Dr. Yong Xu (research fellow in UK)

PERSONAL INFORMATION

Name: Yong Xu, 徐勇; Gender: Male; Date of Birth: Oct. 6th, 1988; Place of Birth: Hefei, China

UK Cellphone: 07484685283; Email: yong.xu.ustc@gmail.com

Home page: http://home.ustc.edu.cn/~xuyong62/

My Google Scholar: https://scholar.google.com/citations?user=nCmKPM4AAAAJ&hl=zh-CN (h-index: 6; number of citations: 225)

Current work

University of Surrey, Guildford, UK Full-time Research Fellow Apr. 2016 – present
Deep learning (DNN/CNN/LSTM, etc) based environmental sound classification and analysis

EDUCATION BACKGROUND

Georgia Institute of Technology, USA
Visiting student, major in Speech Signal Processing Sept. 2014 – Apr. 2015
Research focus: Object function optimization for deep neural networks (DNN) based Speech Enhancement

University of Science and Technology of China (USTC), National Engineering Laboratory of Speech and Language Information Processing, China
Ph.D. Candidate, major in Speech Signal Processing Sept. 2012 – Jun. 2015
Research focus: Deep neural networks (DNN) based Speech Enhancement and Speech Recognition

Master, major in Speech Signal Processing Sept. 2010 - Jul. 2012
Research focus: Large-scale Speech Recognition and Spoken Term Detection for Out-of-Vocabulary (OOV) word

Anhui University, China
Bachelor, major in Communication Engineering Sept. 2006 – Jul. 2010

PUBLICATIONS

Patent:
[1] Speech separation method and system, US patent, US 20160189730A1
DU Jun, Xu Yong, TU Yanhui, Dai Li-rong, Wang Zhigu, HU Yu, Liu Qingfeng, June 2016

Journal papers:
[3] Fully Deep Neural Networks Incorporating Unsupervised Feature Learning for Audio Tagging
Yong Xu, Qiang Huang, Wenwu Wang, Peter Foster, Siddharth Sigta, Philip J. B. Jackson, Mark D. Plumbley, submitted to IEEE/ACM Transactions on Audio, Speech, and Language Processing, July 2016

[2] A Regression Approach to Speech Enhancement Based on Deep Neural Networks.
Yong Xu, Jun Du, Li-Rong Dai and Chin-Hui Lee, accepted to IEEE/ACM Transactions on Audio, Speech, and Language Processing, P. 7-19, Vol. 23, No. 1, IEEE/ACM trans. on audio, speech, language processing, 2015

[1] An Experimental Study on Speech Enhancement Based on Deep Neural Networks.
Yong Xu, Jun Du, Li-Rong Dai and Chin-Hui Lee, IEEE signal processing letters, p. 65-68, vol. 21, no. 1, January 2014

Conference papers:
[17] Fully DNN-based Multi-label regression for audio tagging.
Yong Xu, Qiang Huang, Wenwu Wang, Philip J B Jackson, Mark D Plumbley, accepted by DCASE2016 workshop, July 2016

[16] Hierarchical learning for DNN-based acoustic scene classification
Yong Xu, Qiang Huang, Wenwu Wang, Mark D. Plumbley, accepted by DCASE2016 workshop, July 2016

[15] Joint training of DNNs by incorporating an explicit dereverberation structure for distant speech recognition.
Tian Gao, Jun Du, Yong Xu, Cong Liu, Li-Rong Dai, Chin-Hui Lee, submitted to ICASSP2016.

[14] Multi-objective learning and Mask-based Post-processing for Deep Neural Network based Speech Enhancement.
Yong Xu, Jun Du, Zhen Huang, Li-Rong Dai, Chin-Hui Lee, accepted, Interspeech2015, Dresden, Germany
[13] DNN-Based Speech Bandwidth Expansion and Its Application to Adding High Frequency Missing Features for Automatic Speech Recognition of Narrowband Speech.
Kehuang Li, Zhen Huang, Yong Xu and Chin-Hui Lee, accepted, Interspeech2015, Dresden, Germany

[12] Dynamic Noise Aware Training for Speech Enhancement Based on Deep Neural Networks.
Yong Xu, Jun Du, Li-Rong Dai and Chin-Hui Lee, Interspeech2014, Singapore

[11] Improving Deep Neural Network Based Speech Enhancement in Low SNR Environments. (Best paper candidate)
Tian Gao, Jun Du, Yong Xu, Cong Liu, Li-Rong Dai, Chin-Hui Lee, accepted, LVA/ICA 2015, Liberec, Czech Republic

[10] Robust Speech Recognition with Speech Enhanced Deep Neural Networks
Jun Du, Qing Wang, Tian Gao, Yong Xu, Li-Rong Dai and Chin-Hui Lee, Interspeech2014, Singapore

[9] Cross-language Transfer Learning for Deep Neural Network Based Speech Enhancement
Yong Xu, Jun Du, Li-Rong Dai and Chin-Hui Lee, ISCSLP2014, Singapore

[8] Speech Separation Based on Improved Deep Neural Networks with Dual Outputs of Speech Features for both Target and Interfering Speakers
Yanhui Tu, Jun Du, Yong Xu, Lirong Dai and Chin-Hui Lee, ISCSLP2014, Singapore

[7] Speech separation of a target speaker based on deep neural networks.
Jun Du, Yanhui Tu, Yong Xu, Li-Rong Dai and Chin-Hui Lee, P. 532 – 536, ICSP2014, Hangzhou, China

[6] Deep neural network based speech separation for robust speech recognition.
Yanhui Tu, Jun Du, Yong Xu, Lirong Dai and Chin-Hui Lee, P. 532 – 536, Hangzhou, China

[5] Global Variance Equalization for Improving Deep Neural Network Based Speech Enhancement.
Yong Xu, Jun Du, Li-Rong Dai and Chin-Hui Lee, to be appeared at ChinaSIP2014, Xi’an, China

[4] Spoken Term Detection for OOV Terms Based on Phone Fragment.
Yong Xu, Wu Guo, Shan Su and Li-Rong Dai, ICALIP2012, Shanghai, China

[3] Improved Spoken Term Detection by Template-based Confidence Measure.
Shan Su, Wu Guo, Yong Xu and Li-Rong Dai, ICALIP2012, Shanghai, China

[2] A hybrid fragment / syllable-based system for improved OOV term detection.
Yong Xu, Wu Guo and Li-Rong Dai, ISCSLP2012, Hong Kong

[1] Spoken term detection for OOV terms based on tri-phone confusion matrix.
Yong Xu, Wu Guo and Li-Rong Dai, ISCSLP2012, Hong Kong

RESEARCH EXPERIENCE

**University of Surrey, Guildford, UK**

Full-time Research Fellow

Deep learning (DNN/CNN/LSTM, etc) based environmental sound classification and analysis.

**IFLYTEK co., ltd, Hefei, China**

Full-time Researcher

Deep learning (DNN/CNN/RNN, etc) based speech recognition, speech enhancement and speech dereverberation. Now focus on the task for distant speech recognition.

**Georgia Institute of Technology, USA**

Visiting Student

Deep neural networks based speech enhancement and used for the automatic speech recognition (ASR), and my advisor is Prof. Chin-Hui Lee.

**Bosch - research center, CA, USA**

Short Internship

Deep neural networks based speech enhancement and used for the automatic speech recognition (ASR)

**National Engineering Laboratory of Speech and language information processing, USTC, China**

DNN based speech enhancement, cooperated with Prof. Chin-Hui Lee (Georgia Tech)

**IFLYTEK co., ltd, Hefei, China**

Internship

I developed a Large Vocabulary Continuous Speech Recognition (LVCSR) system trained on 2300h English speech database, and built a baseline for OOV term detection. MLE, DT, Tandem systems were built.

**IFly Speech Lab, USTC, Hefei, China**

Graduate student

Sept. 2010 – Jul. 2012
Working on Spoken Term Detection (STD) for Out-Of-Vocabulary (OOV) words, I use the tri-phone confusion matrix and a hybrid fragment / syllable system to improve the performance of OOV term detection.

IFly Speech Lab, USTC, Hefei, China  **Undergraduate student** Mar. 2010 – Jul. 2010
I did the project of my undergraduate thesis about room acoustic impulse response.

**ACADEMIC ACTIVITIES**

| Event                                      | 2015.09          |
|--------------------------------------------|------------------|
| Volunteer of INTERSPEECH2015, Dresden, Germany | 2015.09          |
| Oral in INTERSPEECH2015, Dresden, Germany   | 2015.09          |
| Visiting student in Georgia Institute of Technology, Atlanta, GA, USA | 2014.09-2015.04 |
| Short Intern in Bosch-research center, Palo Alto, CA, USA | 2014.10-2014.11 |
| Visitor in GlobalSIP2014, Atlanta, USA      | 2014.12          |
| Poster in INTERSPEECH2014, Singapore        | 2014.09          |
| Oral in ChinaSIP2014, Xi’an, China          | 2014.07          |
| Oral in ISCSLP2014, Singapore               | 2014.09          |
| Oral in ISCSLP2012, Hong Kong               | 2012.09          |
| Oral in ICALIP2012, Shanghai, China         | 2012.05          |

**HONORS AND AWARDS**

| Event                                                        | 2016.07          |
|--------------------------------------------------------------|------------------|
| Fund of early career researcher in University of Surrey, UK   | 2016.07          |
| “monthly star” of IFLYTEK                                     | 2015.08          |
| Huayu scholarship of USTC                                     | 2014.10          |
| China Scholarship Council-joint PhD program in Georgia Tech, USA | 2014.06          |
| National Scholarship                                          | 2009.10          |
| Provincial-level Merit Student                                | 2008.05          |
| First-class annual scholarship four times                     | 2006-2010        |

**SKILLS**

Program skills: familiar with C/C++, Python, Perl, Matlab, GPU programing on Windows or Linux
Language skills: CET-4 (542), CET-6 (521), GRE, Fluent English speaking and writing, (Mandarin is my native language)
Others: familiar with HTK toolkit, QuickNet, SRILM, some toolkits of DNN, such as Kaldi, currentRNN, LSTM

**Hobbies**

Cycling, Running, Traveling, Basketball, etc.