Jaiminkumar Ashokbhai Bhoi

+13864568222 ajaymin28@gmail.com Oviedo, FL, USA - 32765
in/ajaymin28 github.com/ajaymin28 𝕏 ajaymin28.github.io

Professional Experience

Graduate Research Assistant *Center for Research in Computer Vision(CRCV)* **FL, USA** *05/2024 - 01/2025*

- **Proposed a novel solution** for Dynamic Scene Graph Generation (DSGG) with MLLMs, demonstrating a **10-40% performance improvement using just 5-10% of training data** across varying top-K metrics, while maintaining the recall-precision balance.
- Continuously evaluated and experimented with SOTA models and methods for generating DSGG with next token prediction.
- Efficiently **finetuned SOTA MLLMs**(Video-LLAVA, LLaVa-OneVision, InternVL2) with Flash Attention using High Performance Computing(HPC) on Action Gnome and VidVRD datasets.
- **Benchmarked** and analyzed model's performance **demonstrating reduced predicate perplexity** after finetuning as well as maintaining performance for long-trail predicates.
- **Created** a novel dataset of **2D/3D shapes and real images** to evaluate reasoning limits in **MLLMs/VLMs** (LLaVA-One-Vision, InternVL3, Qwen2-VL), revealing critical **biases and failure cases** in geometric and in-context understanding of SOTA VLMs.
- **Generated 14k synthetic images** and **50k VQA queries** for robust, scalable evaluation of multimodal models.
- **Collected 200 real images** to support **shape/object matching, odd-one-out, and rotation reasoning** evaluation tasks.

Individual Contributor *University of Central Florida* **FL, USA** *12/2023 - 04/2024*

- **Built** an automated system for tibia/femur angle measurement using segmentation with **Segment Anything Model and a light weight classifier**, reducing manual review time of 3D scans(DICOMS) of the patient by **35%**

Software Engineer / Research Associate *Tata Consultancy Services* **Bangalore, India** *06/2018 - 09/2023*

- Designed and implemented a **video analytics solution** to **prevent losses in retail self-checkout environments**, addressing an industry-wide annual loss of **$90B**.
- **Executed and Delivered** three projects to production: **IVI system**, **Computer Vision on QC RB500 board**, and **Container Image Analytics** that directly impacted human lives and **saved billions in USD**.
- Developed and deployed **Computer Vision algorithms** that **saved $4M** in container repair and cleaning costs, and **reduced lead time from 12 to 1 day** for 10% of repair volume while ensuring high accuracy and performance.
- Fine-tuned deep learning models on **production image datasets** using **TensorFlow**, achieving **over 90% accuracy** in defect and quality inspection for **image classification, object detection, and segmentation** tasks.
- Built a **Continuous Learning Framework (CLF)** with customized **AzureML Ops**, reducing retraining efforts by **80%** and accelerating iteration cycles with **human-in-the-loop feedback** by developing custom annotation tools.
- Deployed scalable multi-model APIs with **Flask/RestX** and **Docker** on **Azure Kubernetes**, leveraging auto-scaling to efficiently process **10k+ high-quality images per hour** through optimized **ONNX** hierarchical chained inference.
- Deployed 4 Computer Vision solutions Dlib **face recogntion**, **PosNet**-based theft detection, **YOLO v3** ticket-switch detection, and queue counting in **C++** on the **Qualcomm RB5 board**, demonstrating edge AI capabilities.
- **Compiled and optimized OpenCV and Dlib** with **OpenBLAS** for **ARM**, achieving a **4x performance boost** on edge hardware.
- **Quantized** models to **TFLite** for efficient on-device inference, **minimizing model size** with negligible accuracy loss.
- Developed and **optimized** real-time **computer vision pipelines** for deployment on edge.
- **Leveraged Android delegates** (**DSP**, **CPU**, **GPU**, **NNAPI**) to execute multiple Computer Vision solutions concurrently, **increasing throughput** of edge devices.

Education

Master's in Computer Vision *(University of Central Florida)* **FL, USA** *08/2023 - 04/2025*
Bachelor's in Computer Engineering *(A. D. Patel Institute of Technology)* **GJ, India** *04/2014 - 03/2018*

Skills

Python ,C++ **PyTorch**, TensorFlow, **OpenCV**, **Computer Vision**, **Image recognition**, Pandas, Classification, Detection, Segmentation, NumPy, **Machine Learning**, **NLP**, Deep Learning, Transformers, Quantization, **LLM**, Software Engineering, Deployment, **MLOps**, **Slurm**, **Docker**, **AzureML**, feature extraction, Edge deployment, **REST APIs**, Git, SQL, Web Development.

Patents & Research

- What can Off-the-Shelves Large Multi-Modal Models do for Dynamic Scene Graph Generation? (ongoing)
- Method and system to detect a text from multimedia content captured at a scene. (**US-12333832-B2**)
- An efficient ensemble-based deep learning model for the diagnosis of cervical cancer (ISCAIE-22)

Projects

Can Visual Language Models understand shapes? *02/2025 - current*
Video Understanding (Using Neuro Symbolic AI) *02/2024 - current*
Human Activity Recognition on Static Images (HAR) *08/2023 - 12/2023*
Container Image Analytics (CIA) *02/2021 - 06/2023*
Computer Vision on Qualcomm RB5 Development Board *06/2020 - 02/2021*