

文豪

年龄： 26

电话： 173-5651-1427

邮箱： 812367510@qq.com



教育经历

合肥工业大学	电气与自动化工程学院	智能电网信息工程专业	本科	2016.09 - 2020.06
• 获得校一等奖学金（前3%）、校三等奖学金（前14%）各一次				
浙江大学	航空航天学院	航空宇航科学与技术专业	硕博连读	2020.09 - 2025.06
• 获得学业优秀奖助学金一等奖一次				

项目经历

基于人工智能的喷涂行为识别技术开发（埃夫特智能装备股份有限公司） AI算法开发 2021.06 - 2022.06

- 任务目标：在工厂数字化升级的背景下,家具制造行业存在人工计数成本高、工人行为数据缺乏统计等问题,需要通过AI算法对家具厂工人的行为进行识别,帮助提升工厂工作效率。
- 工作内容：研发过程中先尝试了基于3D CNN的端到端的动作识别算法进行开发,但由于动作的模糊性和多样性导致模型泛化难度高。经过对于项目需求以及实际场景的分析,最终开发了以目标检测为中心的位置判别算法完成项目,通过构建时序状态字典,以人物交互产生的位置变动间接识别上料、下料、打磨、休息等动作。并基于识别结果统计工人的做工计数、有效工作时长。帮助工厂生成工人工作画像。这一方案将动作识别准确率提高了30%以上并且降低了90%以上的数据集制作成本。
- 工作成果：实现了95%以上的动作分类准确率。著有一项发明专利《一种面向智慧工厂的视频动作识别方法及系统》（已授权）。

基于骨架关键点的智慧养老关键模块开发（正元智慧集团股份有限公司） AI算法开发 2024.03 - 2024.10

- 任务目标：老龄化社会对智慧养老的需求日益增加,传统监控设备仅提供视频回放功能,缺乏实时预警能力,无法低成本的检测老人摔倒、长时间不动或离开监控范围等危险情况。特别是在独居场景下,时间延误可能导致严重后果。并且由于要在低算力边缘计算设备上运行,模型轻量化和实时性要求较高。
- 工作内容：设计了一套基于单帧和多帧的算法来跟踪匹配运动的骨架与人脸信息,并监控跌倒、长时间不动和离开监控区域等行为。系统使用骨架索引字典和状态字典来记录每个老人在一段时间之内的状态,并设计了一套基于关节点运动和位置关系的判据来判断老人有无发生异常行为。通过Socket方式与边缘计算单元进行通信。
- 工作成果：实现了94.4%的类平均分类准确率。匹配跟踪与动作识别模块的延迟小于30ms。协助开发了一整套包含摄像头、边缘计算单元Orin、和算法服务器构成的异常行为检测系统。著有一项软著《基于时序状态字典的多判据异常行为识别软件V1.0》。

智慧警务视频AI智能识别产品开发（杭州屹道科技有限公司） AI算法开发 2024.06 - 至今

- 任务目标：在警务交接场景中,枪械拉枪检验是关键安全环节,但人工检查成本高并且容易因疏忽或疲劳导致漏检。传统视频监控无法自动识别此类规范动作的执行情况,存在较大安全隐患。并且由于场景以及动作的多样性,容易导致算法拟合困难。
- 工作内容：通过基于传感器伪标签的方式降低了100%的人工标注成本,然后微调ir-CSN网络实现了端到端的拉枪动作识别。通过设计数据集的构建规范,以及编写伪标签生成数据集的脚本,构建了全流程自动化的数据集生产管线。通过基于对样例的分析,每条样本选择为连续时长为3秒的帧,起始时间为动作开始的前0.5秒。此外我们还通过时间平移以及倒放这种非典型的数据增强方式将样本数量扩展了10倍,并且通过测试将正负样本的比例设置为12:1,将动作漏检率降低了70%。
- 工作成果：实现了90%以上的拉枪检测准确率。著有一项发明专利《一种基于姿态估计的多判据动作识别方法》（实质审查阶段）。

论文

基于相位谱重构图的运动表征 (sci, 已接收, 一作) 2022.08-2022.12

- 工作内容：本文主要针对动作识别提出了一种新的辅助图像特征——相位谱重构图（phase spectrum reconstruction map）。该特征能有效突出动作相关信息,同时抑制外观冗余信息。并且与帧差法以及光流法不同的是,相位谱重构图是基于单帧图像获取的,具有更好的特征稳定性。

- 工作成果：在UCF101数据集以及HMDB51数据集上分别实现了96.6%以及77.8%的准确率。在稳定性方面，能够有效抵抗镜头缩放以及移动、镜头切换、慢速移动、长距离帧间隔的攻击。

引入物体信息增强的骨架动作识别方法 (sci, 已接收, 一作)

2023-04-2023.10

- 工作内容：针对基于骨架的动作识别方法在处理人物交互动作方面的不足，提出在骨架图中引入交互对象节点以补充交互对象信息的缺失，并且提出了随机节点攻击 (Random Node Attack) 的数据增强方法，解决因引入物体信息而导致的过拟合问题。
- 工作成果：所提出的ST-VGCN 框架成功捕获了人类动作与交互物体的关系，大幅提高了对人物交互动作的识别能力。在NTU RGB+D 60 CS数据集中达到了96.7%的准确率，超过之前最先进的结果 (94.1%)。构建了一个工业环境下包含人物交互信息的动作识别数据集JXGC 24。

基于姿态估计特征图增强的骨架动作识别方法 (sci, 在审, 一作)

2024.04-2024.09

- 工作内容：提出一种基于姿态估计网络的特征图增强的骨架表征 (PFMESR)，通过利用姿态估计网络内部的副产物特征图，以及极低的计算代价增强了骨骼特征在外观表示方面的不足。
- 工作成果：PFMESR以增加0.4%计算量的代价在骨骼动作识别任务中表现出显著性能提升，在NTU RGB+D 60、NTU RGB+D 120与UAV-Human上都实现了先进的效果。尤其是在需要人体外观信息的任务 (如 Action-Identity Recognition) 中表现出重要价值 (分类性能提升2%-10%)。

基于扩散模型的文档信息隐藏方法 (CCF-A, 在审, 共同一作)

2024.10-2024.11

- 工作内容：首次在文本水印算法中引入扩散模型，利用字符关键点指导字符的局部编辑，将信息嵌入到字符结构中。
- 工作成果：DiffMark 实现了自动化信息嵌入与提取的pipeline。并且在主流数据集上的 PSNR 提高了 5 dB, 提取准确率提升超过 5%，表现优于现有最先进方法。DiffMark 在隐蔽性、鲁棒性以及跨字形的泛化能力方面全面优于传统方法，展示了较高的实用价值。

个人技能

-
- 熟悉linux平台, 熟悉python语言的算法程序开发, 熟悉深度学习框架Pytorch。
 - 熟悉深度学习理论, 熟悉计算机视觉领域常见方向(动作识别, 目标检测等)的落地算法开发。
 - 具有基本的英语专业文献阅读能力(CET-6)。