SUYOUN KIM

Research Scientist ⋄ Meta, AI Speech
suyounkim@meta.com ⋄ http://suyoun.kim

RESEARCH INTERESTS

Speech Recognition, Deep Learning, Machine Learning, and Conversational AI

PH.D. THESIS

End-to-End Speech Recognition on Conversations, Carnegie Mellon University (2019)

Committee: Florian Metze, Richard M. Stern, Bhiksha Raj, Michael L. Seltzer, and Shinji Watanabe

EDUCATION

Carnegie Mellon University (2019)
Ph.D. in Electrical and Computer Engineering
Advisors: Richard M. Stern, and Florian Metze

Carnegie Mellon University (2014)
M.S. in Computer Science, Language Technologies Institute
Advisor: Madhavi Ganapathiraju

Georgia Institute of Technology (2011)
M.S. in Computer Science
Dual dgree with Korea University

Konkuk University (2005)
B.S. in Engineering

AWARDS

Center for Machine Learning and Health Fellowship in Digital Health, (2017 - 2018)

Samsung Graduate Fellowship, (2010 - 2011)

PROFESSIONAL EXPERIENCE

Meta, AI Speech (2019 - Present)
Research Scientist

Carnegie Mellon University (2012 - 2019)
Graduate Research Assistant

Abridge AI, Speech team (Summer 2018)
Research Intern, responsible for research on end-to-end speech recognition system on medical conversation dataset

Microsoft Research (MSR), Speech and Dialog Research Group (Summer 2017)
Research Intern, responsible for research on end-to-end speech recognition system
Collaboration with Mike Seltzer, Jinyu Li, and Rui Zhao

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

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 Intern

PUBLICATIONS

Suyoun Kim, Ke Li, Lucas Kabela, Rongqing Huang, Jiedan Zhu, Ozlem Kalinli, Duc Le, "Joint Audio/Text Training for Transformer Rescorer of Streaming Speech Recognition", *(in Findings of EMNLP, 2022)*.

Suyoun Kim, Duc Le, Weiyi Zheng, Tarun Singh, Abhinav Arora, Xiaoyu Zhai, Christian Fuegen, Ozlem Kalinli, and Michael L Seltzer, "Evaluating User Perception of Speech Recognition System Quality with Semantic Distance Metric", *(in INTERSPEECH, 2022)*. [selected to give the oral presentation]

Duc Le, Akshat Shrivastava, Paden Tomasello, **Suyoun Kim**, Aleksandr Livshits, Ozlem Kalinli, Michael L Seltzer, "Deliberation Model for On-Device Spoken Language Understanding", *(in INTERSPEECH, 2022)*. [selected to give the oral presentation]

Suyoun Kim, Abhinav Arora, Duc Le, Ching-Feng Yeh, Christian Fuegen, Ozlem Kalinli, and Michael L Seltzer, "Semantic Distance: A New Metric for ASR Performance Analysis Towards Spoken Language Understanding", *(in INTERSPEECH, 2021)*.

Duc Le, Mahaveer Jain, Gil Keren, **Suyoun Kim**, Yangyang Shi, Jay Mahadeokar, Julian Chan, Yuan Shangguan, Christian Fuegen, Ozlem Kalinli, Yatharth Saraf, and Michael L Seltzer, "Contextualized Streaming End-to-End Speech Recognition with Trie-Based Deep Biasing and Shallow Fusion", *(in INTERSPEECH, 2021)*.

Suyoun Kim, Yuan Shangguan, Jay Mahadeokar, Antoine Bruguier, Christian Fuegen, Michael L Seltzer, and Duc Le, "Improved Neural Language Model Fusion for Streaming Recurrent Neural Network Transducer", *(in ICASSP, 2021)*.

Chunxi Liu, Frank Zhang, Duc Le, **Suyoun Kim**, Yatharth Saraf, and Geoffrey Zweig, "Improving RNN transducer based ASR with auxiliary tasks", *(in SLT, 2021)*.

Suyoun Kim, Siddharth Dalmia, and Florian Metze, "Cross-Attention End-to-End ASR for Two-Party Conversations", *(in INTERSPEECH, 2019)*.

Suyoun Kim, Siddharth Dalmia, and Florian Metze, "Gated Embeddings in End-to-End Speech Recognition for Conversational-Context Fusion", *(in ACL, 2019)*.

Suyoun Kim, and Florian Metze, "Acoustic-to-Word Models with Conversational Context Information", *(in NAACL, 2019)*.

Suyoun Kim, and Florian Metze, "Dialog-context aware end-to-end Speech Recognition", *(in SLT, 2018)*.

Suyoun Kim*, Siddharth Dalmia*, and Florian Metze, "Situation Informed End-to-End ASR for CHiME-5 Challenge", *(in CHiME Workshop, 2018)*.

Suyoun Kim, Michael L. Seltzer, Jinyu Li, and Rui Zhao, "Improved Training for Online End-to-End Speech Recognition Systems", *(in INTERSPEECH, 2018)*.

Suyoun Kim, and Michael L. Seltzer, "Towards Language-universal End-to-End Speech Recognition", *(in ICASSP, 2018)*. [**selected to give the oral presentation**]

Shinji Watanabe, Takaaki Hori, **Suyoun Kim**, John R Hershey, and Tomoki Hayashi, "Hybrid CTC/ Attention Architecture for End-to-End Speech Recognition", *(IEEE Journal of Selected Topics in Signal Processing, 2017)*.

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

TEACHING EXPERIENCE

Carnegie Mellon University, Spring 2018
Teaching Assistant for Prof. Rohit Negi, Estimation, Detection and Learning

Carnegie Mellon University, Fall 2015
Teaching Assistant for Prof. Ian Lane, and Prof. 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