

Dunyu Road 600  
310030 Hangzhou  
China

+86 136 0259 1515

jackywang28@outlook.com

wangzedong@westlake.edu.cn

jacky1128.github.io



# Zedong Wang

Google Scholar

Twitter

OpenReview

LinkedIn

GitHub

Citations: 83, H-index: 4

Contribute: 3, Stars: 1.1k

## Short Biography

I am an HK-born AI researcher. I received my B.Eng. degree in Electronic and Information Engineering from [Huazhong University of Science and Technology \(HUST\)](#), in June, 2023. My research interests center around Multi-modal Learning and Multi-task Scene Understanding. I am also interested in Computer Vision from 3 aspects: (i) Mixup Augmentation and label-efficient learning (Data-level); (ii) Efficient network architecture design (Network-level); (iii) Generation with Vector Quantization, Diffusion Models and more (Framework-level). Currently, I am a visiting student at [Westlake University](#) under [Chair Prof. Stan Z. Li](#) (IEEE Fellow, IAPR Fellow). I am also a research intern at [HKUST](#) advised by [Prof. Dan Xu](#). At [HUST](#), I was fortunate to work on few-shot semantic segmentation under [Prof. Xinggang Wang](#). Prior to that, I conducted an internship at [SIAT-MMLab](#), Shenzhen Institute of Advanced Technology (SIAT), Chinese Academy of Sciences (CAS).

## Education and Degrees

- 2019 – 2023 **B.Eng. in Electronic and Information Engineering**, *Huazhong University of Science and Technology*.
- Multiple High Quality Computer Vision Research Experience | **Undergrad Supervisor: Prof. Xinggang Wang**.
  - Graduation Thesis**: Efficient Visual Backbone Architecture Design | **Grade: 92/100 (First-Tier, Full Novelty Scores)** | **Advisor: Prof. Xinggang Wang**
  - High GPA in **AI-related** core courses (**90.0/100 in Average**): Introduction to Green Communications (**95/100**), Engineering Training (**94/100**), Multimedia Retrieval (**93/100**), Graduation Thesis (**92/100**), Software Project (**92/100**), Principles and Applications of Sensors (**90/100**), Python programming (**87/100**), Capstone Project in Machine Intelligence (**87/100**), Deep Learning and Computer Vision (**87/100**), Machine Learning (**85/100**) etc.

## Research Experience (Links are provided)

- Feb. 2024 **Please View My [Homepage](#) for The Latest Updates!**
- Mar. 2024 **Research Intern**, *The Hong Kong University of Science and Technology (HKUST)*.  
**Advisor: Prof. Dan Xu** | **Research Topic**: Multi-modal Learning and Multi-task Scene Understanding.
- Jan. 2024 **Invited as a Reviewer for ICML 2024, ECCV 2024, ICLR 2024 (TinyPapers), and ICPR 2024.**
- Dec. 2022 **Ph.D. Offer**, *AI Division, School of Engineering, Westlake University*.
- Sep. 2022 – Present **Visiting Student** (representation learning, ai4science), *CAIRI AI Lab (Chair Prof. Stan Z. Li Lab), Westlake University*.  
(i) Data Mixing and Label-Efficient Learning. | [SemiReward \(ICLR 2024\)](#); [SAMix](#); [OpenMixup \(549 GitHub stars\)](#).  
(ii) Efficient Network Architecture Design on Vision and Beyond. | [MogaNet \(ICLR 2024, 114 GitHub stars\)](#).  
(iii) Representation Learning Framework. | [OpenSTL \(NeurIPS 2023, 508 GitHub stars\)](#); [Masked Modeling Survey](#).  
(iv) Vector Quantized Vocabulary Learning for Multi-Species DNA Foundation Model. | [VQDNA \(to ICML2024\)](#).
- Jul. 2022 – Sep. 2022 **Summer Research Studentship**, *School of Engineering, Westlake University*.  
**Advisor: Chair Prof. Stan Z. Li (only 2 selected out of 100+ applicants)** | **Research Topic**: Representation Learning.
- Sep. 2021 – Jun. 2022 **Research Intern**, *HUST Vision Lab, School of EIC, Huazhong University of Science and Technology*.  
**Advisor: Prof. Xinggang Wang** | **Research Topic**: Efficient Visual Recognition & Few-shot Semantic Segmentation.
- Jul. 2021 – Sep. 2021 **Visiting Student**, *MMLab, Shenzhen Institute of Advanced Tech. (SIAT), Chinese Academy of Sciences*.  
**Advisor: Dr. Bin Fu** | **Research Topic**: Semantic Segmentation.
- Sep. 2020 – Apr. 2021 **Research Intern**, *Key Lab of Digital Earth Science, Chinese Academy of Sciences*.  
**Advisor: Dr. Xiaoping Du** | **Research Topic**: High Resolution Remote Sensing Building Semantic Segmentation.

## Languages and Skills

Chinese (native), English (fluent). **IELTS 7.5 (2023)** overall grades, **CET-4 646** overall grades.  
Python DL Libraries, PyTorch, Git, Anaconda, Linux (basic),  $\text{\LaTeX}$ , All-round Research Skills.

## Publications (\*: Equivalent Contribution. †: Corresponding Author. Links are provided)

### ICLR 2024 **MogaNet: Multi-order Gated Aggregation Network.**

Siyuan Li\*, **Zedong Wang\***, Zicheng Liu, Cheng Tan, Haitao Lin, Di Wu, Zhiyuan Chen, Jiangbin Zheng, Stan Z. Li<sup>†</sup>

- **The first** network backbone design through the lens of multi-order game-theoretic interaction, which portrays inter-variable interaction effects w.r.t. varying scale of context via game theory.
- The representative interactions are emphasized by the proposed spatial gated aggregation and channel reallocation architecture, leading to better representation learning capacity.
- Impressive scalability and superior performance with a more efficient use of model parameters than state-of-the-art ViTs and ConvNets scaling from 5M to 180M on various vision benchmarks.
- **Spontaneously forwarded by world-wide media (Twitter, Zhihu, Wechat) with high appraisal.**
- **Two Weak Accept in CVPR'23, One Strong Accept in ICCV'23, all positive ratings 6668 in ICLR'24**

### ICLR 2024 **SemiReward: A General Reward Model for Semi-supervised Learning.**

Siyuan Li\*, Weiyang Jin\*, **Zedong Wang**, Fang Wu, Zicheng Liu, Cheng Tan, Stan Z. Li<sup>†</sup>

- **The first** online-optimizable reward model that predicts reward scores to filter out high-quality pseudo labels for semi-supervised representation learning (both classification and regression).
- State-of-the-art across **12 classification and regression** semi-supervised learning benchmarks.

### To ICML 2024 **OpenMixup: Open Mixup Toolbox for Visual Representation Learning.**

Siyuan Li\*, **Zedong Wang\***, Zicheng Liu\*, Di Wu, Stan Z. Li<sup>†</sup>

- **The first** comprehensive mixup visual classification benchmark. where 16 representative mixup algorithms are impartially evaluated from scratch across 12 visual classification datasets, ranging from classical iconic scenarios to fine-grained, long-tail, and scenic cases.
- **The first** standardized mixup-based vision model design and training codebase framework OpenMixup for customized visual classification.
- **Interesting observations** are derived through extensive empirical analysis on various scenarios.
- **Spontaneously retweeted by Prof. Sebastian Raschka (Twitter) with high appraisal.**
- **Spontaneously reported by Lightning AI official account (Twitter) as 'weakly highlights in AI'. (the same session as pytorch's departure of facebook)**

### To ECCV 2024 **Boosting Discriminative Visual Representation Learning with Scenario-Agnostic Mixup.**

Siyuan Li\*, Zicheng Liu\*, **Zedong Wang\***, Di Wu, Zihan Liu, Stan Z. Li<sup>†</sup>

- A unified online-optimizable mixup framework that **first addresses the two remaining critical issues** at once: **(i)** Drastic performance variation over different scenarios caused by trivial solutions; **(ii)** Self-supervised learning (SSL) dilemma for online-optimizable mixup policies.
- To reduce the computational cost from online training, a pre-trained version is presented.
- Exceptional performance and generalizability across **12 SL and SSL** image benchmarks.

### NeurIPS 2023 **OpenSTL: A Comprehensive Benchmark of Spatio-Temporal Predictive Learning.**

Cheng Tan, Siyuan Li, Zhangyang Gao, Wenfei Guan, **Zedong Wang**, Zicheng Liu, Lirong Wu, Stan Z. Li<sup>†</sup>

- **The first** comprehensive benchmarking study for spatio-temporal predictive learning that categorized prevalent approaches into recurrent-based and recurrent-free models.
- **The first** modular and extensible framework implementing various state-of-the-art methods. Impartial evaluations and analysis are conducted across various domains, including synthetic moving object trajectory, human motion, driving scenes, traffic flow, weather forecasting.
- Surprisingly, we find that recurrent-free models achieve a good balance between efficiency and performance than recurrent models.
- **All positive ratings in NeurIPS 2023 (6,6,7,7) | Accepted as Poster by NeurIPS 2023**

### To ICML 202 **Unleashing the Power of Vector Quantization for Multi-Species Genomic Sequence Modeling.**

Siyuan Li\*, **Zedong Wang\***, Zicheng Liu, Stan Z. Li<sup>†</sup>