

Batu Ozturkler

650-680-6618 | batumozturkler@gmail.com | batuozt.github.io | [Google Scholar](#)

RESEARCH AND PRODUCT INTERESTS

I am interested in developing agentic AI systems with meaningful product impact. My recent work spans **agentic model orchestration for AI coding, long-horizon tool-using workflows, LLM post-training, calibrated confidence and hallucination detection, and probabilistic reasoning**. During my Ph.D., I also worked on diffusion models, robustness under distribution shifts, memory-efficient learning, self-supervised learning, and medical and computational imaging.

EXPERIENCE

Senior Research Scientist

Aug 2026 – Present

Research Scientist 2

Mar 2026 – Jul 2026

Microsoft CodeAI

Mountain View, CA

- **Agentic Model Orchestration and Routing for GitHub Copilot:** Trained small coordinator LLMs to compose tool-using workflows and route each step to the appropriate model. Developed prompt-cache-preserving routing for long-horizon, multi-turn conversations; matched the always-strong baseline within 0.3 percentage points on SWE-Bench Verified while reducing model cost by 54%.
- **Secret Scanner:** Led development of a fine-tuned ModernBERT credential scanner deployed across GitHub's internal and external products at the scale of all GitHub push traffic, reducing inference cost by over 98% and latency by over 90% while achieving higher precision than multi-LLM workflows.

Applied Scientist 2

May 2024 – Feb 2026

Microsoft

Mountain View, CA

- **Confidence Estimation in LLMs:** Developed post-training methods for calibrated confidence estimation using auxiliary predictors and RL-based optimization, improving hallucination detection and reliability in production systems including vision-language models for structured extraction from long documents.
- **LLM Post-training:** Co-developed multilingual post-training for GPT-4o across mid-training, SFT, DPO, and RLHF; built large-scale benchmarks, scalable evaluation pipelines, and synthetic-data generation systems that improved multilingual model performance.
- **Azure AI Foundry Agents:** Designed agentic workflows with tool calling, retrieval, and search; generated synthetic evaluation data and built benchmarking pipelines. Optimized tool-calling prompts to deliver 30% faster responses without reducing response quality.

Postdoctoral Scholar

Mar 2024 – May 2024

Research Assistant

Sep 2019 – Jan 2024

Stanford University

Stanford, CA

- Developed memory-efficient training methods for high-dimensional MRI reconstruction and self-supervised-learning-based augmentation to improve robustness under signal-to-noise and motion distribution shifts.
- Developed equivalent convex formulations of non-convex neural networks to provide robustness, interpretability, and convergence guarantees.

Research Intern

Sep 2022 – Apr 2023

NVIDIA

Santa Clara, CA

- Improved the robustness and efficiency of diffusion models for MRI reconstruction under distribution shifts using SURE and variational inference.

Research Intern

Jun 2022 – Aug 2022

Microsoft Research

Redmond, WA

- Developed a probabilistic inference method that improved LLM reasoning and robustness to prompt design; characterized failure modes of in-context learning.

EDUCATION

Stanford University

Sep 2019 – Jan 2024

Ph.D. in Electrical Engineering; GPA: 3.99/4.00

Stanford, CA

– Advisors: Prof. John Pauly and Prof. Mert Pilanci

– Thesis: Efficient and robust deep learning for medical imaging and natural language processing

Stanford University

Sep 2019 – Jun 2021

M.S. in Electrical Engineering; GPA: 3.99/4.00

Stanford, CA

Middle East Technical University (METU)

Aug 2015 – Jun 2019

B.S. in Electrical Engineering; GPA: 3.96/4.00, Rank: 3/413

Ankara, Turkey

SELECTED PUBLICATIONS

1. **Ozturkler, B.M.**, Malkin, N., Wang, Z., and Jojic, N. ThinkSum: Probabilistic reasoning over sets using large language models. **ACL 2023**. arXiv:2210.01293.
2. Sahiner, A., Ergen, T., **Ozturkler, B.M.**, Pauly, J., Mardani, M., and Pilanci, M. Scaling Convex Neural Networks with Burer-Monteiro Factorization. **ICLR 2024**.
3. Sahiner, A., Ergen, T., **Ozturkler, B.M.**, Pauly, J., Mardani, M., and Pilanci, M. Unraveling Attention via Convex Duality: Analysis and Interpretations of Vision Transformers. **ICML 2022**.
4. **Ozturkler, B.M.**, Mardani, M., Vahdat, A., Kautz, J., and Pauly, J. RED-diff: Regularization by Denoising Diffusion Process for MRI Reconstruction. **NeurIPS 2023 Workshop on Deep Learning and Inverse Problems**.
5. Desai, A.D.*, Gunel, B.*, **Ozturkler, B.M.**, et al. VORTEX: Physics-Driven Data Augmentations for Consistency Training for Robust Accelerated MRI Reconstruction. **MIDL 2022 Best Paper Award**.
6. **Ozturkler, B.M.**, Liu, C., Eckart, B., Mardani, M., Song, J., and Kautz, J. SMRD: SURE-based Robust MRI Reconstruction with Diffusion Models. **MICCAI 2023** (early accept, top 15% of all submissions).

TECHNICAL SKILLS

Languages: Python, MATLAB, R, C/C++

Deep Learning: PyTorch, TensorFlow, Keras, JAX

Systems and Cloud: Linux, Git, Docker, Kubernetes, Google Cloud Platform, Microsoft Azure, Weights & Biases

SELECTED HONORS AND SERVICE

- **NSF Graduate Research Fellowship (NSF GRFP)**
- **Stanford EE Departmental Fellowship**
- **Reviewer:** ACL, NeurIPS, ICML, ICLR, MICCAI, IEEE Transactions on Medical Imaging
- **Program Committee:** ICML Workshop on Knowledge and Logical Reasoning in the Era of Data-Driven Learning