

Sarthak Kumar Maharana

✉ skm200005@utdallas.edu
🌐 <https://sarthaxxxx.github.io/>
in [sarthak9811](#)
🐙 [sarthaxxxx](#)

Research Interests

Short summary: Computer Vision and Machine Learning. Core interests: Multimodal Learning, Inference-time Learning, Data-driven ML, Robustness, Generalization

Education

- 2023-Present **PhD in Computer Science**, *The University of Texas at Dallas*, Richardson, USA, 4.0/4.0 (core)
Advisor: Dr. Yunhui Guo; Research Focus: Multimodal Learning, Data-Efficient Deep Learning, Computer Vision, Machine Learning.
- 2021-2023 **MS in Electrical Engineering**, *University of Southern California (USC)*, Los Angeles, USA, 3.85/4.0
Track: Machine Learning and Data Science.
- 2016-2020 **BTech in Electrical and Electronics Engineering**, *International Institute of Information Technology (IIIT)*, Bhubaneswar, India, 8.32/10, Thesis: 9.85/10.0
Thesis: *Acoustic-to-articulatory inversion of dysarthric speech by using cross-corpus acoustic-articulatory data.*

Experiences

- May'25- **Dolby Laboratories**, Sunnyvale, USA
Aug'25 *PhD Research Intern* — Sight Experiences Lab (Advanced Technology Group)
○ Proposed a novel model-agnostic framework for robust continual audio-visual learning. Patent filed.
- May'22- **USC's Mark and Mary Stevens Neuroimaging and Informatics Institute**, *Neuro Imaging Computing Research*, Los Angeles, USA
July'23 *Student Researcher* — Advisor: Dr. Yonggang Shi
○ Developed a tool to perform automatic tractography of the brainstem using d-MRI images.
○ Leveraged image registration and label fusion methods to automatically generate the anatomical ROIs.
- Dec'21- **Signal Analysis and Interpretation Lab (SAIL)**, *USC*, Los Angeles, USA
Dec'22 *Student Researcher* — Advisor: Dr. Shrikanth (Shri) Narayanan
○ Speaker recognition from rt-MRI videos, based on an unsupervised disentanglement representation learning scheme.
- Dec'19- **Signal Processing and Interpretation (SPIRE) Lab**, *Indian Institute of Science (IISc)*, Bengaluru, India
Sep'20 *Student Researcher* — Advisor: Dr. Prasanta Kumar Ghosh
○ Studied acoustic-to-articulatory inversion (AAI) for dysarthric speech at low-resource data conditions involving Indian languages, using joint learning and multi-task training. [ICASSP'21]
○ Conditioned the jointly-trained AAI model with x-vectors to study its benefits on the AAI performance of dysarthric subjects.

Publications/Preprints

AVROBUSTBENCH: Benchmarking the Robustness of Audio-Visual Recognition Models at Test-Time

Sarthak Kumar Maharana, Saksham Singh Kushwaha, Baoming Zhang, Adrian Rodriguez, Songtao Wei, Yapeng Tian, and Yunhui Guo.

[Under Review]

BATCLIP: Bimodal Online Test-Time Adaptation for CLIP

Sarthak Kumar Maharana, Baoming Zhang, Leonid Karlinsky, Rogerio Feris, and Yunhui Guo.

IEEE/CVF International Conference on Computer Vision (ICCV) 2025

[Paper][Code][Project]

SELECT: A Submodular Approach for Active LiDAR Semantic Segmentation

Ruiyu Mao, **Sarthak Kumar Maharana**, Xulong Tang, and Yunhui Guo.

[Under Review]

PALM: Pushing Adaptive Learning Rate Mechanisms for Continual Test-Time Adaptation

Sarthak Kumar Maharana, Baoming Zhang, and Yunhui Guo.

Proceedings of the 39th AAAI Conference on Artificial Intelligence 2025 (AAAI)

[Paper][Code][Project]

Variational Diffusion Unlearning: A Variational Inference Framework for Unlearning in Diffusion Models

Subhodip Panda, MS Varun, Shreyans Jain, **Sarthak Kumar Maharana**, and Prathosh AP.

NeurIPS Safe Generative AI Workshop 2024

[Paper]

STONE: A Submodular Optimization Framework for Active 3D Object Detection

Ruiyu Mao, **Sarthak Kumar Maharana**, Rishabh K Iyer, and Yunhui Guo.

Advances in Neural Information Processing Systems 2024 (NeurIPS)

[Paper][Code]

Not Just Change the Labels, Learn the Features: Watermarking Deep Neural Networks with Multi-View Data

Yuxuan Li, **Sarthak Kumar Maharana**, and Yunhui Guo.

European Conference on Computer Vision 2024 (ECCV)

[Paper] [Code]

Acoustic-to-articulatory inversion for dysarthric speech: Are pre-trained self-supervised representations favorable?

Sarthak Kumar Maharana, Krishna Kamal Adidam, Shoumik Nandi, and Ajitesh Srivastava.

2024 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) Workshops - Self-supervision in Audio, Speech, and Beyond.

[Paper] [Poster]

Acoustic-to-articulatory inversion for dysarthric speech by using cross-corpus acoustic-articulatory data

Sarthak Kumar Maharana, Aravind Illa, Renuka Mannem, Yamini Belur, Preetie Shetty, Veeramani Preethish Kumar, Seena Vengalil, Kiran Polavarapu, Nalini Atchayaram, and Prasanta Kumar Ghosh.

2021 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)

[Paper] [Poster] [Talk]

Selected Projects

- Oct'22- **Understanding Multi-Modal Speaker Recognition via Disentangled Representation Learning**, USC
Dec'22
- Presented an adversarial invariance approach to address multimodal speaker recognition, robust to various sources of variability present in videos and speech. [Code], [Report]
- Nov'22- **Understanding Linguistic Patterns for Text-Based Speaker Classification**,
Dec'22 USC
- Studied various text feature extraction methodologies using pre-trained models and classification algorithms and compared them to build a computationally efficient system targeted for text-based speaker classification. [Code], [Report]
- Apr'22- **The Effect of Conditioning of Trigonometric Transformations of Dates with**
May'22 **Meteorological Data in Forest Fires Prediction: An Experimental Study**,
USC
- Studied the effects of conditioning a trigonometric transformation of dates with meteorological data, that would aid in predicting the occurrence of forest fires in Algeria. [Code]
- Jan'19- **Single Image Haze Removal using Dark Channel Prior**, IIIT-Bh
Mar'19
- The dark channel is based on the following observation relating to outdoor images: In most of the non-sky patches, at least a certain color channel has at least one dark channel i.e. it has low intensities tending to zero.
 - Redesigned an algorithm regarding this. [Link].

Skills

- Languages *Advanced*: Python, MATLAB; *Familiar*: C++, Java, Bash
- ML Libraries *Advanced*: PyTorch, Keras, TensorFlow; *Intermediate*: OpenCV, scikit-learn
- Softwares *Intermediate*: RStudio, ITK-SNAP, BrainSuite
- Others Git, \LaTeX ; OS: Unix, Windows

Talks

PALM: Pushing Adaptive Learning Rate Mechanisms for Continual Test-Time Adaptation

AAAI'25 (Oral)

Acoustic-to-articulatory inversion of dysarthric speech by using cross-corpus acoustic-articulatory data

ICASSP'21

Teaching Experience

- Aug'23- **Artificial Intelligence, Operating Systems**, *Teaching Assistant*, UTD — Develop-
Dec'23 ing course materials, grading assignments, and holding doubt-clearing sessions.
- Jan'22- **A Computational Introduction to Deep Learning**, *Grader and Mentor*, USC —
May'22 Grading assignments, holding office hours, monitoring online forums, and project grading.

Academic Service

Workshop Organizer

- 1st Workshop on Multimodal Continual Learning, ICCV'25
- 2nd Workshop on Test-Time Adaptation: Putting Updates to the Test!, ICML'25

Invited Reviewer

- CVPR'25, ICLR'25, NeurIPS-W'24, BMVC'24, CVPR-W'24, ECCV'24, AAAI'24.

Volunteer Work

- Aug'23- **CORD.ai** — Helping build CORD.ai, an AI research community, as a research advisor and volunteer.
- Oct'21- **USC IEEE Graduate Society** — Attended group meetings, strengthened academic and social growth of the members, and hosted workshops.
- May'23
- Aug'20- **PyCon India 2020** — Content writer for social media handles, helped the promotions team, and created virtual swags.
- Oct'20
- Jul'18- **International Association of Engineers (IAENG)** — Student member.

Awards

- 2024 **Oxford Summer School in Machine Learning 2024**, Accepted to the MLx Representation Learning and Generative AI track
- 2023 **Computer Science PhD Fellowship**, The University of Texas at Dallas
- 2020 **Governing Body Merit Scholarship**, Award of INR 15k, *Academic year: 2019-2020*, IIIT-Bh
- 2019 **Dean's List** [Link], IIIT-Bh
- 2019 **Summer Research Fellowship**, *Indian Academy of Sciences (10% selection rate)* [Link]