NAN WU

Room 733, 60 5th Ave, New York, NY, 10011
nan.wu@nyu.edu ◇ +1 9174985729 ◇ homepage ◇ google scholar

RESEARCH INTERESTS

Multi-modal learning with application on medical imaging, clinical notes understanding and text-to-speech.

EDUCATION

New York University, Center for Data Science
Ph.D. student, supervised by Professor Krzysztof J. Geras and Professor Kyunghyun Cho, *September 2018 -*
M.S. in Data Science *September 2016 - May 2018*

University of Science and Technology of China, Special Class of Gifted Young
B.S. in Statistics and B.A. in Business Administration *September 2012 - June 2016*

AWARDS

Google PhD Fellowship in Machine Perception, Speech Technology and Computer Vision *2020 -*
Most Innovative Healthcare Solution, MIT CSAIL AI in Healthcare Summit 2020. *October 2020*
Travel Funding, MIDAS Symposium 2019. *November 2019*
Best Paper Award, AI for Social Good workshop, ICML 2019. *June 2019*
Travel Funding, AI for Social Good workshop, ICML 2019. *June 2019*
Full Graduate Scholarship, NYU. *2018 -*
Roadie Award for one of the most popular abstracts at the RSNA conference. *2018*
Outstanding Student Scholarship, Top 10% in USTC. *November 2015*
Alumni Outstanding Student Scholarship, 6 out of 1027 in School of Gifted Young, USTC. *June 2015*
Outstanding Student Scholarship, Top 10% in USTC. *November 2013*
Outstanding Freshman Scholarship, Top 5% in USTC. *September 2012*
National Olympiad in Biology, Hebei Province, Silver Medal *2011*
National Olympiad in Physics, Hebei Province, Bronze Medal *2011*
National Olympiad in Informatics, Hebei Province, Silver Medal. *2010 and 2011*

SELECTED PUBLICATIONS & PREPRINTS

SLAM: A Unified Encoder for Speech and Language Modeling via Speech-Text Joint Pre-Training. *A. Bapna, Y. Chung,* ***N. Wu****, A. Gulati, Y. Jia, J. H. Clark, M. Johnson, J. Riesa, A. Conneau, Y. Zhang.* arXiv. 2021.

Artificial Intelligence System Reduces False-Positive Findings in the Interpretation of Breast Ultrasound Exams. *Y. Shen, F. E. Shamout, J. R. Oliver, J. Witowski, K. Kannan, J. Park,* ***N. Wu****, C. Huddleston, S. Wolfson, A. Millet, R. Ehrenpreis, D. Awal, C. Tyma, N. Samreen, Y. Gao, C. Chhor, S. Gandhi, C. Lee, S. Kumari-Subaiya, C. Leonard, R. Mohammed, C. Moczulski, J. Altabet, J. Babb, A. Lewin, B. Reig, L. Moy, L. Heacock, K. J. Geras.* Nature Communications. 2021.

Reducing False-Positive Biopsies with Deep Neural Networks that Utilize Local and Global Information in Screening Mammograms. ***N. Wu****, Z. Huang, Y. Shen, J. Phang, J. Park, T. Makino, S Kim, K. Cho, L. Heacock, L. Moy, K. J. Geras.* Journal of Digital Imaging. 2021.

Weakly-supervised High-resolution Segmentation of Mammography Images for Breast Cancer Diagnosis. *K. Liu, Y. Shen,* ***N. Wu****, J. Chffdowski, C. Fernandez-Granda, K. J. Geras.* MIDL 2021.

An Artificial Intelligence System for Predicting the Deterioration of COVID-19 Patients in the Emergency Department. *F. E Shamout, Y. Shen,* ***N. Wu****, A. Kaku, J. Park, T. Makino, S. Jastrzbski, D. Wang, B. Zhang, S. Dogra, M. Cao, N. Razavian, D. Kudlowitz, L. Azour, W. Moore, Y. W Lui, Y. Aphinyanaphongs, C. Fernandez-Granda, K. J. Geras.* NPJ digital medicine, 2021.

An Interpretable Classifier for High-resolution Breast Cancer Screening Images Utilizing Weakly Supervised Localization. *Y. Shen,* ***N. Wu****, J. Phang, J. Park, K. Liu, S. Tyagi, L. Heacock, S Kim, L. Moy, K. Cho, K. J. Geras.* Medical image analysis, 2021.

Improving the Ability of Deep Neural Networks to Use Information from Multiple Views in Breast Cancer Screening. ***N. Wu****, J. Park, S. Jastrzebski, L. Moy, K. Cho, K. J. Geras.* MIDL 2020.
Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. ***N. Wu****, J. Phang, J. Park, Y. Shen, Z. Huang, M. Zorin, S. Jastrzebski, T. Fvry, J. Katsnelson, E. Kim, S. Wolfson, U. Parikh, S. Gaddam, L. Young Lin, K. Ho, J. D. Weinstein, B. Reig, Y. Gao, H. Toth, K. Pysarenko, A. Lewin, J. Lee, K. Airola, E. Mema, S. Chung, E. Hwang, N. Samreen, S. Kim, L. Heacock, L. Moy, K. Cho, K. J. Geras.* IEEE Transactions on Medical Imaging, 2019.
Large-scale Classification of Breast MRI Exams using Deep Convolutional Networks. *S. Gong, M. Muckley,* ***N Wu****, T. Makino, S. Kim, L. Heacock, L. Moy, F. Knoll, K. J. Geras* Medical Imaging meets NeurIPS 2019.
Breast Density Classification with Deep Convolutional Neural Networks. ***N. Wu****, K. J. Geras, Y. Shen, J. Su, S. Kim, E. Kim, S. Wolfson, L. Moy, K. Cho.* ICASSP 2018.

INTERNSHIPS

Google Research Remote
Research intern *June 2021 - September 2021*
- Joint multi-modal pre-training with speech and text for text normalization towards end-to-end text-to-speech.
- Supervised by Jason Riesa and Jonathan H. Clark.

FAIRWAYiQ New York
Data Science consultant *June 2019 - January 2020*
- Worked as data consultant for a startup in intelligent golf course management. Helped with building automatic task tracking and efficiency improvement system based on operational data and real-time GPS signals.

AIG, Science. New York
Computer Vision and NLP Research Assistant *May 2017 - December 2017*
- Developed networks for car damage detection, especially for generating saliency map. Built connection between real image and 3D projection with generative models.
- Compared performances of various embedding models for insurance data and explored the transition of word space under a growing context space.

Bayesquare Foundation Inc. New York
Machine Learning Research Intern *November 2016 - May 2017*
- Built classification model for financial entities using features about the degree of their marginality using machine learning models, including random forests models.

Haloband LLC. Shanghai
Growth hacker (Marketing Researcher) Intern *March 2016 - May 2016*
- Built a business model of the product and applied ideas in growth hacking to analyze user types and activities;
- Discovered patterns related with advertising strategies and website designs.

RESEARCH EXPERIENCE

CILVR Lab, New York University. New York
Research Intern *September 2017 - August 2018*
- Supervisors: Kyunghyun Cho and Krzysztof J. Geras.
- Worked on projects related with medical image analysis, including building models for breast cancer classification, utilizing global context to image patch classification with DNNs and incorporating patient characteristics features in DNNs for breast cancer classification.

Key Laboratory of Brain Function and Disease , Chinese Academy of Sciences. Hefei
Undergraduate Research Intern *November 2014 - July 2016*
- Supervisor: Xiaochu Zhang.
- Built feature extraction and classification toolbox of real time brain signal. Designed a Neurofeedback platform on addiction via a real-time display of craving status for potential treatment.

- Junjie Bu, Ru Ma, Nan Wu (co-first author), Xiaochu Zhang. (2015). Smoking cue-reactivity cortical connectivity reflected in EEG network in smokers. Invited for oral presentation in The 18th National Academic Congress of Psychology, Tianjin, China.
- Xiaochu Zhang, Junjie Bu, Nan Wu. BrainART. Copyright of Computer Software (ID: 2016SR202558).

Shanghai Institute of Biological Sciences, Chinese Academy of Sciences. Shanghai
Summer Research Intern, *June 2013 - July 2013*

- Supervisor: Yi-Feng Zhang.
- Assisted in experiments in the Laboratory of Retinal Neural Circuits and presented paper at weekly team meeting.

TALKS

CS 523: Research Seminar in Computer Vision and Healthcare, Stanford University. *May 2021*
Medical Imaging with Deep Learning (MIDL) 2020, Montreal Canada. *July 2020*
PhD Fellowship Summit 2020, Google. *July 2020*
The 2019 Data Science Symposium, Michigan Institute for Data Science, University of Michigan. *November 2019*
Invited lecture for Capstone Project course, NYU Center for Data Science. *September 2019*
AI for Social Good workshop, ICML 2019. *June 2019*
CILVR Seminar, NYU Computer Science. *April 2019 and February 2018*

TEACHING & GRADING EXPERIENCE

Natural Language Processing with Representation Learning, NYU Center for Data Science. *Fall 2021*
Computer Vision, Computer Science Department at NYU. *Fall 2021*
Predictive Modeling with Sports Data, NYU Center for Data Science. *Spring 2021*
Responsible Data Science, NYU Center for Data Science. *Spring 2021*
Introduction to Data Science, NYU Center for Data Science. *Fall 2019 & 2020*
Data Management and Analysis, Computer Science Department at NYU. *Fall 2018*
Machine Learning and Computational Statistics, NYU Center for Data Science. *Spring 2018*
Decision Analytics for Sports, NYU Stern. *Fall 2017*
Data Driven Decision Making, NYU Stern. *Spring 2016*

REVIEWING

NeurIPS 2021, ICML 2020, KDD 2020, IEEE Transactions on Medical Imaging, Transactions on Pattern Analysis and Machine Intelligence, NMR in Biomedicine, Information Fusion.
Serving as Program Committee Member for the ML4H workshop at NeurIPS 2020

ACTIVITIES

Women in Data Science (WiDS) NYU Center for Data Science. *2020- 2021*
Member of Executive Board.
NGIST 2020, summer school in Bioinformatics for high school students. *June 2020*
Google NYC PhD Summit. *January 2020*
2018 MIT-Mount Sinai NYC Grand Hack. *November 2018*
Mentor in Rehabilitation & Human Performance Track.
Challenge Cup, National College Student Business Plan Competition. *December 2015*
Team Member & Keynote Speaker, honored with Gold Medal.
China Youth Innovation and Entrepreneurship Venture Competition. *November 2015*
Team Member & Keynote Speaker.
APRU-University of Malaya Program: Developing Future Global Leaders of the Pacific Rim. *August 2015*
Student Ambassador of USTC.
Women soccer team of School for Gifted Young, USTC. *2012 - 2016*
Member & Captain at 2015.
Student Union of School for Gifted Young, USTC. *2012 - 2014*
Member & President at 2014.

TECHNICAL SKILLS

Programming Languages	Python, C, Java, MATLAB, R, SQL, SAS, Latex
Framework	PyTorch, Tensorflow, Keras, Hadoop, Spark, PySpark

LANGUAGES

Mandarin, mother tongue
English, fluent