# Aysan Aghazadeh

📞 (+1) 412 916 2139    ✉ aysan.aghazadeh@gmail.com
⌗ aysanaghazadeh    in 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 Learning      Generative AI      Natural Language processing

## Publications

- *'Benchmarking VLMs' Reasoning About Persuasive Atypical Images'*, Under Review at WACV.
- *'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, Node.js, Vue.js, Express.js
**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*

- Conducting research on text-to-image generation models and their application in generating creative and persuasive images (ex. Advertisement images)
- Conducting extensive research on reasoning applied to advertisement images. Proposing a zero-shot approach for reasoning on advertisement images, improving the semantical reasoning ability of the model.
- 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 and pioneered deep learning approaches for object detection and medical image segmentation, driving innovation within the company.
- Led the development of diverse methodologies and curated a specialized dataset for medical image segmentation.
- Conducted groundbreaking research on transfer learning and semi-supervised learning, with a primary focus on their applications in medical image segmentation.
- Proposed a memory-efficient model for high-density object detection, 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

**Advertisement Image Generation** | *PyTorch, Text-to-Image Models(T2I), RLHF, LoRA*　　　　　　　Ongoing
- Generation of creative advertisement images needs common knowledge to understand the abstract message. It also needs creativity to visualize persuasive images. Due to these two reasons, current T2I models fail to generate persuasive advertisement images. We introduced a image generation method for advertisement image generation using step-by-step description generated by Fine-tuned LLMs with AI feedback (RLAIF)- Mistral, LLaMA3, and Phi3 as the input of T2I model-PixArt, SDXL, AuraFlow
- Introduced new evaluation metrics to evaluate the persuasiveness, creativity, and text-image alignment in T2I tasks utilizing different MLLMs like InternVL, and LLaVA.

**Reasoning on Advertisement Images** | *Python, Multimodal LLM(MLLM), VLMs, LLM, Multimedia*　　submitted
- Proposed a method exploiting MLLM (e.g., LLaVA, GPT4v), and LLMs (e.g., Vicuna, GPT4/3.5) to effectively verbalize and understand the underlying message of ad images in a zero-shot manner, resulting in significantly improving the performance by 30%.
- Rhetoric visual media (e.g., ad images) creators often use unusual portrayals of objects (atypicality) to convey a message to the users. Understanding atypicality in ad image requires complex reasoning and external knowledge. We benchmarked three novel tasks for understanding atypicality in ad images and evaluated the performance of MLLMs (e.g., LLaVA, GPT4-V, MiniGPT4, InstructBLIP), VLMs (CLIP), and LLM (e.g., GPT, Vicuna).
- Developed a novel method to extract atypicality from the ad image by incorporating effective prompting strategies, which then was exploited by LLMs for complex reasoning tasks, boosting the ad image understanding performance.

**Exploring reasoning capabilities of VLMs and LLMs** | *Vision-Language Models (VLM), Large Language Models (LLM)*
- Analyzed the performance of VLMs (e.g., BLIP-2) and LLMs (e.g., FlanT5) in complex reasoning tasks such as Theory of Mind (ToM), Riddle Sense, and Social Interaction Question Answering, etc. We showed the LLMs' performance on the same complex reasoning task is better overall.
- Evaluated the robustness of VLMs and LLMs to the more complex forms of the tasks by adding an irrelevant sentence including a wrong option to the question. We showed that VLMs' are more robust than LLMs.

**Re-ranking the answers of common sense question answering** | *Python, PyTorch, NLTK, ALBERT, Transformers*
- We proposed a novel method for re-ranking the GPT answers to the common sense question answers to have the more common answer in the forefront, using the ALBERT model as a multichoice question-answer to choose between every two answers. This resulted in an impressive increase of 12.29% in the ranking score.

**Exploring Domain Shift in Abstractive Summarization** | *Python, BART, PEGASUS, Transformers*
- Designed and Developed various pipelines for abstractive summarization tasks utilizing language models, such as BART and PEGASUS. Evaluated both models trained on one dataset on another dataset using both human evaluation and automatic metrics to show the domain shift in the summarization task. We show that both model have a drop in their performance when we evaluate them on another dataset.

**Lip Reading** | *Python, PyTorch, Transformers*
- Enhanced the accuracy of generated text in the lip reading task by incorporating sentiment features for the words with similar pronunciations.

**Multimodal Sentiment Analysis** | *Python*
- Enhanced the performance of the multimodal transformer in sentiment analysis by incorporating facial expression features alongside visual and audio features, resulting in a reduction in Mean Squared Error.

## Extra Curricular & Leadership

**Member of Scientific Student Chapter**　　　　　　　　　　　　　　　　　　Jan. 2017 – March 2018
*Amirkabir University of Technology, Computer Engineering Department*　　　　　　　　　　*Tehran, Iran*
- Me and my teammates held 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.