

中文引用格式:左明成,焦文华. 面向煤矿井下作业场景的安全帽佩戴识别算法[J]. 中国安全科学学报,2024,34(3):237-246.

英文引用格式:ZUO Mingcheng, JIAO Wenhua. Helmet-wearing recognition algorithm for coal mine underground operation scenarios[J]. China Safety Science Journal,2024,34(3):237-246.

面向煤矿井下作业场景的安全帽佩戴识别算法*

左明成 助理研究员, 焦文华** 研究员
(中国矿业大学 人工智能研究院, 江苏 徐州 221008)

中图分类号: X936

文献标志码: A

DOI: 10.16265/j.cnki.issn1003-3033.2024.03.1985

基金项目:国家自然科学基金资助(62303465);山东省自然科学基金资助(ZR2022LZH017);企事业单位委托项目(JAI2301)。

【摘要】 为提高煤矿井下作业场景的工人安全帽佩戴识别准确度,提出一种人体姿态分析的安全帽佩戴识别算法和机器视觉系统优化方法。首先,选择单步多框目标检测(SSD)作为多目标识别的基础模型,利用网络压缩模型(SqueezeNet)削减改进模型参数,以形成高效的识别模型,在提高对工人安全帽的识别准确度的同时,维持识别准确度与计算速度之间的平衡;然后,引入多人姿态估计算法,定位每个工人的人体关节点,判断工人的复杂行为状态;最后,基于多目标识别与多人姿态估计的融合模型,提取目标人员的上肢关节点,结合其与安全帽包围框的空间拓扑结构关系,判定安全帽的佩戴情况,选取3000条图像数据验证所提方法的有效性。结果表明:机器视觉系统优化方法能优化硬、软件配置,提高系统性能。该方法识别准确度达到91.1%,明显优于最新的安全帽佩戴识别算法,能够满足煤矿井下作业的工人安全帽佩戴识别准确度要求。

【关键词】 煤矿井下; 安全帽佩戴; 识别算法; 姿态分析; 模型优化

Helmet-wearing recognition algorithm for coal mine underground operation scenarios

ZUO Mingcheng, JIAO Wenhua

(Institute of Artificial Intelligence, China University of Mining & Technology,
Xuzhou Jiangsu 221008, China)

Abstract: To improve the accuracy of miners' helmet-wearing recognition in coal mine underground operations, a helmet-wearing recognition algorithm was proposed based on human posture analysis and machine vision system optimization methods. First, Single Shot MultiBox Detector (SSD) was used as the basic model of multi-target identification. The squeezed neural network (SqueezeNet) was used to reduce the model parameters to develop an efficient recognition model, which improved the recognition accuracy of the miners' helmet and maintained the balance between the recognition accuracy and the calculation speed. Then, a multi-person posture estimation algorithm was used to locate the joint points of each miner and determine the miners' complex behavioral status. Finally, the upper limb nodes of the target were extracted based on fusion model of multi-target recognition and multi-person posture estimation, and then the helmet-

* 文章编号:1003-3033(2024)03-0237-10; 收稿日期:2023-09-09; 修稿日期:2023-12-22

** 通信作者:焦文华(1975—),男,山东济宁人,博士,研究员,主要从事机器视觉及其矿山应用研究。E-mail: wjiao@cumt.edu.cn。

wearing condition was determined by the spatial topological relationship between the upper limb nodes and the helmet frame. Moreover, 3 000 pieces of image data were selected to validate the proposed method's performance. The results indicated that the machine vision system can optimize the hardware and software configuration to improve the system's performance. Furthermore, the identification accuracy reached up to 91.1%, which was much better than that of the latest helmet-wearing recognition algorithm. Therefore, the proposed system in this study can meet the requirements of helmet-wearing recognition accuracy for underground coal miners.

Keywords: coal mine underground; helmet-wearing; recognition algorithm; posture analysis; model optimization

0 引言

煤矿工人井下作业处于黑暗、封闭、复杂多变的地下环境,佩戴安全帽能够有效降低其头部受到落石、器械、摔打、电击等有害因素带来的危害。监测工人的安全帽佩戴情况是煤矿井下作业的重要环节,而且,由于工人数量庞大,采用机器视觉自动化监测安全佩戴情况是高效、必要的措施。

采用机器视觉识别安全帽佩戴情况的关键之一是确定图像检测目标。目前,图像目标的检测模式分为2种:①整体标记安全帽和工人脸部(或侧脸和后头),识别安全帽的佩戴情况;②分别标记安全帽和人体(或人体部位),通过识别人体与安全帽的空间关系识别安全帽的佩戴情况。其中,人体目标和安全帽提取的主要难点是处理微小目标识别准确度过低、提高计算效率和应对复杂背景等问题。吴天舒等^[1]为改善单步多框目标检测(Single Shot MultiBox Detector, SSD)模型对微小目标识别准确度过低的缺陷,在SSD算法中增加了转置卷积(Transposed Convolution, TC)等优化算法,用于卷积输出的特征图。赵红成等^[2]在一种一阶段卷积神经网络(Convolutional Neural Network, CNN)模型—YOLO(You Only Look Once)中引入注意力机制,将通道和空间注意力模块分别集成到YOLOv3的主干网络和监测层的残差结构中,使模型专注于安全帽特征学习。颜银慧等^[3]为了改善SSD的资源消耗问题,采用小尺寸的卷积核替换大尺寸的卷积核,有效优化了模型结构,减少了模型参数,实现了识别准确度与计算速度2方面的平衡。张萌^[4]等采用轻量化移动网络(Mobile Convolutional Neural Network, MobileNet)替换YOLOv4的主干特征提取网络,实现动态视频状态下安全帽的快速检测。相对而言,获取安全帽和人体头部空间关系的主要难点是在小目标识别结果下获取判别信息。刘群等^[5]采用3阶

段识别思路,使用帧差法、椭圆模型法等算法检测图像区域内是否包含人体目标,将人体目标区域分成头部区域和其他区域,提取头部区域的图颜色特征、边缘特征等关键特征,利用统计分析法、对比法或神经网络法判断是否包含安全帽。毕林等^[6]提出基于神经网络的安全帽佩戴检测算法,为每1个训练数据打上戴安全帽、不戴安全帽、默认背景图3种标记,分别采用不同卷积层、池化层、全连层的CNN训练数据,并使用训练好的神经网络模型判断新输入的工人图像是否佩戴安全帽。这些算法默认工人处于直立状态,并未考虑人体弯腰、后仰等复杂姿态。考虑到人体姿态的影响,张博等^[7]采用姿态估计网络(Open Pose Neural Network, OPN)定位人体头颈部位置并自动截取其周围小范围的子图像,通过颈部和安全帽的空间位置判断佩戴情况。王雨声等^[8]考虑复杂姿态下的安全帽佩戴识别问题,通过肤色判断人体脸部,用于对比头部与安全帽的空间关系,一定程度上解决特定场景下的安全帽佩戴识别问题。总体而言,上述算法虽然能较好解决特定场景下的安全帽佩戴识别问题,但难以完全适用于煤矿井下环境。

鉴于此,笔者拟根据煤矿生产的实际需求,提出人体姿态分析的安全帽佩戴识别算法和机器视觉系统优化方法;通过识别人的重要关节信息分析人体姿态,综合所获人脸信息(或部分信息)和姿态信息,推断头部位置,降低仅通过脸部或颈部确定头部位置的不确定性;采用网络压缩模型(Squeezed Neural Network, SqueezeNet)优化模型结构,实现计算效率和辨识精度的平衡,以期为煤矿井下作业场景的安全帽佩戴识别提供理论基础。

1 考虑人体姿态的安全帽佩戴识别

采用改进SSD^[9]和多人姿态估计方法(Multi-

people Posture Estimation Method, AlphaPose)^[10] 的模型融合方法,提出一种高效的安全帽佩戴识别算法和系统应用优化方法。首先,基于 SqueezeNet 优化 SSD 的模型结构,形成 SqueezeNet-SSD 模型,兼顾识别准确度与计算效率;然后,基于 SqueezeNet-SSD 和 AlphaPose 的模型融合方法,通过综合所获人脸信息(或部分信息)和姿态信息,推断头部位置;最后,根据不同的人体姿态,确定适用的安全帽佩戴识别算法。此外,针对所提方法中的参数设置与调整问题,提出一种安全帽佩戴异常指数表示方法,用于指导机器视觉系统的硬件调试和软件参数优化,提高系统的实际应用性能和效率。

1.1 安全帽佩戴识别算法

1.1.1 基于 SqueezeNet 的 SSD 模型优化

图 1 为煤矿井下不同作业区域的安全帽佩戴识别,大量的安全帽识别对模型效率具有很高要求。由 WOMG 等^[11]的研究可知:10 个火模型能提供强大的目标识别功能。因此,考虑在原始的 SSD 结构中使用火模型替换复杂的卷积层,减少 SSD 模型中 3×3 尺寸的卷积核,以及 3×3 尺寸卷积核的输入通道数量,并保证 SSD 模型的空间拓扑结构不变性,实现识别性能与计算精度间的相对平衡。对于 SSD 优化后的特征提取网络,每个火模型中不同尺寸(1×1 或 3×3)的卷积核数量如图 2 所示。图中,火模块主要包含压缩操作 S 和扩展操作 E;卷积层:3×3×57 为 3×3 尺寸的卷积层,卷积核个数为 57;最大池化层:3×3×1 为 3×3 尺寸的池化层,卷积核个数为 1。该网络的执行过程为:

- 1) 输入特征图。压缩层以特征图作为输入,该特征图通常是通过卷积层提取的。
- 2) 压缩操作。使用 1×1 卷积核处理输入特征图,将输入的通道数量压缩为较小的值,减少参数数量和计算量,同时保留重要的特征信息。
- 3) 扩展操作。首先,采用 1×1 卷积核扩展通道数量和压缩后的特征图;然后,采用 3×3 卷积核捕捉更大范围空间信息,增强特征表达能力。
- 4) 输出特征图。输出经过压缩和扩展后得到的特征图,使其具有较小的通道数量和更丰富的特征表示。

SSD 优化后的特征提取网络具有如下特点:

- 1) 设置 10 个火模型层。该设置可增加网络的深度、多层抽象、非线性能力和特征重用,学习到更复杂和抽象的特征表示,提高模型的性能。
- 2) 以第 1 个火模型模块的卷积核个数为基础,



图 1 不同作业区域的安全帽佩戴识别
Fig. 1 Identification of helmet-wearing in different work areas

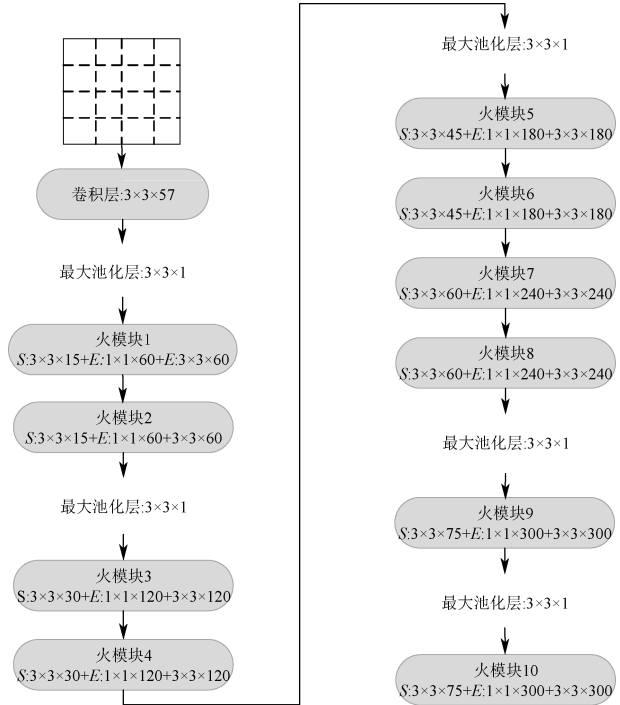


图 2 SSD 优化后的特征提取网络
Fig. 2 Feature extraction network of optimized SSD

每隔 2 个模块,压缩层的卷积核数量加倍,其变化特征为 60、120、180、240、300。该设置下,每个卷积核将对输入特征图能提取更多样化的特征信息,学习到更复杂和抽象的特征表示,提高网络对不同类别和特征的区分能力;此外,减少欠拟合的风险,并提高模型的泛化能力。

3) 根据文献^[11],扩展层中 1×1 的卷积核数量和 3×3 的卷积核数量相同,扩展层 1×1 的卷积核数量和 3×3 卷积核数量是同层压缩层中卷积核数量的 4 倍。该设置可以增加网络的参数化能力和表达能力,提高模型计算性能。

采用 1 组辅助卷积特征层和卷积预测器,优化 SSD 中多尺度检测网络,图 3 为 SqueezeNet-SSD 多尺度检测网络。如图 3 所示,网络结构的作用是替换 SSD 网络末端的 6 个复杂卷积层,火模型 4、火模型 8、火模型 9、火模型 10 分别对应于图 2 中特征提

取网络的结构,线段代表2个层之间具有连接关系。该网络结构,尤其是卷积层2-1、2-2、3-1、3-24层的卷积核个数,借鉴文献[11]中的设置。连接到最终预测层的数据层,每层都需要计算出2组数据,分别为位置误差(Localization Loss, Loc)和置信度误差(Confidence Loss, Con)。Loc卷积核个数 N_L 和Con卷积核个数 N_c 的确定方式:

$$N_L = N_p \times 4 \quad (1)$$

$$N_c = N_p \times N_c \quad (2)$$

式中: N_p 为先验框的个数; N_c 为预测目标种类数量(这里指人和安全帽)。火模型4、Loc卷积核3-2层中 $N_p = 4$,火模型8、火模型9、火模型10、卷积核2-1、2-2、3-1层中 $N_p = 6$ 。采用这种设置是因为试验中发现其可同时实现网络的最大压缩率和80%以上的预测精确率。按实际需求,设置 $N_c = 2$ 。位置信息包含边界框的中心坐标 (p, q) 及宽 w 、高 h ,表示为 (p, q, w, h) ;置信度信息中,如果检测目标共有 Γ 个类别,模型需要预测 $\Gamma + 1$ 个置信度值,其中第1个置信度是不含目标只属于背景的评分,当第1个置信度最高时,表示边界框中不包含目标。模型的实际预测值是相对于先验框的转换值。记先验框位置为 $d = (d^p, d^q, d^w, d^h)$,对应预测的包围框为 $b = (b^p, b^q, b^w, b^h)$,那么边界框预测值 l 可编码为 b 相对于 d 的转换值:

$$l^p = \frac{b^p - d^p}{d^w}, l^q = \frac{b^q - d^q}{d^h},$$

$$l^w = \ln\left(\frac{b^w}{d^w}\right), l^h = \ln\left(\frac{b^h}{d^h}\right) \quad (3)$$

式中: l^p 和 l^q 分别为转换边界框的左下角横、纵坐标; l^w 和 l^h 分别为边界框的宽度和高度。从预测值 l 中解码得到边界框真实值 b 的方法如下:

$$b^p = d^w l^p + d^p, b^q = d^h l^q + d^q,$$

$$b^w = d^w \exp(l^w), b^h = d^h \exp(l^h) \quad (4)$$

优化后尺度检测网络的总体结构及每层的参数数量见表1。其中火模型4-Loc和火模型4-Con分别为火模型4中计算Loc和Con的部分。该网络的损失函数 $L(x, \varepsilon, l, g)$ 为Loc (x, l, g) 与Con (x, ε) 的加权和:

$$L(x, \varepsilon, l, g) = \frac{1}{N}(\text{Loc}(x, l, g) + m\text{Con}(x, \varepsilon)) \quad (5)$$

式中: x 为图像数据; N 为先验框的正样本数量; m 为权重系数; ε 为类别的置信度预测值; l 为先验框的所对应边界框的位置预测值; g 为标注的位置参

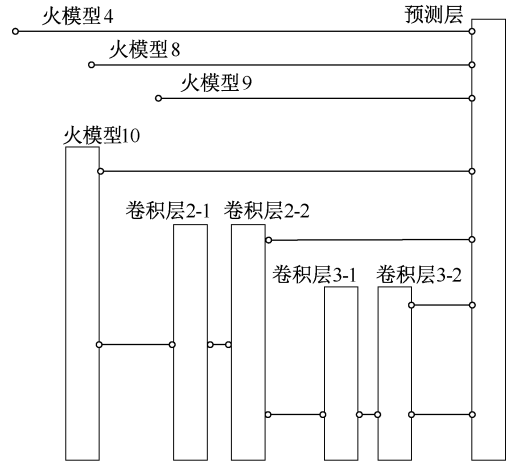


图3 SqueezeNet-SSD多尺度检测网络
Fig. 3 Multi-scale detection network of SqueezeNet-SSD

数。 $x_{ij}^p \in \{1, 0\}$ 为指示参数;当 $x_{ij}^p = 1$ 时为第 i 个先验框与第 j 个标签匹配,并且类别为 p 。

$$L_L(x, l, g) = \sum_{i \in \Xi} \sum_{m \in \{p, q, w, h\}} x_{ij}^m s_{L_1}(l_i^m - \hat{g}_j^m) \quad (6)$$

式中: $\hat{g}_j^p = \frac{g_j^p - d_i^p}{d_i^w}$, $\hat{g}_j^q = \frac{g_j^q - d_i^q}{d_i^h}$, $\hat{g}_j^w = \ln\left(\frac{g_j^w}{d_i^w}\right)$, $\hat{g}_j^h = \ln\left(\frac{g_j^h}{d_i^h}\right)$; g_j^p 为第 j 个先验框的 p 值,其余同类型参数具有相同的表示方法; Ξ 为正样本,即与任一标签匹配的先验框; $s_{L_1}(x)$ 为平滑 L_1 损失函数:

$$s_{L_1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{否则} \end{cases} \quad (7)$$

$$\text{Con}(x, \varepsilon) = - \sum_{i \in \Xi} x_{ij}^p \ln(\hat{\varepsilon}_i^p) - \sum_{i \in \Omega} \ln(\hat{\varepsilon}_i^0) \quad (8)$$

式中: Ω 为负样本,即未能与任何标签匹配的先验框; $\hat{\varepsilon}_i^p = \frac{\exp(\varepsilon_i^p)}{\sum_p \exp(\varepsilon_i^p)}$, ε_i^p 为第 i 个先验框属于类别 p 的置信度预测值; ε_i^0 为第 i 个先验框属于背景的置信度预测值。

表1 优化后SSD多尺度检测网络拓扑结构

Tab. 1 Topology of multi-scale detection network in optimized SSD

层	尺寸与个数	层	尺寸与个数
卷积层2-1	3×3×51	火模型9-Loc	3×3×24
卷积层2-2	3×3×46	火模型9-Con	3×3×12
卷积层3-1	3×3×55	火模型10-Loc	3×3×24
卷积层3-2	3×3×85	火模型10-Con	3×3×12
层	尺寸与个数	层	尺寸与个数
火模型4-Loc	3×3×16	卷积层2-2-Loc	3×3×24
火模型4-Con	3×3×8	卷积层2-2-Con	3×3×12

续表 1

火模型 8-Loc	3×3×24	卷积层 3-2-Loc	3×3×16
火模型 8-Con	3×3×12	卷积层 3-2-Con	3×3×8

1.1.2 SqueezeNet 和 AlphaPose 的融合

根据人体包围框的坐标和安全帽包围框的坐标,只能识别工人处于直立状态下的安全帽佩戴情况。假设 1 个工人正弯腰工作,此时安全帽包围框和人体包围框的空间位置如图 4a 所示;但当 1 个工人手拿安全帽站立时,其安全帽包围框和人体包围框的空间位置如图 4b 所示。图 4a 与图 4b 所示的情况是截然相反的,但其空间关系却一致,如图 4c 所示。因此,仅根据 2 个目标包围框的坐标信息无法判断工人复杂行为下的安全帽佩戴情况。

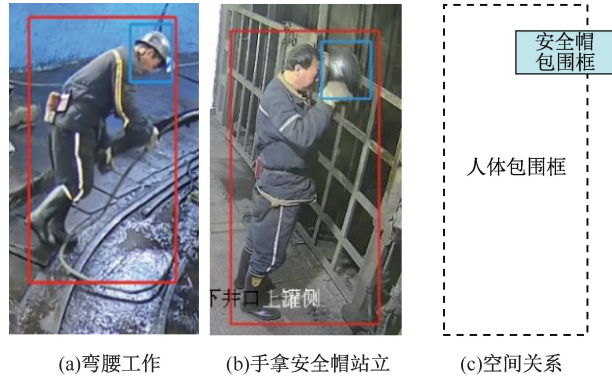


图 4 2 种典型情况下的包围框空间关系
Fig. 4 Spatial relationship of encirclement frames for two typical cases

鉴于此,笔者加入人体姿态估计算法 AlphaPose 来获取人体的关节点信息,以提高对工人复杂姿态下的安全帽佩戴识别准确度。在 SqueezeNet-SSD 和 AlphaPose 的模型融合中,SqueezeNet-SSD 模型主要用来识别图像中工人的安全帽并定位安全帽在图像中的最小包围框,进而获得每个目标的类别、识别准确度、以及目标的像素最小包围框,其效果如图 5 所示。AlphaPose 用来估计人体姿态,获得图像中每个人的重要关节点,如左眼、右眼、鼻子、左耳、右耳、左肩、右肩、左肘、右肘、左腕、右腕、左臀、右臀、左膝、右膝、左脚踝、右脚踝,其效果如图 6 所示。

根据推断出的头部区域、安全帽最小包围框、肩膀、臀部的相对关系,识别安全帽的佩戴情况。为了方便描述,将第 k 个人的头部信息 H^k 分为鼻子坐标 (N_x^k, N_y^k) 、左眼坐标 (η_x^k, η_y^k) 、右眼坐标 (μ_x^k, μ_y^k) 、左耳坐标 (ξ_x^k, ξ_y^k) 、右耳坐标 (o_x^k, o_y^k) ,那么 H^k 为

$$H^k = \{N_x^k, N_y^k, \eta_x^k, \eta_y^k, \mu_x^k, \mu_y^k, \xi_x^k, \xi_y^k, o_x^k, o_y^k\} \quad (9)$$

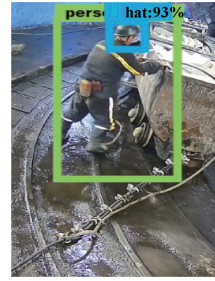


图 5 SSD 识别人体与安全帽

Fig. 5 Identification of human body and helmet by SSD



图 6 AlphaPose 标识人体关节点

Fig. 6 Human joint points identification by AlphaPose

同样,第 k 个人的左肩坐标为 (ρ_x^k, ρ_y^k) 、右肩坐标为 (ρ_x^k, ρ_y^k) 、左臀坐标为 (σ_x^k, σ_y^k) 、右臀坐标为 (σ_x^k, σ_y^k) 。那么,人体坐标为

$$\{\rho_x^k, \rho_y^k, \rho_x^k, \rho_y^k, \sigma_x^k, \sigma_y^k, \sigma_x^k, \sigma_y^k\} \quad (10)$$

将 SqueezeNet-SSD 识别出来的第 i 个安全帽的最小包围框记为 ψ^i 。 ψ^i 的右上角和左下角坐标分别记为 $(\psi_{Ux}^i, \psi_{Uy}^i)$ 和 $(\psi_{Dx}^i, \psi_{Dy}^i)$,即

$$\psi^i = \{\psi_{Ux}^i, \psi_{Uy}^i, \psi_{Dx}^i, \psi_{Dy}^i\} \quad (11)$$

值得注意的是,AlphaPose 可利用已知的面部特征和周围环境的上下文信息推断被遮挡部分的位置。例如:如果眼睛被遮挡,AlphaPose 可利用鼻子和嘴巴的位置,以及头部的朝向来推断眼睛的位置;当人脸特征因煤灰导致无法被全部识别时,AlphaPose 可利用人体姿势和动作的信息来帮助推断被遮挡部分的位置和姿势。

1.1.3 算法流程

以式(9)一式(11)所表征的数据为输入,识别 K 个工人的安全帽佩戴情况。对于第 k 个工人,若 H^k 的点全部在 ψ^i 之内,那么第 k 个人戴了第 i 个帽子;否则,判断是否满足条件 CA,若满足,则第 k 个人戴了第 i 个帽子;若仍不满足条件 CA,判断是否满足条件 CB,若满足,则第 k 个人戴了第 i 个帽子;否则,第 k 个人未戴第 i 个帽子。

以摄像头视角为参考观察人体行为,上述情况

CA 为人体左倾情况的判断条件,包括人体向左倾斜和向左弯腰。此时需要同时满足 2 个条件: $\sigma_x^k - \rho_x^k \geq \delta$ 且 H^k 的点全部位于 $\{\psi_{U_x}^i, \psi_{U_y}^i, \varphi \times \psi_{D_x}^i - \psi_{U_x}^i, \varphi \times \psi_{D_y}^i - \psi_{U_y}^i\}$ 之内。这 2 个条件分别为头部位于左肩膀的左上方和头部位于帽子扩大 $\varphi - 1$ 倍的包围框内。 δ 为肩膀与臀部两侧关节的横向像素距离阈值; φ ($\varphi \geq 1$) 控制模型的预测精度,当 $\varphi = 1$ 时,计算精度最高。由于计算过程中存在难以避免的误差,如头部的包围框预测, $\varphi = 1$ 会导致过多的误检,因此,需要在模型的应用过程中探寻适合的 φ 值。同理,情况 CB 为人体右倾情况的判断条件,包括人体向右倾斜和向右弯腰。此时需要同时满足 2 个条件: $\sigma_x^k - \rho_x^k \geq \delta$ 且 H^k 的点全部位于 $\{\varphi \times \psi_{U_x}^i - \psi_{D_x}^i, \psi_{U_y}^i, \psi_{D_x}^i, \varphi \times \psi_{D_y}^i - \psi_{U_y}^i\}$ 之内。这 2 个条件分别为头部位于右肩膀的右上方和头部位于帽子扩大 $\varphi - 1$ 倍的包围框内。

计算肩膀与臀部两侧关节的横向像素距离,需要获取人体臀部到头部的长度 h_1, cm ; 脚到臀部的长度 h_2, cm ; 上半身与腿所在方向的夹角,即倾斜度 θ 。根据文献[12],臀部到肩膀的距离 h_3 与 h_1 以及 h_2 的关系可表示为

$$h_3 = (h_1 + h_2) \times 0.382 \quad (12)$$

考虑到远距离拍摄条件下的人体细节像素点较少,此处设置判定人体存在倾斜或弯腰情况的 θ 阈值为 30° 。此时, δ 的计算方式为

$$\delta = h_3 \times \cos 60^\circ \quad (13)$$

值得注意的是,工人在摄像头下的行为具有非常强的不确定性,导致在一些特殊情况下对人体姿态的捕捉不够精准。针对该问题,可以在现场部署 2 个拍摄方向垂直的摄像头,通过获取多源信息,综合判断工人的安全帽佩戴情况。如何融合多源信息,提高煤矿井下工人安全帽佩戴情况的识别准确度不在考虑范围之内。

1.2 安全帽佩戴异常指数

为评估工人违反安全帽佩戴规范的程度,统计该工人未戴安全帽的次数和持续时长。一旦系统检测到某位工人未佩戴安全帽,将开始记录视频。在视频记录过程中,记视频帧采样的间隔为 ε , 当该工人戴上安全帽或离开摄像仪视野,则停止记录视频。将该次工人违规的持续时长记为 τ , 第 t 次视频帧采样的检测结果记为 γ_t 。那么,该次工人违规的异常程度量化值,称为异常指数,可记为:

$$\vartheta = \sum_{t=1}^{\frac{\tau}{\varepsilon}+1} \gamma_t \quad (14)$$

式中: γ_t 为检测结果标记,如果工人佩戴安全帽, $\gamma_t = 0$; 否则, $\gamma_t = 1$; $\frac{\tau}{\varepsilon} + 1$ 为采样点数量, $1 \leq t \leq \frac{\tau}{\varepsilon} + 1$ 。进而,第 k 个工人在第 m 个 τ 时间段内的安全帽佩戴异常指数为:

$$v_m^k = \sum_{t=1}^{\frac{\tau_m^k}{\varepsilon}+1} \gamma_t \quad (15)$$

式中: τ_m^k 为第 k 个工人的第 m 个记录时间段; γ_t 为第 k 个工人的第 m 个记录时间段中第 t 个采样点的检测结果标记。在统计周期内,记第 k 个工人的视频记录数量为 ω_k , 那么,第 k 个工人在统计周期内的安全帽佩戴异常指数为:

$$\vartheta_k = \sum_{m=1}^{\omega_k} \sum_{t=1}^{\frac{\tau_m^k}{\varepsilon}+1} \gamma_t \quad (16)$$

进而,系统记录的所有安全帽佩戴异常指数为:

$$\vartheta = \sum_{k=1}^{U_1} \sum_{m=1}^{\omega_k} \sum_{t=1}^{\frac{\tau_m^k}{\varepsilon}+1} \gamma_t \quad (17)$$

式中 U_1 为系统记录的异常工人总数。为了方便度量, ϑ 可被进一步处理为:

$$\vartheta = \frac{\vartheta}{\max(\sum_{m=1}^{\omega_k} \sum_{t=1}^{\frac{\tau_m^k}{\varepsilon}+1} \gamma_t) U_2} \quad (18)$$

式中 U_2 为所有井下作业人数,且满足 $1 \leq U_1 \leq U_2$;

$\max(\sum_{m=1}^{\omega_k} \sum_{t=1}^{\frac{\tau_m^k}{\varepsilon}+1} \gamma_t)$ 为 $\sum_{m=1}^{\omega_k} \sum_{t=1}^{\frac{\tau_m^k}{\varepsilon}+1} \gamma_t$ 中的最大值,所以

$$\max(\sum_{m=1}^{\omega_k} \sum_{t=1}^{\frac{\tau_m^k}{\varepsilon}+1} \gamma_t) U_2 \geq \sum_{k=1}^{U_1} \sum_{m=1}^{\omega_k} \sum_{t=1}^{\frac{\tau_m^k}{\varepsilon}+1} \gamma_t \quad (19)$$

即 $0 \leq \vartheta \leq 1$ 。

当 $\vartheta = 0$ 时,系统中没有记录的违规行为;当 $\vartheta = 1$ 时,所有工人均违规,且违规数量相同。可以看出,影响 ϑ 的原因较多,包括检测出未佩戴安全帽的情况和误检安全帽佩戴结果的情况。影响安全帽佩戴结果误检的因素包括硬件和软件因素,其中,硬件因素包括光照条件、物体遮挡、拍摄角度、岩石振动、安全帽颜色和污渍;软件因素包括参数设置,如 δ, φ 和神经网络参数等。所以,降低异常指数 ϑ 可以作为提高安全帽佩戴规范和系统识别安全帽佩戴能力的模型优化指标。告警人工确认界面,如图 7 所示。系统按照识别结果自动存储异常佩戴结果,并由人工再次确认。确认完成后,系统自动计算

得到 ϑ 值。在实际工作中,决策者可结合系统计算得到的 ϑ 值,优化井下作业中采用机器视觉监测安全帽佩戴的各个环节,提高安全隐患的预防能力。



图 7 告警人工确认界面

Fig. 7 Alarm manual confirmation interface

2 姿态分析的安全帽配戴识别算法

为验证所提算法的有效性,共设置 4 组试验。其中,第 1 组试验对比典型的安全帽佩戴识别算法,验证算法的安全帽佩戴识别性能;第 2 组试验验证 ϑ 值对机器视觉应用过程中硬件、软件调试的指导意义;第 3、4 组试验分别验证 SqueezeNet-SSD 模型的性能和考虑姿态分析的必要性。

2.1 对比算法参数和运行环境设置

1) 对比算法。最新的目标检测算法包括 YOLOv8 检测算法^[13]和 YOLOv7 检测算法^[14];最新的复杂姿态下的安全帽佩戴识别算法,包括增强的 Faster R-CNN 算法^[7]和融合姿态分析的 YOLOv4^[8];昏暗生产环境下的安全帽佩戴识别算法,包括改进 YOLOv5 检测算法^[15]和多策略 YOLOv5 检测算法^[16]、Faster R-CNN 算法^[17]和 Vibe-DPM 算法^[18]。上述对比算法的参数与原文保持一致。所提算法中的 AlphaPose 参数,与原文先保持一致; $\varphi=2, \delta=50$ 。

2) 性能准确度验证。

$$\alpha = \left(1 - \frac{\phi_1 + \phi_2}{\phi} \right) \times 100\% \quad (20)$$

式中: α 为识别准确率,%; ϕ 为所有人数; ϕ_1 为误检戴安全帽的人数; ϕ_2 为误检没戴安全帽的人数。

3) 图像预处理。在分析图像数据前,先对图像做增强处理,生成颜色对比图。图 8 为在黑暗工作场景下拍摄的图像,由于工人身着深色制服,与昏暗的背景非常相近。利用图像增强算法来增强前景目标,增加目标与背景颜色的区别度,如图 9 所示。

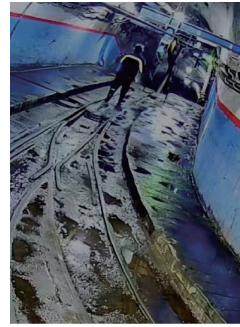


图 8 昏暗的原始图像

Fig. 8 Darkened original image

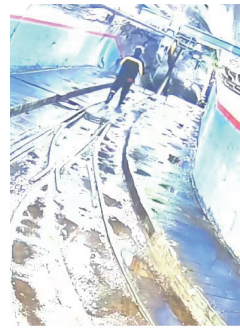


图 9 增强的新图像

Fig. 9 Enhanced new image

2.2 安全帽佩戴识别算法的有效性

挑选煤矿生产场景下的 3 000 张图像,用于对比所提算法与其他算法的识别准确度。选取 70% 的图像数据作为训练数据,从剩下的 30% 图像数据中不定量随机选择测试数据,用于测试所有算法。按照该算法重复执行 30 次,计算所有算法的平均识别准确度。为提高测试难度,着重选择含有复杂人体姿态的 1 256 张图像,其在所有图像数据中的比例为 42%,结果见表 2。

表 2 试验对比分析

Tab. 2 Experimental comparative analysis

算法	识别准确率/%	行为检测效果
所提算法	91.1	可检测全部行为
YOLOv8	86.8	仅站立、立坐
YOLOv7	85.1	仅站立、立坐
增强的 Faster R-CNN	86.6	可检测全部行为
融合姿态分析的 YOLOv4	87.9	可检测全部行为
改进 YOLOv5	87.2	仅站立、立坐
多策略 YOLOv5	85.7	仅站立、立坐
Faster R-CNN	84.3	仅站立、立坐
Vibe-DPM	80	仅站立、立坐

1) 所提算法的识别准确度达到 91.1%,明显高于 YOLOv8 的 86.8%,YOLOv7 的 85.1%,增强 Faster R-CNN 的 86.6%,融合姿态分析 YOLOv4 的 87.9%,

改进 YOLOv5 的 87.2%，多策略 YOLOv5 的 85.7%，Faster R-CNN 的 84.3% 和 Vibe-DPM 的 80%。

2) 针对昏暗生产环境下安全帽佩戴识别问题所提出的特定算法,改进 YOLOv5 和多策略 YOLOv5 均比 YOLOv7 的识别准确度高,改进 YOLOv5 甚至高于 YOLOv8。

3) 增强的 Faster R-CNN 和融合姿态分析的 YOLOv4,因考虑了复杂姿态下的安全帽佩戴识别,均具有较高的识别准确度,其中,融合姿态分析的 YOLOv4 仅次于文中所提算法,增强的 Faster R-CNN 具有非常接近 YOLOv8 的识别准确度。

4) 远距离下所获图像的像素点较少,使当前最新算法 YOLOv8 和 YOLOv7 的优势并不明显。而识别复杂人体姿态下的安全帽佩戴情况,不仅仅是一个目标检测问题,还必须加入人体姿态分析,提高判别精度。

2.3 安全帽佩戴异常指数的有效性

由上述试验可知:所提算法仍无法准确识别测试数据中 8.9% 的安全帽佩戴情况。因此,文中统计系统应用后 100 天的 ϑ 值变化,和根据 ϑ 值调试系统的硬件、软件设置的过程。在此调试周期内,首先根据识别结果寻找系统硬件布置的缺陷,包括灯光布置、摄像机角度、拍摄背景、边缘计算设备算力等方面,通过摄像机架设调整、补光、调整算力等方式提高识别性能。然后,调整系统的软件参数设置,包括 Squeeze 优化后的 SSD 超参数、 δ 、 φ 和 AlphaPose 参数等。在整个调试结果中,第 1—16 天为硬件调试阶段,此后为软件调试阶段。

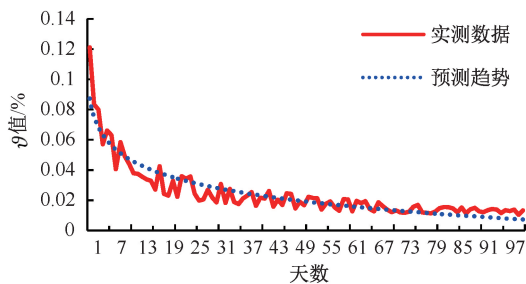


图 10 系统应用后的 ϑ 值变化情况

Fig. 10 Changes in ϑ after system application

如图 10 所示的 ϑ 值变化和预测趋势可知:硬件调试能实现模型优化,明显提高安全帽的佩戴识别准确度,而调试软件对提升系统性能的作用相对较小。逐渐降低的 ϑ 值波动,说明系统调试效果逐渐趋于稳定,其中, ϑ 值的波动主要来源于各种无法解决的外部因素,如作业场景的变化、工人数量的变化等情况。

此外, ϑ 值最终稳定在 0.02 左右,说明系统每天仍旧存在一些识别异常记录。从分类上,这些异常记录极少数为未佩戴安全帽的情况,而绝大部分是佩戴了安全帽却未能识别的情况。由于实际场景的限制,这些误识别结果无法完全避免,如工人弯腰工作时,安全帽被身体或设备遮挡,导致机器视觉失效。

2.4 姿态分析的安全帽佩戴识别其他性能

1) SqueezeNet-SSD 模型的性能。SqueezeNet-SSD 模型兼顾了识别准确度和训练速度。同样采用煤矿试验数据中的 3 000 张图像,分别对 SqueezeNet-SSD 与原始 SSD 模型、经过煤矿数据微调的 SSD 模型(利用煤矿数据针对性训练的 SSD 模型)在精度与速度上加以对比,试验结果见表 3。从精度上,原始 SSD 模型的标签库中没有安全帽标签,只能识别出人体,造成检测精度较低;而利用标记好的煤矿数据微调原始 SSD 模型后,一定程度上提高识别准确度;经过参数削减的 SSD 模型能达到 91.1% 的精度。从速度上,未经过优化的 SSD 模型包含大量参数,在测试 GPU 上每秒钟能计算 33 张图像;而优化后的 SqueezeNet-SSD 模型相对于原始 SSD 模型,被削减近 4 倍的参数,处理速度提升到每秒钟计算 100 张图像。因为 SqueezeNet-SSD 只识别煤矿中的工人与安全帽 2 个目标,所以 SqueezeNet-SSD 模型较于微调后的 SSD 模型精度有所提高。从整体上可以看出,SqueezeNet-SSD 能有效平衡计算精度与处理速度。

表 3 3 个模型在目标识别方面的性能对比

Tab. 3 Comparison of the performance of the three models in target recognition

模型	精度/%	尺寸/MB	速度/(帧·s ⁻¹)
原始 SSD 模型	45	82.1	33
微调 SSD 模型	72.7	82.1	33
SqueezeNet-SSD	91.1	21.1	100

2) 考虑姿态分析的必要性。图 11 为不考虑姿态分析、考虑安全帽与人体包围框位置和考虑姿态分析 3 种情况下的安全帽佩戴识别结果。图 11a 为 SqueezeNet-SSD 模型识别出来的人体和 安全帽结果;图 11b 为未加入 AlphaPose 模型,仅根据 2 个目标包围框的重合度来判断工人是否正确佩戴安全帽的识别结果,此时算法认为人体已经佩戴安全帽,属于误判;图 11c 为加入了 AlphaPose 模型后,融合安全帽的坐标数据以及人体关节的数据(主要是头部的位 置)。加入 AlphaPose 模型后,算法可识别出人体未佩戴安全帽,并用矩形框圈出。



图 11 安全帽佩戴检测结果

Fig. 11 Test results of helmet-wearing test

3 结 论

1) 采用 SqueezeNet 替换 SSD 模型中的卷积模块,在保持较高识别准确度的前提下,降低模型规模,提高检测效率,能快速反馈昏暗环境、多人作业的复杂场景下安全帽佩戴识别结果,实现对安全隐患的实时排查。

2) 考虑人体姿态分析的安全帽佩戴识别方法,优化了图像目标检测流程,避免已有算法中对检测目标的姿态限制,进一步提高工人作业过程中的安

全隐患自动化监测能力,有利于降低事故发生率和人力成本。

3) 提出一种安全帽佩戴异常指数评价方法,提高特定场景下安全帽佩戴识别系统的准确评估性,引导快速制定高效的安全帽佩戴监测方案,进一步增强监测系统的性能。

4) 对于新一代煤矿机器视觉监测系统硬件更新,文中采用的人体姿态分析模型计算量较大,耗费时间较长,无法适用于更高帧率的采样设备,在后续研究中将通过优化模型继续提高检测速度。

参 考 文 献

- [1] 吴天舒, 张志佳, 刘云鹏. 基于改进 SSD 的轻量化小目标检测算法[J]. 红外与激光工程, 2018, 47(7):37-43.
WU Tianshu, ZHANG Zhijia, LIU Yunpeng. Lightweight small target detection algorithm based on improved SSD[J]. Infrared and Laser Engineering, 2018, 47(7):37-43.
- [2] 赵红成, 田秀霞, 杨泽森, 等. 改进 YOLOv3 的复杂施工环境下安全帽佩戴检测算法[J]. 中国安全科学学报, 2022, 32(5):194-200.
ZHAO Hongcheng, TIAN Xiuxia, YANG Zesen, et al. Improved safety helmet wearing detection algorithm of YOLOv3 in complex construction environment[J]. China Safety Science Journal, 2022, 32(5):194-200.
- [3] 颜银慧, 周正锦, 向涛. 基于 SqueezeNet 的改进 SSD 模型及其应用[J]. 中国交通信息化, 2018(增1): 189-191.
YAN Yinhui, ZHOU Zhengjin, XIANG Tao. Improved SSD model and its application based on SqueezeNet[J]. China Transportation Informatization, 2018(S1): 189-191.
- [4] 张萌, 韩豫, 刘泽锋. 深度学习下建筑工人高空安全防护装备检测方法[J]. 中国安全科学学报, 2022, 32(5): 140-146.
ZHANG Meng, HAN Yu, LIU Zefeng. Detection method of high-altitude protective equipment for construction workers based on deep learning[J]. China Safety Science Journal, 2022, 32(5):140-146.
- [5] 刘群, 刘鸿涛, 田小锋. 一种适用于变电站内智能视频安全帽识别算法的研究与应用[J]. 黑龙江科技信息, 2016(31): 26-27.
LIU Qun, LIU Hongtao, TIAN Xiaofeng. Research and application of an intelligent video safety hat recognition algorithm

- in substation[J]. Heilongjiang Science and Technology Information, 2016(31): 26–27.
- [6] 毕林, 谢伟, 崔君. 基于卷积神经网络的矿工安全帽佩戴识别研究[J]. 黄金科学技术, 2017, 25(4):73–80.
BI Lin, XIE Wei, CUI Jun. Research on miners' safety helmet wear recognition based on convolutional neural network[J]. Gold Science and Technology, 2017, 25(4):73–80.
- [7] 张博, 宋元斌, 熊若鑫, 等. 融合人体关节点的安全帽佩戴检测[J]. 中国安全科学学报, 2020, 30(2):177–182.
ZHANG Bo, SONG Yuanbin, XIONG Ruoxin, et al. Helmet wear detection with human body junction[J]. China Safety Science Journal, 2020, 30(2):177–182.
- [8] 王雨生, 顾玉宛, 庄丽华, 等. 复杂姿态下的安全帽佩戴检测方法研究[J]. 计算机工程与应用, 2022, 58(1): 190–196.
WANG Yusheng, GU Yuwan, ZHUANG Lihua, et al. Research on helmet wearing detection method under complex posture[J]. Computer Engineering and Applications, 2022, 58(1):190–196.
- [9] LIU Wei, DRAGOMIR A, DUMITRU E, et al. SSD: single shot multibox detector[C]. European Conference on Computer Vision, Amsterdam, Springer International Publishing, 2016:21–37.
- [10] FANG Haoshu, XIE Shuqin, TAI Yuwing, et al. RMPE: regional multi-person pose estimation[C]. International Conference on Computer Vision, 2017:2 334–2 343.
- [11] WOMG A, SHAFIEE M, LI Francis, et al. Tiny SSD: a tiny single-shot detection deep convolutional neural network for real-time embedded object detection[C]. IEEE Conference on Computer and Robot Vision, 2018:95–101.
- [12] KIIRE S. Effect of leg-to-body ratio on body shape attractiveness[J]. Archives of Sexual Behavior, 2016, 45: 901–910.
- [13] LOU Haitong, DUAN Xuehu, GUO Junmei, et al. DC-YOLOv8: small-size object detection algorithm based on camera sensor[J]. Electronics, 2023, 12(10): DOI: 10.3390/electronics12102323.
- [14] WANG Chienyao, BOCHKOVSKIY A, LIAO Mark. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:7 464–7 475.
- [15] LI Xingke, WANG Yikai, QU Jiaxin, et al. Intelligent monitoring method of helmet wearing identification in power plants[C]. Asia Conference on Power and Electrical Engineering, 2023:1 699–1 703.
- [16] GUAN Yaqi, LI Wenqiang, HU Tianyu, et al. Design and implementation of safety helmet detection system based on YOLOv5[C]. Asia Conference on Computers and Communications, 2021:69–73.
- [17] 王忠玉. 智能视频监控下的安全帽佩戴检测系统的设计与实现[D]. 北京: 北京邮电大学, 2018.
WANG Zhongyu. Design and Implementation of helmet wearing detection system under intelligent Video surveillance[D]. Beijing: Beijing University of Posts and Telecommunications, 2018.
- [18] 杜思远. 变电站人员安全帽佩戴识别算法研究[D]. 重庆: 重庆大学, 2017.
DU Siyuan. Research on safety helmet wearing recognition algorithm for substation personnel[D]. Chongqing: Chongqing University, 2017.



作者简介: 左明成 (1992—),男,山东烟台人,博士,助理研究员,主要从事深度学习及其矿山应用研究。E-mail: Mingcheng.Zuo@cumt.edu.cn。