



Access provided by:
University of Surrey

Sign Out



Access provided by:
University of Surrey

Sign Out

All



ADVANCED SEARCH

Author Resources

Submission Guidelines

Submit Manuscript

Author Center

Browse Journals & Magazines > IEEE/ACM Transactions on Audio... ?

Become a Reviewer



IEEE/ACM Transactions on Audio, Speech, and Language Processing

Open Access Publishing Options



Submit Manuscript



Add Title To My Alerts



Add to My Favorites



Home Meet the Editor

Popular

Early Access

Current Volume

All Issues

About Journal

Editor-in-Chief



Popular Documents - June 2024 Popular Article Feed

Paris Search results 50 most frequently accessed documents for this publication.

University of Illinois at Urbana-Champaign, USA

paris@illinois.edu


Showing 1-50 of 50

Refine

Date

June 2024






- End-to-End Speech Recognition: A Survey** 

Rohit Prabhavalkar; Takaaki Hori; Tara N. Sainath; Ralf Schlüter; Shinji Watanabe

Publication Year: 2024 , Page(s): 325 - 351

Cited by: Papers (18)


Abstract **HTML**  

- End-to-End Speech Recognition: A Survey** 

Rohit Prabhavalkar; Takaaki Hori; Tara N. Sainath; Ralf Schlüter; Shinji Watanabe

IEEE/ACM Transactions on Audio, Speech, and Language Processing



Year: 2024 | Volume: 32


- Conv-TasNet: Surpassing Ideal Time–Frequency Magnitude Masking for Speech Separation** 

Yi Luo; Nima Mesgarani

Publication Year: 2019 , Page(s): 1256 - 1266

Cited by: Papers (980) | Patents (1)

Abstract **HTML**  

Conv-TasNet: Surpassing Ideal Time–Frequency Magnitude Masking for Speech Separation 

Yi Luo; Nima Mesgarani

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2019 | Volume: 27 | Issue: 8


An Overview of Voice Conversion and Its Challenges: From Statistical Modeling to Deep Learning 

Berrak Sisman; Junichi Yamagishi; Simon King; Haizhou Li

Publication Year: 2021 , Page(s): 132 - 157

Cited by: Papers (139)


✓ Abstract **HTML**  

An Overview of Voice Conversion and Its Challenges: From Statistical Modeling to Deep Learning 

Berrak Sisman; Junichi Yamagishi; Simon King; Haizhou Li

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2021 | Volume: 29


Music ControlNet: Multiple Time-Varying Controls for Music Generation 

Shih-Lun Wu; Chris Donahue; Shinji Watanabe;

Nicholas J. Bryan

Publication Year: 2024 , Page(s): 2692 - 2703

✓ Abstract **HTML**  


Music ControlNet: Multiple Time-Varying Controls for Music Generation 

Shih-Lun Wu; Chris Donahue; Shinji Watanabe;
Nicholas J. Bryan
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2024 | Volume: 32

Speech Enhancement and Dereverberation With Diffusion-Based Generative Models 

Julius Richter; Simon Welker; Jean-Marie Lemerrier;
Bunlong Lay; Timo Gerkmann
Publication Year: 2023 , Page(s): 2351 - 2364
Cited by: Papers (30)

▼ Abstract **HTML**  

Speech Enhancement and Dereverberation With Diffusion-Based Generative Models 

Julius Richter; Simon Welker; Jean-Marie Lemerrier;
Bunlong Lay; Timo Gerkmann
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2023 | Volume: 31

HuBERT: Self-Supervised Speech Representation Learning by Masked Prediction of Hidden Units 

Wei-Ning Hsu; Benjamin Bolte; Yao-Hung Hubert Tsai;
Kushal Lakhotia; Ruslan Salakhutdinov; Abdelrahman Mohamed
Publication Year: 2021 , Page(s): 3451 - 3460
Cited by: Papers (627)

▼ Abstract **HTML**  


HuBERT: Self-Supervised Speech Representation Learning by Masked Prediction of Hidden Units 

Wei-Ning Hsu; Benjamin Bolte; Yao-Hung Hubert Tsai;
Kushal Lakhotia; Ruslan Salakhutdinov;
Abdelrahman Mohamed
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2021 | Volume: 29

AudioLDM 2: Learning Holistic Audio Generation With Self-Supervised Pretraining 

Haohe Liu; Yi Yuan; Xubo Liu; Xinhao Mei; Qiuqiang Kong;
Qiao Tian; Yuping Wang; Wenwu Wang; Yuxuan Wang;
Mark D. Plumbley
Publication Year: 2024 , Page(s): 2871 - 2883
Cited by: Papers (4)

▼ Abstract **HTML**  

AudioLDM 2: Learning Holistic Audio Generation With Self-Supervised Pretraining 

Haohe Liu; Yi Yuan; Xubo Liu; Xinhao Mei; Qiuqiang Kong;
Qiao Tian; Yuping Wang; Wenwu Wang; Yuxuan Wang;
Mark D. Plumbley
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2024 | Volume: 32


Golden Gemini is All You Need: Finding the Sweet Spots for Speaker Verification 

Tianchi Liu; Kong Aik Lee; Qionqiong Wang; Haizhou Li

Publication Year: 2024 , Page(s): 2324 - 2337

Cited by: Papers (1)

✓ Abstract **HTML**  

- Golden Gemini is All You Need: Finding the Sweet Spots for Speaker Verification** 

Tianchi Liu; Kong Aik Lee; Qionqiong Wang; Haizhou Li
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2024 | Volume: 32

- Disentangling Prosody Representations With Unsupervised Speech Reconstruction** 

Leyuan Qu; Taihao Li; Cornelius Weber;
Theresa Pekarek-Rosin; Fuji Ren; Stefan Wermter
Publication Year: 2024 , Page(s): 39 - 54









✓ Abstract **HTML**  

- Disentangling Prosody Representations With Unsupervised Speech Reconstruction** 

Leyuan Qu; Taihao Li; Cornelius Weber;
Theresa Pekarek-Rosin; Fuji Ren; Stefan Wermter
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2024 | Volume: 32

- CMGAN: Conformer-Based Metric-GAN for Monaural Speech Enhancement** 

Sherif Abdulatif; Ruizhe Cao; Bin Yang
Publication Year: 2024 , Page(s): 2477 - 2493

- ▼ Abstract **HTML**   
- CMGAN: Conformer-Based Metric-GAN for Monaural Speech Enhancement**
- Sherif Abdulatif; Ruizhe Cao; Bin Yang
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2024 | Volume: 32
-
- Accented Text-to-Speech Synthesis With Limited Data** 
- Xuehao Zhou; Mingyang Zhang; Yi Zhou; Zhizheng Wu; Haizhou Li
Publication Year: 2024 , Page(s): 1699 - 1711
Cited by: Papers (1)
- ▼ Abstract **HTML**   
- Accented Text-to-Speech Synthesis With Limited Data**
- Xuehao Zhou; Mingyang Zhang; Yi Zhou; Zhizheng Wu; Haizhou Li
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2024 | Volume: 32
-
- PANNs: Large-Scale Pretrained Audio Neural Networks for Audio Pattern Recognition** 
- Qiuqiang Kong; Yin Cao; Turab Iqbal; Yuxuan Wang; Wenwu Wang; Mark D. Plumbley
Publication Year: 2020 , Page(s): 2880 - 2894
Cited by: Papers (426)

▼ Abstract **HTML**  

PANNs: Large-Scale Pretrained Audio Neural Networks for Audio Pattern Recognition 

Qiuqiang Kong; Yin Cao; Turab Iqbal; Yuxuan Wang;
Wenwu Wang; Mark D. Plumbley

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2020 | Volume: 28


An Overview of Deep-Learning-Based Audio-Visual Speech Enhancement and Separation 

Daniel Michelsanti; Zheng-Hua Tan; Shi-Xiong Zhang; Yong Xu;
Meng Yu; Dong Yu; Jesper Jensen

Publication Year: 2021 , Page(s): 1368 - 1396

Cited by: Papers (124)

▼ Abstract **HTML**  

An Overview of Deep-Learning-Based Audio-Visual Speech Enhancement and Separation 

Daniel Michelsanti; Zheng-Hua Tan; Shi-Xiong Zhang;
Yong Xu; Meng Yu; Dong Yu; Jesper Jensen

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2021 | Volume: 29

AudioLM: A Language Modeling Approach to Audio Generation 


Zalán Borsos; Raphaël Marinier; Damien Vincent;
Eugene Kharitonov; Olivier Pietquin; Matt Sharifi;
Dominik Roblek; Olivier Teboul; David Grangier;

Marco Tagliasacchi; Neil Zeghidour

Publication Year: 2023 , Page(s): 2523 - 2533

Cited by: Papers (57)

✓ Abstract **HTML**  

- AudioLM: A Language Modeling Approach to Audio Generation** 

Zalán Borsos; Raphaël Marinier; Damien Vincent;

Eugene Kharitonov; Olivier Pietquin; Matt Sharifi;

Dominik Roblek; Olivier Teboul; David Grangier;

Marco Tagliasacchi; Neil Zeghidour

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2023 | Volume: 31


-
- Audio Super-Resolution With Robust Speech Representation Learning of Masked Autoencoder** 

Seung-Bin Kim; Sang-Hoon Lee; Ha-Yeong Choi;

Seong-Whan Lee

Publication Year: 2024 , Page(s): 1012 - 1022

✓ Abstract **HTML**  





- Audio Super-Resolution With Robust Speech Representation Learning of Masked Autoencoder** 





Seung-Bin Kim; Sang-Hoon Lee; Ha-Yeong Choi;


Seong-Whan Lee

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2024 | Volume: 32

- Pre-Training With Whole Word Masking for Chinese BERT** 
Yiming Cui; Wanxiang Che; Ting Liu; Bing Qin; Ziqing Yang
Publication Year: 2021 , Page(s): 3504 - 3514
Cited by: Papers (486)
- Abstract **HTML**  
- Pre-Training With Whole Word Masking for Chinese BERT** 
Yiming Cui; Wanxiang Che; Ting Liu; Bing Qin; Ziqing Yang
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2021 | Volume: 29

- SpatialNet: Extensively Learning Spatial Information for Multichannel Joint Speech Separation, Denoising and Dereverberation** 
Changsheng Quan; Xiaofei Li
Publication Year: 2024 , Page(s): 1310 - 1323
Cited by: Papers (1)
- Abstract **HTML**  
- SpatialNet: Extensively Learning Spatial Information for Multichannel Joint Speech Separation, Denoising and Dereverberation** 
Changsheng Quan; Xiaofei Li
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2024 | Volume: 32

ASVspoof 2021: Towards Spoofed and Deepfake Speech Detection in the Wild 

Xuechen Liu; Xin Wang; Md Sahidullah; Jose Patino;
Héctor Delgado; Tomi Kinnunen; Massimiliano Todisco;
Junichi Yamagishi; Nicholas Evans; Andreas Nautsch;
Kong Aik Lee

Publication Year: 2023 , Page(s): 2507 - 2522

Cited by: Papers (33)


▼ Abstract **HTML**   

ASVspoof 2021: Towards Spoofed and Deepfake Speech Detection in the Wild 

Xuechen Liu; Xin Wang; Md Sahidullah; Jose Patino;
Héctor Delgado; Tomi Kinnunen; Massimiliano Todisco;
Junichi Yamagishi; Nicholas Evans; Andreas Nautsch;
Kong Aik Lee

IEEE/ACM Transactions on Audio, Speech, and Language
Processing


Year: 2023 | Volume: 31

Masked Modeling Duo: Towards a Universal Audio Pre-Training Framework 

Daisuke Niizumi; Daiki Takeuchi; Yasunori Ohishi;
Noboru Harada; Kunio Kashino

Publication Year: 2024 , Page(s): 2391 - 2406

▼ Abstract **HTML**  


Masked Modeling Duo: Towards a Universal Audio Pre-Training Framework 

Daisuke Niizumi; Daiki Takeuchi; Yasunori Ohishi;
Noboru Harada; Kunio Kashino
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2024 | Volume: 32

Convergence and Performance Analysis of Classical, Hybrid, and Deep Acoustic Echo Control 

Ernst Seidel; Pejman Mowlae; Tim Fingscheidt
Publication Year: 2024 , Page(s): 2857 - 2870

▼ Abstract **HTML**  


Convergence and Performance Analysis of Classical, Hybrid, and Deep Acoustic Echo Control 

Ernst Seidel; Pejman Mowlae; Tim Fingscheidt
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2024 | Volume: 32

SoundStream: An End-to-End Neural Audio Codec 

Neil Zeghidour; Alejandro Luebs; Ahmed Omran; Jan Skoglund;
Marco Tagliasacchi
Publication Year: 2022 , Page(s): 495 - 507
Cited by: Papers (89)

▼ Abstract **HTML**  


SoundStream: An End-to-End Neural Audio Codec 

Neil Zeghidour; Alejandro Luebs; Ahmed Omran;
Jan Skoglund; Marco Tagliasacchi
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2022 | Volume: 30

 Analyzing the Robustness of Vision & Language Models 

Alexander Shirnin; Nikita Andreev; Sofia Potapova;
Ekaterina Artemova
Publication Year: 2024 , Page(s): 2751 - 2763

✓ Abstract **HTML**  


 Analyzing the Robustness of Vision & Language Models 

Alexander Shirnin; Nikita Andreev; Sofia Potapova;
Ekaterina Artemova
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2024 | Volume: 32

 Supervised Speech Separation Based on Deep Learning: An Overview 

DeLiang Wang; Jitong Chen
Publication Year: 2018 , Page(s): 1702 - 1726
Cited by: Papers (871)

✓ Abstract **HTML**  

- Supervised Speech Separation Based on Deep Learning: An Overview** 

DeLiang Wang; Jitong Chen

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2018 | Volume: 26 | Issue: 10


- On Semi-Blind Source Separation-Based Approaches to Nonlinear Echo Cancellation Based on Bilinear Alternating Optimization** 

Xianrui Wang; Yichen Yang; Andreas Brendel; Tetsuya Ueda;

Shoji Makino; Jacob Benesty; Walter Kellermann; Jingdong Chen

Publication Year: 2024 , Page(s): 2973 - 2987

▼ Abstract **HTML**  

- On Semi-Blind Source Separation-Based Approaches to Nonlinear Echo Cancellation Based on Bilinear Alternating Optimization** 

Xianrui Wang; Yichen Yang; Andreas Brendel; Tetsuya Ueda;

Shoji Makino; Jacob Benesty; Walter Kellermann;

Jingdong Chen

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2024 | Volume: 32









- A Consolidated Perspective on Multimicrophone Speech Enhancement and Source Separation** 

Sharon Gannot; Emmanuel Vincent; Shmulik Markovich-Golan;

Alexey Ozerov

Publication Year: 2017 , Page(s): 692 - 730

Cited by: Papers (371) | Patents (2)


- [Abstract](#) [HTML](#)   
- A Consolidated Perspective on Multimicrophone Speech Enhancement and Source Separation**
- Sharon Gannot; Emmanuel Vincent;
Shmulik Markovich-Golan; Alexey Ozerov
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2017 | Volume: 25 | Issue: 4
-
- A Time-Frequency Attention Module for Neural Speech Enhancement** 
- Qiquan Zhang; Xinyuan Qian; Zhaoheng Ni; Aaron Nicolson;
Eliathamby Ambikairajah; Haizhou Li
Publication Year: 2023 , Page(s): 462 - 475
Cited by: Papers (16)
- [Abstract](#) [HTML](#)   
- A Time-Frequency Attention Module for Neural Speech Enhancement**
- Qiquan Zhang; Xinyuan Qian; Zhaoheng Ni; Aaron Nicolson;
Eliathamby Ambikairajah; Haizhou Li
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2023 | Volume: 31
-
- The LOCATA Challenge: Acoustic Source Localization and Tracking** 
- Christine Evers; Heinrich W. Löllmann; Heinrich Mellmann;
Alexander Schmidt; Hendrik Barfuss; Patrick A. Naylor;

Walter Kellermann

Publication Year: 2020 , Page(s): 1620 - 1643

Cited by: Papers (95)

▼ Abstract **HTML**  

- The LOCATA Challenge: Acoustic Source Localization and Tracking** 

Christine Evers; Heinrich W. Löllmann; Heinrich Mellmann;
Alexander Schmidt; Hendrik Barfuss; Patrick A. Naylor;
Walter Kellermann

IEEE/ACM Transactions on Audio, Speech, and Language
Processing


Year: 2020 | Volume: 28

-
- DiffProsody: Diffusion-Based Latent Prosody Generation for Expressive Speech Synthesis With Prosody Conditional Adversarial Training** 

Hyung-Seok Oh; Sang-Hoon Lee; Seong-Whan Lee

Publication Year: 2024 , Page(s): 2654 - 2666

▼ Abstract **HTML**  

- DiffProsody: Diffusion-Based Latent Prosody Generation for Expressive Speech Synthesis With Prosody Conditional Adversarial Training** 

Hyung-Seok Oh; Sang-Hoon Lee; Seong-Whan Lee

IEEE/ACM Transactions on Audio, Speech, and Language
Processing


Year: 2024 | Volume: 32

-
- Statistical Analysis for Speaker Recognition Evaluation With Data Dependence and Three Score Distributions** 

Jin Chu Wu; Raghu N. Kacker

Publication Year: 2024 , Page(s): 1 - 14

▼ Abstract **HTML**  

- Statistical Analysis for Speaker Recognition Evaluation With Data Dependence and Three Score Distributions** 

Jin Chu Wu; Raghu N. Kacker

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2024 | Volume: 32

-
- Cooperative Scene-Event Modelling for Acoustic Scene Classification** 

Yuanbo Hou; Bo Kang; Andrew Mitchell; Wenwu Wang;

Jian Kang; Dick Botteldooren

Publication Year: 2024 , Page(s): 68 - 82

Cited by: Papers (2)

▼ Abstract **HTML**  


- Cooperative Scene-Event Modelling for Acoustic Scene Classification** 

Yuanbo Hou; Bo Kang; Andrew Mitchell; Wenwu Wang;

Jian Kang; Dick Botteldooren

IEEE/ACM Transactions on Audio, Speech, and Language Processing


Year: 2024 | Volume: 32

-
- Improving Speaker Verification With Noise-Aware Label Ensembling and Sample Selection: Learning and Correcting Noisy Speaker Labels** 

Zhihua Fang; Liang He; Lin Li; Ying Hu

Publication Year: 2024 , Page(s): 2988 - 3001

✓ Abstract **HTML**  

- Improving Speaker Verification With Noise-Aware Label Ensembling and Sample Selection: Learning and Correcting Noisy Speaker Labels 

Zhihua Fang; Liang He; Lin Li; Ying Hu

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2024 | Volume: 32

- A Large-Scale Evaluation of Speech Foundation Models** 

Shu-wen Yang; Heng-Jui Chang; Zili Huang; Andy T. Liu; Cheng-I Lai; Haibin Wu; Jiatong Shi; Xuankai Chang; Hsiang-Sheng Tsai; Wen-Chin Huang; Tzu-hsun Feng; Po-Han Chi; Yist Y. Lin; Yung-Sung Chuang; Tzu-Hsien Huang; Wei-Cheng Tseng; Kushal Lakhotia; Shang-Wen Li; Abdelrahman Mohamed; Shinji Watanabe; Hung-yi Lee
Publication Year: 2024 , Page(s): 2884 - 2899

✓ Abstract **HTML**  

A Large-Scale Evaluation of Speech Foundation Models 

Shu-wen Yang; Heng-Jui Chang; Zili Huang; Andy T. Liu;
Cheng-I Lai; Haibin Wu; Jiatong Shi; Xuankai Chang;
Hsiang-Sheng Tsai; Wen-Chin Huang; Tzu-hsun Feng;
Po-Han Chi; Yist Y. Lin; Yung-Sung Chuang;
Tzu-Hsien Huang; Wei-Cheng Tseng; Kushal Lakhota;
Shang-Wen Li; Abdelrahman Mohamed; Shinji Watanabe;
Hung-yi Lee

IEEE/ACM Transactions on Audio, Speech, and Language
Processing


Year: 2024 | Volume: 32

 Waveform-Domain Speech Enhancement Using Spectrogram Encoding for Robust Speech Recognition 

Hao Shi; Masato Mimura; Tatsuya Kawahara

Publication Year: 2024 , Page(s): 3049 - 3060

▼ Abstract **HTML**  

 Waveform-Domain Speech Enhancement Using Spectrogram Encoding for Robust Speech Recognition 

Hao Shi; Masato Mimura; Tatsuya Kawahara

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2024 | Volume: 32


 Speaker Distance Estimation in Enclosures From Single-Channel Audio 

Michael Neri; Archontis Politis; Daniel Aleksander Krause;

Marco Carli; Tuomas Virtanen

Publication Year: 2024 , Page(s): 2242 - 2254

▼ Abstract **HTML**  

- Speaker Distance Estimation in Enclosures From Single-Channel Audio** 

Michael Neri; Archontis Politis; Daniel Aleksander Krause;
Marco Carli; Tuomas Virtanen

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2024 | Volume: 32


-
- EfficientTTS 2: Variational End-to-End Text-to-Speech Synthesis and Voice Conversion** 

Chenfeng Miao; Qingying Zhu; Minchuan Chen; Jun Ma;
Shaojun Wang; Jing Xiao

Publication Year: 2024 , Page(s): 1650 - 1661

Cited by: Papers (1)

▼ Abstract **HTML**  

- EfficientTTS 2: Variational End-to-End Text-to-Speech Synthesis and Voice Conversion** 

Chenfeng Miao; Qingying Zhu; Minchuan Chen; Jun Ma;
Shaojun Wang; Jing Xiao

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2024 | Volume: 32


-
- Acoustic Imaging With Circular Microphone Array: A New Approach for Sound Field Analysis** 

Marco Olivieri; Amy Bastine; Mirco Pezzoli; Fabio Antonacci;

Thushara Abhayapala; Augusto Sarti

Publication Year: 2024 , Page(s): 1750 - 1761

▼ Abstract **HTML**  

- Acoustic Imaging With Circular Microphone Array: A New Approach for Sound Field Analysis** 

Marco Olivieri; Amy Bastine; Mirco Pezzoli; Fabio Antonacci;
Thushara Abhayapala; Augusto Sarti

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2024 | Volume: 32


-
- BYOL for Audio: Exploring Pre-Trained General-Purpose Audio Representations** 

Daisuke Niizumi; Daiki Takeuchi; Yasunori Ohishi;
Noboru Harada; Kunio Kashino

Publication Year: 2023 , Page(s): 137 - 151

Cited by: Papers (11)

▼ Abstract **HTML**  

- BYOL for Audio: Exploring Pre-Trained General-Purpose Audio Representations** 

Daisuke Niizumi; Daiki Takeuchi; Yasunori Ohishi;
Noboru Harada; Kunio Kashino

IEEE/ACM Transactions on Audio, Speech, and Language
Processing


Year: 2023 | Volume: 31

-
- Unified Cross-Modal Attention: Robust Audio-Visual Speech Recognition and Beyond** 

Jiahong Li; Chenda Li; Yifei Wu; Yanmin Qian

Publication Year: 2024 , Page(s): 1941 - 1953

[Abstract](#) [HTML](#)  

-
- Unified Cross-Modal Attention: Robust Audio-Visual Speech Recognition and Beyond**
- 

Jiahong Li; Chenda Li; Yifei Wu; Yanmin Qian

IEEE/ACM Transactions on Audio, Speech, and Language Processing


Year: 2024 | Volume: 32

-
-
- Pretraining and Adaptation Techniques for Electrolaryngeal Speech Recognition**
- 

Lester Phillip Violeta; Ding Ma; Wen-Chin Huang; Tomoki Toda

Publication Year: 2024 , Page(s): 2777 - 2789


[Abstract](#) [HTML](#)  

-
- Pretraining and Adaptation Techniques for Electrolaryngeal Speech Recognition**
- 

Lester Phillip Violeta; Ding Ma; Wen-Chin Huang; Tomoki Toda

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2024 | Volume: 32

-
-
- Selective Acoustic Feature Enhancement for Speech Emotion Recognition With Noisy Speech**
- 







Seong-Gyun Leem; Daniel Fulford; Jukka-Pekka Onnela;

David Gard; Carlos Busso


Publication Year: 2024 , Page(s): 917 - 929

Cited by: Papers (1)

[Abstract](#) [HTML](#)  

- Selective Acoustic Feature Enhancement for Speech Emotion Recognition With Noisy Speech** 
Seong-Gyun Leem; Daniel Fulford; Jukka-Pekka Onnela;
David Gard; Carlos Busso
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2024 | Volume: 32
-
- Text-Inductive Graphone-Based Language Adaptation for Low-Resource Speech Synthesis** 
Takaaki Saeki; Soumi Maiti; Xinjian Li; Shinji Watanabe;
Shinnosuke Takamichi; Hiroshi Saruwatari
Publication Year: 2024 , Page(s): 1829 - 1844
- ▼ Abstract **HTML**  
- Text-Inductive Graphone-Based Language Adaptation for Low-Resource Speech Synthesis** 
Takaaki Saeki; Soumi Maiti; Xinjian Li; Shinji Watanabe;
Shinnosuke Takamichi; Hiroshi Saruwatari
IEEE/ACM Transactions on Audio, Speech, and Language Processing
Year: 2024 | Volume: 32
-
- Wavelet Multiresolution Analysis Based Speech Emotion Recognition System Using 1D CNN LSTM Networks** 
Aditya Dutt; Paul Gader
Publication Year: 2023 , Page(s): 2043 - 2054
Cited by: Papers (2)

▼ Abstract **HTML**  

- Wavelet Multiresolution Analysis Based Speech Emotion Recognition System Using 1D CNN LSTM Networks** 

Aditya Dutt; Paul Gader

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2023 | Volume: 31

-
- Music Source Separation With Band-Split RNN** 

Yi Luo; Jianwei Yu

Publication Year: 2023 , Page(s): 1893 - 1901

Cited by: Papers (19)

▼ Abstract **HTML**  

- Music Source Separation With Band-Split RNN** 

Yi Luo; Jianwei Yu

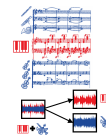
IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2023 | Volume: 31


-
- Source Separation of Piano Concertos Using Musically Motivated Augmentation Techniques** 

Yigitcan Özer; Meinard Müller

Publication Year: 2024 , Page(s): 1214 - 1225



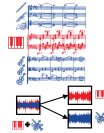
▼ Abstract **HTML**  

Source Separation of Piano Concertos Using Musically Motivated Augmentation Techniques 

Yigitcan Özer; Meinard Müller

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2024 | Volume: 32



Automated Data Augmentation for Audio Classification 

Yanjie Sun; Kele Xu; Chaorun Liu; Yong Dou; Huaimin Wang;
Bo Ding; Qinghua Pan

Publication Year: 2024 , Page(s): 2716 - 2728

✓ Abstract **HTML**  

Automated Data Augmentation for Audio Classification 

Yanjie Sun; Kele Xu; Chaorun Liu; Yong Dou; Huaimin Wang;
Bo Ding; Qinghua Pan

IEEE/ACM Transactions on Audio, Speech, and Language Processing

Year: 2024 | Volume: 32


Binaural Sound Source Distance Estimation and Localization for a Moving Listener 

Daniel Aleksander Krause; Guillermo García-Barrios;
Archontis Politis; Annamaria Mesaros

Publication Year: 2024 , Page(s): 996 - 1011

Cited by: Papers (1)

▼ Abstract **HTML**   

Binaural Sound Source Distance Estimation and Localization for a Moving Listener 

Daniel Aleksander Krause; Guillermo García-Barrios;
Archontis Politis; Annamaria Mesaros

IEEE/ACM Transactions on Audio, Speech, and Language
Processing


Year: 2024 | Volume: 32

Speech Dereverberation With Frequency Domain Autoregressive Modeling 

Anurenjan Purushothaman; Debottam Dutta; Rohit Kumar;
Sriram Ganapathy

Publication Year: 2024 , Page(s): 29 - 38

▼ Abstract **HTML**  

Speech Dereverberation With Frequency Domain Autoregressive Modeling 

Anurenjan Purushothaman; Debottam Dutta; Rohit Kumar;
Sriram Ganapathy

IEEE/ACM Transactions on Audio, Speech, and Language
Processing

Year: 2024 | Volume: 32


Unsupervised Music Source Separation Using Differentiable Parametric Source Models 

Kilian Schulze-Forster; Gaël Richard; Liam Kelley;
Clement S. J. Doire; Roland Badeau


Publication Year: 2023 , Page(s): 1276 - 1289

Cited by: Papers (7)

▼ Abstract **HTML**  


Unsupervised Music Source Separation Using Differentiable Parametric Source Models 

Kilian Schulze-Forster; Gaël Richard; Liam Kelley;
Clement S. J. Doire; Roland Badeau
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2023 | Volume: 31

Deep Prior-Based Audio Inpainting Using Multi-Resolution Harmonic Convolutional Neural Networks 

Federico Miotello; Mirco Pezzoli; Luca Comanducci;
Fabio Antonacci; Augusto Sarti
Publication Year: 2024 , Page(s): 113 - 123
Cited by: Papers (1)

▼ Abstract **HTML**  

Deep Prior-Based Audio Inpainting Using Multi-Resolution Harmonic Convolutional Neural Networks 

Federico Miotello; Mirco Pezzoli; Luca Comanducci;
Fabio Antonacci; Augusto Sarti
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2024 | Volume: 32

Hyperbolic Pre-Trained Language Model 

Weize Chen; Xu Han; Yankai Lin; Kaichen He; Ruobing Xie;
Jie Zhou; Zhiyuan Liu; Maosong Sun

Publication Year: 2024 , Page(s): 3101 - 3112

▼ Abstract **HTML**  

Hyperbolic Pre-Trained Language Model 

Weize Chen; Xu Han; Yankai Lin; Kaichen He; Ruobing Xie;
Jie Zhou; Zhiyuan Liu; Maosong Sun
IEEE/ACM Transactions on Audio, Speech, and Language
Processing
Year: 2024 | Volume: 32

[CHANGE
USERNAME/PASSWORD](#)

[PAYMENT OPTIONS
VIEW PURCHASED
DOCUMENTS](#)

[COMMUNICATIONS
PREFERENCES
PROFESSION AND
EDUCATION](#)


[TECHNICAL INTERESTS](#)

[US & CANADA: +1 800
678 4333](#)

[WORLDWIDE: +1 732
981 0060](#)

[CONTACT & SUPPORT](#)



[About IEEE Xplore](#) [Contact Us](#) [Help](#) [Accessibility](#) [Terms of Use](#) [Nondiscrimination Policy](#) [IEEE Ethics Reporting](#)  [Sitemap](#)
[IEEE Privacy Policy](#)

IEEE Account

- » [Change Username/Password](#)
- » [Update Address](#)

Purchase Details

- » [Payment Options](#)
- » [Order History](#)
- » [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](#)

[About IEEE Xplore](#) [Contact Us](#) [Help](#) [Accessibility](#) [Terms of Use](#) [Nondiscrimination Policy](#) [Sitemap](#) [Privacy & Opting Out of Cookies](#)

A not-for-profit organization, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.
© Copyright 2024 IEEE - All rights reserved. Use of this web site signifies your agreement to the terms and conditions.