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 — Expected: Jan 2026
PhD in Computer Science — *Pittsburgh, PA*

- Advisor: Adriana Kovashka
- Committee: Adriana Kovashka, Boqing Gong, Xiange Lorraine Li, Milos Hauskrecht
- Thesis: *Compositional Gaps in Object Representations for Generative and Discriminative Models*

Amirkabir University of Technology — Sep 2015 – May 2020
B.Sc. in Software Engineering — *Tehran, Iran*

Selected Peer-Reviewed Publications

- **Role Bias in Text-to-Image Diffusion Models: Diagnosing and Mitigating Compositional Failures through Intermediate Decomposition**. **S. Malakouti**, A. Kovashka. **Under Review**
- **Benchmarking VLMs' Reasoning About Persuasive Atypical Images**. **S. Malakouti***, A. Aghazadeh*, A. Khandelwal, A. Kovashka. **WACV 2025**
- **Incorporating Geo-Diverse Knowledge into Prompting for Increased Geographical Robustness in Object Recognition**. K. Buettner, **S. Malakouti** *(major contributor)* , X.L. Li, A. Kovashka. **CVPR 2024**
- **Semi-Supervised Domain Generalization for Object Detection via Language-Guided Feature Alignment**. **S. Malakouti**, A. Kovashka. **BMVC 2023**
- **DeepTreeNetworks: A New Symbolic Deep Architecture**. **S. Malakouti***, Z. Ahmadi*, S. Kramer. **DeCoDeML Workshop, ECML PKDD 2019**

Experience

Graduate Research Assistant — Sep 2022 - Present
University of Pittsburgh — *Pittsburgh, PA*

- **Research Focus:** Developing robust ML methods for compositional understanding and against domain shift, focusing on Foundational Models such as vision-language models (VLMs), Multimodal LLMs, and text-to-image diffusion models
- **Publications:** Publication at well-known venues (CVPR, WACV, BMVC)

Research contribution:

- Benchmarked T2I models on **long-tailed cross-cultural generation**, revealing significant limitations and biases in representing developing countries. Designed a novel **Q-former-based framework** utilizing cultural descriptions for improved cross-cultural generation
- Proposed **RoleBench** revealing significant **directional role bias** in T2I models; developed **LLM-guided decomposition framework** reducing bias by **15.2 points** and achieving **>70% human preference**
- Curated **PersuasiveAdVLM**, first benchmark for MLLMs on persuasive ads with unusual objects, revealing significant **reasoning gaps** between MLLMs and LLMs and **lack of visual reasoning in MLLMs**
- Proposed **atypicality-aware chain-of-thought** prompting method, improving zero-shot reasoning on unusual ads by **40%**
- Designed **novel soft prompt learning** method adapting CLIP for cross-cultural object recognition by distilling LLM knowledge, achieving **state-of-the-art performance**
- Developed **multi-scale contrastive** method learning robust visual features through consistency objectives in language space, improving cross-domain object detection by **12%** without target domain data

Applied Scientist Intern — May 2024 - Sep 2024
Prime Video, Amazon — *New York, NY*

- Led research on multimodal content understanding & duplicate detection. Developed a novel **CLIP-based fusion model** and multimodal LLM (Claud, InternVL) chain-of-thought, achieving $>$ 10% improvement. **S3, SageMaker**

Applied Research Intern — May 2023 - Aug 2023
Search Science, eBay — *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**

Computer Vision Intern
May 2022 - September 2022
Apple — *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
July 2018 - Sep 2019
Johannes Gutenberg University — *Mainz, Germany*

· Designed an efficient symbolic deep network using differentiable decision trees, effective on imbalanced data.

Technical Skills

Programming Languages	Python, Java, MATLAB, SQL, C, R, JavaScript, HTML/CSS
ML Tools	PyTorch, TensorFlow, Keras, Scikit-learn, DL4j, Weka, Numpy, Pandas
AI & CV Methods	CNNs, RNNs, Transformers & Attention Mechanism, Vision-Language Models (VLMS), Large Language Models (LLMs), Multimodal LLMs (MLLMs), Text-to-Image (T2I) Generative Models, Diffusion Models, Contrastive Learning, Semi-Supervised, Domain Adaptation/Generalization (e.g., Pseudo Labeling, Student-Teacher, Consistency),
Big Data & Databases	Hadoop, Spark, S3, MySQL, MongoDB, SQLite
Other	Data Engineering, Object-Oriented Design, MVC, Problem-Solving

Other Projects

- **Compositional Generalization of T2I models** — Ongoing

  · Benchmarking T2I models to test their compositional generalization power for faithfully generating complex prompts. Developing **curriculum-based method** to enhance the generalizability of T2I models

- **Cross-Cultural Creative ads Generation** — Ongoing

  · Developing **cross-cultural theory-based metrics** and reward models with **feedback-based editing** to improve cultural relevance and effectiveness across diverse regions

- **Weakly Supervised Object Detectors (WSOD) Robustness Toward Domain Shift**

  · Underlined more reliance of WSOD on domain-specific features compared to fully supervised models. Developed **consistency regularization** method with style transfer, improved detection on unseen domains by **2%**

- **Multimodal Transformer Fusion For Depression Prediction**

  · Developed novel approach for depression severity prediction through **multimodal transformer fusion** of unaligned video, language, and audio features

- **MuST for Semi-Supervised Medical Image segmentation**

  Proposed **feature-space consistency** method for brain lesion segmentation, achieving **SOTA performance** with only **3% labeled data**

Professional Services

Conference Reviewer: CVPR, ICCV, ECCV, NeurIPS, AAAI, EMNLP, WACV
Co-Organizer: Demographic Diversity in Computer Vision Workshop, CVPR 2025

Honors & Awards

· **Doctoral Consortium**, Winter Conference on Applications of Computer Vision (WACV), 2025
· **Outstanding Reviewer Award**, European Conference on Computer Vision (ECCV), 2024
· **Travel Award**, Department of Computer Science 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
Jan 2017 - March 2018
Computer Engineering, Amirkabir University of Technology — *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.