Jiwei Yang

13005149668 | yangjiwei@tsinghua.vin | github.com/NoctYang | noctyang.github.io

Education

School of Vehicle and Mobility, Tsinghua University 2025.11 - 2026.06
Jointly Trained Student, Prof. Jun Li's Research Group

School of Information Science and Technology, Great Bay University 2025.06 - 2025.08
Visiting Student, YU Vision Group

School of Artificial Intelligence, Guangzhou Institute of Science and Technology 2022.09 - 2026.06
Bachelor's Degree in Software Engineering

Research Experience

Prof. Jun Li's Research Group, Tsinghua University 2026.06 - Present
Research Assistant
Research on action-oriented world models and autonomous driving planning; completed Auto-JEPA and SDRA, both submitted to AAAI 2027.

Center for Brain Inspired Computing Research (CBICR), Tsinghua University 2025.08 - 2025.10
Embodied Intelligence Algorithm Engineer
Explored multimodal perception, language understanding, and action planning through model and algorithm experiments in embodied intelligence.

Publications

- **Auto-JEPA: A Latent World Model of Continuous Intent for End-to-End Autonomous Driving**
  First Author | Submitted to AAAI 2027 | arxiv.org/abs/2607.29031
  🌟 **Cited by Yann LeCun.**
  - Proposed Auto-JEPA to learn continuous future driving intent via joint-embedding prediction, addressing reliance on dense future-scene prediction.
  - Achieved **91.3 PDMS on NAVSIM v1** and **89.1 EPDMS on NAVSIM v2**.
  - Semantic occlusion of dynamic traffic-participant regions induced 2.97 times the intent change caused by random occlusion.
  - Occluding vehicles relevant to future driving significantly changed predicted intent and selected trajectories; occluding irrelevant vehicles left both largely unchanged.
- **SDRA: Spatiotemporally Decoupled VLM-Enhanced Reinforcement Learning for Energy-Efficient Heavy-Vehicle Driving**
  First Author | Submitted to AAAI 2027
  - Proposed a VLM-enhanced, spatiotemporally decoupled reinforcement learning framework for heavy vehicles in complex highway scenarios.
  - Designed spatial and temporal agents for lane-level intent decisions and cruise-coast control, respectively. A VLM integrates road images and structured traffic states to provide scene understanding and energy-efficient driving triggers.
  - Experiments using real heavy-vehicle data showed a 6.7% reduction in fuel consumption with optimized trajectories.
- **OP-RTDETR: Real-Time Outdoor Parking Space Detection with Grouped Convolutions and Group-Level Contrastive Denoising**
  First Author | SCI Q2 Journal
  - Proposed OP-RTDETR based on RT-DETR to improve the speed and accuracy of real-time outdoor parking space detection.
  - Used a GC-VanillaNet grouped-convolution backbone for efficient inference, with group-level contrastive denoising and grouped one-to-many matching to improve accuracy.
  - On PKLot, OP-RTDETR achieved higher inference speed and detection accuracy than the mainstream models evaluated.
- **CHyFN: A Consistency-Aware Hybrid Feature Network for Robust Object Recognition in Complex Scenes**
  Co-first Author (Listed Second)
  - Proposed a unified recognition framework for weakly discriminative features arising from background and instance similarity.
  - Fused global structure and local texture features across diffusion timesteps; used contextual consistency to suppress spurious target-background correlations.
  - Conducted experiments on camouflaged object detection, fine-grained image classification, and generated-image detection, achieving competitive localization and classification performance.

Research Interests

My research focuses on learning and reasoning about the real world with world models, particularly real-world dynamics and interactions: understanding how the world evolves and how agents interact, beyond predicting the next observation.

Honors and Awards

First Prize, Programming Finals, 5th National College Computer Ability Challenge
National-level Project Approval, Undergraduate Innovation and Entrepreneurship Training Program, 2024