SIQI (LOLA) WANG

Cambridge, MA | (530)-231-0945 | lolawanggz@gmail.com | LinkedIn | Portfolio

EDUCATION

Harvard T.H. Chan School of Public Health, Boston, MA — Aug 2023 - May 2024 (Expected)
MS in Health Data Science — GPA: 3.9/4.0
Coursework: Artificial Intelligence, Advanced Computer Vision, Quantitative Methods for Natural Language Processing, Statistical Learning, Computing for Big Data, Systems Development for Computational Science

University of California, Davis, Davis, CA — Sep 2019 - Jan 2023
B.S. Statistics | *B.A Biological Sciences*, Minor: Computer Science, Dean's Honor List — GPA: 3.8/4.0

SKILLS

Languages: Python (Pytorch, TensorFlow, OpenCV, scikit-learn), R (dplyr, tidyverse, ggplot2, Shiny), SQL
Tools: AWS, Azure, Git, Linux, Docker, Spark, Hive, MATLAB, Tableau, Excel, Slurm/cluster

PROFESSIONAL EXPERIENCE

Comprehensive CT Segmentation Processing Pipeline — Oct 2023 - Present
Research Assistant | Harvard T.H. Chan School of Public Health (CELEHS Lab)

- Built an **end-to-end** image processing **pipeline** for large-scale data, achieving **0.82** size Dice score for target segmentation.
- Leveraged **LLM** to extract keywords from unstructured text and combined text-image data for downstream predictive analytics using multimodal models (**CLIP**), improving inference accuracy by **18%**.
- Evaluated a **weakly supervised** segmentation model that incorporated tumor size information enhance model performance.

Self-supervised Anomaly Detection for Brain MRI — Oct 2023 - Present
Research Assistant | Harvard T.H. Chan School of Public Health (Beam Lab)

- Developed a **foundational model** for **self-supervised anomaly detection**, enhancing efficiency and inference speed by leveraging symmetry-based masking in the preprocessing step to reduce computational overhead.
- Improved model inference performance, achieving **98.71% sensitivity** at a 95% threshold, utilizing high-dimensional embeddings from difference maps and Mahalanobis distance thresholding for precise anomaly classification.
- Fine-tuned the model on the **BraTS** and **ADNI** datasets, conducting **scaling tests and few-shot learning** to evaluate generalizability and improve adaption to new datasets.

Philips North America, Cambridge, US — May 2024 – Aug 2024
Medical Image Analysis R&D Intern | Image Guided Therapy Department

- Developed and optimized an end-to-end deep learning pipeline integrating **YOLOv8** and **SAM** for real-time segmentation.
- Achieved a **Dice coefficient of 0.85** by leveraging **transfer learning** and limited annotated data, improving inference accuracy.
- Contributed to the creation of a **beta product** with advanced features, integrating advanced features and collaborating with **product and engineering teams** to transition the model to proof-of-concept deployment.

Microbiome-Driven Model for Breast Cancer Treatment Response Prediction — Aug 2023 - Apr 2024
Research Assistant | Guangdong Provincial People's Hospital

- Developed deep learning models for automated image segmentation and feature extraction, using 3D U-Net to process large-scale medical imaging datasets, achieving Dice coefficients of **0.88**.
- Built a **multi-modal** fusion model combining image embeddings and structured microbiome data, improving prediction accuracy (AUC **0.91**)
- Implemented feature attribution using Shapley analysis to enhance **model interpretability**, identifying 19 key features that drive predictions, enabling trustworthy AI insights for downstream decision-making.

Kuaishou Technology, Beijing, China — Mar 2023 - Jul 2023
Data Science Intern | E-commerce Department

- Analyzed **merchant (20K+)** revenue metrics and **customer (700K+)** behavior patterns, using **SQL** and **Python** to optimize business strategies, ultimately increasing **revenue growth by 28%**.
- Developed and deployed **A/B testing** frameworks to optimize **exposure distribution models**, working with engineering and product teams to refine real-time systems and enhance operational efficiency, leading to a **15%** improvement in user engagement.
- Designed and maintained 7 **automated Power BI dashboards**, providing real-time analytics on key performance metrics like GMV, enhancing **performance monitoring** and decision support for business leaders.
- Applied **data analysis** to tackle high return rates, performing in-depth root cause analysis and using data-driven insights to propose model improvements that boosted **merchant satisfaction by 35%**.