

YUFENG WANG

+1 6315385209 | airscker@gmail.com | www.linkedin.com/in/yufeng-wang-142a3a263/

EDUCATION

Stony Brook University, New York

Aug 2023 – Present

Computer Science Department, Ph.D. Student, Supervised by [Prof. Haibin Ling](#).

University of Science and Technology of China (USTC), Hefei, China

Sep 2019 – Jul 2023

Physics Department, Bachelor of Science.

EXPERIENCES

Stanford University, Center for Artificial Intelligence in Medicine and Imaging (AIMI), *ML Research Assistant*.

- **ML/DL Pipeline Development:** Developed deep learning pipelines for analyzing large clinical and medical image datasets, aimed at diagnosing heart disease.
- **Multi-Domain Collaboration:** Collaborated with various labs on study design, data collection, and report writing, effectively communicating and visualizing quantitative analysis results to interdisciplinary collaborators.

2025 IEEE Symposium on Computational Intelligence in Natural Language Processing and Social Media, *Reviewer*.

- **Scientific Literature Review:** Contributing to the evaluation and advancement of cutting-edge research in NLP.

SKILLS

Proficient: Python, R, C, PyTorch, Docker, HTML, CSS; MATLAB, Wolfram Mathematica.

Prior Experience: C++, CUDA, TensorFlow 2.0; SolidWorks, SPSS, COSMOL.

SELECTED PROJECTS

AI-enhanced Cardiac Magnetic Resonance Imaging and Cardiovascular Disease Diagnosis (Computer Vision, Medical Imaging, Image Processing, Statistics, Pattern Recognition)

Mar 2022 – July 2023

Stanford University, University of Science and Technology of China

- Processed 80,000 *Cardiac Magnetic Resonance Image* (CMR) sequences (including short-axis and four-chamber views) from datasets spanning the UK, China, and the USA, for AI model training and analysis.
- Implemented the *nnU-Net* model for semi-supervised segmentation of the heart region of interest (ROI) in CMR images, achieving a Dice Coefficient exceeding 95%.
- Developed the entire deep learning model training pipeline using the *Video Swin Transformer* (VST) model as the backbone to detect cardiac anomalies from CMR ROI images, achieving a top F1-score of 99.1% on the hold-out testing set.

Diffusion Model Based X-ray Absorption Spectrum - Protein Structure Interpretation (Computer Vision, Statistics, Image Processing, Diffusion Model)

Mar 2023 – Present

Brookhaven National Laboratory, Stony Brook University

- Constructed a database of *X-ray Absorption Spectra* (XAS) and corresponding protein structures from scientific literature, providing a valuable dataset for training deep learning models to learn the relationships between XAS and protein structures.
- Explored *Graph Neural Network* (GNN) and *Transformer*-based models to generate *X-ray Absorption Spectra* (XAS) from crystal periodic structures obtained from the Materials Project, with the goal of predicting XAS for protein structures in our dataset using transfer learning.
- Currently developing a *Diffusion Model*-based multi-modal approach to reconstruct protein structures using XAS spectra from our dataset.

Multi-modal AI-enabled Materials Informatics Collection for Structure-Property Relationships (Natural Language Processing, Computer Vision, Pattern Recognition, Physics Simulation)

Oct 2023 – Present

Stony Brook University

- Fine-tuned *YOLOv6* and *BERT/LLAMA2* models to recognize and annotate curve plots in literature and extract textual molecular information from scientific reports, creating a comprehensive dataset that integrates both visual and textual data.
- Generated a precise simulated database for various organic materials and their electrochemical properties using *Density Functional Theory* (DFT), integrating this database with the database extracted from literature for AI model training.
- Developed a *Graph Neural Network* (GNN) model to predict electrochemical properties using the combined dataset, achieving an R^2 coefficient exceeding 99.1% between ground truths and GNN predictions.

PUBLICATIONS

1. **Screening and diagnosis of cardiovascular disease using artificial intelligence-enabled cardiac magnetic resonance imaging.** *Nature Medicine* (2024): 1-10.
2. **An X-ray absorption spectrum database for iron-containing proteins.** *Nature Scientific Data*, Under Review.