# QIUQIANG KONG

qiuqiangkong@gmail.com

## EDUCATION

**University of Surrey, UK** *May. 2016 - Sep. 2019*
*Ph.D. of Electronic engineering*; Advisor: Professor Mark D. Plumbley
Centre for Vision, Speech and Signal Processing (CVSSP)

**Queen Mary University of London, UK** *May. 2015 - Apr. 2016*
*Ph.D. Electronic engineering*; Advisor: Professor Mark D. Plumbley
Centre for Digital Music (C4DM)

**South China University of Technology** *Sep. 2012 - Apr. 2015*
*Master of Electronic engineering*

**South China University of Technology** *Sep. 2008 - Apr. 2012*
*Bachelor of Electronic engineering*

## RESEARCH INTEREST

My research interest is sound understanding and audio processing with machine learning methods. My research topics include audio tagging, sound event detection, source separation of real-world sounds, general audio and music signal processing.

## EXPERIENCE

**ByteDance, California, U.S.** *Research intern* *May. 2019 - Sep. 2019*
**AudioAnalytic, Cambridge, UK** *Research intern* *Apr. 2017 - Sep. 2017*

## PUBLICATION

- Z. Ren, J. Han, N. Cummins, Q. Kong, M. D. Plumbley, and B. W. Schuller. "Multi-instance Learning for Bipolar Disorder Diagnosis using Weakly Labelled Speech Data," in *Proceedings of the 9th International Conference on Digital Public Health*, 2019, pp. 79-83. *doi: 10.1145/3357729.3357743*
- Y. Fu, K. Xu, H. Mi, Q. Kong, D. Wang, H. Wang, and T. Hong. "Multi Model-Based Distillation for Sound Event Detection," *IEICE Transactions on Information and Systems*, vol. 102, no. 10, pp. 2055-2058, 2019. *doi: 10.1587/transinf.2019EDL8062*
- Q. Kong, C. Yu, Y. Xu, T. Iqbal, W. Wang, and M. D. Plumbley. "Weakly labelled audioset tagging with attention neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 11, pp. 1791-1802, 2019. *doi: 10.1109/TASLP.2019.2930913*
- Q. Kong, Y. Xu, W. Wang, P. J. B. Jackson, and M. D. Plumbley. "Single-Channel Signal Separation and Deconvolution with Generative Adversarial Networks," *in Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2019, pp. 2747-2753. *doi: 10.24963/ijcai.2019/381*
- K. Xu, B. Zhu, Q. Kong, H. Mi, B. Ding, D. Wang, and H. Wang. "General audio tagging with ensembling convolutional neural networks and statistical features," *The Journal of the Acoustical Society of America*, vol. 145, no. 6, pp. 521-527, 2019. *doi: 10.1121/1.5111059*
- Y. Cao, T. Iqbal, Q. Kong, M. Galindo, W. Wang, and M. D. Plumbley. "Polyphonic Sound Event Detection and Localization using a Two-Stage Strategy," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2019, pp. 30-34. *doi: 10.33682/4jhy-bj81*

- C. Zor, M. Awais, J. Kittler, M. Bober, S. Husain, Q. Kong, and C. Kroos. "Divergence Based Weighting for Information Channels in Deep Convolutional Neural Networks for Bird Audio Detection," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 3052-3056. *doi: 10.1109/ICASSP.2019.8682483*
- Q. Kong, Y. Xu, T. Iqbal, Y. Cao, W. Wang, and M. D. Plumbley. "Acoustic Scene Generation with Conditional Samplernn," *in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 925-929. *doi: 10.1109/ICASSP.2019.8683727*
- Z. Ren, Q. Kong, J. Han, M. D. Plumbley, and B. W. Schuller. "Attention-based Atrous Convolutional Neural Networks: Visualisation and Understanding Perspectives of Acoustic Scenes," *in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 56-60. *doi: 10.1109/ICASSP.2019.8683434*
- Y. Hou, Q. Kong, S. Li, and M. D. Plumbley. "Sound Event Detection with Sequentially Labelled Data Based on Connectionist Temporal Classification and Unsupervised Clustering," *in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 46-50. *doi: 10.1109/ICASSP.2019.8683627*
- Q. Kong, Y. Xu, I. Sobieraj, W. Wang, and M. D. Plumbley. "Sound Event Detection and TimeFrequency Segmentation from Weakly Labelled Data," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 27, no. 4, pp. 777-787, 2019. *10.1109/TASLP.2019.2895254*
- Y. Hou, Q. Kong, and S. Li. "A Comparison of Attention Mechanisms of Convolutional Neural Network in Weakly Labeled Audio Tagging," *in Proceedings of the 6th Conference on Sound and Music Technology (CSMT)*, 2019, pp. 85-96. *doi.org/10.1007/9789811387074_8*
- J. Jiang, M. Hoogendoorn, Q. Kong, D. M. Roijers, and N. Gilbert. "Predicting Appliance Usage Status In Home Like Environments," *In IEEE International Conference on Digital Signal Processing (DSP)*, 2018, pp. unavailable. *doi: 10.1109/ICDSP.2018.8631580*
- Y. Hou, Q. Kong, J. Wang, and S. Li. "Polyphonic audio tagging with sequentially labelled data using CRNN with learnable gated linear units," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2018, pp. 78-82.
- D. Wang, L. Zhang, C. Bao, K. Xu, B. Zhu, and Q. Kong. "Weakly supervised CRNN system for sound event detection with large-scale unlabeled in-domain data," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2018, 93-97.
- T. Iqbal, Y. Xu, Q. Kong, and W. Wang. "Capsule routing for sound event detection," *in European Signal Processing Conference (EUSIPCO)*, 2018, pp. 2255-2259. *doi: 10.23919/EUSIPCO.2018.8553198*
- S. Wei, K. Xu, D. Wang, F. Liao, H. Wang, and Q. Kong. "Sample Mixed-Based Data Augmentation For Domestic Audio Tagging," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2018, 93-97.
- Q. Kong, T. Iqbal, Y. Xu, W. Wang, and M. D. Plumbley. "DCASE 2018 Challenge baseline with convolutional neural networks," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2018, 217-221.
- A. Zermini, Q. Kong, Y. Xu, M. D. Plumbley, and W. Wang. "Improving reverberant speech separation with binaural cues using temporal context and convolutional neural networks," *in International Conference on Latent Variable Analysis and Signal Separation (LVA-ICA)*, 2018, pp. 361-371. *doi: 10.1007/978-3-319-93764-9_34*
- Q. Kong, Y. Xu, W. Wang, and M. D. Plumbley. "A joint separation-classification model for sound event detection of weakly labelled data," *in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 321-325. *doi: 10.1109/ICASSP.2018.8462448*
- Q. Kong, Y. Xu, W. Wang, and M. D. Plumbley. "Audio set classification with attention model: A probabilistic perspective," *in IEEE International Conference on Acoustics, Speech and Signal*

*Processing (ICASSP)*, 2018, pp. 316-320. *doi: 10.1109/ICASSP.2018.8461392*

- Y. Xu, Q. Kong, W. Wang, and M. D. Plumbley. "Large-scale weakly supervised audio classification using gated convolutional neural network," *in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 121-125. *doi: 10.1109/ICASSP.2018.8461975*
- C. Yu, K. S. Barsim, Q. Kong, and B. Yang. "Multi-level attention model for weakly supervised audio classification," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2018, pp. 188-192.
- T. Iqbal, Q. Kong, M. D. Plumbley, and W. Wang. "General-purpose audio tagging from noisy labels using convolutional neural networks," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2018, pp. 212-216.
- Z. Ren, Q Kong, K. Qian, M. D. Plumbley, and B. Schuller. "Attention-based convolutional neural networks for acoustic scene classification," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2018, pp. 39-43.
- Q. Kong, Y. Xu, and M. D. Plumbley. "Joint detection and classification convolutional neural network on weakly labelled bird audio detection," *in IEEE European Signal Processing Conference (EUSIPCO)*, 2017, pp. 1749-1753. *doi: 10.23919/EUSIPCO.2017.8081509*
- Y. Xu, Q. Kong, Q. Huang, W. Wang, and M. D. Plumbley. "Convolutional gated recurrent neural network incorporating spatial features for audio tagging," *in IEEE International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 3461-3466. *doi: 10.1109/IJCNN.2017.7966291*
- Q. Kong, Y. Xu, W. Wang, and M. D. Plumbley. "A joint detection-classification model for audio tagging of weakly labelled data," *in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 641-645. *doi: 10.1109/ICASSP.2017.7952234*
- I. Sobieraj, Q. Kong, and M. D. Plumbley. "Masked non-negative matrix factorization for bird detection using weakly labeled data," *in IEEE European Signal Processing Conference (EUSIPCO)*, 2017, pp. 1769-1773. *doi: 10.23919/EUSIPCO.2017.8081513*
- Q. Kong, I. Sobieraj, W. Wang, and M. Plumbley. "Deep neural network baseline for DCASE challenge 2016," *in Workshop on the Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2016, pp. 50-54.

## REVIEW

- IEEE/ACM Transactions on Audio Speech and Language Processing
- IEEE Transactions on Multimedia
- IEEE Signal Processing Letters
- ACM Transactions on Knowledge Discovery from Data (TKDD)
- Neurocomputing
- Neural Networks
- IEEE Intelligent Transportation Systems Society
- The Multidisciplinary Open Access Journal
- The Detection and Classification of Acoustic Scenes and Events (DCASE) Workshop
- The European Association for Signal Processing (EURASIP)
- Conference on Sound and Music Technology (CSMT)

## PATENT

- Large scale music fingerprint and retrieval (CN103853836), bought by KUGOU, one of the largest music company in China in 2014.

## HIGHLIGHT

- Win the Reproducible System Award for the DCASE2019 challenge Task 3.
- Nominated as the PGR student of the year of CVSSP/EEE of University of Surrey in 2019.
- Ranked the 3rd (out of 558 systems) in the private leaderboard of the DCASE2018 Challenge Task 2: General purpose audio tagging of Freesound.
- Invited to give a talk of deep learning in Audio Engineering Society (AES) Convention 2018, Milan, Italy.
- Ranked the 1st (out of 29 system) in the DCASE2017 Challenge Task 4: Large-scale weakly supervised sound event detection for smart cars sponsored by Google.