Videsh **Suman**

(+1) (413)409-1993 | videsh.suman@gmail.com | sumanvid97

Career Summary

- Technically strong Machine Learning/Computer Vision Engineer seeking a challenging position to create impactful AI solutions
- **Strengths:** Deep Learning, Computer Vision, Perception, Object Recognition, 3D Reconstruction, Mobile Deployment, NLP

Education

University of Massachusetts (UMass) Amherst *Sep '19 - May '21*

M.S. in Computer Science (GPA: $4/4$)

Key Courses: 3D Vision, Computer Vision, Advanced NLP, Probabilistic Graphical Models, Optimization, Data Science Algorithms

Indian Institute of Technology (IIT) Bombay *Aug '15 - May '19*

B. Tech. in Engineering

Key Courses: Deep Learning, Machine Learning I & II, Reinforcement Learning, Medical Imaging, Data Structures & Algorithms

Work Experience

Software Engineer - CV/ML | Leia Inc. *Jan '22 - Feb '23*

Depth Estimation, Headtracking, Mobile/Cloud Deployment, Profiling

- Shipped AI-driven R&D projects on the cloud and the flagship LumePad^2 tablet to actualize immersive 3D experience
- Deployed depth estimation + view synthesis pipeline for 2D to 3D conversion of generative AI content for LeiaDream app
- Formulated reliable metrics using optical flow and transform invariance to effectively measure the flickering in video frames
- Enhanced existing video depth estimation with temporal consistency losses, improving temporal consistency by **30%**
- Worked on headtracking data collection on the cloud (AWS S3) and data wrangling for training head-position estimation model
- Implemented an attention-based seq2seq time-series forecasting pipeline for accurate and real-time headtracking on LumePad^2
- Optimized depth estimation using pruning and SNPE quantization, achieving **4x** reduction in model size and **2.3x** faster inference
- Created a custom profiling tool to monitor performance metrics of headtracking and 2D-3D content conversion apps on LumePad^2

Computer Vision Research Intern | Fyusion Inc. *July '21 - Jan '22*

Neural Rendering for 3D Car Exterior Visualization

- Proposed and built upon NeRF to perform novel view synthesis for challenging car exterior visualization with varying illumination
- Investigated differentiable rendering techniques for learning implicit shape representation and camera pose refinement

Student Researcher | GAMMA Lab, UMD *June '20 - Dec '20*

Driving Scene Understanding and Risk Assessment

- Constructed an end-to-end driving framework of perception stack + graph CNNs for learning traffic interactions among road users
- Achieved overall mAP scores within **3%** of state-of-the-art on driver action and cause recognition benchmarks of Honda Dataset
- Improved the existing risk assessment benchmark for pedestrians by **6%** through vulnerability modeling of road users

Research Intern | Microsoft Research India *May '19 - Jul '19*

Generative ML for Realistic Network Simulation

- Devised a novel network simulation pipeline to mimic the network behavior learned from Skype call logs
- Built LSTM-based state space framework to generate network traces on realism parameters like packet delay/loss, cross-traffic
- Accomplished high confidence on the realism quality by A/B testing with ns-2 for congestion-control algorithms

Skills & Achievements

- **Programming:** Python, C/C++, Java, MATLAB, R, SQL, Android Studio, Bash, Git, LaTeX
- **ML/CV Tools:** PyTorch, PyTorch Lightning, Tensorflow, Keras, Numpy, Scikit-learn, Pandas, PySpark, OpenCV, Open3D, Blender
- **MLOps Tools:** AWS, Azure, GCP, MLFlow, Grid AI, Weights & Biases, Sagemaker, Docker, SNPE, Torchscript
- **Scholastic Achievements:** KVPY Fellowship (2015) from IISc Bangalore; national rank of 1490/140K in JEE Advanced (2015)

Key Projects

3D Vision and Neural Rendering *UMass, Mar '20 - May '20*

- Developed PointNet for point cloud alignment (**76%** accuracy), and DeepSDF for 3D surface reconstruction from point clouds
- Performed efficient gradient-based camera pose optimization for 3D scenes from their implicit NeRF representations

Discriminative Adversarial Search for Text Summarization [report] *UMass, Oct '20 - Dec '20*

- Demonstrated the effectiveness of discriminative adversarial beam reranking for text summarization on CNN DailyMail dataset
- Implemented discriminator-driven beam reranking with UniLM for generating human-like (i.e. longer & more novel) summaries

Semi-Supervised Learning for Vision and Language Reasoning [report] [code] *UMass, Oct '19 - Dec '19*

- Investigated the direction of SSL via self-training, mixup regularization and MixMatch algorithms for NLVR2 dataset using LXMERT
- Concluded the **limitation** (**63%** accuracy) of this learning paradigm of leveraging unlabeled training data for a multimodal task

Super-Resolution of Indian Rainfall Projections | UG Thesis *IIT Bombay, Sep '18 - Apr '19*

- Leveraged encoder-decoder CNNs for 10× super-resolution to predict regional rainfall projections from climate parameters
- Improved state-of-the-art by achieving overall MSE of **5** mm/day rainfall for climate data across seven zones of Indian landmass