Sivaramakrishnan Subramanian

Seeking Full-time roles in Computer Vision from Dec 2023
(571) 352-0457 | sivarams@cs.cmu.edu | Site | LI | Scholar

Education

Carnegie Mellon University, Robotics Institute **Pittsburgh, PA**
Master of Science in Computer Vision, GPA: ***4.17/4.00*** *Dec 2023*
Coursework: 3D vision, Learning-based vision, Geometric vision, Math for robotics, HRI, CV & ML
Teaching Assistant: Machine Learning, TA (392 students, Spring 2023); Head TA (487 students, Fall 2023)

Anna University, SSN **Chennai, IN**
Bachelors in Electrical Engineering, **Coursework:** Adv. Control Systems, DSP, OOP *May 2018*
Awards: Research grant equiv. $25k from *Dept. of Science & Technology, Govt. of India*; undergrad grant from SSN trust

Experience

Perception Intern **Waymo** *May 2023 - Aug 2023*
- Addressed object understanding problems in the perception **long-tail** by leveraging large foundational vision-language models (**VLMs**) to outperform production models
- Implemented 3D Point Cloud Transformer (**PCT**) in JAX, achieving **>0.80 F1** for mission-critical objects
- Extended Google's internal VLM with PCT and implemented distributed multi-pod **TPU** fine-tuning of the fused **5.3B** parameter model with **T5x** backend

Research Assistant, Xu Lab **School of Computer Science, CMU** *Aug 2022 - Dec 2022*
- Investigated **controllable-GAN** pipelines for 3D scene representation of Cryo-EM cells in tomography images
- Mentored 2 interns on self-supervised **domain adaptation** for detecting cell organelle in 3D tomogram slices

Senior Data Scientist, R&D Div **AppOrchid** *Oct 2019 - Jul 2022*
- Led a **12-member** ML team for ContractAI, a flagship product offering in CLM space, driving **$34MM** revenue in FY 2021 (**Spot award**: Employee of the Quarter)
- Prototyped custom layout analysis model on **YOLOv4** backbone for metadata extraction, achieving **0.73 mAP** on PubLayNet dataset and **8x** higher throughput than Caffe2 baseline
- Developed analysis-by-synthesis pipeline for underline (**U-Net**) and signature (**spatial pyramid pooling**) detection, reaching **0.93 precision** on Tobacco800 dataset at **34 fps**, a 15x speedup
- Designed cascade mask R-CNN + **BiFPN** model for financial document tabular extraction. Attained **0.48 AP** on ICDAR-cTDaR task with a **3x** latency cut post FP-quantization

Computer Vision Engineer **Soliton Tech** *Jun 2018 - Oct 2019*
- Engineered **KLT** tracker pipeline to monitor air voids in glass rods using **GigE** vision cameras with **16 msec** latency, deploying in Corning Inc's **$2MM pilot** system (**Spot award**: High Impact)
- Optimized **stroke-width transform** with operator approximation for font attribute extraction, cutting inference time from 40 to **0.6 sec/page**

Projects

Capstone: Stereo Depth Prediction **Air Lab, CMU** *Jan 2023 - Present*
- Developing 360° **multi-view** stereo depth prediction pipeline for drones with 6-pair **fisheye** cameras *(site)*
- Acquired dataset of **28** AirSim envs, studying **deformable**/spherical convolutions to handle spatial variance

IKEA-GAFR: 3D Reconstruction **Learning for 3D Vision** *Jan 2023 - May 2023*
- Built a **single-view** point cloud reconstruction pipeline with **joint part segmentation** for IKEA sketches, using a novel **combined loss** of point-wise cross-entropy, Chamfer distance, and MAE *(github)*
- Identified furniture assembly steps using **YOLOv5** detector trained on IKEA-Manual dataset with **0.807 mAP**

Skills

- **Languages**: Python, C++, MATLAB, LabView
- **ML Frameworks**: JAX/Flax, PyTorch, Tensorflow, Chainer
- **Tools**: OpenCV, AWS, REST API microservices, Docker, PostgreSQL, Django, FFmpeg, CUDA, Shell, XML