Hetian Guo

Ph.D. in AI, Jilin University | Research: VLA policies, predictive world models, visuo-tactile manipulation

hetianguoml@gmail.com +86 186 8239 8096 Homepage GitHub

Education

Jilin University · Ph.D. in Artificial Intelligence Aug 2026–Present

Research: embodied AI, VLA, world-action models, dexterous-hand control; Advisor: Prof. Zipei Fan

University of Georgia · M.S. in Engineering (Mechanical Engineering) 2024–2025

Research: intelligent transportation, autonomous driving, hardware-software co-simulation; Advisor: Prof. Yunli Shao

Southern University of Science and Technology · B.Eng. in Computer Science 2020–2024

Research: intelligent transportation, spatiotemporal data mining, traffic forecasting; Advisor: Prof. Xuan Song

Research and Industry Experience

AI Algorithm Intern · Avant Robotics (Shenzhen Daohe Tongtai Robotics) Shenzhen, China Jan 2026–Present

- Deployed Sonic and GMR + AMO motion pipelines on Unitree G1 and integrated Inspire dexterous hands; collected multi-task real-robot teleoperation data and fine-tuned/deployed Psi-0 and GR00T N1.7.
- Studied SeMoCo language-conditioned motion generation (AAAI 2027 submission); replayed generated whole-body motion on physical Unitree G1 through Sonic.

Research Intern · LocationMind Inc. Tokyo, Japan Aug–Oct 2023

- Studied waterway-network extraction, trend/seasonality decomposition, and static/dynamic spatial relations in North Sea and Baltic Sea AIS data; contributed to a marine traffic forecasting model published in *Transportation Research Part E* (2025).

Selected Research and Publications

PLaW-VLA: Predictive Latent World Modeling for Vision-Language-Action Policies

Co-first author | CoRL 2026 accepted

- Predicts future states in a pretrained visual latent space, giving VLA policies task-relevant dynamics rather than pixel-reconstruction targets.
- Joint predictive pretraining and action modeling achieve 97.4% success on LIBERO and 72.7% on LIBERO-Plus.

BeliefVLA: Implicit Belief-State Driven Vision-Language-Action Models for Robust Manipulation

Second author | IROS 2026 accepted

- Infers continuous implicit belief states from observation history to inject temporal memory and dynamics priors into VLA policies.
- Uses V-JEPA2 predictive features instead of single-frame reactive control, improving robustness to occlusion and observation perturbations.

SeMoCo: A Semantic-First Motion Codec for Motion Language Modeling

Co-first author | AAAI 2027 submitted

- Separately quantizes motion semantics and kinematic residuals into hierarchical tokens that preserve both meaning and reconstruction detail.
- A dual-axis generator models semantic progression and kinematic refinement; Ω-MotionVerse unifies multi-source motion data.

TacDyn-WAM: Learning Implicit Tactile Dynamics in a Heterogeneous Visuo-Tactile World Action Model

ICLR 2027 submitted

- Learns dynamics-aware tactile representations and predicts multi-horizon contact changes without reconstructing tactile pixels.
- Couples visual and tactile experts for future-state and action modeling, achieving 81.5% average success on UniVTAC.

Earlier Research in Intelligent Transportation Systems (ITS)

HHGNN: Heterogeneous Hypergraph Neural Network for Traffic Agents Trajectory Prediction in Grouping Scenarios

First author | ICRA 2024

- Models higher-order interactions in grouped traffic scenes with heterogeneous hypergraphs for agent trajectory prediction.

A Decomposable Multi-Fusion Spatio-Temporal Marine Traffic Flow Forecasting Algorithm: Taking the North Sea and Baltic Sea Region as an Example

Second author | *Transportation Research Part E*, 2025

- Fuses waterway networks and multi-scale spatiotemporal relations to forecast marine traffic flow in the North Sea and Baltic Sea.

Research and Engineering Skills

Research methods: VLA / world-action models, latent world prediction, visuo-tactile learning, motion generation, and policy fine-tuning.

Robot systems: Bimanual/humanoid deployment and teleoperation data; hands-on with Astribot S1, Realman, ARX R5, AgileX PIPER, and Unitree G1.

Implementation: Python, C++, PyTorch, ROS / ROS2, Transformers, Diffusers; Sonic and GMR + AMO.