

GAUTAM RAJENDRAKUMAR GARE

Inference-Time Adaptation & Alignment of Vision–Language Models

@ ggare@andrew.cmu.edu in gautam-gare ggare-cmu.github.io Pittsburgh, PA

Research Summary: I make **frozen vision–language models (VLMs)** do new things without touching their weights, treating adaptation as an inference-time problem: optimizing prompts (**DetPO, ECCV 2026**), selecting image-conditioned attributes, and reading activations (**ACL Findings 2026**). I study when and why multimodal models generalize (**ICML 2024 Oral**) and prove these methods in the high-stakes setting of medical imaging (5 ISBI papers, **4 U.S. patents**). **On the job market for research scientist roles** (graduating late 2026).

EDUCATION

Ph.D. in Robotics (Candidate)

Carnegie Mellon University

Aug 2021 – Late 2026 (expected) Pittsburgh, PA

- Advisors: Prof. John Galeotti, Prof. Deva Ramanan
- Focus: Inference-time adaptation and alignment of VLMs; generalization; medical imaging AI.

M.S. in Electrical and Computer Engineering

Carnegie Mellon University

Dec 2020 Pittsburgh, PA

- GPA: 4.00/4.00 - Courses: SLAM, CV, ML, Optimization, Deep Learning

B.E. in Electronics and Communication

BMS College of Engineering

May 2016 Bangalore, India

RESEARCH EXPERIENCE

Soft-Prompt Tuning of Vision–Language Models for Few-Shot Object Detection

Robotics Institute, CMU

Nov 2025 – Present

- Designed a soft-prompting framework that learns continuous prompt embeddings via gradient-based optimization, enabling efficient task adaptation without modifying the underlying VLM weights.
- Outperformed LoRA fine-tuning in the 10-shot setting, demonstrating that lightweight prompt optimization can match or exceed parameter-efficient fine-tuning methods at low data regimes.

DetPO: Iterative Prompt Optimization for Few-Shot Object Detection (ECCV 2026)

Robotics Institute, CMU

Sep 2025 – Jan 2026

- Developed an iterative prompt optimization framework using false-positive and false-negative feedback to improve few-shot object detection; accepted at **ECCV 2026**.
- Achieved state-of-the-art results with contemporary VLMs (e.g., Qwen3, Gemini3) on the RF20 object detection dataset, outperforming specialist models such as GroundingDINO under a black-box optimization setting.

Distribution-Conditioned Attribute Selection for VLMs

Robotics Institute, CMU

2026

- Showed LLM-generated class descriptors are label-conditioned, not image-conditioned: removing class names collapses ImageNet accuracy from 59.5% to 15.5%.
- Selecting attributes from target images in CLIP’s embedding space outperforms CoOp prompt tuning with one image per class, fitting in under a minute instead of 14 hours.

Causal Feature Selection & Natural Experiments

Robotics Institute, CMU

📅 Jan 2023 – Aug 2025

- Designed a Markov-blanket-based causal discovery algorithm for feature selection, improving accuracy and interpretability on medical datasets.
- Showed real-world datasets contain natural experiments exploitable via causal inference.

Generic Human-Interpretable Lung Ultrasound Biomarkers

Biomedical Image Guidance Lab, CMU

📅 2021 – 2025

- Developed interpretable LUS biomarkers enabling causal and contrastive learning for diagnosis.
- Showed biomarker-based features match deep features in downstream performance.

SELECTED PUBLICATIONS

- **Gare, G. R.***, N. Peri*, M. Popov, S. Jain, J. Galeotti, D. Ramanan. (*equal contribution)
DetPO: In-Context Learning with Multi-Modal LLMs for Few-Shot Object Detection.
European Conference on Computer Vision (ECCV), 2026. [paper] [project] [code]
- **Gare, G.R.**, J. Shi, Z. Lin, D. Pathak, J. Galeotti, D. Ramanan.
Attributes Should Come from Images, Not Class Names: Distribution-Conditioned Attribute Selection for VLMs.
PFATCV Workshop at ECCV, 2026. **(Oral)** [paper] [project]
- R. H. Abhinandan, J. Galeotti, D. Ramanan, **G.R. Gare**.
Ask Twice, Look Twice: Prompt Echoing Resolves the Question-First Paradox in Vision-Language Models.
eXCV Workshop at ECCV, 2026. [paper] [project]
- T. Chai, C. Mitra, B. Huang, **G.R. Gare**, Z. Lin, A. Arbelle, L. Karlinsky, R. Feris, T. Darrell, D. Ramanan, R. Herzig.
Activation reward models for few-shot model alignment.
Findings of the Association for Computational Linguistics (ACL), 2026. [paper]
- J. Shi, **G.R. Gare**, J. Tian, S. Chai, Z. Lin, A. Vasudevan, D. Feng, F. Ferroni, S. Kong.
LCA-on-the-Line: Benchmarking Out-of-Distribution Generalization with Class Taxonomies.
International Conference on Machine Learning (ICML), 2024. **(Oral)** [paper] [project]
- **Gare, G.R.**, J. M. Galeotti, M. C. Mozer, D. Ramanan, N. R. Ke.
Do Real-World Datasets Contain Natural Experiments? An Empirical Study Using Causal Feature Selection.
arXiv:2606.03251, 2026. [paper]
- **Gare, G.R.**, T. Fox, P. Lowery, K. Zamora, H. Tran, L. Hutchins, D. Montgomery, A. Krishnan, D. Ramanan, R. Rodriguez, B. deBoisblanc, J. Galeotti.
Human-Aligned Learning of Interpretable Biomarkers for Lung Ultrasound Diagnosis.
IEEE International Symposium on Biomedical Imaging (ISBI), 2026. [paper]
- J. Armouti, N. Madaan, R. Panda, T. Fox, L. Hutchins, A. Krishnan, R. Rodriguez, B. DeBoisblanc, D. Ramanan, J. Galeotti, **G.R. Gare**.
LEARNER: Learning Granular Labels from Coarse Labels using Contrastive Learning.
IEEE International Symposium on Biomedical Imaging (ISBI), 2026. [paper]
- M. Shivaram, **G.R. Gare**, L. Hutchins, J. Duplantis, T. Deiss, T.N. Gomes, T. Tran, K.H. Patel, K. Kirby, T.H. Fox, A. Krishnan, D. Ramanan, B. DeBoisblanc, R. Rodriguez, J. Galeotti.
Label uncertainty for ultrasound segmentation.
IEEE International Symposium on Biomedical Imaging (ISBI), 2026. **(Oral)** [paper]
- **G.R. Gare**, H.V. Tran, B.P. deBoisblanc, R.L. Rodriguez, J.M. Galeotti. *Weakly supervised contrastive learning for better severity scoring of lung ultrasound.* **IEEE International Symposium on Biomedical Imaging (ISBI)**, 2025. [paper]
- **Gare, G.R.**, T. Fox, B. Chansangavej, A. Krishnan, R. Rodriguez, B. deBoisblanc, D. Ramanan, J. Galeotti.
Improving Model's Interpretability and Reliability using Biomarkers.
Bern Interpretable AI Symposium (BIAS), 2023. [paper]
- **Gare, G.R.**, J. Li, R. Joshi, R. Magar, M. P. Vaze, M. Yousefpour, R. Rodriguez, J. Galeotti.
W-Net: Dense and diagnostic semantic segmentation of subcutaneous and breast tissue in ultrasound.
Medical Image Analysis (MIA), 76:102326, 2022. [paper]
- **G.R. Gare**, A. Schoenling, V. Philip, H.V. Tran, B.P. deBoisblanc, R.L. Rodriguez, J.M. Galeotti. *Dense pixel-labeling for reverse-transfer and diagnostic learning on lung ultrasound for COVID-19 and pneumonia detection.* **IEEE International Symposium on Biomedical Imaging (ISBI)**, 2021, pp. 1406–1410. [paper]

PROFESSIONAL EXPERIENCE

AI Research Intern Robert Bosch LLC (Bosch Research)

May 2026 – Aug 2026 Pittsburgh, PA

- Researched grounding of vision–language–action (VLA) models for autonomous-vehicle applications, connecting language-conditioned action reasoning to visual scene evidence.

Computer Vision Research Intern DawnLight Inc.

Jun 2021 – Aug 2021 Palo Alto, CA

- Developed efficient multimodal action recognition networks (RGB + IR) for clinical monitoring.

Senior Data Scientist & Software Engineer Sling Media R&D

Jul 2016 – Jul 2019 Bangalore, India

- Led analytics pipeline for user churn modeling and AirTV Player Android App interface optimization.
- Improved image caching, reducing startup time by 20%.

Summer Research Fellow IIST, Shibpur

Jun 2015 – Jul 2015 India

- Built real-time tele-operated robotic arm system with RGB-D feedback.

SELECTED PATENTS

- US12154354B2 – **System and Method for Labeling Ultrasound Data**, granted 2024.
- US20240177000A1 – **System and Method for Training Machine-Learning Models with Probabilistic Confidence Labels**, 2024.
- US20240177445A1 – **System and Method for Direct Diagnostic and Prognostic Semantic Segmentation of Images**, 2024.
- US20240249498A1 – **System, Method, and Computer Program Product for Extracting Features from Imaging Biomarkers with ML Models**, 2024.

FELLOWSHIPS & AWARDS

- Liang Zhao Endowed Fellowship, CMU (2025)
- CMLH Translational Fellowship, CMU (2022)
- NSF XSEDE Compute Grant (130,500 GPU hours, 2020)
- Neocortex EUP Compute Award, NSF PSC (2020)

TEACHING & SERVICE

- Guest Lecturer – 16-720 Computer Vision (Fall 2026)
- Research Mentor – supervised the prompt-echoing VLM study (arXiv:2607.15565, 2026)
- Research Mentor – supervised the lung-ultrasound CHF-readmission study (arXiv:2605.18878, 2026)
- Teaching Assistant – 16-720 Computer Vision (Spring 2022, Fall 2021)
- Reviewer – ECCV (2026), eXCV Workshop at ECCV (2026), MICCAI (2022–2024), HSMR (2023)
- Invited Panelist – LLCOVID19 MICCAI Workshop (2021)
- Mentorship – CMU AI Mentoring Program (2021)

MEDIA COVERAGE

- Megan Harris, “*Seeing Beneath The Surface*”, **LINK Magazine**, School of Computer Science, Carnegie Mellon University, Summer 2022 Issue. [article]
- “*CMU and others propose a highly practical ‘model generalization’ metric, improving out-of-distribution accuracy by 6% | ICML 2024 Oral*”, **New Zhiyuan**, December 26, 2024, covering *LCA-on-the-Line*. [article]