Paritosh Mittal

paritosm@andrew.cmu.edu
Contact No.: +1-(412)-726-1168
www.github.com/paritoshmittal12
www.linkedin.com/in/paritoshiitg

Seeking **full time position** in Computer Vision, Machine Learning or Deep Learning

EDUCATION

Carnegie Mellon University, School of Computer Science — Pittsburgh, PA
Master of Science in Computer Vision (GPA: 4.17) — 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

Perception Team Autonomous Driving, NVIDIA — Santa Clara, CA
Computer Vision Software Intern — May 2021 - Aug 2021
Developed Pytorch based pipeline to predict the height of road using large scale HD data. Stereo image pairs were used as input and lidar points as ground truth.
- Enabled distributed training and evaluation to achieve 55x speed-up without loss in accuracy
- Experimented with GroupNorm, FPN, U-Nets to reduce mean absolute test error from 6cm to under 2cm

Advanced Technology Labs, Samsung Research — Bangalore, IN
Senior Engineer - Machine Learning — March 2020 - Dec 2020
Conceptualized and implemented a new multi-sketch and position based input method for gallery image retrieval. Customized a Resnet like model for sketch recognition. Optimized and deployed the model on device.

Engineer - Machine Learning — July 2018 - March 2020
Developed a deep neural network to remove fence and reflections. Created a novel pre-processing stage using stereo images to generate guidance. Method eliminates the need for optical flow or depth computation.
- Designed customized loss function and achieved 0.968 F1 score for fence segmentation
- Trained the neural network using adversarial, perceptual, and L2 losses to generate realistic HD output

PUBLICATIONS

Non-sequential Autoregressive Priors for 3D Completion, Reconstruction and Generation [Link]
[Submitted 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).

[Patent] A method for enhanced recognition of continuous inputs from multiple modalities
Undergoing US filing : **Mittal P**, *Sahu P, Pandotra H, Veera V, Venkatesan V*
Invented an AI system that uses geometric position and scene understanding to disambiguate continuous inputs.

[Patent] Method and multi-vision imaging device for segmenting obstruction [Link]
Undergoing US filing; ArXiv Pre-print : **Mittal P**, *Venkatesan S, Veera V, De A*
Innovated a method of using stereo images to generate additional input as real-time guidance to neural model.

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 - Ongoing
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 an autoregressive approach which can generate latent features based on 'non-sequential' inputs
- Proposed a unifying framework to combine shape priors with domain-specific encoders to perform 3D shape completion, single-view reconstruction and language based generation

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