

Fei Xia

👤 Staff Research Scientist @ Google DeepMind

✉ xf1280@gmail.com 🌐 <https://fxia22.github.io>

RESEARCH STATEMENT

My research interests lie in **Computer Vision** and **Robotics**. In particular, I am interested in machine learning for long-horizon robotics tasks, including reinforcement learning, learning from demonstration, and simulation to real world transfer. Most recently, I have been exploring using foundation models for robot decision making.

EDUCATION

Stanford University, Stanford, CA, USA 2016.9 - 2021.9
PhD, Department of Electrical Engineering, Advisor: Silvio Savarese and Leo Guibas
Tsinghua University, Beijing, China 2012.8 - 2016.7
Bachelor of Engineering, Department of Automation

Georgia Institute of Technology, Atlanta, GA, USA 2014.8 - 2014.12
Exchange Student in the School of Electrical and Computer Engineering

SELECTED PUBLICATIONS

See [my google scholar](#) for an up-to-date list.

(*indicates equal contributions)

- [1] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, **Fei Xia**, Ed Chi, Quoc V Le, Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. *NeurIPS 2022*.
- [2] Michael Ahn, ..., Vincent Vanhoucke, **Fei Xia** (co-corresponding author), ..., Andy Zeng (alphabetically listed), [Do As I Can, Not As I Say: Grounding Language in Robotic Affordances](#), *CoRL (Oral)*, 2022.
- [3] Danny Driess, **Fei Xia**, Mehdi SM Sajjadi, Corey Lynch, Aakanksha Chowdhery, Brian Ichter, Ayzaan Wahid, et al. [PaLM-E: An Embodied Multimodal Language Model](#). *ICML 2023*.
- [4] **Fei Xia**, Amir R Zamir, Zhiyang He, Alexander Sax, Jitendra Malik, Silvio Savarese. Gibson env: Real-world perception for embodied agents. *CVPR 2018*.
- [5] Wenlong Huang*, **Fei Xia***, Ted Xiao*, Harris Chan, Jacky Liang, Pete Florence, Andy Zeng, et al. [Inner Monologue: Embodied Reasoning through Planning with Language Models](#). *CoRL 2022*.
- [6] Anthony Brohan, Noah Brown, Justice Carbajal, Yevgen Chebotar, Joseph Dabis, Chelsea Finn, ..., **Fei Xia**, ... [RT-1: Robotics Transformer for Real-World Control at Scale](#). *RSS 2023*.
- [7] Anthony Brohan, Noah Brown, ..., **Fei Xia**, ... Brianna Zitkovich. [RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control](#). *CoRL 2023*.
- [8] Jacky Liang, Wenlong Huang, **Fei Xia**, Peng Xu, Karol Hausman, Brian Ichter, Pete Florence, Andy Zeng. [Code as policies: Language model programs for embodied control](#). *ICRA 2023*.
- [9] Google Team, Petar Georgiev, Vincent I Lei, Ryan Burnell, Liang Bai, Abhanshu Gulati, Greg Tanzer, ..., **Fei Xia**, ... Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. *arXiv 2024*.
- [10] Aditya Padalkar, Alex Pooley, Ajinkya Jain, Alex Bewley, Alex Herzog, Alex Irpan, ..., **Fei Xia**, ... Open x-embodiment: Robotic learning datasets and rt-x models. *arXiv 2023*.
- [11] Panos Achlioptas, Ahmed Abdelreheem, **Fei Xia**, Mohamed Elhoseiny, Leonidas Guibas. [ReferIt3D: Neural Listeners for Fine-Grained 3D Object Identification in Real-World Scenes](#). *ECCV 2020*.
- [12] Xinke Deng, Arsalan Mousavian, Yu Xiang, **Fei Xia**, Timothy Bretl, Dieter Fox. PoseRBPF: A rao-blackwellized particle filter for 6d object pose estimation. *RSS 2019*.
- [13] **Fei Xia**, William B Shen, Chengshu Li, Priya Kasimbeg, Martin E Tchapmi, Alexander Toshev, et al. Interactive Gibson Benchmark: A Benchmark for Interactive Navigation in Cluttered Environments. *RA-L 2020*.

- [14] Wenhao Yu, Nimrod Gileadi, Chuyuan Fu, Sean Kirmani, Kuang-Huei Lee, Montse Gonzalez Arenas, ..., **Fei Xia**. [Language to Rewards for Robotic Skill Synthesis](#). *CoRL 2023*.
- [15] Chengshu Li, **Fei Xia**, Roberto Martín-Martín, Michael Lingelbach, Sanjana Srivastava, Bokui Shen, Kent Vainio, et al. iGibson 2.0: Object-centric Simulation for Robot Learning of Everyday Household Tasks. *CoRL 2021*.
- [16] Boyuan Chen, **Fei Xia**, Brian Ichter, Kanishka Rao, Keerthana Gopalakrishnan, Michael S Ryoo, Austin Stone, et al. [Open-vocabulary Queryable Scene Representations for Real World Planning](#). *ICRA 2023*.
- [17] William B Shen, **Fei Xia**, Chengshu Li, Roberto Martín-Martín, Linus Fan, Gang Wang, et al. iGibson 1.0: A simulation environment for interactive tasks in large realistic scenes. *IROS 2021*.
- [18] Allen Z Ren, Anushri Dixit, Alexandra Bodrova, Sumeet Singh, Stephen Tu, Noah Brown, ..., **Fei Xia**, et al. [Robots That Ask For Help: Uncertainty Alignment for Large Language Model Planners](#). *CoRL 2023 (Best student paper award)*.
- [19] Chengshu Li, Ruohan Zhang, Jiankai Wong, Cem Gokmen, Sanjana Srivastava, Roberto Martín-Martín, ..., **Fei Xia**, et al. Behavior-1k: A benchmark for embodied ai with 1,000 everyday activities and realistic simulation. *CoRL 2023*.
- [20] Suvir Mirchandani, **Fei Xia**, Pete Florence, Brian Ichter, Danny Driess, Montserrat Gonzalez Arenas, Kanishka Rao, Dorsa Sadigh, Andy Zeng. Large Language Models as General Pattern Machines. *CoRL 2023*.
- [21] Tianhe Yu, Ted Xiao, Austin Stone, Jonathan Tompson, Anthony Brohan, Su Wang, Jaspiar Singh, Clayton Tan, Jodilyn Peralta, Brian Ichter, Karol Hausman, **Fei Xia**. [Scaling Robot Learning with Semantically Imagined Experience](#). *RSS 2023*.
- [22] **Fei Xia***, Zhiyang He*, Amir Zamir*, Sasha Sax, Jitendra Malik, Silvio Savarese. [Embodied Real-World Active Perception](#). *CVPR 2018 (spotlight Oral, Nvidia Pioneering Research Award)*.
- [23] Chengshu Li*, **Fei Xia***, Roberto Martín-Martín*, Or Litany, Alexander Toshev, Silvio Savarese. [ReLMoGen: Leveraging Motion Generation in Reinforcement Learning for Mobile Manipulation](#). *RSS workshop, ICRA 2021*.
- [24] Chengshu Li, **Fei Xia**, Roberto Martín-Martín, Silvio Savarese. [HRL4IN: Hierarchical Reinforcement Learning for Interactive Navigation with Mobile Manipulators](#). *CoRL 2019*.
- [25] Noriaki Hirose, Amir Sadeghian, **Fei Xia**, Silvio Savarese. [VUNet: Traversability Estimation via Dynamic Scene View Synthesis](#). *IEEE RA-L, and ICRA 2019*.
- [26] Noriaki Hirose, **Fei Xia**, Roberto Martín-Martín, Amir Sadeghian, Silvio Savarese, [Deep Visual MPC-Policy Learning for Navigation](#). *IEEE RA-L, and IROS 2019*.
- [27] Kevin Chen, Juan Pablo de Vicente, Gabriel Sepulveda, **Fei Xia**, Alvaro Soto, Marynel Vazquez, Silvio Savarese. [A Behavioral Approach to Visual Navigation with Graph Localization Networks](#). *RSS 2019*.
- [28] Martin J. Zhang, **Fei Xia**, James Zou. [AdaFDR: a Fast, Powerful and Covariate-Adaptive Approach to Multiple Hypothesis Testing](#). *Best Paper Award at RECOMB 2019*.
- [29] Martin J. Zhang, **Fei Xia**, James Zou. [Fast and covariate-adaptive method amplifies detection power in large-scale multiple hypothesis testing](#). *Nature Communications*.
- [30] Soheil Feizi, Changho Suh, **Fei Xia** and David Tse. [Understanding GANs: the LQG Setting](#).
- [31] **Fei Xia***, Martin Zhang*, James Zou, David Tse. [NeuralFDR: learning decision threshold from hypothesis features](#). *NIPS 2017*.
- [32] Qiao Liu, **Fei Xia**, Qijin Yin, Rui Jiang. [Chromatin accessibility prediction via a hybrid deep convolutional neural network](#). *Bioinformatics, 2017*.
- [33] Govinda Kamath*, Ilan Shomorony*, **Fei Xia***, Thomas Courtade, David Tse. [HINGE: Long-Read Assembly Achieves Optimal Repeat Resolution](#). *Genome Research Vol 27 2017*.
- [34] Ilan Shomorony, Govinda Kamath, **Fei Xia**, Thomas Courtade and David Tse, [Partial DNA Assembly: A Rate-Distortion Perspective](#). *ISIT 2016*.
- [35] Anastasia Dubrovina, **Fei Xia**, Panos Achlioptas, Mira Shalah, Leonidas Guibas. [Composite Shape Modeling via Latent Space Factorization](#). *ICCV 2019*.

AWARDS

- 2023** ICRA 2023 Outstanding Robot Learning Paper
- 2022** CoRL 2022 Special Innovation Award
- 2019** Qualcomm Innovation Fellowship (\$100k)
- 2019** Best Paper Award at RECOMB 2019.
- 2018** Nvidia Pioneering Research Award at CVPR 2018.
- 2016** Stanford Graduate Fellowship (Michael J. Flynn Fellow), Stanford University
- 2015** Chang Jiong Scholarship (Highest honor in Dept. of Automation, Tsinghua University, 1/560)
- 2014** Fang Chongzhi Scholarship (Highest honor in Dept. of Automation, Tsinghua University, 1/560)
- 2013** National Southwest Associated University Scholarship (1/560)

EXPERIENCES

Google DeepMind, Mountain View, CA, USA 2021.10 - Present
Robotics Team
Staff Research Scientist, Tech Lead Manager.

- I work on robot semantic planning in complex and unstructured environments. I co-lead SayCan effort [2], which is a system that uses Large Language Models (LLMs) for robot planning based on natural language human instructions. I also work on closed-loop planning with LLMs [5] and combining Vision-Language Models (VLMs) and LLMs for semantic scene understanding [16].
- I lead research efforts on using LLMs for robot planning and reasoning. Our recent work RT-2 [7] shows that we can train a single model that can both perceive and act, directly outputting robot actions from visual inputs. We also developed PaLM-E [3], an embodied multimodal language model that can perceive the world and reason about it.
- SayCan is reported by more than 260 media outlets including [TechCrunch](#), [WIRED](#), [Reuters](#), [Washington Post](#), [The Verge](#), and [CNET](#).

Stanford University, Stanford, CA, USA 2016.12 - 2021.9
Stanford Vision and Learning Group, Stanford AI Lab
Research Assistant, Advisors: **Prof. Silvio Savarese**, **Prof. Leo Guibas**

- My PhD thesis is on Large Scale Simulation for Embodied Perception and Robot Learning. My mission is to build intelligent embodied agents that can interact with complex and unstructured real-world environments, with applications to home robotics. I have been approaching this problem from 3 aspects: 1) Large scale and transferrable simulation for Robotics. 2) Learning algorithms for long-horizon tasks. 3) Combining geometric and semantic representation for environments.
- *I developed state-of-the-art large-scale simulation infrastructure that replicates the complexity of real world*
- Developed [Gibson Env](#), a robotics simulator for easy transfer to real-world. First robotics simulator that enables real-world perception. Used neural network to do real-time rendering for generating photo-realistic video stream.
- Implemented a pixel level domain adaptation mechanism to map real-world images and neural network generated images to a common space for transferring to real world.
- Gibson Env[22] was published in CVPR 2018 and won Nvidia Pioneering Research Award. Within a year, more than 10 publications in the robotics community successfully trained navigation policies in Gibson Environment and 3 publications transferred the policy to real robots. Up till now, Gibson Env has been cited more than 400 times.
- I am the lead developer of [iGibson](#) (Interactive Gibson), an interactive simulation environment of large scale virtualized realistic scenes for robot learning.
- *I developed learning algorithms that work with the simulation infrastructure*
- Used Gibson Environment for multiple robot learning projects, including Visual Trajectory following [26][25], Mobile Manipulation [23][24], Interactive Navigation [13], Topological Navigation [27] and demonstrated state of the art performance in those navigation tasks.

Google Inc. , Mountain View, CA, USA

2020.6 - 2021.1

Robotics at Google

Research Intern, Host: **Dr. Alexander Toshev, Dr. Brian Ichter**

Combining Classical Motion Planning and RL for Navigation and Coarse Manipulation.

- Part of the results are published in [23].

Co-organized **CVPR workshop on Embodied AI** and **iGibson Challenge** with as a Stanford-Google collaboration.

Nvidia Research, Seattle, WA, USA

2018.6-2018.9

Seattle Robotics Lab

Research Intern, Manager: **Prof. Dieter Fox**

Intuitive Physics Modelling for Real-World Object Interactions

- Developed methods to predict pose changes after physical interactions for real world objects.

Fast Rendering for Doing Render-and-Compare on Particles

- Developed a fast rendering engine that supports CUDA-OpenGL interoperation that enables rendering of a large set of images concurrently and doing rendering-and-compare with real images. Details in [12].

Stanford University, Stanford, CA, USA

2015.7 - 2016.12

Information Systems Laboratory, Department of Electrical Engineering

Research Assistant, Advisor: **Prof. David Tse**

- *Prior to robotics, I worked on Computational Biology and Statistical Machine Learning, which gives me an interdisciplinary background.*

HINGE: A de novo Sparse String Graph Assembler for PacBio Reads

- Generated **finished** assembly at accuracy 99.9% for *E.Coli* based on sparse string graph methods, with details in publication [33][34].
- Extended NSG(Not-So-Greedy) algorithm to a regime when triple repeats are all-bridged and interleaved repeats are bridged, i.e. information-theoretic bound for perfect assembly.

NeuralFDR: learning decision threshold from hypothesis features

- Proposed a learning based method for FDR control. Developed mirroring method for FDP prediction. NeuralFDR has provable performance in FDP control.
- Implemented the method and tested on RNASeq and GWAS datasets. Details can be found in paper [31].
- Our follow-up work of NeuralFDR, which replaced the neural network with a mixture of gaussian [29][28], showed better results and won Best Paper Award in RECOMB 2019.

TECHNICAL STRENGTHS

Deep Learning Software Stacks

Programming Languages

Additional Skills

PyTorch, Tensorflow, TF & PyTorch CUDA module development

Proficient with C/C++, Python, MATLAB, Java

Robot sensing and control, ROS, MPI, OpenMP, CUDA

last updated: 11/08/2024