

台谈

13025474644 | taiyan2020@ia.ac.cn



教育经历

中国科学院大学 (GPA 3.77/4.00) 2020年09月 - 2024年01月
人工智能 硕士 人工智能学院 北京

- 中国科学院自动化研究所 模式识别国家重点实验室
- 指导老师：王金桥
- 研究方向：多模态、语义分割、目标检测、异常检测

南昌航空大学 2013年09月 - 2017年06月
自动化 本科 南昌

科研经历

- Yan Tai***, Weichen Fan*, Zhao Zhang, Feng Zhu, Rui Zhao, Ziwei Liu. *Link-Context Learning in Multimodal LLMs* (Waiting for submission to AAAI2024)
 - 定义新的链接上下文学习概念并初步验证其在多模态大模型中的作用，即**training-free few-shot**学习未知类别
 - 提出数据集 ISEKAI，用于测试模型对未知类别的感知能力
- Bingke Zhu, **Yan Tai**, Yingying Chen, Wei Zhou, Ming Tang, Jinqiao Wang. *NextInd: Next Generation Pre-Trainer for Industrial Image Representation* (Waiting for Submission)
 - 开源新的大规模工业品数据集 Ind-2M，包含约220万张图像，涵盖多种工业场景和产品
 - 提出基于对比学习的大规模预训练模型**NextInd**，在工业缺陷检测任务中进行**fine-tune**可有效提升检测精度
- Yan Tai**, Bingke Zhu, Yingying Chen, Ming Tang, Jinqiao Wang. *Pointrefine: Patch-attention Based Small Objects Segmentation Refiner* (Waiting for Submission)
 - 提出新的“采样-优化”过程以及逐级优化训练方案，在逐级还原分辨率的过程中实现Coarse-to-Fine的优化流程
 - 提出即插即用的优化模块，可对已完成训练的分割模型实现低成本优化，且在小面积分割任务中效果显著
- 王金桥，陈盈盈，朱炳科，**台谈**，图像语义分割方法、装置、电子设备和存储介质，2021/12，中国

工作与实习经历

北京市商汤科技开发有限公司 2023年05月 - 至今
见习研究员 智慧城市与商业事业群 北京

- Sensetime-Monolith 多模态专题**：对多模态大模型 (MLLM) 能力进行评估、验证和优化
 - 长尾目标：探索Open-Vocabulary Detection(OVD)方案配合MLLM在长尾目标检测中的Zero-Shot能力边界
 - 复杂事件：探索In-Context Learning(ICL)方案配合MLLM在复杂事件研判中的Few-Shot能力边界
 - 统一模型：训练支持直接输出检测坐标的多模态大模型，支持采用Prompt方法快速适配多种学术及业务场景
 - 已完成工作：
 - 组合训练OVD+MLLM Zero-Shot方案，在31个业务事件中分别有8个超过和5个接近发版精度
 - 提出新的LCL方法并准备论文发表，业务上在8个事件中实现**training-free 8-shot**超越发版精度

- LCL应用于检测模型，训练支持few-shot精度提升的多模态大模型

中国科学院自动化研究所

2020年07月 - 至今

硕士研究生 模式识别国家重点实验室

北京

- **中公高科-双阶段路面病害分割**：解决路面分割任务中标注质量低、泛化性差、分割精度低等问题
 - 设计异常检测模型，实现高召回率路面病害检测
 - 设计语义分割模型，融合双阶段特征和结果，设计新的数据扩增和低质量标注监督方案
 - 提出可插拔的分割优化模块，针对小面积病害进行针对性的采样和优化
 - 算法已应用于全国多地高速公路病害检测任务中，且在多种评测指标上达到SOTA水平
- **华为云-轻量级人像抠图**：设计可在移动端实时运行的全自动人像抠图模型
 - 设计轻量级分割模型，快速高效地预测粗略分割结果
 - 提出保留边缘特性的羽化模块，对输入图像和分割结果直接生成Alpha Matte
 - 算法已应用于视频会议背景更换、证件照背景底色更换、表情包自动换脸等任务
- **其他视觉相关项目，如欣旺达电池印刷检测、欧派超晶板检测等**

北京阿丘科技有限公司

2017年09月 - 2019年09月

算法工程师 图像算法部

深圳

- **通用工业机器视觉软件VIDI**
 - 实现图像采集、模板匹配、边缘测量、卡尺工具、直线/圆拟合、条码扫描、颜色识别、仿射矩阵计算等视觉功能算法
 - 基于C#实现界面交互，将上述功能模块化为可拖选的控件，用户通过手动拖选控件组织逻辑和设置参数可自定义方案
- **通用工业AI检测软件AIDI**：面向工业检测项目基于Keras训练分类、分割模型

荣誉奖项

2022北京大数据技能大赛 —— 漆面检测赛道（一等奖）	2022
三好学生-中国科学院大学人工智能学院	2021
NAIC-“华为-昇腾杯” AI+遥感影像分割（17/2207）	2020

个人技能

- 熟悉深度学习相关工具、框架等
- 熟悉 C++/C#/Python 编程及 Linux 常用指令
- 持续跟进科研进展，快速复现相关工作