ZHENNING ZHANG

LinkedIn: linkedin.com/in/zhenning-zhang | Portfolio: zhenningzhang.net | GitHub: github.com/mizaimao

PROFESSIONAL SUMMARY

Machine Learning Engineer with 8 years building deep learning and related systems for drug discovery, medical imaging, ML platforms and tooling. Specializes in multimodal models combining cell painting microscopy with chemical structures (IC50 prediction), 3D volumetric segmentation with transformer architectures, and graph-based analysis of multiplex immunofluorescence.

EXPERIENCE

Platform Engineer — Unlearn.AI, San Francisco, CA (Remote) *08/2026 – Present*

- Migrate clinical trial ML training from a Dagster stack to Databricks, working under change control.
- Conduct ongoing validation of migrated data and retrained models against the released reference models.

Machine Learning Engineer — AstraZeneca, Gaithersburg, MD *07/2019 – 07/2026*

Drug Discovery ML

- Built multimodal deep learning models for IC50 prediction by fusing cell painting microscopy with SMILES chemical representations across compound libraries totaling 39K compounds (customized cell painting transformer + ChemBERTa).
- Designed, built, and maintained the interactive visualization tool that served as the wet-lab stakeholder interface (self-hosted and deployed internally). Underlying work built on unsupervised graph neural network embeddings (Deep Graph Infomax) over multiplex immunofluorescence, extended to supervised WSI-level survival prediction and classification.

Tooling & Infrastructure

- Designed and built a unified data ingestion and preprocessing platform for biomedical imaging at scale, now adopted as the team's standard data layer: 16 multi-center clinical trial datasets totaling 150K CT volumes ingested, with automated QC and mapping of thousands of annotation masks across diverse file standards (DICOM, NIfTI), plus model serving and retraining feeding downstream training and inference.
- Built web-based visualization suites for high-dimensional embedding interpretation (heatmaps, histograms, archetype analysis) used by R&D labs to interpret model outputs.

Medical Imaging & Segmentation

- Built an interactive Transformer-backed (SwinUNETR) segmentation toolset for 3D volumetric data with text-guided prompts and in-place guided mode; halved annotation time for 4 radiologists and deployed to 12+ internal users across R&D (radiologists, ML engineers, PMs, partner labs).
- Architected patch-based unsupervised pipelines for 3D imaging datasets exceeding GPU memory, applying HDBSCAN and UMAP over learned embeddings to surface morphological structure across large multi-cohort studies.

Software Engineer — Ann Arbor Algorithms, Ann Arbor, MI *06/2018 – 06/2019*

- Built end-to-end containerized (Docker) deep learning pipelines covering ingestion, preprocessing, training, and visualization.
- Trained classification, 3D bounding-box detection, and anomaly identification models on multimodal datasets of millions of points using ResNet and U-Net variants; trained XGBoost models for structured metadata classification.
- Developed web service frameworks for cohort data management with interactive dashboards over complex distributions.

SYSTEMS & INFRASTRUCTURE (PERSONAL PROJECTS)

- **Self-Hosted ML Inference Platform —** a self-hosted multi-model inference stack: several LLMs served concurrently on a GPU-passthrough Proxmox host behind one OpenAI-compatible endpoint, fronting a RAG stack and multiple agent services. Tuned for concurrency and resident-model management, with GPU-utilization measurement under load.
- **Automated Data Pipeline —** a multi-source job ingestion and classification pipeline (SQLite backend) collecting postings across 100+ sources and routing them through several local or cloud LLMs for automated triage, with concurrent pulling and rate-limit handling.
- **Trader Advisor —** an agentic stock research application with a hand-rolled tool-calling loop (eleven tools, no framework) running on locally hosted LLMs via Ollama, with a Streamlit dashboard viewable on Hugging Face Spaces.

PUBLICATIONS

Peer-Reviewed Journal

- Y. Guan, X. Wang, H. Li, **Z. Zhang**, X. Chen, O. Siddiqui, S. Nehring, X. Huang. "Detecting asymmetric patterns and localizing cancers on mammograms." *Patterns* 1, no. 7 (2020).

Preprint

- C. Innocenti, **Z. Zhang**, B. Selvaraj, I. Gaffney, M. Frangos, J. Cohen-Setton, et al. "An unsupervised graph embeddings approach to multiplex immunofluorescence image exploration." *bioRxiv* 2021.06.09.447654 (preprint).

Conference Posters / Abstracts

- G. Hughes, M. Patwari, Y. Wei, M. Parker, J. Parkin, **Z. Zhang**, A. Filippov. "Automatic contrast phase classification of polyphasic CT scans." Poster, *AACR Annual Meeting 2026*. Abstract: *Cancer Research* 86, Suppl. 7, 2777.
- M. Patwari, Y. Wei, M. Xu, **Z. Zhang**, K. Sidiropoulos, B. Selvaraj, G. Hughes, et al. "Fast, interactive, AI-assisted 3D lung tumour segmentation." Poster, *AACR Annual Meeting 2024*. Abstract: *Cancer Research* 84, Suppl. 6, 887.
- Y. Wei, B. Selvaraj, M. Patwari, Q. Li, M. Xu, K. Sidiropoulos, **Z. Zhang**, et al. "Improving non-small cell lung cancer segmentation on a challenging dataset." Poster, *AACR Annual Meeting 2023*. Abstract: *Cancer Research* 83, Suppl. 7, 5427.
- L. Fedden, **Z. Zhang**, K. Baykaner, Q. Li, L. Bordeaux. "DIME-CT: Self-supervised learning for medical image analysis using patch-based embeddings." Poster, *AACR Annual Meeting 2022*. Abstract: *Cancer Research* 82, Suppl. 12, 1937.

SKILLS

Programming: Python; SQL; some C++

ML / AI: PyTorch, Lightning, Weights & Biases, Hugging Face, scikit-learn, UMAP, MLX, Ollama

Medical AI / Imaging: MONAI, nnU-Net, 3D Slicer, DICOM, pydicom, NiBabel, NIfTI, ITK / SimpleITK

Infrastructure & MLOps: Docker, Kubernetes, AWS, Ray, Slurm, CI/CD, model serving, multi-GPU / distributed training, GPU optimization, Ollama / local inference, Proxmox / KVM (PCIe passthrough), Git

EDUCATION

M.S. Bioinformatics — University of Michigan, Ann Arbor, MI 08/2016 – 04/2018

B.S. Biomedical Sciences — University of Macau, Macau SAR, China 08/2012 – 06/2016