

# 基于机器视觉的施工人员危险行为监测与识别技术

## ——以某矿坑公园为例

张红卫<sup>1</sup>, 王冬松<sup>1</sup>, 陈文斌<sup>1</sup>, 王赓睿<sup>2</sup>, 徐俊豪<sup>2</sup>

<sup>1</sup>江苏泰禾建设工程有限公司, 江苏 徐州

<sup>2</sup>中国矿业大学力学与土木工程学院, 江苏 徐州

收稿日期: 2024年3月25日; 录用日期: 2024年4月22日; 发布日期: 2024年4月30日

### 摘要

超过80%的建筑安全事故源于施工人员的危险行为, 而传统人工管理方式的滞后性限制了监测效果, 迫切需要引入智能算法以提高监测精度和实时性。本文首先通过事故报告的统计数据对施工危险行为进行了评估和分类。以某矿坑公园为实证案例, 提出了一套监测点优化布置方案。在此基础上, 采用YOLO v5和CNN-LSTM算法实现了危险行为的精准识别。研究结果表明, 监测点布置方案经优化后, 各覆盖参数均取得不同程度的提升, 监测覆盖率提高了2.17%, 崖壁监测覆盖长度增加了40.33%。在状态类行为识别方面, 采用YOLO v5算法实现了高达97.2%的识别精度。对于动作类行为, 引入CNN-LSTM算法, 其准确率达到95.8%, 视频帧数达到30帧/s, 满足实时监测需求。该研究在实际应用场景中为施工人员危险行为监测提供了有效的技术支持。

### 关键词

监测点优化, 深度学习, 行为识别

# Construction Worker Hazardous Behavior Monitoring and Recognition Technology Based on Machine Vision

## —A Case Study of a Certain Quarry Park

Hongwei Zhang<sup>1</sup>, Dongsong Wang<sup>1</sup>, Wenbin Chen<sup>1</sup>, Gengrui Wang<sup>2</sup>, Junhao Xu<sup>2</sup>

<sup>1</sup>Jiangsu Taihe Construction Engineering Co., Ltd., Xuzhou Jiangsu

<sup>2</sup>School of Mechanics and Civil Engineering, China University of Mining and Technology, Xuzhou Jiangsu

Received: Mar. 25<sup>th</sup>, 2024; accepted: Apr. 22<sup>nd</sup>, 2024; published: Apr. 30<sup>th</sup>, 2024

文章引用: 张红卫, 王冬松, 陈文斌, 王赓睿, 徐俊豪. 基于机器视觉的施工人员危险行为监测与识别技术[J]. 计算机科学与应用, 2024, 14(4): 298-307. DOI: 10.12677/csa.2024.144099

## Abstract

More than 80% of the construction safety accidents are caused by the hazardous behavior of construction personnel, and the lag of traditional manual management limits the monitoring effect. To improve the monitoring accuracy and real-time performance, there is an urgent need to introduce intelligent algorithms. First of all, this paper evaluates and classifies the construction hazardous behavior through the statistical data of the accident report. Taking a certain quarry park as an empirical case, a set of optimal layout scheme of monitoring points is put forward. On this basis, the accurate identification of hazardous behavior is realized by using YOLO v5 and CNN-LSTM algorithm. The results show that after the layout of the monitoring site is optimized, the coverage parameters are improved in varying degrees, the monitoring coverage rate is increased by 2.17%, and the cliff monitoring coverage length is increased by 40.33%. In the aspect of state behavior recognition, the recognition accuracy of 97.2% is achieved by using YOLO v5 algorithm. For action behavior, the CNN-LSTM algorithm is introduced, and its accuracy is 95.8%, and the number of video frames can reach 30 frames/s, which meets the needs of real-time monitoring. This study provides effective technical support for hazardous behavior monitoring of construction workers in practical application scenarios.

## Keywords

Optimization of Monitoring Points, Deep Learning, Behavior Recognition

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

施工管理在建筑工程领域扮演至关重要的角色，超过 80% 的建筑安全事故源于人的主动或无意识的不安全行为[1] [2]。然而，传统的人工管理方式已明显滞后于时代要求。为进一步规范施工人员的行为，降低安全事故发生率，亟需采用先进的监测和预警技术，以适应施工管理智能化的要求[3]。

传统监测方法在处理施工人员不安全行为方面存在对人工监测的过度依赖，效率低下以及在复杂多变的施工环境中适应性受到限制的问题。为克服这些挑战，引入智能算法成为提高施工人员危险行为监测与识别效果的必然选择。Ludl 等人[4]将场景内人群的动作特征编码为人体姿态图像，在姿态识别的基础上实现了实时动作检测。徐守坤等人[5]采用 YOLO v3 对施工人员是否佩戴安全帽进行检测，并根据语义规则生成了基于检测结果的图像描述语句。林创鲁等人[6]采用 YOLO v4 对自动扶梯出入口场景下乘客的异常行为进行识别和预警。王琼等人[7]提出一种改进的密集轨迹算法，通过选择性搜索算法剔除人体行为轨迹的干扰，提高了人体行为识别分类结果的准确性。上述研究所涉及的检测目标皆处于视觉角度较佳的情境，然而，随着工程的不断推进，施工场地内的环境日益复杂，监测盲区不断增加。在此施工环境中，为有效监测施工人员的行为，合理布置监控点和应用精准算法缺一不可。

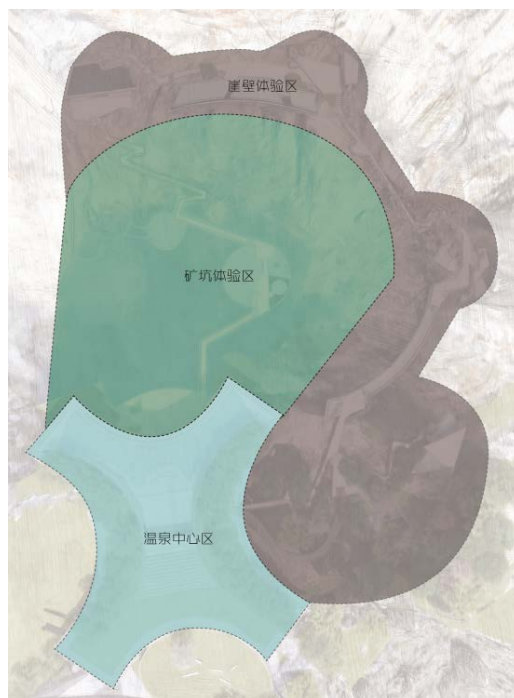
本文旨在讨论基于机器视觉的施工人员危险行为监测与识别技术，首先基于事故报告对关键的施工危险行为进行了统计分析，然后针对矿坑工程特殊的地形条件优化监控摄像头的布置方案，最后分别采用 YOLO v5 和 CNN-LSTM 对施工人员状态类和动作类的行为进行了监测识别。本文的研究内容为改善施工工地的整体安全性提供先进的解决方案。

## 2. 工程概况

某矿坑公园是一座多功能城市开放公园(图 1)，以其独特设计和广泛的功能性而备受关注。公园主体包括植物文化交流中心、民俗文化交流中心、地质文化交流中心、景观长廊以及温泉体验中心五个部分，总建筑面积达 12192.57 m<sup>2</sup>。该工程的独特之处在于依附崖壁而建，具有显著地形落差，为公园赋予了层次感和多样化的自然景观。然而，大幅度的地形变化同样为施工现场人员危险行为的监测工作带来挑战。



(a) 矿坑公园场地模型



(b) 公园分区示意图

Figure 1. A certain quarry park  
图 1. 某矿坑公园

## 3. 基于深度学习的智能检测技术

### 3.1. YOLOv5

YOLO v5 (You Only Look Once version 5)是一种前沿的深度学习目标检测框架,采用单一卷积神经网络

络进行全图的前向推理, 通过密集的预测框架实现对图像中对象的实时检测。YOLO v5 由骨干网络、颈部网络和检测头网络组成[8], 如图 2 所示, 引入深度残差网络(ResNet)和焦点损失函数(Focal Loss)等技术以提高针对小尺寸和高密度目标的检测精度[9]。对于实时处理施工人员密集、目标尺寸较小的情境, YOLO v5 其轻量化设计和高度优化的计算架构成为施工现场目标检测的理想选择。

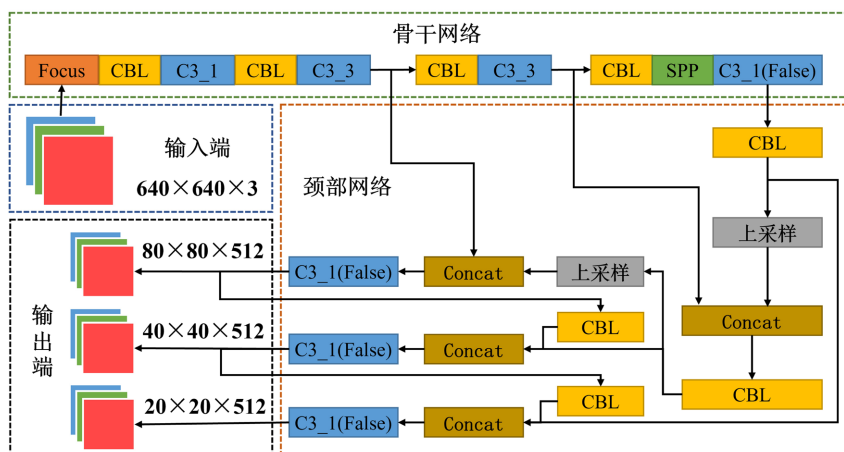


Figure 2. Architecture diagram of YOLO v5  
图 2. YOLO v5 架构图

### 3.2. CNN-LSTM

在施工场景中, 佩戴安全帽与摘除安全帽的行为因其动作顺序的改变而具有截然不同的安全含义。时空网络(Convolutional Neural Networks-Long short-term Memory, CNN-LSTM)能充分处理具有时空关联性的序列数据, 为有效识别诸如此类具有时序性的行为提供了解决方案。CNN-LSTM 通过融合卷积神经网络和长短时记忆网络[10], 不仅能够捕捉场景中对象的静态特征, 如安全帽的识别, 还能追踪特征随时间变化的动态性。基于 CNN-LSTM 的危险行为识别模型如图 3 所示, 首先利用卷积神经网络对预处理后的视频帧进行计算, 以提取其特征表示; 其次, 通过 LSTM 模型捕捉动作特征的时序演变, 将其嵌入到 CNN 结构的 Softmax 分类器中, 从而实现行为的准确分类。

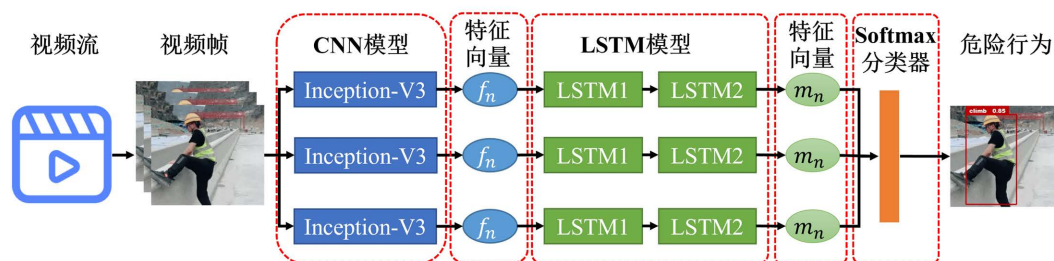


Figure 3. Unsafe behavior recognition model based on CNN LSTM  
图 3. 基于 CNN-LSTM 的危险行为识别模型

### 3.3. 数据库

通过在施工环境中模拟和在线搜集的方式获取了共计 4956 张图像, 具体呈现了施工人员佩戴安全帽的情景。明确定义了三个类别, 分别为未佩戴安全帽, 正确佩戴安全帽和已佩戴安全帽但不规范。通过对原始数据集采用翻转、旋转、缩放、裁剪等增强技术进行扩充, 将图像尺寸设置为  $640 \times 640$  像素, 最

终生成了 23,784 张图像, 其中 70% 的数据作为 YOLO v5 的训练集, 剩余 30% 的数据作为测试集。同样在施工环境中模拟翻越护栏的动作, 获取了共计 100 段不同时长的视频素材。在视频素材中遍历截取视频帧, 步长为 2, 帧数为 32。共计获得 9356 段视频帧, 其中 70% 作为 CNN-LSTM 的训练数据, 剩余 30% 作为测试数据。

## 4. 基于机器视觉和深度学习的施工危险行为识别

### 4.1. 施工危险行为统计分析

施工现场的危险因素主要来源于施工人员的不安全行为、施工机械的异常状态及不良环境条件[11]。根据住建部办公厅发布的 2017~2023 年房屋市政工程生产安全事故情况通报, 据不完全统计, 近 7 年来全国房屋市政工程生产安全事故总计 4780 起, 其中高处坠落 2532 起, 占 52.97%; 物体打击 725 起, 占 15.17%; 坍塌 415 起, 占 8.68%; 起重伤害 346 起, 占 7.24%。

结合统计数据 and 事故报告分析, 由施工人员的危险行为导致的工程事故总结如表 1 所示。由表可知, 对于高空坠落事故, 因个人安全意识不足而出现的一系列危险行为高达 58%。由于施工现场施工人员在设备、材料等物品的存放方面存在不规范行为, 并伴随着人员活动和设备、材料运输过程中的碰撞, 从而引发了物体打击事故的发生。而起重事故和坍塌事故通常源于机械质量问题、机械安装及操作人员的违章操作, 其发生具有不可预测性, 超过监测及时响应的速度。因此, 为预防和减少伤亡, 亟需对施工人员的行为进行实时监测。基于统计分析的结果, 对施工人员的危险行为进行监测等级的评估, 其中 L1 表示最高级别, L6 最低级别, 详见表 1。

**Table 1.** Statistics on construction hazard behaviors and types of safety accidents

**表 1.** 施工危险行为及安全事故类型统计

事故种类	相关的施工危险行为	比例/%	等级
高空坠落	施工区域内未佩戴安全帽	14	L3
	高处作业未系挂安全带	44	L1
	在未设置防护设施的高处进行临边作业	30	L2
物体打击	未使用安全通道	40	L4
	施工区域内未佩戴安全帽	15	L3
起重伤害	乘坐吊物升降装置	15	L5
	吊物下行走、站立	10	L6
	高处作业未系挂安全带	5	L1
坍塌事故	施工区域未佩戴安全帽	5	L3

### 4.2. 危险行为图采方案

#### 4.2.1. 监测点优化布置

鉴于矿坑公园依托矿山修复工程进行, 施工现场可划分为山底(矿坑体验区和温泉中心区)和山腰(崖壁体验区)两部分, 呈现出复杂的地形高度差, 导致监测范围难以覆盖全部场地, 为解决此难题, 拟采用优化布置摄像头测点与无人机巡航图采相结合的策略。结合 4.1 节对施工危险行为等级的评估结果和矿坑公园地形情况, 明确以下摄像头布置思路:

- 1) 场地呈现出中央低、四周高的地形特征, 其中西南方向相对较低, 而东北方向相对较高。首先考

考虑将摄像头设置在东北角，朝向西南方向进行观测。这样的布置方案能够有效利用地形高差，提高图像采集的覆盖面积；

2) 在矿坑体验区与崖壁体验区连接部分(包括连接梯和管道滑梯)，由于高度差较大，摄像头的采集效果收到限制。因此，为确保有效的图像采集，采用无人机进行图像补采，而不考虑摄像头的布置；

3) 场地内多为崖壁，地势高差显著，且行车道路有限，因此与道路相关的区域监测直接包括在施工区域监测中；

4) 摄像头的可视范围为以  $79^\circ$  开口角和 50 m 半径为参数的扇形区域，如图 4 所示。需要注意的是，为保证设备架设和线路铺设的顺利进行，摄像头不应直接布置在场地边界。

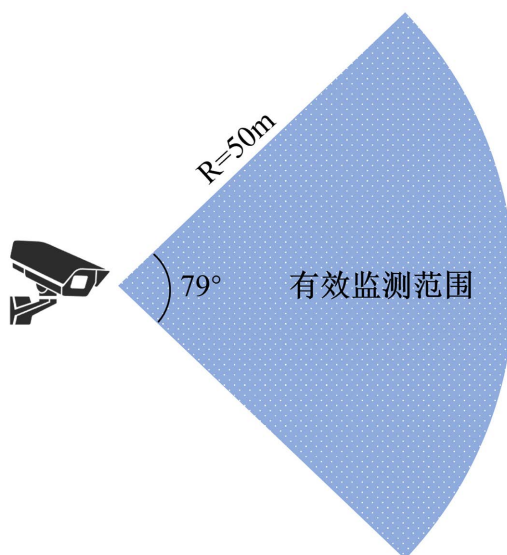


Figure 4. The surveillance visual range of the cameras

图 4. 摄像头的监测可视范围

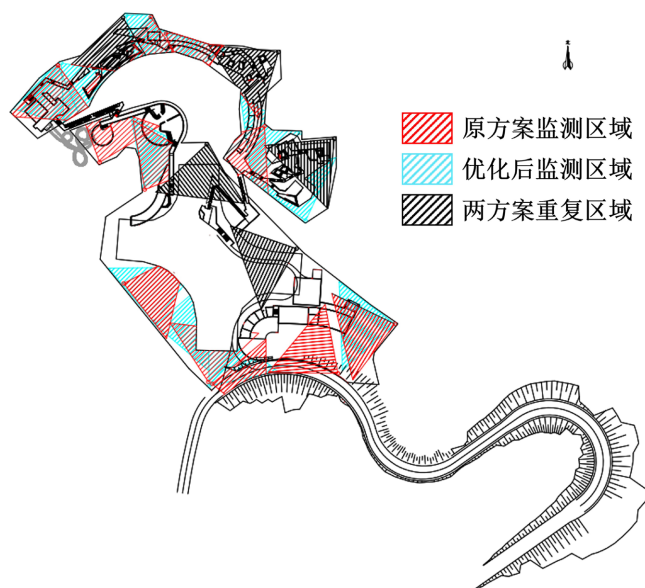


Figure 5. Comparison of the plan for the positioning of camera measurement points

图 5. 摄像头的测点布置方案对比

矿坑项目的初始监测方案如图 5 红色区域所示, 主要关注温泉中心区和矿坑体验区的施工, 同时覆盖进出道路的监测。原监测方案共设置 16 个监测点, 其中崖壁体验区布设 7 处, 矿坑体验区和温泉中心区共 9 处。根据上述思路对原监测方案进行优化, 最终在崖壁体验区布设 9 处摄像头, 矿坑体验区和温泉中心区共布设 7 处摄像头, 优化后的布置方案如图 5 中蓝色区域所示。两种监测方案的覆盖参数如表 2 所示, 优化后监测覆盖面积相较于初始方案增加了 552.42 m<sup>2</sup>, 覆盖率提升了 2.17%。鉴于优化方案侧重于在崖壁体验区布置监控, 而该区域呈现出整体细长的特征, 使得监控区域容易发生重叠, 因此重复覆盖率有所上升。优化布置方案在崖壁监测覆盖长度上增加了 63.186 m, 较原监测方案提高了 40.33%, 这很大程度上提高了对于高度落差区域危险行为的监测能力。综上所述, 优化后的测点布置方案具有良好的可行性和监测效果。

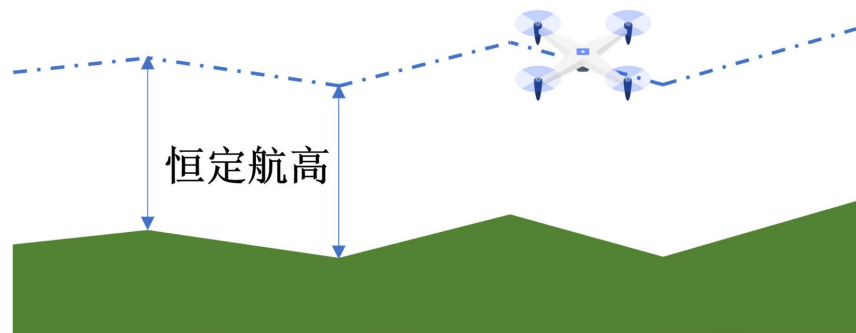
**Table 2.** Comparison of the coverage parameters between two monitoring plans

**表 2.** 监测方案的覆盖参数对比

对比参数	原监测点布置方案		优化后监测点布置方案	
	崖壁体验区	矿坑体验区和温泉中心区	崖壁体验区	矿坑体验区和温泉中心区
场地面积		25467.6 m <sup>2</sup>		25467.6 m <sup>2</sup>
摄像头数	9	7	7	9
监测覆盖面积		9132.33 m <sup>2</sup>		9684.75 m <sup>2</sup>
	3790.94 m <sup>2</sup>	5341.39 m <sup>2</sup>	5041.23 m <sup>2</sup>	4643.52 m <sup>2</sup>
覆盖率		35.86%		38.03%
重复监测面积		278.53 m <sup>2</sup>		520.56 m <sup>2</sup>
	229.01 m <sup>2</sup>	49.52 m <sup>2</sup>	520.56 m <sup>2</sup>	-
重复覆盖率		3.05%		5.37%
崖壁覆盖长度		156.664 m		219.850 m

#### 4.2.2. 无人机定点巡航

由于地形起伏显著, 导致摄像头的可视范围内仍存在部分盲区, 拟采用无人机巡航以弥补监测范围的不足。鉴于地势的复杂性且为获得地面分辨率一致的图像数据, 传统的定高度飞行策略不再适用, 需采用与地表恒定航高的仿地飞行方式进行图像补采[12], 如图 6 所示。



**Figure 6.** Schematic diagram of UAV terra-following flight

**图 6.** 无人机仿地飞行示意图

仿地飞行的关键在于首先要获取目标区域的高程信息, 在场地上以 160 m 的固定高度进行区域预扫

飞行以获取高程模型数据。根据《数字低空摄影规范》的要求，仿地飞行的相对航高设置为 20 m，航向重叠率为 70%，旁向重叠率为 60%，航线规划如图 7 所示。

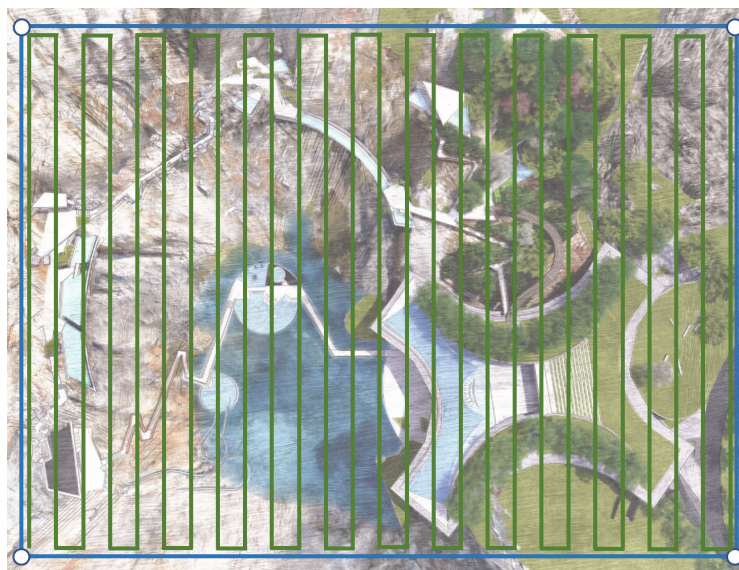


Figure 7. Route planning for terra-floolwing flights  
图 7. 仿地飞行的航线规划

### 4.3. 危险行为识别结果分析

根据 4.1 节统计分析可知，常见施工危险行为可分为两类，一类是未佩戴安全帽、未系挂安全绳等状态类行为，此类危险行为采用 YOLO v5 算法对场景中物体进行高效检测和定位，即可满足要求。然而，另一类以摘下安全帽、安全绳为代表的动作类行为，仅通过场景中是否含有关键物体(如安全帽、安全绳等)无法判定施工人员行为是否存在潜在危险性。时空网络(Convolutional Neural Networks-Long short-term Memory, CNN-LSTM)凭借在时空关联性方面的优越性，能够捕捉施工环境中物体之间的复杂关系。因此，引入该网络以达到人体动作特征的辨识效果。

#### 4.3.1. 基于 YOLO v5 的施工人员安全帽佩戴异常检测



Figure 8. Recognition effectiveness of wearing safety helmets  
图 8. 是否规范佩戴安全帽的识别效果



施工人员未规范佩戴安全帽是工程施工过程中典型的危险行为，以此为例进行验证。在网络训练过程中，设置学习率为 0.001，批处理大小为 32，训练迭代次数为 3000。试验结果表明，基于 YOLO v5 的物体检测与识别方法在对是否规范佩戴安全帽的人体进行有效识别方面表现卓越。在本实验中，该方法的准确率达到 97.2%。由图 8 可知，YOLO v5 算法在处理重叠影像方面的出色判断能力，且在连续的视频流中能够精确地进行目标的识别、分类和跟踪标记。

#### 4.3.2. 基于 CNN-LSTM 的危险动作识别

翻越护栏是施工场地中一种典型的违反安全规程的行为，可能引发人身伤害继而导致工程进度受阻。为验证 CNN-LSTM 算法在动作类行为识别方面的有效性。采用独热编码对翻越行为和未翻越行为进行标注，其中 0 表示未翻越，1 表示翻越。基于 CNN-LSTM 算法的行为识别效果如图 9 所示。将 CNN-LSTM 模型和其他三类以光流特征为基础的行为识别模型对比，即定向光流直方图(Histograms of Oriented Optical Flow, HOF)，运动边界直方图(Motion Boundary Histograms, MBH)和 HOF-MBH 复合模型。四种模型的识别效果如表 3 所示，HOF、MBH 和 HOF-MBH 复合模型需要手动提取特征，其自动化水平相较于 CNN-LSTM 算法较差。在实际应用场景中，算法的识别效率至关重要，CNN-LSTM 具有最高的识别帧率，达到 30 帧/s。相较于其他模型，CNN-LSTM 模型对动作的连续性识别效果最好，对于相似动作产生误判的几率最低，准确率高达 95.8%。



Figure 9. Recognition effectiveness of climbing  
图 9. 翻越护栏行为的识别效果

Table 3. Comparison of the effects of behavior recognition models  
表 3. 行为识别模型的效果对比

算法	HOF	MBH	HOF-MBH	CNN-LSTM
识别准确率/%	83.1	84.5	88.3	95.8
识别帧率	16.3	17.6	22.7	30.3

## 5. 结论

本文以某矿坑公园为具体案例，通过对施工危险行为的统计分析，明确了监测点的优化布置原则，基于 YOLO v5 和 CNN-LSTM 实现了状态类和动作类施工危险行为的精准识别。研究结论如下：

1) 基于对施工危险行为统计分析的结论，并结合矿坑公园的地形特点，提出的监测点优化方案相较于原方案在各覆盖参数上均取得不同程度的改善，监测覆盖率提升了 2.17%，崖壁监测覆盖长度增加了

40.33%。优化方案显著增强了高度落差大处危险行为的监测能力。

2) 在状态类行为的识别方面, YOLO v5 表现出较高的识别精度, 在判断场景中工人是否佩戴安全帽的准确率达到 97.2%, 并且在处理重叠影像时表现出良好的鲁棒性。对于动作类行为的识别, CNN-LSTM 较其他三种行为识别模型表现出最高的准确率, 达到 95.8%。此外, 该模型在视频帧数方面能达到 30 帧/s, 满足实时监测的要求。

## 参考文献

- [1] 龙丹冰, 魏君豪, 杨成. 知识图谱改进的施工行为安全风险与危险位置识别算法[J]. 中国安全科学学报, 2021, 31(12): 10-16. <https://doi.org/10.16265/j.cnki.issn1003-3033.2021.12.002>
- [2] Yang, J.J., Ye, G., Xiang, Q.T., et al. (2021) Insights into the Mechanism of Construction Workers' Unsafe Behaviors from an Individual Perspective. *Safety Science*, **133**, Article ID: 105004. <https://doi.org/10.1016/j.ssci.2020.105004>
- [3] 刘泽锋, 韩豫, 李文涛, 等. 视觉联合交集下施工危险区域侵入行为动态感知方法[J]. 安全与环境工程, 2023, 30(4): 18-25. <https://doi.org/10.13578/j.cnki.issn.1671-1556.20220727>
- [4] Ludl, D., Gulde, T. and Curio, C. (2019) Simple yet Efficient Real Time Pose Based Action Recognition. 2019 *IEEE Intelligent Transportation Systems Conference (ITSC)*, Auckland, 27-30 October 2019, 581-588. <https://doi.org/10.1109/ITSC.2019.8917128>
- [5] 徐守坤, 倪楚涵, 吉晨晨. 基于 YOLOv3 的施工场景安全帽佩戴的图像描述[J]. 计算机科学, 2020 47(8): 233-240.
- [6] 林创鲁, 叶亮, 李刚, 等. 基于深度学习的自动扶梯乘客异常行为识别方法研究[J]. 自动化与信息工程, 2022, 43(6): 1-6.
- [7] 王琼, 王旭, 刘云麟, 等. 基于改进密集轨迹算法的人体行为识别[J]. 计算机仿真, 2022, 39(12): 284-289+356.
- [8] Zhou, J.C., Jiang, P., Zou, A.R., et al. (2021) Ship Target Detection Algorithm Based on Improved YOLOv5. *Journal of Marine Science and Engineering*, **9**, Article 908. <https://doi.org/10.3390/jmse9080908>
- [9] Zhu, L.L., Geng, X., Li, Z., et al. (2021) Improving YOLO v5 with Attention Mechanism for Detecting Boulders from Planetary Images. *Remote Sensing*, **13**, Article 3776. <https://doi.org/10.3390/rs13183776>
- [10] Yao, H.X., Tang, X.F., Wei, H., et al. (2018) Modeling Spatial-Temporal Dynamics for Traffic Prediction. ArXiv abs/1803.01254: n pag.
- [11] 姜佩奇, 梁斌杰, 刘辉, 等. 工程场区施工人员靠近风险区实时预警方法[J]. 水利规划与设计, 2023(12): 123-129.
- [12] 郑朝峰, 王东野. 仿地飞行技术在山区大比例尺地形图测绘中的应用研究[J]. 科技创新与生产力, 2023, 44(9): 142-144.