

Liyang Song

song.liy@northeastern.edu | Boston, MA | liyangsong.github.io | github.com/LiyangSong | linkedin.com/in/ly-song

Research Interests

Robust computer vision · Grounded multimodal learning · Long-form / egocentric video understanding

Domain adaptation & distribution shifts · Reliable & trustworthy ML

AI for healthcare & computational medicine (video-based physiological monitoring, infant development)

Education

Northeastern University - M.S. in Data Science (*Khoury College of Computer Sciences*)

May 2025

Advisor: **Dr. Michael Wan** (*Institute for Experiential AI*)

- **GPA:** 3.96/4.00

- **Coursework:** Machine Learning, Deep Learning, Distributed Systems, Computer Systems, Database Systems

- **Teaching Assistant:** CS5600: Computer Systems (Spring 2024)

University of Nottingham - M.S. in International Business (*Merit*)

Nov. 2017

Xiamen University - B.S. in Marine Chemistry

Jul. 2015

Publications

- **Song, L.**, Bishnoi, H., Manne, S. K. R., Ostadabbas, S., Taylor, B. J., Wan, M. “**Overcoming Small Data Limitations in Video-Based Infant Respiration Estimation**”. *2026 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Tucson, AZ, Mar. 6-10, 2026. (to appear; [arXiv: 2512.06888](https://arxiv.org/abs/2512.06888))

Research Experience

Deep Learning Researcher Intern | Research Assistant

Sep. 2024 - Present

Institute for Experiential AI, Northeastern University

Boston, MA

Advisor: **Dr. Michael Wan** (*Institute for Experiential AI*)

- **Led a project on video-based infant respiration estimation (AIR-400)**, designing temporal search and inference pipelines for long, noisy clinical video under small-data and distribution-shift constraints.
- **Designed an infant-specific pipeline** combining ROI detection, optical-flow-based motion representations, and spatiotemporal neural networks, and evaluated across infant and adult datasets under consistent subject-level splits.
- **Systematically analyzed robustness factors under severe distribution shifts**, such as classical vs. deep optical flow, infant–adult domain mismatch, and frame-rate variability, and used these insights to refine the final model design.
- Demonstrated that body-level ROIs **consistently improved model accuracy** for three of four spatiotemporal architectures on AIR-400, improving MAE by roughly 0.5–1.0 bpm and boosting correlation by ~0.1, highlighting the importance of infant-specific spatial priors.
- **Built a reproducible benchmark** with standardized cross-validation, dataset-specific post-processing (e.g., bandpass filtering, PSD-based loss), and subject-level evaluation; the pipeline and artifacts support fair comparison of new models and preprocessing variants.
- Implemented a **scalable training and inference stack in PyTorch**, including multi-GPU training (DDP), Slurm-based experiment management, caching of preprocessed data, and scripted evaluation, enabling large sweeps of models and preprocessing configurations.
- Explored the EU DREAM dataset on **robot-assisted child therapy** using 3D body skeletons, head pose, and gaze reconstructed from multi-camera JSON logs in an egocentric interaction setting, and computed body-symmetry and gaze-stability statistics to explore early autism screening and engagement-related hypotheses.
- Performed second-round annotation and **inter-rater agreement analysis** on a large infant sleep/arousal dataset; our coder pair achieved consistently above-average agreement across all metrics (F1 AUC, F1 at 0.2/0.5/0.8, and Cohen’s Kappa at 10s/60s/180s), ranking among the top annotators and helping validate the labeling protocol for downstream modeling.

Teaching Experience

Teaching Assistant - CS5600: Computer Systems

Spring 2024

Northeastern University

- Led recitations and office hours on computer systems topics (C programming, concurrency, memory management); graded programming assignments and provided feedback to 80+ students.

Industry Experience

Machine Learning Engineer Intern

Jun. 2024 - Aug. 2024

DTonomy - AI Security Automation

Cambridge, MA

- Developed an **LLM-based security assistant** using LangChain and OpenAI/Gemini APIs to summarize alerts and generate recommendations within an existing security automation platform.
- Built Express.js REST APIs to integrate the assistant with existing Node-RED flows, and connected to **AWS CloudTrail, Azure Monitor, and VPC logs** to support multi-cloud incident analysis.

Full Stack Developer

Nov. 2023 - Apr. 2024

DemocracyLab - "Tech-for-Good" Non-profit Platform

Seattle, WA

- Contributed to a Django/React platform for matching volunteers with civic-tech projects, focusing on role-based access, project workflow, and reliability at scale.
- Integrated **PostgreSQL (AWS RDS)**, S3 media storage, and Redis caching; helped improve API responsiveness and system stability under load tests.

Data Engineer

Jun. 2018 - Dec. 2020

AstraZeneca - Global Top 5 Pharmaceutical Company

Shanghai, China

- Built PySpark ETL pipelines processing tens of millions of daily sales records from 40k+ institutions and 100+ products, with logging and retry mechanisms for production robustness.
- Optimized **MySQL** queries and schemas (index tuning, partitioning), improving performance of analytics dashboards used by business stakeholders.

Open-Source Projects

AIR-400 - Infant Respiration Video Benchmark

[GitHub](#)

- Maintainer of an **open-source research codebase and experiment suite** for infant respiration estimation from thermal and RGB videos, with modular components for preprocessing, feature extraction, and spatiotemporal model training/evaluation.
- Provided **one-command scripts** to reproduce all WACV 2026 results and run inference, along with **configuration-driven experiment management**, subject-level cross-validation utilities, and links to trained models and dataset downloads for AIR-125 / AIR-400 / COHFACE.

DREAM-ASD-Toolkit - Multimodal Skeleton & Gaze Analysis

[GitHub](#)

- Implemented a toolkit for parsing and visualizing **3D child skeletons, head pose, and gaze** from robot-assisted and therapist-only therapy sessions in the DREAM dataset, with scripts for session-level visualizations and feature summaries (limb symmetry, motion, posture stability).
- Integrating gaze vectors with an approximate scene model of interaction targets (robot, therapist, tabletop) to estimate attention distributions, gaze stability, and shifts between **social and task-focused attention**.

Honors

Alfond Scholars Initiative Award (\$25,000, 2022-2024); Graduation with Merit (2017); Excellent Student Leader (2014); National Student Grant (2012-2013)

Skills

Languages: Python, Java, C/C++, SQL, TypeScript, JavaScript, R

Machine Learning & Vision: PyTorch, NumPy, SciPy, scikit-learn, Pandas, OpenCV, Matplotlib

Data & Systems: PySpark, MySQL, PostgreSQL, Linux, Slurm (HPC), Docker, Git

Backend & Web: Python (FastAPI, Django), Java (Spring Boot, Spring Cloud), Node.js / Express.js, React, LangChain

Cloud & DevOps: AWS (EC2, S3, RDS, SQS, Lambda), Redis, Jenkins, JMeter