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 (computer science) **FL, USA** *08/2023 - 05/2025*
University of Central Florida

Bachelor's in Computer Engineering **GJ, India** *04/2014 - 03/2018*
A. D. Patel Institute of Technology

Professional Experience

Research Assistant *Center for Research in Computer Vision(CRCV), University of Central Florida* **FL, USA** *05/2024 - present*
- Achieved SOTA performance in video scene graph generation by designing and implementing a multimodal LLM framework, yielding a 12% accuracy improvement.

Individual Contributor *College of Medicine, University of Central Florida* **FL, USA** *12/2023 - 04/2024*
- Developed an automated angle measurement system for tibia and femur alignment using segmentation (SAM) and classification, reducing manual inspection time for 3D patient data by 35%.

Software Engineer/Research Associate *Cognitive AI, Tata Consultancy Services* **Bangalore, India** *06/2018 - 06/2023*
- Designed, developed, and deployed multiple deep learning models for quality assurance usecases.
- Designed a Continuous Learning Framework (CLF) on top of Azure ML Services using MLOps, capable of automated training and deploying multiple models within the ecosystem.
- Led a developers team in a Scrum environment, collaborating with the product manager to deliver a container image analytics solution consisting of 10+ deep learning models.
- Designed and developed production-grade computer vision solutions on Azure Cloud and Edge, reduced cloud service costs by 60% by distilling models to a smaller size.

Skills & Interests

- **Programming & Tools:** Python, TensorFlow, PyTorch, OpenCV, Docker, Kubernetes, Flask, RESTful APIs, Git version control, SQL, Javascript,Android Studio, Software Engineering
- **AI/ML Techniques:** Computer Vision, Object Detection/Segmentation/Tracking, Image Processing, Multimodal models, Machine Learning, Deep Learning, Video Scene Graph Generation, Self-Supervised Learning(SSL), algorithm development, Natural language processing, neural networks, transformers, Large Language Models(LLM), Prompt Engineering, Chain of Thoughts, RAG
- **Platforms & Frameworks:** Cloud computing , AzureML, NVIDIA Jetson, ARM64, Android, Linux, Edge Computing
- **Soft Skills:** Team Leadership, Analytical, Agile Development,Problem solving, Cross-Functional Collaboration

Patents & Publications

- What can Off-the-Shelves Large Multi-Modal Models do for Dynamic Scene Graph Generation?
- An Efficient Ensemble-Based Deep Learning Model for the Diagnosis of Cervical Cancer: (ISCAI-2022)
- Aerial Video Analytics based dynamic Non-linear distance measurement between on-ground objects (Patent Filed)
- Method and system to detect a text from multimedia content captured at a scene (Patent Published, 2023)

Projects & Research

Video Understanding (using Neurosymbolic AI) *Github* *02/2024 - present*
- Generated zero-shot dynamic scene graph using prompt engineering.
- Finetuned Video LMMs using LORAs for customized and improved scene graph generation.
- Investigated perplexities for the predicted scene graph triplets.

DumbVLMs (Dataset benchmark for 2D/3D shapes understanding) *Github* *02/2024 - present*
- Constructed a synthetic dataset using pyrender for various shapes with various combinations to evaluate SOTA VLMs(LLaVA-one-vision, Qwen-VL-2, Paligemma, InternVL-2.5) capabilities.

Container Image Analytics (CIA) *02/2021 - 06/2023*
- Fine-tuned deep learning models including VGG16, MobileNetV2, and RCNN, achieving over 90% accuracy on production data.
- Designed and implemented a Continuous Learning Framework (CLF) using AzureML, significantly reducing manual training efforts by 80%.
- Developed REST APIs for deep learning models, leveraging AzureML, Docker, Flask/RestX, and Azure Kubernetes. These APIs efficiently handle 10,000+ requests per hour with auto-scaling capabilities, resulting in optimized cloud resource usage and a 60% reduction in cloud service costs.
- Engineered and delivered multiple user interfaces for data collection and validation purposes, utilizing C# Blazor, HTML/CSS, and jQuery-based web pages, reducing customers' efforts and providing AI-aided judgment.
- Quantized models for ARM-based processors and developed a Flutter app for edge inference, utilizing Method channels for cross-device code development for speed and efficiency of the models.

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

- Spearheaded the design and development of various computer vision solutions including face detection (utilizing Dlib), self-checkout theft detection (employing PosNet), barcode-switching detection (leveraging YoloV3), queue counting, and person tracking heatmap generation, all implemented in C++.
- Engineered Java Native Interfaces (JNI) to facilitate seamless communication between C++ and Java components for Android applications.
- Optimized performance on ARM64 devices by compiling custom Android libraries for OpenCV, Qualcomm Snappy, and Tensorflow, and integrating OpenBLAS for compiling DLIB, resulting in a notable performance enhancement of up to 4x.
- Successfully quantized deep learning models to enable their deployment on Android edge devices with minimal loss of accuracy.
- Demonstrated expertise in leveraging various Android delegates including DSP, CPU, GPU, and NNAPI to concurrently execute four computer vision solutions on edge devices, showcasing adept multi-threading capabilities.