

设计研究与应用

深度学习在安防技术中的应用研究

柴伟

(中国人民银行乌鲁木齐中心支行, 新疆乌鲁木齐 830002)

摘 要: 在安防系统中随着数字化硬件设备的不断推广使用, 传统的目标检测等视频分析技术已不能满足多种多样复杂场景的需要, 智能化视频分析技术已成为安防监控系统中重要的辅助手段, 利用 YOLOv5 算法建立适合发行库安全管理中人员未同操作、睡觉等场景的模型, 通过建立模型可自动判断人员异常行为, 通过与安防系统中各子系统之间的联动, 主动发出预警, 进一步提升安防系统效能, 对安防系统建设及应用具有重要意义。

关键词: 深度学习; 视频分析; 安全管理; 大数据

中图分类号: TP391

文献标识码: A

DOI: 10.3969/j.issn.1003-6970.2023.02.030

本文著录格式: 柴伟.深度学习在安防技术中的应用研究[J].软件,2023,44(02):126-128

Application Research of Deep Learning in Security Technology

CHAI Wei

(The People's Bank of China Urumqi Central Sub-branch, Urumqi Xinjiang 830002)

【Abstract】: With the continuous promotion and use of digital hardware equipment in security systems, video analysis technology has become an important auxiliary means in video surveillance systems. Traditional video analysis technologies such as target detection can no longer meet the needs of a variety of complex scenes. With the development of intelligent technology, deep learning has gradually become the mainstream research direction. In scenarios with big data, the use of YOLOv5 algorithm to establish models suitable for various complex scenarios is applied in security systems, which is of great significance to the construction and application of security systems.

【Key words】: deep learning; video analysis; security management; big data

0 引言

安防系统由视频监控系统、报警系统、门禁系统等子系统组成, 一般通过管理平台对所有子系统进行管理, 使每个独立的子系统能够联动, 充分发挥每个子系统作用。安防系统主要是通过硬件设备实现目标检测与预警, 但硬件设备检测手段和范围较为局限, 需要被动触发才能进行告警, 在安防系统中需要对人员异常进行分析和检测, 目前只能依靠人工进行判定, 人员工作强度较大, 因此利用深度学习算法对安防系统进行辅助管理, 突破硬件设备短板和局限性, 利用已有数据建立人员异常行为检测模型, 从而进一步提升安防系统应对复杂场景能力。

1 深度学习背景及应用现状

深度学习^[1]是通过建立多层次网络模拟人类大脑神经网络结构提取数据特征, 总结数据规律的一类算法总称。

自 20 世纪 40 年代至 60 年代模拟生物学习的计算模型, 到 20 世纪 80、90 年代用多层次网络将大量神经元连接在一起实现智能行为, 创造了反向传播算法, 在序列建模中取得重要进展, 但到 2006 年左右, 因计算复杂度高, 硬件设备难以满足需要, 热度降低。直至近年来, 随着计算机硬件设备运算能力的不断提升, 大数据得到广泛应用, 深度学习在安防等多个行业得到广泛应用。

在安防领域中, 深度学习主要用于人脸识别、车辆牌照识别、特定场景异常检测等方面, 尤其在图像领域, 从识别图片中的物体, 到逐步应用于视频中异常场景检测与识别, 对视频中目标的动作和行为分析以及目标的跟踪。随着视频数据获取手段的多样化, 视频数据数量及质量的不断提高, 深度学习算法在视频分析中的准确率得到很大提升^[2,3]。

目前基于深度学习的图像领域中目标检测算法主要

作者简介: 柴伟 (1994—), 女, 新疆昌吉人, 硕士研究生, 中国人民银行乌鲁木齐中心支行副主任科员, 研究方向: 人工智能、安防技术。

分为2类：(1) TWO-stage 算法，以 Faster R-CNN 为代表，检测目标分为2个阶段，首先生成包含待检测物体的预选框，其次调整边界框；(2) ONE-stage 算法，以 SSD、YOLO 为代表，直接利用深度学习网络提取特征来预测图像中的物体类别和位置，进行细粒度的物体检测。

2 安防技术现状及存在问题

2.1 安防技术现状

目前安防系统使用率较高的子系统为视频监控系统，在金融行业等安全等级要求较高的场所结合报警系统、门禁系统进行使用，但现有视频监控系统主要用于视频的实时预览及录像存储，存储的视频数据主要用于后期追溯使用，在特殊应用场景中，视频监控系统需要人工对异常事件及人员的异常行为进行判断和监督，无法做到系统自动对人员进行行为检测，对人工依赖性较高。报警系统通过使用一种或多种不同类型探测器，通过对声音、光线及震动等变化进行检测报警，大多需要被动触发，由人工对产生的报警事件进行判断和处置。门禁系统对人员出入进行管控，对人员认证采用刷卡、指纹、虹膜识别、人脸识别等，仅是对人员身份进行认证。视频监控系统与报警系统、门禁系统之间关联性较低，每个系统都需要人工对异常事件进行判断和处置，安防系统中所有子系统之间联动效率不高。

2.2 存在问题

2.2.1 安防系统无主动预警功能

目前安防系统中视频监控主要用于实时预览及录像回放功能，实时视频画面需要大量的人员对异常情况进行人为判断，录像回放只是作为事后追溯的手段，报警系统中红外、震动等探测器依靠外界被动触发，门禁系统用于人员身份核实和出入信息的记录，各子系统之间关联度较低，联动效果较差，无法在重要安防领域中达到安防系统自动对异常情况预警的目的。

2.2.2 视频数据量大，但利用率较低

随着视频监控高清化、数字化发展，以 100 路视频通道存储 3 个月计算，视频数据总量达到 200TB，占用大量存储空间，实际工作中视频数据往往只作为一种事后追溯的手段，未能发挥数据量级较大的优势，并且在调阅录像资料时，需要人工对视频数据进行排查，浪费大量人力、物力。

3 深度学习与安防技术结合框架

3.1 深度学习优势

3.1.1 节约人力成本

在视频监控系统中使用深度学习算法，替代了传统

的人工监督方式，在人员做出异常行为时，使用深度学习算法建立的模型进行识别并联动报警系统进行提醒^[4,5]。通过模型对视频画面中异常情况进行准确定位^[6-8]，改善人工浏览视频易产生错判、效率低的情况，进一步解放人力，降低人为识别产生的错误，提升安防系统的智能化水平。

3.1.2 发挥安防系统主动预警作用

目前安防系统已有数据量级较大，通过将数据进行分类，建立适合不同应用场景的模型，有针对性地对特定场景异常情况进行预测，可达到较高的识别率，进一步提升安防系统主动预警的能力。

3.1.3 发挥视频监控系统实时性特点

人员目标检测与在报警系统联动，电子大屏弹图等综合应用。人脸识别与人的行动轨迹结合，提供预警和行为预测，为情报信息研判和安全形势分析提供参考。例如在发行库等重要安防领域，利用现有视频监控设备及视频数据，在发行库内人员有异常行为时实时进行分析并对利用语音、文字等形式进行提醒，大大提高工作效率，减轻人员劳动强度，实现利用技术手段监督制约人员行为动作。

3.2 分析算法

3.2.1 YOLO v5 目标检测算法

YOLO v5 目标检测算法分为输入端、Backbone、Neck、Prediction 4 个部分，输入端采用 Mosaic 数据增强、自适应锚框计算、自适应图片缩放；Backbone 主要由 Focus 结构和 CSP 结构组成；Neck 由 FPN+PAN 结构组成；Prediction 由 CIOU_Loss 组成。

Backbone 是在不同图像细粒度上聚合形成图像特征的卷积神经网络。Neck 是一系列混合和组合图像特征的网络层，将图像特征传递到预测层。Prediction 是对图像特征进行预测，生成边界框并预测类别。

Mosaic 数据增强采用 4 张图片，随机缩放、随机裁剪、随机排布的方式进行拼接，进一步丰富检测数据集，特别是随机缩放增加了很多小目标，使网络鲁棒性更好。其次进行自适应锚框计算，设置初始设定的长宽锚框，在训练中，网络在初始锚框的基础上输出预测框，进而与真实框 Groundtruth 进行比较，计算两者差距，在反向更新，迭代网络参数。自适应图片缩放是将不同长宽的图片统一缩放到一个标准尺寸。

Focus 是在图片进入 Backbone 前，每隔一个像素取一个值，获得 4 个独立的特征层，将这 4 个特征层进行堆叠，此时就将宽高维度上的信息转换到了通道维度，输入通道扩充了四倍，进行特征扩充，再进行特征

的提取。

CSP 是将基础层的特征映射划分为两部分, 然后通过跨阶段层次结构进行合并, 在减少计算量的同时保证准确率。

FPN 是自顶向下的, 将高层的特征信息通过上采样的方式进行传递融合, 得到进行预测的特征图, 在此基础上增加一个自底向上的特征金字塔 (PAN), 通过结合, FPN 自顶向下传达强语义特征, 而特征金字塔自底向上传达强定位特征, 从不同的 Backbone 对不同的检测层进行参数聚合。

CIoU_Loss 在预测框与标注框的重叠面积、中心点距离、长宽比作为损失函数, 用来更新网络参数。

3.2.2 人员未同操作事件的判断

在发行库的管理要求中, 需要发行库管理人员 (2 人或 3 人) 在同一区域进行操作在通过, 视频帧数据使用 YOLO v5 进行目标检测, 对一定时间范围内的帧序列进行判定, 建立视频帧判定结果缓冲区, 缓冲区内存储目标检测结果 (人数), 在设置的一定人数阈值 (如: 2 人或 3 人) 判定, 对缓冲区结果并进行统计, 所在一定时间内 (如: 5min) 目标检测结果始终小于设定阈值, 则判定为人员未进行同操作。

3.2.3 睡觉事件的判断

在发行库管理中, 需要守卫值班人员对发行库内异常情况进行处置, 针对守卫值班人员的异常情况主要为睡觉事件, 因人员睡觉时动作和行为相对静态, 通过建立相邻两个视频帧对比结果缓冲区, 对每个视频帧数据使用 YOLO v5 进行目标检测, 对检测结果进行相似度分析, 若前后两帧相似度较高, 则证明人员动作和行为未产生变化, 若在一定时间段内 (如: 15min), 检测结果相似度较高, 则判定该时间段内人员睡觉。

3.3 结合框架

本次深度学习算法在安防中应用框架如图 1 所示, 用于对发行库业务相关人员进行目标检测, 利用目标检测算法结果判定对人员未同操作事件和睡觉事件, 达到对异常行为进行分析通过报警系统发出提醒并预警。通过前端摄像头采集的图像与后端深度学习算法结合, 弥补硬件设备被动触发的缺点, 达到主动识别检测与预警, 增强安防系统主动识别与防御能力。

4 结语

通过使用深度学习算法克服了安防系统依赖人工和硬件设备被动防御缺点, 使安防系统向检测、预警、主

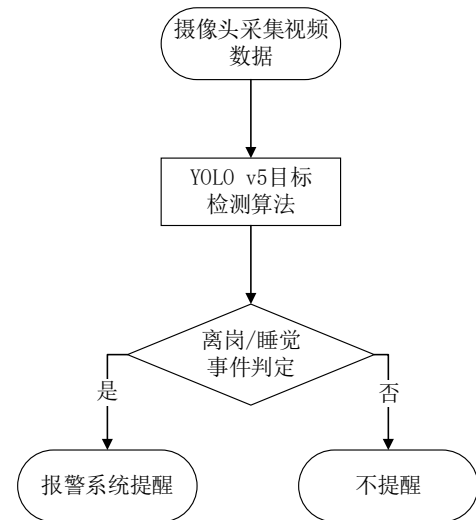


图 1 深度学习与安防系统结合框架

Fig.1 Combination framework of deep learning and security system

动防御转变, 安防系统智能化程度更高, 降低人员工作强度, 深度学习结合发行库等重要安防领域实际工作情景, 对人员未同操作、睡觉事件进行自动判定, 通过安防各子系统之间相互联动, 发出预警提示, 进一步提升发行库安全管理效率, 推动安防系统向自动化、智能化方向发展。

参考文献

- [1] 罗荣,王亮,肖玉杰.深度学习技术应用现状分析与发展趋势研究[J].计算机教育,2019(10):19-22.
- [2] 黄凯奇,陈晓棠,康运锋,等.智能视频监控技术综述[J].计算机学报,2015,38(6):1093-1118.
- [3] 任彦鹏.深度学习在视频安防监控系统的应用[D].成都:电子科技大学,2020.
- [4] 万里晴.基于深度学习的视频动作识别方法研究[D].北京:北京交通大学,2019.
- [5] 廖煊龙.面向智能视频监控的目标检测和行为识别技术研究[D].成都:电子科技大学,2019.
- [6] 徐博.基于深度学习的视频目标跟踪研究[D].大连:大连理工大学,2020.
- [7] 张洁.基于深度学习的监控场景异常检测方法研究[D].合肥:中国科学技术大学,2020.
- [8] 史璐璐.深度学习及其在视频目标跟踪中的应用研究[D].南京:南京邮电大学,2019.