

Xianbang Wang

(+1) 617-251-7258 | xbwang@mit.edu | peppaking8.github.io | [GitHub](#) | [X](#) | [LinkedIn](#)
IMO 2024 Gold Medal | Putnam Top 17 | Lead author of MiniT2I (nearly 500 GitHub stars)
Mathematics | *Machine Learning* | *Training Systems (JAX/TPU)*

EDUCATION

Massachusetts Institute of Technology
S.B. candidate, Mathematics with Computer Science (18C)
→ Expected graduation: May 2027

Sep 2025 – Present
GPA: 5.0/5.0

IIS (Yao Class), Tsinghua University
Pre-college Program in Computer Science

Sep 2023 – Jul 2025
GPA: 4.0/4.0

HONORS

Gold Medal, International Mathematical Olympiad (IMO) 2024 — one of six members of the China national team
Top 17, William Lowell Putnam Mathematical Competition (86th, 2025)
Silver Medal ×2, China National Olympiad in Informatics (NOI), 2021 & 2022 | 17th, USACO Platinum Open 2022

RESEARCH & ENGINEERING EXPERIENCE

MIT — Undergraduate Researcher, He Vision Group
Advisor: Prof. Kaiming He

Feb 2025 – Present
Cambridge, MA

- **MiniT2I — Project lead & first author.** Built a text-to-image foundation model from scratch: data pipeline, training recipe, multi-host JAX/TPU training, evaluation, and open-source release (**nearly 500 combined GitHub stars** across the official JAX/PyTorch repos; community ports to ComfyUI and tinygrad).
- **Training infrastructure.** My training codebase became the group's shared foundation: FSDP/HSDP sharding over pjit/GSPMD, a TPU orchestrator that resumes through preemptions and dead nodes, and Pallas attention kernels delivering **1.5× end-to-end nanoGPT speedup**.
- Through systematic ablations, the **258M-parameter B/16 model reaches 0.87 GenEval / 84.2 DPG-Bench** at 512×512, **on par with ~2B-parameter models such as SD3-Medium**; released models scale to 912M (L/16: 0.883 / 85.9), and internal runs up to **2B parameters** showed continued scaling gains.
- **BiFlow — Co-first author; CVPR 2026 Highlight.** Co-developed normalizing flows with a learned reverse map, removing the analytic-inverse constraint; the paper reports up to **100× faster sampling** than its causal-decoding counterpart.

Adobe Research — Research Intern, NextCam Group

Jun – Aug 2026

- Studied scaling of DMD distillation for image/video diffusion: with the student fixed, **scaling the teacher and critic together at matched size** stabilizes training with consistent gains, while scaling either alone hurts; conditioning the critic on student timesteps improves multi-step DMD. **Submitted to ICLR 2027.**

JD.com — Research Intern, LLM Post-training

Jan – Feb 2026

- Built the post-training pipeline for an on-device companion LLM in one week (DPO on pairwise preferences, SFT for persona alignment); wrote onboarding docs for later interns.

MIT — Undergraduate Researcher, HAN Lab
Advisor: Prof. Song Han

Sep – Dec 2025
Cambridge, MA

- Integrated **SANA-Video (INT4/SVDQuant)** and **LongLive (FP8)** into **Nunchaku** with quality/latency benchmarks; coauthor of both ICLR 2026 papers (SANA-Video: Oral).

PUBLICATIONS & RELEASES

5 peer-reviewed papers at CVPR / ICML / ICLR 2026 | 450+ Google Scholar citations

- *MiniT2I: A Minimalist Baseline for Text-to-Image Generation.* **Technical report, 2026.** [[blog](#) / [code](#)]
X. Wang, H. Zhao, Y. Lu, K. Zhou, L. Ma, K. He.
- *Bidirectional Normalizing Flow: From Data to Noise and Back.* **CVPR 2026 Highlight.** [[paper](#)]
Y. Lu*, Q. Sun*, X. Wang*, Z. Jiang, H. Zhao, K. He.
- *One-step Latent-free Image Generation with Pixel Mean Flows.* **ICML 2026.** [[paper](#)]
Y. Lu*, S. Lu*, Q. Sun*, H. Zhao*, Z. Jiang, X. Wang, T. Li, Z. Geng, K. He.
- *SANA-Video: Efficient Video Generation with Block Linear Diffusion Transformer.* **ICLR 2026 Oral.** [[paper](#)]
Junsong Chen*, Yuyang Zhao*, Jincheng Yu*, R. Chu, Junyu Chen, S. Yang, X. Wang, Y. Pan, D. Zhou, H. Ling, H. Liu, H. Yi, H. Zhang, M. Li, Y. Chen, H. Cai, S. Fidler, P. Luo, S. Han, E. Xie.
- *LongLive: Real-time Interactive Long Video Generation.* **ICLR 2026.** [[paper](#)]
S. Yang, W. Huang, R. Chu, Y. Xiao, Y. Zhao, X. Wang, M. Li, E. Xie, Yingcong Chen, Y. Lu, S. Han, Yukang Chen.
- *H³DP: Triply-Hierarchical Diffusion Policy for Visuomotor Learning.* **ICLR 2026.** [[paper](#)]
Y. Lu*, Y. Tian*, Z. Yuan*, X. Wang, P. Hua, Z. Xue, H. Xu.

* Equal contribution.

TALKS & SKILLS

Invited talk, Stanford Vision & Learning Lab, Aug 2026: *Generative Modeling in Pixel Space: Why Remove Autoencoders?*
Skills: Python, C++, CUDA | JAX/Flax, PyTorch, Pallas | Multi-host TPU (v5p/v6e), FSDP/HSDP, GSPMD