

ZHIJIAN YANG

📍 Bellevue, WA 98004 📞 267-581-9947 ✉ zhijianyyy@gmail.com 🏠 zhijian-yang.com
🌐 [linkedin.com/in/zhijian-yang](https://www.linkedin.com/in/zhijian-yang) 🐙 github.com/zhijian-yang 🎓 scholar.google.com/zhijian-yang

RESEARCH INTEREST

I am a researcher specializing in AI for healthcare, with a current focus on vision-language foundation models and their applications in the medical field. I recently completed my PhD at the University of Pennsylvania (May 2024), where my work centered on developing deep generative models for medical image analysis and multimodal disease modeling. My research has been published or accepted in Nature Medicine, ICLR, CVPR, and Nature Communications.

EDUCATION

University of Pennsylvania

Ph.D. in Applied Mathematics and Computational Science; GPA: 4.0

Aug. 2019 – May 2024

Philadelphia, PA

New York University

B.A. in Honors Mathematics and Computer Science; GPA: 3.823

Sep. 2015 – May 2019

New York, NY

RESEARCH EXPERIENCE

GE Healthcare

AI Research Scientist at Foundation AI Team

May 2024 – Now

Bellevue, WA

Efficient Prompting and Preference Optimization for Segmentation (Intern Project Mentor)

Accepted to CVPR 2025

- Delivered effective text-based prompts to SAM for medical image segmentation through pretrained Vision-Language Models
- Enabled segmentation model tuning using preference optimization, allowing annotators to rank multiple segmentation outputs from the model without needing to provide full segmentation masks
- Enhanced workflow efficiency through the optimized prompting and annotation processes, while preserving segmentation quality

Center for Biomedical Image Computing and Analytics

Graduate Researcher; Supervisor: Christos Davatzikos

August 2019 – May 2024

Philadelphia, PA

Patient Clustering based on Disease-related Imaging Patterns (Smile-GAN)

Published in Nature Communications

- Created a cutting-edge **deep clustering method** that effectively groups patients based on distinct disease-related imaging patterns
- Utilized **generative model** with categorical latent variables to estimate various disease effects on human brains, while minimizing the impact of disease-unrelated confounding effects
- Validated the model's performance using the Alzheimer's Disease dataset and performed statistical and survival analyses to test the clinical significance of model-derived imaging pattern types

Interpretable Representation of Disease-related Imaging Patterns (Surreal-GAN)

Published in ICLR 2022

- Developed a state-of-the-art **representation learning** method based on Smile-GAN, effectively summarizing complex disease patterns into concise and interpretable latent representations that indicate severity of different underlying pathologies
- Designed theoretically valid regularization mechanisms that guarantee the disentanglement of imaging patterns, as well as the positive correlations between latent representations and disease pattern severity

Five Dominant Dimensions of Brain Aging Derived through Surreal-GAN

Published in Nature Medicine

- Extended the Surreal-GAN methodology by introducing a correlation structure among the R-indices in the representation latent space, thereby capturing interactions among multiple underlying neuropathological processes
- Applied the Surreal-GAN model to a large cohort of aging participants, derived representations of five dominant dimensions of neuroanatomical brain aging, and associated them with cognitive, clinical, lifestyle, and genetic measures

Multimodal Disease Subtyping with Imaging and Genetic Signatures (Gene-SGAN)

Published in Nature Communications

- Built a novel **multi-modal learning** method based on Smile-GAN, which clusters patients using both imaging and genetic data, facilitating more accurate and informative disease subtype identification
- Leveraged a unique approach combining GAN and **Variational Inference** to disentangle associated and unassociated imaging and genetic variations into separate latent variables through iterative training
- Conducted genome-wide association tests with multinomial logistic regression to detect significant SNP-subtype associations

RIPS program, Institute for Pure and Applied Mathematics

Research Intern; Supervisor: Blerta Shtylla and Tom Chou

June 2018 – August 2018

Los Angeles, CA

Radiomics-Based Quality of Life Prediction for Cancer Patients

Published in Computer in Biology and Medicine

- Used deep learning and statistical methods to predict patients' quality of life after prostate radiation therapy using CT images and radiation treatment plans
- Performed data augmentation through image registration and leveraged transfer training for patient classification

PROFESSIONAL AND LEADERSHIP EXPERIENCE

GE Healthcare

January 2024 – Now

AI Scientist

Bellevue, WA

- Developed a specialized vision-language pretraining framework tailored for medical applications, enabling foundation models with robust capabilities to interpret and summarize medical imaging and report data.
- Adapted and tested the foundation models across diverse tasks such as disease diagnosis, tumor detection, and visual question answering, aiming to improve efficiency and optimize healthcare service quality.
- Mentored an intern to enhance medical image segmentation through more efficient prompting and annotation workflows, leveraging pretrained vision-language models and preference optimization.

Center for Biomedical Image Computing and Analytics

January 2022 – May 2024

Machine Learning Seminar Organizer

Philadelphia, PA

- Proposed topics based on the latest machine learning research directions and invited experienced speakers from different labs
- Hosted weekly seminar for twenty attendees and led discussions on potential intersections between the presented topics and biomedical research

PricewaterhouseCoopers China

June 2016 – July 2016

Assurance Intern

Nanjing, China

- Participated in the IPO audit for a financial leasing company with a three billion CNY market cap
- Reviewed financial statements and feasibility reports to perform credit reviews and suggested credit ratings
- Performed control tests and substantial procedures by sending confirmation requests, sampling, and recalculating financial data

SELECTED PUBLICATIONS

Paper in Conference Proceedings

- **Yang,Z.**, Wen,J.,and Davatzikos,C. Surreal-GAN: Semi-Supervised Representation Learning via GAN for uncovering heterogeneous disease-related imaging patterns. *International Conference on Learning Representations (ICLR)* (2022)

Paper Accepted to Conference Proceedings

- Konwer,A., **Yang,Z. (mentor)**, Bas,E., et al. Enhancing SAM with Efficient Prompting and Preference Optimization for Semi-supervised Medical Image Segmentation. *Accepted to The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2025)

Papers in Peer Reviewed Journals

- **Yang,Z.**,Wen,J., Erus,G., et al. Brain aging patterns in a large and diverse cohort of 49,482 individuals *Nature Medicine* (2024)
- **Yang,Z.**,Wen,J., Abdulkadir,A., et al. Gene-SGAN: a method for discovering disease subtypes with imaging and genetic signatures via multi-view weakly-supervised deep clustering *Nature Communications* (2024)
- **Yang,Z.**, Nasrallah,I.M., Shou,H., et al. A deep learning framework identifies dimensional representations of Alzheimer's Disease from brain structure. *Nature Communications* (2021)
- **Yang,Z.**, Olszewski,D., He,C., et al. Machine learning and statistical prediction of patient quality-of-life after prostate radiation therapy. *Computers in Biology and Medicine* (2021)
- Duggan,W., **Yang,Z.**, Cui,Y., et al. Proteomic analyses reveal plasma EFEMP1 and CXCL12 as biomarkers and determinants of neurodegeneration *Alzheimer's and Dementia* (2024)
- Wen,J., **Yang,Z.**, Nasrallah,I.M., et al. Genetic and clinical correlates of two neuroanatomical AI dimensions in the Alzheimer's disease continuum *Translational Psychiatry* (2024)
- Wen,J., Zhao,B., **Yang,Z.**, et al. The genetic architecture of multimodal human brain age *Nature communications* (2024)
- Chintapalli,S., Wang,R., **Yang,Z.**, et al. Generative models of MRI-derived neuroimaging features and associated dataset of 18,000 samples *Scientific Data* (2024)
- Wang,R., Bashyam,V., **Yang,Z.**, et al. Applications of Generative Adversarial Networks in Neuroimaging and Clinical Neuroscience *Neuroimage* (2023)

PRESS COVERAGE

Nature News: "Five ways the brain can age: 50,000 scans reveal possible patterns of damage"

HONORS AND AWARDS

Benjamin Franklin Fellowship , graduate fellowship at University of Pennsylvania	2019-Now
Research in Industrial Projects for Students (RIPS) , undergraduate research fellowship at UCLA	2018
Summer Undergraduate Research Experience (SURE) , undergraduate research fellowship at New York University	2017
Phi Beta Kappa Honor Society , Beta Chapter of New York	2019
Dean's List , New York University	2015-2019

MENTORING EXPERIENCE

Brandon Theodorou; Summer intern at GE Healthcare, PhD student at UIUC	May. 2024 - Aug. 2024
Aishik Konwer; Summer intern at GE Healthcare, PhD student at Stony Brook University	May. 2024 - Aug. 2024

SOFTWARE PACKAGES

SurrealGAN: A python implementation of Surreal-GAN for semi-supervised representation learning. [PyPI Link](#)
SmileGAN: A python implementation of Smile-GAN for semi-supervised clustering. [PyPI Link](#)
GeneSGAN: A python implementation of Gene-SGAN for semi-supervised multi-view clustering. [PyPI Link](#)

TECHNICAL SKILLS

Programming Languages: Python, Matlab, Java, R
Software and Frameworks: Scikit-learn, Pytorch, TensorFlow
Language: English, Chinese

SERVICE

Reviewer: ICLR 2025; Radiology-AI; Medical Image Analysis; Scientific Report; MICCAI 2023; IEEE-ISBI 2023