

刘汉鹏 Hanpeng Liu

华中科技大学 计算机科学与技术 博士（直博） · 2027 届（2027 年 6 月毕业）

求职意向：多模态大模型 / 视觉基础模型 算法研究员

电话：+86 183-7263-7700 · hanpengliu@hust.edu.cn · 1359947339@qq.com

个人主页 · [Google Scholar](#) · [GitHub](#)



研究概述

- 研究方向：视觉-语言表征学习，覆盖图文对比预训练、CLIP 后训练、视觉自监督预训练与 MLLM 视觉编码器。
- 论文成果：发表 CCF-A 会议论文 4 篇（NeurIPS $\times 3$ 、CVPR $\times 1$ ），一作/共一论文 4 篇；理想汽车实习期间将图文预训练扩展到 128 卡 A800、十亿级数据。

教育背景

华中科技大学 计算机科学与技术 博士（直博） 导师：何琨教授 2022.09 – 2027.06（预计）

华中科技大学 计算机科学与技术 工学学士 专业排名 23/373（前 7%）；校优秀毕业生 2018.09 – 2022.06

实习经历

理想汽车（Li Auto） 研究型实习生 2024.09 – 至今

- 搭建视觉与视觉-语言基础模型的分布式训练与评测平台（DDP / FSDP / DeepSpeed）；将 CLIP 式图文预训练扩展到 128 张 A800、十亿级图文数据（保留梯度的 all-gather 汇聚负样本，WebDataset 流式读取）。
- 实习期间完成 ITO（NeurIPS 2026，一作）、ComCLIP（NeurIPS 2026，共同通讯）与 iGVLM（在投，一作）。

科研成果

按研究方向分组；完整作者列表见 Google Scholar；CCF 等级依据 2022 版目录

视觉-语言模型：图文预训练、CLIP 后训练与 MLLM 视觉编码

- ITO: Multi-View Alignment and Training-Time Fusion for Image-Text Pretraining. NeurIPS 2026 [CCF-A · 一作]
多视图对齐 + 训练期融合正则，推理开销不变；3M 至 1B 数据规模零样本稳定超越 CLIP（26 项平均 +1.2 至 +5.6）
- Rethinking Contrastive Loss in CLIP Post-training: A Complementary Framework with Frozen Text Encoder. NeurIPS 2026 [CCF-A · 共同通讯]
揭示后训练遗忘主要源于 InfoNCE 温度设置不当；单 epoch 方案线性探测 48.99 vs. 42.28（对比 CLIP-Refine）
- iGVLM: Dynamic Instruction-Guided Vision Encoding for Question-Aware Multimodal Understanding. 在投 [一作]
AdaLN 指令调制的双分支视觉编码器 + 诊断基准 MM4；MMStar 较 LLaVA-1.5 提升 3.6 至 4.5

视觉自监督预训练

- Separators in Enhancing Autoregressive Pretraining for Vision Mamba. arXiv 2026 [一作]
分隔符拼接多图做长序列自回归预训练；STAR-B 的 ImageNet-1K top-1 达 83.5%，检测与鲁棒性优于 ARM
- Boosting the Patch-based Self-Supervised Learning through Past-to-Present Smoothing. Pattern Recognition 2025 [CCF-B · 中科院一区 · 一作]

高效视觉模型与知识蒸馏

- You Only Need Less Attention at Each Stage in Vision Transformers. CVPR 2024 [CCF-A · 共同一作]
- Residual Learning for Self-Knowledge Distillation. IEEE Trans. Big Data 2025 [CCF-C · 共同一作]
- Hyperspherical Dataset Distillation via Contrastive Embedding Alignment. ICME 2025 [CCF-B]
- Grouped Logit Distillation Enhanced with Superclass Awareness for Efficient Knowledge Transfer. ECAI 2024 [CCF-B]
- Class Discriminative Knowledge Distillation. IEEE TETCI 2025 [SCI]
- Knowledge Distillation via Token-Level Relationship Graph Based on the Big Data Technologies. Big Data Research 2024

图表示学习

- Leveraging Contrastive Learning for Enhanced Node Representations in Tokenized Graph Transformers. NeurIPS 2024 [CCF-A]

专业技能

- 编程与框架：Python、PyTorch、timm、Hugging Face Transformers、OpenCLIP
- 训练经验：DDP / FSDP / DeepSpeed ZeRO、WebDataset 流式数据；LLaVA 式 MLLM 训练（Vicuna、Qwen2/2.5）

荣誉奖项与学术服务

- 华中科技大学三好学生、知行奖学金（2023、2024）；校优秀毕业生（2022）；中国研究生数学建模竞赛全国三等奖（2022）；编译系统设计赛（毕昇杯）全国三等奖，队长（2021）；全国大学生数学竞赛一等奖（2020）
- 审稿人：NeurIPS、ICLR、CVPR、AAAI；Pattern Recognition、IEEE Trans. Cybernetics