

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. Zhengyang Chen, Bing Han, Shuai Wang and Yanmin Qian (Corresponding author) . Attention-based Encoder-Decoder End-to-End Neural Diarization with Embedding Enhancer. Submitted to **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, 2023.
2. Wei Wang and Yanmin Qian (Corresponding author) . Universal Cross-Lingual Data Generation for Low Resource ASR. Submitted to **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, 2023.
3. Xun Gong, Yu Wu, Jinyu Li, Shujie Liu, Rui Zhao, Xie Chen and Yanmin Qian (Corresponding author) . Advanced Long-Form Speech Recognition with Factorized Neural Transducer. Submitted to **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, 2023.
4. Bing Han, Zhengyang Chen, Yanmin Qian (Corresponding author) . Self-Supervised Learning with Cluster-Aware-DINO for High-Performance Robust Speaker Verification. Submitted to **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, 2023.
5. Zhengyang Chen, Shuai Wang, Bing Han, Yanmin Qian (Corresponding author) . Unsupervised Domain Adaptation for Robust Speaker Verification. Submitted to **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, 2023.
6. 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. Submitted to **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, 2023.
7. 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.
8. Shaoxiong Lin, Wangyou Zhang and Yanmin Qian. Two-Stage Single-Channel Speech Enhancement with Multi-Frame Filtering. **Applied Sciences**, vol. 1, 1-14, 2023.
9. 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.
10. 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.

-
11. **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.
 12. 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.
 13. **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.
 14. **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.
 15. 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.
 16. 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.
 17. 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.
 18. 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.
 19. **Yanmin Qian**, Hu Hu, Tian Tan. Data Augmentation using Generative Adversarial Networks for Robust Speech Recognition. **Speech Communication**, vol. 114, 1-9, 2019.
 20. 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.
 21. **Yanmin Qian**, Xu Xiang. Binary Neural Networks for Speech Recognition. **Frontiers of Information Technology & Electronic Engineering**, 20 (5), 701-715, 2019.
 22. **Yanmin Qian**, Xuankai Chang, Dong Yu. Single-Channel Multi-talker Speech Recognition with Permutation Invariant Training. **Speech Communication**, vol. 104, 1-11, 2018.
 23. Zhehuai Chen, **Yanmin Qian (Corresponding author)** , Kai Yu. Sequence Discriminative Training for Deep Learning based Keyword Spotting. **Speech Communication**, vol. 102, 100-111, 2018.
 24. 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.
 25. 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.
 26. **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)
 27. 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.
28. **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.
 29. 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.
 30. **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.
 31. **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.
 32. 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.
 33. **Yanmin Qian**, Nanxin Chen, Kai Yu. Deep Features for Automatic Spoofing Detection. **Speech Communication**, vol. 85, 43-52, 2016.
 34. 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.
 35. **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.
 36. **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.
 37. **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.
 38. **Yanmin Qian**, Jia Liu. Cross-entropy OSF-based voice activity detection algorithm. **Tsinghua Science and Technology**, 2009, 49(10):87-90.
 39. **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.
 40. Hong Liu, **Yanmin Qian**, Jia Liu. English speech recognition system on chip. **Tsinghua Science and Technology**, 2011, 16(1):95-99.
 41. 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.
 42. 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.
 43. 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:

44. Bing Han, Zhengyang Chen, Bei Liu and **Yanmin Qian**. Unsupervised Domain Adaptation with Pseudo Labeling for Robust Speaker Verification. Submitted to **ASRU**, Taipei, Taiwan, 2023.
45. Hang Shao, Wei Wang, Xun Gong and **Yanmin Qian**. Whisper-KDQ: A Lightweight Whisper via Guided Knowledge Distillation and Quantization for Efficient ASR. Submitted to **ASRU**, Taipei, Taiwan, 2023.
46. Wangyou Zhang, Lei Yang and **Yanmin Qian**. Exploring Time-Frequency Domain Target Speaker Extraction For Causal and Non-Causal Processing. Submitted to **ASRU**, Taipei, Taiwan, 2023.
47. Wangyou Zhang, Kohei Saijo, Zhong-Qiu Wang, Shinji Watanabe and **Yanmin Qian**. Toward Universal Speech Enhancement for Diverse Input Conditions. Submitted to **ASRU**, Taipei, Taiwan, 2023.
48. Shaoxiong Lin, Chao Zhang and **Yanmin Qian**. Improving Speech Enhancement Using Audio Tagging Knowledge from Pre-trained Representations and Multi-task Learning. Submitted to **ASRU**, Taipei, Taiwan, 2023.
49. Chang Chen, Xun Gong and **Yanmin Qian**. Efficient Text-Only Domain Adaptation for CTC-based ASR. Submitted to **ASRU**, Taipei, Taiwan, 2023.
50. Dongning Yang, Wei Wang and **Yanmin Qian**. FAT-HuBERT: Front-end Adaptive Training of Hidden-unit BERT for Distortion-Invariant Robust Speech Recognition. Submitted to **ASRU**, Taipei, Taiwan, 2023.
51. 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. Submitted to **ASRU**, Taipei, Taiwan, 2023.
52. Xinyuan Wu, Yifei Wu, Chenda Li and **Yanmin Qian**. Adaptive Alignment for Audio-Visual Speech Separation. Submitted to **APSIPA**, Taipei, Taiwan, 2023.
53. 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. Submitted to **NeurIPS**, USA, 2023.
54. Zhiqiang Lv, Bing Han, Zhengyang Chen, **Yanmin Qian**, Jiawei Ding and Jia Liu. Unsupervised Anomalous Detection based on Unsupervised Pretrained Models. Accepted by DCASE-Challenge, 2023.
55. 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. Accepted by **WASPAA**, New Paltz, USA, 2023.
56. Bei Liu and **Yanmin Qian**. ECAPA++: Fine-Grained Deep Embedding Learning for TDNN based Speaker Verification. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
57. Bei Liu, Haoyu Wang and **Yanmin Qian**. Extremely Low Bit Quantization for Mobile Speaker Verification Systems Under 1MB Memory. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
58. Bei Liu and **Yanmin Qian**. Reversible Neural Networks for Memory-Efficient Speaker Verification. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
59. Haoyu Wang, Bei Liu, Yifei Wu and **Yanmin Qian**. Adaptive Neural Network Quantization for Lightweight Speaker Verification. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
60. 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. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
61. 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. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
62. 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. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
 63. Linfeng Yu, Wangyou Zhang, Chenda Li and **Yanmin Qian**. Overlap Aware Continuous Speech Separation without Permutation Invariant Training. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
 64. Wangyou Zhang and **Yanmin Qian**. Weakly-Supervised Speech Pre-training: A Case Study on Target Speech Recognition. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
 65. Wei Wang and **Yanmin Qian**. UniSplice: Universal Cross-Lingual Data Splicing for Low-Resource ASR. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
 66. 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. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
 67. Zhilong Zhang, Wei Wang and **Yanmin Qian**. Fast and Efficient Multilingual Self-Supervised Pre-training for Low-Resource Speech Recognition. Accepted by **InterSpeech**, Dublin, Ireland, 2023.
 68. Shaoxiong Lin, Wangyou Zhang and **Yanmin Qian**. Two Stage Single-Channel Speech Enhancement with Multi-Frame Filtering. Accepted by **NCMMSC**, Hefei, China, 2022.
 69. 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.
 70. 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.
 71. 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.
 72. Bing Han, Zhengyang Chen and **Yanmin Qian**. Exploring Binary Classification Loss for Speaker Verification. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 73. 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.
 74. 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.
 75. 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.
 76. 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.
 77. 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.
 78. Leying Zhang, Zhengyang Chen and **Yanmin Qian**. Adaptive Large Margin Fine-tuning for Robust Speaker Verification. In proceedings of **ICASSP**, Rhodes Island, Greece, 2023.
 79. 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.
80. 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.
 81. 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.
 82. 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.
 83. Shaoxiong Lin, Wangyou Zhang and **Yanmin Qian**. Two Stage Single-Channel Speech Enhancement with Multi-Frame Filtering. Accepted by **NCMMSC**, Hefei, China, 2022.
 84. 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.
 85. 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.
 86. Houjun Huang and **Yanmin Qian**. Speaking Style Compensation on Synthetic Audio for Robust Keyword Spotting. In proceedings of **ISCSLP**, Singapore, 2022.
 87. 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.
 88. 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.
 89. 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.
 90. 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.
 91. 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.
 92. 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.
 93. 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.
 94. Bei Liu, Zhengyang Chen and **Yanmin Qian**. Attentive Feature Fusion for Robust Speaker Verification. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 286-290.
 95. 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.
 96. 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:

97. 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.
98. 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.
99. Xun Gong, Zhikai Zhou and **Yanmin Qian**. Knowledge Transfer and Distillization from Autoregressive to Non-Autoregressive Speech Recognition. In proceedings of **InterSpeech**, Incheon, Korea, 2022: 2618-2622.
100. 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.
101. 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.
102. 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.
103. 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.
104. 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.
105. Bing Han, Zhengyang Chen and **Yanmin Qian**. Local Information Modeling with Self-Attention for Speaker Verification. In proceedings of **ICASSP**, Singapore, 2022: 6727-6731.
106. 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.
107. 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.
108. 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.
109. 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.
110. 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.
111. Zhikai Zhou, Tian Tan and **Yanmin Qian**. Punctuation Prediction for Streaming On-Device Speech Recognition. In proceedings of **ICASSP**, Singapore, 2022: 7277-7281.
112. 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.
113. Zhengxi Liu and **Yanmin Qian**. Basis-MelGAN: Efficient Neural Vocoder Based on Audio Decomposition. In

-
- proceedings of **InterSpeech**, Brno, Czechia, 2021: 2222-2226.
114. 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.
115. 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.
116. 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.
117. 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.
118. 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.
119. 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.
120. 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.
121. Christoph Boeddeker, Wangyou Zhang, Tomohiro Nakatani, Keisuke Kinoshita, Tsubasa Ochiai, Marc Delcroix, 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.

-
122. 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.
 123. 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.
 124. 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.
 125. Wangyou Zhang, Christoph Boeddeker, Shinji Watanabe, Tomohiro Nakatani, Marc Delcroix, 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.
 126. 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.
 127. 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.
 128. 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.
 129. 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.
 130. 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.
 131. 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.
 132. 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.
 133. 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.
 134. Wangyou Zhang and **Yanmin Qian**. Learning Contextual Language Embeddings for Monaural Multi-talker Speech Recognition. In proceedings of **InterSpeech**, Shanghai, China, 2020: 304-308.
 135. 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.
 136. 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.
 137. Zhengyang Chen, Shuai Wang and **Yanmin Qian**. Multi-modality Matters: A Performance Leap on VoxCeleb. In proceedings of **InterSpeech**, Shanghai, China, 2020: 2252-2256.
 138. 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.

-
139. 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.
 140. Chenda Li and **Yanmin Qian**. Deep Audio-Visual Speech Separation with Attention Mechanism. In proceedings of **ICASSP**, Barcelona, Spain, 2020: 7314-7318.
 141. 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.
 142. 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.
 143. 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.
 144. 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.
 145. 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.
 146. 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.
 147. 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.
 148. 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.
 149. Yefei Chen, Shuai Wang, **Yanmin Qian** and Kai Yu. End-to-End Speaker-Dependent Voice Activity Detection. In proceedings of **NCMMSC**, Xining, China, 2019.
 150. 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.
 151. 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.
 152. 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.
 153. 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.
 154. 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.
 155. 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.
 156. 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.
 157. 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.
158. 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.
159. 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.
160. 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.
161. 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.
162. 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.
163. 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.
164. 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.
165. 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.
166. 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.
167. 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.
168. 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.
169. 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.
170. 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.
171. 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.
172. 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.
173. 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.
174. 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.
175. **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.
176. Qi Liu, **Yanmin Qian** and Kai Yu. Future Vector Enhanced LSTM Language Model for LVCSR. In proceedings of **ASRU**, Okinawa, Japan, 2017: 104-110.
177. 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.
178. 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.
179. 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.
180. 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.
181. Cheng Chang, Huifeng 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.
182. 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.
183. Dong Yu, Xuankai Chang and **Yanmin Qian**. Recognizing Multi-talker Speech with Permutation Invariant Training. In proceedings of **InterSpeech**, Stockholm, Sweden, 2017: 2456-2460.
184. Shuai Wang, **Yanmin Qian** and Kai Yu. What Does the Speaker Embedding Encode? In proceedings of **InterSpeech**, Stockholm, Sweden, 2017: 1497-1501.
185. Xu Xiang, **Yanmin Qian** and Kai Yu. Binary Deep Neural Networks for Speech Recognition. In proceedings of **InterSpeech**, Stockholm, Sweden, 2017: 533-537.
186. 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.
187. 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.
188. **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.
189. **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.
190. 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.
191. Yimeng Zhuang, Sibong Tong, Maofan Yin, **Yanmin Qian**, Kai Yu. Multi-Task Joint-Learning for Robust Voice Activity Detection. In proceedings of **ISCSLP**, Tianjin, China, 2016:1-5.
192. 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.
193. 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.
194. 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.
195. 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.
196. 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.
197. **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.
198. **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.
199. 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.
200. 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.
201. 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.
202. 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.
203. **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.
204. Mengxiao Bi, **Yanmin Qian**, Kai Yu. Very Deep Convolutional Neural Networks for LVCSR. In proceedings of **InterSpeech**, Dresden, Germany, 2015: 3259-3263.
205. 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.
206. Nanxin Chen, **Yanmin Qian**, Kai Yu. Multi-Task Learning for Text-dependent Speaker Verification. In proceedings of **InterSpeech**, Dresden, Germany, 2015: 185-189.
207. 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.
208. 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.

-
209. 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.
210. **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.
211. 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.
212. 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.
213. 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.
214. Sibotong, Nanxin Chen, **Yanmin Qian** and Kai Yu. Evaluating VAD For Automatic Speech Recognition. In proceedings of ICSP, Hangzhou, 2014: 2308-2314.
215. Suliang Bu, **Yanmin Qian** and Kai Yu. A Novel Dynamic Parameters Calculation Approach For Model Compensation. In proceedings of InterSpeech, Singapore, 2014: 2744-2748.
216. 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.
217. Jianwei Niu, **Yanmin Qian** and Kai Yu. Acoustic Emotion Recognition using Deep Neural Network. In proceedings of ISCSLP, Singapore, 2014: 128-132.
218. 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.
219. 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.
220. 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.
221. 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.
222. **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.
223. **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.
224. 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.
225. **Yanmin Qian**, Jia Liu. Articulatory feature based multilingual MLPs for low-resource speech recognition. In proceedings of InterSpeech, Portland, Oregon, USA, 2012: 2602-2605.
226. **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.
227. 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.

-
228. **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.
229. **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.
230. 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.
231. 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.
232. **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.
233. **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.