Xiaoguang Zhu

apollozxg@gmail.com · +1 (408) 931-0061 · https://xiaoguangzhu.org/

Education

Shanghai Jiao Tong University, Shanghai — Apr. 2017 – Jun. 2022

PhD Information and Communication Engineering
Mentor: Prof. Peilin Liu.

Shanghai Jiao Tong University, Shanghai — Sep. 2014 – Mar. 2017

Master Electronics and Communication Engineering
Mentor: Prof. Peilin Liu.

Shanghai Jiao Tong University, Shanghai — Sep. 2010 – Jun. 2014

Bachelor Electronic Science and Technology

Professional Experience

University of California, Davis, Davis — April 2024 – Present

Postdoctoral Scholar, DataLab: Data Science and Informatics.
Mentors: Prof. Vladimir Filkov (Mar. 2025 – Present), Prof. Chen-Nee Chuah (Apr. 2024 – Feb. 2025), Prof. Uma Srivatsa, and Prof. Nipavan Chiamvimonvat.

Research Interest

Causal Deep Learning for Health, Multimodal Learning, Multimodal Large Language Models, Representation Learning, Model Generalization, Action Recognition, EEG Signal Recognition

Research Experience

Causal Multimodal Learning for Robust Medical AI — April. 2024 – Present

This research introduces causal principles into multimodal learning to improve robustness in medical AI tasks such as diagnosis, disease progression, and medical reasoning.

• Proposed a causal debiasing framework for medical multimodal representation learning with missing modalities, tackling the challenges of distribution bias and missingness bias
• Developed Causal Hidden Markov Model for interpretable disease progression prediction using hyperbolic latent space modeling for Alzheimer's disease
• Designed a counterfactual-based DPO-based Reinforcement Learning framework for MLLM medical VQA, improving the reliability of diagnostic reasoning

Foundation Model Adaptation in Low-Data Medical Settings — April. 2024 – Present

This line of work focuses on adapting vision-language foundation models to clinical scenarios with limited supervision and strong domain gaps.

• Source-free domain adaptation with curriculum-guided reliable pseudo-labeling
• Adaptive weighted parameter fusion for CLIP in class-incremental learning, mitigating forgetting and preserving discriminability
• Led a systematic survey of MLLMs in medicine and nursing, categorizing architectures, datasets, and applications for real-world deployment guidance

Deep Learning for Cardiac Disease — April. 2024 – Present

This research aims to classify cardiac disease subtypes and evaluate the treatment effects. To identify the pathway from multimodal data that triggers cardiac diseases, such as atrial fibrillation, causal discovery and causal deep learning methods are proposed.

• Deep causal discovery and causal inference with metabolomic data for atrial fibrillation.

• Treatment effect estimation in medication treatment analysis
• Disease pathway analysis based on multimodal data

Multimodal Representation Learning
Dec. 2017 – Present

To address the generalization performance degradation problem in open-world image retrieval, graph-based network and uncertainty-aware metric learning methods are proposed for representation learning for image retrieval. Moreover, mentor postgraduates in the NSAI lab of Fudan University to work on anomaly detection research.

• Content-based image retrieval
• Cross-modal image retrieval
• Metric learning for image retrieval

EEG Signal Recognition
Sep. 2021 – Sep. 2023

Deep learning models of graph neural networks and neural stochastic differential equations are proposed for classifying EEG signals. The state-of-the-art performance is achieved on public benchmarks. Causal models are used to analyze causal relationships through EEG signal channels to enhance generalization ability through individual people.

• Deep models for EEG signal recognition
• Casual inference-based domain generalization for EEG models

Human Action Analysis
Jul. 2018 – Jun. 2024

A novel multi-modal fusion network is proposed to take advantage of the complementarity of the skeleton and RGB modalities, which achieves competitive accuracy over SOTA but with lower computation. Deep generative models and domain adaptation methods are proposed to enhance the generalization capability of action recognition models.

• Multi-modal video action recognition
• Anomaly detection for action recognition
• Domain adaptation for action recognition

GNSS Receiver Test and Evaluation
Jul. 2015 – Dec. 2016

Implemented an algorithm to replay high-fidelity GNSS signals by modeling the error between real signals and recorded signals. Introduced a test method for GNSS receiver evaluation on various scenarios.

• GNSS test and evaluation method with high-fidelity playback
• Machine learning method for scenarios classification with GNSS signals

ARTICLES

• **Xiaoguang Zhu***, Cayde Wang, Kartik Patwari, Lianlong Sun, Chen-Nee Chuah, Chengxin Pang. Causality-Aware Direct Preference Optimization for Aligning Medical Vision Language Models. ***2026 Conference on Computer Vision and Pattern Recognition (CVPR)***. Under review.
• **Xiaoguang Zhu***, Uma Srivatsa, Nipavan Chiamvimonvat, Vladimir Filkov. Graph-Based Neural Controlled Diffusion Equations for Atrial Fibrillation Outcome Prediction Using Intracardiac Electrogram. ***2026 IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)***. Under review.
• Dongjie Chen, Kartik Patwari, Zhengfeng Lai, **Xiaoguang Zhu**, Sen-ching Samson Cheung, Chen-Nee Chuah. Empowering Source-Free Domain Adaptation via MLLM-Guided Reliability-Based Curriculum Learning. ***2026 Winter Conference on Applications of Computer Vision (WACV)***. Accepted.
• **Xiaoguang Zhu***, Lianlong Sun, Yang Liu, Pengyi Jiang, Uma Srivatsa, Nipavan Chiamvimonvat, Vladimir Filkov. Causal Debiasing Medical Multimodal Representation Learning with Missing Modalities. ***IEEE Transactions on Knowledge and Data Engineering***. Under review.
• Yang Liu, Hongjin Wang, Zepu Wang, **Xiaoguang Zhu**, Jing Liu, Peng Sun, Rui Tang, Jianwei Du, Victor C.M. Leung, Liang Song. CRCL: Causal Representation Consistency Learning for Anomaly Detection in Surveillance Videos. ***IEEE Transactions on Image Processing***, Volume 34, April 2025, Pages 2351–2366.
• Yang Liu, Siao Liu, **Xiaoguang Zhu**, Hao Yang, Jielin Li, Juncen Guo, Liangyu Teng, Dingkang Yang, Yan Wang, Jing Liu. Privacy-Preserving Video Anomaly Detection: A Survey. ***IEEE Transactions on Neural Networks and Learning Systems***, 2025.

• Juncen Guo, **Xiaoguang Zhu***, Liangyu Teng, Hao Yang, Jing Liu, Yang Liu, Liang Song. Adaptive Weighted Parameter Fusion with CLIP for Class-Incremental Learning. ***IEEE International Conference on Multimedia & Expo (ICME)***, 2025.

• Shivam Rajendra Rai Sharma, **Xiaoguang Zhu**, Luca Cerny Oliveira, Kartik Patwari, La Rissa Vasquez, David Garcia, Louise Nicole C. Sevilla, Brittany N. Dugger, Chen-Nee Chuah. Benchmarking Parameter Efficient Adaptation of Vision Language Models on Pathology. ***NeurIPS 2025 Workshop for Imageomics***.

• Luca Cerny Oliveira, Kartik Patwari, **Xiaoguang Zhu**, Sen-Ching Cheung, Brittany Dugger, Chen-Nee Chuah. Co-HSF: Resource-Efficient One-Shot Semi-Supervised Adaptation of Histopathology Foundation Models. ***AAAI 2025 Spring Symposium on AI for Health***. Accepted.

• Jing Liu, Linxiao Gong, Juncen Guo, Jingyi Wu, Lianlong Sun, Yulai Bi, Kartik Patwari, Boan Chen, Lichi Zhang, Wei Zhou, Yang Liu, **Xiaoguang Zhu***, Chen-Nee Chuah, Bala Rajaratnam. Multimodal Large Language Models in Medicine and Nursing: A Survey. ***IEEE Reviews in Biomedical Engineering.*** Under review.

• Ye Zhu, **Xiaoguang Zhu**, Yuzhang Shang, Zhenghao Zhao, Yan Yan. Supplementing Missing Visions via Dialog for Scene Graph Generations. ***2024 IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)***, April 2024, Pages 3375–3379.

• Chengfang Li, **Xiaoguang Zhu**, Xinhua Zeng, Liang Song. Graph-based Networks with Channel Selection for EEG Signal Learning. ***Biomedical Signal Processing and Control***, Volume 87, January 2024, Pages 105524.

• Jing Liu, **Xiaoguang Zhu**, Yang Liu, Wei Zhu, Liang Song. Distributional and Spatial-Temporal Robust Representation Learning Based on Multimodal Sensing. ***Pattern Recognition***, Volume 140, August 2023, Pages 109568.

• Ziming Wang, Han Yu, **Xiaoguang Zhu**, Zengwen Li, Changxue Chen, Liang Song. Learning 3D Human Pose and Shape Estimation Using Uncertainty-Aware Body Part Segmentation. ***2023 IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)***, April 2023, Pages 1–5.

• Hanqi Wang, **Xiaoguang Zhu**, Tao Chen, Chengfang Li, Liang Song. Rethinking Saliency Map: A Context-Aware Perturbation Method to Explain EEG-Based Deep Learning Model. ***IEEE Transactions on Biomedical Engineering***, Volume 70, October 2022, Pages 1462–1472.

• Donglai Wei, **Xiaoguang Zhu**, Yang Liu, Jing Liu, Xinhua Zeng. MSAF: Multimodal Supervise-Attention Enhanced Fusion for Video Anomaly Detection. ***IEEE Signal Processing Letters***, Volume 29, October 2022, Pages 2178–2182.

• Di Li, **Xiaoguang Zhu**, Liang Song. Mutual Match for Semi-Supervised Online Evolutive Learning. ***Applied Intelligence***, May 2022.

• Jiameng Pan, **Xiaoguang Zhu**, Peilin Liu. Generating Adaptive Targeted Adversarial Examples for Content-Based Image Retrieval. ***2022 International Joint Conference on Neural Networks (IJCNN)***, July 2022, Pages 1–9.

• Donglai Wei, Chengeng Liu, Yang Liu, Jing Liu, **Xiaoguang Zhu**, Xinhua Zeng. Look, listen and pay more attention: fusing multi-modal information for video violence detection. ***2022 IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)***, May 2022, Pages 1980–1984.

• Yang Liu, Jing Liu, **Xiaoguang Zhu**, Donglai Wei, Xiaohong Huang, Liang Song. Learning task-specific representation for video anomaly detection with spatial-temporal attention. ***2022 IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)***, May 2022, Pages 2190–2194.

• Yang Liu, Jing Liu, Mengyang Zhao, Dingkang Yang, **Xiaoguang Zhu**, Liang Song. Learning appearance-motion normality for video anomaly detection. ***2022 IEEE International Conference on Multimedia & Expo (ICME)***, July 2022, Pages 1–6.

• **Xiaoguang Zhu**, Haoyu Wang, Peilin Liu, Zhantao Yang, Jiuchao Qian. Graph-based reasoning attention pooling with curriculum design for content-based image retrieval. ***Image and Vision Computing***, Volume 115, November 2021, Pages 104289.

• **Xiaoguang Zhu**, Haoyu Wang, Peilin Liu, Zhantao Yang, Jiuchao Qian. Skeleton sequence and RGB frame based multi-modality feature fusion network for action recognition. ***ACM Transactions on Multimedia Computing, Communications, and Applications***, Volume 18, Issue 3, October 2021, Pages 1–24.

• **Xiaoguang Zhu**, You Wu, Zhantao Yang, Peilin Liu. Uncertainty-aware domain adaptation for Action Recognition. ***2021 28th International Conference on Neural Information Processing (ICONIP)***, December 2021, Pages 494–506.

• **Xiaoguang Zhu**, Siran Huang, Wenjing Fan, Yuhao Cheng, Huaqing Shao, Peilin Liu. SDAN: Stacked diverse attention network for video action recognition. ***2021 IEEE International Symposium on Circuits and Systems (ISCAS)***, May 2021, Pages 1–5.

• **Xiaoguang Zhu**, Jiuchao Qian, Haoyu Wang, Peilin Liu. Curriculum enhanced supervised attention network for person re-identification. ***IEEE Signal Processing Letters***, Volume 27, September 2020, Pages 1665–1669.
• Zhantao Yang, **Xiaoguang Zhu**, Jiuchao Qian, Peilin Liu. Dark-aware network for fine-grained sketch-based image retrieval. ***IEEE Signal Processing Letters***, Volume 28, December 2020, Pages 264–268.
• Minting Pan, **Xiaoguang Zhu**, Yongfu Li, Jiuchao Qian, Peilin Liu. MRNet: A keypoint guided multi-scale reasoning network for vehicle re-identification. ***2020 27th International Conference on Neural Information Processing (ICONIP)***, November 2020, Pages 469–478.
• Guiyu Liu, Jiuchao Qian, Fei Wen, **Xiaoguang Zhu**, Rendong Ying, Peilin Liu. Action recognition based on 3d skeleton and RGB frame fusion. ***2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)***, November 2019, Pages 258–264.
• Junyu Dai, Jiuchao Qian, Zheng Tao, Junhong Chen, **Xiaoguang Zhu**, Huaqing Shao, Zheng Gong, Peilin Liu. A system integrating speech interaction and vision sensing applying in smart home scenario. ***2019 IEEE International Symposium on Circuits and Systems (ISCAS)***, May 2019, Pages 1–5.
• **Xiaoguang Zhu**, Xin Chen, Rendong Ying, Shanshan Zhan, Yuze Wang, Xiaoran Fang, Fei Su, Di He, Peilin Liu, Wenxian Yu. An innovative remote testing system for GNSS receiver based on IF sampled signal library of various scenarios. ***2015 International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015)***, September 2015, Pages 250–257.

⚙ Other Professional Experiences

Invited Reviewer: CVPR 2021, CVPR 2022, CVPR 2023, CVPR 2024, CVPR 2025, ICCV 2021, ICCV 2023, ICCV 2025, ECCV 2022, ECCV 2024, AAAI 2025, BMVC 2020, BMVC 2021, ACMMM 2025, IEEE Signal Processing Letters, ACM Computer Surveys, Artificial Intelligence Reviews
Technical Program Committee: 2021-2024 International Conference on Networking System of AI (INSAI)
Visiting Fellowship: Lab of networking system of AI, Fudan University, 2021-2022
Guest Editor: Special Issue "Advancing Action Recognition: Novel Approaches, Techniques and Applications" of the Journal of Imaging
Teaching Assistant: Digital Circuits, Shanghai Jiao Tong University, 2015

♡ Awards

Graduate National Scholarship	2015
RongChang Scholarship, Shanghai Jiao Tong University (10/year)	2017-2019
Shanghai Jiao Tong University Outstanding Student (5%)	2013, 2016, 2020
Outstanding Graduates of Shanghai	2022