

# ZHIJIAN YANG

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

## 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

### 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

*AI Scientist*

**January 2024 – Now**

*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.

### Center for Biomedical Image Computing and Analytics

*Machine Learning Seminar Organizer*

**January 2022 – May 2024**

*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)

### 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)
- 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)
- Duggan,W., **Yang,Z.**, Cui,Y., et al. Proteomic analyses reveal plasma EFEMP1 and CXCL12 as biomarkers and determinants of neurodegeneration *Alzheimer's Dementia* (2024)

## HONORS AND AWARDS

---

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

## SERVICE

---

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