Aysan Aghazadeh

(+1) 412 916 2139 aysan.aghazadeh@gmail.com
aysanaghazadeh Aysan Aghazadeh aysanaghazadeh.github.io

Education

University of Pittsburgh Sep. 2021 – Dec. 2025
Ph.D. in Computer Science, Advisor: Dr. Adriana Kovashka *Pittsburgh, PA*

Amirkabir University of Technology Sep. 2014 – July 2019
B.Sc. in Computer Engineering, *Tehran, Iran*
Outstanding courses Computer Vision, Machine Learning, Artificial Intelligence, Statistics, Data Mining, Database Design, Signals and Systems, Human-Computer Interaction

Online courses Deep Learning, Machine Learning, NLP (Stanford), Linear Algebra (Imperial College London)

Interests

Computer Vision Multimodal Reasoning Generative AI Foundational Models (MLLMs, VLMs, LLMs)

Publications

- *'CAP: Evaluation of Persuasive and Creative Image Generation'*, **Submitted to CVPR 2025**
- *'Benchmarking VLMs' Reasoning About Persuasive Atypical Images'*, **WACV 2025.**
- *'A Distributed Approximate Nearest Neighbor Method for Real-Time Face Recognition'*

Technical Skills

Languages Python, Java, MATLAB, C/C++, SQL, R
ML & Deep Learning PyTorch, Transformers, NLTK, OpenCV, Numpy, Scikit-learn, Pandas, Tensorflow, Keras
Cloud Services Amazon AWS, Oracle Cloud
Web Programming HTML/CSS, Javascript, Flask, jQuery
Database MongoDB, MySQL
Tools Git, Docker, LaTeX, Postman, RapidMiner Studio, ImageJ, ITK-SNAP
Misc Data Cleaning, MVC, Problem Solving

Experience

Graduate Research Assistant Sep. 2021 - Present
University of Pittsburgh *Pittsburgh, PA*
- Conducted research on **text-to-image generation models**, focusing on their application in creating creative images.
- Conducted research on the **evaluation of text-to-image models** and introduced metrics to evaluate **creativity**, **persuasiveness**, and **alignment of images with implicit messages**.
- Conducting research on **complex and multi-step reasoning** applied to atypical and unusual images (ex. advertisement images). Proposed a zero-shot approach for reasoning on atypical images, improving the **semantical reasoning**. Introduced three novel complex reasoning tasks on unusual images.
- Conducted research on various aspects of **common sense question-answering**. Investigated evaluation methods, datasets, and cutting-edge approaches to advance the understanding and application of common sense in AI systems.

Deep Learning - Computer Vision Intern May 2022 – Aug. 2022
Cellanome *Palo Alto, CA*
- Initiated deep learning approaches for **object detection and medical image segmentation** and improved the accuracy of the image segmentation by 30%.
- Led the development of diverse methodologies and created a specialized dataset for medical image segmentation.
- Conducted groundbreaking research on transfer learning and semi-supervised learning, primarily focusing on their applications in medical image segmentation.
- Proposed a memory-efficient model for high-density instance segmentation, significantly advancing the company's capabilities in this domain.

Machine Learning Engineer, Intern Jul. 2017 – Sep. 2017
Tarafdari *Tehran, Iran*
- Developed clustering-based methods for tag optimization and removing redundancy. Streamlined data organization and improved user experience by retaining a single representative tag within each cluster.

Presentations

- *(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**
- *(Invited Talk)* Introduction to Few-shot learning on Medical Images, Explainable Deep Few-shot Learning on the Oracle Cloud and its Application in Medical Imaging Informatics tutorial, **ISVC'23**

Projects

Creative Image Generation from Abstract Messages Ongoing
PyTorch, Text-to-Image Models(T2I) - Stable Diffusion, SDXL, AuraFlow, PixArt, Image-Reward-, RLHF, RLAIF, LoRA

- Introduced a method for **creative and persuasive image generation** from implicit messages by fine-tuning T2I models with AI feedback (RLAIF).
- Introduced a contrastive learning-based loss to train the diffusion models to generate images based on abstract text.

Evaluation of Persuasiveness, Creativity, and Abstract Alignment in T2I Models Submitted
PyTorch, T2I models, RLAIF, LoRA, Abstract Reasoning, Large Language Models (LLMs), Multimodal LLMs(MLLMs)

- Introduced new evaluation metrics to evaluate the **persuasiveness** and **creativity** in T2I tasks utilizing different MLLMs. Improved the agreement with human annotations by 0.40 and 0.51 out of 1 for creativity and persuasiveness scores, respectively.
- Introduced a new evaluation metric for **abstract text and image alignment** utilizing MLLMs and fine-tuned LLMs with **Contrastive Preference Optimization**. Improved the agreement with human annotation for text-image alignment by 0.59 out of 1 compared to baseline metrics.
- Proposed a zero-shot method for creative and persuasive image generation from abstract text. The creativity, persuasiveness, and alignment scores improved by 25%, 18.5%, and 20%, respectively.

Complex Reasoning on Atypical Images WACV 2025
MLLMs, VLMs, LLMs, Compositional Reasoning, Chain-of-thought Reasoning

- Proposed an **atypicality-aware chain-of-thought verbalization for multi-step reasoning** on atypical advertisement images, improving the performance by 30%.
- Benchmarked three novel **complex reasoning tasks** on rhetorical content.
- Showed that LLMs given the verbalization of the image outperform MLLMs in complex reasoning tasks.
- Introduced an extended dataset based on the PittAd dataset to be **semantically challenging**.

Reasoning Capabilities of VLMs and LLMs
PyTorch, Transformers, VLMs, LLMs

- Designed an evaluation pipeline to compare the performance of VLMs (e.g., BLIP-2) and the corresponding LMs (e.g., FlanT5) in **complex reasoning** tasks such as Theory of Mind (ToM), Riddle Sense, and Social Interaction Question Answering, etc. Highlighted the superiority of LLMs' performance in complex reasoning tasks.
- Evaluated the **robustness** of VLMs and LLMs to the more complex forms and showed that VLMs are more robust than LLMs.

Re-ranking the answers of common sense question answering
Python, PyTorch, Ranking Evaluation, Answer Ranking

- Proposed a novel method for **re-ranking** the GPT-generated answers to the **common-sense questions** to have the more frequent responses in the forefront. Fine-tuned the ALBERT to choose between every two answers. Increased the **ranking score** by 13%, reducing the gap between the response and **oracle score**.

Exploring Domain Shift in Abstract Summarization
PyTorch, Transformers, Language Models (LMs)

- Designed and Developed various pipelines for abstract **summarization tasks** utilizing language models, such as BART and PEGASUS. Highlighted the drop in the performance of both models when evaluating the model on the unseen datasets.

Extra Curricular & Leadership

Member of Scientific Student Chapter Jan. 2017 – March 2018
Amirkabir University of Technology, Computer Engineering Department *Tehran, Iran*

- Organized over 70 national and international events, collaborated internationally with Technische Universität München, Germany, and KTH Royal Institute of Technology, Sweden.
- I was the head of "AUT DMC" executive team, the first Data Mining Contest at AUT.
- Our team was awarded the best organization of the year in 2018.