

YIYANG (OLIVER) NAN

+1 217 904 9338 || nanyiyang@gmail.com || <https://nanyyyyyy.github.io>

EDUCATION

Brown University Providence, RI
M.S. in Computer Science May 2024

University of Michigan - Ann Arbor Ann Arbor, MI
B.S. in Data Science, Statistics, Mathematics of Finance Dec 2021

PROFESSIONAL EXPERIENCE

Cohere For AI Remote
Research Scholar Jul. 2024 – Present
Mentors: Sara Hooker, Matthias Gallé, Beyza Ermis

Large Multimodal Model (pre and post) training and evaluation; Data Curation; Data Synthesis

PUBLICATIONS AND MANUSCRIPTS

Conference Papers

Nihal V. Nayak, **Yiyang Nan**, Avi Trost, Stephen H. Bach. "[Learning to Generate Instruction Tuning Datasets for Zero-Shot Task Adaptation](#)". *ACL Findings*, 2024

Tianyu Yang, **Yiyang Nan**, Lisen Dai, Zhenwen Liang, Yapeng Tian, Xiangliang Zhang. "[SaSR-Net: Source-Aware Semantic Representation Network for Enhancing Audio-Visual Question Answering](#)". *EMNLP Findings*, 2024

Weiguo Pian, **Yiyang Nan**, Shijian Deng, Shentong Mo, Yunhui Guo, Yapeng Tian. "[Continual Audio-Visual Sound Separation](#)". *NeurIPS*, 2024.

Journal Papers

Hanrui Zhang, Ziyang Wang, **Yiyang Nan**, Bulat Zagidullin, Daiyao Yi, Jing Tang, Yuanfang Guan. "[Harmonizing across Datasets to Improve the Transferability of Drug Combination Prediction](#)". *Communications Biology*, 2023.

Shijian Deng, Erin E. Kosloski, Siddhi Patel, Zeke Aharon Barnett, **Yiyang Nan**, Alexander Kaplan, Sisira Aarukapalli, William T. Doan, Matthew Wang, Harsh Singh, Pamela R Rollins, Yapeng Tian. "[Hear Me, See Me, Understand Me: Audio-Visual Autism Behavior Recognition](#)". *IEEE Transactions on Multimedia*, 2024.

RESEARCH EXPERIENCE

BATS Lab, Brown University Providence, RI
Advisor: Prof. Stephen Bach Sep. 2022 – Present

Researched the generation of synthetic instruction-tuning data conditioned on specialized domain context and task type attributes.

Participated in [2022 DARPA AI for Critical Mineral Assessment Competition](#) and developed an adapted CLIP model with visual prompts to extract map features based on USGS map legend symbols.

CVMC Lab, University of Texas - Dallas Dallas, TX
Advisor: Prof. Yapeng Tian Mar. 2023 – May 2024

Collaborated on introducing continual learning problem in the context of audio-visual sound separation and investigated its issue of catastrophic forgetting.

Designed a network with source-wise learnable tokens to disentangle sound sources, extract semantically representations from inputs, and enhance performance in the Audio-Visual Question Answering

Introduced the most extensive dataset currently available for recognizing autism-related behaviors in children; Conducted benchmarking across various foundation models.

Guan Lab, University of Michigan

Advisor: Prof. Yuanfang Guan

Ann Arbor, MI

Aug. 2021 – May 2023

[Ranked in 3rd place in 2022 Heart Failure: Microbiome FINRISK DREAM Challenge](#); Utilized a non-parametric ranking to transform the host phenotype data and modeled the survival risk for heart failure with ensembles of kernel-based and tree-based models.

[Ranked in 4th place in 2022 Cough Diagnostic Algorithm for Tuberculosis \(CODA TB\) DREAM Challenge](#); Designed an LightGBM based approach with engineered cough sounds audio features and clinical information to predict the presence of Tuberculosis.

Participated in [2022 NeurIPS Weakly Supervised Cell Segmentation data challenge](#); Implemented the pipeline of Mask-RCNN with customized non-maximum suppression for whole-slide cell segmentation.

Proposed the standardization for the different concentration settings in different studies enables pharmacodynamics of monotherapies in machine learning models.

SKILLS

Programming: Python (Proficient), R (Proficient), C++ (Intermediate), Java (Basic), SQL (Basic)

Languages: English (Fluent), Mandarin (Native)