

Mahsa (Zahra) Anvarian

📍 Calgary, AB, Canada

📞 (+1) 825-733-4066 • ✉️ mahsawz@gmail.com • 🌐 mahsawz.github.io
in linkedin.com/in/mahsa-anvarian • 🏠 github.com/mahsawz • 🎓 Google Scholar

SUMMARY

Machine learning researcher and engineer with 4+ years building and evaluating deep learning systems across computer vision, geospatial ML, and NLP, spanning industry ML engineering and academic research. Four publications, three as first or co-first author. Completing an M.Sc. in Computer Science at the University of Calgary, with a focus on the parts of ML that decide whether results hold up outside the lab: how datasets are built, how evaluation is designed, and whether models generalize to data they have not seen.

SKILLS

Programming Languages: Python, SQL, JavaScript, Bash

Machine Learning & Deep Learning: PyTorch, TensorFlow, Keras, Scikit-learn, HuggingFace Transformers, OpenCV, NLTK, Hazm

Data Analysis & Visualization: Pandas, NumPy, SciPy, Matplotlib, Seaborn, Tableau, Power BI

Geospatial & Scientific Computing: Rasterio, Polyscope, SRTM/DEM Processing

Tools & Platforms: Git, AWS, PostgreSQL, Linux, Jupyter, Google Colab, Jira, Selenium, $\LaTeX$

Web Development: HTML, CSS, JavaScript, React

Languages: Persian (Native); English (Full Professional Proficiency)

EDUCATION

M.Sc. in Computer Science

University of Calgary • CGPA: 3.93/4

Calgary, Canada

Jan 2024 - Expected Jan 2027

B.Sc. in Computer Engineering

Iran University of Science and Technology • Ranked 4th nationally (QS)

Tehran, Iran

Sep 2016 - Jun 2021

EXPERIENCE

Computer Vision Researcher

[Vision Research Lab](#), University of Calgary

Calgary, AB, Canada

May 2024 - Present

- Demonstrated that random tile-based splitting understated DEM RMSE by 10% through spatial data leakage, establishing region-based evaluation as a more reliable protocol for assessing real-world generalizability.
- Designed frequency-aware loss functions combining wavelet, terrain, and geometric cues (slope, aspect, TRI, TPI), improving fine-detail reconstruction by 6% on TRI RMSE.
- Collected, built, and cleaned spatially disjoint, terrain-diverse datasets of $\sim 26K$ paired 120 m/30 m SRTM DEM tiles, spanning mountainous, moderate, and low-relief regions; using **Python**, **Rasterio**, and **NumPy**.
- Adapted, trained, and benchmarked SOTA super-resolution architectures in **PyTorch** under controlled experimental protocols.
- Fully Developed the full research pipeline end-to-end; geospatial acquisition, preprocessing, patch generation, training, evaluation, and reproducible experiment management; and created a **Polyscope**-based Python package for interactive 3D inspection of DEM tiles.

Teaching Assistant

University of Calgary

Calgary, AB, Canada

May 2024 - Apr 2026

Courses: [DATA 201](#) | [DATA 601](#) | [DATA 608](#) | [CPSC 433](#)

Excellent TA Award, 2025/2026

- Taught tutorials and labs for four data science and AI courses over four terms, reaching 15+ students from

first-year undergraduates to graduate level; distilled deep learning, model evaluation, and distributed data processing into material each cohort could act on.

- Mentored 10+ student teams through semester-long course projects end-to-end; scoping, big-data processing pipelines, and cloud deployment on **AWS**; running weekly checkpoints and unblocking design, debugging, and scalability issues from proposal through final delivery.
- Ran office hours supporting **Python**, **TensorFlow**, **Pandas**, **Tableau**, and **Power BI** workflows, diagnosing broken pipelines and overfitting issues under time pressure, and partnered with faculty to keep tutorial content aligned with lectures.

Machine Learning Engineer

Dadmatach Company

Tehran, Iran

Sep 2020 - Oct 2022

- Built the first large-scale Persian dataset for MBTI personality-trait prediction, collecting and processing 1.55M+ tweets from 938 users using **Python** and the Twitter API, later expanded to 6.1M+ tweets from 3,876 users.
- Developed data pipelines for user validation, tweet collection, Persian-language filtering, anonymization, preprocessing, and class balancing using **Pandas**, **NumPy**, **NLTK**, and **Hazm**.
- Fine-tuned **ParsBERT** using **HuggingFace Transformers** and developed a combined fastText-embedding, BiLSTM, and attention-based model across four MBTI classification tasks in **TensorFlow/Keras**, benchmarked against BERT, Doc2Vec, CNN, and LSTM baselines.
- Achieved 66.71% average accuracy across the four MBTI tasks, outperforming the majority-class baseline by approximately 13 points on P/J and 7 points on T/F.
- Applied **LIME** to interpret model predictions and surface the influential Persian words and user-level linguistic patterns driving each trait, turning a black-box classifier into an auditable one.

SELECTED PROJECTS

License Plate Detection

 github.com/mahsawz/License-Plate-Detection

15-configuration ablation over architecture (Xception, ResNet50, MobileNet), optimizer, and augmentation in **Keras** for three-class license-plate **classification** on an imbalanced 2,727-image dataset; HSV lighting **augmentation** reached 99.64% accuracy (0.99 macro-F1), while synthesized minority-class data degraded every architecture.

PUBLICATIONS

- [1] **Z. Anvarian** and F. Maleki, “Methodological Factors Affecting the Generalizability of Deep Learning-Based Digital Elevation Model Super-Resolution,” in *Earth Precision Agriculture Conference (EPA)*, 2026.
- [2] P. Taghipour*, A. Khodadadi*, F. Rezaei†, **Z. Anvarian**†, R. C. Sotero, and S. Madadi, “Beyond Mean Pooling: Statistical Meta-Aggregation for Subject-Level EEG Classification in Parkinson’s Disease,” in *The 10th International Conference on Biomedical Engineering and Bioinformatics (ICBEB)*, 2026.
- [3] S. Fatehi*, **Z. Anvarian***, Y. Madani, M. J. Mahditabar, and S. Eetemadi, “MBTI Personality Prediction Approach on Persian Twitter,” in *The 6th WiNLP Workshop at Empirical Methods in Natural Language Processing (EMNLP)*, 2022.
- [4] M. M. Abdollahpour*, **Z. Anvarian***, S. Fatehi, and S. Eetemadi, “ParsTSet: A Persian Dataset for Personality Prediction on Twitter,” in *The Fourth Annual West Coast NLP (WeCNLP) Summit*, 2021.

* / † denote equal contribution among authors sharing the same symbol within an entry.

CERTIFICATIONS

DeepLearning.AI

All Grade Achieved: 100%

- Deep Learning Specialization [\[Certificate\]](#) Sep 2020
- Convolutional Neural Networks in TensorFlow [\[Certificate\]](#) Oct 2020
- Introduction to TensorFlow for AI, Machine Learning, and Deep Learning [\[Certificate\]](#) Oct 2020
- Natural Language Processing in TensorFlow [\[Certificate\]](#) Oct 2020