Chen Liu

(570)939-7703 | chen.liu.cl2482@yale.edu | chenliu-1996.github.io | github.com/ChenLiu-1996 | linkedin.com/in/chenliu1996 | 3rDjnykAAAAJ

Education

Yale University, Ph.D. in Computer Science (AI/ML Specialization). Advised by Smita Krishnaswamy. *New Haven, CT. Aug 2022 ~ Dec 2027*

Research Areas I study how to **organize latent representations** of neural networks through the lens of ♠ manifold learning, with the goal of understanding and utilizing the structure and geometry emerged in learned representations. My works are often applied to **high-dimensional** data such as images, videos, molecules, or proteins. I also explore ♡ dynamics modeling, ♣ multimodal fusion, and ♢ self-supervised learning.

Teaching Deep Learning on Graphs (Rex Ying). AI Foundation Models (LLM) (Arman Cohan). Deep Learning (my advisor).

Columbia University, M.S. in Electrical Engineering *New York, NY. Aug 2018 ~ Dec 2019*

Nikola Tesla Electrical Engineering Scholar — "to the most exceptional applicants", top 10% among those admitted.

Bucknell University, B.S. in Electrical Engineering *Lewisburg, PA. Aug 2014 ~ May 2018*

Experience

Ph.D. Research @ Yale University *New Haven, CT. Aug 2022 ~ Dec 2027*

- ♠ ♢ **Observation-driven improvement on transformers' optimization objective:** Discovered a model-scale-specific embedding condensation effect in transformers, and designed a series of embedding dispersion objectives to improve LLM generalization. Paper #1
- ♠ **Biological sequence optimization on a plausible landscape:** Instead of generation from scratch, proposed a property-guided on-manifold Langevin dynamics to optimize from an existing biological sequence, with all intermediate steps being plausible. Paper #2
- ♠ ♣ ♢ **Multimodal learning of biological data:** Several projects that combine biological modalities such as sequence, structure, and properties. Designed a novel cancer-wildtype contrastive learning objective to establish a new SOTA in cancer immunogenicity. Paper #3
- ♠ ♡ **Forecasting spatial-temporal progression over irregularly-sampled sparse "videos":** Proposed position-parameterized neural differential equations to flow the trajectories of latent representations of irregularly-sampled longitudinal medical images. Paper #5
- ♠ ♢ **Unsupervised multigranular image segmentation:** Extracted latent representation of images with intra-image contrastive learning and a local reconstruction objective, and coarse-grained into multiscale segments using diffusion condensation. Paper #11

Senior Research Scientist @ GE Healthcare *San Ramon, CA. Aug 2021 ~ Jul 2022*

Research and development of deep learning solutions in medical imaging.

- **Keypoint detection:** Designed an adversarial objective to improve detection of anatomical landmarks in X-ray images. Patent #P1
- **Image classification:** Classified X-ray images on whether they contain unwanted external objects. Patent #P2

Research Software Engineer @ Matic *Palo Alto, CA. Jan 2021 ~ Jun 2021*

In a team of 3, developed from scratch a SLAM system using only visual input in Rust language, running 30 times faster than the SOTA ORB-SLAM.

Research Assistant, Fully Funded by Grant @ Columbia University Medical Center *New York, NY. Dec 2019 ~ Nov 2020*

- **Image-to-image translation & downstream analysis:** Image synthesis to bypass harmful contrast agents in MRI. Paper #13
- **Signal processing & signal registration:** Designed and developed a software for MR spectroscopy processing. Paper #15

Achievements and Services

2022~ **Outstanding/Top Reviewer,** NeurIPS 2024 (top 9%), ICML 2022 (top 10%).

2021~ **Reviewer,** Conferences NeurIPS, ICLR, ICML, MICCAI, ICASSP. Journals TMLR, IEEE TNNLS, IEEE TIP.

— [Git] **CitationMap, the first free tool** to visualize your Google Scholar citations. Personal project with >600 stars.

Skills

Research Deep Learning, Manifold Learning, Self-Supervised Learning, Computer Vision, Geometric Learning, Medical Image Analysis.

Programming Python (PyTorch, NumPy, Matplotlib, Pandas, TensorFlow, etc.), LaTeX, Git.

Selected Publications and Patents

Titles may be paraphrased. ★ First or co-first author ♠ Manifold learning ♡ Dynamics ♣ Multimodal ♢ Self-supervised ✦ LLM

1. ★ ♠ ♢ ✦ [Prelim] "Dispersing embeddings in Transformer layers improves generalization of language models". *TAG-DS 2025 (**Talk**).*
2. ★ ♠ [PDF] "RNAGenScape for property-guided mRNA sequence optimization and interpolation". *ICML GenBio 2025 (**Oral**).*
3. ★ ♠ ♣ ♢ [PDF] [Git] "ImmunoStruct enables multimodal deep learning for immunogenicity prediction". ***Nature Machine Intelligence**.*
4. ★ ✦ [PDF] "Creativity or brute force? Using Brainteasers as a window into the problem-solving abilities of LLMs". ***NeurIPS** 2025.*
5. ★ ♠ ♡ ♢ [PDF] [Slides] [Git] "ImageFlowNet: forecasting trajectories of disease progression in sparse videos". *ICASSP 2025 (**Oral**).*
6. ★ ♠ ♢ [PDF] [Slides] [Git] "DiffKillR: killing and recreating diffeomorphisms for annotation in microscopy images". *ICASSP 2025 (**Oral**).*
7. ✦ [PDF] "CourtReasoner: can LLM agents reason like judges?" ***EMNLP** 2025.*
8. ♠ ♢ [PDF] "GAGA for warped Riemannian metric learning and generative modeling on manifolds". ***AISTATS** 2025.*
9. ♠ ♡ [PDF] "Revealing dynamic temporal trajectories and underlying regulatory networks with Cflows". ***Nature Genetics**, under review.*
10. ♡ ♣ [PDF] "Integrating static and dynamic variables from EHR for time-to-death prediction using ODE-RNN". *Scientific Reports* 2025.
11. ★ ♠ ♢ [PDF] [Git] "CUTS: contrastive and topology learning for multigranular unsupervised image segmentation". ***MICCAI** 2024.*
12. ★ ♠ [PDF] [Git] "Diffusion Spectral Entropy for mutual information estimation in deep neural nets". *IEEE CISS 2024.*
13. ★ ♣ [PDF] [Git] "Deep learning of MRI contrast enhancement in aging and Alzheimer's disease brains". *Frn. Aging Neuros. 2022.*
14. ★ [PDF] [Git] "Segmentation with an edge-enhancement approach preserves cell shape features". *IEEE EMBC 2022.*
15. ★ [PDF] [Git] "JET: automated J-difference-edited magnetic resonance spectra processing". *ISMRM 2021.*

P1 [Patent] "Quantification and measurements of X-ray image by estimating key anatomical locations". *US Patent App.* GE Healthcare.

P2 [Patent] "X-ray lead marker detection system for X-ray imaging system". *US Patent.* GE Healthcare.