## 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, Visual Recognition, Perception, Neural Rendering, 3D Reconstruction, Mobile Deployment, NLP, LLMs

# **Education**

## University of Massachusetts (UMass) Amherst

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

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.

DEPTH ESTIMATION, HEADTRACKING, MOBILE/CLOUD DEPLOYMENT, PROFILING

- Shipped AI-driven R&D projects on the cloud and the flagship LumePad<sup>2</sup> 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, reducing video flickering by 30%
- Optimized depth estimation using pruning and SNPE quantization, achieving 4x reduction in model size and 2.3x faster inference
- Worked on headtracking data collection on the cloud (AWS S3) and data wrangling for extracting training data •
- Implemented an attention-based seq2seq time-series forecasting pipeline for accurate and real-time headtracking on LumePad<sup>2</sup>
- Created a custom profiling tool to monitor performance metrics of headtracking and 2D-3D content conversion apps on device

#### **Computer Vision Research Intern | Fyusion Inc.**

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

DRIVING SCENE UNDERSTANDING AND RISK ASSESSMENT

- Constructed an end-to-end driving framework of perception stack + graph CNNs for learning traffic interactions
- Achieved overall recognition scores in top 3% on Honda Dataset by effective feature extraction and learning pairwise attention
- Improved the existing risk assessment benchmark for pedestrians by 6% through vulnerability modeling of road users

#### **Research Intern | Microsoft Research India**

GENERATIVE ML FOR REALISTIC NETWORK SIMULATION

- Devised a novel network simulation pipeline to mimic the network behavior learned from Skype call logs
- Developed scripts for extracting training data from call logs, utilizing them to train LSTM models to generate network traces
- Accomplished high statistical confidence on the realism metrics by A/B testing with ns-2 for congestion-control

## Skills & Achievements

- Programming: Python, C/C++, Java, MATLAB, R, SQL, Android Studio, Bash, Git, LTFX
- 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**

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

## Discriminative Adversarial Transformers for Text Summarization [report]

- Led research to finetune BERT for discriminating between LLM-generated and human-like (i.e. longer & novel) summaries
- Demonstrated the benefit of discriminative BERT in UniLM's beam reranking algorithm for realistic text summarization

## Semi-Supervised Learning for Image+Text Data [report] [code]

- Researched the efficacy of SSL algorithms like self-training, mixup and MixMatch on multimodal BERT-like transformer
- Concluded the limitation (63% accuracy) of leveraging large-scale unlabeled training with images+textual data

## Super-Resolution of Indian Rainfall Projections | UG Thesis

- Leveraged encoder-decoder CNNs for 10× super-resolution to predict regional rainfall projections from climate parameters
- Gained 37% improvement over state-of-the-art by deep learning and parameter sharing across seven climatic zones

June '20 - Dec '20

May '19 - Jul '19

UMass, Mar '20 - May '20

UMass. Oct '20 - Dec '20

UMass, Oct '19 - Dec '19

IIT Bombay, Sep '18 - Apr '19

July '21 - Jan '22

Sep '19 - May '21

Aug '15 - May '19

Jan '22 - Feb '23