Yilun Du

Harvard Science and Engineering Complex
150 Western Ave, 6.138
Allston, MA 02134

ydu@seas.harvard.edu
https://yilundu.github.io/

Research Overview

My research is driven by the goal of developing intelligent decision-making agents. My research focuses on using generative AI to construct models of the world, using planning on top of the models as a way to determine the actions of the agent. A unique challenge of constructing world models 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, reasoning, multimodal perception, and planning. My work has further shown how such developed techniques can be broadly applied to settings in both the 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

Assistant Professor Summer 2025	Harvard University Kempner Institute and Computer Science
Senior Research Scientist 2024-2025	Google Deepmind Research Scientist on the Deepmind Robotics Team
Graduate Researcher 2019-2024	Computer Science and Artificial Intelligence Lab, MIT LIS Group (PIs: Tomás Lozano-Pérez and Leslie Pack Kaelbling)
Graduate Researcher 2019-2024	Brain and Cognitive Science Department, MIT Computational Cognitive Science Lab (PI: Joshua Tenenbaum)
Research Advisor 2023-2024	Pika Labs Text-to-Video Synthesis / Running Research Projects

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 ICLR	2024
Best Paper Award at NeurIPS Instruction Following Workshop	2023
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
International Biology Olympiad Gold Medal	2014, 2015

Publications

The * sign denotes equal contribution.

Peer-Reviewed Papers

C95 Shalev Lifshitz, Sheila A. McIlraith, <u>Yilun Du</u>. "Multi-Agent Verification: Scaling Test-Time Compute with Multiple Verifiers" *Conference on Language Modeling (CoLM)*, 2025.

C94 Haoyu Zhen*, Qiao Sun*, Hongxin Zhang, Junyan Li, Siyuan Zhou, <u>Yilun Du</u>, Chuang Gan. "Tesser-Act: Learning 4D Embodied World Models" *International Conference on Computer Vision (ICCV)*, 2025.

C93 Runhan Huang*, Shaoting Zhu*, <u>Yilun Du</u>, Hang Zhao. "MoELoco: Mixture of Experts for Multitask Locomotion" *International Conference on Intelligent Robots and Systems (IROS)*, 2025.

C92 Yanbo Wang, Justin Dauwels, <u>Yilun Du</u>. "Compositional Scene Understanding through Inverse Generative Modeling" *International Conference in Machine Learning (ICML)*, 2025.

C91 Kiwhan Song*, Boyuan Chen*, Max Simchowitz, <u>Yilun Du</u>, Russ Tedrake, Vincent Sitzmann. "History-Guided Video Diffusion" *International Conference in Machine Learning (ICML)*, 2025.

C90 Shenyuan Gao, Siyuan Zhou, <u>Yilun Du</u>, Jun Zhang, Chuang Gan. "Learning Adaptable World Models with Latent Actions" *International Conference in Machine Learning (ICML)*, 2025.

C89 Zishen Wan, Jiayi Qian, Yuhang Du, Jason Jabbour, <u>Yilun Du</u>, Yang (Katie) Zhao, Arijit Raychowhury, Tushar Krishna, Vijay Reddi. "Generative AI in Embodied Systems: System-Level Analysis of Performance, Efficiency and Scalability" *IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, 2025.

C88 Sigmund H. Hoeg, <u>Yilun Du</u>, Olav Egeland. "Streaming Diffusion Policy: Fast Policy Synthesis with Variable Noise Diffusion Models" *International Conference on Robotics Automation (ICRA)*, 2025.

C87 Yanwei Wang, Lirui Wang, <u>Yilun Du</u>, Balakumar Sundaralingam, Xuning Yang, Yu-Wei Chao, Claudia Perez-D'Arpino, Dieter Fox, Julie Shah. "Inference-Time Policy Steering through Human Interactions" *International Conference on Robotics Automation (ICRA)*, 2025.

C86 Siddharth Ancha, Sunshine Jiang, Travis Manderson, Laura Brandt, <u>Yilun Du</u>, Phillip Osteen, Nicholas Roy. "Anomalies-by-Synthesis: Anomaly Detection using Generative Diffusion Models for Off-Road Navigation" *International Conference on Robotics Automation (ICRA)*, 2025.

C85 Vighnesh Subramaniam*, <u>Yilun Du</u>*, Joshua B. Tenenbaum, Antonio Torralba, Shuang Li, Igor Mordatch. "Multiagent Finetuning: Self Improvement with Diverse Reasoning Chains" *International Conference on Learning Representations (ICLR)*, 2025.

C84 Yunhao Luo, <u>Yilun Du</u>. "Grounding Video Models to Actions through Goal Conditioned Exploration" *International Conference on Learning Representations (ICLR)*, 2025.

C83 Calvin Luo, Zilai Zeng, <u>Yilun Du</u>, Chen Sun. "Solving New Tasks by Adapting Internet Video Knowledge" *International Conference on Learning Representations (ICLR)*, 2025.

C82 Ying Fan, <u>Yilun Du</u>, Kannan Ramchandran, Kangwook Lee. "Looped Transformers for Length Generalization" *International Conference on Learning Representations (ICLR)*, 2025.

C81 Wenhao Chai*, Enxin Song*, <u>Yilun Du</u>, Chenlin Meng, Vashisht Madhavan, Omer Bar-Tel, Jenq-Neng Hwang, Saining Xie, Christopher Manning. "AuroraCap: Efficient, Performant Video Detailed Captioning and a New Benchmark" *International Conference on Learning Representations (ICLR)*, 2025.

C80 Hongxin Zhang*, Zeyuan Wang*, Qiushi Lyu*, Zheyuan Zhang, Sunli Chen, Tianmin Shu, <u>Yilun Du</u>, Chuang Gan. "COMBO: Compositional World Models for Embodied Multi-Agent Cooperation" *International Conference on Learning Representations (ICLR)*, 2025.

C79 Julia Briden, <u>Yilun Du</u>, Enrico Zucchelli, Richard Linares. "Compositional Diffusion Models for Powered Descent Trajectory Generation with Flexible Constraints" *IEEE Aerospace*, 2025.

C78 Aviv Netanyahu, <u>Yilun Du</u>, Antonia Bronars, Jyotish Pari, Joshua B. Tenenbaum, Tianmin Shu, Pulkit Agrawal. "Few-Shot Task Learning through Inverse Generative Modeling" *Neural Information Processing Systems (NeurIPS)*, 2024.

C77 Boyuan Chen, Diego Marti Monso, <u>Yilun Du</u>, Max Simchowitz , Russ Tedrake, Vincent Sitzmann. " Diffusion Forcing: Next-token Prediction Meets Full-Sequence Diffusion" *Neural Information Processing Systems (NeurIPS)*, 2024.

C76 Yinan Zhang, Eric Tzeng, <u>Yilun Du</u>, Dmitry Kislyuk. " Large-scale Reinforcement Learning for Diffusion Models" *International Conference on Computer Vision (ICCV)*, 2024.

C75 Jie Yie, Andrew Luo, <u>Yilun Du</u>, Anoop Cherian, Tim K. Marks, Jonathan Le Roux, Chuang Gan. " Disentangled Acoustic Fields For Multimodal Physical Scene Understanding" *International Conference on Intelligent Robots and Systems (IROS)*, 2024.

C74 Yiqing Xu, Jiayuan Mao, <u>Yilun Du</u>, Tomas Lozano-Perez, Leslie Kaelbling, David Hsu. " "Set It Up!": Functional Object Arrangement with Compositional Generative Models" *Robotics Science and Systems (RSS)*, 2024.

C73 Lirui Wang, Alan Zhao, <u>Yilun Du</u>, Ted Adelson, Russ Tedrake. "PoCo: Policy Composition from and for Heterogeneous Robot Learning" *Robotics Science and Systems (RSS)*, 2024.

C72 <u>Yilun Du</u>, Leslie Kaelbling. "Compositional Generative Modeling: A Single Model is Not All You Need" *International Conference on Machine Learning (ICML)*, 2024.

C71 <u>Yilun Du</u>, Shuang Li, Antonio Torralba, Joshua B. Tenenbaum, Igor Mordatch. "Improving Factuality and Reasoning in Language Models through Multiagent Debate" *International Conference on Machine Learning (ICML)*, 2024.

C70 <u>Yilun Du</u>*, Jiayuan Mao*, Joshua B. Tenenbaum. "Learning Iterative Reasoning through Energy Diffusion" *International Conference on Machine Learning (ICML)*, 2024.

C69 Yunhao Luo, Chen Sun, Joshua B. Tenenbaum, <u>Yilun Du</u>. "Potential Based Diffusion Motion Planning" *International Conference on Machine Learning (ICML)*, 2024.

C68 Jocelin Su*, Nan Liu*, Yanbo Wang*, Joshua B. Tenenbaum, <u>Yilun Du</u>. "Compositional Image Decomposition with Diffusion Models" *International Conference on Machine Learning (ICML)*, 2024.

C67 Siyuan Zhou, <u>Yilun Du</u>, Jiaben Chen, Yandong Li, Dit-Yan Yeung, Chuang Gan. "RoboDreamer: Learning Compositional World Models for Robot Imagination" *International Conference on Machine Learning (ICML)*, 2024.

C66 Sherry Yang, Jacob Walker, Jack Parker-Holder, <u>Yilun Du</u>, Jake Bruce, Andre Barreto, Pieter Abbeel, Dale Schuurmans. "Video as the New Language for Real-World Decision Making" *International Conference on Machine Learning (ICML)*, 2024.

C65 Haoyu Zhen, Xiaowen Qiu, Peihao Chen, Jincheng Yang, Xin Yan, <u>Yilun Du</u>, Yining Hong, Chuang Gan. "3D-VLA: A 3D Vision-Language-Action Generative World Model" *International Conference on Machine Learning (ICML)*, 2024.

C64 <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" *International Conference on Learning Representations (ICLR)*, 2024.

C63 Kevin Black*, Michael Janner*, <u>Yilun Du</u>, Ilya Kostrikov, Sergey Levine. "Training Diffusion Models with Reinforcement Learning" *International Conference on Learning Representations (ICLR)*, 2024.

C62 Tailin Wu*, Takashi Maruyama*, Long Wei*, Tao Zhang*, <u>Yilun Du</u>*, Gianluca Iaccarino, Jure Leskovec. "Compositional Generative Inverse Design" *International Conference on Learning Representations (ICLR)*, 2024.

C61 Mengjiao Yang*, <u>Yilun Du</u>*, Bo Dai, Dale Schuurmans, Joshua B. Tenenbaum, Pieter Abbeel. "Probabilistic Adaptation of Text-to-Video Models" *International Conference on Learning Representations (ICLR)*, 2024

C60 Mengjiao Yang, <u>Yilun Du</u>, Kamyar Ghasemipour, Jonathan Tompson, Dale Schuurmans, Pieter Abbeel. "Learning Interactive Real-World Simulators" *International Conference on Learning Representations (ICLR)*, 2024.

C59 Po-Chen Ko, Jiayuan Mao, <u>Yilun Du</u>, Shao-Hua Sun, Joshua B. Tenenbaum. "Learning to Act from Actionless Video through Dense Correspondences" *International Conference on Learning Representations (ICLR)*, 2024.

C58 Yichen Li, <u>Yilun Du</u>, Chao Liu, Chao Liu, Francis Williams, Michael Foshey, Benjamin Eckart, Jan Kautz, Joshua B. Tenenbaum, Antonio Torralba, Wojciech Matusik. "Learning to Jointly Understand Visual and Tactile Signals" *International Conference on Learning Representations (ICLR)*, 2024.

C57 Hongxin Zhang, Weihua Du, Jiaming Shan, Qinhong Zhou, <u>Yilun Du</u>, Joshua B. Tenenbaum, Tianmin Shu, Chuang Gan. "Building Cooperative Embodied Agents Modularly with Large Language Models" *International Conference on Learning Representations (ICLR)*, 2024.

C56 Qinhong Zhou, Sunli Chen, Yisong Wang, Haozhe Xu, Weihua Du, Hongxin Zhang, <u>Yilun Du</u>, Joshua B. Tenenbaum, Chuang Gan. "BHAZARD Challenge: Embodied Decision Making in Dynamically Changing Environments" *International Conference on Learning Representations (ICLR)*, 2024.

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 Computer 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 Compouter 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.

Journal

J5 Yiqing Xu, Jiayuan Mao, Linfeng Li, <u>Yilun Du</u>, Tomas Lozáno-Pérez, Leslie Pack Kaelbling, David Hsu. ""Set It Up": Functional Object Arrangement with Compositional Generative Models" *International Journal on Robotics Research (IJRR)* 2025

J4 Robert McCarthy, Daniel CH Tan, Dominik Schmidt, Fernando Acero, Nathan Herr, <u>Yilun Du</u>, Thomas G Thuruthel, Zhibin Li. ""owards generalist robot learning from internet video: A survey" *Journal on Artificial Intelligence Research (IJRR)* 2025

J3 Julen Urain, Ajay Mandlekar, <u>Yilun Du</u>, Mahi Shafiullah, Danfei Xu, Katerina Fragkiadaki, Georgia Chalvatzaki, Jan Peters. "Deep generative models in robotics: A survey on learning from multimodal demonstrations" *IEEE Transactions on Robotics (TRO)* 2025

J2 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" *International Journal on Robotics Research (IJRR)* 2025

J1 Cheng Chi*, Zhenjia Xu*, Siyuan Feng , Eric Cousineau , <u>Yilun Du</u>, Benjamin Burchfiel , Russ Tedrake, Shuran Song. "Diffusion Policy: Visuomotor Policy Learning via Action Diffusion" *International Journal on Robotics Research (IJRR)* 2024

Technical Reports

P3 Google Deepmind. "Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities"

P2 Gemini Robotics Team. "Gemini Robotics: Bringing AI into the Physical World"

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

Theses

T1 <u>Yilun Du</u> 2024. "Learning Generalizable Systems by Learning Composable Energy Landscapes". PhD Thesis, Massachusetts Institute of Technology.

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

Oregan State University
Building Embodied Intelligence with Compositional World Models, January 2026

NeurIPS Workshop on Embodied World Models
Building Intelligent Robots with World Models, December 2025

NeurIPS Tutorial on Beyond Autoregressive Models
Autoregressive World Models, December 2025

MIT FutureTech
AI Automation with Generative World Models, November 2025

Amazon Robotics
Robot Manipulation with World Models, November 2025

Tufts ECE Seminar
Learning Compositional Models of the World, September 2025

Nvidia
Multiagent Reasoning with Language Models, August 2025

Toyota Research Insitute ML and Robotics Seminar
Flexible Decision Making with Compositional Constraints, August 2025

Meta MRS AI Seminar
Multiagent Reasoning with Language Models, July 2025

Microsoft Research Seminar of Sampling and Generative Models
Inference Time Reasoning with Generative Models, July 2025

CVPR Tutorial on Continual Robot Learning from Humans
Continual Robot Learning with Compositional Generative Models, June 2025

RSS Workshop on Leveraging Foundation Models for Interactive Robot learning
Scaling Robot Learning with Video Models, June 2025

CVPR Tutorial on Robotics 101: An Odyssey from A Vision Perspective
Learning Compositional Models of the World, June 2025

CVPR Workshop on 3D-LLM/VLA
Build 4D Models of the World, June 2025

CVPR Workshop on Foundation Models Meet Embodied Agents
Learning Compositional Models of the World, June 2025

ICRA Workshop on Foundation Models and Neuro-Symbolic AI for Robotics
Generalizable Foundation Models through Compositional Generation, May 2025

New England NLP Symposium
Constructing Generalizable Multimodal Models through Compositional Generation, April 2025

Weizmann AI and Vision Seminar
Learning Compositional Models of the World, March 2025

Shanghai AI Lab
Learning Compositional Models of the World, Nov 2024

ICML Workshop on Multimodal Foundation Models in Embodied AI
Compositional Foundation Models, July 2024

RSS Workshop on Semantic Reasoning and Goal Understanding in Robotics
Panelist, July 2024

RSS Workshop on GenAI and HRI
Build Customizable Generative Models through Compositional Generation, July 2024

Generalizing Outside the Training Distribution through Composition Generation
UCLA, November 2024
KAIST, October 2024
Vanderbilt Machine Learning, September 2024
NUS Tea Time, July 2024
CMU MLD, April 2024
UMich CSE, April 2024
UT Austin, April 2024
UBC, April 2024
UT Austin, April 2024
Princeton, April 2024
Brown, April 2024
NYU, March 2024
Columbia, March 2024
UPenn, March 2024
University of Washington, March 2024
Georgia Tech, March 2024
University of Toronto, March 2024
Harvard, March 2024
UCSD, Feb 2024
Boston Dynamics AI Institute, Feb 2024
ASU Frontiers in GenAI Seminar, Jan 2024
Genesis Therapeutics, Jan 2024
JHU, Nov 2023
KAIST, Nov 2023
CMU, AI Seminar, Nov 2023
Brown University, Robotics Seminar, Oct 2023
Cornell University, Robotics Seminar, Oct 2023
UC Berkeley, Oct 2023
UT Austin, Oct 2023
Stanford, Vision Seminar, July 2023

NeurIPS Tutorial on Language Models Meet World Models
Panelist, Dec 2023

Learning a Visual Imagination
BuzzRobot, Dec 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

Katherine Lin	2023-
Zhengyang Liang	2023-
Yunhao Luo	2023-
Potential Based Diffusion Motion Planning	ICML 2024
Pochen Ko	2023-
Learning to Act from Actionless Videos through Dense Correspondences	ICLR 2024
Bharat Runwal	2022-
A Benchmark for Learning to Continually Generate	In Submission
Alpner Canberk	2023
Jocelin Su	2022-2023
Compositional Image Decomposition with Diffusion Models	ICML 2024
Zachery Akner	2022-2023
Megan Wei - Now PhD Student at Brown	2022-2023
Ethan Chun	2022-2023
Local Neural Descriptor Fields: Locally Conditioned Object Representations for Manipulation	ICRA 2023
Hongyi Chen - Now PhD Student at CMU	2022-2023
Planning with Sequence Models through Iterative Energy Minimization	ICLR 2023
Qingquan Bao - Now Master Student at UPenn	2022-2023
Embodied Depth Prediction	In Submission
Nan Liu - Now Research Engineer at Google	2021-2023
Learning to Compose Visual Relations	NeurIPS 2021
Compositional Visual Generation with Composable Diffusion Models	ECCV 2022
Unsupervised Compositional Concepts Discovery with Text-to-Image Generative Models	ICCV 2023
Jose Itturalde	2022
Visiblity-Aware Navigation Among Movable Objects	ICRA 2023
Vabhihav Singh	2022
Wake-Sleep Energy Based Models for Continual Learning	In Submission
Yihong Sun - Now PhD Student at Cornell	2021-2022
Alex Zhou - Now PhD Student at Stanford	2020-2021
3D Shape Generation and Completion through Point-Voxel Diffusion	ICCV 2021

PhD Students

Yichen Li	2022-2023
Learning to Jointly Understand Visual and Tactile Signals	ICLR 2024
Siyuan Zhou	2022-
Adaptive Online Replanning with Diffusion Models	NeurIPS 2023
Armand Comas	2021-

Professional Service

Conference Refereeing

Neural Information Processing Systems (NeurIPS)	2021-25
International Conference in Machine Learning (ICML)	2021-25
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)	2022-24
SIGGRAPH Asia	2024
Conference on Robot Learning (CoRL)	2023-24
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

Workshop on Robotics World Modeling at CoRL 2025	September 2025
Workshop on Robot Planning in the Era of Foundation Models at RSS 2025	June 2025
Workshop on World Models in Physical World at ICML 2025	June 2025
Workshop on Multi-Agent AI in the Real World at AAAI 2025	March 2025
Workshop on Compositional Learning at NeurIPS 2024	Dec 2024
Workshop on Safe Generative AI at NeurIPS 2024	Dec 2024
Workshop on Geometric and Algebraic Structure in Robot Learning at RSS 2024	July 2024
Generative Models for Decision Making Workshop at ICLR 2023	May 2024
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