



Jun Du

Associate Professor

Department of Electronic Engineering and Information Science

University of Science and Technology of China

Phone: (+86)-0551-63603303

Office: Room 502, West Building of Science and Technology Laboratory

Address: No. 443, Huangshan Road, Hefei, Anhui, 230027, China

E-mail: jundu@ustc.edu.cn

Research Area:

1. Audio and Speech Signal Processing: speech enhancement, speech separation, voice activity detection, speech dereverberation, robust speech recognition, speaker diarization, speech emotion recognition, acoustic scene classification, sound event localization and detection, few-shot bioacoustic event detection
2. Computer Vision: handwriting recognition, text detection, OCR, math equation recognition, object detection in aerial images
3. Multimodality: audio-visual emotion recognition, audio-visual speech enhancement, audio-visual speech recognition, audio-visual speaker diarization, audio-visual voice activity detection, scene text visual question answering

Education

Ph. D degree, 2004-2009, University of Science and Technology of China (USTC)

Bachelor degree, 2000-2004, USTC

Work Experience

2013-present, USTC, Associate Professor,

2010-2013, Microsoft Research Asia (MSRA), Speech Group, Associate Researcher

2009-2010, iFlytek Research, Research Manager

Awards and Honors

- 2023, the first place of [DASR task of CHiME-7 challenge](#)
- 2023, the first place of [SELD task of DCASE challenge](#)
- 2023, the first place of [FSBED task of DCASE challenge](#)
- 2022, the first prize of Wu Wenjun AI S&T Award (吴文俊科技进步一等奖)
- 2022, the best paper award for ISCSLP
- 2022, the first place of [SELD task of DCASE challenge](#)
- 2022, the first place of [FSBED task of DCASE challenge](#)
- 2021, the first place of [22 tasks of OpenASR challenge](#)
- 2020, the first place of [DIHARD-III challenge](#)
- 2020, the first place of [SELD task of DCASE challenge](#)
- 2020, the first place of Track 1 of [CHiME-6 challenge](#)
- 2020, the first place of ICFHR Competition on [OffRaSHME](#)
- 2020, [APSIPA Distinguished Lecturer](#) (2020-2021)
- 2019, the first place of Task1 and Task2 of [ICDAR CROHME](#)
- 2019, the first place of Task1 of [CVPR ODAI challenge](#)
- 2018, the first prize for S&T award of Anhui Province (安徽省科技进步一等奖)
- [2018 IEEE Signal Processing Society Best Paper Award](#)
- 2018, the first place of all four tasks of [CHiME-5 challenge](#)
- 2018, the first place of all three tasks of [ICPR MTWI contest](#)
- 2018, the first place of all two tasks of [ICPR ODAI contest](#)
- 2018, the first place of two tasks of [ICHFR competition](#)
- 2018, the second place of [DIHARD-I challenge](#)
- 2016, the first place of all three tasks of [CHiME-4 challenge](#)
- 2015, the third place of [CHiME-3 challenge](#)
- 2015, the finalist of best student paper award for INTERSPEECH
- 2015, the finalist of best student paper award for LVA/ICA
- 2012, MSRA technology transfer award for contribution in Windows Phone 8
- 2012, Microsoft Ship-It award for Chinese handwriting recognition
- 2012, Microsoft Ship-It award for camera-based translation

2011, Microsoft Ship-It award for mouse-hovering function in Bing dictionary

2008, the finalist of best student paper award for ISCSLP

2006, the finalist of best student paper award for INTERSPEECH

2006, [the finalist of best student paper award for ISCSLP](#)

Teaching

Fall 2023, INY5204, Digital Signal Processing II

Spring 2023, 006205, Signal Modeling and Algorithm Practice

Fall 2022, INY5204, Digital Signal Processing II

Fall 2022, 006205, Signal Modeling and Algorithm Practice

Spring 2022, INFO7405P, Speech Signal and Information Processing

Fall 2021, INY5204, Digital Signal Processing II

Fall 2021, 006205, Signal Modeling and Algorithm Practice

Fall 2020, INY5204, Digital Signal Processing II

Fall 2020, 006205, Signal Modeling and Algorithm Practice

Spring 2020, INY6302, Speech Signal and Information Processing

Fall 2019, INY5204, Digital Signal Processing II

Fall 2019, 006M02, Statistical Data Modeling

Fall 2018, INY5204, Digital Signal Processing II

Fall 2018, 006M02, Statistical Data Modeling

Spring 2018, INY6302, Speech Signal and Information Processing

Fall 2017, INY5204, Digital Signal Processing II

Fall 2017, 006M02, Statistical Data Modeling

Fall 2016, INY5204, Digital Signal Processing II

Fall 2016, 006M02, Statistical Data Modeling

Fall 2016, 601011, Freshman Seminar “Science and Society”

Spring 2016, INY6302, Speech Signal and Information Processing

Fall 2015, 006M02, Statistical Data Modeling

Fall 2014, 006M02, Statistical Data Modeling

Fall 2014, 601011, Freshman Seminar “Science and Society”

Fall 2013, 006M02, Statistical Data Modeling

Summer 2013, 006M02, Statistical Data Modeling

Projects

2022-2025, NSFC (General Program), Grant No. 62171427

2017-2020, NSFC (General Program), Grant No. 61671422

2017-2020, NSFC (United Fund) Subproject, Grant No. U1613211

2014-2016, NSFC (Youth Program), Grant No. 61305002

2017-2020, Key S&T Special Project of Anhui Province, Grant No. 17030901005

2013-2014, S&T Project of Anhui Province, Grant No. 13Z02008-4

2014-2016, NSF of Anhui Province, Grant No. 1408085QF101

2018-2020, USTC Research Funds

2018-2020, USTC Fundamental Research Funds for Central University

2022-2023, China Mobile Collaboration Project

2022-2023, Media Collaboration Project

2022-2023, Alibaba Collaboration Project

2022-2023, SenseTime Collaboration Project

2022-2023, OPPO Collaboration Project

2021-2022, Ximalaya Collaboration Project

2020-2021, Microsoft Research Asia Collaborative Research Grant

2017-2018, Microsoft Research Asia Collaborative Research Grant

2015-2016, Microsoft Research Asia Collaborative Research Grant

2017-2018, Samsung (Beijing) R&D Collaboration Project

2017-2018, Wechat Collaboration Project

2020-2021, Huawei Noah's Ark Collaboration Project

2018-2019, Huawei Noah's Ark Collaboration Project

2017-2018, HiSilicon Collaboration Project

2020-2021, Tencent Collaboration Project

2019-2020, Tencent Collaboration Project

2018-2019, Tencent Collaboration Project

Patent (Issued)

2021, Method, apparatus and storage medium for recognizing character US10949701

2016, Actionable content displayed on a touch screen, US9329692

2016, Translating language characters in media content, US9251144

2015, Rotation-free recognition of handwritten characters, US8977042

Service

Challenges and Data Collection Subcommittee Member of IEEE SPS (2023.08-)

Chair of APSIPA SLA TC (2023-)

IEEE Senior Member (2021.02)

SLTC Member of IEEE Signal Processing Society (2019.12-)

Associate Editor of IEEE/ACM TASLP (2018.7-2022.7)

Vice Chair of APSIPA SLA TC (2020-2022)

Secretary of APSIPA SLA TC (2019-2020)

APSIPA SLA TC Member (2017-2019)

Session Chair and Grand Challenge Chair for ICASSP 2022/2023

Publicity Chair for INTERSPEECH 2020

Session Chair for IEEE SLT 2021

Session Chair for APSIPA 2018-2022

Special Session Organizer for INTERSPEECH 2018/2019

Special Session Chair for ISCSLP 2018

Award Committee Member for MLSP 2017

Session Chair for MLSP 2017

Session Chair for ISCSLP 2016

Session Chair for ChinaSIP 2014/2015

Publications

I have published more than **200** journal and conference papers with **8300+** Google Scholar citations (the full publication can refer to [Google Scholar website](#)). I listed the selected work here.

1. Qing Wang, **Jun Du***, Hua-Xin Wu, Jia Pan, Feng Ma, and Chin-Hui Lee, “A four-stage data augmentation approach to ResNet-Conformer based acoustic modeling for sound event localization and detection,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, Vol. 31, pp.1251-1264, 2023.
2. Mao-Kui He, **Jun Du***, Qing-Feng Liu, and Chin-Hui Lee, “ANSD-MA-MSE: Adaptive neural speaker diarization using memory-aware multi-speaker embedding,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2023.
3. Shu-Tong Niu, **Jun Du***, Lei Sun, Yu Hu, and Chin-Hui Lee, “QDM-SSD: Quality-aware dynamic masking for separation-based speaker diarization,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, Vol. 31, pp.1037-1049, 2023.
4. Xueyang Zhang, Shuxian Wang, **Jun Du***, Genwei Yan, Jigang Tang, Tian Gao, Xin Fang, Jia Pan, and Jianqing Gao, “Frame-level embedding learning for few-shot bioacoustic event detection,” IEEE International Conference on Multimedia and Expo (ICME), pp.750-755, 2023.
5. Shutong Niu, **Jun Du***, Qing Wang, Li Chai, Huaxin Wu, Zhaoxu Nian, Lei Sun, Yi Fang, Jia Pan, and Chin-Hui Lee, “An experimental study on sound event localization and detection under realistic testing conditions,” **International Conference on Acoustics, Speech and Signal Processing (ICASSP)**, 2023.
6. Qing Wang, **Jun Du***, Zhaoxu Nian, Shutong Niu, Li Chai, Huaxin Wu, Jia Pan, and Chin-Hui Lee, “Loss function design for DNN-based sound event localization and detection on low-resource realistic data,” **International Conference on Acoustics, Speech and Signal Processing (ICASSP)**, 2023.
7. Yajian Wang, **Jun Du***, Hang Chen, Qing Wang, and Chin-Hui Lee, “Deep segment model for acoustic scene classification,” **INTERSPEECH**, pp.4177-4181, 2022.
8. Chenxi Wang, Hang Chen, **Jun Du***, Baocai Yin, and Jia Pan, “Multi-task joint learning for embedding aware audio-visual speech enhancement,” **International Symposium on Chinese Spoken Language Processing (ISCSLP)**, pp.255-259, 2022. (**Best Paper Award for ISCSLP 2022**).

9. Hengshun Zhou, **Jun Du***, Yuanyuan Zhang, Qing Wang, Qing-Feng Liu, and Chin-Hui Lee, "Information fusion in attention networks using adaptive and multi-level factorized bilinear pooling for audio-visual emotion recognition," **IEEE/ACM Transactions on Audio, Speech and Language Processing**, Vol. 29, pp.2617-2629, 2021.
10. Yanhui Tu, **Jun Du***, Tian Gao, Chin-Hui Lee, "A multi-target SNR-progressive learning approach to regression based speech enhancement," **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2020.
11. Jianshu Zhang, **Jun Du***, Yongxin Yang, Yi-Zhe Song, Lirong Dai, "A tree-structured decoder for image-to-markup generation," **International Conference on Machine Learning**, 2020.
12. Jianshu Zhang, **Jun Du***, Yongxin Yang, Yi-Zhe Song, Lirong Dai, "SRD: A tree structure based decoder for online handwritten mathematical expression recognition," **IEEE Transactions on Multimedia**, 2020.
13. Zi-Rui Wang, **Jun Du***, "Joint architecture and knowledge distillation in CNN for Chinese text recognition," **Pattern Recognition**, 2020.
14. Yixing Zhu, **Jun Du***, "TextMountain: accurate scene text detection via instance segmentation," **Pattern Recognition**, 2020.
15. Yixing Zhu, **Jun Du***, Xuqing Wu, "Adaptive period embedding for representing oriented objects in aerial images," **IEEE Transactions on Geoscience and Remote Sensing**, 2020.
16. Jia Pan, Genshun Wan, **Jun Du***, Zhongfu Ye, "Online speaker adaptation using memory-aware networks for speech recognition," **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2020.
17. Zi-Rui Wang, **Jun Du***, Jia-Ming Wang, "Writer-aware CNN for parsimonious HMM-based offline handwritten Chinese text recognition," **Pattern Recognition**, 2020.
18. Jun Qi, **Jun Du**, Sabato Marco Siniscalchi, Xiaoli Ma, Chin-Hui Lee, "On mean absolute error for deep neural network based vector-to-vector regression," **IEEE Signal Processing Letters**, 2020.
19. Jianshu Zhang, **Jun Du***, Lirong Dai, "Radical analysis network for learning hierarchies of Chinese characters," **Pattern Recognition**, 2020.
20. Li Chai, **Jun Du***, Qing-feng Liu, Chin-Hui Lee, "Using Generalized Gaussian Distributions to Improve Regression Error Modeling for Deep Learning-Based

- Speech Enhancement,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2019.
21. Yanhui Tu, **Jun Du***, Chin-Hui Lee, “Speech Enhancement Based on Teacher-Student Deep Learning Using Improved Speech Presence Probability for Noise-Robust Speech Recognition,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2019.
 22. Lei Sun, **Jun Du***, Tian Gao, Yi Fang, Feng Ma, Chin-Hui Lee, “A Speaker-Dependent Approach to Separation of Far-Field Multi-Talker Microphone Array Speech for Front-End Processing in the CHiME-5 Challenge,” **IEEE Journal of Selected Topics in Signal Processing**, 2019.
 23. Jianqing Gao, **Jun Du***, and Enhong Chen, “Mixed-Bandwidth Cross-Channel Speech Recognition via Joint Optimization of DNN-Based Bandwidth Expansion and Acoustic Modeling,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2019.
 24. Yanhui Tu, **Jun Du***, Lei Sun, etc., “An iterative mask estimation approach to deep learning based multi-channel speech recognition,” **Speech Communication**, 2019.
 25. Yixing Zhu, Chixiang Ma, **Jun Du***, “Rotated cascade R-CNN: A shape robust detector with coordinate regression,” **Pattern Recognition**, 2019.
 26. Qing Wang, **Jun Du***, and Li-Rong Dai, “A Multi-Objective Learning and Ensembling Approach to High-Performance Speech Enhancement with Compact Neural Network Architectures,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2018.
 27. Jianshu Zhang, **Jun Du***, and Li-Rong Dai, “Track, Attend and Parse (TAP): An End-to-end Framework for Online Handwritten Mathematical Expression Recognition,” **IEEE Transactions on Multimedia**, 2018.
 28. Jianshu Zhang, **Jun Du***, etc., “Watch, attend and parse: An end-to-end neural network based approach to handwritten mathematical expression recognition,” **Pattern Recognition**, 2017.
 29. Yannan Wang, **Jun Du***, Li-Rong Dai, and Chin-Hui Lee, “A gender mixture detection approach to unsupervised single-channel speech separation based on deep neural networks,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2017.

30. **Jun Du**, Yong Xu, “Hierarchical Deep Neural Networks for Multivariate Regression,” **Pattern Recognition**, 2017.
31. Yanhui Tu, **Jun Du**^{*}, Li-Rong Dai, and Chin-Hui Lee, “An information fusion framework with multi-channel feature concatenation and multi-perspective system combination for the deep-learning-based robust recognition of microphone array speech,” **Computer Speech and Language**, 2017.
32. Tian Gao, **Jun Du**^{*}, Li-Rong Dai, and Chin-Hui Lee, “A unified DNN approach to speaker-dependent simultaneous speech enhancement and speech separation in low SNR environments,” **Speech Communication**, 2017.
33. **Jun Du**, Yan-Hui Tu, Li-Rong Dai, and Chin-Hui Lee, “A regression approach to single-channel speech separation via high-resolution deep neural networks,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, 2016.
34. Yong Xu, **Jun Du**^{*}, Li-Rong Dai, and Chin-Hui Lee, “A regression approach to speech enhancement based on deep neural networks,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, Vol. 23, No. 1, pp.7-19, 2015. **(2018 IEEE Signal Processing Society Best Paper Award)**
35. Yong Xu, **Jun Du**^{*}, Li-Rong Dai, and Chin-Hui Lee, “An experimental study on speech enhancement based on deep neural networks,” **IEEE Signal Processing Letters**, Vol. 21, No. 1, pp.65-68, 2014.
36. **Jun Du** and Qiang Huo, “An irrelevant variability normalization approach to discriminative training of multi-prototype based classifiers and its applications for online handwritten Chinese character recognition,” **Pattern Recognition**, Vol. 47, No. 12, pp.3959-3966, 2014.
37. **Jun Du** and Qiang Huo, “An improved VTS feature compensation using mixture models of distortion and IVN training for noisy speech recognition,” **IEEE/ACM Transactions on Audio, Speech and Language Processing**, Vol. 22, No. 11, pp.1601-1611, 2014.
38. **Jun Du** and Qiang Huo, “A discriminative linear regression approach to adaptation of multi-prototype based classifiers and its applications for Chinese OCR,” **Pattern Recognition**, Vol. 46, No. 8, pp.2313-2322, 2013.
39. **Jun Du**, Yu Hu, and Hui Jiang, “Boosted mixture learning of Gaussian mixture hidden Markov models based on maximum likelihood for speech recognition,” **IEEE Trans. on Audio, Speech and Language Processing**, Vol. 19, No. 7.,

pp.2091-2100, 2011.

40. **Jun Du**, and Qiang Huo, “A feature compensation approach using high-order vector Taylor series approximation of an explicit distortion model for noisy speech recognition,” **IEEE Trans. on Audio, Speech and Language Processing**, Vol. 19, No. 8, pp.2285-2293, 2011.
41. Qing Wang, **Jun Du**^{*}, Xiao Bao, Zi-Rui Wang, Li-Rong Dai, and Chin-Hui Lee, “A universal VAD based on jointly trained deep neural networks,” Proc. INTERSPEECH 2015. (**Best student paper finalist**)
42. Tian Gao, **Jun Du**^{*}, Yong Xu, Cong Liu, Li-Rong Dai, and Chin-Hui Lee, “Improving deep neural network based speech enhancement in low SNR environments,” Proc. LVA/ICA 2015. (**Best student paper finalist**)
43. **Jun Du**, Qiang Huo, and Yu Hu, “Evaluation of a feature compensation approach using high-order vector Taylor series approximation of an explicit distortion model on Aurora2, Aurora3, and Aurora4 tasks,” Proc. ISCSLP, 2008, pp.81-84. (**Best student paper finalist**)
44. **Jun Du**, Peng Liu, Frank Soong, Jian-Lai Zhou, and Ren-Hua Wang, “Minimum divergence based discriminative training,” Proc. INTERSPEECH, 2006, pp.2410-2413. (**Best student paper finalist**)
45. **Jun Du**, Peng Liu, Frank Soong, Jian-Lai Zhou, and Ren-Hua Wang, “Noisy speech recognition performance of discriminative HMMs,” Proc. ISCSLP, 2006, pp.358-369. (**Best student paper finalist**)