Yuliang(Leo) Chen

805-637-4650 | yuc204@ucsd.edu | GitHub | Portfolio | San Diego, CA

Education

University of California, San Diego — San Diego, California
M.S. in Data Science — *Sep 2023 – Jun 2025*
GPA: 3.97/4.00

University of California, Santa Barbara — Santa Barbara, California
B.S. in Statistics – Data Science | B.S. in Mathematics — *Sep 2018 – Jun 2022*
GPA: 3.43/4.00

Relevant coursework: Computer Vision, Natural Language Processing, Data Mining, Algorithm, Distributed Computing, Advanced Machine Learning, Interpretable ML, Probability Theory, Convex Optimization, Linear Algebra, Abstract Algebra, Mathematical Analysis

Publication

Toward Foundation Model for Multivariate Wearable Sensing of Physiological Signals
Yunfei Luo, **Yuliang Chen (Co-first author)**, Asif Salekin, Tauhidur Rahman.*(submitted to ICLR 2025)*

Experience

Graduate Student Researcher — San Diego, California
Advisor: Jingjing Zou — *July 2024 – Present*

- Implemented self-supervised learning algorithms to tackle cardiovascular disease research challenges.
- Constructed dashboards to expose representational behavior in MIM models trained on synthesized dataset.
- Established theoretical guarantees for MIM models to ensure their robustness and effectiveness in downstream classification tasks.

Graduate Student Researcher — San Diego, California
Advisor: Tauhidur Rahman — *Mar 2024 – Present*

- Developed a self-supervised learning algorithm to produce learned wearable signal representation for use in few-shot and zero-shot healthcare application.
- Proposed modality- and number-agnostic fusion for representation learning with arbitrary wearable signal input.
- Trained foundation models distributed across multiple GPUs on a Kubernetes-based cluster using large-scale, self-curated datasets.
- Researched fine-grained health acoustic event detection from audio.

Micro Ingredients — Montclair, California
Supply Chain Analyst — *Oct 2022 – Sep 2023*

- Led a team of three to implement and deploy a Random Forest model for demand forecasting, achieving 83% accuracy and improving product availability by 27%.
- Collaborated with software engineers to establish a robust database system using Amazon RDS Aurora.
- Developed an interactive Python-based Streamlit dashboard to provide data-driven insights for strategic decision-making.

Projects

E-StyTR2: Efficient Image Style Transfer with Transformers | *GitHub*
Spring 2024

- Developed efficient attention-based fusion for style transfer.
- Evaluated common neural style transfer models (CycleGAN, Pix2Pix, Stable Diffusion, StyTR2) using Fréchet inception distance (FID).

Vivid Panels: Deep Neural Networks for Manga Colorization | *GitHub*
Spring 2024

- Researched GAN- and diffusion-based generative models for large-scale image inpainting and colorization.

- Developed a CRAFT-based manga panel extractor to bridge the distribution gap for effective use of a pre-trained CycleGAN backbone.

VitT: Vision-Topological Transformer for Medical Image Classification | *GitHub*
Spring 2024

- Researched topological tools and algorithm (persistent homology and persistence diagrams) for characterizing feature under complex data (image, graph, point sets).
- Designed a dual-branch topology-transformer framework for empirical improvement of medical image classification.

Foundation Model On Retinal Images using Masked Autoencoders | *GitHub*
Fall 2023, Winter 2024

- Curated large-scale (100k+) medical image datasets for foundation model training and evaluation.
- Developed a foundation model for fundus images using Masked Autoencoders (MAE) for medical image segmentation and classification.

Image-to-Image Retrieval with CLIP | *GitHub*
Winter 2024

- Evaluated common vision models (ResNets, Inception, VGG) for image representation extraction.
- Developed an image-to-image retrieval algorithm using the CLIP image encoder to enhance semantic retrieval.

Skills

Languages: Python, SQL, R, C/C++

Technologies: PyTorch, TensorFlow, Kubernetes, Docker, Ray, Git, Jupyter Notebook, OpenCV, MySQL