AmirArshia Hemmat

Computer Engineering Student

TVG, Oxford University

arshia@robots.ox.ac.uk ArshiaHemmat

Education

University of Isfahan *(UI)* *Isfahan, IRAN*

B.S. in Computer Engineering *Sept. 2019 - Jul. 2024*

- **GPA: 3.83 out of 4** or **17.94 out of 20**
- **GPA of last four semesters: 4 out of 4**
- Thesis topic: *Graph Fusion in Multi-modal Networks with Zero-shot Evaluation*
- Top 5% Academic ranking among +120 students

Research Interests

Multi-Modal LLM, Trustworthy AI, Generative AI, Zero-shot Learning

Publications

Hidden in Plain Sight: Evaluating Abstract Shape Recognition in Vision-Language Models *NeurIPS 2024*

Arshia Hemmat, Adam Davies, Tom A. Lamb, Jianhao Yuan, Philip Torr, Ashkan Khakzar, Francesco Pinto *Accepted*

- Generating a comprehensive dataset with Diffusion Models to audit failures of Multi-Modal LLMs
- Assessing the cutting-edge models like GPT-4o, Gemini, Llava and ...
- Conducting various experiments (Zero and Few shots and Human evaluation) on the VLMs
- Paper webpage

Survey on LLM in Requirement Engineering *Frontier Journal*

Arshia Hemmat, Mohammadreza Sharbaaf, Shekoufeh Kolahdouz-Rahimi, Kevin Lano, *Accepted*

- Paper webpage

MEENA (PersianMMMU): Multimodal-Multilingual Educational Exams for N-level Assessment *Under Review (COLM 2025)*

Omid Ghahroodi, Arshia Hemmat, Marzia Nouri, Seyed Mohammad Hadi Hosseini, Doratossadat Dastgheib, Mohammad Vali Sanian, Alireza Sahebi, Reihaneh Zohrabi, Mohammad Hossein Rohban, Ehsaneddin Asgari, Mahdieh Soleymani Baghshah *Submitted*

- introduced the first large-scale Persian benchmark for vision-language models evaluating scientific reasoning, problem-solving, and human-level understanding
- compiled 7,500 Persian and 3,000 English multimodal questions with rich metadata covering difficulty, descriptive answers, and student success rates
- conducted extensive experiments comparing GPT-4, Gemini, and other VLMs under zero-shot, few-shot, and hallucination detection settings

RAG-Driven Video QA with Adaptive Chunking: Leveraging a Bilingual Educational Dataset for the Educational Video QA *29th CSICC*

Arshia Hemmat, Mohammad Hassan Heydari, Kianoosh Vadaei, Melika Shirian, Afsaneh Fatemi *Accepted*

Leveraging Retrieval-Augmented Generation for University Knowledge Retrieval *15th IKT*

Arshia Hemmat, Kianoosh Vadaei, Mohammad Hassan Heydari, Afsaneh Fatemi *Accepted (Oral)*

- proposed a two-stage Retrieval-Augmented Generation (RAG) pipeline combined with Persian Large Language Models (LLMs)
- created a comprehensive benchmark dataset.
- Paper Link

Context Awareness Gate For Retrieval Augmented Generation *15th IKT*

Mohammad Hassan Heydari, Arshia Hemmat, Erfan Naman, Afsaneh Fatemi *Accepted*

• Context Awareness Gate (CAG): A mechanism that dynamically adjusts the input prompts of large language models (LLMs) based on the necessity of external context retrieval.
• Vector Candidates Method: A statistical, LLM-independent, and scalable approach that enhances semantic search.

Paper Link

CLIP Exhibits Improved Compositional Generalization Through Representation Disentanglement. *Arxiv*

Reza Abbasi, Amirarshia Hemmat, Mahdi Samiei, Mohammad Hossein Rohban, Mahdieh Soleymani Baghshah *Preprint*

• Investigation of CLIP's Compositional Generalization
• Impact of Training Data Diversity and Representation Disentanglement

Paper Link

Advanced Mutation Testing with Zero and Few-Shot Evaluation Using GPT-v4 *29th CSICC*

Arshia Hemmat, Fatame Aghababaei, Mohammadreza Sharbaf *Under Review*

Research Experiences and Academic Visiting

VIDEO-LLM UNCERTAINTY CALIBRATION AND TRAINING DATA EXTRACTION FROM DIFFUSION MODELS *Torr Vision Group - Oxford University*

Prof. Philip Torr *Sept 2024- Present*

• Evaluating the calibration of VLMs, and Generating a comprehensive MCQ in video
• Assessing the Memorization of Diffusion Models
• Will be submitted to NeurIPS 2025

VISION-LANGUAGE BENCHMARK IN EDUCATIONAL DOMAIN *RIML - Sharif University Technology*

Prof. Mohammad Hossein Rohban *Aug 2024- Present*

• Designing a comprehensive dataset of Different education levels from elementary to high school in the form of MCQs
• Categorizing based on different Cognitive tasks

VISION-LANGUAGE BENCHMARK DEVELOPMENT FOR SHAPE VS. SCENE REASONING *Torr Vision Group - Oxford University*

Prof. Philip Torr *March 2024- Aug 2024*

• **Published in NeurIPS 2024**

3D VESSEL SEGMENTATION USING GEOMETRIC APPROACHES *CAMP Chair - Technical University of Munich*

Prof. Nassir Navab *Sep 2023- Mar 2024*

• Utilizing Graph Neural Network and Fourier Transform to find the disconnectivity
• Working with 3D images and trying to preserve the continuity of vessels

COUNTERFACTUAL MODELING USING VISION-LINGUISTIC MODELS *CAMP chair - Technical University of Munich*

Prof. Nassir Navab *Aug 2023- Nov 2023*

• Visiting student at CAMP Chair
• Counterfactual Modeling in Medical Domain
• Utilizing cutting-edge multi-modal models like DRAGON and Geometric Deep Learning techniques specifically Graph modeling

ROBUSTNESS IN VISION-LANGUAGE MODELS *RIML - Sharif University of Technology*

Prof. Mohammad Hossein Rohban *Feb 2023 - Oct 2023*

• Disentanglement Analysis in contrastive models like CLIP and working on Robustness and OoD of different contrastive models

ZERO-SHOT LEARNING IN MEDICAL DOMAIN *ILS Lab- University of Isfahan*

Dr. Peyman Adibi *Aug 2021 - Nov 2022*

- Leveraging Geometric Learning (Manifold learning) in the Medical Domain (Chest analysis -ChestX-ray8 dataset)
- Utilizing the image-image modalities to bridge the gap between Seen and Unseen Classes

Honors & Awards

2023	**TUM Summer Research Fellowship**, Technical University of Munich	*Munich, Germany*
2023	**Top 5% Academic Ranking**, University of Isfahan	*Isfahan, Iran*
2022	**2nd place (Among +50 teams)**, National Data Science Competition	*Isfahan, Iran*
2020	**22nd in ICPC Asia Tehran**, ICPC West Asian Internet Online Programming Contest	*Tehran, Iran*
2020	**3rd place (Among +80 teams)**, Poytek business summer school	*Isfahan, Iran*
2015	**Math Kangaroo diploma**, Diploma of Honor in International Kangaroo Mathematics Competition	*Tehran, Iran*

Teaching Assistant

2023	**Deep Learning Course Graduate Course**, Dr. Emad FatemiZadeh	*SUT*
2023	**Machine Learning Course**, Dr. Peyman Adibi	*UI*
2023	**Knowledge Systems Course Graduate Course**, Dr. Reza Ramezani	*UI*
2023	**Medical Image Processing Course**, Dr. Emad FatemiZadeh	*SUT*
2023	**Advanced Programming Course**, Dr. Amin Fazli	*SUT*
2023	**Linear Algebra Course**, Dr. Peyman Adibi	*UI*
2023	**Artificial Intelligence Course**, Dr. Hossein Karshenas	*UI*
2022	**Engineering Probability and Statistics Course**, Dr. Abolfazl Motahari	*SUT*
2022	**Fundamental Programming Course**, Dr. Amin Fazli	*SUT*
2022	**Artificial Intelligence Course**, Dr. Hossein Karshenas	*UI*
2022	**Advanced Programming Course**, Dr. Behrouz Shahgholi	*UI*
2022	**Computer Architecture, (Two Times)**, Dr. Zohre Beiky	*UI*
2021	**Object Oriented Analysis and Design**, Dr. MohammadReza Sharrbaaf & Dr. Bahman Zamani	*UI*
2021	**Fundamentals of Robotics Course**, Dr. HamidReza Baraadran	*UI*
2021	**Data structures Course**, Dr. Reza Ramezani	*UI*
2021	**Advanced Programming Course, (Three Times)**, Dr. Reza Ramezani	*UI*
2020	**Fundamental Programming Course**, Dr. Hossein Mahvash Mohammadi	*UI*
2020	**Machine Learning**, Amin Mousavi	*Taarlab (UT)*

Associations

Quera *Tehran*

Data Team course instructor *Jun. 2024 - Oct. 2024*

- Designing courses for the Data team, including those on LLM, Deep Learning, and NLP where's one of the biggest and most famous Computer Science learning.

BigData Lab *UI*

Board Member *2023 - Present*

- Research on Retrieval Augmented Generation (RAG) and submitted **three papers** under supervision of Dr. Afsaneh Fatemi

Others *SUT - UI*

Chair Man, Organizer, Co-organizer, Advisor *2020 - PRESENT*

- Founder and Chairman of Community for Artificial Intelligence at University of Isfahan
- Chairman and development head of Rasta association at Sharif University of Technology
- Organizer of 5th, 3rd and Co-organizer of 2nd **TEDx** University of Isfahan

Language

Native **Persian**,
Advanced **English**, TOEFL Score: 93 (Best score: 96)

References

1. **Prof. Philip Torr**

Full Professor, Oxford University

Department of Engineering Science, email: philip.torr@eng.ox.ac.uk, phone: 01865 283059

2. **Prof. Nassir Navab**

Full Professor, Technical University of Munich

Chair of Computer Aided Medical Procedures, email: navab@cs.tum.edu, phone: +49 (89) 289 - 17057

3. **Dr. Mohammad Hossein Rohban**

Associate Professor, Sharif University of Technology

Department of Computer Engineering, email: Rohban@sharif.edu, phone: +98-21-66165786