

MENGQI (AMBER) ZHANG

☎ 858-319-5828 ✉ mq.zhang7@outlook.com 🔗 mengqi-z-9a83341bb 🌐 mq-zhang1.github.io

RESEARCH INTERESTS

My research focuses on **robust and generalizable systems** by grounding generative and multimodal models in **physical structure and dynamics**. I develop **video-action models** with video prediction priors, and **diagnostic benchmarks** for world models' physical consistency and generalization.

EDUCATION

University of California, Irvine

Ph.D. Student in Computer Science | Advisor: Judy Hoffman

Jan. 2026 – Present

California, USA

Georgia Institute of Technology

Ph.D. Student in Computer Science | Advisor: Judy Hoffman

Aug. 2024 – Dec. 2025

Georgia, USA

University of California, San Diego

M.S. in Machine Learning | Advisors: Xiaolong Wang, Zhuowen Tu

Sept. 2021 – Jun. 2023

California, USA

Zhejiang University

B.E. in Biosystem Engineering | Advisor: Tao Lin

Sept. 2017 – Jun. 2021

Zhejiang, China

PUBLICATIONS

(*) denotes equal contribution

Mengqi Zhang, Sahil Khose, Simar Kareer, Yuchen Song, Unnat Jain, Judy Hoffman, "DeVA: Decoupled Video-Action Model with physical guidance for robot policy learning" **In submission** [[arXiv](#)] [[Project & Demo](#)]

Sahil Khose, **Mengqi Zhang**, Prithvijit Chattopadhyay, Judy Hoffman, "WFM-Eval: Interpretable Error Diagnostics for Video World Models in Robotics" **In submission** [[Paper](#)] [[Project](#)]

Mengqi Zhang, Sahil Khose, Fiona Ryan, Judy Hoffman, "SPOT: Structured Prompting with Object-centric Tokens for open-world scene graphs" **CVPR 2026**, MUSI Workshop [[OpenReview](#)]

Mengqi Zhang, Qi Li, Mehmet Saygin Seyfioglu, Karim Bouyarmane, "Style-Instructed Mask-Free Virtual Try-On" **CVPR 2026**, AI4CC Workshop (**Oral**) [[arXiv](#)] [[Project](#)]

Minghua Liu*, Chong Zeng*, Xinyue Wei, Ruoxi Shi, Linghao Chen, Chao Xu, **Mengqi Zhang**, Zhaoning Wang, Xiaoshuai Zhang, Isabella Liu, Hongzhi Wu, Hao Su. "MeshFormer: High-Quality Mesh Generation with 3D-Guided Reconstruction Model" **NeurIPS 2024** (**Oral**) [[arXiv](#)] [[Project](#)]

Mengqi Zhang*, Yang Fu*, Zheng Ding, Sifei Liu, Zhuowen Tu, Xiaolong Wang. "HOIDiffusion: Generating Realistic 3D Hand-Object Interaction Data" **CVPR 2024** [[arXiv](#)] [[Project](#)]

Zheng Ding*, **Mengqi Zhang***, Jiajun Wu, Zhuowen Tu. "Patched Denoising Diffusion Models For High-Resolution Image Synthesis" **ICLR 2024** [[arXiv](#)] [[Project](#)]

PROJECTS

Point2GS: Generalized Plug-and-Play 3D Gaussian Splatting

Mengqi Zhang*, Yang Fu*, Judy Hoffman, Sifei Liu, and Xiaolong Wang

Developed and evaluated a feed-forward 3D reconstruction pipeline combining point-cloud geometry, multi-view image features, and Gaussian prediction, with cross-dataset generalization tests.

EXPERIENCE

Amazon Inc.

WA, USA

Applied Research Scientist

May 2025 – Aug. 2025

- Contributed to the development and delivery of virtual try-on capabilities in **Amazon Nova Canvas**.
- Developed a **mask-free virtual try-on model** with natural-language control over garment types and wearing styles, using an **attention-guided training objective** to improve garment alignment; achieved **10–60% lower FID** than the strongest reported baselines across two benchmarks.

UC Irvine / Georgia Tech

CA / GA, USA

Research Assistant (Advisor: Judy Hoffman)

Aug. 2024 – Present

- Developed DeVA, a **decoupled VAM** transferring pretrained video models' visual and physical priors to robot policies; achieved **SOTA across three manipulation benchmarks** while processing **up to 20× fewer training examples** than the baseline.
- Co-developed **WFM-Eval**, a **diagnostic benchmark** decomposing video world-model quality into three error axes with object-level VLM grounding; benchmarked **five world models across two robot datasets**, showing model rankings reverse across datasets and that its hallucination metric—unlike FVD/FID—predicts **downstream policy success** on LIBERO.
- Built SPOT for spatially and semantically grounded **scene graphs** with open-source VLMs, enabling cross-domain and zero-shot 3D transfer.
- Developed Point2GS for **3D Gaussian reconstruction** from uncalibrated images and point clouds, generalizing zero-shot across object and scene datasets.

Hillbot Inc.

CA, USA

Machine Learning Engineer

Feb. 2024 – Jun. 2024

- Contributed to **text-to-3D generation** in MeshFormer by developing text-conditioned multi-view diffusion models for 3D object reconstruction.
- Built an **automated 3D human asset pipeline** in Blender to create diverse synthetic data for human avatar generation.

University of California, San Diego

CA, USA

Research Assistant

Dec. 2021 – Jul. 2024

Advisors: Xiaolong Wang, Zhuowen Tu, and Zhiting Hu

- Developed HOIDiffusion for **controllable hand-object interaction synthesis** via 3D-grounded diffusion, decoupling geometry and appearance to improve 6D object pose estimation.
- Proposed Patch-DM for **resolution-agnostic diffusion** and **high-resolution synthesis** beyond training resolution; achieved best FID among patch-based baselines on six datasets.
- Explored **fine-grained vision-language alignment** and controllable text generation through multi-instance learning and hierarchical reinforcement learning.

Institute of Network Technology, Chinese Academy of Sciences

Beijing, China

Machine Learning Engineer Intern

Jun. 2020 – Sept. 2020

- Built a **social media avatar classifier** with approximately 95% accuracy in anomaly detection.
- Developed and deployed **video-based violence detection** for real-time monitoring.

Academic Service

CVPR '25'26, NeurIPS '25'26, ICLR '26, ICCV '25, CoRL '26, ECCV '24, WACV '25

Honors & Awards (Selected)

Zhejiang University Outstanding Graduate Student Award	2021
1st Place (Advanced Group) , ASABE Robotics Student Design Competition	2020
Zhejiang Province Excellence Student Award	2020
1st Prize , Mathematics Competition of Chinese College Students	2019