Tian XIA
Machine Learning Scientist

+1 609 375 5853
tianx@bcm.edu
linkedin.com/in/tian-xia-7a6639148
Personal website
github.com/no1summer
2300 Old Spanish Trl, Houston, TX 77054 - USA

I am currently a cross-disciplinary researcher with expertise in machine learning, imaging analysis, and computational biology. My Ph.D. research focused on developing quantitative measurements for 3D+time images. It equips me with the knowledge and experience of image acquisition, processing and analysis in both temporal, spatial and frequency domain. With a solid background in math, physics, computer science and biology, I have the skills and expertise to develop and apply Machine Learning/ Deep Learning to address biological questions and beyond. I have developed several models that have been deployed in both industrial environment or for academical purpose, in a wide range of regression, classification, clustering, segmentation, and more.

SKILLS

Programming	Python, R, MATLAB, C, Bash, SQL, Git, LaTex
Data Science	Pandas, Scikit-Learn, PyTorch, 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

EXPERIENCE

Jun 2024 – Jan 2024 | **Merck - CBGx | CAMBRIDGE SITE - Co-op/intern**
Merck Research Laboratories – MENTOR : DR. REBECCA SENFT

- Fine tuning self-supervised **transformer**-based **deep learning** image feature extraction method DINOv2 for classifying cells with different treatment
- Classify RNA-seq data of patients' for different locations with **machine learning**
- Developed a **machine learning** pipeline for determining the suitable marker for Optical Pooled **CRISPR Screening** (OPS)
- Established a CellProfiler and Cellpose (deep learning) pipeline to segment and quantify the cell images, which is integrated with Nextflow and deployed to High Performance Clustering (HPC).
- Experience with processing **scRNA-seq** data from fastq file
- Build and train **GPT2** from scratch and generate Shakespeare-like dialogue
- Explore different methods for integration of **multiomic/multimodality** data including scRNA-seq, ATAC-seq and imaging data.
- Fine tuning **scGPT** for multiomic/multimodality integration

Multiomic/Multimodality Integration | Machine Learning | Deep Learning | LLM | Scikit-learn | Pytorch | Nextflow | HPC

Nov 2024 – Dec 2019 | **Baylor College of Medicine - LARINA'S LAB | IMAGING SCIENCE - Graduate Student**
Department of Integrative Physiology – MENTOR : DR. IRINA LARINA

- Constructed an **3D segmentation** procedure by fine-tuning pretrained **3D Swin Transformer** to quantify follicle volumes during the mouse ovulation process.
- Build an **semi-supervised deep learning** model for 3D segmentation, by introducing cross-attention block and mutual learning to the classic V-net backbone.
- Developing AI coach for tennis player based on **pose recognition** and **graph convolution network** from self-recording videos.
- 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

Department of Computer Science – Data Science Project, Statistical Machine Learning

- 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 *minimum Redundancy Maximum Relevance, Random Forest, Generalized Linear Model, Principle Component Analysis, Statistical Test*.
- 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

Apr 2019 – Sep 2018

Massachusetts Institute of Technology - Weinberg's Lab | Cancer Biology - Research Assistant

Department of Biology – Mentor : Dr. Robert Weinberg

- Help build a genetically defined syngeneic mouse model of ovarian cancer (Published at **Cancer Discovery**)

Molecular Biology | Gene Editing | CRISPR | Drug Testing

Jul 2018 – May 2018

Princeton University - Kang's Lab | Cancer Biology - Research Assistant

Department of Biology – Mentor : Dr. Yibin Kang

- Study the phenotype of mir200 knockout in the mouse model by Immunochemistry (IHC) imaging.

Immunohistochemistry Imaging | Statistical Analysis

Feb 2018 – Jul 2017

Zhejiang University - Shao's Lab | Cancer Biology - Research Assistant

School of Medicine – Mentor : Dr. Jimin Shao

- Study the IL-6, p-stat3, Fra-1, Nanog pathway in the progression and metastasis of colon cancer by immunofluorescence (IF) imaging. (Published at **Oncogene**)

Immunofluorescence Imaging | Cancer Research

Education

2024 **Baylor College of Medicine**
Ph.D. Quantitative & Computational Biosciences
Concentration : Image Processing and Analysis, Machine/Deep Learning, Multimodality Integration

2019 **Zhejiang University**
B.S. Pharmaceutical Sciences, GPA 3.97
Concentration : Molecular Biology, Cancer Biology, Computational Biology, Drug Discovery

Publications

In vivo volumetric depth-resolved imaging of cilia metachronal wave with dynamic optical coherence tomography 2023
Tian Xia, Kohei Umezu, Deirdre Scully, Shang Wang, Irina Larina
Optica
Imaging Processing | Spatial and Temporal Imaging | Dynamic Signal Processing | Fourier Transform | Phase

Dynamic volumetric imaging and cilia beat mapping in the mouse male reproductive tract with optical coherence tomography 2022
Kohei Umezu, Tian Xia, Irina Larina
Biomedical Optics Express
Imaging Analysis | Volumetric 3D Imaging | Dynamic Signal Processing | Reproductive Biology

Tracking spermatozoa movement toward the egg with functional optical coherence tomography 2022
Tian Xia, Kohei Umezu, Shang Wang, Irina Larina
Dynamics and Fluctuations in Biomedical Photonics XIX
Object Detection | Dynamic Signal Processing | Denoising

The inflammatory cytokine IL-6 induces FRA1 deacetylation promoting colorectal cancer stem-like properties 2019
Tingyang Wang, Ping Song, Tingting Zhong, Xianjun Wang, Xueping Xiang, Qian Liu, Haiyi Chen, Tian Xia, ..., Riccardo Fodde, Jimin Shao
Oncogene
Cancer Immunology Pathway | Immunofluorescence Imaging | Imaging Analysis

Position of Responsibility

2024-present | **BOE - Biomedical Optics Express**
Invited Reviewer
› Top reviewer with high reviewer score and fast response

2023-present | **JOSA A - Journal of the Optical Society of America A**
Invited Reviewer
› Top reviewer with high reviewer score and fast response

2022-present | **Cats of Houston - Stray Cat Adoption Platform**
Co-founder and Photographer
› Photograph and post kitten for adoption to reduce stray cat in Houston. Help more than 50 kittens to be adopted.

Achievements & Recognitions

2023 **Travel Award for Invited Talk** SPIE Photonic West 2023, San Francisco
2022 **Second Place for Poster Presentation** Texas Forum of Reproductive Sciences 2022, Houston
2018 **First-Class Scholarship for Outstanding Students** (Top 1%) Zhejiang University
2018 **The President's Scholarship** Zhejiang University
2018 **Championship of Men's Singles** Tennis Competition, Zhejiang University

Public Media Exposure

Imaging advance poised to provide new insights into reproduction and infertility 2023
Optica News Phys Org AAAS

Researchers develop new OCT method to directly image cilia dynamics in living organisms 2023
Medical Life Sciences News

OCT for In Vivo Imaging of Cilia Dynamics 2023
Optics and Photonics News

Team develops imaging method to capture previously inaccessible coordination of tiny hair-like cilia 2023
From the Labs at Baylor College of Medicine Twitter

OCT-based approach examines human physiology 2023
Laser Focus World