

Zihao Jing

Western University

Ontario, Canada — zjing29@uwo.ca — Homepage — Github — Google Scholar

I am a second-year Master’s student in the Department of Computer Science at Western University, supervised by Prof. Pingzhao Hu. My research interests focus on Large Language Models (LLMs) and multimodal reasoning, with a focus on improving LLM generalization and reliability for socially impactful domains and AI-for-Science. More broadly, I am open to other advanced LLM research across agent, efficient training, fairness, safety, and etc.

EDUCATION

Western University, M.Sc. in Computer Science, supervised by Prof. Pingzhao Hu

Sep. 2024–Present

Beihang University, B.Eng. in Software Engineering, Beijing, China

Sep. 2020–Aug. 2024

SELECTED PUBLICATIONS

Top-Tier Conferences

[1] **Zihao Jing**, Qiu hao Zeng, Ruiyi Fang, Yan Sun, Boyu Wang, Pingzhao Hu. *Entropy-Guided Dynamic Tokens for Graph-LLM Alignment in Molecular Understanding*. In **ICLR 2026**. [Code] [Paper]

- We propose EDT-Former with entropy-guided dynamic query to solve *structural information loss* in graph-LLM alignment, where fixed-length Q-Former queries collapse substructure context for variable-size entities.
- We propose a connector-only frozen-backbone alignment pipeline via a *Dynamic Query Transformer* to solve heavy LLM fine-tuning cost/overfitting in coarse multimodal fusion, enabling efficient training with the LLM frozen.
- Applied and achieves SOTA results on MoleculeQA and molecule-oriented Mol-Instructions benchmarks, and best accuracy on 10/10 MoleculeNet/TDC property benchmarks under matched settings.

[2] **Zihao Jing**, Yan Sun, Yan Yi Li, Sugitha Janarthanan, Alana Deng, Pingzhao Hu. *Structure-Aware Fusion with Progressive Injection for Multimodal Molecular Representation Learning*. In **NeurIPS 2025**. [Code] [Paper]

- We propose *MuMo*, a foundation-style multimodal fusion model to solve brittle token–structure alignment in multimodal representation learning, where unreliable 3D inputs make naive 1D–3D fusion unstable;
- We propose *asymmetric Progressive Injection* to solve modality collapse and feature domination under early fusion, progressively injecting the structural prior into the token stream while preserving modality-specific modeling; this enables stable training and scalable long-context dependency propagation with an efficient sequence backbone.
- Ranks 1st on 22/29 downstream tasks with a +2.7% average improvement over strong baselines.

[3] **Zihao Jing**, Qiu hao Zeng, Ruiyi Fang, Yan Yi Li, Yan Sun, Boyu Wang, Pingzhao Hu. *Scaling-Aware Adapter for Structure-Grounded LLM Reasoning*. **Under review**. [Code] [Paper]

- We propose *Scaling-Aware Patching* to solve fixed-length connector bottlenecks when grounding LLMs on variable-size spatial graphs, where constant query budgets over-compress large entities and over-allocate small ones; an instruction-conditioned gate selects anchors and grows soft patches so the structural token budget scales with input complexity.
- We propose a *Geometry Grounding Adapter* to solve structural hallucination in structure-conditioned generation, where sequence-style encodings lack verifiable spatial evidence;
- Achieves the best overall performance on 3 benchmark groups across both all-atom modality reasoning tasks.

Other Highlights

[4] Junqin Huang, Zhongjie Hu, **Zihao Jing**, Mengya Gao, Yichao Wu. *Piccolo2: General Text Embedding with Multi-Task Hybrid Loss Training*. **SenseTime Technical Report, 2024** [SenseNova] [Code] [Report]

- General-purpose text embedding model trained with an efficient multi-task hybrid loss across diverse tasks.
- Achieves top-1 performance on the CMTEB ranking in 2024.05, with scalable embedding dimensions via MRL training.

[5] **Zihao Jing**, Yuxi Long, Ganlin Feng. *Pruning for Generalization: A Transfer-Oriented Spatiotemporal Graph Framework*. **Under review at Workshop**. [Code] [Preprint]

- Transfer-oriented traffic forecasting framework to improve cross-city generalization under limited data.
- Entropy-based graph pruning removes boundary nodes dominated by external factors, improving MAE by over 13%.

[6] Yan Sun, Lianghong Chen, **Zihao Jing**, Yan Yi Li, Dongkyu Kim, et al. *Generative AI for the Design of Molecules: Advances and Challenges*. In **Journal of Chemical Information and Modeling**. [Paper]

- Comprehensive review of generative AI for biomolecular design, including diffusion and optimization-based approaches.
- Highlights key challenges such as multimodal integration, data scarcity, and autonomous AI for drug discovery.

RESEARCH INTERNSHIPS

SenseTime

LLM Research Intern

2023.09–2024.06

- **Text embeddings (Piccolo2)**: Trained general-purpose embedding models with multi-task hybrid-loss objectives; built end-to-end training/evaluation pipelines and led iterative optimization of a generative embedding LLM; achieved top-1 ranking on C-MTEB (May 2024).
- **Domain LLM adaptation (100B)**: Fine-tuned a 100B-parameter LLM for vertical livestream marketing; drove data/recipe iteration and productionized the model for deployment at Sina Weibo.

- **LLM research, engineering & scaling:** Gained hands-on experience with large-scale pretraining/fine-tuning codebases (SenseNova series), hyperparameter tuning, experiment tracking, and reproducible training workflows on multi-GPU infrastructure.

Jina AI

AI Research Intern

2023.04–2023.09

- **LLM engineering:** Improved RunGPT interface and contributed solutions to the **Llama** open-source ecosystem.
- **Applied LLM analytics:** Implemented LLM-based denoising and sentiment pipeline for Budweiser public-opinion analysis; reduced operational cost by >13%.
- **Model commercialization:** Led evaluation/tuning of a super-resolution model; executed performance testing and produced pricing recommendations.

HONORS AND AWARDS

- | | |
|---|------|
| • Silver Prize , Feng Ru Cup Science and Innovation Competition (University-level) | 2023 |
| • Third Prize , 13th National Undergraduate Mathematics Competition (National-level) | 2021 |
| • Third Prize , 32nd Beijing Undergraduate Mathematics Competition (Province-level) | 2021 |
| • Outstanding Student Leader Award (University-level) | 2021 |
| • Third Prize , Physics Academic Competition (University-level) | 2021 |
| • Honorable Mention (H Prize) , Mathematical Contest In Modeling | 2020 |

ACADEMIC SERVICE

Conference Reviewer ICLR 2026 (5 papers), ICML 2026 (6 papers)

Workshop Reviewer Time Series in the Age of Large Models (TSALM), ICLR 2026

Journal Reviewer ACM Transactions on Knowledge Discovery from Data (TKDD)

TECHNICAL SKILLS

Skills narrative. Trained in Software Engineering, I built a broad programming foundation, then transitioned to full-stack development. Industry projects, spanning production management systems and the Shanghai Highway (G60) AI monitor&prediction system, strengthened my engineering practice, service maintenance, and high-throughput deployment, including containerized workflows and Kubernetes.

In my third undergraduate year, I shifted toward AI research through an internship at Jina AI, developing applied models (e.g., segmentation and sentiment analysis) and consolidating modern ML engineering skills (NumPy, PyTorch, Transformers) and professional collaborative practices.

I then joined SenseTime as an intern, where I worked on LLM training (SenseNova series), building practical experience in pretraining and fine-tuning, codebase organization, and hyperparameter optimization.

During my M.Sc. in CS, Western University, I have focused on advanced LLM for spatial geometry understanding, further deepening my large-scale training stack (multi-GPU/multi-node), HPC scheduling, and reproducible container-based experimentation. My technical skills are summarized as follows:

- **Large Language Models.** Pretraining and fine-tuning; multimodal fine-tuning/alignment (e.g., EDT-Former, Cuttlefish); training/evaluation pipelines.
- **HPC & Systems.** Linux administration; Slurm; containerization (Docker, Singularity); parallel computing.
- **Distributed Training.** PyTorch DDP/FSDP; multi-GPU/multi-node training; DeepSpeed; profiling/monitoring.
- **Full-Stack Engineering.** C/C++, Java, Python, MATLAB; SQL; Linux shell; HTML5/CSS, JavaScript; front-end/back-end development (websites, applications, mini-programs).
- **Collaboration.** Git; clean/modular code; reproducible workflows (environment/version pinning); documentation.

MENTORSHIP

Teaching Assistant, COMPSCI 3305 Operating System	26 Winter
---	-----------

Teaching Assistant, COMPSCI 2211 Software Tools and Systems Programming	25 Fall
---	---------

Teaching Assistant, COMPSCI 3305 Operating System	25 Winter
---	-----------

Teaching Assistant, COMPSCI 2211 Software Tools and Systems Programming	24 Fall
---	---------

- I enjoy mentoring junior students, inspired by the mentorship I received from my advisor and senior peers, and find strong fulfillment in teaching and guiding others.