Orestis Zambounis

me@orestisz.com · Swiss Citizen

github.com/orestis-z, linkedin.com/in/orestis-z
Deep Learning, Computer Vision, Robotics,
Systems & Control, Distributed Systems

Experience

2023 - Present (1 year) · **Senior Machine Learning Engineer** · QSC · Zurich, CH · Remote

- Ported all vision ML models to ONNX and **TensorRT**, **tripling** pipeline speed and reducing VRAM usage by **15%**
- Implemented batched inference, increasing system speed by **30%** on resource-constrained hardware
- Led efforts to prototype innovative CV/ML experiences, staying up-to-date with the latest research
- Redesigned ML architecture for modularity and flexibility, and led efforts to clean up technical debt
- Co-managed and mentored the ML team, integrating teams and enforcing best practices
- Built monitoring tools for vision pipeline performance, enhancing observability and alerting capabilities

2021 - 2023 (2 years) · **Machine Learning Engineer** · Seervision (ETHZ Spin-off, acq. by QSC) · Zurich, CH · Remote

- Optimized real-time person detection and pose estimation pipeline, reducing latency by **24%**, VRAM usage by **45%**, and increased accuracy by **10%**
- Designed, prototyped, tuned, and deployed a face recognition system with a false-positive rate below **5%**
- Drove real-time inference optimization efforts, **tripling** the number of supported systems per hardware unit
- Collaborated with the product team to prototype and experiment with CV/ML systems for novel user experiences and features
- Received recognition for achieving the highest business impact among all engineers in 2022
- Enhanced expertise in **ROS**, **C++**, **Python**, **PyTorch**, **OpenCV**, **CUDA**, **Docker**, CI/CD, SDLC, and monitoring

2020 - 2021 (6 months) · **MLOps Engineer** · benshi.ai (funded by BMGF) · Barcelona, ES · Hybrid

- Designed and maintained a scalable end-to-end data pipeline, and led development using **Databricks**, **Spark**, and CI/CD
- Managed the ML model lifecycle, from data ingestion to deployment, utilizing **Pandas**, **MLflow**, **Azure**, **Docker**, and **Kubernetes**

2019 - 2020 (1 year) · **Full-Stack & Machine Learning Engineer** · Freelancer · Remote

- Designed, prototyped, trained, and deployed a CNN-based face predictor using **TensorFlow** and **scikit-learn**, achieving an **18%** improvement in accuracy and optimizing it for low-latency inference
- Designed a cross-platform architecture using **Cordova**, **React Native**, **React.js**, and **Electron** for the frontend; deployed scalable microservices to **AWS** with **Python**/**Flask**, **PostgreSQL**, and proxies

2016 - 2017 (1 year) · **Control Systems Engineer, Intern** · Rapyuta Robotics (ETHZ Spin-off) · Tokyo, JP · On-site

- Achieved a **55x** speedup of **NumPy**-heavy simulation iterations and open-sourced the **Python** package PyJet
- Designed energy estimators for a multicopter using a Kalman Filter (EKF), **Python**, **SciPy**, and **C++**
- Enhanced a setpoint tracking controller and performed sensor tests for a multicopter using **C++** and **Python**

Education

2018 - 2019 (7 months) · **Imperial College London** · Master's Thesis · London, UK

- Pioneered an online multi-task CNN deep learning architecture for object instance prediction, human pose estimation, instance masking, as well as multi-person tracking
- Trained the visual cue-matching head of a Siamese network using Mask R-CNN outputs on the MOT dataset
- Implemented a CNN using **Caffe2** and **Python** including custom operators with **CUDA C/C++**

2017 - 2019 (2 years) · **ETH Zurich** · MSc Robotics, Systems & Control · Zurich, CH

- Showed that an additional depth input channel improved the segmentation accuracy of Mask R-CNN by **31%**
- Designed a time-efficient training strategy using data augmentation, pretraining on synthetic RGB-D data, and fine-tuning on real-world data
- Submitted paper to CoRL 2018 and leveraged knowledge in **TensorFlow**, **Keras**, **OpenCV**, and **Python**

2012 - 2016 (3.5 years) · **ETH Zurich** · BSc Mechanical Engineering · Zurich, CH

- Graduated with more than two standard deviations above the average (top **5%**)
- Implemented balancing maneuvers for the Omnicopter to demonstrate its 6DoF flying versatility
- Derived system dynamics, synthesized non-linear attitude control algorithms, simulation environment, and a Kalman filter using quaternions, **C++** and **MATLAB** / **Simulink**