# Jaiminkumar Ashokbhai Bhoi

📕 +13864568222 🛛 ajaymin28@gmail.com 🕋 Oviedo, FL, USA - 32765 in/ajaymin28 🖓 github.com/ajaymin28  ${\mathbb X}$  ajaymin28.github.io

## Education

Master's in Computer Vision (University of Central Florida) **Bachelor's in Computer Engineering** (A. D. Patel Institute of Technology)

## **Professional Experience**

**Graduate Research Assistant** Center for Research in Computer Vision(CRCV) FL. USA 05/2024 - 01/2025 • Designed and implemented an MLLM-based method for dynamic video scene graph generation, improving performance by

- 10-40% for different top-Ks and achieving state-of-the-art performance on benchmark datasets. 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 - 06/2023

- 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.
- **Delivered** 60+ sprints, aligning feature delivery with the product roadmap through close **collaboration with Product Manager**, Scrum Masters, and Subject Matter Experts.
- Mentored and led two junior developers to successfully deliver a high-impact production project that resulted in significant cost savings in computer vision-based quality assurance and maintenance use cases.

## Skills

Python ,C++ PyTorch, TensorFlow, OpenCV, Computer Vision, Video Analytics, Classification, Detection, Segmentation, Tracking, Machine Learning, Deep Learning, Transformers, Quantization, LLMs, Prompt Engineering, Software Engineering, Deployment, MLOps Docker, Kubernetes, AzureML, NVIDIA Jetson, Edge deployment, REST APIs, Git, SQL, Web Development

## **Projects & Research**

#### Video Understanding (Using neurosymbolic AI approach)

- 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.
- 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.

#### **DumbVLMs** (Visual Language Models)

- 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.

## Container Image Analytics (CIA)

- 02/2021 06/2023 • 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.
- Fine-tuned deep learning models on production image datasets using TensorFlow, achieving over 90% accuracy in defect and quality inspection for classification, 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.
- 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.

• Distilled and quantized for **ARM processors**, and developed a cross-platform MVP in **Flutter** for edge deployment.

- **Computer Vision on Qualcomm RB5 Development Board** 06/2020 - 02/2021 • Designed and implemented a video analytics solution to prevent losses in retail self-checkout environments, addressing an industry-wide annual loss of \$90B.
- Deployed 4 Computer Vision solutions Dlib face detection, PosNet-based theft detection, YOLOv3 ticket-switch detection, and queue counting in C++ on the Qualcomm RB5 board, advancing 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.
- Implemented JNI bridges between C++ and Java for Android integration, enabling seamless native-to-Java communication.
- Leveraged Android delegates (DSP, CPU, GPU, NNAPI) to execute multiple Computer Vision solutions concurrently, increasing throughput of edge devices.

#### 02/2024 - current

02/2025 - 05/2025

FL, USA 08/2023 - 04/2025 GJ, India 04/2014 - 03/2018