

Tian XIA

+1 609 375 5853 @ steveissummer@gmail.com in linkedin.com/in/tian-xia-7a6639148 github.com/no1summer

- Expertise in data analysis and image analysis
- Proficient in acquisition, processing and analysis of 3D image and video
- Proficient in developing ML/DL models in both industrial and academic environments

EDUCATION

- | | |
|-----------|---|
| 2019-2024 | Baylor College of Medicine
Ph.D. Quantitative & Computational Biosciences |
| 2015-2019 | Zhejiang University
B.S. Pharmaceutical Sciences, GPA 3.97 |

EXPERIENCE

- | | |
|----------------------|---|
| Now
July 2025 | Yale University - PREV UTHealth - Postdoc
<i>Center for AI and Genome Informatics – MENTOR : DR. DEGUI ZHI</i> <ul style="list-style-type: none">> Whole Slide Imaging (WSI)/Pathology :<ul style="list-style-type: none">- Develop a multimodal fusion approach leveraging WSI-guided genomic cross-attention- Evaluate the impact of various foundation models and design choices on WSI and genomic embeddings for multimodal fusion> MRI :<ul style="list-style-type: none">- Developing a longitudinal VAE-derived MRI endophenotype that better reflects Alzheimer Disease drug treatment response than amyloid PET intensity and correlates more strongly with cognitive outcomes- Developing a AI-agent system to automatically identify optimal feature extraction models for brain MRI in imaging genetics research- Optimizing the best backbone for feature extraction on brain MRI surface meshes with GCN- Developed a pipeline leveraging contrastive learning MocoV2 and cross-attention to learn mutual imaging features from multimodal T1/T2 MRI, showing superior performance in genetic discovery than single modality (<i>under review by Nature Biomedical Engineering</i>)- Build an arena for benchmark the performance of different deep learning models in imaging genetic discovery, including CNN, ViT (accepted by the NeurIPS 2025 Workshop), Swin Transformer, MAE, VAE, and other self-supervised models like DINOv3 (in preparation for Nature Methods)- Developed a multi-modal predictive model for Alzheimer's disease using features extracted from MRI images, polygenic risk score and multi-omic datasets, showing SOTA performance- Familiar with multi-GPU optimization of 3D computer vision architectures. <div><div>Deep learning</div><div>Multi-modality</div><div>Multi-organ</div><div>Multi-omics</div><div>Contrast learning</div><div>Vision Transformer</div><div>MRI</div><div>MONAI</div></div> |
| Jun 2024
Jan 2024 | Merck - CBGX CAMBRIDGE SITE - Co-op/intern
<i>Merck Research Laboratories – MENTOR : DR. REBECCA SENFT</i> <ul style="list-style-type: none">> Cell Image Analysis :<ul style="list-style-type: none">- Developed an image processing and analysis pipeline for the Optical Pooled Screening, including segmenting cells using deep-learning based method and quantify cell feature using CellProfiler followed by ML classification algorithm to classify cells with different treatments- Fine tuned self-supervised transformer-based deep learning method based on DINOv2, reaching 3% accuracy increase compared with the previous method with proper image preprocess- Automated the screening workflow with Nextflow, integrated MLOps practices for reproducible pipeline execution, version control, model tracking, and scalable deployment on High Performance Computing (HPC).> RNA-seq Data Tissue Region Identification :<ul style="list-style-type: none">- Applied supervised machine learning logistic regression with optimizing hyperparameter tuning. Generally it distinguished between different regions of patient sample well, except those very close to each other <div><div>Machine Learning</div><div>Deep Learning</div><div>LLM</div><div>Scikit-learn</div><div>Pytorch</div><div>Nextflow</div><div>HPC</div><div>MLOps</div></div> |

Nov 2024 Dec 2019	Baylor College of Medicine - LARINA'S LAB IMAGING SCIENCE - Graduate Student <i>Department of Integrative Physiology – MENTOR : DR. IRINA LARINA</i> > 3D Segmentation with Vision Transformer : - Fine-tuned pretrained 3D Swin Transformer to quantify follicle volumes during the mouse ovulation process, but given the training image modality difference, the performance is limited. - Built a semi-supervised deep learning model, utilized both labeled and unlabeled data, and introducing cross-attention block and contrast learning to the classic V-net backbone, reaching >90% Dice similarity coefficient with 10% of labeled data. > Developed a quantitative imaging method of cilia metachronal wave in mouse fallopian tube with optical coherence tomography in vivo (Published at Optica, IF=10.4). > Established a dynamic image signal processing procedure to track spermatozoa movement toward the egg (Invited oral presentation at SPIE, 2022). Computer Vision Deep Learning Object Detection Segmentation Pose Recognition CNN RNN GCN Fourier Transform Phase Analysis
Dec 2023 Jan 2020	Rice University - CLASS MACHINE LEARNING - Visiting Student <i>Department of Computer Science – DATA SCIENCE PROJECT, STATISTICAL MACHINE LEARNING</i> > Constructed a machine learning pipeline to identify genomic signatures in age-related macular degeneration. Reduced the number of feature from >18000 to <100 using feature selection techniques, including <i>minimum Redundancy Maximum Relevance, Random Forest, Generalized Linear Model, Principle Component Analysis, Statistical Test</i>. > Built a Python package for identifying the possible genes related to the disease from machine learning feature selection perspective. > Created a classification network (customized Resnet50 with Ensemble strategy) to distinguish fine-grained food images. > Construct a generalized linear model to identify individuals with the high risk of stroke with more than 90% accuracy. Data Science Machine Learning Computational Biology Pandas Scikit-learn R Version Control Data Visualization

SKILLS

Programming	Python, R, MATLAB, C, Bash, SQL, Git, LaTeX, AWS Cloud Computing, docker
ML/DL	Scikit-Learn, PyTorch, Pandas, AnnData, Seurat
Imaging Analysis	OpenCV, Numpy, Scipy, ImageJ, Scikit-Image, Matlab Image Processing Toolbox
Data Visualization	Matplotlib, Seaborn, ggplot2
Research	Data Science, Machine Learning, Deep Learning, Imaging Processing and Analysis, Computational Biology
Soft Skills	Critical Thinking, Active Learning, Time Management, Communication

PUBLICATIONS

Genetic lens for reproducibility of self-supervised learning derived imaging endophenotypes across data sets 2026 Tian Xia, Saiful Sheikh Muhammad Islam, Ziqian Xie, Xingzhong Zhao, Degui Zhi Nature Methods in review Genetic Analysis Deep learning MRI	
Learning heritable multimodal brain representation via contrastive learning 2026 Tian Xia, Xingzhong Zhao, Saiful Sheikh Muhammad Islam, Kamil Khan Mohammed, Ziqian Xie, Degui Zhi Science Advances in review Deep learning Constrast learning MRI Genetic Discovery	
Vision transformer autoencoders captures local and non-local features in brain imaging to reveal novel genetic associations 2025 Samia Islam*, Tian Xia*, Ziqian Xie, Degui Zhi Nature Communication Biology; NeurIPS 2025 Workshop Imageomics Deep learning MRI Image Analysis GWAS	
In vivo volumetric depth-resolved imaging of cilia metachronal wave with dynamic optical coherence tomography 2023 Tian Xia, Kohei Umezue, Deirdre Scully, Shang Wang, Irina Larina Optica Imaging Processing Spatial and Temporal Imaging Dynamic Signal Processing Fourier Transform Phase	