

MOHAMED SAMIR EL SAYED

Perception Engineer

✉ mohamedsamirtkd@gmail.com 📞 +20 1270184543

🌐 Portfolio | Mohamed Samir 📍 Alexandria, Egypt | Open to Relocation

🐙 GitHub | mohamedsamirx 🌐 LinkedIn | Mohamed Samir

Experience

Computer Vision Engineer

March 2025 - Present

Anovate.Ai 🌐

- Architected and engineered a Full **Perception system** using ROS2 stack, integrating 6 ZED cameras and ultrasonics with sub-millimeter calibration precision (**Kalibr**). Fine-tuned a Mask-RT-DETR model trained on a custom 5x augmented dataset (**0.89 mAP**) and deployed via **C++/TensorRT** on Jetson Orin Nano for a 15x inference speedup. Implemented a robust 3D pipeline utilizing stereo depth with classical enhancement to mitigate outdoor sunlight interference, fusing segmentation masks to compute precise 3D Oriented Bounding Boxes (OBB) for robotic manipulation.
- Led 4D virtual patient reconstruction using **Gaussian Splatting and face avatar models** (INSTA, Fate-Avatar); developed a novel pipeline for accurate 3D/4D dental model (intraoral scan) alignment and tracking within dynamic facial avatars, employing 3D Slicer, PnP, and Metrical Tracker algorithms.

Research assistant

February 2025 - Present

Connected Autonomous Vehicles Lab, The American University in Cairo (AUC) 🌐

- Architected and engineered the full stack for a golf cart conversion, developing a **high-performance C++** server that integrates 6 cameras (GStreamer/DeepStream) for parallel acquisition and executes dynamic-batch **YOLOv11 TensorRT** inference with custom **CUDA pre/post-processing**. Integrated a complete drive-by-wire system via **WebRTC** (libdatachannel) and a custom Android app (Kotlin), managing real-time sensor streaming and vehicle actuation via ESP32.
- Conducted advanced research on **high-resolution segmentation**, developing a hybrid architecture that integrates **Mamba and Implicit Neural Representations (INR)** to tackle scale variation. Benchmarked the model on the DIS5K dataset, achieving **state-of-the-art** boundary precision by effectively modeling long-range dependencies without the memory cost of traditional upsampling methods.

Perception Engineer Intern

July 2024 - September 2024

Bright-skies 🌐

- Developed and integrated a Box Fitting algorithm with a lightweight 3D deep learning model for yaw angle prediction in the CVAT annotation tool. The model achieved a Mean Average Precision (mAP) of **89.6%** at an IoU threshold of **0.7** on the KITTI dataset. This enhancement enables 2D bounding box annotation on point clouds from the bird's-eye view, automatically fitting a 3D bounding box and predicting the yaw angle, optimized for CPU use. The tool reduced annotation time by **40%**.

Perception Engineer Intern

September 2023 - December 2023

Bright-skies 🌐

- Fine-tuned various SLAM (Simultaneous Localization and Mapping) algorithms, including Point-LIO, Fast-LIO-2, LIO-SAM, and others. Utilized data from cameras, LiDAR, and GPS sensors to generate precise maps of the vehicle environment. Transformed the generated maps into the XODR format for HD map correction, ensuring compatibility with autonomous systems. Integrated the SLAM solutions with the ROS2 framework for efficient map generation and real-time processing.
- Developed and optimized dynamic object filtering algorithms, including ERASOR, Removert, and RANSAC, utilizing LiDAR point cloud data. These algorithms were employed to extract and filter road features and dynamic obstacles from the vehicle environment map, enhancing the accuracy and reliability of the map with a **15%** improvement in map accuracy for autonomous navigation.

Education

Bachelor of Computing and Data Science

Graduated: June 2025

Alexandria University

GPA: 3.76/4.0

(Excellence with Honor)

Skills

- **Programming Languages:** - Python - Modern C++ - CUDA
- **Deep Learning:** - Computer Vision - 2D/3D Object Detection - Instance/Semantic Segmentation - 2D/3D Object Tracking - Training Deep Learning Models
- **Robotics:** - SLAM (Simultaneous Localization and Mapping) - Automatic HD Mapping
- **Frameworks and Tools:** - ROS/ROS2 - Foxglove Studio - PyTorch - ONNX - TensorRT - PaddleX/PaddleDetection - CVAT - CARLA - LaTeX - Git - GitHub
- **Libraries:** - OpenCV - Open3D - PCL - Eigen - scikit-learn - Albumentations
- **Additional Skills:** - Problem-Solving - Communication Skills - Presentation Skills - Time Management

Key Projects

- Contributed to the open-source YOLOs-CPP project by developing high-performance C++ headers for real-time pose estimation using YOLOv8 and YOLOv11. Added support for Oriented Bounding Box (OBB) format in YOLOv8 and YOLOv11, implemented segmentation and detection support for YOLOv9, and integrated YOLOv12 detection. Additionally, fixed several post-processing functions across multiple files, improved the build script, and optimized ONNX Runtime integration for efficient inference. 🔄
- Developed and published a high-performance inference pipeline for YOLOv11 and YOLOv12 models using the TensorRT C++ API and CUDA-based preprocessing. Engineered TensorRT engines from ONNX models, achieving a **4x speedup** in inference time compared to CPU. Contributed this work to the official YOLOv12 repository, integrating Google Colab support for easy deployment. 🔄🔄🔄
- Contributed to the YOLOv10 official repository by deploying a custom-trained model into multiple 2D object tracking algorithms, including BoTSORT, DeepOCSORT, OCSORT, HybridSORT, ByteTrack, and StrongSORT, utilizing the BoxMOT library for efficient tracking performance. 🔄
- Developed an optimized inference pipeline for the YOLOv11, v12 model using the TensorRT C++ API. Engineered a TensorRT engine from an exported ONNX model and implemented efficient C++ code to enable real-time inference. Integrated CUDA-based preprocessing techniques to further enhance performance, achieving about **4x speedup** in inference time compared to standard CPU-based execution. 🔄🔄
- Developed an optimized inference pipeline for the YOLOv9 model integrated with the ByteTrack multi-tracking algorithm. Implemented in C++, this pipeline utilized the TensorRT engine for accelerated performance, achieving about **3x speedup** in inference time compared to CPU execution. Incorporated CUDA preprocessing to enhance tracking efficiency and accuracy.
- Fine-tuned Mask2Former for multi-class semantic segmentation, achieving mIoU values up to **65%** on challenging datasets. Trained state-of-the-art models including U-Net, U-Net++, and SegFormer with mIoU performance ranging from **60% to 85%**, leveraging tools like Hugging Face Transformers and TorchVision for diverse segmentation tasks.

Certificates & Courses

- Self-driving Car Nondegree - **Udacity**.
- Learn Deep Point Clouds: Apex Predator Techniques for 3D Deep Learning Engineers - **Think Autonomous**.
- CUDA Programming Masterclass with C++ - **Udemy**.
- Deep Learning Specialization - **Deep Learning.AI (Coursera)**.
- Self-Driving Cars Specialization - **University of Toronto (Coursera)**.
- Deep Learning for Computer Vision - **Coursat.ai, Udemy**.
- Visual Perception - **Columbia University (Coursera)**.

Languages

- **English:** Fluent (IELTS 7.0)
- **Arabic:** Native
- **French:** Fluent