

Quanling Zhao

quzhao@ucsd.edu | Website | Google Scholar

Research Interests

Efficient and theoretically grounded machine learning; randomized and kernel methods; efficient LLM/VLM inference; vector symbolic architectures and hyperdimensional computing.

Education

Ph.D. in Computer Science

University of California San Diego

2024–Present

CA, USA

B.S. in Computer Science

University of California San Diego

2023

CA, USA

Research Experience

System Energy Efficiency Lab

Graduate Researcher / Advisor: Tajana Rosing

2024–Present

UCSD

- Lead the development of efficient and theoretically grounded machine learning methods across statistical learning, representation learning, and efficient AI systems.
- Drive multiple research directions from initial problem formulation through algorithm design, theoretical analysis, experimental validation, and publication.
- Mentor junior researchers and coordinate collaborative research across academic and industry teams.

IBM Research

Research Collaboration

2025–Present

- Lead collaborative research with IBM on database systems and semiconductor manufacturing, with a focus on efficient, memory-centric, and online intelligent processing.
- Developed new HDC-based systems for in-storage database acceleration on emerging memory devices and online discovery of recurring semiconductor defect patterns, resulting in publications at DAC 2026 and ASP-DAC 2027.

Intel Corporation

Research Collaboration

2026–Present

- Lead collaborative research with Intel researchers on an accurate and ultra-fast anomaly detection algorithm; developed RGLD, a randomized global-local density estimator submitted to ICLR 2027.

UC San Diego, Computer Science & Engineering

Undergraduate Researcher / Advisor: Tajana Rosing

2021–2023

UCSD

- Developed new learning methods based on hyperdimensional computing for federated and multimodal learning, targeting efficient intelligence on distributed and resource-constrained systems.
- Research contributed to publications and recognition with UCSD CSE’s Excellence in Research Award, one of two recipients in the graduating class.

Publications

* Equal contribution.

Manuscripts Under Review

1. **Quanling Zhao**, Jiaying Yang, Tianqi Zhang, Ziyang Hao, Fatemeh Asgarinejad, Flavio Ponzina, Tajana Rosing. “Superposed Latent Autoencoder.” *Submitted to ICLR 2027*. [arXiv]
2. **Quanling Zhao**, Jiaying Yang, Ye Tian, Josh Victoria, Zhijun Wang, Pietro Mercati, Onat Gungor, Tajana Rosing. “RGLD: Randomized Global-Local Density Estimation for Tabular Anomaly Detection.” *Submitted to ICLR 2027*. [arXiv]
3. **Quanling Zhao**, Nilesh Prasad Pandey, Ye Tian, Tajana Rosing. “Superposed Inference for Hyperdimensional Computing.” *Submitted to ICLR 2027*.

Peer-Reviewed Publications

4. **Quanling Zhao**, Ava Emami, Tsuyoshi Ide, Rajiv Joshi, Steve Holmes, Michael Passow, Tajana Rosing. “Online Open-World Wafer Defect Discovery using Hyperdimensional Computing.” Asia and South Pacific Design Automation Conference (ASP-DAC), 2027.
5. Nilesh Prasad Pandey, Jason Kong, Lanxiang Hu, **Quanling Zhao**, Yujie Zhao, Onat Gungor, Hao Zhang, Tajana Rosing. “AgentKVShift: Efficient KV Cache Reuse for Agentic Memory Systems.” Conference on Neural Information Processing Systems (NeurIPS), 2026. [arXiv]
6. You Hak Lee, Keming Fan, Xiaofan Yu, **Quanling Zhao**, Flavio Ponzina, Tajana Rosing. “FeUHD: Unsupervised Federated Learning using Hyperdimensional Computing.” IEEE Transactions on Circuits and Systems for Artificial Intelligence (TCASAI), 2026. [Paper]
7. **Quanling Zhao***, Yanru Chen*, Runyang Tian, Sumukh Pingre, Weihong Xu, Augusto Vega, Steven Holmes, Saransh Gupta, Tajana Rosing. “HDDb: Efficient In-Storage SQL Database Search Using Hyperdimensional Computing on

Ferroelectric NAND Flash.” Design Automation Conference (**DAC**), 2026. [arXiv]

8. Le Zhang*, **Quanling Zhao***, Run Wang, Shirley Bian, Onat Gungor, Flavio Ponzina, Tajana Rosing. “Offload Rethinking by Cloud Assistance for Efficient Environmental Sound Recognition on LPWANs.” ACM Conference on Embedded Networked Sensor Systems (**SenSys**), 2025. [Paper]
9. **Quanling Zhao**, Anthony Thomas, Ari Brin, Xiaofan Yu, Tajana Rosing. “Bridging the Gap between Hyperdimensional Computing and Kernel Methods via the Nyström Method.” AAAI Conference on Artificial Intelligence (**AAAI**), 2025. [Paper]
10. **Quanling Zhao**, Xiaofan Yu, Shengfan Hu, Tajana Rosing. “MultimodalHD: Federated Learning Over Heterogeneous Sensor Modalities using Hyperdimensional Computing.” Design, Automation, and Test in Europe (**DATE**), 2024. [Paper]
11. Xiaofan Yu, Ludmila Cherkasova, Harsh Vardhan, **Quanling Zhao**, Emily Ekaireb, Xiyuan Zhang, Arya Mazumdar, Tajana Rosing. “Async-HFL: Efficient and Robust Asynchronous Federated Learning in Hierarchical IoT Networks.” ACM/IEEE Conference on Internet of Things Design and Implementation (**IoTDI**), 2023. [Paper]

Workshop Papers, Posters, and Demos

1. Run Wang*, Shirley Bian*, Xiaofan Yu, **Quanling Zhao**, Le Zhang, Tajana Rosing. “Poster: Resource-Efficient Environmental Sound Classification Using Hyperdimensional Computing.” ACM Conference on Embedded Networked Sensor Systems (**SenSys**), 2024. [Paper]
2. **Quanling Zhao**, Anthony Thomas, Ari Brin, Xiaofan Yu, Tajana Rosing. “Unleashing Hyperdimensional Computing with Nyström Method based Encoding.” NeurIPS Workshop on ML with New Compute Paradigms (**MLNCP@NeurIPS**), 2023. [Paper]
3. **Quanling Zhao**, Xiaofan Yu, Tajana Rosing. “Poster Abstract: Attentive Multimodal Learning on Sensor Data using Hyperdimensional Computing.” ACM/IEEE International Conference on Information Processing in Sensor Networks (**IPSN**), 2023. [Paper]
4. **Quanling Zhao**, Kai Lee, Jeffrey Liu, Muhammad Huzaifa, Xiaofan Yu, Tajana Rosing. “FedHD: Federated Learning with Hyperdimensional Computing.” ACM Annual International Conference on Mobile Computing and Networking (**MobiCom**) Demo, 2022. [Paper]
5. Emily Ekaireb, Xiaofan Yu, Kazim Ergun, **Quanling Zhao**, Kai Lee, Muhammad Huzaifa, Tajana Rosing. “ns3-fl: Simulating Federated Learning with ns-3.” Workshop on ns-3 (**WNS3**), 2022. [Paper]

Invited Talks

1. “RGLD: Randomized Global-Local Density Estimation for Tabular Anomaly Detection”
Intel Invited Talk forthcoming, 2026
2. “Connecting Kernel Methods and Hyperdimensional Computing for Capable and Efficient Learning”
CoCoSys Center Seminar July 2026
3. “Towards Generalized Learning in Hyperdimensional Computing: Density Estimation and Hybrid Models”
CoCoSys Center Seminar July 2025

Mentoring

Ava Emami (BS 2027 at UCSD) - Online Learning	Publication: ASP-DAC’27
Ari Brin (BS 2024 at UCSD) - Kernel Approximation	Publication: AAAI’25
Run Wang (BS 2026 at UCSD) - Efficient Edge Intelligence	Publication: SenSys’24
Shirley Bian (BS 2026 at UCSD) - Efficient Edge Intelligence	Publication: SenSys’24

Professional Service

Conference service: ICLR 2027 – Reviewer; NeurIPS 2026 – Reviewer; ICML 2026 – Reviewer; AAAI 2026, 2027 – Program Committee; RAAAI@NeurIPS 2026 – Reviewer; IJCNN 2025 – Reviewer; MLNCP@NeurIPS 2024 – Reviewer.

Journal service: IEEE/ACM Transactions on Networking – Reviewer; IEEE Access – Reviewer; IEEE Transactions on Cognitive Communications and Networking – Reviewer; IEEE Transactions on Emerging Topics in Computing – Reviewer; IEEE Transactions on Neural Networks and Learning Systems – Reviewer; IEEE Journal of Biomedical and Health Informatics – Reviewer.

Research Programs & Sponsored Projects

- **CoCoSys** — SRC/DARPA JUMP 2.0 Center for the Co-Design of Cognitive Systems.
- **PRISM** — SRC/DARPA JUMP 2.0 Center for Processing with Intelligent Storage and Memory.
- **TILOS** — NSF AI Institute for Learning-Enabled Optimization at Scale.

Technical Skills

Programming:	Python, C/C++
Machine Learning:	PyTorch, scikit-learn, NumPy, SciPy, pandas
LLM/VLM:	Hugging Face Transformers, PEFT/LoRA, vLLM, quantization, KV-cache optimization
Systems/Tools:	Linux, Git, Docker, CUDA, LaTeX