

Soravit Beer Changpinyo

Education

University of Southern California *Los Angeles, CA* 2012-2018
M.S. and Ph.D. in Computer Science
Advisor: Prof. Fei Sha
GPA: 3.97/4.00

Brown University *Providence, RI* 2008-2012
B.S. in Computer Science
magna cum laude with Honors
Advisor: Prof. Erik B. Sudderth
GPA: 4.00/4.00

Professional Experience

University of Southern California, Los Angeles, CA *Research Assistant* 08/2012-09/2018
Zero-shot, transfer, and multi-task learning with applications to computer vision and natural language processing. Similarity and Metric Learning.

Google Research, New York City, NY *Software Engineering Intern* Summer 2017
Large-scale temporal and higher-order analysis of the discourse of topics on the Web

Google Research, Mountain View, CA *Software Engineering Intern* Summer 2016
Convolutional neural network compression.

Brown University, Providence, RI *Research Assistant* Fall 2011 & Spring 2012
Bayesian nonparametric methods for image attribute and topic modeling.

Information Trust Institute, Urbana, IL *Research Intern* Summer 2011
Keyword search over semi-structured information sources.

Brown University, Providence, RI *Research Assistant* Summer 2010
Transactional memory on embedded devices. Karen T. Romer Undergraduate Research Award.

Awards

Outstanding Reviewer Award, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2020. Top 3.8%.

Best Reviewer Award, Neural Information Processing Systems (NeurIPS) 2019. Top 8.5%.

Outstanding Reviewer Award, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2019. Top 8.5%.

Annenberg Graduate Fellowship, University of Southern California. 2012-2016.

Invited Talks

“Unlocking New Dimensions: The Power of Multimodal Large Language Models”

Mahidol University Bangkok, Thailand	10/2024
SCBX Bangkok, Thailand	10/2024

“Toward Vision and Richer Language(s)”

Chulalongkorn University	10/2024
Vidyasirimedhi Institute of Science and Technology Rayong, Thailand	09/2024
Advances in Language and Vision Research (ALVR), ACL Workshop Bangkok, Thailand	08/2024
<u>VizWiz Grand Challenge</u> , CVPR Workshop Seattle, WA	06/2024

“All You May Need for VQA are Image Captions”

<u>Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)</u> Seattle, WA	07/2022
The Ohio State University Virtual	06/2022

“Tightly Connecting Vision and Language”

<u>Microsoft Research</u> Virtual	08/2021
Vidyasirimedhi Institute of Science and Technology Rayong, Thailand	11/2020
<u>USC/ISI NL Seminar</u> Marina del Rey, CA	12/2019

“Pushing the Limits of Featurization in Image Captioning”

<u>Conceptual Captions Challenge</u> , CVPR Workshop Long Beach, CA	06/2019
---	---------

“Zero-Shot Learning for Object Recognition in the Wild”

Amazon Seattle, WA	01/2018
Bloomberg New York City, NY	01/2018

“Synthesized Classifiers for Zero-Shot Learning”

The French Institute for Research in Computer Science and Automation (Inria) Lille, France	10/2016
The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Las Vegas, NV	06/2016

Teaching Experience

Graduate Teaching Assistant

Machine Learning (CSCI567)	Fall 2017
Special Topics in Machine Learning (CSCI599)	Spring 2017

Undergraduate Teaching Assistant

Introduction to Machine Learning (CSCI1950f)	Spring 2012
Scientific Computing and Problem Solving (CSCI0040)	Spring 2011
Models of Computation (CSCI0500)	Fall 2010

Professional Service

Senior Program Committee or Area Chair

International Conference on Learning Representations (ICLR) 2025
 International Conference on Artificial Intelligence and Statistics (AISTATS) 2025
 North American Chapter of the Association for Computational Linguistics (NAACL) 2025

Program Committee or Reviewer

Neural Information Processing Systems (NeurIPS) 2016, 2019, 2020, 2021, 2023
 International Conference on Machine Learning (ICML) 2019, 2020
 International Conference on Learning Representations (ICLR) 2021, 2022, 2023, 2024
 Computer Vision and Pattern Recognition (CVPR) 2018, 2019, 2020, 2021, 2022, 2023
 International Conference on Computer Vision (ICCV) 2019
 European Conference on Computer Vision (ECCV) 2018
 ACL Rolling Review (ARR) Jan, Mar, Jun, Sept 2022, Dec (NAACL) 2023
 North American Chapter of the Association for Computational Linguistics (NAACL) 2021, 2024
 Empirical Methods in Natural Language Processing (EMNLP) 2019
 Annual Meeting of the Association for Computational Linguistics (ACL) 2019
 AAAI Conference on Artificial Intelligence (AAAI) 2017, 2019
 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD) 2014
 International Journal of Computer Vision (IJCV) 2022
 Transactions on Knowledge and Data Engineering (TKDE) 2021
 IEEE Signal Processing Letters (SPL) 2019
 Pattern Recognition (PR) 2018
 IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2017, 2018, 2022
 Computer Vision and Image Understanding (CVIU) 2017

Publications

Refereed Journal Articles

- [J3] Yu-Chuan Su, **Soravit Changpinyo**, Xiangning Chen, Sathish Thoppay, Cho-Jui Hsieh, Lior Shapira, Radu Soricut, Hartwig Adam, Matthew Brown, Ming-Hsuan Yang, Boqing Gong. 2.5D Visual Relationship Detection. *Computer Vision and Image Understanding (CVIU)*, 2022.
- [J2] Ben McCamish, Vahid Ghadakchi, Arash Termehchy, Behrouz Touri, Eduardo Cotilla Sanchez, Liang Huang, **Soravit Changpinyo**. A Game-theoretic Approach to Data Interaction. *ACM Transactions on Database Systems (TODS)*, 2020.
- [J1] **Soravit Changpinyo***, Wei-Lun Chao*, Boqing Gong, Fei Sha. Classifier and Exemplar Synthesis for Zero-Shot Learning. *International Journal of Computer Vision (IJCV)*, 2020.

Refereed Conference Proceedings

- [C24] Xi Chen, Josip Djolonga, Piotr Padlewski, Basil Mustafa, **Soravit Changpinyo**, Jialin Wu, Carlos Riquelme Ruiz, Sebastian Goodman, Xiao Wang, Yi Tay, Siamak Shakeri, Mostafa Dehghani, Daniel Salz, Mario Lucic, Michael Tschannen, Arsha Nagrani, Hexiang Hu, Mandar Joshi, Bo Pang, Ceslee Montgomery, Paulina Pietrzyk, Marvin Ritter, AJ Piergiovanni, Matthias Minderer, Filip Pavetic, Austin Waters, Gang Li, Ibrahim Alabdulmohsin, Lucas Beyer, Julien Amelot, Kenton Lee, Andreas Peter Steiner, Yang Li, Daniel Keysers, Anurag Arnab, Yuanzhong Xu, Keran Rong, Alexander Kolesnikov, Mojtaba Seyedhosseini, Anelia Angelova, Xiaohua Zhai, Neil Houlsby, Radu Soricut. PaLI-X: On Scaling up a Multilingual Vision and Language Model. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.
- [C23] Michal Yarom*, Yonatan Bitton*, **Soravit Changpinyo**, Roei Aharoni, Jonathan Herzig, Oran Lang, Eran Ofek, Idan Szpektor. What You See is What You Read? Improving Text-Image Alignment Evaluation. *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. *Patent pending*.
- [C22] Yang Chen, Hexiang Hu, Yi Luan, Haitian Sun, **Soravit Changpinyo**, Alan Ritter, Ming-Wei Chang. Can Pre-trained Vision and Language Models Answer Visual Information-Seeking Questions? *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2023.
- [C21] **Soravit Changpinyo**, Linting Xue, Michal Yarom, Ashish V. Thapliyal, Idan Szpektor, Julien Amelot, Xi Chen, Radu Soricut. MaXM: Towards Multilingual Visual Question Answering. *Findings of the Association for Computational Linguistics: Empirical Methods in Natural Language Processing (EMNLP-Findings)*, 2023.
- [C20] Jihyung Kil, **Soravit Changpinyo**, Xi Chen, Hexiang Hu, Sebastian Goodman, Wei-Lun Chao, Radu Soricut. PreSTU: Pre-Training for Scene-Text Understanding. *IEEE International Conference on Computer Vision (ICCV)*, 2023.
- [C19] Paul Voigtlaender, **Soravit Changpinyo**, Jordi Pont-Tuset, Radu Soricut, Vittorio Ferrari. Connecting Vision and Language with Video Localized Narratives. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
Highlight, top 2.5%
- [C18] Arjun R. Akula, Brendan Driscoll, Pradyumna Narayana, **Soravit Changpinyo**, Zhiwei Jia, Suyash Damle, Garima Pruthi, Sugato Basu, Leonidas Guibas, William T. Freeman, Yuanzhen Li, Varun Jampani. MetaCLUE: Towards Comprehensive Visual Metaphors Research. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
- [C17] Xi Chen, Xiao Wang, **Soravit Changpinyo**, AJ Piergiovanni, Piotr Padlewski, Daniel Salz, Sebastian Goodman, Adam Grycner, Basil Mustafa, Lucas Beyer, Alexander Kolesnikov, Joan Puigcerver, Nan Ding, Keran Rong, Hassan Akbari, Gaurav Mishra, Linting Xue, Ashish Thapliyal, James Bradbury, Weicheng Kuo, Mojtaba Seyedhosseini, Chao Jia, Burcu Karagol Ayan, Carlos Riquelme, Andreas Steiner, Anelia Angelova, Xiaohua Zhai, Neil Houlsby, Radu Soricut. PaLI: A Jointly-Scaled Multilingual Language-Image Model. *International Conference on Learning Representations (ICLR)*, 2023.
Oral presentation, Notable top 5%
- [C16] Khyathi Raghavi Chandu, Piyush Sharma, **Soravit Changpinyo**, Ashish Thapliyal, Radu Soricut. Denoising Large-Scale Image Captioning from Alt-text Data using Content Selection Models. *International Conference on Computational Linguistics (COLING)*, 2022
- [C15] Nan Ding, Xi Chen, Tomer Levinboim, **Soravit Changpinyo**, Radu Soricut. PACTran: PAC-Bayesian Metrics for Estimating the Transferability of Pretrained Models to Classification Tasks. *European Conference on Computer Vision (ECCV)*, 2022.

Oral presentation, 2.7% acceptance rate

- [C14] **Soravit Changpinyo***, Doron Kukliansky*, Idan Szpektor, Xi Chen, Nan Ding, Radu Soricut. All You May Need for VQA are Image Captions. *North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022.

Oral presentation

- [C13] Arjun Akula, Varun Jampani, **Soravit Changpinyo**, Song-Chun Zhu. Robust Visual Reasoning via Language Guided Neural Module Networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- [C12] Tai-Yu Pan*, Cheng Zhang*, Yandong Li, Hexiang Hu, Dong Xuan, **Soravit Changpinyo**, Boqing Gong, Wei-Lun Chao. On Model Calibration for Long-Tailed Object Detection and Instance Segmentation. *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- [C11] Arjun Akula, **Soravit Changpinyo**, Boqing Gong, Piyush Sharma, Song-Chun Zhu, Radu Soricut. CrossVQA: Scalably Generating Benchmarks for Systematically Testing VQA Generalization. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2021.

Oral presentation

- [C10] **Soravit Changpinyo**, Jordi Pont-Tuset, Vittorio Ferrari, Radu Soricut. Telling the What while Pointing to the Where: Multimodal Queries for Image Retrieval. *IEEE International Conference on Computer Vision (ICCV)*, 2021.
- [C9] Cheng Zhang*, Tai-Yu Pan*, Yandong Li, Hexiang Hu, Dong Xuan, **Soravit Changpinyo**, Boqing Gong, Wei-Lun Chao. MosaicOS: Simple and Effective Use of Object-Centric Images for Long-Tailed Object Detection. *IEEE International Conference on Computer Vision (ICCV)*, 2021.
- [C8] **Soravit Changpinyo**, Piyush Sharma, Nan Ding, Radu Soricut. Conceptual 12M: Pushing Web-Scale Image-Text Pre-Training To Recognize Long-Tail Visual Concepts. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- [C7] Jordi Pont-Tuset, Jasper Uijlings, **Soravit Changpinyo**, Radu Soricut, Vittorio Ferrari. Connecting Vision and Language with Localized Narratives. *European Conference on Computer Vision (ECCV)*, 2020.

Oral spotlight presentation, 5.1% acceptance rate

- [C6] **Soravit Changpinyo**, Bo Pang, Piyush Sharma, Radu Soricut. Decoupled Box Proposal and Featurization with Ultrafine-Grained Semantic Labels Improve Image Captioning and Visual Question Answering. *Conference on Empirical Methods in Natural Language Processing and International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.
- [C5] **Soravit Changpinyo**, Hexiang Hu, Fei Sha. Multi-Task Learning for Sequence Tagging: An Empirical Study. *International Conference on Computational Linguistics (COLING)*, 2018.
- [C4] **Soravit Changpinyo**, Wei-Lun Chao, Fei Sha. Predicting Visual Exemplars of Unseen Classes for Zero-shot Learning. *IEEE International Conference on Computer Vision (ICCV)*, 2017.
- [C3] Wei-Lun Chao*, **Soravit Changpinyo***, Boqing Gong, Fei Sha. An Empirical Study and Analysis of Generalized Zero-Shot Learning for Object Recognition in the Wild. *European Conference on Computer Vision (ECCV)*, 2016.

Oral spotlight presentation, 4.9% acceptance rate

- [C2] **Soravit Changpinyo***, Wei-Lun Chao*, Boqing Gong, Fei Sha. Synthesized Classifiers for Zero-Shot Learning. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

Oral presentation, 3.9% acceptance rate

- [C1] **Soravit Changpinyo***, Kuan Liu*, Fei Sha. Similarity Component Analysis. *Advances in Neural Information Processing Systems (NIPS)*, 2013.

Pre-prints and Technical Reports

- [P3] Gemini Team, Google. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. *Technical Report*, 2024.
Part of **Core Contributors** within *Multimodal-Vision*.
- [P2] Gemini Team, Google. Gemini: A Family of Highly Capable Multimodal Models. *Technical Report*, 2023.
Part of **Core Contributors** within *Multimodal-Vision*.
- [P1] **Soravit Changpinyo**, Mark Sandler, Andrey Zhmoginov. The Power of Sparsity in Convolutional Neural Networks. *arXiv*, 2017.

Theses

- [T2] **Soravit Changpinyo**. Modeling, Learning, and Leveraging Similarity. Ph.D. Thesis, University of Southern California, 2018.
- [T1] **Soravit Changpinyo**. Learning Image Attributes using the Indian Buffet Process. Sc.B. Honors Thesis, Brown University, 2012.