

Yanmin Qian's List of Publication

Thesis (2012s): Yanmin Qian. Study on New Speech Recognition Technology under Low Data Resource Conditions. PHD Thesis in Tsinghua University.

Book:

1. 俞栋, 邓力, 俞凯, 钱彦旻. 著作: 人工智能: 语音识别理解与实践. 电子工业出版社, 2020: ISBN 978-7-121-38143-0
2. 俞凯, 钱彦旻. 译著: 解析深度学习-语音识别实践. 电子工业出版社, 2016: ISBN 978-7-121-28796-1

Book Chapters:

1. Khe Chai Sim and Yanmin Qian. Adaptation of Deep Neural Network Acoustic Models for Robust Automatic Speech Recognition, in New Era for Robust Speech Recognition: Exploiting Deep Learning. Springer, 219-243, 2017: ISBN 978-3-319-64680-0
2. Pawel Swietojanski, Yulan Liu and Yanmin Qian. Distant Speech Recognition, in Natural Speech Technology. Cambridge University Press 2017.

Journal Papers:

1. Leying Zhang, Tingxiao Zhou, Haiyang Sun, Mengxiao Bi and Yanmin Qian (Corresponding author) . DeepASMR: LLM-based Zero-Shot ASMR Speech Generation for Anyone of Any Voice. Submitted to **IEEE Transactions on Audio, Speech, and Language Processing**, 2026.
2. Xun Gong, Wei Wang, Hang Shao and Yanmin Qian (Corresponding author) . Holistic 3-Stage Text-Only Domain Adaptation for E2E ASR: Architecture, Fine-Tuning, and Decoding. Submitted to **IEEE Transactions on Audio, Speech, and Language Processing**, 2025.
3. Leying Zhang, Wangyou Zhang, Chenda Li and Yanmin Qian (Corresponding author) . Scale This, Not That: Investigating Key Dataset Attributes for Efficient Speech Enhancement Scaling. Submitted to **IEEE Signal Processing Letters**, 2025.
4. Zhengyang Chen, Shuai Wang, Bing Han, Yanmin Qian (Corresponding author) . Unsupervised Domain Adaptation for Robust Speaker Verification. Submitted to **IEEE Transactions on Audio, Speech, and Language Processing**, 2025.
5. Pingyi Fan, Anbai Jiang, Shuwei Zhang, Xinhua Zheng, Zhiqiang Lv, Bing Han, Wenrui Liang, Junjie Li, Wei-Qiang Zhang, Yanmin Qian, Xie Chen, Jia Liu. FISHER: A Foundation Model for Multi-Modal Industrial Signal Comprehensive Representation. Submitted to **IEEE Transactions on Industrial Informatics**, 2025.
6. Bei Liu and Yanmin Qian (Corresponding author) . Memory-Efficient Training for Deep Speaker Embedding Learning in Speaker Verification. **IEEE Transactions on Audio, Speech, and Language Processing**, vol. 34, 138-152, 2026.
7. Yoshiki Masuyama, Xuankai Chang, Wangyou Zhang, Samuele Cornell, Zhong-Qiu Wang, Nobutaka Ono, Yanmin Qian and Shinji Watanabe. An End-to-end Integration of Speech Separation and Recognition with Self-Supervised Learning Representation. **Computer Speech & Language**, vol. 95, 101813, 2026.
8. Bing Han, Anbai Jiang, Xinhua Zheng, Wei-Qiang Zhang, Jia Liu, Pingyi Fan and Yanmin Qian (Corresponding author) . Exploring Self-Supervised Audio Models for Generalized Anomalous Sound Detection. **IEEE Transactions on Audio, Speech, and Language Processing**, vol. 33, 4126-4141, 2025.
9. Bei Liu, Tianteng Gu, Haoyu Wang and Yanmin Qian (Corresponding author) . MixPQ: Joint Pruning and Quantization for Speech and Language Foundation Models Compression. **IEEE Transactions on Audio, Speech, and Language Processing**, vol. 33, 4098-4112, 2025.
10. Yidi Jiang, Ruijie Tao, Zhengyang Chen, Yanmin Qian and Haizhou Li. Target Speech Detection with Multimodal Prompts. **IEEE Transactions on Audio, Speech, and Language Processing**, vol. 33, 2788-2802, 2025.

-
11. **Yanmin Qian**, Chenda Li, Wangyou Zhang and Shaoxiong Lin. Contextual Understanding with Contextual Embeddings for Multi-talker Speech Separation and Recognition in A Cocktail Party. **npj Acoustics**, vol. 1, no. 1, 1-12, 2025.
 12. Xuankai Chang, Shinji Watanabe, Marc Delcroix, Tsubasa Ochiai, Wangyou Zhang and **Yanmin Qian** (**Corresponding author**) . Modular-based End-to-End Distant Speech Processing: a case study of far-field automatic speech recognition. **IEEE Signal Processing Magazine**, vol. 41, no. 6, 39-50, 2025.
 13. Junyi Ao, Mehmet Sinan Yıldırım, Ruijie Tao, Meng Ge, Shuai Wang, **Yanmin Qian** and Haizhou Li. USED: Universal Speaker Extraction and Diarization. **IEEE Transactions on Audio, Speech, and Language Processing**, vol. 33, 96-110, 2025.
 14. Shuai Wang, Zhengyang Chen, Kong Aik Lee, **Yanmin Qian** (**Corresponding author**) and Haizhou Li. Overview of Speaker Modeling and Its Applications: From the Lens of Deep Speaker Representation Learning. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 32, 4971-4998, 2024.
 15. Shuai Wang, Zhengyang Chen, Bing Han, Hongji Wang, Chengdong Liang, Binbin Zhang, Xu Xiang, Wen Ding, Johan Rohdin, Anna Silnova, **Yanmin Qian** and Haizhou Li. Advancing Speaker Embedding Learning: Wespeaker Toolkit for Research and Production. **Speech Communication**, vol. 162, 103104, 2024.
 16. Bei Liu, Haoyu Wang and **Yanmin Qian** (**Corresponding author**) . Towards Lightweight Speaker Verification via Adaptive Neural Network Quantization. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 32, 3771-3784, 2024.
 17. Jiahong Li, Chenda Li, Yifei Wu and **Yanmin Qian** (**Corresponding author**) . Unified Cross-Modal Attention: Robust Audio-Visual Speech Recognition and Beyond. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 32, 1941-1953, 2024.
 18. Xun Gong, Yu Wu, Jinyu Li, Shujie Liu, Rui Zhao, Xie Chen and **Yanmin Qian** (**Corresponding author**) . Advanced Long-Content Speech Recognition with Factorized Neural Transducer. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 32, 1803-1815, 2024.
 19. Zhengyang Chen, Bing Han, Shuai Wang and **Yanmin Qian** (**Corresponding author**) . Attention-Based Encoder-Decoder End-to-End Neural Diarization With Embedding Enhancer. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 32, 1636-1649, 2024.
 20. Wei Wang and **Yanmin Qian** (**Corresponding author**) . Universal Cross-Lingual Data Generation for Low Resource ASR. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 32, 973-983, 2024.
 21. Bing Han, Zhengyang Chen, **Yanmin Qian** (**Corresponding author**) . Self-Supervised Learning with Cluster-Aware-DINO for High-Performance Robust Speaker Verification. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 32, 529-541, 2024.
 22. Bei Liu, Zhengyang Chen, **Yanmin Qian** (**Corresponding author**) . Depth-First Neural Architecture with Attentive Feature Fusion for Efficient Speaker Verification. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 31, 1825-1838, 2023.
 23. Shaoxiong Lin, Wangyou Zhang and **Yanmin Qian**. Two-Stage Single-Channel Speech Enhancement with Multi-Frame Filtering. **Applied Sciences**, vol. 1, 1-14, 2023.
 24. Wangyou Zhang, Xuankai Chang, Christoph Boeddeker, Tomohiro Nakatani, Shinji Watanabe and **Yanmin Qian** (**Corresponding author**) . End-to-End Dereverberation, Beamforming, and Speech Recognition in A Cocktail Party. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 30, 3173-3188, 2022.
 25. Sanyuan Chen, Chengyi Wang, Zhengyang Chen, Yu Wu, Shuijie Liu, Zhuo Chen, Jinyu Li, Naoyuki Kanda,

-
- Takuya Yoshioka, Xiong Xiao, Jian Wu, Long Zhou, Shuo Ren, **Yanmin Qian**, Yao Qian, Jian Wu, Michael Zeng and Furu Wei. WavLM: Large-Scale Self-Supervised Pre-Training for Full Stack Speech Processing. **IEEE Journal of Selected Topics in Signal Processing**, vol. 16, 1505-1518, 2022.
26. **Yanmin Qian**, Xun Gong, Houjun Huang. Layer-Wise Fast Adaptation for End-to-End Multi-Accent Speech Recognition. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 30, 2842-2853, 2022.
 27. Chenda Li, Zhuo Chen, **Yanmin Qian (Corresponding author)** . Dual-Path Modeling with Memory Embedding Model for Continuous Speech Separation. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 30, 1508-1520, 2022.
 28. **Yanmin Qian**, Zhikai Zhou. Optimizing Data Usage for Low-Resource Speech Recognition. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 30, 394-403, 2022.
 29. **Yanmin Qian**, Zhengyang Chen and Shuai Wang. Audio-Visual Deep Neural Network for Robust Person Verification. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 29, 1079-1092, 2021.
 30. Jichen Yang, Hongji Wang, Rohan Kumar Das and **Yanmin Qian (Corresponding author)** . Modified Magnitude-Phase Spectrum Information for Spoofing Detection. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 29, 1065-1078, 2021.
 31. Shuai Wang, Yexin Yang, Zhanghao Wu, **Yanmin Qian (Corresponding author)** and Kai Yu. Data Augmentation using Deep Generative Models for Embedding based Speaker Recognition. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 28, 2598-2609, 2020.
 32. Wangyou Zhang, Xuankai Chang, **Yanmin Qian (Corresponding author)** and Shinji Watanabe. Improving End-to-End Single-Channel Multi-Talker Speech Recognition. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 28, 1385-1394, 2020.
 33. Zhehuai Chen, Wenlu Zheng, Yongbin You, **Yanmin Qian**, Kai Yu. Label Synchronous Decoding for Speech Recognition. **Chinese Journal of Computers**, 42 (7), 1511-1523, 2019.
 34. **Yanmin Qian**, Hu Hu, Tian Tan. Data Augmentation using Generative Adversarial Networks for Robust Speech Recognition. **Speech Communication**, vol. 114, 1-9, 2019.
 35. Shuai Wang, Zili Huang, **Yanmin Qian (Corresponding author)** , Kai Yu. Discriminative Neural Embedding Learning for Short-Duration Text-Independent Speaker Verification. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 27, no. 11, 1686-1696, 2019.
 36. **Yanmin Qian**, Xu Xiang. Binary Neural Networks for Speech Recognition. **Frontiers of Information Technology & Electronic Engineering**, 20 (5), 701-715, 2019.
 37. **Yanmin Qian**, Xuankai Chang, Dong Yu. Single-Channel Multi-talker Speech Recognition with Permutation Invariant Training. **Speech Communication**, vol. 104, 1-11, 2018.
 38. Zhehuai Chen, **Yanmin Qian (Corresponding author)** , Kai Yu. Sequence Discriminative Training for Deep Learning based Keyword Spotting. **Speech Communication**, vol. 102, 100-111, 2018.
 39. Heinrich Dinkel, **Yanmin Qian (Corresponding author)** , Kai Yu. Investigating Raw Wave Deep Neural Networks for End-to-End Speaker Spoofing Detection. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 26, no. 11, 2002-2014, 2018.
 40. Tian Tan, **Yanmin Qian (Corresponding author)** , Hu Hu, Ying Zhou, Wen Ding, Kai Yu. Adaptive Very Deep Convolutional Residual Network for Noise Robust Speech Recognition. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 26, no. 8, 1393-1405, 2018.
 41. **Yanmin Qian**, Chao Wen, Xuankai Chang, Shuai Wang, Dong Yu. Past Review, Current Progress, and

-
- Challenges Ahead on the Cocktail Party Problem. **Frontiers of Information Technology & Electronic Engineering**, 19 (1), 40-63, 2018. (Invited Survey Paper)
42. Ruinian Chen, Ying Zhou and **Yanmin Qian (Corresponding author)** . Emotion Recognition using Support Vector Machine and Deep Neural Network. Accepted by Springer Lecture Notes in Computer Science (LNCS), 2018.
 43. **Yanmin Qian**, Nanxin Chen, Heinrich Dinkel, Zhizheng Wu. Deep Feature Engineering for Noise Robust Spoofing Detection. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 25, no. 10, 1942-1955, 2017.
 44. Zhehuai Chen, Yimeng Zhuang, **Yanmin Qian (Corresponding author)** , Kai Yu. Phone Synchronous Speech Recognition with CTC Lattices. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 25, no. 1, 90-101, 2017.
 45. **Yanmin Qian**, Mengxiao Bi, Tian Tan, Kai Yu. Very Deep Convolutional Neural Networks for Noise Robust Speech Recognition. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 24, no. 12, 2263-2276, 2016.
 46. **Yanmin Qian**, Tian Tan and Dong Yu. Neural Network Based Multi-Factor Aware Joint Training for Robust Speech Recognition. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 24, no. 12, 2231-2240, 2016.
 47. Tian Tan, **Yanmin Qian (Corresponding author)** and Kai Yu. Cluster Adaptive Training for Deep Neural Network Based Acoustic Model. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, vol. 24, no. 3, 459-468, 2016.
 48. **Yanmin Qian**, Nanxin Chen, Kai Yu. Deep Features for Automatic Spoofing Detection. **Speech Communication**, vol. 85, 43-52, 2016.
 49. Yuan Liu, **Yanmin Qian (Corresponding author)** , Nanxin Chen, Tianfan Fu, Ya Zhang and Kai Yu. Deep Feature for Text-dependent Speaker Verification. **Speech Communication**, vol. 73, 1-13, 2015.
 50. **Yanmin Qian**, Jia Liu, Michael T Johnson. Efficient embedded speech recognition for very large vocabulary Mandarin car-navigation systems. **IEEE Transactions on Consumer Electronic**, 2009, 55(3):1496-1500.
 51. **Yanmin Qian**, Yuxiang Shan, Linfang Wang, Jia Liu. Improvement comparison of different lattice-based discriminative training methods in Chinese-Monolingual and Chinese-English Bilingual speech recognition. **ACTA AUTOMATICA SINICA**, 2012, 38(7):1162-1168.
 52. **Yanmin Qian**, Ji Xu, Jia Liu. Multi-Stream posterior features and combining Subspace Gmms for low resource LVCSR. **Chinese Journal of Electronics**, 2013, 22(2): 291-295.
 53. **Yanmin Qian**, Jia Liu. Cross-entropy OSF-based voice activity detection algorithm. **Tsinghua Science and Technology**, 2009, 49(10):87-90.
 54. **Yanmin Qian**, Jia Liu. Optimized data selection strategy based unsupervised acoustic modeling for low data resource speech recognition. **Tsinghua Science and Technology**, 2013, 53(7): 1001-1004+1010.
 55. Hong Liu, **Yanmin Qian**, Jia Liu. English speech recognition system on chip. **Tsinghua Science and Technology**, 2011, 16(1):95-99.
 56. Yan Deng, Weiqiang Zhang, **Yanmin Qian**, Jia Liu. Language recognition based on acoustic diversified phone recognizers and phonotactic feature fusion. **IEICE Transactions on Information and Systems**, 2011, E94-D(3):679-689.
 57. Yan Deng, Weiqiang Zhang, **Yanmin Qian**, Jia Liu. Time-Frequency cepstral features and combining

discriminative training for phonotactic language recognition. **Journal of Computers**, 2011, 6(2):178-183.

58. Hua Yuan, **Yanmin Qian**, Junhong Zhao, Jia Liu. Mispronunciation detection with an optimized detection network and multi-layer perception based features. **Tsinghua Science and Technology**, 2012, 52(4):557-560+570.

Conference Papers:

59. Wangyou Zhang, Zhengyang Chen, Chenda Li and **Yanmin Qian**. TS-HuBERT: From Weakly-Supervised To Self-Supervised Speech Pre-training For Target-Speaker Speech Processing. Submitted to **InterSpeech**, Sydney, Australia, 2026.
60. Haiyang Sun, Shujie Hu, Shujie Liu, Lingwei Meng, Hui Wang, Bing Han, Yifan Yang, Yanqing Liu, Sheng Zhao, Yan Lu and **Yanmin Qian**. Zero-Shot Streaming Text to Speech Synthesis with Transducer and Auto-Regressive Modeling. Submitted to **InterSpeech**, Sydney, Australia, 2026.
61. Chenda Li, Wei Wang, Samuele Cornell, Bing Han, Leying Zhang, Zhengyang Chen, Shinji Watanabe and **Yanmin Qian**. Speaking Guided by Listening: Unsupervised Text-to-Speech Generative Model Guided by End-to-End Speech Recognition. Submitted to **InterSpeech**, Sydney, Australia, 2026.
62. Haiyang Sun, Bing Han, Zheng Lian, Leying zhang, Chenda Li, Chenyang Le, Ye Bai, Yi Zhao and **Yanmin Qian**. ContextSpeech: A Large-Scale Real-Human Speech Corpus with Context-Aware Descriptions. Submitted to **InterSpeech**, Sydney, Australia, 2026.
63. Wangyou Zhang, Kohei Saijo, Samuele Cornell, Robin Scheibler, Chenda Li, Zhaoheng Ni, Anurag Kumar, Marvin Sach, Wei Wang, Yihui Fu, Xiaohuai Le, Zhihang Sun, Xiaobin Rong, Rong Chao, Shinji Watanabe, Tim Fingscheidt and **Yanmin Qian**. Datasets and Learnings in the 2024 URGENT Speech Enhancement Challenge. Submitted to **InterSpeech**, Sydney, Australia, 2026.
64. Jiahe Wang, Lei Yang, Chenda Li, Siyi Zhao, Wei Wang, Lufen Tan, Xie Chen and **Yanmin Qian**. Efficient Online Diffusion-based Models for Speech Enhancement. Submitted to **InterSpeech**, Sydney, Australia, 2026.
65. Chenda Li, Qinzhe Hu, Shujie Liu, Yan Lu and **Yanmin Qian**. TF-MOE: Time-Frequency Mixture-of-Experts for Efficient Speech Enhancement and Separation. Submitted to **InterSpeech**, Sydney, Australia, 2026.
66. Xin Zhou, Chenda Li, Hongyu Wang, Wangyou Zhang, Shuai Wang, Qiang Zhou and **Yanmin Qian**. Bridging Speech Separation and Target Speech Extraction: A Universal Model with Encoder-Decoder Attractors. Submitted to **InterSpeech**, Sydney, Australia, 2026.
67. Xuanyu Wang, Wangyou Zhang, Wei Wang, Jianqing Gao and **Yanmin Qian**. NISE: Noise-Injected Speech Enhancement. Submitted to **InterSpeech**, Sydney, Australia, 2026.
68. Bing Han, Chushu Zhou, Yifan Yang, Wei Wang, **Yanmin Qian**. Representation-Regularized Convolutional Audio Transformer for Audio Understanding. Submitted to **IJCAI**, Bremen, Germany, 2026.
69. Haiyang Sun, Chenyang Le, Wei Wang, Leying Zhang, Chuang Li, Bing Han, Chenda Li, Mengxiao Bi and **Yanmin Qian**. Beyond Sentence-level Labels: Integrating Conversational Context and Personal Experience for Natural Emotional Expression. Submitted to **ACL**, San Diego, California, USA, 2026.
70. Wei Wang, Wangyou Zhang, Chenda Li, Jiahe Wang, Marvin Sach, Kohei Saijo, Samuele Cornell, Yihui Fu, Zhaoheng Ni, Mengxiao Bi, Tim Fingscheidt, Shinji Watanabe and **Yanmin Qian**. UrgentMOS: Unified Multi-Metric and Preference Learning for Robust Speech Quality Assessment. Submitted to **ACL**, San Diego, California, USA, 2026.
71. Wen Huang, Yuchen Mao and **Yanmin Qian**. A Data-Centric Approach to Generalizable Speech Deepfake Detection. Submitted to **ACL**, San Diego, California, USA, 2026.
72. Tianteng Gu, Bo Xiao, Ke Zeng and **Yanmin Qian**. NorSA: Accelerate LLM Decoding via Normalized Sparse

-
- Activation. Submitted to **ICLR**, Rio de Janeiro, Brazil, 2026.
73. Marvin Sach, Yihui Fu, Kohei Saijo, Wangyou Zhang, Samuele Cornell, Robin Scheibler, Chenda Li, Zhaoheng Ni, Anurag Kumar, Wei Wang, **Yanmin Qian**, Shinji Watanabe and Tim Fingscheidt. 2025 URGENT Speech Enhancement Challenge Multilingual P.808 Listening Tests: Approach and Results. Accepted by **ICASSP**, Barcelona, Spain, 2026.
 74. Yuchen Mao, Wen Huang and **Yanmin Qian**. Localizing Speech Deepfakes Beyond Transitions via Segment-aware Learning. Accepted by **ICASSP**, Barcelona, Spain, 2026.
 75. Tianhua Li, Chenda Li, Wei Wang, Xin Zhou, Xihui Chen, Jianqing Gao and **Yanmin Qian**. SLM-SS: Speech Language Model for Generative Speech Separation. Accepted by **ICASSP**, Barcelona, Spain, 2026.
 76. Jiahe Wang, Hongyu Wang, Wei Wang, Lei Yang, Chenda Li, Wangyou Zhang, Lufen Tan and **Yanmin Qian**. MEANSE: Efficient Generative Speech Enhancement with Mean Flows. Accepted by **ICASSP**, Barcelona, Spain, 2026.
 77. Hongyu Wang, Chenda Li, Xin Zhou, Shuai Wang and **Yanmin Qian**. USE: A Unified Model for Universal Sound Separation and Extraction. Accepted by **AAAI**, Singapore, 2026.
 78. Haoyu Wang, Bei Liu, Hang Shao, Bo Xiao, Ke Zeng, Guanglu Wan and **Yanmin Qian**. OOQ: Outlier-Oriented Quantization for Efficient Large Language Models. In proceedings of **ASRU**, Honolulu, Hawaii, USA, 2025.
 79. Kaicheng Luo, Xuefei Gong, Yutao Sun, Jinling He, Yujie Hou, Xiaoyang Xing, Huiyan Li, Bing Han and **Yanmin Qian**. StellarTTS: Sparse Temporal Embedding for Low-Latency and Robust Speech Synthesis. In proceedings of **ASRU**, Honolulu, Hawaii, USA, 2025.
 80. Haibin Yu, Jiayi Zhou, Wei Wang, Zhiming Wang, Huijia Zhu and **Yanmin Qian**. Advancing Controllable Music Generation with Latent Rectified Flow Guided by Rhythm and Harmony. In proceedings of **ASRU**, Honolulu, Hawaii, USA, 2025.
 81. Wei Wang, Wangyou Zhang, Chenda Li, Jiatong Shi, Shinji Watanabe and **Yanmin Qian**. Improving Speech Enhancement with Multi-Metric Supervision from Learned Quality Assessment. In proceedings of **ASRU**, Honolulu, Hawaii, USA, 2025.
 82. Jiahe Wang, Chenda Li, Wei Wang, Wangyou Zhang, Samuele Cornell, Marvin Sach, Robin Scheibler, Kohei Saijo, Yihui Fu, Zhaoheng Ni, Anurag Kumar, Tim Fingscheidt, Shinji Watanabe and **Yanmin Qian**. URGENT-PK: Perceptually-Aligned Ranking Model Designed for Speech Enhancement Competition. In proceedings of **ASRU**, Honolulu, Hawaii, USA, 2025.
 83. Chenda Li, Wangyou Zhang, Wei Wang, Robin Scheibler, Kohei Saijo, Samuele Cornell, Yihui Fu, Marvin Sach, Zhaoheng Ni, Anurag Kumar, Tim Fingscheidt, Shinji Watanabe and **Yanmin Qian**. Less is More: Data Curation Matters in Scaling Speech Enhancement. In proceedings of **ASRU**, Honolulu, Hawaii, USA, 2025.
 84. Leying Zhang, Yao Qian, Xiaofei Wang, Manthan Thakker, Dongmei Wang, Jianwei Yu, Haibin Wu, Yuxuan Hu, Jinyu Li, **Yanmin Qian** and Sheng Zhao. CoVoMix2: Advancing Zero-Shot Dialogue Generation with Fully Non-Autoregressive Flow Matching. In proceedings of **NeurIPS**, San Diego, California, USA, 2025.
 85. Tianteng Gu, Bei Liu, Bo Xiao, Ke Zeng, Jiacheng Liu and **Yanmin Qian**. DenoiseRotator: Enhance Pruning Robustness for LLMs via Importance Concentration. In proceedings of **NeurIPS**, San Diego, California, USA, 2025.
 86. Chenyang Le, Bing Han, Jinshun Li, Chen Songyong and **Yanmin Qian**. SimulMEGA: MoE Routers are Advanced Policy Makers for Simultaneous Speech Translation. In proceedings of **NeurIPS**, San Diego, California, USA, 2025.

-
87. Tingxiao Zhou, Leying Zhang, Zhengyang Chen and **Yanmin Qian**. Training Text-to-Speech Model with Purely Synthetic Data: Feasibility, Sensitivity, and Generalization Capability. In proceedings of **NCMMSC**, Zhenjiang, China, 2025.
 88. Shaoxiong Lin, Wangyou Zhang, Leying Zhang and **Yanmin Qian**. Improving SI-SNR Loss with Frequency Weighting for Speech Enhancement. In proceedings of **NCMMSC**, Zhenjiang, China, 2025.
 89. Xin Zhou, Chenda Li and Yanmin Qian. MT-Skim: Skipping Memory Transformer for Speech Separation. In proceedings of **NCMMSC**, Zhenjiang, China, 2025.
 90. Tianhua Li, Chenda Li and Yanmin Qian. Multi-talker Diarization and Speech Transcription with End-to-end Speech LLM. In proceedings of **NCMMSC**, Zhenjiang, China, 2025.
 91. Wen Huang, Yanmei Gu, Zhiming Wang, Huijia Zhu and **Yanmin Qian**. SpeechFake: A Large-Scale Multilingual Speech Deepfake Dataset Toward Cutting-Edge Speech Generation Methods. In proceedings of **ACL**, Vienna, Austria, 2025: 9985-9998.
 92. Yinfeng Xia, Huiyan Li, Chenyang Le, Manhong Wang, Yutao Sun, Xingyang Ma and **Yanmin Qian**. MFLA: Monotonic Finite Look-ahead Attention for Streaming Speech Recognition. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 4408-4412.
 93. Jiahong Li, Yiwen Shao, Jianheng Zhuo, Chenda Li, Liliang Tang, Dong Yu and **Yanmin Qian**. Efficient Multilingual ASR Finetuning via LoRA Language Experts. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 1138-1142.
 94. Wangyou Zhang, Kohei Saijo, Samuele Cornell, Robin Scheibler, Chenda Li, Zhaoheng Ni, Anurag Kumar, Marvin Sach, Wei Wang, Yihui Fu, Shinji Watanabe, Tim Fingscheidt and **Yanmin Qian**. Lessons Learned from the URGENT 2024 Speech Enhancement Challenge. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 853-857.
 95. Yihan Liu, Zhengyang Chen, Leying Zhang and **Yanmin Qian**. E2E-BPVC: End-to-End Background-Preserving Voice Conversion via In-Context Learning. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 1378-1382.
 96. Wen Huang, Xuechen Liu, Xin Wang, Junichi Yamagishi and **Yanmin Qian**. From Sharpness to Better Generalization for Speech Deepfake Detection. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 5338-5342.
 97. Haoxiang Hou, Xun Gong, Wangyou Zhang, Wei Wang and **Yanmin Qian**. Ranking and Selection of Bias Words for Contextual Bias Speech Recognition. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 5183-5187.
 98. Chenyang Le, Yinfeng Xia, Huiyan Li, Manhong Wang, Yutao Sun, Xingyang Ma and **Yanmin Qian**. Novel Parasitic Dual-Scale Modeling for Efficient and Accurate Multilingual Speech Translation. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 908-912.
 99. Tianteng Gu, Bei Liu, Haoyu Wang and **Yanmin Qian**. Ultra-Low Bit Post-Training Quantization of Large Speech Models via K-Means Clustering and Mixed Precision Allocation. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 1988-1992.
 100. Xun Gong, Anqi Lv, Wangyou Zhang, Zhiming Wang, Huijia Zhu and **Yanmin Qian**. BR-ASR: Efficient and Scalable Bias Retrieval Framework for Contextual Biasing ASR in Speech LLM. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 4043-4047.
 101. Siyi Zhao, Wei Wang and **Yanmin Qian**. Lightweight Front-end Enhancement for Robust ASR via Frame

-
- Sampling and Sub-Band Pruning. In proceedings of **InterSpeech**, Rotterdam, The Netherlands, 2025: 3409-3413.
102. Leying Zhang, Wangyou Zhang, Zhengyang Chen and **Yanmin Qian**. Advanced Zero-Shot Text-to-Speech for Background Removal and Preservation with Controllable Masked Speech Prediction. In proceedings of **ICASSP**, Hyderabad, India, 2025: 1-5.
103. Tianteng Gu, Bei Liu and **Yanmin Qian**. Efficient Pruning for Large-Scale Seq2Seq Speech Models without Back-Propagation. In proceedings of **ICASSP**, Hyderabad, India, 2025: 1-5.
104. Haitian Lu, Gaofeng Chen, Liuping Luo, Leying Zhang, **Yanmin Qian** and Pengyuan Zhang. SLIDE: Integrating Speech Language Model with LLM for Spontaneous Spoken Dialogue Generation. In proceedings of **ICASSP**, Hyderabad, India, 2025: 1-5.
105. Zhengyang Chen, Bing Han, Shuai Wang, Yidi Jiang and **Yanmin Qian**. Flow-TSVAD: Target-Speaker Voice Activity Detection via Latent Flow Matching for Speaker Diarization. In proceedings of **ICASSP**, Hyderabad, India, 2025: 1-5.
106. Wen Huang, Yanmei Gu, Zhiming Wang, Huijia Zhu and **Yanmin Qian**. Generalizable Audio Deepfake Detection via Latent Space Refinement and Augmentation. In proceedings of **ICASSP**, Hyderabad, India, 2025: 1-5.
107. Wei Wang, Siyi Zhao and **Yanmin Qian**. Advancing Non-intrusive Suppression on Enhancement Distortion for Noise Robust ASR. In proceedings of **ICASSP**, Hyderabad, India, 2025: 1-5.
108. Bing Han, Wen Huang, Zhengyang Chen, Anbai Jiang, Pingyi Fan, Cheng Lu, Zhiqiang Lv, Jia Liu, Wei-Qiang Zhang and **Yanmin Qian**. Data-Efficient Low-Complexity Acoustic Scene Classification via Distilling and Progressive Pruning. In proceedings of **ICASSP**, Hyderabad, India, 2025: 1-5.
109. Bing Han, Long Zhou, Shujie Liu, Lingwei Meng, **Yanmin Qian**, Yanqing Liu, Sheng Zhao, Jinyu Li and Furu Wei. VALL-E R: Robust and Efficiency Zero-Shot Text-to-Speech via Monotonic Alignment. In proceedings of **NeurIPS Workshop**, Vancouver, Canada, 2024.
110. Leying zhang, Yao Qian, Long Zhou, Shujie Liu, Dongmei Wang, Xiaofei Wang, Midia Yousefi, **Yanmin Qian**, Jinyu Li, Lei He, Sheng Zhao and Michael Zeng. CoVoMix: Advancing Zero-Shot Speech Generation for Human-like Multi-talker Conversations. In proceedings of **NeurIPS**, Vancouver, Canada, 2024: 100291-100317.
111. Chenyang Le, Yao Qian, Dongmei Wang, Long Zhou, Shujie Liu, Xiaofei Wang, Midia Yousefi, **Yanmin Qian**, Jinyu Li and Michael Zeng. TransVIP: Speech to Speech Translation System with Voice and Isochrony Preservation. In proceedings of **NeurIPS**, Vancouver, Canada, 2024: 89682-89705.
112. Wangyou Zhang, Robin Scheibler, Kohei Saijo, Samuele Cornell, Chenda Li, Zhaoheng Ni, Jan Pirklbauer, Marvin Sach, Shinji Watanabe, Tim Fingscheidt and **Yanmin Qian**. NeurIPS 2024 Competition Proposal: URGENT Challenge. In proceedings of **NeurIPS 2024 Competition Track**, Vancouver, Canada, 2024.
113. Leying Zhang, Yao Qian, Linfeng Yu, Heming Wang, Hemin Yang, Shujie Liu, Long Zhou, **Yanmin Qian**. DDTSE: Discriminative Diffusion Model for Target Speech Extraction. In proceedings of **SLT**, Macao, China, 2024: 299-306.
114. Hang Shao, Bei Liu, Wei Wang, Xun Gong and **Yanmin Qian**. DQ-Whisper: Joint Distillation and Quantization for Efficient Multilingual Speech Recognition. In proceedings of **SLT**, Macao, China, 2024: 245-251.
115. Chenda Li, Samuele Cornell, Shinji Watanabe and **Yanmin Qian**. Diffusion-based Generative Modeling with Discriminative Guidance for Streamable Speech Enhancement. In proceedings of **SLT**, Macao, China, 2024: 338-345.
116. Zhengyang Chen, Shuai Wang, Mingyang Zhang, Xuechen Liu, Junichi Yamagishi and **Yanmin Qian**.

-
- Disentangling the Prosody and Semantic Information with Pre-trained Model for In-Context Learning based Zero-Shot Voice Conversion. In proceedings of **SLT**, Macao, China, 2024: 708-714.
117. Jiahe Wang, Shuai Wang, Junjie Li, Ke Zhang, **Yanmin Qian** and Haizhou Li. Enhancing Speaker Extraction Through Rectifying Target Confusion. In proceedings of **SLT**, Macao, China, 2024: 354-361.
118. Xihu Zheng, Anbai Jiang, Bing Han, **Yanmin Qian**, Pingyi Fan, Jia Liu and Wei-Qiang Zhang. Improving Anomalous Sound Detection via Low-Rank Adaptation Fine-Tuning of Pre-Trained Audio Models. In proceedings of **SLT**, Macao, China, 2024: 979-984.
119. Tingxiao Zhou, Leying Zhang and **Yanmin Qian**. Knowledge Distillation from Discriminative Model to Generative Model with Parallel Architecture for Speech Enhancement. In proceedings of **ISCSLP**, Beijing, China, 2024: 179-183.
120. Wen Huang, Bing Han, Zhengyang Chen, Shuai Wang and **Yanmin Qian**. Prototype and Instance Contrastive Learning for Unsupervised Domain Adaptation in Speaker Verification. In proceedings of **ISCSLP**, Beijing, China, 2024: 383-387.
121. Haoxiang Hou, Xun Gong and **Yanmin Qian**. ConMamba: A Convolution-Augmented Mamba Encoder Model for Efficient End-to-End ASR Systems. In proceedings of **ISCSLP**, Beijing, China, 2024: 711-715.
122. Zhengyang Chen, Shuai Wang, Bing Han and **Yanmin Qian**. Combining the Self-supervised Learning and Adversarial Training based Domain Adaptation for Speaker Verification. In proceedings of **ISCSLP**, Beijing, China, 2024: 701-705.
123. Xin Zhou, Wangyou Zhang, Chenda Li and **Yanmin Qian**. Insights from Hyperparameter Scaling of Online Speech Separation. In proceedings of **ISCSLP**, Beijing, China, 2024: 561-565.
124. Siyi Zhao, Wei Wang and **Yanmin Qian**. Band-wise Front-end Distortion Suppression for Robust Speech Recognition. In proceedings of **ISCSLP**, Beijing, China, 2024: 681-685.
125. Tianteng Gu, Bei Liu, Hang Shao and **Yanmin Qian**. SparseWAV: Fast and Accurate One-Shot Unstructured Pruning for Large Speech Foundation Models. In proceedings of **InterSpeech**, Kos Island, Greece, 2024: 4498-4502.
126. Xun Gong, Anqi Lv, Zhiming Wang and **Yanmin Qian**. Contextual Biasing Speech Recognition in Speech-based Large Language Model. In proceedings of **InterSpeech**, Kos Island, Greece, 2024: 257-261.
127. Wangyou Zhang, Robin Scheibler, Kohei Saijo, Samuele Cornell, Chenda Li, Zhaoheng Ni, Jan Pirklbauer, Marvin Sach, Shinji Watanabe, Tim Fingscheidt and **Yanmin Qian**. URGENT Challenge: Universality, Robustness, and Generalizability for speech Enhancement. In proceedings of **InterSpeech**, Kos Island, Greece, 2024: 4868-4872.
128. Wangyou Zhang, Kohei Saijo, Jee-weon Jung, Chenda Li, Shinji Watanabe and **Yanmin Qian**. Beyond Performance Plateaus: A Comprehensive Study on Scalability in Speech Enhancement. In proceedings of **InterSpeech**, Kos Island, Greece, 2024: 1740-1744.
129. Zhengyang Chen, Xuechen Liu, Erica Cooper, Junichi Yamagishi and **Yanmin Qian**. Generating Speakers by Prompting Listener Impressions for Pre-trained Multi-Speaker Text-to-Speech Systems. In proceedings of **InterSpeech**, Kos Island, Greece, 2024: 4428-4432.
130. Anbai Jiang, Bing Han, Zhiqiang Lv, Yufeng Deng, Wei-Qiang Zhang, Xie Chen, **Yanmin Qian**, Jia Liu and Pingyi Fan. AnoPatch: Towards Better Consistency in Machine Anomalous Sound Detection. In proceedings of **InterSpeech**, Kos Island, Greece, 2024: 107-111.
131. Shuai Wang, Ke Zhang, Shaoxiong Lin, Junjie Li, Xuefei Wang, Meng Ge, Jianwei Yu, **Yanmin Qian** and

-
- Haizhou Li. WeSep: A Scalable and Flexible Toolkit Towards Generalizable Target Speaker Extraction. In proceedings of **InterSpeech**, Kos Island, Greece, 2024: 4273-4277.
132. Bing Han, Junyu Dai, Xuchen Song, Weituo Hao, Xinyan He, Dong Guo, Jitong Chen, Yuxuan Wang and **Yanmin Qian**. InstructME: An Instruction Guided Music Edit Framework with Latent Diffusion Models. In proceedings of **IJCAI**, Jeju, Korea, 2024: 5835-5843.
133. Wen Huang, Anbai Jiang, Bing Han, Xinhua Zheng, Yihong Qiu, Wenxi Chen, Yuzhe Liang, Pingyi Fan, Wei-Qiang Zhang, Cheng Lu, Xie Chen, Jia Liu and **Yanmin Qian**. Semi-supervised Acoustic Scene Classification with Test-Time Adaptation. In proceedings of **ICMEW**, Niagra Falls, Canada, 2024: 1-5.
134. Yuzhe Liang, Wenxi Chen, Anbai Jiang, Yihong Qiu, Xinhua Zheng, Wen Huang, Bing Han, **Yanmin Qian**, Pingyi Fan, Wei-Qiang Zhang, Cheng Lu, Jia Liu and Xie Chen. Improving Acoustic Scene Classification via Self-Supervised and Semi-Supervised Learning with Efficient Audio Transformer. In proceedings of **ICMEW**, Niagra Falls, Canada, 2024: 1-6.
135. Wangyou Zhang, Jee-weon Jung and **Yanmin Qian**. Improving Design of Input Condition Invariant Speech Enhancement. In proceedings of **ICASSP**, Seoul, Korea, 2024: 10696-10700.
136. Bing Han, Zhiqiang Lv, Anbai Jiang, Wen Huang, Zhengyang Chen, Yufeng Deng, Jiawei Ding, Cheng Lu, Wei-Qiang Zhang, Pingyi Fan, Jia Liu and **Yanmin Qian**. Exploring Large Scale Pre-trained Models for Robust Machine Anomalous Sound Detection. In proceedings of **ICASSP**, Seoul, Korea, 2024: 1326-1330.
137. Wen Huang, Bing Han, Shuai Wang, Zhengyang Chen and **Yanmin Qian**. Robust Cross-domain Speaker Verification with Multi-level Domain Adapters. In proceedings of **ICASSP**, Seoul, Korea, 2024: 11781-11785.
138. Shuai Wang, Qibing Bai, Qi Liu, Jianwei Yu, Zhengyang Chen, Bing Han, **Yanmin Qian** and Haizhou Li. Leveraging In-the-Wild Data for Effective Self-Supervised Pretraining in Speaker Recognition. In proceedings of **ICASSP**, Seoul, Korea, 2024: 10901-10905.
139. Yidi Jiang, Zhengyang Chen, Ruijie Tao, Liqun Deng, **Yanmin Qian**, Haizhou Li. Prompt-driven Target Speech Diarization. In proceedings of **ICASSP**, Seoul, Korea, 2024: 11086-11090.
140. Hang Shao, Bei Liu, Bo Xiao, Ke Zeng, Guanglu Wan and **Yanmin Qian**. One-Shot Sensitivity-Aware Mixed Sparsity Pruning for Large Language Models. In proceedings of **ICASSP**, Seoul, Korea, 2024: 11296-11300.
141. Linfeng Yu, Wangyou Zhang, Chenpeng Du, Leying Zhang, Zheng Liang and **Yanmin Qian**. Generation-based Target Speech Extraction with Speech Discretization and Vocoder. In proceedings of **ICASSP**, Seoul, Korea, 2024: 12612-12616.
142. Bing Han, Zhengyang Chen, Bei Liu and **Yanmin Qian**. Unsupervised Domain Adaptation with Pseudo Labeling for Robust Speaker Verification. In proceedings of **NCMMSC**, Suzhou, China, 2023.
143. Xinyuan Wu, Yifei Wu, Chenda Li and **Yanmin Qian**. Adaptive Alignment for Audio-Visual Speech Separation. In proceedings of **NCMMSC**, Suzhou, China, 2023.
144. Wangyou Zhang, Lei Yang and **Yanmin Qian**. Exploring Time-Frequency Domain Target Speaker Extraction For Causal and Non-Causal Processing. In proceedings of **ASRU**, Taipei, Taiwan, 2023.
145. Wangyou Zhang, Kohei Saijo, Zhong-Qiu Wang, Shinji Watanabe and **Yanmin Qian**. Toward Universal Speech Enhancement for Diverse Input Conditions. In proceedings of **ASRU**, Taipei, Taiwan, 2023.
146. Shaoxiong Lin, Chao Zhang and **Yanmin Qian**. Improving Speech Enhancement Using Audio Tagging Knowledge from Pre-trained Representations and Multi-task Learning. In proceedings of **ASRU**, Taipei, Taiwan, 2023.
147. Chang Chen, Xun Gong and **Yanmin Qian**. Efficient Text-Only Domain Adaptation for CTC-based ASR. In

-
- proceedings of **ASRU**, Taipei, Taiwan, 2023.
148. Dongning Yang, Wei Wang and **Yanmin Qian**. FAT-HuBERT: Front-end Adaptive Training of Hidden-unit BERT for Distortion-Invariant Robust Speech Recognition. In proceedings of **ASRU**, Taipei, Taiwan, 2023.
149. Yuhao Liang, Mohan Shi, Fan Yu, Yangzhe Li, Shiliang Zhang, Zhihao Du, Lei Xie, **Yanmin Qian**, Jian Wu, Zhuo Chen, Kong Aik Lee, Zhijie Yan and Hui Bu. The Second Multi-Channel Multi-Party Meeting Transcription Challenge (M2MeT 2.0): A Benchmark for Speaker-Attributed ASR. In proceedings of **ASRU**, Taipei, Taiwan, 2023.
150. Chenyang Le, Yao Qian, Long Zhou, Shujie Liu, **Yanmin Qian**, Michael Zeng and Xuedong Huang. ComSL: A Composite Speech-Language Model for End-to-End Speech-to-Text Translation. In proceedings of **NeurIPS**, New Orleans, USA, 2023.
151. Zhiqiang Lv, Bing Han, Zhengyang Chen, **Yanmin Qian**, Jiawei Ding and Jia Liu. Unsupervised Anomalous Detection based on Unsupervised Pretrained Models. In proceedings of **DCASE-Challenge**, 2023.
152. Yoshiki Masuyama, Xuankai Chang, Wangyou Zhang, Samuele Cornell, Zhong-Qiu Wang, Nobutaka Ono, **Yanmin Qian** and Shinji Watanabe. Exploring the Integration of Speech Separation and Recognition with Self-Supervised Learning Representation. In proceedings of **WASPAA**, New Paltz, USA, 2023.
153. Bei Liu and **Yanmin Qian**. ECAPA++: Fine-Grained Deep Embedding Learning for TDNN based Speaker Verification. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 3132-3136.
154. Bei Liu, Haoyu Wang and **Yanmin Qian**. Extremely Low Bit Quantization for Mobile Speaker Verification Systems Under 1MB Memory. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 1973-1977.
155. Bei Liu and **Yanmin Qian**. Reversible Neural Networks for Memory-Efficient Speaker Verification. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 3127-3131.
156. Haoyu Wang, Bei Liu, Yifei Wu and **Yanmin Qian**. Adaptive Neural Network Quantization for Lightweight Speaker Verification. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 5331-5335.
157. Chenda Li, Yao Qian, Zhuo Chen, Naoyuki Kanda, Dongmei Wang, Takuya Yoshioka, **Yanmin Qian** and Michael Zeng. Adapting Multi-Lingual ASR Models for Handling Multiple Talkers. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 1314-1318.
158. Zhengyang Chen, Bing Han, Xu Xiang, Houjun Huang, Bei Liu and **Yanmin Qian**. Build A SRE Challenge System: Lessons From VOXSRC 2022 and CNSRC 2022. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 3202-3206.
159. Zhengyang Chen, Bing Han, Shuai Wang and **Yanmin Qian**. Attention-based Encoder-Decoder Network for End-to-End Neural Speaker Diarization with Target Speaker Attractor. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 3552-3556.
160. Linfeng Yu, Wangyou Zhang, Chenda Li and **Yanmin Qian**. Overlap Aware Continuous Speech Separation without Permutation Invariant Training. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 3512-3516.
161. Wangyou Zhang and **Yanmin Qian**. Weakly-Supervised Speech Pre-training: A Case Study on Target Speech Recognition. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 3517-3521.
162. Wei Wang and **Yanmin Qian**. UniSplice: Universal Cross-Lingual Data Splicing for Low-Resource ASR. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 2253-2257.
163. Wei Wang, Xun Gong, Hang Shao, Dongning Yang and **Yanmin Qian**. Text Only Domain Adaptation with Phoneme Guided Data Splicing for End-to-End Speech Recognition. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 3347-3351.

-
164. Zhilong Zhang, Wei Wang and **Yanmin Qian**. Fast and Efficient Multilingual Self-Supervised Pre-training for Low-Resource Speech Recognition. In proceedings of **InterSpeech**, Dublin, Ireland, 2023: 2248-2252.
 165. Bing Han, Wen Huang, Zhengyang Chen and **Yanmin Qian**. Improving DINO-based Self-Supervised Speaker Verification with Progressive Cluster-aware Training. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 166. Yifei Wu, Chenda Li and **Yanmin Qian**. Light-Weight VisualVoice: Neural Network Quantization on Audio-Visual Speech Separation. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 167. Haoyu Wang, Bei Liu, Yifei Wu, Zhengyang Chen and **Yanmin Qian**. Low-Bit Neural Network Quantization for Speaker Verification. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 168. Bing Han, Zhengyang Chen and **Yanmin Qian**. Exploring Binary Classification Loss for Speaker Verification. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 169. Chenda Li, Yao Qian, Zhuo Chen, Dongmei Wang, Takuya Yoshioka, Shujie Liu, **Yanmin Qian** and Michael Zeng. Target Sound Extraction with Variable Cross-Modality Clues. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 170. Chenda Li, Yifei Wu and **Yanmin Qian**. Predictive SKIM: Contrastive Predictive Coding for Low-Latency Online Speech Separation. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 171. Haibin Yu, Yuxuan Hu, Yao Qian, Ma Jin, Linqun Liu, Shujie Liu, Yu Shi, **Yanmin Qian**, Edward Lin and Michael Zeng. Code-Switching Text Generation and Injection in Mandarin-English ASR. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 172. Hang Shao, Tian Tan, Wei Wang, Xun Gong and **Yanmin Qian**. Joint Discriminator and Transfer based Fast Domain Adaptation for End-to-End Speech Recognition. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 173. Jiahong Li, Chenda Li, Yifei Wu and **Yanmin Qian**. Robust Audio-Visual ASR with Unified Cross-Modal Attention. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 174. Leying Zhang, Zhengyang Chen and **Yanmin Qian**. Adaptive Large Margin Fine-tuning for Robust Speaker Verification. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 175. Hongji Wang, Chengdong Liang, Shuai Wang, Zhengyang Chen, Binbin Zhang, Xu Xiang, Yanlei Deng and **Yanmin Qian**. Wespeaker: A Research and Production Oriented Speaker Embedding Learning Toolkit. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 176. Wei Wang and **Yanmin Qian**. HuBert-AGG: Aggregated Representation Distillation of Hidden-Unit Bert for Robust Speech Recognition. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 177. Xun Gong, Wei Wang, Hang Shao, Xie Chen and **Yanmin Qian**. Factorized AED: Factorized Attention-based Encoder-Decoder for Text-Only Domain Adaptive ASR. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 178. Xun Gong, Yu Wu, Jinyu Li, Shujie Liu, Rui Zhao, Xie Chen and **Yanmin Qian**. LONGFNT: Long-Form Speech Recognition with Factorized Neural Transducer. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 179. Shaoxiong Lin, Wangyou Zhang and **Yanmin Qian**. Two Stage Single-Channel Speech Enhancement with Multi-Frame Filtering. In proceedings of **NCMMSC**, Hefei, China, 2022.
 180. Zhikai Zhou, Shuang Cao, Zhengyang Chen, Bei Liu, Ming Xia, Hong Jiang and **Yanmin Qian**. Medical Difficult Airway Detection Using Speech Technology. In proceedings of **ISCSLP**, Singapore, 2022: 349-353.
 181. Wei Wang, Wangyou Zhang, Shaoxiong Lin and **Yanmin Qian**. Text-Informed Knowledge Distillation for Robust Speech Enhancement and Recognition. In proceedings of **ISCSLP**, Singapore, 2022: 334-338.
 182. Houjun Huang and **Yanmin Qian**. Speaking Style Compensation on Synthetic Audio for Robust Keyword

-
- Spotting. In proceedings of **ISCSLP**, Singapore, 2022: 448-452.
183. Bowen Qu, Chenda Li, Jinfeng Bai and **Yanmin Qian**. Improving Speech Separation with Knowledge Distilled from Self-Supervised Pre-trained Models. In proceedings of **ISCSLP**, Singapore, 2022: 329-333.
184. Tao Liu, Xu Xiang, Zhengyang Chen, Bing Han, Kai Yu and **Yanmin Qian**. The X-Lance Speaker Diarization System for the Conversational Short-Phrase Speaker Diarization Challenge 2022. In proceedings of **ISCSLP**, Singapore, 2022: 498-501.
185. Gaofeng Cheng, Yifan Chen, Runyan Yang, Qingxuan Li, Zehui Yang, Lingxuan Ye, Pengyuan Zhang, Qingqing Zhang, Lei Xie, **Yanmin Qian**, Kong Aik Lee, Yonghong Yan. The Conversational Short-Phrase Speaker Diarization (CSSD) Task: Dataset, Evaluation Metric and Baselines. In proceedings of **ISCSLP**, Singapore, 2022: 488-492.
186. Robin Scheibler, Wangyou Zhang, Xuankai Chang, Shinji Watanabe and **Yanmin Qian**. End-to-End Multi-speaker ASR with Independent Vector Analysis. In proceedings of **SLT**, Doha, Qatar, 2022: 496-501.
187. Zhengyang Chen, Yao Qian, Bing Han, **Yanmin Qian** and Michael Zeng. A Comprehensive Study on Self-Supervised Distillation for Speaker Representation Learning. In proceedings of **SLT**, Doha, Qatar, 2022: 599-604.
188. Yen-Ju Lu, Xuankai Chang, Chenda Li, Wangyou Zhang, Samuele Cornell, Zhaoheng Ni, Yoshiki Masuyama, Brian Yan, Robin Scheibler, Zhong-Qiu Wang, Yu Tsao, **Yanmin Qian**, Shinji Watanabe. ESPnet-SE++: Speech Enhancement for Robust Speech Recognition, Translation, and Understanding. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 5458-5462.
189. Tao Liu, Shuai Fan, Xu Xiang, Hongbo Song, Shaoxiong Lin, Jiaqi Sun, Tianyuan Han, Siyuan Chen, Binwei Yao, Sen Liu, Yifei Wu, **Yanmin Qian** and Kai Yu. MSDWILD: Multi-Modal Speaker Diarization Dataset in the Wild. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 1476-1480.
190. Bei Liu, Zhengyang Chen and **Yanmin Qian**. Attentive Feature Fusion for Robust Speaker Verification. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 286-290.
191. Bei Liu, Zhengyang Chen and **Yanmin Qian**. Dual Path Embedding Learning for Speaker Verification with Triplet Attention. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 291-295.
192. Bei Liu, Zhengyang Chen, Shuai Wang, Haoyu Wang, Bing Han and **Yanmin Qian**. DF-ResNet: Boosting Speaker Verification Performance with Depth-First Design. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 296-300.
193. Bing Han, Zhengyang Chen and **Yanmin Qian**. Self-Supervised Speaker Verification using Dynamic Loss-Gate and Label Correction. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 4780-4784.
194. Leying Zhang, Zhengyang Chen and **Yanmin Qian**. Enroll-aware Attentive Statistics Pooling for Target Speaker Verification. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 311-315.
195. Xun Gong, Zhikai Zhou and **Yanmin Qian**. Knowledge Transfer and Distillation from Autoregressive to Non-Autoregressive Speech Recognition. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 2618-2622.
196. Wangyou Zhang, Zhuo Chen, Naoyuki Kanda, Shujie Liu, Jinyu Li, Sefik Emre Eskimez, Takuya Yoshioka, Xiong Xiao, Zhong Meng, **Yanmin Qian** and Furu Wei. Separating Long-Form Speech with Group-Wise Permutation Invariant Training. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 5383-5387.
197. Fan Yu, Shiliang Zhang, Pengcheng Guo, Yihui Fu, Zhihao Du, Siqi Zheng, Weilong Huang, Lei Xie, Zheng-Hua Tan, DeLiang Wang, **Yanmin Qian**, Kong Aik Lee, Zhijie Yan, Bin Ma, Xin Xu and Hui Bu. Summary on The ICASSP 2022 Multi-Channel Multi-Party Meeting Transcription Grand Challenge. In proceedings of **ICASSP**, Singapore, 2022: 9156-9160.

-
198. Wei Wang, Xun Gong, Yifei Wu, Zhikai Zhou, Chenda Li, Wangyou Zhang, Bing Han and **Yanmin Qian**. The SJTU System for MultiModal Information based Speech Processing Challenge 2021. In proceedings of **ICASSP**, Singapore, 2022: 9261-9265.
 199. Bei Liu, Haoyu Wang, Zhengyang Chen, Shuai Wang and **Yanmin Qian**. Self-Knowledge Distillation via Feature Enhancement for Speaker Verification. In proceedings of **ICASSP**, Singapore, 2022: 7542-7546.
 200. Bing Han, Zhengyang Chen, Bei Liu and **Yanmin Qian**. MLP-SVNET: A Multi-Layer Perceptrons based Network for Speaker Verification. In proceedings of **ICASSP**, Singapore, 2022:7522-7526.
 201. Bing Han, Zhengyang Chen and **Yanmin Qian**. Local Information Modeling with Self-Attention for Speaker Verification. In proceedings of **ICASSP**, Singapore, 2022: 6727-6731.
 202. Chenda Li, Lei Yang, Weiqin Wang and **Yanmin Qian**. SKIM: Skipping Memory LSTM for Low-Latency Real-Time Continuous Speech Separation. In proceedings of **ICASSP**, Singapore, 2022: 681-685.
 203. Wei Wang, Shuo Ren, Yao Qian, Shujie Liu, Yu Shi, **Yanmin Qian** and Michael Zeng. Optimizing Alignment of Speech and Language Latent Spaces for End-to-End Speech Recognition and Understanding. In proceedings of **ICASSP**, Singapore, 2022: 7802-7806.
 204. Yifei Wu, Chenda Li, Jinfeng Bai, Zhongqin Wu and **Yanmin Qian**. Time-Domain Audio-Visual Speech Separation on Low Quality Videos. In proceedings of **ICASSP**, Singapore, 2022: 256-260.
 205. Zhengyang Chen, Sanyuan Chen, Yu Wu, Yao Qian, Chengyi Wang, Shuijie Liu, **Yanmin Qian** and Michael Zeng. Large-Scale Self-Supervised Speech Representation Learning for Automatic Speaker Verification. In proceedings of **ICASSP**, Singapore, 2022: 6147-6151.
 206. Zhikai Zhou, Wei Wang, Wangyou Zhang and **Yanmin Qian**. Exploring Effective Data Utilization for Low-Resource Speech Recognition. In proceedings of **ICASSP**, Singapore, 2022: 8192-8196.
 207. Zhikai Zhou, Tian Tan and **Yanmin Qian**. Punctuation Prediction for Streaming On-Device Speech Recognition. In proceedings of **ICASSP**, Singapore, 2022: 7277-7281.
 208. Wangyou Zhang, Jing Shi, Chenda Li, Shinji Watanabe and **Yanmin Qian**. Closing the Gap Between Time-domain Multi-channel Speech Enhancement on Real and Simulation Conditions. In proceedings of **WASPAA**, New Paltz, USA, 2021: 146-150.
 209. Zhengxi Liu and **Yanmin Qian**. Basis-MelGAN: Efficient Neural Vocoder Based on Audio Decomposition. In proceedings of **InterSpeech**, Brno, Czechia, 2021: 2222-2226.
 210. Bing Han, Zhengyang Chen, Zhikai Zhou and **Yanmin Qian**. The SJTU System for Short-Duration Speaker Verification Challenge 2021. In proceedings of **InterSpeech**, Brno, Czechia, 2021: 2332-2336.
 211. Leying Zhang, Zhengyang Chen and **Yanmin Qian**. Knowledge Distillation from Multi-Modality to Single-Modality for Person Verification. In proceedings of **InterSpeech**, Brno, Czechia, 2021: 1897-1901.
 212. Xun Gong, Yizhou Lu, Zhikai Zhou and **Yanmin Qian**. Layer-Wise Fast Adaptation for End-to-End Multi-Accent Speech Recognition. In proceedings of **InterSpeech**, Brno, Czechia, 2021: 1274-1278.
 213. Yifei Wu, Chenda Li, Song Yang, Zhongqin Wu and **Yanmin Qian**. Audio-Visual Multi-Talker Speech Recognition in a Cocktail Party. In proceedings of **InterSpeech**, Brno, Czechia, 2021: 3021-3025.
 214. Xian Shi, Fan Yu, Yizhou Lu, Yuhao Liang, Qiangze Feng, Daliang Wang, **Yanmin Qian** and Lei Xie. The Accented English Speech Recognition Challenge 2020: Open Datasets, Tracks, Baselines, Results and Methods. In proceedings of **ICASSP**, Toronto, Canada, 2021: 6918-6922.
 215. Chenda Li, Zhuo Chen, Yi Luo, Cong Han, Tianyan Zhou, Keisuke Kinoshita, Marc Delcroix, Shinji Watanabe and **Yanmin Qian**. Dual-path Modeling for Long Recording Speech Separation in Meetings. In proceedings of

-
- ICASSP**, Toronto, Canada, 2021: 5739-5743.
216. Chenpeng Du, Bing Han, Shuai Wang, **Yanmin Qian** and Kai Yu. SYNAUG: Synthesis-based Data Augmentation for Text-Dependent Speaker Verification. In proceedings of **ICASSP**, Toronto, Canada, 2021: 5829-5833.
217. Christoph Boeddeker, Wangyou Zhang, Tomohiro Nakatani, Keisuke Kinoshita, Tsubasa Ochiai, Marc Delcrois, Naoyuki Kamo, **Yanmin Qian** and Reinhold Haeb-Umbach. Convolutional Transfer Function Invariant SDR Training Criteria for Multi-Channel Reverberant Speech Separation. In proceedings of **ICASSP**, Toronto, Canada, 2021: 8428-8432.
218. Houjun Huang, Xu Xiang, Yexin Yang, Rao Ma and **Yanmin Qian**. AISPEECH-SJTU Accent Identification System for the Accented English Speech Recognition Challenge. In proceedings of **ICASSP**, Toronto, Canada, 2021: 6254-6258.
219. Houjun Huang, Xu Xiang, Fei Zhao, Shuai Wang and **Yanmin Qian**. Unit Selection Synthesis based Data Augmentation for Fixed Phrase Speaker Verification. In proceedings of **ICASSP**, Toronto, Canada, 2021: 5849-5853.
220. Tian Tan, Yizhou Lu, Rao Ma, Sen Zhu, Jiaqi Guo and **Yanmin Qian**. AISPEECH-SJTU ASR System for the Accented English Speech Recognition Challenge. In proceedings of **ICASSP**, Toronto, Canada, 2021: 6413-6417.
221. Wangyou Zhang, Christoph Boeddeker, Shinji Watanabe, Tomohiro Nakatani, Marc Delcrois, Keisuke Kinoshita, Tsubasa Ochiai, Naoyuki Kamo, Reinhold Haeb-Umbach and **Yanmin Qian**. End-to-end Dereverberation, Beamforming, and Speech Recognition with improved numerical Stability and Advanced Frontend. In proceedings of **ICASSP**, Toronto, Canada, 2021: 6898-6902.
222. Wei Wang, Zhikai Zhou, Yizhou Lu, Hongji Wang, Chenpeng Du and **Yanmin Qian**. Towards Data Selection on TTS Data for Children's Speech Recognition. In proceedings of **ICASSP**, Toronto, Canada, 2021: 6888-6892.
223. Zhengyang Chen, Shuai Wang and **Yanmin Qian**. Self-Supervised Learning based Domain Adaptation for Robust Speaker Verification. In proceedings of **ICASSP**, Toronto, Canada, 2021: 5834-5838.
224. Chenda Li, Yi Luo, Cong Han, Jinyu Li, Takuya Yoshioka, Tianyan Zhou, Marc Delcroix, Keisuke Kinoshita, Christoph Boeddeker, **Yanmin Qian**, Shinji Watanabe and Zhuo Chen. Dual-path RNN for Long Recording Speech Separation. In proceedings of **SLT**, Shenzhen, China, 2021: 865-872.
225. Chenpeng Du, Hao Li, Yizhou Lu, Lan Wang and **Yanmin Qian**. Data Augmentation for End-to-end Code-Switching Speech Recognition. In proceedings of **SLT**, Shenzhen, China, 2021: 194-200.
226. Xun Gong, Zhengyang Chen, Yexin Yang, Shuai Wang, Lan Wang and **Yanmin Qian**. Speaker Embedding Augmentation with Noise Distribution Matching. In proceedings of **ISCSLP**, Hongkong, China, 2021: 1-5.
227. Shuai Wang, Yexin Yang, **Yanmin Qian** and Kai Yu. Revisiting the Statistics Pooling Layer in Deep Speaker Embedding Learning. In proceedings of **ISCSLP**, Hongkong, China 2021: 1-5.
228. Hongji Wang, Heinrich Dinkel, Shuai Wang, **Yanmin Qian** and Kai Yu. Dual-Adversarial Domain Adaptation for Generalized Replay Attack Detection. In proceedings of **InterSpeech**, Shanghai, China, 2020: 1086-1090.
229. Wangyou Zhang, Aswin Shanmugam Subramanian, Xuankai Chang, Shinji Watanabe and **Yanmin Qian**. End-to-End Far-Field Speech Recognition with Unified Dereverberation and Beamforming. In proceedings of **InterSpeech**, Shanghai, China, 2020: 324-328.
230. Wangyou Zhang and **Yanmin Qian**. Learning Contextual Language Embeddings for Monaural Multi-talker Speech Recognition. In proceedings of **InterSpeech**, Shanghai, China, 2020: 304-308.
231. Chenda Li and **Yanmin Qian**. Listen, Watch and Understand at the Cocktail Party: Audio-Visual-Contextual

-
- Speech Separation. In proceedings of **InterSpeech**, Shanghai, China, 2020: 1426-1430.
232. Zhengyang Chen, Shuai Wang and **Yanmin Qian**. Adversarial Domain Adaptation for Speaker Verification using Partially Shared Network. In proceedings of **InterSpeech**, Shanghai, China, 2020: 3017-3021.
233. Zhengyang Chen, Shuai Wang and **Yanmin Qian**. Multi-modality Matters: A Performance Leap on VoxCeleb. In proceedings of **InterSpeech**, Shanghai, China, 2020: 2252-2256.
234. Yizhou Lu, Mingkun Huang, Hao Li, Jiaqi Guo and **Yanmin Qian**. Bi-encoder Transformer Network for Mandarin-English Code-switching Speech Recognition using Mixture of Experts. In proceedings of **InterSpeech**, Shanghai, China, 2020: 4766-4770.
235. Xuankai Chang, Wangyou Zhang, **Yanmin Qian**, Jonathan Le Roux and Shinji Watanabe. End-to-End Multi-Speaker Speech Recognition with Transformer. In proceedings of **ICASSP**, Barcelona, Spain, 2020: 6129-6133.
236. Chenda Li and **Yanmin Qian**. Deep Audio-Visual Speech Separation with Attention Mechanism. In proceedings of **ICASSP**, Barcelona, Spain, 2020: 7314-7318.
237. Yexin Yang, Shuai Wang, Xun Gong, **Yanmin Qian** and Kai Yu. Text Adaptation for Speaker Verification with Speaker-Text Factorized Embeddings. In proceedings of **ICASSP**, Barcelona, Spain, 2020: 6454-6458.
238. Zhengyang Chen, Shuai Wang, **Yanmin Qian** and Kai Yu. Channel Invariant Speaker Embedding Learning with Joint Multi-task and Adversarial Training. In proceedings of **ICASSP**, Barcelona, Spain, 2020: 6574-6578.
239. Xuankai Chang, Wangyou Zhang, **Yanmin Qian**, Jonathan Le Roux and Shinji Watanabe. MIMO-Speech: End-to-End Multi-Channel Multi-Speaker Speech Recognition. In proceedings of **ASRU**, Sentosa, Singapore, 2019: 237-244.
240. Mingkun Huang, Yizhou Lu, Lan Wang, **Yanmin Qian** and Kai Yu. Exploring Model Units and Training Strategies for End-to-End Speech Recognition. In proceedings of **ASRU**, Sentosa, Singapore, 2019: 524-531.
241. Peiyao Sheng, Zhoulin Yang and **Yanmin Qian**. GANS for Children: A Generative Data Augmentation Strategy for Children Speech Recognition. In proceedings of **ASRU**, Sentosa, Singapore, 2019: 129-135.
242. Wangyou Zhang, Man Sun, Lan Wang and **Yanmin Qian**. End-to-End Overlapped Speech Detection and Speaker Counting with Raw Waveform. In proceedings of **ASRU**, Sentosa, Singapore, 2019: 660-666.
243. Xu Xiang, Shuai Wang, Houjun Huang, **Yanmin Qian** and Kai Yu. Margin Matters: Towards More Discriminative Deep Neural Network Embeddings for Speaker Recognition. In proceedings of **APSIPA**, Lanzhou, China, 2019: 1652-1656.
244. Hao Li, Zhehuai Chen, Qi Liu, **Yanmin Qian** and Kai Yu. OOV Words Extension for Modular Neural Acoustics-to-Word Model. In proceedings of **NCMMSC**, Xining, China, 2019.
245. Yefei Chen, Shuai Wang, **Yanmin Qian** and Kai Yu. End-to-End Speaker-Dependent Voice Activity Detection. In proceedings of **NCMMSC**, Xining, China, 2019.
246. Wangyou Zhang, Ying Zhou and **Yanmin Qian**. Robust DOA Estimation based on Convolutional Neural Network and Time-Frequency Masking. In proceedings of **InterSpeech**, Graz, Austria, 2019: 2703-2707.
247. Wangyou Zhang, Xuankai Chang and **Yanmin Qian**. Knowledge Distillation for End-to-End Monaural Multi-talker ASR System. In proceedings of **InterSpeech**, Graz, Austria, 2019: 2633-2637.
248. Jiaqi Guo, Yongbin You, **Yanmin Qian** and Kai Yu. Joint Decoding of CTC based Systems for Speech Recognition. In proceedings of **InterSpeech**, Graz, Austria, 2019: 2205-2209.
249. Chenda Li and **Yanmin Qian**. Prosody Usage Optimization for Children Speech Recognition with Zero Resource Children Speech. In proceedings of **InterSpeech**, Graz, Austria, 2019: 3446-3450.
250. Hongji Wang, Heinrich Dinkel, Shuai Wang, **Yanmin Qian** and Kai Yu. Cross-domain Replay Spoofing Attack

-
- Detection Using Domain Adversarial Training. In proceedings of **InterSpeech**, Graz, Austria, 2019: 2938-2942.
251. Shuai Wang, Johan Rohdin, Lukas Burget, Oldrich Plchot, **Yanmin Qian**, Kai Yu and Jan Honza Cernocky. On the Usage of Phonetic Information for Text-independent Speaker Embedding Extraction. In proceedings of **InterSpeech**, Graz, Austria, 2019: 1148-1152.
252. Yexin Yang, Hongji Wang, Heinrich Dinkel, Zhengyang Chen, Shuai Wang, **Yanmin Qian** and Kai Yu. The SJTU Robust Anti-spoofing System for the ASVspoof 2019 Challenge. In proceedings of **InterSpeech**, Graz, Austria, 2019: 1038-1042.
253. Zhanghao Wu, Shuai Wang, **Yanmin Qian** and Kai Yu. Data Augmentation using Variational Autoencoder for Embedding based Speaker Verification. In proceedings of **InterSpeech**, Graz, Austria, 2019: 1163-1167.
254. Xuankai Chang, **Yanmin Qian**, Kai Yu and Shinji Watanabe. End-to-End Monaural Multi-Speaker ASR System Without Pretraining. In proceedings of **ICASSP**, Brighton, United Kingdom, 2019: 6256-6260.
255. Shuai Wang, Yexin Yang, Tianzhe Wang, **Yanmin Qian** and Kai Yu. Knowledge Distillation for Small Foot-Print Deep Speaker Embedding. In proceedings of **ICASSP**, Brighton, United Kingdom, 2019: 6021-6025.
256. Yexin Yang, Shuai Wang, Man Sun, **Yanmin Qian** and Kai Yu. Generative Adversarial Networks based X-vector Augmentation for Robust Probabilistic Linear Discriminant Analysis in Speaker Verification. In proceedings of **ISCSLP**, Taipei, Taiwan, 2018.
257. Shuai Wang, Zili Huang, **Yanmin Qian** and Kai Yu. Deep Discriminant Analysis for i-vector based Robust Speaker Recognition. In proceedings of **ISCSLP**, Taipei, Taiwan, 2018.
258. Peiyao Sheng, Zhuolin Yang, Hu Hu, Tian Tan and **Yanmin Qian**. Data Augmentation using Conditional Generative Adversarial Networks for Robust Speech Recognition. In proceedings of **ISCSLP**, Taipei, Taiwan, 2018.
259. Lianwu Chen, Meng Yu, **Yanmin Qian**, Dan Su and Dong Yu. Permutation Invariant Training of Generative Adversarial Network for Monaural Speech Separation. In proceedings of **InterSpeech**, Hyderabad, India, 2018: 302-306.
260. Jun Wang, Jie Chen, Dan Su, Lianwu Chen, Meng Yu, **Yanmin Qian** and Dong Yu. Deep Extractor Network for Target Speaker Recovery from Single Channel Speech Mixtures. In proceedings of **InterSpeech**, Hyderabad, India, 2018: 307-311.
261. Xuankai Chang, **Yanmin Qian** and Dong Yu. Monaural Multi-Talker Speech Recognition with Attention Mechanism and Gated Convolutional Networks. In proceedings of **InterSpeech**, Hyderabad, India, 2018: 1586-1590.
262. Mingkun Huang, Yongbin You, Zhehuai Chen, **Yanmin Qian** and Kai Yu. Knowledge Distillation for Sequence Model. In proceedings of **InterSpeech**, Hyderabad, India, 2018: 3703-3707.
263. Shuai Wang, Heinrich Dinkel, **Yanmin Qian** and Kai Yu. Covariance Based Deep Feature for Text-dependent Speaker Verification, In proceedings of **IScIDE**, Lanzhou, China, 2018: 228-239.
264. Zili Huang, Shuai Wang and **Yanmin Qian**. Joint I-Vector with End-to-End System for Short Duration Text-Independent Speaker Verification. In proceedings of **ICASSP**, Calgary, Alberta, Canada, 2018: 4870-4873.
265. Hu Hu, Tian Tan and **Yanmin Qian**. Generative Adversarial Networks based Data Augmentation for Noise Robust Speech Recognition. In proceedings of **ICASSP**, Calgary, Canada, 2018: 5044-5048.
266. Wen Ding, Tian Tan and **Yanmin Qian**. Fast Adaptation on Deep Mixture Generative Network Based Acoustic Modeling. In proceedings of **ICASSP**, Calgary, Canada, 2018: 5944-5948.
267. Ying Zhou and **Yanmin Qian**. Robust Mask Estimation by Integrating Neural Network-Based and Clustering-

-
- Based Approaches for Adaptive Acoustic Beamforming. In proceedings of **ICASSP**, Calgary, Canada, 2018: 536-540.
268. Xuankai Chang, **Yanmin Qian** and Dong Yu. Adaptive Permutation Invariant Training with Auxiliary Information for Monaural Multi-Talker Speech Recognition. In proceedings of **ICASSP**, Calgary, Canada, 2018: 5974-5978.
269. Tian Tan, **Yanmin Qian** and Dong Yu. Knowledge Transfer in Permutation Invariant Training for Single-channel Multi-talker Speech Recognition. In proceedings of **ICASSP**, Calgary, Canada, 2018: 5714-5718.
270. Shuai Wang, **Yanmin Qian** and Kai Yu. Focal KL-Divergence based Dilated Convolutional Neural Networks for Co-channel Speaker Identification. In proceedings of **ICASSP**, Calgary, Canada, 2018: 5339-5343.
271. **Yanmin Qian**, Tian Tan, Hu Hu and Qi Liu. Noise Robust Speech Recognition On Aurora4 By Humans And Machines. In proceedings of **ICASSP**, Calgary, Canada, 2018: 5604-5608.
272. Qi Liu, **Yanmin Qian** and Kai Yu. Future Vector Enhanced LSTM Language Model for LVCSR. In proceedings of **ASRU**, Okinawa, Japan, 2017: 104-110.
273. Yue Wu, Tianxing He, Zhehuai Chen, **Yanmin Qian** and Kai Yu. Multi-view LSTM Language Model with Word-synchronized Auxiliary Feature for LVCSR. In proceedings of **CCL**, Nanjing, China, 2017: 398-410.
274. Xiaowei Jiang, Shuai Wang, Xu Xiang and **Yanmin Qian**. Integrating Online i-vector into GMM-UBM for Text-dependent Speaker Verification. In proceedings of **APSIPA**, Kuala Lumpur, Malaysia, 2017: 1628-1632.
275. Ruinian Chen, Ying Zhou and **Yanmin Qian**. Emotion Recognition Using Support Vector Machine and Deep Neural Network. In proceedings of **NCMMSC**, Lianyungang, China, 2017: 459-464.
276. Kaiyu Shi, Xuan Liu and **Yanmin Qian**. Speech Emotion Recognition Based on SVM and GMM-HMM Hybrid System. In proceedings of **NCMMSC**, Lianyungang, China, 2017: 470-474.
277. Cheng Chang, Hui Feng Zhang, Zhangxuan Gu and **Yanmin Qian**. Fusion Model for Speech Emotion Recognition with Low Level Descriptor Features. In proceedings of **NCMMSC**, Lianyungang, China, 2017: 487-492.
278. Zhehuai Chen, **Yanmin Qian**, and Kai Yu. A Unified Confidence Measure Framework Using Auxiliary Normalization Graph, In proceedings of **IScIDE**, Dalian, China, 2017: 123--133.
279. Dong Yu, Xuankai Chang and **Yanmin Qian**. Recognizing Multi-talker Speech with Permutation Invariant Training. In proceedings of **InterSpeech**, Stockholm, Sweden, 2017: 2456-2460.
280. Shuai Wang, **Yanmin Qian** and Kai Yu. What Does the Speaker Embedding Encode? In proceedings of **InterSpeech**, Stockholm, Sweden, 2017: 1497-1501.
281. Xu Xiang, **Yanmin Qian** and Kai Yu. Binary Deep Neural Networks for Speech Recognition. In proceedings of **InterSpeech**, Stockholm, Sweden, 2017: 533-537.
282. Heinrich Dinkel, **Yanmin Qian** and Kai Yu. Small-footprint Convolutional Neural Network for Spoofing Detection. In proceedings of **IJCNN**, Anchorage, Alaska, USA, 2017: 3086-3091.
283. Heinrich Dinkel, **Yanmin Qian** and Kai Yu. End-to-End Spoofing Detection with Raw Waveform CLDNNs. In proceedings of **ICASSP**, New Orleans, Louisiana, USA, 2017: 4860-4864.
284. **Yanmin Qian**, Tian Tan. The SJTU ChiME-4 system: Acoustic noise robustness for real single or multiple microphone scenarios. In proceedings of **CHiME4**, San Francisco, California, USA, 2016: 70-72.
285. **Yanmin Qian**, Philip C. Woodland. Very Deep Convolutional Neural Networks for Robust Speech Recognition. In proceedings of **SLT**, San Diego, California, USA, 2016: 481-488.
286. Pavel Korshunov, Sébastien Marcel, Hannah Muckenhirn, Andre Ricardo Goncalves, Alan Godoy Souza Mello,

-
- Ricardo Paranhos Velloso Violato, Flavio Olmos Simoes, Marcus de Assis Angeloni, Jose Augusto Stuchi, Heinrich Dinkel, Nanxin Chen, **Yanmin Qian**, Dipjyoti Paul, Goutam Saha, and Md Sahidullah. Overview of BTAS 2016 Speaker Anti-spoofing Competition (BTAS2016), In proceedings of BTAS, Niagara Falls, Buffalo, New York, USA, 2016: 1-6.
287. Yimeng Zhuang, Sibotong, Maofan Yin, **Yanmin Qian**, Kai Yu. Multi-Task Joint-Learning for Robust Voice Activity Detection. In proceedings of **ISCSLP**, Tianjin, China, 2016:1-5.
288. Pierre Lanchantin, Mark Gales, Penny Karanasou, Xunying Liu, **Yanmin Qian**, Linlin Wang, Phil Woodland and Chao Zhang. Selection of Multi-Genre Broadcast Data for the Training of Automatic Speech Recognition. In proceedings of **InterSpeech**, San Francisco, America, 2016: 3057-3061.
289. Yimeng Zhuang, Xuankai Chang, **Yanmin Qian**, Kai Yu. Unrestricted Vocabulary Keyword Spotting Using LSTM-CTC. In proceedings of **InterSpeech**, San Francisco, America, 2016: 938-942.
290. Linlin Wang, Chao Zhang, Phil Woodland, Mark Gales, Penny Karanasou, Pierre Lanchantin, Xunying Liu, **Yanmin Qian**. Improved DNN-based segmentation for multi-genre broadcast audio. In proceedings of **ICASSP**, Shanghai, China, 2016: 5700-5704.
291. Xie Chen, Xunying Liu, **Yanmin Qian**, Mark J. F. Gales, Philip C. Woodland. CUED-RNNLM-An open-source toolkit for efficient training and evaluation of recurrent neural network language models. In proceedings of **ICASSP**, Shanghai, China, 2016: 6000-6004.
292. Tian Tan, **Yanmin Qian**, Dong Yu, Souvik Kundu, Liang Lu, Khe Chai SIM, Xiong Xiao and Yu Zhang. Speaker-aware training of LSTM-RNNs for acoustic modeling. In proceedings of **ICASSP**, Shanghai, China, 2016: 5280-5284.
293. **Yanmin Qian**, Tian Tan, Dong Yu and Yu Zhang. Integrated adaptation with multi-factor joint-learning for far-field speech recognition. In proceedings of **ICASSP**, Shanghai, China, 2016: 5770-5774.
294. **Yanmin Qian**, Tian Tan and Dong Yu. An investigation into using parallel data for far-field speech recognition. In proceedings of **ICASSP**, Shanghai, China, 2016: 5725-5729.
295. Souvik Kundu, Gautam Mantena, **Yanmin Qian**, Tian Tan, Marc Delcroix and Khe Chai Sim. Joint acoustic factor learning for robust deep neural network based automatic speech recognition. In proceedings of **ICASSP**, Shanghai, China, 2016: 5025-5029.
296. Penny Karanasou, Mark J. F. Gales, Pierre Lanchantin, Xunying Liu, **Yanmin Qian**, Linlin Wang, Philip C. Woodland and Chao Zhang. Speaker Diarisation and Longitudinal Linking in Multi-Genre Broadcast Data. In proceedings of **ASRU**, Scottsdale, Arizona, USA, 2015: 660-666.
297. Pierre Lanchantin, Mark Gales, Penny Karanasou, Xunying Liu, **Yanmin Qian**, Linlin Wang, Phil Woodland and Chao Zhang. The Development of the Cambridge University Alignment Systems for the Multi-Genre Broadcast Challenge. In proceedings of **ASRU**, Scottsdale, Arizona, USA, 2015: 647-653.
298. Philip C. Woodland, Xunying Liu, **Yanmin Qian**, Chao Zhang, Penny Karanasou, Pierre Lanchantin, Linlin Wang and Mark J. F. Gales. Cambridge University Transcription Systems for the Multi-Genre Broadcast Challenge. In proceedings of **ASRU**, Scottsdale, Arizona, USA, 2015: 639-646.
299. **Yanmin Qian**, Maofan Yin, Yongbin You and Kai Yu. Multi-Task Joint-Learning of Deep Neural Networks for Robust Speech Recognition. In proceedings of **ASRU**, Scottsdale, Arizona, USA, 2015: 310-316.
300. Mengxiao Bi, **Yanmin Qian**, Kai Yu. Very Deep Convolutional Neural Networks for LVCSR. In proceedings of **InterSpeech**, Dresden, Germany, 2015: 3259-3263.
301. Nanxin Chen, **Yanmin Qian**, Heinrich Dinkel, Bo Chen, Kai Yu. Robust Deep Feature for Spoofing Detection -

-
- The SJTU System for ASVspoof 2015 Challenge. In proceedings of **InterSpeech**, Dresden, Germany, 2015: 2097-2101.
302. Nanxin Chen, **Yanmin Qian**, Kai Yu. Multi-Task Learning for Text-dependent Speaker Verification. In proceedings of **InterSpeech**, Dresden, Germany, 2015: 185-189.
303. Wengong Jin, Tianxing He, **Yanmin Qian**, Kai Yu. Paragraph Vector based Topic Model for Language Model Adaptation. In proceedings of **InterSpeech**, Dresden, Germany, 2015: 3516-3520.
304. Yongbin You, **Yanmin Qian**, Kai Yu. Local Trajectory Based Speech Enhancement for Robust Speech Recognition With Deep Neural Network. In proceedings of ChinaSIP, Chengdu, China, 2015: 5-9.
305. Yongbin You, **Yanmin Qian**, Tianxing He, Kai Yu. An Investigation on DNN-Derived Bottleneck Features for GMM-HMM Based Robust Speech Recognition. In proceedings of ChinaSIP, Chengdu, China, 2015: 30-34.
306. **Yanmin Qian**, Tianxing He, Wei Deng and Kai Yu. Automatic Model Redundancy Reduction for Fast Back-Propagation for Deep Neural Networks in Speech Recognition. In proceedings of IJCNN, Killarney, Killarney, Ireland, 2015:1-6.
307. Suliang Bu, Yunxin Zhao, **Yanmin Qian** and Kai Yu. A Novel Static Parameter Calculation Method for Model Compensation. In proceedings of **ICASSP**, Brisbane, Australia, 2015: 4510-4514.
308. Tianxing He, Xu Xiang, **Yanmin Qian** and Kai Yu. Recurrent Neural Network Language Model with Structured Word Embeddings for Speech Recognition. In proceedings of **ICASSP**, Brisbane, Australia, 2015: 5396-5400.
309. Tian Tan, **Yanmin Qian**, Maofan Yin, Yimeng Zhuang and Kai Yu. Cluster Adaptive Training For Deep Neural Network. In proceedings of **ICASSP**, Brisbane, Australia, 2015: 4325-4329.
310. Sibotong, Nanxin Chen, **Yanmin Qian** and Kai Yu. Evaluating VAD For Automatic Speech Recognition. In proceedings of ICSP, Hangzhou, 2014: 2308-2314.
311. Suliang Bu, **Yanmin Qian** and Kai Yu. A Novel Dynamic Parameters Calculation Approach For Model Compensation. In proceedings of **InterSpeech**, Singapore, 2014: 2744-2748.
312. Tianfan Fu, **Yanmin Qian**, Yuan Liu and Kai Yu. Tandem Deep Features for Text-Dependent Speaker Verification. In proceedings of **InterSpeech**, Singapore, 2014: 1327-1331.
313. Jianwei Niu, **Yanmin Qian** and Kai Yu. Acoustic Emotion Recognition using Deep Neural Network. In proceedings of **ISCSLP**, Singapore, 2014: 128-132.
314. Yuan Liu, Tianfan Fu, Yuchen Fan, **Yanmin Qian** and Kai Yu. Speaker Verification with Deep Features. In proceedings of IJCNN, Beijing, China, 2014: 747-753.
315. Suliang Bu, **Yanmin Qian**, Khe Chai Sim, Yongbin You, and Kai Yu. Second Order Vector Taylor Series Based Robust Speech Recognition. In proceedings of **ICASSP**, Florence, Italy, 2014: 1788-1792.
316. Tianxing He, Yuchen Fan, **Yanmin Qian**, Tian Tan, and Kai Yu. Reshaping Deep Neural Network for Fast Decoding by Node-pruning. In proceedings of **ICASSP**, Florence, Italy, 2014:245-249.
317. Wei Deng, **Yanmin Qian**, Yuchen Fan, Tianfan Fu and Kai Yu. Stochastic Data Sweeping for Fast DNN Training. In proceedings of **ICASSP**, Florence, Italy, 2014: 240-244.
318. **Yanmin Qian**, Kai Yu and Jia Liu. Combination of Data Borrowing Strategies for Low-Resource LVCSR. In proceedings of **ASRU**, Olomouc, Czech Republic, 2013: 404-409.
319. **Yanmin Qian** and Jia Liu. MLP-HMM Two-Stage Unsupervised Training for Low-Resource Languages on Conversational Telephone Speech Recognition. In proceedings of **InterSpeech**, Lyon, France, 2013: 1816-1820.
320. Daniel Povey, Mirko Hannemann, Gilles Boulianne, Lukáš Burget, Arnab Ghoshal, Miloš Janda, Martin Karafiát, Stefan Kombrink, Petr Motlíček, **Yanmin Qian**, Korbinian Riedhammer, Karel Veselý, Ngoc Thang Vu.

-
- Generating exact lattices in the WFST framework. In proceedings of **ICASSP**, Kyoto, Japan, 2012: 4213-4216.
321. **Yanmin Qian**, Jia Liu. Articulatory feature based multilingual MLPs for low-resource speech recognition. In proceedings of **InterSpeech**, Portland, Oregon, USA, 2012: 2602-2605.
322. **Yanmin Qian**, Jia Liu. Cross-Lingual and ensemble MLPs strategies for low-resource speech recognition. In proceedings of **InterSpeech**, Portland, Oregon, USA, 2012: 2582-2585.
323. Daniel Povey, Arnab Ghoshal, Gilles Boulianne, Lukas Burget, Ondrej Glembek, Nagendra Goel, Mirko Hannemann, Petr Motlicek, **Yanmin Qian**, Petr Schwarz, Jan Silovsky, Georg Stemmer, Karel Vesely. The Kaldi speech recognition toolkit. In proceedings of **ASRU**, Honolulu, Hawaii, USA, 2011.
324. **Yanmin Qian**, Ji Xu, Daniel Povey, Jia Liu. Strategies for using MLP based features with limited target-language training data. In proceedings of **ASRU**, Honolulu, Hawaii, USA, 2011:354-358.
325. **Yanmin Qian**, Daniel Povey, Jia Liu. State-Level data borrowing for low-resource speech recognition based on Subspace GMM Models. In proceedings of **InterSpeech**, Florence, Italy, 2011:553-556.
326. Ji Xu, **Yanmin Qian**, Jia Liu. Utilizing auxiliary data in phoneme recognition based on articulatory feature. In proceedings of **ICCSN**, Xi'an, China, 2011:363-366.
327. Yan Deng, Weiqiang Zhang, **Yanmin Qian**, Jia Lsiu. Integration of complementary phone recognizers for phonotactic language recognition. In proceedings of **ICICA**, Tangshan, China, 2010:237-244.
328. **Yanmin Qian**, Jia Liu. Mandarin-English bilingual phone modeling and combining MPE based Discriminative training for cross-language speech recognition. In proceedings of **ISCSLP**, Tainan, Taiwan, 2010:103-108.
329. **Yanmin Qian**, Jia Liu. Phone model and combining discriminative training for Mandarin-English bilingual speech recognition. In proceedings of **ICASSP**, Dallas, TX, USA, 2010:4918-4921.