Jaiminkumar Ashokbhai Bhoi

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

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*

Professional Experience

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

- **Designed and implemented** an MLLM framework for video scene graph generation, improving accuracy by **12%** 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 **Segment Anything Model-based segmentation and optional classification heads**, reducing manual review time of 3D patient scans by **35%**

Software Engineer / Research Associate *Cognitive AI, 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**, demonstrating breadth in applied computer vision solutions
- **Delivered** 32+ sprints, aligning feature delivery with the product roadmap through close collaboration with **Product Manager**.
- **Automated** test workflows using **sim and real hardware** in a **multi-node test environment**, accelerating Dev and QA cycles.
- **Mentored and led** two junior developers to deploy **computer vision solutions**, reducing quality inspection costs by up to **60%**

Skills & Interests

- **Python**, **PyTorch**, **TensorFlow**, **OpenCV**, **Computer Vision**, **Deep Learning**, **Image Processing**, Multimodal Models, **Object Detection**, Segmentation, Tracking, Pruning, **Quantization**, Transformers, **Prompt Engineering**, **Large Language Models (LLMs)**, Docker, Kubernetes, Flask, **REST APIs**, Git, SQL, Android Studio, JavaScript, C++, Linux, ARM64, Edge Computing, **AzureML**, NVIDIA Jetson, Team Leadership, Agile Development, **Problem Solving**, Analytical Thinking

Projects & Research

Video Understanding (Using neurosymbolic AI approach) *02/2024 - current*

- **Engineered** a pipeline to generate **zero-shot scene graphs** using **prompt engineering and chain-of-thought reasoning** with **SOTA Video LMMs**, enabling explainable video understanding
- **Fine-tuned Video-LMMs** on **HPC** and integrated **LoRAs** to improve object-predicate accuracy in temporal scenes
- **Benchmarked** models on **Action Gnome** and **VidVRD** video datasets, demonstrating reduced predicate perplexity after fine-tuning, validating improved reasoning using the proposed method.

DumbVLMs *02/2025 - 05/2025*

- **Created** a benchmarking dataset of **2D/3D shapes and real images** to evaluate reasoning limits in **MLLMs** (LLaVA-One-Vision, InternVL3, Qwen2-VL, etc.), revealing **biases and failure cases** in geometric and in-context understanding
- **Generated 14k synthetic images** with **50k VQA queries**, providing robust and scalable evaluation of multimodal models.
- **Collected 200 real images** replicating similar settings for **shape matching, odd-one-out, and rotation reasoning** tasks.

Container Image Analytics (CIA) *02/2021 - 06/2023*

- **Fine-tuned** deep learning models on **production image datasets** using **TensorFlow**, achieving over **90% accuracy** and improving defect and quality inspection across **classification, detection, and segmentation** tasks
- **Built** a **Continuous Learning Framework (CLF)** using customized **AzureML Ops**, reducing retraining efforts by **80%** and accelerating iteration cycles with **human-in-the-loop feedback**
- **Deployed** scalable model APIs using **Flask/RestX** and **Docker** on **Azure Kubernetes**, auto-scaling to process **10k+ high-quality images per hour** to meet business needs.
- **Distilled and quantized** models for **ARM processors**, and developed a **cross-platform MVP** in **Flutter** for edge deployment.

Computer Vision on Qualcomm RB500 Development Board *06/2020 - 02/2021*

- **Deployed 4 computer vision solutions** (**Dlib face detection**, **PosNet-based theft detection**, **YoloV3 ticket-switch detection**, and queue counting) in **C++** on the **Qualcomm RB500 board**, enhancing CV capabilities at the edge.
- **Implemented JNI bridges** between **C++ and Java** for **Android integration**, enabling seamless native-to-Java communication.
- **Compiled and optimized OpenCV and Dlib** with **OpenBLAS** for **ARM**, resulting in **4x performance boost** on edge hardware.
- **Quantized** deep learning models to **TFLite** for **on-device inference**, minimizing model size with negligible accuracy loss.
- **Leveraged Android delegates** (**DSP**, **CPU**, **GPU**, **NNAPI**) to execute four Computer Vision solutions concurrently, increasing throughput of the edge devices.