

KATERINA FRAGKIADAKI

Associate Professor

Machine Learning Department, Carnegie Mellon University

5000 Forbes Avenue, Pittsburgh, PA, 15213

Tel:2675289476

katef@cs.cmu.edu, www.cs.cmu.edu/~katef

Employment

EECS, UC Berkeley	PostDoctoral Fellow,	2013–2015
Google Research	PostDoctoral Fellow,	Oct. 2015–December 2016
Assistant Professor	MLD, CMU	September 2016–July 2023
Associate Professor	MLD, CMU	July 2023–present

Education

National Technical University of Athens	EECS	Diplomat 2007
University of Pennsylvania	CIS	M.S. 2011
University of Pennsylvania	CIS	PhD 2013

Selected Publications

Peer-reviewed Publications

- CVPR, ICCV and ECCV are top tier conferences in Computer Vision and Machine Learning with H-index 158, 89, 98, respectively, and acceptance rates $\sim 25 - 30\%$.
- NeurIPS is a top tier conference in Machine Learning with H-index 101 and acceptance rate $\sim 20\%$.
- ICLR is a top tier conference in Deep Learning with acceptance rate $\sim 30\%$.
- AAAI is a top tier conference in Artificial Intelligence with acceptance rate in 2019 (when we submitted) 16%.
- CoRL is the main (and top-tier) conference in the intersection of Robotics and Machine Learning with acceptance rate $\sim 30\%$.
- RSS is an established conference in Robotics with acceptance rate $\sim 20 - 30\%$.
- ICRA is an established conference in Robotics with acceptance rate $\sim 40\%$.

(statistics obtained from [conference rankings](#), [conference acceptance rates](#))

1. Wen-Hsuan Chu*, Lei Ke*, **Katerina Fragkiadaki**, DreamScene4D: Dynamic Multi-Object Scene Generation from Monocular Videos, NeurIPS 2024
2. Gabriel Sarch, Lawrence Jang, Michael J. Tarr, William W. Cohen, Kenneth Marino, **Katerina Fragkiadaki**, ICAL: Continual Learning of Multimodal Agents by Transforming Trajectories into Actionable Insights, NeurIPS 2024
3. Tsung-Wei Ke* , Nikolaos Gkanatsios*, **Katerina Fragkiadaki**, 3D Diffuser Actor: Policy Diffusion with 3D Scene Representations, CoRL 2024
4. Yufei Wang, Zhou Xian, Feng Chen, Tsun-Hsuan Wang, Yian Wang, **Katerina Fragkiadaki**, Zckory Erickson, David Held, Chuang Gan, 2024, RoboGen: Towards Unleashing Infinite Data for Automated Robot Learning via Generative Simulation. ICML 2024
5. Pushkal Katara, Xian Zhou and **Katerina Fragkiadaki**, 2023, Scaling Up Simulation Environments for Robotic Skill Learning with Foundational Generative Models, ICRA 2023
6. Ayush Jain, Pushkal Katara, Nikolaos Gkanatsios, Adam W. Harley, Gabriel Sarch, Kriti Aggarwal, Vishrav Chaudhary, **Katerina Fragkiadaki**, 2024, ODIN: A Single Model for 2D and 3D Segmentation, CVPR spotlight

7. Brian Yang, Huangyuan Su, Nikolaos Gkanatsios , Tsung-Wei Ke, Ayush Jain, Jeff Schneider and **Katerina Fragkiadaki**, 2024, Diffusion-ES: Gradient-free Planning with Diffusion for Autonomous Driving and Zero-Shot Instruction Following, *CVPR*
8. Tsung-Wei Ke and Nikolaos Gkanatsios and **Katerina Fragkiadaki**, 2024, 3D Diffuser Actor: Policy Diffusion with 3D Scene Representations, in submission
9. Wen-Hsuan Chu, Lei Ke and **Katerina Fragkiadaki**, 2024, DreamScene4D: Dynamic Multi-Object Scene Generation from Monocular Videos, in submission
10. Gabriel Sarch, William Cohen, Kenneth Marino, Michael J. Tarr and **Katerina Fragkiadaki**, 2024, Continual Learning of Multimodal Agents by Transforming Trajectories into Actionable Insights, in submission
11. Gabriel Sarch, Yue Wu, Michael J. Tarr, **Katerina Fragkiadaki**, Open-Ended Instructable Embodied Agents with Memory-Augmented Large Language Models, *EMNLP findings 2023*
12. Gabriel Sarch, Mike Tarr, 2023, **Katerina Fragkiadaki**, Leila Wehbe, 2023, Brain Dissection: fMRI-trained Networks Reveal Spatial Selectivity in the Processing of Natural Images, *NeurIPS*
13. Test-time Adaptation of Discriminative Models via Diffusion Generative Feedback Mihir Prabhudesai, Tsung-Wei Ke, Alexander Li, Deepak Pathak, and **Katerina Fragkiadaki**, 2023, *NeurIPS*
14. Gabe Sarch, Yue Wu, Mike Tarr and **Katerina Fragkiadaki**, 2023, Open-Ended Instructable Embodied Agents with Memory-Augmented Large Language Models, *EMNLP*
15. Theophile Gervet, Zhou Xian, Nikolaos Gkanatsios, and **Katerina Fragkiadaki**, 2023, Act3D: 3D Feature Field Transformers for Multi-Task Robotic Manipulation, *CoRL*
16. Zhou Xian, Nikolaos Gkanatsios, Theophile Gervet, and **Katerina Fragkiadaki**, 2023, Chained-Diffuser: Unifying Trajectory Diffusion and Keypose Prediction for Robotic Manipulation, *CoRL*
17. Nikolaos Gkanatsios, Ayush Jain, Zhou Xian, Yunchu Zhang, Christopher Atkeson, **Katerina Fragkiadaki**, 2023, Energy-based Models are Zero-Shot Planners for Compositional Scene Rearrangement, *RSS*
18. Mihir Prabhudesai, Anirudh Goyal, Sujoy Paul, Sjoerd van Steenkiste, Mehdi S. M. Sajjadi, Gaurav Aggarwal, Thomas Kipf, Deepak Pathak, and **Katerina Fragkiadaki**, 2023, Test-time Adaptation with Slot-Centric Models, *ICML*
19. Adam W. Harley, Zhaoyuan Fang, Jie Li, Rares Ambrus and **Katerina Fragkiadaki**, 2023, A Simple Baseline for BEV Perception Without LiDAR, *ICRA*
20. Xian Zhou, Fish Tung, Antonio Torralba **Katerina Fragkiadaki**, Chuang Gang, 2023, FluidLab: A Differentiable Environment for Benchmarking Complex Fluid Manipulation, *International Conference to Learning Representations(ICLR)*
21. Nikolaos Gkanatsios, Mayank Singh, Shubham Tulsiani, **Katerina Fragkiadaki**, 2023, Analogical Networks for Memory-Modulated 3D visual Parsing, *International Conference to Learning Representations(ICLR)*
22. Ayush Jain, Nikolaos Gkanatsios, Ishita Mediratta and **Katerina Fragkiadaki**, 2022, Bottom Up Top Down Detection Transformers for Language Grounding in Images and Point Clouds, *IEEE European Conference on Computer Vision (ECCV)*
23. Adam W Harley, Zhaoyuan Fang and **Katerina Fragkiadaki**, 2022, Particle Videos Revisited: Tracking Through Occlusions Using Point Trajectories, *IEEE European Conference on Computer Vision (ECCV)*

24. Gabriel Sarch, Zhaoyuan Fang, Adam W. Harley, Paul Schydlo, Michael J. Tarr, Saurabh Gupta, **Katerina Fragkiadaki**, 2022, TIDEE: Tidying Up Novel Rooms using Visuo-Semantic Commonsense Priors, *IEEE European Conference on Computer Vision (ECCV)*
25. Zhou Xian, Shamit Lal, Hsiao-Yu Tung, Emmanouil Antonios Platanios and **Katerina Fragkiadaki**, 2021, HyperDynamics: Meta-Learning Object and Agent Dynamics with Hypernetworks, *International Conference to Learning Representations(ICLR)*
26. Mihir Prabhudesai, Shamit Lal, Darshan Patil, Hsiao-Yu Tung, Adam W Harley and **Katerina Fragkiadaki**, 2021, Disentangling 3D Prototypical Networks for Few-Shot Concept Learning *International Conference to Learning Representations(ICLR)*
27. Zhaoyuan Fang, Ayush Jain, Gabriel Sarch, Adam W. Harley and **Katerina Fragkiadaki**, 2021, Move to See Better: Towards Self-Supervised Amodal Object Detection, *British Machine Vision Conference (BMVC)*
28. Shamit Lal, Mihir Prabhudesai, Ishita Mediratta, Adam W. Harley and **Katerina Fragkiadaki**, 2021, CoCoNets: Continuous Contrastive 3D Scene Representations, *Conference on Computer Vision and Pattern Recognition (CVPR)*
29. Adam W. Harley, Yiming Zuo, Jing Wen, Ayush Mangal, Shubhankar Potdar, Ritwick Chaudhry and **Katerina Fragkiadaki**, 2021, Track, Check, Repeat: An EM Approach to Unsupervised Tracking *Conference on Computer Vision and Pattern Recognition (CVPR)*
30. Jingyun Yang, Hsiao-Yu Tung, Yunchu Zhang, Gaurav Pathak, Ashwini Pokle, Christopher G Atkeson and **Katerina Fragkiadaki**, 2021, Visually-Grounded Library of Behaviors for Generalizing Manipulation Across Objects, Configurations and Views *Conference on Robotic Learning (CoRL)*
31. Hsiao-Yu Fish Tung, Zhou Xian, Mihir Prabhudesai, Shamit Lal, **Katerina Fragkiadaki**, 2020, 3D-OES: Viewpoint-Invariant Object-Factorized Environment Simulators, *Conference on Robotic Learning (CoRL)*
32. Yihui He, Rui Yan, **Katerina Fragkiadaki** and Shoou-I Yu, 2020, Epipolar Transformers *Conference on Computer Vision and Pattern Recognition (CVPR)*
33. Adam W. Harley, Shrinidhi K. Lakshmikanth, Paul Schydlo and **Katerina Fragkiadaki**, 2020, Tracking Emerges by Looking Around Static Scenes, with Neural 3D Mapping, *IEEE European Conference on Computer Vision (ECCV)*
34. Mihir Prabhudesai, Hsiao-Yu Fish Tung, Syed Ashar Javed, Maximilian Sieb, Adam W. Harley and **Katerina Fragkiadaki**, Embodied Language Grounding with Implicit 3D Visual Feature Representations, 2020, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
35. Adam W. Harley, Shrinidhi K. Lakshmikanth, Fangyu Li, Xian Zhou, Hsiao-Yu Fish Tung and **Katerina Fragkiadaki**, Learning from Unlabelled Videos Using Contrastive Predictive Neural 3D Mapping, 2020, *International Conference to Learning Representations(ICLR)*
36. Maximilian Sieb, Zhou Xian, Audrey Huang, Oliver Kroemer and **Katerina Fragkiadaki**, 2019, Graph-structured Visual Imitation, *Conference on Robotic Learning (CoRL)*
37. Adam W. Harley, Shih-En Wei, Jason Saragih and **Katerina Fragkiadaki**, 2019, Image Disentanglement and Uncooperative Re-entanglement for High-Fidelity Image-to-Image Translation *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*
38. Hsiao-Yu Fish Tung, Ricson Cheng and **Katerina Fragkiadaki**, 2019, Learning Spatial Common Sense with Geometry-Aware Recurrent Networks, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (oral)

39. Arpit Agarwal, Katharina Muelling and **Katerina Fragkiadaki**, 2019, Model Learning for Look-ahead Exploration in Continuous Control, *Association for the Advancement of Artificial Intelligence (AAAI)* (oral)
40. Maximilian Sieb and **Katerina Fragkiadaki**, 2018, Data Dreaming for Object Detection: Learning Object-Centric State Representations for Visual Imitation, *IEEE RAS International Conference on Humanoid Robots* (oral)
41. Ricson Cheng, Arpit Agarwal and **Katerina Fragkiadaki**, 2018, Reinforcement Learning of Active Vision for Manipulating Objects under Occlusions, *Conference on Robotic Learning (CoRL)* (spotlight)
42. Ricson Cheng, Ziyang Wang and **Katerina Fragkiadaki**, 2018, Geometry-Aware Recurrent Neural Networks for Active Visual Recognition, *Neural Information Processing Systems (NeurIPS)*
43. Hsiao-Yu Fish Tung, Adam W. Harley, Liang-Kang Huang and **Katerina Fragkiadaki**, 2018, Reward Learning from Narrated Demonstrations, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
44. Chris Ying and **Katerina Fragkiadaki**, 2017, Depth-adaptive Computational Policies for Efficient Visual Tracking, *Energy Minimization Methods in Computer Vision and Pattern Recognition (EMM-CVPR)*
45. Hsiao-Yu Fish Tung, Hsiao-Wei Tung, Ersin Yumer and **Katerina Fragkiadaki**, 2017, Self-supervised Learning of Motion Capture, *Neural Information Processing Systems (NeurIPS)* (spotlight)
46. Hsiao-Yu Fish Tung, Adam W. Harley, William Seto and **Katerina Fragkiadaki**, 2017, Adversarial Inverse Graphics Networks: Learning 2D-to-3D Lifting and Image-to-Image Translation from Unpaired Supervision, *International Conference on Computer Vision (ICCV)*
47. Jitendra Malik, Pablo Arbeláez, João Carreira, **Katerina Fragkiadaki**, Ross Girshick, Georgia Gkioxari, Saurabh Gupta, Bharath Hariharan, Abhishek Kar and Shubham Tulsiani, 2016, The three R's of Computer Vision: Recognition, Reconstruction and Reorganization, *Pattern Recognition Letters*
48. Katerina Fragkiadaki, Pulkit Agrawal, Sergey Levine and Jitendra Malik, 2016, Learning Visual Predictive Models of Physics for Playing Billiards, *International Conference to Learning Representations (ICLR)*
49. Joao Carreira, Pulkit Agrawal, **Katerina Fragkiadaki** and Jitendra Malik, 2016, Human Pose Estimation with Iterative Error Feedback, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
50. Panna Felsen, **Katerina Fragkiadaki**, Jitendra Malik and Alyosha Efros, 2015, Learning Feature Hierarchies from Long Term Trajectory Associations in Videos, *Transfer and Multitask Learning NeurIPS Workshop*
Katerina Fragkiadaki, Sergey Levine, Panna Felsen and Jitendra Malik, 2015, Recurrent Network Models for Human Dynamics, *IEEE International Conference on Computer Vision (ICCV)*
51. **Katerina Fragkiadaki**, Pablo Arbelaez, Panna Felsen and Jitendra Malik, 2015, Learning to Segment Moving Objects in Videos, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
52. **Katerina Fragkiadaki**, Marta Salas, Pablo Arbelaez and Jitendra Malik, 2014, Grouping-based Low-Rank Trajectory Completion and 3D Reconstruction, *Neural Information Processing Systems (NeurIPS)*

53. **Katerina Fragkiadaki**, Han Hu and Jianbo Shi, 2013, Pose from Flow and Flow from Pose, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
54. **Katerina Fragkiadaki**, Weiyu Zhang, Geng Zhang and Jianbo Shi, 2012, Two Granularity Tracking: Mediating Trajectory and Detection Graphs for Tracking under Occlusions *European Conference on Computer Vision (ECCV)*
55. **Katerina Fragkiadaki**, Geng Zhang and Jianbo Shi, 2012, Video Segmentation by Tracing Discontinuities in a Trajectory Embedding, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
56. **Katerina Fragkiadaki**, Weiyu Zhang, Jianbo Shi and Elena Bernardis, 2012, Structural-flow Trajectories for Unravelling 3D Tubular Bundles *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*
57. **Katerina Fragkiadaki** and Jianbo Shi, 2011, Detection-free Tracking: Exploiting Motion and Topology for Segmenting and Tracking under Entanglement, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
58. **Katerina Fragkiadaki** and Jianbo Shi, 2010, Figure-ground Image Segmentation Helps Weakly-supervised Learning of Objects *European Conference on Computer Vision(ECCV)*

Other Notable Published Products

1. Sudheendra Vijayanarasimhan, Susanna Ricco, Cordelia Schmid, Rahul Sukthankar and **Katerina Fragkiadaki**, 2017, SfM-Net: Learning of Structure and Motion from Video, *arXiv* (369 citations)
2. **Katerina Fragkiadaki**, Jonathan Huang, Alex Alemi, Sudheendra Vijayanarasimhan, Susanna Ricco and Rahul Sukthankar, 2017, Motion Prediction Under Multimodality with Conditional Stochastic Networks, *arXiv*

Awards and Recognition

- JPMorgan Chase Career Development Professorship 2023
- DARPA Young Investigator award 2021
- AFOSR Young Investigator award 2020
- Amazon faculty award 2020, 2023
- TRI faculty award 2020
- NSF CAREER award 2020
- UPMC faculty Award 2019
- Sony faculty award 2019
- Google Faculty Award 2018
- Best PhD Thesis, Computer and Information Science Department, University of Pennsylvania, 2013
- Program Chair ICLR 2024

Teaching

- Instructor, Fall 2017, 10-808, Language Grounding for Vision and Control. This is a seminar course that covers recent advances in language grounding, language and vision, instruction following, and learning from narrated demonstrations.

- Co-instructor, Spring 2017, 10-703, Deep Reinforcement Learning and Control. This is a graduate course that covers recent advances in deep reinforcement learning and deep robotic learning. It is focused on robotic learning methods from visual sensory input, self-supervision of visual representations for action, visually-grounded model-based reinforcement learning. The course is very successful, so since we introduced it in 2017 we have to teach it every semester. 10-703 is the graduate version with typically 150 students, and 10-403 is the undergraduate version with typically 35 students.
- Co-instructor, Fall 2018, 10-703, Deep Reinforcement Learning and Control.
- Instructor, Spring 2019, 10-403, Deep Reinforcement Learning and Control. This is the undergraduate version of 10-703.
- Instructor, Fall 2019, 10-703, Deep Reinforcement Learning and Control.
- Instructor, Spring 2020, 10-403, Deep Reinforcement Learning and Control.
- Instructor, Fall 2020, 10-703, Deep Reinforcement Learning and Control.
- Instructor, Spring 2021, 10-403, Deep Reinforcement Learning and Control.
- Co-instructor, Fall 2021, 10-703, Deep Reinforcement Learning and Control.
- Instructor, Spring 2022, 10-403, Deep Reinforcement Learning and Control.
- Instructor, Fall 2022, 10-703, Deep Reinforcement Learning and Control.
- Instructor, Spring 2023, 10-403, Deep Reinforcement Learning and Control.
- Instructor, Fall 2023, 10-703, Deep Reinforcement Learning and Control.
- Instructor, Spring 2024, 10-403, Deep Reinforcement Learning and Control.

Advising

- **Current PhD students:**

Mihir Prabhudesai, Robotics Institute (RI), (co-advised with Deepak Pathak)

Xian Zhou, RI

Gabe Sarch, Neuroscience Institute and Center for the Neural Basis of Cognition (CNBC), (co-advised with Mike Tarr)

Brian Yang, RI (co-advised with Jeff Schneider)

Nikos Gkanatsios, RI

Wen-hsuan Chu, RI

Ayush Jain, RI

- **Current M.S. students:**

Pushkal Katara, MSR, RI

Alex Swerdlow, MSR, RI

- **Graduated students:**

Yunchu Zhang, Masters in Robotics (MSR), RI, currently at Ph.D. in UW

Zhaoyuan Fang, MSR, RI, currently in Google

Ayush Jain, MSR, RI, currently at Ph.D. in CMU

Mayank Singh, MSR, RI, currently at Apple
Adam Harley, PhD RI May 2022, currently a PostDoc with Leo Guibas
Fish Tung, PhD MLD, August 2021, then PostDoc with Josh Tenenbaum, now at Tesla
Jingyun Yang, Masters in Machine Learning (MSML) MLD, Dec. 2021, currently a PhD student at Stanford CS
Jing Wen, Masters in Language Technology (MLT) Language Technologies Institute (LTI), Aug 2021, currently a PhD student at University of Illinois Urbana-Champaign (UIUC) CS
Fangyu Li, MSR RI, currently at NVIDIA
Hao Zhu, research associate, currently a PhD student at University of Washington (UW) CS
Niles Christensen, MSML, MLD, Sept 2019
Shamit Lal, Masters in Computer Vision (MSCV), RI, May 2020
Shrinidhi K. Lakshmikanth, research associate
Darshan Patil, CMU undergraduate, May 2020, currently at PhD, Mila
Gaurav Pathak, research associate, currently at MSR, RI, CMU
Yiming Zuo, MSR RI, June 2021, currently a PhD student at Princeton CS
Arpit Agarwal, MSR RI, currently a PhD student at RI, CMU
Max Sieb, MSR RI, currently at COVARIANT.AI
Rui Huang, Masters, ECE, currently at Apple
Ziyan Wang, MSR, RI, currently a PhD student at MLD, CMU
Henry Huang, MSML, MLD, currently at Bloomberg
Chris Ying, MSML, MLD, currently at Google Brain
Ricson Cheng, undergraduate, CSD, then research scientist at Uber
Yijie Wang, MSML, MLD, currently at TuSimple

- **Graduate Thesis Committee Service:**
Shikhar Bahl, PhD RI, CMU
Rohan Choudhury, PhD RI, CMU
Julen Urain, Ph.D., TU Darmstadt
Tom Monnier, Ph.D., ENPC
Dian Chen, PhD, CS, UT Austin
Chaoyang Wang, PhD, RI, CMU
Jayanth Koushik, PhD, CNBC, CMU
Shaojie Bai, PhD, MLD, CMU
Siddharth Ancha, PhD MLD, CMU
Emilio Parisotto, PhD MLD, CMU

Calvin Murdock, PhD MLD, CMU
Achal Dave, PhD, RI, CMU
Devin Schwab, PhD, MLD, CMU
Tarasha Khurana, MSR, RI, CMU
Andrew (Drew) Jaegle, PhD, CIS, University of Pennsylvania
Konstantinos Moustakas, CS, University of Patras
Avinash Siravuru, PhD Mechanical Engineering, CMU
Brendan Miller, MSR RI, CMU
Akash Sharma, MSR RI, CMU
Anjana Kakecochi Nellithimaru MSR, RI, CMU
Robert Li, MSR, RI, CMU
Tanya Marwah, MSR, RI, CMU

Grants and Gifts

- CAREER: Geometry, Physics and Semantics from Motion: Learning Expressive and Space-Aware Video Representations Source of Support: National Science Foundation Total Award Amount: \$546,457 Period of Performance: 1/1/2020 to 12/31/2024 Number of Person-Months: 1.0 SU This project proposes visual feature representations that decompose RGB and RGB-D streams into scene appearance and camera and object motion (PI).
- FY2021 DARPA Young Faculty Award (YFA): From 3D Perception to Spatial Intelligence: Self-supervising Hierarchical Neural 3D Perceptual Simulators for Transferring Learned Physics and Policies across Views, Objects, Scenes and Tasks Direct Sponsor: DEFENSE ADVANCED RESEARCH PROJECTS AGENCY (DARPA) Sponsor Award #: D21AP10115 MIPR HR0011153938 \$500K (PI).
- AFOSR Young Investigator Award (YIP): Air Force Office of Scientific Research Solicitation Number: AFOSR FOA-AFRL-AFOSR- 2020-0003 YIP Title: Common Sense for Superhuman Perception: From Vibration to Vision to Verbal Multiscale Multimodal Intelligence Total Award Amount: \$445,180 Period of Performance: 1/1/2021 to 12/31/2023 (PI).
- Amazon Faculty research award 2021. Total Award Amount: \$100K. Work on libraries of behaviour of generalizing manipulation across objects scenes, views and contexts (PI).
- Google faculty award 2018 on learning instructable agents (single year one student support). Total Award Amount: \$90K (PI).
- Sony faculty awards 2019. Geometry-aware Deep Visual Feature Learning for Semantic Scene Understanding. Source of Support: Sony Corporation. Total Award Amount: \$150,000. Period of Performance: 11/15/2019 to 11/14/2020. Number of Person-Months: 1.2 CY. This project aims to develop trainable neural network architectures that lift 2D RGB frames into 3D feature tensors and into geometrically consistent 3D deep feature maps of the scene. (PI)
- UPMC Faculty award 2019. Work for mining structured knowledge from medical notes Total Award Amount: \$150K (PI).

- Title: Embodied Symbiotic Assistants that See, Act, Infer and Chat. Source of Support: Amazon Alexa. Prize Period of Performance: 1/24/2022 to 4/28/2023. Number of Person-Months: 1.1 CY. Our approach focuses on building abstractions of objects and behaviours that can be re-used to assist the human user into expressing desires in a more succinct and natural way, understand novel objects and perform novel tasks (PI).
- Title: Sony Chef. Source of Support: Sony Corporation. Total Award Amount: \$ 3,800,000. Period of Performance: 4/2/2018 to 4/1/2022. Number of Person-Months: 2.0 CY. This project uses robotics research to focus on optimizing food preparation, cooking and delivery (co-PI).
- Leveraging Advanced Algorithms, Autonomy, and Artificial Intelligence (A4I) to Enhance National Security and Defense: Mobile Multiview Video Understanding with 3D- Bottlenecked Neural Networks. Source of Support: Army Research Laboratory. Total Award Amount: \$1,282,655. Period of Performance: 4/1/2019 to 3/31/2022. Award Number: W911NF1820218. Number of Person-Months: 2.0 SU. This project will accelerate research and development to enable the autonomous collection, processing, and exploitation of operational data into real-time, actionable intelligence (co-PI).
- SAS: INT: Smart And Autonomous Systems For Repair And Improvisation. Source of Support: National Science Foundation. Total Award Amount: \$670,000. Period of Performance: 6/1/2019 to 5/31/2023. Award Number: IIS1849287. Number of Person-Months: 0.5 SU. This project focuses on enabling robots to understand instructions from users, to fix errors and repair broken processes, objects, and tools, and to improvise and find new and better ways to do tasks (co-PI).
- MESS: Model Building, Exploration, Social Systems. Source of Support: University of California, Berkeley / DARPA. Total Award Amount: \$1,320,000. Period of Performance: 9/19/2019 to 9/18/2023. Location of Project: Carnegie Mellon University. Number of Person-Months: 1.2 CY. This project aims to build models that specifically focus on the emergence of object properties from passively and actively collected video sequences (co-PI).
- Exploiting Physical Interactions to Increase the Understanding of Dynamics and Environment Properties for Enhanced Autonomy Source of Support: Army Research Laboratory. Total Award Amount: \$2,723,035. Period of Performance: 10/1/2019 to 9/30/2024. Number of Person-Months: 1.0 CY. This project aims to develop a general framework to enhance autonomy of intelligent agents by actively learning the physical commonsense and predicting the movement and behaviors of complex objects (co-PI).
- Adobe gifts for self-supervised learning of motion capture Total Award Amount: \$30K.
- Collaboration with Software Engineering Institute on developing domain adaptation methods and tracking models for the MAVEN project (two student support) (\$200K per year). This support is no longer active.
- Air Force Office of Scientific Research. Award Number: FA9550-20-1-042. Title: Open Eyes Common Sense: Machine Reasoning using Visual Simulation. Total Award Amount: \$434,330. Period of Performance: 9/28/2020 to 9/27/2023 (PI).
- Facebook, Inc. Title: Collecting Robot-Object Interaction Data at Scale. Total Award Amount: \$104,000. Period of Performance: 11/19/2020 to 12/31/2021 (PI).
- Stanford University / Toyota Research Institute. Title: Semi-supervised Learning for Robust, Multi-object 3D Detection and Tracking. Total Award Amount: \$254,204. Period of Performance: 4/1/2021 to 3/31/2024 (co-PI).

- Source of Support: Toyota Research Institute Title: Learning Traffic Screenplays through 3D Object Detection and Tracking Award No.: Technical Point of Contact: Eric Krotkov, 4440 El Camino Real, Los Altos, CA, 94022, eric.krotkov@tri.globa Total Award Amount: \$199,057 Period of Performance: 4/1/2021 to 3/31/2024 Location of Project: Carnegie Mellon University Number of Person-Months: 0.45 SU Objectives: The overall goal of this project is to develop novel, robust, multi-object detection and tracking algorithms as an integrated module.
- Title: End-to-end Memory-modulated Visuomotor Learning Award No.: 903783 Source of Support: Google Technical Point of Contact: William Cohen, 412-901-4879, wcohen@google.com Total Award Amount: \$80,000 Period of Performance: 11/23/2022 to 11/22/2024 Location of Project: Carnegie Mellon University Number of Person-Months: 0.01 CY Objectives: This project aims to develop end-to-end models to retrieve and adapt memories for perception and action-related tasks.
- Title: An Interactive Continual Multi-modal Learner based on Memory-Modulated Neural Analogical Reasoning Award No.: N000142312415 Source of Support: Office of Naval Research Technical Point of Contact: BEHZAD KAMGARPARSI, MATH COMPUTERS AND INFO SCI (MCIS) DIV, 875 N. Randolph Street, Arlington VA 22203-1995, BEHZAD.KAMGARPARSI.CIV@US.NAVY.MIL Total Award Amount: \$509,722 Period of Performance: 5/1/2023 to 4/30/2026 Location of Project: Carnegie Mellon University Number of Person-Months: 1.2 CY Objectives: This project proposes an analogical framework for knowledge representation, perception and language grounding that encodes domain knowledge explicitly, in a collection of structured sensory experiences at different levels of spatial and temporal abstraction, in addition to implicitly, as network parameters.
- Title: Analogical World Models: Perception, Action and Language Grounding through Analogical Prediction Award No.: FA9550-23-1-0257 Source of Support: Air Force Office of Scientific Research Technical Point of Contact: Dr. R. Doug Riecken, AFOSR/RTA, (703) 696-9736, richard.riecken@us.af.mil Total Award Amount: \$531,798 Period of Performance: 6/1/2023 to 5/31/2026 Location of Project: Carnegie Mellon University Number of Person-Months: 1.2 CY Objectives: This project proposes embodied perception systems for surveillance and human-robot collaboration that learn world models of space, physics and geometry to help their sensory interpretations generalize across domains and improve during deployment.

(For more details on funding support please contact Nancy McCarthy: namccart@andrew.cmu.edu)

Service

CMU service:

- MSML admission committee 2019.
- Chair of MSML admissions committee 2020, 2021.
- Director of Masters in Machine Learning (MSML) 2020-2022.
- Chair of the re-appointment committee for an Assistant Professor in MLD, CMU. 2022
- MLD hiring committee 2023.
- Chair of PhD. admissions committee 2024.

External Professional Service:

- Program Chair for ICLR 2024

- Senior Area Chair for ICLR 2025, CVPR 2025
- Best paper award committee 3DV 2022, 3DV 2024
- Area Chair for: CVPR 2018, ICLR 2019, ICML 2019, CVPR 2020, ICML 2020, NeurIPS 2020, NeurIPS 2021, ICML 2021, ICLR 2021, CVPR 2021, ICCV 2021, 3DV 2021, ICML 2022, ECCV 2022, NeurIPS 2022, CVPR 2023, ICML 2023, NeurIPS 2023, CVPR 2024, Associate Editor for TPAMI
- Journal referee for: International Journal of Computer Vision 2011, Image and Vision Computing 2011, Digital Signal processing 2011, Machine Vision and Applications 2012
- NSF Panelist for Robust Intelligence: Machine Learning Core Program Panel, March 2020
- External reviewer for: IEEE Conference on Computer Vision and Pattern Recognition 2010, 2011, IEEE European Conference on Computer Vision 2010, 2018,2016, NeurIPS 2018, 2017, 2016, IEEE International Conference on Computer Vision 2011, 2013, 2015, 2017, 2019, IEEE Asian Conference on Computer Vision 2012, CVPR 2022, ICLR 2022, CoRL 2022
- (Co-)Organizer of the 11th Perceptual Organization for Computer Vision Workshop, CVPR 2016: "The role of feedback in Recognition and Segmentation". (Co-)organizer of the 2nd International Workshop on Video Segmentation, ECCV 2016: "How can segmentation drive learning?" Co-organizer of the 1st and 2nd "Learning from unlabelled video" workshop in CVPR 2019, CVPR 2020. (Co-)organizer of "Minds versus machines: how far are we from the common sense of a toddler?" CVPR 2020. (Co-)Organizer of [ECCV 2020 workshop on Perception Through Structured Generative Models](#) (Co-)Organizer of [CVPR 2021 Workshop on 3D Vision and Robotics](#), multiple other workshops in CVPR, ECCV, ICCV, CoRL

Community service, Outreach:

- **Diversity in STEM** I co-organized alongside [Jonathan Reynolds](#) from the CMU outreach office the first [AI4LL](#) summer school in CMU in July 2018. AI4ALL is a summer residency program which aims at increasing diversity in Artificial Intelligence by familiarizing students from under-represented backgrounds (ethnic minorities, rural areas, low-SES, females, non-binary) with the field of Artificial Intelligence, its goals, its practices, and its successes so far. Last year was a success: all participants indicated at the end of the program that they were planning to pursue AI as their college major, which was not the case in the beginning of the program. CMU AI4ALL is organized every summer since then.

I was a mentor in [WiML](#) (Women in Machine Learning) in NeurIPS 2018 and in [WiCV](#) (Women in Computer Vision) in CVPR 2019, 2020, 2022.

My first PhD student, Dr. Fish Tung, is a female. She received numerous prestigious fellowships during her PhD (Siemens FutureMakers, Seibel, Yahoo InMind) and is currently a PostDoctoral research in the labs of Prof. John Tanenbaum and Dan Yamin. I have advised a female Masters student, Audrey Huang, and Ishita Mediratta who is currently at Facebook research.

- **Undergraduate research** My lab has hosted undergraduate students on [SN Bose scholarship](#)(Dakshit Agrawal, Ishu Garg) as well as on PKU-Tsinghua CMU summer internship program (Mengze Zhu, Qiyuan Dong, Chufan Wu, Jiayuan Gu), many of which entered then Masters and PhD programs, e.g., Jiayuan Gu is now a PhD candidate in UCSD.

I have worked closely with a CMU CSD undergraduate student, Ricson Cheng. Ricson was awarded the [CRA outstanding undergraduate research award](#) runner-up, being within [top 8](#) in North America.

This award program recognizes undergraduate students in North American colleges and universities who show outstanding potential for research in an area of computing research.

I hosted Mihir Prabhudesai for 1.5 years as a research assistant in my lab right after completing his undergraduate degree in India, and he is now a PhD student in RI, CMU.

- I gave a talk at [Achievement Rewards for College Scientists \(ARCS\) Foundation](#) that provides fellowships to support graduate students in CMU.
- I participated in a panel in Shadyside high school to discuss the future of AI and how to prepare middle and high schoolers for it.

Invited and keynote talks

- [British Machine Vision Conference \(BMVC\) 2021](#), keynote speaker
- [International Conference on 3D Vision \(3DV\) 2021](#), keynote speaker
- ONR MURI workshop on self-supervising perception, 2023
- [Future of Machine Learning Symposium](#), ISTA, 2024, keynote speaker
- International Conference on Machine Vision Applications (MVA) 2025, keynote speaker
- Invited speaker in numerous CVPR, ECCV, ICCV, RSS, CoRL, ICRA workshops (please, see my webpage for details)
- Invited speaker in numerous AI seminars, ML reading groups and summer schools in different Universities and industries, including MIT robotics seminar, UW, UoT, UTAustin, UIUC, Google Computational Photography, Uber, Waymo, Stanford Vision Group, etc.