals: Emotion and Synthetic Voice Evaluation

王毅慧 (北京师范大学); haocheng Lu (北京师范大学); gaowu Wang (北京师范大学)*

Sat-M-P-1-31: Acoustic Features of Consonants and Vowels in Standard Spoken Chinese among Dong and Miao Learners in Guizhou

Yuan JIA (Chinese Academy of Social Sciences); Mingshuai Yin (University Of Chinese Academy Of Social Sciences (UCASS))*

Sat-M-P-1-32: Interaction of Stress Pattern and Tone Type in Mandarin Tertiary EFL Learners' Perception of English Heteronym Lexical tress

Weijing Zhou (Yangzhou University)*; Nolan Francis (University of Cambridge); Xinru Su (Yangzhou University)

Sat-M-P-1-33: Acquisition of Focus Prosody in English Dialogue Produced by Chinese Tertiary EFL Learners

Xinyue Zhang (Yangzhou University)*; Weijing Zhou (Yangzhou University); Francis Nolan (University of Cambridge); Hui Gong (Yangzhou University); Li Liu (Yangzhou University)

Sat-M-P-1-34: Acquisition of English Plosive Clusters Articulated by Mandarin College Students: A Survey Study

Ziye Cao (Yangzhou University)*; Weijing Zhou (Yangzhou University); Francis Nolan (University of Cambridge); Xinru Su (Yangzhou University); Yan Huang (Yangzhou University)

Sat-M-P-1-35: Fluency, Dis-fluency and Learner-related Factors of English Narrative Passage Read by Chinese College Students

Zhazhan Tian (Yangzhou University)*; Weijing Zhou (Yangzhou University); Francis Nolan (University of Cambridge); Jingying Xie (Yangzhou University); Mingyue Zhu (Yangzhou University)

Sat-M-P-1-36: Categorical Perception of Mandarin Vowel in Children Aged 4 to 7 Years with Autism

Zhaoxuan QIAN (北京师范大学); junlin Huang (北京师范大学); Gaowu Wang (北京师范大学)*

Sat-M-P-1-37: 普通话智力障碍儿童上声变调能产性实验研究

Rongrui Ren (上海外国语大学中文学院/国际文化交流学院)*; Li Yu (上海外国语大学中文学院/国际文化交流学院); Haifeng Qi (上海外国语大学中文学院/国际文化交流学院)

Sat-M-P-1-38: The Influence of Prosodic Prominence and Syntactic Cues on Focus Realization and Perception: In the Case of Minnan Dialect

Ding Yidan (Tianjin University)*

Sat-M-P-1-39: 印尼学习者对二语汉语声调与语调的感知研究

Min Liu (Jinan University)*

Sat-M-P-1-40: 基于机器学习建模的汉语声调感的关键时间窗口研究

陈月 (南京航空航天大学)*

Sat-M-P-1-41: Focus marked by pitch, but not heard: Contrastive focus production by Mandarin-speaking children with cochlear implants

Feng Xu (Macquarie University)*; Ping Tang (Nanjing University of Science and Technology); Katherine Demuth (Macquarie University); Nan Xu Rattanasone (Macquarie University)

Sat-M-P-1-41: Focus marked by pitch, but not heard: Contrastive focus production by Mandarin-speaking children with cochlear implants

Feng Xu (Macquarie University)*; Ping Tang (Nanjing University of Science and Technology); Katherine Demuth (Macquarie University); Nan Xu Rattanasone (Macquarie University)

Sat-A-PLEN: 迈向可靠的任务型大语言模型智能体

Room: 商务楼二楼-华怡厅

Type: Plenary

October 18, 2:00 PM to 2:45 PM

Presenter: 俞凯老师

Chair: 毛启容老师

Sat-A-O-1: Speech Synthesis and Conversion

Room: 商务楼三楼-明都厅

Type: Oral

October 18, 3:00 PM to 4:30 PM

Chair: 陈诺老师

Sat-A-O-1-1: 基于知识蒸馏的低延迟神经网络声码器

Hui-Peng Du (University of Science and Technology of China)*; Yang Ai (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China)

Sat-A-O-1-2: CL-EDiff: Cross-lingual emotional TTS system based on diffusion model

Zhihua Huang (School of Computer Science and Technology, Xinjiang University, Urumqi,)*; Haiyang Zhou (School of Computer Science and Technology, Xinjiang University, Urumqi,); Bowen Li (School of Computer Science and Technology, Xinjiang University, Urumqi,)

Sat-A-O-1-3: High-Quality and Low-Complexity Neural Speech Bandwidth Extension in MDCT Spectral Domain

Yuan Tian (University of Science and Technology)*; Yang Ai (University of Science and Technology); Ye-Xin Lu (University of Science and Technology); Hui-Peng Du (University of Science and Technology); Zhen-Hua Ling (University of Science and Technology)

Sat-A-O-1-4: Training Text-to-Speech Model with Purely Synthetic Data: Feasibility, Sensitivity, and Generalization Capability

Tingxiao Zhou (Shanghai Jiao Tong University)*; Leying Zhang (Shanghai Jiao Tong University); Zhengyang Chen (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)

Sat-A-O-1-5: SynthVC: Leveraging Synthetic Data for End-to-End Low Latency Streaming Voice Conversion

Zhao Guo (NWPU)*; Ziqian Ning (NWPU); Guobin Ma (NWPU); Lei Xie (NWPU)

Sat-A-O-1-7: ZSEmo-MTVITS: A Zero-Shot Cross-Lingual Emotional Speech Synthesis Model for Mandarin and Tibetan Based on VITS

weizhao zhang (Northwest Normal University)*; Mengjuan WANG (Northwest Normal University); Hongwu YANG (Northwest Normal University)

Sat-A-YF: Young Forum 2

Room: 商务楼二楼-华怡厅

Type: Young Forum

October 18, 3:00 PM to 4:30 PM

Chair: 李明老师

Sat-A-YF-1: 面向开放域与可解释多模态情感理解

Presenter: 连政老师, 中国科学院自动化研究所多模态人工智能系统全国重点实验室副研究员

Sat-A-YF-2: 基于听觉注意力解码的神经导向语音分离方法研究

Presenter: 范存航老师, 安徽大学计算机科学与技术学院副教授, 博士生导师

Sat-A-YF-3: 面向通用场景的语音增强研究: URGENT 系列挑战赛回顾

Presenter: 张王优老师, 上海交通大学人工智能学院助理教授

Sat-A-SS-1: AI 语音赋能生命健康: 无障碍沟通与人机协同创新

Room: 研发楼 401

Type: Special Session

October 18, 3:00 PM to 4:30 PM

Organizer: 刘李老师, 吴锡欣老师, 燕楠老师, 王甦菁老师

Chair: 刘李老师

Sat-A-SS-1-1: 双重视角下的 “AI+ 辅具”

Presenter: 王甦菁老师 (中国科学院心理研究所)

Sat-A-SS-1-2: 基于大语言模型的构音障碍语音重建

Presenter: 吴锡欣老师 (香港中文大学)

Sat-A-SS-1-3: 基于语音分析的轻度认知障碍筛查系统的方法研究

Presenter: 燕楠老师 (中科院先进技术研究院)

Sat-A-SS-2: 复杂场景下的通用声音增强与分离

Room: 研发楼 402

Type: Special Session

October 18, 3:00 PM to 4:30 PM

Organizer: 涂卫平老师, 任延珍老师, 杨玉红老师

Chair: 涂卫平老师

Sat-A-SS-3: 音色属性强度检测竞赛 (NCMMSC2025-VtaD)

Room: 研发楼 308

Type: Special Session

October 18, 3:00 PM to 4:30 PM

Organizer: 陈丽萍老师, Kong Aik Lee 老师, 凌震华老师

Chair: 陈丽萍老师

Sat-A-O-2: Speech Coding and Speech Enhancement

Room: 商务楼三楼-明都厅

Type: Oral

October 18, 4:45 PM to 6:15 PM

Chair: 范存航老师

Sat-A-O-2-1: Improving SI-SNR Loss with Frequency Weighting for Speech Enhancement

Shaoxiong Lin (ShanghaiJiaoTongUniversity)*; Wangyou Zhang (ShanghaiJiaoTongUniversity); Leying Zhang (ShanghaiJiaoTongUniversity); Yanmin Qian (ShanghaiJiaoTongUniversity)

Sat-A-O-2-2: Research on Speech Enhancement Algorithm of Hearing Aid based on Capsule Network

Hairong Jia (Taiyuan University of Technology)*; Jiajie Pan (Taiyuan University of Technology); Zhengjie Lv (Taiyuan University of Technology)

Sat-A-O-2-3: VoCodec: 一种基于发音属性量化的低比特率流式神经网络语音编解码器

Xiao-Hang Jiang (University of Science and Technology of China)*; Yang Ai (University of Science and Technology of China); Rui-Chen Zheng (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China)

Sat-A-O-2-4: Integrating Multi-Scale Acoustic Features with State Space Model for Speech Separation

Wang Xiang (Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University); Jian Zhou (Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University); Yujie Chen (Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University); Shuai Fan (AISpeech Ltd); Qiang Zhou (AISpeech Ltd); Zhao Lv (Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University); Cunhang Fan (Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University)*

Sat-A-O-2-5: 基于知识蒸馏的高质量低复杂度流式语音编码技术

En-Wei Zhang (University of Science and Technology of China)*; Hui-Peng Du (University of Science and Technology of China); Xiao-Hang Jiang (University of Science and Technology of China); Yang Ai (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China)

Sat-A-O-2-6: A Neural Speech Codec for Noise Robust Speech Coding

Jiayi Huang (Shanghai Jiao Tong University); Zeyu Yan (Shanghai Jiao Tong University); Wenbin Jiang (Hangzhou Dianzi University); He Wan (SHANGHAI AERO MEASUREMENT CONTROLLING RESEARCH INSTITUTE); Fei Wen (Shanghai Jiao Tong University)*

Sat-A-SS-4: 中文连续视觉语音识别挑战赛 (CNVSRC 2025)

Room: 研发楼 308

Type: Special Session

October 18, 4:45 PM to 6:15 PM

Organizer: 王东老师, 李蓝天老师, 李科老师

Chair: 王东老师

Sat-A-SS-5: 第二届半监督智能语音与语言技术研讨会

Room: 研发楼 401

Type: Special Session

October 18, 4:45 PM to 6:15 PM

Organizer: 张鹏远老师, 欧智坚老师

Chair: 张鹏远老师

Sat-A-SS-5-1: 基于检索增强的语音识别技术以及挑战性语音任务建设

Presenter: 秦勇老师 (南开大学)

Sat-A-SS-5-2: Recent Advances in Machine Voiceprint for Robust Machine Anomalous Sound Detection

Presenter: 钱彦旻老师 (上海交通大学)

Sat-A-SS-5-3: 基于自监督学习方法的声纹识别

Presenter: 李明老师 (昆山杜克大学)

Sat-A-SS-5-4: Towards Zero-shot MoE Speaker Adaptation of Speech Foundation Models for Dysarthric and Elderly Speech Recognition

Presenter: 刘循英老师 (香港中文大学)

Sat-A-SS-5-5: Phonemes as Speech-Language Interface for Data-Efficient Multilingual and Crosslingual Speech Recognition

Presenter: 欧智坚老师 (清华大学)

Sat-A-SS-6: 开源论坛 OASIS · 开源音频语音处理分论坛**Room:** 研发楼 402**Type:** Special Session**October 18, 4:45 PM to 6:15 PM****Organizer:** 王帅老师, 谢磊老师, 陈谐老师**Chair:** 王帅老师**Sat-A-P-1: Detection and Classification of Acoustic Scenes and Events****Room:** 商务楼二楼-海报区**Type:** Poster**October 18, 16:45 PM to 18:15 PM****Sat-A-P-1-1: 全双工语音识别优化的回声消除模型算法**

Yuxi Lai (Multimedia Product Center, Beijing OPPO Mobile Communication Co., Ltd.)*; Ziteng Wang (Multimedia Product Center, Beijing OPPO Mobile Communication Co., Ltd.); Mou Wang (Multimedia Product Center, Beijing OPPO Mobile Communication Co., Ltd.)

Sat-A-P-1-2: Revisiting Target Signal Definitions in Distortionless Superdirective Beamforming for Reverberant Speech Enhancement

Kang Chen (Wuhan University)*

Sat-A-P-1-3: FAE-DPF: Frequency-Attention Enhanced Neural Audio Codec with Transient Adaptability

zhu heng (Samsung)*

Sat-A-P-1-4: Improving Anomalous Sound Detection with Top-M Pseudo-Labeling

Zhang chen (昆山杜克大学)*; Zhang yucong (昆山杜克大学); Li ming (昆山杜克大学)

Sat-A-P-1-5: 基于词间时间感知和检索增强的阿尔兹海默症检测

mingyang gu (Tianjin University)*; zunsheng tan (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences); kai li (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences); bin wen (Tianjin University); tianrui wang (Tianjin University); longbiao wang (Tianjin University); jianwu dang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences)

Sat-A-P-1-6: Methods and Progress in Group Emotion Recognition: A Review

Xiaocui Zhou (Tianjin Normal University)*; Xuefei Liu (Tianjin Normal University); Jizhou Cui (ShanghaiTech University); Heng Xie (Beijing Institute of Technology); Zhengqi Wen (Department of Automation, BNRist, Tsinghua University); Ruibo Fu (Institute of Automation, Chinese Academy of Sciences); Yukun Liu (University of Chinese Academy of Sciences); Jianhua Tao (Department of Automation, BNRist, Tsinghua University); Ziping Zhao (Tianjin Normal University)

Sat-A-P-1-7: HuBERT and LLM Fusion for Suicide Risk Detection: A Multi-Modal Approach to Spontaneous Speech Analysis

Haiyang Ye (Jiangnan University)*; Yu Chengbo (Jiangnan University); Yihui Wang (Jiangnan University); Yu Fang (Jiangnan University); Yuxin Li (Jiangnan University)

Sat-A-P-1-8: Robust Keyword Spotting via Environment-Aware Speech Enhancement and Adaptive Modeling

Wen Wen (shanghai Jiao Tong University)*; Xiaoyu Gu (AISpeech Ltd); Haoyu Li (shanghai Jiao Tong University); Yu Xi (Shanghai jiaotong university); Qiang Zhou (AISpeech Ltd); Shizhang Tang (AISpeech Ltd); Minghang Chu (AISpeech Ltd); Ziqi Gong (AISpeech Ltd); Kai Yu (Shanghai jiaotong university)

Sat-A-P-1-9: 促音条件下 F0 轨迹中断对日语声调重音知觉的影响——日语母语者与汉语母语者的比较研究

刘玥 (上智大学)*

Sat-A-P-1-10: Bridging the Modality Gap: A Low-Cost Framework for Weakly Labeled TextMusic Datasets

Chao Wang (China Telecom Artificial Intelligence Technology Co.Ltd)*; Jie Lian (China Telecom Artificial Intelligence Technology Co.Ltd); Yuxin Zhang (China Telecom Artificial Intelligence Technology Co.Ltd); Jie Li (China Telecom Artificial Intelligence Technology Co.Ltd); Yongxiang Li (China Telecom Artificial Intelligence Technology Co.Ltd)

Sat-A-P-1-11: Emotional Resonance: Exploring the Emotional Effects of AI-Generated and Real Music

Cuiping ZHU (Beijing Institute of Technology); Lin Shen (Beijing Institute of Technology); Zhonghao Zhao (Beijing Institute of Technology); Ruobing Li (Beijing Institute of Technology); Haodong Lu (Beijing Institute of Technology); Jingwen Xu (Beijing Institute of Technology); Kun Qian (Beijing Institute of Technology)*; Bin Hu (Beijing Institute of Technology); Björn Schuller (Technical University of Munich); Yoshiharu Yamamoto (The University of Tokyo)

Sat-A-P-1-12: 多模态情感识别中的门控特征选择和专家建模

Weixiang Xu (Hunan University)*; Tengda Yin (Hunan University); Zixing Zhang (Hunan University)

Sat-A-P-1-13: A Two-Stage Band-Split Mamba-2 Network For Music Source Separation

Jinglin Bai (Inner Mongolia University)*; Yuan Fang (Inner Mongolia University); Jiajie Wang (SenseTime); Zixuan Xue (Inner Mongolia University); Xueliang Zhang (Inner Mongolia University)

Sat-A-P-1-14: 无离散音频编码注入的独立全双工对话大模型

Wenyi Yu (Tsinghua)*; Siyin Wang (Tsinghua); Chao Zhang (Tsinghua)

Sat-A-P-1-15: METC: Promoting Self-correcting LLMs In Speech Summarization

xuejian zhao (Xinjiang University)*; xiaozhe qi (Xinjiang University); hongfu jia (Xinjiang University); liang he (Tsinghua University)

Sat-A-P-1-16: EchoVoices: Preserving Generational Voices and Memories for Seniors and Children

HAIYING Xu (nankai)*; Haoze Liu (nankai); MingShi Li (nankai); Siyu Cai (nankai); GuangXuan Zheng (nankai); Yuhang Jia (nankai); Jinghua Zhao (nankai); Yong Qin (nankai)

Sat-A-P-1-17: Knowledge Augmented Finetuning Matters in Both RAG and Agent Based Dialog Systems

Yuxuan Wu (Tsinghua University); Yucheng Cai (Tsinghua University); Yi Huang (China Mobile Research Institute); Junlan Feng (China Mobile Research Institute); Zhijian Ou (Tsinghua University)*

Sat-A-P-1-18: 基于青年男性听音被试的非同时期普通话语音听觉辨别影响因素研究

WANG DANLIN (Beijing Police College)*; KE QINGQING (Changsha Tianxin District People's Court); CAO HONGLIN (Key Laboratory of Evidence Science (China University of Political Science and Law), Ministry of Education)

Sat-A-P-1-19: The Roles of Context and Prosody in Ironic Speech Comprehension: A Behavioral and ERP Investigation

胡涵 (南京师范大学)*; 顾文涛 (南京师范大学)

Sat-A-P-1-20: EEG as a Biometric Identifier: The Impact of Electrode Arrangement, Brain Areas, and Frequency Bands

Yiming Wang (University of Science and Technology of China); Jiaxin Chen (University of Science and Technology of China); Jiayang Han (University of Science and Technology of China); Ziyu Zhang (University of Science and Technology of China); Yin-Long Liu (University of Science and Technology of China); Rui Feng (University of Science and Technology of China); Xiuyuan Liang (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China); Jiahong Yuan (University of Science and Technology of China)*

Sat-A-P-1-21: Interface challenges in pitch accent processing among Chinese EFL learners: Neural and behavioral insights

纪晓丽 (Jiangsu University of Science and Technology)*; 王灿 (Jiangsu University of Science and Technology); 朱雨涵 (Jiangsu University of Science and Technology)

Sat-A-P-1-22: 一种基于双通道加权注意网络的听觉注意力检测方法研究

徐辉杭 (安徽大学); 李璐 (安徽大学); 李佳朕 (安徽大学); 董兴广 (安徽大学); 吕钊 (安徽大学); 范存航 (安徽大学)*

Sat-A-P-1-23: 汉语焦点重音感知的神经机制初探

YALEI LI (Nanjing Normal University)*; GU Wentao (南京师范大学); Zhang Wei (Nanjing University)

Sat-A-P-1-24: 六盘水方言单字调实验研究

李语 (西北民族大学)*

Sat-A-P-1-25: 山东高青方言塞音过渡音征研究

黄慧翔 (华中师范大学文学院)*; 程烨 (安徽大学新闻传播学院); 王师佳 (华中师范大学外国语学院)

Sat-A-P-1-26: 山东惠民方言两字连读变调实验研究

王瑞圆 (北京语言大学)*

Sat-A-P-1-27: 福州方言陈述和疑问语气边界调分析

Jiaqi Shi (ShiJiaqi)*

Sat-A-P-1-28: 盐城方言两字连读变调实验研究

王瑞圆 (北京语言大学)*

Sat-A-P-1-29: 莱州方言擦音声学空间格局研究

Xiaoqing Liu (曲阜师范大学)*; Wei Zhang (曲阜师范大学)

Sat-A-P-1-30: Effect of focus on pitch and duration in Shantou Min

Han Yu (College of Chinese Language and Culture, Jinan University)*; Maolin Wang (College of Chinese Language and Culture, Jinan University)

Sat-A-P-1-31: 同一声调系统内微降与直降、平调的对立模式研究——以西南官话六枝方言为例
左创 (兰州大学)*

Sat-A-P-1-32: 昌黎城关方言单字调的青年发音与青老年发音差异研究
李婧 (北京师范大学)*; 曹梦雪 (中国社会科学院语言研究所)

Sat-A-P-1-33: 基于离散表示的大规模语音生成技术综述
Wu Zhongyang (Guangxi Normal University for Nationalities)

Sat-E-FS: 基金分享沙龙活动
Room: 商务楼三楼·明都厅
Type: Fundings Session
October 18, 8:30 PM to 9:30 PM
Organizer: 刘瑞, 秦霄羽
Chair: 刘瑞老师

Sunday, October 19

Sun-M-PLN-1: VoxCPM: 面向高拟真语音生成的高效大模型架构探索与应用

Room: 研发楼 401

Type: Plenary

October 19, 9:00 AM to 9:45 AM

Presenter: 刘知远老师

Chair: 张劲松老师

Sun-M-PLN-2: 《机器合成普通话水平测评等级标准及测评大纲》研制与实践

Room: 研发楼 401

Type: Plenary

October 19, 9:45 AM to 10:30 AM

Presenter: 李慧老师

Chair: 张劲松老师

Sun-M-O-1: Automatic Speech Recognition

Room: 商务楼三楼-明都厅

Type: Oral

October 19, 10:45 AM to 12:15 PM

Chair: 秦勇老师

Sun-M-O-1-1: Baseline Systems for Chinese Continuous Visual Speech Recognition Challenge 2025

Zehua Liu (Beijing University of Posts and Telecommunications); Xiaolou Li (Beijing University of Posts and Telecommunications); Chen Chen (Tsinghua University); Lantian Li (Beijing University of Posts and Telecommunications); Dong Wang (Tsinghua University)*

Sun-M-O-1-2: Zero- and One-Shot Data Augmentation for Sentence-Level Dysarthric Speech Recognition in Constrained Scenarios

shiyao wang (Nankai University)*; shiwan zhao (Nankai University); jiaming zhou (Nankai University); yong qin (Nankai University)

Sun-M-O-1-3: Ideal-LLM: Integrating Dual Encoders and Language-Adapted LLM for Multilingual Speech-to-Text

Hongfei Xue (NWPU)*; Wei Ren (Chongqing Changan Automobile Co., Ltd); Xuelong Geng (NWPU); Kun Wei (NWPU); Longhao Li (NWPU); Qijie Shao (NWPU); Linju Yang (Chongqing Changan Automobile Co., Ltd); Kai Diao (Chongqing Changan Automobile Co., Ltd); Lei Xie (NWPU)

Sun-M-O-1-4: 基于端到端语音大模型的说话人日志与多人语音转写

Tianhua Li (Shanghai Jiao Tong University); Chenda Li (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)*

Sun-M-O-1-5: Unified Semantic-Speaker Encoding and Selective Reinforcement Learning for Target Speaker ASR

yinlong Zhang (Tsinghua)*; yiru Zhang (Xiaomi Inc.); hang Su (Xiaomi Inc.); lichun Fan (Xiaomi Inc.); jian Luan (Xiaomi Inc.); zhiyong Wu (Tsinghua)

Sun-M-O-1-6: Adaptive Multi-source Fusion for Uyghur ASR Error Correction

Qi Du (Xinjiang University)*; Ya Huang (Xinjiang University); Yongchao Li (Xinjiang University); Nurmemet Yolwas (Xinjiang University)

Sun-M-SF: Student Forum 1
Room: 研发楼 402
Type: Student Forum
October 19, 10:45 AM to 12:15 PM
Chair: 洪青阳老师

10:45-11:07
Sun-M-SF-1: 音色复刻与通用音频生成
Presenter: 江子越 (浙江大学)

11:07-11:29
Sun-M-SF-2: 面向声学场景事件检测的音频图建模表征策略
Presenter: 肖飞扬 (哈尔滨工程大学)

11:29-11:51
Sun-M-SF-3: 构音障碍语音识别中的发音态势特征提取与多特征协同激活
Presenter: 林羽钦 (福州大学)

11:51-12:15
Sun-M-SF-4: 面向情感支持对话的社交模拟研究
Presenter: 曹亚茹 (中科院深圳先进技术研究院)

Sun-M-YF: Young Forum 3
Room: 研发楼 401
Type: Young Forum
October 19, 10:45 AM to 12:15 PM
Chair: 陈霏老师

Sun-M-YF-1: 复杂声学场景下的多任务通用声源分离
Presenter: 李晨达老师, 上海交通大学听觉认知与计算声学实验室助理研究员

Sun-M-YF-2: 跨语言国际音标识别研究
Presenter: 王天恒老师, 南开大学文学院博士生

Sun-M-YF-3: 解耦驱动的音频离散化表征及多任务统一构建
Presenter: 张世磊老师, 中国移动九天人工智能研究院语音技术负责人, 北京大学多媒体信息处理全国重点实验室研究员, 中关村试听产业技术创新联盟第三届理事会理事

Sun-M-P-1: Speech Coding and Speech Enhancement
Room: 研发楼四楼-海报区
Type: Poster
October 19, 10:45 AM to 12:15 PM

Sun-M-P-1-1: 融合相位估计的声码器语音增强算法

Wen Zhang (Hangzhou Dianzi University)*; Wenbin Jiang (Hangzhou Dianzi University); Yang Zhang (Hangzhou Dianzi University); Kaiying Wu (Hangzhou Dianzi University); Xuanhao Cai (Hangzhou Dianzi University)

Sun-M-P-1-2: 结合信号处理与神经网络的回声消除研究

zhan zongyi zhan zongyi (xiamen university)*

Sun-M-P-1-3: MelGenSE: Generative Speech Enhancement on Mel-Spectrogram via Conditional Flow Matching

Yang Zhang (Hangzhou Dianzi University)*; Wenbin Jiang (Hangzhou Dianzi University); Kaiying Wu (Hangzhou Dianzi University); Wen Zhang (Hangzhou Dianzi University)

Sun-M-P-1-4: 基于人耳掩蔽效应的下行语音增强

Yuanwu Liu (Inner Mongolia University)*; Zixuan Xue (Inner Mongolia University); Xueliang Zhang (Inner Mongolia University)

Sun-M-P-1-5: Cross-domain Challenges in Speech Enhancement: A Solution Driven by WavLM

Yating Zhang (Dalian Maritime University); Yilin Pan (Dalian Maritime University)*; Yijia Zhang (Dalian Maritime University); Mingyu Lu (Dalian Maritime University)

Sun-M-P-1-6: MT-SkiM: Skipping memory Transformer for speech separation

Xin Zhou (Shanghai Jiao Tong University)*; Chenda Li (Shanghai Jiao Tong University); Yanmin Qian (Shanghai Jiao Tong University)

Sun-M-P-1-7: Implementation and Analysis of Neural Network Quantization in Speech Enhancement

Mou Wang (Institute of Acoustics)*; Zhuo Li (OPPO); Jinfu Wang (OPPO); Xin Liu (OPPO); Feiran Yang (Institute of Acoustics)

Sun-M-P-1-8: 基于信道信噪比自适应语义通信的远程语音增强方法

Haiyan Guo (Nanjing University of Posts and Telecommunications)*

Sun-M-P-1-9: Taming Latents in Your VQ-VAE

Zhihan Li (Shanghai Jiao Tong University)*; Yiwei Guo (Shanghai Jiao Tong University); Bohan Li (Shanghai Jiao Tong University); Hankun Wang (Shanghai Jiao Tong University); Xie Chen (Shanghai Jiao Tong University); Kai Yu (Shanghai Jiao Tong University)

Sun-M-P-1-10: Extract and Diffuse: Latent Integration for Improved Diffusion-based Speech and Vocal Enhancement

Yudong Yang (Tsinghua University)*; Zhan Liu (Tsinghua University); Wenyi Yu (Tsinghua University); Guangzhi Sun (University of Cambridge Department of Engineering); Qiuqiang Kong (CUHK); Chao Zhang (Tsinghua University)

Sun-M-P-1-11: Dual-Path Net with Full-Band Spectral Modeling for Monaural Noisy-Reverberant Speech Enhancement

Guowu Tan (Beijing Institute of Technology)*; Xiang Xie (Beijing Institute of Technology); Zhaoqi Cui (Beijing Institute of Technology); Xinyu Yang (Beijing Institute of Technology)

Sun-M-P-1-12: HiStyle: Hierarchical Style Embedding Prediction for Text-Prompt-Guided Controllable Speech Synthesis

Ziyu ZHANG (Northwestern Polytechnical University)*; Hanzhao Li (Northwestern Polytechnical University); Jingbin Hu (Northwestern Polytechnical University); Wenhao Li (Northwestern Polytechnical University); Lei Xie (Northwestern Polytechnical University)

Sun-M-P-1-13: 多阶段微调 HuBERT 构音障碍语音修复系统研究

蔡坤良 (深圳大学)*; 郑能恒 (深圳大学)

Sun-M-P-1-14: Re-Sonance: A Dysarthric Asynchronous Real-Time Speech Conversion System Based on a Three-Stage Cascaded ASR-LLM-TTS Architecture

Yuxuan Wu (Southeast University)*; Yifan Xu (Southeast University); Junkun Wang (Southeast University); Jiayong Jiang (Southeast University); Xin Zhao (Southeast University); Zhaojie Luo (Southeast University)

Sun-M-P-1-15: Enhancing Teochew Speech Synthesis via Two-Stage Fine-Tuning of CosyVoice 2

Shuang Yao (Beijing Language and Culture University)*; Dengfeng Ke (Beijing Language and Culture University); Jinsong Zhang (Beijing Language and Culture University)

Sun-M-P-1-16: DiffEditor: Enhancing Speech Editing with Semantic Enrichment and Acoustic Consistency

Yang Chen (Nankai University)*; Yuhang Jia (Nankai University); Shiwan Zhao (Nankai University); Haoran Li (Nankai University); Jiarong Kang (Nankai University); yong qin (Nankai University)

Sun-M-P-1-17: 面向表现力语音合成的语音高层次理解方法

kaifeng yun (Tsinghua University)*; zeyu jin (Tsinghua University); xiaoyu qin (Tsinghua University); jia jia (Tsinghua University); zhiyong wu (Tsinghua University)

Sun-M-P-1-18: Research on Vision Accessible System for Blind and Low-vision People

Qixin Wang (Tsinghua University)*; Songtao Zhou (Tsinghua University); Zeyu Jin (Tsinghua University); Shikun Sun (Tsinghua University); Qinghuan Lan (Tsinghua University); Chenglin Guo (New York University); Xiaoyu Qin (Tsinghua University)

Sun-M-P-1-19: Temporally Consistent Teeth Restoration for Talking Heads

Tianze Zhou (Tsinghua University)*; Zhiyong Wu (Tsinghua University)

Sun-M-P-1-20: EGGLab: A Praat-based program for Electroglossography signal processing

Van Dat Ta (Peking University)*

Sun-M-P-1-21: 基于语料库的乌尔都语依存距离研究

zhu qi (国防科技大学外国语学院)*

Sun-M-P-1-22: When LLMs Encounter Human in the Annotation of Psychiatric Data: Decent Yet Long Way to Go

Tianyi Sun (Shanghai Jiao Tong University); Kunyao Lan (Shanghai Jiao Tong University); Yanli Ding (Shanghai Mental Health Center); Yiming Yan (Shanghai Mental Health Center); Yunyun Han (Tianqiao and Chrissy Chen Institute); Tao Yang (Tianqiao and Chrissy Chen Institute); Haiyang Geng (Tianqiao and Chrissy Chen Institute); Mengyue Wu (Shanghai Jiao Tong University)*; Jianhua Chen (Shanghai Mental Health Center)

Sun-M-P-1-23: A Lightweight Multimodal Model for Mandarin Retelling Scoring

吕雅轩 (吕雅轩); Aishan Wumaier (Xinjiang University)*

Sun-M-P-1-24: Evaluating the Effects of Low-Rank Adaptation on Low-Resource Urdu Language ASR Benchmarks

Muhammad Sharif (Institute of Automation Chinese Academy of Sciences); Zeeshan Abbas (College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China); Sardar Shan Ali Naqvi (School of artificial intelligence and automation, Huazhong University of Science and Technology); Bin Liu (Institute of Automation Chinese Academy of Sciences)*; Cheng-Lin Liu (Institute of Automation Chinese Academy of Sciences)

Sun-M-P-1-25: Study of the focal effect on pitch in Dongguan Cantonese

Ziyu Xiong (Key Laboratory of Linguistics, Chinese Academy of Social Sciences, Beijing, China)*; Maolin Wang (College of Chinese Language and Culture, Jinan University, Guangzhou, China)

Sun-M-P-1-26: Dialect-Specific Modes Influence Second Language Production: Evidence From Bidialectal Shanghai-Mandarin Chinese Learners of English Within the Second Language Linguistic Perception Model

Xiaoluan Liu (East China Normal University)*

Sun-M-P-1-27: Dementia Detection via Speech Temporal Sequences with Shifted Windows

Yue Wang (Dalian Maritime University); Yilin Pan (Dalian Maritime University)*; Yijia Zhang (Dalian Maritime University); Mingyu Lu (Dalian Maritime University)

Sun-M-P-1-28: 融合 LLM 解码的构音障碍语音识别模型

李浩鹏 (深圳大学)*; 颜森 (深圳大学); 郑能恒 (深圳大学)

Sun-M-P-1-29: 老龄化对化州粤语升调产出和感知的影响研究

Meier Wang (Foshan Open University)*; Qiudong Li (Foshan Open University); Qin Wang (Foshan Open University)

Sun-M-P-1-30: 基于共振峰参数与唇形特征的普通话四呼无监督聚类分析

Liwen Jing (中山大学)*; Bei Yang (中山大学)

Sun-M-P-1-31: Prosodic Speech Rate and Auditory Word Valence: Competition with Facial Emotion in Working Memory Processing

Jieyu Chen (The Hong Kong Polytechnic University)*; Yujia Tian (The Hong Kong Polytechnic University); Jing Qi (The Hong Kong Polytechnic University); Ran Tao (The Hong Kong Polytechnic University)

Sun-M-P-1-32: Age-related differences in production cues of vowel length contrast: the case of Northern Vietnamese

Van Dat Ta (Peking University)*

Sun-M-P-1-33: 基于互信息的粤语声调功能负载研究

Xiaolin Wu (Beijing Language and Culture University); Xiaowang Liu (Beijing Language and Culture University); Jinsong Zhang (Beijing Language and Culture University)*; Wenfa Bai (Beijing Language and Culture University); Dengfeng Ke (Beijing Language and Culture University); Yanlu Xie (Beijing Language and Culture University)

Sun-M-P-1-34: 越南语声调音位功能负载研究

Phat Bach (Chinese Language and Chinese Studies Department, School of Information Science, Beijing Language and Culture University); Xiaolin Wu (School of Information Science, Beijing Language and Culture University); Jinsong Zhang (School of Information Science, Beijing Language and Culture University)*

Sun-M-P-1-35: 基于互信息的粤语声母功能负载研究

Xiaolin Wu (Beijing Language and Culture University)*; Jinsong Zhang (Beijing Language and Culture University)

Sun-M-P-1-36: 播音员与普通发音人的轻声音节噪音特征对比分析

胡明夏 (北京语言大学); Jinsong Zhang (Beijing Language and Culture University)*

Sun-M-P-1-37: The Phonetic Modification and Facial Movements Made During Mandarin Vowel and Tone Production in Noise

曹幸怡 (南京理工大学)*

Sun-A-O-1: Speech Technology**Room:** 商务楼三楼-明都厅**Type:** Oral**October 19, 1:15 PM to 2:45 PM****Chair:** 刘李老师**Sun-A-O-1-1: 基于并行幅度和相位谱估计的神经网络语音分离**

Fei Liu (University of Science and Technology of China)*; Yang Ai (University of Science and Technology of China); Zhen-Hua Ling (University of Science and Technology of China)

Sun-A-O-1-2: 大语言模型和采样边缘化驱动的基于音素的语音识别

Te Ma (Xinjiang University); Nanjie Li (Xinjiang University); Hao Huang (Xinjiang University); Zhijian Ou (Tsinghua University)*

Sun-A-O-1-3: ConSeCodec: Semantically Robust Tokens via Consistency Training

Guanyu Chen (Shanghai Jiao Tong University)*; Yiwei Guo (Shanghai Jiao Tong University); Kai Yu (Shanghai Jiao Tong University)

Sun-A-O-1-4: REA-TTS: Retrieval-Augmented Expressive Audiobook Text-to-Speech Generation with Contrastive Language-Audio Learning

Qihang Lu (Beijing University of Posts and Telecommunications)*; Bingsong Bai (Beijing University of Posts and Telecommunications); Jinlong Xue (Beijing University of Posts and Telecommunications); Ya Li (Beijing University of Posts and Telecommunications); Yingming Gao (Beijing University of Posts and Telecommunications)

Sun-A-O-1-5: ADFMoE: A Universal Audio Deepfake Detection Model for Speech and Singing Voice

Shengbo Cai (Shenzhen International Graduate School, Tsinghua University)*; Zhiyong Wu (Shenzhen International Graduate School, Tsinghua University)

Sun-A-SF: Student Forum 2**Room:** 研发楼 402**Type:** Student Forum**October 19, 1:15 PM to 2:45 PM****Chair:** 万继康老师

13:15-13:38

Sun-A-SF-1: 球面谐波域的选择性空间主动噪声控制系统

Presenter: 张华为 (澳大利亚国立大学)

13:38-14:10

Sun-A-SF-2: 利用机器特定滤波器组提取的时频调制特征进行机械异常声音检测

Presenter: 李凯 (中国科学院深圳先进技术研究院)

14:10-14:32

Sun-A-SF-3: Integration of Old and New Knowledge for Generalized Intent Discovery: A Consistency-driven Prototype-Prompting Framework

Presenter: 魏笑 (天津大学)

14:32-14:45

Sun-A-SF-4: 大模型时代的语音修复新范式：从“无所不能”到“深得人心”

Presenter: 张峻安 (香港中文大学 (深圳))

Sun-A-SS-1: 人机语音交互中的多模态智能建模与个性化感知

Room: 研发楼 401

Type: Special Session

October 19, 1:15 PM to 2:45 PM

Organizer: 沈莹老师

Chair: 沈莹老师

Sun-A-SS-1-1: 面向嘈杂环境的多声源音视频导航

Presenter: 师展博老师 (同济大学)

Sun-A-SS-1-2: 基于 Fourier-KAN 的连续音频信号表征

Presenter: 李林飞老师 (同济大学)

Sun-A-SS-1-3: 基于多模态语义线索的对话情绪原因抽取

Presenter: 梁乔老师 (同济大学)

Sun-A-SS-1-4: 基于深度学习逆卷积的房间冲激响应盲估计

Presenter: 陆嘉骐老师 (同济大学)

Sun-A-SS-1: 2025 CCF 先进音频技术竞赛

Room: 研发楼 308

Type: Special Session

October 19, 1:15 PM to 2:45 PM

Organizer: 卜辉老师, 凌震华老师, 朱梦尧老师

Chair: 卜辉老师

Sun-A-O-2: Spoof Speech Detection & Others

Room: 商务楼三楼-明都厅

Type: Oral

October 19, 3:00 PM to 4:30 PM

Chair: 李晨达老师

Sun-M-O-2-1: A Hierarchical Fusion Modeling from Perception to Prediction with Personalized Features for Multimodal Depression Detection

yingying zhou (Beijing University of Posts and Telecommunications)*; han chi (Beijing University of Posts and Telecommunications); jingyao xue (Beijing University of Posts and Telecommunications); yiming gao (Beijing University of Posts and Telecommunications); minchi hu (Beijing University of Posts and Telecommunications); yuhua wen (Beijing University of Posts and Telecommunications); qifei li (Beijing University of Posts and Telecommunications); yingming gao (Beijing University of Posts and Telecommunications); ya li (Beijing University of Posts and Telecommunications)

Sun-M-O-2-2: Speech Synthesis with Low-quality Data Using Noise-Robust Diffusion Autoencoders

Yuheng Lu (Tianjin University)*; Qiuyu Liu (Tianjin University); Cheng Gong (Institute of Artificial Intelligence (TeleAI), China Telecom); Chunyu Qiang (Tianjin University); Tianrui Wang (Tianjin University); Haoyu Wang (Tianjin University); Yu Jiang (Tianjin University); Xixin Wu (The Chinese University of Hong Kong); Xuanchen Li (Tianjin University); Longbiao Wang (Tianjin University); Jianwu Dang (Chinese Academy of Sciences)

Sun-M-O-2-3: Traceable TTS: Toward Watermark-Free TTS with Strong Traceability

Yuxiang Zhao (Shanghai Jiao Tong University)*; Yunchong Xiao (Shanghai Jiao Tong University); Yushen Chen (Shanghai Jiao Tong University); Zhikang Niu (Shanghai Jiao Tong University); Shuai Wang (School of Intelligence Science and Technology, Nanjing University); Kai Yu (Shanghai Jiao Tong University); Xie Chen (Shanghai Jiao Tong University)

Sun-M-O-2-4: Audio Deepfake Detection via Dual Branch Classifier with Self-Supervised Pre-Trained Model

Yuxuan Cao (Jiangsu University); Lijian Gao (Jiangsu University); Qirong Mao (Jiangsu University)*

Sun-M-O-2-5: 基于视觉大语言模型的局部伪造音频检测

王蓉 (中国科学院自动化研究所)*; 顾浩 (中国科学院自动化研究所); 何佳毅 (中国科学院自动化研究所); 易江燕 (清华大学); 陶建华 (清华大学); 李冠君 (中国科学院自动化研究所); 温正棋 (清华大学); 刘斌 (中国科学院自动化研究所); 曾思丁 (中国科学院自动化研究所)

Sun-M-O-2-6: 基于改进低阶适应的开放词汇语音情感识别

Kangwei LIU (Jiangsu University)*; Hanyu DING (Jiangsu University); Lijian GAO (Jiangsu University); Qirong MAO (Jiangsu University)

Sun-A-O-3: Speech, Paralinguistics and Language Analysis

Room: 研发楼 402

Type: Oral

October 19, 3:00 PM to 4:30 PM

Chair: 高迎明老师

Sun-A-O-3-1: When AI Speaks, Do We Follow? Phonetic Entrainment in Human-AI Dialogues

Qixin Li (BNU)*; Haocheng Lu (北京师范大学文学院); Gaowu Wang (北京师范大学文学院)

Sun-A-O-3-2: Exploring Gender Bias in Alzheimer' s Disease Detection: Insights from Mandarin and Greek Speech Perception

Liu He (University of Science and Technology of China)*; Yuanchao Li (University of Edinburgh); Rui Feng (University of Science and Technology of China); Xinran Han (University of Minnesota); Yin-long Liu (University of Science and Technology of China); Yuwei Yang (Jiangsu Normal University); Zude Zhu (Jiangsu Normal University); Jiahong Yuan (University of Science and Technology of China)

Sun-A-O-3-3: Cross-cultural Emotion Recognition by Chinese EFL learners Based on AI Voice

zhangchenxi zhu (Institute of Linguistics, Chinese Academy of Social Sciences)*

Sun-A-O-3-4: Neural Scoring: A Refreshed End-to-End Approach for Speaker Verification in Complex Conditions

Wan Lin (CSLT, BNRist, Tsinghua University); Junhui Chen (Beijing University of Posts and Telecommunications); Tianhao Wang (Beijing University of Posts and Telecommunications); Zhenyu Zhou (Beijing University of Posts and Telecommunications); Lantian Li (Beijing University of Posts and Telecommunications); Dong Wang (CSLT, BNRist, Tsinghua University)*

Sun-A-O-3-5: Adapting LLMs for Speech-Based Suicidal Risk Detection in The 1st SpeechWellness Challenge

Xuanchen Li (Tianjin University)*; Chenrui Cui (Tianjin University); Tianrui Wang (Tianjin University); Zikang Huang (Tianjin University); Yu Jiang (Tianjin University); Yuheng Lu (Tianjin University); Longbiao Wang (Tianjin University); Jianwu Dang (Chinese Academy of Sciences)

Sun-A-O-3-6: Serial-Parallel Dual-Path Architecture for Speaking Style Recognition

Guojian Li (Northwestern Polytechnical University)*; Qijie Shao (Northwestern Polytechnical University); Zhixian Zhao (Northwestern Polytechnical University); Shuiyuan Wang (Northwestern Polytechnical University); Zhonghua Fu (Northwestern Polytechnical University); Lei Xie (Northwestern Polytechnical University)

Sun-A-O-4: Audio Processing

Room: 研发楼 401

Type: Oral

October 19, 3:00 PM to 4:30 PM

Chair: 李明老师

Sun-A-O-4-1: Automatic Speech Evaluation Method Leveraging Deep Feature Fusion

XuDong Pang (新疆大学); Wenwen Lu (新疆大学); Silajiaihemaiti Silajiaihemaiti (新疆大学); Aishan Wumaier (新疆大学)*

Sun-A-O-4-2: Aishell1Mix: Towards Robust Mandarin Speech Separation with Scalable Audio Language Models

Zijian Huang (THBI)*

Sun-A-O-4-3: 昆曲唱词歌声合成的深度学习方法与优化

Huali Zhou (Wuhan University); Ming Li (Duke Kunshan University)*

Sun-A-O-4-4: StreamFlow: Streaming Flow Matching with Block-wise Guided Attention Mask for Speech Token Decoding

Dake Guo (Northwestern Polytechnical University)*; Jixun Yao (npu); Linhan Ma (npu); He Wang (npu); Lei Xie (npu)

Sun-A-O-4-5: Study of the Low-Rank Minimum Variance Distortionless Response Beamformer for Speech Enhancement

Hanchen Pei (Wuhan University)*

Sun-A-P-1: Automatic Speech Recognition

Room: 研发楼四楼-海报区

Type: Poster

October 19, 15:00 PM to 16:30 PM

Sun-A-P-1-1: Curriculum Reinforcement Learning for Robust Low-Resource Chinese Dialect Speech Recognition

Leijing Hou (China Mobile Research Institute)*; Leijing Hou (China Mobile Research Institute); Yingying Gao (China Mobile Research Institute); Min Liu (China Mobile Research Institute); Huihui Li (China Mobile Research Institute); Xiqing Zhang (China Mobile Research Institute); Nannan Wang (China Mobile Research Institute); Shilei Zhang (China Mobile Research Institute)

Sun-A-P-1-2: NC-KWS: Few-Shot Class-Incremental Keyword Spotting Based on Neural Collapse

Jin Li (Hangzhou Dianzi University)*; Wenbin Jiang (Hangzhou Dianzi University); Yitao Tian (Hangzhou Dianzi University); Zhuoyang Li (Hangzhou Dianzi University)

Sun-A-P-1-3: Environment-Aware Robust Keyword Spotting in Realistic Scenarios

Xiaoyu Gu (aispeech)*; Wen Wen (shanghai Jiao Tong University); Haoyu Li (shanghai Jiao Tong University); Yu Xi (Shanghai jiaotong university); Qiang Zhou (AISpeech Ltd); Shizhang Tang (AISpeech Ltd); Minghang Chu (AISpeech Ltd); Ziqi Gong (AISpeech Ltd); Kai Yu (Shanghai jiaotong university)

Sun-A-P-1-4: 一种端到端音视频理解大语言模型

Changli Tang (Tsinghua University)*; Chao Zhang (Tsinghua University)

Sun-A-P-1-5: 跨域说话人验证技术综述

Pan zhongjia (Xiamen University)*; Li lin (Xiamen University); Hong qingyang (Xiamen University)

Sun-A-P-1-6: BiAAT-ICGF: A Bidirectional Attention Aggregation and Iterative Graph Fusion Framework for Enhanced Multimodal Emotion Recognition

Zhangcheng Yang (College of Information Science and Technology, Donghua University)*; Xuebin Zhang (College of Information Science and Technology, Donghua University); Longting Xu (College of Information Science and Technology, Donghua University)

Sun-A-P-1-7: 基于 Whisper 的藏语方言语音识别研究

Like Ma (Northwest Minzu University), Guanyu Li (Northwest Minzu University)*

Sun-A-P-1-8: Multilevel and Granular L2 Pronunciation Assessment Using Stress-Based Suprasegmental Features and Proficiency Adaptation

wenqian Bao (Beijing Language and Culture University)*; jingsong zhang (Beijing Language and Culture University)

Sun-A-P-1-9: 复杂声学条件下音频-文本对比模型的鲁棒性研究

Shuai Zhang (Northwestern Polytechnical University)*; Xiaolei Zhang (Northwestern Polytechnical University)

Sun-A-P-1-10: Orthogonal Disentanglement with Projected Feature Alignment for Multimodal Emotion Recognition in Conversation

Xinyi Che (Sichuan University); Wenbo Wang (Harbin Institute of Technology); Jian Guan (Harbin Engineering University)*; Qijun Zhao (Sichuan University)

Sun-A-P-1-11: HACT-Net: Hierarchical Alignment and Cross-Modal Transformer for Personality Recognition

Hanzhe Xu (Tianjin Normal University)*; Xuefei Liu (Tianjin Normal University); Jizhou Cui (ShanghaiTech University); Heng Xie (Beijing Institute of Technology); Zhengqi Wen (Beijing National Research Center for Information Science and Technology); Ruibo Fu (Institute of Automation, Chinese Academy of Sciences); Yukun Liu (University of Chinese Academy of Sciences); Jianhua Tao (Department of Automation, BNRist, Tsinghua University); Ziping Zhao (Tianjin Normal University)

Sun-A-P-1-12: 一种理解空间音频的多模态大语言模型

Changli Tang (Tsinghua University)*; Chao Zhang (Tsinghua University)

Sun-A-P-1-13: MELON: Multimodal Emotion Understanding with LLMs

Jing Peng (Tsinghua University, Hunan University)*; Yaoxun Xu (Tsinghua University); Yunrui Cai (Tsinghua University); Zixing Zhang (Hunan University); Zhiyong Wu (Tsinghua University)

Sun-A-P-1-14: 鼻音化的频谱特征与口元音鼻音程度研究

Chu Zhou (Ningbo University)*; Yanzhi Yang (Coventry University)

Sun-A-P-1-15: 基于梯度一致的说话人识别系统迁移攻击

Mengyao Liu (Northwestern Polytechnical University)*; Xiaolei Zhang (Northwestern Polytechnical University)

Sun-A-P-1-16: Exploring the Safety of Speech Tokenizer for Privacy Protection against Illegal LLM-based Speech Synthesis

Zhisheng Zhang (Tsinghua University); Zhiyong Wu (Tsinghua University)*

Sun-A-P-1-17: SpkGuard : Backdoor Defense for Speaker Recognition Based on Adversarial Perturbation

tuan fang (Ningbo University); Diqun Yan (Ningbo University of Finance and Economics)*

Sun-A-P-1-18: 隐性模仿情境下的普通话声调趋同

Xiyao Lu (Beijing Language and Culture University, SAIT Lab); Qian Li (Chinese Academy of Social Sciences, Institute of Linguistics)*

Sun-A-P-1-19: PQSR: A Speech Corpus of Polar Questions and Spontaneous Responses in Standard Chinese with Complex Intentions Annotated

Yingyi Luo (Institute of Linguistics, Chinese Academy of Social Sciences)*; Yue Huang (Institute of Linguistics, Chinese Academy of Social Sciences); Qingke Sun (Institute of Linguistics, Chinese Academy of Social Sciences); Shuwen Chen (Institute of Linguistics, Chinese Academy of Social Sciences)

Sun-A-P-1-20: 用于自然梯度法的改进的经验费舍信息矩阵近似

Wenyi Yu (Tsinghua)*; Chao Zhang (Tsinghua)

Sun-A-P-1-21: 面向无监督聚类的渐进优化策略研究

wu heng (厦门大学)*

Sun-A-P-1-22: IDKD:Further Improve the Target Speaker and Non-target Speaker Separation in Knowledge Distillation for Automatic Speaker Verification

maosheng chen (shanghaijiaotong university)*

Sun-A-P-1-23: 声纹识别中的池化技术综述

陈永鑫 (陈永鑫)*

Sun-A-P-1-24: 普通话单元音前五条共振峰的话者区分能力研究

Yijing Lei (China University of Political Science and Law)*; Honglin Cao (China University of Political Science and Law)

Sun-A-P-1-25: NE-PADD: Leveraging Named Entity Knowledge for Robust Partial Audio Deepfake Detection via Attention Aggregation

Huhong Xian (Inner Mongolia University)*; Rui Liu (Inner Mongolia University); Berrak Sisman (Johns Hopkins University); Haizhou Li (The Chinese University of Hong Kong)

Sun-A-P-1-26: ISSE: Informative Sample Selection via Experts for Unsupervised Domain Adaptation in Partially Deepfake Audio Localization

Siding Zeng (Institute of Automation, Chinese Academy of Sciences)*; Jiangyan Yi (Department of Automation, Tsinghua University); Jiayi He (Institute of Automation, Chinese Academy of Sciences); Guan Jun Li (Institute of Automation, Chinese Academy of Sciences); Zhenqi Wen (Department of Automation, Tsinghua University); Jianhua Tao (Department of Automation, BNRist, Tsinghua University)

Sun-A-P-1-27: A Multi-Subspace Attention Approach for Robust Speech Spoofing Detection in Silence-Trimming Conditions

Wang Haoyang (Anhui University)*; Jian Zhou (Anhui University); Hon Keung Kwan (Windsor University)

Sun-A-P-1-28: DWMHA: Dynamic Window Multi-Head Attention for Speech Emotion Recognition

yongqi Liu (Xinjiang University); Hankiz Yilahun (Xinjiang University)*; Askar Hamdulla (Xinjiang University); Gulnur Arkin (Xinjiang University of Finance and Economics)

Sun-A-P-1-29: Speech BandWidth Extension Incorporated with Speaker Characteristics

pan xu (Northwestern Polytechnical University)*; zhongyu zhang (iFLYTEK Xi'an Superbrain Information Technology Co., Ltd.); zhonghua fu (Northwestern Polytechnical University;iFLYTEK Xi'an Superbrain Information Technology Co., Ltd.)

Sun-A-P-1-30: UniDaugMamba: A Unimodal Data-augmented Mamba for Speech-Based Depression Detection

Han Chi (Beijing University of Posts and Telecommunications)*; Tianze Xu (Beijing University of Posts and Telecommunications); Yingying Zhou (Beijing University of Posts and Telecommunications); Ya Li (Beijing University of Posts and Telecommunications); Yingming Gao (Beijing University of Posts and Telecommunications)

Sun-A-P-1-31: 基于大模型的汉语韵律边界检测研究

韩梓萌 (北京语言大学)*; 柯登峰 (北京语言大学); 张劲松 (北京语言大学); 解焱陆 (北京语言大学)

Organizer List

名誉主席 (Honorary Chair)

- 俞凯 - 上海交通大学
- 艾斯卡尔·艾木都拉 - 新疆大学

大会主席 (Conference Chairs)

- 郑方 - 清华大学
- 张劲松 - 北京语言大学
- 毛启容 - 江苏大学

程序委员会主席 (Program Committee Chairs)

- 贾珈 - 清华大学
- 吴志勇 - 清华大学深圳研究生院
- 高利剑 - 江苏大学

组织委员会主席 (Organizing Committee Chairs)

- 熊子瑜 - 中国社科院语言研究所
- 陈锦富 - 江苏大学
- 施静 - 江苏大学

工业联络主席 (Industrial Liaison Chairs)

- 谢磊 - 西北工业大学
- 陈谐 - 上海交通大学
- 朱璇 - 三星电子中国研究院
- 武执政 - 香港中文大学 (深圳)
- 成科扬 - 江苏大学

宣传主席 (Publicity Chairs)

- 凌震华 - 中国科学技术大学
- 卜辉 - 语音之家
- 柳娜 - 江苏大学

出版主席 (Publication Chairs)

- 黄公平 - 武汉大学
- 李雅 - 北京邮电大学
- 贾洪杰 - 江苏大学

科普教育主席 (Educational Chairs)

- 王东 - 清华大学
- 徐明星 - 清华大学
- 朱璇 - 三星电子中国研究院

教程报告主席 (Tutorial Chairs)

- 韩纪庆 - 哈尔滨工业大学
- 李美竹 - 江苏大学

特殊议题主席 (Special Session Chairs)

- 王东 - 清华大学
- 张晓雷 - 西北工业大学
- 范存航 - 安徽大学

青年论坛主席 (Young Forum Chairs)

- 欧智坚 - 清华大学
- 李倩 - 中国社会科学院语言研究所
- 毕凡 - 江苏大学

基金分享论坛主席 (Fund Forum Chairs)

- 刘瑞 - 内蒙古大学
- 秦霄羽 - 清华大学

学生论坛主席 (Student Forum Chairs)

- 洪青阳 - 厦门大学
- 何文文 - 江苏大学
- 万继康 - 江苏大学

财务主席 (Financial Chairs)

- 张超 - 清华大学
- 徐佩锋 - 江苏大学