

# HAIWEN DIAO

## Postdoctoral Research Fellow

Multimedia Laboratory @ Nanyang Technological University ([MMLab@NTU](#))

College of Computing and Data Science, Nanyang Technological University

Phone: (+86) 185-2544-5313 ♦ Email: [r1228240468@gmail.com](mailto:r1228240468@gmail.com) ♦ [Homepage](#) ♦ [Google Scholar](#)

## Research Interests

---

My primary research interests are *Vision-and-Language Domain*, *Transfer Learning*, and *Large Multi-Modality Models*. Specifically, my research in the vision-and-language domain strives to build an efficient and powerful system that recognizes visual-semantic perception and contextualizes fine-grained correspondence across multiple modalities:

1. **Vision-Language Retrieval:** [SGRAF \(AAAI'21\)](#), [RCAR \(TIP'23\)](#), [DBL \(TIP'24\)](#), [GSSF \(TIP'24\)](#)
2. **Efficient Transfer Learning:** [UniPT \(CVPR'24\)](#), [SHERL \(ECCV'24\)](#), [ReSoRA \(2024\)](#), [KARST \(ICASSP'24\)](#)
3. **Multi-Modality Perception:** [EVE \(NIPS'24\)](#), [EVEv2 \(2024\)](#), [DenseFusion \(NIPS'24\)](#), [Infinity-MM \(2024\)](#)
4. **Multi-Modality Generation:** [NOVA \(ICLR'25\)](#), [MoTrans \(ACMMM'24\)](#)
5. **Multi-Modality Unification:** [ETT \(2025\)](#)

## Experience

---

**Nanyang Technological University (NTU), Singapore**

**Apr. 2025 — Present**

*Postdoctoral Research Fellow*

- Thesis: [Large Fundamental Models, Inference-Time Scaling and Reasoning](#)
- Supervisor: [Prof. Ziwei Liu](#) (MMLab@NTU)

**Beijing Academy of Artificial Intelligence (BAAI), Beijing**

**Sep. 2023 — Mar. 2025**

*Research Internship*

- Thesis: [Large Multi-Modality Models for Understanding and Generation](#)
- Director: [Dr. Xinlong Wang](#) (Large Multimodal Model Center)

**Hong Kong University of Science and Technology (HKUST), Hongkong**

**Jan. 2023 — Aug. 2023**

*Remote Cooperation*

- Thesis: [Parameter-Efficient and Memory-Efficient Transfer Learning](#)
- Director: [Prof. Long Chen](#) (Computer Science and Engineering)

**Tencent AI Lab, Shenzhen**

**Jun. 2020 — Mar. 2021**

*2020 Tencent AI Lab Rhino-Bird Focused Research Program*

- Thesis: [Image-Text Retrieval, Cross-Modal Boosting and Metric Learning](#)
- Director: [Dr. Ying Zhang](#), [Dr. Lin Ma](#) (Computer Vision Center)

## Education

---

**Dalian University of Technology (DLUT), Dalian**

**Sep. 2021 — Mar. 2025**

*Ph.D. in Signal and Information Processing*

- Thesis: [Parameter-Efficient Transfer Learning, Large Multi-Modality Models](#)
- Supervisor: [Prof. Huchuan Lu](#)

**Dalian University of Technology (DLUT), Dalian**

**Sep. 2018 — Jun. 2021**

*M.E. in Information and Communication Engineering*

- Thesis: [Person Re-Identification, Vision-and-Language Domain](#)
- Supervisor: [Prof. Huchuan Lu](#)

**Dalian University of Technology (DLUT), Dalian**

**Sep. 2014 — Jun. 2018**

*B.E. in Electronic Information Engineering*

- GPA: Top3% (School of Information and Communication Engineering)

## Awards & Hornors

---

## Awards

- Silver Award of the 9th China International College Students' 'Internet+' Innovation and Entrepreneurship Competition (Industrial Track) (Project Leader) (Top 120/421M) *Ministry of Education, 2023*
- Excellence Award of 2020 Tencent AI Lab Rhino-Bird Focused Research Program (9/27) *Tencent AI Lab, 2021*

## Honors

- Excellent Graduate of Dalian City (Bachelor's degree) *Dalian City, 2018*
- Lingshui Scholarship (First Prize) (Top 5%) *Dalian City, 2016*
- China National Scholarship (Top 1%) *Ministry of Education, 2015*
- Awarded Learning Excellence Award (First Prize) (Top 5%) *DLUT, 2015/16*

## Publications

### Conference Publications .....

- **Autoregressive Video Generation without Vector Quantization**  
Haoge Deng\*, Ting Pan\*, **Haiwen Diao\***, Zhengxiong Luo\*, Yufeng Cui, Huchuan Lu, Shiguang Shan, Yonggang Qi, Xinlong Wang ☒  
*International Conference on Learning Representations (ICLR), 2025. [Code]*
- **KARST: Multi-Kernel Kronecker Adaptation with Re-Scaling Transmission for Visual Classification**  
Yue Zhu\*, **Haiwen Diao\***, Shang Gao\*, Long Chen, Huchuan Lu ☒  
*IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2025. [Code]*
- **Unveiling Encoder-Free Vision-Language Models**  
**Haiwen Diao\***, Yufeng Cui\*, Xiaotong Li, Yuezhe Wang, Huchuan Lu ☒, Xinlong Wang ☒  
*Advances in Neural Information Processing Systems (NeurIPS), 2024. (spotlight) [Code]*
- **DenseFusion-1M: Merging Vision Experts for Comprehensive Multimodal Perception**  
Xiaotong Li\*, Fan Zhang\*, **Haiwen Diao\***, Yuezhe Wang, Xinlong Wang ☒, Lingyu Duan ☒  
*Advances in Neural Information Processing Systems (NeurIPS), 2024. [Code]*
- **LLMs Can Evolve Continually on Modality for X-Modal Reasoning**  
Jiazuo Yu, Haomiao Xiong, Lu Zhang ☒, **Haiwen Diao**, Yunzhi Zhuge, Lanqing Hong, Dong Wang, Huchuan Lu, You He, Long Chen  
*Advances in Neural Information Processing Systems (NeurIPS), 2024. [Code]*
- **MoTrans: Customized Motion Transfer with Text-driven Video Diffusion Models**  
Xiaomin Li, Xu Jia ☒, Qinghe Wang, **Haiwen Diao**, Mengmeng Ge, Pengxiang Li, You He, Huchuan Lu  
*ACM International Conference on Multimedia (ACMMM), 2024. [Code]*
- **SHERL: Synthesizing High Accuracy and Efficient Memory for Resource-Limited Transfer Learning**  
**Haiwen Diao**, Bo Wan, Xu Jia, Yunzhi Zhuge, Ying Zhang, Huchuan Lu ☒, Long Chen  
*European Conference on Computer Vision (ECCV), 2024. [Code]*
- **UniPT: Universal Parallel Tuning for Transfer Learning with Efficient Parameter and Memory**  
**Haiwen Diao**, Bo Wan, Ying Zhang, Xu Jia, Huchuan Lu ☒, Long Chen  
*IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024. [Code]*
- **Similarity Reasoning and Filtration for Image-Text Matching**  
**Haiwen Diao**, Ying Zhang, Lin Ma, Huchuan Lu ☒  
*AAAI Conference on Artificial Intelligence (AAAI), 2021. (Google citations: 370+) [Code]*

### Journal Publications .....

- **Exploring Dynamic Transformer for Efficient Object Tracking**  
Jiawen Zhu, Xin Chen, **Haiwen Diao**, Shuai Li, Jun-Yan He, Chenyang Li, Bin Luo, Dong Wang, Huchuan Lu ☒  
*IEEE Transactions on Neural Networks and Learning Systems (TNNLS), 2025.*
- **GSSF: Generalized Structural Sparse Function for Deep Cross-modal Metric Learning**  
**Haiwen Diao**, Ying Zhang, Shang Gao, Jiawen Zhu, Long Chen, Huchuan Lu ☒  
*IEEE Transactions on Image Processing (TIP), 2024. [Code]*
- **Deep Boosting Learning: A Brand-new Cooperative Approach for Image-Text Matching**  
**Haiwen Diao**, Ying Zhang, Shang Gao, Xiang Ruan, Huchuan Lu ☒  
*IEEE Transactions on Image Processing (TIP), 2024. [Code]*

- **Plug-and-Play Regulators for Image-Text Matching**  
Haiwen Diao, Ying Zhang, Wei Liu, Xiang Ruan, Huchuan Lu ✉  
*IEEE Transactions on Image Processing (TIP)*, 2023. [\[Code\]](#)

Preprints (Under Review) .....

- **End-to-End Vision Tokenizer Tuning**  
Wenxuan Wang\*, Fan Zhang\*, Yufeng Cui\*, Haiwen Diao\*, Zhuoyan Luo, Huchuan Lu, Jing Liu, Xinlong Wang  
*Under Review.*
- **EVEv2: Improved Baselines for Encoder-Free Vision-Language Models**  
Haiwen Diao\*, Xiaotong Li\*, Yufeng Cui\*, Yueze Wang\*, Haoge Deng, Ting Pan, Wenxuan Wang, Huchuan Lu  
✉, Xinlong Wang ✉  
*Under Review.* [\[Code\]](#)
- **Infinity-MM: Scaling Multimodal Performance with Large-Scale and High-Quality Instruction Data**  
Shuhao Gu, Jialing Zhang, Siyuan Zhou, Kevin Yu, Zhaohu Xing, Idwang, Zhou Cao, Jintao Jia, Zhuoyi Zhang,  
Yixuan Wang, Zhenchong Hu, Bo-Wen Zhang, Jijie Li, D.Liang, YingliZhao, Songjing Wang, Yulong Ao, Xiaotong  
Li, Haiwen Diao, Yufeng Cui, Xinlong Wang, Yaoqi Liu, Fangxiang Feng, Guang Liu ✉  
*Under Review.* [\[Code\]](#)
- **Regularizing Subspace Redundancy of Low-Rank Adaptation**  
Yue Zhu\*, Haiwen Diao\*, Shang Gao\*, Jiazuo Yu, Jiawen Zhu, Yunzhi Zhuge, Shuai Hao, Xu Jia, Lu Zhang,  
Ying Zhang, Huchuan Lu ✉  
*Under Review.*

Patents Declaration .....

- **An Image Retrieval Method and Related Device**  
Huchuan Lu, Haiwen Diao, Ying Zhang, Lin Ma  
CN111914113A, China, 2024.

## Professional Services

---

Journal Reviewer:

TPAMI, IJCV, TIP, TNNLS

Conference Reviewer:

NeurIPS, CVPR, ICCV, ECCV, AAAI, ACMMM

## References in Time Order

---

- **Prof. Huchuan Lu (Doctoral Supervisor)**  
School of Information and Communication Engineering  
Dalian University of Technology (DLUT)  
Dalian, Liaoning, 116024, China  
Email: lhchuan@dlut.edu.cn
- **Prof. Xu Jia (Laboratory Advisor)**  
School of Artificial Intelligence  
Dalian University of Technology (DLUT)  
Dalian, Liaoning, 116024, China  
Email: xjia@dlut.edu.cn
- **Dr. Ying Zhang (Industrial Director)**  
Tencent WeiXin Group (WXG)  
Shenzhen, Guangdong 518000, China  
Email: zydl0907@gmail.com
- **Dr. Lin Ma (Industrial Director)**  
Meituan AI Lab  
Chaoyang, Beijing, 100102, China

Email: forest.linma@gmail.com

- **Prof. Long Chen (Remote Director)**

Department of Computer Science and Engineering  
Hong Kong University of Science and Technology (HKUST)  
Clear Water Bay, Kowloon, Hong Kong, 999077, China  
Email: longchen@cse.ust.hk

- **Dr. Xinlong Wang (Industrial Director)**

Beijing Academy of Artificial Intelligence (BAAI)  
Haidian, Beijing, 100871, China  
Email: wangxinlong@baai.ac.cn

- **Prof. Ziwei Liu (Postdoctoral Supervisor)**

College of Computing and Data Science  
Nanyang Technological University (NTU)  
50 Nanyang Avenue, 639798, Singapore  
Email: ziwei.liu@ntu.edu.sg