Amr Amer

Machine Learning Engineer

M.Sc. Visual Computing, Saarland University

Computer Vision · Multimodal AI · Deep Learning

Contact

Saarbrücken, Germany

+49 162 6667294

amribrahim.amer@gmail.com

Portfolio: amramer.github.io

linkedin.com/in/amr-amer

github.com/amramer

Skills

Programming Languages

Python, Java, C++, SQL

Machine Learning

PyTorch, TensorFlow, Scikit-learn, MONAI (Medical Imaging), Hugging Face Transformers, FastAI

Computer Vision

OpenCV, YOLO (v8/v10), Object Detection, Instance & Semantic Segmentation, Multi-Object Tracking (MOT), 3D Human Mesh Recovery, SMPL-X, PIXIE

Generative & Multimodal AI

Vision-Language Models (CLIP, LLaVA), VQ-VAE, Vision Transformers (ViT), Cross-Modal Transformers, Audio-Visual Modeling

MLOps & Experimentation

Weights & Biases (W&B), TensorBoard, Optuna, Data Pipelines, ETL Workflows, Hyperparameter Tuning, Experiment Tracking

Deployment & Infrastructure

Profile

Machine Learning Engineer with M.Sc. in Visual Computing from Saarland University, research experience at DFKI, and software development experience. Specialized in computer vision, multimodal machine learning, and generative AI, with experience building end-to-end ML pipelines spanning data processing, model training, evaluation, and deployment. Worked on applied AI projects in human-computer interaction and social behavior analysis, medical imaging, autonomous driving, and sports analytics.

Portfolio Website: amramer.github.io

Experience

Research Assistant (Machine Learning & Multimodal AI) — **09/2021 – 12/2025**

German Research Center for Artificial Intelligence (DFKI), Saarbrücken

- Contributed to the **MePheSTO project**, developing machine learning pipelines for psychiatric interaction analysis from video, speech, and biosignal data to derive behavioral markers relevant to schizophrenia and depression.
- Implemented **multimodal models for social behavior understanding**, integrating visual, temporal, and speech-related features to analyze non-verbal communication in recorded human interactions.
- Developed a **Vision Transformer-based generative model** for non-verbal behavior synthesis, generating personality-aware listener responses in dyadic interactions through facial expressions and upper-body motion.
- Worked at the intersection of multimodal behavior analysis, human-computer interaction, and socially interactive AI, contributing to both research and applied AI systems for human behavior understanding.

Software Developer (Backend & Data Systems) — **10/2016 – 08/2020**

General Authority of Civil Aviation (GACA), Jeddah, Saudi Arabia

- Developed and maintained backend services and enterprise applications using **Java** and **REST-based architectures**, supporting operational systems in a production environment.
- Built data-driven application components integrating **relational databases (SQL)**, enabling structured data processing and efficient data access across internal systems.
- Improved system reliability through **debugging, performance optimization, and modular design**, applying clean code practices and contributing to scalable and maintainable software systems.

Software Engineering Intern. — **04/2016 – 09/2016**

Ejada Systems, Alexandria, Egypt

- Developed and maintained enterprise software components using Java and J2EE technologies.
- Built and optimized SQL queries and database-driven modules to support efficient data access and processing.

Docker, FastAPI, Flask, Streamlit, AWS (EC2, S3), NVIDIA CUDA, Multi-GPU Training (DDP), SLURM

Software Engineering
REST APIs, Modular Design, Clean Code, Git, PyTest, Performance Profiling, Reproducible ML Pipelines

Languages

- English — C1 (Fluent)
- German — B1 (Progressing)

Education

Visual Computing (M.Sc.) 10/2020 – 04/2024
Saarland University, Saarbrücken, Germany
Degree: Master of Science, Grade: 1.9
Thesis Website | Thesis Code

- **Master's Thesis:** *Personality-Aware Non-verbal Behavior Generation in Dyadic Interactions* (collaboration with the Max Planck Institute for Informatics and DFKI Saarland).
- Developed a multimodal generative model using **Vision Transformers, VQ-VAE**, and **PIXIE (SMPL-X)** to synthesize realistic listener avatars conditioned on personality traits.
- Built and deployed an end-to-end pipeline for data preprocessing, PyTorch training, and evaluation on **SLURM clusters** with **NVIDIA A100 GPUs**, including hyperparameter optimization.
- Achieved state-of-the-art performance on UDIVA (**FID = 6.15, P-FID = 10.31** for facial expressions; **FID = 43.16, P-FID = 87.73** for body and hand gestures), with user studies showing **86% accuracy** in identifying extroverted vs. introverted avatars.

Computer Engineering (B.Eng.) 09/2011 – 02/2016
Arab Academy for Science and Technology, Egypt
Degree: Bachelor of Engineering, GPA: 3.2/4.0

- Thesis: Sign Language Translator System using Kinect device, implemented in C++ Watch Demo

Projects

- **Badminton-VisionAI**
  Developed ML pipeline for player & shuttle tracking, shot classification, and mini-court projection; achieved **92% shot detection accuracy**.
  (OpenCV · YOLOv8 · SAM · TrackNet · Homography · Docker · Streamlit · Plotly)
  Live Demo Video
- **Personality-Aware Non-verbal Behavior Generation**
  Implemented a Vision Transformer based model to generate personality-conditioned avatars; achieved **86% accuracy** in distinguishing extroverted/introverted traits.
  (PyTorch · Transformers · SMPL-X · CUDA · SLURM · Multi-GPU Training (A100))
  Thesis Website Thesis Document
- **MePheSTO (DFKI Research Project)**
  Developed ML pipeline analyzing video - speech - biosignal data to derive behavioral markers for psychiatric disorders.
  (PyTorch · Transformers · Multimodal Fusion · Enroot · SLURM · TensorBoard)
- **Semantic Segmentation for Autonomous Driving**
  Built segmentation pipeline on BDD100K dataset with optimized preprocessing & training; achieved **mIoU = 0.45** (road: 0.88, vehicles: 0.78) and **30% faster inference**.
  (PyTorch · Fastai · Semantic Segmentation · Optuna · Weights & Biases)
- **3D Brain Tumor Segmentation**
  Developed 3D SegResNet pipeline using MONAI for MRI segmentation; achieved **Mean Dice = 0.78** (WT: 0.90, TC: 0.82) and reduced false positives by **20%**.
  (PyTorch · Dice-Loss · MONAI · 3D CNNs · Sliding Window Inference)

Additional projects and demos: amramer.github.io