Prabhdeep Singh Sethi

E-mail: prabhdes@andrew.cmu.edu | Website: bit.ly/prab | LinkedIn: prabhdeep1999 | Github: Prabhdeep1999 | Mob: (412) 589-8023

Objective: Seeking Full-Time Opportunities in Computer Vision for December 2024

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
Master of Science in Computer Vision (GPA: 4.11/4) — December 2024
Coursework: Advanced Computer Vision, 3D Computer Vision, Multimodal Machine Learning, Reinforcement Learning

Government College of Engineering, Nagpur — Nagpur, India
Bachelor of Engineering in Computer Science (GPA: 9.5/10) — August 2021
Coursework: Operating Systems, Data Structures & Algorithms, Artificial Intelligence, Databases, Object Oriented Programming

SKILLS

Languages and Frameworks: Python, C++, Go, C, Bash; PyTorch, PyTorch3D, TensorFlow, Numpy, OpenCV, Flask
Tools and Platforms: Docker, Git, Triton, HuggingFace, TensorBoard, DeepStream, TensorRT, AIMET, SQL, AWS, Azure

EXPERIENCE

Apple — Sunnyvale, CA
Computer Vision Research Intern — 05/2024 - Current
- Researching 3D reconstruction for AR / VR applications using Gaussian Splatting in the Vision Pro team.

Wobot Intelligence — New Delhi, India
Computer Vision Engineer-II — 02/2022 - 08/2023
- Led a 6 member team to deliver vehicle and person Re-Identification solutions, serving 1M+ cameras and 10,000+ customers.
- Implemented an attribute-based fuzzy search with modified EfficientNet for local and global feature extraction. Further utilized VAE for dimensionality reduction & designed dynamic cosine similarity thresholds using k-means clustering.
- This approach reduced false IDs by 65% and improved Rank-1 of ReID by 35% in our multi-camera object tracking algorithm.
- Achieved 55 HOTA (Higher Order Tracking Accuracy) on MOT17 with custom object tracker, reducing false tickets by 28%.

Solar Industries India Ltd. (Research and Development Lab) — Nagpur, India
Senior Computer Vision Researcher — 08/2021 - 01/2022
- Led Smart Blast Project, achieved fume toxicity detection through background subtraction and color clustering.
- Designed the pipeline for Product Inspection of critical military parts by using a Vision Transformer (ViT) for object detection, achieving 96.5% mAP for detecting 9 such parts. Deployed models using Nvidia Triton for enhanced operational efficiency.

Computer Vision Intern — 01/2020 - 08/2021
- Developed Overspeeding and Automatic Number Plate Recognition solutions using YOLOv4, PaddleOCR for plate extraction, DeepSORT for real-time tracking and relative speed calculation with a margin of error of 10 m/s.

PUBLICATIONS

- S. Jain*, A. Kuthiala*, **P. Sethi**, P. Saxena, *"StyleSplat: 3D Object Style Transfer with Gaussian Splatting"*, Proceedings of the Conference on European Conference on Computer Vision **(ECCV)** Workshops, 2024 *(In Review)* [Project Page] [Paper]
- **P. Sethi***, A. Agrawal*, CMS Lezcano*, I. Heredia*, *"Listen Then See: Video Alignment with Speaker Attention"*, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition **(CVPR)** Workshops, 2024 [Poster] [Paper]
- R. Zawar, **P. Sethi**, R. Roy, *"Jensen-Shannon Divergence in Safe Multi-Agent RL", in* **ICLR**, Tiny Paper Track, 2024 [Paper]

PROJECTS

Generalizable Sparse view 3D Pose and Reconstruction (Gaussian Splatting) [Project Page] — 01/2024 - Current
Advisor: Dr Shubham Tulsiani
- Developed a feed forward network to represent 3D scenes using Gaussian Splatting in sparse settings with unposed cameras.
- Achieved to make it generalizable using transformer blocks to encode the latent 3D representation of the trained category.

GIF Tune (Video Diffusion) [Project Page] — 02/2024 - 04/2024
- GIF-Tune is a one-shot tuning strategy for continuous text-to-GIF synthesis. The model is trained on a single text-GIF pair and can generate GIFs from any text prompt. It uses depth conditioned Stable Diffusion model with 3D temporal attention layers.

UAV Detection (Small Object Detection) | [Code] — 10/2021 - 12/2021
- Enhanced UAV detection via GAN-based augmentation & tiling of infrared streams, using TensorRT Quantized YOLOv5s.
- The solution excelled in Anti-UAV Challenge by ICCV '21, delivering 37 FPS on Jetson TX2 achieving 95.1% mAP.
- Improved bird vs drone classification by 17% through trajectory analysis of flight pattern classified temporally using XGBoost.