AMOGH GUPTA

amogh112@gmail.com | (412)330-7272 | New York, NY
amogh112.github.io | linkedin.com/in/amogh112/

EDUCATION

Columbia University | MS in Computer Science — Dec 2020
Teaching Assistant - Advanced Software Engineering

Indian Institute of Technology Delhi | BTech in Industrial & Production Eng. — May 2019

EXPERIENCE

Amazon — *New York, NY*
Applied Scientist II — *July 2023 - Present*
Applied Scientist I — *March 2021 - June 2023*

· Research and development of 3D computer vision based products [On device deep learning for Augmented Reality(AR), parametric 3D human body model(SMPL-X) reconstruction from images and videos, automating generation of 3D assets from catalog images, camera control in video generation].

· AR - Shoe Virtual Try On Experience

- Launched real time on-device AR experience for virtual shoe try-on (iOS and Android), rolled out across North America and Europe. Co-inventor on 3 patent filings.
- Designed and trained efficient, multi-task deep learning models for on-device 6-DoF pose estimation, semantic segmentation, object detection, keypoints localization.
- Developed end-to-end ML pipeline : 3D data collection and annotation; distributed model training; quantization aware training; model compression and deployment; and cross-platform (iOS/Android) evaluation of quality, latency, and user experience.
- Presented work as an oral presentation at internal computer vision conference.

· 3D Human Body Reconstruction

- Improved prediction of parametric 3D human body mesh model (SMPL-X) from images and video using differentiable rendering, optical flow and optimization based approaches.

· 3D Object Reconstruction and Generation

- Automated 3D eyewear asset generation from 2D catalog images by inferring geometry and texture. Trained detection, classification, segmentation, orientation prediction, image-to-image generative models on synthetic and real-world data. Over 20,000 assets are live in the virtual try-on experience.
- Worked on 3D generative models based on multi-view diffusion predicting geometry and texture.

· Camera Control in Video Generative Models

- Developing techniques to inject camera control in video generative models, enabling controlled synthesis of product shots with desired viewpoints and motion paths.

Columbia University - Graduate Research Assistant — *September 2019 - December 2020*
Advisor - Prof. Carl Vondrick and Prof. Junfeng Yang — *New York, NY*

· Multi-Task Learning Strengthens Adversarial Robustness(Paper, Code) Showed that multi-task learning can make deep learning models robust in 11 computer vision tasks such as semantic segmentation, depth using Cityscapes, Taskonomy datasets. Accepted as oral presentation at ECCV 2020 (European Conference on Computer Vision), and in "Adversarial Robustness in Real World" workshop.

- Steering GANs to learn Robust Representations (Paper, Code) Proposed framework using BigGANs to mitigate spurious correlations in image classification and achieved state-of-the-art Top 1 accuracy of 39.3% (9% improvement over published baseline) on ObjectNet dataset. Presented at CVPR 2021.
- Learning vision-and-language representations Used contrastive learning based approaches to train vision-language transformer models on large image captioning and video datasets for commonsense reasoning and visual question answering.

Carnegie Mellon University - Research Assistant *March 2018 - February 2019*
Advisor - Prof. Jack Mostow *Pittsburgh, PA*

- Formulated affective computing framework for Android based intelligent tutoring system RoboTutor (code), finalist in Global Learning XPRIZE challenge to address lack of teachers in developing countries.
- Deployed computer vision pipeline in Android app using OpenCV, Java Native Interface, dlib libraries to detect and normalise faces, extract features and classify facial action units using FERA dataset.
- Designed user interface for writing tutor using Human Computer Interaction and educational psychology literature in Android, achieved significant learning gains in different writing skills against control group.

Center for Innovation and Business Creation at TU Munich *May 2017 - July 2017*
Digital Product School - Internship Program *Munich, Germany*

- Worked as machine learning engineer on user modeling and context aware hybrid recommender systems.
- Used HCI design approaches to design and implement recommender system interactions in a travel app.

Hi-Tech Robotic Systems - Autonomous Robotics Engineering Intern *December 2016*

- Implemented forward collision, lane change response system as Robot Operating System plugin in C++.
- Estimated odometry of vehicles using pinhole model on a monocular camera.

PUBLICATIONS

Multitask Learning Strengthens Adversarial Robustness - ECCV 2020 (Oral)
Chengzhi Mao, **Amogh Gupta***, V Nitin*, B Ray, S Song, J Yang, Carl Vondrick *Paper Code*

Generative Interventions for Causal Learning - CVPR 2021
Chengzhi Mao, Augustine Cha*, **Amogh Gupta***, Hao Wang, J Yang, Carl Vondrick *Paper Code*

PATENTS GRANTED

Virtual Shoe Try-On - U.S. Patent No. 11,978,174 *Link*
Yuelong Li, Gitika Karumuri, Miriam Bellver, Sunil Hadap, Ashwin Swaminathan, Amogh Gupta, Xin Shen

Reducing False Positives Based on Classification and Segmentation Cues - U.S. Patent No. 12,211,090 *Link*
Gitika Karumuri, Miriam Bellver, Amogh Gupta, Ashwin Swaminathan, Sunil Hadap, Yuelong Li, Xin Shen

SKILLS AND COURSEWORK

Languages: Python [PyTorch, Tensorflow, OpenCV, Scikit-Learn, Numpy], Java, C++, Matlab, SQL
Frameworks: Blender, AWS, Robot Operating System (ROS), Android Studio, Arduino, SolidWorks
Coursework: Computer Vision, Deep Learning Systems Performance, Natural Language Processing, Machine Learning, Speech Recognition, Databases, Analysis of Algorithms, Theoretical Data Science, Robotics, Inclusive Innovation, Embedded Systems Design Project

OTHER PROJECTS

Multi-Task Learning Model Compression for Edge Devices (Prof Parijat Dube) *Slides* *Code*
Analyzed local & global pruning strategies on multi-task deep learning models using Taskonomy dataset.

Speech Recognition (Prof Homayoon Beigi, Columbia University)
Built speech recognition pipeline for Indic languages using Kaldi toolkit.

LearnLab Summer School, CMU (Prof Jack Mostow, CMU)
Inferred metrics-usage, reliability, engagement and rejection from distribution of usage frequency and session duration in tutoring application.

Mahindra Driverless Car Challenge (Prof Sunil Jha, IIT Delhi) *Slides*
Integrated ROS modules for fusing sensor streams to get an accurate odometry, finetuned navigation stack for path planning, localization and lane detection for autonomous navigation of electric car.

ABU Robocon (Prof Kolin Paul, IIT Delhi)
Represented IIT Delhi at national round of university level robotics challenge. Designed and manufactured robots to drive using wind energy without actuation and pole climbing.

AWARDS

Design Innovation Seed Grant IIT Delhi
Awarded INR 100K to deploy embedded systems project in university for energy conservation.

Escalade, National Level Robotics Competition - 2015
Winners at IIT Guwahati with national record timing to complete a university level robotics challenge.

KVPY Fellowship - 2014
Awarded to students pursuing a research career, by Department of Science and Technology and conducted by Indian Institute of Science, Bengaluru.