Chaitanya Chawla

chaitanya1chawla.github.io cchawla@cs.cmu.edu 0412-606-6352 chaitanya1chawla

Education

Carnegie Mellon University Aug 2024 – July 2026
MS in Robotics - GPA: 4.13/4.00 – Advisor: Prof. Guanya Shi
- **Interests:** Learning-based manipulation, Imitation Learning, Reinforcement-learning

Technical University of Munich Oct 2020 — May 2024
BS in Electrical Engineering - GPA: 3.91/4.00 (Top 3%)
- 4x National German Scholarship - Deutschlandstipendium

Publications

R. Qiu*, S. Yang*, X. Cheng*, ***C. Chawla****, T. He, R. Hoque, ... G. Shi, X. Wang. ***Humanoid Policy ~ Human Policy***
Submitted to Robotics Science and Systems 2025

T. Shankar, ***C. Chawla***, J. Oh. ***Translating Agent-Environment Interactions across Humans and Robots***
International Conference on Intelligent Robots and Systems 2024: **Oral Presentation**

Industrial Experience

Student Collaborator Dec 2024 – Present
Manufacturing Futures Institute, CMU (Pittsburgh)
- Designing an end-to-end visual pick-place framework for industrial manipulation. Used Dino-V2 to detect and match visual features such as grasp location on the objects.
- Developed algorithms to perform pick-place maneuvers on Lego-like objects using Yaskawa GP4 manipulator. Modeled a gripper to house an endoscopic camera for closed-loop feedback.

Robotics Software Developer - Part-time July 2023 – July 2024
Roboverse Reply GmBH (Munich, Germany)
- Built a perception pipeline for autonomous analog gauge reading, including data annotation, real-time inference, and post-processing, integrated with Boston Dynamic's Spot via ROS and Docker for monitoring in a factory setup.
- Created a webRTC pipeline using gRPC to transfer point cloud data from Spot's LIDAR sensor to Oculus VR Headset, enabling a remote user to observe Spot's immediate environment in real time.

Research Experience

Sim2Real Policy transfer for Humanoid Manipulation *May 2025 – Present*
- Developing RL policies using large-scale MoCap human-object interaction datasets as imitation rewards
- Developing framework for sim2real transfer of learnt policies using retargeting and teacher-student distillation to a Unitree-G1 robot with Dex3 hands

Cross-embodiment Learning for Humanoid Manipulation Sept 2024 – Mar 2025
- Co-developed HAT (Human Action Transformer), a unified cross-embodiment policy trained on large-scale egocentric human and small-scale humanoid robot data, achieving strong out-of-distribution generalization in dexterous manipulation.
- Designed a unified state-action representation enabling joint training of human and humanoid robot policies via wrist pose and fingertip retargeting, improving generalization across embodiments.

Learning Abstract Representations of Agent-Environment Interactions Aug 2023 – Dec 2023
- Proposed TransAct, a framework for learning temporally abstract interaction representations that jointly model agent behavior and environment dynamics from human and robot demonstrations.
- Developed a factored latent representation with auxiliary objectives (contrastive, Jacobian, and state reconstruction) to enable smooth, composable skill abstractions for zero-shot human-to-robot skill transfer.

Robot-Agnostic Framework for Human-in-the-Loop Feature Extraction Jan 2023 – Aug 2023
- Proposed an algorithmic framework to extract task-specific constraints from human demonstrations
- Generated a knowledge base for task affordances, such as wiping/rubbing, by computing surface normals, surface limits, and object-surface proximity through point cloud data

Selected Projects

Self-supervised fine-tuning for Pre-Grasps through 3D Object Generation
- Proposed a novel 3D latent diffusion framework for text- and image-conditioned voxel generation, integrating efficient VAE and attention-based UNet architectures.
- Enhanced robotic grasp synthesis by fine-tuning pre-trained models on generated 3D assets, demonstrating improved generalization to diverse object geometries.

Learning Dexterous Manipulation from Human Video Pretraining using 3D Point Tracks
- Proposed a pipeline to benchmark pre-training methods using different state representations.
- extract sensorimotor information from videos by lifting the human hand and the manipulated object in a shared 3D space in simulation (IsaacGym), i.e. either 3D point-tracks or 3D meshes.
- Retargeting hand-trajectories to a Franka with a Shadow hand, followed by BC task-specific fine-tuning.