

文章编号: 1671-251X(2023)01-0090-09

DOI: 10.13272/j.issn.1671-251x.2022030098

基于改进时空图卷积神经网络的钻杆计数方法

杜京义¹, 党梦珂¹, 乔磊², 魏美婷³, 郝乐⁴

(1. 西安科技大学 电气与控制工程学院, 陕西 西安 710054;
2. 西安热工研究院有限公司, 陕西 西安 710054; 3. 西安工程大学 管理学院, 陕西 西安 710048;
4. 西安科技大学 安全科学与工程学院, 陕西 西安 710054)

摘要: 针对现有钻杆计数方法存在劳动重复、计数误差较大、未考虑动作的时序信息等问题, 提出了一种基于改进时空图卷积神经网络(MST-GCN)模型的钻杆计数方法。首先, 通过矿用监控摄像头获取井下打钻视频数据, 采用 Alphapose 算法在图像序列中提取人体的关键点信息, 得到单帧图像上的人体骨架和连续图像序列上的骨架序列数据, 进而构建表征人体动作的骨架序列; 然后, 在时空图卷积神经网络(ST-GCN)模型的基础上设计了 MST-GCN 模型, 采用远空间分区策略关注骨架上距离较远的关键点运动信息, 通过注意力机制网络 SENet 融合原空间特征与远空间特征, 从而有效识别骨架序列上的动作类别; 最后, 在打钻视频上利用支持向量机辨识打钻姿势来决定是否保存骨架序列, 若骨架序列长度保存到 150 帧则使用 MST-GCN 模型识别动作类别, 并根据实际打钻时间设置相邻动作的识别间隔, 从而记录动作数量, 实现钻杆计数。实验结果表明: 在自建的数据集上, MST-GCN 模型的识别准确率为 91.1%, 比 ST-GCN、Alphapose-LSTM 和 NST-GCN 动作识别模型的准确率分别提升了 6.2%, 19.0% 和 4.8%, 模型的损失值收敛在 0.2 以下, 学习能力更强。在相同条件的打钻视频上, MST-GCN 模型、人工方法和改进 ResNet 方法的平均误差依次为 0.25, 0.75, 21 次, 说明 MST-GCN 模型的计数效果更好。在真实打钻 1 300 根的现场应用中, MST-GCN 模型的平均误差为 9 根, 误计数结果低, 能够满足实际需求。

关键词: 瓦斯抽采; 钻孔; 钻杆计数; 打钻姿势; 人体骨架; 动作识别; MST-GCN 模型

中图分类号: TD713 文献标志码: A

Drill pipe counting method based on improved spatial-temporal graph convolution neural network

DU Jingyi¹, DANG Mengke¹, QIAO Lei², WEI Meiting³, HAO Le⁴

(1. College of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China;
2. Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710054, China; 3. College of Management,
Xi'an Polytechnic University, Xi'an 710048, China; 4. College of Safety Science and Engineering,
Xi'an University of Science and Technology, Xi'an 710054, China)

Abstract: There are some problems in the existing drill pipe counting methods, such as repeated labor, large counting error, and failure to consider the timing information of actions. In order to solve the above problems, a drill pipe counting method based on an improved multi spatial-temporal graph convolution neural network (MST-GCN) model is proposed. Firstly, the video data of underground drilling is obtained through the mine monitoring camera. The Alphabose algorithm is used to extract the key points of the human body from the image sequence. The human skeleton on a single frame image and the skeleton sequence data on a continuous image sequence are

收稿日期: 2022-03-29; 修回日期: 2022-12-28; 责任编辑: 张强。

基金项目: 陕西省科技厅自然科学基金项目(2018JQ5197); 陕西省重点研发计划项目(2019GY-097)。

作者简介: 杜京义(1965—), 男, 山东淄博人, 教授, 硕士研究生导师, 主要研究方向为检测技术及其自动化, E-mail: 517571853@qq.com。通信作者: 党梦珂(1998—), 男, 陕西武功人, 硕士研究生, 主要研究方向为目标检测与动作识别, E-mail: 2447439418@qq.com。

引用格式: 杜京义, 党梦珂, 乔磊, 等. 基于改进时空图卷积神经网络的钻杆计数方法[J]. 工矿自动化, 2023, 49(1): 90-98.

DU Jingyi, DANG Mengke, QIAO Lei, et al. Drill pipe counting method based on improved spatial-temporal graph convolution neural network[J]. Journal of Mine Automation, 2023, 49(1): 90-98.



扫码移动阅读

obtained. The skeleton sequence representing human actions is built. Secondly, the MST-GCN model is designed based on the spatial-temporal graph convolution neural network (ST-GCN) model. The far space partition strategy is used to focus on the motion information of the key points that are far away from the skeleton. The squeeze and excitation network (SENet) is used to fuse the original space features and the far space features, so as to effectively identify the action categories on the skeleton sequence. Finally, support vector machine is used to identify the drilling pose on the drilling video to decide whether to save the skeleton sequence. If the sequence length is saved to 150 frames, the MST-GCN model is used to identify the action category. The identification interval of adjacent actions is set according to the actual drilling time, so as to record the number of actions and realize the drill pipe counting. The experimental results show that the recognition accuracy of the MST-GCN model is 91.1% on the self-built data set, which is 6.2%, 19.0% and 4.8% higher than that of ST-GCN, Alphapose-LSTM and NST-GCN, respectively. The loss value of the MST-GCN model converges below 0.2, and the learning capability is stronger. On the drilling videos under the same conditions, the average error values of the MST-GCN model, the artificial method and the improved ResNet method are 0.25, 0.75 and 21 respectively, which shows that the counting effect of the MST-GCN model is better. The average error of MST-GCN model is 9 and the miscount is low in the field application of drilling 1 300 pieces, which can meet the actual requirements.

Key words: gas extraction; drilling; drill pipe counting; drilling pose; human skeleton; action recognition; MST-GCN model

0 引言

深层煤炭开采过程中存在高瓦斯压力不可控因素,极易诱发煤与瓦斯突出事故,从而对煤矿安全生产造成威胁^[1-2]。一般采用钻孔抽采和排放煤层内的瓦斯,以降低瓦斯浓度,减少甚至消除煤与瓦斯突出灾害^[3-4]。钻孔深度的准确测量是钻孔工作的重要部分,若实际钻孔深度与要求不符,不仅会加剧危险事故的发生,而且会造成人力物力的极大浪费^[5]。实际测量中,通常根据钻机每次打入钻杆长度一致的原则,根据钻杆数量得到钻孔深度。

钻杆计数方法根据计数方式的不同,可以分为传统方法与计算机视觉方法。传统方法主要有人员查验法、行程开关法、电驻波法等。人员查验法是指工作人员通过观看打钻录像来记录钻杆数量的方法,其劳动强度大、效率低。行程开关法^[6]主要是在钻机上安装行程开关和单片机,当行程开关被钻机触发时,单片机会收到电信号进行判断处理,进而得到钻杆数量,但是行程开关容易被误触,单片机可能会受潮而影响计数精度。电驻波法^[7]通过钻杆形成的驻波找到驻波节点,得到钻杆长度,但是有关仪器还未达到矿用的安全要求。计算机视觉方法是指利用视觉图像算法自动分析打钻视频中钻杆数量的方法,该方法因人为干预少、效率高等受到众多研究者的关注。董立红等^[8]提出了一种基于改进 Camshift 算法的钻杆计数方法,通过 Camshift 算法跟踪钻机标志物,并根据标志物运动轨迹上的波峰数量判断出

钻杆数量,但是标志物容易被粉尘污染及被旁物遮挡,进而导致目标丢失甚至跟踪失败。高瑞等^[9]提出了一种基于改进 ResNet 网络的井下钻杆计数方法,通过 ResNet 识别图像上的卸杆与非卸杆动作,并对识别结果进行积分滤波得到置信度曲线,根据下降沿的个数判断出钻杆数量,但是 ResNet 未考虑动作的时序信息,在相似背景的图像上容易发生误识别。党伟超等^[10]提出了一种利用三维卷积神经网络(3D Convolutional Neural Network, 3DCNN)识别钻机卸杆动作的方法,虽然 3DCNN 解决了一些时序问题,但同时也学习了大量的图像冗余信息,导致该方法的场景适应性不强。

针对上述问题,本文提出了一种基于改进时空图卷积神经网络的钻杆计数方法。该方法将打钻过程中作业人员的动作划分为上杆、卸杆和其他 3 种,在时空图卷积神经网络(Spatial-Temporal Graph Convolution Neural Network, ST-GCN)模型^[11]的基础上,采用远空间分区策略与注意力机制设计了一种动作识别精度良好的改进时空图卷积神经网络模型——多空间特征融合时空图卷积神经网络(Multi Spatial-Temporal Graph Convolution Neural Network, MST-GCN)模型,利用 MST-GCN 模型识别出打钻视频中的上杆与卸杆动作,确定打钻数量。

1 打钻过程分析

钻机的打钻过程是一个双向过程,即工人先将多根钻杆依次装上钻机并打入煤层,待打孔完成,再

将钻机拔出的钻杆依次卸下,根据工人上杆和卸杆动作的次数来判断钻杆数量。然而,该过程中还包含与上杆、卸杆无关的动作。因此,本文将所有的动作类型划分为上杆、卸杆和其他 3 种,具体如图 1 所示。



图 1 打钻过程中的 3 种动作

Fig. 1 Three kinds of action during drilling

2 钻杆计数方法原理

基于改进时空图卷积神经网络的钻杆计数方法原理如图 2 所示。



图 2 基于改进时空图卷积神经网络的钻杆计数方法原理

Fig. 2 Principle of drill pipe counting method based on improved multi spatial-temporal graph convolution neural network

(1) 打钻视频采集。通过矿用监控摄像头获取并下打钻视频数据。

(2) 人体骨架数据获取。为了减少图像背景对动作识别的干扰,本文采用高精度人体姿态估计算法——Alphapose 算法^[12]检测工人的身体关键点信息,得到单帧图像上的人体骨架和连续图像序列上的骨架序列数据。

(3) 动作识别。在 ST-GCN 模型的基础上,通过远空间分区策略与注意力机制设计了 MST-GCN 模

型,利用 MST-GCN 模型有效识别骨架序列上的动作类型,提高打钻动作的识别精度。

(4) 动作开始位置捕获。通过支持向量机(Support Vector Machine, SVM)算法辨识工人的打钻姿势,若判断成功则保存 150 帧人体骨架序列,并输入 MST-GCN 模型进行识别。

(5) 打钻数量计算。为减少模型出错的概率,本文将打钻视频手动分割成上杆视频和卸杆视频,并在每种视频上只识别各自对应的动作。在卸杆视频中,MST-GCN 模型每次都从打钻姿势出现的位置预测 1 次动作,若结果为卸杆动作,则卸杆数量加 1,并且间隔一段时间后再重新识别。上杆数量的计算同理。最后对上杆数量和卸杆数量求和,取平均值得到钻杆数量。

3 钻杆计数方法具体实现

3.1 人体骨架数据获取

人体关键点即人体拓扑上的重要骨骼部位,如图 3(a)中的 18 个关键点:鼻、颈、右肩、右肘、右腕、左肩、左肘、左腕、右臀、右膝、右脚踝、左臀、左膝、左脚踝、右眼、左眼及左耳(编号依次为 0, 1, …, 17)。人体骨架是指将 Alphapose 算法检测出的 18 个关键点按图 3(a)中的方式进行连接。人体骨架序列是指在时间轴上将多帧骨架按照相同编号的关键点进行连接的数据,结构如图 3(b)所示。

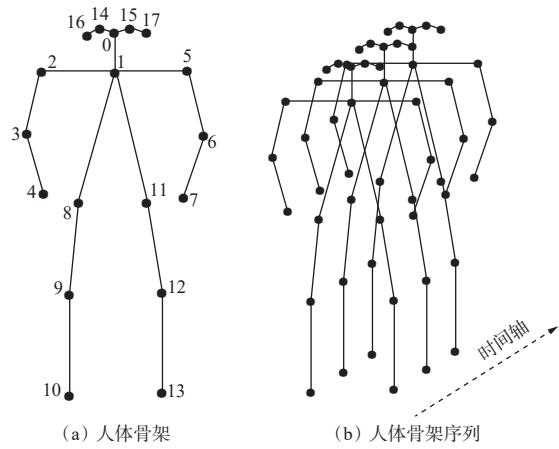


图 3 人体骨架数据

Fig. 3 Skeleton data of human body

3.2 动作识别

3.2.1 MST-GCN 模型

与 ResNet、3DCNN 等模型不同,ST-GCN 模型是一种用于骨架动作识别的序列模型,它利用图卷积神经网络(Graph Convolutional Neural Network, GCN)与时间卷积神经网络(Temporal Convolutional

Neural Network, TCN)在骨架序列上自动提取动作特征,可以减少图像背景对动作识别的影响。但是,ST-GCN模型提取的空间特征仍不够健壮,导致动作识别效果不佳。为此,本文在ST-GCN模型的基础上重新设计了一种MST-GCN模型,其原理如图4所示。它由批量标准化层(BatchNormal)、时空特征提取单元层(总共10个单元,每个单元的时空图卷积核大小均为(15, 3))、平均池化层(av-pool)、全连接层(fc)及Softmax函数组成。其中,BatchNormal层用于规范骨架序列数据的分布,提高模型的泛化能力;时空特征提取单元层在时空维度上提取有效的动作特征,以提高模型精度;av-pool层用于对高维动作特征进行降维,降低模型计算量;fc层与Softmax函数先将特征映射到标签空间,再输出动作的置信度向量。图中(M, C, T, V)为特征图, M 为人数, C 为关键点维数, T 为序列长度, V 为关键点个数。

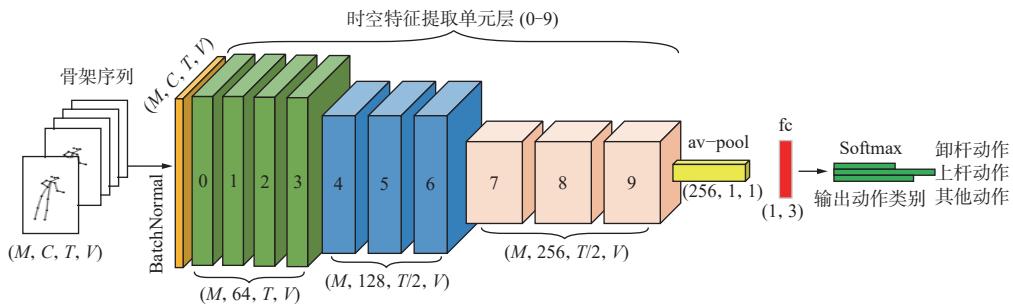


图4 MST-GCN模型原理

Fig. 4 Principle of multi spatial-temporal graph convolution neural network model

3.2.2 空间分区策略

GCN在骨架上提取空间特征时,需要使用分区策略来对关键点的邻域集进行划分。空间分区策略是一种构建GCN的有效策略,其依据身体部位的静止、向心运动和离心运动3种状态将关键点的邻域集划分为3个子集:根节点集、向心点集和离心点集。图5为ST-GCN中采用的空间分区策略(原空间分区策略),其邻域采样范围 $D=1$,“ \times ”为身体重心,“0”为根节点,“1”为向心点,“2”为离心点。

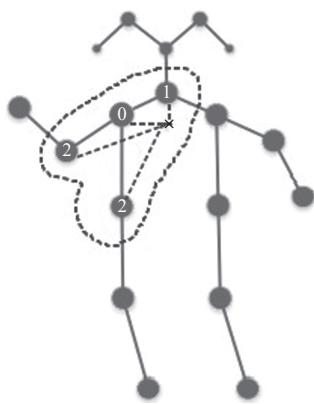


图5 原空间分区策略

Fig. 5 Original spatial partitioning strategy

当GCN采用原空间分区策略时,其3个子集分别被赋予以下权重参数:

$$l_{ti}(v_{tj}) = \begin{cases} 0 & r_j = r_i \\ 1 & r_j < r_i \\ 2 & r_j > r_i \end{cases} \quad (1)$$

式中: $l_{ti}(v_{tj})$ 为全部子集的权重映射函数, $t = 1, 2, \dots$,

接层(fc)及Softmax函数组成。其中,BatchNormal层用于规范骨架序列数据的分布,提高模型的泛化能力;时空特征提取单元层在时空维度上提取有效的动作特征,以提高模型精度;av-pool层用于对高维动作特征进行降维,降低模型计算量;fc层与Softmax函数先将特征映射到标签空间,再输出动作的置信度向量。图中(M, C, T, V)为特征图, M 为人数, C 为关键点维数, T 为序列长度, V 为关键点个数。

$T, i = 0, 1, \dots, 17, v_{tj}$ 为第 t 帧第 $j(j=0, 1, 2)$ 个子集的关键点; r_j 为 v_{tj} 到重心的距离; r_i 为根节点到骨架重心的平均距离。

GCN的实现公式如下:

$$\begin{cases} f_{\text{out}} = \sum_j \Lambda_j^{-\frac{1}{2}} \mathbf{A}_j \Lambda_j^{-\frac{1}{2}} f_{\text{in}} \mathbf{W}_j \\ \sum_j \mathbf{A}_j = \mathbf{A}_0 + \mathbf{A}_1 + \mathbf{A}_2 \end{cases} \quad (2)$$

式中: f_{out} 为GCN的输出结果; $\sum_j \Lambda_j^{-\frac{1}{2}} \mathbf{A}_j \Lambda_j^{-\frac{1}{2}}$ 为归一化的邻接矩阵; Λ_j 为第 j 个子集的对角线矩阵; \mathbf{A}_j 为第 j 个子集的邻接矩阵; f_{in} 为输入特征数据; \mathbf{W}_j 为第 j 个子集的权重矩阵; $\sum_j \mathbf{A}_j$ 为整体邻接矩阵; $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2$ 分别为每个子集的邻接矩阵。

实际打钻过程中,由于工人的身体形态、摄像机角度等原因,会导致同一种作业动作呈现出不同的运动幅度,而采用原空间分区策略只能关注到身体上关键点之间的局部运动,不利于GCN提取空间特征。为了更好地提取空间特征,即关注到身体上较远关键点($D>1$ 时的相邻关键点)之间的运动信息,本文将原空间分区策略中的 D 设置为2,并且将权重参数之间的差值进一步增大,从而得到图6所示的远空间分区策略。此时,3个子集的权重参数被重新定义如下:

$$l'_{ti}(v_{tj}) = \begin{cases} 0 & r_j = r_i \\ 1 & r_j < r_i \\ 4 & r_j > r_i \end{cases} \quad (3)$$

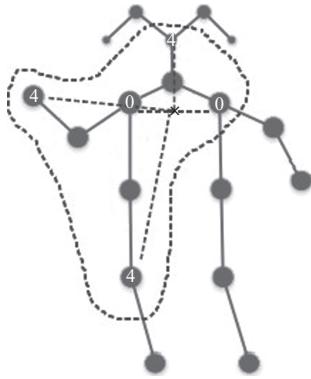


图 6 远空间分区策略

Fig. 6 Remote spatial partitioning strategy

所以, GCN 实现公式更新为

$$\begin{cases} f'_{\text{out}} = \sum_j A_j^{-\frac{1}{2}} A'_j A_j^{-\frac{1}{2}} f_{\text{in}} W_j \\ \sum_j A'_j = A_0 + A'_1 + A'_2 \end{cases} \quad (4)$$

式中: f'_{out} 为采用远空间分区策略的 GCN 输出; A'_j 为第 j 个子集更新后的邻接矩阵; W_j 为第 j 个子集更新后的权重矩阵; $\sum_j A'_j$ 为更新后的整体邻接矩阵; A'_1 , A'_2 分别为向心点集和离心点集更新后的邻接矩阵。

3.2.3 多空间特征融合

SENet(Squeeze and Excitation Networks)^[13] 是一种特征通道注意力机制网络, 它在训练过程中可以使网络自动选择重要的特征。因此, 本文在 ST-GCN 中添加 SENet 模块来融合不同分区策略的空间特征。该模块融合空间特征的原理如图 7 所示, 其中 X_1 为 GCN 采用原空间分区策略的空间特征, X_2 为 GCN 采用远空间分区策略的空间特征。首先将 2 个大小为 (M, C, T, V) 的特征图 X_1 , X_2 通过堆叠模块(CAT)堆叠为 $(M, 2C, T, V)$ 的新特征图; 然后将新特征图依次经过平均池化层、全连接层 1 和全连接层 2 压缩成 $(M, 2C, 1, 1)$, 并与新特征图按通道相乘; 最后使用一个二维卷积将分配权重后的特征图压缩成与输入大小一致的输出特征图。图中 ReLu, Sigmoid 为激活函数。

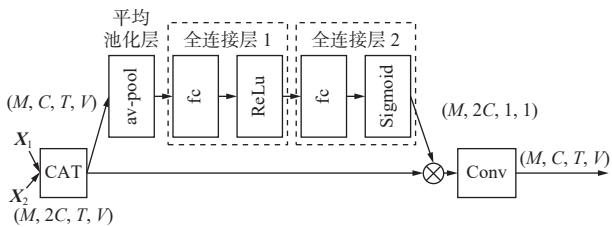


图 7 SENet 模块融合空间特征

Fig. 7 Fusion spatial features of the SENet module

3.2.4 时空特征提取单元

时空特征提取单元是 MST-GCN 模型中的基本

单元, 它在时间轴上将空间特征进行聚合, 从而得到表征动作的时空特征, 其详细结构如图 8 所示。首先将 2 种空间分区策略下提取的空间特征使用 SENet 模块进行特征融合; 其次使用 TCN 提取时间特征; 最后使用残差结构保留原始特征信息, 实现时空特征提取。图中 \odot 为矩阵按位相与, X 为输入特征图, Y 为输出特征图。

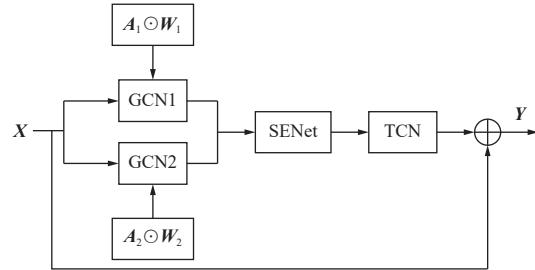


图 8 时空特征提取单元结构

Fig. 8 Structure of spatio-temporal feature extraction unit

3.3 动作开始位置捕获

3.3.1 捕获原理

工人准备取放钻杆时, 身体会呈现出手持钻杆的姿势, 即打钻姿势(图 9)。通过打钻姿势确定上杆、卸杆动作在视频中的开始位置, 不仅可以避免 MST-GCN 模型遍历视频的重复工作, 而且能提高计数效率。因此, 本文通过区分打钻姿势与非打钻姿势来判断工人是否将要进行上杆或卸杆动作。



(a) 打钻姿势



(b) 非打钻姿势

图 9 打钻姿势与非打钻姿势

Fig. 9 Drilling posture and non-drilling posture

3.3.2 打钻姿势识别

辨识人体骨架上所表达的姿势类别信息, 可以通过卷积神经网络在伪图像上自动提取特征实现分类, 也可以通过分类算法对手工特征进行分类^[14]。SVM 是一种快速且有效的监督式二分类算法, 可对人体骨架上由关节夹角组成的特征向量进行分类,

即能够辨识打钻姿势。

辨识打钻姿势的流程如图 10 所示。首先通过余弦公式计算人体骨架上的 10 个关节夹角: θ_0 (右颈角)、 θ_1 (左颈角)、 θ_2 (右肩角)、 θ_3 (左肩角)、 θ_4 (右肘角)、 θ_5 (左肘角)、 θ_6 (右跨角)、 θ_7 (左跨角)、 θ_8 (右膝角)和 θ_9 (左膝角);然后将 10 个夹角数据组成 1 个特征向量(即关节夹角余弦特征),并使用标准化公式对其进行规范;最后将特征向量送入训练好的 SVM 模型进行分类。



图 10 打钻姿势识别流程

Fig. 10 Drill pose recognition process

4 实验与结果分析

为了验证本文方法的可行性与实际效果,对 MST-GCN 模型与钻杆计数进行了实验。实验基于 64 位 Windows 10 操作系统进行,硬件配置:处理器为 Intel(R) Xeon(R) E5-1680 v4,内存为 32 GB,显卡为 NVIDIA TITAN XP 12 GB;软件环境:图形处理加速器为 CUDA10.1 和 cudnn7.6.5,编程语言为 Python3.8,深度学习框架为 Pytorch1.7.1。

4.1 数据集制作

为训练和评估 MST-GCN 模型,制作了 Drill-action 数据集。原始数据来自于河南某煤矿的机巷里段与集中回风巷,该位置画面整体清晰,且包含多段钻机工作视频。使用视频剪辑软件从原始视频中裁剪出 825 个动作,并统一调整成分辨率为 640×480 、帧率为 25 帧/s、时长为 8~10 s 的短视频序列。其中,上杆和卸杆动作视频各 300 个,剩余为其他动作视频。划分数据集时,随机选出 660 个动作作为训练集,剩余为测试集。训练和评估模型前,先采用 Alphapose 算法在动作视频上检测出全部的人体骨架信息,再将其设置成大小为(2, 3, 150, 18)的骨架序列数据。其中,“2”为每帧图像上最多保存 2 个人体平均置信度排序靠前的骨架,“3”为每个关键点维数,“150”为骨架序列的保存长度,18 为关键点保存个数。人体平均置信度为

$$c = \frac{\sum_{i=0}^{17} c_i}{18} \quad (5)$$

式中 c_i 为第 i 个关键点的置信度。

4.2 MST-GCN 模型实验

4.2.1 模型训练

采用随机梯度下降法(Stochastic Gradient Descent,

SGD)作为训练优化器,并设置动量因子为 0.9,权重衰减系数为 0.0001,Batchsize 为 32,模型迭代次数为 100,初始学习率为 0.01,在第 40,60,80 次迭代时将学习率依次衰减 0.1 倍。MST-GCN 与 ST-GCN 模型的训练损失曲线如图 11 所示。

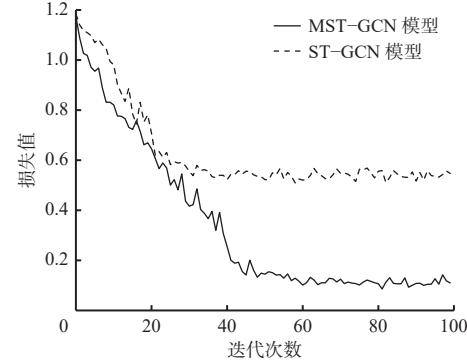


图 11 MST-GCN 与 ST-GCN 模型的训练损失曲线

Fig. 11 Training loss curves of MST-GCN model and ST-GCN model

由图 11 可看出,在第 28 次迭代时,ST-GCN 模型的损失值开始收敛,但 MST-GCN 模型仍在下降,并且最终损失值稳定在 0.2 以下,说明 MST-GCN 模型具有更强的学习能力。

4.2.2 模型评估

为验证对 ST-GCN 模型改进的有效性,在 Drill-action 数据集上进行了消融实验,并采用 3 种类别动作的整体准确率作为评估指标,结果见表 1。表中 ADD 为常见的特征相加融合方法。

表 1 消融实验结果比较

Table 1 Comparison of ablation experiment results

基准模型	编号	分区策略		特征融合		准确率/%
		原空间分区策略	远空间分区策略	ADD	SENet	
ST-GCN	1	√	✗	✗	✗	84.9
	2	✗	√	✗	✗	83.1
	3	√	√	√	✗	87.9
	4	√	√	✗	√	91.1

比较表 2 中“1”,“2”和“3”的结果可知,同时采用 2 种空间分区策略的准确率最高,说明远空间分区策略可以改善模型仅使用原空间分区策略的特征提取不足问题;比较“3”和“4”结果可知,采用 SENet 进行特征融合可进一步提高模型的准确率。

为进一步证明 MST-GCN 模型的识别能力,将文献[15-16]中的模型与之对比,结果见表 2。可看出 MST-GCN 模型较 Alphapose-LSTM(Long Short Term Memory, 长短期记忆)和 NST-GCN(新分区策略的 ST-GCN)模型的准确率分别提高了 19.0% 和 4.8%。

表 2 不同模型的动作识别结果

Table 2 Action recognition results of different models

模型方法	准确率/%
Alphapose-LSTM ^[15]	72.1
NST-GCN ^[16]	86.3
MST-GCN	91.1

4.3 钻杆计数实验

4.3.1 计数方法比较

为说明本文方法采用 MST-GCN 模型的计数效果更好, 在相同条件下的打钻视频(共 2 段打钻视频, 真实钻杆数量分别为 80 根和 100 根, 且上杆、卸杆次数与钻杆数量保持一致)上对 3 种计数方法进行比较, 结果见表 3。表中人工方法是指实验室成员通过观看录像进行计数, “0”表示该方法未实现此功能。平均误差是指 2 段视频上的平均计数误差, 如人工方法平均计数误差为 $[(80+80)\times 0.5-80]+[(102+101)\times 0.5-100]\times 0.5$ 次。

表 3 3 种计数方法的动作识别结果

Table 3 Action recognition results of three counting methods

打钻 视频	真实钻 杆/根	人工方法		改进ResNet ^[9]		本文方法	
		上杆/次	卸杆/次	上杆/次	卸杆/次	上杆/次	卸杆/次
1	80	80	80	0	88	79	80
2	100	102	101	0	134	98	102
平均误差		0.75		21		0.25	

从表 3 可看出, 3 种计数方法与真实数量的平均误差依次为 0.75, 21, 0.25 次, 显然本文方法的误差最小。经分析可知, 人工方法因为实验室成员出现了注意力不集中的情况, 从而造成误计数。改进 ResNet 方法因为只能识别出卸杆动作的某个简单状态, 时序信息太少, 所以误计数比较严重。

4.3.2 现场计数测试

为验证本文方法在煤矿井下的应用效果, 在河南某煤矿搭建了实验平台进行测试。现场采集 2 台钻机的打钻录像, 视频画面分辨率为 1280×720, 目标分辨率为 150×300~500×500, 目标之间的遮挡堆叠面积低于目标 15%, 具体如图 12 所示。总共涉及 18 处钻孔, 共 1300 根钻杆, 钻孔编号 1—9 为 1 号钻机的施工位置, 编号 10—18 为 5 号钻机的施工位置。

计算某钻孔的钻杆数量 H_n (n 为钻孔个数, $1 \leq n \leq 18$)时, 根据以下步骤实现:

(1) 根据台账记录的上杆、卸杆时间范围, 从录像机中取出关于该钻孔的上杆、卸杆视频。

(2) 在上杆视频中使用本文方法只识别上杆动作, 并根据经验设置 2 次上杆动作的识别间隔为

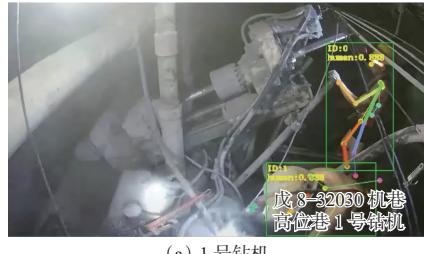


图 12 现场测试打钻画面

Fig. 12 Field test drilling screen

60 s, 打钻姿势阈值为 0.4, 上杆动作的判断阈值为 0.4, 上杆次数为 U_n 。同理在卸杆视频中只识别卸杆动作, 并根据经验设置 2 次卸杆动作的识别间隔为 8 s, 打钻姿势阈值为 0.4, 卸杆动作的判断阈值为 0.45, 卸杆次数为 D_n 。

(3) 待上杆和卸杆视频均被识别完后, 计算该钻孔的钻杆数量。

$$H_n = \frac{U_n + D_n}{2} \quad (6)$$

所有钻孔的钻杆数量 H 为

$$H = \left\lceil \sum_{n=1}^{18} H_n \right\rceil \quad (7)$$

式中“ \lceil ”为计算结果向上取整。

经过上述步骤, 计算出 18 个钻孔的所有钻杆数量, 结果见表 4。比较表 4 中最后一行数据可知, 单独以上杆、卸杆次数作为钻杆数量时, 它们与真实钻杆数量的误差分别为 46 根和 -28 根。其中上杆数量比真实钻杆数量偏大, 主要是因为上杆过程中的动作类型比较复杂, 模型误识别为上杆动作的概率大; 卸杆数量比真实钻杆数量偏小, 主要是因为存在多根钻杆连在一起被卸下的情况, 而本文方法只能识别为 1 个卸杆动作。不过, 根据二者平均值计算的钻杆数量与真实数量的误差为 9 根, 一定程度上降低了误差, 结果满足实际应用需求。

5 结论

(1) 提出了一种基于改进时空图卷积神经网络的钻杆计数方法。采用 Alphapose 算法在图像序列上得到表征人体动作的骨架序列数据; 在 ST-GCN

表4 全部钻孔的实验结果

Table 4 The results of all drilling experiments

钻孔编号	真实钻杆/根	识别结果		
		上杆/次	卸杆/次	均值/次
1	40	40	38	39
2	45	45	45	45
3	43	43	43	43
4	53	53	53	53
5	47	50	43	46.5
6	56	54	54	54
7	46	46	45	45.5
8	106	119	100	109.5
9	106	113	99	106
10	80	79	80	79.5
11	80	82	79	80.5
12	80	84	80	82
13	80	83	81	82
14	80	82	80	81
15	80	83	79	81
16	80	78	75	76.5
17	98	105	102	103.5
18	100	107	96	101.5
总计	1 300	1 346	1 272	1 309

模型的基础上,采用远空间分区策略关注骨架上距离较远的关键点运动信息,与注意力机制网络SENet融合2种空间特征,设计了MST-GCN模型,可以有效提升动作识别精度;在打钻视频上利用SVM辨识打钻姿势,利用MST-GCN模型识别动作类别,并根据实际打钻时间设置相邻动作的识别间隔,从而记录上杆、卸杆动作数量,实现钻杆计数。

(2)为了验证MST-GCN模型的有效性,在自建的数据集上进行了实验,并与ST-GCN、Alphapose-LSTM及NST-GCN模型进行了对比,结果表明:MST-GCN、ST-GCN、Alphapose-LSTM及NST-GCN模型的识别准确率分别为91.1%、84.9%、72.1%和86.3%,MST-GCN模型的识别能力更好。在相同条件的打钻视频上,MST-GCN模型、人工方法和改进ResNet方法的计数结果平均误差依次为0.25,0.75,21次,说明MST-GCN模型的计数效果更好。在真实打钻1 300根的现场应用中,MST-GCN模型的平均误差为9根,误计数结果低,说明了基于改进时空图卷积神经网络的钻杆计数方法的可靠性。

(3)该方法也存在一些不足,比如在频繁发生多

根钻杆连接在一起被卸下的情况下,该方法可能会导致少计数问题。下一步工作将对此进行深入研究,以使方法适用于更复杂的工作环境。

参考文献(References):

- [1] 梁运培, 郑梦浩, 李全贵, 等. 我国煤与瓦斯突出预测与预警研究综述[J/OL]. 煤炭学报: 1-24[2022-12-08]. DOI: 10.13225/j.cnki.jccs.2022.0965.
LIANG Yunpei, ZHENG Menghao, LI Quangui, et al. A review on prediction and early warning methods of coal and gas outburst[J/OL]. Journal of China Coal Society: 1-24[2022-12-08]. DOI: 10.13225/j.cnki.jccs.2022.0965.
- [2] PAN Xiaokang, CHENG Hao, CHEN Jie, et al. An experimental study of the mechanism of coal and gas outbursts in the tectonic regions[J]. *Engineering Geology*, 2020, 279: 105883. DOI: 10.1016/j.enggeo.2020.105883.
- [3] 谢和平, 周宏伟, 薛东杰, 等. 我国煤与瓦斯共采: 理论、技术与工程[J]. 煤炭学报, 2014, 39(8): 1391-1397.
XIE Heping, ZHOU Hongwei, XUE Dongjie, et al. Theory, technology and engineering of simultaneous exploitation of coal and gas in China[J]. Journal of China Coal Society, 2014, 39(8): 1391-1397.
- [4] 李东前. 煤矿瓦斯防治技术研究[J]. *当代化工研究*, 2021(10): 99-100.
LI Dongqian. Study on coal mine gas prevention and control technology[J]. *Modern Chemical Research*, 2021(10): 99-100.
- [5] 李树刚, 包若羽, 张天军, 等. 本煤层瓦斯抽采钻孔合理密封深度确定[J]. 西安科技大学学报, 2019, 39(2): 183-188, 216.
LI Shugang, BAO Ruoyu, ZHANG Tianjun, et al. Determining the rational sealing depth for horizontal gas drainage borehole[J]. Journal of Xi'an University of Science and Technology, 2019, 39(2): 183-188, 216.
- [6] 孙志飞, 吴银成, 胡云. 钻杆长度测量方法[J]. 工矿自动化, 2015, 41(3): 51-53.
SUN Zhifei, WU Yincheng, HU Yun. Length measurement methods of drill pipe[J]. Industry and Mine Automation, 2015, 41(3): 51-53.
- [7] 徐钊, 房咪咪, 周红伟, 等. 基于电驻波的锚杆长度无损测量方法[J]. *工矿自动化*, 2013, 39(9): 112-115.
XU Zhao, FANG Mimi, ZHOU Hongwei, et al. Non-destructive measurement method of anchor stock length based on electricity standing wave[J]. *Industry and*

Mine Automation, 2013, 39(9): 112-115.

- [8] 董立红, 王杰, 岸向阳. 基于改进Camshift算法的钻杆计数方法[J]. 工矿自动化, 2015, 41(1): 71-76.
DONG Lihong, WANG Jie, SHE Xiangyang. Drill counting method based on improved Camshift algorithm[J]. Industry and Mine Automation, 2015, 41(1): 71-76.
- [9] 高瑞, 郝乐, 刘宝, 等. 基于改进ResNet网络的井下钻杆计数方法[J]. 工矿自动化, 2020, 46(10): 32-37.
GAO Rui, HAO Le, LIU Bao, et al. Research on underground drill pipe counting method based on improved ResNet network[J]. Industry and Mine Automation, 2020, 46(10): 32-37.
- [10] 党伟超, 姚远, 白尚旺, 等. 煤矿探水卸杆动作识别研究[J]. 工矿自动化, 2020, 46(7): 107-112.
DANG Weichao, YAO Yuan, BAI Shangwang, et al. Research on unloading drill-rod action identification in coal mine water exploration[J]. Industry and Mine Automation, 2020, 46(7): 107-112.

- [11] YAN Sijie, XIONG Yuanjun, LIN Dahua. Spatial temporal graph convolutional networks for skeleton-based action recognition[C]. Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, 2018: 5361-5368.

- [12] FANG Haoshu, XIE Shuqin, TAI Yuwing, et al. RMPE: regional multi-person pose estimation[C]. Proceedings of the IEEE International Conference on

(上接第 35 页)

- [12] 陈利成, 陈建宏. 基于数据填补-机器学习的煤与瓦斯突出预测效果研究[J]. 中国安全生产科学技术, 2022, 18(9): 69-74.

CHEN Licheng, CHEN Jianhong. Study on prediction effect of coal and gas outburst based on data imputation and machine learning[J]. *Journal of Safety Science and Technology*, 2022, 18(9): 69-74.

- [13] 郑晓亮, 来文豪, 薛生. MI和SVM算法在煤与瓦斯突出预测中的应用 [J]. *中国安全科学学报*, 2021, 31(1): 75-80.

ZHENG Xiaoliang, LAI Wenhao, XUE Sheng. Application of MI and SVM in coal and gas outburst prediction[J]. *China Safety Science Journal*, 2021,

Computer Vision, Venice, 2017: 2334-2343.

- [13] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018: 7132-7141.
- [14] KE QiuHong, BENNAMOUN M, AN Senjian, et al. A new representation of skeleton sequences for 3D action recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017: 3288-3297.
- [15] 卫少洁, 周永霞. 一种结合Alphapose和LSTM的人体摔倒检测模型 [J]. *小型微型计算机系统*, 2019, 40(9): 1886-1890.
WEI Shaojie, ZHOU Yongxia. Human body fall detection model combining Alphapose and LSTM[J]. *Journal of Chinese Computer Systems*, 2019, 40(9): 1886-1890.
- [16] 杨世强, 李卓, 王金华, 等. 基于新分区策略的 ST-GCN人体动作识别[J/OL]. 计算机集成制造系统: 1-16[2022-03-29]. <http://kns.cnki.net/kcms/detail/11.5946.TP.20211022.1500.014.html>.
YANG Shiqiang, LI Zhuo, WANG Jinhua, et al. ST-GCN human action based on new partition strategy[J/OL]. Computer Integrated Manufacturing Systems: 1-16[2022-03-29]. <http://kns.cnki.net/kcms/detail/11.5946.TP.20211022.1500.014.html>.

31(1): 75-80.

- [14] LI Zhuoxuan, SHI Xinli, CAO Jinde, et al. CPSO-XGBoost segmented regression model for asphalt pavement deflection basin area prediction[J]. *Science China (Technological Sciences)*, 2022, 65(7): 1470-1481.
- [15] 任万兴, 郭庆, 石晶泰, 等. 基于标志气体统计学特征的煤自燃预警指标构建[J]. *煤炭学报*, 2021, 46(6): 1747-1758.
REN Wanxing, GUO Qing, SHI Jingtai, et al. Construction of early warning indicators for coal spontaneous combustion based on statistical characteristics of index gases[J]. *Journal of China Coal Society*, 2021, 46(6): 1747-1758.