

Sheng Liu

Interests: Robust Machine (Deep) Learning for Medicine

✉ shengliu@nyu.edu

homepage

☎ +1 (917) 803-6914

EDUCATION

New York University New York, NY

Ph.D. candidate at the **Center for Data Science**

Sept. 2018 – present.

New York University New York, NY

Master of Science in Data Science; GPA:3.9

Sept. 2016 – May. 2018

Shanghai University; University of California, San Diego

Bachelor of Science, Mathematic; GPA:3.9

Sept. 2012 – Jul. 2016

SELECTED PUBLICATIONS

Machine Learning

S. Liu, X. Zhang, N. Sekhar, K., Y. Wu, P. Singhal, C. Fernandez-Granda.

Avoiding spurious correlations via logit correction [[ArXiv](#)].

preprint 2022.

H. Huang*, H. Kang*, **S. Liu**, O. Salvado, T. Rakotoarivelo, D. Wang, T. Liu.

PADDLES: Phase-Amplitude Spectrum Disentangled Early Stopping for Learning with Noisy Labels [[ArXiv](#)].

preprint 2022.

J. Zhou, C. You, X. Li, K. Liu, **S. Liu**, Q. Qu, Z. Zhu.

Are All Losses Created Equal: A Neural Collapse Perspective [[ArXiv](#)].

NeurIPS 2022.

S. Liu, Z. Zhu, Q. Qu, C. You.

Robust Training under Label Noise by Over-parameterization [[ArXiv](#)].

ICML 2022 (*Spotlight presentation*).

S. Liu*, A. Kaku*, W. Zhu*, M. Leibovich*, S. Mohan*, B., L. Zanna, N. Razavian, C. Fernandez-Granda.

Deep Probability Estimation [[ArXiv](#)].

ICML 2022 (*Spotlight presentation*).

L. Yi, **S. Liu**, Q. She, A. I. McLeod, B. Wang.

On Learning Contrastive Representations for Learning with Noisy Labels [[ArXiv](#)].

CVPR 2022.

S. Liu*, K. Liu*, W. Zhu, Y. Shen, C. Fernandez-Granda.

Adaptive Early-Learning Correction for Segmentation from Noisy Annotations [[ArXiv](#)].

CVPR 2022 (*Oral presentation*).

S. Liu*, X. Li*, Y. Zhai, C. You, Z. Zhu, C. Fernandez-Granda, Q. Qu.

Convolutional Normalization: Improving Deep Convolutional Network Robustness and Training [[ArXiv](#)].

NeurIPS 2021.

S. Liu, J. Niles-Weed, N. Razavian, C. Fernandez-Granda.

Early-Learning Regularization Prevents Memorization of Noisy Labels [[ArXiv](#)].

NeurIPS 2020.

B. Bernstein, **S. Liu**, C. Papadaniil, C. Fernandez-Granda.

Sparse Recovery Beyond Compressed Sensing: Separable Nonlinear Inverse Problems [[ArXiv](#)].

IEEE Transactions on Information Theory.

Medical Applications

S. Liu, A. Masurkar, H. Rusinek, J. Chen, B. Zhang, W. Zhu, C. Fernandez-Granda, N. Razavian.

Generalizable deep learning model for early Alzheimer's disease detection from structural MRIs [[ArXiv](#)].

Nature Scientific Reports 2022

S. Liu, C. Yadav, C. Fernandez-Granda, N. Razavian.

On the design of convolutional neural networks for automatic detection of Alzheimer's disease [[ArXiv](#)].

NeurIPS 2019 Machine Learning for Healthcare (ML4H) workshop.

S. Liu, M. Cheng, H. Brooks, W. Mackey, D. Heeger, E. Tabak, C. Fernandez-Granda.

Time-Series Analysis via Low-Rank Matrix Factorization Applied to Infant-Sleep Data [[ArXiv](#)].

NeurIPS 2019 Machine Learning for Healthcare (ML4H) workshop.

PROFESSIONAL EXPERIENCES

Google Research *New York*

Sparse Overparameterization for Learning with Noisy Signals

Jan. 2022 – May. 2022

- Investigated the learning dynamic of over-parameterization models.
- Utilized the implicit bias of optimization algorithm to achieve parameter sparseness.
- Proposed a method to disentangle the clean signals from the sparse noises.

EE department, Umich *Ann Arbor, MI*

Weakly-supervised Segmentation on Nature Images

Jan.. 2021 – Aug. 2021

- Reframed a prevailing weakly supervised segmentation pipeline into segmentation with noisy annotations.
- Developed a novel approach to perform segmentation with noisy annotations.
- Conducted comprehensive experiments on benchmark datasets such as PASCAL VOC and MS COCO.

Amazon Science *Seattle, WA*

Improving Robustness of RNN-T on Rare Entities Recognition

May. 2021 – Sept. 2021

- Proposed a method to utilize text to speech based synthetic data for rare entities to train RNN-T.
- Designed a regularization term that provides consistent predictions of the encoder network.
- Achieve a relative reduction of $\sim 5\%$ in word error rate (WER) without degradation on general traffic.

AWARDS

Travel Grant. *ICML 2022*

July 2022

Travel Funding. *ML4H workshop at NeurIPS 2019*

Dec. 2019

Full Graduate Scholarship. *NYU CDS*

2018-2023

Best Research Project. *NYU CDS*

Feb. 2019

Outstanding Graduate in Shanghai. *Shanghai Municipal Education Commission*

May 2016

Chinese High School Physics Olympiad, Bronze Medal. *Chinese Physical Society, China*

2011

TALKS

TrustML young scientist seminars (UTokyo) . *Fast & slow: robust learning for probability estimation* *Sept. 5, 2022*

CVPR 2022. *Segmentation with noisy annotations*

June 12, 2022

Tsinghua University AI Time Seminar. *Robust Learning with Label Noise.*

Nov. 21, 2021

INFORMS Annual Meeting 2020. *Optimization Methods for Machine Learning.*

Nov. 11, 2020

SAIL 2020. *AI for Alzheimer's Automatic Detection.*

Oct. 2020

MAD Seminar, NYU Courant and CDS. *Separable Nonlinear Inverse Problems.*

Apr. 2020; Feb. 2018

TEACHING

Mathematical Tools for Data Science. *NYU Courant Institute*

2018 Spring

Probability and Statistics for Data Science. *NYU Center for Data Science*

2017 Fall

Signal Processing and Harmonic Analysis. *NYU Courant Institute*

2018 Fall

Signal Processing and Harmonic Analysis. *Data Mining in R, NYU Stern*

2017, 2018 Summer

TECHNICAL SKILLS

Programming Languages: Python, MATLAB, R, C, C++, SQL, Java, SAS, Latex

Tools & Libraries: Pytorch, Keras, TensorFlow, Pandas, nltk, Scikit-Learn, OpenCV

LANGUAGES

Mandarin: mother tongue

English: fluent