DISHANI LAHIRI

(412) 579-3618 | ✉ dishanil@andrew.cmu.edu | ⏿ github.com/dishanil | 🔗 linkedin.com/dishani-lahiri

EDUCATION

Carnegie Mellon University - School of Computer Science — Pittsburgh, PA
Master of Science in Computer Vision — *Dec 2023*
Courses: Generative Computer Vision, Learning for 3D Vision, Machine Learning, — CGPA: **4.17/4.00**
Visual Learning and Recognition, Advanced Natural Language Processing (Audit)
Graduate Teaching Assistant: Advanced CV (Prof. David Held), Machine Learning (Prof. Matt Gormley)

PROFESSIONAL EXPERIENCES

Carnegie Mellon Univeristy — Pittsburgh, PA
Research Collaborator @KLab, advised by **Prof. Kris Kitani** — *Jan 2023 - Present*

- Researching to build a low-power **visual-inertial localization** module for AR glasses
- Developing an **uncertainty estimation** method using an ensemble of IDOL-like models for inertial odometry
- Researched and implemented a PoseNet-like model employing Vision Transformers for visual odometry

Slingshot AI — New York City, NY
ML Research Intern — *May 2023 - Aug 2023*

- Optimized **text-to-image model** with Stable Diffusion 2.1 by foregoing the need for regularization images and user-defined captions per image; improved the training code by Diffusers
- Improved **consistency of facial features** in generations across ages for **Generative Aging** using **CycleGAN** by incorporating background through pre- and post-processing steps and improving **Identity encoders**
- Ideated and implemented reliable **text style transfer** model by fine-tuning **LLaMA2 7B** model on a dataset curated using GPT 3.5-turbo; generated text maintained content of the original text

Samsung Research and Development Institute — Bengaluru, India
Senior Software Engineer, Visual Intelligence Team — *March 2021 - July 2022*
Software Engineer, Visual Intelligence Team — *June 2019 - Feb 2021*
Software Engineer Intern, 5G Communication Team — *May 2018 - July 2018*

- Led the development of an **encoder-decoder based residual architecture** to denoise, retain texture, and demosaic low-light scene images to produce highly enhanced sRGB images; **deployed solution** in the Samsung Flagship S22 series and "Expert RAW" Application
- Experimented with CycleGAN with a modified weighted cycle-consistency loss function, Knowledge Distillation, and Quantization-Aware Training
- Implemented an on-device execution framework in C++ to run DL models on GPU, NPU, and DSP of edge devices; **minimized latency by 4x** and **memory footprint by 16x**; **deployed** framework in Flagship devices

PUBLICATIONS

S2RF: Semantically Stylized Radiance Fields | Paper | Code | Project page — *March 2023 - May 2023*
Lahiri, D., Panse, N., Kumar, M. In Proceedings of the 2023 International Conference on Computer Vision (**ICCV**) Workshop on AI for 3D Content Creation

- Researched and implemented a method to render a radiance field by optimizing **Plenoxels** for semantic and instance-based **stylization**; captures high-frequency details and maintains 3D-consistent structure
- Used **Nearest Neighbor Feature Matching (NNFM)** loss on masks generated by Segment Anything Model (**SAM**) on object detections by Detection Transformer (**DETR**)

Abnormal human action recognition using average energy images | Paper — *August 2017 - Dec 2017*
Lahiri, D., Dhiman, C., Vishwkarma, D.K. In Proceedings of the 2017 Conference on Information and Communication Technology (**CICT**).

- Developed a solution to **detect** humans in the image using **Feature-Pyramids with Fast-RCNN** and **predict** Abnormal Human Activities from depth images; achieved **92.5% accuracy** on UR Fall Dataset; **reduced latency by 80%** using average energy images
- Released a **dataset** for abnormal human activities of fainting, headache, and chest pain

TECHNICAL SKILLS

Python, PyTorch, C/C++, Git, OpenAI API, Hugging Face, OpenCV, Scikit-Learn, Generative AI, CMake