Aviral Agrawal

Email : aviralagrawal124@gmail.com | Mo. : (412) 390-6740 | Website : aviral-agrawal.github.io

EDUCATION

Carnegie Mellon University Pittsburgh, PA
Master of Science in Computer Vision Dec 2024

- **CGPA** : **4.18/4.0**
- Graduate Researcher : Real-time use of 3D Gaussian Splatting for Robot tasks with Prof. Jeffrey Ichnowski Feb 2024 - Present
- Fall 23 (TA) 07-300 : "Research and Innovation in Computer Science" Aug 2023 – Dec 2023
- **Coursework** : Learning for 3D Vision, Geometry for Vision, Learning based Image Synthesis, Multi-Modal Machine Learning, Advanced NLP, Robot Learning, Advanced Computer Vision

Birla Institute of Technology and Science, Pilani Pilani, India
Bachelor of Engineering in Computer Science | Minor in Finance July 2020

- **CGPA : 9.18/10** (Graduated in *DISTINCTION Division)*
- **Coursework** : Machine Learning, Data mining, Data Structures and Algorithms, Object Oriented Programming, Operating System, Database systems, Compiler Construction, Computer Networks

PUBLICATIONS

KOROL: Learning Visualizable Object Feature with Koopman Operator Rollout for Manipulation
CoRL-2024 — Link Nov, 2024

- The paper proposes a learning based dexterous manipulation framework using Koopman operator that utilizes object features predicted by a spatial and frequency domain CNN-based feature extractor to auto-regressively advance system states. We evaluate our approach on simulated and real-world robot tasks, with results showing that it outperformed the model-based imitation learning NDP by 8% and the image-to-action Diffusion Policy by 16%

MaterialFusion: Enhancing Inverse Rendering with Material Diffusion Priors
3DV-2025— Link Aug, 2024

- In this paper we tackle the intrinsic challenge of disentangling albedo and material properties from input images by incorporating a 2D prior on texture and material properties of 3D objects using a StableMaterial, a diffusion model which is trained on albedo, material, and relit image data derived from BlenderVault, a dataset of approximately ~12K artist-designed synthetic Blender objects containing high quality material assets

Listen Then See: Video Alignment with Speaker Attention
CVPR-2024 Proceedings — Link Jun, 2024

- The paper proposes a cross-modal alignment and subsequent representation fusion approach to help the Visual Question Answering task's secondary modalities to work in tandem with the primary modality. We achieve state-of-the-art results (**82.06%** accuracy) on the Social IQ2.0 dataset for the task of Socially Intelligent Question Answering

Clear-Splatting: Learning Residual Gaussian Splats for Transparent Object Manipulation
ICRA-2024 RoboNeRF workshop (**Spotlight Presentation**) — Link May, 2024

- The paper proposes a method to leverage the scene-prior to first learn a *Background Splat* and subsequently learns a *Residual Splat* with the transparent object and the background combined. We also introduce *Depth Pruning* to address floaters. We achieve upto **67.09%** lower RMSE and upto **87.80%** lower MAE for depth estimation with transparent object compared to NeRF baselines

Development of a Machine Learning Based Model for Damage Detection, Localization and Quantification to Extend Structure Life
CIRP-2021 — Link Mar, 2021

- The paper proposes the transformation of a physical structure's mechanical response features to visual features which are fed to a CNN network delivering upto 85% better prediction than the previous state-of-the-art

PATENTS

NVIDIA Santa Clara, CA

- Filed - 18/830219

Samsung Research Institute Bangalore — Click patent numbers for status Bengaluru, India

- Published - US20230281458A1, US20230362367A1, WO2023167465A1, WO2023167514A1
- RQ filed - 202241042992, 202241072050
- Provisional filing - 202241072649, 202241073249

EXPERIENCE (3+ YOE)

NVIDIA — Santa Clara, CA
AI Video Intern — May 2024 – Aug 2024

- Designed novel AI-based Super Resolution and Video denoising model-based coding tool for AV2 video-compression standard
- The new model was 50x smaller than the previous model and resulted in better visual quality (VMAF) while maintaining the PSNR
- Applied for a patent for the newly innovated model

SAMSUNG RESEARCH INSTITUTE BANGALORE — Bengaluru, India
Senior Engineer (Computer Vision) — Jan 2021 - Jun 2023

- Owned the AI-based replacement of Video compression In-Loop filter, achieved 10% bd-rate gain
- Curated data using quantization range resultant artifacts-based binning for model generalization
- Innovated a novel training strategy for a better performing smaller network than a conventional more complex network. Worked on device deployment feasibility by reducing model multiply-and-accumulation operations
- Innovated and developed the model for AI-based Transform and Quantization as well for AI-hybrid VVC codec by creating novel architecture and an amalgamation of multiple loss functions to achieve energy compaction and suitability for entropy encoding

ORACLE — Bengaluru, India
DevOps Engineer — Nov 2020 - Jan 2021

- Created and owned a FLASK-based web-service and the deployment as part of the OCI Exascale team
- Applied the Flask application for synchronous resource management for team-based shared resources

Couture.ai — Bengaluru, India
Data Engineer — Aug 2020 - Oct 2020

- Responsible for building generalized ML usecase pipelines using Couture.ai API for end-user black-box deployment

AMAZON — Bengaluru, India
Research Engineer Intern — Jan 2020 - Jul 2020

- Developed a Reverse Geocoding module using Named Entity Recognition, custom clustering tree, beam search, and reference data-based filtering to output an address. Model deployed in Amazon India marketplace
- Created an Address Classifier using a multi-branch CNN architecture resulting in 6% better prediction AU-ROC than previously used LSTM model. Model deployed in Amazon middle- east marketplace

SAMSUNG RESEARCH INSTITUTE BANGALORE — Bengaluru, India
Student Trainee — May 2019 - Jul 2019

- Researched methods to improve digital image zoom by leveraging a multi-focal lens array system feeding a custom-Unet model to fuse input images and produce a single image with better zoom legibility
- Achieved 1.5 dB Peak-Signal-to-Noise-Ratio (PSNR) improvements over baseline method

SKILLS

PyTorch, Python, C++, C, Pytest, PySpark, Flask, AWS

AWARDS & SCHOLARSHIPS

- **Samsung Star IP Award - Star Young Innovator** : Most IPs created within 1 year of joining — Apr 2023
- **Samsung Excellence Award** : Exceptional research to market and ecosystem building activities — Feb 2022
- **Bengalathon** : Felicitated by the WEST BENGAL GOVERNMENT for winning the hackathon — Dec 2019
- **Scholarship America** : Received scholarship (thrice) for holistically meritorious students — 2019, 18, 17