Yiqiu Shen

832-693-4010 | ys1001@nyu.edu | https://seyiqi.github.io | https://github.com/seyiqi

Research Interests

Areas: deep learning, medical imaging, computer vision, interpretability, weakly supervised learning, algorithmic fairness

My research objective is to build explainable AI systems for medical imaging to provide accurate diagnoses and improve patient care. Towards this goal, my research has centered on weakly supervised learning, interpretable multi-modal learning, and algorithmic fairness.

Education

New York University, Center for Data Science — *New York, NY*

PhD in Data Science — *Aug. 2018 - May. 2023*

- Supervised by Professor **Krzysztof J. Geras** and Professor **Kyunghyun Cho**

New York University, Center for Data Science — *New York, NY*

M.S. in Data Science, GPA: 4.0/4.0 — *Sep. 2016 - Jun. 2018*

Rice University — *Houston, TX*

B.S. in Computer Science, GPA: 3.9, Cum Laude — *Aug. 2010 - May. 2014*

Journal Publications

Yiqiu Shen, Laura Heacock, Jonathan Elias, Keith D Hentel, Beatriu Reig, George Shih, Linda Moy. "ChatGPT and other large language models are double-edged swords" In *Radiology*, 2023. **[paper link]**

Yiqiu Shen, Farah E. Shamout, Jamie R. Oliver, Jan Witowski, Kawshik Kannan, et. al. "Artificial Intelligence System Reduces False-Positive Findings in the Interpretation of Breast Ultrasound Exams" In *Nature Communications*, 2021. **[paper link]**

Yiqiu Shen, Nan Wu, Jason Phang, Jungkyu Park, Kangning Liu, et. al. "An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization" In *Medical Image Analysis*, 2021. **[paper link]**

Farah E. Shamout*, **Yiqiu Shen***, Nan Wu*, Aakash Kaku*, Jungkyu Park*, et. al. "An artificial intelligence system for predicting the deterioration of COVID-19 patients in the emergency department" In *npj Digital Medicine*, 2021. **[paper link]**

Nan Wu, Zhe Huang, **Yiqiu Shen**, Jungkyu Park, et. al. "Reducing false-positive biopsies with deep neural networks that utilize local and global information in screening mammograms" In *Journal of Digital Imaging*, 2021. **[paper link]**

Nan Wu, Jason Phang, Jungkyu Park, **Yiqiu Shen**, Zhe Huang, et. al. "Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening" In *IEEE Transactions on Medical Imaging*, 2020. **[paper link]**

Kevin Leung, Bofei Zhang, Jimin Tan, **Yiqiu Shen**, Krzysztof J. Geras, et. al. "Prediction of Total Knee Replacement and Diagnosis of Osteoarthritis by Using Deep Learning on Knee Radiographs: Data from the Osteoarthritis Initiative" In *Radiology*, 2020. **[paper link]**

Thomas Schaffter, Diana S. M. Buist, Christoph I. Lee, et. al. "Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms" In *JAMA Network Open*, 2020. **[paper link]**

Conference Publications

Kangning Liu, Weicheng Zhu, **Yiqiu Shen**, Sheng Liu, Narges Razavian, Krzysztof J Geras, Carlos Fernandez-Granda. "Multiple Instance Learning via Iterative Self-Paced Supervised Contrastive Learning " In *Conference on Computer Vision and Pattern Recognition*, 2023. **[paper link]**

Sheng Liu, Kangning Liu, Weicheng Zhu, **Yiqiu Shen**, Carlos Fernandez-Granda. "Adaptive Early-Learning Correction for Segmentation from Noisy Annotations " In *Conference on Computer Vision and Pattern Recognition*, 2022. (oral) **[paper link]**

Kangning Liu, **Yiqiu Shen**, Nan Wu, Jakub Chłędowski, Carlos Fernandez-Granda, Krzysztof J Geras. "Weakly-supervised High-resolution Segmentation of Mammography Images for Breast Cancer Diagnosis" In *Medical Imaging with Deep Learning*, 2021. (oral) **[paper link]**

Yiqiu Shen, Nan Wu, Jason Phang, Jungkyu Park, Gene Kim, Linda Moy, Kyunghyun Cho, Krzysztof J Geras. "Globally-Aware Multiple Instance Classifier for Breast Cancer Screening" In *International Workshop on Machine Learning in Medical Imaging*, 2019. (oral) **[paper link]**

Jungkyu Park, Jason Phang, **Yiqiu Shen**, Nan Wu, S Gene Kim, Linda Moy, Kyunghyun Cho, Krzysztof J Geras. "Screening Mammogram Classification with Prior Exams" In *Medical Imaging with Deep Learning*, 2019. **[paper link]**

Nan Wu, Krzysztof J. Geras, **Yiqiu Shen**, Jingyi Su, S. Gene Kim, Eric Kim, Stacey Wolfson, Linda Moy, Kyunghyun Cho. "Breast Density Prediction with Deep Convolutional Neural Networks" In *IEEE International Conference on Acoustics, Speech and Signal Processing* , 2018. (oral) **[paper link]**

Work Experience

Google
Palo Alto, CA

Research Scientist Intern, Google Health — *May. 2022 - Aug. 2022*

- Led the development of a novel fairness-ware AI system that detects breast cancer on mammography images.
- Utilized contrastive learning and importance re-weighting to advance algorithmic fairness of an EfficientDet object detection model.
- Improved detection performance (mean average precision) by 3.4% on minority patient cohorts that are under-represented in the training data.
- Reduced existing disparity (ΔEqualized Odds) by 56.7% on multiple protected attributes including ethnicity, age, and breast density.
- Provided statistical analysis to support prospective studies and CE mark regulatory filing (completed in July 2022).

Two Sigma Investments
Houston, TX

Software Development Engineer, Alpha Capture Team — *Jul. 2014 - Jun. 2016*

- Maintained a platform that extracts trading signals from market sentiment data by collecting trading ideas from brokers and traders.
- Developed a real-time dashboard to monitor analytical statistics of the Alpha Capture platform.
- Built a scalable reporting system that periodically evaluates the performance of brokers and traders.
- Performed exploratory data analysis on both internal and external datasets to provide data-backed insights into various business questions and expedite the team's decision-making process.

Indeed.com
Austin, TX

Software Engineer Intern — *May. 2012 - Aug. 2012*

- Developed a spell check application for users' search queries based on the Noisy Channel Model.
- This application is equipped with an n-gram language model and a Bayesian error model trained on click stream data hosted on a Hadoop server cluster.
- This application has been deployed to production, supports 10 languages, and achieves a 71.7% click through rate.

Talks

Weakly supervised learning for breast cancer screening

AWS Research — *Sep. 2022*

Deep learning systems for screening mammograms and breast ultrasound interpretation

Google Health — *Nov. 2021*

Deep learning to assist radiologists in breast ultrasound interpretation

Medical Imaging meets NeurIPS — *Dec. 2020*

Power prognosis of COVID-19 patients using AI

Translational Research in Progress Seminar, NYU Langone Health — *Nov. 2020*

Interpretable breast cancer diagnosis with weakly supervised learning

Ezra AI, Inc — *Mar. 2020*

Department Research Seminar, NYU Center for Data Science — *Dec. 2019*

Globally-aware multiple instance classifier for breast cancer screening

CILVR Seminar, NYU Computer Science — *Nov. 2019*

10th International Workshop on Machine Learning in Medical Imaging, MICCAI 2019 — *Oct. 2019*

Machine Learning in Radiology, NYU Langone Medical Center — *Sep. 2019*

Professional Service

Journal Reviewer: Medical Image Analysis, npj Digital Medicine, Biotechnology Advances, npj Breast Cancer, European Journal of Radiology, QIMS, NMR in Biomedicine, Artificial Intelligence In Medicine, Ophthalmic Research, Applied Artificial Intelligence

Conference Reviewer: ICLR 2021-2023, MICCAI 2020-2022, ICML 2020-2023, NeurIPS 2019-2023, CVPR 2023

Awards & Certificates

Travel Grant Award, Machine Learning for Health Workshop at NeurIPS 2019 *Dec. 2019*
Best Paper Award, AI for Social Good Workshop at ICML 2019 *Jun. 2019*
Best Interdisciplinary Project, NYU Center for Data Science Academy Award *April 2019*
Roadie Award for the most popular abstracts, RSNA 2018 *Nov. 2018*
Social Impact Award, NYU Center for Data Science Academy Award *Feb. 2017*
Passed CFA Level II Exam *Jun. 2016*
Moncrief Undergraduates Summer Fellowship, University of Texas at Austin *May. 2013*

Teaching Experience

DS-GA 1003: Machine Learning and Computational Statistics *New York, NY*

Teaching Assistant, Center for Data Science, New York University *Jan. 2020 - May. 2020*

- Instructor: Prof. Christopher Policastro and Prof. He He
- `https://nyu-ds1003.github.io/spring2022/#home`

DS-GA 3001: Modeling Time Series Data *New York, NY*

Teaching Assistant, Center for Data Science, New York University *Jan. 2018 - May. 2018*

- Instructor: Prof. Cristina Savin
- `https://github.com/charlieblue17/timeseries2018`

DS-GA 1011: Natural Language Processing with Representation Learning *New York, NY*

Teaching Assistant, Center for Data Science, New York University *Sep. 2017 - Dec. 2017*

- Instructors: Prof. Sam Bowman and Prof. Kyunghyun Cho
- `https://github.com/nyu-mll/DS-GA-1011-Fall2017`