

Furkan Küçük

Machine Learning and Artificial Intelligence Engineer/Researcher

[in linkedin.com/in/kucukfurkan](https://www.linkedin.com/in/kucukfurkan) [📞 \(+41\) 772607662](tel:+41772607662) [@ furkankucuk@outlook.com](mailto:furkankucuk@outlook.com)
[🐙 furkankucuk.com](https://github.com/iamkucuk) [🐙 github.com/iamkucuk](https://github.com/iamkucuk) [📍 1004 Lausanne, Switzerland](https://www.google.com/maps/place/1004+Lausanne,+Switzerland)
[🇹🇷](#) Turkish Nationality, Swiss Work Permit



Machine Learning and Artificial Intelligence Engineer/Researcher with strong academic foundation and 5+ years of industry experience in delivering scalable, real-world AI solutions. Expert in computer vision, deep learning, large language models (LLMs), generative AI, and multimodal foundation models. Passionate about bridging research and production by engineering maintainable, production-grade systems with measurable impact. Proven experience in developing and deploying foundation models, distributed inference architectures, and multi-node training pipelines on cloud platforms.

🔧 TECHNICAL SKILLS

Programming Languages	Python, C/C++, CUDA (GPU), TypeScript, JAVA, MATLAB, $\LaTeX$
ML Inference & Deployment	TensorRT, ONNX, DeepStream, Docker, TensorBoard
ML/AI Frameworks	PyTorch, Torchscript, OpenCV, MMCV, Keras, Scikit-Learn, AI-SDK, Mastra, LangChain
MLOps & Cloud	AWS, GCP, MLFlow, DVC, W&B, ClearML, Git, Unity, Roboflow, CVAT
LLMs & Orchestration	LLMs, vLLM, LangChain, LangGraph, Langsmith, Langflow, n8n
Data Streaming	Kafka, Kafka Connect, RedPanda, gRPC, RTSP, InfluxDB, QuestDB, Grafana
Data Science Libraries	Numpy, Pandas, scipy, sci-opt, Matplotlib, Seaborn
Specializations	LLMs, AI Agents, Computer Vision, Object Detection, OCR, Super Resolution, Domain Adaptation, Diffusion Models, ControlNet, LoRA, Distributed HPC workloads (training and inference)

🏢 WORK EXPERIENCE

11/2025 04/2025	Lead AI Solutions Architect, İKAS, Türkiye - Directed AI integration for products and internal workflows across multiple teams - Designed and deployed 6+ commercial-grade computer vision tools for image generation/editing - Built 2 agentic systems for documentation and chatbot automation AI Integration Computer Vision Agentic Systems MLOps LLMs Multi-Modal Models
04/2025 11/2024	Lead Data Scientist, EPAM SYSTEMS, Türkiye - Led EPAM's Computer Vision initiatives in Turkey - Collaborated with Zalando on production-grade VTON pipeline, advancing base generator and generative super-resolution models in a 3-component architecture - Conducted multi-node training and deployment on Kubernetes clusters Computer Vision Diffusion Models Generative Super-Resolution VTON MLOps Kubernetes Multi-Node Training
11/2024 10/2023	Head of Artificial Intelligence, POLIARK INC., Türkiye - Developed LLM-based agents and integrated graph models, diffusion models and MIQP optimizations - Built custom ControlNet and LoRA pipelines for high-accuracy visual outputs in architectural design - Engineered floor plan vectorization from CAD images via segmentation, graph processing and optimization - Trained task-specific LLMs for architectural design and space planning - Managed GPU/CPU deployment pipelines across AWS and GCP LLM Agents Computer Vision ControlNet LoRA Instance Segmentation Graph Processing MIQP AWS GCP MLOps
09/2023 01/2020	Machine Learning Researcher in Computer Vision, DATABOSS SECURITY & ANALYTICS, Türkiye - Conducted research on video object detection, image enhancement, super-resolution, dehazing, and OCR - Built distributed real-time CV deployment systems utilizing TensorRT and DeepStream - Led MLOps infrastructure buildout with DBC, experiment tracking and orchestration - Contributed to design, development and deployment of from-scratch distributed inference architecture; conducted multi-node training - Improved object detection and super-resolution performance by 80% and 30% respectively Computer Vision TensorRT DeepStream MLOps Object Detection OCR Super Resolution Dehazing Multi-Node Training Distributed Inference
11/2018 05/2018	Research Engineer, TOBB UNIVERSITY OF ECONOMICS AND TECHNOLOGY, Türkiye - Led "Collaborative Caching in 5G Networks" with HAVELSAN, resulting in patents in EU, US and Turkey - Built LTE simulations using NS-3 5G Networks NS-3 Research Patent Development Probability Theory Caching Algorithms



LANGUAGE SKILLS

Turkish



English (TOEFL iBT : 98/120)



French



EDUCATION

09/2026
01/2022

Master of Science in Computer Engineering, MIDDLE EAST TECHNICAL UNIVERSITY, Türkiye

GPA : 3.93/4.00 (Summa Cum Laude)

- One-year preparatory program in undergraduate Computer Engineering (Operating Systems, Algorithms, Data Structures, Discrete Computational Structures)
- Research on advanced Computer Vision techniques under Asst. Prof. G. Cinbiş
- Focus on Domain Adaptation, Generative models, and compression

Computer Vision Domain Adaptation Video Compression Diffusion Models

05/2018
09/2013

Bachelor of Engineering in Electrical and Electronics Engineering, TOBB ETU, Türkiye

GPA : 3.27/4.00

- Contributed to radar systems and 5G communication networks research
- Collaborated with Dr. T. Girici and Dr. A.M. Yüksel Turgut on caching algorithms
- Strong foundation in electrical engineering and telecommunications

5G Networks Radar Systems Telecommunications Caching Algorithms Probability Theory



PROJECTS

EULERIAN VIDEO MAGNIFICATION IN CUDA

github.com/iamkucuk/eulerian-video-magnification-cuda

Hand-written CUDA C++ implementation of MIT's Eulerian Video Magnification, up to ~1,650x faster than the NumPy reference implementation on an RTX 3090 (FP16, kernel time). Six backends (CUDA, Metal, Vulkan, OpenCL, PyTorch, NumPy).

CUDA C++ Metal Vulkan OpenCL PyTorch pybind11 GPU Benchmarking

ResuVL

github.com/iamkucuk/tiny-resume-parser huggingface.co/furkankucuk

Three fine-tuned vision-language models, ResuVL-S, ResuVL-M and ResuVL-L (0.26B, 0.51B and 0.87B parameters), that read a CV page image and write its content as JSON, with no PDF text layer. On 204 held-out pages ResuVL-L scores 0.845 field F1 and reads a page in about a minute on a laptop CPU.

Vision-language models Fine-tuning Knowledge distillation Document AI PyTorch



PUBLICATIONS & PATENTS

A Use Frequency-based Cooperative Caching Method for Multi-layer Network Structures (2022)

Patent Granted : TR : WO2020142046A1 - US : US11343350B2 - EU : EP3906658NWA1

F. Küçük, T. Girici, A. M. Yüksel, M. K. Cetin

Collaborative Hierarchical Caching in ICN-Based 5G Networks (2019)

Third International Balkan Conference on Communications and Networking, BALKANCOM'19

F. Küçük, T. Girici, A. M. Yüksel, M. K. Cetin

Fair Beam Allocation in Millimeter-Wave Multiuser Transmission (2017)

WSA 2017; 21th International ITG Workshop on Smart Antennas, pp. 1-5, Berlin, Germany

F. Karababa, F. Küçük, T. Girici



CONFERENCES & SEMINARS

CVPR '23 Vision-based InduStrial InspectiON (January 2023 - June 2023)

Vancouver, Canada

1st workshop on Vision-based InduStrial InspectiON, in collaboration with Apple