
Education

- 2024–2029 : **Ph.D., Computer Science**, *University of Maryland, College Park, Maryland, USA*,
Advisors: Prof Dinesh Manocha.
Research Focus: Speech and Audio Understanding
- 2020–2023 : **M.S., Computer Science**, *Indian Institute of Technology, Chennai, India*, **CGPA:** 9.12/10,
Advisors: Prof S.Umesh, Prof Hema A.Murthy.
Research Focus: Building General Purpose Audio Classification Models
- 2016–2020 : **B.E.**, *Army Institute Of Technology, Pune, India*, **CGPA:** 8.76/10.

Selected Publications

Journal Articles

- 2022 Sreyan Ghosh*, *Ashish Seth**, and S Umesh. Decorrelating Feature Spaces for Learning General-Purpose Audio Representations. *IEEE Journal of Selected Topics in Signal Processing (IEEE J-STSP)*, 2022.

In Conference Proceedings

- 2024 *Ashish Seth**, Ramaneswaran S, S Sakshi, Sonal Kumar, Sreyan Ghosh, and Dinesh Manocha. EH-MAM: Easy-to-Hard Masked Acoustic Modeling for Self-Supervised Speech Representation Learning. *accepted to (EMNLP) (Main (Oral))*, 2024.
- 2024 *Ashish Seth**, Sreyan Ghosh*, S. Umesh, and Dinesh Manocha. FusDom: Combining In-Domain and Out-of-Domain Knowledge for Continuous Self-Supervised Learning. *accepted in (ICASSP)*, 2024.
- 2024 *Ashish Seth**, Sreyan Ghosh*, S. Umesh, and Dinesh Manocha. Stable Distillation: Regularizing Continued Pre-training for Low-Resource Automatic Speech Recognition. *accepted in (ICASSP)*, 2024.
- 2024 Sreyan Ghosh*, *Ashish Seth**, Sonal Kumar, Utkarsh Tyagi, Chandra Kiran Reddy Evuru, S. Ramaneswaran, S Sakshi, Oriol Nieto, Ramani Duraiswami, and Dinesh Manocha. CompA: Addressing the Gap in Compositional Reasoning in Audio-Language Models. *accepted in (ICLR)*, 2024.
- 2024 Sreyan Ghosh, Sonal Kumar, *Ashish Seth*, Chandra Kiran Reddy Evuru, Utkarsh Tyagi, S Sakshi, Oriol Nieto, Ramani Duraiswami, and Dinesh Manocha. GAMA: A Large Audio-Language Model with Advanced Audio Understanding and Complex Reasoning Abilities. *accepted to (EMNLP) (Main)*, 2024.
- 2023 *Ashish Seth**, Mayur Hemani*, and Chirag Agarwal. DeAR: Debiasing Vision-Language Models with Additive Residuals. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
- 2023 *Ashish Seth**, Sreyan Ghosh*, S. Umesh, and Dinesh Manocha. SLICER: Learning Universal Audio Representations Using Low-Resource Self-Supervised Pre-Training. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023.
- 2023 Sreyan Ghosh*, *Ashish Seth**, S. Umesh, and Dinesh Manocha. MAST: Multiscale Audio Spectrogram Transformers. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023.

In Workshops

- 2023 *Ashish Seth**, Sreyan Ghosh*, S. Umesh, and Dinesh Manocha. UNFUSED : UNsupervised Finetuning Using Self supervised Distillation. In *2023 workshop on Self-supervision in Audio, Speech and Beyond, (ICASSP)*, 2023.
- 2022 Sreyan Ghosh*, *Ashish Seth**, Deepak Mittal, Maneesh Singh, and S. Umesh. DeLoRes: Decorrelating Latent Spaces for Low-Resource Audio Representation Learning. In *2022 workshop on Self-supervised Learning for Audio and Speech Processing, (AAAI)*, 2022.

Professional Work Experience

MTS (Member of Technical Staff) @ Adobe India

2023–2024 **Building AI driven Services for B2B Marketing.**

- Worked as a software engineer for building AI driven Services for B2B (Business-to-Business) marketing software.
- Led the team in building deep learning models that capture customer behaviors by combining customer and salesperson feedback.

Project Associate @ IIT (Indian Institute Of Technology) Madras

2020 – 2023 **Building ASR and TTS models for Indian Languages.**

- Build Automatic Speech Recognition (ASR) and Text-to-Speech Models for various Indian languages
- Presented our findings at many international conferences e.g. ICASSP, InterSpeech, AAAI, etc.

Building end-to-end pipeline for video-to-video translation.

- Developed tools and APIs for video-to-video translation for an e-learning platform like NPTEL (National Program on Technology Enhanced Learning) from English to several regional Indian languages.
- Led the team in building SOTA machine translation systems for various educational domains for Indian languages and collecting diverse data for building such models

Advisor: Prof. Hema A. Murthy, Prof. S. Umesh

MDSR (Media Data Science and Research) Intern @ Adobe India

May 2022 – **Mitigating Societal Bias in Large Foundation Models.**

- Aug 2022 ◦ Led the team that proposed a novel approach for debiasing Vision-Language models such as CLIP, BLIP, etc, against various protected attributes like race, gender, and age.

Mentors : Dr. Chirag Agarwal, Mayur Hemani

Research and Teaching Experience

Research Scholar @ GAMMA Lab, University Of Maryland

2024 – **Speech and Audio Understanding.**

- present ◦ I am currently working on building complex Q/A reasoning framework for enhancing and evaluating speech and audio understandability in large foundation models.

Advisor : Prof. Dinesh Manocha

Teaching Assistant @ University Of Maryland

fall 2024 **TA for Graduate NLP (CMSC 723).**

Research Scholar @ SPRING Lab, IIT Madras

2021 – 2023 **Towards General-Purpose Audio Representation Learning.**

- Developed novel self-supervised frameworks to train audio classification models under low resource settings (in terms of data and model size)
- We proposed an open-source benchmark called *LAPE* (Low-Resource Audio Pretraining and Evaluation) which contains multiple speech and non-speech classification tasks.

2022 – 2023 **Domain Adaptation in Speech Recognition Models (ASR).**

- Proposed novel solutions to tackle domain mismatch while fine-tuning a pre-trained SSL models.

Advisors : Prof. S. Umesh, Prof. Hema A. Murthy

Academic Achievements & Recognitions

- 2023 Best Project Award, at Garage Week, Adobe
- 2023 Best Demo Award at Hack-X, Adobe
- 2023 Awarded *IEEE Signal Processing Society Travel Grant*
- 2023 Best Presentation/Demo Award in *Speech and Language Technology (SLT) Hackathon*, Qatar
- 2021 Runner up in *Multilingual and Code-Switching ASR Challenges for Low Resource Indian Languages* in Interspeech

Community Service

- Reviewer IEEE-TNNLS (2024), TMLR(2024), NeurIPS (2024), ICASSP (2024, 2023), EMNLP(2024), ACL (2023)

Programming Skills

- Programming Languages Python (Framework: PyTorch, Keras, Numpy, Pandas, Sklearn, Matplotlib, Seaborn), C, C++, Java (Framework: Spring, Scala)
- Web Technologies HTML 5, PHP, JSP, Javascript, ReactJS
- Database SQL, MySQL, MongoDB