

Sina Malakouti

✉ siinamalakouti@gmail.com or sem238@pitt.edu | 🏠 sinamalakouti.github.io | 📧 sinamalakouti | 🌐 sinamalakouti

Legal status in the US: Permanent Resident (Green Card holder)

Education

University of Pittsburgh

PhD in Computer Science

Aug 2020 - Oct 2025

Pittsburgh, PA

- Advisor: Adriana Kovashka

Amirkabir University of Technology

B.Sc. in Software Engineering

Sep 2015 - May 2020

Tehran, Iran

- Advisors: Maryam Amir Haeri and Saeedeh Montazi

Interests

Computer Vision Multimodal Learning Foundational Models (LLM, Vision-Language) Generative AI Robustness

Publications

- *Benchmarking VLMs' Reasoning About Persuasive Atypical Images*, **WACV'25**
- *Incorporating Geo-Diverse Knowledge into Prompting for Increased Geographical Robustness in Object Recognition*, **CVPR'24**
- *Semi-Supervised Domain Generalization for Object Detection via Language-Guided Feature Alignment*, **BMVC'23**
- *A MuST for Consistency Regularization in Semi-Supervised Medical Image Segmentation*
- *DeepTreeNetworks: A New Symbolic Deep Architecture*, DeCoDeML workshop, **ECML PKDD'19**

Presentations & Talks

- *(Invited Talk)* Introduction to Labeled-Efficient Deep Learning Approaches, From Few to None: Exploring Few-Shot, One-Shot, and Zero-Shot Deep Learning in Clinical Settings tutorial, **BHI'23**
- *DeepTreeNetworks: A New Symbolic Deep Architecture*. DeCoDeML workshop, **ECML PKDD'19**

Technical Skills

Programming Languages Python, Java, MATLAB, SQL, C/C++, R

ML & Deep Learning PyTorch, DL4j, Scikit-learn, Weka, Keras, Tensorflow, Numpy, Pandas

Big data Hadoop, Spark

Web Programming JavaScript, Vue.js, Node.js, Express.js, jQuery, HTML/CSS, Flask, Jetty

Database MySQL, MongoDB, SQLite

Misc Data Engineering and Cleaning, Object Oriented, MVC, Problem-Solving

Experience

Graduate Research Assistant

University of Pittsburgh

Aug 2020 - Present

Pittsburgh, PA

- Researching on making ML methods more **robust** and capable of understanding and **reasoning** about complex visual scenes with a focus on **vision-language** and **large language models**.

Applied Scientist Intern

Prime Video, Amazon

May 2024 - Sep 2024

New York, NY

- Devised and executed a research plan (from data acquisition, cleaning to proposing and executing solution) to tackle a new problem of using multimodal data (text and image) for content understanding and duplicate detection.
- Developed two methods: CLIP-based vision-language model (VLM) with new data fusion and generative VLM (InternVL/Claude-3) with chain-of-thought, achieving >10% improvement and highlighting limitations, future possibilities and improvements. **S3, SageMaker**; *To be submitted*

Applied Research Intern

Search Science, eBay

May 2023 - Aug 2023

San Jose, CA

- Employed **vision-language models (CLIP)** and a novel **transformer-based Mixture-of-Modality-Experts fusion** model, significantly boosting results on search and ranking tasks. **PyTorch, Spark, Hadoop**

Machine Learning Image Processing Intern

Image Signal Processing (ISP), Apple

May 2022 - September 2022

Cupertino, CA

- Developed efficient models for computer vision and Image Processing tasks, achieving enhanced performance and efficiency over state-of-the-art methods and baselines. **Python, PyTorch, and Matlab.**

Machine Learning Research Assistant, Intern

Johannes Gutenberg University

July 2018 - Sep 2019

Mainz, Germany

- Proposed a novel efficient symbolic deep architecture with differentiable decision trees, achieving superior performance on imbalanced data. **Java, DL4j, Weka**

Machine Learning Engineer, Intern

Shahid Rajaei Hospital & Research Center

June 2019 - Sep 2019

Tehran, Iran

- Developed ML pipeline, predicting pulmonary complication with 20% improvement. **Python, scikit-learn, Flask.**

Selected Projects

- **Improved abstract reasoning and spatial reasoning for Diffusion Models** Fall 2024- ongoing
PyTorch, Generative AI, Diffusion Models, Reasoning
 - Investigating and addressing limitation of text-to-image diffusion models in spatial reasoning and understanding abstract concepts for accurate representation of scene, objects, attributes and relations.
- **Multi-Modal Reasoning for Understanding Advertisement Images** WACV'25
PyTorch, Large Language Models (LLM), Multimodal Large Language Modeling (MLLM)
 - Benchmarked 3 novel tasks and proposed a novel semantically hard negative generation method to assess MLLM (e.g., LLaVA, InternVL, InstructBLIP, GPT4-V) understanding of complex visual reasoning data. Developed an atypicality-aware verbalization strategy that mitigates MLLM's lack of reasoning ability, significantly improving ad image understanding in a zero-shot manner.
- **Domain Robustness with Soft Prompting in Vision-Language Object Recognition** CVPR'24
LLM, Parameter Efficient Finetuning (PEFT), Domain Robustness, Vision-Language Models (VLM)
 - Proposed a novel distillation-based approach leveraging LLMs' extensive world knowledge to learn generalized soft prompts in a few-shot manner, enhancing cross-geography generalization.
- **Cross-Domain Descriptive Multi-Scale Learning for Object Detection** BMVC'23
Contrastive Learning, Vision-Language Pre-training (VLP), Domain Robustness, Object Detection
 - Developed a novel multi-scale method by proposing a contrastive consistency objective to enforce descriptive consistency in the language feature space, preserving essential semantic information and improving object detection performance by up to 12%.
- **MuST for Semi-Supervised Medical Image segmentation**
Python, PyTorch, Data Augmentation, Consistency Regularization, Semantic Segmentation
 - Proposed a novel consistency regularization framework for brain lesion segmentation with feature-space augmentation. Achieved novel performance by only having 3% labeled data.

Professional Services

Conference Reviewer: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024
European Conference on Computer Vision (ECCV), 2024
Winter Conference on Applications of Computer Vision (WACV), (2022, 2024-2025)
Empirical Methods in Natural Language Processing (EMNLP), 2022
Association for the Advancement of Artificial Intelligence (AAAI), 2024

Honors & Awards

- Outstanding Reviewer Award, European Conference on Computer Vision (ECCV), 2024
- Department of Computer Science Travel Award, University of Pittsburgh (2023)
- Full SCI Fellowship, University of Pittsburgh (2020)
- Honored as an outstanding student, Amirkabir University of Technology (2015-2020)

Extra Curricular & Leadership

President of Student Scientific Chapter

Computer Engineering, Amirkabir University of Technology

Jan 2017 - March 2018

Tehran, Iran

- Organized 70+ national and international contests, talks, and workshops in collaboration with Technische Universität München, Germany, and KTH Royal Institute of Technology, Sweden.