

Delong Chen (陈德龙)

chendelong.world | delong.chen@connect.ust.hk | [Google Scholar](#) | [Github](#)

Education

Hong Kong University of Science and Technology (HKUST) Ph.D. Student in ECE Department, advised by Prof. Pascale Fung	Sept. 2023 – Present Hong Kong, China
Hohai University B.Eng. in Computer Science, advised by Prof. Fan Liu	Sept. 2017 – Jun. 2021 Nanjing, China
University of British Columbia Summer Program: Linguistics and computation for natural language processing	Jul. 2018 – Aug. 2018 Vancouver, Canada

Work Experience

Research Intern Visual instruction tuning for multimodal LLMs	Xiaobing.ai (Microsoft XiaoIce) Beijing, China	Mar. 2023 – Sept. 2023
Research Intern Vision-language learning for embodied AI	Microsoft Research Asia (MSRA) Beijing, China	Oct. 2022 – Jan. 2023
Research Intern Large-scale vision language pretraining	MEGVII Research (Face++) Beijing, China	Oct. 2021 – Oct. 2022
Research Assistant Vision-language learning for remote sensing, few-shot learning	Hohai University Nanjing, China	Sept. 2021 – Sept. 2023

Awards and Prizes

- **Best Paper** at AAAI 2023 Inaugural Summer Symposium Series - AI x Metaverse
- **Best Demo** Award in IEEE International Conference on Multimedia and Expo (ICME) 2021
- **Best Dataset** Paper Award in Long-Tailed Distribution Learning Workshop, IJCAI 2021
- **Best Presentation** Award in International Conference on Big Data and Artificial Intelligence (BDAI) 2021
- **First-Class Outstanding Thesis** in Jiangsu Province
- **Outstanding Thesis** of Hohai University
- **Outstanding Graduate** of Hohai University
- **Person of the Year** of the in Jiangsu Province (nomination award)
- **Person of the Year** of Hohai University in 2019
- **Third Prize** of 2018 the 8th China Software Cup, East China Division Finals (team Leader)

Publications

* joint first author ✉ corresponding author

- [1] **Delong Chen**, Jianfeng Liu, Wenliang Dai, Baoyuan Wang
Visual Instruction Tuning with Polite Flamingo
Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI), 2024. [[github](#)] [[arxiv](#)]
- [2] **Delong Chen**, Zhao Wu, Fan Liu, Zaiquan Yang, Huaxi Huang, Ying Tan, Erjin Zhou.
ProtoCLIP: Prototypical Contrastive Language Image Pretraining
IEEE Transactions on Neural Networks and Learning Systems (T-NNLS), 2023. [[github](#)] [[arxiv](#)] [[doi](#)]
- [3] **Delong Chen**, Fan Liu, Zewen Li, Feng Xu
VirtualConductor: Music-driven Conducting Video Generation System
IEEE International Conference on Multimedia & Expo (ICME) 2021. [[video](#)] (Best Demo)

-
- [4] **Delong Chen**, Fan Liu, Xiaoyu Du, Ruizhuo Gao, Feng Xu.
MEP-3M: A Large-scale Multi-modal E-Commerce Products Dataset
IJCAI-21 Workshop on Long-Tailed Distribution Learning, 2021. [[paper](#)] (Best Dataset Paper)
- [5] **Delong Chen**, Fan Liu, Zheqi Zhang, Xiaomin Lu, Zewen Li
Significant Wave Height Prediction based on Wavelet Graph Neural Network
The 4th International Conference on Big Data and Artificial Intelligence, 2021. [[arxiv](#)] [[doi](#)] (Best Presentation)
- [6] **Delong Chen**, Ruizhi Zhou, Yanling Pan, Fan Liu.
A Simple Baseline for Adversarial Domain Adaptation-based Unsupervised Flood Forecasting
ArXiv Technical Report, 2022. [[github](#)] [[arxiv](#)]
- [7] Fan Liu, **Delong Chen** ✉, Ruizhi Zhou, Sai Yang, Feng Xu
Self-Supervised Music Motion Synchronization Learning for Music-Driven Conducting Motion Generation
Journal of Computer Science and Technology (JCST), 2022 [[github](#)] [[video](#)] [[doi](#)]
- [8] Fan Liu*, **Delong Chen***, Zhangqingyun Guan, Xiaocong Zhou, Jiale Zhu, Jun Zhou
RemoteCLIP: A Vision Language Foundation Model for Remote Sensing
ArXiv Technical Report, 2023. [[github](#)] [[arxiv](#)]
- [9] Fan Liu*, **Delong Chen***, Xiaoyu Du, Ruizhuo Gao, Feng Xu
MEP-3M: A Large-scale Multi-modal E-commerce Product Dataset
Pattern Recognition, 2023. [[github](#)] [[doi](#)]
- [10] Fan Liu*, **Delong Chen***, Fei Wang, Zewen Li, Feng Xu
Deep Learning based Single Sample Face Recognition: A Survey
Artificial Intelligence Review, 2022. [[arxiv](#)] [[doi](#)]
- [11] Xinyu Zhou*, **Delong Chen***, Yudong Chen
Towards Joint Modeling of Dialogue Response and Speech Synthesis based on Large Language Model
The 6th International Conference on Natural Language and Speech Processing (ICNLSP), 2023. [[github](#)] [[arxiv](#)]
- [12] Fan Liu, Tianshu Zhang, Wenwen Dai, Wenwen Cai, Xiaocong Zhou, **Delong Chen** ✉
Few-shot Adaptation of Multi-modal Foundation Models: A Survey
Under Review, 2023 [[preprint](#)]
- [13] Fan Liu, **Delong Chen**, Jun Zhou, Feng Xu
A Review of Driver Fatigue Detection and Its Advances on the Use of RGB-D Camera and Deep Learning
Engineering Applications of Artificial Intelligence, 2022. [[doi](#)]
- [14] Fan Liu, Sai Yang, **Delong Chen**, Huaxi Huang, Jun Zhou
Few-shot Classification Guided by Generalization Error Bound
Pattern Recognition, 2023. [[doi](#)]
- [15] Sai Yang, Fan Liu, **Delong Chen**, Jun Zhou
Few-shot Classification via Ensemble Learning with Multi-Order Statistics
Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI), 2023. [[arxiv](#)] (Oral)
- [16] Zhuoran Zhao, Jinbin Bai, **Delong Chen**, Debang Wang, Yubo Pan
Taming Diffusion Models for Music-driven Conducting Motion Generation
AAAI 2023 Inaugural Summer Symposium Series – AI x metaverse. [[github](#)] [[arxiv](#)] [[doi](#)] (Best Paper)

Academic Services

-
- **Reviewer:** ACM Multimedia, Artificial Intelligence Review, IEEE Transactions on Intelligent Transportation Systems, Journal of Digital Imaging