Zhengfeng (Jeff) Lai, Ph.D.

Tel: +1 (530)-574-9480 | Email: jeff_lai@apple.com | LinkedIn

Current Position

Machine Learning Researcher, Apple AI/ML Nov. 2023 - Now
Team: Data & Machine Learning Innovation *Cupertino, CA*

- Foundation vision model, multimodal large language model.
- Data-centric approaches for large-scale pre-training.

Education

University of California, Davis Davis, CA
Ph.D. in Electrical and Computer Engineering *Sept. 2019 – Nov. 2023*

Zhejiang University Hangzhou, China
B.Eng. in Information Engineering *Sept. 2015 – June 2019*

Research Interest: Multimodal understanding, semi-supervised learning, computer vision, language-vision model, unsupervised domain adaptation, medical imaging, and AI healthcare.
Dissertation Advisors: Prof. Chen-Nee Chuah, Prof. Sen-ching Cheung, and Prof. Brittany N. Dugger

Preprints

- **Z. Lai**, H. Zhang, W. Wu, H. Bai, A. Timofeev, X. Du, Z. Gan, J. Shan, C-N. Chuah, Y. Yang, M. Cao, "From Scarcity to Efficiency: Improving CLIP Training via Visual-enriched Captions", under review by **ECCV 2024**.
- **Z. Lai**, J. Chauhan, B. Dugger, C-N. Chuah, "Path-CLIP: Efficient Adaptation of CLIP for Pathology Image Analysis with Limited Data", under review by **ECCV 2024**.

Published/Accepted Papers

- **Z. Lai**, J. Chauhan, D. Chen, B. Dugger, S-C. Cheung,C-N. Chuah, "Semi-Path: An Interactive Semi-supervised Learning Framework for Gigapixel Pathology Image Analysis", to be appear in **Elsevier Smart Health Journal**.
- **Z. Lai**, H. Bai, H. Zhang, X. Du, J. Shan, Y. Yang, C-N. Chuah, M. Cao, "Empowering Unsupervised Domain Adaptation with Large-scale Pre-trained Vision-Language Models," **WACV 2024**.
- **Z. Lai**, S. Vesdapunt, N. Zhou, J. Wu, X. Li, C. Huynh, C-N. Chuah, "PADCLIP: Pseudo-labeling with Adaptive Debiasing in CLIP for Unsupervised Domain Adaptation," **ICCV 2023**.
- **Z. Lai**, C. Wang, H. Gunawan, S-C. Cheung, and C-N. Chuah, "Smoothed Adaptive Weighting for Imbalanced Semi-Supervised Learning: Improve Reliability Against Unknown Distribution Data," **ICML 2022**. (**Won Participation Grant**)
- **Z. Lai**, C. Wang, S-C. Cheung, and C-N. Chuah, "SaR: Self-adaptive Refinement on Pseudo Labels for Multiclass-Imbalanced Semi-supervised Learning," 2022 **CVPR** Workshop on Learning with Limited Labelled Data for Image and Video Understanding. (**Best Paper Award**)
- **Z. Lai**, L. Cerny Oliveira, R. Guo, W. Xu, Z. Hu, K. Mifflin, C. DeCarlie, S-C. Cheung, C-N. Chuah, and B. N. Dugger, "BrainSec: Automated Brain Tissue Segmentation Pipeline for Scalable Neuropathological Analysis," **IEEE Access**, 2022.
- **Z. Lai***, C. Wang*, L. Cerny Oliveira, B. Dugger, S-C. Cheung, C-N. Chuah, "Joint Semi-supervised and Active Learning for Segmentation of Gigapixel Pathology Images with Cost-Effective Labeling," **ICCV 2021** Workshop.
- **Z. Lai**, C. Wang, Z. Hu, B. Dugger, S-C. Cheung, C-N. Chuah, "A Semi-supervised Learning for Segmentation of Gigapixel Histopathology Images from Brain Tissues," 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2021.

- **Z. Lai**, P. Vadlaputi, D. J. Tancredi, M. Garg, R. I. Koppel, M. Goodman, W. Hogan, N. Cresalia, S. Juergensen, E. Manalo, S. Lashminrusimha, C-N. Chuah, and H. Siefkes, "Enhanced Critical Congenital Cardiac Disease Screening by Combining Interpretable Machine Learning Algorithms," 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2021.
- L. Cerny Oliveira, **Z. Lai**, D. Harvey, K. Nzenkue, L-W. Jim, C. DeCarlie, C-N. Chuah, and B. N. Dugger, "Pre-analytic variable effects on segmentation and quantification machine learning algorithms for amyloid beta analyzes on digitized human brain slides," **Journal of Neuropathology and Experimental Neurology**, 82(3):212-220, Feb 2023.
- L. Cerny Oliveira, **Z. Lai**, H. Siefkes, C-N. Chuah, "Generalizable Semi-supervised Learning Strategies for Multiple Learning Tasks using 1-D Biomedical Signals," **NeurIPS** Workshop on Learning from Time Series for Health, Dec 2022.
- L. Cerny Oliveira, **Z. Lai**, W. Geng, H. Siefkes, C-N. Chuah, "A Machine Learning Driven Pipeline for Automated Photoplethysmogram Signal Artifact Detection", 1st Workshop on Artificial Intelligence and Internet of Things for Digital Health (AIIOT4DH) at IEEE CHASE, Dec 16-17, 2021.
- K. Doshi, G. Rehm, P. Vadlaputi, **Z. Lai**, S. Lakshminrusimha, C-N. Chuah, and H. M Siefkes, "A Novel System to Collect Dual Pulse Oximetry Data for Critical Congenital Heart Disease Screening Research," Journal of Clinical and Translational Science, pp. 1-25, October 2020.
- C. Linghu, C. Wang, N. Cen, J. Wu, **Z. Lai**, J. Song, "Rapidly Tunable and Highly Reversible Bio-inspired Dry Adhesion for Transfer Printing in Air and a Vacuum," Soft Mater, 2019.

Industrial Experience

ML Research Intern at Apple March 2023 – Sept. 2023
Team: AI/ML *Cupertino, CA*
- Foundation vision model, multimodal understanding.
- Data-centric approaches for large-scale pre-training.

Applied Scientist Intern at Amazon June 2022 – Sept. 2022
Team: Lab126 *Sunnyvale, CA*
- Unsupervised domain adaption with CLIP.
- Causal inference and counterfactual reasoning.

PhD Data Science Intern at Electronic Arts June 2021 – Sept. 2021
Advisor: Jason Park *Redwood city, CA*
- Designed a few-shot toxic object detection pipeline with transfer learning.
- Explored zero-shot learning for unseen classes in the dataset.

National Science Foundation's Innovation Corps Program (I-Corps) Jan. 2021 – Mar. 2021
Team with Dr. Heather Siefkes, Dr. Jim Swick, Pranjali Vadlaputi *Washington, D.C.*
- Successfully completed the requirements of I-Corps by finishing more than 100 interviews in pediatric field.
- Named as one of the inventors on a patent application **"Systems and Methods for Classifying Critical Heart Defects."**

Artificial Intelligence Engineer Intern June 2018 – Aug. 2018
DeepThink (Top 10 AI Startup Company in Hangzhou) *Hangzhou, China*
- Implemented an ICO Scan Identification System based on RNN with real-world white books.
- Designed a LSTM model to predict Ethereum trends using blockchain activity data with 65% of accuracy.
- Maintained the Linux Server for 20 members in the group.

Teaching & Mentoring Experience

Lead Teaching Assistant for EEC 193AB Sept. 2019 – Mar. 2021
University of California, Davis *Davis, CA*

- Helped develop and teach EEC 193 AB (AI Systems Senior Design) for two years.
- Independently hosted lab sessions and mentored four teams over 2 academic quarters.
- Designed three lab assignments involving classical ML algorithms (Logistic Regression, SVM), CNN and basic Python in the application of ML on health.

Summer Undergraduate Research Mentor June 2021 – Sept. 2021
University of California, Davis *Davis, CA*

- Mentored one undergraduate on website design for visualizing waveform from photoplethysmogram signals.
- Mentored one undergraduate on waveform artifact detection by using machine learning algorithms.
- Designed three lab assignments involving classical ML algorithms (Logistic Regression, SVM), CNN and basic Python in the application of ML on health.

Professional Services

Reviewer:

- NeurIPS 2021 & 2022 & 2023
- ICML 2022 & 2023
- IEEE Transaction on Image Processing
- CVPR 2022 & 2023 & 2024
- ICCV 2023
- ICLR 2024

Technical Skills

Programming: Python, C/C++, Matlab
Developer Tools: Git, Docker, Google Cloud Platform, VS Code, Visual Studio, PyCharm
Frameworks: PyTorch, Tensorflow, Caffe, OpenCV, Scikit-Learn

Awards

- 2018 Interdisciplinary Contest In Modeling: Meritorious Winner
- Outstanding Senior Design Project Award of UC Davis, 2019
- The Best Senior Design of ISEE, ZJU in 2019: Multiple Objects Detection
- 2019 ZJU Overseas Senior Design Scholarship
- 2022 Smita Bakshi Digital Learning and Teaching Award
- 2022 ICML Participation Award
- 2022 CVPR Workshop **Best Paper Award**
- 2022 AANP R13 Grant Travel Award