Yajie Miao

Carnegie Mellon University
School of Computer Science
Language Technologies Institute

Phone: (412) 951-6032
Email: ymiao@cs.cmu.edu
Homepage: http://www.cs.cmu.edu/~ymiao

Research Interests

- Speech recognition based on deep learning

- Sequence learning using RNN and LSTM models

Education

- Ph.D., School of Computer Science, **Carnegie Mellon University**. 2016 (Expected)

- M.Eng. in Software Engineering, **Tsinghua University**. 2011

- B.Eng. in Electrical Engineering, **Nanjing University of Posts and Telecommunications**. 2008

Awards and Honors

- **Best Paper Nomination**, ASRU 2015

- **Best Poster Award**, SLT 2014

- **Travel Grants**, INTERSPEECH 2014, ICASSP 2016

- **CMU/LTI Departmental Fellowship**, 2011-2016

- **Outstanding Graduate**, Tsinghua University, 2011

- **Outstanding Master Thesis Award**, Tsinghua University, 2011

- **Chiang Chen Scholarship**, Tsinghua University, 2010

- **Toshiba Scholarship**, Tsinghua University, 2009

Research Experience

- Research Intern, **Microsoft Research Redmond**. Redmond, WA. 2015/05-2015/08

- Research Assistant, **Carnegie Mellon University**. Pittsburgh, PA. 2011/08-present

- Visiting Student, **Johns Hopkins University**. Baltimore, MD. 2014/01

- Research Assistant, **Tsinghua University**. Beijing, China. 2008/09-2011/07

Teaching Experience

- Teaching Assistant. 11-775 Large-Scale Multi-media Analysis, **Carnegie Mellon University**. Spring 2015

- Teaching Assistant. 11-751 Speech Recognition and Understanding, **Carnegie Mellon University**. Fall 2013

- Lectures on Discriminative Training, Speech Recognition and Understanding, **Carnegie Mellon University**. 2012-2015

- Teaching Assistant. Data Warehousing and Data Mining, **Tsinghua University**. Fall 2009

Publications

Journal Articles

- **Yajie Miao**, Hao Zhang, Florian Metze. Speaker Adaptive Training of Deep Neural Network Acoustic Models using I-vectors. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23 issue 11*, 2015.   PDF

Conference Papers

- **Yajie Miao**, Jinyu Li, Yongqiang Wang, Shixiong Zhang, Yifan Gong. Simplifying Long Short-Term Memory Acoustic Models for Fast Training and Decoding. To Appear in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2016*   PDF

- **Yajie Miao**, Mohammad Gowayyed, Xingyu Na, Tom Ko, Florian Metze, Alexander Waibel. An Empirical Exploration of CTC Acoustic Models. To Appear in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2016*   PDF

- **Yajie Miao**, Mohammad Gowayyed, Florian Metze. EESEN: End-to-End Speech Recognition using Deep RNN Models and WFST-Based Decoding. *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU) 2015*   **Best Paper Nomination**   PDF

- **Yajie Miao**, Florian Metze. On Speaker Adaptation of Long Short-Term Memory Recurrent Neural Networks. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2015*   PDF

- **Yajie Miao**, Florian Metze. Distance-Aware DNNs for Robust Speech Recognition. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2015*   PDF

- Justin Chiu, **Yajie Miao**, Alan W Black, Alexander Rudnicky. Distributed Representation-based Spoken Word Sense Induction. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2015*   PDF

- Yashesh Gaur, Florian Metze, **Yajie Miao**, Jeffrey Bigham. Using Keyword Spotting to Help Humans Correct Captioning Faster. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2015*   PDF

- Hao Zhang, **Yajie Miao**, Florian Metze. Regularizing DNN Acoustic Models with Gaussian Stochastic Neurons. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2015*   PDF

- Florian Metze, Ankur Gandhe, **Yajie Miao**, Zaid Sheikh, Yun Wang, Di Xu, Hao Zhang, Jungsuk Kim, Ian Lane, Won Kyum Lee, Sebastian Stuker, Markus Muller. Semi-supervised Training in Low-Resource ASR and KWS. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2015*   PDF

- **Yajie Miao**, Lu Jiang, Hao Zhang, Florian Metze. Improvements to Speaker Adaptive Training of Deep Neural Networks. *IEEE Spoken Language Technology Workshop (SLT) 2014*   **Best Poster Award**   PDF

- **Yajie Miao**, Florian Metze. Improving Language-Universal Feature Extraction with Deep Maxout and Convolutional Neural Networks. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2014*   PDF

- **Yajie Miao**, Hao Zhang, Florian Metze. Towards Speaker Adaptive Training of Deep Neural Network Acoustic Models. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2014*   PDF

- **Yajie Miao**, Hao Zhang, Florian Metze. Distributed Learning of Multilingual DNN Feature Extractors using GPUs. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2014*   PDF

- Lu Jiang, **Yajie Miao**, Yi Yang, ZhenZhong Lan, Alexander Hauptmann. Viral Video Style: A Closer Look at Viral Videos on YouTube. *ACM International Conference on Multimedia Retrieval (ICMR) 2014*   PDF

- **Yajie Miao**, Florian Metze, Shourabh Rawat. Deep Maxout Networks for Low-Resource Speech Recognition. *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU) 2013*   PDF

- **Yajie Miao**, Florian Metze. Improving Low-Resource CD-DNN-HMM using Dropout and Multilingual DNN Training. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2013*   PDF

- Jonas Gehring, Wonkyum Lee, Kevin Kilgour, Ian Lane, **Yajie Miao**, Alex Waibel. Modular Combination of Deep Neural Networks for Acoustic Modeling. *Annual Conference of the International Speech Communication Association (INTERSPEECH) 2013*   PDF

- **Yajie Miao**, Florian Metze, Alex Waibel. Subspace Mixture Model for Low-resource Speech Recognition in Cross-lingual Settings. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2013*   PDF

- **Yajie Miao**, Florian Metze, Alex Waibel. Learning Discriminative Basis Coefficients for Eigenspace MLLR Unsupervised Adaptation. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2013*   PDF

- Jonas Gehring, **Yajie Miao**, Florian Metze, Alex Waibel. Extracting Deep Bottleneck Features using Stacked Auto-encoders. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2013*   PDF

- **Yajie Miao**, Chunping Li, Jie Tang, Lili Zhao. Identifying New Categories in Community Question Answering Archives: A Topic Modeling Approach. *ACM Conference on Information and Knowledge Management (CIKM) 2010*   PDF

- **Yajie Miao**, Lili Zhao, Chunping Li, Jie Tang. Automatically Grouping Questions in Yahoo! Answers. *IEEE/WIC/ACM International Conference on Web Intelligence (WI) 2010*   PDF

- **Yajie Miao**, Chunping Li. Enhancing Query-oriented Summarization based on Sentence Wikification. *SIGIR Workshop on Feature Generation and Selection for Information Retrieval, 2010*   PDF

- **Yajie Miao**, Chunping Li. Mining Wikipedia and Yahoo! Answers for Question Expansion in Opinion QA. *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) 2010*   PDF

Toolkits and Datasets

- **EESEN**. A framework to build end-to-end speech recognition systems. **Author and Maintainer**.
  https://github.com/yajiemiao/eesen

- **PDNN**. An open-source toolkit for deep learning research and applications. **Author and Maintainer**.
  https://github.com/yajiemiao/pdnn

- **Kaldi+PDNN**. Fully-fledged recipes to build a wide range of DNN acoustic models based on the PDNN and Kaldi toolkits. **Author and Maintainer**.
  https://github.com/yajiemiao/pdnn

- **Kaldi**. The popular open-source toolkit for speech recognition. **Contributor**
  http://kaldi-asr.org/

- **CMU Viral Videos**. The largest public viral video dataset. **Main Contributor**
  https://sites.google.com/site/cmuviralvideos

Professional Activities

Conference Committee Member

- Technical Program Committee, INTERSPEECH 2016.

Reviewer

- Annual Conference of the International Speech Communication Association (INTERSPEECH). 2016.

- ACM Transactions on Knowledge Discovery from Data. 2015.

- IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2015, 2016.

- IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2015

- Machine Learning, 2015

- EURASIP Journal on Advances in Signal Processing, 2015

- IEEE Transactions on Neural Networks and Learning Systems, 2014

- Journal of Zhejiang University-SCIENCE C, 2014

- National Conference on Communications (NCC), 2014

Presentations and Talks

- *End-to-End Speech Recognition using Deep LSTMs, CTC Training and WFST Decoding*

    - Snapchat, 2016/03
    - TTIC Young Researcher Seminar Series, 2016/02
    - Massachusetts Institute of Technology (MIT), 2015/12

- *Enriching and Simplifying Deep Learning Models for Speech Recognition*

    - Microsoft, 2016/01

- *Simplified LSTM Models for Acoustic Modeling*

    - Oral Presentation at ICASSP. 2016/03
    - Microsoft, 2015/08

- *Incorporating Context Information into Deep Neural Network Acoustic Models*

    - Mitsubishi Electric Research Laboratories (MERL), 2015/12
    - Ph.D. Thesis Proposal, 2015/05
    - Sphinx Lunch Seminar, Carnegie Mellon University. 2015/04

- *Speaker Adaptive Training of Deep Neural Networks using I-vectors*

    - Facebook, 2015/04
    - Oral Presentation at INTERSPEECH. 2014/09
    - Sphinx Lunch Seminar, Carnegie Mellon University. 2014/08

Last updated: May 1, 2016