

基于香橙派的矿井电梯安全检测系统的研究与设计

赵 盛, 张智齐, 蔡昌新

长江大学电子信息与电气工程学院, 湖北 荆州

收稿日期: 2024年9月10日; 录用日期: 2024年10月9日; 发布日期: 2024年10月17日

摘 要

矿井电梯作为矿业施工现场不可或缺的垂直运输设备, 其运行的安全性直接关系到矿工的生命安全。因此, 开发一种高效、准确的矿井电梯安全检测系统显得尤为重要。基于以上要求, 提出设计并实现了一套混合检测控制系统, 将香橙派和STM32F1系列处理器结合, 利用深度学习进行实时目标检测与跟踪, 并且使用改进YOLOv4进行目标的安全帽检测, 以提升矿工进入电梯时的安全帽佩戴率, 保证矿井安全。

关键词

香橙派, STM32, 深度学习, 目标检测与跟踪, 改进YOLOv4, 矿井安全

Research and Design of Mine Elevator Safety Detection System Based on Orange Pie

Sheng Zhao, Zhiqi Zhang, Changxin Cai

College of Electronic Information and Electrical Engineering, Yangtze University, Jinzhou Hubei

Received: Sep. 10th, 2024; accepted: Oct. 9th, 2024; published: Oct. 17th, 2024

Abstract

As an indispensable vertical transportation equipment in mining construction sites, the safety of mine elevators is directly related to the life safety of miners. Therefore, developing an efficient and accurate safety detection system for mine elevators is particularly important. Based on the above requirements, a hybrid detection and control system was proposed and implemented, which combines Orange Pie and STM32F1 series processors, uses deep learning for real-time object detection and tracking, and uses improved YOLOv4 for safety helmet detection of targets to improve the safety

helmet wearing rate of miners entering elevators and ensure mine safety.

Keywords

Orange Pi, STM32, Deep Learning, Object Detection and Tracking, Improved YOLOv4, Mine Safety

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

长期以来,我国安全工作的重点都放在防范重特大事故上,对安全帽的使用重视不够,认为安全帽不会引发大事故,思想上存在麻痹松懈意识,在工作中因安全帽质量问题或使用不当造成的人员伤亡的事故时有发生[1]。由于人员的不安全行为导致的矿山事故时有发生,所以保证井下人员的安全作业是矿山企业安全管理的重要内容[2]。

伴随着工业 4.0 的到来,矿井人员的日常开采工作及监督管理也进入信息化、智能化阶段[3],但是这样的系统的智能化是指通过 WEB 浏览器可管理、查看、检索、回放任意关键场所视频图像[3][4],跟传统点对点的监控相比其视角仍然固定,无法进行实时的目标检测以及跟踪。

随着深度学习的发展及其在目标检测等图像任务上的成功应用[5],利用深度学习自动提取特征的功能,追踪技术已经发展出一种能够精确判别并具有高度泛化能力的目标特征,这一进步超越了传统的手动特征提取方法。得益于其高级的高维特征提取功能,目标追踪算法的性能也得到了显著提高。

综上所述,本文为克服视角固定,无法进行实时检测以及跟踪目标进入矿井前安全帽佩戴情况,设计了一种系统,其整合了多个组件,其一使用嵌入式部署改进 yolov4 网络以实现高效的安全帽佩戴检测功能,其二使用 OpenCV 的脸部样本深度学习文件并搭配其配套的 DNN 模块,实现改进目标跟踪算法以实现跟踪功能,将两者融合成为一种混合检测系统,实现对目标的实时检测跟踪、控制和警报为一体的矿井电梯安全检测系统。

2. 安全检测系统的总体设计

系统架构中,香橙派 5plus 充当上位机,是平台的主控模块,可以运行 Ubuntu 操作系统,基于此,本设计中用香橙派 5plus 开发平台用 python 编程语言作为系统开发程序。如图 1 所示,通过加载与其连接的广角摄像头获取到现场实时画面,经过香橙派处理后,再通过串口与下位机主控通讯,实现控制摄像头转向模块的舵机调整角度、预警模块的蜂鸣器警报以及显示模块的 OLED 屏幕显示的功能。

同时下位机主控还设计有两个电源输入口,可同时接入两个电源进行 DC 供电与 USB 供电的双供电模式,最大接入电源为 25 V 直流电源,电源输入到稳压模块后,稳压到 5 V 电压,同时给上位机以及下位机进行供电,从而保障了整个硬件系统的稳定运行。

在本文中上位机的主要功能就是实现人脸识别跟踪的主要控制以及检测人员是否佩戴安全帽。通过使用 OpenCV 加载深度神经网络(DNN)模块实现的人脸检测和追踪,香橙派能够准确地识别人脸并跟踪其移动。一旦识别到人脸,系统将计算旋转角度,并通过下位机的控制板,控制摄像头云台舵机的旋转使得单个人脸或者多个人脸识别框的几何中心一直位于画面中央,这样能够保证最大数量的人员处于画面中。同时在识别人脸的过程中,通过并行处理加载改进 yolov4 模型,对安全帽佩戴进行检测。

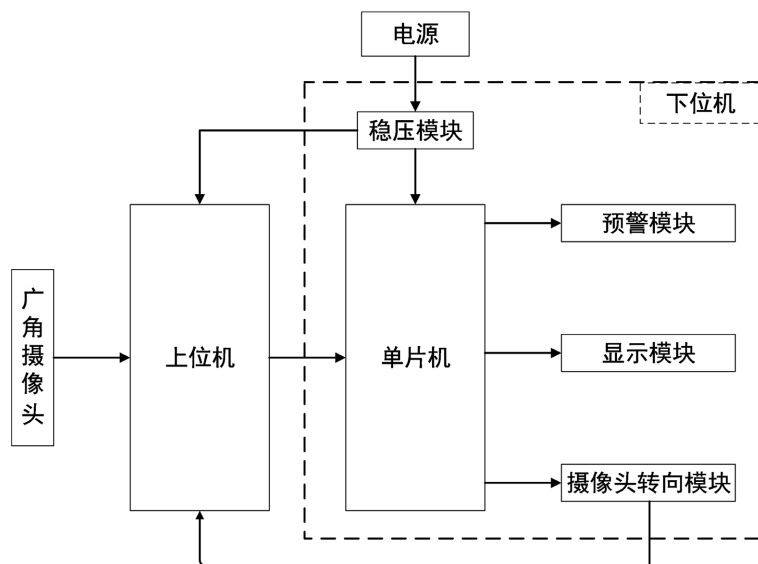


Figure 1. Hardware system architecture
图 1. 硬件系统架构

此外,下位机还负责控制蜂鸣器的报警以及 OLED 屏幕的显示,以实现全面的安全监控和报警功能。因此,该系统充分利用了香橙派的处理能力和 DNN 技术,实现了高效的人脸实时检测追踪控制。

3. 硬件系统设计

3.1. 上位机平台

如图 2 所示上位机采用的是香橙派 5plus 的 8G 版本,它是国产的一个开放源代码的硬件平台,其搭载的 Soc 处理器是基于 ARM 架构打造。其可兼容 Linux 系统绝大部分的开源软件资源,可以直接使用,同时还拥有丰富的 GPIO 接口,支持多种深度学习框架的配置与安装,极大提高了平台的扩展性[6]。

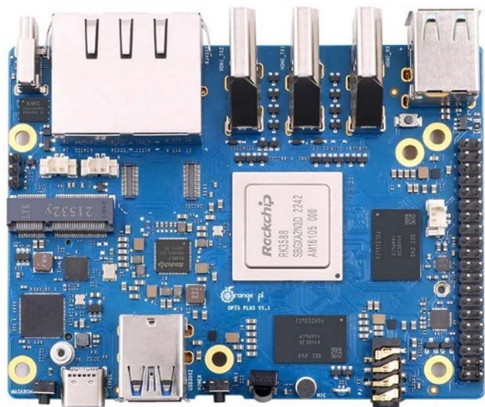


Figure 2. Orange Pi 5plus platform
图 2. 香橙派 5plus 平台

3.2. 下位机平台

下位机的主控采用了 STM32F103C8T6 型号的单片机开发板。此主控通过自主设计 PCB,将单片机最小系统开发板与多个功能模块相连接,从而响应上位机下达的一系列控制指令。

如图 3 所示, 香橙派作为上位机与自主设计的下位机融合成一个整体, 共同构成了安全检测系统, 为本设计的实现提供了有利保障。

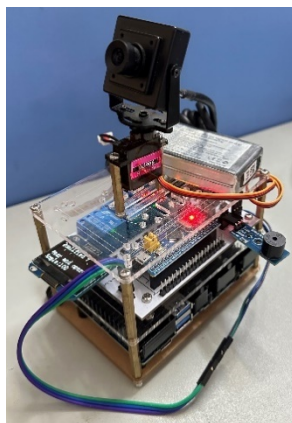


Figure 3. Schematic diagram of the overall R&D platform
图 3. 整体研发平台示意图

4. 软件系统设计

4.1. 上位机软件设计

本文使用 OpenCV 加载实时画面, 利用其内嵌的 DNN 功能加载深度学习的模型, 对图像的画面进行提取判断, 当识别到人脸后会对其进行画框, 随后根据其运动轨迹, 进行跟踪, 始终保持人脸单框或者多框几何中心位于画面中心。在识别人脸后, 还采用 OpenCV 的级联分类器, 通过特定的训练集训练的分类器判断是否佩戴好安全帽, 经过一些的数据采集分析处理后, 将以上所有信息所给定的判断再通过串口发送指令给下位机。

在本文的设计中, 上位机主要采用 Python 编程, 通过分析摄像头捕获的图像和视频来确定摄像头实时跟踪的方向。在目标跟踪方面, 采用 OpenCV 4.0 版本以后新加入的 DNN 模块, 加载 OpenCV 官方深度学习训练好的人脸模型进行目标识别跟踪。在安全帽识别方面, 则是采用直接的 Linux 版本的 pytorch 加载训练好的改进 yolov4 模型进行识别。

4.2. 下位机软件设计

下位机负责执行所有上位机传达的指令, 类似于机器人的开发控制板, 需要提前写好底层代码。在本文中, 下位机采用 atio 函数, 用于将字符串转换成整数, 配合指针运算, 可将任何角度传入舵机控制函数中去, 进行 0~180 度全方位的角度调整。在整个下位机底层代码设计中, 为了准确无误地接收来自上位机的报文, 在接收时以特殊字符“@”为接收的起始点, 直到最后一个字符串“\r\n”的接收完毕, 从而关闭串口准备下一段信号的接收。

5. 算法设计

5.1. 图像获取与分析

在图像采集阶段, 主要通过接入的摄像头进行操作。由于是混合识别, 需要将获取的图像进行分流, 一部分用于人脸识别跟踪, 一部分用于安全帽的识别。为了拥有更好的香橙派对于目标识别跟踪的观测体验, 本文使用网线连接香橙派并且调用后台窗口, 实时观看系统对于目标识别跟踪的情况。

5.2. 安全帽识别

YOLO (You Only Look Once)系列算法常常被用于目标检测中去,是目前已知主流算法之一,一般来说由主干网络(Backbone)、颈部(Neck)也就是特征融合部分和头部(Head)构成。在本文的设计中,为了准确识别安全帽,同时考虑到每个人带上安全帽的颜色以及形状也大不相同,为了提高识别的泛化性,为此采用改进 YOLOv4 对施工人员进行安全帽佩戴检测。

YOLOv4 [7]算法是 YOLO 系列第 4 代算法 YOLOv3 [8]的改进版本。YOLOv4 算法在检测精度和检测速度方面相对 YOLOv3 都具有显著优势,这主要归功于其在多个方面的优化和增强。YOLOv4 算法的网络结构如图 4 所示。

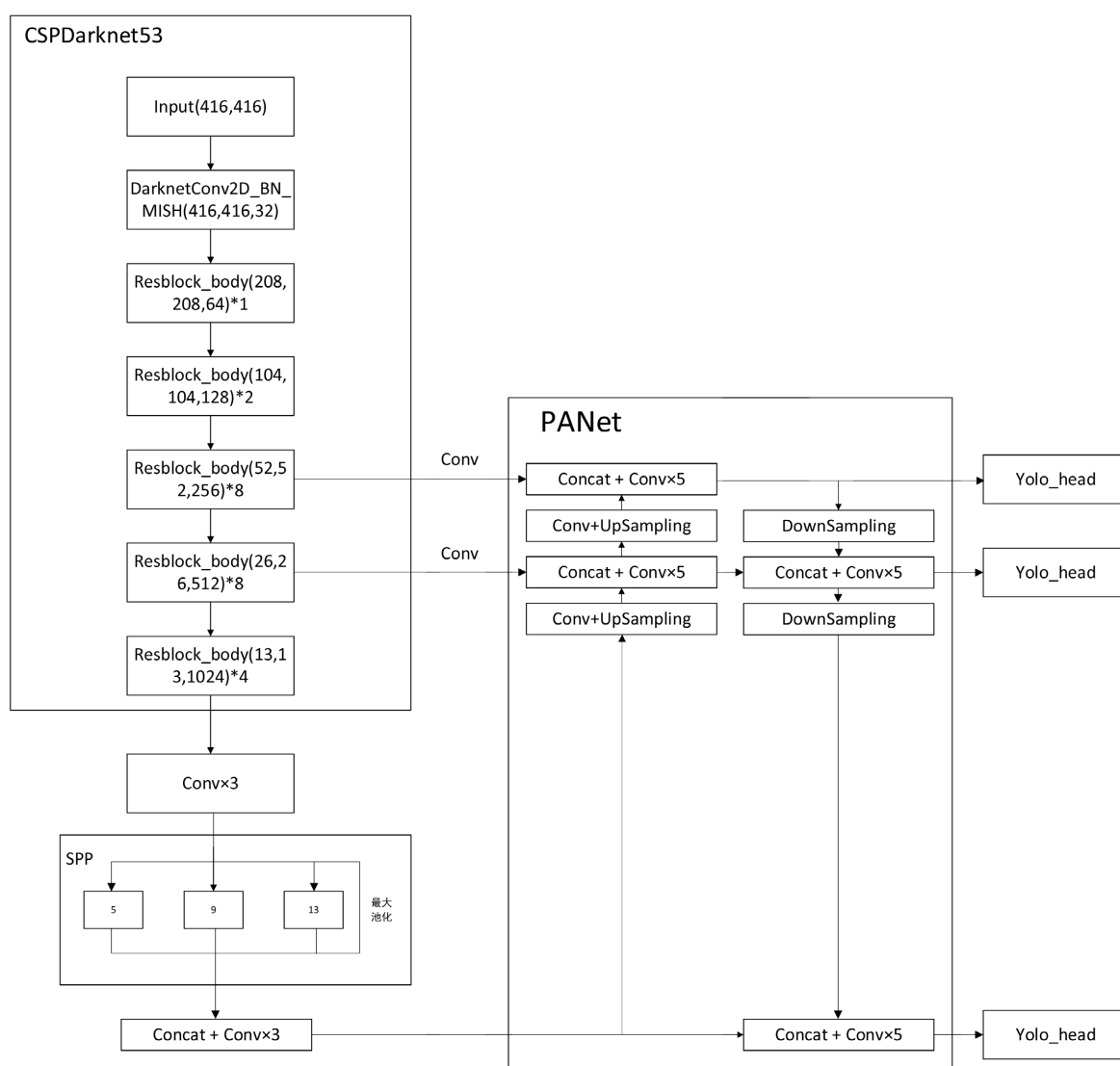


Figure 4. YOLOv4 algorithm structure diagram
图 4. YOLOv4 算法结构图

5.2.1. 对注意力模块的改进

目前几大常用的主流注意力机制有 SE 注意力机制、MSCA 注意力机制以及 CBAM 注意力机制等。

本文检测的目标为移动目标,需要动态的调整每个位置的重要性,捕捉空间层次的特征信息,又因为安全帽与人脸共同组成了“hat”的标签,人物在移动时还得动态的调整每个通道的权重,增强有用特征的表达,抑制无用特征的影响,针对以上需求预备采用由 Woo 等提出的 CBAM (Convolutional Block Attention Module) 注意力机制[9]。但是,在目标距离摄像头较远时也就是画面中检测目标较小时,画面过于杂乱,导致容易检测不到或误判[10]。为此,借鉴 yolov4 中的 SPP 模块中多特征融合的思维[11],将其加入到 CBAM 注意力模块中。这样通过不同尺寸的卷积核,可以捕捉到不同尺度的特征,从而对输入数据的多尺度信息进行建模,并且不同尺度的特征包含了更多的信息,融合后的特征表示得更加丰富和细化,提升了模型对复杂场景的理解能力[12]。最后,再加入维度注意力模块,这样在多特征融合和前 CBAM 注意力机制的基础上,进一步对特征图进行细化处理。基于以上的改进方法,得到一种全新的注意力模块 MCBAMD (Multi-scale convolution for CBAM with Dimensions attention), 其结构如图 5 所示。

首先,输入的特征图向量 x 进入 MCBAMD 注意力模块中,先通过多特征融合模块[13]得到 F_{merged} , 其结构如图 6 所示,具体公式如式(1)所示:

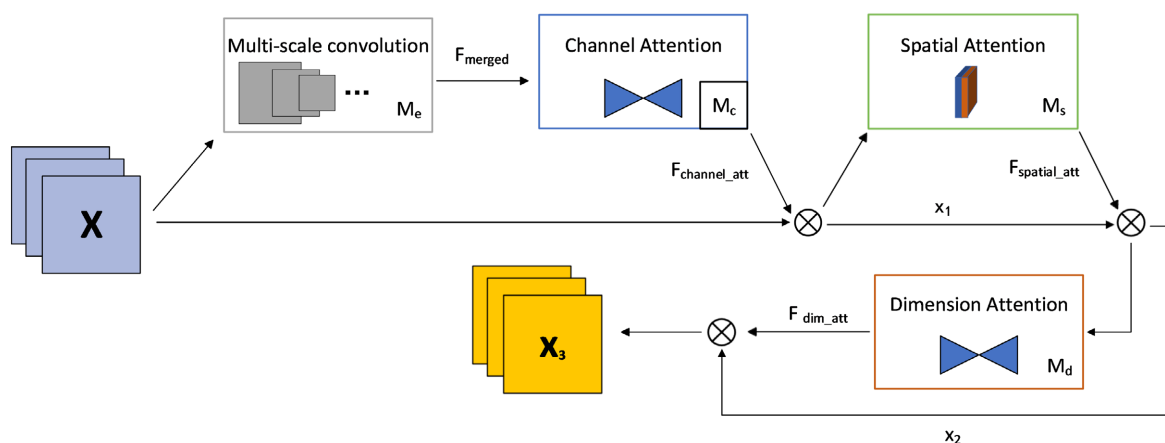


Figure 5. MCBAMD attention structure

图 5. MCBAMD 注意力结构

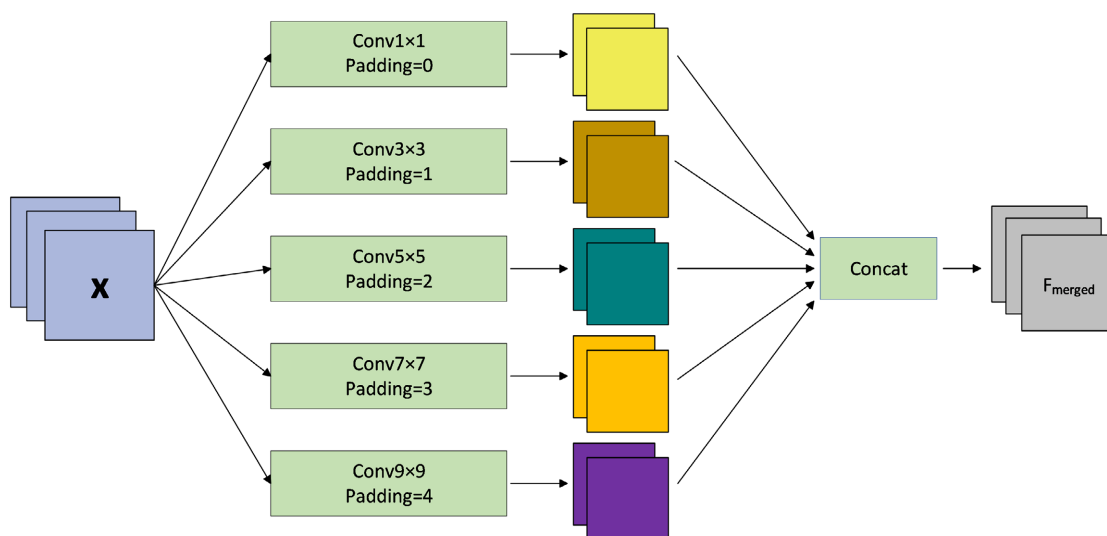


Figure 6. Multi-feature fusion module network structure

图 6. 多特征融合模块网络结构

$$F_{merged} = \text{concat}(conv_1(x), conv_3(x), conv_5(x), conv_7(x), conv_9(x)) \quad (1)$$

(1)式中, $conv_k(x)$ 表示卷积核大小为 k ($k = 1, 3, 5, 7, 9$) 的卷积操作, 通过拼接这些卷积层, 最终输出 F_{merged} 的特征图。

MCBAMD 注意力机制是对 CBAM 注意力机制的一种改进, 其通道注意力和空间注意力模块与原 CBAM 中的通道注意力以及空间注意力模块相同, 这里就不展开赘述了, 下面介绍维度注意力模块。

维度注意力模块[14]采用了一种类似于 MLP 的结构, 通过两个 1×1 的卷积层一个 ReLU 激活函数对输入特征进行加权的调整。其结构如图 7 所示。

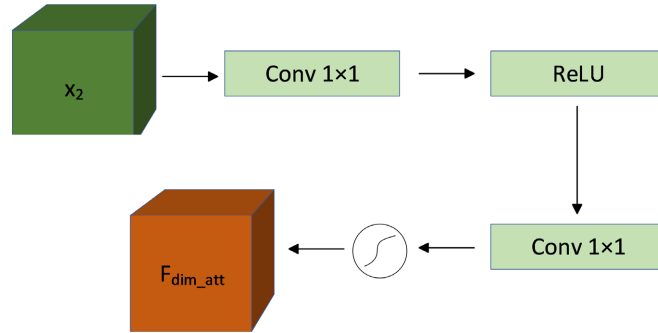


Figure 7. Dimensional attention module structure
图 7. 维度注意力模块结构

x_2 为输入特征, 设 W_{d1} 、 W_{d2} 分别为两个卷积层的权重矩阵, δ 为 Sigmoid 输入函数, 那么其公式如式(2)所示:

$$F_{dim_att} = x_2 \odot \delta(W_{d2} * \text{ReLU}(W_{d1} * x_2)) \quad (2)$$

这样, 维度注意力模块就能有效地增强了输入特征图的表达能力, 使得模型能够更好地关注重要特征, 提高整体识别性能。

5.2.2. 对 YOLOV4 算法颈部网络的改进

在安全帽的检测中, 要检测工人是否佩戴安全帽往往是重中之重, 一旦事故发生, 只要头部不受到建筑物压迫, 或许还能有一线生机, 检测画面中的背景环境噪音是不必要的特征。检测时的环境较为复杂, 使得网络难以注意到重要信息, 给模型带来性能影响。为了解决上述问题, 本文提出了一种对网络颈部的改进方法, 如图 8 所示, 即在上采样的操作之后添加 MCBAMD 注意力机制(Multi-scale convolution for CBAM with Dimensions attention)。

由于 PANet 网络中[15], 上采样的过程会对图像的质量造成影响, 特别的当物体离摄像头较远且在画面中显得较小时, 脸部容易在画面中显黑, 使得人脸与安全帽组合形成的特征容易被误判, 对此 MCBAMD 注意力机制中的多特征融合有利于补做不同尺度的图像信息, 特别是比较小的特征图, 通道注意力模块以及空间注意力模块的组合使得模型能够更有效的利用关键特征以及空间维度上更加精准, 最后加入维度注意力模块, 强调了重要维度特征, 进一步提高了特征表示的有效性, 提高模型在复杂环境中对安全帽是否佩戴的检测精度。

5.2.3. 实验结果与分析

(1) 实验环境与数据

本次实验是在自筹组装的主机上搭建环境并进行训练, 硬件配置为 CPU (i9-13900k)、GPU (Nvidia

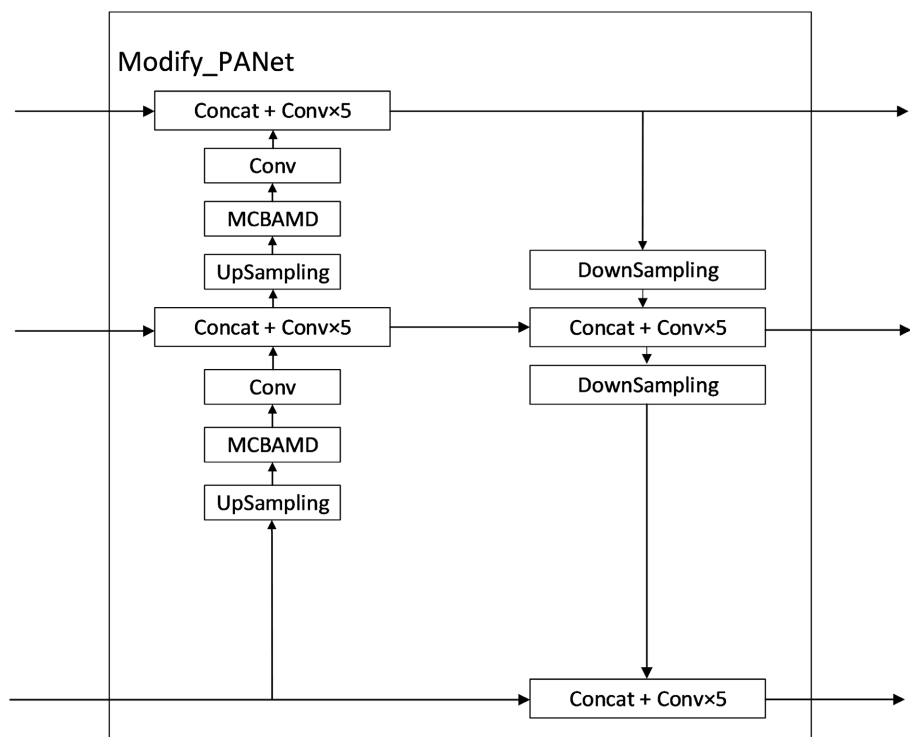


Figure 8. Improved PANet network
图 8. 改进 PANet 网络

RTX 4090D)、主板(Asus Z690-A); 操作系统为 Windows11 专业版; 软件配置为 Anaconda、Visual Studio Code; 编程语言为 python, 深度学习框架为 pytorch。

实验所用的安全帽数据集, 为自建安全帽佩戴检测数据集, 是由编写的爬虫程序在谷歌和百度上收集的施工现场图片, 然后再与 SCUT-HEAD 中的一部分图片融合而成。所有图像采用 labeling 软件标注类别并且整理为 PASCL VOC 格式, 图像分为两个类别, 第一个正样本类别, 是佩戴安全帽标记为“hat”, 第二个负样本类别, 是未佩戴安全帽标记为“person”。该安全帽数据集一共有 7000 张图片, 并且按照 9:1 划分训练集与测试集。

(2) 网络训练

本文采用安全帽数据集进行网络训练, 训练时的输入图片设置为 416×416 , 优化器为 Adam, 下降方式为 step。由于本文中改进网络的主体结构没有变, 深度学习中的迁移学习的思想仍然适用于本文。由此, 网络训练分为两个阶段, 第一阶段为预训练阶段也称粗训练阶段, 采用 yolov4_voc_weights 作为预训练权重, 目的是以最快的速度来降低损失值并且提取目标特征的同时不让权值具有任意性, 第二阶段为正常训练阶段也称细训练阶段, 在第一阶段后, 调整学习率变小, 逐渐优化模型, 提升模型的检测性能和准确率。总训练周期设定为 140 个 epoch, 训练过程分为两个阶段: 第一阶段使用 20 个 epoch, batch size 设为 12, 学习率设置为 0.001; 第二阶段则延续 120 个 epoch, batch size 调整为 4, 初始学习率为 0.001, 并在训练过程中逐渐降低至最小学习率 0.000001, 以便进行更细致的调整。如图 9 所示为改进 yolov4 训练时的 loss 下降曲线, 可以看出最终曲线在 0.0057 处收敛。

(3) 模型性能评估

本文对模型性能的评估采用精确率(Precision)、召回率(Recall)、AP (Average Precision)、mAP (mean Average Precision)作为检测指标, 其公式分别如式(3)、(4)、(5)、(6)所示。另外为了体现模型的动态性能

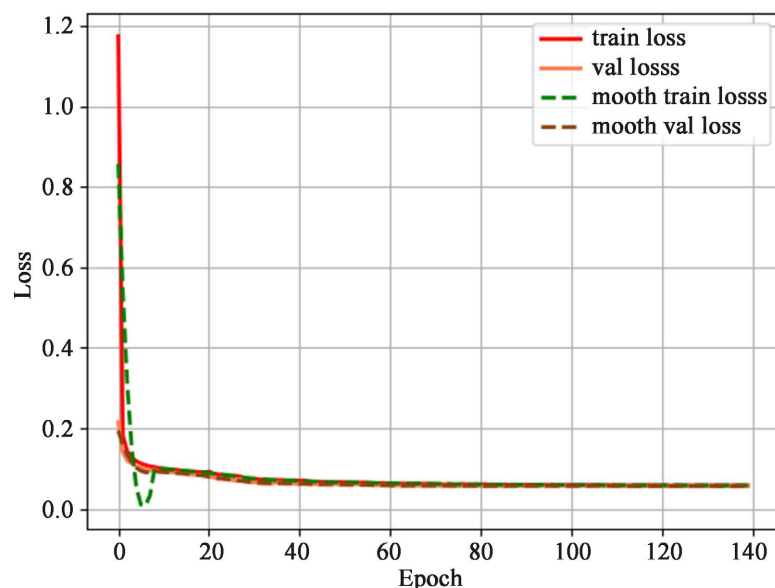


Figure 9. Loss reduction curve of improved YOLOv4

图 9. 改进 yolov4 的 loss 下降曲线

加入了处理单张图片的 FPS 值，并且检测 100 次求平均值。

$$P_{recision} = \frac{TP}{TP + FP} \quad (3)$$

$$R_{ecall} = \frac{TP}{TP + FN} \quad (4)$$

式中，TP (Ture Positive)是真正例，FP (False Positive)是假正例，FN (False Negatives)是假负例。

$$AP = \int_0^1 P(R) dR \quad (5)$$

式中： $P(R)$ 是精确率 P 在召回率 R 下的精度值。AP 值反映了目标检测算法的鲁棒性和泛化能力，AP 值越高对单个目标的检测能力越强。

$$mAP = \frac{\sum_{i=1}^C AP_i}{C} \quad (6)$$

式中， C 为类别总数； AP_i 是第 i 个类别的 AP 值；mAP 值反应了对于不同类别目标的平均覆盖率。

为了验证 yolov4 改进的有效性，将 yolov3 以及 yolov4 以相同训练方式以及训练集训练后与改进 yolov4 进行对照。如表 1 所示。

由表 1 可知，相比于原 yolov4，改进 yolov4 的精确率、AP、mAP 均略有所上升。这表明，在相同的训练集下，本文提出的改进检测方法能够识别出更多佩戴安全帽和未佩戴安全帽的目标，并且其综合性能最佳，效果更好。虽然召回率略有下降，但是为了提高精确率，模型可能会变得保守，这导致一些实际的正类样本可能被漏掉，从而召回率下降，这种精确率与召回率的权衡关系通常会导致精确率和召回率不一定同时提高，因此精确率上升但是召回率略有下降是正常现象。相比于前面提到的静态检测指标，动态检测方面的 FPS 值的提升更加巨大，相比于原 yolov4，改进 FPS 值提升达 78.6%，这说明在实时检测方面，改进 yolov4 检测速度更快。虽然改进 yolov4 的 FPS 值不如 yolov3，但是众所周知，yolov4 是 yolov3 的改进版本，yolov4 权衡了检测精度与检测速度之间的关系，本文在原 yolov4 上不仅加强了检

测精度，同时加强了检测速度，使之完全满足部署在嵌入式上面进行实时检测。

Table 1. Horizontal comparison test results
表 1. 横向对比实验结果

算法	精确率(Precision)		召回率(Recall)		AP		mAP	FPS
	hat	person	hat	person	hat	person		
yolov3	92.66%	81.58%	83.24%	69.02%	91.48%	77.55%	84.51%	107.06
yolov4	94.18%	88.43%	88.87%	84.46%	94.16%	88.84%	91.50%	32.74
改进 yolov4	95.3%	89.33%	86.40%	83.31%	94.33%	88.85%	91.59%	58.46

为了进一步验证改进 yolov4 的实时检测性能，本文取用一段工地入口处的监控视频，并分别用 yolov3、yolov4 以及改进 yolov4 对进出口处人员佩戴安全帽的情况进行检测，如图 10 所示，为检测的具体情况截图，图 10(a)~图 10(c)从左到右分别为 yolov3、yolov4、改进 yolov4 对于视频实时检测情况的截图。

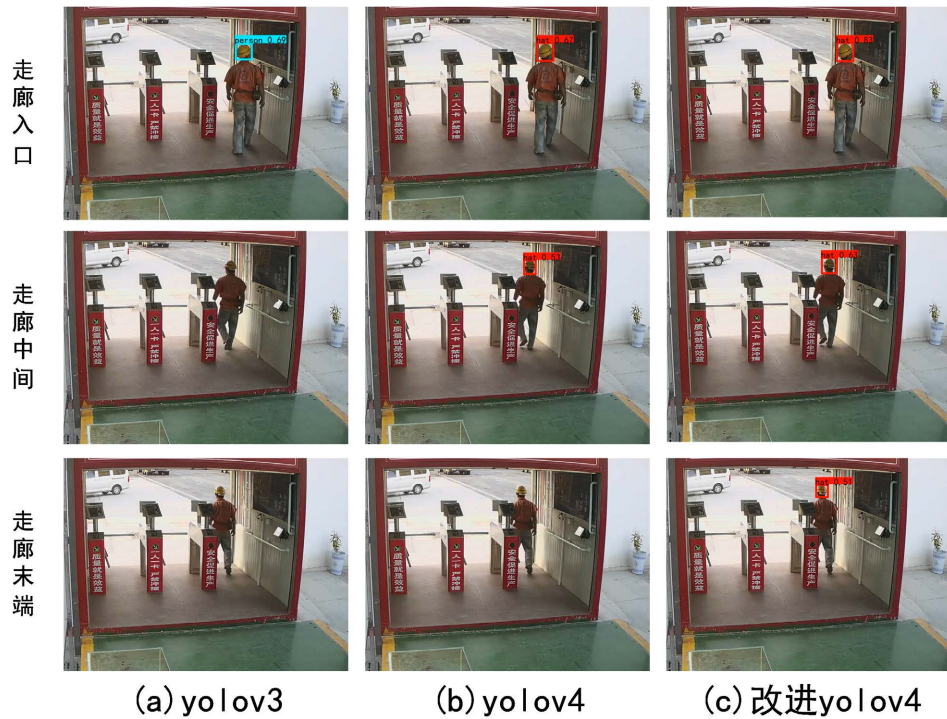


Figure 10. Screenshot of real-time monitoring video detection
图 10. 监控视频实时检测画面截图

从图 10 中可以看出，yolov3 在入口处误判，将佩戴安全帽的正样本判断为“person”的负样本，其他时间没有检测到人员佩戴，而 yolov4 没有出现入口处的误判行为，这更加验证了 yolov4 相较于 yolov3 对于检测速度与检测精度做了平衡，但是 yolov4 模型在走廊末端的最后一刹那，由于检测目标过小没有检测到。反观改进 yolov4，不仅全程全部正确识别检测到佩戴安全帽的正样本，而且在入口处预测值为 0.83，大于 yolov4 在入口处的预测值 0.67。同时，在人员行至走廊中间时，预测结果显示，改进后的 YOLOv4 模型在性能上优于原始模型。这表明本文提出的改进方法确实有效，能够更好地挖掘图像的深层特征信息，提升了检测能力。

5.3. 目标跟踪

传统的识别中, 会面临两大问题。第一, 摄像头总是被固定在一个比较固定的位置, 当人从画面中走过, 即检测结束。在实际的生产过程中, 若采取传统的方法进行检测, 工人很可能在进入电梯前或者离开电梯后不戴帽子, 只是在进入电梯时佩戴安全帽, 为此本设计采取目标的实时跟踪, 将生产安全执行到最大化, 确保工地的安全生产。第二, 在传统的识别项目中, 摄像头被固定后, 无论多大广角, 总是有一部分人不在画面中, 或者一部分人在画面中另一部分人在画面边缘不利于安全帽检测, 为此将多个人脸尽可能框定在画面以检测安全帽佩戴很有必要。所以, 本文的目标跟踪直接决定了上述安全帽的识别的最大数量的多少以及持续检测的时间。

本文针对首要问题, 采用深度神经网络(DNN)技术与连续帧同步融合策略[16]-[18]。使用 OpenCV 内嵌的 DNN 模块加载 OpenCV 官方训练好的人脸识别权重文件, 在画面中识别出人脸特征, 再运用 OpenCV 内嵌的 Rectangle 功能对目标周围描绘一个矩形边框, 以实现对该特定实体的视觉锁定, 确保其在后续动态过程中保持可见状态。紧接着, 依据已绘制的矩形区域, 可精确提取其以像素为单位的坐标值 x 与 y 。因为本文设计的云台只涉及左右移动跟踪, 因此只需要 x 的坐标即可。本文中对于目标跟踪首先是采用背景差分法算出人物移动速度, 再根据速度是否达到某一个转动角度所对应的阈值来进行动作。对于速度的运算如式(7)、式(8)所示:

$$d_k = |f_k(x) - b_k(x)| \quad (7)$$

$$v_k = \frac{d_k}{t_k} \quad (8)$$

式中: d_k 为移动的像素距离, px; $f_k(x)$ 为当前帧的坐标值, px; $b_k(x)$ 为背景帧的坐标值, px; v_k 为移动的速度, px/frame; t_k 是当前帧与背景帧之间的帧差值, frame。为了连续检测与运动, 本文将系统强制检测连续两帧的移动距离, 也就是当前帧与前一帧人物移动距离, 强制使得 $t_k = 1$ frame, 这样有利于系统进行实时检测跟踪。

面对第二个问题, 即对多个目标进行有效追踪并最大限度地将其纳入画面范围, 本文采用多目标的平均坐标值替换单个目标的坐标的方法进行处理[18], 公式如式(9)所示:

$$d_{ka} = |f_k(x_a) - b_k(x_a)| \quad (9)$$

式中: d_{ka} 为连续帧之间多个目标的平均坐标值移动的像素距离, px; $f_k(x_a)$ 为当前帧中多个目标的平均坐标值, px; $b_k(x_a)$ 为背景帧中多个目标的平均坐标值, px。这样, 即使画面中不断有多个目标连续变换位置, 系统均可精准依据人物移动速度进行目标追踪, 并确保尽可能多的目标处于画面范围内。

6. 部署实验结果

在室内将灯光全开, 对整个系统进行模拟实验验证。主要模拟其在工地中露天阳光曝晒下的工况能否稳定运行, 整体部署效果如图 11 所示。图 11(a)展示了一个目标未佩戴安全帽的情况, 图 11(b)展示了一个目标佩戴安全帽的情况, 图 11(c)展示了两个目标中一个佩戴安全帽而另一个未佩戴安全帽的情形, 图 11(d)则展示了两个目标均佩戴安全帽的场景。

从实验结果可以看出, 当画面中只有一个人物时, 摄像头, 识别框, 以及画面都在正中心, 当人物数量增至两个及以上时, 可以明显看出摄像头左转, 将两个人像识别框的平均几何中心保持在画面中心。同时采用改进 yolov4 对于安全帽佩戴情况进行检测, 极大提高了对于单个或多个人员安全帽佩戴情况检测的可靠性与准确性。

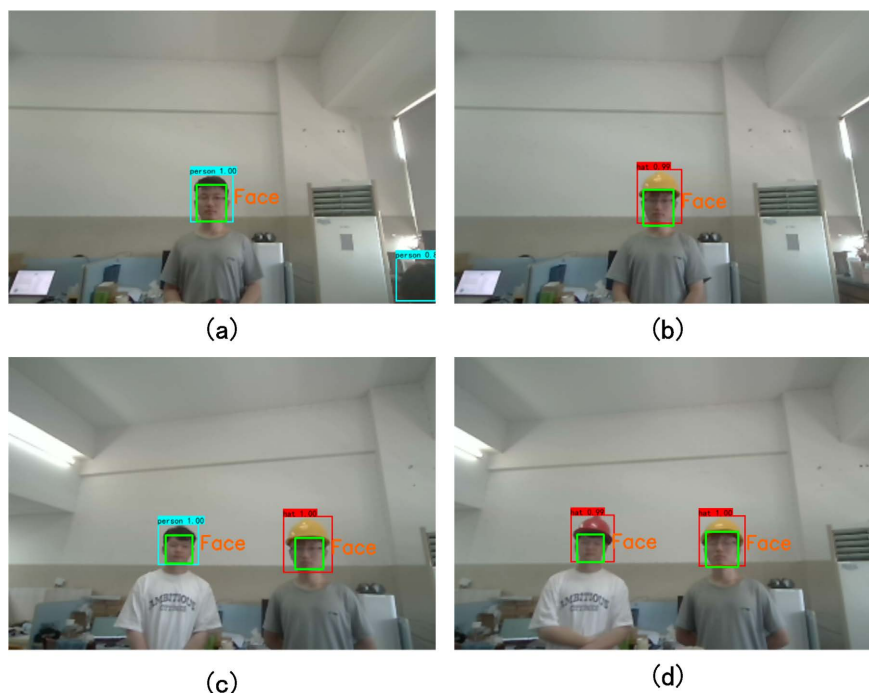


Figure 11. Real-time target tracking and detection effect diagram
图 11. 实时目标跟踪以及检测效果图

7. 总结和展望

本文深入探讨了矿井电梯安全事故频发的现状，并针对此现状进行了分析，从而确认了研究的实际可行性。本研究从安全检测的总体架构设计、硬件系统构建、软件系统设计以及核心算法研发等四大维度，全面阐述了研究的具体过程与实施方法。

首先，总体架构设计合理，先是给出了总体的硬件选取搭建思路以及软件系统的设计；其次对于硬件系统构建，详细给出上位机选用以及下位机设计思路，并且精确到具体型号；另外，对于软件的设计，针对上位机给出如何构建检测系统以及下位机对于上位机指令的接收与动作具体方法；最后，算法部分为安全帽识别算法与目标追踪算法，在安全帽识别算法部分详细给出了基于 yolov4 算法的改进方法、横向对比实验结果以及监控实时检测的对比结果，在目标追踪算法部分，给出了具体的算法公式以及实际上机部署结果。

经部署结果可以看出，安全检测系统可以在对图像进行获取并处理后，通过深度学习的方法，将目标检测算法与跟踪控制算法相结合，以及配合硬件系统，可以实现对于单一或者多个目标进行实时跟踪检测，验证了该系统可以较好地实时跟踪检测目标，该系统在工程电梯安全检测中将有很好的应用。

就整个系统的设计而言，使用了大量软硬件结合的方法，配合深度学习，可以很好地完成安全检测任务，在当今世界大力发展人工智能的趋势中，对于进一步完善智慧工地技术，防范矿井安全事故的发生，进行高度泛化目标特征的安全检测具有深远意义。

参考文献

- [1] 张苏, 傅贵, 殷文韬, 等. 基于行为安全的安全帽事故原因分析及预防[J]. 煤矿安全, 2014, 45(4): 229-232.
- [2] 张瑜, 冯仕民, 杨赛烽, 等. 矿工不安全行为影响因素本体构建与推理研究[J]. 煤矿安全, 2019, 50(5): 300-304.
- [3] 张喜萍. 基于工业以太网的全数字矿井视频监控系统[J]. 煤矿安全, 2018, 49(12): 112-114.

-
- [4] 吴文臻. 矿井主要通风机在线监控技术[J]. 煤矿安全, 2016, 47(7): 109-111.
 - [5] 李明, 鹿朋, 朱美强, 等. 基于改进 YOLO-tiny 的闸板阀开度检测[J]. 煤炭学报, 2021, 46(S2): 1180-1190.
 - [6] 陈浩安, 李晖, 黄瑞, 等. 基于树莓派 4B 的无人机动态追踪平台设计[J]. 电子测量技术, 2024, 47(6): 182-189.
 - [7] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. <https://arxiv.org/abs/2004.10934>
 - [8] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. <https://arxiv.org/abs/1804.02767>
 - [9] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. *15th European Conference*, Munich, 8-14 September 2018, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
 - [10] Eggert, C., Brehm, S., Winschel, A., Zecha, D. and Lienhart, R. (2017) A Closer Look: Small Object Detection in Faster R-CNN. *2017 IEEE International Conference on Multimedia and Expo (ICME)*, Hong Kong, 10-14 July 2017, 421-426. <https://doi.org/10.1109/icme.2017.8019550>
 - [11] Wang, X., Wang, S., Cao, J. and Wang, Y. (2020) Data-Driven Based Tiny-YOLOv3 Method for Front Vehicle Detection Inducing SPP-Net. *IEEE Access*, **8**, 110227-110236. <https://doi.org/10.1109/access.2020.3001279>
 - [12] Kim, K., Kim, P., Chung, Y. and Choi, D. (2018) Performance Enhancement of Yolov3 by Adding Prediction Layers with Spatial Pyramid Pooling for Vehicle Detection. *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Auckland, 27-30 November 2018, 1-6. <https://doi.org/10.1109/avss.2018.8639438>
 - [13] Yu, F. and Koltun, V. (2015) Multi-Scale Context Aggregation by Dilated Convolutions.
 - [14] Veit, A., Belongie, S. and Edlund, S. (2016) Residual Networks Behave like Ensembles of Relatively Shallow Networks. <https://arxiv.org/abs/1605.06431>
 - [15] Ke, S.-T., Chen, M.-H., Zheng, Z.-X., *et al.* (2022) Generate a Super-Resolution Reconstruction of OCT Retinal Images by an Adversarial Network. *Chinese Journal of Lasers*, **49**, 90-98.
 - [16] 屈晶晶, 辛云宏. 连续帧间差分与背景差分相融合的运动目标检测方法[J]. 光子学报, 2014, 43(7): 0710002.
 - [17] 陈旭, 孟朝晖. 基于深度学习的目标视频跟踪算法综述[J]. 计算机系统应用, 2019, 28(1): 1-9.
 - [18] 翟济云, 周鑫, 王从庆. 基于多特征融合的运动目标检测[J]. 电光与控制, 2017, 24(7): 23-27.