

Wenhui Zhu

School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ

✉ wzhu59@asu.edu

☎ 602-773-9980

🐙 Github

🔍 Google Scholar

🌐 LinkedIn

Research Interests

Large Multimodal Model, Machine Learning, Deep Learning, Vision-Language Model, Conversational generative AI, transfer-learning, fine-tuning, Generative model, Weakly Supervised Learning, and Medical Image Analysis.

Education

Arizona State University

Doctoral student in computer engineering GPA: 3.89/4.0

Tempe, AZ

Aug. 2022 – Present

Arizona State University

Master in computer engineering

Tempe, AZ

Aug. 2020 – May. 2022

Northern Arizona University & CQUP

Bachelor's degree in Electrical Engineering

Flagstaff, AZ & Chongqing, China

Aug. 2014 – May. 2019

Overview

- Self-motivated and productive. In the last few years, I have 28 papers and 14 papers as first author (including the co-author). Some of them are published in top venues, such as **ICML**, **ECCV**, **MICCAI**, **CVPR**, **WACV**, **AAAI**, **JAMA**, and **IPMI**.
- Background in mathematic optimal transport learning, topology preserving, computational imaging, computational holography, image & signal processing.
- Programming Languages: Python (preferred), Java, C
- Deep Learning framework: Pytorch, Tensorflow, OpenGP, Python Optimal Transport, OpenCV, Panda.
- Industrial experience as front-end (Linktoun.com) and full-stack (Shanghai TongRui Network Technology Co., Ltd) engineer before my PhD. Familiar with Vue, CSS, Node, React, SQL, TypeScript, HTML, Java, Spring, KOA, Next.js, RestfulAPI, and GitLab CI/CD.

Awards

- **Outstanding Contribution Award in the MICCAI UWF4DR Challenge 2024** - Task 1 and Taks 3
- **Two Third Place Award in the MICCAI UWF4DR Challenge 2024** - Task 1 and Taks 3 (*out of 22 teams*)
- **Two Third Place Award in the MICCAI MMAC Challenge 2023** - Task 1 and Task 3 (*out of 59 teams*)
- **SCAI Doctoral Fellowship** - Spring 2024 Semester

Selected Publication

Video Generation

We introduce a novel data processing pipeline that reconstructs and renders human avatar movements from dynamic camera videos.

- Yang, Z., Shan, M., Farazi, M., Zhu, W., Chen, Y., Dong, X., & Wang, Y. (2024). AMG: Avatar Motion Guided Video Generation. arXiv preprint arXiv:2409.01502.

Work-in-Progress: audio guided the video generation.

Multimodal

We propose a novel formulation of *multimodal variational autoencoders (VAEs) using barycenters*, extending the traditional product and mixture of experts (PoE and MoE) approaches. We better capture *modality-specific and modality-invariant representations* by incorporating the *Wasserstein barycenter*.

- Qiu, P., Zhu, W., Kumar, S., Chen, X., Yang, J., Sun, X., Razi, A., ... & Sotiras, A. (2024). Multimodal Variational Autoencoder: a Barycentric View. **AAAI 2025**.

Work-in-Progress: Gaussian Mixture Multimodal VAE and Gromov-Wasserstein Barycenter joint distribution.

Generative Model

Since the first year of my PhD, I have researched generative models, *from GANs to diffusion models*, with applications in medical imaging (e.g., unpaired image denoising, retinal image reconstruction from fMRI), collaborating closely with Mayo Clinic. I developed and expanded an *OT-based GAN framework integrating optimal transport theory for clinical use*. Recently, we shifted our focus to diffusion models, including the *Schrödinger Bridge and Denoising Diffusion Bridge Models*, to address denoising in unpaired medical images. The outcomes are as follows.

- **Zhu, W.**, Qiu, P., Dumitrescu, O. M., Sobczak, J. M., Farazi, M., Yang, Z., ... & Wang, Y. (2023, June). OTRE: where optimal transport guided unpaired image-to-image translation meets regularization by enhancing **IPMI**.
- Dong, X., Vasa, V. K., **Zhu, W.**, Qiu, P., Chen, X., Su, Y., ... & Wang, Y. (2024). CUNSB-RFIE: Context-aware Unpaired Neural Schrödinger Bridge in Retinal Fundus Image Enhancement. **WACV 2025**.

Work-in-Progress: Denoising Diffusion Bridge Models, Schrödinger Bridge.

Computational pathology - Whole Slide Images

We employed *Multiple Instance Learning (MIL)* for histological WSI classification and localization, proposing *a novel aggregation method* and deeply through the dropout mechanism in MIL. The outcomes are as follows.

- **Zhu, W.**, Chen, X., Qiu, P., H., Sotiras, A., ... & Wang, Y. (2024). DGR-MIL: Exploring Diverse Global Representation in Multiple Instance Learning for Whole Slide Image Classification. **ECCV2024**.

Work-in-Progress: Ready to submit to a **CVPR** paper about permutation invariant effective in MIL.

Time Series Analysis

I cooperate with Xiwen Chen and Peijie Qiu to focus on multivariate time series analysis. Firstly, We reformulate multivariate time series classification (MTSC) as *a weakly supervised* problem and propose a *time-aware MIL framework*. The second work *enhances Transformer performance* in time forecasting by introducing a novel attention mechanism with *Sequence Complementors*, validated through information theory. The outcomes are as follows.

- Chen, X., Qiu, P., **Zhu, W.**, Li, H., Wang, H., Sotiras, A., ... & Razi, A. (2024). TimeMIL: Advancing Multivariate Time Series Classification via a Time-aware Multiple Instance Learning. **ICML2024**.

Work-in-Progress: Modeling Temporal Data as Continuous Functions with Diffusion Model.

Physics-driven Computational Imaging

We developed an untrained physics-based low-level computer vision framework for *image reconstruction under uncertain forward model parameters*, demonstrating its effectiveness in digital holography and compressed sensing applications.

- Chen, X., **Zhu, W.**, Qiu, P., & Razi, A. (2024). Imaging Signal Recovery Using Neural Network Priors Under Uncertain Forward Model Parameters. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1420-1429). (**PBDL@CVPR2024**, Oral).

General Medical Image Analysis

In past PhD research, I worked on various projects in medical image analysis, including *medical image segmentation, image denoising, computer-aided diagnosis for retinal diseases, and self-supervised learning in Medical images*. The outcomes are as follows.

- **Zhu, W.**, Chen, X., Qiu, P., Farazi, M., ... & Wang, Y. (2024). SelfReg-UNet: Self-Regularized UNet for Medical Image Segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention. (**MICCAI2024**)
- Qian, B., Sheng, B., Chen, H., Wang, X., ... & **Zhu, W.**, Wang, Y. X. (2024). A competition for the diagnosis of myopic maculopathy by artificial intelligence algorithms. **JAMA ophthalmology**.

Professional Activities

- Program Committees for AAAI
- Invited reviewer by ICLR, AAAI, CVPR, WACV, MICCAI, MedA, IEEE TMI, and IEEE ISBI.
- Teaching Assistant for CSE 220 Programming for Computer Engineering, CSE 570 Advanced Computer Graphics.
- Research Assistant in Geometry Systems Laboratory lab, Arizona State University.

Full Peer-reviewed Publication List

- Dumitrascu, O. M., Li, X., **Zhu, W.**, Woodruff, B. K., Nikolova, S., Sobczak, J., ... & Wang, Y. (2024). Color Fundus Photography and Deep Learning Applications in Alzheimer Disease. *Mayo Clinic Proceedings: Digital Health*, 2(4), 548-558.
- Xiong, Y., **Zhu, W.**, Wang, Y., & Lu, Z. L. (2024). Visual Inputs Reconstructing through Enhanced 3T fMRI Data from Optimal Transport Guided Generative Adversarial Network. *Journal of Vision*, 24(10), 1478-1478.
- Vasa, V. K., Qiu, P., **Zhu, W.**, Xiong, Y., Dumitrascu, O., & Wang, Y. (2024). Context-Aware Optimal Transport Learning for Retinal Fundus Image Enhancement. *WACV*.
- Yang, Z., Shan, M., Farazi, M., **Zhu, W.**, Chen, Y., Dong, X., & Wang, Y. (2024). AMG: Avatar Motion Guided Video Generation. *arXiv preprint arXiv:2409.01502*.
- Wang, H., **Zhu, W.**, Qin, J., Li, X., Dumitrascu, O., Chen, X., ... & Razi, A. (2024). RBAD: A Dataset and Benchmark for Retinal Vessels Branching Angle Detection. *IEEE BHI and JBHI*.
- **Zhu, W.**, Chen, X., Qiu, P., Sotiras, A., Razi, A., & Wang, Y. (2024). DGR-MIL: Exploring Diverse Global Representation in Multiple Instance Learning for Whole Slide Image Classification. *ECCV 2024*.
- Likassa, H. T., Chen, D. G., Chen, K., Wang, Y., & **Zhu, W.** (2024). Robust PCA with Lw, and L2, 1 Norms: A Novel Method for Low-Quality Retinal Image Enhancement. *Journal of Imaging*, 10(7), 151.
- **Zhu, W.**, Chen, X., Qiu, P., Farazi, M., Sotiras, A., Razi, A., & Wang, Y. (2024). SelfReg-UNet: Self-Regularized UNet for Medical Image Segmentation. *MICCAI 2024*.
- Yang, Z., Su, Y., Rouhi, R., Farazi, M., **Zhu, W.**, Chen, Y., ... & Lepore, N. (2024, May). Pre-Training Graph Attention Convolution for Brain Structural Imaging Biomarker Analysis and Its Application to Alzheimer's Disease Pathology Identification. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)* (pp. 1-5). IEEE.
- Chen, X., Qiu, P., **Zhu, W.**, Li, H., Wang, H., Sotiras, A., ... & Razi, A. (2024). TimeMIL: Advancing Multivariate Time Series Classification via a Time-aware Multiple Instance Learning. *ICML 2024*.
- Dumitrascu, O., Li, X., Sobczak, J., **Zhu, W.**, Caselli, R., Woodruff, B., & Wang, Y. (2024, April). Retinal Color Fundus Photography and Deep Learning can Discriminate Alzheimer's Disease (P8-9.006). In *Neurology* (Vol. 102, No. 17 supplement 1, p. 2716). Hagerstown, MD: Lippincott Williams & Wilkins.
- Xiong, Y., **Zhu, W.**, Lu, Z. L., & Wang, Y. (2024). Reconstructing Retinal Visual Images from 3T fMRI Data Enhanced by Unsupervised Learning. *ISBI 2024*.
- Qian, B., Sheng, B., Chen, H., Wang, X., Li, T., Jin, Y., ... & **Zhu, W.**, Wang, Y. X. (2024). A competition for the diagnosis of myopic maculopathy by artificial intelligence algorithms. *JAMA ophthalmology*.
- **Zhu, W.**, Qiu, P., Chen, X., Li, X., Lepore, N., Dumitrascu, O. M., & Wang, Y. (2024). nnMobileNet: Rethinking CNN for Retinopathy Research. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2285-2294).
- Chen, X., **Zhu, W.**, Qiu, P., & Razi, A. (2024). Imaging Signal Recovery Using Neural Network Priors Under Uncertain Forward Model Parameters. In *Proceedings of the IEEE/CVF Conference*

- on Computer Vision and Pattern Recognition (pp. 1420-1429).
- Li, H., **Zhu, W.**, Chen, X., & Wang, Y. (2023, October). Prediction of Spherical Equivalent with Vanilla ResNet. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 66-74). Cham: Springer Nature Switzerland.
 - **Zhu, W.**, Qiu, P., Chen, X., Li, H., Wang, H., Lepore, N., ... & Wang, Y. (2023, October). Beyond mobilenet: An improved mobilenet for retinal diseases. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 56-65). Cham: Springer Nature Switzerland.
 - **Zhu, W.**, Qiu, P., Chen, X., Dumitrascu, O. M., & Wang, Y. (2023). Pdl: Regularizing multiple instance learning with progressive dropout layers. arXiv preprint arXiv:2308.10112.
 - Farazi, M., Yang, Z., **Zhu, W.**, Qiu, P., & Wang, Y. (2023, June). Tetcnn: Convolutional neural networks on tetrahedral meshes. In International Conference on Information Processing in Medical Imaging (pp. 303-315). Cham: Springer Nature Switzerland.
 - **Zhu, W.**, Qiu, P., Dumitrascu, O. M., Sobczak, J. M., Farazi, M., Yang, Z., ... & Wang, Y. (2023, June). OTRE: where optimal transport guided unpaired image-to-image translation meets regularization by enhancing. In International Conference on Information Processing in Medical Imaging (pp. 415-427). Cham: Springer Nature Switzerland.
 - Wu, J., Su, Y., Chen, Y., **Zhu, W.**, Reiman, E. M., Caselli, R. J., ... & Alzheimer's Disease Neuroimaging Initiative. (2023). A Surface-Based Federated Chow Test Model for Integrating APOE Status, Tau Deposition Measure, and Hippocampal Surface Morphometry. *Journal of Alzheimer's Disease*, 93(3), 1153-1168.
 - **Zhu, W.**, Qiu, P., Farazi, M., Nandakumar, K., Dumitrascu, O. M., & Wang, Y. (2023, April). Optimal transport guided unsupervised learning for enhancing low-quality retinal images. In 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI) (pp. 1-5). IEEE.
 - **Zhu, W.**, Qiu, P., Lepore, N., Dumitrascu, O. M., & Wang, Y. (2023, March). Self-supervised equivariant regularization reconciles multiple-instance learning: Joint referable diabetic retinopathy classification and lesion segmentation. In 18th International Symposium on Medical Information Processing and Analysis (Vol. 12567, pp. 100-107). SPIE.
 - Wu, J., Su, Y., **Zhu, W.**, Jalili Mallak, N., Lepore, N., Reiman, E. M., ... & Alzheimer's Disease Neuroimaging Initiative. (2023). Improved prediction of amyloid- and tau burden using hippocampal surface multivariate morphometry statistics and sparse coding. *Journal of Alzheimer's Disease*, 91(2), 637-651.
 - Farazi, M., **Zhu, W.**, Yang, Z., & Wang, Y. (2023). Anisotropic multi-scale graph convolutional network for dense shape correspondence. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 3146-3155).
 - Dumitrascu, O. M., **Zhu, W.**, Qiu, P., Nandakumar, K., & Wang, Y. (2022, October). Automated Retinal Imaging Analysis for Alzheimer Disease Screening. In IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI).

Under Review

- Chen, X., Qiu, P., **Zhu, W.**, Li, H., Wang, H., Sotiras, A., ... & Razi, A. (2024). Sequence Complementor: Complementing Transformers For Time Series Forecasting with Learnable Sequences. **Submitted to AAAI 2025.**
- Chen, X., **Zhu, W.**, Qiu, P., Li, H., Wang, H., Sotiras, A., ... & Razi, A. (2024). FIC-TSC: Learning Time Series Classification with Fisher Information Constraint. **Submitted to WWW 2025**
- Dong, X., Vasa, V. K., **Zhu, W.**, Qiu, P., Chen, X., Su, Y., ... & Wang, Y. (2024). CUNSB-RFIE: Context-aware Unpaired Neural Schrödinger Bridge in Retinal Fundus Image Enhancement. **Submitted WACV.**
- Qiu, P., Zhu, W., Kumar, S., Chen, X. Yang, J., Sun, X., Razi, A., ... & Sotiras, A. (2024).

