

Pardis Taghavi

Mountain View, CA 979-326-9823 taghavi.pardis@gmail.com
<https://linkedin.com/in/pardis-taghavi> <https://github.com/PardisTaghavi>

RESEARCH FOCUS

Ph.D. researcher focused on controllable video world models, efficient inference for diffusion/flow-matching models, sparse attention, and autonomous-driving simulation. Research experience spans DiT-based video generation, conditioning and controllability of world models, dynamic-scene generation, and physical AI.

EXPERIENCE

Graduate Researcher Jan 2022 – Present
Texas A&M University – *supervised by Shu Kong and Zhengzhong Tu* College Station, TX

- Developing controllable video world models on top of NVIDIA Cosmos-Predict2.5 for autonomous-driving simulation, conditioning generation on ego trajectory, dynamic-object motion, and scene structure to improve temporal consistency and controllability.
- Researching efficient inference for DiT-based video generation models through sparse attention, block selection, and flow-preserving sparsity, targeting faster video world-model sampling with minimal quality degradation.
- Distilled large vision-language and foundation models into efficient student perception models for in-vehicle GPUs using stage-wise semi-supervised training, reducing FLOPs by 70% while maintaining or improving performance.
- Led the perception team for the AVA Project, building a modular multi-sensor autonomy pipeline integrating LiDAR, cameras, and IMU for rural and urban driving environments.
- Developed SwinMTL, a shared-encoder multi-task model for depth estimation and semantic segmentation, reducing compute by 40% while improving Cityscapes mIoU by 3%; ranked among the top 4 teams awarded MCity access. (Code)

AI Engineer Intern May 2025 – Sep 2025
Rivian Palo Alto, CA

- Developed and evaluated a GRPO-based RL post-training framework for in-vehicle ASR, improving robustness over supervised fine-tuning under out-of-distribution noisy driving conditions.
- Proposed a curriculum learning schedule for RL optimization, stabilizing training and improving generalization under challenging in-vehicle acoustic conditions.

EDUCATION

Ph.D. Mechanical Engineering, Texas A&M University; GPA: 4.0/4.0 2022 – Present
B.S. Mechanical Engineering, Politecnico di Torino; GPA: 110/110, *summa cum laude* 2018 – 2022

SELECTED PUBLICATIONS

The Pulse of Motion: Measuring Physical Frame Rate from Visual Dynamics. *[Under review]*. arXiv
SwinMTL: A Shared Architecture for Simultaneous Depth Estimation and Semantic Segmentation in Scene Understanding. *[IROS 2024]*. arXiv
Training a Student Expert via Semi-Supervised Foundation Model Distillation. *[CVPRW 2026]*. arXiv
NaviDriveVLM: Decoupling High-Level Reasoning and Motion Planning for Autonomous Driving. *[Under review]*. arXiv

TECHNICAL SKILLS

Video/World Models: NVIDIA Cosmos, Wan, HunyuanVideo, Diffusers, DiT, Flow Matching
Acceleration: CUDA, FlashAttention, TensorRT, vLLM, HPC/Slurm
Frameworks: PyTorch, JAX, TensorFlow, Hugging Face, OpenMMLab
AV/Robotics: ROS, CARLA, LiDAR, Multi-Sensor Fusion Languages: Python, C++

HONORS AND AWARDS

TOPolito Scholarship awarded to the top 10 students of the engineering school. 2018 – 2022
NSF sub-award for remote autonomous vehicle perception testing at MCity facility. 2024

ACADEMIC SERVICE

Program Committee Member, CVPR Workshop on Multi-Agent Embodied Intelligent Systems 2026
Reviewer: CVPR, ECCV, NeurIPS