

# Amartya Dutta

Blacksburg, VA, USA 24060

Email: [amartya@vt.edu](mailto:amartya@vt.edu)

[Personal Website](#)

[LinkedIn](#)

[Google Scholar](#)

[GitHub](#)

## RESEARCH INTERESTS

Computer Vision, LLMs & Multimodal Models, Bioinformatics

## EDUCATION

### Virginia Tech, Blacksburg, USA

PhD, Computer Science

March 2025 - Present

Master of Science, Computer Science; GPA: 3.9/4

Aug 2022 - Feb 2025

### Indian Institute of Information Technology (IIIT) Guwahati, India

Bachelor of Technology, Computer Science and Engineering; GPA: 8.6/10

July 2017 - May 2021

## WORK EXPERIENCE

### Graduate Research Assistant, COMPASS Centre

May 2025 – Present

- Working on building foundation models for drug repurposing.

### Graduate Research Assistant, Virginia Tech

Aug 2022 – Dec 2024

- Developed a **novel zero-shot approach** for **Scene Graph Relationship Prediction** using **VLMs**, reframing Predicate Classification as an **MCQA task** and surpassing trained baselines by at least 7% for **balanced relationship prediction**. Designed an **open-ended relationship generation framework** to eliminate answer-choice biases and improve contextual understanding.
- Performed comparative analysis of **Weakly Supervised Semantic Segmentation** approaches, highlighting the superiority of saliency maps over CAMs and introducing stochastic aggregation to enhance saliency effectiveness.

### Augmented Reality Developer Intern, Amply

Dec 2019 - Mar 2020

- Developed interactive Augmented Reality (AR) portals using AR-Core in Unity3D, enabling secure interactions with virtual objects to create immersive AR tours for client companies

### Virtual Reality Developer Intern, IIT Guwahati

May 2019 - Jul 2019

- Designed and developed an interactive Virtual Reality (VR) tour using Unity3D, focusing on smooth navigation within the virtual environment to enhance user experience (UX) in VR.

## PUBLICATIONS

[P1] **Amartya Dutta**, Kazi Sajeed Mehrab, Medha Sawhney, Abhilash Neog, Anuj Karpatne et al. “Open World Scene Graph Generation using Vision Language Models”. [CVPR 2025 Workshop](#), [ICML 2025 Workshop](#) **Paper**

[P2] Abhilash Neog, Medha Sawhney, Kazi Sajeed Mehrab, Sepideh Fatemi, Mary E. Lofton, **Amartya Dutta**, Anuj Karpatne et al. “Toward Scientific Foundation Models for Aquatic Ecosystems”. [ICML 2025 Workshop](#)

[P3] Medha Sawhney, Abhilash Neog, Mridul Khurana, **Amartya Dutta**, Arka Daw, Anuj Karpatne. “Physics-guided Diffusion Neural Operators for Solving Forward and Inverse PDEs”. [CVPR 2025 Workshop Poster](#)

[P4] Sepideh Fatemi, Abhilash Neog, Emma Marchisin, **Amartya Dutta**, Anuj Karpatne et al. “Scientific Equation Discovery using Modular Symbolic Regression via Vision-Language Guidance”. [CVPR 2025 Workshop Poster](#)

[P5] M. Maruf, Arka Daw, **Amartya Dutta**, Jie Bu, Anuj Karpatne. “Beyond Discriminative Regions: Saliency Maps as Alternatives to CAMs.” **Paper**

[P6] **Amartya Dutta**, Rajat Kanti Bhattacharjee, Ferdous Ahmed Barbhuiya. “Efficient Detection of Lesions During Endoscopy.” [ICPR Workshop 2021](#) **Paper**

[P7] **Amartya Dutta**, Kamaljiyoti Nath. “Learning via LSTM for Railway Bridge strains.” [ICDSMLA 2020](#) **Paper**

[P8] **Amartya Dutta**, Ferdous Ahmed Barbhuiya “Predicting Popularity of Images Over 30 Days.” **Paper**

## PROJECTS

**SEAL-0 (Search Engine Augmented Language) (Ongoing):** Creating a dataset of complex questions that challenge state-of-the-art LLMs. This project evaluates how these models handle queries requiring up-to-date knowledge and complex reasoning by leveraging search engines for real-time information.

**Evaluating Model Reasoning and Hallucinations in Medical LLMs:** This project investigates factual error propagation in open-source medical LLMs (e.g., BioMistral, Asclepius) and documents their datasets for transparency. By highlighting performance variations, it aims to guide the development of safer, more reliable language models for healthcare. [GitHub](#)

**Visualizing the Spotify Soundscape:** This project visualizes the Spotify Top 50 Tracks of 2023 through an interactive, HTML-based dashboard. Using D3.js and Plotly.js, it enables dynamic, data-driven exploration of each track’s popularity and attributes. [GitHub](#)

**Predicting Popularity of Flickr Images (ICIP 2021):** This project predicts how popular a Flickr image will be over 30 days, even before it's uploaded. By analyzing user and image social features alongside image visuals, it models engagement based on two factors: scale and shape. Using these factors, the method forecasts the daily engagement sequence. See publication [P8]. [GitHub](#)

## TEACHING EXPERIENCE

---

**CS 2064: Intermediate Programming in Python:** Graduate Teaching Assistant, Virginia Tech.

**CS 5644: Machine Learning with Big Data:** : Graduate Teaching Assistant, Virginia Tech.

**CS 5805: Machine Learning:** : Graduate Teaching Assistant, Virginia Tech.

## ACADEMIC SERVICES & AWARDS

---

**Reviewer:** AAAI-24-Imageomics-Workshop

**Reviewer:** CVPR 2025 Workshop, Computer Vision in the Wild (CVinW)

**Program Committee:** CVPR 2025 Workshop, CVinW

**3rd Position, IEEE ICIP Image Popularity Prediction Challenge (Oct 2020):** [View Results](#)

## SKILLS SUMMARY

---

**Languages:** Python, L<sup>A</sup>T<sub>E</sub>X, C++, Shell

**Frameworks/Tools:** PyTorch, TensorFlow, Keras, Scikit, OpenCV, Numpy, Pandas, Git, Unity3D, vLLM, LangChain, FAISS, OpenAI API, Hugging Face Transformers & APIs