Gheffari Youcef Soufiane

+213552598134 | gheffari.youcef.soufiane@gmail.com | linkedin.com/in/gheffari-youcef-soufiane | github.com/youcefgheffari

Summary

Postgraduate student specializing in Artificial Intelligence, with practical experience in deep learning, computer vision, and natural language processing.
Proficient in designing, developing, and deploying machine learning systems using Python, PyTorch, and TensorFlow.
Demonstrates strong analytical and problem-solving abilities, a research-oriented mindset, and a proven capacity for collaboration in multidisciplinary environments.
Currently seeking challenging opportunities to contribute to innovative research and the development of impactful AI-driven solutions.

Technical Proficiencies

Programming Languages: Python, SQL, JavaScript, C

Development Tools & Platforms: Git, Jupyter, Google Colab, VS Code, PyCharm, Linux (Ubuntu/Debian)

Software Engineering: Object-Oriented Programming, Git Version Control, Debugging

AI and Machine Learning Toolkits

Deep Learning Frameworks: PyTorch, TensorFlow, Keras

Machine Learning Libraries: Scikit-learn, NumPy, Pandas, Seaborn, Matplotlib

Computer Vision Tools: OpenCV, YOLOv8, Panda3D

Natural Language Processing (NLP): NLTK, Hugging Face Transformers, Wav2Vec2, LangChain

3D and Visual Computing: Gaussian Splatting, Open3D (basic), Panda3D

Experiment Tracking & Visualization: Weights & Biases (W&B)

Education

University Center of Naama — Naama, Algeria
Bachelor of Science in Computer Science — *Oct 2020 – May 2023*
Core focus on algorithms, data structures, software engineering, and mathematical foundations.

Completed projects in systems programming, databases, and web development.

University of Science and Technology of Oran (USTOMB) — Oran, Algeria
Master of Science in Artificial Intelligence and Its Applications — *Oct 2023 – May 2025*
Specialized in deep learning, computer vision, and natural language processing.

Completed advanced coursework in neural networks, reinforcement learning, and probabilistic models.

Thesis: Arabic Speech Emotion Recognition System (Wav2Vec2, CNN, Transformer, LSTM).

Master's Thesis Project

Arabic Speech Emotion Recognition System
Master's Thesis — MSc in Artificial Intelligence and its Applications — *Oct 2023 – May 2025*

Developed a deep learning system to classify emotions in Arabic speech using a combination of acoustic and contextual features.

Utilized the **Egyptian Arabic Speech Emotion (EYASE) database**, comprising labeled emotional utterances in Modern Standard Arabic.

Preprocessed raw audio and extracted features using the pre-trained **Wav2Vec2** model.

Implemented and evaluated multiple architectures:

* CNN-based classifier: **Accuracy: 75.2%**
* CNN- Bidirectional-LSTM with attention model: **Accuracy: 84.6%**
* CNN-Transformer : **Accuracy: 93.4%**

Conducted comparative analysis and tuning to optimize F1-score and inference speed across models.

Results indicated the Transformer-based model outperformed CNN and LSTM in both accuracy and generalization.

Engineering Projects

Brain Tumor Classification System

Deep Learning Project *Python, PyTorch, Scikit-learn, CNN, Autoencoder*

Designed and implemented a pipeline for classifying brain tumors using MRI scans.

Applied CNNs and autoencoders for feature extraction and tumor classification.

Combined traditional descriptors (HOG, LBP, SIFT, ORB) with deep learning features.

Used PCA to reduce dimensionality and improve classification performance.

Achieved high accuracy in distinguishing glioma, meningioma, pituitary tumors, and healthy images.

Group Activity Recognition

Deep Learning Project *Python, PyTorch, Scikit-learn, CNN, LSTM*

Designed and implemented a hierarchical deep learning pipeline for recognizing group activities in videos.

Applied CNNs for person-level feature extraction and LSTMs for modeling group-level temporal dynamics.

Trained and evaluated on the Collective Activity Dataset (Crossing, Waiting, Queueing, Walking, Talking).

Achieved accurate group activity classification through a two-stage CNN+LSTM architecture.

Multi-Object Tracking + 3D Reconstruction in Driving Scenes

Advanced Computer Vision Project *Python, PyTorch, YOLOv8, DeepSORT, TUM RGB-D, TensoRF*

Built a full object-centric 3D reconstruction pipeline combining detection, tracking, pose fusion, and volume optimization.

Used **YOLOv8 (COCO)** for object detection and **DeepSORT** for real-time multi-object tracking.

Cropped tracked object images and associated them with **groundtruth camera poses** from the TUM freiburg1_desk dataset.

Trained a dummy **TensoRF-style 3D tensor volume** per object using image supervision and pose alignment.

Visualized learned volumetric slices and exported intermediate `tensor.pt` files for downstream 4D reconstruction.

Research Projects

DrivingRecon 4D Scene Reconstruction

4D Reconstruction and Neural Rendering *Python, PyTorch, OpenCV, Panda3D*

Implemented a simplified version of the DrivingRecon paper using multi-view driving video.

Built a 4D Gaussian rendering pipeline with depth estimation, 3D encoding, and temporal attention.

Achieved real-time scene reconstruction and novel view synthesis from synthetic or real data.

Semi-Supervised Semantic Segmentation with Self-Correction (CVPR 2020 Reproduction)

Computer Vision Paper Reproduction *Python, PyTorch, Scikit-learn, CNN*

Re-implemented and trained dual-network segmentation system using PyTorch.

Integrated ancillary and primary segmentation networks with self-correction (linear & CNN-based).

Achieved segmentation performance competitive with supervised baselines.

Applied on Pascal VOC 2012 dataset using weak and full annotations.

LLM-based Visual Grounding in 3D Driving Scenes

AI Research Project — Self-Initiated *Python, PyTorch, Detectron2, GPT-4, Panda3D*

Built a complete 3D scene understanding pipeline using RGB-D frames from the TUM dataset, leveraging object detection, depth backprojection, and visual reasoning.

Constructed dynamic scene graphs capturing spatial relationships and queried them using GPT-4 to identify objects based on free-form natural language input.

Mapped GPT-identified objects back to 3D point clouds and visualized them interactively in real-time using Panda3D.

Exported point clouds and scene graphs for demo-ready `.ply` and `.json` formats, enabling portfolio visualization and reproducibility.

Achieved real-time 3D semantic understanding and language-based scene exploration in autonomous driving contexts.

Certifications

Machine Learning Process A-Z, 365 Data Science

Machine Learning Algorithms A-Z, 365 Data Science

Machine Learning Specialization, Coursera — Andrew Ng (DeepLearning.AI)

Data Science & Machine Learning Foundations, Maven Analytics

Languages

Arabic — Native

French — Fluent

English — Professional Working Proficiency

Core Research Domains

Computer Vision and 3D Perception: Representation learning from images, videos, and depth data; neural rendering; geometric deep learning; object detection, segmentation, and 3D scene reconstruction.

Multimodal Learning: Joint modeling of vision, language, and audio to enhance cross-modal understanding, robustness, and generalization in complex environments.

Natural Language Processing (NLP): Focus on semantic understanding, contextual embeddings, and Arabic NLP in low-resource settings.

Speech Processing: Emotion recognition and automatic speech recognition (ASR), integrated as supporting modalities in multimodal systems.

Applied AI Focus Areas

Vision-Language Models (VLMs): Leveraging models like CLIP, GPT-4V, and Flamingo for visual grounding, image captioning, referring expression comprehension, and multimodal reasoning.

Explainable AI (XAI): Designing interpretable and trustworthy AI systems for sensitive domains such as healthcare, autonomous driving, and safety-critical applications.

Real-Time Perception Systems: Efficient CV + NLP pipelines optimized for deployment in autonomous systems, including object-centric 3D tracking and multimodal scene understanding.

Research Vision

Advance the development of **multimodal foundation models** capable of reasoning jointly over vision, language, and audio with minimal supervision and enhanced generalization.

Explore **4D dynamic scene understanding** and spatiotemporal perception for real-time applications in autonomous driving, robotics, and AR/VR.

Design **interpretable and reliable AI systems** tailored for high-stakes domains such as healthcare, autonomous systems, and AI-assisted decision making.

Bridge the gap between **vision-language understanding and embodied AI**, enabling agents to perceive, communicate, and act intelligently in real-world 3D environments.

Promote inclusive AI through research in **low-resource languages** and culturally diverse multimodal datasets, with a focus on Arabic NLP and multilingual scene understanding.