

IEEE Transactions on Audio, Speech and Language Processing

Title  
History

Home Popular Early Access Current Volume All Volumes About Journal

Popular Documents - April 2025

Popular Article Feed

Includes the 50 most frequently accessed documents for this publication.

Export Email Selected Results

Showing 1-50 of 50

Date

April 2025



- ☐ **Multi-Domain Emotion Recognition Enhancement: A Novel Domain Adaptation Technique for Speech-Emotion Recognition**
- Ammar Amjad; Sucharita Khuntia; Hsien-Tsung Chang; Li-Chia Tai
- Publication Year: 2025 , Page(s): 528 - 541
- Abstract

HTML
- ☐ **Neural Codec Language Models are Zero-Shot Text to Speech Synthesizers**
- Sanyuan Chen; Chengyi Wang; Yu Wu; Ziqiang Zhang; Long Zhou; Shujie Liu; Zhuo Chen; Yanqing Liu; Huaming Wang; Jinyu Li; Lei He; Sheng Zhao; Furu Wei
- Publication Year: 2025 , Page(s): 705 - 718
- Cited by: [Papers \(14\)](#)
- Abstract

HTML
- ☐ **Personalized and Controllable Voice Style Transfer With Speech Diffusion Transformer**
- Ha-Yeong Choi; Sang-Hoon Lee; Seong-Wan Lee
- Publication Year: 2025 , Page(s): 922 - 934
- Abstract

HTML
- ☐ **Separate Anything You Describe**
- Xubo Liu; Qiuqiang Kong; Yan Zhao; Haohe Liu; Yi Yuan; Yuzhuo Liu; Rui Xia; Yuxuan Wang; Mark D. Plumbley; Wenwu Wang
- Publication Year: 2025 , Page(s): 458 - 471
- Cited by: [Papers \(2\)](#)
- Abstract

HTML
- ☐ **Vec-Tok Speech: Speech Vectorization and Tokenization for Neural Speech Generation**
- Xinfu Zhu; Yuanjun Lv; Yi Lei; Tao Li; Wendi He; Hongbin Zhou; Heng Lu; Lei Xie
- Publication Year: 2025 , Page(s): 1243 - 1254
- Cited by: [Papers \(1\)](#)
- Abstract

HTML
- ☐ **Adaptive Multimodal Graph Integration Network for Multimodal Sentiment Analysis**
- Peizhu Gong; Jin Liu; Xiliang Zhang; Xingye Li; Lai Wei; Huihua He
- Publication Year: 2025 , Page(s): 23 - 36
- Abstract

HTML
- ☐ **Acoustic Prompt Tuning: Empowering Large Language Models With Audition Capabilities**
- Jinhua Liang; Xubo Liu; Wenwu Wang; Mark D. Plumbley; Huy Phan; Emmanouil Benetos
- Publication Year: 2025 , Page(s): 949 - 961
- Cited by: [Papers \(1\)](#)



✓ Abstract [HTML](#)  

☐ **Memory-Tuning: A Unified Parameter-Efficient Tuning Method for Pre-Trained Language Models** 

Wang Qi; Rui Liu; Yuan Zuo; Fengzhi Li; Yong Chen; Junjie Wu

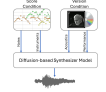
Publication Year: 2025 , Page(s): 1 - 10

✓ Abstract [HTML](#)  

☐ **Multi-Aspect Conditioning for Diffusion-Based Music Synthesis: Enhancing Realism and Acoustic Control** 

Ben Maman; Johannes Zeitler; Meinard Müller; Amit H. Bermano

Publication Year: 2025 , Page(s): 68 - 81



✓ Abstract [HTML](#)   

☐ **The Codefake Dataset and Countermeasures for the Universally Detection of Deepfake Audio** 

Yuankun Xie; Yi Lu; Ruibo Fu; Zhengqi Wen; Zhiyong Wang; Jianhua Tao; Xin Qi; Xiaopeng Wang;

Yukun Liu; Haonan Cheng; Long Ye; Yi Sun

Publication Year: 2025 , Page(s): 386 - 400

Cited by: [Papers \(1\)](#)

✓ Abstract [HTML](#)  

☐ **Zero-Shot Cross-Lingual Text-to-Speech With Style-Enhanced Normalization and Auditory Feedback Training Mechanism** 

Chung Tran; Chi Mai Luong; Sakriani Sakti

Publication Year: 2025 , Page(s): 1479 - 1492

✓ Abstract [HTML](#)  

☐ **Imposing Correlation Structures for Deep Binaural Spatio-Temporal Wiener Filtering** 

Marvin Tammen; Simon Doclo

Publication Year: 2025 , Page(s): 1278 - 1292

✓ Abstract [HTML](#)  

☐ **Complex-bin2bin: A Latency-Flexible Generative Neural Model for Audio Packet Loss Concealment** 

Carlo Aironi; Leonardo Gabrielli; Samuele Cornell; Stefano Squartini

Publication Year: 2025 , Page(s): 199 - 210

✓ Abstract [HTML](#)  

☐ **Beamforming in the Short-Time Fourier Transform Domain via Dimensionality Reduction** 

Wei Liu; Jacob Benesty; Gongping Huang; Jingdong Chen

Publication Year: 2025 , Page(s): 1730 - 1742

✓ Abstract [HTML](#)  

☐ **Cross-Modal Knowledge Distillation With Multi-Stage Adaptive Feature Fusion for Speech Separation** 

Cunhang Fan; Wang Xiang; Jianhua Tao; Jiangyan Yi; Zhao Lv

Publication Year: 2025 , Page(s): 935 - 948

✓ Abstract [HTML](#)  

☐ **ASR Error Correction Using Large Language Models** 

Rao Ma; Mengjie Qian; Mark Gales; Kate Knill

Publication Year: 2025 , Page(s): 1389 - 1401

✓ Abstract [HTML](#)  

☐ **Cross-Document Distillation via Graph-Based Summarization of Extracted Essential Knowledge** 

Luca Ragazzi; Gianluca Moro; Lorenzo Valgimigli; Riccardo Fiorani

Publication Year: 2025 , Page(s): 518 - 527

|                                                                                                                                                            |  |      |                                                                                     |                                                                                       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------|--|------|-------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |     |       |
| <input type="checkbox"/> <b>Self-Supervised Reflective Learning Through Self-Distillation and Online Clustering for Speaker Representation Learning</b>    |  |      |                                                                                     |    |
| Danwei Cai; Zexin Cai; Ze Li; Ming Li                                                                                                                      |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 1535 - 1550                                                                                                              |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |    |      |
| <input type="checkbox"/> <b>A Closer Look at Class-Incremental Learning for Multi-Label Audio Classification</b>                                           |  |      |                                                                                     |    |
| Manjunath Mulimani; Annamaria Mesaros                                                                                                                      |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 1293 - 1306                                                                                                              |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |    |      |
| <input type="checkbox"/> <b>Fast Neural Vocoder With Fundamental Frequency Control Using Finite Impulse Response Filters</b>                               |  |      |                                                                                     |    |
| Yamato Ohtani; Takuma Okamoto; Tomoki Toda; Hisashi Kawai                                                                                                  |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 1 - 14                                                                                                                   |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  |      |    |      |
| <input type="checkbox"/> <b>MN-Net: Speech Enhancement Network via Modeling the Noise</b>                                                                  |  |      |                                                                                     |    |
| Ying Hu; Qin Yang; Wenbing Wei; Li Lin; Liang He; Zhijian Ou; Wenzhong Yang                                                                                |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 1208 - 1219                                                                                                              |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |    |      |
| <input type="checkbox"/> <b>Audio-Visual Target Speaker Extraction With Selective Auditory Attention</b>                                                   |  |      |                                                                                     |    |
| Ruijie Tao; Xinyuan Qian; Yidi Jiang; Junjie Li; Jiadong Wang; Haizhou Li                                                                                  |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 797 - 811                                                                                                                |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |  |    |
| <input type="checkbox"/> <b>An Approach to Microphone Array Geometry Transform</b>                                                                         |  |      |                                                                                     |  |
| Chao Pan; Jingdong Chen; Jacob Benesty                                                                                                                     |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 869 - 882                                                                                                                |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |  |    |
| <input type="checkbox"/> <b>UnitCorrect: Unit-Based Mispronunciation Correcting System With a DTW-Based Detection</b>                                      |  |      |                                                                                     |  |
| Hyun-Woo Bae; Hyung-Seok Oh; Seung-Bin Kim; Seong-Whan Lee                                                                                                 |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 1753 - 1762                                                                                                              |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |  |    |
| <input type="checkbox"/> <b>VM-ASR: A Lightweight Dual-Stream U-Net Model for Efficient Audio Super-Resolution</b>                                         |  |      |                                                                                     |  |
| Ting-Wei Zhang; Shanq-Jang Ruan                                                                                                                            |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 666 - 677                                                                                                                |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |  |    |
| <input type="checkbox"/> <b>Multimodal Audio-Based Disease Prediction With Transformer-Based Hierarchical Fusion Network</b>                               |  |      |                                                                                     |  |
| Jinjin Cai; Ruiqi Wang; Dezhong Zhao; Ziqin Yuan; Victoria McKenna; Aaron Friedman; Rachel Foot; Susan Storey; Ryan Boente; Sudip Vhaduri; Byung-Cheol Min |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 1170 - 1182                                                                                                              |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |  |    |
| <input type="checkbox"/> <b>Utterance Alignment of Language Models for Effective User Simulation in Task-Oriented Dialogues</b>                            |  |      |                                                                                     |  |
| Xiang Luo; Jin Wang; Xuejie Zhang                                                                                                                          |  |      |                                                                                     |                                                                                       |
| Publication Year: 2025 , Page(s): 483 - 493                                                                                                                |  |      |                                                                                     |                                                                                       |
| <input type="checkbox"/> Abstract                                                                                                                          |  | HTML |  |    |

- ☐ **USED: Universal Speaker Extraction and Diarization**   
Junyi Ao; Mehmet Sinan Yıldırım; Ruijie Tao; Meng Ge; Shuai Wang; Yanmin Qian; Haizhou Li  
Publication Year: 2025 , Page(s): 96 - 110  
[v Abstract](#) [HTML](#)  
- ☐ **Nonparallel Spoken-Text-Style Transfer for Linguistic Expression Control in Speech Generation**   
Daiki Yoshioka; Yusuke Yasuda; Tomoki Toda  
Publication Year: 2025 , Page(s): 333 - 346  
[v Abstract](#) [HTML](#)  
- ☐ **Enhancing Multimodal Object-Entity Relation Extraction via Multi-Aspect Contrastive Learning in Large Multimodal Models**   
Yaming Zhang; Jianfei Yu; Wenya Wang; Li Yang; Jia Yang; Rui Xia  
Publication Year: 2025 , Page(s): 1220 - 1229  
[v Abstract](#) [HTML](#)  
- ☐ **A Two-Step Attention-Based Feature Combination Cross-Attention System for Speech-Based Dementia Detection**   
Yilin Pan; Bahman Mirheidari; Daniel Blackburn; Heidi Christensen  
Publication Year: 2025 , Page(s): 896 - 907  
[v Abstract](#) [HTML](#)  
- ☐ **Zero-Shot Prompting for LLM-Based Machine Translation Using In-Domain Target Sentences**   
Baijun Ji; Xiangyu Duan; Yue Zhang; Kaixin Wu; Min Zhang  
Publication Year: 2025 , Page(s): 251 - 261  
[v Abstract](#) [HTML](#)  
- ☐ **Validity and Robustness of Denoisers: A Proof of Concept in Speech Denoising**   
Masoud Ebrahimi; Qamar Alfalouji; Mina Basirat  
Publication Year: 2025 , Page(s): 650 - 665  
[v Abstract](#) [HTML](#)  
- ☐ **Enhancing Speech Emotion Recognition With Conditional Emotion Feature Diffusion and Progressive Interleaved Learning Strategy**   
Yang Liu; Xin Chen; Zhichao Peng; Yongwei Li; Xingfeng Li; Peng Song; Masashi Unoki; Zhen Zhao  
Publication Year: 2025 , Page(s): 1787 - 1800  
[v Abstract](#) [HTML](#)  
- ☐ **Weakly-Supervised Depression Detection in Speech Through Self-Learning Based Label Correction**   
Yanfei Sun; Yuanyuan Zhou; Xinzhou Xu; Jin Qi; Feiyi Xu; Zhao Ren; Björn W. Schuller  
Publication Year: 2025 , Page(s): 748 - 758  
[v Abstract](#) [HTML](#)  
- ☐ **Multitask Frequency-Time Feature-Wise Linear Modulation for Piano Transcription With Pedal**   
Qi Wang; Mingkuan Liu; Maoshen Jia  
Publication Year: 2025 , Page(s): 1049 - 1062  
[v Abstract](#) [HTML](#)  
- ☐ **Subband Adaptive Filtering Based on the Squared Sine Error Criterion: Analysis and Applications**   
Zhikai Gong; Yingsong Li; Xinqi Huang; Liping Li; Zhixiang Huang; Rodrigo. C. de Lamare; Yingying Zhu  
Publication Year: 2025 , Page(s): 1 - 11  
[v Abstract](#)  
- ☐ **Physics-Informed Identification and Interpolation of the Directivity of Sound Sources**   
Mathias Lemke; Arne Hölter; Stefan Weinzierl

Publication Year: 2025 , Page(s): 372 - 385

[v](#) Abstract [HTML](#)  ☐ **Predict-and-Update Network: Audio-Visual Speech Recognition Inspired by Human Speech Perception** 

Jiadong Wang; Xinyuan Qian; Haizhou Li

Publication Year: 2025 , Page(s): 11 - 22

Cited by: [Papers \(1\)](#)[v](#) Abstract [HTML](#)  ☐ **TTSslow: Slow Down Text-to-Speech With Efficiency Robustness Evaluations** 

Xiaoxue Gao; Yiming Chen; Xianghu Yue; Yu Tsao; Nancy F. Chen

Publication Year: 2025 , Page(s): 693 - 704

Cited by: [Papers \(1\)](#)[v](#) Abstract [HTML](#)  ☐ **MetaR: Few-Shot Named Entity Recognition With Meta-Learning and Relation Network** 

Bingli Sun; Kaiqi Gong; Wenxin Li; Xiao Song

Publication Year: 2025 , Page(s): 974 - 986

[v](#) Abstract [HTML](#)  ☐ **DiffuSyn: A Diffusion-Driven Framework With Syntactic Dependency for Aspect Sentiment Triplet Extraction** 

Qiu Hua Yi; Xiangjie Kong; Linan Zhu; Chenwei Zhang; Guo Jiang Shen

Publication Year: 2025 , Page(s): 842 - 854

[v](#) Abstract [HTML](#)  ☐ **Strategic Optimization for Worst-Case Augmentation and Classification** 

Jen-Tzung Chien; Mahdin Rohmatillah; Chang-Ting Chu

Publication Year: 2025 , Page(s): 111 - 123

[v](#) Abstract [HTML](#)  ☐ **Joint Music Segmentation and Clustering Based on Self-Attentive Contrastive Learning of Multifaceted Self-Similarity Representation** 

Tsong-Ping Chen; Kazuyoshi Yoshii

Publication Year: 2025 , Page(s): 1267 - 1277

[v](#) Abstract [HTML](#)  ☐ **Optimal Transport With Class Structure Exploration for Cross-Domain Speech Emotion Recognition** 

Ruiteng Zhang; Jianguo Wei; Xugang Lu; Junhai Xu; Yongwei Li; Di Jin; Lin Zhang; Wenhuan Lu

Publication Year: 2025 , Page(s): 37 - 53

[v](#) Abstract [HTML](#)  ☐ **Residual Echo Suppression Using Dual-Stream Interactive Transformers With Selective Multi-Scale Encoding** 

Kai Xie; Ziye Yang; Jie Chen; Mengyao Zhu

Publication Year: 2025 , Page(s): 1103 - 1116

[v](#) Abstract [HTML](#)  ☐ **Towards High-Quality and Efficient Speech Bandwidth Extension With Parallel Amplitude and Phase Prediction** 

Ye-Xin Lu; Yang Ai; Hui-Peng Du; Zhen-Hua Ling

Publication Year: 2025 , Page(s): 236 - 250

Cited by: [Papers \(3\)](#)[v](#) Abstract [HTML](#)  

☐

**SF-Speech: Straightened Flow for Zero-Shot Voice Clone**

Xuyuan Li; Zengqiang Shang; Hua Hua; Peiyang Shi; Chen Yang; Li Wang; Pengyuan Zhang

Publication Year: 2025 , Page(s): 1706 - 1718

Abstract

HTML

☐

**Improving End-to-End Sign Language Translation via Multi-Level Contrastive Learning**

Biao Fu; Liang Zhang; Peigen Ye; Pei Yu; Cong Hu; Xiaodong Shi; Yidong Chen

Publication Year: 2025 , Page(s): 1230 - 1242

Abstract

HTML

☐

**Learnable Counterfactual Attention for Music Classification**

Yi-Xing Lin; Jen-Chun Lin; Wen-Li Wei; Jia-Ching Wang

Publication Year: 2025 , Page(s): 570 - 585

Abstract

HTML

[Submit Manuscript](#)

[Submission Guidelines](#)

[Become a Reviewer](#)

[Open Access Publishing Options](#)

[Add Title To My Alerts](#) [Add to My Favorites](#)

Contact Information

Editor-in-Chief

Mark Hasegawa-Johnson  
University of Illinois  
Urbana, IL  
USA  
[jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)

[Editorial Board](#)

[Author Center](#)

[IEEE Transactions on Audio, Speech and Language Processing](#)

[Information for Authors](#)

[Publishing Policies](#)

IEEE Personal Account

CHANGE USERNAME/PASSWORD

Purchase Details

PAYMENT OPTIONS

VIEW PURCHASED DOCUMENTS

Profile Information

COMMUNICATIONS PREFERENCES

PROFESSION AND EDUCATION

TECHNICAL INTERESTS

Need Help?

US & CANADA: +1 800 678 4333

WORLDWIDE: +1 732 981 0060

CONTACT & SUPPORT

Follow

f

@

in

▶

X

About IEEE *Xplore* | Contact Us | Help | Accessibility | Terms of Use | Nondiscrimination Policy | IEEE Ethics Reporting  | Sitemap | IEEE Privacy Policy

A public charity, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.

© Copyright 2025 IEEE - All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies.