

APOORVA KULKARNI

apoorvak@umd.edu [◇ GitHub](#) [◇ LinkedIn](#) [◇ Google Scholar](#)

SUMMARY

MS CS student at University of Maryland (graduating 2027) pursuing research in multimodal foundation models

EDUCATION

University of Maryland - College Park **2025–2027**
Master of Science in Computer Science **(3.93/4.0)**

PES University, Bangalore **2020–2024**
B.Tech in Computer Science and Engineering **9.69/10 (3.97/4.0)**
Awards: Prof. CNR Rao Scholarship (top 2% of cohort in every semester)

SELECTED PUBLICATIONS

- [1] **A Closer Look at Failure Modes in Temporal Understanding of Large Audio-Language Models**
Apoorva Kulkarni, Kaousheik J, Sreyan G, Sarah Wiegrefe, Dinesh Manocha, Ramani Duraiswami, **Under Review - ACL 2026**
- [2] **FESTA: Functionally Equivalent Sampling for Trust Assessment of Multimodal LLMs**
Debarpan Bhattacharya*, Apoorva Kulkarni*, Sriram Ganapathy, **Findings of EMNLP 2025**
- [3] **Benchmarking and Confidence Evaluation of Audio-LLMs For Temporal Reasoning**
Debarpan Bhattacharya*, Apoorva Kulkarni*, Sriram Ganapathy, **INTERSPEECH 2025**
- [4] **The Second DISPLACE Challenge: DIarization of SPeaker and LAnguage in Conversational Environments**
Shareef Babu Kalluri, Prachi Singh, Pratik Roy Chowdhuri, Apoorva Kulkarni, et al., **INTERSPEECH 2024**

EXPERIENCE

Research Assistant Sept 2025 - present

GAMMA Lab - Audio, UMD College Park

- Built a 1,027-question benchmark to diagnose temporal reasoning failures in Large Audio-Language Models, conducted mechanistic attention analysis revealing that models achieve only 58.6-68.75% accuracy due to imprecise attention distribution, and demonstrated layer-specific interventions improve performance
- Leveraging reinforcement learning to enhance temporal alignment and timestamping precision in LALMs

Research Assistant Jan 2024 – Jun 2025

LEAP Lab, Indian Institute of Science (IISc)

- Proposed FESTA, a black-box unsupervised sampling method for uncertainty estimation in multimodal LLMs via consistency and sensitivity measures. Achieved 30% AUROC improvement on visual and audio reasoning benchmarks
- Co-organized DISPLACE-2024, a special session on multilingual, multispeaker speech tasks at INTERSPEECH 2024. Developed baseline Automatic Speech Recognition models and an evaluation portal for large-scale challenge submissions.

Intern May 2023 – Aug 2023

Adobe India

- Built and integrated B2B e-commerce features (product request and Q&A modules) for Adobe Commerce, increasing customer engagement and reducing support needed for product queries

SELECTED PROJECTS

Evaluation of Automated Poster Generation Sept 2025 – Dec 2025

- Created a 15-metric evaluation rubric and 93-poster benchmark to assess VLM-as-a-Judge reliability and model-specific biases for automated poster generation, validated across multiple VLMs and human participants. Results showed that InternVL-3 achieved highest human alignment.

Chain-of-Thought in vision-language models (VLMs) Jan 2025 – May 2025

- Benchmarked various prompting strategies on spatial reasoning datasets. Designed a two-stage question decomposition pipeline for VLMs, achieving an 8% relative accuracy gain over a Chain-of-Thought baseline.

TECHNICAL SKILLS

Libraries	PyTorch, pandas, numpy, scikit-learn, HuggingFace, NLTK, Keras
Languages	Python, C, Java, JavaScript, HTML/CSS
Machine Learning	pre-training, efficiency, fine-tuning, post-training, reinforcement learning, interpretability