

# Sidhartha Reddy Potu

+1 (917) 914-1526 | sp7835@nyu.edu | GitHub | LinkedIn | Portfolio

## EDUCATION

### New York University, Tandon School of Engineering

*Master of Science in Computer Science; GPA: 3.83/4*

New York, NY

Sep 2023 – May 2025

### Indian Institute of Technology (IIT), Indore

Indore, IN

*Bachelor of Technology in Electrical Engineering, Minor in Astronomy; GPA: 8.49/10*

Jul 2019 – May 2023

## EXPERIENCE

### Graduate Student Researcher, CILVR, NYU

New York, NY

*3D Computer Vision under Prof. David Fouhey*

Jan 2024 – Present

- Incorporated Colmap SfM and HaMeR transformer to get global 3D trajectories for EgoExo4D and Epic-Kitchens.
- Implemented temporal smoothing, minimized velocity and jerk for stable hand motion in egocentric videos.
- Developed visualization workflows with Rerun Viewer and scaled experiments on NYU HPC using SLURM.

### Course Assistant, NYU

New York, NY

*Computer Vision, ML courses*

Sep 2024 – May 2025

- Advised 50+ students in Learning with Large Language and Vision Models (Courant) under Prof. Saining Xie.
- Mentored 400+ students in Deep Learning (Tandon) under Prof. Chinmay Hegde.
- Managed 50+ students in Advanced Topics in Computer Vision (Tandon) under Prof. David Fouhey.

### B.Tech Thesis, PRIA Lab, IIT Indore

Indore, IN

*Multimodal Crowd Counting using Vision Transformers [\[Report\]](#)*

Jun 2022 – Dec 2022

- Proposed a novel vision transformer-based approach for crowd counting after an in-depth review of 30+ papers.
- Trained a ViT based model in weakly supervised setting, achieving MAE 15.78 and MSE 30.70.
- Created a Pyramid Vision Transformer (PVT v2) model, achieving MAE 19.218 and MSE 32.940.
- Ranked top 10 of 250+ B.Tech theses; nominated for Best Thesis award at IIT Indore..

### Research Assistant, NLP Lab, IIT Indore

Indore, IN

*Multi-Modal Hashtag Recommendation*

Oct 2021 – Apr 2022

- Modeled hashtag generation for micro-videos using LSTMs with self-, co-, and hierarchical attention.
- Enhanced feature fusion with BERT, Word2Vec, ResNet152, InceptionV3, VGG16, MFCCs, and spectrograms.

## PROJECTS

### Point Tracking in Egocentric Videos

Feb 2025 – May 2025

- Adapted DINOv2 and CroCoV2 backbones for point tracking in egocentric videos.
- Evaluated on the EgoPoints benchmark, reporting tracking accuracy at 8 pix, 16 pix, and 24 pix thresholds.

### Fine-Tuning Medical QA Models

Oct 2024 – Dec 2024

- Optimized Meta LLaMA 3.2 3B for medical QA using LoRA, QLoRA, and GaLore.
- Achieved best USMLE accuracies: Step 1 - 42% (GaLore), Step 2 - 38% (QLoRA), Step 3 - 45% (GaLore).
- Improved MedQuAD F1 scores to 70% (QLoRA) and 69% (GaLore) with efficient fine-tuning.

### VioletPass:Ticket Booking Platform

Oct 2024 – Dec 2024

- Designed a robust, scalable ticket booking platform on AWS to prevent double booking.
- Implemented Redis distributed locks with TTL for seat reservations and PostgreSQL for finalization.
- Integrated QR code-based authentication for secure and efficient event check-ins.

### AnimeVerse - Recommendation

Apr 2024 – May 2024

- Developed a personalized anime recommendation system, analyzing broadcast types, genres, and anime scores.
- Analyzed 300k user-anime interactions with PySpark and trained ALS, TF-IDF, and ChromaDB models.

## TECHNICAL SKILLS

**Programming:** Python, C/C++, MATLAB, Javascript, HTML, CSS, SQL, Linux, Pyspark, Hadoop

**ML:** TensorFlow, PyTorch, Keras, Pandas, Numpy, Matplotlib, OpenCV, Scikit-learn, Jupyter Notebook

**Development:** MERN Stack, Firebase, Flask, Git, Flutter, GCP, AWS, ReactJS, Django, Redux, Bootstrap, SQL