Suyoun Kim

465 Stierlin Rd, Apt 47, Mountain View, CA 94043, USA
suyoun@cmu.edu ● +1 (412) 692-0503 ● http://www.suyoun.kim

EDUCATION

Carnegie Mellon University   Aug 2014 – Present
Ph.D. in Electrical and Computer Engineering
Research Interests: **Speech Recognition, Deep Learning, Machine Learning**
Advisers: Professors Richard M. Stern and Ian Lane

Carnegie Mellon University   Aug 2012 – Aug 2014
M.S. in Computer Science (Language Technologies Institute)
Research areas: Machine Learning, Computational Biology

Georgia Institute of Technology   Jan 2010 – May 2011
M.S. in Computer Science; Dual degree with Korea University

Konkuk University   Mar 2001 – Feb 2005
B.S. in Multimedia

RESEARCH EXPERIENCE

Carnegie Mellon University, Graduate Research Assistant   Aug 2012 – Present
1) Proposed a **joint CTC-Attention end-to-end speech recognition model** that directly transcribes speech to text in one step. The shared encoder is trained by both CTC and attention model objectives simultaneously within the **multi-task learning** framework. Achieved a relative improvement in CER of 6.6-10.3% on the WSJ and CHiME-4 tasks and accelerated the learning of the desired alignment.

2) Proposed an **attention-based RNN acoustic model** that directly combines **multi-channel** audio to generate phonetic states without requiring explicit signal preprocessing for speech enhancement. Embedded an **attention mechanism** within acoustic model to automatically tune its attention to a more reliable input source. Achieved a relative improvement in WER of 17% on the CHiME-3 challenge task.

3) Proposed a noise robust DNN acoustic model that incorporates background **noise context features learned from deep bottleneck network**. Achieved a relative improvement in WER of 0.41% in the mismatched condition.

4) Proposed a **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. Evaluated on a lip-reading task, and achieved an improvement in accuracy of 3.85%.

PUBLICATIONS

1) Suyoun Kim, Takaaki Hori, and Shinji Watanabe, "Joint CTC-Attention based End-to-End Speech Recognition using Multi-task Learning", *(submitted to ICASSP, 2017)*.
2) Suyoun Kim, and Ian Lane, "Recurrent Models for Auditory Attention in Multi-Microphone Distant Speech Recognition", *(in INTERSPEECH, 2016)*.
3) Suyoun Kim, Bhiksha Raj, and Ian Lane, "Environmental Noise Embeddings for Robust Speech Recognition", *(in arXiv, 2016)*.
4) Suyoun Kim, and Ian Lane, "Recurrent Models for Auditory Attention in Multi-Microphone Distance Speech Recognition," *(ICLR Workshop, 2016)*.
5) Seungwhan Moon, Suyoun Kim, and Haohan Wang, "Multimodal Transfer Deep Learning with an Application in Audio-Visual Recognition," *(NIPS Workshop, 2015)*.

WORK EXPERIENCE

Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA   May 2016 – Aug 2016
Research Intern, Speech & Audio, Host: Dr. Shinji Watanabe
Developed a novel End-to-End Speech Recognition System.

Samsung Electronics Co., Ltd., Suwon, South Korea   Jan 2005 – Jul 2012
Software Engineer, Visual Display Division
Developed Internet Protocol Set-top Box software based on Embedded Linux system.
Participated in an Academic Training Program (about 100 out of 100,000 employees are sponsored to attend graduate school).

Samsung Software Membership, Seoul, South Korea   Jan 2004 – Jan 2005
Intern, Internship program sponsored by Samsung Electronics.
Developed 3D mobile game.