YANG FU

yafu@ucsd.edu https://oasisyang.github.io Yang Fu

EDUCATION

Ph.D. in Electrical and Computer Engineering, University of California San Diego
Advisor: Prof. Xiaolong Wang

M.S. in Electrical and Computer Engineering, University of Illinois, Urbana-Champaign
Thesis: Towards Accurate Person Re-Identification by Deep Learning
Advisor: Prof. Thomas S. Huang and Prof. Humphrey Shi

B.S. in Electrical and Computer Engineering, Beihang University
Thesis: Detection of optic nerve sheath pulsation in ultrasound video

RESEARCH PROJECTS

Advancing Large MultiModal Models with 3D Representation

Data Curation & Benchmark Development: 1) Designed data pipelines to create the Open Spatial Dataset (OSD) by generating 8.7M spatial QA pairs from 3D scene graphs. 2) Additionally, developed the SpatialRGPT-Bench and SR-3D-Bench benchmarks to evaluate VLM spatial cognition.

Massive Transformer-based Architecture Design: 1) Developed the SplatDistill framework that distills 2D foundational vision features into 3DGS to enable better 3D understanding for extreme long videos. 2) Developed a unified multimodal transformer that integrates radar, LiDAR, camera, and mapping data (*i.e.*, surfel and 3D Gaussian Splatting), achieving state-of-the-art performance on the Waymo Open Dataset, and demonstrated clear scaling-law behavior through large-scale experiments. 3) Contributed to the SR-3D that masters kinds of spatial reasoning by embedding canonical 3D positional information into visual tokens across both single and multi-view inputs.

Post-Training Recipes: 1) Implemented multi-stage SFT recipes, including new modality(*i.e.*, depths, point clouds) alignment and instruct fine-tuning, by blending diverse datasets to enhance spatial reasoning. 2) Employed GRPO to further improve the spatial reasoning capability with region-based CoT reasoning.

3D Scene Reconstruction and Novel View Synthesis

COLMAP-Free 3D Reconstruction: Developed COLMAP-Free 3DGS, which simultaneously optimizes camera poses and reconstructs a scene, removing the dependency on SfM systems. This was extended to sparse-view scenarios using a Construct-Optimize approach.

Generalizable Feedforward Neural Fields: Proposed Neural-Prior to reconstruct scenes from sparse views by conditioning a neural field on retrieved priors. Also introduced MonoNeRF, a system that learns a generalizable NeRF from monocular videos without known camera poses.

Object Pose and Shape Understanding

Category-Level 6D Pose Estimation: Pioneered ”in-the-wild” category-level 6D pose estimation by creating the Wild6D dataset, developing a semi-supervised method to reduce annotation needs, and later a self-supervised method to eliminate dependency on synthetic data.

Generative 3D Hand-Object Interaction: Co-developed HOIDiffusion, a diffusion-based model for generating realistic and diverse 3D hand-object interaction data, complete with articulated hand poses and object shapes.

Scaling 3D Object Learning: Introduced Wild-RGBD, a large real-world dataset of 20k iPhone-captured RGB-D videos of 8.5k objects across 46 categories, with masks, metric-scale camera poses, and reconstructed point clouds—providing strong 3D supervision to facilitate frontier research in 3D foundation models.

CONFERENCE PUBLICATIONS AND PREPRINTS

1. **Yang Fu**, Sifei Liu, Hongxu Yin, Benjamin Eckar, Jan Kautz, Xiaolong Wang, Arash Vahdat, Chao Liu. Bridging 2D Vision Language Models to 3D World via Feature Distillation. *Under review*.
2. An-Chieh Cheng, **Yang Fu**, Yukang Chen, Zhijian Liu, Xiaolong Li, Subhashree Radhakrishnan, Jan Kautz, Song Han, Yao Lu, Hongxu Yin, Xiaolong Wang, Sifei Liu. 3D Aware Region Prompted Vision Language Model *Under review*.
3. Ri-Zhao Qiu*, Yafei Hu*, Ge Yang, Yuchen Song, **Yang Fu**, Jianglong Ye, Jiteng Mu, Ruihan Yang, Nikolay Atanasov, Sebastian Scherer, Xiaolong Wang. Learning Generalizable Feature Fields for Mobile Manipulation. *International Conference on Intelligent Robots and Systems*, 2025.
4. An-Chieh Cheng, Hongxu Yin, **Yang Fu**, Qiushan Guo, Ruihan Yang, Jan Kautz, Xiaolong Wang, Sifei Liu. SpatialRGPT: Grounded Spatial Reasoning in Vision Language Model. *Conference on Neural Information Processing Systems*, 2024.
5. Kaiwen Jiang, **Yang Fu**, Mukund Varma T, Yash Belhe, Xiaolong Wang, Hao Su, Ravi Ramamoorthi. A Construct-Optimize Approach to Sparse View Synthesis without Camera Pose. *ACM SIGGRAPH*, 2024.
6. **Yang Fu**, Sifei Liu, Amey Kulkarni, Jan Kautz, Alexei A. Efros, Xiaolong Wang. COLMAP-Free 3D Gaussian Splatting. **Highlight presentation** at *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.
7. Hongchi Xia*, **Yang Fu***, Sifei Liu, Xiaolong Wang. RGBD Objects in the Wild: Scaling Real-World 3D Object Learning from RGB-D Videos. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024. (* equal contribution)
8. Mengqi Zhang*, **Yang Fu***, Zheng Ding, Sifei Liu, Zhuowen Tu, Xiaolong Wang. HOIDiffusion: Generating Realistic 3D Hand-Object Interaction Data. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024. (* equal contribution)
9. **Yang Fu**, Shalini De Mello, Xueting Li, Amey Kulkarni, Jan Kautz, Xiaolong Wang, Sifei Liu. 3D Reconstruction with Generalizable Neural Fields using Scene Priors. *International Conference on Learning Representations*, 2024.
10. **Yang Fu**, Ishan Misra, Xiaolong Wang. MonoNeRF: Learning Generalizable NeRFs from Monocular Videos without Camera Poses. *International Conference on Machine Learning*, 2023.
11. Kaifeng Zhang, **Yang Fu**, Shubhankar Borse, Hong Cai, Fatih Porikli, Xiaolong Wang. Self-supervised Geometric Correspondence for Category-level 6D Object Pose Estimation in the Wild. *International Conference on Learning Representations*, 2023.
12. **Yang Fu**, Xiaolong Wang. Category-Level 6D Object Pose Estimation in the Wild: A Semi-Supervised Learning Approach and A New Dataset. *Conference on Neural Information Processing Systems*, 2022.
13. Yuzhe Qin, Yueh-Hua Wu, Shaowei Liu, Hanwen Jiang, Ruihan Yang, **Yang Fu**, Xiaolong Wang. DexMV: Imitation Learning for Dexterous Manipulation from Human Videos. *European Conference on Computer Vision*, 2022.
14. **Yang Fu**, Sifei Liu, Shalini De Mello, Umar Iqbal, Humphrey Shi, Jan Kautz. Learning to Track Instances without Video Annotations. **Oral presentation** at *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.
15. **Yang Fu**, Linjie Yang, Ding Liu, Thomas Huang, Humphrey Shi. Comprehensive Feature Aggregation for Video Instance Segmentation. *AAAI Conference on Artificial Intelligence*, 2021.
16. **Yang Fu**, Yunchao Wei, Guanshuo Wang, Yuqian Zhou, Honghui Shi, Thomas Huang. Self-similarity Grouping: A Simple Unsupervised Cross-Domain Adaptation Approach for Person Re-Identification. **Oral presentation** at *IEEE/CVF International Conference on Computer Vision*, 2019.
17. **Yang Fu**, Xiaoyang Wang, Yunchao Wei, Thomas S.Huang. STA: Spatial-Temporal Attention for Large-Scale Video-based Person Re-Identification. *AAAI Conference on Artificial Intelligence*, 2019.

18. **Yang Fu**, Yunchao Wei, Yuqian Zhou, Honghui Shi, Gao Huang, Xinchao Wang, Thomas.S Huang. Horizontal Pyramid Matching for Person Re-identification. *AAAI Conference on Artificial Intelligence*, 2019.

HONORS AND FELLOWSHIPS

Roblox Graduate Fellowship Finalist 2024
Qualcomm Innovation Fellowship 2023
UCSD Gradudate Fellowship 2021

WORKING EXPERIENCE

Research Intern, *Waymo*, USA June, 2025 - present

Research Intern (Part time), *Nvidia Research*, USA Sep, 2023 - Jun, 2025

Research Intern, *Nvidia Research*, USA June, 2022 - March, 2023

Research Assistant, *UC San Diego*, USA Sep, 2021 - present

Research Intern, *Nvidia Research*, USA May, 2020 - Dec, 2020

TEACHING EXPERIENCE

Teaching Assistant San Diego, C.A.

- ECE 285 March, 2023 - Jun, 2023
- ECE 176 Jan, 2024 - March, 2024 & Jan, 2025 - March, 2025

SERVICE

Manuscript Reviewer:

TIP, T-PAMI, CVPR, ICCV, ECCV, NeuRIPS, ICML, ICLR, MICCAI

Organizing Committee:

Workshop on Large-scale Video Object Segmentation Challenge, *CVPR* 2022
Workshop on Large-scale Video Object Segmentation Challenge, *CVPR* 2021.