

Divyam Goel

divyamg@andrew.cmu.edu | +14129099325 | dv-fenix.github.io | linkedin.com/in/divyamgoel3 | github.com/dv-fenix

Education

Carnegie Mellon University, PhD in Robotics Aug 2026 – Present

- **Advisor:** Andrea Bajcsy

Carnegie Mellon University, MS in Robotics Aug 2024 – May 2026

- **GPA:** 4.05/4.0
- **Advisors:** David Held & Zackory Erickson
- **Thesis:** Red and Blue Teaming for Robust Manipulation Under Geometric Variations

IIT Roorkee, BTech in Electronics and Communications Engineering July 2018 – May 2022

- **GPA:** 3.9/4.0
- **Thesis:** Multimodal Deep Representation Learning and its Applications in Image Generation and Embodied AI

Research Experience

Robotics Institute, Carnegie Mellon University Aug 2024 – Jun 2026

Graduate Student Researcher — Red and blue teaming for robust robotic manipulation

Advisors: David Held & Zackory Erickson

- Studied how failures under plausible geometric variation reveal the generalization limits of learned manipulation policies, developing a geometric red-teaming framework that searches over continuous object deformations to systematically discover failure-inducing shapes. Our work was accepted at CoRL 2025 as an Oral Presentation (Top 5.7%).
- Demonstrated that geometric vulnerabilities discovered in simulation transfer to the real world across grasping and insertion tasks, building end-to-end evaluation pipelines spanning differentiable shape representations, gradient-free failure search, simulation, and physical experiments on xARM6 and Franka Panda robots.
- Developed a diffusion-based generative model that learns from prior red-teaming experience to produce failure-inducing geometric variations for unseen objects, enabling scalable targeted data generation for improving the robustness of robot grasping policies.

Simon Fraser University July 2023 – July 2024

Research Assistant — Learning generative scene layouts for 3D indoor environments

Advisor: Angel Xuan Chang

- Led a research direction on graph-conditioned 3D layout generation, designing a multi-relational, directed scene-graph extractor over 3D-FRONT with relation priors and pruning policies (e.g., `on`, `above`, `is_facing`, `next_to`, `close_to`).
- Implemented GCN-conditioned diffusion baselines with classifier-free guidance and floor-plan conditioning via pre-trained image encoders, and ran ablations on positional encodings, attention schemes, and rotation representations to stabilize training in low-data 3D layout regimes.
- Developed differentiable, ergonomics-aware objective functions (e.g., reach, line-of-sight, adjacency/accessibility) and integrated them into the generative pipeline as auxiliary training and evaluation signals to target functional, human-centered layouts beyond standard FID/KID metrics.
- Contributed to SemLayoutDiff by building the training and evaluation stack: standardizing the evaluation paradigm and rendering protocols, benchmarking baselines (ATISS, DiffuScene, LayoutGPT) under matched settings. Our work was accepted to 3DV 2026.

GIST Vision Lab July 2021 – Dec 2022

BTech Thesis — Multimodal Representation Learning for Image Generation and Embodied AI

Advisors: Jonghyun Choi (SNU) & Dongyeop Kang (Univ. of Minnesota)

- Designed language-grounded controllers for embodied instruction following on TEACH (EDH) using hierarchical learning: a meta-controller schedules low-level navigation/manipulation skills from dialogue and visual context, trained end-to-end atop the Episodic Transformer with an auxiliary multimodal grounding loss; accepted at the EmbodiedAI workshop (CVPR 2022).
- Developed SMART, a multimodal recurrent transformer with gated memory and VQ-VAE image tokens for temporally consistent story visualization, substantially improving FID/FSD on the Pororo-SV benchmark.
- Proposed a framework for zero-shot story visualization by combining CLIP-guided diffusion with a Temporal Guidance Network that biases generation toward temporal coherence without task-specific training data.

The University of British Columbia

June 2021 – Sept 2022

MITACS Globalink Research Intern (GRI) — Efficient Adaptation of LLMs for Software Engineering

Advisor: Fatemeh H. Fard

- Developed MODE-X, an adapter-based cross-modal model that inserts lightweight language- and task-specific adapters into RoBERTa to transfer representations from natural language to C/C++, Java, and Python for cloze test and code clone detection, published and presented at the IEEE/ACM International Conference on Program Comprehension (ICPC) 2022.
- Demonstrated that MODE-X achieves competitive performance with CodeBERT on CodeXGLUE cloze-test and clone-detection benchmarks while being up to 60–140× more parameter-efficient for task-specific tuning and over 15× more memory-efficient, enabling practical deployment of code models in resource-constrained settings.
- Constructed and released the Super Code Clone Detection-88 (SCD-88) dataset for Python code clone detection, along with open-sourced adapters, scripts, and results, establishing a new benchmark for adapter-based models on source code.

Indian Institute of Technology Roorkee

Feb 2021 – July 2021

Undergraduate Researcher — Machine Learning methods for Hyperspectral Unmixing

Advisor: Saurabh Khanna

- Proposed a Semi-NMF-inspired proximal regularization scheme for training autoencoders, deriving penalties from the KKT conditions of the constrained Semi-NMF formulation of hyperspectral unmixing so that the network remains consistent with the linear mixture model. Our work was accepted to the 30th National Conference on Communications (NCC 2024).
- Developed a general regularized-training framework for both linear and shallow autoencoders, together with Semi-NMF-specific criteria (normalized fit and abundance errors) and synthetic / semi-synthetic testbeds, providing a reusable recipe for physics-informed neural factorizations in hyperspectral unmixing.

Carnegie Mellon University

Sep 2020 – Jan 2021

Undergraduate Independent Study — Decoding and Error Analysis of Neural Machine Translation

Advisor: Sachin Kumar (Ohio State)

- Analyzed the neighbourhood structure of word embedding spaces induced by continuous-output NMT models, linking local geometry to systematic errors on homonyms, rare words, and semantically close confounders.
- Developed embedding-space decoding strategies (e.g., neighbourhood-constrained search and context-sensitive re-ranking) to curtail homonym and near-synonym prediction errors without modifying the underlying translation model.

Indian Institute of Technology Roorkee

June 2020 – Sep 2020

Undergraduate Researcher — Syntax-aware Graph Neural Models for Hate Speech Detection

Advisor: Raksha Sharma

- Proposed SyLSTM, a syntax-aware BiLSTM–GCN architecture that integrates dependency parse trees with word embeddings to jointly model semantic and syntactic cues for fine-grained hate speech detection.
- Designed a graph transformation pipeline to convert dependency parses into GCN-ready graphs with inverse edges and self-loops, enabling the use of sentence structure to reduce bias from pejorative terms used in non-offensive contexts.
- Established new state-of-the-art results on multiple hate/offense benchmarks while using over 10× fewer parameters than BERT. Our work was accepted at the SocialNLP workshop, NAACL 2022.

Publications

SemLayoutDiff: Semantic Layout Generation with Diffusion Model for Indoor Scene Synthesis

Xiahao Sun, Divyam Goel, Angel Chang

International Conference on 3D Vision (3DV) 2026

Geometric Red-Teaming for Robotic Manipulation

Divyam Goel, Yufei Wang, Tiancheng Wu, Guixiu Qiao, Pavel Piliptchak, David Held†, Zackory Erickson†

Conference on Robot Learning (CoRL) 2025 (Oral Presentation — Top 5.7%)

Semi-NMF Regularization-Based Autoencoder Training for Hyperspectral Unmixing

Divyam Goel, Saurabh Khanna

National Conference on Communications (NCC) 2024

Leveraging Dependency Grammar for Fine-Grained Offensive Language Detection using Graph Convolutional Networks

Divyam Goel, Raksha Sharma

SocialNLP @ NAACL 2022

Language Guided Meta-Control for Embodied Instruction Following

Divyam Goel, Kunal Pratap Singh, Jonghyun Choi

Embodied AI @ CVPR 2022

On the cross-modal transfer from natural language to code through adapter modules

Divyam Goel†, Ramansh Grover†, Fatemeh H. Fard

International Conference on Program Comprehension (ICPC) 2022

Awards

- “Geometric Red-Teaming for Robotic Manipulation”, was accepted as an **Oral Presentation (Top 5.7%)** at CoRL 2025.
- I was awarded the travel grant for CoRL 2025 held in South Korea.
- I was awarded the IIT Roorkee Heritage Foundation Excellence Award for 2021 for excellence in academia and leadership.
- I was selected as a MITACS scholar, a research grant awarded by the Mitacs-SICI partnership to top international undergraduates for a research internship under the supervision of Canadian university faculty members.
- I was awarded the IIT Roorkee Heritage Foundation Excellence Award for 2019 for excellence in academia and leadership.
- I was among 10 students selected at IIT Roorkee for the Quantum Computing summer training organized by QuLabs.
- I was awarded “The Special Award for Academic Excellence”, given to top-3 students at DPS Gurgaon for continued academic excellence through middle and high school.

Industry Experience

Machine Learning Engineer (L2), Adobe – Bengaluru, India Nov 2023 – July 2024

- Worked on text-to-video diffusion models for video inpainting, analyzing failure modes where standard samplers (e.g., DDIM) produced physically implausible dynamics (such as water ripples moving against the current after an edit).
- Designed Sequential Monte Carlo (SMC)-based sampling strategies that treat generation as search in the model’s latent space, using proposal distributions, conditional score updates, and resampling to bias trajectories toward physically consistent motion without retraining the underlying model.
- Built a diagnostic workflow for generative failure analysis that distinguishes inference-induced artifacts from genuine data/model limitations, to prioritize when to deploy advanced samplers across multiple video editing scenarios.

Deep Learning Researcher, Rephrase.ai (acquired by Adobe) – Bengaluru, India June 2022 – Nov 2023

- Developed a multimodal generative pipeline for lip-synced talking-head avatars, integrating text, speech, and facial representations to produce high-fidelity cross-lingual lip sync for 100+ languages.
- Designed a diffusion-based text-to-speech model for high-fidelity, zero-shot voice cloning, achieving an order-of-magnitude reduction in training time compared to prior production TTS models while removing dependence on third-party APIs.
- Engineered and deployed a containerized AWS microservice (ECS, S3, Lambda) that orchestrates video rendering pipelines for large-scale personalized video generation, with automated scaling and job scheduling for 10^5 + videos per day.

Research and Development Intern, Pando Labs – Chennai, India Aug 2020 – Nov 2020

- Formulated 3D bin packing and capacitated vehicle routing in the freight domain as sequential decision-making problems, and designed off-policy deep RL agents to optimize packing and routing under realistic capacity and geometry constraints.
- Built simulated freight environments to evaluate policy performance against production heuristics, incorporating shipment sizes, route constraints, and cost models to approximate real-world operating conditions.
- Achieved ~85% packing efficiency on 3D bin packing and ~50% improvement in cost efficiency on capacitated vehicle routing relative to existing solutions, demonstrating the practical value of RL-based policies in operations research.

Selected Projects

Uncertainty-Gated World-Model Shared Control for Autonomous Racing Spring 2026

Course Project - Embodied AI Safety (16-886)

Instructor: Andrea Bajcsy

- Studied risk-aware shared control in autonomous racing, asking whether predicted failure risk and model uncertainty can determine when a learned assistant should intervene in an otherwise competent nominal controller.
- Developed a LiDAR-based DreamerV3 assistant that models the dynamics of executed blended actions and predicts nominal-controller actions during imagined rollouts, enabling consistent world-model training under shared control.
- Showed that full-action assistance improves normalized progress from 0.36 to 0.65 over the nominal controller and increases success at a braking-critical turn from 10% to 70%; further found that uncertainty-based deference can suppress useful intervention precisely in rare regions where assistance is most needed.

Learning SoccerNet Camera Calibration with Geometry Priors

Fall 2025

Course Project - Geometry-based Methods in Vision (16-822)

Instructor: Michael Kaess

- Studied how uncertainty from learned visual detectors can be incorporated into geometric optimization for robust camera calibration from challenging broadcast soccer footage.
- Reformulated point-and-line pose refinement as a heteroscedastic nonlinear least-squares problem, using calibrated detector confidence to control the influence of individual geometric constraints in a Levenberg–Marquardt solver.
- Calibrated neural confidence using isotonic regression and improved the SoccerNet calibration benchmark’s final score from 62.25 to 64.13 while increasing completeness from 84.44% to 86.24%.

Neural Primitive Assembly for 3D CAD Reconstruction

Spring 2025

Course Project — Learning for 3D Vision (16-825)

Instructors: Shubham Tulsiani & Ioannis Gkioulekas

- Developed a neural framework for 3D shape reconstruction that represents objects as compositions of simple primitives using differentiable Boolean logic, avoiding hand-crafted composition rules.
- Designed and implemented two learned assembly modules, FuzzyBoolean and GumbelBoolean, that allow the network to infer complex union / intersection / subtraction operations directly from 3D training data.
- Built a training and evaluation pipeline showing that the learned Boolean assemblies produce high-fidelity, interpretable reconstructions of CAD models, with structured primitive decompositions suitable for downstream design and editing tools.

Technologies

Languages: Python, C++, Matlab

Libraries: Isaac Gym, PyBullet, PyTorch, PyTorch3D, Lightning, Tensorflow, PyTorch Geometric, JAX, Deoxys

Tools: Git, Vim, LaTeX, Bash, HTML/CSS, Docker

Hardware: Franka Emika Panda, xArm6, 3D-printing

Invited Talks

Intent Lab, Carnegie Mellon University: “Geometric Red-Teaming for Robotic Manipulation”

Stella Yu’s group, University of Michigan: “Geometric Red-Teaming for Robotic Manipulation”

Academic Services

External Reviewer

- **Conferences:** CoRL 2026, NeurIPS 2026
- **Workshops:** ICRA 2026 Workshop on Manipulation Robustness

Volunteering

CMU PAIR | Mentor

Fall 2025

- Mentoring CMU CS freshmen to navigate research in LLMs and robotics, with an emphasis on assistive robots and HRI.
- Strategized personalized pathways to help mentees identify and secure entry points into CMU labs and projects aligned with their specific research interests.

AI4ALL Ignite Fellowship | Mentor

Summer 2025

- Mentored a national cohort of undergraduate students from historically excluded backgrounds in AI.
- Guided students through the development of a technical AI portfolio project, providing expert feedback on research direction, methodology, and implementation.

IEEE Student Chapter, IIT Roorkee | Core Member

Feb 2020 - May 2022

- Organized paper reading sessions and invited talks to promote student interest in electronics and computing.

Team Shiksha Kendra | Student Teacher

Jun 2014 - May 2017

- Taught science and history to a batch of 60 underprivileged middle school students, utilizing interactive teaching methods to enhance student comprehension and retention.
- Significantly contributed to an increase in average classroom attendance from 65% to approximately 80% .

Extra Curriculurs

Vision and Language Group – IIT Roorkee

Dec 2020 — May 2022

Core Member

- Organized work-shops, paper reading sessions, and tutorials for undergrads at IIT Roorkee on self-supervised learning, autoregressive transformers and VQ-VAEs.
- Mentored 12 freshmen through their first project in Computer Vision – Semantic Segmentation of Human Faces.

Watchout! – IIT Roorkee

Aug 2018 — May 2022

Chief News Coordinator

- Led a team of over 75 editors, designers and web-developers at the official student media body of IIT Roorkee.
- Authored critical articles on recent developments within the campus community, contributing to informed discourse and campus awareness. Also wrote abstract articles, engaging readers with diverse topics and perspectives.
- Conducted interviews with esteemed guest lecturers, enhancing the media body’s reputation for insightful content.

- Bronze medalist in the Institute Sports Trophy.
- Participated in both inter and intra college sporting events.

References

- **Prof. David Held**, Associate Professor, Robotics Institute, Carnegie Mellon University. [dheld@andrew.cmu.edu]
- **Prof. Zackory Erickson**, Assistant Professor, Robotics Institute, Carnegie Mellon University. [zackory@cmu.edu]
- **Prof. Angel Xuan Chang**, Associate Professor, School of Computing Science, Simon Fraser University. [angelx@sfu.ca]
- **Prof. Jonghyun Choi**, Associate Professor, Dept. of Electrical and Computer Engineering, Seoul National University. [jonghyunchoi@snu.ac.kr]
- **Prof. Fatemeh Fard**, Assistant Professor, The University of British Columbia. [fatemeh.fard@ubc.ca]
- **Prof. Saurabh Khanna**, Assistant Professor, Electronics and Communications Engineering, IIT Roorkee. [sakhanna@ece.iitr.ac.in]
- **Harsh Agarwal**, Senior Machine Learning Manager, ASML, Adobe. [harsh.agarwal.ri20@gmail.com]