Sreyan Ghosh

+1 240 899 7685 | sreyang@umd.edu | LinkedIn | GitHub | Website | Google Scholar | College Park, USA

Education

University of Maryland (Advised by Dr. Dinesh Manocha) — College Park, USA
Ph.D. in Computer Science; ***GPA: 3.9/4.00*** — *Aug 2022 – May 2027*
M.S. in Computer Science; ***GPA: 3.9/4.00*** — *Aug 2022 – May 2024*

Christ University — Bangalore, India
B.Tech in Computer Science and Engineering; ***GPA: 8.7/10*** — *June 2016 – Jun 2020*

Conference Papers (in chronological order)

AdVerb: Visually Guided Audio Dereverberation
Sanjoy Chowdhury*, *Sreyan Ghosh**, Subhrajyoti Dasgupta, Anton Ratnarajah, Utkarsh Tyagi, Dinesh Manocha
ICCV 2023

MMER: Multimodal Multi-task Learning for Speech Emotion Recognition
Sreyan Ghosh, Utkarsh Tyagi, S Ramaneswaran, Harshvardhan Srivastava, Dinesh Manocha
InterSpeech 2023

ACLM: A Selective-Denoising based Generative Data Augmentation Approach for Low-Resource Complex NER
*Sreyan Ghosh**, Utkarsh Tyagi*, Manan Suri, Sonal Kumar, S Ramaneswaran, Dinesh Manocha
ACL 2023

BioAug: Conditional Generation based Data Augmentation for Low-Resource Biomedical NER
*Sreyan Ghosh**, Utkarsh Tyagi*, Sonal Kumar*, Dinesh Manocha
SIGIR 2023

MAST: Multiscale Audio Spectrogram Transformers
*Sreyan Ghosh**, Ashish Seth*, S. Umesh, Dinesh Manocha
IEEE ICASSP 2023

SLICER: Learning universal audio representations using low-resource self-supervised pre-training
Ashish Seth*, *Sreyan Ghosh**, S. Umesh, Dinesh Manocha
IEEE ICASSP 2023

data2vec-aqc: Search for the right Teaching Assistant in the Teacher-Student training setup
Lodagala V S V Durga Prasad*, *Sreyan Ghosh**, S. Umesh
IEEE ICASSP 2023

PADA: Pruning Assisted Domain Adaptation for Self-Supervised Speech Representations
Lodagala V S V Durga Prasad, Sreyan Ghosh, S. Umesh
IEEE SLT 2022

CCC-wav2vec 2.0: Clustering aided Cross Contrastive Self-supervised learning of speech representations
Lodagala V S V Durga Prasad, Sreyan Ghosh, S. Umesh
IEEE SLT 2022

Span Classification with Structured Information for Disfluency Detection in Spoken Utterances
Sreyan Ghosh, Sonal Kumar, Yaman Kumar Singla, Rajiv Ratn Shah, S. Umesh
Interspeech 2022

DeToxy: A Large-Scale Multimodal Dataset for Toxicity Classification in Spoken Utterances
Sreyan Ghosh, Samden Lepcha, Sakshi, Rajiv Ratn Shah, S. Umesh
Interspeech 2022

End-to-end Named Entity Recognition from English Speech
Hemant Yadav, Sreyan Ghosh, Yi Yu, Rajiv Ratn Shah
Interspeech 2022

Journal and Workshop Papers (in chronological order)

Decorrelating Feature Spaces for Learning General Purpose Audio Representations

*Sreyan Ghosh**, Ashish Seth*, S. Umesh

IEEE JSTSP Special Issue on Self-Supervised Learning for Speech and Audio Processing

DeLoRes: Decorrelating Latent Spaces for Low-Resource Audio Representation Learning

*Sreyan Ghosh**, *Ashish Seth**, S. Umesh

SAS Workshop @ AAAI 2022

Cisco at SemEval-2021 Task 5: What's Toxic?: Leveraging Transformers for Multiple Toxic Span Extraction from Online Comments

Sreyan Ghosh, Sonal Kumar

SemEval-2021 @ ACL 2021

Cisco at AAAI-CAD21 shared task: Predicting Emphasis in Presentation Slides using Contextualized Embeddings

Sreyan Ghosh, Sonal Kumar, Harsh Jalan , Hemant Yadav , Rajiv Ratn Shah

CAD-21 @ AAAI 2022

Professional Work Experience

NVIDIA Bangalore, India

Research Scientist *April 2022 – August 2022*

- Worked as a senior solutions architect in the professional services team at NVIDIA. Responsible for delivering deep-learning based NLP solutions to NVIDIA's premier customers around the globe.
- Contributed to AI R&D at NVIDIA. Published 2 papers at **IEEE SLT 2022**.

Cisco Systems Bangalore, India

Software Engineer II *Aug 2020 – March 2022*

- Worked as a senior software engineer in the automation and orchestration team under the Customer Experience BU. Built network assurance solutions for Cisco's telecom customers, leveraging state-of-the-art algorithms for anomaly detection at scale. Built a critical component in Cisco's first telemetry-based network assurance solution.
- Lead the development of an AI-based network security system for one of Cisco's telecom customers.
- Was part of the AI team that developed Cisco's first contact center solution, leveraging state-of-the-art NLP algorithms.
- Contributed to AI R&D at Cisco by representing Cisco at various conferences.

Research and Teaching Experience

Gamma Lab @ UMCP College Park, Maryland, USA

Research Assistant *Fall 2022 – Present*

- My primary research focuses on low-resource (labeled data and compute) learning with applications in speech, NLP or vision. In this area, I solve problems using self-supervised learning, synthetic data augmentation, etc.
- Advised by Dr. Dinesh Manocha.

University of Maryland College Park, Maryland, USA

Teaching Assistant *Fall 2022 – Fall 2022*

- TA for Introduction to NLP (CMSC 470) for Fall 2022

University of Maryland College Park, Maryland, USA

Teaching Assistant *Fall 2022 – Fall 2022*

- TA for Introduction to NLP (CMSC 470) for Fall 2022

Speech Lab, Indian Institute of Technology Madras Chennai, India

Project Associate (Research) *June 2021 – August 2022*

- Worked under the supervision of **Dr. S. Umesh** in the area of self-supervised learning for Speech and Audio processing. Exploring techniques to devise lighter-weight models and efficient algorithms to make supervised and self-supervised learning in speech and audio amenable to resource-constrained scenarios (both data and compute). Paper accepted to **SAS Workshop at AAAI 2022**, **Interspeech 2022**, **IEEE JSTSP Special Issue** and **IEEE SLT 2022**.

MIDAS Labs, IIIT-Delhi Delhi, India

Research Assistant (part-time) *December 2019 – August 2022*

- Worked under the supervision of **Dr. Rajiv Ratn Shah** in the areas of Speech and Natural Language Processing. Worked on building ASR systems for low-resource Indian Languages (mono and multi-lingual) and Indian Accented English which served as a critical component for other systems built by the lab. Worked in the field of content moderation in modalities of both speech and text. Currently exploring multi-modal techniques for identifying disfluencies in spoken utterances. Published papers at **AAAI 2021**, **ACL 2021** and **Interspeech 2022**.

Internships

Adobe Seattle, WA, USA

Research Scientist Intern *May 2023 – Present*

- Working as a research scientist intern at Adobe in the multi-modal understanding group using Large Language Models.

Google Summer of Code Remote

Open Source Developer *April 2022 – August 2022*

- Working on building deep learning based NLP (speech and text) notebooks using Tensorflow and Keras.
- Link to PRs and code contributed on personal website.

Cisco Systems Bangalore, India

Software Developer Intern *January 2020 – June 2020*

- Worked on a project, End-to-End Named Entity Recognition from English Speech under the guidance of **Dr. Rajiv Ratn Shah** as part of my bachelors thesis. Paper accepted at **Interspeech 2020**.

MIDAS Labs, IIIT-Delhi Delhi, India

Research Intern *January 2020 – June 2020*

- Worked on building a VOIP (Voice Over IP) Traffic Analyzer to detect anomalous SIP messages using machine learning.

Noodle.ai Bangalore, India

Data Science Intern *December 2019 – December 2019*

- Worked on multivariate time-series anomaly detection in high-frequency IoT sensor data obtained from steel manufacturing machines.

TEG Analytics Bangalore, India

Data Science Intern *April 2019 – May 2019*

- Worked under the healthcare intelligence division to provide insights from insurance plan enrollment data, for private insurance companies in the US.
- Used Machine Learning and Deep Learning techniques to predict plan enrollment for insurance companies.

Awards & Achievements

Recognised by Cisco CX CTO and higher management on multiple occasions for my research and innovation initiatives.

Awarded the **Graham Bell Award** for being one of the most competitive undergraduates to have graduated in the year 2020.

Winner of Cisco Collab Hacks 2020.

Winner of P&G Global Innovation Challenge 2020.

Appeared on the cover page of Analytics India Magazine twice for winning national level hackathons in 2020 (TEG Analytics and Uber Hackathon)

Winner of Hindustan Unilever BFS Technology Hackathon.

Winner of various inter-college and intra-college hackathons sponsored by MNCs and the Government (Including a bronze medal at Kaggle).

Awards & Achievements

Recognised by Cisco CX CTO and higher management on multiple occasions for my research and innovation initiatives.

Awarded the **Graham Bell Award** for being one of the most competitive undergraduates to have graduated in the year 2020.

Winner of Cisco Collab Hacks 2020.

Winner of P&G Global Innovation Challenge 2020.

Appeared on the cover page of Analytics India Magazine twice for winning national level hackathons in 2020 (TEG Analytics and Uber Hackathon)

Winner of Hindustan Unilever BFS Technology Hackathon.

Winner of various inter-college and intra-college hackathons sponsored by MNCs and the Government (Including a bronze medal at Kaggle).

Major Projects

LAPE | *GitHub*

- LAPE is an easy-to-use toolkit for audio processing. In its initial release, LAPE supports Self-Supervised Learning (SSL)-based Upstream Pre-training and Downstream Fine-tuning. LAPE, originally introduced in this paper, integrates all our research on low-resource audio processing in one unified framework. We open-source LAPE to promote more research in this space.

Intelligent Minutes of the Meeting

- Used Automatic Speech Recognition and Speaker Diarization together with transformer-based text summarization and Named-Entity-Recognition to create a web app for intelligent Minutes of the Meetings.
- Application has capabilities to take meeting recordings and output summary, diarized transcripts, and key points from the meeting, including dated items and to-dos.

Skills

Programming: Python, Java, C, MySQL, HTML, CSS

Frameworks: PyTorch, Keras, Tensorflow, Django, Flask, Spark

Tools: GIT, Android, Tableau, Power BI, AWS, GCP, Rest API, Docker, K8s

Concepts: Speech and Natural Language Processing, Software Development, Functional programming, Object-oriented programming, Machine Learning, Deep Learning, Image Processing, Cloud Computing

Community Service

Reviewer for: EMNLP 2023. ACL 2023, ICASSP 2023, InterSpeech 2023, AAAI 2023, ACL 2021, HASOC 2021, Electronic Letters (IET)

Team Lead and Co-founder for: Neuron, Christ University's first AI club focused on research, served as the first Vice President of the club.

Lecturer of: SLP at University of Buffalo, New York.