Hannes Mathis Stählin

Lachen, Switzerland | +41 79 154 55 24 | hannes.mathis.staehlin@gmail.com
linkedin.com/in/hannes-staehlin | github.com/HanStaehlin | hannesstaehlin.ch

Education

ETH Zürich | **M.Sc. Electrical Engineering and Information Technology** — Feb. 2024 – Aug. 2026
GPA: 5.54/6.00 | *Focus: Signal Processing and Machine Learning* — *Zürich, Switzerland*
- Master's Thesis at **Stanford University** (Grade: 6.00/6.00): New methods to recover signal from extremely noisy radar data, first-author paper in review at IEEE TGRS (2026)
- Exchange semester at **Seoul National University**, GPA: 4.07/4.30 (Sep. 2024 – Jan. 2025)

ETH Zürich | **B.Sc. Electrical Engineering and Information Technology** — Sep. 2020 – Jan. 2024
GPA: 5.25/6.00 | *One year full-time on Swissloop, ETH's Hyperloop focus project* — *Zürich, Switzerland*
- Bachelor's Thesis (Grade: 5.75/6.00): High-voltage power supply for a Hyperloop pod [Link]

Experience

IBM Research — Jul. 2026 – Present
Research Intern, ML Inference Systems — *Rüschlikon, Switzerland*
- Analyzed prefix caching and eviction strategies for hybrid models in vLLM on up to 64 H200 GPUs to find where performance is lost, as co-author of a paper in preparation for MLSys
- Built a measurement protocol for a shared GPU cluster that cuts run-to-run noise 24×

CVG Lab, ETH Zürich & Google Research — Aug. 2025 – Jan. 2026
Semester Project in 3D Vision — *Zürich, Switzerland*
- Extended an open-vocabulary 2D foundation model to 3D object detection from a single camera with a parallel depth branch, keeping its 2D abilities at minimal extra latency
- Trained the depth branch by distilling dense depth from DepthPro, calibrated with sparse LiDAR
- Ran all experiments in PyTorch and achieved competitive monocular 3D detection on KITTI

Ergon Informatik — Feb. 2025 – Aug. 2025
Machine Learning and AI Intern — *Zürich, Switzerland*
- Developed three methods to generate synthetic patient data with LLMs and VLMs
- Built a scalable LLM training and evaluation platform on Kubeflow and Azure Kubernetes
- Fine-tuned local LLMs (QLoRA, Unsloth) on synthetic data for medical data extraction. The models run on a smartphone and come close to frontier models on specific tasks
- Built an AI lab-table parser (FastAPI, Angular) for a medical SaaS client, now in production
- Presented the work as oral talks at RSECon25 (UK) and the EnhanceR Symposium 2025 (CH)

ETH Zürich — Sep. 2022 – Aug. 2024
Teaching Assistant, Department of Computer Science — *Zürich, Switzerland*
- Led weekly recitations for up to 120 students in Stochastic and Machine Learning, Computer Science II (Python), and Computer Science II (C++), including grading and exam invigilation
- Achieved excellent end-of-semester student ratings and high attendance in optional classes

Swissloop — Sep. 2022 – Aug. 2023
Electrical Engineer, Focus Project of ETH Zürich — *Dübendorf, Switzerland*
- One of five full-time electrical engineers in a 35-student Hyperloop team, which won 4 of 6 technical prizes at European Hyperloop Week 2023, including the "Complete System Award"
- Designed and built an 820 V / 250 kW battery and its safety concept, accepted without issues
- Developed PCBs (schematics and layout) for the Battery Management System and programmed embedded software for microcontroller and FPGA in C, C++ and SystemVerilog

Kantonsschule Zug — Aug. 2022 – Sep. 2022
Substitute Mathematics Teacher — *Zug, Switzerland*
- Explained abstract topics to students of different skill levels and kept the classes engaged

Skills, Languages and Interests

- **ML & Tools:** PyTorch, Hugging Face, QLoRA, Kubeflow, Docker, Kubernetes, Slurm (HPC), Git
- **Programming:** Python, C++, C, MATLAB, SystemVerilog, SQL
- **Spoken Languages:** German/Swiss German (Native), English (Cambridge C2 Proficiency), French (B2), Spanish (A2)
- **Interests:** Tennis, Sailing, Scuba Diving, Piano, Aviation (SPHAIR), Radio