Yilun Du

Massachusetts Institute of Technology
32 Vassar Street, 32-331
Cambridge MA 02139 USA

yilundu@mit.edu
https://yilundu.github.io/

Research Overview

My research is driven by the goal of developing AI agents that interact in the physical world, focusing on generative AI as a toolbox for solving such problems. A unique challenge in applying generative AI in this setting is the lack of available decision-making data and the necessity to generalize well to previously unseen situations. My work addresses this by proposing compositional generative modeling, where simple generative models are learned and composed together to construct complex generative models that generalize beyond the narrow amount of available data. My work has illustrated how such compositionality solves various problems in constructing AI agents ranging from complex scene understanding, trajectory planning, multimodal perception, and hierarchical planning. My work has further shown how such developed techniques can be broadly applied to settings in sciences and engineering.

Research Areas: Generative AI, Machine Learning, Computer Vision, Robotics, AI for Science.

Education

Ph.D. EECS 2020-2024	Massachusetts Institute of Technology (MIT) Advisors: Leslie Kaelbling, Tomás Lozano-Pérez, and Joshua B. Tenenbaum Thesis: *Learning to Generate Compositionally*
M.S. EECS 2019-2020	Massachusetts Institute of Technology (MIT) Advisors: Leslie Kaelbling, Tomás Lozano-Pérez, and Joshua B. Tenenbaum Thesis: *Implicit Learning with Energy-Based Models*
B.S. 2015-2019	Massachusetts Institute of Technology (MIT) EECS, GPA: 5.0 / 5.0

Employment

Graduate Researcher 2019-Now	Computer Science and Artificial Intelligence Lab, MIT LIS Group (PIs: Tomás Lozano-Pérez and Leslie Pack Kaelbling)
Graduate Researcher 2019-Now	Brain and Cognitive Science Department, MIT Computational Cognitive Science Lab (PI: Joshua Tenenbaum)
Research Advisor 2023-Now	Pika Labs Research Advisor on Text-to-Video Synthesis
Research Intern Summer 2023	Google Robotics Research Intern Working on Video Models for Robotics
Student Researcher 2022-2023	Google Brain Student Researcher Working on Video Models for Decision Making
Research Intern Summer 2022	Deepmind Research Intern Working on Compositional Generative Modeling

Visiting Researcher 2020-2022	Facebook AI Research Visiting Researcher on the FAIR Proteins (Generative Biology) Team
Research Intern Summer 2019	Facebook AI Research Research Intern on the FAIR Proteins (Generative Biology) Team
Research Fellow Summer/Fall 2018	OpenAI Research Fellow Working on Generative Models

Awards and Honors

Outstanding Paper Award at ICRA RAP4Robots Workshop	2023
Qualcomm Fellowship Finalist	2022
Outstanding Paper Award at NeurIPS Workshop on Controllable Generative Modeling	2021
National Science Foundation (NSF) Graduate Research Fellowship	2020-2024
Open Philanthropy Fellowship Finalist	2020

Publications

The * sign denotes equal contribution.

Peer-Reviewed Papers

C55 Anurag Ajay*, Seungwook Han*, <u>Yilun Du</u>*, Shuang Li, Abhi Gupta, Tommi Jaakkola, Joshua B. Tenenbaum, Leslie Kaelbling, Akash Srivastava, Pulkit Agrawal. "Compositional Foundation Models for Hierarchical Planning" *Neural Information Processing Systems (NeurIPS)*, 2023.

C54 <u>Yilun Du</u>*, Mengjiao Yang*, Bo Dai, Hanjun Dai, Ofir Nachum, Joshua B. Tenenbaum, Dale Schuurmans, Pieter Abbeel. "Learning Universal Policies via Text-Guided Video Generation" *Neural Information Processing Systems (NeurIPS)*, 2023.

C53 Siyuan Zhou, <u>Yilun Du</u>, Shun Zhang, Mengdi Xu, Yikang Shen, Wei Xiao, Dit-Yan Yeung, Chuang Gan. "Adaptive Online Replanning with Diffusion Models " *Neural Information Processing Systems (NeurIPS)*, 2023.

C52 Cameron Smith, <u>Yilun Du</u>, Ayush Tewari, Vincent Sitzmann. "FlowCam: Training Generalizable 3D Radiance Fields without Camera Poses via Pixel-Aligned Scene Flow " *Neural Information Processing Systems (NeurIPS)*, 2023.

C51 Yining Hong, Haoyu Zhen, Peihao Chen, Shuhong Zheng, <u>Yilun Du</u>, Zhenfang Chen, Chuang Gan. "3D-LLM: Injecting the 3D World into Large Language Models" *Neural Information Processing Systems (NeurIPS)*, 2023.

C50 Tsun-Hsuan Wang, Juntian Zheng, Pingchuan Ma, <u>Yilun Du</u>, Byungchul Kim, Andrew Spielberg, Joshua B. Tenenbaum, Chuang Gan, Daniela Rus. "DiffuseBot: Breeding Soft Robots With Physics-Augmented Generative Diffusion Models" *Neural Information Processing Systems (NeurIPS)*, 2023.

C49 Shengzhuang Chen, Long-Kai Huang, Jonathan Richard Schwarz, <u>Yilun Du</u>, Ying Wei. "Secure Out-of-Distribution Task Generalization with Energy-Based Models" *Neural Information Processing Systems (NeurIPS)*, 2023.

C48 Zhutian Yang, Jiayuan Mao, <u>Yilun Du</u>, Jiajun Wu, Joshua B. Tenenbaum, Tomas Lozano-Perez, Leslie Kaelbling. "Compositional Diffusion-Based Continuous Constraint Solvers" *Conference on Robot Learning (CoRL)*, 2023.

C47 Nan Liu*, <u>Yilun Du</u>*, Shuang Li*, Joshua B. Tenenbaum, Antonio Torralba. "Unsupervised Compositional Concepts Discovery with Text-to-Image Generative Models" *International Conference on Compouter Vision (ICCV)*, 2023.

C46 <u>Yilun Du</u>, Conor Durkan, Robin Strudel, Joshua B. Tenenbaum, Sander Dieleman, Rob Fergus, Jascha Sohl-Dickstein, Arnaud Doucet, Will Grathwohl. "Reduce, Reuse, Recycle: Compositional Generation with Energy-Based Diffusion Models and MCMC" *International Conference on Machine Learning (ICML)*, 2023.

C45 Armand Comas, <u>Yilun Du</u>, Christian Fernandez, Sandesh Ghimire, Mario Sznaier, Joshua B. Tenenbaum, Octavia Camps. "Inferring Relational Potentials in Interacting Systems" *International Conference on Machine Learning (ICML)*, 2023.

C44 Jiahui Fu, <u>Yilun Du</u>, Kurran Singh, Joshua B. Tenenbaum, John J. Leonard. "NeuSE: Neural SE(3)-Equivariant Embedding for Consistent Spatial Understanding with Objects" *Robotics Science and Systems (RSS)*, 2023.

C43 Weiyu Liu, <u>Yilun Du</u>, Tucker Hermans, Sonia Chernova, Chris Paxton. "StructDiffusion: Language-Guided Creation of Physically-Valid Structures using Unseen Objects" *Robotics Science and Systems (RSS)*, 2023.

C42 Cheng Chi, Siyuan Feng, <u>Yilun Du</u>, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, Shuran Song. "Diffusion Policy: Visuomotor Policy Learning via Action Diffusion" *Robotics Science and Systems (RSS)*, 2023.

C41 <u>Yilun Du</u>, Cameron Smith, Ayush Tewari, Vincent Sitzmann. "Learning to Render Novel Views from Wide-Baseline Stereo Pairs" *Computer Vision and Pattern Recognition (CVPR)*, 2023.

C40 Yining Hong, Chunru Lin, <u>Yilun Du</u>, Zhenfang Chen, Joshua B. Tenenbaum, Chuang Gan. "3D Concept Learning and Reasoning from Multi-View Images" *Computer Vision and Pattern Recognition (CVPR)*, 2023.

C39 Anurag Ajay*, <u>Yilun Du</u>*, Ahbi Gupta*, Joshua B. Tenenbaum, Tommi S. Jaakkola, Pulkit Agrawal. "Is Conditional Generative Modeling all You Need for Decision-Making?". *International Conference on Learning Representations (ICLR)*, 2023.

C38 Shuang Li*, <u>Yilun Du</u>*, Joshua B. Tenenbaum, Antonio Torralba, Igor Mordatch. "Composing Ensembles of Pre-trained Models via Iterative Consensus". *International Conference on Learning Representations (ICLR)*, 2023.

C37 Hongyi Chen*, <u>Yilun Du</u>*, Yiye Chen*, Joshua B. Tenenbaum, Patricio Antonio Vela. "Planning with Sequence Models through Iterative Energy Minimization" *International Conference on Learning Representations (ICLR)*, 2023.

C36 Prafull Sharma, Ayush Tewari, <u>Yilun Du</u>, Sergey Zakharov, Rares Ambrus, Adrien Gaidon, William T. Freeman, Fredo Durand, Joshua B. Tenenbaum, Vincent Sitzmann. "Seeing 3D Objects in a Single Image via Self-Supervised Static-Dynamic Disentanglement". *International Conference on Learning Representations (ICLR)*, 2023.

C35 Ethan Chun, <u>Yilun Du</u>, Anthony Simeonov, Tomas Lozano-Perez, Leslie Kaelbling. "Local Neural Descriptor Fields: Locally Conditioned Object Representations for Manipulation". *International Conference on Robotics Automation (ICRA)*, 2023.

C34 Jose Iturralde*, Aiden Curtis*, <u>Yilun Du</u>, Leslie Kaelbling, Tomas Lozano-Perez. "Visiblity-Aware Navigation Among Movable Objects". *International Conference on Robotics Automation (ICRA)*, 2023.

C33 Anthony Simeonov*, <u>Yilun Du</u>*, Yen-Chen Lin, Alberto Rodriguez, Leslie Kaelbling, Tomas Lozano-Perez, Pulkit Agrawal. "SE(3)-Equivariant Relational Rearrangement with Neural Descriptor Fields" *Conference on Robot Learning (CoRL)*, 2022.

C32 Yen-Chen Lin, Pete Florence, Andy Zheng, Johnathon T. Barron, <u>Yilun Du</u>, Wei-Chiu Ma, Anthony Simeonov, Alberto Rodriguez, Phillip Isola. "MIRA: Mental Imagery for Robotic Affordances" *Conference on Robot Learning (CoRL)*, 2022.

C31 Andrew Luo, <u>Yilun Du</u>, Michael J. Tarr, Joshua B. Tenenbaum, Antonio Torralba, Chuang Gan. "Learning Neural Acoustic Fields" *Neural Information Processing Systems (NeurIPS)*, 2022.

C30 Yining Hong, <u>Yilun Du</u>, Chunru Lin, Joshua B. Tenenbaum, Chuang Gan. "3D Concept Grounding on Neural Fields" *Neural Information Processing Systems (NeurIPS)*, 2022.

C29 Shuang Li, Xavier Puig, Chris Paxton, <u>Yilun Du</u>, Clinton Wang, Linxi Fan, Tao Chen, De-An Huang, Ekin Akyurek, Anima Anandkumar, Jacob Andreas, Igor Mordatch, Antonio Torralba, Yuke Zhu. "Pre-Trained Language Models for Interactive Decision-Making" *Neural Information Processing Systems (NeurIPS)*, 2022.

C28 Nan Liu*, Shuang Li*, <u>Yilun Du</u>*, Antonio Torralba, Joshua B. Tenenbaumn. "Compositional Visual Generation with Composable Diffusion Models" *European Computer Conference in Vision (ECCV)*, 2022.

C27 <u>Yilun Du</u>, Tomas Lozano-Perez, Leslie Kaelbling. "Learning Object Based State Estimators for Household Autonomy" *International Conference on Intelligent Robots and Systems (IROS)*, 2022.

C26 Jiahui Fu, <u>Yilun Du</u>, Kurran Singh, Joshua B. Tenenbaum, John J. Leonard. "Robust Change Detection Based on Neural Descriptor Fields" *International Conference on Intelligent Robots and Systems (IROS)*, 2022.

C25 Michael Janner*, <u>Yilun Du</u>*, Joshua B. Tenenbaum, Sergey Levine. "Planning with Diffusion for Flexible Behavior Synthesis" *International Conference on Machine Learning (ICML)*, 2022.

C24 <u>Yilun Du</u>, Shuang Li, Joshua B. Tenenbaum, Igor Mordatch. "Learning Iterative Reasoning through Energy Minimization" *International Conference on Machine Learning (ICML)*, 2022.

C23 Rylan Schaeffer, <u>Yilun Du</u>, Gabrielle Liu, Ila Fiete. "Streaming Inference for Infinite Feature Models" *International Conference on Machine Learning (ICML)*, 2022.

C22 Shuang Li, <u>Yilun Du</u>, Gido M. van de Ven, Antonio Torralba, Igor Mordatch. "Energy-Based Models for Continual Learning" *Conference on Lifelong Learning Agents (CoLLA)*, 2022.

C21 Rylan Schaeffer, Gabrielle Liu, <u>Yilun Du</u>, Scott Linderman, Ila Fiete. "Streaming Inference for Infinite Non-Stationary Clustering" *Conference on Lifelong Learning Agents (CoLLA)*, 2022.

C20 Klaus Greff, Francois Belleetti, Lucas Beyer, Carl Doersch, <u>Yilun Du</u>, Daniel Duckworth, David J. Fleet, Dan Gnanapragasam, Florian Golemo, Charles Herrmann, Thomas Kipf, Abhijit Kundu, Dmitry Lagun, issam Laradji, Derek Liu, Hinning Meyer, Yishu Miao, Derek Nowrouzezahrai, Cengiz Oztireli, Etienne Pot, Noha Radwan, Daniel Rebain, Sara Sabour, Mehdi Sajjadi, Matan Sela, Vincent Sitzmann, Austin Stone, Deqing Sun, Suhani Vora, Ziyu Wang, Tianhao Wu, Kwang Moo Yi, Fangcheng Zhong, Andrea Tagliasacchi. "Kubric: A Scalable Dataset Generator" *Computer Vision and Pattern Recognition (CVPR)*, 2022.

C19 Anthony Simeonov*, <u>Yilun Du</u>*, Andrea Tagliasacchi, Joshua B. Tenenbaum, Alberto Rodriguez, Pulkit Agrawal, Vincent Sitzmann. "Neural Descriptor Fields: SE(3)-Equivariant Object Representations for Manipulation". *International Conference on Robotics Automation (ICRA)*, 2022.

C18 <u>Yilun Du</u>, Shuang Li, Yash Sharma, Joshua B. Tenenbaum, Igor Mordatch. "Unsupervised Learning of Compositional Energy Concepts" *Neural Information Processing Systems (NeurIPS)*, 2021.

C17 Nan Liu*, Shuang Li*, <u>Yilun Du</u>*, Joshua B. Tenenbaum, Antonio Torralba. "Learning to Compose Visual Relations" *Neural Information Processing Systems (NeurIPS)*, 2021.

C16 <u>Yilun Du</u>, Katie Collins, Joshua B. Tenenbaum, Vincent Sitzmann. "Learning Signal-Agnostic Manifolds of Neural Fields" *Neural Information Processing Systems (NeurIPS)*, 2021.

C15 Joseph Suarez, <u>Yilun Du</u>, Clare Zhu, Igor Mordatch, Phillip Isola. "The Neural MMO Platform for Massively Multiagent Research" *Neural Information Processing Systems Track on Datasets and Benchmarks (NeurIPS)*, 2021.

C14 <u>Yilun Du</u>, Chuang Gan, Phillip Isola. "3D Shape Generation and Completion through Point-Voxel Diffusion" *International Conference on Computer Vision (ICCV)*, 2021.

C13 <u>Yilun Du</u>, Yinan Zhang, Hong-Xing Yu, Joshua B. Tenenbaum, Jiajun Wu. "Neural Radiance Flow for 4D View Synthesis and Video Processing" *International Conference on Computer Vision (ICCV)*, 2021.

C12 Linqi Zhou, <u>Yilun Du</u>, Jiajun Wu. "Curious Representation Learning for Embodied Intelligence" *International Conference on Computer Vision (ICCV)*, 2021.

C11 Shuang Li, <u>Yilun Du</u>, Antonio Torralba, Josef Sivic, Bryan Russell. "Weakly Supervised Human-Object Interaction Detection in Video via Contrastive Spatiotemporal Regions" *International Conference on Computer Vision (ICCV)*, 2021.

C10 <u>Yilun Du</u>, Shuang Li, Joshua B. Tenenbaum, Igor Mordatch. "Improved Contrastive Divergence Training of Energy Based Models" *International Conference on Machine Learning (ICML)*, 2021.

C9 <u>Yilun Du</u>, Kevin Smith, Tomer Ulman, Joshua B. Tenenbaum, Jiajun Wu. "Unsupervised Discovery of 3D Physical Objects from Video" *International Conference on Learning Representations (ICLR)*, 2021.

C8 <u>Yilun Du</u>, Shuang Li, Igor Mordatch. "Compositional Visual Generation with Energy Based Models" *Neural Information Processing Systems (NeurIPS)*, 2020.

C7 Anthony Simeonov, <u>Yilun Du</u>, Beomjoon Kim, Francois Hogan, Joshua B. Tenenbaum, Pulkit Agrawal, Alberto Rodriguez. "A Long Horizon Planning Framework for Manipulating Rigid Pointcloud Objects" *Conference on Robot Learning (CoRL)*, 2020.

C6 <u>Yilun Du</u>, Joshua Meier, Jerry Ma, Rob Fergus, Alexander Rives. "Energy-based models for atomic-resolution protein conformations". *International Conference on Learning Representations (ICLR)*, 2020.

C5 Xingyou Song, Yiding Jiang, <u>Yilun Du</u>, Behnam Neyshabur. "Observational Overfitting in Reinforcement Learning". *International Conference on Learning Representations (ICLR)*, 2020.

C4 <u>Yilun Du</u>, Toru Lin, Igor Mordatch. "Model Based Planning with Energy Based Models" *Conference on Robot Learning (CoRL)*, 2019.

C3 <u>Yilun Du</u>, Igor Mordatch. "Implicit Generation and Generalization with Energy Based Models" *Neural Information Processing Systems (NeurIPS)*, 2019.

C2 <u>Yilun Du</u>, Karthik Narasimhan. "Task-Agnostic Dynamics Priors for Deep Reinforcement Learning" *International Conference on Machine Learning (ICML)*, 2019.

C1 <u>Yilun Du</u>, Zhijian Liu, Hector Basevi, Ales Leonardis, William T. Freeman, Joshua B. Tenenbaum, Jiajun Wu. "Learning to Exploit Stability for 3D Scene Parsing" *Neural Information Processing Systems (NeurIPS)*, 2018.

Preprints

P5 <u>Yilun Du</u>, Sherry Yang, Pete Florence, Fei Xia, Ayzaan Wahid, Brian Ichter, Pierre Sermanet, Tianhe Yu, Pieter Abbeel, Joshua B. Tenenbaum, Leslie Kaelbling, Andy Zeng, Jonathan Tompson. "Video Language Planning" arXiv 2023

P4 <u>Yilun Du</u>, Shuang Li, Antonio Torralba, Joshua B. Tenenbaum, Igor Mordatch. "Improving Factuality and Reasoning in Language Models through Multiagent Debate" arXiv 2023

P3 Kevin Black*, Michael Janner*, <u>Yilun Du</u>, Ilya Kostrikov, Sergey Levine. "Training Diffusion Models with Reinforcement Learning" arXiv 2023

P2 Mengjiao Yang, <u>Yilun Du</u>, Kamyar Ghasemipour, Jonathan Tompson, Dale Schuurmans, Pieter Abbeel. "Learning Interactive Real-World Simulators" arXiv 2023

P1 Robert Verkuil*, Ori Kabeli*, <u>Yilun Du</u>, Basile Wicky, Lukas Milles, Justas Dauparas, David Baker, Sergey Ovchinnikov, Tom Sercu, Alexander Rives. "Language Models Generalize Beyond Natural Proteins" bioRxiv 2023

Teaching

Experience

6.867 Machine Learning (Graduate) Teaching Assistant, Massachusetts Institute of Technology	Fall 2021
6.036/6.862 Machine Learning (Undergraduate/Graduate) Teaching Assistant, Massachusetts Institute of Technology	Spring 2018
6.036 Machine Learning (Undergraduate) Teaching Assistant, Massachusetts Institute of Technology	Fall 2017

Invited Talks

Learning to Generate Compositionally
CMU, AI Seminar, Nov 2023
Brown University, Robotics Seminar, Oct 2023
Harvard University, Oct 2023
Cornell University, Robotics Seminar, Oct 2023
UC Berkeley, Oct 2023
UT Austin, Oct 2023
Stanford, Vision Seminar, July 2023

Towards Multi-AI Intelligent Systems
University of Tokyo, IRCN Keynote Speech, Oct 2023

Eye on AI Podcast
AI Self-Improvement from Multiagent Debate, Oct 2023

Riding the Wave of Generative Modeling for Decision Making
University of Montreal, Robot Learning Seminar, Oct 2023

Planning with Diffusion Models
Donders Institue, Feb 2023
Deepmind, Aug 2022

Implicit Learning with Energy Based Models
Nuro, October 2022
Tencent, September 2022
CMU, August 2022

Generally Intelligent Podcast
Modularity and Energy Based Models, Nov 2021

Mentoring

Research Mentoring

Undergraduates and Master Students

Pochen Ko	2023-
Zhengyang Liang	2023-
Yunhao Luo	2023-
Yinan Zhang - Now Research Engineer at Pinterest	2021,2023-
Bharat Runwal	2022-
Alpner Canberk	2023
Jocelin Su	2022-2023
Zachery Akner	2022-2023
Megan Wei - Now PhD Student at Brown	2022-2023
Ethan Chun	2022-2023
Hongyi Chen - Now PhD Student at CMU	2022-2023
Qingquan Bao - Now Master Student at UPenn	2022-2023
Nan Liu - Now Research Engineer at Google	2021-2023
Jose Itturalde	2022
Jinge Ma	2022
Vabhihav Singh	2022
Yihong Sun - Now PhD Student at Cornell	2021-2022
Alex Zhou - Now PhD Student at Stanford	2020-2021

PhD Students

Siyuan Zhou	2022-
Yichen Li	2022-
Armand Comas	2021-
Yining Hong	2022-2023
Jiahui Fu	2021-2023
Andrew Luo	2022

Professional Service

Conference Refereeing

Neural Information Processing Systems (NeurIPS)	2021-23
International Conference in Machine Learning (ICML)	2021-23
International Conference in Learning Representations (ICLR)	2021-24

Computer Vision and Pattern Recognition (CVPR)	2021-23
International Conference in Computer Vision (ICCV)	2021,23
European Conference in Computer Vision (ECCV)	2022
Special Interest Group in Graphics (SIGGRAPH)	2022-23
Conference on Robot Learning (CoRL)	2023
Robotics: Science and Systems (RSS)	2023
International Conference on Robotics and Automation (ICRA)	2022-23
International Conference on Intelligent Robots and Systems (IROS)	2022-23

Workshop Organization

Foundation Models for Decision Making Workshop at NeurIPS 2023	Dec 2023
Foundation Models for Decision Making Workshop at NeurIPS 2022	Dec 2022

Seminar Organization

MIT Embodied Intelligence Seminar	2021-2023