SRINIVAS VENKATANARAYANAN

+1 386-456-8731 | San Diego, CA, USA | srinivasvishal7@gmail.com |
https://linkedin.com/in/srinivas-v-9234703a/ | Vi-Sri (Srinivas Venkatanarayanan) · GitHub

SUMMARY

Experienced ML Research Engineer and Graduate Research Student with 6+ years of professional experience in building and leading scalable ML, Computer Vision, and Voice systems on cloud and edge platforms. Proficient in applied research and MLOps, deploying distributed and real-time ML services, with significant embedded AI experience. Proven success in academic and industrial research labs, specializing in Computer Vision research and Camera Systems. Motivated problem-solver with a strong grasp of business constraints and objectives.

EDUCATION

University of Central Florida **August 2023 - May 2025**
Master of Science (MS), Computer Science (Specialization in Computer Vision)

Anna University **August 2013 - May 2017**
Bachelor of Engineering (BE), Electronics and Communication Engineering

Certifications: AWS Associate Solution Architect, AWS Certified Machine Learning - Specialty

SKILLS & INTERESTS

Skills: Python, C++, Rust, Golang, Nim, Computer Vision, Tensorflow, Pytorch, Keras, OpenCV, OpenMVS, ZeroMQ, CUDA, Data Analysis, Data Science, Data Structures, Algorithms, Django, Docker, Elasticsearch, JavaScript, Jupyter notebooks, Git, HTML/CSS, Java, Airflow, MLFlow, Kubeflow, Containerization, Linux/Unix, MySQL, OpenCL, OpenGL, Kubernetes, MATLAB, Unity, Go, NoSQL, FastAPI, GCP CE, AWS Sagemaker, AWS EC2, AWS Greengrass, AWS Lambda, AWS S3, AWS ECR, AWS EKS, AWS EBS, Node.js

Interests: Applied Research, AI/ML systems, Computer vision, NLP, MLOps, Machine learning, Real-time systems, Software Engineering, ML Platform engineering, Distributed systems

PROFESSIONAL EXPERIENCE

Qualcomm Technologies, San Diego, USA
ML Research Intern **May 2024 - Present**

- Researched on Closed Loop Neural Illumination Correction, Deep White Balance, and Neural Enhancement models for image and videos for sequentially modeling complex non-linear illuminations in arbitrary scenes.
- Successfully implemented closed loop Illumination aware neural optimizer system on Camera SoC for color correction.
- Researched on model optimization for sparsity to enable efficient deployment on low-power Qualcomm SoCs.

UCF Center for Research in Computer Vision Lab, Orlando, USA
Graduate Research Assistant **August 2023 - Present**

- Researched on Sample hardness based adaptive routers using aleatoric entropy heuristics for motion aware Activity foundation models generating movement descriptions.
- Improved the accuracy of the Cloth adversarial Re-ID Learning model by 5%.
- Spearheaded the implementation of state-of-the-art deep learning models for Video Person Re-Identification and achieved a significant reduction in false positives by 15% using disentangling appearance biases by using gait, pose, clothes and silhouettes.
- Researched for Novel View Synthesis and View transfer with combined LoRA on Dreambooth diffusion, able to achieve the goal of being able to transfer views like style using ZipLoRA.
- Researched on using Block motion estimation in compressed MPEG-4 domain to be leveraged for training task adaptive models for video classification.

Toyota Connected, Bangalore, India
Lead Machine learning Engineer - 1 **February 2020 - August 2023**

- Led the implementation team for building MLOps framework for domain and intent arbitration engine inspired from Alexa's Hyprank for Toyota's own In-Car Voice assistant launched in over 1000 head units of 2022 Toyota Hilux before full scale adoption.
- Successfully built out-of-distribution utterance detection using Dirichlet Prior Networks which reduced external service API calls by 70%.
- Established and led the computer vision team in Toyota Connected India, from inception to a high-performing unit of 30 members, driving breakthrough innovations and contributing to vision capabilities for Toyota's global connected and driver experience technologies.
- Spearheaded a team of 9 engineers in the design, development, and deployment of a vision based augmented navigation assistance system for Heads-Up Displays and tested on Lexus RX series.
- Implemented a real-time pipeline for synchronizing extrinsic properties of car camera with AR virtual camera and transforming to driver’s viewpoint using Head tracking for rendering 3D models in a given location in heads-up display.
- Initiated and headed the OSS contribution initiative, developed a low code image augmentation library used by 10+ engineers inside the organization.
- Led the team for building the Automated Mapping platform (AMP) project for HD maps in collaboration with Uber ATG, scaled the system to process and recognize live road assets in 200+ videos from Toyota’s fleet cameras for updating maps.
- Built vehicle inspection and ANPR based service automation system, installed in over 200 Toyota dealerships and 800 devices so far across India.

HCL Technologies, Chennai, India
Member Technical Staff **July 2019 - January 2020**

- Architected the entire AI stack, web services, and deployment pipelines for a Musculo-skeletal injury severity scoring engine to assist Radiologists to reduce average patient lead time by 240 minutes a week.
- Ideated and wrote tools for robust and scalable web architecture for model serving with Kubeflow and Seldon Core
- Optimized the throughput of models from 17 Requests per second to 46 requests per second.
- Led the initiative for building Background Task Worker tools with Celery and RabbitMQ Message broker to fetch X-Ray images from PACS (Picture Archiving and Communication System) Server and push to the optimized AI stack optimized AI stack which improved the processing time for TIFF image queues by 120%.

E-con Systems, Chennai, India
Computer Vision Engineer **December 2017 - April 2019**

- Implemented algorithms for unstructured package volume estimation using 3D point cloud reconstruction from two Stereo Cameras, filed a patent on the same.
- Implemented CUDA kernels for applying triangulation and stereo reprojection for stereo cameras.
- Created a GAN pipeline for augmented generations of out of domain images to improve our anomaly detection pipeline.
- Led the development of a repository of custom OpenCL kernels for building an internal ML inference engine for IMX8 Vivante GPUs.

Zoho Corporation, Chennai, India
Member Technical Staff **May 2017 - November 2017**

- Led the initiative for building a Software inventory Detection agent for detecting Redundant Softwares in client machines which reduced redundancies in software inventory by 50%.

Zoho Corporation, Chennai, India
Software engineering Intern **August 2016 - April 2017**

- Developed an API explorer and monitoring tool for ManageEngine Desktop Central to increase traceability and reliability of service usage.

PROJECTS

Zokyo

- Toyota-Connected-India/zokyo: Zokyo is a MLOps friendly image augmentation library written in python built with modularity and extensibility in mind. Specifically crafted for automotive deep learning development. (github.com).
- Zokyo is a MLOps friendly image augmentation library written in python for integrating with in-house Kubeflow pipelines.

PACL-Clip

- Vi-Sri/PACL-CLIP: Unofficial implementation of Patch Aligned Contrastive learning for Zero Shot Open Vocabulary segmentation (github.com).
- Unofficial implementation of Patch Aligned Contrastive learning for Zero Shot Open Vocabulary segmentation. Includes slurm scripts.

Panoptes

- Vi-Sri/panoptes-service: Web service and aws deployment for Natural language video search engine (github.com).
- Implementation of contextual moment and time retrieval engine for long form videos using natural language and deploying on AWS Sagemaker inference.

Adabins-FOHIS

- Vi-Sri/Adabins-FOHIS: Generating realistic Synthetic fog using FOHIS and Adabins (github.com).
- A tool to generate realistic Synthetic fog in monocular images using FOHIS and monocular depth maps from Adabins.

Fragilista

- Vi-Sri/fragilista: An active learning library for computer vision built from a notion wherein, what the model doesnt comprehend, it doesnt need. (github.com).
- An active learning library for computer vision which uses uncertainty methods to sample and re-label while training the model with lesser than 10% of the required data.

Graph LLM Agents for Grounded Database Retrieval

- Vi-Sri/Graph-Language-Retrieval: Natural Language Retrieval from Language graphs (github.com).
- A Natural language retrieval agent for retrieving information from graphs extracted from language and structure databases.

Vision Guided Cerebral Networks for Thought based Image Retrieval

- Vi-Sri/CerebralSignalNetworks: An exploration of experiments based on EEG signals used for vision tasks - Medical Image Computing (github.com)
- Understanding cross modal relationships between EEGs and Images, a descriptive analysis for building a next generation image retrieval from thoughts.

Mobilenet ONNX implementation on Android NDK

- Vi-Sri/MobilenetOnnxNDKAndroid: An experiment for NDK C++ implementation of mobilenet for Onnx Runtime (github.com)
- Implementing ONNX runtime inference for mobilenet object detection model on Android Native NDK.

Retail Eye : Retail theft and Ticket switch prevention system

- Vi-Sri/RetailEye: Retail Ticket switch Verification engine (github.com)
- Implementing State machine based anomalous pose detection and barcode switch detection system.