

# Yiren Jian

## PhD Student

Computer Science, Dartmouth College

✉ yiren.jian.gr@dartmouth.edu | 🏠 yiren-jian.github.io | 📷 yiren-jian

My research spans a wide range of topics within the field of machine learning, focusing on computer vision, natural language processing and computational science. Specifically, my work involves the pre-training of generative models for image-text and video-text tasks, the investigation of multi-modal sentence embeddings, the exploration of few-shot learning in both vision and language domains, as well as the application of AI in scientific contexts.

## Professional Experience

### Research Scientist Intern, ByteDance/TikTok AML

Bellevue, WA

Develop generative video-language models for video summarization.

June. 2023 - Sept. 2023

Advisor: Yunzhe Tao

### Research Scientist Intern, Snap Research

Palo Alto, CA

Develop neural speakers for affective image captioning.

June. 2022 - Sept. 2022

Advisor: Panos Achlioptas

### Research Scientist Intern, NEC Labs America

Princeton, NJ

Develop physical-augmented ML models for precision immunotherapy.

June. 2021 - Sept. 2021

Advisor: Renqiang (Martin) Min

## Education

### Dartmouth College

Hanover, NH, USA

Doctor of Philosophy - Computer Science

Sept. 2018 - Jan. 2024 (expected)

Advisor: Soroush Vosoughi

### The George Washington University

Washington DC, USA

Master of Science - Physics

Sept. 2015 - May. 2017

Advisor: Chen Zeng

### Huazhong University of Science and Technology

Wuhan, China

Bachelor of Science - Physics

Sept. 2011 - May. 2015

Thesis advisor: Yi Xiao

## Publications

### [P17] SimVLG: Simple and Efficient Pretraining of Visual Language Generative Models

preprint, 2023

Yiren Jian, Tingkai Liu, Yunzhe Tao, Soroush Vosoughi, Hongxia Yang

### [P16] Bootstrapping Vision-Language Learning with Decoupled Language Pre-training

Advances in Neural Information Processing Systems, 2023

Yiren Jian, Chongyang Gao, Soroush Vosoughi

### [P15] Knowledge from Large-Scale Protein Contact Prediction Models Can Be Transferred to the Data-Scarce RNA Contact Prediction Task

preprint, 2023

Yiren Jian<sup>†</sup>, Chongyang Gao, Chen Zeng, Yunjie Zhao, Soroush Vosoughi<sup>†</sup>

### [P14] Evaluating Native-like Structures of RNA-protein Complexes Through the Deep Learning Method

Nature Communications, 2023

Chengwei Zeng<sup>#</sup>, Yiren Jian<sup>#</sup>, Soroush Vosoughi, Chen Zeng, Yunjie Zhao

### [P13] Non-Linguistic Supervision for Contrastive Learning of Sentence Embeddings

Advances in Neural Information Processing Systems, 2022

**Yiren Jian**, Chongyang Gao, Soroush Vosoughi

### [P12] T-Cell Receptor-Peptide Interaction Prediction with Physical Model Augmented Pseudo-Labeling

In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022

**Yiren Jian**, Erik Kruus, Martin Renqiang Min

### [P11] Embedding Hallucination for Few-shot Language Fine-tuning

In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

**Yiren Jian**, Chongyang Gao, Soroush Vosoughi

### [P10] Contrastive Learning for Prompt-based Few-shot Language Learners

In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

**Yiren Jian**, Chongyang Gao, Soroush Vosoughi

### [P9] Label Hallucination for Few-Shot Classification

Proceedings of the 36th AAAI Conference on Artificial Intelligence, 2022

**Yiren Jian**, Lorenzo Torresani

### [P8] MetaPix: Domain Transfer for Semantic Segmentation by Meta Pixel Weighting

Image and Vision Computing, 2021

**Yiren Jian**, Chongyang Gao

### [P7] Task Meta-Transfer from Limited Parallel Labels

Meta-Learning Workshop at Neural Information Processing Systems, 2020

**Yiren Jian**, Karim Ahmed, Lorenzo Torresani

### [P6] DIRECT: RNA Contact Predictions by Integrating Structural Patter

BMC Bioinformatics, 2019

**Yiren Jian**, Xiaonan Wang, Jiadi Qiu, Huiwen Wang, Zhichao Liu, Yunjie Zhao, Chen Zeng

### [P5] Trace, Machine Learning of Signal Images for Trace-Sensitive Mass Spectrometry: A Case Study from Single-Cell Metabolomics

Analytical Chemistry, 2019

Zhichao Liu, Erika P. Portero, **Yiren Jian**, Yunjie Zhao, Rosemary M. Onjiko, Chen Zeng, Peter Nemes

### [P4] Design of Tat-activated Cdk9 Inhibitor

Journal of Peptide Research and Therapeutics, 2019

Yunjie Zhao, Hao Chen, Chenghang Du, **Yiren Jian**, Haotian Li, Yi Xiao, Mohammed Saifuddin, Fatah Kashanchi, and Chen Zeng

### [P3] Rbind: Computational Network Method to Predict the Binding Sites of RNA Molecules

Bioinformatics, 2018

Kaili Wang<sup>#</sup>, **Yiren Jian**<sup>#</sup>, Huiwen Wang, Chen Zeng and Yunjie Zhao

### [P2] Computational Study of Non-catalytic T-loop Pocket on CDK Proteins for Drug Development

Chinese Physics B, 2017

Huiwen Wang, Kaili Wang, Zeyu Guan, **Yiren Jian**, Ya Jia, Fatah Kashanchi, Chen Zeng, Yunjie Zhao

### [P1] Network Analysis Reveals the Recognition Mechanism for Dimer Formation of Bulb-type Lectins

Scientific Report, 2017

Yunjie Zhao<sup>#</sup>, **Yiren Jian**<sup>#</sup>, Zhichao Liu, Hang Liu, Qin Liu, Chanyou Chen, Zhangyong Li, Lu Wang, H. Howie Huang, Chen Zeng

## Patents

---

T-cell receptor repertoire selection prediction with physical model augmented pseudo-labeling (NEC Labs America)

## Academic Services

---

- 2024 **Invited Reviewer**, International Conference on Learning Representations
- 2024 **Invited Reviewer**, Annual AAAI Conference on Artificial Intelligence
- 2023 **Invited Reviewer**, Northern European Journal of Language Technology
- 2023 **Invited Reviewer**, IEEE/CVF Winter Conference on Applications of Computer Vision
- 2023 **Invited Reviewer**, Conference on Neural Information Processing Systems
- 2023 **External Reviewer**, Annual Meeting of the Association for Computational Linguistics

## Teaching Experience

---

- 2019, 2021 **Graduate Teaching Assistant**, Deep Learning (graduate-level course) *Dartmouth*
- 2021-2023 **Graduate Teaching Assistant**, Machine Learning (graduate-level course) *Dartmouth*
- 2018 **Graduate Teaching Assistant**, Machine Learning (graduate-level course) *Dartmouth*
- 2015-2016 **Graduate Teaching Assistant**, University Physics (undergraduate course) *GWU*