

# YIMING XIE

· Phone: (+1)781-873-9028 · E-mail: ymxyimingxie@gmail.com · Website: <https://ymingxie.github.io>

## EDUCATION

---

### Northeastern University

Boston, MA, U.S.

- *Ph.D. Student, Khoury College of Computer Sciences*
- *Advisor: Huaizu Jiang*

*Sep. 2021 - Jun. 2026 (expected)*

### Zhejiang University

Hangzhou, China

- *Bachelor of Opto-Electronics Information Science and Engineering*
- *Advisor: Xiaowei Zhou*

*Sep. 2015 - June. 2019*

## WORK EXPERIENCE

---

### 3D Team, Stability AI

Remote

*Research Scientist Intern*

*Feb 2024 - Dec 2024*

Worked with Dr. Varun Jampani, Dr. Chun-Han Yao, and Dr. Vikram Voleti.

Worked on 4D generation.

### Reality Labs Research, Meta

Redmond, WA, U.S.

*Research Scientist Intern*

*May 2022 - Dec. 2022*

Worked with Dr. Julian Straub, Prof. Georgia Gkioxari, and Dr. Minh Phuoc Vo.

Worked on multi-view 3D object detection.

### State Key Laboratory of CAD&CG, Zhejiang University

Hangzhou, China

*Research Assistant*

*Mar. 2021 - Aug. 2021*

Worked with Prof. Xiaowei Zhou.

Worked on planar surface reconstruction.

### MIG-3D&AR Group, SenseTime

Hangzhou, China

*Research Intern*

*Nov. 2018 - Feb. 2021*

Worked with Jiaming Sun and Prof. Xiaowei Zhou.

Worked on 3D scene reconstruction, 3D object detection and tracking, and scene flow estimation.

## PREPRINTS

---

- Xiaogang Peng\*, **Yiming Xie\***, Zizhao Wu, Varun Jampani, Deqing Sun, Huaizu Jiang.  
HOI-Diff: Text-Driven Synthesis of 3D Human-Object Interactions using Diffusion Models.  
*In Submission*, 2024.  
<https://neu-vi.github.io/HOI-Diff/>

**Note:** \* above denotes equal contribution

## PUBLICATIONS

---

Google Scholar

### Conferences

- Zichong Meng, **Yiming Xie**, Xiaogang Peng, Zeyu Han, Huaizu Jiang.  
Rethinking Diffusion for Text-Driven Human Motion Generation.  
*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.  
<https://neu-vi.github.io/MARDM/>

- **Yiming Xie\***, Chun-Han Yao\*, Vikram Voleti, Huaizu Jiang, Varun Jampani.  
SV4D: Dynamic 3D Content Generation with Multi-Frame and Multi-View Consistency.  
*International Conference on Learning Representations (ICLR)*, 2025.  
<https://sv4d.github.io/>
- Lei Zhong, **Yiming Xie**, Varun Jampani, Deqing Sun, Huaizu Jiang.  
SMooDi: Stylized Motion Diffusion Model.  
*The European Conference on Computer Vision (ECCV)*, 2024.  
<https://neu-vi.github.io/SMooDi/>
- **Yiming Xie**, Varun Jampani, Lei Zhong, Deqing Sun, Huaizu Jiang.  
OmniControl: Control Any Joint at Any Time for Human Motion Generation.  
*International Conference on Learning Representations (ICLR)*, 2024.  
<https://neu-vi.github.io/omnicontrol/>
- **Yiming Xie**, Huaizu Jiang, Georgia Gkioxari, Julian Straub.  
Pixel-Aligned Recurrent Queries for Multi-View 3D Object Detection.  
*International Conference on Computer Vision (ICCV)*, 2023.  
<https://ymingxie.github.io/parq/>
- **Yiming Xie**, Matheus Gadelha, Fengting Yang, Xiaowei Zhou, Huaizu Jiang.  
PlanarRecon: Real-time 3D Plane Detection and Reconstruction from Posed Monocular Videos.  
*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.  
<https://neu-vi.github.io/planarrecon/>
- **Yiming Xie\***, Jiaming Sun\*, Siyu Zhang, Linghao Chen, Xiaowei Zhou.  
You Don't Only Look Once: Constructing Visuospatial Working Memory for Integrated 3D Object Detection and Tracking.  
*International Conference on Computer Vision (ICCV)*, 2021.  
(**Strong Accept \* 2 in final rating**)  
<https://zju3dv.github.io/udolo/>
- **Yiming Xie\***, Jiaming Sun\*, Linghao Chen, Xiaowei Zhou.  
NeuralRecon: Real-Time Coherent 3D Reconstruction from Monocular Video.  
*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.  
(**Oral and Best Paper Candidate**)  
<https://zju3dv.github.io/neuralrecon/>
- Jiaming Sun\*, Linghao Chen\*, **Yiming Xie**, Siyu Zhang, Qinhong Jiang, Xiaowei Zhou, Hujun Bao.  
Disp R-CNN: Stereo 3D Object Detection via Shape Prior Guided Instance Disparity Estimation.  
*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.  
<https://arxiv.org/abs/2004.03572>

**Note:** \* above denotes equal contribution

## Journals

- Xi Chen, Jiaming Sun, **Yiming Xie**, Hujun Bao, Xiaowei Zhou.  
NeuralRecon: Real-Time Coherent 3D Scene Reconstruction from Monocular Video.  
*IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2024.  
<https://ieeexplore.ieee.org/document/10508111>
- Linghao Chen\*, Jiaming Sun\*, **Yiming Xie**, Siyu Zhang, Qing Shuai, Qinhong Jiang, Guofeng Zhang, Hujun Bao, Xiaowei Zhou.  
Shape Prior Guided Instance Disparity Estimation for 3D Object Detection.  
*IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2021.  
<https://ieeexplore.ieee.org/document/9419782>
- Fangrui Zhu, **Yiming Xie**, Weidi Xie, Huaizu Jiang.  
Diagnosing Human-object Interaction Detectors.  
*International Journal of Computer Vision (IJCV)*, 2025.  
<https://neu-vi.github.io/Diag-HOI/>

## PROFESSIONAL SERVICE

---

- Conference Reviewer: CVPR, ICCV, ECCV, NeurIPS, ICLR, ICML, SIGGRAPH
- Journal Reviewer: TVCG, TCSVT, EAAI

## TEACHING EXPERIENCE

---

- Teaching assistant of CS 7150, Deep Learning, Northeastern University *Spring 2022/2023*

## TALKS

---

|   |                  |
|---|------------------|
| Dynamic 3D Content Generation. GAMES  | <i>Aug. 2024</i> |
| Dynamic 3D Content Generation with Multi-Frame and Multi-View Consistency. BAAI | <i>Aug. 2024</i> |
| NeuralRecon: Real-Time 3D Reconstruction from Monocular Video. CSIG-3DV         | <i>Aug. 2021</i> |
| Real-Time 3D Reconstruction from Monocular Video. AI Time                       | <i>Apr. 2021</i> |

## SKILLS

---

**Programming Languages:** C, C++, Python, LaTeX  
**Technical:** Pytorch, Git, Linux, Blender, Final cut pro

## HONORS AND AWARDS

---

|                           |             |
|---------------------------|-------------|
| Apple Scholars in AI/ML   | <i>2024</i> |
| CVPR Best Paper Candidate | <i>2021</i> |