

复杂场景下 AI 视频监控系统异常行为识别的精度优化技术研究

徐煜洋

西北师范大学 甘肃兰州 730000

摘要: AI 视频监控系统在复杂场景下对异常行为识别精度的优化具有重大意义。随着监控需求的增长,系统面临诸多挑战,如复杂环境中识别鲁棒性差、海量异构数据整合困难等。本文围绕复杂场景下 AI 视频监控系统异常行为识别精度优化技术展开研究,从数据预处理、模型优化、特征提取与融合以及环境适应性增强等方面进行理论探讨,旨在提升系统在复杂场景下对异常行为识别的精度,为相关技术的发展提供理论支撑,助力 AI 视频监控系统在更多复杂场景中高效应用,增强安全保障与管理效率。

关键词: AI 视频监控系统; 复杂场景; 异常行为识别; 精度优化

1. 引言

AI 视频监控系统在数字化时代已然成为确保公共安全、增进管理效率的关键途径。其应用之广泛,自智慧城市的大规模布设,至工业生产的精密调控,乃至校园、医院等场地的安全防护皆有体现。作为 AI 视频监控系统核心功能的异常行为识别,其精度高低,与系统实用价值紧密相关。在公共场所,对异常行为的精确辨识,可及时察觉潜在安全威胁,诸如暴力冲突、非法闯入等,为安保人员迅速采取应对举措争取宝贵时机,有效防范危险事件的发生或恶化。于工业生产场景里,对工人异常操作行为的准确识别,有助于及时匡正错误,规避生产事故,保障生产流程的顺畅与安全。深入钻研复杂场景下 AI 视频监控系统异常行为识别精度的优化技术,极为迫切,此不仅是技术演进的必然走向,更是契合社会安全与高效管理需求的关键要点。

2. 数据预处理技术

2.1 图像增强

在复杂场景状况之下,视频图像质量往往遭受多种因素的作用^[1]。光照不均衡的情形出现时,部分区域呈现出过亮抑或过暗的状态,致使细节有所遗失。直方图均衡化的采用能够凭借对图像灰度值的重新分配,实现灰度动态范围的拓展,使得图像整体对比度获得增强,让原本隐匿于暗处或者过亮区域的细节能够清晰地展现出来。针对于由低光照所引发的图像模糊现象,自适应直方图均衡化(CLAHE)具备更为良好的效果。CLAH E 依据图像局部区域的特性,针对不同的小块分别施行直方图均衡化,规避全局直方图均衡

化有可能导致的过度增强或者噪声放大等问题,切实提升局部细节的清晰度,为后续异常行为的识别奠定更为清晰的图像基础。

2.2 去噪处理

在视频采集阶段,噪声干扰是无法规避的状况,高斯噪声以及椒盐噪声等均属此类。中值滤波算法在椒盐噪声的去除领域展现出卓越成效。该算法把图像里每个像素点的数值,更替成其邻域像素值的中值,能够高效地滤除椒盐噪声所带来的脉冲干扰,与此同时,对于图像边缘等关键信息,亦能做到较好的留存^[2]。而对于高斯噪声,高斯滤波借助高斯函数针对邻域像素实施加权平均操作。依据高斯分布的特质,离中心像素距离越近的点,被赋予越高的权重。如此,在实现噪声平滑的同时,维持图像的平滑过渡,降低对图像原本结构的破坏程度。借由有效的去噪处置,能够削减噪声对异常行为特征提取工作造成的干扰,提升识别的精准度。

2.3 数据标准化

数据源自不同类型的监控设备采集,在分辨率和色彩空间等维度上存在显著差异^[3]。针对分辨率统一之需求,双线性插值算法或是双三次插值算法可供采用。借助线性抑或三次函数关系,依据周边像素点灰度值或者色彩值,计算目标分辨率下崭新像素点之数值,实现图像缩放之目的,以确保所有数据于相同分辨率基础之上开展处理。至于色彩空间转换,惯常之 RGB 色彩空间被转换为 HSV(色调、饱和度、明度)色彩空间。HSV 模型与人类视觉感知特性契合度更高,在颜色信息分析及处理方面更具便利性,促使不同来源数据

于色彩维度具备一致性，为后续模型学习与识别提供便利。

3. 模型优化技术

3.1 模型选择与改进

识别复杂场景下异常行为的需求面前，深度学习模型各呈优劣。卷积神经网络（CNN）于图像特征提取展现强大效能，卷积层借不同卷积核萃取图像局部特征，池化层对特征图实施降维，削减计算量并强化特征的平移不变特性^[4]。以复杂场景行人异常行为识别为例，改进的 CNN 模型在传统架构之上增添卷积层，加深网络深度，可以学习更复杂、抽象的特征，提升对细微异常行为特征的捕捉能力。循环神经网络（RNN）及其变体——长短期记忆网络（LSTM），在处理时序数据方面专长凸显，于连续视频帧行为序列分析具备优势。异常行为若涉动作的时间先后顺序及持续时长，LSTM 凭借门控机制化解 RNN 处理长期依赖之难题，能够有效记忆与运用历史信息，精准识别具备时间序列特征的异常行为模式。

3.2 模型训练优化

数据增强是扩充训练数据集以及提升模型泛化能力的重要途径之一。于图像翻转过程中，图像的水平翻转或者垂直翻转，能够增添数据的多样性，促使模型习得不同方向的行为特征，进而强化对目标姿态变化的适应性。图像旋转，即依照特定角度对图像实施旋转，以模拟实际场景里目标或许呈现的不同角度，从而丰富训练数据的角度信息。裁剪操作，是从原始图像中随机裁剪出尺寸各异的区域，使得模型接触目标在不同局部区域的特征，提升对目标位置与尺度变化的鲁棒性。迁移学习可对于大规模通用数据集上预训练的模型参数加以有效利用。把于 ImageNet 等通用图像数据集上预训练的模型，针对复杂场景下的异常行为识别任务予以微调。预训练模型已然学习到众多通用的图像特征，在新任务上进行微调时，模型能够更快收敛，削减训练时间与数据需求，同时规避在小数据集上训练可能产生的过拟合问题，提高模型在复杂场景下的识别性能。

4. 特征提取与融合技术

4.1 时空特征提取

对异常行为识别而言，于复杂场景准确抽取视频里时空特征意义重大。空间特征提取方面，3D 卷积神经网络（3DCNN）于传统 2DCNN 之上，增添时间维度卷积操作^[4]。其 3D 卷积核不仅于图像二维平面滑动以抽取空间特征，且

沿时间维度对连续视频帧卷积，也可以直接习得视频时空联合特征。对于打架斗殴这类涉及多人动作交互异常行为识别时，3DCNN 可同步捕捉不同人物空间位置、姿态及随时间变动动作序列，有效抽取此等异常行为关键特征。光流法亦为一种关键时空特征提取方式。它借由计算视频相邻帧间像素点运动信息，获取光流场，光流场中向量表征像素点运动方向与速度。奔跑、追逐等具显著运动特征异常行为，光流法能精准捕捉目标物体运动轨迹与速度变化，为异常行为识别供给重要时空线索。

4.2 多模态特征融合

监控数据于复杂场景之下，常涵盖多样模态信息，诸如视频图像、音频及传感器数据等。人物外观特征、动作姿态特征等，可自视频图像模态予以提取；在音频模态中，通过对声音频率、强度等特征的分析，可以识别出异常的呼喊声、撞击声等；在某些场景中，温度传感器、压力传感器数据等传感器数据，可辅助判定异常情况是否存在，例如温度异常升高或暗示火灾隐患。早期融合是于数据采集阶段便对多模态数据施行合并处理，像将视频图像与同步采集的音频数据，在预处理前依特定方式组合，随后统一输入至后续处理模型之中，此方式可使模型从起始便习得多模态数据的联合特征。至于晚期融合，先分别针对不同模态数据开展独立处理与特征提取，获取各模态的分类结果或者特征表示，接着对这些结果或特征进行融合决策，比如借由投票机制或者加权融合等途径确定最终的异常行为判断，该方式可充分施展各模态数据的优势，提升异常行为识别的准确性与可靠性。

5. 环境适应性增强技术

5.1 光照与天气适应性

视频监控系统于复杂场景里，光照条件以及天气状况变化对其影响彰显出显著性。自适应光照补偿算法被采用以应对光照变化，此算法能够针对图像的光照强度与分布状况予以实时监测，并依据预先设定的光照模型对图像开展动态调整^[5]。当光照处于过强态势时，图像整体亮度需降低，以此避免过曝区域信息的丢失；而当光照过弱，图像亮度要增强，使细节可见度得以提升。针对不同天气条件，雨、雪天气下，借助图像去雨、去雪算法，借由构建雨、雪噪声模型，将雨滴、雪花干扰从视频图像中予以去除，使清晰的背景以及目标物体图像得以恢复。对于大雾天气，应选用图像去雾

算法如基于暗通道先验的去雾方式，通过剖析图像局部区域的统计特性，对大气光值和透射率进行估计，进而对图像实施去雾处理，提高图像清晰度，保障系统在各类光照与天气条件下皆可稳定提取异常行为特征，提升识别精度。

5.2 遮挡处理

在复杂场景当中，遮挡作为一种常见问题，对异常行为识别精度有着严重影响。多视角信息融合方法会借助多个处于不同角度的摄像头来采集视频数据，一个视角若出现遮挡状况，完整的目标信息或许能从其他视角获取。通过对多视角视频数据展开联合分析，将不同视角下的特征予以融合，因遮挡导致的信息缺失可得到有效弥补，对被遮挡目标异常行为的识别能力也得以提高。人体结构模型方法，会先构建人体骨骼结构模型，当遮挡发生，依据人体结构的先验知识以及未被遮挡部分的特征，对被遮挡部分的位置与姿态加以推测，以此对完整的人体动作展开估计与识别，降低遮挡对异常行为识别产生的影响，增强系统于复杂遮挡场景下的鲁棒性。

5.3 动态场景下的目标跟踪优化

在复杂场景下，目标的动态移动状态频繁呈现，并且背景干扰、目标遮挡等因素易于对其产生影响，这种情形致使传统跟踪算法出现跟踪漂移或者丢失等问题。一种可行的改进跟踪算法是与相关滤波相结合的算法。该算法通过构建自适应相关滤波模板，能够实时更新目标的特征信息。前一帧目标的位置与特征，该算法能够加以利用，以此对滤波模板参数展开动态调整，增进对目标外观变化的适应性。比如，当目标因移动使得角度、尺度发生改变之时，滤波模板可以迅速匹配新的目标特征，进而让跟踪误差得以减少。

卡尔曼滤波预测机制的结合，对跟踪稳定性有进一步提升之效。借由构建目标运动模型，卡尔曼滤波依循历史运动轨迹，对目标下一帧的潜在位置加以预测，复结合当前帧的观测资讯，对预测成果予以修正，从而形成“预测-修正”的闭环跟踪模式。此模式在应对目标短暂遮挡情境时颇具成效，当目标处于被遮挡状态，借助预测模型维系跟踪态势，待目标再度现身，能迅速恢复精准跟踪，为后续异常行为识别供应连续且稳定的目标运动轨迹数据，规避因跟踪中断引发的识别漏判状况。

5.4 多场景下的模型自适应调整

对于视频监控系统而言，其在应用过程中，频繁面临

于各类复杂场景间转换之情形，诸如自室内商场场景过渡至室外交通场景。在此类场景转换时，场景环境、目标类型以及行为模式方面，均呈现出明显的差异性。鉴于此，固定的模型参数难以契合多场景识别之需求。一种可行途径是构建基于对场景特征予以感知的模型自适应调整机制。此机制通过对当前场景关键特征（例如光照强度、背景复杂度、目标密度等）的提取，实现场景特征与模型参数间映射关系的建立。

当系统检测到场景切换时，会依据当前场景的特征自动调用预设的参数调整策略。于目标密度较高之交通场景里，模型对小目标的检测权重增大，目标检测算法的IoU阈值得以优化，目标重叠引发的误识别随之减少。在背景复杂的室外场景中，模型对背景噪声的抑制能力增强，特征提取层的感受野大小获调整，目标特征与背景特征的区分度进而提升。引入在线学习机制，让模型在新场景应用进程里，依据实际识别结果持续微调参数，借助积累场景特定数据优化模型性能，逐步提升对新场景的适应能力，实现多场景下异常行为识别精度的稳定保持。

6. 结语

本文聚焦于复杂场景下AI视频监控系统异常行为识别之精度优化技术，从数据预处理、模型优化、特征提取与融合、环境适应性增强这四个关键维度展开理论阐述。数据预处理，为识别供给高质量数据根基；模型优化，促使算法核心性能提升。特征提取与融合，对关键信息捕捉予以强化。环境适应性增强，于复杂多变场景里保障系统具备鲁棒性。复杂场景下识别精度不足问题，经这些技术从不同层面化解，进而构建起完备的精度优化技术体系。

参考文献：

- [1] 赵创. 视频监控下异常行为检测系统设计 [D]. 燕山大学, 2023. DOI:10.27440/d.cnki.gysdu.2023.002079.
- [2] 付星原. 基于异常行为检测的智能视频监控系统的设计与实现 [D]. 北京邮电大学, 2022. DOI:10.26969/d.cnki.gbydu.2022.002591.
- [3] 曾婷, 黄东军. 智能视频监控系统异常行为检测算法研究综述 [J]. 计算机测量与控制, 2021, 29 (07): 1-6+20. DOI:10.16526/j.cnki.11-4762/tp.2021.07.001.
- [4] 姜珊. 视频监控系统中的行人异常行为检测方法研究 [D]. 沈阳理工大学, 2021. DOI:10.27323/d.cnki.

gsgyc.2021.000030.

[5] 陈正晓 . 基于视频监控的异常行为检测算法设计与

嵌入式系统实现 [D]. 山东大学 , 2020. DOI:10.27272/d.cnki.

gshdu.2020.004816.

作者简介: 徐煜洋 (2004年11月-), 男, 籍贯: 浙江省台州市, 民族: 汉族, 职称: 无, 学历: 本科, 研究方向: 软件工程