

Aswin Shanmugam Subramanian

Sr Applied Scientist, Microsoft

📍 Redmond, WA, USA

✉️ sas91@outlook.com

🌐 sas91.github.io

EDUCATION

- 2016-2022 **Doctor of Philosophy**, *Major: Electrical Engineering*
Dept. of Electrical and Computer Engineering, The Johns Hopkins University, USA.
Advisor: Dr. Shinji Watanabe.
Thesis: A Synergistic Combination of Signal Processing and Deep Learning for Robust Speech Recognition (January 2022).
- 2016-2017 **Master of Science in Engineering**, *Electrical & Computer Engineering*
The Johns Hopkins University, USA. CeDiD: 1857-C2C7-A8N8
- 2012-2015 **Master of Science (by Research)**, *Computer Science & Engineering*
Indian Institute of Technology Madras, India.
Advisor: Dr. Hema A. Murthy.
- 2008-2012 **Bachelor of Technology**, *Information Technology*
SSN College of Engineering, Anna University, India.

EXPERIENCE

- Oct 2022 - Present **Sr Applied Scientist**, *Microsoft*, AI Cognitive Services group
📍 Redmond, WA, USA
- Working on E2E multi-talker speech recognition.
- Sep 2021 - Sep 2022 **Research Scientist**, *Mitsubishi Electric Research Laboratories*
📍 Cambridge, MA, USA
- Worked on joint target speaker diarization and separation for meeting transcription.
 - Worked on audio separation, and anomaly detection.
- Fall 2017 - Summer 2021 **Research Assistant**, *Johns Hopkins University*, Advisor: Prof. Shinji Watanabe
📍 Baltimore, MD, USA
- Devised a paradigm called *directional ASR* that can train source localization with ASR objectives.
 - Participated in CHiME-5 and CHiME-6 challenges and our team finished second in both.
 - Devised a technique to estimate speech enhancement hyper-parameters using ASR objectives.
 - Developed a novel single channel speech enhancement method, and contributed to make Kaldi CHiME-4 recipe state-of-art.
- Summer 2020 & Summer 2019 **NLP Research Intern**, *Tencent AI Lab*, Mentors: Dr. Chao Weng & Dr. Dong Yu
📍 Bellevue, WA, USA
- 2020 - devised DNN approaches for multi-source localization and showed its effectiveness in cutting down ASR word error rates by about a factor of two.
 - 2019 - implemented a novel target speech extraction method with end-to-end speech recognition objectives.
- Summer 2018 **Research Intern**, *NTT Communication Sciences Lab*, Mentor: Dr. Marc Delcroix
📍 Kyoto, Japan
- Experimented with speaker and environment adaptation techniques for end-to-end noise robust speech recognition.

July 2015 - **Network Software Engineer, Intel**

August 2016 ♥ Bangalore, India

- IPv6 module for Axxia network accelerators and automation of performance benchmarks.

June 2012 - July **Project Associate, IIT Madras**, , Advisor: Prof. Hema A. Murthy

2015 ♥ Chennai, India

- Member of the TTS consortium that developed a common framework for HMM based speech synthesis of 13 Indian languages.
- Participated in Blizzard Challenge - 2014 & 2015.
- Developed an automatic segmentation tool for the TTS consortium.

TEACHING

Fall '17 & '18 **Course Assistant, Digital Signal Processing**, Johns Hopkins University

Spring '18 - '20 **Course Assistant, Information Extraction from Speech and Text**, Johns Hopkins University

KEY PUBLICATIONS[†]

1. Darius Petermann, Gordon Wichern, Aswin Subramanian, Jonathan Le Roux, “**Hyperbolic Audio Source Separation**,” in Proc. of **IEEE ICASSP 2023**, pp. 1-5.
2. Aswin Shanmugam Subramanian, Chao Weng, Shinji Watanabe, Meng Yu, and Dong Yu, “**Deep Learning based Multi-Source Localization with Source Splitting and its Effectiveness in Multi-Talker Speech Recognition**,” **Computer Speech and Language**, vol. 75, 2022.
3. Aswin Shanmugam Subramanian, Chao Weng, Shinji Watanabe, Meng Yu, Yong Xu, Shi-Xiong Zhang, and Dong Yu, “**Directional ASR: A New Paradigm for E2E Multi-Speaker Speech Recognition with Source Localization**,” in Proc. of **IEEE ICASSP 2021**, pp. 8433-8437.
4. Chenda Li, Jing Shi, Wangyou Zhang, Aswin Shanmugam Subramanian, Xuankai Chang, Naoyuki Kamo, Moto Hira, Tomoki Hayashi, Christoph Boeddeker, Zhuo Chen, and Shinji Watanabe, “**ESPNET-SE: End-to-End Speech Enhancement and Separation Toolkit Designed for ASR Integration**,” in Proc. of **IEEE SLT 2021**, pp. 785–792.
5. Wangyou Zhang, Aswin Shanmugam Subramanian, Xuankai Chang, Shinji Watanabe, and Yanmin Qian, “**End-to-End Far-Field Speech Recognition with Unified Dereverberation and Beamforming**”, in Proc. of **ISCA INTERSPEECH 2020**, pp. 324-328.
6. Aswin Shanmugam Subramanian, Chao Weng, Meng Yu, Shi-Xiong Zhang, Yong Xu, Shinji Watanabe, and Dong Yu, “**Far-Field Location Guided Target Speech Extraction using End-to-End Speech Recognition Objectives**,” in Proc. of **IEEE ICASSP 2020**, pp. 7299-7303.
7. Aswin Shanmugam Subramanian, Xiaofei Wang, Murali Karthick Baskar, Shinji Watanabe, Toru Taniguchi, Dung Tran, and Yuya Fujita, “**Speech Enhancement Using End-to-End Speech Recognition Objectives**,” in Proc. of **IEEE WASPAA 2019**, pp. 229–233.
8. Aswin Shanmugam Subramanian, Szu-Jui Chen, and Shinji Watanabe, “**Student-Teacher Learning for BLSTM Mask-based Speech Enhancement**,” in Proc. of **ISCA INTERSPEECH 2018**, pp. 3249–3253.
9. Szu-Jui Chen, Aswin Shanmugam Subramanian, Hainan Xu, and Shinji Watanabe, “**Building state-of-the-art distant speech recognition using the CHiME-4 challenge with a setup of speech enhancement baseline**,” in Proc. of **ISCA INTERSPEECH 2018**, pp. 1571–1575.
10. S Aswin Shanmugam, and Hema Murthy, “**A Hybrid Approach to Segmentation of Speech Using Group Delay Processing and HMM Based Embedded Reestimation**,” in Proc. of **ISCA INTERSPEECH 2014**, pp. 1648–1652.

[†]For full list of publications please visit <https://sas91.github.io/publication>