

ZHENNING ZHANG

LinkedIn: [linkedin.com/in/zhenning-zhang](https://www.linkedin.com/in/zhenning-zhang) | Portfolio: zhenningzhang.net

PROFESSIONAL SUMMARY

Machine Learning Engineer with 5+ years at AstraZeneca building deep learning systems for drug discovery and medical imaging. 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. M.S. Bioinformatics, University of Michigan.

EXPERIENCE

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

Drug Discovery ML

- Built multimodal deep learning models for IC50 prediction in PyTorch, fusing cell painting microscopy with SMILES chemical representations across compound libraries totaling 39K compounds.
- Designed and deployed multi-channel image analysis pipelines using Graph Neural Networks to capture spatial cell-cell relationships in multiplex immunofluorescence for downstream phenotype classification.

Medical Imaging & Segmentation

- Built an interactive Transformer-backed segmentation toolset for 3D volumetric data with text-guided prompts; halved annotation time and deployed to 12+ internal users across R&D.
- 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.

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 datasets totaling 150K CT volumes ingested, with automated mapping of thousands of annotation masks across diverse file standards (DICOM, NIfTI).
- Built web-based visualization suites for high-dimensional embedding interpretation (heatmaps, histograms, archetype analysis) used by R&D labs to interpret model outputs.
- Maintain and upgrade multiple internal deep learning repositories supporting cross-team R&D workflows.

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.
- Developed web service frameworks for cohort data management with interactive dashboards over complex distributions.
- Trained XGBoost and gradient boosting models for structured metadata classification.

Teaching Assistant — Arkansas State University AI Campus 11/2018 – 03/2019

- Mentored cross-functional student teams building applied AI systems: NLP keyword search, image captioning, and gradient boosting regression.

SELECTED PROJECTS

- **Multi-Agent Stock Analysis Pipeline.** Built a three-agent system (analyst → challenger → synthesizer) running on locally hosted Gemma via Ollama, with a Streamlit dashboard. Designed agent roles to surface

disagreement and adversarial critique before final synthesis. Optimized inference and memory across Apple Silicon and Linux GPU environments.

- **GPU Compute Environment Engineering.** Architected Proxmox / KVM virtualization with PCIe passthrough for direct GPU access in specialized ML workloads and cross-platform development.

SELECTED 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).

Selected 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 (Expert), C++, R

ML / AI: PyTorch (Expert), Keras, scikit-learn, Hugging Face, HDBSCAN, UMAP, MLX, Ollama

Medical AI / Imaging: MONAI, nnU-Net, DICOM, NIfTI, ITK / SimpleITK

Infrastructure & MLOps: Docker, Kubernetes, Git, Proxmox, KVM, Dash, Agile

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