Cheng Perng Phoo

Ph.D. Candidate, Cornell University

107 Hoy Road, 325, Bill and Melinda Gates Hall, Ithaca, NY, USA 14850

cpphoo@cs.cornell.edu | www.cs.cornell.edu/∼cpphoo

Education

Cornell University, USA *Aug 2017 - May 2024 (expected)*

Ph.D. Candidate in Computer Science, GPA 3.90 / 4.00

Advisor: Bharath Hariharan

University of Michigan, Ann Arbor, USA *Sep 2014 - May 2017*

B.S. in Computer Science and Pure Mathematics, GPA 3.78/4.00

Research Interest

My research lies in the intersection of computer vision and machine learning. Specifically, I work on building perception systems that are broadly useful for all problem domains (e.g., remote sensing, medical imagery, self-driving vehicles). Toward this goal, I have identified three major problems: label efficiency, deployment to novel domains, and trustworthiness. Most of my past works have focused on the first two problems. I am currently investigating how to use multimodal sensory input to improve the label efficiency of perception models and how we could build trustworthy specialist perception models from large-scale frontier models.

Publications

(* Equal Contributions)

Remote Sensing Vision-Language Foundation Models without Annotations via Ground Remote Alignment

- Utkarsh Mall*, **Cheng Perng Phoo***, Meilin Kelsey Liu, Carl Vondrick, Bharath Hariharan, Kavita Bala
- International Conference on Learning Representations (ICLR), 2024.
- Summary: We use ground images as intermediary to connect satellite imagery to natural language (encoded using CLIP), yielding VLMs without textual annoations.

Pre-training LiDAR-based 3D Object Detectors through Colorization

- Tai-Yu Pan, Chenyang Ma, Tianle Chen, **Cheng Perng Phoo**, Katie Z Luo, Yurong You, Mark Campbell, Kilian Q Weinberger, Bharath Hariharan, Wei-Lun Chao
- International Conference on Learning Representations (ICLR), 2024.
- Summary: We pre-train a point cloud detector by tasking it to fill in the missing colors within the point cloud.

Reward Finetuning for Faster and More Accurate Unsupervised Object Discovery

- Katie Z Luo*, Zhenzhen Liu*, Xiangyu Chen*, Yurong You, Sagie Benaim, **Cheng Perng Phoo**, Mark Campbell, Wen Sun, Bharath Hariharan, Kilian Q. Weinberger
- Conference on Neural Information Processing Systems (NEURIPS), 2023.
- Summary: We reframe object discovery as an RL problem and design a reward function to enable faster and more accurate discovery of objects in driving scenes without human supervision.

Emergent Correspondence from Image Diffusion

- Luming Tang*, Menglin Jia*, Qianqian Wang*, **Cheng Perng Phoo**, Bharath Hariharan
- Conference on Neural Information Processing Systems (NEURIPS), 2023.
- Summary: Features from off-the-shelf image diffusion models could be used to identify semantic and geometric correspondence without further training.

Distilling from Similar Tasks for Transfer Learning on a Budget

- Kenneth Borup, **Cheng Perng Phoo**, Bharath Hariharan.
- IEEE/CVF International Conference on Computer Vision (ICCV), 2023.
- Summary: We construct label- and compute-efficient models by identifying and distilling from suitable pre-trained models.

Unsupervised Adaptation from Repeated Traversals for Autonomous Driving

- Yurong You*, **Cheng Perng Phoo***, Katie Luo*, Travis Zhang, Wei-Lun Chao, Bharath Hariharan, Mark Campbell, Kilian Q. Weinberger.
- Conference on Neural Information Processing Systems (NEURIPS), 2022.
- Summary: Unlabeled LiDAR scans from repeated traversals could be used to disambiguate foreground and background objects, yielding cleaner signals for self-training adaptation.

Learning to Detect Mobile Objects from LiDAR Scans Without Labels

- Yurong You*, Katie Luo*, **Cheng Perng Phoo**, Wei-Lun Chao, Wen Sun, Bharath Hariharan, Mark Campbell, Kilian Q. Weinberger.
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022.
- Summary: Comparing unlabeled LiDAR scans from multiple traversals on the same location could uncover dynamic LiDAR points that could be used to train a mobile object detector in an unsupervised/self-supervised manner.

Task2Sim: Towards Effective Pre-training and Transfer from Synthetic Data

- Samarth Mishra, Rameswar Panda, **Cheng Perng Phoo**, Chun-Fu Richard Chen, Leonid Karlinsky, Kate Saenko, Venkatesh Saligrama, Rogerio S. Feris.
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022.
- Summary: Different downstream tasks require different representations pre-trained on synthetic data generated using different configurations (lightings, object poses, etc). We use reinforcement learning to learn a policy that maps a compact task representation to the appropriate synthetic data configuration.

Coarsely-labeled Data for Better Few-shot Transfer

- **Cheng Perng Phoo**, Bharath Hariharan.
- IEEE/CVF International Conference on Computer Vision (ICCV), 2021.
- Summary: Coarsely-labeled data can be cheap to acquire and can be used to learn a better representation for few-shot learning.

Self-training for Few-shot Transfer across Extreme Task Differences

- **Cheng Perng Phoo**, Bharath Hariharan.
- International Conference on Learning Representations (ICLR), 2021. Oral Presentation. [53 / 2997 submissions]
- Summary: We can build strong neural representation for novel domains by (self-)training students to replicate pseudo-labels produced by a teacher from another, unrelated problem domain.

Predicting risk of sport-related concussion in collegiate athletes and military cadets: a machine learning approach using baseline data from the CARE Consortium Study

- Joel Castellanos, **Cheng Perng Phoo**, James T. Eckner, Lea Franco, Steven P. Broglio, Mike McCrea, Thomas McAllister, and Jenna Wiens.
- Sports medicine (2020): 1-13.
- Summary: Baseline tests conducted on college athletes and military cadets before each semester could contain information for identifying athletes/military cadets who are at a higher risk of experiencing a concussion.

Perceiving Signs for Navigation Guidance in Spaces Designed for Humans

- Claire Liang*, **Cheng Perng Phoo***, Laasya Renganathan, Yingying Yu, Bharath Hariharan, Hadas Kress-Gazit.
- Closing the Academia to Real-World Gap in Service Robotics Workshop. Robotics Science and Systems Conference (RSS), 2020.
- Summary: We use a combination of computer vision and natural language processing techniques to extract information from signs in human-designed spaces to guide robots in navigation tasks.

Heart Sound Classification based on Temporal Alignment Techniques

- José Javier González Ortiz, **Cheng Perng Phoo**, Jenna Wiens.
- Computing in Cardiology Conference (CinC), 2016.
- Summary: We use temporal alignment techniques such as dynamic time warping to extract features from heart sound recordings for identifying patients at risk of adverse cardiovascular outcomes.

Research Experiences

Graduate Research Assistant
Cornell University

Cornell Graphics and Vision Group
Aug 2018 - Present

- Advisor: Professor Bharath Hariharan
- Researched how to build perception models for various problem domains.
- Investigated how to adapt visual-language models trained on internet imagery (e.g. CLIP and ViperGPT) to the remote sensing domain where annotations are sparse (ICLR 2024).
- Investigated how LiDAR scans from different past traversals could be used to improve camera-based 3D object detectors (in submission).
- Investigated how to use multiple large-scale pre-trained foundation models to create efficient models that could be trained with less annotated data and resources (ICCV 2023).
- Researched how to use unlabeled LiDAR scans from repeated traversals to adapt 3D object detectors to a novel domain (NEURIPS 2022).
- Investigated how to use unlabeled LiDAR scans to discover mobile objects in an unsupervised manner (CVPR 2022).
- Researched how class ontologies and coarsely labeled data could be used to improve few-shot classification (ICCV 2021).
- Researched how to adapt a pre-trained representation for few-shot classification in a novel domain distinct from the pre-training domain (ICLR 2021. Awarded Oral Presentation).

Research Intern @ FAIR Accel
Menlo Park, CA

Meta Fundamental AI Research (FAIR)
Jun 2022 - Aug 2022

- Collaborated with Dr Rama Kovvuri, Dr Effrosyni Mavroudi, Dr Kevin Liang, Dr Huiyu Wang, Dr Jing Wang on modeling object state changes in egocentric videos
- Reformulated the object state change classification problem in Ego4D challenge from a clip-level problem (i.e. is there an object state change in a clip?) to an object-level problem (i.e. does object x goes through a state change in the clip?)
- Proposed a novel state change embedding that could capture different degrees of state changes for an object (e.g how much has a door changed from close to open?)

Research Intern
Remote

MIT-IBM Watson AI Lab
Jun 2021 - Dec 2021

- Collaborated with Dr Rogerio Feris, Professor Kate Saenko, Dr. Chun-Fu (Richard) Chen, Dr.Rameswar Panda on open-set semi-supervised learning (i.e. how to learn a classifier with unlabeled data that are out of distribution).
- Researched the use of low-level features from convolutional neural networks for open-set semi-supervised classification.
- Investigated the use of dynamic neural networks for open-set semi-supervised classification.
- Investigated the interplay between transfer learning and open-set semi-supervised learning (i.e. how a pre-trained model can be used for better open-set semi-supervised learning).

Academic Services

Conference Reviewer

- Computer Vision and Pattern Recognition (CVPR) 2022, 2023
- European Conference on Computer Vision (ECCV) 2022
- International Conference on Computer Vision (ICCV) 2023
- Conference on Neural Information Processing Systems (NEURIPS) 2023
- International Conference on Learning Represenations (ICLR) 2024

Ph.D. Application Reviewer

- Computer Science, Cornell Univeristy 2023

Teaching Experiences

CS4780/5780: Machine Learning for Intelligent Systems *Cornell Univerisity*
Teaching Assistant for Kilian Weinberger, Chris De Sa *Spring 2018*
Awarded **Outstanding TA.**

CS4786/5786: Machine Learning for Data Science *Cornell Univerisity*
Teaching Assistant for Karthik Sridharan *Fall 2017*

EECS445: Introduction to Machine Learning *University of Michigan, Ann Arbor*
Instructional Aide for Jenna Wiens *Winter 2017*

EECS203: Discrete Mathematics *University of Michigan, Ann Arbor*
Instructional Aide *Fall 2015, Fall 2016, Winter 2016*

Skills

Programming Languages: Python, MATLAB, C/C++
Machine Learning: PyTorch, Tensorflow, scikit-learn, NumPy, SciPy, Pandas
Others: Bash, LaTeX, Linux

Languages

Mandarin: Native Language
English, Malay: Fluent (speaking, reading, writing)