Aswin Shanmugam Subramanian

PhD Student, Johns Hopkins University

Baltimore, USA
+1-443-986-6765
aswin@jhu.edu
sas91.github.io

EDUCATION

2016-Present **Doctor of Philosophy**, *Electrical & Computer Engineering*.
Johns Hopkins University, USA.
Advisor: Dr. Shinji Watanabe.

2016-2017 **Master of Science**, *Electrical & Computer Engineering*, GPA: 3.7/4.
Johns Hopkins University, USA.

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

Fall 2017 - Present **Research Assistant**, *Johns Hopkins University*, Advisor: Prof. Shinji Watanabe.
Baltimore, MD, USA

- Fall 2020 - devised a paradigm called *directional ASR* that can train source localization with ASR objectives [1]. Contributed to speech enhancement functionality in ESPnet [3].
- Spring 2020 - Led speech separation efforts for the JHU team in CHiME-6 challenge and our team finished second in "diarization+ASR" track. Collaborated with SJTU on joint dereverberation & beamforming [4].
- Fall 2018 & Spring 2019 - worked on speech dereverberation in collaboration with Yahoo Japan [7]. Devised a technique to estimate speech enhancement hyper-parameters using ASR objectives [6].
- Spring 2018 - participated in CHiME-5 challenge as a member of the JHU-Hitachi team and contributed to the neural denoising module. Our team was placed in top 2.
- Fall 2017 - developed a novel single channel speech enhancement method [8], and contributed to make Kaldi CHiME-4 recipe state-of-art [9].

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 [2].
- 2019 - Implemented a novel target speech extraction method with end-to-end speech recognition objectives [4].

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 - August 2016 **Network Software Engineer**, *Intel*.
Bangalore, India

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

June 2012 - July 2015 **Project Associate**, *IIT Madras,* , Advisor: Prof. Hema A. Murthy.
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 [10].

SKILLS

Programming Languages Python, C, C++, Java, C#, Shell, Perl, MATLAB

Toolkits Kaldi, ESPnet, HTK, HTS, Festival, Chainer, Pytorch

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.

PROFESSIONAL SERVICES

- Reviewer - Interspeech 2015-2020, ICASSP 2017-2021, Speech Communication, APSIPA 2020.
- PC Member - Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL), 2020.

KEY PUBLICATIONS[†]

1. 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**," accepted in ***IEEE ICASSP 2021***.
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**," ***arXiv:2102.07955***.
3. 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.
4. 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.
5. 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.
6. 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.
7. Toru Taniguchi, Aswin Shanmugam Subramanian, Xiaofei Wang, Dung Tran, Yuya Fujita, and Shinji Watanabe, "**Generalized Weighted-Prediction-Error Dereverberation with Varying Source Priors for Reverberant Speech Recognition**," in Proc. of ***IEEE WASPAA 2019***, pp. 288–292.
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`