Paritosh Mittal

paritoshmittal12@gmail.com | Contact No.: +1-(412)-726-1168
paritoshmittal12.github.io | www.linkedin.com/in/paritoshiitg

EDUCATION

Carnegie Mellon University, School of Computer Science — Pittsburgh, PA
Master of Science in Computer Vision (GPA: 4.18) — Feb 2021 - May 2022
Teaching Assistant: Introduction to Computer Vision (16-720 B) - Instructor: Kris Kitani

Indian Institute of Technology - Guwahati — Guwahati, IN
Bachelor of Technology in Computer Science and Engineering, Minor: Product Design — 2014 - 2018

WORK EXPERIENCE

Foundation Models, Tesla Autopilot — Palo Alto, CA
Sr. Machine Learning Engineer — Mar 2024 - Ongoing
Develop & deploy large foundation vision model unifying all perception tasks using shared intermediate features

- replaced 32 individual networks by a single large NN producing 250+ outputs encompassing all perception
- developed quantization friendly architecture with HW aware design to optimally to process inputs at 36Hz
- improved 3D perception geometry and F1 (overall by 10% and curbs etc. 16%) by using additional conditioning e.g. 3D estimated keypoints, maps etc. and additional co-training tasks

Fleet Learning, Tesla Autopilot — Palo Alto, CA
Machine Learning Engineer, ML — June 2022 - March 2024
Develop NN for 3D multi-cam based geometry estimation and collision avoidance system. [Talk Link]

- Shipped models to predict geometry & attributes of drivable space, arbitrary obstacles to 1M+ Tesla's
- Develop algorithms to auto-collect & label 1M+ videos of adverse scenes/rain/fog etc. Resolve over 80% FPs and enable self driving / active safety in low visibility cases
- Gravity align the predictions to improve recall and vehicle safety around slopes, hills and bumpy roads etc.

Perception Team Autonomous Driving, NVIDIA — Santa Clara, CA
Computer Vision Software Intern — May 2021 - Aug 2021
Developed Pytorch based NN to predict the height of road with 2cm MAE using 1M+ stereo HD image data.

- Enabled distributed training and evaluation to achieve 55x speed-up without loss in accuracy

Advanced Technology Labs, Samsung Research — Bangalore, IN
Senior Engineer - Machine Learning — July 2018 - Dec 2020
Formulated a new multi-sketch & position based input method for gallery image retrieval. Awarded a US patent
Developed a NN model to remove fence & reflections using stereo images & pixel disparity resulting in US patent.

PUBLICATIONS

AutoSDF: Shape Priors for 3D Completion, Reconstruction and Generation [Link]
[CVPR - 2022]: ***Mittal P****, Cheng Y*, Singh M, Tulsiani S*
Developed a transformer based autoregressive prior over 3D shapes and deployed it for shape completion. Proposed a framework to combine the proposed prior with pre-trained task specific encoders (ResNet, BERT).

Photo-realistic Emoticon Generation from Multi-modal input [Link]
ACM IUI - 2020 : ***Mittal P****, Aggarwal K, Vatsalya V, Sahu P, Singh V, Mitra S, Venkatesan S, Veera V*
Utilized a generative adversarial network to create realistic and personalized emoticons from user drawn sketches.

Image Memorability: The role of Depth and Motion[Link]
ICIP - 2018 : Sathisha B, ***Mittal P****, Sur A*
Illustrated using deep learning the influence of motion and depth information in predicting image memorability.

ACADEMIC PROJECTS

Autoregressive Conditional generation of realistic 3D objects — Aug 2021 - Apr 2022
MSCV Capstone Project; Advisors : Shubham Tulsiani (CMU), Maneesh Singh (Verisk Analytics) [Link]
Working towards generating high resolution and diverse 3D objects based on partial conditioning

- Developed VQ-VAE based method to map volumetric 3D SDFs to discrete and compact latent space
- Innovated a framework to combine shape priors with domain-specific encoders for conditional generation

Multi-modal multi-hop source retrieval using Graph Convolutions — Aug 2021 Nov 2021
Instuctor : Prof. Louis-Philippe Morency; [Link]

- Experiment with multi-modal arch. like CLIP, VinVL, VLP etc. for semantic feature extraction on WebQA

Computer Vision Course Projects — Feb 2021 - Apr 2021
Instructor : Prof. Deva Ramanan

- Developed an AR application using stereo matching, homography and RANSAC [Link]
- Implemented methods for 3D reconstruction using fundamental matrix and bundle adjustment [Link]

TECHNICAL SKILLS

- **Programming Languages**: Python, Java, C/C++, C#
- **Miscellaneous**: PyTorch, OpenCV, Tensorflow, Pandas, MySQL