# Ashish Seth

(+91) 9370475238
ashishseth16226@gmail.com
My Webpage
Github Linkedin


## Education

2020–2023 : **Master, Computer Science & Engineering**, *Indian Institute of Technology*, Madras, **CGPA**: 9.12/10, **Advisors**: Prof S.Umesh, Prof Hema A.Murthy.
**Research Focus**: Self-Supervised Learning, Audio Classification, Speech Recognition (ASR)

2016–2020 : **Bachelor of Engineering, Computer Science & Engineering**, *Army Institute Of Technology*, Pune, **CGPA**: 8.76/10.

## 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***, Sreyan Ghosh*, S. Umesh, and Dinesh Manocha. Stable distillation: Regularizing continued pre-training for low-resource automatic speech recognition. *Under Review in **(ICASSP)***, 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. *Under Review 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. *Under Review in **(ICLR)** (Avg Score: 6.5)*, 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: Lerning universal audio representations using low-resource self-supervised pre-training. In *IEEE International Conference on Acoustics, Speech and Signal Processing **(ICASSP)***, 2023.

2023 Anusha Prakash, Arun Kumar, **Ashish Seth**, and et al. Technology Pipeline for Large Scale Cross-Lingual Dubbing of Lecture Videos into Multiple Indian Languages. In ***(INTERSPEECH)***, 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.

2022 Anish Bhanushali, Grant Bridgman, Deekshitha G, Prasanta Ghosh, Pratik Kumar, Saurabh Kumar, Adithya Raj Kolladath, Nithya Ravi, Aaditeshwar Seth, **Ashish Seth**, and et al. Gram Vaani ASR Challenge on spontaneous telephone speech recordings in regional variations of Hindi. In ***(INTERSPEECH)***, 2022.

2021 Mari Ganesh Kumar, Jom Kuriakose, Anand Thyagachandran, Arun Kumar A*, **Ashish Seth***, Lodagala V.S.V. Durga Prasad*, Saish Jaiswal, Anusha Prakash, and Hema A. Murthy. Dual Script E2E Framework for Multilingual and Code-Switching ASR. In ***(INTERSPEECH)***, 2021.

### 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.

### Arxiv Pre-prints

2023 **Ashish Seth***, Lodagala V S V Durga Prasad*, Sreyan Ghosh*, and S. Umesh. Analyzing the factors affecting usefulness of self-supervised pre-trained representations for speech recognition. 2023.

2023 Sreyan Ghosh, Sandesh V Katta, **Ashish Seth**, and Srinivasan Umesh. Deep clustering for general-purpose audio representations. 2023.

## Professional Work Experience

### MTS (Member of Technical Staff) @ Adobe India

2023–present **Building API and Services for B2B marketing software**.

- Working as a software engineer in building API and Services for B2B (Business-to-Business) marketing software.
- Led the team in building deep learning models that capture customer behaviors and recommend products 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 of educational lectures into regional Indian Languages**.

- Was part of the team that built tools and APIs for video-to-video translation for an e-learning platform 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: Dr. Hema A. Murthy, Dr. S. Umesh

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

May 2022 – Aug 2022 **Developed novel debiasing technique for Vision-Language models**.

- 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.
- We presented our findings at an international conference, CVPR 2023 in Vancouver, Canada

Mentors : Dr. Chirag Agarwal, Mayur Hemani

## Research Experience

### Research Scholar @ SPRING Lab, IIT Madras

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

- Under the supervision of *Prof. S. Umesh*, I 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 contain multiple speech and non-speech classification task. Under this project, we also open-source all our models and codes for the research community.
- We published our findings at many international conferences and journals e.g. ICASSP-2023, InterSpeech-2022, AAAI-2022, IEEE J-STSP-2022.

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

- Under the supervision of *Prof. S. Umesh* and *Prof. Hema A. Murthy*, I proposed novel solutions to tackle domain mismatch while fine-tuning a pre-trained ASR model.
- Currently our work is under review for ICASSP 2024.

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

### External Collaborator @ GAMMA Lab, University Of Maryland

2022 – 2023 **Compositional reasoning in Foundational Models**.

- Under the supervision of *Prof. Dinesh Manocha*, I collaborated with *GAMMA* lab to develop an open-source benchmark for assessing compositional reasoning in Audio-Language Models
- Led the team for building modular approaches for training Audio-Language models (ALMs), enabling them to comprehend complex compositions of acoustic events in real-world acoustic scenes.
- This work is currently under review for *ICLR 2024* (avg score is 6.5).

## Academic Achievements & Recognitions

2023 Presented 4 conference papers in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece

2023 Awarded *IEEE Signal Processing Society Travel Grant* to attend ICASSP

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

## 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 |