

Yao DU 杜尧

Email: ydubf@connect.ust.hk [Google Scholar](#) [GitHub](#) [Personal Page](#)

Phone: +86 16638700460 (Wechat) +852 54875006



Research Summary: I am a Ph.D. candidate at HKUST working on multimodal foundation models, focusing on [physically grounded perception](#), [numerical understanding](#), and [quantitative reasoning](#) in vision-language and MLLM systems. My research develops principled learning objectives and evaluation frameworks to improve how models perceive, represent, and reason about physical quantities and numerical distributions in real-world environments.

Education

PhD, The Hong Kong University of Science and Technology

Major: Electronic and Computer Engineering

2022.09 – Present

Advisor: [Prof. Xiaomeng Li](#)

MPhil, The Hong Kong Polytechnic University

Major: Structural Health Monitoring

2019.09 – 2022.06

Advisor: [Prof. Yong Xia](#)

BEng, Northwestern Polytechnical University

GPA: 91.88/100 Rank: 1/53

2015.09 – 2019.06

Research Directions

Quantitative Perception, Multimodal Reasoning, and AI for Physical Systems

- **[Multimodal Quantitative Reasoning](#):** Improving numerical understanding and physical quantity perception in vision-language and multimodal foundation models.
- **[Medical and Scientific AI](#):** Developing data-efficient and reliable models for medical imaging and scientific applications, guided by domain knowledge and inductive biases.
- **[AI for Physical Systems](#):** Integrating AI models with real-world sensing systems for sensing data analysis, anomaly detection, and deployment in infrastructure monitoring.

Publications

[Quantitative Reasoning and Evaluation for Multimodal Foundation Models](#)

Numerically grounded perception, quantitative reasoning, and distribution awareness in VLM/MLLMs.

1. Injecting Distributional Awareness into MLLMs via Reinforcement Learning for Deep Imbalanced Regression.

Yao Du, Shanshan Song, Xiaomeng Li. **ICML 2026 (First author)**

Identifies the limitations of supervised fine-tuning for numerical regression, and introduces GRPO with batch-aware supervision for deep imbalanced regression together with a newly constructed DIR benchmark.

2. Teach CLIP to Develop a Number Sense for Ordinal Regression.

Yao Du, Qiang Zhai, Weihang Dai, Xiaomeng Li. **ECCV 2024 (First author)**

Proposes a unified framework that maps numbers to linguistic concepts, combining coarse-to-fine regression with distance-aware cross-modal contrastive learning to improve ordinal reasoning in VLMs.

3. Semi-Supervised Contrastive Learning for Deep Regression with Ordinal Rankings from Spectral Seriation.

Weihang Dai, Yao Du, et al. **NeurIPS 2023 (Second author)**

Studies semi-supervised deep regression by leveraging cross-sample relationships and noisy unlabeled outputs to extract informative supervision signals.

Medical Image Analysis: Echocardiography and Computational Pathology
Parameter-efficient fine-tuning and multimodal representation learning for medical imaging.

1. **CardiacCLIP: Video-based CLIP Adaptation for LVEF Prediction in a Few-shot Manner.**

Yao Du, Jiarong Guo, Xiaomeng Li. **MICCAI 2025 (First author)**

Extends CLIP-based ordinal regression to echocardiography, with a focus on multi-scale representation learning and frame-level relationship modeling.

2. **Beyond H&E: Unlocking Pathological Insights with Polarization Imaging.**

Yao Du, et al. **BIBM 2025 (First author)**

Constructs a paired HE and polarization imaging dataset through a customized acquisition pipeline, and demonstrates the effectiveness of multimodal fusion for enhancing pathology representation learning.

3. **MUST: Multi-Style Virtual Staining with Incomplete Pairs.**

Jiaxin Zhuang, Yao Du, et al. **TMI 2026 (Second author)**

Explores multi-style virtual staining using diffusion models under incomplete pairing, modeling cross-modality relationships in pathology.

4. **CardioInfuser: Multi-scale Semantic Infusion for Few-shot Echocardiogram Video Assessment.**

Yao Du, Marawan Elbatel, Jiewen Yang, Xiaomeng Li. **Under Review 2026 (First author)**

Investigates representation alignment strategies for echocardiogram model enhancement through external representation distillation.

Artificial Intelligence of Things (AIoT) / Structural Health Monitoring (SHM)
Infrastructure sensor anomaly detection, edge computing, and AIoT systems.

1. **Convolutional Neural Network-based Data Anomaly Detection Considering Class Imbalance with Limited Data.**

Yao Du, Lingfang Li, et al. **Smart Structures and Systems, 2022 (First author)**

Applies AI models to anomaly detection in SHM under limited and imbalanced data settings.

2. **Data Anomaly Detection through Semi-supervised Learning Aided by Customized Data Augmentation Techniques.**

Xiaoyou Wang*, Yao Du*, et al. **Structural Control and Health Monitoring, 2023 (Co-first author)**

Investigates data augmentation and semi-supervised learning strategies for anomaly detection under label scarcity in structural health monitoring.

3. **Wireless IoT Monitoring System in Hong Kong-Zhuhai-Macao Bridge and Edge Computing for Anomaly Detection.**

Xiaoyou Wang, Wanglin Wu, Yao Du, et al. **IEEE Internet of Things Journal, 2023**

Develops an end-to-end wireless IoT monitoring framework with real-world deployment for the Hong Kong-Zhuhai-Macao Bridge.

Professional Service

Conference Reviewer: CVPR2025, ICCV2025, NeurIPS2025, ICML2026, BMVC2026, MICCAI2026

Journal Reviewer: IEEE Transactions on Neural Networks and Learning Systems (TNNLS)

Awards

National Scholarship, Northwestern Polytechnical University, 2016, 2017, 2018

Outstanding Reviewer, CVPR 2025

Gold Reviewer Award, ICML 2026