Melika Shirian

Email | Google Scholar | GitHub | LinkedIn

Personal Profile

I am a B.S student in Computer Engineering at the University of Isfahan, driven by a strong interest in artificial intelligence and machine learning, especially self-supervised representation learning. I seek to build my expertise to help improve the effectiveness and reliability of AI technologies, particularly in educational and medical domains.

Education

University of Isfahan **Sep. 2021 – Present**

B.S. in Computer Engineering

- **Overall GPA:** 3.77/4 (18.17/20)
- **GPA of last four semesters:** 3.91/4
- **Thesis topic:** A Unified Self-Supervised Learning Toolbox (20/20)
- Top 5% academic ranking among 120+ students

Relevant A* Courses: Fundamentals of Computer Science and Programming (20/20), Advanced Programming (20/20), Data Structures (20/20), Fundamentals of Machine Learning (20/20), Fundamentals of Computational Intelligence (20/20), Fundamentals of Data Mining (20/20), Special Topics – Deep Learning (20/20), IoT Device Programming (20/20), Algorithm Design (18.75/20 (highest score among students))

Research Interest

Multimodal and Self-Supervised Representation Learning — Video reasoning and 3D Visual Understanding — SciML — Robust AI — Graph Representation Learning — Privacy-Preserving ML — Low-Resource Language Processing — Geometric Deep Learning

Publications

From Scenes to Semantics: PersianCLEVR for Bilingual 3D Visual Reasoning NeurIPS 2025

K. Vadaei, **M. Shirian (Co-First Author)**, A. Hemmat, MH. Heydari, A. Mamanpoosh, A. Fatemi

Status: Published (Spotlight), VLM4RWD Workshop.

Adaptive Chunking for VideoRAG Pipelines with a Newly Gathered Bilingual Educational Dataset CSICC 2025

M. Shirian (Co-First Author), A. Hemmat, K. Vadaei, MH. Heydari, A. Fatemi

Status: Published (Oral).

When 1+1 $\neq$ 2: Benchmarking Civil-Time Reasoning in Large Language Models Under review (ACL 2026)

K. Vadaei, **M. Shirian (Co-First Author)**, MH. Shaker Ardakani, A. Mamanpoosh, M. Moqadas, A. Rismanchian, A. Fatemi, A. Hemmat

PrismSSL: One Interface, Many Modalities; A Single-Interface Library for Multimodal Self-Supervised Learning Preprint

M. Shirian (Co-First Author), K. Vadaei, K. Majlessi, A. Ebrahimi, A. Hemmat, P. Adibi, H. Karshenas

Status: Preprint (arXiv:2511.17776).

Research Experience

Adaptive Chunking for VideoRAG with a Gathered Dataset | *University of Isfahan* **Sep.2024 - Jan.2025**

- The project introduces content-aware video segmentation for efficient retrieval and question answering over slide-based educational videos without retraining large models, with my segmentation idea playing a key role.
- I designed and worked on implementing and testing slide-change detection based on similarity drops between sampled frames, using a mix of pixel-level metrics (such as SSIM) and embedding-based representations (such as CLIP).
- I helped select representative slides using entropy-based measures to improve the quality of segmented video chunks.
- I also contributed to building the benchmark by identifying suitable videos. The benchmark is continuously expanding in size and is being extended to support multiple languages.

MK_SSL Unified Self-Supervised Toolbox | *ILS Lab - University of Isfahan* [Link] **May.2025 - Sep.2025**

- I completed my bachelor's thesis as a joint research project under two faculty supervisors on a unified self-supervised learning framework across multiple modalities, later co-released as an open-source Python package.

- I studied widely used SSL models and implementing them from scratch in PyTorch, covering contrastive and reconstruction-based encoder–decoder methods, audio SSL on waveform and spectrogram inputs, multimodal representation learning, and low-resource fine-tuning using LoRA.

Function Calling with Large Language Models for Home Assistants | *research intern Hoopad* **Summer 2024**

- The project aimed to optimize home assistants for local execution using offline, lightweight models on embedded systems.
- I evaluated offline STT/TTS models like Vosk for on-device inference. I also worked with prompt engineering and function calling methods to improve response reliability.

PersianCLEVR benchmark for Bilingual 3D Visual Reasoning | *University of Isfahan* **Feb.2025 - Nov.2025**

- I contributed to developing a bilingual 3D visual reasoning benchmark to evaluate vision–language models on compositional reasoning tasks and identify their limitations.
- My focus was mainly on dataset preprocessing, selecting suitable models, and defining appropriate evaluation strategies, such as lexical metrics, and LLM-as-judge scoring protocols.
- I gained hands-on experience evaluating multimodal models in a bilingual setting, analyzing how language and 3D spatial variations affect reasoning.

Benchmarking Civil-Time Reasoning in LLMs | *University of Isfahan* **Nov.2025 - Jan.2026**

- The project aim was to introduce a benchmark to evaluate large language models on precise civil-time reasoning such as DST changes, non-hour offsets, and date-line crossings.
- I mainly contributed in fine-tuning and evaluating open-source language models with token-level training strategies to study civil-time reasoning.

Music-Image Retrieval | *University of Isfahan* **Sep.2025 - Oct.2025**

- I used wav2vec 2.0–style for music understanding by redesigning the encoder to capture musical structure beyond speech.
- I developed a multimodal music–image alignment approach by integrating music representations with visual encoders (e.g., BLIP) to model semantic and contextual correspondence between audio and image.

Academic Service

Teaching Assistant

I have volunteered as a teaching assistant for more than 15 course sections, supporting students through one-on-one guidance, designing projects and assignments, and contributing to exam preparation and grading. I have also independently taught classes in place of the instructor, developing strong teaching and communication skills. I was a TA in **Fundamentals of Computer Science and Programming** (3 sections), **Advanced Programming** (3 sections), **Data Structures** (1 section), **Computer Architecture** (3 sections), **Algorithm Design** (1 section), **IoT Device Programming** (1 section), **Computer skills workshop** (3 sections), **Microprocessors and Assembly Language** (1 section), and **Applications of Artificial Intelligence in Cybersecurity** (1 section).

Peer Review

I was a reviewer in VLM4RWD Workshop (NeurIPS), evaluated 2 submissions via OpenReview, focusing on technical quality, experimental rigor, and relevance to vision–language and multimodal reasoning.

Association

- **TEDx (University of Isfahan)** – I Collaborated as a speaker coach in 2024 and I was on the executive team in 2025, gaining hands-on experience in teamwork, communication, and event organization.
- **RASTA Association (Sharif University of Technology)** – Cryptography mentor at the 2024 summer school, teaching and working directly with students.
- **7th International Conference on Internet of Things and Its Applications (2023)** – Executive team member and workshop technical coordinator.

Skills

- **Technical Languages:** Python, Java, C, C++, LaTeX
- **Libraries and Frameworks:** PyTorch, Transformers, Pandas, Scikit-learn, NumPy, Shell (Linux)
- **Tools:** Git, Weights and Biases (W&B)
- **Languages:** IELTS 7.5
- **Other Skills:** Waterpolo notional champion, lifeGuard, Kangaroo math certificate