

Manushree Vasu

+1 (404) 451-9763 ✦ manushree635@gmail.com

[LinkedIn](#) ✦ [GitHub](#) ✦ [Google Scholar](#)

RESEARCH INTERESTS

Representation Learning, Generative Models, Multimodal Learning, Vision Language Models, Interpretability

EDUCATION

Boston University

Ph.D. in Computer Science

Boston, MA, USA

2025 – Present

- Advisor: [Prof. Deepti Ghadiyaram](#)

Georgia Institute of Technology

M.S. in Computer Science

Atlanta, GA, USA

2023 – 2025

- Thesis: *Extending Multimodal Large Language Models Beyond Single Modality* [T1]
- Advisors: [Prof. Humphrey Shi](#), [Prof. Judy Hoffman](#)

Manipal Institute of Technology

B.Tech in Computer Science and Engineering

Manipal, India

2019 – 2023

RESEARCH EXPERIENCE

Boston University

Sep 2025 – Present

Graduate Research Assistant | Advisor: [Prof. Deepti Ghadiyaram](#)

- Developed a cross-resolution distillation framework that enhances fine-grained image reconstruction in variational autoencoders by leveraging information from high-resolution teacher latents to low-resolution student latents. [W1]
- Designed a frequency-aware supervision objective that complements conventional pixel and perceptual losses by explicitly preserving high-frequency image details, enabling improved reconstruction quality.

SHI Labs, Georgia Institute of Technology

Aug 2023 – May 2025

Graduate Student Researcher | Advisor: [Prof. Humphrey Shi](#)

- Improved the smoothness of diffusion model latent spaces, yielding gains on editing and interpolation tasks. [C2]
- Built generalist MLLMs spanning image, audio, and video, and characterized cross-modal interaction and interference. [T1]

Adobe Research, India

Jun 2024 – Aug 2024

Research Intern | Advisor: [Balaji Krishnamurthy](#)

- Evaluated alignment between user intent and LLM output under complex, multi-requirement prompts, and built a tool that attributes misalignment to specific prompt components. [C1]

Neural Garage, Bangalore, India

Feb 2023 – Jun 2023

Machine Learning Intern | [Report](#) | [Recommendation](#)

- Developed the voice-cloning module for [VisualDub](#), applying text-to-speech synthesis across six languages, and improved existing lip-sync architectures by synthesizing neutral-expression videos.

IBM Research, Bangalore, India

Jun 2022 – Sep 2022

Summer Research Intern | Advisor: [Dr. Diptikalyan Saha](#) | [Report](#)

- Worked with the Trusted AI team on interpretability repair of ML models, porting explanations from a target model onto an already-deployed model according to user needs.

PUBLICATIONS

* denotes equal contribution

Thesis

T1. Extending Multimodal Large Language Models Beyond Single Modality

M.S. Thesis, Georgia Tech | [Thesis](#) | Advisors: [Prof. Humphrey Shi](#), [Prof. Judy Hoffman](#)

- Designed a lightweight, jointly-trained 7B multimodal LLM that outperforms larger specialist and generalist models on vision and audio benchmarks by reducing cross-modal interference through a simplified architecture and curated data.

Conference Papers

C1. CoPrompter: User-Centric Evaluation of LLM Instruction Alignment for Improved Prompt Engineering

IUI 2025 | [Paper](#)

- **Manushree Vasu***, et al.
- Automated the detection of misalignment in complex task instructions, enabling fine-grained assessment of the instruction-following ability of LLMs across individual prompt components.

C2. Smooth Diffusion: Crafting Smooth Latent Spaces in Diffusion Models

Poster

CVPR 2024 | [Paper](#) | [GitHub](#) | [Website](#)

- Jiayi Guo, Xingqian Xu, Yifan Pu, Zanlin Ni, Chaofei Wang, **Manushree Vasu**, Shiji Song, Gao Huang, Humphrey Shi
- Exposed the non-smoothness of diffusion latent spaces and introduced a class of diffusion models that are both smooth and high-performing at multiple downstream tasks such as image editing, interpolation, etc.

Workshop Papers

W1. SHARP-VAE: Steerable High-Frequency-Aware Reconstruction Preservation for VAEs

WiML Workshop, ICML 2026 | Advisors: [Prof. Deepti Ghadiyaram](#), [Prof. Vicky Kalogeiton](#)

- Proposed a cross-resolution distillation framework that transfers high-frequency information from high-resolution latent representations to low-resolution VAEs, improving fine-grained detail reconstruction.

W2. Efficient Gender Debiasing of Pre-trained Indic Language Models

Proceedings

Workshop on Deployable AI at AAAI 2023 | [Paper](#) | [Code](#)

- Neeraja Kirtane*, **Manushree Vasu***, Aditya Kane*
- Quantified association bias between profession and gender in Indic-language models, and mitigated it by efficiently fine-tuning on a novel dataset.

W3. Continual Visual Question Answering for Disaster Response Systems

Poster

Tackling Climate Change with Machine Learning at NeurIPS 2022 | [Paper](#) | [Code](#)

- Aditya Kane*, **Manushree Vasu***, Sahil Khose*
- Developed a continual-learning setup for visual question answering in disaster response using experience replay, and analyzed zero-shot CLIP performance against supervised training on pre-trained CLIP embeddings.

W4. Extraction of Color Information from Images for Generation of Colored Outlines and Sketches

Oral

MLCD, DGM, CtrlGen, and New in ML at NeurIPS 2021 | [Paper](#) | [Code](#)

- **Manushree Vasu***, Sameer Saxena*, Parna Chowdhury*, M Varma*, Harsh Rathod*, Ankita Ghosh, Sahil Khose
- Developed a modified conditional GAN for colored-sketch generation and an unsupervised k-means approach for rendering colored outlines from images.

W5. A Studios Approach to Semi-Supervised Learning

Poster

ICBINB at NeurIPS 2021 | [Paper](#) | [Code](#)

- Sahil Khose*, Shruti Jain*, **Manushree Vasu***
- Ran an ablation study of distillation for real-time deployment in a semi-supervised setting on CIFAR-10, comparing EfficientNet-b5, ResNet18, and MobileNet-V3-Large across four label splits to quantify the benefit of model compression.

TEACHING

Computer Vision (CS 6476 / 4476), Georgia Institute of Technology

Fall 2023 – Spring 2025

Head Teaching Assistant | Instructor: Prof. Humphrey Shi

4 semesters

- Led the teaching staff across four semesters: coordinated grading, designed and revised assignments, and held weekly office hours.

SELECTED PROJECTS

Assessing the Visual Capacity of Vision-Language Models

[Report](#)

Advanced Computer Vision | Instructor: [Prof. James Hays](#)

Spring 2024

- Analyzed the visual components of MLLMs, focusing on the alignment of visual features to the language model and measuring robustness to visual perturbations such as augmentations and pixel attacks.

Machine Unlearning with Selective Task Arithmetic and LoRA

[Video](#)

Machine Learning with Limited Supervision | Instructor: [Prof. Judy Hoffman](#)

Fall 2023

- Developed a method using the Fisher information matrix to selectively update task vectors of pre-trained models, retaining overall performance while forgetting targeted data, and adapted LoRA for efficient forgetting of small sample sets in both text and vision.

Out-of-Distribution Detection using XAI Techniques for ViTs

[Report](#)

Explainable AI | Instructor: [Prof. Sonia Chernova](#)

Fall 2023

- Developed a weighted locality-score method using attention flow and gradient attention maps to detect out-of-distribution samples for Vision Transformers.