Meilong Xu

+1-631-615-9328 | meixu@cs.stonybrook.edu | https://melon-xu.github.io/ | Google Scholar | LinkedIn

Research Summary

CS Ph.D. candidate at Stony Brook University specializing in **LLM/VLM post-training**, **generative models**, and **agentic multimodal systems**. Industry experience includes physics-grounded video generation at **Amazon Inc.** and distributed multimodal model training with Megatron-LM for production video moderation at **TikTok Inc.** Publications include papers at **CVPR**, **NeurIPS**, **ICLR**, and **ECCV**. Expected graduation: **May 2027**.

Experience

- **Amazon Inc.** – Prime Video *May 2026 - Present*
  Applied Scientist Intern Seattle, WA
  - Developed **VISTA**, an **agentic framework for physics-grounded video generation** that compiles prompts into executable physics programs and conditions frozen video-diffusion models on simulated trajectories via depth control. Improved aggregate scores on **Physics-IQ** and **PhyGenBench** by **8.9%** and **3.3%**, respectively, relative to the strongest evaluated baseline on each benchmark.
  - Designed a **three-stage verification process** for semantic consistency, simulation behavior, and final states, with **failure attribution and bounded self-repair**; integrated scene and event parsing, physics compilation, and capability routing to validate physical plans before video generation.
  - Reduced **video-generation calls by 70.6%** versus a generate–evaluate–replan baseline on Physics-IQ by completing iterative verification and repair in simulation space before RGB synthesis.

- **TikTok Inc.** – Data-TnS-Algo *May 2025 - Sep. 2025*
  Research Scientist Intern Bellevue, WA
  - Developed **RB-FT**, a **two-stage VLM post-training method** that first fine-tunes models on self-generated rationales and then on classification labels. Improved classification accuracy over direct SFT by **6.6% on SmartHome-LLM** and **4.1% on MultiHateClip**; **first-author preprint**.
  - Contributed to **continued pre-training** of a multimodal foundation model (Falcon-7B + SigLIP) for video moderation using **Megatron-LM** on multi-GPU clusters. The moderation workflow combined chain-of-thought prompting, single-token classification, and model ensembling. The model's integration into the product line saved **tens of millions of dollars**.
  - Designed a **multitask multimodal supervision pipeline** that generates captioning, OCR, VQA, and grounding supervision from public and proprietary video, supporting pre-training and task-specific adaptation of the deployed moderation model.

- **Stony Brook University** *Sep. 2022 - Present*
  Graduate Research Assistant Stony Brook, NY
  - Developed **diffusion models** for cell and tissue layout generation (**CVPR 2025 Oral**); post-trained **VLMs using SFT and RL** for biomedical image reasoning.
  - Designed **topology-aware losses** using persistent homology to preserve anatomical structure in semi-supervised segmentation; co-first-author paper at **NeurIPS 2025** and first-author paper at **ECCV 2024**.

Education

- **Stony Brook University, Advisor: Chao Chen** *Sep. 2022 - Present*
  Ph.D. Candidate in Computer Science Stony Brook, NY
  - GPA: 3.88/4.00; expected graduation: **May 2027**

- **Wuhan University, Advisor: Jiayi Ma** *Sep. 2018 - Jun. 2022*
  Bachelor of Engineering in Electronic Information Engineering Wuhan, China
  - GPA: 3.90/4.00; rank: top 1%

Publications

C=Conference, J=Journal, W=Workshop; *Equal contribution

[C.1] **Meilong Xu**, Saumya Gupta, Xiaoling Hu, Chen Li, Shahira Abousamra, Dimitris Samaras, Prateek Prasanna, Chao Chen. **TopoCellGen: Generating Histopathology Cell Topology with a Diffusion Model**. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**)*, 2025 **(Oral Presentation, 0.7% acceptance rate)**.

[C.2] **Meilong Xu***, Xiaoling Hu*, Shahira Abousamra, Chen Li, Chao Chen. **MATCH: Multi-faceted Adaptive Topo-Consistency for Semi-Supervised Histopathology Segmentation**. *The Thirty-Ninth Annual Conference on Neural Information Processing Systems (**NeurIPS**)*, 2025.

[C.3] **Meilong Xu**, Xiaoling Hu, Saumya Gupta, Shahira Abousamra, Chao Chen. **Semi-supervised Segmentation of Histopathology Images with Noise-Aware Topological Consistency**. *European Conference on Computer Vision (**ECCV**)*, 2024.

[C.4] Jiamu Bai, Xin Yu, **Meilong Xu**, Weitao Lu, Xin Pan, Kiwan Maeng, Daniel Kifer, Jian Wang, Yu Wang. **Towards Better Optimization For Listwise Preference in Diffusion Models**. *International Conference on Learning Representations (**ICLR**)*, 2026.

[C.5] Chen Li, **Meilong Xu**, Xiaoling Hu, Weimin Lyu, Chao Chen. **Multi-Channel Uncertainty-Weighted Score Matching for Conditional Diffusion in Medical UDA**. *European Conference on Computer Vision (**ECCV**)*, 2026.

[J.1] Wentao Huang, **Meilong Xu**, Xiaoling Hu, Shahira Abousamra, Aniruddha Ganguly, Saarthak Kapse, Alisa Yurovsky, Prateek Prasanna, Tahsin Kurc, Joel Saltz, Michael L. Miller, Chao Chen. **RankByGene: Gene-Guided Histopathology Representation Learning Through Cross-Modal Ranking Consistency**. *IEEE Transactions on Medical Imaging (**TMI**)*, 2026.

[J.2] **Meilong Xu**, Linfeng Tang, Hao Zhang, Jiayi Ma. **Infrared and visible image fusion via parallel scene and texture learning**. *Pattern Recognition*, 2022.

[W.1] **Meilong Xu**, Nate Anderson, Richard M. Levenson, Prateek Prasanna, Chao Chen. **A Topological Comparison of the Fluorescence Imitating Brightfield Imaging and H&E Imaging**. *International Workshop on Topology- and Graph-Informed Imaging Informatics (**TGI**^{3} **@ MICCAI**)*, 2024.

Preprints

[1] **Meilong Xu**, Di Fu, Jiaxing Zhang, Gong Yu, Jiayu Zheng, Xiaoling Hu, Dongdi Zhao, Feiyang Li, Chao Chen, Yong Cao.
RB-FT: Rationale-Bootstrapped Fine-Tuning for Video Classification. *arXiv preprint*, 2025.

[2] Qingqiao Hu, Weimin Lyu, **Meilong Xu**, Kehan Qi, Xiaoling Hu, Saumya Gupta, Jiawei Zhou, Chao Chen.
LoC-Path: Learning to Compress for Pathology Multimodal Large Language Models. *arXiv preprint*, 2025.

[3] **Meilong Xu**, Qingqiao Hu, Xiaoling Hu, Shahira Abousamra, Xin Yu, Weimin Lyu, Kehan Qi, Dimitris Samaras, Chao Chen.
Topo-R1: Detecting Topological Anomalies via Vision-Language Models. *arXiv preprint*, 2026.

Skills

- **Programming Languages:** Python, C, C++
- **Frameworks & Libraries:** PyTorch, Hugging Face, Megatron-LM, DeepSpeed
- **LLM/VLM Post-Training:** Supervised Fine-Tuning (SFT), Reinforcement Learning (RLHF, GRPO), Direct Preference Optimization (DPO), On-Policy Distillation, Quantization, LoRA/PEFT
- **Agentic Systems:** Multi-agent orchestration, tool use and function calling, chain-of-thought planning, verification and bounded self-repair loops
- **Distributed Training & Inference:** DDP, vLLM, Slurm, AWS (EC2, S3)
- **Mathematical & Statistical Tools:** Topological Data Analysis, Persistent Homology

Academic Service

Conference Reviewer: CVPR, NeurIPS, ICLR, ICML, ICCV, ECCV, ACL, EMNLP.
Journal Reviewer: TPAMI, TMI, TNNLS.

Honors and Awards

- **NeurIPS 2025 Scholar Award** *Oct. 2025*
- **National Scholarship**, Wuhan University (Top 1% of undergraduates nationwide) *2019-2020*
- **First Class Scholarship**, Wuhan University (Top 6% of undergraduates) *2018-2021*