

# Omar Hassan

github.com/omar-A-hassan

Email : es-omar.hagag2027@alexu.edu.eg

LinkedIn : linkedin.com/in/omar-hassan

## SKILLS

- **ML/AI:** PyTorch, TensorFlow, Scikit-learn, AutoML (H2O), geomloss, PEFT, SFT, RL, ONNX, CTranslate2
- **Languages & Tools:** Python, REST/FastAPI, Gradio, Pydantic, Git, DVC, LFS, Docker, Conda

## EDUCATION

- **Alexandria university** Alexandria, EG  
*Bachelor of Engineering in Biomedical Engineering; GPA: 3.13* Aug. 2022 – Dec. 2027

## EXPERIENCE

- **SmartCI center FOE** Alexandria, EG  
*TinyML research intern* Oct 2025 - Present
  - **Caption-LSTM:** Researching the use of Lstms in Vision tasks such as image captioning and deployment on resource constrained devices as a lightweight alternative to Vision transformers.
- **Made in Alexandria (M.I.A) Robotics** Alexandria, EG  
*AI Engineer* Sep 2025 - Present
  - **Benchmark Nllb200:** Added benchmarking support of ARZ Egyptian Arabic text translation to English and vice versa to (LM Evaluation Harness) with Meta's Nllb200 measuring BLEU, CHRF and BERT scores.
  - **Benchmark Whisper:** Expanded upon an already established Speech to Text models evaluation and benchmarking framework (Pico-Voice) to benchmark OpenAI's Whisper capabilities in Transcribing ARZ Egyptian Arabic speech into Arabic and English text measuring RTF, WER and BLEU scores.
- **Scientific innovation dynamics (SID)** Alexandria, EG  
*Applied ML and AI researcher* July 2024 - Sep 2025
  - **Pytorch data pipeline:** Transferred a mainly Scipy signal processing pipeline applying various filters, preprocessing techniques and normalization to PCG signals to a Pytorch dataset class leveraging Gpu acceleration (P100) speeding up processing times from 5 hours on 8000 PCG's to 40 minutes .
  - **PCG Data collection:** Collected and Curated 8000 PCG audio recordings spanning multiple sources that later got fed to the Pytorch data pipeline.
  - **CNN MFCC baseline model:** Built a baseline CNN classifier for unhealthy heartsounds detection using MFCCS achieving an F1 score of 0.7.
  - **Microfluidic live cell separation:** Designed end to end a Microfluidic cell separator using live cell imagery detecting the target cell using Yolov11 replacing cell staining and tagging.
  - **Yolov11 baseline:** Pretrained Yolov11 on 1.8 mil+ phase contrast cell imagery, achieving mAP50 0.702, mAP50–95 0.436, Precision 0.813 and Recall 0.581. Considering the imbalanced dataset and high overlap of cells in most imagery.
- **ESRI North Africa** Cairo, EG  
*GIS Intern* June 2025 - July 2025
  - **Presto-GeoAI:** Fine tuned Presto a foundational GeoAI model for the task of Classifying 3 crop types across Africa as my internship project and as part of a global competition on Zindi, achieving position 37 on the leaderboard, finally I added an export layer that integrates the model's output into Esri's ArcGIS suite where you can view the results as a feature layer.
- **Faculty of Dentistry STDF Grant Proposal** Alexandria, EG  
*Research Intern* May 2025 - June 2025
  - **AI-Assisted Modeling and Optimization of 3D Bioprinted Grafts for Wound Healing:** Proposed a novel AI solution to model patient and task specific Bioink formulations and wound healing scaffolds ensuring maximum compatibility and potency.

## PROJECTS

---

- **Shato Robotic Assistant:** A voice controlled robotic assistant that can navigate it's environment. Built using microservices architecture and at it's core a Gemma 3 270M LLM fine tuned on a custom instruction dataset (SFT lora with unsloth).
- **XAI with CNNs:** Fine tuned pre-trained ResNet-34 and MobileNetV2 on Caltech 101 dataset. Implemented Adversarial Attacks, Systematically evaluated model robustness against Fast Gradient Sign Method (FGSM) attacks using torchattacks. applied Grad-CAM and vanilla gradient saliency mapping for explainability. Implemented and evaluated defensive strategies against adversarial attacks.
- **Virtual Cell challenge 2025 CAPE-model:** Cape is a virtual cell model aimed at accurately predicting cell state after Crispr gene perturbation. I was responsible for implementing multiple loss functions including wasserstien loss using geomloss.
- **ACE-Hallucinations:** Built hallucination-resistant LLM for Russian hallucination detection (codeforces) using Agentic Context Engineering (ACE) framework. Trained Gemma-3-270M as Generator with Gemini Flash 2.5 as Reflector/Curator on 380 SberQuAD-derived examples that were prepared/transformed with Gemini Flash 2.5 to learn 83 anti-hallucination strategies. Achieved 0% hallucination rate with conservative refusal behavior, deployed via llama.cpp BF16 GGUF for GPU inference.