

WANCONG ZHANG

[Github](#) / [Website](#) / [LinkedIn](#) / Wancongzhang@gmail.com / (917) 484-0079

Education

- PhD in Computer Science** *09/2023-Present*
New York University – Courant Institute, New York, NY
- MS in Computer Science – GPA: 3.9/4.0** *09/2018-09/2020*
New York University – Courant Institute, New York, NY

Research / Employment

- PhD Candidate, NYU** *09/2023-Present*
Advised by Prof. Yann LeCun
- Exploring joint embedding prediction architecture (JEPA) for control. [Preprint](#)
 - Designed an efficient probing benchmark to evaluate the fitness of unsupervised visual representations for reinforcement learning. Applied it to systematically improve upon the state-of-art pretraining setup for Atari. [Paper](#)
- Senior Deep Learning Researcher, AssemblyAI** *02/2021-08/2023*
- Spearheaded the creations of state-of-the-art [Conformer-1](#), [Conformer-2](#), and realtime automatic speech recognition (ASR) models.
 - Applied self supervised learning to train ASR models for low resource languages.
 - Led teams of up to 10 researchers and engineers in designing and deploying large-scale, multi-node training pipelines on both local compute clusters and Google Cloud.
- M.S. Researcher, NYU** *09/2019-06/2020*
Advised by Profs Sam Bowman & Rajesh Ranganath
- Proposed MixUp algorithms for sentence classification using transformers, regularizing model training and reducing calibration error by 50%. [Preprint](#)
 - Built customized, fused CUDA kernels for a novel gateless RNN architecture. [Blog](#) [Git](#)
- Research Intern, University of Toronto** *06/2019-09/2019*
Advised by Prof. Marzyeh Ghassemi
- Implemented a multi-modal representation learner for healthcare data that utilizes lab vitals and medical notes to solve clinical tasks. [Paper](#)
- Software Engineer, Guardant Health** *07/2017-07/2018*
- Fullstack web development and distributed systems
- Research Assistant, Harvard & Stanford Universities** *10/2013-10/2016*
- Studied the molecular mechanisms of stem cell differentiation. [Paper](#)

Publications

Learning from Reward-Free Offline Data: A Case for Planning with Latent Dynamics Models

V Sobal*, W Zhang*, K Cho, R Balestriero, T Rudner, Y LeCun

Under review at ICML 2025 [[Preprint](#)]

Light-weight probing of unsupervised representations for Reinforcement Learning

W Zhang, A Chen, V Sobal, Y LeCun, N Carion
Reinforcement Learning Conference 2024 [[Paper](#)]

Conformer-1: Robust ASR via Large-Scale Semisupervised Bootstrapping

W Zhang*, L Chkhetiani*, F Ramirez*, Y Share, A Vanzo, M Liang, S Martin, G Oexle, R Bousbib,
T Pesach, M Nguyen, D Pulliam, D Donato
arXiv preprint 2024 [[Preprint](#)]

A Comprehensive EHR time series pre-training benchmark

M McDermott, B Nestor, W Zhang, P Szolovitz, A Goldenberg, M Ghassemi
The ACM Conference on Health, Inference, and Learning 2021 [[Paper](#)]

**Mixup training leads to reduced overfitting and improved calibration for the
transformer architecture**

W Zhang, I Vaidya
arXiv preprint 2021 [[preprint](#)]

Teaching

Graduate Student Instructor: [Deep Learning](#) (NYU)

Fall 2024