

# Ziqi Pang

ziqip2@illinois.edu | <https://ziqipang.github.io/> | [Google Scholar](#)

**Researcher** and **Builder** for **Multi-modal Intelligence**:

Experienced in Perception, VLM, Visual Generation, and Multi-modal Agents

## SELECTED WORKS

---

- **Multi-modal Computer Use Agent**: [Lux](#). Data, modeling, agent, and product. 72.3% success rate on OnlineMind2Web and 53.8% on OSWorld, Claude Sonnet-4.5 level with much less resources.
- **Autoregressive Visual Generation**: [RandAR](#) (CVPR 2025 Oral). Flexible pre-training of autoregressive generation enables diverse generalization to novel tasks.
- **(Long) Video Understanding**: [MR. Video](#) and [XComp](#) (both NeurIPS 2025). Agentic and post-training of VLMs for context reduction in long video understanding. Fewer tokens, higher accuracy.
- **3D and Video Perception**: [PF-Track](#) (CVPR 2021) and [SimpleTrack](#). Widely adopted 3D tracking systems and models in the autonomous driving industry. Reducing tracking errors by >90%.

## EDUCATION

---

<b>University of Illinois Urbana-Champaign (UIUC)</b> Doctor of Philosophy in Computer Science	<i>September 2021 - May 2026 (Expected)</i>
<b>Peking University (PKU)</b> Bachelor of Science in Computer Science, Cum Laude	<i>September 2016 - June 2020</i> GPA: 3.74/4.00, Top 15%
<b>Carnegie Mellon University (CMU)</b> Summer Research Assistant	<i>May 2019 - September 2019</i>

## BUILT

---

- OpenAGI Foundation** *2025 - April 2026*
- Founding Research Scientist, Training Computer-Use Agent: [Lux](#)
- (1) Agent Framework: Building Lux-powered WebAgent achieves 72.3% on OnlineMind2Web (Outperforming Google's CUA at 69% and most other competitors)
  - (2) Agentic Modeling: Post-training (SFT+RL) for Lux, achieves 58.8% success rate on OSWorld, performance close to Sonnet-4.5 with much smaller scales.
  - (3) Data: Organizing Annotators, Building Data Cleaning/Annotation Pipelines, Synthetic Data Generation Workflows.
  - (4) Infra: Building inhouse RL&Eval sandboxes and scenarios.
  - (5) Product: Designing and building real-time assistive products based on Computer Use ([Lux](#)).

## RESEARCH EXPERIENCE

---

- University of Illinois Urbana-Champaign, Ph.D.** *September 2021 - May 2026*
- [Multi-modal Video Understanding and 3D Perception](#).
- (1) Principled context management for long video understanding ([MR. Video](#) - NeurIPS 2025).
  - (2) Learnable token compression in VLMs for long video understanding ([XComp](#) - NeurIPS 2025).
  - (3) Language-instructed video segmentation with unified global-local reasoning ([GLUS](#) - CVPR 2025).
  - (4) Improved memory bank for long and challenging video segmentation ([RMem](#) - CVPR 2024).
  - (5) First long-term temporal fusion and data flywheel for 3D HD mapping ([MV-Map](#) - ICCV 2023).
  - (6) Continuous motion prediction mimicking real-world traffic ([Streaming Forecasting](#) - IROS 2023).
  - (7) End-to-end vision-only 3D perception, 10% errors than previous ones ([PF-Track](#) - CVPR 2023).

◦ [Generative Models and Unified Models](#).

- (1) Improving representation-guided visual generation (JERI - In Submission to ICML 2026).
- (2) Autoregressive visual generation with task generalization ([RandAR](#) - CVPR 2025 Oral).
- (3) Diffusion model for perception ([ADDP](#) - ICLR 2025).
- (4) Diffusion-based image generation from graph connections ([InstructG2I](#) - NeurIPS 2024).
- (5) Transformers from LLMs as visual encoders ([LM4VE](#) - ICLR 2024 Spotlight).

◦ [Benchmarking and Evaluation](#).

- (1) Agentic PPT editing ([PPTArena](#) - In Submission to ECCV 2026).
- (2) VLM for agriculture ([AgMMU](#) - NeurIPS 2025, D&B).

**Meta FAIR**, Research Intern

*May 2025 - August 2025*

- [SAM3](#). Building hierarchical vocabulary for SAM3 training.

**NVIDIA**, Research Intern

*May 2024 - August 2024*

- [Online HD Map Prediction](#). For the autonomous driving group at NVIDIA, we conducted research on building scalable generative pre-training for online high-definition (HD) map predictions..

**Toyota Research Institute**, Research Intern

*May 2022 - August 2022*

- [3D tracking and motion forecasting from multiple cameras](#). End-to-end multi-object tracking (MOT) and motion prediction decrease tracking errors ([ID-Switches](#)) by more than 90% on nuScenes compared to previous state-of-the-arts ([PF-Track](#) - CVPR 2023).

**TuSimple**, AI Residency for Perception in Self-driving

*June 2020 - August 2021*

- [LiDAR-based 3D perception for autonomous driving](#). Public projects:

- (1) Sparsity-inspired outdoor 3D detection, first transformer-based 3D detector, widely followed baseline since then ([SST](#) - CVPR 2022).
- (2) A widely used and robust 3D multi-object tracking framework ([SimpleTrack](#) - ECCVW 2022).
- (3) Data flywheel and object auto-labeling from single-object tracking ([LiDAR-SOT](#) - IROS 2021).

## SELECTED PUBLICATIONS

---

**MR. Video: “MapReduce” is the Principle for Long Video Understanding** [[MR. Video](#)]

*Ziqi Pang, Yu-Xiong Wang*

**NeurIPS 2025**

**One Token per Highly Selective Frame: Towards Extreme Compression for Long Video Understanding** [[XComp](#)]

*Zheyu Zhang, Ziqi Pang, Shixing Chen, Xiang Hao, Vimal Bhat, Yu-Xiong Wang*

**NeurIPS 2025**

**RandAR: Decoder-only Autoregressive Visual Generation in Random Orders** [[RandAR](#)]

*Ziqi Pang\*, Tianyuan Zhang\*, Fujun Luan, Yunze Man, Hao Tan, Kai Zhang, Willian T. Freeman, Yu-Xiong Wang*

**CVPR 2025 (Oral)**

**GLUS: Global-Local Reasoning Unified into A Single Large Language Model for Video Segmentation** [[GLUS](#)]

*Lang Lin\*, Xueyang Yu\*, Ziqi Pang\*, Yu-Xiong Wang*

**CVPR 2025**

**Aligning Generative Denoising with Discriminative Objectives Unleashes Diffusion for Visual Perception** [[ADDP](#)]

*Ziqi Pang\*, Xin Xu\*, Yu-Xiong Wang*

**ICLR 2025**

**InstructG2I: Synthesizing Images from Multimodal Attributed Graphs** [[InstructG2I](#)]

*Bowen Jin, Ziqi Pang, Bingjun Guo, Yu-Xiong Wang, Jiaxuan You, Jiawei Han*

**NeurIPS 2024**

**RMem: Restricted Memory Banks Improve Video Object Segmentation** [[RMem](#)]

*Junbao Zhou\*, Ziqi Pang\*, Yu-Xiong Wang*

**CVPR 2024 (Winner at ECCV'24 VOTst Challenge)**

**Frozen Transformers in Language Models Are Effective Visual Encoder Layers** [[LM4VE](#)]

*Ziqi Pang, Ziyang Xie\*, Yunze Man\*, Yu-Xiong Wang*

**ICLR 2024 (Spotlight)**

**MV-Map: Offboard HD-Map Generation with Multi-view Consistency** [[MV-Map](#)]

*Ziyang Xie\*, Ziqi Pang\*, Yu-Xiong Wang*

**ICCV 2023**

**Streaming Motion Forecasting for Autonomous Driving** [[Streaming Forecasting](#)]

*Ziqi Pang, Deva Ramanan, Mengtian Li, Yu-Xiong Wang*

**IROS 2023**

**Standing Between Past and Future: Spatio-Temporal Modeling for Multi-Camera 3D Multi-Object Tracking** [[PF-Track](#)]

*Ziqi Pang, Jie Li, Pavel Tokmakov, Dian Chen, Sergey Zagoruyko, Yu-Xiong Wang*

**CVPR 2023**

**Embracing Single Stride 3D Object Detector with Sparse Transformer** [[SST](#)]

*Lue Fan, Ziqi Pang, Tianyuan Zhang, Yu-Xiong Wang, Hang Zhao, Feng Wang, Naiyan Wang, Zhaoxiang Zhang*

**CVPR 2022**

**SimpleTrack: Understanding and Rethinking 3D Multi-object Tracking** [[SimpleTrack](#)]

*Ziqi Pang, Zhichao Li, Naiyan Wang*

**ECCV Workshop 2022, Patented 2023**

**Model-free Vehicle Tracking and State Estimation in Point Cloud Sequences** [[LiDAR-SOT](#)]

*Ziqi Pang, Zhichao Li, Naiyan Wang*

**IROS 2021**

## PREPRINTS

---

**Unlocking the Full Potential of Small Data with Diverse Supervision** [[SmallData](#)]

*Ziqi Pang\*, Zhiyuan Hu\*, Pavel Tokmakov, Yu-Xiong Wang, Martial Hebert*

**Arxiv Preprint 2021**

**Immortal Tracker: Tracklet Never Dies** [[ImmortalTracker](#)]

*Qitai Wang, Yuntao Chen, Ziqi Pang, Naiyan Wang, Zhaoxiang Zhang*

**Arxiv Preprint 2021**

## PATENTS

---

**Multiple target tracking method and apparatus, calculating device and storage medium**

*Ziqi Pang, Zhichao Li, Naiyan Wang*

US Patent App. 17/816,239, 2023

## SERVICES

---

**Teaching Assistants** for **CS 446 (Machine Learning)** and **CS 445 (Computational Photography)** at University of Illinois Urbana-Champaign (UIUC), and **ICS (Introduction to Computer System)** at Peking University (PKU).

**Reviewer** for CVPR, ICCV, ECCV, NeurIPS, ICLR, ICML, RA-L, ICRA, IROS.

## **AWARDS AND SCHOLARSHIPS**

---

**Outstanding Graduate** at Peking University *June 2020*

**Peking University Scholarship** at Peking University (Top 10%) *September 2018*

**Kwung-hua Scholarship** at Peking University (Top 5%) *September 2017*