

Mahdi Mohseni

📍 United Kingdom ✉ mohseni-m@ulster.ac.uk ☎ +44 7353 799267 in LinkedIn 📄 GitHub

Professional Summary

Passionate and self-taught AI and Python developer with over **four** years of industry experience. Specializing in **Foundation models**, **Generative AI**, and **Multi-modal models**, with a strong track record of contributing to open-source projects. Having robust coding skills, excellent communication skills, and a commitment to advance the field of AI through meaningful academic and research contributions. Looking to leverage expertise in a Ph.D. program to contribute to the academic community and drive innovation in artificial intelligence.

Education

- | | | |
|------------|---|---------------------|
| PhD | University of Ulster , United Kingdom | Sep 2025 – Present |
| | <ul style="list-style-type: none"> • Field: Medical Artificial Intelligence & Clinical Data Systems • Thesis: Designing Agentic AI Architecture for Multi-modal Medical Data Integration and Clinical Decision Support • Supervisor: Dr. Steven Watterson, Medical AI Research Group • Research Area: UK National Health Service (NHS) digital health infrastructure and federated learning for healthcare data | |
| MS | Shahid Beheshti University , Telecommunication Engineering | Oct 2020 – Oct 2023 |
| | <ul style="list-style-type: none"> • GPA: 3.65/4.0 (excluding one course due to absence) • Thesis: Enhancing 3D Object Detection in Autonomous Vehicles using Sensor Fusion • Achievements: Achieved an 11% improvement in inference speed while maintaining state-of-the-art model accuracy • Advisor: Dr. Esfandiar Mehrshahi | |
| BS | Isfahan University , Electrical Engineering | Oct 2016 – Oct 2020 |
| | <ul style="list-style-type: none"> • GPA: 3.33/4.0 • Thesis: Evaluating the Performance of Classical Machine Learning and Deep-based Approaches in Handwriting Recognition • Advisor: Dr. Mohammad Farzan Sabahi | |

Experience

- | | |
|---|--|
| Zebracat , AI Research Engineer | Berlin, Germany
Dec 2023 – Sep 2025 |
| <ul style="list-style-type: none"> • Working remotely with an international Generative AI company in Video Generation context • Lead AI research projects focusing on deep learning and computer vision applications. • Develop and implement advanced algorithms for image and video processing. | |
| Revivoto / AlhomeDesign , AI Research Engineer | Vancouver, CA
Jul 2022 – Dec 2023 |
| <ul style="list-style-type: none"> • Working remotely with an international Generative AI company in Image Editing in real estate agent photography industry • Contributed to the development of virtual staging and image enhancement tools. • Improved model performance and optimized inference speeds. | |
| Freelancer , ML Engineer | Isfahan, Iran Jul 2019 – Jul 2022 |
| <ul style="list-style-type: none"> • Developed custom models and provided AI consulting services. | |

- Completed various AI and machine learning projects for diverse clients.

Research Interests

- Agentic-AI
- Deep Learning
- Computer Vision
- 3D Computer Vision
- Generative AI
- Foundation Models
- Diffusion-based Models
- Multimodal Deep Learning
- Natural Language Processing
- Large Language Models
- Reinforcement Learning
- Robotics
- Extended Reality
- Neuroscience
- Medical Image

Technical Skills

Deep Learning Frameworks

- PyTorch, TensorFlow/Keras, PyTorch-Lightning

Python Libraries

- OpenCV, OpenMM, Lavis, HuggingFace (Diffusers, Transformers, Timm, Datasets), Pandas, Scikit-Learn, Shapely, Seaborn, FastAPI, Scrapy, Streamlit, Xformers, FlashAttention, BitsandBytes

Machine Learning

- Linear Regression, SVM, K-NN, Random Forests

Image Processing Algorithms

- Image Filtering, Histogram Equalization, Edge Detection (Canny, Sobel, LoG), Line Detection (Hough Transform), Morphological Operations

Computer Vision

- Image Classification, Multi-label Classification, Segmentation (Instance, Semantic), Object Detection, Corner Detection, GAN, U-Net, Attention Mechanisms, Vision Transformers, Depth Estimation, Camera Parameter Estimation, Lens Distortion Correction, Style Transfer, Image Editing

3D Computer Vision

- Vanishing Point Detection, 3D Layout Estimation, Panorama Image Layout Estimation, 3D Object Detection (Single Image, Sensor Fusion)

Generative AI (Text-to-Image Models)

- Image Generation, Conditional Image Generation, Image Editing, Stable Diffusion, ControlNet, Instruct-pix2pix, Inpainting, Prompt Tuning

Generative AI (Video Generation)

- Video Generation, Infinite Zoom, Infinite Rotation, Video Upscaling

Natural Language Processing (NLP) / Large Language Models (LLM)

- Text Classification, Text Summarization, Prompt Engineering, Text Watermarking, Image Information Retrieval using LLMs, Chatbots, PEFT, LoRA Fine-Tuning, RLHF, RLAI, PPO

Multi-Modal Models

- Visual Question Answering, Image Captioning, Image Tagging, Pix2Seq, BLIP, CLIP, Flamingo, RAM, Text2Tag, VirConvNet

Foundation Models

- SAM and HQ-SAM, CLIP, Grounding-Dino, Grounded-SAM, RAM, Painter, SegGPT, SEEM, VLPART, Kandinsky, PixArt-alpha, Tiny-LLaMA, Phi

MLOps

- TorchServe, Active Learning

DevOps

- Linux, Windows, Docker, Docker-Compose, Git, GitHub, GitLab, Bash Scripting

Databases

- MongoDB, MinIO

Management Systems

- Agile, Scrum, Trello, Jira, Kanban

Soft Skills

- Fast Learner, Motivated, Team Player, Creative, Reliable, Eager to Learn Cutting-Edge Technologies

Projects

CIFAR-10 Classification (Custom Model and Scratch Training)

- Developed a complete training pipeline using PyTorch, including data preparation, custom ResNet model creation, and training loop development.
- Implemented optimization techniques such as SGD, AdamW, AdaBelief, cosine annealing, mixed precision, weight decay, and gradient clipping.
- Created and deployed an API using FastAPI.

Binary Classifier (Transfer Learning from DINO-V2)

- Trained over 50 binary classifiers to provide image tag labels through a general foundation model tagger filter (RAM).
- Developed an Image Tagger using a DINO-V2 backbone, currently utilized in the real estate industry.

Day to Night Transformation and Vice Versa using IEGAN

- Fine-tuned an image-to-image translator with IEGAN architecture to successfully transform images between day and night.

Scene Classification for Real Estate Agent Photos

- Developed a few-shot learning model using Vision Transformer (ViT) architecture.
- Distilled knowledge from the large ViT model to a GPUNet architecture, reducing parameters while maintaining high accuracy.

Panorama Image Layout Estimation

- Utilized and fine-tuned the LED2Net model integrated with PanoAnnotator tools.
- Created an API for 360 image layout estimation and converted the PyQt5 program to an executable file.

Camera Lens Distortion Correction

- Implemented both deep-based (K-matrix estimation) and classical (checkerboard) approaches for camera lens distortion correction.

Panorama Image Inpainting

- Developed an image inpainting method using equirectangular to cuboid perspective image transformer.
- Applied various inpainting techniques including LaMa, ZITS, and Stable Diffusion variants.

Synthetic Image Dataset Generation

- Created synthetic image datasets using Stable Diffusion for various computer vision tasks, focusing on in-context learning and prompt engineering.

Virtual Staging Project

- **Camera Intrinsic and Extrinsic Parameter Estimation:** Developed methods using classic and deep-based approaches for indoor scenes.
- **Indoor Scene Corner Detection:** Created an LSTM-CNN model and utilized the Pix2Seq architecture.
- **Light Source Detection:** Implemented YOLOv7 model and led a labeling team.
- **Areal Light Source Segmentation:** Fine-tuned models using MMsegmentation platform with Swin and Upernet models.
- **3D Reconstruction from Single Image:** Developed methods using various techniques including GeoLayout and PlaneRCNN.
- **Fully Automated Virtual Staging:** Utilized Stable Diffusion to furnish indoor scenes automatically (Launched in Production).
- **Object Removal:** Implemented Stable Diffusion inpainting model for object removal in indoor scenes (Launched in Production).
- **Image Enhancement:** Fine-tuned a Stable Diffusion fork to provide an image enhancement model (Launched in Production).
- **Day to Dusk Transformation:** Fine-tuned a Stable Diffusion fork to convert outdoor scenes from day to dusk (Launched in Production).
- **Interior Design Automation:** Fine-tuned a Stable Diffusion fork to fully automate indoor scene design changes (Launched in Production).

Additional Projects

- **3D Object Detection Using Sensor Fusion in Self-Driving Cars:** Improved VirConvNet, achieving first rank in KITTI benchmark and an 11% increase in inference speed.
- **Blood Pressure Estimation Using PPG Signal:** Trained a custom ResNet-1D model to estimate systolic and diastolic blood pressure from PPG signals using the UCI dataset.
- **Reimagine Service Implementation:** Developed an end-to-end service similar to Reimagine independently.
- **Watermark Removal and Segmentation:** Created suitable datasets and implemented segmentation and in-painting approaches for watermark removal.
- **Story to Video Generation:** Developed a pipeline to generate videos from user-provided ideas and styles using LLMs, text-to-image, image-to-video, and text-to-speech models.
- **Infinite Zoom Video Generation:** Modified inpainting models to out-painting models for generating infinite zoom videos based on user ideas and styles.
- **Lip-sync and Avatar Generation:** Developed models for lip-syncing and avatar creation.

Language Proficiency

IELTS

- **Overall:** 7
- **Speaking:** 6.5
- **Listening:** 8
- **Writing:** 6
- **Reading:** 6.5

References

Dr. Esfandiar Mehrshahi

- **Position:** Associate Professor, Shahid Beheshti University
- **Email:** mehr@sbu.ac.ir 

Dr. Samaneh Shojaei

- **Position:** Assistant Professor, IROST
- **Email:** s.shojaie@irost.ir 

Note: Many projects are not publicly accessible due to company policies. However, results are available at [AI Home Design](#)  and [Zebracat](#) . For verification, please contact the site owner via email.