Phone: (530)407-0090
Email: jchen378@buffalo.edu
Page: puar-playground.github.io

Jian Chen

Ph.D. Candidate

Google Scholar: Jian Chen
GitHub: puar-playground
LinkedIn: jian-chen-1a0b9a11b

I am a Ph.D. candidate in Computer Science at the University at Buffalo, expecting to graduate soon. My research focuses on large language models, generative models, and multimodal representation learning. I have published extensively at top-tier conferences, including NeurIPS, ICLR, EMNLP, and CVPR. I am seeking full-time opportunities where I can leverage my expertise to drive innovation and contribute to the advancement of technologies.

Research Interests

Large Multimodal Language Models, Generative Models, Representation Learning

Education

Ph.D. in Computer Science, *University at Buffalo, USA* 09/2018 — Present
M.S. in Electrical Enginnering, *Drexel University, USA* 09/2015 — 06/2017
B.S. in Applied Mathematics, *Hunan University, China* 09/2011 — 07/2015

Skills

Tools and Programming Languages: Python, PyTorch, C++, R, Matlab, LaTeX, Markdown

Selected Research Projects

Intern Research Scientist / Mentor: Ruiyi Zhang / Adobe **05/2024 — 11/2024**

- Developed LoCAL, a method that enables multimodal large language models for multi-page document understanding. LoCAL is fine-tuned based on the Phi-3-V, Paligemma, and InternVL2. With only 4B parameters, LoCAL rivals Google's Gemini-1.5-pro and Claude 3 Opus, achieving state-of-the-art accuracy in open-source models. (submitted to ICLR 2025 [1])
- Created the Multi-Modal Reading (MMR) benchmark for evaluating Vision-Language Models' reading ability and tested multiple baseline models. (submitted to ACL ARR 2024 [2])

Research Assistant / Advisor: Changyou Chen / University at Buffalo **09/2022 — 06/2023**

- Developed TextLap, a novel model for text-based layout planning, and created the InsLap dataset for document layout generation. Fine-tuned large language models to achieve state-of-the-art performance on graphical design benchmarks. EMNLP 2024 [**?**]
- Contributed to the development of LLaVA-Read, responsible for part of the experimental work. LLaVA-Read is a multimodal large language model that uses dual visual encoders and a visual text encoder to enhance the understanding of text-rich images. (submitted to ICLR 2025 [3])
- Contributed to the creation of the document dataset used for training LaRA, a multimodal language model that enables reading capability through OCR input. CVPR 2024 [4]
- Developed LACE, a continuous diffusion model using constrained optimization for layout generation and editing. LACE is a unified model that integrates multiple conditional inputs, achieving state-of-the-art performance across several datasets. ICLR 2024 [5]
- Developed LRA-Diffusion, a method for learning from noisy labels by framing data labeling as a conditional generation problem. LRA-Diffusion employs a diffusion model to learn the label generation process, using pre-trained encoders as conditional inputs to mitigate overfitting. It has consistently ranked first on three leaderboards on Papers with Code. NeurIPS 2023 [6]

Research Assistant / Jacobs School of Medicine and Biomedical Sciences / UB **09/2018 — 09/2022**

- Designed TimedHN, a causal inference model capturing gene dependencies during cancer progression. The model uses binary mutation profiles as states and progression time as a latent variable, constructing a continuous-time Markov chain. An efficient optimization algorithm leveraging data sparsity was also developed to improve performance. [7]
- Proposed a neural network architecture replacing convolution with differentiable approximate string matching for microbial DNA and RNA sequence representation learning. This led to AsMac, an efficient large-scale sequence comparison method. [8]

Professional and Volunteer Service

Peer Reviewer

- Conference on Computer Vision and Pattern Recognition (CVPR) 2025
- International Conference on Learning Representations (ICLR) 2025
- International Conference on Machine Learning (ICML) 2024
- Transactions on Machine Learning Research (TMLR) 2024
- IEEE Transactions on Emerging Topics in Computational Intelligence (TETCI) 2024

Volunteer Experience

ACM Conference on Bioinformatics, Computational Biology, and Health Informatics — Niagara Falls, NY, 09/2019
WE16, Society of Women Engineers's Annual Conference — Philadelphia, PA, 10/2016
Workshop on tensor optimization and Application — Changsha, Hunan, China, 05/2015

Awards

Best Graduation Thesis: "Low-rank tensor optimization for video image recovery." — Hunan University, 2015

Selected Publications

1. **Jian Chen**, Ruiyi Zhang, Yufan Zhou, Tong Yu, Jiuxiang Gu, Ryan A. Rossi, Changyou Chen, and Tong Sun. LoRA-Contextualizing Adaptation of Large Multimodal Models for Multi-page Document Understanding. 2024. under review
2. **Jian Chen**, Ruiyi Zhang, Yufan Zhou, Ryan Rossi, Jiuxiang Gu, and Changyou Chen. MMR: Evaluating Reading Ability of Large Multimodal Models. *arXiv preprint arXiv:2408.14594*, 2024
3. **Jian Chen**, Ruiyi Zhang, Yufan Zhou, Jennifer Healey, Jiuxiang Gu, and Changyou Chen. TextLap: Customizing Language Models for Text-to-Layout Planning. In *EMNLP Findings*, 2024
4. Ruiyi Zhang, Yufan Zhou, **Jian Chen**, Jiuxiang Gu, Changyou Chen, and Tong Sun. LLaVA-Read: Enhancing Reading Ability of Multimodal Language Models. *arXiv preprint arXiv:2407.19185*, 2024
5. Ruiyi Zhang, Yanzhe Zhang, **Jian Chen**, Yufan Zhou, Jiuxiang Gu, Changyou Chen, and Tong Sun. TRINS: Towards Multimodal Language Models That Can Read. *CVPR*, 2024
6. **Jian Chen**, Ruiyi Zhang, Yufan Zhou, and Changyou Chen. Towards Aligned Layout Generation via Diffusion Model with Aesthetic Constraints. *ICLR*, 2024
7. **Jian Chen**, Ruiyi Zhang, Tong Yu, Rohan Sharma, zhiqiang xu, Tong Sun, and Changyou Chen. Label-Retrieval-Augmented Diffusion Models for Learning from Noisy Labels. *NeurIPS*, 2023
8. **Jian Chen**. Timed hazard networks: Incorporating temporal difference for oncogenetic analysis. *PLOS ONE*, 2023
9. **Jian Chen**, Le Yang, Lu Li, Steve Goodison, and Yijun Sun. Alignment-free Comparison of Metagenomics Sequences via Approximate String Matching. *Bioinformatics Advances*, 2022
10. Jiayu Qin, **Jian Chen**, Rohan Sharma, Jingchen Sun, and Changyou Chen. A probability contrastive learning framework for 3d molecular representation learning. In *NeurIPS*, 2024