

中文引用格式:郝秦霞,张家千. 矿井输送带运输区矿工不安全行为识别模型[J]. 中国安全科学学报, 2025, 35(10): 98-105.

英文引用格式:HAO Qinxia, ZHANG Jiaqian. Identification model of miners' unsafe behaviors in coal mine conveyor belt [J]. China Safety Science Journal, 2025, 35(10): 98-105.

矿井输送带运输区矿工不安全行为识别模型^{*}

郝秦霞 副教授, 张家千^{**}

(西安科技大学 通信与信息工程学院, 陕西 西安 710054)

中图分类号:X936

文献标志码:A

DOI: 10.16265/j.cnki.issn1003-3033.2025.10.1778

资助项目:陕西省重点研发计划项目(2024GX-YBXM-526)。

【摘要】 为提升矿井输送带运输区矿工不安全行为识别的准确性与实时性,解决现有基于人工监控手段实时性差、误检率高的问题,提出一种融合图像特征与人体骨骼特征的双流时空融合网络(DS-SFNet)。首先,针对井下低光照、粉尘干扰环境,设计亚像素卷积块注意力模块(SPCBAM),通过亚像素卷积与深度可分离卷积优化特征表达;其次,针对 OpenPose 模型计算资源消耗大的问题,采用 MobileNet v3 网络重构其主干特征提取网络,并引入空洞卷积与跨层连接;最后,构建融入注意力机制的分层特征融合模块,通过时空对齐与互补性建模深度融合图像特征与骨骼轨迹特征。结果表明:DS-SFNet 模型在 51 种人类动作数据库(HMDB51)和佛罗里达大学 101 类视频数据集(UCF101)上的识别准确率分别为 76.4%和 97.9%,较 SlowFast 模型分别提升 1.5%和 1.1%;在包含攀爬、跨越、倚靠、手搭 4 类行为的自建煤矿数据集中,平均识别准确率达 92.3%;MobileNet v3 重构的 OpenPose 模型参数量仅为原始视觉几何组网络(VGG19)的 11.5%,推理速度提升 3 倍以上;模型单帧处理时间为 38.7 ms,参数量为 57.3 M。

【关键词】 输送带运输; 不安全行为; 注意力机制; OpenPose 模型; 特征融合

Identification model of miners' unsafe behaviors in coal mine conveyor belt

HAO Qinxia, ZHANG Jiaqian

(College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an Shaanxi 710054, China)

Abstract: To improve the accuracy and real-time performance of identifying unsafe behaviors of miners in the mine belt transportation area, and to address the problems of poor real-time performance and high false detection rate in existing manual monitoring methods, a dual-stream spatiotemporal fusion network (DS-SFNet) that integrated image features and human skeleton features was proposed. First, challenges such as low illumination and dust interference in underground environments were addressed by designing a sub-pixel convolutional block attention module (SPCBAM), which combined with sub-pixel convolution and depth wise separable convolution to optimize feature representation. Second, to mitigate the high computational resource consumption of the OpenPose model, its backbone feature extraction network was reconstructed using MobileNet v3 by incorporating dilated convolutions and cross-layer connections.

* 文章编号:1003-3033(2025)10-0098-08; 收稿日期:2025-05-11; 修稿日期:2025-07-22

** 通信作者:张家千(1999—),男,辽宁葫芦岛人,硕士研究生,主要研究方向为煤矿安全。E-mail:849699608@qq.com。

Finally, a hierarchical feature fusion module was constructed to deeply integrate image features and skeletal trajectory features through spatiotemporal alignment and complementary modeling. The results demonstrate a recognition accuracy of 76.4% on HMDB51 (Human Motion Database 51) and 97.9% on UCF101 (University of Central Florida 101), outperforming the SlowFast model by 1.5% and 1.1%, respectively. On a self-built coal mine dataset containing four unsafe behaviors (climbing, crossing, leaning, and hand-leaning), the average recognition accuracy reaches 92.3%. The MobileNet v3-reconstructed OpenPose model reduces parameters to 11.5% of the original Visual Geometry Group 19 (VGG19) network while increasing inference speed by over 3 times. The complete framework achieves a single-frame processing time of 38.7 ms and a parameter count of 57.3 M.

Keywords: conveyor belt transportation; unsafe behavior; attention mechanism; OpenPose model; feature fusion

0 引言

在煤矿安全生产中,输送带运输区域的违规操作屡禁不绝,人员不安全行为已成为煤矿安全事故频发的主要诱因^[1-2]。矿井输送带运输区主要存在 4 类高风险的不安全行为:攀爬、跨越、手搭和倚靠输送带机。《煤矿及重点非煤矿山重大灾害风险防控建设工作总体方案》明确强调,需构建基于人工智能的视频辅助监管系统,以实现矿山关键区域和重点部位重大风险的实时监测和识别。然而,常规基于图像特征的识别方法在低质量图像条件下准确率显著下降,难以满足实际井下复杂环境的行为识别需求。因此,开展精确识别矿工的不安全行为研究,对于降低事故发生率和强化煤矿安全生产监管具有极其重要的意义。

在行为识别领域,现有技术主要分为基于图像特征、基于人体骨骼特征的单模态方法,以及多模态识别方法 3 大类。在基于图像特征的研究中,HAO Wenchao 等^[3]采用长短期神经网络模型实现井下视频行为识别;DENG Lujuan 等^[4]通过 3D 卷积神经网络方法获得人员行为识别的优异性能。在基于人体骨骼特征的方法中,CAO Xiangang 等^[5]结合 OpenPose 人体姿态估计模型和时空图卷积神经网络 (Spatial Temporal Graph Convolutional Networks, ST-GCN) 模型,通过精准提取人体关键点信息,可有效克服复杂背景和光照变化干扰,显著提升多关节交互行为的识别精度;NAN Yahui 等^[6]则基于轻量化模型 MobileNet 在人体行为和面部表情识别任务中实现精度和计算效率的双重提升。在多模态识别方法中,杨奉展等^[7]提出热红外与可见光图像融合的露天矿区障碍物检测方法,可显著降低低能见度环境下无人矿卡的误报与漏检率;饶天荣等^[8]采用

交叉注意力机制实现跨模态特征融合,能捕捉序列间的复杂关系;马天^[9]、王宇^[10]等构建了图像和人体骨骼特征协同的多模态识别模型。

在煤矿井下复杂环境及资源受限条件下,现有行为识别模型存在识别率低、环境适应性差及多特征融合不足等问题,因此,笔者拟提出一种融合人体姿态与带式输送机运行状态信息的轻量化识别模型,通过设计亚像素卷积注意力模块 (Sub-Pixel Convolutional Block Attention Module, SPCBAM) 与分层融合策略,优化特征表达与推理速度,最终形成可在井下复杂条件下高效稳定运行的不安全行为识别模型,以期预防输送带区域安全事故提供可靠的智能化监管支撑。

1 DS-SFNet 行为识别模型构建

1.1 双流时空融合网络模型架构

双流时空融合网络 (Dual-stream Spatiotemporal Fusion Network, DS-SFNet) 可提取视频图像特征和人体骨骼特征,利用特征融合模块进行特征融合,识别输送带区域不安全行为的类别,并建立 DS-SFNet 模型,其架构如图 1 所示。

1.2 图像特征提取

为降低计算量,视频片段时长控制在 2~4 s,选取连续的 16 帧/s 采用 50 层三维残差网络 (3D Residual Network 50, 3DResNet50) 模型提取图像特征 $F_v = \{f_{v1}, f_{v2}, \dots, f_{vm}\}$, 其中 $f_{vk} (k = 1, 2, \dots, m)$ 为第 k 个视频片段的图像特征。基于卷积注意力模块 (Convolutional Block Attention Module, CBAM) 构建 SPCBAM 注意力机制模型如图 2 所示。

首先,经亚像素卷积模块提取 CBAM 输出的特征进行重塑和上采样,从而将低分辨率特征图转换

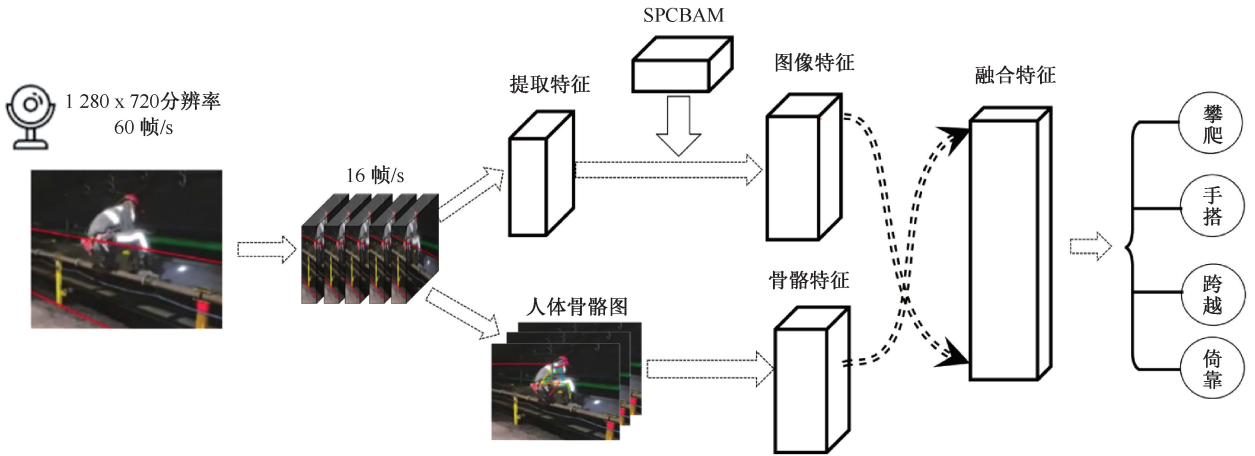


图1 DS-SFNet 模型架构

Fig. 1 DS-SFNet model architecture

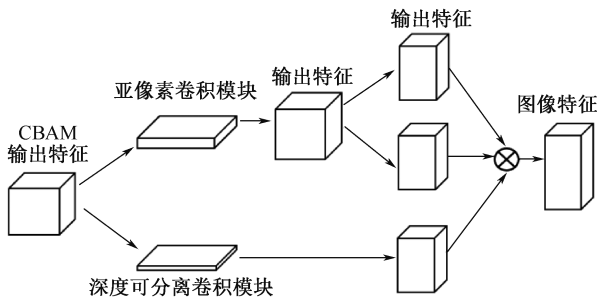


图2 SPCBAM 注意力机制模型

Fig. 2 SPCBAM attention mechanism model

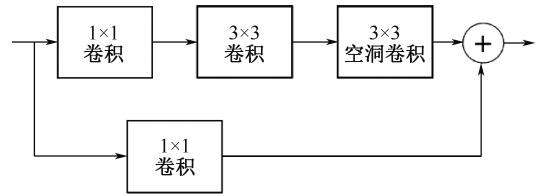


图3 改进卷积结构

Fig. 3 Improved convolutional structure

为高分辨率特征图。

其次,将上采样后的特征图沿通道方向分割为2个子特征图,保留特征图的多样性,为后续的特征融合提供更多的可能性。为增强特征的表达能力和减少计算量,SPCBAM 模块引入深度可分离卷积,通过2个步骤:①在通道维度上深度卷积;②在空间维度上采用1个 1×1 的卷积核卷积,得到更丰富的图像特征。

最后,将3个通道输出特征加权融合,生成最终的高分辨率特征图 F_i 。这种融合方式能综合利用不同层次的特征信息,使 SPCBAM 模块在保持原始通道数的同时,实现更高的空间分辨率,从而提升模型识别的准确率。

1.3 改进的轻量级 OpenPose 模型

为解决传统 OpenPose 模型复杂度高、参数量大等问题,采用 MobileNet v3-small 网络替代视觉几何组网络 (Visual Geometry Group 19, VGG19), 构建轻量级 OpenPose 模型,其改进卷积结构如图3所示。

采用层级卷积组合替代原始 7×7 卷积核,通过 1×1 卷积与双重 3×3 卷积串行,重构特征提取模块。

其中,末级 3×3 卷积引入扩张率为2的空洞卷积机制,可有效补偿深层网络中感受野衰减的问题。

构建基于 1×1 卷积的跨层短路连接通道,通过建立梯度传播的辅助路径,显著缓解网络深度增加导致的梯度弥散问题。

在维持模型识别准确度的前提下,改进后的轻量级 OpenPose 模型解决了原 7×7 卷积核运算开销大的问题,大幅削减模型所需的参数量,从而有效提升模型的推理速度和识别效率。轻量级 OpenPose 模型结构如图4所示。

1.4 人体骨骼特征提取

基于轻量级 OpenPose 模型,提取人体骨骼的序列数据(图5),并采用 ST-GCN 模型提取人体骨骼特征,对动态的不安全行为建模。

人体骨骼特征 F_s 计算如下:

$$F_s = \sum_{t=1}^T \sum_{j=1}^V F^{(L)}(v_{ij}) \quad (1)$$

式中: T 为时间维度; V 为空间维度; t 为动作序列的帧索引,用于遍历序列中的每一帧; j 为单帧内的关节索引,用于遍历帧中的所有关节; L 为网络层深度; $F^{(L)}$ 为第 t 帧第 j 个关节在第 L 层的时空特征向量; v_{ij} 为节点标志符,对应第 t 帧的第 j 个关节。

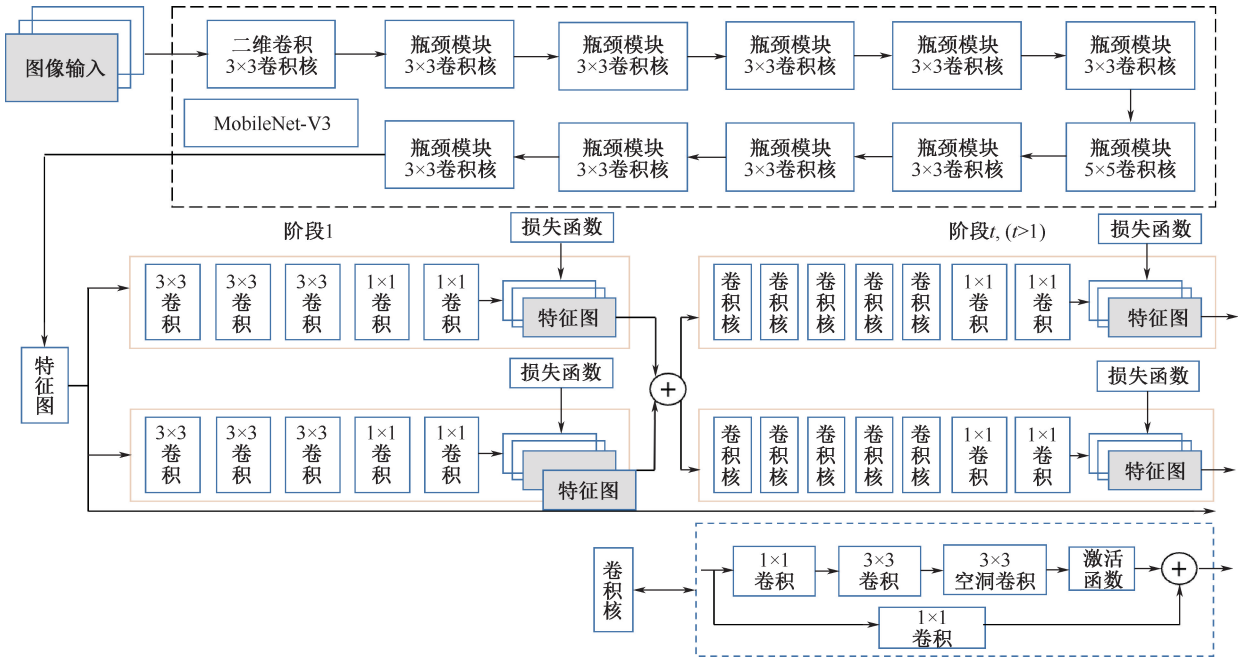


图 4 轻量级 OpenPose 模型结构

Fig. 4 Lightweight OpenPose model structure

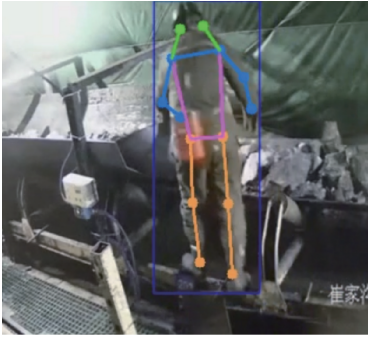


图 5 人体骨骼序列

Fig. 5 Human skeleton sequence

1.5 特征融合

为提升模型的环境感知能力和特征互补性,采用分层融合策略,借助特征融合模块,使图像特征和骨骼特征的深层信息交互,实现行为特征的全面表征,从而缓解井下复杂环境中单一特征表达缺失、信息利用率低和环境适应能力差等问题。特征融合流程如图 6 所示。

特征融合的具体步骤如下:

1) 将预处理得到的 16 帧/s 图像按时间顺序堆叠为 $3 \times 16 \times 224 \times 224$ 的张量,输入 3DResNet50-SPCBAM 网络,得到图像特征 F_i :

$$F_i \in R^{C \times T \times H \times W} \quad (2)$$

式中: R 为图像像素连续值的实数集; C 为特征图的通道数量,即每个空间位置上的特征向量的维度;

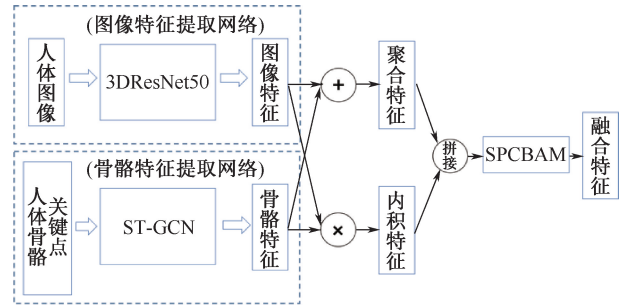


图 6 特征融合流程

Fig. 6 Flow of feature fusion

H 为特征图的空间高度; W 为特征图的空间宽度。

2) 骨骼特征的提取与对齐。基于轻量级 OpenPose 模型,从同一 16 帧/s 图像中提取每帧的 2D 关节点坐标 N ,生成骨骼序列 $S \in R^{T \times N \times 2}$ 。通过线性插值,确保骨骼序列与图像帧在时间维度上严格对齐。随后,将骨骼序列输入 ST-GCN 网络,输出骨骼特征 F_s :

$$F_s \in R^{C \times T \times N} (C = 512, T = 16, N = 17) \quad (3)$$

3) 统一图像特征和骨骼特征的维度。①沿空间维度 ($H \times W$) 平均池化图像特征 F_i ,得到图像 $F'_i \in R^{C \times T}$;②使用 1D 卷积将骨骼特征 F_s 的通道数从 256 映射至 512,得到骨骼特征 F'_s :

$$F'_s \in R^{C \times T} (C = 512, T = 16) \quad (4)$$

4) 在 (C, T) 维度上对齐图像特征与骨骼特征。①内积操作能捕捉 2 个向量间的相似性或相关性,通过计算图像特征与骨骼特征间的点积,衡量二者

的关联程度,得到点积特征向量 $F_d = F'_i \cdot F'_s$ 。
②聚合方法使重要特征在融合过程中占据更大比重,有助于提高融合效果。通过该操作得到聚合特征 $F_w = F'_i + F'_s$ 。
③拼接内积特征向量 F_d 和聚合特征向量 F_w , 得到 $F_c = \text{Concat}(F_d, F_w)$, 其中, Concat 函数表示在通道维度上拼接多个特征图。
④采用改进的 SPCBAM 注意力模块优化处理拼接后的特征 F_c , 最终得到融合后的特征 F_f :

$$F_f = \text{SPCBAM}(F_c) \tag{5}$$

5) 融合后的特征 F_f , 经全连接层线性变换为类别得分, 再通过 Softmax 函数将其归一化为概率分布, 从而得到各行为类别的预测概率。

2 不安全行为识别模型分析

2.1 数据集与试验设置

选取 51 种人类动作数据库 (Human Motion Database 51, HMDB51)^[11] 和佛罗里达大学 101 个人类动作类别视频数据集 (University of Central Florida 101, UCF101)^[12] 以及陕西省崔家沟煤矿井下监控视频数据作为试验数据集。其中, 井下监控视频数据集包含 840 段时长为 2~4 s 的视频, 涵盖攀爬、跨越、手搭和倚靠输送带 4 类不安全行为, 每类各有 210 段视频, 样本分布均匀, 部分井下数据集如图 7 所示。数据集分为训练集和测试集, 比例为 7 : 3, 采用准确率、精确率、召回率、 F_1 -Score、每秒传输帧数 (Frames Per Second, FPS)、参数数量和计算复杂度等指标开展结果评价。



图 7 部分煤矿数据集
Fig. 7 Partial coal mine dataset

试验在 Ubuntu20 操作系统上, 采用 3060 显卡进行图像处理。

2.2 SPCBAM 性能对比试验

为系统验证 SPCBAM 模块的性能增益, 在自建

煤矿井下行为数据集上构建多基线对比试验体系。试验设计 4 组递进式对比方案, 基础对照组 (第 1 组) 采用原始 3DResNet50 架构, 未集成任何注意力机制; 第 2 组试验将经典 CBAM 模块嵌入残差网络, 构建通道与空间联合注意力框架; 第 3 组试验引入挤压激励 (Squeeze-and-Excitation, SE) 注意力机制, 通过通道重要性重校准增强特征表达; 试验组 (第 4 组) 则采用构建的 SPCBAM 模块, 形成时空耦合的注意力增强架构。4 组性能对比结果见表 1。

表 1 性能对比试验

Table 1 Performance comparison test				%
模型	准确率	精确率	召回率	F_1 -Score
无注意力机制	80.2	78.7	78.6	78.6
SE	82.1	81.5	81.7	81.6
CBAM	83.8	83.8	83.1	83.5
SPCBAM	85.6	85.4	85.3	85.3

试验结果表明: 文中构建的 SPCBAM 架构在自建煤矿场景数据集上能够实现 85.6% 的行为识别精度, 与 SE 相比, SPCBAM 使识别准确率提升 3.5%, 同时, 在精确率、召回率和 F_1 -Score 值上分别获得 3.9%、3.6% 和 3.7% 的相对增益, 验证了时空耦合注意力机制在动态行为建模中的有效性。SE 通过全局平均池化和全连接层学习通道依赖性, 虽能强化关键通道特征, 但因缺乏空间维度的关注, 在复杂场景中的性能表现受到限制。与 CBAM 相比, SPCBAM 的准确率较 CBAM 模型提升 1.8%, 精确率、召回率、 F_1 -Score 分别提升 1.6%、2.1% 和 1.8%。SPCBAM 在特征表征能力上优于 CBAM, 且该优势在复杂场景中更为显著。尽管 CBAM 能协同增强通道和空间维度的特征表达, 但在矿井下等复杂环境中的性能仍不及 SPCBAM。

2.3 4 种特征提取网络模型对比试验

对比 VGG19-OpenPose、MobileNet-V1-OpenPose、MobileNet-V2-OpenPose、MobileNet-V3-OpenPose 这 4 种特征提取网络在煤矿数据集上的性能, 结果见表 2。由表 2 可知: 相较于 VGG19-OpenPose 网络, MobileNet-OpenPose 网络在精确率、召回率和 F_1 -Score 这 3 个指标上均有提升, 验证了 MobileNet 架构在兼顾效率提升的同时可实现性能优化。基于 MobileNet 的特征提取算法精确率均达到 88% 以上, 其中, MobileNet-V3-OpenPose 的识别精确率最高, 其识别处理速度同样最快, 较原 VGG19-OpenPose 网络提升 3 倍以上; 在参数数量方面, 改进后的 MobileNet-V3-OpenPose 与原始网络相比, 减少约

8.7 倍,大幅降低的参数量确保了所用网络模型的轻量性。

表 2 4 种特征提取网络性能对比

Table 2 Comparative performance evaluation of four feature extraction networks			
特征提取网络	精确率/ %	FPS/ (帧·s ⁻¹)	参数量/ M
VGG19-OpenPose	87.06	9~11	151
MobileNet-V1-OpenPose	88.23	21~24	18.7
MobileNet-V2-OpenPose	91.35	23~27	17.9
MobileNet-V3-OpenPose	93.68	30~34	17.4

矿工不安全行为识别可视化结果如图 8 所示,由图 8 可知:MobileNet-V3-OpenPose 模型可有效识别矿工手搭输送带机和跨越输送带机 2 类不安全行为,其识别准确率分别为 88%和 89%。



图 8 行为识别可视化结果

Fig. 8 Behavior recognition visualization results

2.4 多特征融合行为识别试验

为评估 SPCBAM 复合模型各部分对整体性能的贡献,并揭示不同模块在行为识别任务中的重要性,在公共数据集 UFC101 和 HMDB51 上开展消融试验,结果见表 3。

表 3 SPCBAM 复合模型的消融试验结果
Table 3 Ablation test results of SPCBAM composite model

图像特征提取网络	人体骨骼特征提取网络	SPCBAM 注意力机制	特征融合模块	准确率/%	
				HMDB51	UCF101
√	×	×	×	63.8	87.2
√	×	√	×	65.4	88.9
×	√	×	×	50.7	79.6
×	√	√	×	52.3	81.6
√	√	×	×	64.5	87.4
√	√	√	×	71.2	92.9
√	√	√	√	76.4	97.9

注“√”和“×”符号分别表示采纳该模块、不采纳该模块。

由表 3 可知:引入 SPCBAM 注意力机制后,基于图像特征网络的识别准确率在 HMDB51 和 UCF101 数据集上分别提升 1.6%和 1.7%;基于人体骨骼特征网络的识别准确率分别提升 1.6%和 2.0%,验证了 SPCBAM 注意力机制可有效提升模型识别准确率。此外,引入特征融合模块后,模型在相同数据集上的准确率分别提升了 5.2%和 5.0%,表明特征融合模块能够充分利用图像信息和骨骼信息的互补性,进而优化行为识别任务的性能表现。

为评估模型在不同行为类别的识别准确率,选取三维卷积神经网络(3D Convolutional Neural Networks, C3D)、3DResNet50、OpenPose + ST-GCN、SlowFast、DS-SFNet 模型,在自建煤矿数据集上开展对比试验,结果见表 4。

表 4 不同模型在各类别的准确率
Table 4 Accuracy of different models in various categories

模型	平均准确率	% categories			
		攀爬	跨越	倚靠	手搭
C3D	78.5	75.0	80.2	76.3	82.5
3DResNet50	84.1	82.5	86.8	81.7	85.4
OpenPose+ST-GCN	85.7	84.3	87.5	83.6	87.4
SlowFast	89.1	87.2	90.6	85.8	92.9
DS-SFNet	92.3	90.8	94.1	91.5	92.8

由表 4 可知:基于图像特征的 C3D 模型因环境敏感性,在“攀爬”和“倚靠”等小幅度动作识别中存在显著局限;3DResNet50 借助残差模块使得准确率较 C3D 有所提升,其“跨越”动作识别率达 86.8%,但“倚靠”动作仍受限于单一图像特征对姿态细节的表征缺陷,准确率仅 81.7%;基于骨骼特征的 OpenPose+ST-GCN 在动态动作识别中表现较优,“跨越”动作识别准确率达 87.5%,但因缺乏环境感知能力,“倚靠”动作识别准确率仅 83.6%;相较之下,DS-SFNet 模型通过多特征融合策略,在“跨越”和“手搭”动作的识别率分别达到 94.1%和 92.8%,较 SlowFast 模型对应指标分别提升 3.6%和 0.9%,验证了图像环境特征与骨骼运动特征的互补优势。

为进一步评估各模型在煤矿井下输送带运输区矿工不安全行为识别任务中的效率,对比各模型在处理单帧时间、FPS、参数量和计算复杂度上的性能表现,结果见表 5。

由表 5 可知:DS-SFNet 模型在煤矿井下输送带运输区矿工不安全行为识别任务中表现优异,兼具高效率和低复杂度特性。具体来看,DS-SFNet 模型

表 5 各模型性能对比

Table 5 Performance comparison of each model

模型	处理单 帧时间/ ms	FPS/ (帧·s ⁻¹)	参数 量/M	计算复 杂度/ GFLOPs
C3D	35.1	28.6	78.4	400
3DResNet50	45.3	22.2	46.2	150
OpenPose+ST-GCN	40.6	24.8	138.7	280
SlowFast	55.2	18.3	60.5	210
DS-SFNet	38.7	25.8	57.3	136

的处理单帧时间、FPS 优于 C3D、3DResNet50 和 SlowFast 模型,仅略低于 OpenPose+ST-GCN;在参数量方面,该模型显著低于 C3D 和 OpenPose+ST-GCN,表明其更高的参数利用效率;此外,其计算复杂度远低于 C3D 和 OpenPose+ST-GCN,在计算资源消耗上展现明显优势。综上,DS-SFNet 模型在维持较高帧率的同时,大幅降低了参数量和计算复杂度,实现了效率与性能的均衡优化,适用于实时视频分析等对计算资源敏感的应用场景。

3 结 论

1) SPCBAM 注意力机制通过亚像素卷积与深

度可分离卷积的协同优化,显著增强模型在煤矿井下复杂环境下的特征提取能力。消融试验结果显示:该模块是提升模型在低质视频中对不安全行为识别鲁棒性的关键。

2) 轻量化人体骨骼特征提取网络,在保持 OpenPose 模型关键点检测精度的前提下,将参数量降至原始模型的 11.5%,推理速度提升 3 倍以上,实现了模型轻量化与高效性的平衡。3) DS-SFNet 模型充分利用了图像上下文信息与人体骨骼运动信息的互补性,在自建煤矿数据集上,该模型对“跨越”、“手搭”等动态行为的识别准确率显著优于单一模态模型,验证了多特征融合策略在复杂行为表征上的优势。

4) 本文主要针对预设的 4 类明显不安全行为,对于更细微的违规动作(如操作姿势不当)识别能力仍有待加强。未来工作将探索对连续、隐蔽性不安全行为的时序建模,并进一步优化模型,以适应边缘计算设备的部署需求。

参 考 文 献

[1] TIAN Shuicheng, WANG Yajuan, LI Hongxia, et al. Analysis of the causes and safety countermeasures of coal mine accidents: a case study of coal mine accidents in China from 2018 to 2022[J]. Process Safety and Environmental Protection, 2024, 187: 864–875.

[2] 王大龙,王冰山,曹睿,等. 煤矿安全管理行为交互机制与模型[J]. 中国安全科学学报,2025,35(5):32–38.
WANG Dalong, WANG Bingshan, CAO Rui, et al. Interaction mechanism and model of coal mine safety management behaviors[J]. China Safety Science Journal, 2025, 35(5): 32–38.

[3] HAO Wenchao, JIANG Haiyan, SONG Qinghui, et al. A multi modal fusion coal gangue recognition method based on IBWO-CNN-LSTM[J]. Scientific Reports, 2024, 14: DOI: 10.1038/S41598-024-80811-6.

[4] DENG Lujuan, FU Ruochong, SUN Qian, et al. Abnormal behavior recognition based on feature fusion C3D network[J]. Journal of Electronic Imaging, 2023, 32(2): DOI: 10.1117/1.JEI.32.2.021605.

[5] CAO Xiangang, ZHANG Chiyu, WANG Peng, et al. Unsafe mining behavior identification method based on an improved ST-GCN[J]. Sustainability, 2023,15(2): DOI: 10.3390/SU15021041.

[6] NAN Yahui, JU Jianguo, HUA Qingyi, et al. A-mobileNet: an approach of facial expression recognition[J]. Alexandria Engineering Journal, 2022, 61(6): 4 435–4 444.

[7] 杨奉展,顾清华,李少博,等. 露天矿低能见度下多模态融合障碍物检测[J]. 中国安全科学学报,2025,35(5): 195–203.
YANG Fengzhan, GU Qinghua, LI Shaobo, et al. Multimodal fusion-based obstacle detection in low-visibility open-pit

mines[J]. China Safety Science Journal, 2025, 35(5):195-203.

[8] 饶天荣,潘涛,徐会军. 基于交叉注意力机制的煤矿井下不安全行为识别[J]. 工矿自动化,2022,48(10):48-54.
RAO Tianrong, PAN Tao, XU Huijun. Unsafe action recognition in underground coal mine based on cross-attention mechanism[J]. Journal of Mine Automation, 2022, 48(10): 48-54.

[9] 马天,姜梅,杨嘉怡,等. 基于多特征融合时差网络的带式输送机区域违规行为识别[J]. 工矿自动化,2024,50(7):115-122.
MA Tian, JIANG Mei, YANG Jiayi, et al. Recognition of violations in belt conveyor area based on multi-feature fusion for time-difference network[J]. Journal of Mine Automation, 2024, 50(7):115-122.

[10] 王宇,于春华,陈晓青,等. 基于多模态特征融合的井下人员不安全行为识别[J]. 工矿自动化,2023,49(11):138-144.
WANG Yu, YU Chunhua, CHEN Xiaoqing, et al. Recognition of unsafe behaviors of underground personnel based on multi modal feature fusion[J]. Journal of Mine Automation, 2023, 49(11): 138-144.

[11] KUEHNE H, JHUANG H, GARROTE E, et al. HMDB: a large video database for human motion recognition[C]. 2011 International Conference on Computer Vision, 2011:2 556-2 563.

[12] SOOMRO K, ZAMIR A R, SHAH M. UCF101: a dataset of 101 human actions classes from videos in the wild[J]. Computer Science, 2012: DOI:10.48550/arXiv.1212.0402.



作者简介： 郝秦霞 （1980—）,女,陕西西安人,博士,副教授,主要从事物联网应用、矿山安全方面的研究。E-mail:haoqinxia@xust.edu.cn。