

# 郝天宇

电话: 18831215660

邮箱: [htianyu429@gmail.com](mailto:htianyu429@gmail.com)

主页: [htkxy.github.io](http://htkxy.github.io)



## 教育背景

西北工业大学 (985 / 211 / 双一流)

计算机学院 计算机科学与技术专业 学士

2022.09 – 至今

- 学业成绩: GPA: 3.918 / 4.1 (专业8 / 192) ; CET-6: 559
- 主要荣誉: 校级一、二等奖学金
- 未来规划: 直博保研至上海交通大学计算机学院, 研究方向为图像生成与编辑, 2026年9月入学;

## 论文发表

### 1. In-context Region-based Drag: Drag Any Region to Any Shape

共同第一作者 在投 ECCV 2026 Conference (CCF-B)

- **概述:** 本工作提出一种基于扩散模型的区域级图像拖拽编辑方法 ICRDrag, 在 **in-context learning** 框架下, 联合输入源图像、源区域掩码和目标区域掩码, 实现“任意区域到任意形状”的精确可控编辑, 有效克服传统点级拖拽方法的歧义性问题。方法设计了**两种关键注意力正则机制**, 包括**图像-掩码注意力一致性**以增强跨模态对齐能力, 以及**源-目标注意力对应**以提升区域语义映射与结构一致性; 同时构建了大规模 **Paired Region Dataset (PRD)** 以支持区域级编辑任务的训练与评估。在该项目中, 我主要负责多种 baseline 方法的**复现与系统性对比实验**, 完成**数据筛选与清洗流程**, 并参与**评估指标与用户实验设计**, 验证所提方法在编辑精度与视觉质量上的显著提升。

## 项目经历

### 面向工业缺陷检测的小样本缺陷图像生成项目 项目负责人

2025.10 – 至今

- **概述:** 针对工业异常检测中缺陷样本稀缺、传统数据增强方法生成结果真实感不足且缺陷区域控制不精确的问题, 我参与设计并实现了一个基于**视觉基础模型与流匹配生成框架的小样本缺陷图像生成方法**。项目以正常图像和参考缺陷图像为输入, 将缺陷合成建模为**上下文感知的局部编辑任务**, 构建了双分支缺陷特征提取模块用于学习细粒度缺陷语义, 并基于 **FLUX.1 Kontext** 设计了结合 inpaint-aware flow matching、缺陷特征匹配和正常区域保持约束的生成训练目标, 同时提出 **Generate--Select--Refine** 的三阶段掩码生成流程以提升缺陷区域与目标图像的对齐质量。在本项目中, 我主要负责方法设计、模型训练与实验分析, 完成了缺陷特征提取模块、生成目标函数和掩码生成流程的实现, 搭建并整理了与 baseline 的对比实验和消融实验, 验证了方法在缺陷生成真实感、多样性、掩码对齐及下游异常检测性能上的提升。

### 国家发明专利 《一种用于无人机视觉语言导航任务的数据增广方法》 学生一作

2024.08 – 2024.11

- **概述:** 针对无人机视觉语言导航任务中数据稀缺、自然语言导航指令质量差以及长路径下“关键动作稀疏、无关语义冗余”等问题, 提出一种基于**双尺度图Transformer (DUET) 模型和大语言模型**的层次化指令生成模型。通过启发式搜索生成路径, 结合仿真器获取视觉观察数据, 利用 ViT-B/16 和 BERT 提取多模态特征, 并通过动态合并策略减少冗余, 利用 DUET 的双尺度视觉表征能力和大语言模型的上下文推理能力, 生成高质量导航指令, 经 BLEU 等指标筛选后形成增广数据。现有方法依赖人工或低质量自动生成数据, 而本方法通过融合全局/局部视觉表征与语言模型语义空间, 显著提升指令的简洁性和关键动作密度。实验表明, 生成指令的多样性和准确性优于主流方法, 有效缓解“长平”问题。

### 科研训练项目: 资源受限场景下的大模型微调训练技术研究与实践

2025.03 – 2025.06

- **概述:** 围绕资源受限环境下大模型 (如 LLaMA) 微调中的显存优化问题, 深入研究高效微调方法的原理与工程实现。建立了 Decoder-only Transformer 显存开销评估模型, 量化各类内存需求, 为策略优化提供理论支持; 系统调研了 LoRA、QLoRA、激活值重计算与参数卸载等主流方法, 从算法和系统两个层面分析其优缺点; 最终基于 GSM8K 数据集在 LLaMA3.2-3B 模型上完成 5 种微调策略对比实验, 评估各方法在显存开销、训练时长与收敛质量方面的表现。我主要负责显存开销建模、微调方法调研与对比实验设计。建立 Transformer 显存评估模型, 量化内存需求; 系统分析 LoRA、QLoRA 等主流方法的技术原理与适用性; 在 LLaMA3.2-3B 上实现 5 种微调策略, 完成对比实验并评估训练性能与资源效率。

## 竞赛经历

美国大学生数学建模竞赛

国际级一等奖 (M 奖)

2025 年 2 月

全国大学生数学建模竞赛

国家级二等奖

2024 年 9 月

“启智杯”机器视觉设计大赛

优秀奖

2024 年 4 月

第 15 届全国大学生数学竞赛

省级一等奖

2023 年 12 月