Khlaifia Bilel

Khlaifiabilel@icloud.com · +216 51663575 · linkedin.com/in/khlaifiabilel · github.com/khlaifiabilel · Tunisia

Summary

ML Engineer and Startup CTO with a track record of delivering production-grade RL and CV systems in geospatial and UAV domains. Co-founded a $400K-funded AI startup (NextAV), deployed real-time inference on Jetson edge devices, and scaled a high-performing 8-engineer team. NVIDIA-certified instructor with hands-on MLOps and cloud deployment across Azure, AWS, and GCP.

Key Achievements

- Co-founded NextAV and raised $400K in pre-seed funding; led 8-engineer ML team across CV/RL deployments.
- Commercialized 10x Super-resolution and methane detection models adopted by enterprise clients in energy and resource sectors.
- Designed and scaled MLOps workflows (CI/CD, DVC), reducing deployment time from 6 weeks to 3 days and boosting retraining efficiency by 85%.
- Deployed RL agents achieving 87% success in UAV autonomous landing (AirSim, PPO/SAC).

Experience

Co-founder and CTO, NextAV (AI Geospatial Startup)

- Spearheaded technical strategy for geospatial AI, leveraging >20TB satellite imagery for 10x Sentinel-2 Super-resolution and methane detection products; managed an 8-engineer ML team and secured $400K pre-seed.
- Engineered distributed ML pipelines using Kubernetes and serverless architectures (AWS/GCP), reducing data processing time by 40% and increasing model training efficiency by 2.5x.
- Delivered advanced Deep Learning solutions (U-Nets, ViTs, GANs with PyTorch/TensorFlow), improving detection accuracy by 18% for critical geospatial features.
- Established CI/CD practices with GitHub Actions and DVC, shortening deployment cycles from weeks to <3 days for rapid production feedback.
- Optimized Computer Vision models for cloud GPUs (AWS SageMaker, GCP Vertex AI) and NVIDIA Jetson, achieving sub-45ms real-time inference for edge applications.
- Mentored engineers through NVIDIA DLI certification pathways to institutionalize deep learning best practices across the team.

Previous Roles: AI Consultant, Startupbootcamp (Advised AI startups on tech roadmaps and MLOps); Autonomous Systems Strategist, AVIONAV (Developed autonomous navigation roadmap for ultralight aircraft).

Projects

- **SI-SR-GAN: 10x Satellite Image Super-resolution**: Built a GAN-based pipeline for 10× upscaling of Sentinel-2 imagery, training on 20TB+ data using DGX A100s for precise land surface monitoring.
- **TimeSeries Forecasting: Temporal Fusion Transformer Model**: Developed a calibrated Temporal Fusion Transformer (TFT) in PyTorch for precise density and lidar forecasting; created a Microsoft Fabric dashboard for hydraulic control optimization.
- **RL UAV Engine: Reinforcement Learning for UAV Control**: Engineered PPO/SAC agents in AirSim, achieving 87% landing success in simulated UAV landings under variable conditions.
- **RedTail UAV: Autonomous Visual Navigation**: Deployed YOLOv5 and visual SLAM for GPS-denied UAV navigation on Jetson AGX (ROS, DeepStream).
- **PyTorch Core Contribution: CUDA Autograd Leak Analysis**: Diagnosed and reported a PyTorch autograd memory leak

with custom CUDA operations; supplied a minimal C++ extension reproducer, contributing to core library debugging.

Skills

- **Languages:** Python (Expert), C++, Bash, SQL
- **Frameworks:** PyTorch, TensorFlow, Keras, AirSim, Unreal Engine, ROS, Hugging Face
- **ML Domains:** Computer Vision (Object Detection, Segmentation, Super-resolution, SLAM), RL (PPO, SAC, DQN), Geospatial AI
- **DevOps/Tools:** Docker, Kubernetes, GitHub Actions, DVC, Weights & Biases (W&B), TensorBoard
- **Cloud Platforms:** Azure (3 yrs), AWS (production), GCP (pilot deployments)
- **Hardware:** NVIDIA Jetson (Orin, AGX), NVIDIA DGX, TensorRT, DeepStream SDK

Education

Private Engineering School of Gafsa

- Engineering Degree, Software Engineering for Information Systems.
- Thesis: Security investment mapping using MITRE ATT&CK and NVIDIA RAPIDS/A100.

Publications

Khlaifia, B., et al. *Few-shot crop mapping using transformers and transfer learning with Sentinel-2 time series.* ISPRS Congress 2022 (Oral Presentation). Link

- *ISPRS 2022 (Oral): Few-shot crop classification via transformers on Sentinel-2; demonstrated accuracy with minimal labeling.*

Khlaifia, B., et al. *Multi-horizon forecasting with enhanced Temporal Fusion Transformers for environmental monitoring systems.* In preparation.

- *In prep: Forecasted environmental variables with 3.7× faster inference and 42% RMSE reduction using optimized Temporal Fusion Transformers (H100).*

Certifications and Awards

- NVIDIA DLI Certified Instructor (since 2019)
- Jetson AI Specialist & Ambassador
- Team Lead, CanSat Winner, AAS 2019
- Speaker: La French Tech, GITEX, Stars of Science