

# Haomin Wang

[✉ kiyotakawang@sjtu.edu.cn](mailto:kiyotakawang@sjtu.edu.cn) | [/github.com/hmwang2002](https://github.com/hmwang2002) | [hmwang2002.github.io](https://hmwang2002.github.io)  
| [G Google Scholar](#)

*Ph.D. Student in Artificial Intelligence, focusing on Multimodal LLMs & SVG Generation.*



## 🎓 Education

### Shanghai Jiao Tong University

*Ph.D. in School of Artificial Intelligence*

Sep. 2025 – Present

*Shanghai, China*

### Nanjing University

*B.Eng. in Software Engineering*

GPA: 4.579/5.00 | Rank: 9/254 (Top 3.54%) | English: CET-4 636, CET-6 578

Sep. 2021 – Jun. 2025

*Nanjing, China*

## 💼 Research Experience

### Shanghai AI Laboratory

*Research Intern, Large Language Model Center*

Jul. 2024 – Present

*Shanghai, China*

## 📄 Publications

- [ICLR 2026] **InternSVG: Towards Unified SVG Tasks with Multimodal Large Language Models** *First Author*
  - Proposed a unified modeling framework for SVG understanding, editing, and generation tasks driven by MLLMs.
  - Constructed **SAgoge**, the largest multimodal SVG training dataset to date (16M samples), and introduced the **SArena** benchmark for comprehensive evaluation.
  - Achieved state-of-the-art performance via SVG-specific tokens and subword-based embedding initialization, resulting in the **InternSVG** framework.
- [NeurIPS 2025] **Point or Line? Using Line-based Representation for Panoptic Symbol Spotting in CAD Drawings** *Co-first Author*
  - Proposed **VecFormer**, a Transformer-based framework replacing point-based primitives with line segments to preserve geometric continuity in CAD drawings.
  - Achieved new SOTA on FloorPlanCAD benchmark with **91.1 PQ**, outperforming previous methods by significant margins (9.6 – 21.2 PQ).
- [NeurIPS 2025] **ArchCAD-400K: A Large-Scale CAD Drawings Dataset and New Baseline for Panoptic Symbol Spotting**
- [ICLR 2026] **InternSpatial: A Comprehensive Dataset for Spatial Reasoning in Vision–Language Models**
- [Technical Report] **InternVL3.5: Advancing Open-Source Multimodal Models in Versatility, Reasoning, and Efficiency**
- [Technical Report] **InternVL3: Exploring Advanced Training and Test-Time Recipes for Open-Source Multimodal Models**

## 🏆 Honors & Awards

- **Outstanding Graduate Award**, Nanjing University 2025
- **Outstanding Student Award**, Nanjing University 2023
- **China Merchants Bank Scholarship**, Nanjing University 2022 – 2023
- **Ruli Scholarship**, Nanjing University 2021 – 2022