Shrikant Arvavasu

734-596-8050 ashrikant39@gmail.com https://www.linkedin.com/in/ashrikant39/ ashrikant39.github.io/

Summary

Computer Vision Engineer with 2 years of experience in 3D perception, LiDAR-based sensor fusion and spatio-temporal fusion using deep learning with Pytorch. Strong foundations in 3D geometry, photogrammetry and LiDAR perception. Experienced in building scalable perception pipelines, training multi-task deep networks, and deploying models for real-world environments.

Education

University of Michigan **Aug 2022 – May 2024**
Master of Science in Electrical and Computer Engineering *Ann Arbor, Michigan*
Specialization: Signal & Image Processing and Machine Learning *GPA: 3.97/4.0*

National Institute of Technology Karnataka **Aug 2018 – April 2022**
Bachelor of Technology in Electronics and Communication Engineering *Karnataka, India*
Honors: Machine Learning and Signal Processing *GPA: 3.87/4.0*

Experience

Kim's Lab, University of Michigan **Ann Arbor, Michigan**
Research Assistant (Computer Vision and GenAI) *May 2023 - Present*

- Integrated bounding box-based late temporal fusion method using a dynamic sparse-voxel transformer model to process NuScenes LiDAR input at 20 Hz, combining past detected boxes with LiDAR points, simulating real-time constraints in autonomous vehicle setups.
- Finetuned the state of the art BEVFusion for 3D bounding box detection by only utilizing **50%** of the lidar beams, achieving a mAP of 0.601 and NDS of 0.63 on **NuScenes dataset** using subsampled point clouds.
- Currently developing a Stream-PETR based framework for effective multi-sensor temporal fusion with camera, LiDAR and IMU.
- Developed novel diffusion sampling algorithms for inverse imaging problems, enhancing the quality and fidelity of the samples of **latent diffusion models**, achieving an FID score of 37.2, **an improvement of 17.2%** over the baseline model.

Skylark Labs **Dover, Delaware**
Machine Learning Intern (Representation Learning and Computer Vision) *June 2023 – August 2023*

- Designed a framework using a pre-trained RegNet model to achieve a **recall of 65% in self-learning** new categories by storing **multi-scale quantized features** to recognize pre-trained classes.
- Trained the neural network with a vector-quantized feature extractor to learn efficient multi-scale features of objects in natural scenes, enhancing the accuracy of the model by **12%** to detect objects from newly learned classes.
- Implemented the system to work on a **single core of a CPU** to run at about **3 fps** while storing features of new classes encountered.

Burris Lab, University of Michigan **Ann Arbor, Michigan**
Research Assistant (Computer Vision and Medical Imaging) *August 2022 – April 2023*

- Trained an **attention-UNET**-based model for aortic segmentation, enhancing the accuracy and efficiency of the Vascular Deformation Mapping pipeline, resulting in an **improvement of 3%** in the F1-score, particularly around aortic walls.
- Implemented corrections to an Elastix-based CT **Registration Pipeline**, improving the elastic registration performance of the pipeline for large deformations in the aortic walls. The corrections resulted in the detection of tissue growth by an improved **recall of 8%.**

SixSense Corporation **Jalan Besar, Singapore**
Computer Vision Intern (Automatic Augmentation and Multiclass Classification) *January 2022 – April 2022*

- Worked on detecting and classifying defects in semiconductor chips using Faster RCNN.
- Trained a stochastic **automatic augmentation** framework based on **Fast AutoAugment** on a ResNet50 model to techniques for several public datasets like CIFAR-100 and in-house datasets which improved the average accuracy by 2.3%. Integrated the automatic augmentation to the defect detection pipeline, improving the **accuracy by 1.4%.**

Projects

Autolabelling Driving Scenes in Cityscapes using DiNO and Segment Anything **November 2024**

- Developed an ML pipeline to auto-label the Cityscapes dataset by combining DINO (for 2D RoI extraction) with Segment Anything Model for semantic labels of road markings.
- Trained a YOLOv11-seg model for semantic segmentation of road markings, achieving robust performance with minimal manual annotation.
- Exported the trained model to ONNX format and integrated it into a C++-based executable for real-time road marking detection, suitable for deployment on embedded platforms.

Block-Based Compressed Sensing for Natural Images and Videos **January 2023**

- Innovated a block-based compressed sensing approach for natural images and videos, leveraging deep learning inspired by the insights from the paper "Video Compressed Sensing Using a Convolutional Neural Network."
- Trained the model and achieved a compression factor of 0.1 on non-keyframes of videos of KITTI Dataset.

Technical Skills

Languages/OS: Python, C, C++, Cuda, MATLAB, Shell Scripting, Linux
Developer Tools: OpenCV, SLURM, Git, Open3D, ROS, Docker
Machine Learning Tools: Pytorch, Pytorch-Lightning, MMDetection3D, Pandas
Technologies/Concepts: Deep Learning, Image Processing, Point Cloud Processing, 3D Computer Vision , Lidar-Voxel Algorithms, Sensor Fusion in AV systems, Generative AI, Diffusion Models
Soft Skills: Professional Communication, Collaborative Working, Interpersonal Communication, Team Oriented, Fast Paced Learner