

Education

- August 2015 **Northwestern Polytechnical University**, Xi'an, China
– July 2019 B.Eng. in Information Engineering
- August 2019 **University of Michigan**, Ann Arbor, MI
– April 2021 M.Eng. in Electrical and Computer Engineering
- August 2021 **Rensselaer Polytechnic Institute**, Troy, NY
– Present Ph.D (Expected) in Electrical Engineering, Computer, and Systems Engineering

Publications and/or Manuscripts

- January 2025 **Theoretical Analysis of Contrastive Learning in Vision-Language Model Pre-Training: The Role of Synthetic Text Captions for Feature Alignment**
Jiawei Sun, Shuai Zhang, Hongkang Li, Meng Wang
- Sep 2023 **How do skip connections affect Graph Convolutional networks with graph sampling? A theoretical analysis on generalization**
Jiawei Sun, Hongkang Li, Meng Wang
- April 2020 **Modelling learning in *C. elegans* chemosensory and locomotive circuitry for T-maze navigation**
Bennet G. Sakelaris, Zongyu Li, Jiawei Sun, Shurjo Banerjee, Victoria Booth, Eleni Gourgou

Research Experiences

- January 2024 **Contrastive Learning with Data Misalignment: Feature Purity, Training Dynamics and Theoretical Generalization Guarantees,**
–January 2025 Advisor: Associate Professor [Meng Wang](#)
Theoretical analysis of VLM training with nonlinear activations, highlighting the role of synthetic captions
- Provided the first theoretical analysis of VLM training dynamics with synthetic captions
 - Demonstrated that synthetic captions reduce spurious correlations and enhance generalization and zero-shot performance
 - Validated theoretical insights through experiments on BLIP and other state-of-the-art VLMs
- Sep 2021 **Deep Learning Theory of Graph Convolutional networks,**
– Sep 2023 Advisor: Associate Professor [Meng Wang](#)
Theoretical analysis on two-hidden-layer GCNs with skip-connections
- Present the first generalization analysis of two-hidden-layer GCNs with skip connections using graph sampling
 - Reveal that skip connections lead to different sampling requirements across layers
 - Validate theoretical results on deep GCNs for benchmark datasets

- Sep 2019 **Image Processing to decipher C. elegans locomotion in mazes,**
- April 2020 Advisor: Research Scientist [Eleni Gourgou](#)
- Working on finding the motion trails of the elegans
- Use the Chan-Vese active contour method and SVD to extract contours of the maze
 - By the Procrustes Transformation method, a T-shape polygon is rotated and shifted to have maximal overlap with the extracted contour
 - Apply the Frame Difference method to find motion trails of the elegans
 - Use the K-Nearest Neighbor (KNN) algorithm to smooth the motion trails
- Jan 2020 **Deep Neural Network for Spectrum Unfolding,**
- April 2020 Advisor: Professor [Alfred Hero](#)
- Working on Recurrent Neural Network (RNN) algorithm
- Propose the RNN architecture that mimics project gradient descent method from optimization theory
 - Complete the Recurrent Neural Network code by Pytorch
- March 2018 **Differential Microphones Arrays based on Differential Equation,**
- June 2018 Advisor: Professor [Jie Chen](#) and Professor [Lijun Zhang](#)
- Worked on Differential Microphones Arrays based on Differential Equation
- Proved that the polynomial of sinusoidal function is the solution of a differential equation and the differential equation corresponding to LDMA and CDMA are same
- May 2017 – **Distributed PCA by the Primal-Dual Method of Multipliers (PDMM),**
- October 2017 Advisor: Professor [Jie Chen](#)
- Worked on Distributed Optimization Algorithm
- Distributed PCA method can be obtained by simply approximating the global correlation matrix via the Average Consensus Algorithm subroutine, so matrices are divided in columns
 - Eigenvalue decomposition of the correlation matrix and reduced its dimension to p-dim by PDMM algorithm
 - Programmed in Matlab to accomplish Distributed PCA

Selected Awards and Honors

- November 2018 **Honorable Mention of the International Mathematical Contest in Modeling**
- November 2017 **First Prize Scholarship,** Northwestern Polytechnical University
Top 15% in 200 students
- November 2016 **National Scholarship,** Northwestern Polytechnical University
0.2% national wide)

Selected Course Projects

- March 2018 **Communication System Design**
Achieved communication between two computers. Achieved source coding by ASIC code and adopted 2FSK modulation based on MATLAB
- January 2018 **Development of Microphone Orientation System**
Accomplished acoustic localization by Conventional Beamforming method and Direction of Arrival (DOA) location method

Skills

Programming Languages: Matlab, Julia, Python, \LaTeX , PyTorch