Berk Sezer

Undergraduate Researcher in Robotics, Vision, and Applied and Theoretical Machine Learning
berk.sezer@metu.edu.tr | berksezer9.github.io | Ankara, Turkey

Publications

Same Actions, Different Context: Investigating Appropriateness of Robot Actions Across Environments, Embodiments, and Explanations
Under review at the IEEE International Conference on Robotics and Automation (ICRA 2027, Main Conference).
Fethiye Irmak Doğan*, **Berk Sezer***, Nikhil Churamani, Yiming Luo, Hatice Gunes (*Equal contribution)

DROD: Dual-Ratio Ordered Dropout for Federated Continual Learning on Heterogeneous Devices
arXiv preprint (2026).
Berk Sezer, Özgür Soysal | Project Page

An Impossibility Result for ETF Target Geometry, and an Equinorm Remedy for Imbalanced Analytic Continual Learning
Accepted to the NeurIPS 2026 Workshop on Symmetry and Geometry in Neural Representations (NeurReps, Extended Abstracts Track).
Berk Sezer | Project Page

Gaze4HRI: Zero-shot Benchmarking Gaze Estimation Neural-Networks for Human-Robot Interaction
Accepted to the 20th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2026, Main Conference).
Berk Sezer, Ali Görkem Küçük, Erol Şahin, Sinan Kalkan | Project Page

Education

Middle East Technical University (METU) — Ankara, Turkey
B.Sc. in Computer Science (4th Year), ***CGPA: 3.89/4.00****, ranked 7th/144 (Top 5%)* — *Oct 2023 - Present*

Advanced Coursework Highlight: **Graduate-level Deep Learning (Grade: 4.0/4.0)**

Research & Engineering Experience

University of Cambridge, Department of Computer Science and Technology — Cambridge, UK
Undergrad Researcher, Affective Intelligence and Robotics (AFAR) Lab — *July – September 2026*

Full-time internship under the supervision of Prof. Hatice Gunes, initially exploring Continual and Federated Learning architectures, but later focusing on Cross-Domain Generalisation. Co-led the benchmarking of a social-appropriateness of robotic actions dataset across CNN and VLM-LoRA models, covering pipeline implementation, experimental validation, and manuscript preparation resulting in a co-first-author ICRA 2027 submission.

METU ROMER (METU Robotics Center) — Ankara, Turkey
Undergrad Researcher & AI–Robotics Engineer — *Sep 2024 – Present*

Home Robotics Project (Beko R&D Partnership) – *AI–Robotics Engineer* *(Part-Time)*
The Home Robotics Project aims to develop assistive robots for domestic environments. I worked under Profs. Şahin, Kalkan, and Cinbis, and have stepped back to focus on my ICRA 2027 and CVPR 2027 papers.

- **VLA Models:** Trained Vision–Language–Action (VLA) models in real-world and simulated settings.
- **Data Pipeline:** Led the architecture and implementation of our data collection pipeline to train VLA models.

USTA Human-Robot Interaction Project – *Undergrad Researcher*
The USTA Project advances Human–Robot Interaction (HRI) to enable collaborative robotic manipulators (cobots) to operate more naturally alongside humans. My work focuses on vision-based perception, specifically gaze estimation for HRI.

- **Gaze Estimation:** As first author of the IEEE FG 2026 main conference paper ($\sim 35\%$ acceptance rate), *"Gaze4HRI: Zero-shot Benchmarking Gaze Estimation Neural-Networks for Human-Robot Interaction"*, I led data collection, code implementation, experiment design, and manuscript writing.

- **Blink Estimation:** Finalizing the manuscript for CVPR 2027 on a novel blink estimation framework that I led. Contributed the core of the framework by applying temporal-order-based pre-training to enhance blink detection accuracy, which on its own substantially improves the majority of existing SOTA models as well as our own new model. On MPEblink 2.0, it beats prior SOTA on the main metric, Blink-AP, by over 80%.

Independent Research

- **DROD:** Dual-Ratio Ordered Dropout, a method for federated continual learning on heterogeneous devices, where clients run different-sized models and their data keeps drifting to new classes. Instead of shrinking a residual block by one width, it cuts the block along two independent ratios, one for the shortcut and one for the branch, so each compute budget becomes a curve of sub-models rather than a single point. At matched parameters and FLOPs it beats every baseline we compared on Split-CIFAR-100, and it stays ahead of the strongest baselines across continual-learning methods (ER, DER++ and BiC), by 8–14% relative. My co-author, a Stanford PhD student, and I had to step back from a full publication project due to other commitments (ICRA 2027, CVPR 2027, etc.).
- **ETF Geometry:** I proved that Neural-Collapse ETF targets are identical to one-hot targets in analytic continual learning, and that progressively expanding them causes forgetting. Equalizing weight norms (equinorm) instead rescues long-tailed data, lifting tail-class accuracy 5× at 100:1 imbalance.

METU Department of Computer Engineering — Ankara, Turkey
Senior Project Member, MIMIC-Sound (team of 5, advisor: Prof. Sinan Kalkan) — *Sep 2026 – Present*

MIMIC-Sound is an AI tutor that shows how to physically change a sound, not just score it. Project Page

- **Idea & Prototype:** Conceived and prototyped the project to address a gap in modern AI: machines can sing, speak and imitate any voice, but people still need clear vocal instructions. It adds articulatory motion as a new modality for mimicry, inferring it from audio alone, turning the gap to the target into a plain instruction built from sounds the learner already makes, and playing the corrected move on a vocal-tract synthesizer.

Industry Experience

Altima Sistem ve Yazılım Teknolojileri Ltd. — Ankara, Turkey
Full-Stack Engineer — *Aug 2022 – Jun 2023 (Full-Time)*

Built and deployed a full-stack LLM + RAG application giving legally valid, citation-linked answers from legal texts for the Union of Turkish Survey Engineers (HKMO), and built the core backend of randevuyla.com.

Technical Skills

Robotics & Data: ROS 2, sensor fusion, Vision–Language–Action (VLA) models, large-scale data collection.

Computer Vision & ML: Deep learning, self-supervised pre-training, continual learning, federated learning, real-time 3D perception (RGB–D), face and gesture detection, gaze estimation.

Speech & Audio: Phone recognition, articulatory inversion, formant analysis, vocal-tract synthesis.

AI Assistants: VLM/LLM assistants with Retrieval-Augmented Generation (RAG) and structured outputs.

Programming & DevOps: Python, C, C++, PHP, JavaScript/TypeScript; Unix, HPC servers, Docker.

Web Systems: Laravel, PostgreSQL, MySQL, Redis, MongoDB, Vue.js, HTML, CSS, TypeScript.

References

Prof. Sinan Kalkan
Professor of Computer Engineering, METU
skalkan@metu.edu.tr

Prof. Erol Şahin
Director of ROMER, METU
erol@metu.edu.tr

Prof. Hatice Gunes
Director of AFAR Lab, University of Cambridge
hg410@cam.ac.uk

Prof. Gökberk Cinbis
Associate Professor of Computer Engineering, METU
gcinbis@metu.edu.tr

Please feel free to contact any of the individuals listed above for inquiries. (symbols link to Google Scholar profiles.)