

XIAO WANG

✉ wang3702@purdue.edu · in Xiao Wang · GitHub · Website · Google Scholar

EDUCATION

Purdue University, West Lafayette, U.S. Aug, 2018 – Present

Ph.D. in Computer Science GPA: **3.91/4.0**

Area: Computational Biology, Machine Learning, Self-Supervised Learning

Xi'an JiaoTong University, Xi'an, China Aug, 2014 – June, 2018

B.S. in Computer Science GPA: 90.57/100; Rank: **1/170**

Area: Intelligent Transportation, Machine Learning

Graduated with highest honors, Top 10 Outstanding Student (10/4000+)

RESEARCH INTERESTS

Computational Biology

- 1) Macromolecular structure modeling and evaluation from experimental data (eg. Cryo-EM).
- 2) Macromolecular structure prediction and evaluation from sequence (gene) information.
- 3) Function of macromolecule prediction via sequence and structure information.

Machine Learning

- 1) Development of *novel, efficient* and *general* Self-Supervised Learning (SSL) algorithms (eg. adversarial SSL, asymmetrical SSL).
- 2) Applications of self-supervised learning on biology, medicine, health.

HONORS AND AWARDS

NSF MolSSI Software Fellowship (10 across U.S, \$80,000) 2022

Chiang Chen Overseas Fellowship (10 across China, \$50,000) 2018

HIWIN Outstanding Student Scholarship (Top 0.3% , \$10,000 CNY) 2018

Top 10 Outstanding Undergraduate of Xi'an Jiaotong University 2017

National Scholarship (Top 1%, \$ 8,000 CNY) 2016

EXPERIENCE

Facebook AI Research, Menlo Park, U.S May, 2021 – Aug, 2021

Research Scientist Intern Advisor: Dr. Xinlei Chen, Dr. Yuandong Tian, Haoqi Fan

Asymmetrical self-supervised learning

Kihara Bioinformatics Lab, West Lafayette, U.S Aug, 2018 – Present

Ph.D. Research Assistant Advisor: Prof. Daisuke Kihara

Macromolecule structure prediction, modeling and evaluation via machine learning algorithms

JD AI Research, Mountain View, U.S May, 2020 – Dec, 2020

Research Scientist Intern Advisor: Dr. Jingen Liu, Prof. Jiebo Luo

Temporal video event segmentation via self-supervised learning

Futurewei AI Lab, Bellevue, U.S May, 2019 – August, 2019

Research Scientist Intern Advisor: Prof. Guojun Qi, Prof. Jiebo Luo

Pre-task self-supervised learning and adversarial self-supervised learning.

Tsinghua University, Beijing, China

June, 2016 – Jan, 2019

Research Intern Advisor: Prof. Li-Li, Prof. Fei-Yue Wang

Car-following behavior prediction and traffic simulation by deep learning.

Institute of Automation, Chinese Academy of Sciences, Beijing, China June, 2016 – June, 2017

Research Intern Advisor: Dr. Xiwei Liu, Prof. Fei-Yue Wang

Intelligent education by deep learning.

SELECTED PUBLICATIONS

Computational Biology

Pre-print

- 1 **Xiao Wang**, Genki Terashi, & Daisuke Kihara. “Cryo-READ: DNA/RNA dE novo Atomic structure modeling in cryo-EM maps with deep learning”. *Science*. (2022). (Submitted).
- 2 Genki Terashi, **Xiao Wang**, Sai Raghavendra Maddhuri Venkata Subramaniya & Daisuke Kihara. “Deep-Mainmast: De novo Protein Structure Modeling for cryo-EM with Deep Learning”. *Nature Methods*. (2022). (Submitted).
- 3 Genki Terashi, **Xiao Wang** & Daisuke Kihara. “Protein Model Refinement for Cryo-EM Maps Using DAQ score”. *Acta Crystallographica Section D: Structural Biology*. (2022). (Submitted). [Paper] [GitHub][Colab]

Published

- 1 **Xiao Wang***, Genki Terashi*, Sai Raghavendra Maddhuri Venkata Subramaniya, John J. G. Tesmer Daisuke Kihara. “Residue-Wise Local Quality Estimation for Protein Models from Cryo-EM Maps.” *Nature Methods*. (2022). [Paper][GitHub] [Colab] [Integrated by Scipon for wider use in structural biology](#).
- 2 **Xiao Wang**, Eman Alnabati, Tunde W Aderinwale, Sai Raghavendra Maddhuri Venkata Subramaniya, Genki Terashi & Daisuke Kihara. “Emap2sec+: Structure Detection in Intermediate Resolution Cryo-EM Maps Using Deep Learning”. *Nature Communications*. (2021). [Paper] [GitHub][CodeOcean] [Integrated by SBGrid and Scipon for wider use in structural biology](#).
- 3 **Xiao Wang**, Sean T, Flannery, Daisuke Kihara. “Protein Docking Model Evaluation by Graph Neural Networks”. *Frontiers in Molecular Biosciences (FMOLB)*. (2021). [Paper] [GitHub]
- 4 **Xiao Wang**, Genki Terashi, Charles W. Christoffer, Mengmeng Zhu, and Daisuke Kihara, “Protein Docking Model Evaluation by 3D Deep Convolutional Neural Networks”. *Bioinformatics* 36: 2113-2118 (2020). [Paper] [GitHub][Platform] [Recognized as representative deep-Learning-based method to directly tackle key bottleneck of structure quality assessment by CAPRI community](#).
- 5 **Xiao Wang***, Mizu Kittaka*, Yilin He, Yiwei Zhang, Yasuyoshi Ueki, and Daisuke Kihara. “OC_Finder: A deep learning-based software for osteoclast segmentation, counting, and classification”. *Frontiers in Bioinformatics*. (2021). [Paper] [GitHub] [Colab]

Self-Supervised Learning

Pre-print

- 1 **Xiao Wang**, Yuhang Huang, Dan Zeng, Guo-Jun Qi. “CaCo: Both Positive and Negative Samples are Directly Learnable via Cooperative-adversarial Contrastive Learning”. *IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE T-PAMI)*. (2022). (Major Revision). [Paper] [GitHub]
- 2 **Xiao Wang**, Ying Wang, Ziwei Xuan, Guo-Jun Qi. “AdPE: Adversarial Positional Embeddings For MIM Pretraining of Transformers”. *The Eleventh International Conference on Learning Representations (ICLR 2023)*. (Submitted).
- 3 **Xiao Wang**, Jingen Liu, Tao Mei, Jiebo Luo. CoSeg: “Cognitively Inspired Unsupervised Generic Event Segmentation”. *IEEE Transactions on Neural Networks and Learning Systems (IEEE TNNLS)*. (Submitted). [Paper]

Published

- 1 **Xiao Wang***, Haoqi Fan*, Yuandong Tian, Daisuke Kihara, Xinlei Chen. “On the Importance of Asymmetry for Siamese Representation Learning”. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022)*. [Paper] [GitHub] [First asymmetrical self-supervised learning approach](#).

- 2 **Xiao Wang**, Guo-Jun Qi. “Contrastive Learning with Stronger Augmentations”. *IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE T-PAMI)*. (2022). [Paper] [GitHub]
- 3 **Xiao Wang***, Qianjiang Hu*, Wei Hu, Guo-Jun Qi. AdCo: “Adversarial Contrast for Efficient Learning of Unsupervised Representations from Self-Trained Negative Adversaries”. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2021)*. [Paper] [GitHub]. [Pioneering work in adversarial self-supervised learning.](#)
- 4 Guo-Jun Qi, Liheng Zhang, Feng Lin, **Xiao Wang**. “Learning Generalized Transformation Equivariant Representations via Autoencoding Transformations”. *IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE T-PAMI)*. (2020). [Paper] [GitHub]
- 5 **Xiao Wang**, Daisuke Kihara, Jiebo Luo, Guo-Jun Qi. “Enaet: Self-trained ensemble autoencoding transformations for semi-supervised learning”. *IEEE Transactions on Image Processing (IEEE TIP)*. (2020). [Paper] [GitHub]

Intelligent Transportation (before Ph.D.)

- 1 **Xiao Wang**, Rui Jiang, Li Li, Yi-Lun Lin, and Fei-Yue Wang. “Long memory is important: A test study on deep-learning based car-following model”. *Physica A: Statistical Mechanics and its Applications*, 514: 786-795, (2019) [Paper]
- 2 **Xiao Wang**, Rui Jiang, Li Li, Yilun Lin, Xinhua Zheng, and Fei-Yue Wang. “Capturing car-following behaviors by deep learning”. *IEEE Transactions on Intelligent Transportation Systems (IEEE T-ITS)*, PP(99):1–11, (2018). [Paper] [Nominated for George N. Saridis Best Transactions Paper Award.](#)

* denotes equal contribution.

PRESENTATIONS & TALKS

- 1 Invited Talk: “The GRU-based Car-Following Model.”, **Xiao Wang**, Rui Jiang, Li Li, Yilun Lin, Xinhua Zheng, and Fei-Yue Wang. Institute of Automation, Chinese Academy of Sciences, Aug 15, 2022.
- 2 Invited Talk: “How to achieve adversarially self-supervised learning?”, **Xiao Wang**, Qianjiang Hu, Yuhang Huang, Wei Hu, Zeng Dan, Guo-Jun Qi. TechBeat, Aug 4, 2022.
- 3 Invited Talk: “De novo DNA/RNA structure modeling from cryo-EM maps by deep learning.”, **Xiao Wang**, Genki Terashi, Daisuke Kihara. The Molecular Sciences Software Institute, Virginia Tech, June 29, 2022.
- 4 Presentation: “On the Importance of Asymmetry for Siamese Representation Learning.”, **Xiao Wang***, Haoqi Fan*, Yuandong Tian, Daisuke Kihara, Xinlei Chen. CVPR 2022 Representation Learning Session. June 22, 2022.
- 5 Invited Talk: “Adversarial self-supervised learning”, **Xiao Wang**, Qianjiang Hu, Wei Hu, Guo-Jun Qi. Huawei Central Research. May 16th, 2022.
- 6 Invited Talk: “Pre-task and contrastive self-supervised learning”, **Xiao Wang**, OPPO AI Research, March 11th, 2022.
- 7 Presentation: “Deep learning-based local quality estimation for protein structure models from cryo-EM maps”, Genki Terashi*, **Xiao Wang***, Sai Raghavendra Maddhuri Venkata Subramaniya, John J. G. Tesmer, Daisuke Kihara, 66th Biophysical Society Annual Meeting, San Francisco, Feb. 19-23, 2022
- 8 Invited Talk: “Pre-task and contrastive self-supervised learning”. **Xiao Wang**, Meta AI, Feb 22th, 2022.
- 9 Invited Talk: “Pre-task and contrastive self-supervised learning”. **Xiao Wang**, Microsoft Research, Dec 10th, 2021.
- 10 Invited Talk: “Introduction of self-supervised learning”. **Xiao Wang**, Google Brain, Nov 29th, 2021.
- 11 Invited Talk: “Pre-task and contrastive self-supervised learning”. **Xiao Wang**, VISA AI Research, Nov 12th, 2021.
- 12 Presentation: “The OC_Finder: A deep learning-based software for in vitro osteoclast counting”, **Xiao Wang**, Mizuho Kittaka, Yilin He, Yiwei Zhang, Daisuke Kihara, Yasuyoshi Ueki, 2021 Annual Meeting of The American Society for Bone and Mineral Research, October 1-4, 2021.
- 13 Presentation: “Emap2sec+: Detecting protein and DNA/RNA structures in cryo-EM maps of intermediate resolution using deep learning.”, **Xiao Wang**, Eman Alnabati, Tunde W. Aderinwale, Sai Raghavendra Maddhuri Venkata Subramaniya, Genki Terashi, Daisuke Kihara, 71st Annual Meeting of the American Crystallographic Association, July 30 - Aug 5, 2021.
- 14 Presentation: “Asymmetrical self-supervised learning.”, **Xiao Wang***, Haoqi Fan*, Yuandong Tian, Daisuke Kihara, Xinlei Chen. Facebook AI CV group, July 24, 2021.

- 15 Presentation: “Adversarial Contrast for Efficient Learning of Unsupervised Representations from Self-Trained Negative Adversaries”, **Xiao Wang**, Qianjiang Hu, Wei Hu, Guo-Jun Qi. CVPR 2021 Representation Learning Session. June 17, 2022.
- 16 Invited Talk: “Adversarial self-supervised learning”, **Xiao Wang**, Qianjiang Hu, Wei Hu, Guo-Jun Qi, Microsoft Research Asia, May 15, 2021.
- 17 Presentation: “Detecting protein and DNA/RNA structures in cryo-EM maps of intermediate resolution using deep learning”, **Xiao Wang**, Eman Alnabati, Tunde W. Aderinwale, Sai Raghavendra Maddhuri Venkata Subramaniya, Genki Terashi, Daisuke Kihara, The Hitchhiker’s Guide to the Biomolecular Glaxy 2021, May 12-13, 2021.
- 18 Presentation: “Detecting protein and DNA/RNA structures in cryo-EM maps of intermediate resolution using deep learning”, **Xiao Wang**, Eman Alnabati, Tunde W. Aderinwale, Sai Raghavendra Maddhuri Venkata Subramaniya, Genki Terashi, Daisuke Kihara, 2021 65th Biophysical Society Annual Meeting, February 22-26, 2021.
- 19 Presentation: “Emap2sec+: Detecting protein and DNA/RNA structures in cryo-EM maps of intermediate resolution using deep learning”, **Xiao Wang**, Eman Alnabati, Tunde Aderinwale, Sai Raghavendra Maddhuri Venkata Subramaniya, Genki Terashi, Daisuke Kihara, Frontiers in Cryo-Electron Microscopy, Keystone Symposia, Feb. 3-4, 2021.
- 20 Invited Talk: “Adversarial self-supervised learning.”, **Xiao Wang**, Qianjiang Hu, Wei Hu, Guo-Jun Qi. Facebook AI Research, Jan 20, 2021.
- 21 Presentation: “Protein Docking Model Evaluation by Graph Neural Networks”, **Xiao Wang**, Sean T. Flannery, Daisuke Kihara, Structural and Computational Biology and Biophysics (SCBB) Graduate Student Symposium of Purdue, Dec 16, 2020.
- 22 Presentation: “Emap2sec+: Detecting protein and DNA/RNA structures in cryo-EM maps of intermediate resolution using deep learning”, **Xiao Wang**, Eman Alnabati, Tunde Aderinwale, Sai Raghavendra Maddhuri Venkata Subramaniya, Genki Terashi, Daisuke Kihara, 5th Annual Southern California Cryo-EM Symposium, Oct 30, 2020.
- 23 Guest Lecture: “Review of car-following model”, **Xiao Wang**, Institute of Automation, Tsinghua University, April 21, 2020.
- 24 Presentation: “Protein docking model evaluation by 3D convolutional neural networks”, **Xiao Wang**, Genki Terashi, Charles W Christoffer, Mengmeng Zhu, Daisuke Kihara, invited talk by Aggregate Intellect - AI.SCIENCE, Mar 24, 2020. [Video].

PROFESSIONAL SERVICES

Reviewer of IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE-T-PAMI)

Reviewer of IEEE Transactions on Image Processing (IEEE-TIP)

Reviewer of Pattern Recognition

Reviewer of Bioinformatics

Reviewer of IEEE Transactions on Intelligent Transportation Systems (IEEE-T-ITS)

Reviewer of IEEE Transactions on Multimedia (IEEE-TMM)

Reviewer of ACM Transactions on Knowledge Discovery from Data (IEEE-TKDD)

Reviewer of IEEE Intelligent Systems

Reviewer of Frontiers in Bioinformatics

Reviewer of Conference on Neural Information Processing Systems (NurIPS)

Reviewer of Conference on Computer Vision and Pattern Recognition (CVPR)

Reviewer of International Conference on Computer Vision (ICCV)

Reviewer of European Conference on Computer Vision (ECCV)

TEACHING & MENTORING EXPERIENCE

Teaching

Computing for Life Sciences, Purdue
Java Programming, XJTU

2020 Fall
 2017 Spring

Guest Lecture

Traffic Flow Theory, Tsinghua
Molecular Sciences, Virginia Tech

April, 2020
July, 2020

Mentoring

Grace Su, undergraduate from Columbia University	2021-2022
Yunhan Huang, undergraduate student from Purdue University	2021-2022
Yuhang Huang, undergraduate from Shanghai University	2021-2022
Rohan Raghavan Narasimha, M.S. graduate student from Purdue University	2020-2021
Sean T Flannery, M.S. graduate student from Purdue University	2020-2021
Qianjiang Hu, undergraduate from Peking University	2020-2021
Yilin He, undergraduate summer intern from Shandong University	2019
Yiwei Zhang, undergraduate summer intern from Rensselaer Polytechnic Institute	2019

OPEN SOURCE PROJECT

DAQ-Refine

DAQ-Refine is a protein structure refinement tool by DAQ-score and ColabFold.
Available at: <https://github.com/kiharalab/DAQ-Refine> and <https://bit.ly/DAQ-Refine>.

Asym-Siam

Asym-Siam experimentally verified the importance of asymmetry for Siamese Representation Learning with obvious improvement.
Available at: <https://github.com/facebookresearch/asym-siam>.

DAQ

DAQ is a software accesses the quality of protein models built from cryo-Electron Microscopy (EM) maps, which uses local map density features captured by deep learning.
Available at: <https://github.com/kiharalab/DAQ> and <https://bit.ly/daq-score>.

CaCo

CaCo is a state-of-the-art cooperative-adversarial contrastive learning method where both positive and negative samples are directly learnable.
Available at: <https://github.com/maple-research-lab/caco>.

OC_Finder

OC_Finder is a computational tool using deep learning for fully automated osteoclast segmentation, classification, and counting.
Available at: https://github.com/kiharalab/OC_Finder.

CLSA

CLSA is a general contrastive learning framework by introducing the information from stronger augmentation.
Available at: <https://github.com/maple-research-research-lab/CLSA>.

GNN_DOVE

GNN_DOVE is a software can evaluate the quality of protein-docking models using graph neural networks by reformulating protein structures as graphs.
Available at: https://github.com/kiharalab/GNN_DOVE.

GNN_Pocket

GNN_Pocket is a software that can detect protein pockets with graph neural network.
Available at: https://github.com/kiharalab/GNN_Pocket.

AdCo

AdCo is an algorithm for effective self-supervised learning through adversarial training of negative examples.
Available at: <https://github.com/maple-research-lab/AdCo>.

Emap2sec+

Emap2sec+ is a software detects local structure information of proteins and DNA/RNA in cryo-EM maps using deep learning.

Available at: <https://github.com/kiharalab/Emap2secPlus> and <https://doi.org/10.24433/CO.7165707.v1>.

Attention_AD

Attention_AD is a software that can distinguish active and inactive peptides for gene expression using Long Short Term Memory (LSTM).

Available at: https://github.com/kiharalab/Attention_AD.

EnAET

EnAET is a software that benefits semi-supervised learning via self-trained ensemble auto-encoding transformations.

Available at: <https://github.com/maple-research-lab/EnAET>.

DOVE

DOVE is a software can evaluate the quality of protein-docking models using 3D neural networks.

Available at: <https://github.com/kiharalab/DOVE> and <https://kiharalab.org/dove>.

SKILLS

- Expertise: Pytorch, Python, Tensorflow, Matlab, Java
- Capable: SQL, C, C++, Fortran, Java Web, Latex, Android