Shaurya Dewan

Website | Linkedin | Github | Google Scholar | P: +1 (412) 287-8594 | shauryadewanmanu@gmail.com

EDUCATION

Carnegie Mellon University   Pittsburgh, PA
Master of Science in Computer Vision | GPA: 4.17 / 4.00   *Dec 2024*
Coursework: Deep Learning Systems, Multimodal Machine Learning, Learning for 3D Vision

International Institute Of Information Technology - Hyderabad   Hyderabad, India
Bachelor of Technology (Honors) in Computer Science and Engineering | CGPA: 9.12 / 10.00   *May 2023*
Teaching Assistantships: Automata Theory, Mobile Robotics
Coursework: Machine Data and Learning, Digital Image Processing, Computer Vision, Statistical Methods in AI

PUBLICATIONS

DiffusionPID: Interpreting Diffusion via Partial Information Decomposition (Project Page)
NeurIPS 2024 (Under Review)   *May 2024*
- Led a research project (**first author**) on interpreting diffusion models by adapting the information theoretic concept of **Partial Information Decomposition (PID)** and conducted an extensive analysis to identify the model's various failure modes and biases.

StableSemantics: A Synthetic Language-Vision Dataset of Semantic Representations in Naturalistic Images (Project Page)
NeurIPS 2024 (Under Review)   *Jun 2024*
- **Co-first author** on a paper introducing a comprehensive dataset comprising 224K human-curated prompts, over 2 million synthetic diffusion-generated images, and **10 million semantic maps** corresponding to individual nouns. Includes benchmarks on several tasks.

Canonical Fields: Self-Supervised Learning of Pose-Canonicalized Neural Fields (Project Page)
CVPR 2023 (Highlight)   *Jun 2023*
- **Published a CVPR 2023 highlight paper** as a **co-first author** in collaboration with **Brown University** (Prof. Srinath Sridhar) to develop a pipeline, CaFi-Net, for self-supervised canonicalization of NeRF-generated fields in PyTorch-lighting.
- Created a dataset of **1300 trained NeRF models**. This involved designing a Blender-based simulator to capture images of ShapeNet models, experimenting with the NeRF codebase to obtain clean models, and sampling the density fields from the trained models.
- Extended the notion of equivariance to and enabled the operation of 3D point cloud networks on NeRFs.

RESEARCH

Physical Perception Lab @ CMU | Prof. Shubham Tulsiani   *Nov 2023 – Present*
Generalizable 3D Gaussian Splatting in the Sparse Setting (Ongoing)
- Currently leading a project on learning a model that can predict attributes of 3D gaussians in a single-forward pass without any optimization for a new scene at inference from sparse data (4-5 views) such that these gaussians can be rendered from any novel view.

Robotics Research Center @ IIIT-H | Prof. Madhava Krishna   *May 2021 – May 2023*
Pose-Shape Optimizer Pipeline (Code)
- Recreated a pipeline in MATLAB and Ceres that reconstructs and localizes a vehicle in 3D given a monocular 2D image.
- This involved deforming a 3D wireframe (obtained from vehicle shape priors using PCA) by estimating and optimizing for the pose and shape deformation such that the wireframe aligned with the vehicle's key points when reprojected to 2D.
- Further applied multi-view optimization over the shape to stabilize it across sequences of video frames.
- Made use of a deep stacked hourglass CNN model to estimate vehicle key points in the images.

Precog Research Group @ IIIT-H | Prof. Ponnunrangam Kumaraguru   *Jan 2022 – Dec 2022*
Synthetic Data Generation | Infosys
- Assembled a relationship-preserving multi-table synthetic data generation system using a hierarchy-preserving variant of CTGAN from the SDV framework in Python which has since been **integrated into the Infosys Enterprise Data Privacy Suite (iEDPS)**. Employed classification and regression models such as Random Forest, Lasso, Ridge, etc. to model and enforce relationships between attributes.

WORK EXPERIENCE

NVIDIA   Santa Clara, CA
Robotics Software Intern - Robotics Perception   *May 2024 – Aug 2024*
- Enhanced the performance of the ESS 4.0 stereo perception model with better data (real data), more diverse augmentations, and architectural experiments. Improvements have since been integrated into the pipeline and will be part of ESS 4.1.
- Boosted training time by 10x with distributed training and explored alternate evaluation methods to better quantify model performance.

PROJECTS

GaussCraft: Language Driven Segmentation and Editing in 3D Using Gaussian Splatting | CMU   *Jan 2024 - Apr 2024*
- Introduced a method capable of segmenting and editing any desired object in a given 3DGS scene solely based on users text inputs.

Needle Deep Learning Library (Code) | CMU   *Sep 2023 - Dec 2023*
- Developed a PyTorch-like Deep Learning library from scratch including **auto-differentiation** and **CUDA programming** of linear operations such as matrix multiplication, convolution, etc. Supports fundamental models like CNN, LSTM, Transformer, etc.

TECHNICAL SKILLS

Languages: Python, C++, C, MATLAB, R, Octave, Bash, x86   **Tools**: Docker, Blender, Anaconda, Jupyter, Git, Latex
Libraries: PyTorch, OpenCV, Open3D, scikit-learn, Pandas   **Databases**: MongoDB, SQL