Suyoun Kim

Ph.D. Student
Carnegie Mellon University
Electrical and Computer Engineering

Phone: 412-692-0503
Email: suyoun@cmu.edu
Homepage: http://suyoun.kim

Research Interests

Speech Recognition, Deep Learning, and Machine Learning

Education

Carnegie Mellon University 2014 - Present

Ph.D. in Electrical and Computer Engineering

Advisors: Richard M. Stern, and Ian Lane

Carnegie Mellon University, 2012 - 2014

M.S. in Computer Science, Language Technologies Institute

Advisor: Madhavi Ganapathiraju

Georgia Institute of Technology, 2010 - 2011

M.S. in Computer Science

Dual degree with Korea University

Konkuk University, 2001 - 2005

B.S. in Multimedia

Awards

Samsung Graduate Fellowship, 2010 - 2011

Academic Training Program for top rated employees

Professional Experience

Microsoft Research (MSR), Speech Speech and Dialog Research Group, Summer 2017

Research Intern, responsible for research on speech recognition system

Collaboration with Mike Seltzer

Carnegie Mellon University, Electrical and Computer Engineering, 2014 - Present

Research Assistant, responsible for research on speech recognition

Mitsubishi Electric Research Laboratories (MERL), Speech & Audio Lab., Summer 2016

Research Intern, responsible for research on End-to-end speech recognition system

Collaboration with Shinji Watanabe, and Takaaki Hori

Carnegie Mellon University, School of Computer Science, LTI, 2012 - 2014

Research Assistant, responsible for research on computational biology, protein protein interaction, and drug repositioning

Samsung Electronics, Visual Display Division, 2005 - 2012

Software Engineer, responsible for development of Internet Protocol Set-top Box software on embedded linux system

Samsung Software Membership, 2004 - 2005

Software Engineer, responsible for development of 3D game for mobile phone

Publications

Suyoun Kim, and Ian Lane, "End-to-End Speech Recognition with Auditory Attention for Multi-Microphone Distance Speech Recognition", *(submitted to INTERSPEECH, 2017)*.

Suyoun Kim, Takaaki Hori, and Shinji Watanabe, "Joint CTC-Attention based End-to-End Speech Recognition using Multi-task Learning", *(in ICASSP, 2017)*. **[selected to give the oral presentation]**

Suyoun Kim, and Ian Lane, "Recurrent Models for Auditory Attention in Multi-Microphone Distant Speech Recognition", *(in INTERSPEECH, 2016)*.

Suyoun Kim, Bhiksha Raj, and Ian Lane, "Environmental Noise Embeddings for Robust Speech Recognition", *(in arXiv, 2016)*.

Suyoun Kim, and Ian Lane, "Recurrent Models for Auditory Attention in Multi-Microphone Distance Speech Recognition," *(ICLR Workshop, 2016)*.

Seungwhan Moon, **Suyoun Kim**, and Haohan Wang, "Multimodal Transfer Deep Learning with an Application in Audio-Visual Recognition" *(NIPS Workshop, 2015)*.

Dae Hyun Kim, **Suyoun Kim**, and Sung Kyu Lim, "Impact of nano-scale through-silicon vias on the quality of today and future 3D IC designs," in *Proceedings of the System Level Interconnect Prediction Workshop. IEEE Press, 2011*.

Patent Applications

"Attention-based Neural Networks for Multi-Microphone Speech Recognition," Provisional Patent Application 2016-127

Research Projects

End-to-end Speech Recognition | CMU, 2016 - Present
Conducted research on deep neural network models for end-to-end speech recognition system

End-to-end Speech Recognition | MERL, summer 2016
Proposed a joint CTC and Attention based end-to-end speech recognition model within the multi-task learning framework

Auditory Attention for Multi-microphone processing Distant Speech Recognition | CMU, 2015
Conducted research on multi-channel processing within acoustic model by attention-based recurrent neural networks in distant speech recognition scenarios

Noise Robust Speech Recognition | CMU, 2014
Proposed a noise robust model that incorporates background noise context features (environmental noise embeddings) learned from deep bottleneck network

Transfer Learning, Multimodal deep learning | CMU, 2014
Conducted research on multimodal deep learning framework that can transfer knowledge obtained from a single-modal DNN to a different modality, and developed a model that learns the analogy-preserving embeddings between audio and video representation

Protein Protein Interaction Prediction, Drug Repositioning | CMU, 2012
Developed the transfer learning model to overcome the scarcity of labeled data in human protein-protein interaction prediction by transferring generic features across species.

Teaching Experience

Carnegie Mellon University, Fall 2015
Teaching Assistant for Ian Lane, and Florian Metze, Speech Recognition and Understanding

Relevant Courses

Convex Optimization, Deep Learning, Machine Learning, Machine Learning for Signal Processing, Speech Recognition and Understanding, Fundamentals of Signal Processing, Algorithms for Natural Language Processing, Language and Statistics, Information Retrieval, Software Engineering

Last updated: April 1, 2017