Carl Vondrick

YM Associate Professor of Computer Science
Columbia University
618 Schapiro CEPSR, 530 West 120th St, New York, NY 10027

vondrick@cs.columbia.edu
cs.columbia.edu/∼vondrick
Google Scholar

Fields of Specialization

Computer Vision; Machine Learning.

Appointments and Employment

2025–Present	**Apple**, Research Scientist
2018–Present	**Columbia University**, Department of Computer Science
2024–Present	YM Associate Professor (with tenure)
2023–2024	YM Associate Professor
2023	Associate Professor
2018–2022	Assistant Professor
2022–2023	**Cruise**, Visiting AI Faculty
2022–2023	**Snap**, Research Design Consultant
2017–2019	**Google**, Research Scientist

Education

2017 **Massachusetts Institute of Technology**
Ph.D. in Computer Science
Advisor: Antonio Torralba
Thesis: Predictive Vision
Minor: Cognitive Science

2013 **Massachusetts Institute of Technology**
M.Sc. in Computer Science
Advisor: Antonio Torralba
Thesis: Visualizing Object Detection Features

2011 **University of California, Irvine**
B.Sc. in Computer Science
Advisor: Deva Ramanan
Thesis: Crowdsourced Video Annotation
Summa Cum Laude

Awards and Honors

2024	PAMI Young Researcher Award
2021	National Science Foundation Early Career Development (CAREER) Award ($550k)
2021	Toyota Research Institute Young Faculty Award ($750k)
2018	Amazon Research Award ($100k)
2018	Best Paper Finalist at CVPR
2015	Google Ph.D. Fellowship in Machine Perception

2011 National Science Foundation Graduate Research Fellowship

Awards and Honors of Lab Members

2026 Apple Ph.D. Fellowship to Sudhakar (2 years)
2024 CAIRFI Ph.D. Fellowship to Menon (1 year)
2024 Apple Ph.D. Fellowship to Tendulkar (2 years)
2023 National Science Foundation Graduate Research Fellowship to Geng (3 years)
2022 National Science Foundation Graduate Research Fellowship to Sudhakar (3 years)
2022 Microsoft Ph.D. Fellowship to Surís (2 years)
2021 Amazon CAIT Ph.D. Fellowship to Chiquier (2 years)
2021 National Science Foundation Graduate Research Fellowship to Mani (3 years)
2020 CRA Honorable Mention to Epstein
2020 National Science Foundation Graduate Research Fellowship to Menon (3 years)
2020 Presidential Fellowship to Menon

Publications

Trainees from my group are underlined. The authorship convention in my field is to order by decreasing contribution, with the advisor often appearing last.

Conference Papers (Peer Reviewed)

C1 Sreehari Rammohan, Yousef Yassin, Anthony Fuller, Junfeng Wen, **Carl Vondrick**, Evan Shelhamer. *LookThere! Sparse Vision by Reinforced Selection*. NeurIPS 2026, h5-index 245.

C2 Sreehari Rammohan, Huy Ha, **Carl Vondrick**. *A^2: Smaller Self-Supervised ViTs Localize Better than Larger Ones*. NeurIPS 2026, h5-index 245.

C3 Santhosh Kumar Ramakrishnan, **Carl Vondrick**, Raja Giryes, Philipp Krähenbühl, Vladlen Koltun. *Do multimodal models imagine electric sheep?* NeurIPS 2026, h5-index 245.

C4 Arjun Mani, **Carl Vondrick**, Richard Zemel. *Few-Shot Design Optimization by Exploiting Auxiliary Information*. ICML 2026, h5-index 237.

C5 Arsha Nagrani, Sachit Menon, Ahmet Iscen, Shyamal Buch, Ramin Mehran, Nilpa Jha, Anja Hauth, Yukun Zhu, **Carl Vondrick**, Mikhail Sirotenko, Cordelia Schmid, Tobias Weyand. *MINERVA: Evaluating Complex Video Reasoning*. ICCV 2025, h5-index 184.

C6 David S. Hayden, Mao Ye, Timur Garipov, Gregory P. Meyer, **Carl Vondrick**, Zhao Chen, Yuning Chai, Eric Wolff, Siddhartha S. Srinivasa. *Generative Data Mining with Longtail-Guided Diffusion*. ICML 2025, h5-index 237.

C7 Utkarsh Mall, Cheng Perng Phoo, Mia Chiquier, Bharath Hariharan, Kavita Bala, **Carl Vondrick**. *DiSciPLE: Learning Interpretable Programs for Scientific Visual Discovery*. CVPR 2025, h5-index 356.

C8 Ruoshi Liu, Huy Ha, Mengxue Hou, Shuran Song, **Carl Vondrick**. *Self-Improving Autonomous Underwater Manipulation*. ICRA 2025.

C9 Ruoshi Liu, Alper Canberk, Shuran Song, **Carl Vondrick**. *Differentiable Robot Rendering*. CoRL 2024. *Oral presentation*

C10 Junbang Liang, Ruoshi Liu, Ege Ozguroglu, Sruthi Sudhakar, Achal Dave, Pavel Tokmakov, Shuran Song, **Carl Vondrick**. *Dreamitate: Real-World Visuomotor Policy Learning via Video Generation*. CoRL 2024.

C11 Sachit Menon, Richard Zemel, **Carl Vondrick**. *Whiteboard-of-Thought: Thinking Step-by-Step Across Modalities*. EMNLP 2024.

C12 Alper Canberk, Maksym Bondarenko, Ege Ozguroglu, Ruoshi Liu, **Carl Vondrick**. *EraseDraw: Learning to Draw Step-by-Step via Erasing Objects from Images*. ECCV 2024, h5-index 197.

C13 Sumit Sarin, Utkarsh Mall, Purva Tendulkar, **Carl Vondrick**. *How Video Meetings Change Your Expression*. ECCV 2024, h5-index 197.

C14 Sruthi Sudhakar, Ruoshi Liu, Basile Van Hoorick, **Carl Vondrick**, Richard Zemel. *Controlling the World by Sleight of Hand*. ECCV 2024, h5-index 197.
Oral presentation

C15 Basile Van Hoorick, Rundi Wu, Ege Ozguroglu, Kyle Sargent, Ruoshi Liu, Pavel Tokmakov, Achal Dave, Changxi Zheng, **Carl Vondrick**. *Generative Camera Dolly: Extreme Monocular Dynamic Novel View Synthesis*. ECCV 2024, h5-index 197.
Oral presentation

C16 Mia Chiquier, Utkarsh Mall, **Carl Vondrick**. *Evolving Interpretable Visual Classifiers with Large Language Models*. ECCV 2024, h5-index 197.

C17 Haozhe Chen, **Carl Vondrick**, Chengzhi Mao. *SelfIE: Self-Interpretation of Large Language Model Embeddings*. ICML 2024, h5-index 237.

C18 Ege Ozguroglu, Ruoshi Liu, Dídac Surís, Dian Chen, Achal Dave, Pavel Tokmakov, **Carl Vondrick**. *pix2gestalt: Amodal Segmentation by Synthesizing Wholes*. CVPR 2024, h5-index 356.

C19 Chengzhi Mao, **Carl Vondrick**, Hao Wang, Junfeng Yang. *Raidar: geneRative AI Detection viA Rewriting*. ICLR 2024, h5-index 253.

C20 Haozhe Chen, Junfeng Yang, **Carl Vondrick**, Chengzhi Mao. *Interpreting and Controlling Vision Foundation Models via Text Explanations*. ICLR 2024, h5-index 253.

C21 Rundi Wu, Ruoshi Liu, **Carl Vondrick**, Changxi Zheng. *Sin3DM: Learning a Diffusion Model from a Single 3D Textured Shape*. ICLR 2024, h5-index 253.

C22 Utkarsh Mall, Cheng Perng Phoo, Meilin Liu, **Carl Vondrick**, Bharath Hariharan, Kavita Bala. *Remote Sensing Vision-Language Foundation Models without Annotations via Ground Remote Alignment*. ICLR 2024, h5-index 253.

C23 Matt Deitke et al. *Objaverse-XL: A Universe of 10M+ 3D Objects*. NeurIPS 2023, h5-index 245.

C24 Dídac Surís, Sachit Menon, **Carl Vondrick**. *ViperGPT: Visual Inference via Python Execution for Reasoning*. ICCV 2023, h5-index 184.
Oral presentation (5% acceptance rate)

C25 Ruoshi Liu, Rundi Wu, Basile Van Hoorick, Pavel Tokmakov, Sergey Zakharov, **Carl Vondrick**. *Zero-1-to-3: Zero-shot One Image to 3D Object*. ICCV 2023, h5-index 184.
Oral presentation

C26 Mia Chiquier, **Carl Vondrick**. *Muscles in Action*. ICCV 2023, h5-index 184.

C27 Arjun Mani, Ishaan Preetam Chandratreya, Elliot Creager, **Carl Vondrick**, Richard Zemel. *SurfsUp: Learning Fluid Simulation for Novel Surfaces*. ICCV 2023, h5-index 184.

C28 Ruoshi Liu, Chengzhi Mao, Purva Tendulkar, Hao Wang, **Carl Vondrick**. *Landscape Learning for Neural Network Inversion*. ICCV 2023, h5-index 184.

C29 Hongge Chen, Zhao Chen, Greg Meyer, Dennis Park, **Carl Vondrick**, Ashish Shrivastava, Yuning Chai. *SHIFT3D: Synthesizing Hard Inputs For Tricking 3D Detectors*. ICCV 2023, h5-index 184.

C30 Chengzhi Mao, Lingyu Zhang, Abhishek Joshi, Junfeng Yang, Hao Wang, **Carl Vondrick**. *Robust Perception through Equivariance*. ICML 2023, h5-index 237.

C31 Ruoshi Liu, **Carl Vondrick**. *Humans as Light Bulbs: 3D Human Reconstruction from Thermal Reflection*. CVPR 2023, h5-index 356.

C32 Ruoshi Liu, Sachit Menon, Chengzhi Mao, Dennis Park, Simon Stent, **Carl Vondrick**. *What You Can Reconstruct from a Shadow*. CVPR 2023, h5-index 356.

C33 Basile Van Hoorick, Pavel Tokmakov, Simon Stent, Jie Li, **Carl Vondrick**. *Tracking through Containers and Occluders in the Wild*. CVPR 2023, h5-index 356.

C34 Purva Tendulkar, Dídac Surís, **Carl Vondrick**. *FLEX: Full-Body Grasping Without Full-Body Grasps*. CVPR 2023, h5-index 356.

C35 Chengzhi Mao, Revant Teotia, Amrutha Sundar, Sachit Menon, Junfeng Yang, Xin Wang, **Carl Vondrick**. *Doubly Right Object Recognition: A Why Prompt for Visual Rationales*. CVPR 2023, h5-index 356.

C36 Sachit Menon, **Carl Vondrick**. *Visual Classification via Description from Large Language Models*. ICLR 2023, h5-index 253.
Oral presentation (5% acceptance rate)

C37 Chengzhi Mao, Scott Geng, Junfeng Yang, Xin Wang, **Carl Vondrick**. *Understanding Zero-Shot Adversarial Robustness for Large-Scale Models*. ICLR 2023, h5-index 253.

C38 Hui Lu, Mia Chiquier, **Carl Vondrick**. *Private Multiparty Perception for Navigation*. NeurIPS 2022, pp. 3318–3328, h5-index 245.

C39 Dídac Surís, **Carl Vondrick**. *Representing Spatial Trajectories as Distributions*. NeurIPS 2022, pp. 13731–13744, h5-index 245.

C40 Sachit Menon, David Blei, **Carl Vondrick**. *Forget-me-not! Contrastive Critics for Mitigating Posterior Collapse*. UAI 2022, pp. 1360–1370.

C41 Basile Van Hoorick, Purva Tendulkar, Dídac Surís, Dennis Park, Simon Stent, **Carl Vondrick**. *Revealing Occlusions with 4D Neural Fields*. CVPR 2022, pp. 3011–3021, h5-index 356.
Oral presentation (3% acceptance rate)

C42 Dídac Surís, Dave Epstein, **Carl Vondrick**. *Globetrotter: Connecting Languages by Connecting Images*. CVPR 2022, pp. 16474–16484, h5-index 356.
Oral presentation (3% acceptance rate)

C43 Chengzhi Mao, Kevin Xia, James Wang, Hao Wang, Junfeng Yang, Elias Bareinboim, **Carl Vondrick**. *Causal Transportability for Visual Recognition*. CVPR 2022, pp. 7521–7531, h5-index 356.

C44 Dídac Surís, **Carl Vondrick**, Bryan Russell, Justin Salamon. *It's Time for Artistic Correspondence in Music and Video*. CVPR 2022, h5-index 356.

C45 Will Price, **Carl Vondrick**, Dima Damen. *UnweaveNet: Unweaving Activity Stories*. CVPR 2022, pp. 13770–13779, h5-index 356.

C46 Xingyu Fu, Ben Zhou, Ishaan Preetam Chandratreya, **Carl Vondrick**, Dan Roth. *There is a Time and Place for Reasoning Beyond the Image*. ACL 2022, h5-index 157.
Oral presentation

C47 Mia Chiquier, Chengzhi Mao, **Carl Vondrick**. *Real-Time Neural Voice Camouflage*. ICLR 2022, h5-index 253.
Oral presentation (1% acceptance rate)

C48 Chengzhi Mao, Lu Jiang, Mostafa Dehghani, **Carl Vondrick**, Rahul Sukthankar, Irfan Essa. *Discrete Representations Strengthen Vision Transformer Robustness*. ICLR 2022, h5-index 253.

C49 Boyuan Chen, Mia Chiquier, Hod Lipson, **Carl Vondrick**. *The Boombox: Visual Reconstruction from Acoustic Vibrations*. CoRL 2021, pp. 1067–1077.

C50 Chengzhi Mao, Mia Chiquier, Hao Wang, Junfeng Yang, **Carl Vondrick**. *Adversarial Attacks are Reversible with Natural Supervision*. ICCV 2021, pp. 661–671, h5-index 184.

C51 Basile Van Hoorick, **Carl Vondrick**. *Dissecting Image Crops*. ICCV 2021, pp. 9741–9750, h5-index 184.

C52 Dídac Surís, Ruoshi Liu, **Carl Vondrick**. *Learning the Predictability of the Future*. CVPR 2021, pp. 12607–12617, h5-index 356.

C53 Chengzhi Mao, Amogh Gupta, Augustine Cha, Hao Wang, Junfeng Yang, **Carl Vondrick**. *Generative Interventions for Causal Learning*. CVPR 2021, pp. 3947–3956, h5-index 356.

C54 Dave Epstein, **Carl Vondrick**. *Learning Goals from Failure*. CVPR 2021, pp. 11194–11204, h5-index 356.

C55 Ruilin Xu, Rundi Wu, Yuko Ishiwaka, **Carl Vondrick**, Changxi Zheng. *Listening to Sounds of Silence for Speech Denoising*. NeurIPS 2020, pp. 9633–9648, h5-index 245.

C56 Chengzhi Mao, Amogh Gupta, Vikram Nitin, Baishakhi Ray, Shuran Song, Junfeng Yang, **Carl Vondrick**. *Multitask Learning Strengthens Adversarial Robustness*. ECCV 2020, pp. 158–174, h5-index 197.
Oral presentation (2% acceptance rate)

C57 Alex Andonian, Camilo Fosco, Mathew Monfort, Allen Lee, **Carl Vondrick**, Rogerio Feris. *We Have So Much In Common: Modeling Semantic Relational Set Abstractions in Videos*. ECCV 2020, pp. 18–34, h5-index 197.

C58 Dídac Surís, Dave Epstein, Heng Ji, Shih-Fu Chang, **Carl Vondrick**. *Learning to Learn Words from Visual Scenes*. ECCV 2020, pp. 434–452, h5-index 197.

C59 Dave Epstein, Boyuan Chen, **Carl Vondrick**. *Oops! Predicting Unintentional Action in Video*. CVPR 2020, pp. 919–929, h5-index 356.

C60 Chengzhi Mao, Ziyuan Zhong, Junfeng Yang, **Carl Vondrick**, Baishakhi Ray. *Metric Learning for Adversarial Robustness*. NeurIPS 2019, pp. 480–491, h5-index 245.

C61 Chen Sun, Austin Myers, **Carl Vondrick**, Kevin Murphy, Cordelia Schmid. *VideoBERT: A Joint Model for Video and Language Representation Learning*. ICCV 2019, pp. 7464–7473, h5-index 184.

C62 Hassan Akbari, Svebor Karaman, Surabhi Bhargava, Brian Chen, **Carl Vondrick**, Shih-Fu Chang. *Multi-level Multimodal Common Semantic Space for Image-Phrase Grounding*. CVPR 2019, pp. 12476–12486, h5-index 356.

C63 Chen Sun, Abhinav Shrivastava, **Carl Vondrick**, Rahul Sukthankar, Kevin Murphy, Cordelia Schmid. *Relational Action Forecasting*. CVPR 2019, pp. 273–283, h5-index 356.

C64 **Carl Vondrick**, Abhinav Shrivastava, Alireza Fathi, Sergio Guadarrama, Kevin Murphy. *Tracking Emerges by Colorizing Videos*. ECCV 2018, pp. 391–408, h5-index 197.

C65 Hang Zhao, Chuang Gan, Andrew Rouditchenko, **Carl Vondrick**, Josh McDermott, Antonio Torralba. *The Sound of Pixels*. ECCV 2018, h5-index 197.

C66 Chen Sun, Abhinav Shrivastava, **Carl Vondrick**, Kevin Murphy, Rahul Sukthankar, Cordelia Schmid. *Actor-centric Relation Network*. ECCV 2018, pp. 318–334, h5-index 197.

C67 Chunhui Gu et al. *AVA: A Video Dataset of Spatio-temporally Localized Atomic Visual Actions*. CVPR 2018, pp. 6047–6056, h5-index 356.
Spotlight presentation

C68 Adria Recasens, **Carl Vondrick**, Aditya Khosla, Antonio Torralba. *Following Gaze in Video*. ICCV 2017, pp. 1435–1443, h5-index 184.

C69 **Carl Vondrick**, Antonio Torralba. *Generating the Future with Adversarial Transformers*. CVPR 2017, pp. 1020–1028, h5-index 356.

C70 **Carl Vondrick**, Hamed Pirsiavash, Antonio Torralba. *Generating Videos with Scene Dynamics*. NeurIPS 2016, pp. 613–621, h5-index 245.

C71 Yusuf Aytar, **Carl Vondrick**, Antonio Torralba. *SoundNet: Learning Sound Representations from Unlabeled Video*. NeurIPS 2016, pp. 892–900, h5-index 245.

C72 **Carl Vondrick**, Hamed Pirsiavash, Antonio Torralba. *Anticipating Visual Representations from Unlabeled Video*. CVPR 2016, h5-index 356.
Spotlight presentation

C73 **Carl Vondrick**, Deniz Oktay, Hamed Pirsiavash, Antonio Torralba. *Predicting Motivations of Actions by Leveraging Text*. CVPR 2016, h5-index 356.

C74 Lluis Castrejon, Yusuf Aytar, **Carl Vondrick**, Hamed Pirsiavash, Antonio Torralba. *Learning Aligned Cross-Modal Representations from Weakly Aligned Data*. CVPR 2016, pp. 2940–2949, h5-index 356.

C75 **Carl Vondrick**, Hamed Pirsiavash, Aude Oliva, Antonio Torralba. *Learning Visual Biases from Human Imagination*. NeurIPS 2015, pp. 289–297, h5-index 245.

C76 Adria Recasens, Aditya Khosla, **Carl Vondrick**, Antonio Torralba. *Where are they looking?* NeurIPS 2015, h5-index 245.

C77 Hamed Pirsiavash, **Carl Vondrick**, Antonio Torralba. *Assessing the Quality of Actions*. ECCV 2014, pp. 556–571, h5-index 197.

C78 **Carl Vondrick**, Aditya Khosla, Tomasz Malisiewicz, Antonio Torralba. *HOGgles: Visualizing Object Detection Features*. ICCV 2013, pp. 1–8, h5-index 184.
Oral presentation (3% acceptance rate)

C79 Xiangxin Zhu, **Carl Vondrick**, Deva Ramanan, Charless C. Fowlkes. *Do We Need More Training Data or Better Models for Object Detection?* BMVC 2012, vol. 3, no. 5.

C80 **Carl Vondrick**, Deva Ramanan. *Video Annotation and Tracking with Active Learning*. NeurIPS 2011, pp. 28–36, h5-index 245.

C81 Sangmin Oh et al. *A Large-scale Benchmark Dataset for Event Recognition*. CVPR 2011, pp. 3153–3160, h5-index 356.

C82 **Carl Vondrick**, Deva Ramanan, Donald Patterson. *Efficiently Scaling Up Video Annotation with Crowdsourced Marketplaces*. ECCV 2010, pp. 610–623, h5-index 197.

Journal Papers (Peer Reviewed)

J1 Boyuan Chen, Robert Kwiatkowski, **Carl Vondrick**, Hod Lipson. *Full-Body Visual Self-Modeling of Robot Morphologies*. Science Robotics 2022, vol. 8.

J2 Boyuan Chen, **Carl Vondrick**, Hod Lipson. *Visual Behavior Modelling for Robotic Theory of Mind*. Scientific Reports 2021, vol. 11, pp. 1–14, h5-index 200.

J3 Mathew Monfort et al. *Moments in Time Dataset: one million videos for event understanding*. PAMI 2019, vol. 42, pp. 502–508, h5-index 149.

J4 Yusuf Aytar, Lluis Castrejon, **Carl Vondrick**, Hamed Pirsiavash, Antonio Torralba. *Cross-Modal Scene Networks*. PAMI 2017, vol. 40, pp. 2303–2314, h5-index 149.

J5 **Carl Vondrick**, Aditya Khosla, Hamed Pirsiavash, Tomasz Malisiewicz, Antonio Torralba. *Visualizing Object Detection Features*. IJCV 2016, pp. 1–8.

J6 Xiangxin Zhu, **Carl Vondrick**, Charless C. Fowlkes, Deva Ramanan. *Do We Need More Training Data?* IJCV 2015, vol. 3, no. 5.

J7 **Carl Vondrick**, Donald Patterson, Deva Ramanan. *Efficiently Scaling Up Crowdsourced Video Annotation*. IJCV 2012, vol. 101, pp. 184–204.

Preprints and Technical Reports

P1 Sruthi Sudhakar, Junbang Liang, Sreehari Rammohan, Pavel Tokmakov, Richard Zemel, **Carl Vondrick**. *Robot Critics that Sweat the Small Stuff*. arXiv 2026.

P2 Ege Ozguroglu, Junbang Liang, Ruoshi Liu, Mia Chiquier, Michael DeTienne, Wesley Wei Qian, Alexandra Horowitz, Andrew Owens, **Carl Vondrick**. *New York Smells: A Large Multimodal Dataset for Olfaction*. arXiv 2025.

P3 Junbang Liang, Pavel Tokmakov, Ruoshi Liu, Sruthi Sudhakar, Paarth Shah, Rares Ambrus, **Carl Vondrick**. *Video Generators are Robot Policies*. arXiv 2025.

P4 Chia Hsiang Kao, Wenting Zhao, Shreelekha Revankar, Samuel Speas, Snehal Bhagat, Rajeev Datta, Cheng Perng Phoo, Utkarsh Mall, **Carl Vondrick**, Kavita Bala, Bharath Hariharan. *Towards LLM Agents for Earth Observation*. arXiv 2025.

P5 Mia Chiquier, Orr Avrech, Yossi Gandelsman, Berthy Feng, Katherine Bouman, **Carl Vondrick**. *Teaching Humans Subtle Differences with DIFF-usion*. arXiv 2025.

P6 Ruoshi Liu, Junbang Liang, Sruthi Sudhakar, Huy Ha, Cheng Chi, Shuran Song, **Carl Vondrick**. *PaperBot: Learning to Design Real-World Tools Using Paper*. arXiv 2024.

P7 Scott Geng, Revant Teotia, Purva Tendulkar, Sachit Menon, **Carl Vondrick**. *Affective Faces for Goal-Driven Dyadic Communication*. arXiv 2023.

P8 Lingyu Zhang, Chengzhi Mao, Junfeng Yang, **Carl Vondrick**. *Adversarially Robust Video Perception by Seeing Motion*. arXiv 2022.

P9 Sachit Menon, Ishaan Preetam Chandratreya, **Carl Vondrick**. *Task Bias in Vision-Language Models*. arXiv 2022.

P10 Yusuf Aytar, **Carl Vondrick**, Antonio Torralba. *See, Hear, and Read: Deep Aligned Representations*. arXiv 2017.

Teaching

Semester	**Course**
Fall 2018	COMS W4731 Computer Vision I
Spring 2019	COMS E6998 Advanced Computer Vision
Fall 2019	COMS W4731 Computer Vision I
Fall 2020	COMS E6998 Representation Learning
Summer 2021	COMS W4732 Computer Vision II
Fall 2021	COMS E6998 Representation Learning
Spring 2022	COMS W4732 Computer Vision II
Fall 2022	COMS E6998 Representation Learning
Spring 2023	COMS W4732 Computer Vision II
Spring 2024	COMS W4732 Computer Vision II
Fall 2024	COMS E6998 Machine Learning Frontiers
Spring 2025	COMS W4732 Computer Vision II
Fall 2025	COMS E6998 Machine Learning Frontiers

Additional Teaching and Tutorials

2023	Lecture on AI, Computer Vision, and Machine Learning Columbia+ and Columbia Engineering Executive Education; École Polytechnique.
2022–2023	Lecture on Computational Imaging and Vision Dual MBA/Executive MS: Engineering and Applied Science. Co-taught with Shree Nayar.
2022	Computer Vision and Robotics The Online AI Program from Columbia Engineering. Co-taught with Hod Lipson.
2018	Tutorial on Unsupervised Visual Learning International Computer Vision Summer School (ICVSS).

2018 Tutorial on Unsupervised Visual Learning
IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Co-taught with Anelia Angelova and Pierre Sermanet.

Mentoring and Team

Postdoctoral Researchers

2023–2025 **Utkarsh Mall**
Next position: Assistant Professor at MBZUAI

2023–2024 **Chengzhi Mao**
Next position: Assistant Professor at Rutgers

Ph.D. Students Completed

2020–2026 **Sachit Menon**
Thesis: Trustworthy Multimodal Intelligence
Fellowships: CAIRFI Ph.D. Fellowship, National Science Foundation Graduate Research Fellowship, Presidential Fellowship
Next position: Member of Technical Staff at Anthropic

2020–2025 **Mia Chiquier**
Thesis: Discovery Through Bottlenecks in Multimodal Models
Fellowships: Amazon CAIT Ph.D. Fellowship
Next position: Research Scientist at Mistral

2021–2025 **Ruoshi Liu**
Thesis: Generative Computer Vision for Physical Intelligence
Next position: Assistant Professor at UMD

2021–2024 **Basile Van Hoorick**
Thesis: Spatial Reasoning in Dynamic Scenes
Next position: Research Scientist at TRI

2019–2024 **Dídac Surís**
Thesis: Multimodal Representations for Video
Fellowships: Microsoft Ph.D. Fellowship
Next position: Research Scientist at Meta

2018–2023 **Chengzhi Mao**
Thesis: Robust Machine Learning by Integrating Context
Co-advised with: Junfeng Yang
Next position: Assistant Professor at Rutgers

Ph.D. Students in Progress

2021–Present **Arjun Mani**
Co-advised with: Richard Zemel
Fellowships: National Science Foundation Graduate Research Fellowship

2023–Present **Junbang Liang**

2023–Present **Lennart Schulze**
Co-advised with: Matei Ciocarlie

2024–Present **Sreehari Rammohan**

2022–Present **Sruthi Sudhakar**
Co-advised with: Richard Zemel
Fellowships: Apple Ph.D. Fellowship, National Science Foundation Graduate Research Fellowship

Undergraduate Researchers

2019–2022 **Hui Lu**
Topic: Secure Multiparty Visual Recognition
Next position: Facebook

2019–2023 **Ishaan Preetam Chandratreya**
Topic: Predictive Models for Robotics
Next position: PhD in Computer Science at Massachusetts Institute of Technology

2022–2023 **Scott Geng**
Topic: Affective Human Activity Understanding
Awards: National Science Foundation Graduate Research Fellowship
Next position: PhD in Computer Science at University of Washington

2020–2021 **Jillian Ross**
Topic: Visually Grounded Language Models
Next position: PhD in Computer Science at Massachusetts Institute of Technology

2019–2021 **Ruoshi Liu**
Topic: Hyperbolic Machine Learning
Next position: PhD in Computer Science at Columbia University

2018–2020 **Dave Epstein**
Topic: Recognizing Human Activity and Goals from Video
Awards: CRA Honorable Mention
Next position: PhD in Computer Science at University of California, Berkeley

Masters Students

2021–2023 **Revant Teotia**
Topic: Explainable Visual Recognition
Next position: PhD in Computer Science at New York University

2019–2021 **Amogh Gupta**
Topic: Robust Visual Representations
Next position: Amazon Research

Ph.D. Committees (excluding my own students)

Defense	Student	Advisor
Jan 2021	Alireza Zareian	Shih-Fu Chang
Mar 2021	Iretiayo Akinola	Peter Allen
Apr 2021	Chang Xiao	Changxi Zheng
Jul 2021	Yuchi Tian	Baishakhi Ray
Oct 2021	Boxi Xian	Hod Lipson
Jan 2022	Hassan Akbari	Shih-Fu Chang
Apr 2022	Boyuan Chen	Hod Lipson
May 2022	Rob Kwiatkowski	Hod Lipson
Jun 2022	Simone Fobi	Vijay Modi
Jun 2022	Terry Conlon	Vijay Modi
Aug 2022	Henrique Maia	Eitan Grinspun
Dec 2022	Sai Saketh Rambhatla	Abhinav Shrivastava
Feb 2023	Emily Hannigan	Matei Ciocarlie
Feb 2023	Brian Chen	Shih-Fu Chang
Nov 2023	Xudong Lin	Shih-Fu Chang
Dec 2023	Ziyuan Zhong	Baishakhi Ray
Jan 2024	Shiyuan Huang	Shih-Fu Chang

Defense	Student	Advisor
Feb 2024	Ruilin Xu	Shree Nayar
Apr 2024	Kelly Buchanan	John Cunningham and Liam Paninski
Apr 2024	Zhenjia Xu	Shuran Song
May 2024	Cheng Chi	Shuran Song
Jul 2024	Haoxuan You	Shih-Fu Chang
Jul 2024	Charles Carver	Xia Zhou
Aug 2024	Zhecan Wang	Shih-Fu Chang
Aug 2024	Yookoon Park	David Blei
Apr 2025	Yuhang Hu	Hod Lipson
May 2025	Hammad Ayyubi	Shih-Fu Chang
May 2025	Samir Gadre	Shuran Song
Aug 2025	Jingxi Xu	Matei Ciocarlie
Aug 2025	Mengda Xu	Shuran Song and Manuela Veloso
Sep 2025	Jingyi Shi	David Blei
Jul 2026	Huy Ha	Shuran Song
Aug 2026	Judah Goldfeder	Hod Lipson

Funding

NSF is the National Science Foundation; DARPA is the Defense Advanced Research Projects Agency.

Center Grants

2023–2027 **Artificial and Natural Intelligence Institute**
Program: NSF, National AI Institute
Role: Senior Personnel
PI: Richard Zemel
Institutions: Columbia (lead), CUNY, MILA, Google, Tuskegee, Yale, Princeton, UPenn

2022–2027 **Center for Smart Streetscapes**
Program: NSF, Engineering Research Center
Role: Co-lead, situational awareness and computer vision
PI: Andrew Smyth
Institutions: Columbia (lead), Florida Atlantic, Rutgers, University of Central Florida

2021–2026 **Learning the Earth with Artificial Intelligence and Physics**
Program: NSF, Science and Technology Center
Role: Director of Data Science
PI: Pierre Gentine
Institutions: Columbia (lead), NYU, UCI, NCAR, Minnesota, NASA, MILA, Teachers College

Gifts and Awards

2023 **Google, Research Gift**
Role: PI (sole)

2021–2026 **Spatial Awareness for Machine Perception**
Program: NSF, CAREER; *Role:* PI (sole)

2021–2024 **Self-supervised Learning for Machine Situational Awareness**
Program: Toyota Research Institute, Young Faculty Award; *Role:* PI (sole)

2022 **Google, Research Gift**
Role: PI (sole)

2022 **Adobe, Research Gift**
Role: PI (sole)

2018–2019 **Anticipating Human Behaviors from Unlabeled Video**
Program: Amazon, Research Award; *Role:* PI (sole)

Standard Grants

2024–2028 **Programmatic Foundation Models for Visual Analysis on a Planetary Scale**
Program: NSF, Robust Intelligence
Role: Co-PI (site lead)
Team: Kavita Bala (PI), Bharath Hariharan (Co-PI), Carl Vondrick (Co-PI)

2024–2026 **Open-world Predictive Models**
Program: Toyota Research Institute
Role: PI (sole)

2024–2026 **Learning and Executing Programs for Robot Manipulation**
Program: Toyota Research Institute
Role: PI
Team: Carl Vondrick (PI), Yunzhu Li (PI)
Institutions: Columbia, UIUC

2023–2026 **MIRACLE: Multimodal Interactive Conceptual Learning**
Program: DARPA, Environment-driven Conceptual Learning
Role: Co-PI (site lead)
PI: Heng Ji
Institutions: UIUC (lead), Columbia, UCLA, UNC Chapel Hill

2022–2025 **Seeing Science: Using Computer Vision to Explore the Scientific Principles Behind Everyday Objects**
Program: NSF, Emerging Technologies for Teaching and Learning
Role: Co-PI
Team: Lydia Chilton (PI), Carl Vondrick (Co-PI), Paulo Blikstein (Co-PI)
Institutions: Columbia (lead), Teachers College

2022–2025 **Cross-Cultural Harmony through Affect and Response Mediation**
Program: DARPA, Computational Cultural Understanding
Role: Co-PI
PI: Kathleen McKeown
Institutions: Columbia (lead), NYU, UC Davis, Stony Brook

2021–2024 **Hierarchical Representation Learning for Robot Assistants**
Program: NSF, National Robotics Initiative
Role: Co-PI
Team: Shuran Song (PI), Carl Vondrick (Co-PI), Zhou Yu (Co-PI)

2019–2024 **Reasoning About Events with Knowledge**
Program: DARPA, Knowledge-directed Artificial Intelligence Reasoning
Role: Co-PI
PI: Heng Ji
Institutions: UIUC (lead), Columbia, UPenn, UNC Chapel Hill, UColorado Boulder

2019–2023 **Quantifying and Synthesizing Novelty for Videos and Images**
Program: DARPA, Science of Artificial Intelligence and Learning of Novelty
Role: Co-PI (site lead)
Team: Abhinav Shrivastava (PI), Carl Vondrick (Co-PI), Abhinav Gupta
Institutions: UMD (lead), Columbia, CMU

2019–2024 **Multi-modal Embeddings for Machine Commonsense**
Program: DARPA, Machine Commonsense
Role: Co-PI
PI: Ralph Weischedel
Institutions: USC (lead), Columbia, UMass Amherst, UCLA

2019–2024 **Learning Visual Dynamics from Interaction**
Program: NSF, National Robotics Initiative
Role: PI (lead)
Team: Carl Vondrick (PI), Hod Lipson (Co-PI)

2019–2022 **Learning Visually Grounded Language Models**
Program: DARPA, Grounded Artificial Intelligence Language Acquisition
Role: PI (lead)
Team: Carl Vondrick (PI), Shih-Fu Chang (Co-PI), Heng Ji (Co-PI)
Institutions: Columbia (lead), UIUC

2019–2022 **Learning Predictive Representations from Unlabeled Video**
Program: NSF, CRII
Role: PI (sole)

Patents

Granted

Dec 2021 Visual Tracking by Colorization
Abhinav Shrivastava, Alireza Fathi, Sergio Guadarrama, Kevin Murphy, Carl Vondrick.

Pending

System and Method for 3D Reconstruction from Shadows
Ruoshi Liu, Sachit Menon, Chengzhi Mao, Dennis Park, Simon Stent, Carl Vondrick.

A Neural Network Model for Speech Denoising
Changxi Zheng, Ruilin Xu, Rundi Wu, Carl Vondrick, Yuko Ishikawa.

Action Localization using Relational Features
Chen Sun, Abhinav Shrivastava, Cordelia Schmid, Rahul Sukthankar, Kevin Murphy, Carl Vondrick.

Invited Talks and Visiting Lectures

Creative Computer Vision Research

Jun 2025 How to Stand Out in the Crowd? Workshop, at CVPR

Embodied Generative AI

Jun 2025 3D Scene Understanding Workshop, at CVPR
Nov 2024 New York University
Nov 2024 University of Pennsylvania

Making Sense of the Multimodal World

Mar 2024 University of Pittsburgh
Mar 2024 Stanford University
Nov 2023 Cornell University
Nov 2023 Cornell Tech

Nov 2023 Carnegie Mellon University

The Rise of Visual Skills in Large Models

Oct 2023 GeoNet: Unsupervised Adaptation across Geographies, at ICCV
Oct 2023 International Challenge on Compositional and Multimodal Perception, at ICCV
Oct 2023 BigMAC: Big Model Adaptation for Computer Vision, at ICCV
Oct 2023 Workshop on Large-scale Video Object Segmentation, at ICCV
Jun 2023 University of California, Berkeley
Jun 2023 T4V: Transformers for Vision, at CVPR
Jun 2023 Holistic Video Understanding Workshop, at CVPR
May 2023 Stanford University

Emergent 3D Vision

Oct 2023 Visual Object Tracking Challenge, at ICCV
Oct 2023 Frontiers of Monocular 3D Perception, at ICCV

Connecting Vision, Language, and Code

Apr 2023 Data Science Day, at Columbia University

Visual Recognition by Reading

Oct 2022 Workshop on Out of Distribution Generalization, at ECCV
Oct 2022 Workshop on Compositional and Multimodal Perception, at ECCV

Learning the Predictability of the Future

Jul 2022 Samsung AI Research
Jul 2022 Aibee
Jun 2022 Workshop on Robustness in Sequential Data, at CVPR
Feb 2022 Google Research
Oct 2021 Workshop on Applications of Signal Processing to Audio and Acoustics
Sep 2021 Cruise
Feb 2021 Stanford University
Feb 2021 Massachusetts Institute of Technology
Feb 2021 Princeton University
Feb 2021 International Business Machines (IBM)

Audio Privacy

Jun 2022 Workshop on Sound and Vision, at CVPR

Inverting the Neural Network

Jun 2022 Workshop on Visual Perception and Learning in an Open World, at CVPR
Mar 2022 Massachusetts Institute of Technology
Dec 2021 Workshop on Dealing with the Novelty in Open Worlds, at WACV

Learning from Unlabeled Video

Apr 2020 New York University
Mar 2020 Carnegie Mellon University
Mar 2020 University of Pittsburgh

Nov 2019 University of Massachusetts, Amherst
Nov 2018 Butterfly Network
Mar 2018 University of Maryland, College Park

Predictive Vision

Nov 2017 University of Pennsylvania
Nov 2017 Snapchat Research
Nov 2017 University of Southern California
Nov 2017 Workshop on Video Frontiers
May 2017 Rework Summit
Apr 2017 University of California, San Diego
Apr 2017 Cornell University
Mar 2017 University of Texas, Austin
Mar 2017 Columbia University
Mar 2017 Google Research
Mar 2017 Adobe Research
Mar 2017 OpenAI
Feb 2017 Brown University
Feb 2017 University of California, Los Angeles
Feb 2017 NVIDIA
Nov 2016 Rework Summit
Oct 2016 Twitter
Sep 2016 Toyota Technological Institute at Chicago
Sep 2016 Massachusetts Institute of Technology
Aug 2016 Apple
Aug 2016 University of California, Berkeley
Aug 2016 Stanford University
Mar 2016 Boston University
Mar 2016 University of Massachusetts, Boston

Visualizing Object Detection Features

Mar 2016 University of Massachusetts, Boston
Sep 2015 Massachusetts Institute of Technology
Nov 2013 Brown University

Efficient Video Annotation

Jun 2013 CVPR Workshop
Jun 2011 CVPR Workshop

Service

Department Service

2024–2025 **Computing and Facilities Committee, Computer Science**
Chair: 2024, 2025

2023–2024 **Hiring Committee, Computer Science**
Member: 2023, 2024

2019–2023 **PhD Admissions Committee, Computer Science**
Vice Chair: 2021–2023; *Member:* 2019–2021

2018–Present **Distinguished Lectures Committee, Computer Science**
Member: 2018–Present

University Service

2024, 2025 **Empire AI Committee**

2019, 2021 **Reviewer, Ad Hoc Committees for Fellowship Selection**

2021 **Speaker, Data Science Institute Council Annual Meeting**

2021 **Speaker, NIH Workshop at Mailman School of Public Health**

Service to the Discipline

2025–2030 **Board Member, ICLR**

2026 **General Chair, ICLR**

2025 **Senior Program Chair, ICLR**

2018–2024 **Area Chair**
CVPR: 2018, 2019, 2021, 2022, 2024
ICCV: 2021, 2023
ECCV: 2022
NeurIPS: 2019, 2020, 2023
ICLR: 2020

2020–2025 **Review Panelist, National Science Foundation**
IIS: 2020, 2021, 2022, 2023, 2025
NRI: 2020
SBIR: 2020

2011–Present **Reviewer for CVPR, ICCV, ECCV, NeurIPS, ICML, TPAMI and IJCV**

2021, 2022 **Organizer, Workshop on Open-world Vision, CVPR**

2019–2021 **Organizer, Workshop on Learning from Unlabeled Video, CVPR**

2018–2019 **Organizer, Workshop on Self-supervised Learning**
ICML: 2019
CVPR: 2018

Media Coverage

Television and Radio

Outlet	Story
NPR	Computer Binge-Watched TV And Learned To Predict
CNN	New AI Can Predict When Two People Will Kiss
Late Show with Stephen Colbert	Television clip on human action prediction
NPR Marketplace	Algorithms Identify Audio through Video Footage
CBC	Teaching Software to Predict Handshakes, Hugs, and Kisses

Newspapers and Magazines

Outlet	Story
Science	Is Technology Spying on You? New AI Could Prevent Eavesdropping

Outlet	Story
Associated Press (AP)	How Do You Teach Human Interaction to a Robot? Lots of TV
NBC News	Deep Learning: Teaching Computers to Predict the Future
Wired	This AI learned to predict the future by watching loads of TV
Scientific American	Artificial Intelligence Can Predict How Scenes Will Play Out
New Scientist	Binge-watching videos teaches computers to recognise sounds
Week Junior	A machine that learns by listening (children's magazine)
Forbes	MIT Computers Binge-Watch To Learn About Hugs
Newsweek	Artificial Intelligence Algorithms Predicts the Future
ABC News	New AI Can Predict When Two People Will Kiss
Fox News	New Artificial Intelligence Can Predict When You Will Kiss
Popular Science	Algorithm Binge Watches TV to Predict Human Behavior
New Scientist	AI learns to predict the future by watching 2 million videos
Vice Magazine	This Algorithm Taught Itself to Animate a Still Photo
The Verge	Machine Learning's Next Trick is Generating Videos from Photos
Technology Review	Image Experiment Reveals The Building Blocks of Imagination