

# Fangzhou Mu (穆方舟)

Email: [fzmu727@gmail.com](mailto:fzmu727@gmail.com) | Mobile: +1 (608) 332-9265 | [Personal webpage](#) | [Google scholar](#) | [GitHub](#) | [LinkedIn](#)

## SUMMARY

**Research Interests:** Multi-modal Foundation Models, Computer Vision, Machine Learning

- Strong knowledge in multi-modal LLMs, diffusion models, 3D vision, computational photography, and video understanding.
- Extensive experience in the design, implementation, training, and performance optimization of deep neural networks.
- Publication in top-tier vision and ML conferences (CVPR x6, ICLR x2, ICCV x2, NeurIPS x1) and journals (TPAMI x1).

## EDUCATION

**University of Wisconsin-Madison**

**Madison, WI, USA**

Ph.D. in Computer Sciences | Advisor: Prof. [Yin Li](#)

Sept. 2018 – Dec. 2023

M.S. in Computer Sciences & M.S. in Pharmaceutical Sciences

Sept. 2014 – Jun. 2018

**Awarded University of Wisconsin Distinguished Graduate Fellowship**

**Zhejiang University (浙江大学)**

**Hangzhou, China**

B.S. in Biological Sciences (GPA: 3.86/4.0, Ranking: 1/33)

Sept. 2010 – Jun. 2014

**Awarded National Scholarship (top 3%)**

## WORK EXPERIENCE

**Meta**

**Menlo Park, CA, USA**

Research Scientist (Meta Superintelligence Labs)

May 2025 – Present

**Building Native Video Generation Models**

- Implement data pipelines for text-to-video (T2V), image-to-video (I2V), and video-to-video (V2V) generation tasks.
- Develop novel methods for product-driven video edit categorization, data synthesis, and retrieval from billion-scale data.
- Optimize workflows for large-scale video quality filtering, video captioning, and synthetic data generation.
- Collaborate closely with cross-functional teams to drive progress in a rapidly evolving project landscape.

**NVIDIA**

**Santa Clara, CA, USA**

Senior Deep Learning Algorithms Engineer

Jan. 2024 – May 2025

**Performance optimization and cloud deployment of Multi-modal Large Language Models (MLLMs)**

- Analyze and optimize the inference-time performance of MLLMs on bleeding-edge NVIDIA data center GPUs.
- Design and implement production-quality software for MLLM serving as part of NVIDIA Inference Microservices (NIM).
- Introduce key features to leading MLLM optimization and serving frameworks (vLLM, TensorRT-LLM) for Day0 model release.
- Work with Partner and cross-functional teams to deliver commercial NIM containers for frontier MLLMs at speed of light.  
(Meta: [Llama3.2](#), [Llama4](#), Google: [PaliGemma](#), [Gemma3](#), [DePlot](#); Microsoft: [Phi-3.5-vision](#), [Phi-4-multimodal](#); Mistral: Pixtral, Mistral-Small-3.1)

**University of Wisconsin-Madison**

**Madison, WI, USA**

Graduate Research Assistant (Advisor: Prof. [Yin Li](#))

Sept. 2018 – Dec. 2023

**Thread 1: 3D vision with single-photon LiDAR sensors** (with Prof. [Andreas Velten](#) and Prof. [Mohit Gupta](#))

- Invented NeRF-based deep models for high-speed imaging beyond the line of sight (ICCP/TPAMI 22).
- Developed learned spatiotemporal compression of SPAD\* histograms for 3D imaging (ICCV 23, US patent filed).
- Developed 3D reconstruction method with low-cost SPAD sensors and differentiable transient rendering (CVPR 24).  
(\* SPAD: single-photon avalanche diode, an emerging time-of-flight sensor for 3D imaging with single-photon light sensitivity and picosecond-scale time resolution.)

**Thread 2: Scalable and resource-constrained video understanding**

- Proposed scalable Transformer design and efficient training techniques for action detection in long videos (CVPR 24).
- Won top prizes in prestigious video action detection challenges (1<sup>st</sup> prize: EPIC-Kitchens 2023, 2<sup>nd</sup> prize: Ego4D 2022 and 2023).
- Developed a latency-aware scheduler for resource-adaptive video object detection on mobiles (CVPR 22).

**Thread 3: Controllable text-to-image generation with Diffusion Models**

- Proposed a training-free method for the spatial control of Stable Diffusion with any conditions (CVPR 24, NeurIPS 24).
- Introduced adapter-based fine-tuning of Stable Diffusion for grounded text-to-image generation (CVPR 23).

**Other work:** Studied linearized and multi-task fine-tuning for the adaptation of foundation models (ICLR 20, ICLR 24).

**InnoPeak Technology, Inc.**

**Palo Alto, CA, USA**

Research Intern (Manager: Dr. [Yanli Liu](#))

Jun. 2023 – Aug. 2023

**Personalization meets controllability in text-to-image diffusion**

- Develop methods for fast, one-shot personalization of Stable Diffusion without fine-tuning.
- Introduce object pose control and camera viewpoint control to personalized Stable Diffusion.

**Snap Inc.**

**New York, NY, USA**

Research Intern (Managers: Dr. [Jian Wang](#) and Dr. [Yicheng Wu](#))

May 2022 – Aug. 2022

## Image restoration for face captures in the wild

- Developed *StyleGAN inversion* using spatially varying latent codes for *improved facial identity preservation*.
- Invented an *image quality aware diffusion model* with state-of-the-art *face restoration* capability.

### Snap Inc.

Remote

Research Intern (Manager: Prof. [Shree Nayar](#))

May 2021 – Aug. 2021

### 3D Photo Stylization (CVPR 22 oral, mentored by Dr. [Jian Wang](#) and Dr. [Yicheng Wu](#))

- Invented a method for *novel view synthesis* from a single image *in an artistic style* for creative 3D photo browsing.
- Proposed a point cloud stylization approach to facilitate *multi-view consistency* of stylized images.
- Implemented *custom CUDA ops* in PyTorch for *differentiable 3D point cloud rasterization*.

### University of Wisconsin-Madison / Morgridge Institute of Research

Madison, WI, USA

Graduate Research Assistant (Advisor: Prof. [Anthony Gitter](#))

May. 2018 – Aug. 2018

### ML4BIO - Machine learning literacy for biologists

- Developed and delivered an *AI4Science workshop* to a broad audience (Bioinformatics/ISMB 22 and [Software Carpentries](#)).
- Built *open-source software with GUI* to support interactive learning of *sklearn ML models* (pip install ml4bio).

## PUBLICATION

- [1] Carter Sifferman, Yiquan Li, Yiming Li, [Fangzhou Mu](#), Michael Gleicher, Mohit Gupta, Yin Li. **Recovering Parametric Scenes from Very Few Time-of-Flight Pixels**. *ICCV 2025*
- [2] Tiantian Wang, Xinxin Zuo, [Fangzhou Mu](#), Jian Wang, Ming-Hsuan Yang. **Towards 4D Human Video Stylization**. *Computer Vision and Image Understanding 2025*
- [3] Kuan Heng Lin\*, Sicheng Mo\* (equal contribution), Ben Klingher, [Fangzhou Mu](#), Bolei Zhou. **Ctrl-X: Controlling Structure and Appearance for Text-To-Image Generation Without Guidance**. *NeurIPS 2024*
- [4] [Fangzhou Mu](#)\*, Carter Sifferman\* (equal contribution), Sacha Jungerman, Yiquan Li, Zhiyue Han, Michael Gleicher, Mohit Gupta, Yin Li. **Towards 3D Vision with Low-Cost Single-Photon Cameras**. *CVPR 2024*
- [5] [Fangzhou Mu](#)\*, Sicheng Mo\* (equal contribution), Yin Li. **SnAG: Scalable and Accurate Video Grounding**. *CVPR 2024*
- [6] Sicheng Mo\*, [Fangzhou Mu](#)\* (equal contribution), Kuan Heng Lin, Yanli Liu, Bochen Guan, Yin Li, Bolei Zhou. **FreeControl: Training-Free Spatial Control of Any Text-to-Image Diffusion Model with Any Condition**. *CVPR 2024*.
- [7] Zhuoyan Xu, Zhenmei Shi, Junyi Wei, [Fangzhou Mu](#), Yin Li, Yingyu Liang. **Towards Few-Shot Adaptation of Foundation Models via Multitask Finetuning**. *ICLR 2024*
- [8] Yimeng Dou, [Fangzhou Mu](#), Yin Li, Tomy Varghese. **Sensorless end-to-end freehand three-dimensional ultrasound reconstruction with physics guided deep learning**. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 2024*
- [9] Yimeng Dou, [Fangzhou Mu](#), Yin Li, Tomy Varghese. **Sensorless End-to-End Freehand Ultrasound with Physics-Inspired Network**. *IUS 2023*
- [10] Felipe Gutierrez-Barragan, [Fangzhou Mu](#), Andrei Ardelean, Atul Ingle, Claudio Bruschini, Edoardo Charbon, Mohit Gupta, Yin Li, Andreas Velten. **Learned Compressive Representations for Single-Photon 3D Imaging**. *ICCV 2023*
- [11] Yuheng Li, Haotian Liu, Qingyang Wu, [Fangzhou Mu](#), Jianwei Yang, Jianfeng Gao, Chunyuan Li, Yong Jae Lee. **GLIGEN: Open-set Grounded Text-to-image Generation**. *CVPR 2023*
- [12] [Fangzhou Mu](#), Sicheng Mo, Jiayong Peng, Xiaochun Liu, Ji Hyun Nam, Siddeshwar Raghavan, Andreas Velten, Yin Li. **Physics to the Rescue: Deep Non-line-of-sight Reconstruction for High-speed Imaging**. *ICCP/TPAMI 2022*
- [13] Chris S Magnano, [Fangzhou Mu](#), Rosemary S Russ, Milica Cvetkovic, Debora Treu, Anthony Gitter. **An Approachable, Flexible, and Practical Machine Learning Workshop for Biologists**. *ISMB/Bioinformatics 2022*
- [14] [Fangzhou Mu](#), Jian Wang<sup>†</sup>, Yicheng Wu<sup>†</sup>, Yin Li<sup>†</sup> (co-corresponding authors). **3D Photo Stylization: Learning to Generate Stylized Novel Views from a Single Image**. *CVPR 2022 (oral presentation)*
- [15] Ran Xu, [Fangzhou Mu](#), Jayoung Lee, Preeti Mukherjee, Somali Chatterji, Saurabh Bagchi, Yin Li. **SmartAdapt: Multi-branch Object Detection Framework for Videos on Mobiles**. *CVPR 2022*
- [16] Yin Li, Runyu L Greene, [Fangzhou Mu](#), Yu Hen Hu, Robert G Radwin. **Towards Video-based Automatic Lifting Load Prediction**. *HFES 2020*
- [17] [Fangzhou Mu](#), Yingyu Liang, Yin Li. **Gradients as Features for Deep Representation Learning**. *ICLR 2020*
- [18] Tingting Liang, Qi Zhao, Shan He, [Fangzhou Mu](#), Wei Deng, Bingnan Han. **Modeling Analysis of Potential Target of Dolastatin 16 by Computational Virtual Screening**. *Chemical and Pharmaceutical Bulletin 2018*
- [19] Robyn Umans, Hannah Henson, [Fangzhou Mu](#), Chaithanyarani Parupalli, Bensheng Ju, Jennifer Peters, Kevin Lanham, Jessica Plavicki, Michael Taylor. **CNS angiogenesis and barrierogenesis occur simultaneously**. *Developmental Biology 2017*

## AWARDS

### International Research Competitions

1<sup>st</sup> Prize, EPIC-Kitchens Challenges 2024 (Action Detection, Action Recognition, and Audio-based Interaction Detection tracks)

1<sup>st</sup> Prize, EPIC-Kitchens Challenges 2023 (Action Detection track)  
1<sup>st</sup> Prize, Ego4D Challenges 2024 (Moment Queries track)  
2<sup>nd</sup> Prize, Ego4D Challenges 2023 (Moment Queries track)  
2<sup>nd</sup> Prize, Ego4D Challenges 2022 (Moment Queries and Natural Language Queries tracks)

**International Conference Awards**

CVPR 2023 Outstanding Reviewer Award (among 232 of >7,000 reviewers)  
ICLR 2020 Student Travel Award

## PROFESSIONAL SERVICES

**Conference Area Chair:** CVPR 2026

**Conference Reviewer:**

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022, 2023, 2024, 2025)  
IEEE/CVF International Conference on Computer Vision (ICCV) (2023, 2025)  
European Conference on Computer Vision (ECCV) (2022, 2024)  
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (2022, 2023, 2024, 2025, 2026)  
Conference on Neural Information Processing Systems (NeurIPS) (2023, 2024, 2025)  
International Conference on Machine Learning (ICML) (2024, 2025)  
International Conference on Learning Representations (ICLR) (2024, 2025)  
AAAI Conference on Artificial Intelligence (AAAI) (2025)  
International Conference on Artificial Intelligence and Statistics (AISTATS) (2025)  
ACM International Conference on Multimedia (ACM MM) (2025)

**Journal Reviewer:**

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (Impact Factor: 20.8)  
International Journal of Computer Vision (IJCV) (Impact Factor: 11.6)  
IEEE Transactions on Image Processing (TIP) (Impact Factor: 10.8)  
IEEE Transactions on Neural Networks and Learning Systems (TNNLS) (Impact Factor: 10.2)  
IEEE Transactions on Computational Imaging (TCI) (Impact Factor: 4.2)  
The Visual Computers - International Journal of Computer Graphics (Impact Factor: 3.0)  
Ad Hoc Networks (Impact Factor: 4.4)