

# VAISHNAV POTLAPALLI

☎ 332-373-5002

✉ [vp2499@nyu.edu](mailto:vp2499@nyu.edu)

🌐 [linkedin.com/in/vaishnav-potlapalli](https://www.linkedin.com/in/vaishnav-potlapalli)

🐙 [github.com/vaishn9v](https://github.com/vaishn9v)

## Education

New York University Courant Institute of Mathematical Sciences

Sep. 2023 – Dec 2024

*Masters of Science in Computing, Entrepreneurship and Innovation GPA: 3.96/4.0*

New York City, NY

**Relevant Courses:** LLVMs, Big Data and ML Systems, Foundations of Computer Networks

**Honors:** M. Michael Waller Master's Fellowship

## Publications

**PromptIR: Prompting for All-in-One Blind Image Restoration**

NeurIPS 2023

*Vaishnav Potlapalli, Syed Waqas Zamir, Salman Khan, Fahad Shahbaz Khan*

- Proposed an implicit prompt-learning based approach for All-in-One blind Image Restoration. Achieves SoTA performance on multiple image restoration tasks, without any prior degradation information.

**Sketch3T: Test-Time Training for Zero-Shot SBIR**

CVPR 2022

*Aneeshan Sain, Ayan Kumar Bhunia, Vaishnav Potlapalli, Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song*

- Introduced a novel test-time training paradigm for zero-shot sketch-based image retrieval that adapts to new categories and sketch distributions using a single sketch, outperforming state-of-the-art methods.

**MediTables IIIT**

GREC 2021

*Akshay Praveen Deshpande, Vaishnav Potlapalli, Ravi Kiran Sarvadevabhatla*

- Built a new dataset and semantic segmentation model for camera captured medical document images.

## Experience

Floma Inc

January 2025 – Present

*Software Engineer - Machine Learning*

- Developed a computer vision-based tool using an ensemble of **Segmentation** and **Object-detection** models to detect various visual elements in display ads, enabling automatic asset identification and labelling.
- Worked on a **Multimodal LLM-based AI agent** that generates and renders dynamic display ads from text copy and visual assets. Developed an **SVG creation and editing tool** used by the agent to construct display ads in various sizes and formats, enabling seamless customization and editing.

MBZ University of Artificial Intelligence

July 2022 – July 2023

*Research Assistant - Computer Vision (Advisor: Dr. Salman Khan)*

- Proposed and implemented a novel Visual transformer based prompt-learning framework for All-in-one blind Image Restoration / Enhancement called **PromptIR**, which achieved **SoTA** performance improving over previous methods by **0.9 dB** on dehazing, deraining and denoising benchmarks. Work presented as part of **Neurips 2023**
- Adapted computer vision based continual learning techniques **L2P**, **DualPrompt** methods for video action recognition improving performance over previous techniques by over **10%** accuracy and **14% BWF**, on several public benchmarks.
- Studied parameter-efficient finetuning techniques to improve downstream performance of **Multimodal LLM models**.

Dhan AI

April 2020 – April 2022

*Machine learning Engineer*

- Developed an ensemble of **BERT-based Classifiers** to enhance the NER engine, resulting in a **12% accuracy** improvement in internal benchmarks on entity recognition and sentiment classification, significantly improving the primary product of the company, which was a Patient Life Cycle Management Chatbot.
- Rewrote the application testing pipeline to utilize increased **parallelism and Redis cache** to reduce CI/CD time by **60%** that enabled rapid development of new features.

## Projects

**Efficient Mixture-of-Depths(MoD) LLM Inference** | *PyTorch, CUDA, LLMs*

March 2024 – May 2024

- Engineered a Mixture-of-Depths (MoD) transformer on a LLaMa-style baseline (55M parameters, 6 layers) by integrating dynamic token routing with top-k selection, auxiliary loss, and an auxiliary MLP predictor
- Utilized Torch CUDA events for profiling and simulated a novel GPU scheduling policy to boost throughput, lower latency, and validate improvements via ablation studies and perplexity analysis.

**PITCHPAL: AI-Powered Presentation Coach** | *FastAPI, React, LLMs, AI Agents*

October 2024 – December 2024

- Implemented a multi-modal AI Agent combining advanced speech recognition, natural language processing, and computer vision to analyze presentation content and delivery.
- Developed a FastAPI backend for presentation analysis, enabling seamless integration of NLP and CV models, with the AI Agent to evaluate slide content, assess speech quality, and generate actionable feedback.

## Technical Skills

**Computer Vision:** OpenCV, Slam, Object Detection, Semantic Segmentation, Image Restoration, GAN, VAE, CNN, ViT

**Languages/Frameworks:** Python, Java, C/C++, CUDA, TypeScript, Pytorch, TF, Transformers, Django, MySQL