

基于深度学习的园区安防吸烟行为检测^{*}

陈 赛^{1,2} 左云波^{1,2} 郑伊凡^{1,2} 胡 欢² 谷玉海^{1,2}

(1. 北京信息科技大学现代测控技术教育部重点实验室 北京 100192; 2. 北京信息科技大学机电工程学院 北京 100192)

摘要: 针对园区禁烟区时常出现一些流动人员违规抽烟造成的安防隐患问题,提出一种联合人体骨骼关键点检测和改进的YOLOv7烟支检测的吸烟行为深度学习检测方法。该方法首先通过OpenPose提取人体关键点的坐标信息,计算手、鼻子、脖子之间距离的比值,手、肘、肩之间的角度,判别是否满足吸烟姿态。然后联合改进的YOLOv7算法检测图像中是否存在烟支,来最终判断吸烟行为是否存在。其中改进的YOLOv7算法引入了全局注意力机制模块,强化了语义与位置信息,使用转置卷积改进上采样方式,减小了信息丢失,并采用MPDIOL损失函数,增强了回归结果准确性,提升了对烟支小目标的检测精度。通过实验测试,本文方法准确率达到95.45%,可以有效地检测出吸烟行为。

关键词: 深度学习;吸烟行为检测;人体关键点;改进的YOLOv7

中图分类号: TN29 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Detection of smoking behavior in park security based on deep learning

Chen Sai^{1,2} Zuo Yunbo^{1,2} Zheng Yifan^{1,2} Hu Huan² Gu Yuhai^{1,2}

(1. Key Laboratory of Modern Measurement and Control Technology, Ministry of Education, Beijing Information Science and Technology University, Beijing 100192, China; 2. Mechanical Electrical Engineering School, Beijing Information Science and Technology University, Beijing 100192, China)

Abstract: Aiming at the hidden security risks caused by some mobile workers illegally smoking in no-smoking areas of the park, a deep learning method for smoking behavior detection combined with human bone key point detection and improved YOLOv7 cigarette detection was proposed. The method first extracts the coordinate information of the key points of the human body through OpenPose, calculates the ratio of the distance between the hand, nose and neck, and the angle between the hand, elbow and shoulder, and determines whether the smoking posture is met. Then, the improved YOLOv7 algorithm is combined to detect whether there is a cigarette in the image to finally determine whether there is smoking behavior. The improved YOLOv7 algorithm introduces a global attention mechanism module, strengthens the semantic and position information, uses transposed convolution to improve the upsampling method, reduces information loss, and adopts the MPDIOL loss function to enhance the accuracy of the regression results and improve the detection accuracy of small cigarette targets. Through experimental tests, the accuracy of this method reaches 95.45%, which can effectively detect smoking behavior.

Keywords: deep learning; smoking behavior detection; human body key points; improved YOLOv7

0 引言

据相关医学研究,吸烟产生的致癌物质多达70种^[1],其不仅有害健康,而且违规抽烟也极易造成火灾等事故^[2],随着工业、商业、旅游和科技等各类园区的快速兴建,园区的安防问题也成为社会的关注热点^[3]。尤其对于化工类包含很多易燃易爆品的园区^[4],除了规范园区人员的日常吸

烟行为,强化消防设备外,及时监测并警告园区内其他流动人员的违规吸烟行为,也是保障安全的重要环节。

目前,吸烟行为识别的方法主要分为传统方法和基于计算机视觉的方法两大类。其中,传统方法主要有为监督、烟雾报警器、基于WiFi设备^[5]和基于可穿戴传感器^[6]等,但是这些方法都存在显著缺陷。对于人为监督而言,需要大量人力,易疲劳,无法满足24小时实时监督;对于烟雾

报警器而言,在稀薄环境中无法有效检测,特别是在通风良好或面积较大的场所;对于 WiFi 设备的方法而言,需多台接收器工作以获稳定效果,但易受磁场环境干扰,实现困难;对于穿戴设备而言,可以通过手环等检测运动和生理数据判断吸烟,但适用于主动戒烟人群,普及困难。

在计算机视觉方面,最初是基于香烟烟雾特征变化进行吸烟行为检测,潘广贞等^[7]利用混合高斯模型提取实际背景,通过帧差法识别稀薄烟雾区域,然后进一步分析变化规律来进行吸烟检测。汪祖云等^[8]利用吸烟烟雾在 HSV 颜色空间下的特殊性,先初步定位疑似区域,进行面积和质心变化分析,然后再利用 SVM 训练判定吸烟行为。相较传统仪器设备,虽然检测效果有所提升,但仍避免不了烟雾飘散带来的干扰。相反,对于烟支固体目标检测则相对稳定,韩贵金等^[9]通过图像分割初步筛选烟支目标,再用 Faster R-CNN 算法进一步检测,提高了对吸烟行为检测的准确性。杨国亮等^[10]将 Transformer 引入网络颈部,并增加了小目标检测层来改进 YOLOv5s 算法以提高对烟支目标的识别能力,进行吸烟行为检测。由于烟支目标较小,对于户外复杂场景,纯粹的烟支检测,误检现象常会发生,因此,一些复合判定的方法被提出,赵鉴福等^[11]使用 Azure Kinect 深度相机获取 32 个三维人体关键点,通过计算手部-头部关键点距离初判吸烟姿态再联合新增检测尺度的改进的 YOLOv3 算法进行吸烟行为检测,但该方法成本较高,对于不同人物,距离判断泛化能力较弱。徐婉晴等^[12]提出一种利用 AlphaPose 和 RetinaFace 算法分别检测出人体和脸部关键点信息,通过制定手腕-嘴部-眼睛的距离比判断吸烟姿态再结合 YOLOv4 检测烟头的吸烟检测方法,但 AlphaPose 是一种自上而下的算法,若前置的人体边界框的定位不准确,则无法进行人体姿态预测,且未经改进的 YOLOv4 算法对烟支小目标检测能力也较差。孙冰等^[13]提出一种目标关联的吸烟检测方法,采用 YOLOv7 算法先分别检测和定位香烟和场地人员,再通过计算人和烟支检测框中心点距离是否小于设定阈值来判断是否存在吸烟行为。

本文提出了一种基于人体关键点和改进的 YOLOv7 深度学习吸烟行为检测方法,首先采用自底向上的 OpenPose 算法检测人体关键点,在制定手-鼻-脖子距离比例的同时,引入手-肘-肩的角度范围条件来进一步提升对吸烟姿态判断的准确性,然后再结合改进后的 YOLOv7 算法对小目标的较强识别能力来检测烟支,以实现最终更为准确的吸烟行为检测。

1 算法设计

本文方法首先通过 OpenPose 获取每帧图像中的人体关键点及其坐标信息,计算出手部-鼻子和鼻子-脖子之间距离的比值,手-肘-肩之间的角度,通过距离比和角度是否满足预设条件,来对吸烟姿态动作进行判别;同时利用改进

的 YOLOv7 算法对图像做进一步的烟支检测;若同时满足以上两个条件则判断目标人物存在吸烟行为。本文提出的方法整体的实现流程设计如图 1 所示。

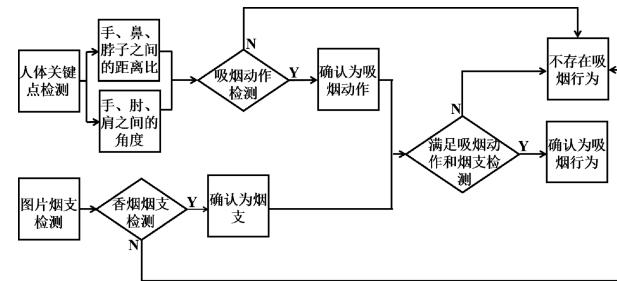


图 1 吸烟检测流程图

Fig. 1 Smoking detection process diagram

1.1 基于 OpenPose 算法的人体关键点检测

OpenPose 算法是一种自底向上的方法,可以不用定位人体边界框,就准确地实现对人的肢体、面部和手指等各个关键点检测,将人体模型^[14]简化为由多个骨骼关键点拼接成的人体骨架,并根据其运动变化进行相应的姿态估计。经 OpenPose 算法检测后可以得到 25 个人体骨骼关键点,这些关键点与人体的对应部位如图 2 所示。

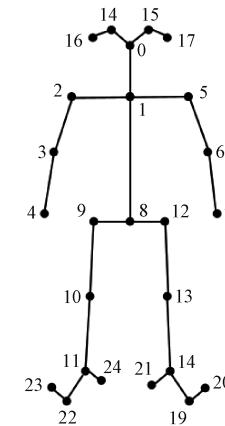


图 2 人体骨架图

Fig. 2 Human skeleton diagram

OpenPose 算法网络的具体结构流程如图 3 所示,首先使用 VGG-19 网络对图像进行卷积操作,提取特征图集 F。然后,将 F 作为第一阶段输入,分为高斯热度图检测和亲和区域两个分支。高斯热度图确定关键节点位置,亲和区域用 PAF(part affinity fields)方法预测节点连接方向^[15]。随后,在 CNN 卷积连续作用下,输出节点位置和连接方向预测结果,并与 F 特征图融合,不断迭代,持续提高精度直至得到人体骨架。图中,用 S 表示每个关节点的热度图, L 表示亲和度向量场, t 表示迭代次数, ρ 和 Φ 分别代表第 1 个分支和第 2 个分支的 CNN 网络结构。

1.2 基于关键点的吸烟姿态分析

通过对日常中的吸烟动作仔细观察,可以发现一定的

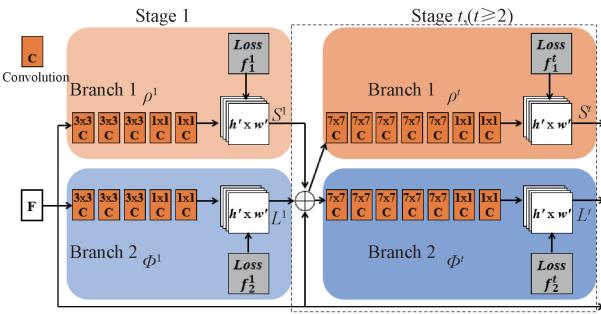


图 3 OpenPose 算法网络结构图

Fig. 3 Network structure of OpenPose algorithm

规律性。人在抽烟的过程中,嘴巴接触到烟嘴前,手腕会慢慢向鼻子靠近,手腕与鼻子的距离会越来越小。在开始抽烟时,手部会在嘴部区域停留一段时间,此时手腕与鼻子的距离基本保持不变。在抽烟结束烟嘴离开嘴巴时,手腕与鼻子的距离会越来越大。接着,不断重复上述规律动作直到停止抽烟,由此,可以将吸烟过程中手腕与鼻子的距离变化关系作为吸烟行为的特征之一。

通过 OpenPose 获取人体关键点的位置坐标信息,计算手腕到鼻子的距离,对应到图 2 中就是,计算右(左)手腕到鼻子的距离即计算关键点 4(7)到关键点 0 之间的距离。因为每个人在抽烟时的习惯用手不同,因此并行计算右(左)手与鼻子之间的距离,其公式为:

$$d = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2} \quad (1)$$

式中: (x_1, y_1) 表示手腕处的位置坐标, (x_2, y_2) 表示鼻子处的位置坐标。

现实中,每个人的身高及肢体长度不同,在监控下,人物距离摄像头的远近不同,大小比例也会不同,所以,如果直接采用手腕到鼻子关键点距离作为判断,则会不准确,且泛化能力较弱。因此,采用一种比值的方法来对数据做进一步处理,对于同一个人,鼻子到脖子的距离是恒定的,手腕到鼻子的距离是变化的,因此选取手腕-鼻子和鼻子-脖子的距离比值会相对稳定:

$$r = \frac{d}{l} \quad (2)$$

设脖子的关键点坐标为 (x_3, y_3) 则 l 为:

$$l = \sqrt{(y_3 - y_2)^2 + (x_3 - x_2)^2} \quad (3)$$

式中: d 表示手腕到鼻子关键点的距离, l 表示鼻子到脖子关键点的距离, r 表示手腕-鼻子关键点距离与鼻子-脖子关键点距离的比值。

事实上,吸烟过程也存在着一定的角度变化规律,吸烟前,手腕会慢慢向鼻子靠近,大臂和小臂的夹角会越来越小。抽烟时,手会在嘴部停留一段时间,此时大臂和小臂的夹角会基本保持不变。抽烟后,大臂和小臂的夹角会越来越大。大臂与小臂夹角的示意图如图 4 所示。其中,大臂与小臂的夹角即手 C、肘 A、肩膀 B 之间的夹角。对应到

图 2 中,右大臂与小臂夹角为:关键点 4 右手腕和关键点 3 右肘的连线与关键点 2 右肩膀和关键点 3 右肘的连线之间的夹角。左大臂与小臂夹角为:关键点 7 左手腕和关键点 6 左肘的连线与关键点 5 左肩和关键点 6 的连线之间的角度。由于不同抽烟者的惯用手不同,因此并行判断右(左)大臂与右(左)小臂之间的位置关系。其大臂与小臂的夹角 β 的计算公式为:

$$\beta = \arccos \frac{b^2 + c^2 - a^2}{2bc} \quad (4)$$

式中: a 表示手腕关键点与肩膀关键点之间的距离, b 表示手腕关键点与肘部关键点之间的距离, c 表示肘部关键点与肩膀关键点之间的距离。

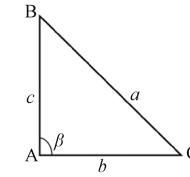


图 4 角度示意图

Fig. 4 Schematic diagram of angle

1.3 YOLOv7 模型结构

YOLOv7 是 YOLO(you only look once)系列算法的一种版本,它使用扩展高效聚合网络(ELAN)、MP 模块、改进的 SPP、PANet 结构和重参数模块 RepConv 结构等技术提升性能。通过锚框和非极大值抑制,实现高效的多尺度目标检测,并解析和筛选结果。其网络结构如图 5 所示。

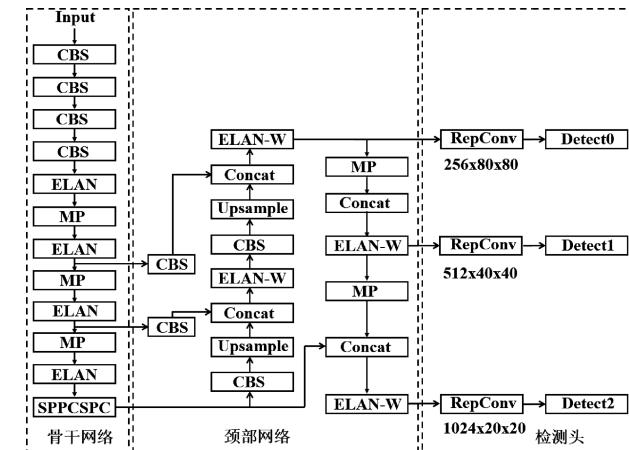


图 5 YOLOv7 网络结构图

Fig. 5 YOLOv7 network structure diagram

1.4 改进的 YOLOv7 烟支检测算法

根据烟支是小目标的特点,在 YOLOv7 网络部分添加全局注意力机制(global attention mechanism, GAM)模块,增强特征中语义和位置信息,并强化特征融合方面的性能;通过转置卷积替换原有的最近邻插值进行上采样,减少烟支这种小目标在上采样过程中的信息损失;最后,引入 MPDIoU 损失函数,解决现有损失函数局限性,增强回归

结果准确性提升整体网络检测精度。改进后的模型结构如图 6 所示。

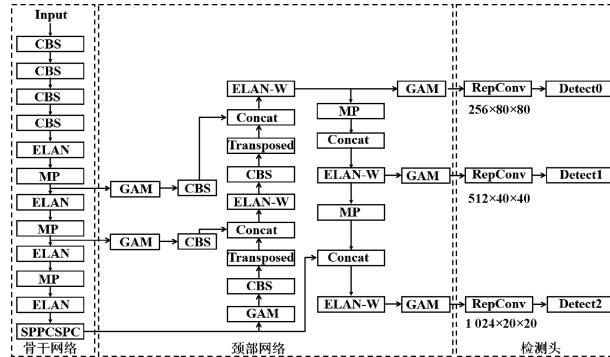


图 6 改进后的 YOLOv7 网络结构图

Fig. 6 Improved YOLOv7 network architecture diagram

1) 加入全局注意力机制

在目标检测任务中,以往的方法忽略了模型特征提取过程中的跨维交互能力,为减少信息分散和放大全局交互表示来增强网络的特征融合能力,引入全局注意力机制(global attention mechanism, GAM)^[16],GAM 注意力机制由通道 M_c 和空间 M_s 子注意力模块构成,该模型可以更全面地利用图像信息,从而实现在复杂的场景和光线等情况下对烟支的准确识别。GAM 注意力机制原理如图 7 所示,给定输入特征映射 F_1 ,中间状态 F_2 和输出 F_3 ,其过程表述如式(5)和(6)所示,其中 \otimes 表示级联。

$$F_2 = M_c(F_1) \otimes F_1 \quad (5)$$

$$F_3 = M_s(F_2) \otimes F_2 \quad (6)$$

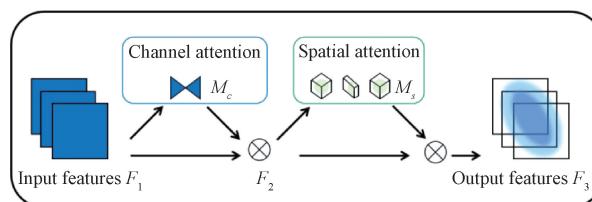


图 7 GAM 注意力机制原理图

Fig. 7 Schematic diagram of GAM attention mechanism

2) 采用转置卷积上采样

YOLOv7 中采用最近邻插值进行上采样,即缺少的像素直接用最邻近的颜色生成,但是对于小目标多,像素数量相对较小的情况,会混合周围的背景信息,影响网络对小目标检测的精度。为了更好地进行上采样,采用转置卷积模块^[17]进行上采样操作。其原理如图 8 所示,取 2 作为卷积核宽度、1 为步长示例,将卷积核张量与输入张量逐个元素相乘,放在相应的位置,以增大图像的高与宽来完成上采样操作。转置卷积可以有效缓解上采样过程中产生的失真问题,在一定程度上保留了原始图像的细节和特征。

3) 引入 MPDIOU 损失

在 YOLOv7 网络中,默认采用的是 CIOU 作为其损失

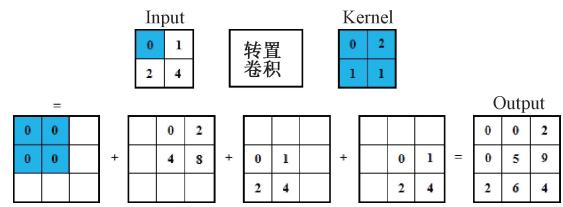


图 8 转置卷积原理示意图

Fig. 8 Schematic diagram of transpose convolution principle

函数,但是当预测框与真实框的长宽比一致而长宽值不相等时,回归损失将失去作用。为此,本文使用最小点距离基交并比(minimum point distance based IoU, MPDIOU)损失函数^[18]对边界框的回归损失重新定义,它是一种基于最小点距离的边界框相似度比较度量,可以最小化总体损失,有效的处理小目标问题,MPDIOU 损失如式(7)所示,原理如图 9 所示。

$$L_{MPDIOU} = 1 - MPDIOU = 1 - IOU + \frac{d_1^2}{h^2 + w^2} + \frac{d_2^2}{h^2 + w^2} \quad (7)$$

式中: d_1^2 和 d_2^2 分别表示预测边界框 P 与真值边界框 G 左上角点和右下角点坐标差值的平方之和,即左上角点和右下角点的距离的平方, h 和 w 表示输入图像的宽和高。

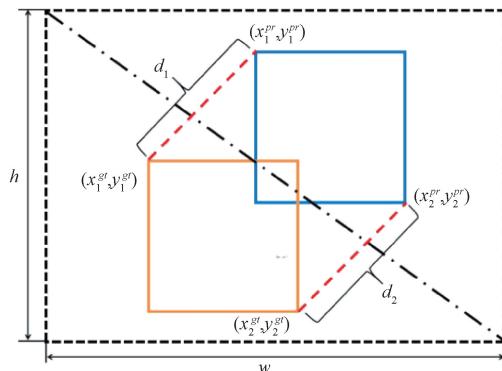


图 9 MPDIOU 损失计算原理图

Fig. 9 Schematic diagram of MPDIOU loss calculation

4) 消融效果对比

本文通过网络和自己拍摄获取共 5 000 张烟支图片数据集,按 8:2 分为训练集和验证集,设置 epochs 为 200, batch-size 为 16, 图片输入大小为 640×640 , 初始学习率为 0.01, 对 YOLOv7 网络模型进行烟支目标检测训练,为验证不同注意力模块对网络性能造成的影响效果,在原网络中相同位置单独引入 GAM、通道空间注意力机制(convolutional block attention module, CBAM)和坐标 CA 注意力机制(coordinate attention, CA),做消融实验,对比结果见表 1。可见,虽然 3 种注意力模块在效果上都有提升,但 GAM 注意力机制,相较于原 YOLOv7 网络模型,精确率(P)、召回率(R)和平均精确率(M_{AP})分别提高了

4.42%、5.01% 和 2.98%，性能提升更为明显。因此，选择 GAM 注意力模块加入 YOLOv7 颈部网络以提升性能。

表 1 注意力机制模块消融效果对比图

Table 1 Comparison of ablation effects of attention mechanism modules

模型	P / %	R / %	M_{AP} / %
YOLOv7	88.20	87.46	92.71
YOLOv7+CBAM	91.86	90.71	94.21
YOLOv7+CA	90.40	90.41	93.12
YOLOv7+GAM	92.62	92.47	95.69

为进一步量化说明不同改进方法对网络性能的影响，对单个改进和整合的所有改进分别进行消融效果对比，表 2 展现了对比实验结果。

表 2 不同改进方法对比

Table 2 Comparison of different improvement methods

模型	P	R	M_{AP}
YOLOv7	88.20	87.46	92.71
YOLOv7+GAM	92.62	92.47	95.69
YOLOv7+转置卷积	91.99	90.38	94.64
YOLOv7+MDIOU	90.31	92.62	95.27
整合所有改进	94.12	94.87	96.85

从表可见，整合之后的改进模型在各个指标方面表现的更为突出，图 10 显示了 YOLOv7 改进前后 M_{AP} 对比曲线图，证明了改进的有效性。

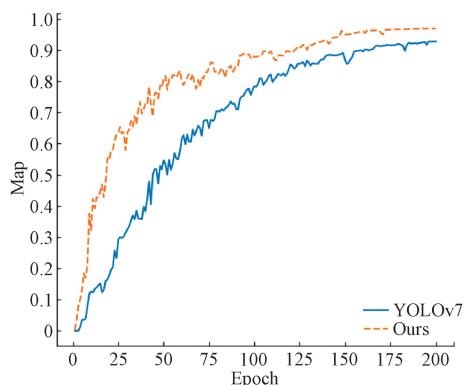


图 10 平均精度对比曲线图

Fig. 10 Comparison diagram of mean average precision

2 实验设计与分析

2.1 实验环境与数据集

实验采用的环境配置如表 3 所示。

由于公开的吸烟动作行为视频数据集比较少，因此，在室内室外多种场景下录制 200 个视频，每个视频时常约

表 3 实验环境配置

Table 3 Experimental environment configuration

参数	配置
CPU	Intel(R) Core(TM) i7-12700
GPU	NVIDIA GeForce RTX 3090
系统环境	Windows 10
语言	Python 3.7
学习框架	Pytorch 1.13.1
加速环境	CUDA 11.6

1 min，其中包含吸烟动作为主和一些非吸烟动作。另外，对 200 个样本视频进行抽帧并随机抽选烟支数据集中图像共计 1 500 张，构成 YOLOv7 烟支检测的测试集。

本文所涉及到的评价指标有，准确率 A 、精确率 P 、召回率 R 、单类别精确率 AP 和平均精确率 M_{AP} 。如公式(8)~(12)所示。

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$P = \frac{TP}{TP + FP} \quad (9)$$

$$R = \frac{TP}{TP + FN} \quad (10)$$

$$A_p = \int_0^1 P(r) dr \quad (11)$$

$$M_{AP} = \frac{\sum_{i=1}^c A_{pi}}{C} \quad (12)$$

式中： TP 表示检测到且确实存在吸烟行为的样本数量， TN 表示检测到不存在且确实不存在吸烟行为的样本数量， FP 表示事实上不存在但被检测为吸烟行为的样本数量， FN 表示事实上存在但被检测为不存在吸烟行为的样本数量， C 表示类别数。

2.2 吸烟姿态准则制定与结果分析

为了使吸烟姿态检测判断结果更为准确，本文将距离比值 r 和角度 β 同时考虑在内。通过对数据的大量分析与观察，发现当人物处于吸烟状态时手腕-鼻子关键点距离与鼻子-脖子关键点距离的比值 r 会小于一个阈值且大臂与小臂的夹角 β 也会小于一个阈值，而未吸烟状态则相反，设比例阈值为 r_0 ，角度阈值为 β_0 ，由此设置吸烟动作判定条件，如果距离比值 $r < r_0$ ，且大臂与小臂的夹角 $\beta < \beta_0$ ，则判定存在吸烟动作，否则判定为不存在吸烟动作。

对检测视频进行逐帧处理，从中选取关键帧进行吸烟动作过程展示，如图 11 所示。图 11(a)对应的是吸烟前手部拿起烟到嘴部的过程，图 11(c)对应吸完烟手部放下的过程。图 11(b)检测为吸烟姿态，其他均是非吸烟姿态。

对实验数据结果做进一步处理，用 x 轴数值表示视频中对应的帧数， y 轴数值表示关键点距离的比值 r ，绘制实时变化曲线图，对视频中人物的吸烟动作进行分析。如图 12 所示。

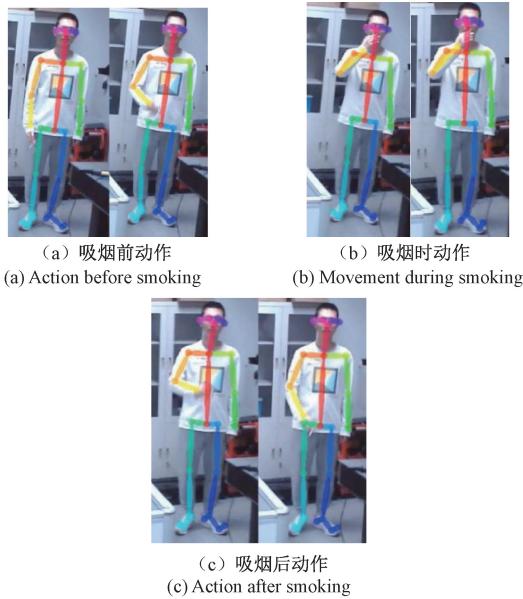
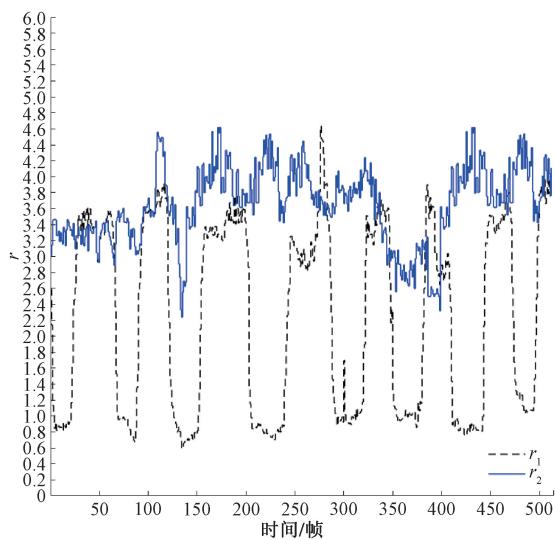


图 11 吸烟动作帧示例

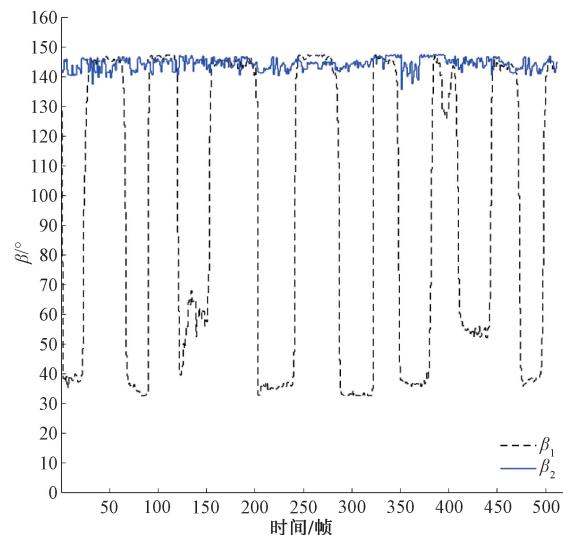
Fig. 11 Example of smoking action frame

图 12 比值 r 随帧数变化曲线图Fig. 12 Curve of ratio r as a function of frame number

右手腕到鼻子的距离比值和左手腕到鼻子的比值分别用 r_1 和 r_2 两条不同的曲线表示。为了验证采用比值方法, 可以有效解决肢体长短与距摄像头远近差异带来的不稳定的影响。样本视频中, 实验对象分别在距离摄像头远近两个不同的位置做相关吸烟动作来进行检测。为了清晰的分辨出吸烟时的曲线变化范围, 准确地确定的阈值, 在每一个位置距离都会做出 3 次吸烟动作。通过分析曲线可知, 当视频中的人物处于吸烟姿态时, 右手比值 r_1 稳定小于 1.2, 处于非吸烟姿态时, 如放下右手, 比值 r_1 普遍大于 3, 与此同时左手处于闲置状态, 手部的随机动作呈现如图中 r_2 曲线变化情况。由此可得处于吸烟状态时, 无论人物距离摄像头远

近, 肢体长度是否相同, 比值始终比较稳定小于 1.2。

同理, 用 x 轴数值表示视频中对应的帧数, y 轴数值表示大小臂的夹角值, 做角度随帧数的实时变化曲线, 如图 13 所示。

图 13 角度 β 随帧数变化曲线图Fig. 13 Curve of angle β with frame rate variation

其中 β_1 表示有大小臂的夹角变化曲线, β_2 表示左大小臂的夹角变化曲线。通过分析观察可得, 当人处于吸烟姿态, 夹角稳定小于 60° , 而当处于非吸烟姿态时, 夹角普遍大于 140° , 与此同时左手处于闲置状态, 角度变化如曲线 β_2 所示。

采用准确率 A 和召回率 R 两个指标来对检测效果进行评估。通过实验发现, 阈值的设定值不同, 其对应效果也会不同。如表 4 可见, 当阈值 r_0 或 β_0 设置过低时, 会漏掉一些吸烟姿态的检测, 当 r_0 设置为 1.2, β_0 设置为 60° 时, 其识别准确率最高, 因此吸烟姿态的判断准则条件设置为: $r < 1.2$ 且 $\beta < 60^\circ$ 。

表 4 不同参数设定效果对比

Table 4 Comparison of effects of different parameter settings

编号	r_0	$\beta_0 / {}^\circ$	$A / \%$	$R / \%$
1	1.2	40	88.14	82.23
2	1.0	60	87.36	81.26
3	1.2	60	92.60	90.94

为了进一步说明本文吸烟姿态检测方法的有效性, 针对相同的样本数据集与目前存在的一些主流方法进行对比, 如表 5 所示。

表中第 1 种方法采用 OpenPose 算法计算手到鼻子的距离, 设定阈值 150 mm, 第 2 种方法 AlphaPose 的距离比例采用与本文相同的关键点部位的比值计算和阈值。由表

表 5 不同方法结果对比

Table 5 Comparison of results from different methods

方法	A / %
距离	85.12
AlphaPose-距离比	90.53
本文吸烟姿态判定方法	92.60

可见,仅依据距离判断的准确率较差,而本文 OpenPose 综合距离比值和角度的方法准确率比第 2 种方法高 3.07%,可以更准确检测出吸烟姿态。

2.3 基于改进的 YOLOv7 烟头检测结果分析

为了更好的评估改进后的 YOLOv7 算法的性能,使用相同的实验设备配置以及实验策略,与主流的 Faster-RCNN、YOLOv5s 等 YOLO 系列目标检测算法相对比,采用 M_{AP} 和 FPS 作为衡量标准,如表 6 所示,由表可知,改进后的 YOLOv7 模型较 Faster-RCNN、YOLOv5s、YOLOv7 和 YOLOv8n 模型 M_{AP} 分别提高了 14.26%、7.43%、4.14% 和 6.01%,在速度上分别提高了 38.6.3 和 1。较其他改进后的网络模型 (YOLOv5s + 小目标检测层、YOLOv7+ Decouple Head 解耦头), M_{AP} 分别提高了 5.4% 和 3.66%,在速度上分别提高了 9 和 7。综上所述,本文提出的改进后的 YOLOv7 算法在 M_{AP} 和速度上优于上述所有算法,验证了改进方法的有效性。

表 6 不同算法对比图

Table 6 Comparison of different algorithms

模型	M_{AP} / %	FPS
Faster-RCNN	82.59	27
YOLOv5s	89.42	59
YOLOv5s+小目标检测层	91.45	56
YOLOv7	92.71	62
YOLOv7+Decouple Head	93.19	58
YOLOv8n	90.84	63
改进的 YOLOv7	96.85	65

为检验改进后的模型在实际场景下的泛化能力和效果,在测试集上分别对前后两个模型进行测试,数据对比如表 7 所示,并随机抽取一组样本进行可视化效果对比,如图 14 和 15 所示。可见,原模型在对烟支目标检测,存在漏检和置信度相对较低的情况,改进后的模型明显提高了识别精度,改善了漏检现象,更适用于现实场景检测。

表 7 YOLOv7 改进前后模型测试效果对比

Table 7 Comparison of model testing results before and after YOLOv7 improvement

模型	A	R
YOLOv7	88.79	86.34
改进的 YOLOv7	93.51	91.03



图 14 原 YOLOv7 可视化检测效果
Fig. 14 Visual detection effect of original YOLOv7



图 15 改进的 YOLOv7 可视化检测效果
Fig. 15 Improved YOLOv7 visual detection effect

2.4 实验结果分析

将视频图像中的人物是否满足设定吸烟姿态准则作为判据 1,将改进的 YOLOv7 是否检测到烟头作为判据 2,当判据 1 和判据 2 同时满足时,则判定存在吸烟行为,吸烟行为检测效果如图 16 所示。



图 16 吸烟行为检测效果图
Fig. 16 Effect of smoking behavior detection

在自建数据集上,分别将吸烟姿态识别算法、改进的 YOLOv7 烟支检测算法和本文提出的联合算法做出了实验对比,如表 8 所示。

可见,与单方面的姿态识别和烟支检测方法相比,本文提出的联合方法准确率更高,除此之外,本文联合方法不仅

表 8 不同算法效果对比

Table 8 Comparison of the effects of different algorithms

算法	A	R
烟支检测	93.51	91.03
吸烟姿态识别	92.60	90.94
本文联合算法	95.45	94.22

可以有效区分喝水、挠头、板书等与吸烟相似动作,同时,也可以大大降低单方面目标检测带来的误识别概率,优势明显,更适用于吸烟行为检测。

为了进一步验证本文方法在吸烟检测方面的优越性,在同样的数据集和实验环境下,将本文方法与其他两种被提出的吸烟行为检测方法做对比,文献[12]基于AlphaPose 和 YOLOv4 的复合吸烟检测方法,文献[13]基于目标关联的吸烟检测方法。结果如表 9,可见,本文方法准确率更高,同样优于其他两种复合判定方法,效果较好。

表 9 不同方法对比结果

Table 9 Comparison results of different methods

检测方法	准确率/%
文献[12]	93.13
文献[13]	91.75
本文方法	95.45

3 结论

针对园区安防吸烟行为检测,传统仪器设备和单目标检测方法效果不佳问题。本文提出一种联合吸烟姿态检测和烟支目标检测的深度学习吸烟行为检测方法。首先通过OpenPose 检测人体关键点做吸烟姿态检测,然后用改进的YOLOv7 算法对烟支目标进行检测,当图像中的人物满足吸烟姿态动作且同时检测到有烟支,则判断存在吸烟行为。实验表明,在自采集的数据集中,本文方法可以有效检测出吸烟行为,在园区视频安防的吸烟检测方面具有重大应用价值和现实意义。

本文方法对单人和离散人群效果较好,但对于扎堆密集人群,难免存在各种遮挡情况,不能得到理想检测效果,在未来工作中,将进一步探究和解决此类问题。

参考文献

- [1] 王辰,肖丹,池慧.《中国吸烟危害健康报告 2020》概要[J].中国循环杂志,2021,36(10):937-952.
WANG CH, XIAO D, CHI H. 2020 report on health hazards of smoking in China: An updated summary [J]. Chinese Circulation Journal, 2021, 36 (10): 937-952.
- [2] 柳浩,李淮江,房腊早.基于多传感器的火灾监测及小型

灭火系统研究[J].电子测量技术,2021,44(5):11-15.

LIU H, LI H J, FANG L Z. Research on fire monitoring and small extinguishing system based on multi-sensor[J]. Electronic Measurement Technology, 2021, 44(5): 11-15.

- [3] 王乐,张明青,陈春晓,等.人脸识别技术在园区安防中的应用研究[J].科技与创新,2022(2):103-105.

WANG L, ZHANG M Q, CHEN CH X, et al. Research on the application of facial recognition technology in park security [J]. Science and Technology & Innovation, 2022(2): 103-105.

- [4] 梁佃军.化工园区火灾场景分析与灭火救援圈的探讨[J].化纤与纺织技术,2022,51(11):65-67.

LIANG T J. Analysis of fire scenarios in chemical industrial parks and exploration of fire extinguishing and rescue circles [J]. Chemical Fiber Textile Technology, 2022, 51(11): 65-67.

- [5] 张云柯.远程智能火灾监控报警及控制系统的设计与实现[D].西安:西安建筑科技大学,2018.

ZHANG Y K. Design and research of remote intelligent fire monitoring and automatic alarm system[D]. Xi'an: Xi'an University of Architecture and Technology, 2018.

- [6] 孙英,刘乃源,余臻伟,等.基于 NFC 的可穿戴传感器中柔性/可拉伸天线的研究进展[J].仪器仪表学报,2020,41(12):122-137.

SUN Y, LIU N Y, YU ZH W, et al. Research progress of flexible/stretchable antennas in wearable sensors based on NFC[J]. Chinese Journal of Scientific Instrument, 2020, 41(12): 122-137.

- [7] 潘广贞,元琴,樊彩霞,等.基于混合高斯模型和帧差法的吸烟检测算法[J].计算机工程与设计,2015,36(5):1290-1294,1336.

PAN G ZH, YUAN Q, FAN C X, et al. Cigarette-smoke detection based on Gaussian mixture model and frame difference [J]. Computer Engineering and Design, 2015, 36(5): 1290-1294,1336.

- [8] 汪祖云,廖惠敏,张日东,等.结合烟雾多特征的出租车司机吸烟行为检测[C].第十二届全国信号和智能信息处理与应用学术会议论文集,2018:368-373.

WANG Z Y, LIAO H M, ZHANG R D, et al. Detection of smoking behavior of taxi drivers based on smoke characteristics[C]. 12th National Conference on Signal and Intelligent Information Processing and Application, 2018: 368-373.

- [9] 韩贵金属,李倩.基于 Faster R-CNN 的吸烟快速检测算法[J].西安邮电大学学报,2020,25(2):85-91.

HAN G J, LI Q. A rapid detection algorithm for smoking based on Faster R-CNN[J]. Journal of Xi'an

- University of Posts and Telecommunications, 2020, 25(2): 85-91.
- [10] 杨国亮,龚志鹏,黄聪,等.基于深度学习的吸烟行为实时检测[J].安全与环境学报,2023,23(10):3696-3705.
YANG G L, GONG ZH P, HUANG C, et al. Real-time smoking behavior detection based on deep learning[J]. Journal of Safety and Environment, 2023, 23 (10): 3696-3705.
- [11] 赵鉴福. 基于人体关键点识别的吸烟行为检测方法研究[D]. 天津:河北工业大学,2022.
ZHAO J F. Research on smoking behavior detection method based on human keypoint recognition [D]. Tianjin: Hebei University of Technology, 2022.
- [12] 徐婉晴,王保栋,黄艺美,等.基于人体骨骼关键点的吸烟行为检测算法[J].计算机应用,2021,41(12):3602-3607.
XU W Q, WANG B D, HUANG Y M, et al. Smoking behavior detection algorithm based on human skeleton key points [J]. Journal of Computer Applications, 2021, 41(12): 3602-3607.
- [13] 孙冰,李好,黄鑫凯,等.基于YOLOv7的人体关联实时吸烟目标检测方法[J].软件工程,2024,27(1):64-67,78.
SUN B, LI H, HUANG X K, et al. Real-time smoking target detection method for human body based on YOLOv7[J]. Software Engineering, 2024, 27(1): 64-67,78.
- [14] MA W Y, ZHANG D P. Human pose estimation based on multi-spectral attention and high resolution network[J]. Journal of Computer-Aided Design & Computer Graphics, 2022, 34(8): 1283-1292.
- [15] CAO ZH, SIMON T, WEI SH E, et al. OpenPose: realtime multi-person 2D pose estimation using part affinity fields[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(1): 172-186.
- [16] LIU Y CH, SHAO Z R, HOFFMANN N. Global Attention mechanism: Retain information to enhance channel-spatial interactions[J]. ArXiv preprint arXiv: 2112.05561,2021.
- [17] 刘继红. 视频监控区人员异常行为视觉检测及智能分析技术研究[D]. 重庆:重庆理工大学,2022.
LIU J H. Research on visual detection and intelligent analysis technology for abnormal behavior of personnel in video surveillance area[D]. Chongqing: Chongqing University of Technology, 2022.
- [18] MA S L, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv preprint arXiv:2307.07662, 2023.

作者简介

陈赛,硕士研究生,主要研究方向为视觉测量技术。

E-mail:1970676099@qq.com

左云波,博士研究生,副研究员,硕士生导师,主要研究方向为机器视觉检测技术。

E-mail: zuoyunbo@bistu.edu.cn