Jian Chen
Ph.D. Candidate

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

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, AI for Music

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 [3]

- 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 [4])

- 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 [5]

- 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 [6]

- 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 [7]

Research Assistant / MBZUAI 10/2023 — 02/2024
- Conducted research on methods for learning knowledge graph representations using conditional generative diffusion models.

Research Assistant / University at Buffalo 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. [8]
- 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. [9, 10]

Research Assistant / Drexel University 07/2017 — 02/2018
- Benchmarked 16S rRNA sequence embedding methods and implemented a k-mer-based method in C++, enhancing computational efficiency. [11]

PROFESSIONAL AND VOLUNTEER SERVICE

Peer Reviewer

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

PUBLICATIONS AND MANUSCRIPTS

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

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*, 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*, 18(3):e0283004, 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. **Jian Chen**, Le Yang, Lu Li, and Yijun Sun. Predicting alignment distances via continuous sequence matching. *bioRxiv*, 2020

11. Stephen Woloszynek, Zhengqiao Zhao, **Jian Chen**, and Gail L Rosen. 16S rRNA sequence embeddings: Meaningful numeric feature representations of nucleotide sequences that are convenient for downstream analyses. *PLoS computational biology*, 15(2):e1006721, 2019

12. Jiayu Qin, **Jian Chen**, Rohan Sharma, Jingchen Sun, and Changyou Chen. A probability contrastive learning framework for 3d molecular representation learning. In *NeurIPS*, 2024

13. Ming Li, Ruiyi Zhang, **Jian Chen**, Jiuxiang Gu, Yufan Zhou, Wanrong Zhu, Tianyi Zhou, and Tong Sun. Docgroundingbench: Evaluating mllms' grounding capabilities on text-rich image QA, 2024

14. Rohan Sharma, Shijie Zhou, **Jian Chen**, Kaiyi Ji, and Changyou Chen. Discriminative adversarial unlearning, 2024

15. Jingchen Sun, Rohan Sharma, **Jian Chen**, Zihao Lin, and Changyou Chen. A unified optimization objective for vision-language model adaptation, 2024