

中文引用格式:张泽辉,张乾隆,徐晓滨,等. 基于视觉的工人高处攀爬不安全行为识别模型[J]. 中国安全科学学报,2025,35(2): 144-151.

英文引用格式:ZHANG Zehui, ZHANG Qianlong, XU Xiaobin, et al. Unsafe behavior recognition model of high climbing workers based on vision[J]. China Safety Science Journal, 2025, 35(2): 144-151.

基于视觉的工人高处攀爬不安全行为识别模型*

张泽辉¹副研究员,张乾隆¹,徐晓滨¹教授,赵祖国²,王海泉³教授,李昊⁴

(1 杭州电子科技大学 中国-奥地利人工智能与先进制造“一带一路”联合实验室,浙江 杭州 310018;

2 中职物联(湖北)信息科技有限公司,湖北 武汉 430014;3 中原工学院 电子信息学院,

河南 郑州 450007;4 宁夏长骏科技咨询有限公司,宁夏 银川 750001)

中图分类号:X910

文献标志码:A

DOI: 10.16265/j.cnki.issn1003-3033.2025.02.0278

基金项目:国家重点研发计划资助项目(2022YFE0210700);国家水运安全工程技术研究中心开放基金资助(A202403);浙江省自然科学基金资助(LTGG24F030004);浙江省“尖兵”,“领雁”研发攻关计划资助项目(2024C03254)。

【摘要】 为精准识别高处作业人员攀爬过程中出现的不安全行为,提出一种基于视觉的工人高处攀爬不安全行为识别模型,由人体姿态估计算法和一维卷积不安全行为识别算法组成。人体姿态估计研究者采用量子化自编码器对人体关键点进行结构化建模,实现人体关键点坐标的检测;结合高处作业安全行为知识,基于一维卷积神经网络(1DCNN)模型构建高处攀爬不安全行为识别算法,并通过实验进行验证。结果表明:该模型在人员无遮挡情况和有部分遮挡情况下,分别达到93.91%和90.34%的精度;与支持向量机(SVM)、K最邻近算法(KNN)相比,该模型具有更强的泛化能力。

【关键词】 计算机视觉; 工人高处攀爬; 不安全行为; 识别模型; 一维卷积神经网络(1DCNN); 人体姿态估计

Unsafe behavior recognition model of high climbing workers based on vision

ZHANG Zehui¹, ZHANG Qianlong¹, XU Xiaobin¹, ZHAO Zuguo², WANG Haiquan³, LI Hao⁴

(1 China-Austria Belt and Road Joint Laboratory on Artificial Intelligence and Advanced Manufacturing,

Hangzhou Dianzi University, Hangzhou Zhejiang 310018, China; 2 Secondary Vocational Internet of

Things (Hubei) Information Technology Co., Ltd., Wuhan Hubei 430014, China; 3 School of

Electronic Information, Zhongyuan University of Technology, Zhengzhou Henan 450007, China;

4 Ningxia Changjun Technology Consulting Co.,Ltd., Yinchuan Ningxia 750001, China)

Abstract: In order to accurately identify unsafe behaviors during the climbing process of high-altitude workers, this paper proposed an unsafe behavior recognizing method for high climbing workers based on vision, which included the human pose estimation and the one-dimensional convolutional unsafe behavior recognition models. Quantized autoencoder was used to structurally model human key points in human pose estimation, enabling the detection of human key point coordinates. Combining with safety behavior

knowledge in high climbing operations, the unsafe behavior recognition model was constructed based on one-dimensional convolutional neural network model, and it was validated by industrial data experiments. Experimental results show that the accuracy of this method is 93.91% and 90.34% on unobstructed and partially obstructed datasets, respectively. Moreover, compared with support vector machines (SVM) and K-nearest neighbor (KNN), this method has stronger generalization capability.

Keywords: computer vision; worker high-altitude climbing; unsafe behavior; recognition model; one-dimensional convolutional neural network (1DCNN); human pose estimation

0 引言

据统计,2023年全国各类生产安全事故共死亡21 242人,其中工矿商贸企业就业人员10万人生产安全事故死亡人数1.244人^[1]。研究表明^[2-3]:绝大多数安全事故是由人的不安全行为造成的,快速、准确地识别工人不安全行为能够有效预防生产安全事故的发生。工业安全事故中,高处坠落是一种常见的事故类型^[4]。然而,传统的视频监控和安全员巡检的方式,难以及时识别和纠正高处作业工人的不安全行为,并且会消耗大量的人力和物力。因此,工业视频监控系统中引入计算机视觉技术,智能分析“人”的生产作业行为,能有效提升企业安全生产水平。

通过计算机视觉技术分析和处理图像或视频等数据,实现计算机对外部世界的感知和理解^[5-7]。张萌等^[8]运用深度学习技术检测工人是否佩戴安全帽与安全带,并在单目标、多目标和小目标检测中取得较高的精度。程淑红等^[9]提出一种基于卷积神经网络(Convolutional Neural Network, CNN)的人员吸烟检测算法,首先,利用回归树方法快速定位嘴部感兴趣区域,随后,采用残差网络检测这一区域的吸烟行为。LI Li等^[10]通过视频图像识别出驾驶员右手和右耳的边界,进而判断驾驶员是否存在分心驾驶行为。ZHANG Xiaojun^[11]利用深度学习和穿戴式传感器,实现人体运动的识别。XIE Yadong等^[12]通过分析驾驶员的吸烟习惯,采用卷积神经网络分析驾驶员手部运动,实现了对吸烟行为的高精度检测。SHEN Jie等^[13]基于卷积神经网络(Convolutional Neural Network, CNN)的人脸检测和边界盒回归方法,检测安全帽是否被正确佩戴。SENER等^[14]提出基于DenseNet121的卷积神经网络模型,根据安全带和身体的相对位置自动识别安全带是否正确佩戴。这些研究主要关注人员安全防护装备的穿戴,侧重于对点状、静态、碎片化的不安全行为的识别,较难满足特定工业流程连贯性、动态

性不安全行为的综合考量,如《危险化学品特殊作业安全规范》(GB30871—2022)^[15]中的高处作业。

鉴于此,笔者拟采用基于量子化编码器和CNN,精准监测工人作业流程,构建一种基于计算机视觉的工人高处攀爬不安全行为识别模型,以期填补现有应对连贯动态工业流程安全监测研究的空白。

1 工人高处攀爬不安全行为识别方法

1.1 不安全行为识别方法框架

根据工人高处攀爬不安全行为识别的需求,提出基于计算机视觉的工人高处攀爬不安全行为识别方法,如图1所示。该方法包括视频采集、人体姿态估计、行为特征处理、不安全行为识别和识别结果输出等5个基本模块。

视频采集模块通过视频监控系统采集工人高处作业行为的信息。该模块能够高效且准确地捕捉工人在高处攀爬过程中的各种行为动作和环境变化,并对采集的动态视频数据执行图像预处理。

人体姿态估计模块采用基于量子化编码器的人体姿态估计模型,检测目标对象的人体关键点坐标,即精准定位人体17个关键点,包括鼻子、左眼、右眼、左耳、右耳、左肩、右肩、左肘、右肘、左手、右手、左髌、右髌、左膝、右膝、左脚和右脚。

行为特征处理模块通过滑动窗口处理人体姿态关键点时间序列数据,并计算人体关键点的速度和加速度。根据GB30871—2022^[15],在攀爬脚手架时,身体与脚手架时刻保持“三点接触”,即在攀爬过程中,四肢中至多只有一个关节可以离开脚手架,视为安全动作;出现“两点接触”的情形,即在攀爬过程中,四肢中的2个关节同时离开脚手架,则视为危险动作。此外,攀爬过程中身体的运动变化,可以通过特定时刻的人体关键点的速度和加速度^[16]表示。

不安全行为识别模块采用一维CNN(One-Dimensional CNN, 1DCNN)不安全行为识别模型识别工人攀爬过程中的不安全行为。以17个关键点

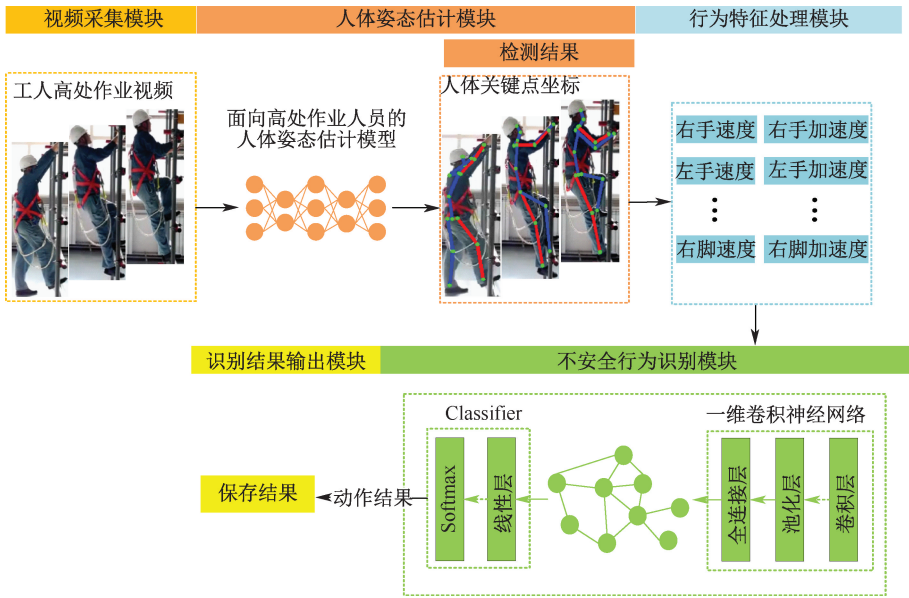


图1 基于计算机视觉的工人高处攀爬不安全行为识别方法

Fig. 1 Unsafe high-altitude climbing action recognition for industrial workers based on vision

坐标的速度和加速度为输入数据,通过数据分析,识别出多种不安全行为,具体包括:双手同时离开脚手架、一只手和一只脚同时离开脚手架以及双脚同时离开脚手架。

识别结果输出模块向各相关人员发出相应的预警信号,以便迅速采取必要的纠正措施,最大程度地降低事故发生的可能性。

1.2 人体姿态估计模型

根据文献[17]中的人体组合标记(Pose as Compositional Tokens, PCT),构建基于量子化自编码器(Vector Quantized Variational Auto-Encoder, VQ-VAE)的人体姿态估计模型,如图2所示。通过训练密码本量化人体各种姿态,并建立人体各个关节之间的关联关系。当人体部分关键点被遮挡时,使用VQ-VAE提取图像中人体关节特征,在训练好的密码本中进行匹配,实现人体姿态的估计。

第1阶段,通过人体关键点坐标来训练VQ-VAE网络^[18],该网络由编码器、密码本和解码器组成,如图2a所示。编码器 $F_e(G)$ 将一个人体姿态转换成 M 个token特征:

$$F_e(G) = (R_1, R_2, \dots, R_M) \quad (1)$$

式中: G 为原始姿态; M 为token特征的数量; F_e 为特征编码器。每个token特征对应于姿态的一个子结构,密码本将所提取的特征离散化,并记录在表格中。

第2阶段,构建一个转换器用于图像特征提取,并转换至密码本所在的特征空间。该转换器由骨干

网络和特征提取网络组成,如图2b所示。骨干网络使用Swin-transformer^[19]和2个残差卷积块^[20]提取图像特征,特征提取网络由线性层、多层感知机和Softmax组成,输出多个token特征。通过以上2个阶段训练编码器、密码本和解码器来构建人体关键点识别的深度学习神经网络。

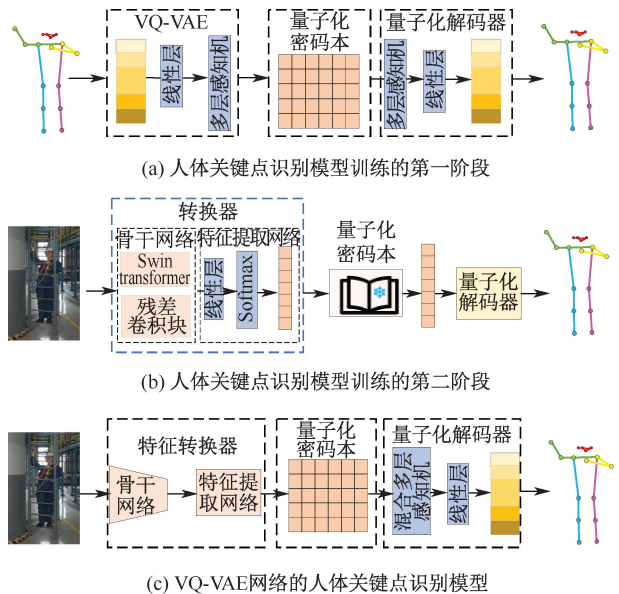


图2 基于VQ-VAE的人体姿态模型

Fig. 2 VQ-VAE based-human pose estimation model

第3阶段,输入待识别视频,将视频中每一帧通过转换器生成多个token特征,将每个特征与密码本中的特征库做最邻近匹配,然后将匹配结果输入到解码器中重构关键点坐标,如图2c所示。在人体

姿态估计模型中,从表达层面对姿态的关系进行建模,并将人体姿态约束在一个低维、紧凑和本质的表达空间,从而高效精准地检测高处作业人员的人体关键点。

1.3 人体关键点数据的预处理

高处作业工人在攀爬过程中,关键点位置随时间变化曲线和对应的关键点速度随时间变化的曲线,如图 3 所示。将人体关键点的速度和加速度作为特征数据输入不安全行为识别模型。由高处攀爬视频连续帧的人体关键点组成的时间序列 $X = \{x_j(1), x_j(2), \dots, x_j(t), \dots, x_j(T)\}$, J 为身体不同关节, t 为关键点所对应时间帧数。选取时间序列的滑动

窗口的大小为 N , 滑动窗口的时间序列为 $[x_j(1), x_j(2), \dots, x_j(N)]$ 。由于输入的格式不同,采用数据重构的方法将训练样本转换为预测模型的输入格式,关键点时间序列数据重构如图 4 所示。

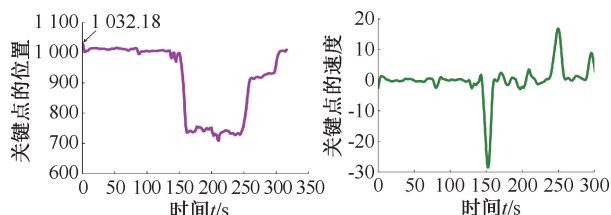


图 3 人体关键点位置和速度的变化曲线

Fig. 3 Human key point position and velocity change curve

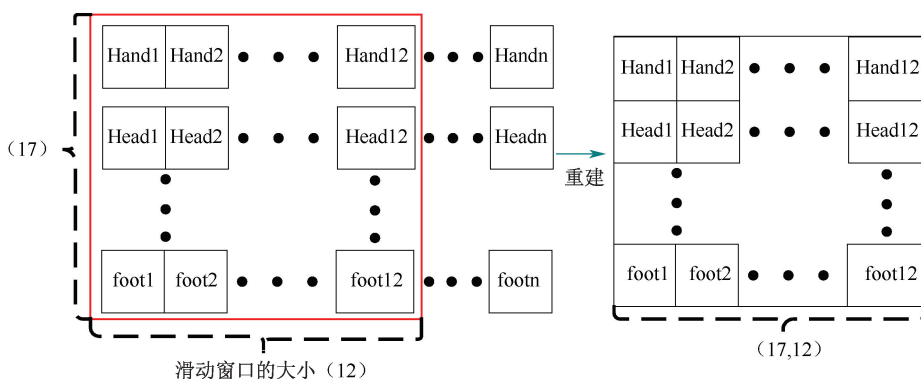


图 4 关键点时间序列数据重构

Fig. 4 Key point time series data reconstruction

获取滑动窗口关键点的速度 $[v_j(1), v_j(2), \dots, v_j(n)]$ 和加速度 $[a_j(1), a_j(2), \dots, a_j(n)]$:

$$v = \frac{x_j(n) - x_j(n-1)}{t_n - t_{n-1}} \quad (2)$$

$$a = \frac{v_j(n) - v_j(n-1)}{t_n - t_{n-1}} \quad (3)$$

式中: v 为人体关键点在相邻帧(帧每秒, Frame Per Second, FPS)图片像素点的位移变化速度, px/fps ; a 为人体关键点在相邻帧图片像素点的位移加速度, px^2/fps ; t 为视频对应的时间帧, fps ; $x_j(n)$ 为关键点在图片像素点的位置, px 。

1.4 不安全行为识别模型

CNN 包含输入层、卷积层、池化层和输出层 4 个层级,如图 5 所示。CNN 通过建立多个滤波器提取输入数据的特征。滤波器对输入层的数据逐层卷积和池化,层层提取输入数据的表征信息。这样的结构能够降低模型的复杂性,减少网络学习的参数,提高网络训练效率。

1DCNN 的输入为一维向量,如时间序列、文本

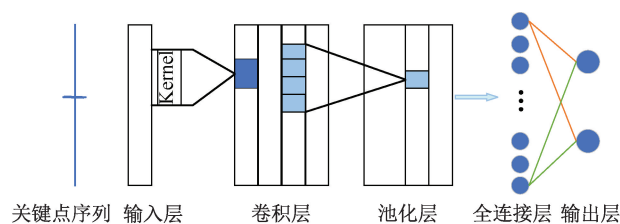


图 5 卷积神经网络

Fig. 5 Convolutional neural network

等,其卷积层使用 $k \times 1$ 大小的卷积核^[21]。在卷积层中,卷积核对前一层输出的特征向量进行卷积操作,利用非线性激活函数构建输出特征向量,每一层的输出均为对多个输入特征的卷积结果。卷积运算过程可以描述为:

$$y_j^l = f\left(\sum_{i \in N_j} z_i^{l-1} \times k_{ij}^l + b_j^l\right) \quad (4)$$

式中: N_j 为输入特征向量; l 为网络的层数; k 为卷积核的大小; b 为网络偏置; z_j^l 为第 l 层输出; z_j^{l-1} 为第 l 层输入; y 为第 i 个通道的第 j 个神经元的输出值。

在 CNN 的前向传播过程中,常用的非线性激活函数为线性整流函数 (Rectified Linear Unit,

ReLU), ReLU 会把输入向量中 ≤0 的特征值输出为 0, >0 的特征值原样输出, 提高网络的稀疏性, 加快模型训练过程中的收敛速度。

池化层用于减小特殊数据的尺寸, 进一步提炼特征信息。在实际应用中常采用最大池化层, 该层选择池化区域中的最大值进行输出, 从而降低可训练参数数目和提高特征的鲁棒性。

全连接层连接上一层池化层, 通过将多维数据展平为一维向量, 输入到全连接层中。全连接层将特征结果提纯, 增强特征的非线性表达能力, 将输出结果送入输出层中。

输出层采用 Softmax 分类器常用于多类别的分类任务, 用作输出层激活函数。将模型的原始输出转换为类别概率分布, 每个类别的概率在 0~1 之间, 所有类别的概率和为 1。

使用交叉熵损失函数训练 1DCNN 模型:

$$L = \sum_{i=1}^K - (h_i \ln(y_i) - (1 - h_i) \ln(1 - y_i)) \tag{5}$$

式中: y_i 为预测结果; h_i 为实际值; K 为输入数据的个数。

2 不安全行为识别模型训练

2.1 高处攀爬数据集构建

利用工业实训现场视频监控系统获取不同角度、光照条件下的工人高处攀爬视频。搜集到 10 组高处攀爬视频数据, 包括 5 组无遮挡和 5 组有遮挡的高处攀爬视频数据, 每组视频抽帧得到 870 张图片, 共得到 8 700 张图片数据。根据 GB30871—2022^[15], 在不安全行为数据集中标注安全和不安全标签。

2.2 识别模型训练

2.2.1 识别模型参数设置

不安全行为识别模型的训练参数设置见表 1。

表 1 模型训练参数设置

Table 1 Model parameter configuration

网络模型参数	参数配置值
检测数	2
输入尺寸	(34, 1)
批大小	50
优化器	Adam
学习率	0.001
卷积层数	2
池化层数	2
损失函数	交叉熵损失
训练轮次	80

2.2.2 训练步骤

1) 采用人体姿态估计模型检测人体关键点, 生成人体关键点坐标位置变化的时间序列, 计算人体关键点坐标变化的速度和加速度, 采用滑动时间窗口生成数据样本。将数据样本随机划分为训练集、验证集和测试集。

2) 使用训练集训练一维卷积不安全行为识别模型, 迭代更新模型的权重参数, 使所构建的模型能够识别工人高处攀爬的不安全行为。训练过程中, 使用验证集调整不安全行为识别模型的超参数。使用测试集评估验证训练好的不安全行为识别模型。模型训练步骤如图 6 所示。

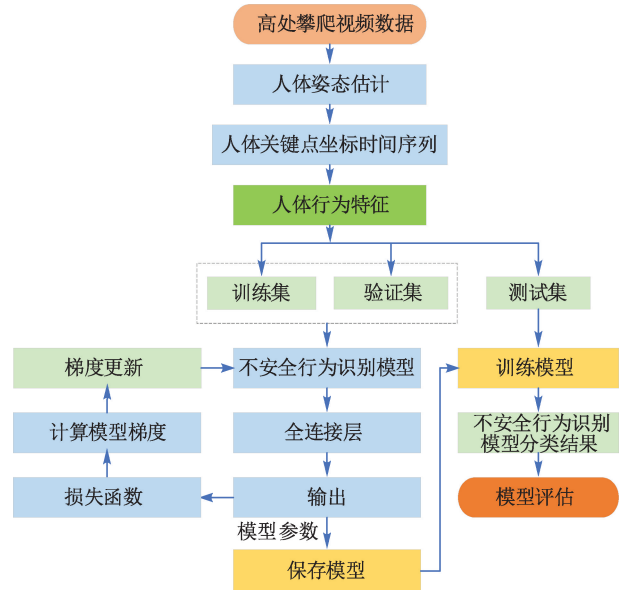


图 6 模型训练的流程

Fig. 6 Workflow of model training process

3 不安全行为识别模型的评估与应用

3.1 不安全行为识别模型评估

根据文献[22-24], 将构建的人体关键点的不安全行为数据集以 8:2 的比例划分训练集和测试集。视频数据集截帧图片后, 训练集和测试集数量分别为 6 960 和 1 740, 且分别包含相同数量的无遮挡和有遮挡的图片。使用处理后的数据训练不安全行为识别模型, 设置训练轮数为 80, 批次大小为 32。训练集和测试集的损失值和准确率曲线分别如图 7 和图 8 所示, 模型经过 30 轮左右趋于收敛。

为评价模型的识别效果, 选取常用评价分类模型性能的 3 个指标: 准确率 A 、精确率 P 和召回率 R , 对应的表达式如下:

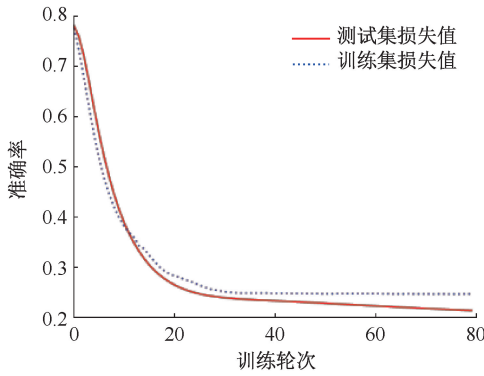


图 7 训练集和测试集损失值变化曲线

Fig. 7 Loss curves of training and testing datasets

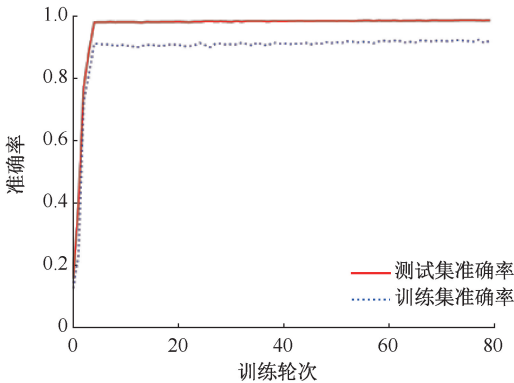


图 8 训练集和测试集准确率变化曲线

Fig. 8 Accuracy curves of training and testing datasets

$$A = \frac{T_p + T_N}{T_p + F_N + F_p + T_N} \quad (6)$$

$$P = \frac{T_p}{T_p + F_p} \quad (7)$$

$$R = \frac{T_p}{T_p + F_N} \quad (8)$$

式中： T_p 为正样本被正确检测的个数； F_p 为负样本被误检为正样本的个数； F_N 为正样本被误检为负样本的个数； T_N 为负样本被正确检测的个数。

图 9 为模型在测试集上分类结果的混淆矩阵。其中，矩阵中的行对应预测类别，列对应实际类别。矩阵的主对角线单元代表正确分类的观测值，副对角线单元表示错误分类的观测值。所有的数据样本被分类为安全和不安全 2 个标签。该模型在测试集的 A 、 P 和 R 分别为 92.18%、96.16% 和 95.32%。在高空攀爬过程中，系统针对出现的不安全行为发出预警。由图 9 可知：系统的安全行为出现误报率为 4.91%，误报概率较低。

PCT-1DCNN 模型、PCT-支持向量机 (Support Vector Machines, SVM) 模型和 PCT- K 最邻近算法

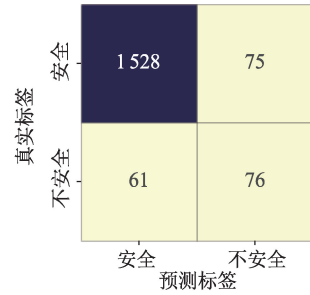


图 9 模型在测试集的混淆矩阵

Fig. 9 Confusion matrix of model in test dataset

(K-Nearest Neighbor, KNN) 模型对不安全行为数据集的处理结果见表 2。PCT-1DCNN 模型在整体测试的 A 、 P 和 R 较 PCT-SVM 模型和 PCT-KNN 模型有所提升，其中 A 分别提升 3.74% 和 3.06%。上述结果表明：相较于其他 2 个模型，PCT-1DCNN 模型在高空攀爬不安全行为识别方面具有更高的精准度。

表 2 不同模型结果对比

Table 2 Comparison results on different models %

检测模型	A	P	R
PCT-SVM	88.44	81.63	73.97
PCT-KNN	89.12	81.85	74.31
PCT-1DCNN	92.18	96.16	95.32

3.2 无遮挡动态视频的不安全行为识别

从测试集中选取无遮挡的数据集，共 870 组数据进行验证。该模型在无遮挡数据测试集中，由式(6)一式(8) 计算 A 、 P 和 R 分别为 93.91%、97.27% 和 96.32%，表明模型能较好识别无遮挡动态视频的不安全行为。无遮挡测试数据集上的混淆矩阵如图 10 所示。无遮挡测试数据的可视化效果如图 11 所示。其中，出现攀爬过程中一只手和一只脚同时离开脚手架的不安全行为。

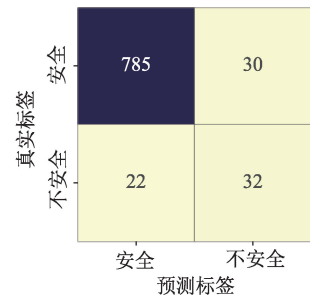


图 10 无遮挡测试数据集上的混淆矩阵

Fig. 10 Confusion matrix of unobscured test dataset

3.3 遮挡动态视频的不安全行为识别

从测试集中取出遮挡的数据集，共 870 组数据进行验证。该模型在遮挡数据测试集中，由



(a) 无遮挡安全行为 (b) 无遮挡不安全行为

图 11 无遮挡测试数据的可视化效果

Fig. 11 Visualization results of proposed method on unobscured test dataset

式(6)一式(8)计算 A 、 P 和 R 在 90.34%、95.01% 和 94.28%。与无遮挡识别任务相比,目标的误检个数有所增加,但模型的准确率并无明显下降,表明基于量子化的人体姿态估计模型能较好识别人体关键点。遮挡测试数据集上的混淆矩阵如图 12 所示。

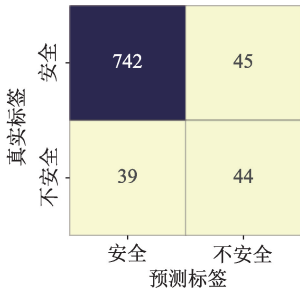


图 12 遮挡数据集上的混淆矩阵

Fig. 12 Confusion matrix of obscured test dataset

3.4 测试结果分析

综合 2 种高处攀爬场景的测试结果,人体姿态

估计算法和 1DCNN 不安全行为识别算法的组合模型,在无遮挡和有遮挡的数据集上均取得较好的识别效果。但不安全行为识别方法存在误检现象,且按无遮挡识别和有遮挡识别呈递增的趋势,主要原因是工人存在较大面积的严重遮挡且视频拍摄距离较远。存在严重遮挡的情况下,模型提取的有用特征较少,对关键点估计的精度有所下降,导致出现不安全行为误检。鉴于此,在后续模型优化中,改进人体姿态模型的结构,要进一步丰富高处攀爬图像数据集,使用存在严重遮挡和复杂场景高处攀爬图像进行模型训练,以提升不安全行为识别模型的鲁棒性;优化视频采集装置的布置方式,以获取更高质量的视频图像。

4 结论

1) 针对工业现场环境复杂和身体遮挡问题,采用量子化人体姿态估计模型检测人体关键点的坐标位置,结合 1DCNN 不安全行为识别模型,可动态监测工人高处攀爬行为。同时,该识别模型在人员无遮挡情况和有部分遮挡情况下,分别达到 93.91% 和 90.34% 精度。

2) 文中利用单一摄像头识别高处作业工人的不安全行为,在识别某些动作时可能存在漏检和误检的情况,后续研究拟采用多角度摄像头,并采用数据信息融合技术提高识别精度。

参考文献

[1] 国家统计局. 中华人民共和国 2023 年国民经济和社会发展统计公报[EB/OL]. (2024-02-29). https://www.stats.gov.cn/xxgk/sjfb/tjgh2020/202402/t20240229_1947923.html.

[2] SABRINA S O, WILLAMES D A S, BIANCA M V. Fatal fall-from-height accidents: statistical treatment using the human factors analysis and classification system-HFACS [J]. Journal of Safety Research, 2023: DOI: 10.1016/j.jsr.2023.05.004.

[3] YANG Jingjing, YE Gui, XIANG Qingting, et al. Insights into the mechanism of construction workers' unsafe behaviors from an individual perspective[J]. Safety Science, 2021: DOI:10.1016/j.ssci.2020.105004.

[4] 郑霞忠,王晓宇,陈述,等. 高处坠落事故的人因失误与干预策略研究[J]. 中国安全生产科学技术,2017,13(6): 139-144.
ZHENG Xiazhong, WANG Xiaoyu, CHEN Shu, et al. Study on human error and intervention strategies of high falling accidents[J]. Journal of Safety Science and Technology, 2017,13(6):139-144.

[5] ZHOU Longfei, ZHANG Lin, KONZ N. Computer vision techniques in manufacturing [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022,53(1):105-117.

[6] 伍麟,郝鸿宇,宋友. 基于计算机视觉的工业金属表面缺陷检测综述[J]. 自动化学报,2024,50(7):1 261-1 283.
WU Lin, HAO Hongyu, SONG You. A review of metal surface defect detection based on computer vision [J]. Acta Automatica Sinica, 2024,50(7):1 261-1 283.

- [7] FANG Weili, DING Lieyun, LOVE P E D, et al. Computer vision applications in construction safety assurance[J]. *Automation in Construction*, 2020; DOI:10.1016/j.autcon.2019.103013.
- [8] 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 safety protective equipment for construction workers based on deep learning[J]. *China Safety Science Journal*, 2022, 32(5): 140–146.
- [9] 程淑红, 马晓菲, 张仕军, 等. 基于多任务分类的吸烟行为检测[J]. *计量学报*, 2020, 41(5): 538–543.
CHENG Shuhong, MA Xiaofei, ZHANG Shijun, et al. Smoking detection algorithm based on multitask classification[J]. *Acta Metrologica Sinica*, 2020, 41(5): 538–543.
- [10] LI Li, ZHONG Boxuan, HUTMACHER J C, et al. Detection of driver manual distraction via image-based hand and ear recognition[J]. *Accident Analysis & Prevention*, 2020; DOI:10.1016/j.aap.2020.105432.
- [11] ZHANG Xiaojun. Application of human motion recognition utilizing deep learning and smart wearable device in sports[J]. *International Journal of System Assurance Engineering and Management*, 2021, 12(4): 835–843.
- [12] XIE Yadong, LI Fan, WU Yue, et al. HearSmoking: smoking detection in driving environment via acoustic sensing on smartphones[J]. *IEEE Transactions on Mobile Computing*, 2021, 21(8): 2 847–2 860.
- [13] SHEN Jie, XIONG Xin, LI Ying, et al. Detecting safety helmet wearing on construction sites with bounding-box regression and deep transfer learning[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2021, 36(2): 180–196.
- [14] SENER A S, INCE I F, BAYDARGIL H B, et al. Deep learning based automatic vertical height adjustment of incorrectly fastened seat belts for driver and passenger safety in fleet vehicles[J]. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 2022, 236(4): 639–654.
- [15] GB 30871—2022, 危险化学品企业特殊作业安全规范[S].
GB 30871—2022, Safety specifications of special work in hazardous chemicals enterprises[S].
- [16] 刘勃, 孔韦韦, 肖家钦, 等. 基于RBF神经网络的跌倒检测算法[J]. *计算机技术与发展*, 2022, 32(6): 167–178.
LIU Bo, KONG Weiwei, XIAO Jiaqin, et al. Fall detection algorithm based on rbf neural network[J]. *Computer Technology and Development*, 2022, 32(6): 167–178.
- [17] GENG Zigang, WANG Chunyu, WEI Yixuan, et al. Human pose as compositional tokens[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023; DOI:10.1109/CVPR52729.2023.00071.
- [18] HU Qiyu, ZHANG Guangyi, QIN Zhijin, et al. Robust semantic communications with masked VQ-VAE enabled codebook[J]. *IEEE Transactions on Wireless Communications*, 2023, 22(12): 8 707–8 722.
- [19] LIU Ze, LIN Yutong, CAO Yue, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021; DOI: 10.1109/ICCV48922.2021.00986.
- [20] YU Jianbo, ZHOU Xingkang. One-dimensional residual convolutional autoencoder based feature learning for gearbox fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(10): 6 347–6 358.
- [21] 池毅, 陈光武. 基于一维卷积神经网络的实时道岔故障诊断[J]. *计算机工程与应用*, 2022, 58(20): 293–299.
CHI Yi, CHEN Guangwu. Real-time turnout fault diagnosis based on one-dimensional convolutional neural network[J]. *Computer Engineering and Applications*, 2022, 58(20): 293–299.
- [22] SONG Yifan, ZHANG Zhang, SHAN Caifeng, et al. Constructing stronger and faster baselines for skeleton-based action recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(2): 1 474–1 488.
- [23] SONG Yifan, ZHANG Zhang, SHAN Caifeng, et al. Richly activated graph convolutional network for robust skeleton-based action recognition [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 31(5): 1 915–1 925.
- [24] ZHAO Baining, LAN Haijuan, NIU Zhenwen, et al. Detection and location of safety protective wear in power substation operation using wear-enhanced YOLOv3 algorithm[J]. *IEEE Access*, 2021; DOI:10.1109/ACCESS.2021.3104731.

作者简介: 张泽辉 (1994—), 男, 湖南衡阳人, 博士, 副研究员, 主要从事计算机视觉、人员行为识别、工业安全管控等方面的研究。E-mail: zhangzehui@hdu.edu.cn。