

储海东,陈振宇,杜建光,等.面向变电站智能安监的行为识别与时空特征决策方法[J].辽宁工程技术大学学报(自然科学版),2026,45(2):242-248.DOI:10.11956/j.issn.1008-0562.20250442

CHU Haidong, CHEN Zhenyu, DU Jianguang, et al. Behavior recognition and spatiotemporal feature decision-making method for intelligent safety monitoring of substations[J]. Journal of Liaoning Technical University(Natural Science), 2026, 45(2): 242-248. DOI: 10.11956/j.issn.1008-0562.20250442

面向变电站智能安监的行为识别与时空特征决策方法

储海东¹, 陈振宇^{2*}, 杜建光², 闫华光^{2,3}, 陈毅⁴, 赵帅⁵

(1. 浙江大学电气工程学院, 浙江杭州 310012; 2. 国家电网有限公司信息通信中心(大数据中心), 北京 100031;

3. 中国电力科学研究院有限公司, 北京 100192; 4. 浙江大学海洋学院, 浙江舟山 316000;

5. 国网浙江省电力有限公司信息通信分公司, 浙江杭州 310012)

摘要:为解决传统电力人员行为识别方法特征提取粒度不足、难以适应变电站复杂视频监控场景等问题,面向电力运维需求开展行为识别技术研究。采用端到端的视频行为识别框架实现对原始监控视频的直接建模,并设计基于时空特征的关键帧提取方法提升推理效率;构建行为分类解码器,提高多类作业动作的判别能力。在真实变电站作业视频数据集上的实验结果表明,所提方法的综合识别率可达93.7%,在识别精度与处理速度方面均显著优于基于支持向量机(SVM)、多层感知机(MLP)等传统图像识别方法。研究结论为提升电力现场智能监控能力提供技术参考。

关键词:视频监控;行为识别;时空特征;关键帧提取;端到端;电力运维

中图分类号:TP277

文献标志码:A

文章编号:1008-0562(2026)02-0242-07

Behavior recognition and spatiotemporal feature decision-making method for intelligent safety monitoring of substations

CHU Haidong¹, CHEN Zhenyu^{2*}, DU Jianguang², YAN Huaguang^{2,3}, CHEN Yi⁴, ZHAO Shuai⁵

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310012, China; 2. State Grid Corporation of China Information & Telecommunication Center (Big Data Center), Beijing 100031, China; 3. China Electric Power Research Institute, Beijing 100192, China; 4. Ocean College, Zhejiang University, Zhoushan 316000, China; 5. Information and Communication Branch, State Grid Zhejiang Electric Power Company Limited, Hangzhou 310012, China)

Abstract: To address the issues of insufficient granularity in feature extraction and limited adaptability to complex video surveillance scenarios in substations inherent in traditional power personnel behavior recognition methods, this study investigates behavior recognition technology tailored to the needs of power operation and maintenance. An end-to-end video behavior recognition framework is adopted to directly model raw surveillance videos, and a key frame extraction method based on spatiotemporal features is designed to improve inference efficiency. A behavior classification decoder is constructed to enhance the discriminative ability for multiple types of operational actions. The experimental results on the real substation operation video dataset show that the proposed method achieves an overall recognition rate of 93.7%, significantly outperforming traditional image recognition methods such as support vector machine (SVM) and multi-layer perceptron (MLP) in both recognition accuracy and processing speed. The research conclusion provides a technical reference for improving the intelligent monitoring ability of power field.

收稿日期: 2025-10-18 修回日期: 2025-11-17 接受日期: 2025-11-25 责任编辑: 张雯

基金项目: 国家电网公司总部科技项目(5700-202490330A-2-1-ZX); 国家自然科学基金项目(62476242)

作者简介: 储海东(1978-), 男, 江苏南通人, 博士研究生, 主要从事智能电网与计算机视觉方面的研究。E-mail: 13639629785@126.com

通信作者: 陈振宇(1985-), 男, 湖南长沙人, 博士, 正高级工程师, 主要从事电力人工智能与大数据、智能电网等方面的研究。

E-mail: czy9907@gmail.com

Keywords: video surveillance; behavior recognition; spatiotemporal features; keyframe extraction; end-to-end; power operation and maintenance

0 引言

变电站作为电力系统的核心组成部分,承担着电压变换、电流分配等重要功能。变电站工作环境复杂,工作人员的操作行为直接影响设备运行的安全性。因工作人员违规操作或异常行为引发的安全事故屡见不鲜,这不仅威胁到操作人员的生命安全,也对电力系统的稳定运行构成严重威胁^[1-2]。

传统的电力人员行为识别方法主要依赖支持向量机(SVM)和多层感知机(MLP)等机器学习算法。这些方法通过对静态图像进行处理,提取颜色、纹理、形状等特征进行分类。然而,这些方法存在明显的局限性:一方面,静态图像无法捕捉行为的时序动态特征,导致细粒度行为识别能力不足;另一方面,传统方法在复杂场景(如光照变化、天气影响)下的识别精度显著下降,难以满足实际视频监控场景的需求^[3-4]。近年来,随着深度学习技术的发展,基于视频的行为识别方法逐渐成为研究热点。FANG等^[5]提出了一种基于改进YOLOv4的轻量级检测模型,通过引入MobileNetV3和高效通道注意力机制(ECA),显著提升了检测变电站人员异常行为的精度和实时性。CHEN等^[6]利用改进的Faster R-CNN算法,结合Retinex图像增强技术,实现了对安全帽佩戴检测的高精度识别,平均精度(mAP)达到94.3%。这些方法虽然在特定任务上取得了显著进展,但仍存在检测速度慢、实时性不足等问题,

且多集中于单一行为(如安全帽检测),难以全面覆盖变电站场景中的多样化行为^[2-6]。

针对上述问题,提出一种面向变电站视频监控场景的新型端到端行为识别算法,将视频特征提取、关键帧选择和行为分类有机结合。构建端到端的视频行为识别框架,直接处理视频流并输出行为识别结果,避免传统方法对静态图像逐帧处理的依赖^[3-4]。这种统一架构充分利用视频的时序信息,从整体上提升复杂场景下对人员细微行为的识别能力。引入基于时空特征的关键帧自适应提取策略,与固定间隔采样或简单帧差阈值等已有方法相比,智能选取具有代表性的帧以去除冗余数据,在保留关键行为信息的同时显著提高推理效率。构建自注意力行为分类解码器的识别模块,通过多层次特征提取与帧间注意力机制有效捕捉细粒度的动作差异,以增强对相似行为的辨别能力,提高复杂场景下行为识别的精度和鲁棒性。

1 面向变电站监控的行为识别算法

1.1 算法总体框架

面向变电站视频监控的行为识别算法总体框架见图1。该框架是一个端到端的视频行为识别系统,可直接处理视频流数据并输出行为识别结果。框架由2个核心模块组成,分别是基于时空特征的关键帧提取模块和行为分类解码器。各模块协同工作,实现从原始视频输入到行为分类输出的完整流程。

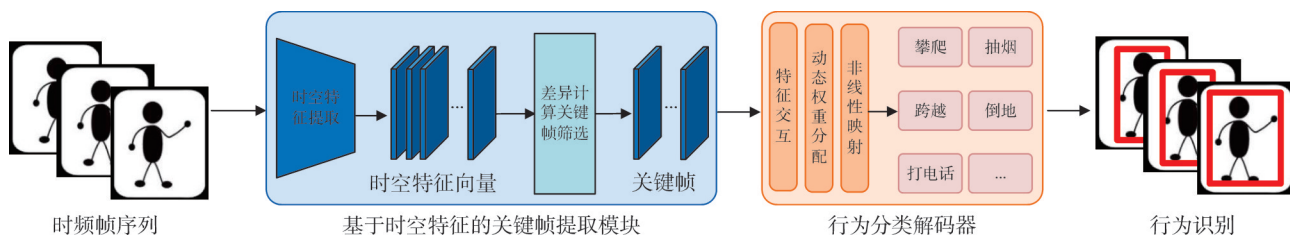


图1 算法总体框架

Fig.1 overall framework of the algorithm

1.2 基于时空特征的关键帧提取模块

关键帧提取是视频行为识别中的关键步骤,其目标是从视频流中提取最具代表性的帧,以减少冗余计算并保留关键行为信息。传统的关键帧提取方法多基于帧间差异或固定时间间隔采样,难以有效捕捉复杂场景下的行为动态特征。针对

这一问题,提出一种基于时空特征的关键帧提取方法,通过结合视频的时空信息,实现对关键帧的智能筛选。

(1) 时空特征提取

时空特征是描述视频中行为动态变化的重要信息。采用3D卷积神经网络(3D-CNN)^[7]提取视

频的时空特征, 结构见图2。3D-CNN通过在时间维度上扩展传统2D卷积核, 能够同时捕捉空间和时间维度上的特征变化。给定输入视频帧序列的3D-CNN卷积操作可表示为

$$F_{i,j,k} = \sum_{m=1}^M \sum_{n=1}^N \sum_{p=1}^P W_{m,n,p} \cdot V_{i+m-1,j+n-1,k+p-1} + b, \quad (1)$$

式中: $F_{i,j,k}$ 为输出特征图张量; $W_{m,n,p}$ 为3D卷积核权重; b 为3D卷积层的偏置向量; $M \times N \times P$ 为卷积核的尺寸; V 为视频3D卷积后的特征张量; i, j, k 为三维卷积索引。

通过多层3D卷积和池化操作, 最终得到视频的时空特征表示 $F = (f_1, f_2, \dots, f_T)$, 其中 f_t ($t = 1, 2, \dots, T$, T 为总帧数) 表示第 t 帧的时空特征向量。

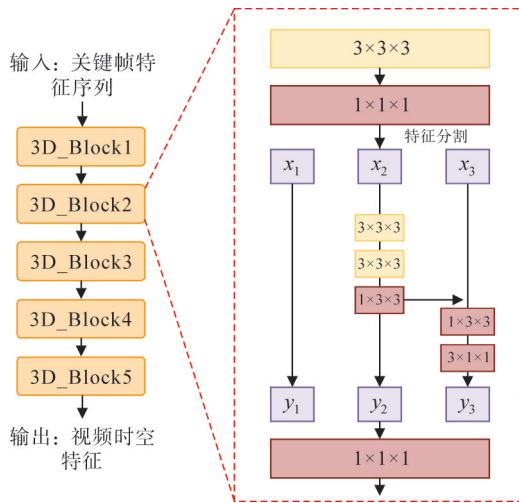


图2 时空特征编码器

Fig.2 temporal-spatial encoder

(2) 帧间特征差异计算

对于时空特征序列 F , 定义第 t 帧与第 $t-1$ 帧之间的特征差异为

$$D_t = \|f_t - f_{t-1}\|_2, \quad (2)$$

式中, $\|\cdot\|_2$ 为欧氏距离。

通过计算所有相邻帧的特征差异, 得到差异序列 $D = \{D_2, D_3, \dots, D_T\}$ 。

(3) 关键帧筛选

基于特征差异序列 D , 设计一种自适应阈值的关键帧筛选方法。差异序列的均值和标准差分别为

$$\mu_D = \frac{1}{T-1} \sum_{t=2}^T D_t, \quad (3)$$

$$\sigma_D = \sqrt{\frac{1}{T-1} \sum_{t=2}^T (D_t - \mu_D)^2}. \quad (4)$$

设定自适应阈值为

$$\tau = \mu_D + \alpha \cdot \sigma_D, \quad (5)$$

式中, α 为可调参数, $\alpha \in [0, 1]$, 用于控制关键帧的筛选严格度。

对于每一帧, 若其差异值 D_t 大于阈值 τ 则将其标记为候选关键帧。

(4) 冗余帧去除

为了进一步优化关键帧集合, 采用 K-means 聚类算法^[8]去除内容相似的冗余帧, 具体步骤如下。

步骤 1 将所有候选关键帧的特征向量 $\{f_{k_1}, f_{k_2}, \dots, f_{k_k}\}$ 作为输入, 其中 K 为候选关键帧数量。

步骤 2 使用 K-means 算法^[8]将特征向量聚类为 C 个簇, 其中 C 为预设的聚类中心数。

步骤 3 对于每个簇, 选择距离簇中心最近的帧为最终关键帧。

通过上述步骤, 最终得到优化后的关键帧集合 $K = \{k_1, k_2, \dots, k_C\}$, 其中 C 为预设的聚类中心数, 显著减少了视频数据量。

1.3 行为分类解码器

设计一种基于注意力机制的行为分类解码器 (behavior classification decoder), 通过多层次的特征交互与动态权重分配, 结合自注意力机制与分组解码策略, 提升电力人员行为识别的精度与鲁棒性。

(1) 解码器架构设计

解码器通过多层次的时空特征提取与动态权重分配, 实现对人体行为的精准分类。解码器结构见图3。

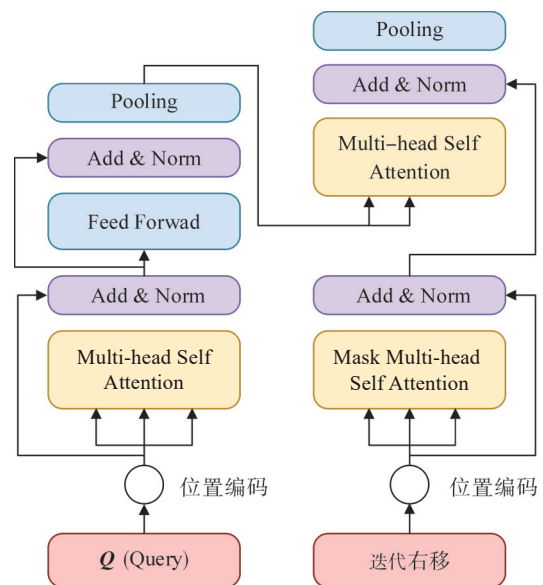


图3 解码器结构

Fig.3 structure of the decoder process

其输入包括两部分:一是由时空特征提取器获得的关键帧的时空特征,二是由可学习参数构成的行为类别查询向量 $\mathbf{Q}$ (Query)。解码器的处理流程分为4个主要阶段:特征交互、动态权重分配、非线性映射与输出生成。

①特征交互模块 (feature interaction module)

该模块通过自注意力机制,捕捉不同行为类别之间的潜在关联性。通过计算查询向量之间的相似度,动态分配权重,从而提取全局特征。其数学表达为

$$g_1(\mathbf{Q}, \mathbf{K}) = g_s \left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}} \right) \mathbf{Q}, \quad (6)$$

式中: $g_1(\mathbf{Q}, \mathbf{K})$ 为 $\mathbf{Q}$ 与 $\mathbf{K}$ 的特征交互操作; $\mathbf{K}$ 为关键向量; $g_s(\cdot)$ 为Softmax函数,用于归一化权重^[9]; d 为向量的维度。

②动态权重分配模块 (dynamic weight allocation module)

在特征交互的基础上,该模块进一步将查询向量与空间特征张量进行交互,动态分配权重以捕捉关键帧之间的空间关联性。其计算公式为

$$g_A(\mathbf{Q}, \mathbf{S}) = g_s \left(\frac{\mathbf{Q}\mathbf{S}^T}{\sqrt{d_f}} \right) \mathbf{S}, \quad (7)$$

式中: $\mathbf{S}$ 为空间特征张量; d_f 为特征维度。

③非线性映射模块 (nonlinear mapping module)

该模块通过多层感知机(MLP)^[10]对特征进行非线性变换,进一步增强特征的表达能力。其数学表达式为

$$g_M(\mathbf{x}) = g_{\text{ReLU}}(\mathbf{W}_1\mathbf{x} + b_1)\mathbf{W}_2 + b_2, \quad (8)$$

式中: $\mathbf{x}$ 表示经过MLP非线性映射后的高阶特征; $\mathbf{W}_1$ 、 $\mathbf{W}_2$ 为可学习权重矩阵; b_1 、 b_2 为偏置项; $g_{\text{ReLU}}(\cdot)$ 为激活函数^[11]。

④输出生成模块 (output generation module)

该模块通过池化操作将高维特征映射到行为类别的概率分布上,生成最终的分类输出。其计算式为

$$g_U(\mathbf{x}) = g_p(\mathbf{x}, d), \quad (9)$$

式中: d_f 为特征维度; $g_p(\mathbf{x}, d)$ 表示沿特征维度进行均值池化^[12]。

(2)多头注意力解码策略

为解决行为类别增多带来的计算复杂度问题,提出一种多头注意力解码策略。传统的解码器中,每个查询向量对应一个行为类别,随着类别数量的增加,计算复杂度 $O(n^2)$ 呈二次方增长。分组解

码策略将行为类别划分为若干组,每组共享一个查询向量,并通过仿射变换与池化操作生成多个类别的输出,从而将计算复杂度 $O(n)$ 降低为线性。

①分组查询 (grouped query)

将行为类别随机划分为若干组,每组包含 g 个类别。每组对应一个查询向量,查询向量的数量为 $K=N/g$,其中 N 为总类别数。

②仿射变换 (affine transformation)

通过仿射变换层,将每个查询向量扩展为 g 个类别得分(logits)。其数学表达式为

$$L_i = \mathbf{W}_h \mathbf{q}_h + b_h, \quad (10)$$

式中: $\mathbf{q}_h$ 为第 h 个查询向量; $\mathbf{W}_h$ 和 b_h 为可学习的权重矩阵和偏置项。

③池化操作 (pooling operation)

在仿射变换之后,沿特征维度进行池化操作,生成最终的分类输出。该操作的计算量为 $N \times d$ 次乘加运算。通过分组解码策略,行为分类解码器在保持高精度的同时,显著减少了计算开销,适用于大规模行为分类任务。

2 实验结果与分析

2.1 数据集

为确保模型在实际电力场景中的泛化能力与鲁棒性,构建高质量的数据集,用于电力人员行为识别任务。该数据集涵盖变电站等典型电力场景,体现了实际工作环境中的复杂性与多样性。

(1)数据集构建

数据采集工作在中国南方电网的多个变电站现场进行,使用高分辨率摄像头捕捉电力人员的日常工作行为。采集过程中,充分考虑了不同光照条件、天气变化以及设备遮挡等因素,以确保数据的多样性和代表性。

每张图像均经过人工标注,标注内容包括人体关键点(如头部、肩部、肘部、膝盖等)以及行为类别标签。标注工作由专业团队完成,并经过多轮校验,确保标注精度达到99%以上。

为了进一步提升数据集的多样性和模型的泛化能力,采用多种数据增强技术,包括随机翻转、亮度调整、对比度调整和饱和度调整等。

(2)数据集统计

私有数据集共包含2万条行为实例,涵盖了变电站场景中电力人员的多种典型行为,见表1。例子数表示在一段较长的视频中该行为出现的实例个数。

表 1 私有数据集样本数量
Tab.1 number of private data set samples

行为类别	例子数	视频总时长/h
攀爬 (climbing)	2 321	213
跨越 (crossing)	2 442	283
打电话 (calling)	2 121	509
倒地 (falling)	2 486	236
托举 (carrying)	2 776	247
抽烟 (smoking)	653	80
未佩戴安全帽 (no helmet)	4 321	599
未穿工装 (no uniform)	2 545	360

2.2 实验设备及参数

实验环境如下：计算平台配备 AMD EPYC 7742 处理器（搭载 1 TB 内存）和 4 块 NVIDIA Tesla V100 GPU（32 GB 显存），操作系统为 CentOS 7.6，深度学习框架基于 TensorFlow 2.5 构建。在模型预训练阶段，采用分阶段训练策略，先优化编码器网络，同时保持解码器参数不变。将输入图像尺寸调整为 416×320 像素，批量大小设为 32。优化器选用 RMSprop 算法^[13]，初始学习率设为 2×10^{-4} ，采用 ReduceLROnPlateau 学习率调度策略^[14]，当验证损失连续 5 个轮次未下降时，将学习率降至当前值的 20%。整个预训练过程持续 250 个轮次，并在训练过程中引入早停机制（patience 为 15）以防止过拟合。在模型微调阶段，采用参数冻结策略，固定编码器权重，仅对解码器进行端到端优化。该阶段将初始学习率设为 5×10^{-6} ，采用循环学习率（cyclic learning rate）调度方法^[15]，批量大小调整为 64，同时设置权重衰减为 5×10^{-7} 以控制模型复杂度。微调过程持续 50 个轮次，并采用梯度裁剪（gradient clipping）^[16]技术，将梯度范数限制在 1.0 以下，确保训练稳定性。此外，在微调阶段加入了标签平滑（label smoothing 为 0.1）和混合精度训练^[17]等技术，以提升模型性能。

2.3 实验结果

通过参数敏感性实验，评估可调参数 α 与 C 。其中， $\alpha\in[0,1]$ 为可调参数，用于控制关键帧的筛选严格度， C 为预设的聚类中心数。对模型识别准确率的影响，结果见表 2。当 $\alpha=0.5$ 、 $C=5$ 时，模型取得最优综合识别性能，综合识别率达到 93.7%。实验结果表明，随着超参数 α 的增大，综合识别率在 α 为 0.5~0.7 时达到最高。当 α 从 0.1 增至 0.5 时，识别率稳步上升； α 继续增大至 0.9 时识别率略有下降，说明过大或过小的 α 均不利于分类性能的提

升。随着参数 C 的增大，综合识别率在 C 为 2~4 时持续提升；当 $C\geq5$ 时性能基本趋于稳定或略微下降。表明适中的查询数量更有助于模型捕获关键信息，而过小或过大的 C 会影响识别效果。模型识别准确率呈现先上升后下降的趋势，表明在当前参数范围内存在使模型性能达到最优的平衡点。

表 2 参数敏感性实验结果
Tab.2 experimental results of parameter sensitivity

α	综合识别率/%				
	$C=2$	$C=3$	$C=4$	$C=5$	$C=6$
0.1	91.8	92.4	92.9	93.1	92.2
0.3	92.8	93.2	93.5	93.6	92.9
0.5	92.4	93.0	93.4	93.7	93.0
0.7	92.9	93.3	93.6	93.5	92.8
0.9	91.9	92.6	93.0	93.2	92.1

将本文提出的模型与传统机器学习算法 K 最近邻算法^[18]（K-nearest neighbors, KNN）、SVM^[19]和深度学习算法 MLP^[20]进行对比实验。本文模型对常见电力人员的 5 类行为识别准确率混合矩阵与 Top-1 精度见图 4 及表 3。

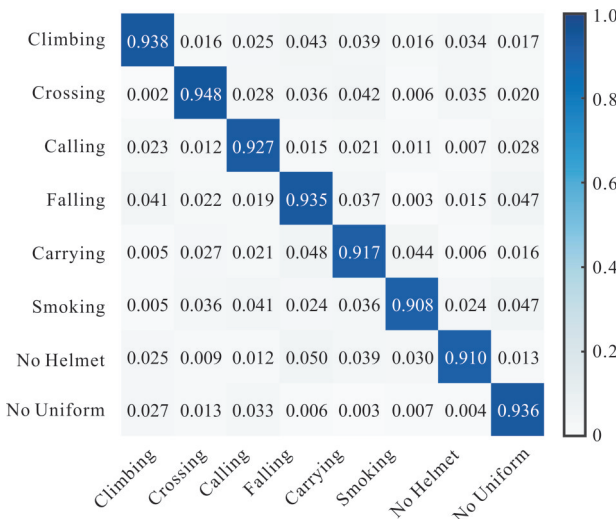


图 4 行为识别准确率混合矩阵

Fig.4 mixed matrix of behavior recognition accuracy

表 3 电力人员行为识别结果
Tab.3 power personnel behavior identification results

行为类别	准确率/%
攀爬(Climbing)	93.8
跨越(Crossing)	94.8
打电话(Calling)	92.7
倒地(Falling)	93.5
托举(Carrying)	91.7
抽烟(Smoking)	90.8
未佩戴安全帽(No Helmet)	91.0
未穿工装(No Uniform)	93.6

算法对比实验结果见表 4。从分类精度来看, 攀爬和跨越行为的辨识度较高, 因此最容易被准确识别; 对抽烟和未佩戴安全帽这类容易混淆的行为识别精度有所降低, 因为此类行为动作所占视觉区域过于狭小。但整体而言, 相对传统分类算法和深度学习算法, 本文所提方法行为识别准确率最高。

为了验证所提方法在行为分类任务中的优越性能, 进行了广泛的消融研究, 涉及不同的编码

器 (Encoder) 和解码器 (Decoder)。从表 4 的数据可以看出, 在视频处理阶段, 关键帧采样方法相较于均匀采样推理速度显著提升而精度不降低。此外, 基于 3D 卷积的时空特征编码器在准确性方面优于 OpenPose, 这一结果同样体现了本文方法在编码器设计上的优势。所提出的解码器在性能上也超越了传统的行为解码器, 进一步证明了新方法的有效性。

表 4 算法对比实验结果
Tab.4 compare experimental results of algorithm

方法	编码器	解码器	Top-1 精度/%	运行时间/ms
本文方法	OpenPose ^[21]	KNN	66.5	321
		SVM	75.2	521
		MLP	83.2	321
		本文解码器	89.5	698
Uniform sampling	OpenPose ^[21]	KNN	66.8	9 312
		SVM	75.3	15 920
		MLP	83.5	9 002
		本文解码器	89.9	23 421
本文方法	本文编码器	KNN	72.4	301
		SVM	86.1	509
		MLP	88.0	310
		本文解码器	93.7	612

对所提方法在视频动作识别领域的公开基准数据集 UCF-101 与 HMDB-51 上, 进行了系统评估, 并与当前主流方法进行充分对比, 实验结果见表 5。

表 5 公开数据集动作识别 Top-1 准确率对比
Tab.5 comparison of Top-1 accuracy for action recognition on public datasets

模型	Top-1 准确率/%	
	UCF-101	HMDB-51
Two-Stream	88.0	59.4
IDT	86.4	61.7
Dynamic Image Networks + IDT	89.1	65.2
TDD+IDT	91.5	65.9
Two-Stream Fusion + IDT	93.5	69.2
Deep Networks Sports 1M pre-training	65.2	69.6
Temporal Segment Networks	94.2	69.4
本文方法	95.2	77.9

UCF-101 包含 101 类人体动作, 共计约 1.3 万个视频片段。HMDB-51 涵盖 51 个动作类别, 约 7 000 个视频, 二者均为评估模型在复杂真实场景下泛化能力的重要基准。相较于经典的时序分段网络 (TSN), 本文方法在 UCF-101、HMDB-51 上 Top-1 准确率分别提升了 1.0 个百分点和 8.5 个百分点。本文方法也优于其他代表性基线, 如 Two-

Stream Fusion 结合改进稠密轨迹特征 (IDT) 的方法以及轨迹描述子特征 (TDD) 与 IDT 的融合方案。实验结果验证了关键帧筛选策略与行为分类解码器设计在应对复杂多变场景时具有更强泛化能力, 能够平衡模型精度与计算效率。

3 结论

- (1) 通过构建端到端的视频行为识别框架, 直接处理视频数据, 摆脱了对静态图像的依赖。设计了基于时空特征的关键帧提取方法, 显著提升了视频推理效率。提出了行为分类解码器, 有效提升了细粒度行为特征的捕捉能力。
- (2) 实验结果表明, 所提算法在变电站工作场景视频数据集上的综合识别率达到 93.7%, 相较于传统图像处理方法, 在识别精度和处理速度方面均实现了显著提升, 验证了所提算法在实际应用中的有效性和工程实用价值。未来将进一步探索该算法在更为复杂多样的电力作业场景中的应用, 并结合边缘计算技术, 搭建更高效、智能的行为识别系统。

参考文献(References):

[1] YANG J, LI C M, HONG D J, et al. Real-time safety behavior

- detection technology of indoors power personnel based on human key points[J]. Journal of Shanghai Jiaotong University (Science), 2024, 29(2): 309-315.
- [2] 王鸿, 邓元实, 常政威, 等. 基于深度学习的电力作业人员行为识别技术[J]. 四川电力技术, 2022, 45(3): 23-28.
WANG Hong, DENG Yuanshi, CHANG Zhengwei, et al. Behavior recognition technology of power operating person based on deep learning[J]. Sichuan Electric Power Technology, 2022, 45(3): 23-28.
- [3] 周震震, 宋云海, 何宇浩, 等. 基于分组查询注意力的可扩展电力人员行为分类方法[J]. 中国电力, 2023, 56(11): 77-85.
ZHOU Zhenzhen, SONG Yunhai, HE Yuhao, et al. Extensible classification method for power personnel behavior based on pose estimation[J]. Electric Power, 2023, 56(11): 77-85.
- [4] 闫云凤, 陈汐, 金浩远, 等. 基于计算机视觉的电力作业人员行为分析研究现状与展望[J]. 高电压技术, 2024, 50(5): 1842-1854.
YAN Yunfeng, CHEN Xi, JIN Haoyuan, et al. Research status and development of computer-vision-based power workers' behavior analysis[J]. High Voltage Engineering, 2024, 50(5): 1842-1854.
- [5] FANG J X, LI X W. Object detection related to irregular behaviors of substation personnel based on improved YOLOv4[J]. Applied Sciences, 2022, 12(9): 4301.
- [6] CHEN S B, TANG W H, JI T Y, et al. Detection of safety helmet wearing based on improved faster R-CNN[C]//2020 International Joint Conference on Neural Networks. July 19-24, 2020. Glasgow, United Kingdom. IEEE, 2020: 1-7.
- [7] ELDEMERY H I, PASHA H H, HAMID N A, et al. Temporal pedestrian behavior prediction using LSTM networks[C]//2025 7th Novel Intelligent and Leading Emerging Sciences Conference. October 25-27, 2025, Giza, Egypt. IEEE, 2025: 250-253.
- [8] IKOTUN A M, EZUGWU A E, ABUALIGAH L, et al. K-means clustering algorithms: a comprehensive review, variants analysis, and advances in the era of big data[J]. Information Sciences, 2023, 622: 178-210.
- [9] HU F. Survey on neural networks in natural language processing [C]//2023 IEEE International Conference on Control, Electronics and Computer Technology. April 28-30, 2023, Jilin, China. IEEE, 2023: 591-594.
- [10] WANJARI K, VERMA P. A review on the applications of machine learning and deep learning algorithms for image recognition[C]//2025 4th International Conference on Sentiment Analysis and Deep Learning. February 18-20, 2025, Bhimdatta, Nepal. IEEE, 2025: 1707-1711.
- [11] SINGH A, JANGID A K, SRINIVASU B. Advancing neural network performance with probabilistic computing for ReLU function[C]//2025 38th International Conference on VLSI Design and 2024 23rd International Conference on Embedded Systems. January 4-8, 2025, Bangalore, India. IEEE, 2025: 237-242.
- [12] GHOLAMALINEZHAD H, KHOSRAVI H. Pooling methods in deep neural networks, a review[EB/OL]. 2020: arXiv: 2009.07485. <https://arxiv.org/abs/2009.07485>
- [13] MURTHY A, RAO P S, PALLAVI N S, et al. Optimizing convolutional neural networks: a comparative study of gradient-descent, Adam, and RMSprop optimizers for accuracy and loss in apple leaf disease detection[C]//2024 Second International Conference on Networks, Multimedia and Information Technology. August 9-10, 2024, Bengaluru, India. IEEE, 2024: 1-6.
- [14] AL-KABABJI A, BENSAAFI F, DAKUA S P. Scheduling techniques for liver segmentation: ReduceLRonPlateau vs OneCycleLR[C]//Intelligent Systems and Pattern Recognition: Second International Conference, ISPR 2022, Recised Selected Papers. March 24-26, 2022, Hammamet, Tunisia. Springer International Publishing, 2022: 204-212.
- [15] LV P, ZHANG Y. Prediction of organic chemical reactions using cyclical learning rate based long-short term memory[C]//2023 3rd International Conference on Mobile Networks and Wireless Communications. December 4-5, 2023, Tumkur, India. IEEE, 2024: 1-4.
- [16] TAN Q C, CHENG T, LI J Z, et al. Differential privacy adaptive gradient clipping method based on hessian matrix[C]//2025 3rd International Conference on Big Data and Privacy Computing. May 30 - June 1, 2025, Fuzhou, China. IEEE, 2025: 96-102.
- [17] CHAUHAN A, TIWARI U, R V N. Post training mixed precision quantization of neural networks using first-order information[C]//2023 IEEE/CVF International Conference on Computer Vision Workshops. October 2-6, 2023, Paris, France. IEEE, 2023: 1335-1344.
- [18] MANZALI Y, BARRY K A, EL FAR M. An improved KNN algorithm based on ensemble methods and correlation[C]//2023 7th IEEE Congress on Information Science and Technology. December 16-22, 2023, Agadir - Essaouira, Morocco. IEEE, 2024: 64-70.
- [19] CAI Q. Human behavior recognition algorithm based on HOG feature and SVM classifier[C]//2019 IEEE 10th International Conference on Software Engineering and Service Science. October 18-20, 2019, Beijing, China. IEEE, 2020: 233-236.
- [20] TOLSTIKHIN I O, HOULSBY N, KOLESNIKOV A, et al. MLP-mixer: an all-MLP architecture for vision[C]//35th Conference on Neural Information Processing Systems. December 6, 2021, Sydney, Australia. NeurIPS, 2021: 24261-72.
- [21] CAO Z, SIMON T, WEI S H, et al. Realtime multi-person 2D pose estimation using part affinity fields[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 1302-1310.