Mia (Miao) Zhang (She/Her)

miaozhng@nyu.edu | LinkedIn | Google Scholar | Website

EDUCATION

New York University – New York, USA *Sep. 2021 – present*
Ph.D. candidate. Computer Science, Tandon school of Engineering (Expected graduation date: December 2025)
Supervisor: Dr. Rumi Chunara
Professional areas: Generalizability and robustness of vision and multi-modal models. Unsupervised/self-supervised learning.

Stanford University – Stanford, CA, USA *Sep. 2018 – Jun. 2020*
M.A., Center for Computer Research in Music and Acoustics (CCRMA)
Supervisor: Dr. Daniel Rubin, Dr. Liangqiong Qu

Beijing University of Posts and Telecommunications – Beijing, China *Sep. 2014 – Jun. 2018*
Supervisor: Dr. Dong Wang, Dr. Lantian Li
B.Eng. Electrical and Computer Engineering, School of Electrical Engineering

PROFESSIONAL EXPERIENCE

Futurewei Technologies – Santa Clara, CA
Research scientist intern | Supervisor: Dr. Masood Mortazavi *May. 2024 – Aug. 2024*
Generative model (DDPM and MAE) based representation learning for image-speech alignment and conditional generation.

Reality Defender – New York City, NY
Research scientist intern | Supervisor: Dr. Gaurav Bharaj *May. 2023 – Aug. 2023*
Spurious correlation modeling for large vision dataset by leveraging LLM, achieving SOTA unsupervised bias mitigation results.

Amazon Web Services, Amazon – Seattle, WA
Software Engineer (full-stack): Built and maintained database migration service for cloud computing. *Sept. 2020 – Sept. 2021*

Engineering Department, Poly (Plantronics) – Santa Cruz, CA *Jun. 2019 – Dec. 2019*
Technology Strategy Intern: Personalized deep learning based speech recognition integration for headphone.

PROJECTS

Vision-speech representation learning with latent diffusion models *May. 2024 - Aug. 2024*
IC Lab, Futurewei Technologies
- Proposed joint learning of single-modal reconstruction and cross-modal alignment over Masked Auto Encoder (**MAE**) and Denoising Diffusion Probabilistic Models (**DDPM**) for image and speech.
- Improved image-speech retrieval and image conditioned audio generation task with better semantics alignment.

Causal data augmentation for improved generalizability of clinical LLM *April. 2024 - present*
The Visualization and Data Analytics Research Center (VIDA), New York University
- Fine-tuned a BERT base **large language models/LLMs** for patient readmission prediction based on clinical notes.
- Augmenting confounding variables of LLM model representation to achieve equal performance across clients.

Fair vision recognition across demographic (race/gender/age) and income groups *Jan. 2024 - present*
- Proposed to infer sensitive information by leveraging and improving pre-trained **large vision-language models**.
- Adaptive re-sampling to mitigate optimization loss across pseudo sensitive groups.

Robustness of digital data source Google Street View *Mar. 2023 - Jan. 2024*

- **Causal mediation analysis** for "big data" driven machine learning model studying effect of built environment on health outcomes. Proposed a new causal model framework that accounts for the mediator of individual-level activity and leads to 4.17 times higher health improvement than the vanilla model.

Mitigating urban and rural disparity in self-supervised learning for geographic images *June. 2022 - May. 2023*

- Mitigating performance disparities for **semantic segmentation** of geographic images in **contrastive self-supervised learning** across urban and rural areas. De-biasing algorithms applied to multi-level visual representations.

Federated learning for medical images that is robust to data heterogeneity *Nov. 2019 - Nov. 2021*

Laboratory of Quantitative Imaging and Artificial Intelligence (QIAI), Stanford University

- Proposed a **heterogeneity-robust federated learning** algorithm: Split Averaging (SplitAVG), which outperforms SOTA methods when participating clients contain highly heterogeneous data.

Speaker verification: Generalizable speaker features from human short speech. *May. 2017 – July. 2018*

Beijing National Research Center for Information Science and Technology, Tsinghua University

- **Machine speaker recognition** (SRE) via deep learning structure (**Convolutional Time-delayed Neural Network**) to learn speaker vocal traits from the raw text-independent speech signals. Publication at ICASSP conference.

PUBLICATIONS

Research Interests:

Multi-modal reasoning and perception (image/text/speech), **Algorithmic fairness and robustness** via dataset bias analysis and mitigation, and causal inference for interpretability.

Conferences:

***Miao Zhang**, Rumi Chunara. "*Mitigating urban-rural disparities in contrastive representation learning with satellite imagery". AAAI/ACM Conference on AI, Ethics, and Society (AIES) 2024.

*Junyuan Zhang, Shuang Zeng, **Miao Zhang**, Runxi Wang, Feifei Wang, Yuyin Zhou, Paul Pu Liang, Liangqiong Qu.* "FLHetBench: Benchmarking Device and State Heterogeneity in Federated Learning". Conference on Computer Vision and Pattern Recognition (CVPR) 2024.

***Miao Zhang**, Harvineet Singh, Lazarus Chok, Rumi Chunara.* "Segmenting across places: The need for fair transfer learning with satellite imagery". Fair, Data-efficient, and Trusted Computer Vision (TCV) 2022.

***Miao Zhang**, Xiaofei Kang, Yanqing Wang, Lantian Li, Zhiyuan Tang, Haisheng Dai, Dong Wang.* "Human and Machine Speaker Recognition based on Short Trivial Events". IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP), pp. 5009-5013. IEEE, 2018.

Journals:

***Miao Zhang**, Salman Rahman, Vishwali Mhasawade, Rumi Chunara.* "Utilizing Big Data Without Domain Knowledge Impacts Public Health Decision Making", Proceedings of National Academy of Sciences (PNAS).

***Miao Zhang**, Liangqiong Qu, Daniel Rubin.* "SplitAVG – A Federated Deep Learning Method to Tackle Data Heterogeneity for Medical Imaging". Journal of the Biomedical and Health Informatics, 2022.

*Liangqiong Qu, Niranjan Balachandar, **Miao Zhang**, Daniel Rubin.* "Handling Data Heterogeneity with Generative Replay in Distributed Deep Learning Models for Medical Imaging". Medical Image Analysis, 2022.

Pre-prints:

***Miao Zhang**, Rumi Chunara. "*Leveraging vision-language model for fair facial attribute classification." [link]

***Miao Zhang**, Zee Fryer, Gaurav Bharaj.* "Common-sense bias discovery and mitigation for classification tasks." [link]

***Miao Zhang**, Rumi Chunara.* "Quantifying greenspace using deep learning in Karachi, Pakistan.". EarthArXiv. (Under review at ACM Journal on Computing and Sustainable Societies) [link]

TECHNICAL SKILLS

Programming languages: Python, JavaScript, R, C++, Matlab, Bash
Software & Tools: Pytorch, OpenCV, Tensorflow, Git, VSCode, Docker

---

SERVICES & AWARDS

Reviewer for WACV2025, ML4H 2024, AISTATS 2023, CVPR-WICV 2023, Nature Communications, Journal of Biomedical Health Informatics, ISPRS Journal of Photogrammetry and Remote Sensing
Guest lecture for CS-GY 6053 Foundation of Data Science, New York University
Natural Language Processing Methods in Policy and Services Research, New York University
Invited presentation at Cohere for AI for a research work in mitigating common sense bias in ML training datasets.
Funding award for university collaboration from Adobe Research.
Future Leader Fellowship, NYU
CRA-WP Grad Cohort for Women Workshop
YOFC Enterprise Scholarship, China, *Top 1%*