Xinyao Rao

Algorithm Engineering
Foundation Models · Autonomous Driving

+86 198 2926 5306 744044292@qq.com github.com/Rtwotwo Google Scholar

Research interests: World Models · Foundation Models · Vision-Language Models
Autonomous Driving · Pre-/Mid-/Post-training

EDUCATION

Nanjing University | Master's studies Sep 2026 - Sep 2029 (expected)

CITE Lab / NJU-3DV · Advisors: Xun Cao and Wei Li

Northwestern Polytechnical University | Undergraduate studies Sep 2022 - Jun 2026

EXPERIENCE

EABOT | Intern Jul 2026 - present · Ends Jan 2027 (expected)

Research focus: autonomous driving and trajectory planning.

RESEARCH CONTRIBUTIONS

DriveSTA | Learning from Scored Trajectory Alternatives for End-to-End Driving

First author · arXiv:2611.27515 · 2026 · Paper · Project · Code

- **Richer supervision.** Constructed scored alternatives beyond one human demonstration through smooth deformations, execution evaluation, and diversity screening: **2.61M targets across 95,998 real scenes**.
- **Temporal planning.** Combined autoregressive proposals, Hungarian matching, and multi-target residual diffusion; shared **JEPA video features** between refinement and rescoring, avoiding online VLM inference.
- **Quality and efficiency.** Scale achieved **95.55 NAVTEST PDMS** and **46.7 HUGSIM HD-Score** (345 episodes). Mean forward latency fell **73% (1,295 to 349 ms)** versus ChainFlow-VLA in matched H20, batch-1 evaluation.

DriveTTO | Search and Selection in Test-Time Trajectory Optimization

First author · arXiv:2611.29205 · 2026 · Paper · Project · Code

- **Training-free search.** Coupled proposal-supported endpoint search with bounded refinement and time-aligned fusion of proposal and control branches, keeping **all planners and scorers frozen**.
- **Reliable selection.** Required both scorers to predict reference-relative gains before branch/fusion updates. Recovered **51.35% of native zero-PDMS cases** on average; net zero-score rate fell from **1.63% to 1.13%**.
- **Generalization.** Reached **95.54 PDMS / 92.30 EPDMS** on NAVTEST and **57.10 EPDMS** on NAVHARD under benchmark-specific configurations. The single-scorer variant improved **all six evaluated planners** on both splits.

Spark-H3 | Adaptive Block Sparse Attention for Video Diffusion Models

Co-author (4th author) · arXiv:2611.24256 · 2026 · Paper · Project · Code

- **Adaptive grouping.** Co-authored a training-free attention method. **Spark-Reblock** recursively groups tokens by attention preference into regular tiles, combining adaptive sparsity with GPU-friendly block structure.
- **Attention calibration. Spark-Reweight** uses a global query anchor to build weighted KV representatives and logit biases; exact and compressed branches share normalization to preserve attention mass.
- **Measured acceleration.** Achieved **2.49× DiT speedup** with **24.81 dB PSNR** versus dense outputs on MiniMax-H3 at 10% exact-block density (20 steps, 345 frames, 1344×768; 4 dense warm-up steps).

TECHNICAL SKILLS

Tools: Python, PyTorch, NumPy, OpenCV, Linux, Git.

Modeling & evaluation: Vision Transformers, DINOv2/LoRA, JEPA, diffusion models, trajectory scoring, NAVSIM/HUGSIM, reproducible ablations, and experiment recovery.