Qifeng Zhou

Arlington, TX | qfchou@gmail.com | qfchou.github.io | Linkedin

Education

University of Texas at Arlington, *Ph.D. in Computer Science (GPA 4.0/4.0)* Sept. 2022 – Present
- Supervisor: Dr. Junzhou Huang

Zhejiang University, *BS in Chu Kochen Honors College (Rank Top 5%)* Sept. 2018 – June 2022

Selected Publications

1 **Q. Zhou** et al., "Contrastive Pretraining for Computational Pathology With Visual Language Models", **ISBI 25**.
2 **Q. Zhou** et al., "PathM3: A Multimodal Multi-Task MIL Framework for WSI Classification and Captioning", **MICCAI 24**.
3 **Q. Zhou** et al., "Histopathology Omni-modal Embedding for Pathology Composed Retrieval", **Under Review**.
4 T. Dang, ..., **Q. Zhou** et al., "Hierarchical Alignment Gene-Enhanced Pathology Representation Learning", **MICCAI 25**.
5 J. Jonnagaddala, ..., **Q. Zhou** et al., "Multimodal analysis of whole slide images in colorectal cancer", **npj DM**.
6 T. Dang, ..., **Q. Zhou** et al., "MFMF: Multiple Foundation Model Fusion Networks for WSI Classification", **BCB 24**.

Experience

Digital Pathology Intern | Genmab | Princeton, NJ Jun. 2025 - Aug. 2025
- Developed a multimodal survival analysis framework that fuses **Whole Slide Images (WSI), Omics, and Clinical data**. Implemented **Cross-Attention** and dynamic fusion strategies (Early/Late) to combine these diverse data types.
- Designed a deep learning pipeline to predict **Folate Receptor Alpha (FRα)** directly from H&E slides, enabling non-invasive biomarker discovery without the need for expensive chemical staining.
- Built and optimized large-scale model training pipelines on **AWS and Databricks**, efficiently processing **TB-scale** multimodal datasets using Python and PyTorch.

Research Projects

Omni-modal MLLM for Pathology Composed Retrieval Sept. 2025 – Present
- Engineered a unified retrieval system that processes interleaved inputs (e.g., Image + Text, Multi-image) to enable complex queries (e.g., "Find tissue like Image A but with Feature B"), defining the Pathology Composed Retrieval task.
- Transformed **Qwen2.5-VL** for retrieval tasks using a two-stage contrastive learning framework to resolve task mismatch. Implemented **Native Resolution encoding** and pathology stain to capture fine-grained morphological.
- Achieved SOTA performance on the Pathology Composed Retrieval benchmark, outperforming existing models (MUSK, CONCH) by over 30% by effectively bridging the modality gap.

MLLM-based Foundation Model for Pathology Vision-language Embedding Sept. 2024 - Apr. 2025
- Built a vision-language foundation model based on **LLaVA-Next**, surpassing pathology CLIP methods (PLIP, QuileNet, PathCLIP) over **10 zero-shot tasks** and **10 datasets**.
- Designed specific prompts to guide the MLLM for representation learning and used LoRA to fine-tune MLLM with over 590,000 pathology image-text pairs. Published this work at **ISBI 2025**.

Generative Cross-Modal Learning for Spatial Transcriptomics Dec. 2024 - Mar. 2025
- Developed a framework to predict spatial transcriptomics gene expression directly from H&E images. Designed a **Gene-Informed Image Encoder** enabling the model to learn gene-gene co-regulation patterns within image features.
- Proposed a **hierarchical alignment** strategy using clustered contrastive learning to align histological phenotypes with molecular profiles, achieving 18% MSE reduction on HER2+ cancer datasets. Published at **MICCAI 2025**.

Abnormality-Aware Multimodal foundation model integrating for WSI classification Apr. 2024 - July 2024
- Built a framework by fusing three foundation models: CONCH (Image), Segment any cell (Cell), and Quilt-LLaVA (Text). Introduced a three-step cross-attention module to integrate multimodal features and an abnormality-aware module based on auto-encoder to identify abnormal instances.
- Achieved SOTA performance with **0.99 AUC** on the SLN-Breast dataset and **0.98 AUC** on TCGA-Lung cancer subtyping, significantly outperforming unimodal baselines. Published at **ACM BCB 2024**.

Multi-instance image-text learning for WSI classification and captioning Nov. 2023 - Feb. 2024
- Developed a multimodal, multi-task MIL framework for Gigapixel WSI classification and captioning, improving **5% accuracy and 0.2 BLEU scores** over SOTA methods in gastric adenocarcinoma dataset.
- Designed a **correlation module** to reduce redundancy and a **Query-Based Transformer** to align visual features with diagnostic texts, bridging a ViT image encoders with a Large Language Model (Flan-T5), resulting in a publication at **MICCAI 2024**.

Skills

Languages/Frameworks: Python, C/C++, PyTorch, Scikit-learn, OpenCV, Hugging Face, Git, Docker, AWS, Databricks
Core Competencies: Multimodal LLM (MLLM), Vision-Language Pretraining, Contrastive Learning, MIL, Generative AI