

ZHIJIAN YANG

📍 3333 Walnut St., Philadelphia, PA 19104 📞 267-581-9947 ✉ yzj@sas.upenn.edu
🌐 [linkedin.com/in/zhijian-yang](https://www.linkedin.com/in/zhijian-yang) 🐙 github.com/zhijian-yang 🎓 scholar.google.com/zhijian-yang

I am a fourth-year Ph.D. student in Applied Math and Computational Science at the University of Pennsylvania, advised by Christos Davatzikos. My research interests lie in machine learning and deep learning, with specific focuses on generative model, representation learning, and multi-modal learning. Throughout my Ph.D. study, I primarily worked on machine learning for healthcare, developing several deep learning models to disentangle disease heterogeneity using medical imaging and genomic data. My research work have been published in Nature Communications and ICLR 2022.

EDUCATION

University of Pennsylvania

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

Sep. 2019 – May 2024 (Expected)

Philadelphia, PA

New York University

B.A. in Honors Mathematics and Computer Science; GPA: 3.8

Sep. 2015 – May 2019

New York, NY

RESEARCH EXPERIENCE

Center for Biomedical Image Computing and Analytics

PhD Student Researcher; Supervisor: Christos Davatzikos

September 2019 – Now

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 Disease Imaging Pattern Representation (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

Disease Subtyping with Imaging and Genetic Signatures (Gene-GAN)

Under Review 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

Center for Biomedical Image Computing and Analytics

Machine Learning Seminar Organizer

January 2022 – Now

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

Sunrise Brokers LLC

Trade Support Intern

February 2017 – May 2017

New York, NY

- Developed a database with information on over one thousand traders and their frequently traded assets
- Built Excel tools that automatically compile daily trading information, significantly increasing the efficiency of OATS reporting

- 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

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)

Papers in Peer Reviewed Journals

- **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)
- Wen,J., Fu,C., Tosun,D., Veturi,Y., **Yang,Z.**, et al. Characterizing heterogeneity in neuroimaging, cognition, clinical symptomatology, and genetics among patients with Late-Life Depression. *JAMA Psychiatry* (2022)
- Wen,J., Varol,E., Sotiras,A., **Yang,Z.**, et al. Multi-scale semi-supervised clustering of brain images: deriving disease subtypes. *Medical image analysis* (2022)
- Wang,R., Bashyam,V., **Yang,Z.**, et al. Applications of Generative Adversarial Networks in Neuroimaging and Clinical Neuroscience *Neuroimage* (2023)

Papers Under Review

- **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 *Under Review in Nature Communications* (2023)
- Wen,J.,**Yang,Z.**, Nasrallah,I.M., et al. Genetic, clinical underpinnings of subtle early brain change along Alzheimer's dimensions *Under Review in Brain* (2022)

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

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