

Mark Endo

PhD Student · Stanford University

Department of Computer Science

Email: markendo@stanford.edu · [Homepage](#) · [Google Scholar](#)

Education

Stanford University, PhD in Computer Science 09/2023 - Present

- Advisor: [Prof. Serena Yeung-Levy](#) and [Prof. Fei-Fei Li](#)

Stanford University, BS in Computer Science 08/2018 - 04/2023

- GPA: 4.06

Summary

My research interests span the areas of computer vision, representation learning, and multimodal learning, with an emphasis on applications in healthcare. I currently focus on foundational advancements in multimodal models to improve learning of interpretable, efficient, and robust representations for solving challenging visual problems ([ICCV '25](#), [arXiv](#)). My earlier work explored similar ideas in domains such as human motion understanding ([ICML '23](#), [MICCAI '22](#), [Nature Machine Intelligence '24](#)) and automated radiology report generation ([ML4H '21](#), [MIDL '21](#), [Patterns '23](#)).

Publications

* indicates equal contribution

- **Mark Endo**, Serena Yeung-Levy. Downscaling Intelligence: [Exploring Perception and Reasoning Bottlenecks in Small Multimodal Models](#). *Under Review 2025*.
- Anita Rau, **Mark Endo**, Josiah Aklilu, Jaewoo Heo, Khaled Saab, Alberto Paderno, Jeffrey Jopling, F. Christopher Holsinger, Serena Yeung-Levy. [Systematic Evaluation of Large Vision-Language Models for Surgical Artificial Intelligence](#). *Under Review 2025*.
- **Mark Endo**, Xiaohan Wang, Serena Yeung-Levy. [Feather the Throttle: Revisiting Visual Token Pruning for Vision-Language Model Acceleration](#). *International Conference on Computer Vision (ICCV) 2025*.
- **Mark Endo**, Favour Nerrise, Qingyu Zhao, Edith V. Sullivan, Li Fei-Fei, Victor W. Henderson, Kilian M. Pohl, Kathleen L. Poston, Ehsan Adeli. [Data-Driven Discovery of Movement-Linked Heterogeneity in Neurodegenerative Diseases](#). *Nature Machine Intelligence 2024*.
- **Mark Endo***, Joy Hsu*, Jiaman Li, Jiajun Wu. [Motion Question Answering via Modular Motion Programs](#). *International Conference on Machine Learning (ICML) 2023*.

- Feiyang Yu*, **Mark Endo***, Rayan Krishnan*, Ian Pan, Andy Tsai, Eduardo Pontes Reis, Eduardo Kaiser Ururahy Nunes Fonseca, Henrique Min Ho Lee, Zahra Shakeri Hossein Abad, Andrew Y. Ng, Curtis P. Langlotz, Vasantha Kumar Venugopal, Pranav Rajpurkar. [Evaluating Progress in Automatic Chest X-Ray Radiology Report Generation](#). *Patterns - Cell Press* 2023.
- **Mark Endo**, Kathleen Poston, Edith Sullivan, Li Fei-Fei, Kilian Pohl, Ehsan Adeli. [GaitForeMer: Self-Supervised Pre-Training of Transformers via Human Motion Forecasting for Few-Shot Gait Impairment Severity Estimation](#). *Medical Image Computing and Computer Assisted Intervention (MICCAI)* 2022.
- **Mark Endo***, Rayan Krishnan*, Viswesh Krishna, Andrew Y. Ng, Pranav Rajpurkar. [Retrieval-Based Chest X-Ray Report Generation Using a Pre-trained Contrastive Language-Image Model](#). *Machine Learning for Health (ML4H)* 2021.
- Soham Gadgil*, **Mark Endo***, Emily Wen*, Andrew Y. Ng, Pranav Rajpurkar. [CheXseg: Combining Expert Annotations with DNN-generated Saliency Maps for X-ray Segmentation](#). *Medical Imaging with Deep Learning (MIDL)* 2021.

Selected Awards/Honors

NSF Graduate Research Fellowship Program	04/2024
Paper featured on the cover of Cell Press' <i>Patterns</i> journal	09/2023
Phi Beta Kappa, Stanford University	06/2023
Special Recognition Award for Outstanding Psychiatric Research, Stanford Department of Psychiatry and Behavioral Sciences	07/2022

Teaching Experience

CS 106A (Programming Methodology) / CS106B (Programming Abstractions), Section Leader	01/2020 - 08/2021
CS198 (Teaching Computer Science), Small Group Leader	09/2020 - 12/2020

Presentations

Harvard Medical AI Research Intensive: Walkthrough of a deep learning codebase.	08/2022
OpenCV Weekly Webinar. Semi-supervised Medical Image Segmentation with CheXseg.	03/2022
Harvard Medical AI: Mark Endo presents Masked Autoencoders Are Scalable Vision Learners.	02/2022
Short Oral Presentation of CheXseg: Combining Expert Annotations with DNN-generated Saliency Maps for X-ray Segmentation.	07/2021