

Matt Deitke

Research Interests: Computer Vision, Artificial Intelligence, Deep Learning, Embodied AI

PERSONAL INFORMATION

✉ EMAIL: mattdeitke@allenai.org
🌐 WEBSITE: mattdeitke.com
🐙 GITHUB: [@mattdeitke](https://github.com/mattdeitke)

CURRENT POSITION



Allen Institute for AI, Seattle, WA
Research & Engineering in Computer Vision
Full-Time Employee
Full-Time Research Intern

Fall 2019 – Present

Summer 2021 – Present

Fall 2019 – Spring 2021

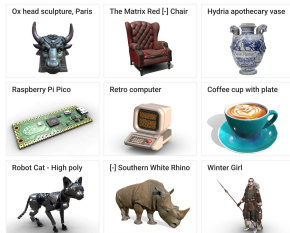


University of Washington, Seattle, WA
Paul G. Allen School of Computer Science & Engineering
Pursuing a B.S. in Computer Science

Fall 2019 – Spring 2023 (Exp.)

GPA: 3.82/4.0

PREPRINTS

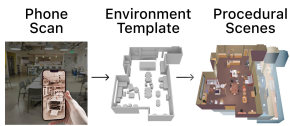


[1] Objaverse: A Universe of Annotated 3D Objects

Matt Deitke, Dustin Schwenk, Jordi Salvador, Luca Weihs, Oscar Michel, Eli VanderBilt, Ludwig Schmidt, Kiana Ehsani, Aniruddha Kembhavi, Ali Farhadi

CVPR 2023 [[arXiv](#)] [[website](#)]

tl;dr: Objaverse is a massive dataset of objects with 800K+ (and growing) 3D models with descriptive captions, tags, and animations. We demonstrate it's potential by training generative models, improving 2D instance segmentation, training open-vocabulary object navigation models, and creating a benchmark for testing the robustness of vision models.



[2] Phone2Proc: Bringing Robust Robots Into Our Chaotic World

Matt Deitke*, Rose Hendrix*, Luca Weihs, Ali Farhadi, Kiana Ehsani, Aniruddha Kembhavi

CVPR 2023 [[arXiv](#)] [[website](#)]

tl;dr: From a 10-minute iPhone scan of any environment, we generated simulated training scenes that semantically match that environment. Training a robot to perform ObjectNav in these scenes dramatically improves sim-to-real performance from 35% to 71% and results in an agent that is remarkably robust to human movement, lighting variations, added clutter, and rearranged objects.

PUBLICATIONS



[3] ProcTHOR: Large-Scale Embodied AI Using Procedural Generation

Matt Deitke, Eli VanderBilt, Alvaro Herrasti, Luca Weihs, Jordi Salvador, Kiana Ehsani, Winson Han, Eric Kolve, Ali Farhadi, Aniruddha Kembhavi, Roozbeh Mottaghi

NeurIPS 2022 Outstanding Paper Award [[arXiv](#)] [[website](#)]

tl;dr: We built a platform to procedurally generate realistic, interactive, simulated 3D environments to dramatically scale up the diversity and size of training data in Embodied AI. We find that it helps significantly with performance on many tasks.



[4] Retrospectives on the Embodied AI Workshop

Matt Deitke, Dhruv Batra, Yonatan Bisk, Tommaso Campari, Angel X. Chang, Devendra Singh Chaplot, Changan Chen, Claudia Pérez D’Arpino, Kiana Ehsani, Ali Farhadi, Li Fei-Fei, Anthony Francis, Chuang Gan, Kristen Grauman, David Hall, Winson Han, Unnat Jain, Aniruddha Kembhavi, Jacob Krantz, Stefan Lee, Chengshu Li, Sagnik Majumder, Oleksandr Maksymets, Roberto Martín-Martín, Roozbeh Mottaghi, Sonia Raychaudhuri, Mike Roberts, Silvio Savarese, Manolis Savva, Mohit Shridhar, Niko Sünderhauf, Andrew Szot, Ben Talbot, Joshua B. Tenenbaum, Jesse Thomason, Alexander Toshev, Joanne Truong, Luca Weihs, Jiajun Wu

ArXiv [arXiv]

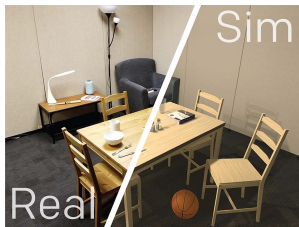
tl;dr: We present a retrospective on the state of Embodied AI research. Our analysis focuses on 13 challenges in visual navigation, rearrangement, and embodied vision-and-language. We discuss the scope of embodied AI research, performance of state-of-the-art models, common modeling approaches, and future directions.



[5] Visual Room Rearrangement

Luca Weihs, **Matt Deitke**, Aniruddha Kembhavi, Roozbeh Mottaghi
CVPR 2021 Oral Presentation [arXiv] [code] [video]

tl;dr: We built a pre-training task where the agent’s goal is to interactively rearrange objects in a room from one state to another. For instance, the agent may have to open the FRIDGE and move the LETTUCE to the COUNTERTOP. Modern deep-RL struggles.

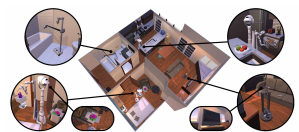


[6] RoboTHOR: An Open Simulation-to-Real Embodied AI Platform

Matt Deitke*, Winson Han*, Alvaro Herrasti*, Aniruddha Kembhavi*, Eric Kolve*, Roozbeh Mottaghi*, Jordi Salvador*, Dustin Schwenk*, Eli VanderBilt*, Matthew Wallingford*, Luca Weihs*, Mark Yatskar*, and Ali Farhadi

CVPR 2020 [website] [arXiv] [video] (~22% acceptance rate)

tl;dr: We rent office buildings in Seattle and turn them into apartment studios with many possible furniture and wall layouts. Each apartment layout is then computationally remodeled by hand to enable a simulated robot to interact with it in video-game-like context. We study how well a robot trained purely in the simulated environments can transfer to reality.



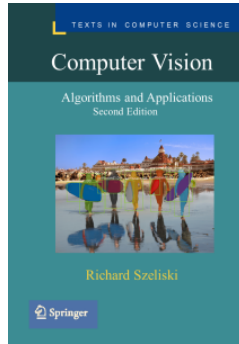
[7] AI2-THOR: An Interactive 3D Environment for Visual AI

Eric Kolve, Roozbeh Mottaghi, Winson Han, Eli VanderBilt, Luca Weihs, Alvaro Herrasti, **Matt Deitke**, Kiana Ehsani, Daniel Gordon, Yuke Zhu, Aniruddha Kembhavi, Abhinav Gupta, Ali Farhadi

ArXiv [website] [arXiv] [video]

tl;dr: We introduce The House Of inteRactions (THOR), a framework for visual AI research. AI2-THOR consists of near photo-realistic 3D indoor scenes, where AI agents can navigate in the scenes and interact with objects to perform tasks. It has enabled research in many areas of AI.

BOOK
CONTRIBUTIONS



[8] Computer Vision: Algorithms and Applications

Richard Szeliski

2nd ed. Springer Science & Business Media, 2022.

Contributions: Drafted initial sections on transformers, VAEs, text-to-image generation, CLIP, and many new works in deep learning. Provided feedback on early drafts. Wrote exercises for Chapters 5 (Deep Learning) and 6 (Recognition). Created several new figures. Added updates to Appendix C on supplementary material (*i.e.* datasets & benchmarks, software, and slides & lectures).

[\[website\]](#)

SELECTED
SOFTWARE



AI2-THOR

2019 – Present

Core Contributor. AI2-THOR is a highly customizable near photo-realistic interactive simulation framework for Embodied AI agents. The backend is in Unity/C# and we provide a Python API to interact with it. 546K+ total installations, 175+ publications have used it for experimentation.

[\[website\]](#) [\[code\]](#) [\[video\]](#) [\[PyPi\]](#)



ALLENACT

2020 – Present

Contributor. ALLENACT is a framework used to train embodied AI agents with reinforcement learning. It provides first-class support to work with PyTorch and multiple embodied AI simulators (*e.g.* AI2-THOR, Habitat, MiniGrid).

[\[website\]](#) [\[code\]](#) [\[PyPi\]](#)



CVPR BUZZ

2021

Lead Developer. Scrapes Twitter and Semantic Scholar to find which conference papers at CVPR 2021 have been discussed the most. Built a front-end to display everything using GraphQL and Gatsby.

[\[website\]](#) [\[code\]](#)



AI2-THOR × COLAB

2021 – Present

Lead Developer. Provides the ability to run AI2-THOR freely in the cloud using Google Colab.

[\[website\]](#) [\[code\]](#) [\[PyPi\]](#)



DLB: Deep Learning Board

2021 – Present

Lead Developer. Early prototype of a low-level deep learning visualization dashboard with React. Solves my issues with TensorBoard & Wandb being too high level to build custom interactive visualizations. Meant to be used with all the big JavaScript visualization libraries, including D3.js, Vega, and Vega-Lite. Only used internally at the moment.

[\[demo\]](#) [\[video\]](#)

RESEARCH
WEBSITES BUILT



ai2thor.allenai.org

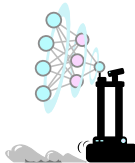
2019 – Present

Lead Developer. The website for AI2-THOR. Contains dozens of pages, my favorites include:

- [Web Demo](#)
- [Publication Tracker](#)
- [iTHOR Documentation](#)

Developed from scratch with Gatsby.

[\[website\]](#)



embodied-ai.org

2020 – Present

Lead Developer. Contains the information for the CVPR Embodied AI workshops. Developed from scratch with Gatsby.

[\[website\]](#) [\[code\]](#)

WORKSHOP
ORGANIZATION



CVPR Embodied AI Workshop

2020, 2021

I've co-organized the Embodied AI workshops at CVPR with researchers from a variety of institutions. The goal of the workshops are to bring together researchers from the fields of computer vision, language, graphics, and robotics to share and discuss the current state of intelligent agents that can see, talk, act, and reason.

[\[CVPR 2020\]](#) [\[CVPR 2021\]](#) [\[CVPR 2022\]](#)

CHALLENGE
ORGANIZATION



Visual Room Rearrangement Challenge

2021, 2022

The goal of this challenge is to build a model to rearrange objects in a room, such that they are restored to a given initial configuration. Held in conjunction with the Embodied AI Workshop at CVPR.

[\[CVPR 2021\]](#) [\[CVPR 2022\]](#)



Sim2Real ObjectNav Challenge

2020, 2021

The goal of this challenge is to encourage researchers to work on Sim2Real transfer, and to create a unified benchmark to track progress over time on the task. Built upon RoboTHOR, held in conjunction with the Embodied AI Workshop at CVPR.

[\[CVPR 2020\]](#) [\[CVPR 2021\]](#)

INVITED TALKS



ProcTHOR: Where We Are and What's Next

- [1] Ali Farhadi's Lab at the University of Washington
- [2] Ludwig Schmidt's Lab at the University of Washington
- [3] Allen Institute for AI's All-Hands Meeting (Primary Talk)
- [4] CVPR Embodied AI Workshop (Challenge Winner Talk)

REVIEWING



Served as a reviewer at:

- [1] CVPR 2023
- [2] ICLR 2023
- [3] CVPR 2022 Embodied AI Workshop

SELECTED
MEDIA
COVERAGE



- [1] Richard Szeliski. **On Computer Vision: Algorithms and Applications**, *RSIP Vision, Computer Vision News*, March, 2022. [\[website\]](#)



- [2] Karen Hao. **An ever-changing room of Ikea furniture could help AI navigate the world**, *MIT Technology Review*, 2020. [\[website\]](#)



- [3] Alan Boyle. **ManipulaTHOR training software from AI2 gives virtual robots a hand — and an arm**, *GeekWire*, 2021. [\[website\]](#)

- [4] Alan Boyle. **AI2 throws down the challenge for robotic scavenger hunt in virtual and real rooms**, *GeekWire*, 2020. [\[website\]](#)

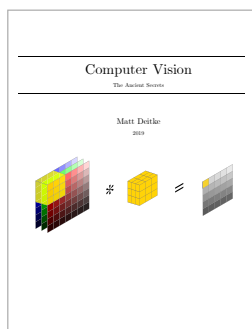


- [5] Jack Clark. **Issue 187: Real world robot tests at CVPR**, *Import AI*, 2020. [\[website\]](#)



- [6] **Edge#116: AI2-THOR is an Open-Source Framework for Embodied AI Research**, *TheSequence*, 2021. [\[website\]](#)

COURSE
TEXTBOOKS



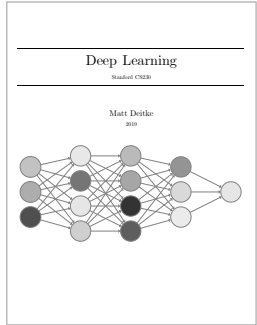
(201 PAGES)

Matt Deitke. Computer Vision: The Ancient Secrets, 2019.

Contents are based on lectures from Joseph Redmon and Ali Farhadi's computer vision course at the University of Washington (CSE 455).

Includes 17 chapters covering early work in computer vision (*e.g.* convolutions, edge detection, corner detection, SIFT, optical flow), an introduction to machine learning and neural networks, and newer deep learning approaches (*e.g.* YOLO, Faster R-CNN, GANs, U-Net).

[\[PDF\]](#)



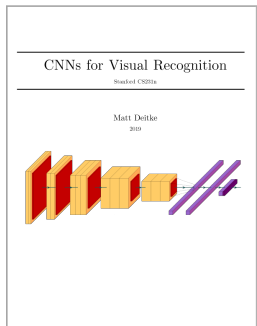
(101 PAGES)

Matt Deitke. Deep Learning, 2019.

Contents are based on lectures from Stanford's CS230 course on deep learning, taught by Andrew Ng.

Includes 7 chapters covering neural networks, optimization techniques, applications to computer vision, and applications to natural language processing. I created over 100 new figures for the book.

[\[PDF\]](#)



(103 PAGES)

Matt Deitke. CNNs for Visual Recognition, 2019.

Contents are based on lectures from Stanford's CS231n course neural networks for computer vision, taught by Fei-Fei Li, Justin Johnson, and Serena Yeung.

Includes 10 chapters on deep learning for computer vision. Covers neural networks, CNNs (and their variants), practical training tips, hardware and software, and vision-and-language.

[\[PDF\]](#)

PAST
EMPLOYMENT



Freelance Developer

2014 – 2019

With 50+ Companies on Computer Graphics, Interface Design, & Visualization



The University of Cincinnati, Cincinnati, OH

2016 – 2018

Part-Time Computer Graphics Developer for the Department of Athletics



The Ohio State University, Columbus, OH

2016 – 2018

Part-Time Computer Graphics Developer for the Department of Athletics



Cleveland Browns, Cleveland, OH

2014 – 2016

Part-Time Computer Graphics Developer

PAST
EDUCATION



Georgia Institute of Technology, Atlanta, GA

Fall 2018 – Spring 2019

Concurrent Enrollment While in High School

GPA: 4.0/4.0



Illinois Institute of Technology, Chicago, IL

Spring 2019

Concurrent Enrollment While in High School

GPA: 4.0/4.0

SELECTED
COURSEWORK



Georgia Institute of Technology

Concurrently in High School:

Built an Independent Study on Machine Learning & Deep Learning Research



University of Washington

Freshman:

CSE 576 Computer Vision (Grad) Steve Seitz, Richard Szeliski, *et al.*
CSE 571 Probabilistic Robotics (Grad) Dieter Fox

Sophomore:

CSE 512 Data Visualization (Grad) Jeffrey Heer
CSE 492M Startup Seminar Madrona Venture Labs

Junior:

CSE 573 Artificial Intelligence (Grad) Hannaneh Hajishirzi
CSE 599D1 Ethics in NLP (Grad) Yulia Tsvetkov
CSE 599A1 Entrepreneurship (Grad) Greg Gottesman, Ed Lazowska
CSE 590R Robotics Colloquium (Grad) Dieter Fox, Maya Cakmak

Senior:

CSE 599D1 Language, Knowledge, & Reasoning (Grad) Yejin Choi

REFERENCES

Ali Farhadi

Professor at the University of Washington, Seattle
Director of AI at Apple

Roozbeh Mottaghi

Research Scientist Manager at Meta AI
Affiliate Faculty at the University of Washington, Seattle

Aniruddha Kembhavi

Director of Computer Vision at the Allen Institute for AI
Affiliate Faculty at the University of Washington, Seattle

Richard Szeliski

Distinguished Scientist at Google Research
Affiliate Faculty at the University of Washington, Seattle

Ludwig Schmidt

Assistant Professor at the University of Washington, Seattle
Research Scientist at the Allen Institute for AI