HSIAO-YU FISH TUNG

sfish0101@gmail.com ⋄ Homepage: https://sfish0101.bitbucket.io/

RESEARCH INTERESTS

Multimodal LLM, LLM-powered Agents, 3D Computer Vision, Robotics, Generative Models, Self-supervised Learning

EMPLOYMENT

Google DeepMind, Mountain View, USA *July 2023-present*
Research Engineer
- Multimodal-LLM for AI agents (Project Astra) and AI for education.

Tesla Inc., Palo Alto, USA *Feb 2023 - May 2023*
Senior Autopilot Machine Learning Research Scientist

MIT, Cambridge, USA *May 2021 - Feb 2023*
Postdoctoral associate in Brain and Cognitive Science,
Advisor: Prof.Josh Tenenbaum (MIT), Dan Yamins (Stanford)

EDUCATION

Carnegie Mellon University, Pittsburgh, USA *September 2015-2021*
PhD in Machine Learning, Advisor: Prof. Katerina Fragkiadaki
Overall GPA: 4.03
• Thesis Title: Learning to See by Moving: Self-supervising 3D Scene Representations for Perception, Control, and Visual Reasoning
Committee Members: Katarina Fragkiadaki, Tom Mitchell, Chris Atkeson, Jitendra Malik

Carnegie Mellon University, Pittsburgh, USA *August 2013-May 2015*
M.Sc. in Machine Learning, *Advisor: Prof. Alexander J. Smola*
Overall GPA: 4.0
• Master thesis: Spectral methods for nonparametric models

National Taiwan University, Taipei, Taiwan *September 2009 -June 2013*
B.S. in Electrical Engineering
Overall GPA: 3.95/4.0, Rank Top 9%

INTERNSHIP EXPERIENCE

2018 Summer, OpenAI Inc.. Mentor: Wojciech Zaremba, Peter Welinder
- Robust state estimation for human-like robot hand.

2017 Summer, Adobe Research.. Mentor: Ersin Yumer
- Self-supervised Learning of Motion Capture [15]. Policy Leaning in physics simulator.

2016 Summer, Google Brain Team, Google Inc.. Mentor: Andrew Dai
- Generative Adversarial Nets (GANs) for Text.

2015 Summer, Parallel Computing Lab, Intel Labs.. Mentor: Shang Li
- Accelerating Long Short Term Memory Network with full-stack optimization.

2014 Summer, Machine Learning Department, CMU. Advisor: Aarti Singh

2012 Summer, Home Entertainment group, MediaTek.

PUBLICATION

Arxiv Preprint

[1] Mihir Prabhudesai, Shamit Lal, **Hsiao-Yu Fish Tung**, Adam W Harley, Shubhankar Potdar and Katerina Fragkiadaki, "3D Object Recognition By Corresponding and Quantizing Neural 3D Scene Representations," CVPR 2020 Machine Common Sense Workshop, https://arxiv.org/pdf/2010.16279.pdf, 2020.

Peer-Reviewed Conference Papers

[1] Haotian Xue, Antonio Torralba, Joshua Tenenbaum, Daniel Yamins, Yunzhu Li, **Hsiao-Yu Fish Tung**, "3D-IntPhys: Towards More Generalized 3D-grounded Visual Intuitive Physics under Challenging Scenes," NeurIPS 2023.

[2] **Hsiao-Yu Fish Tung***, Mingyu Ding*, Zhenfang Chen*, Sirui Tao, Vedang Lal, Daniel Bear, Gan Chuang, Joshua B Tenenbaum, Daniel L. K. Yamins, Judith E. Fan, Kevin Smith, "Physion++: Evaluating Physical Scene Understanding with Objects Consisting of Different Physical Attributes in Humans and Machines," *Cognitive Science Society (CogSci), 2023. NeurIPS 2023.*

[3] Gabriel Sarch, **Hsiao-Yu Fish Tung**, Aria Wang, Jacob Prince, Michael Tarr, "3D View Prediction Models of the Dorsal Visual Stream," *Conference on Cognitive Computational Neuroscience (CCN), 2023.*

[4] Kei Ota, Devesh K Jha, **Hsiao-Yu Fish Tung**, Joshua B Tenenbaum, "Tactile-Filter: Interactive Tactile Perception for Part Mating," *Robotics: Science and Systems (RSS), 2023.*

[5] Zhou Xian, Bo Zhu, Zhenjia Xu, **Hsiao-Yu Fish Tung**, Antonio Torralba, Katerina Fragkiadaki, Chuang Gan, "Fluidlab: A differentiable environment for benchmarking complex fluid manipulation," *International Conference on Learning Representations (ICLR), 2023.*

[6] Kei Ota, **Hsiao-Yu Fish Tung**, Kevin A. Smith, Anoop Cherian, Tim K. Marks, Alan Sullivan, Asako Kanezaki, Joshua B. Tenenbaum, "H-SAUR: Hypothesize, Simulate, Act, Update, and Repeat for Understanding Object Articulations from Interactions," *International Conference on Robotics and Automation (ICRA), 2023.*

[7] Jingyun Yang*, **Hsiao-Yu Fish Tung***, Yunchu Zhang*, Gaurav Pathak, Ashwini Pokle, Christopher G Atkeson, Katerina Fragkiadaki, "Visually-Grounded Library of Behaviors for Manipulating Diverse Objects across Diverse Configurations and Views," *Conference on Robot Learning (CoRL), 2021.*

[8] Daniel M Bear, Elias Wang, Damian Mrowca, Felix J Binder, **Hsiao-Yu Fish Tung**, RT Pramod, Cameron Holdaway, Sirui Tao, Kevin Smith, Li Fei-Fei, Nancy Kanwisher, Joshua B Tenenbaum, Daniel LK Yamins, Judith E Fan, "Physion: Evaluating Physical Prediction from Vision in Humans and Machines," *Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks Track, 2021.*

[9] Mihir Prabhudesai*, Shamit Lal*, Darshan Patil*, **Hsiao-Yu Fish Tung**, Adam W Harley, and Katerina Fragkiadaki, "Disentangling 3D Prototypical Networks for Few-Shot Concept Learning," *International Conference on Learning Representations (ICLR), 2021.*

[10] Zhou Xian, Shamit Lal, **Hsiao-Yu Fish Tung**,Emmanouil Antonios Platanios, and Katerina Fragkiadaki, "HyperDynamics: Generating Expert Dynamics Models by Observation," *International Conference on Learning Representations (ICLR), 2021.*

[11] **Hsiao-Yu Fish Tung***, Xian Zhou*, Mihir Prabhudesai, Shamit Lal and Katerina Fragkiadaki, "3D-OES: View-Invariant Object-factorized Environment Simulators," *Conference on Robot Learning (CoRL), 2020.*

[12] Mihir Prabhudesai*, **Hsiao-Yu Fish Tung***, Syed Ashar Javed*, Maximilian Sieb, Adam W. Harley and Katerina Fragkiadaki, "Embodied Language Grounding with 3D Visual Feature Repre-

sentations," *Conference on Computer Vision and Pattern Recognition (CVPR), 2020.*

[13] Adam W. Harley, Fangyu Li, Shrinidhi K. Lakshmikanth, Zian Zhou, **Hsiao-Yu Fish Tung** and Katerina Fragkiadaki, "Visual Representation Learning with 3D View-Contrastive Inverse Graphics Networks," *International Conference on Learning Representations (ICLR), 2020.*

[14] **Hsiao-Yu Fish Tung***, Ricson Cheng* and Katerina Fragkiadaki, "Learning Spatial Common Sense with Geometry-Aware Recurrent Networks," *Conference on Computer Vision and Pattern Recognition (CVPR), 2019.* (**Oral**)

[15] **Hsiao-Yu Fish Tung**, Adam W. Harley, Liang-Kang Huang and Katerina Fragkiadaki, "Reward Learning from Narrated Demonstrations," *Conference on Computer Vision and Pattern Recognition (CVPR), 2018.*

[16] **Hsiao-Yu Fish Tung**, Hsiao-Wei Tung, Ersin Yumer, and Katerina Fragkiadaki, "Self-supervised learning of motion capture," *Neural Information Processing Systems (NIPS), 2017.* (**Spotlight**)

[17] **Hsiao-Yu Fish Tung**, Adam Harley, William Seto, and Katerina Fragkiadaki, "Adversarial inverse graphics networks: Learning 2d-to-3d lifting and image-to-image translation from unpaired supervision," *International Conference on Computer Vision (ICCV), 2017.*

[18] Dougal J. Sutherland, **Hsiao-Yu Tung**, Heiko Strathmann, Soumyajit De, Aaditya Ramdas, Alex Smola, and Arthur Gretton, "Generative models and model criticism via optimized maximum mean discrepancy," *International Conference on Learning Representations (ICLR), 2017.*

[19] Yining Wang, **Hsiao-Yu Fish Tung**, Alexander J. Smola and Animashree Anandkumar, "Fast and guaranteed tensor decomposition via sketching," *Neural Information Processing Systems (NIPS), 2015.* (**Spotlight**)

[20] **Hsiao-Yu Fish Tung** and Alexander J. Smola, "Spectral Methods for Indian Buffet Process Inference," *Neural Information Processing Systems (NIPS), 2014.*

[21] **Hsiao-Yu Tung**, Wei-Chiu Ma, and Tian-Li Yu, "Novel Traffic Light Timing Adjustment Strategy Based On Genetic Algorithm," *IEEE Congress on Evolutionary Computation (IEEE CEC), 2014.* (**Oral**)

Journal Paper

[J1] **Hsiao-Yu Fish Tung**, Chao-Yuan Wu, Manzil Zaheer, and Alexander J. Smola, "Spectral methods for nonparametric models," CMU MLD Master Thesis. http://arxiv.org/abs/1704.00003

[J2] C.-L. Li, Y.-C. Su, T.-W. Lin, C.-H. Tsai, W.-C. Chang, K.-H. Huang, T.-M. Kuo, S.-W. Lin, Y.-S. Lin, Y.-C. Lu, C.-P. Yang, C.-X. Chang, W.-S. Chin, Y.-C. Juan, **H.-Y. Tung**, J.-P. Wang, C.-K. Wei, Felix Wu, T.-C. Yin, T. Yu, Y. Zhuang, S.-d. Lin, H.-T. Lin, and C.-J. Lin. "Combination of Feature Engineering and Ranking Models for Paper-Author Identification in KDD Cup 2013," *Journal of Machine Learning Research, 2015.*

[J3] W.-S. Chin, Y.-C. Juan, Y.-Zhuang, Felix Wu, **H.-Y. Tung**, T. Yu, J.-P. Wang, C.-X. Chang, C.-P. Yang, W.-C. Chang, K.-H. Huang, T.-M. Kuo, S.-W. Lin, Y.-S. Lin, Y.-C. Lu, Y.-C. Su, C.-K. Wei, T.-C. Yin, C.-L. Li, T.-W. Lin, C.-H. Tsai, S.-d. Lin, H.-T. Lin, and C.-J. Lin. "Effective String Processing and Matching for Author Disambiguation," *Journal of Machine Learning Research, 2014.*

AWARDS AND HONOR

2020 Siebel Scholar

2019 Rising Star in EECS

2019 Siemens FutureMaker Fellowship (Awarded)

2019 Yahoo InMind Fellowship (Awarded)

2018 Qualcomm QinF Finalists

2018 Open Philanthropy Project AI Fellows Finalists

2018 Facebook Fellowship Finallists

2014 2017 NIPS Travel Award Winner

2013 KDD Cup Award, Track1&2 Champion [J2][J3]

2013 NTU Student Outstanding Performance Scholarship

2011 2012 NTU President Award
-Awarded to top 5% of students in each department of National Taiwan University.

2012 Altera Innovate Asia FPGA Design Competition,Outstanding Achievement

TEACHING AND MENTORSHIP EXPERIENCE

Instructor

2019 Summer, CMU AI4ALL. Lecture: Computer Vision

Teaching Assistant

2016 Fall, CMU. 10-601 Introduction to Machine Learning (ML for master students)

2015 Fall, CMU. 10-715 Advanced Introduction to Machine Learning (ML for PhD students)

Lab Mentees
Mihir Prabhudesai (now PhD at CMU), Jingyun Yang (now PhD at Stanford), Haotian Xue (now PhD at Georgia Tech), Kei Ota (PhD at TokyoTech, MERL), Shamit Lal (AWS), Syed Ashar Javed (PathAI), William Seto (NASA), Sirui Tao, Vedang Lal

SELECTED TALKS

2022 Mitsubishi Research Institute (MERL). Learning to See by Moving

2021 Summer, Women in Data Science, Bangalore @ MathWorks. Penalist on the topic – "Should artificial intelligence copy the brain?"

2020 Fall, CMU brAIn seminar, MIT, Stanford NeuroAILab. Learning to See by Moving

2020 Summer, Columbia University and UC Berkeley. Learning Geometry-Aware Visual Representations for Embodied Agents and through Embodied Agents

2020 Spring, NYU Human and Machine Learning Lab. Learning Geometry-Aware Visual Representations for Embodied Agents and through Embodied Agents

2019 Fall, MIT Computer Vision Group. Embodied Visual Recognition

2019 Summer, Oculus Reality Lab. Embodied Visual Recognition

2019 Spring, CMU AI Seminar. Geometry-Aware Recurrent Networks: A visual system for embodied agents.

2018 Summer, OpenAI Summer Open House. Robust Vision-Based State Estimation

2017 Spring, CMU AI Seminar. Adversarial Inversion: Self-supervision with Adversarial Priors

2017 Spring, OpenAI Inc.. Generative models with optimized maximum mean discrepancy and adversarial imagination priors

PROFESSORIONAL SERVICE

Program Committees
Reviewer of ICCV, ECCV, CVPR, NeuIPS, ICML, ICRA, ICLR, IEEE Image Processing

Conference Workshop Organizer:
2022 ECCV: 1st Challenge on Machine Visual Common Sense: Perception, Prediction, Planning
2021 NeurIPS: Physical Reasoning and Inductive Biases for the Real World [Website Link]
2020 CVPR: Minds vs. Machines: How far are we from the common sense of a toddler? [Recorded Talks]

University Committees:
2019 Speaking Skills Committee for the CMU Machine Learning Department
2016-2018 PhD admission committee for the CMU Machine Learning Department
2015-2016 Master admission committee for the CMU Machine Learning Department

Research Community Service:
2017 Mentor in Adobe x Girls Who Code
2014 Staff in Machine Learning Summer School