

SUMMARY

I am a PhD candidate with 5+ years of research experience in **video understanding, multi-modal training and foundation models, incl. LLMs and VLMs**. I have extensive experience in developing Machine Learning and AI models for diverse modalities such as video, text, audio, and wearable sensor data.

EDUCATION

- **Georgia Institute of Technology** Atlanta, GA
PhD in Electrical and Computer Engineering Expected Summer 2025
Advisor: Dr. Irfan Essa Co-advisor: Dr. Justin Romberg.
- **Georgia Institute of Technology** Atlanta, GA
MSc in Electrical and Computer Engineering (Specializing in Machine Learning) Expected Summer 2025
- **PES Institute of Technology** Bangalore, India
Bachelor of Engineering in Electrical and Electronics Engineering May 2015

SELECTED PUBLICATIONS

1. **Apoorva Beedu**, Zhikang Dong, Jason Sheinkopf, and Irfan Essa. Mamba fusion: Learning actions through questioning. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2025
2. **Apoorva Beedu**, Harish Haresamudram, and Irfan Essa. Text descriptions of actions and objects improve action anticipation. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2025
3. Harish Haresamudram, **Apoorva Beedu**, Mashfiqui Rabbi, Sankalita Saha, Irfan Essa, and Thomas Ploetz. Limitations in employing natural language supervision for sensor-based human activity recognition—and ways to overcome them. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025
4. Karan Samel, **Apoorva Beedu**, Nitish Sontakke, and Irfan Essa. Exploring efficient foundational multi-modal models for video summarization. *arXiv preprint arXiv:2410.07405*, 2024
5. Hyeongju Choi, **Apoorva Beedu**, and Irfan Essa. Multimodal contrastive learning with hard negative sampling for human activity recognition. *ICCV 2023 workshop on PerDream: PERception, Decision making and REAsoning through Multimodal foundational modeling*, 2023

PROFESSIONAL EXPERIENCE

- **Facebook Reality Lab** May 2021 - August 2021
Research Intern Atlanta(remote), USA
Host: Dr. Chengde Wan Dr. Robert Wang
 - Developed a model to track a pen, and estimate 6D pose of the pen for Hand-Pen interaction.
- **Microsoft Research** May 2020 - August 2020
Research Intern Atlanta(remote), USA
Host: Dr. Amol Ambardekar Dr. Harpreet Sawhney
 - Developed a model to estimate and refine 6D object poses for large day-to-day objects.
- **NodeIn Robotics** May 2018 - August 2018
Robotics Intern Connecticut, USA
Host: Dr. Suresh Kannan
 - Worked on creating a map for indoor environment.
 - Developed a method to enhance images, and identify cases when feature extractions fail

PROFESSIONAL ACTIVITIES

- Outstanding reviewer for BMVC'24.
- Reviewer for BMVC(2021-24), PerDream2023, VTTA2022

TEACHING EXPERIENCE

- **Graduate Teaching Assistant**

Course: OMSCS: 6476 Computer Vision

August 2017 - Present

Atlanta, USA

PROJECTS

- (2024 - Ongoing) Video summarization using multi-modalities.
- (2024 - Ongoing) unifying datasets for action recognition using LLMs.
- (2017) MBZIRC Competition: challenges involved detecting the right stem valve size, detecting corresponding wrench and rotating the stem for 360°.
- (2017) Traffic Sign Classification using HOG and SVM.

SKILLS

- **Languages:** Python, C++
- **Libraries/Packages:** Numpy, scikit-learn, Scipy, Pandas, OpenCV, Jupyter, Matplotlib.
- **Machine Learning Libraries:** PyTorch
- **Keywords:** Computer Vision, Machine Learning, LLMs, Multi-Modality, VLMs, Video Analysis, Foundation Models, Vision-Language.

MENTORING

- Kara Bethany Liu
- Zhikang Dong, Jason Sheinkopf - Work led to the submission *Mamba fusion: Learning actions through questioning* (ICASSP '25)
- Hyeongju John Choi - Work led to a submission *Multimodal Contrastive Learning with Hard Negative Sampling for Human Activity Recognition* (PerDream@ICCV2023)
- Hrishikesh Kale