

Shanmukha Srinivas Battula

614-653-3204 | battula.12@osu.edu | linkedin.com/in/shanmukh-srinivas | github.com/ssnani

Summary

My current research projects focus on multi-modal speaker extraction using EEG. Experienced researcher specializing in deep learning, signal processing, and hybrid approaches for multi-channel speech enhancement, speaker separation, and localization tasks in noisy and reverberant environments for moving and stationary sources. Proficient in designing, analyzing, and debugging advanced state-of-the-art deep learning models with attention, recurrent, convolutional, and fully connected layers. Experienced with multi-node, multi-GPU, and mixed-precision training.

With 4.5 years of industry experience, I have developed expertise in end-to-end solutions for real-time speech enhancement on edge devices. My focus has been on research and prototyping in high-level languages, followed by efficient low-level language implementation, optimizations, and profiling at both the algorithmic and system levels.

Education

The Ohio State University

Ph.D. in Computer Science, Artificial Intelligence

Columbus, Ohio

Jan 2021 –

Indian Institute of Technology Hyderabad

B.S. in Electrical Engineering and Entrepreneurship

Hyderabad, Telangana

Aug 2013 – May 2017

Areas of Interest

- Deep Learning, Neural Networks, Machine Learning, Supervised Learning, Self-supervised Learning, Digital/Adaptive/Stochastic Signal Processing, Beamformers, Speech, EEG, Image, NLP, High-Performance Computing, Real-time systems

Technical Skills

Programming Languages: Python, C, MATLAB, Assembly

Deep Learning Frameworks: PyTorch, PyTorchLightning

Research Projects

EEG Guided Target Speaker Extraction

- Developing deep learning algorithms that efficiently decode human attention information recorded via EEG to perform attended speaker extraction in multi-speaker environments.
- Utilizing attention based models for EEG feature extraction and fusion.

Frame-Level Speaker Localization in Reverberant and Noisy Environments

- Developed a two-stage localization algorithm combining multi-channel complex spectral mapping models and weighted GCC-PHAT
- Utilized Densely connected Convolutional Recurrent Network (DC-CRN) and SpatialNet (modified transformer) models for enhancement
- Inclusion of novel Inter-channel phase difference in the loss function.

Moving Speaker Separation and Localization

- Developed location-based training algorithm and compared with UPIT for moving speaker separation
- Developed multi-task models for moving speaker separation/enhancement and localization
- Utilized DC-CRN, TF-GridNet, and SpatialNet models
- Dataset simulations contain linear and circular speaker trajectories with reverberation

Experience

AI Research Scientist Intern

Meta Reality Labs

May 2026 – Aug 2026

Redmond, USA

AI Research Scientist Intern

Meta Reality Labs

May 2025 – Aug 2025

Redmond, USA

- Researched on own voice suppression algorithms while maintaining speech enhancement for on-device smart glasses
- Implemented hybrid systems combining deep learning models with beamformers
- Evaluated on internal real product data in terms of perceptual enhancement and own voice suppression

AI Research Scientist Intern

Meta Reality Labs

May 2024 – Aug 2024

Redmond, USA

- Researched on-device speech enhancement and localization algorithms for smart glasses
- Implemented hybrid systems combining deep learning models with beamformers (MCWF, MVDR)
- Evaluated on internal real data in terms of perceptual enhancement and localization

Software Engineer

Qualcomm

June 2019 – Jan 2021

Hyderabad, India

- Worked on on-device deep learning-driven echo cancellation and noise suppression algorithms for smartphones
- Profiling and optimization using intrinsics and assembly for voice recognition and VoIP call applications.

Software DSP Research Engineer

Meeami Technologies

June 2017 – May 2019

Hyderabad, India

- Researched on-device real-time blind source separation algorithms for edge devices.
- Implemented frequency and time domain algorithms such as Independent Component/Vector Analysis, Non-Negative Matrix/Tensor Factorization, Second/Higher Order Statistics based Separation in offline and online scenarios.
- Online variants are developed using online whitening and LMS/RLS updates
- Developed a real-time Interference Detector module utilizing the long-term variability in the spectrum statistics such as mean, variance, median, etc

Intern

Qualcomm

May 2016 – Aug 2016

Hyderabad, India

- Developed Crash analysis (Python) Parser, DDR Stress Analysis (C) code

Publications

- **S. S. Battula**, A. Kumar, A. Bhaumik and D. Williamson, “EASE: Reassessing EEG-assisted target speaker extraction-Reliability and Vulnerability”, (*in review.*)
- **S. S. Battula**, H. Taherian, A. Pandey, D. Wong, B. Xu, and D. L. Wang, “Robust Frame-level Speaker Localization in Reverberant and Noisy Environments by Exploiting Phase Difference Losses”, *ICASSP, 2025*, pp. 1-5.
- **S. S. Battula**, H. Taherian, A. Pandey, D. Wong, B. Xu, and D. L. Wang, “Robust Frame-level Speaker Localization Guided by Multi-channel Speech Enhancement and Inter-channel Phase Difference Losses”, *JASA, 2025*, 158 (1), pp. 790–800.
- E.Grinstein, A. Pandey, C. Li, **S. S. Battula**, J. Azcarreta, J. Donley, S. Lee, A. Aroudi, C. Bilen, “Controlling the Parameterized Multi-channel Wiener Filter using a tiny neural network,” *WASPAA, 2025*.
- A. Pandey, A. Aroudi, C. Bilen, J. Donley, S. Lee, K. Li, J. A. Ortiz, B. Xu, E. Z. Grinstein, **S. S. Battula**, T. Hsieh, and D. Wong, “Ultra Low-Compute Complex Spectral Masking for Multichannel Speech Enhancement”, *Patent in review, 2025*.
- B. Appina, A. Jalli, **S. S. Battula**, S. S. Channappayya, “No Reference Stereoscopic Video Quality Assessment Using Joint Motion and Depth Statistics”, *ICIP, 2018*

Undergraduate Projects

No-Reference Stereoscopic Video Quality Assessment

May 2016 – Aug 2017

- Steerable pyramid decomposition on Extracted Depth and Motion Features
- Model the relation between the Depth and Motion Features in sub-band domain and predict quality using model parameters
- Estimating the depth from motion features and model parameters

Number plate Extraction from an Image

Jan 2016 – Apr 2016

- Locating the number plate region with Sobel Edge detector
- Extracted number plate region is transformed into binary image
- Template matching or SIFT or online OCR is used to extract the information.

Learning 3D structure from a single image

Aug 2015 – Nov 2015

- Conversion from RGB domain to YCbCr domain and Filtering the Image
- Extraction of Features for each patch at multiple scales
- Modelling the relation between depth and features extracted at multiple scales for a patch.
- Predict the depth for a given image using the model parameters and features.

JPEG 2000

Jan 2015 – Apr 2015

- Reversible color transformation on the image
- Discrete Wavelet Transform (DWT) and Quantisation
- Conversion to Bit-planes and Fractional Bit plane coding
- Arithmetic coding to the output of above stage

Extra Circular Activities AND Hobbies

- Graduate and Undergraduate Teaching Assistant, Placement Coordinator
- Peer Reviewer for conferences and journals
- Travel, Reading Books, Playing Games (Cricket, Badminton, Table Tennis, Carroms)