Rutika Moharir

rmoharir@cs.cmu.edu | rutika-moharir | +1 (412) 954-7912

EDUCATION

Carnegie Mellon University, School of Computer Science — Pittsburgh, PA
Master of Science in Computer Vision (Robotics Institute) — Aug. 2022 – Dec. 2023 (expected)

- **Coursework**: Computer Vision, Machine Learning, 3D Vision, Visual Learning and Recognition

Indian Institute of Technology — Dhanbad, India
Bachelor of Technology in Computer Science and Engineering — July 2015 – May 2019

WORK EXPERIENCE

KLab, Robotics Institute, CMU — Pittsburgh, PA
Research Assistant, Advisor: Prof. Kris Kitani — Dec. 2022 – Present

Sparse Image and Dual-IMU based Localization for AR/VR glasses

- Developing a multimodal fusion transformer-based architecture for visual-inertial odometry using inputs from IMU sensors combined with sparse camera images to improve state estimation accuracy.

Apple Inc. — Cupertino, CA
Computer Vision Intern, Vision Products Group — May 2023 – Aug. 2023

Realistic Video Anonymization using Generative Models

- Developed a video anonymization framework using generative model based data replacement techniques, such as object removal using GAN based inpainting and face anonymization using conditional GANs. The developed inpainting algorithm achieved an inference speed of 1.4 FPS on Apple M1 Mac.

Samsung R&D Institute — Bangalore, India
Senior Machine Learning Researcher — May 2019 – Aug. 2022

Scene Text Recognition Network

- Developed a CRNN architecture for OCR using channel and spatial attention based LSTM modules, achieving accuracy of 88.4% on ICDAR-13 with an inference speed of 2.44ms per word on Exynos 990 chipset device.
- Built an orientation classifier to identify horizontal and vertical text blocks using 3D-CNN with Global Average Pooling along width dimension jointly optimized with the recognition network.

Synthetic Data Generation

- Built a data generation pipeline that rendered text onto well defined image regions obtained using graph-cut algorithms.
- Implemented RANSAC with cues from depth estimations to find surface normal for perspective projection of the text onto images.

Joint Learning for Text Localization and Script Identification

- Developed a multi-task dual branch network where features extracted using ShuffleNet backbone were shared using a U-net architecture for real-time text localization and high level script clustering. Presented at IJCNN'21.
- Implemented locality sensitive hashing based projection to replace fully connected layers required for script classification.

Knowledge Distillation for Script Detection

- Achieved 35x compression for VGG-backbone with 2% performance drop using supervised distillation approach where the student network was trained using soft probabilities from the teacher network.

SKILLS

Deep Learning Frameworks: Pytorch, Keras, Tensorflow
Programming Languages : Python, C++, C, Swift, Android
Tools: OpenCV, NumPy, Git, Xcode, Visual-Studio, LaTeX, JIRA, CUDA

PUBLICATIONS / PATENTS

- Methods and systems for performing on-device image to text conversion (US Patent App. 17/859,629) A method for performing on-device image to text conversion which includes image language detection, understanding the text and using contextual and localized lexicon set for post optical character recognition (OCR) correction.
- TeLCoS: OnDevice Text Localization with Clustering of Script Presented in 2021 International Joint Conference on Neural Networks (IJCNN) Shenzhen, China, IEEE Xplore, DOI: 10.1109/IJCNN52387.2021.9533292
- On-Device Spatial Attention based Sequence Learning Approach for Scene Text Script Identification Presented in 6th IAPR International Conference on Computer Vision & Image Processing (CVIP2021)