

文章编号:1671-251X(2020)10-0032-06

DOI:10.13272/j.issn.1671-251x.2020040054

基于改进 ResNet 网络的井下钻杆计数方法

高瑞¹, 郝乐¹, 刘宝¹, 文静怡², 陈宇航¹

(1. 西安科技大学 电气与控制工程学院, 陕西 西安 710054;
2. 西北大学 公共管理学院, 陕西 西安 710127)



扫码移动阅读

摘要:针对现有井下钻杆数量统计方式精度较低、受环境变化影响大等问题,结合卷积神经网络、信号滤波等技术,提出了一种基于改进 ResNet 网络的井下钻杆计数方法。根据视频图像中卸杆动作与非卸杆动作的差异,采用 ResNet-50 网络模型对样本集进行分类训练,判断视频中每一帧图像是否包含卸杆动作;结合线性学习率预热和基于 Logistic 曲线的学习率衰减策略进行学习率更新,以提高模型分类准确率;通过积分法对视频分类置信度进行滤波,并统计置信度曲线下下降沿数量,实现钻杆计数。实验结果表明,预热+衰减的学习率更新策略能够有效提高图像分类模型的分类精度,模型分类检测准确率为 89%。实际应用结果表明,基于改进 ResNet 网络的井下钻杆计数方法可以高效识别视频中的卸杆图像,平均钻杆计数精度为 97%,满足实际应用需求。

关键词:钻孔深度;钻杆计数;图像分类;卷积神经网络;学习率更新;ResNet-50

中图分类号:TD713.3 文献标志码:A

Research on underground drill pipe counting method based on improved ResNet network

GAO Rui¹, HAO Le¹, LIU Bao¹, WEN Jingyi², CHEN Yuhang¹

(1. School of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China; 2. School of Public Administration, Northwest University, Xi'an 710127, China)

Abstract:In view of problems that the existing underground drill pipe quantity statistics method has low accuracy and is easily affected by environmental changes, an underground drill pipe counting method based on improved ResNet network was proposed combining with convolutional neural network, signal filtering and other technologies. According to the difference between unloading action and non-unloading action in video image, the sample set is classified and trained based on the ResNet-50 network model to determine whether each frame of the video contains the unloading action; linear learning rate preheating and Logistic-based learning rate attenuation strategy are combined to update learning rate and improve the accuracy of model classification; the video classification confidence is filtered through the integration method, and the number of falling edges of the confidence curve is counted to realize drill pipe counting. Experimental results show that the preheating + attenuation learning rate update strategy can effectively improve classification accuracy of the image classification model to 89%. The actual application results show that underground drill pipe counting method based on improved ResNet network can efficiently identify unloaded rod images in the video with an average accuracy of 97%, which meets the actual application requirements.

Key words:drilling depth; drill pipe counting; image classification; convolutional neural network; learning rate update; ResNet-50

收稿日期:2020-04-16;修回日期:2020-09-29;责任编辑:胡娴。
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:高瑞(1991—),男,陕西神木人,助理工程师,硕士,主要研究方向为模式识别与神经网络,E-mail:635827438@qq.com。
引用格式: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.

0 引言

目前,中国大部分煤矿企业采用钻孔方式抽采瓦斯,由于钻头在煤层中的路径难以获取,一般通过计算钻杆数量来间接计算钻孔深度,以满足钻孔深度设计要求^[1-3]。传统的钻杆数量人工统计方式存在自动化程度低、受人为因素影响、误差较大等缺点^[4-5]。有学者采用传感器、声波等检测钻杆数量或钻孔深度,如仇海生等^[6]基于声波回声技术测量钻孔深度,但存在测量设备体积较大、检测距离有限等问题。路拓等^[7]基于弹性波理论,在钻杆底端激发、接收弹性波,实现瓦斯抽采钻孔快速抽查,但由于钻孔存在角度偏差,探测精度较低。董立红等^[8]提出一种基于改进 Camshift 算法的钻杆计数方法,通过 Camshift 算法追踪钻杆特征目标,根据追踪路径统计钻杆数量,但该方法对环境光线变化过于敏感,易造成统计结果偏差。

针对上述问题,本文提出一种基于改进 ResNet 网络的井下钻杆计数方法。通过 ResNet 网络判断视频中每一帧图像是否包含卸杆动作,并基于 Logistic 经验曲线自动更新学习率,以提高模型分类准确率;通过积分法对视频分类置信度进行滤波,最后统计置信度曲线下下降沿数量,实现钻杆计数。

1 井下钻杆计数方法原理

基于改进 ResNet 网络的井下钻杆计数方法流程如图 1 所示,包括数据集建立和预处理、图像分类模型选取和训练、钻杆数量统计 3 个部分^[9]。

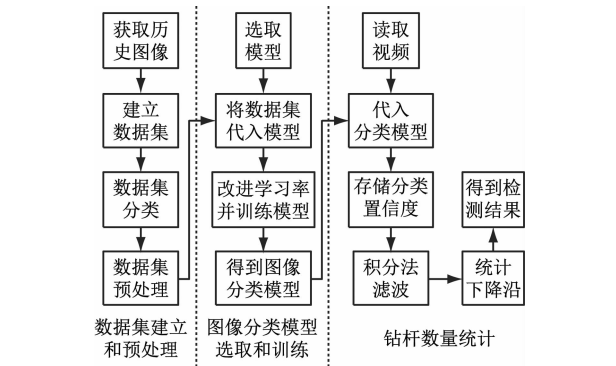


图 1 基于改进 ResNet 网络的井下钻杆计数方法流程
Fig. 1 Process of underground drill pipe counting method based on improved ResNet network

数据集建立和预处理过程:① 获取历史图像并建立数据集;② 按照图像特征,将收集的数据手动分为正在卸杆与非卸杆 2 类;③ 数据集预处理,对数据集进行扩充。

图像分类模型选取和训练过程:① 选取合适的图像分类模型;② 将数据集代入分类模型;③ 对分类模型进行改进和训练,优化模型输出结果。

钻杆数量统计过程:① 将卸杆视频代入分类模型;② 存储模型输出的每帧图像的分类结果与置信度;③ 用积分法对视频分类置信度进行滤波;④ 统计并输出钻杆计数结果。

2 井下钻杆计数方法实现

2.1 图像分类模型构建

与 Camshift 目标跟踪算法相比,深度学习算法能够有效、自动提取图像特征,简化操作流程^[10]。当深度学习模型在 10 层以下时,易出现梯度消失、梯度爆炸现象^[11-12];当模型增至 10 层以上时,由于模型深度加大,在训练中易出现退化现象,导致准确率下降^[13]。K. He 等^[14]提出的残差网络(Residual Network, ResNet)通过残差块与 batchnorm 归一化,有效解决了梯度爆炸、模型退化问题,因此,本文选用 ResNet 构建图像分类模型。为提取图像中卸杆时的有效信息,需要相对较深的网络结构。为保证模型最终的测试精度、减少训练时间,本文使用 ResNet-50 模型对图像进行分类训练。

ResNet-50 结构如图 2 所示。通过 5 段卷积层与全连接层对输入图像进行处理后,再通过分类器判断数据的置信度,以此判断图像类别。使用 ReLU 作为激活函数,与 sigmoid, tanh 等其他激活函数相比,ReLU 收敛速度更快。

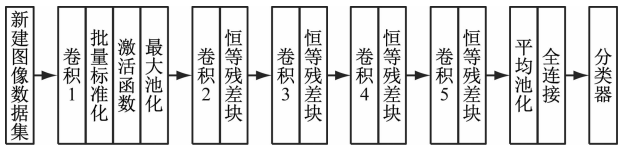


图 2 ResNet-50 结构
Fig. 2 ResNet-50 structure

以 ResNet-50 三层残差块为例,其一般形式如图 3 所示。其中, X 为残差块输入,经过三层卷积运算后,输出 Y :

$$Y = F(X, W_1, W_2, W_3) + WX \tag{1}$$

式中: W_1, W_2, W_3 为每层的学习权重; W 为将输入 X 转换为输出的权值。

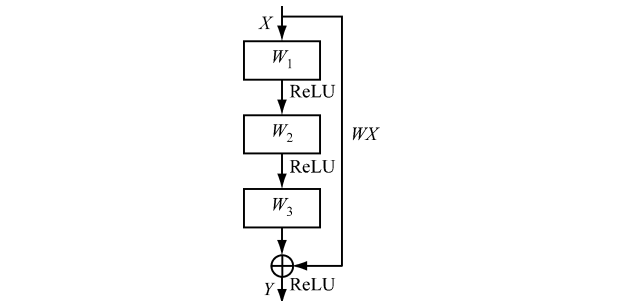


图 3 三层残差块的一般形式
Fig. 3 General form of three-layer residual block

2.2 学习率更新策略

模型训练期间,权重更新的量称为步长或学习

率。学习率是最影响模型训练结果的超参数之一,良好的学习率更新策略能够保证模型更快达到损失的最小值,提高模型分类准确率,还能保证模型不陷入局部最优解,对模型训练结果的导向有重要意义。

学习率设置标准:在训练开始阶段,选用较大的学习率,防止模型陷入局部最优,加快模型收敛速度;随着训练批次增加,学习率逐渐减小,更接近模型最优解。一般而言,可以利用过去的经验(或其他类型的学习资料)直观地设定学习率的最佳值。当已经设定好学习率并训练模型时,只有等学习率随着时间的推移而下降,模型才能最终收敛。然而,随着梯度达到高原,训练损失更难得到改善。针对上述问题,本文结合线性学习率预热和基于 Logistic 曲线的学习率衰减策略进行学习率更新,以提高模型准确率。

模型中所有参数的初始值都是随机值,而初始学习率过高可能会导致模型不稳定。因此,需进行学习率预热,即在开始阶段设置较小的学习率,随着训练批次增加,逐步提高学习率。提取训练中前 10% 的批次进行线性学习率预热,使学习率从 0 开始线性增长至初始学习率 η_1 。预热阶段的学习率更新公式为

$$\eta = \eta_1 i / m \quad (2)$$

式中: η 为第 i ($1 \leq i \leq m$) 批次的学习率; m 为总学习批次的 10%。

在迭代优化的后期,逐步减小学习率的值,有助于算法收敛,更容易接近最优解。因此,需采用一定的学习率衰减策略。目前常用的学习率衰减方法是分段常数衰减,但该方法分段间隔大、更新速度慢^[15]。针对该问题,本文提出利用 Logistic 曲线更新学习率。Logistic 曲线公式为

$$y = \frac{K}{1 + \exp(a + bx)} \quad (3)$$

式中: y 为输出; K, a, b 为未知参数; x 为输入。

学习率太大可能会导致模型出现振荡现象,因此,本文设学习率不超过 1,令 $K=1$ 。当 $a < 0, b > 0$ 时,Logistic 曲线为下降曲线,下降曲线的特点是开始缓慢下降,中间迅速下降,达到某一限度后缓慢下降。

衰减阶段的学习率更新公式为

$$\eta' = \frac{1}{1 + \exp(a + bm')} \quad (4)$$

式中: η' 为第 m' 批次的学习率; $m < m' \leq n'$, n' 为总学习批次。

结合学习率预热与衰减 2 种更新策略,将总学习批次的前 10% 设置为学习率预热部分,后 90% 设

置为 Logistic 衰减部分,得到综合学习率曲线,如图 4 所示。

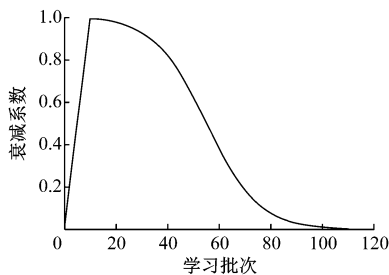


图 4 学习率变化曲线

Fig. 4 Curve of learning rate

2.3 置信度数据滤波

完成模型训练后,将待检测视频输入模型,得到视频中每帧图像的分类置信度,并按照时间顺序存储至 CSV (Comma-Separated Values) 文件中。由于模型检测存在一定误差,难以直接通过分类置信度计算钻杆数量。为便于统计钻杆数量,降低模型检测误差,选用积分法对分类置信度数据进行滤波,将其转换为 0,1 信号。设 CSV 文件中存储的置信度数据为 $f(x)$,连续 5 个数据点进行一次积分,积分公式为

$$\int_{5n}^{5(n+1)} [f(x) - \varphi] dx \quad (5)$$

式中: n 为正整数, $5(n+1) \leq$ 输入视频总帧数; φ 为积分参考值, $0.5 \leq \varphi \leq 0.8$ 。

若式(5)的计算结果大于 0,则该段 5 帧图像为卸杆过程图像,将该 5 帧连续数据记录为 1;若式(5)的计算结果小于 0,则该段 5 帧图像为非卸杆过程图像,将该 5 帧连续数据记录为 0。

为加快检测速率,对积分过程进行简化,以减少滤波计算量。在实际生产中,可采用下式进行滤波:

$$\sum_{x=5n}^{5(n+1)} f(x) - 5 \times 0.5 \quad (6)$$

通过式(6)将输入视频或图像的分类置信度全部转换为 0 或 1,最后计算置信度波形下降沿数量,即为该段视频中包含的卸杆数量。

3 实验分析

3.1 数据集制作

提取原始采集图像中有代表性的 1 500 张图像作为钻杆分类的原始数据集,并统一转换为 224×224 的图像。其中,工人正在卸杆的图像作为分类数据集的第 1 类,样本数量为 670;其余时刻图像作为分类数据集的第 2 类,样本数量为 830。

为了提高训练模型的准确率,需增加数据集样本数量,采用图像翻转、图像旋转与亮度增强操作对样本进行预处理。以第 1 类图像为例,预处理后的

图像如图 5 所示。

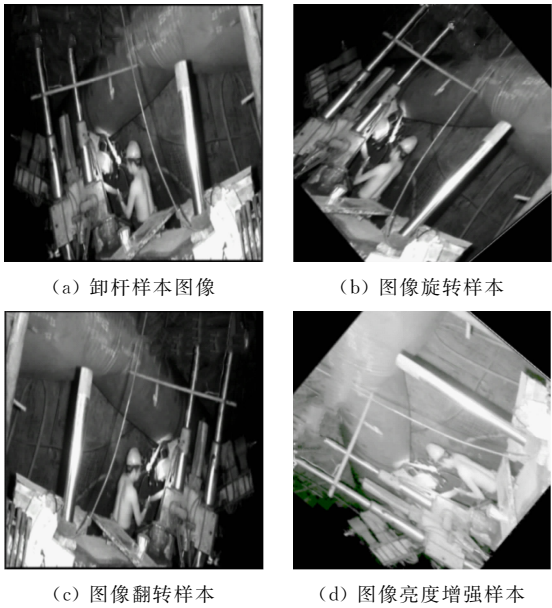


图 5 图像预处理结果

Fig. 5 Image preprocessing results

3.2 学习率更新策略验证

3.2.1 学习率衰减经验公式

由式(4)可得,参数 b 的变化关系到曲线中间缓冲区陡峭或平滑,参数 a 的变化关系到曲线的左移或右移。若 b 值相对较大,会导致曲线陡峭,与分段学习率衰减曲线相似;若 b 值相对较小,则输出结果变化相对较小。为此,在无衰减学习率为 0.002、总学习批次 n' 为 100 的情况下,以衰减曲线过学习批次中点为前提(即输出为 0.5、输入为 $n'/2$ 时, $a = -50b$),先对参数 b 取值,讨论曲线梯度对模型输出的影响。共进行 50 组测试,选取其中有代表性的 3 组进行分析,并与无衰减和分段衰减情况比较,见表 1。

表 1 参数 b 对模型分类结果的影响

Table 1 Influence of parameter b on model classification results

衰减种类	检测准确率	平均损失
无衰减	0.818	0.255
分段衰减	0.839	0.189
Logistic($a = -25, b = 0.5$)	0.836	0.393
Logistic($a = -5, b = 0.1$)	0.857	0.089
Logistic($a = -1, b = 0.02$)	0.845	0.123

由表 1 可知,当 $b = 0.1$ 时,即 $b = 10/n'$ 的情况下,模型准确率相比无衰减情况提升 3.9%,并且能够有效降低模型损失。若 a 值相对较大,会导致曲线左移,处于低学习率批次增多,反之,会导致高学习率批次增多。为此,本文先选定参数 $b = 10/n'$,对参数 a 的取值进行探讨。共进行 20 组测试,选取其

中有代表性的 5 组结果进行分析,见表 2。

表 2 参数 a 对模型分类结果的影响

Table 2 Influence of parameter a on model classification results

a	b	检测准确率	平均损失
-3	0.1	0.844	0.122
-4	0.1	0.877	0.036
-5	0.1	0.857	0.089
-6	0.1	0.864	0.078
-7	0.1	0.848	0.104

由表 2 可知,在 b 值固定的基础上,当 $a = -6 \sim -4$ 时,即 $(-n'/b/2) - 1 < a < (-n'/b/2) + 1$, $-2a \approx n'b$ 的情况下,分类模型能够取得较优的分类结果。

经过多次实验并比较 a, b 取值对模型分类结果的影响,发现分类模型能够取得较优的分类结果时, b 值主要集中在 $5/n' \sim 20/n'$ 内。基于上述实验结果,结合式(4),可得出预热阶段的学习率更新经验公式:

$$\begin{cases} \frac{5}{n'} \leq b \leq \frac{20}{n'} \\ -2a \approx n'b \\ \eta = \frac{1}{1 + \exp\left(-\frac{n'}{2}b + bm'\right)} \end{cases} \quad (7)$$

3.2.2 模型分类结果

采用衰减+预热 2 种策略训练模型后,模型训练的准确率与损失如图 6 所示。从图 6 可看出,模型经过训练后,最终损失下降至 0.1 以内,训练准确率达到 0.98 以上,说明该模型具有较高的准确率。

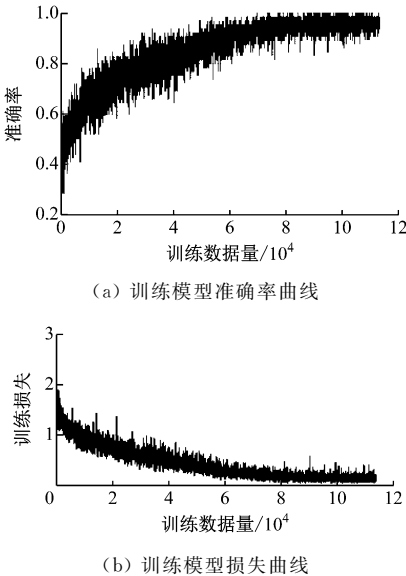


图 6 训练模型准确率及损失曲线

Fig. 6 Accuracy and loss curves of training model

在完成模型训练后,通过 1 900 组卸杆实验对 3 种学习率更新策略的效果进行对比,结果见表 3。由表 3 可知,预热+衰减的更新策略效果最好,分类检测准确率为 0.890,即 89%。

表 3 学习率更新策略效果对比

Table 3 Comparison of effect of learning rate updating strategy		
学习率更新策略	检测准确率	平均损失
无预热、无衰减	0.818	0.255
预热	0.864	0.313
预热+衰减	0.890	0.029

3.3 钻杆计数结果

将训练完成的模型保存至本地,模型接收普通视频与单帧图像输入。人工卸杆平均耗时 4~5 s/根,为加快模型检测速度,同时,不影响卸杆过程的完整性,将视频的帧率由 24 帧/s 降至 2 帧/s。

将卸杆视频输入分类模型后,得出图像置信度并以百分比形式在图像上标注,然后通过积分法进行滤波,得出分类结果。图像分类过程如图 7 所示。

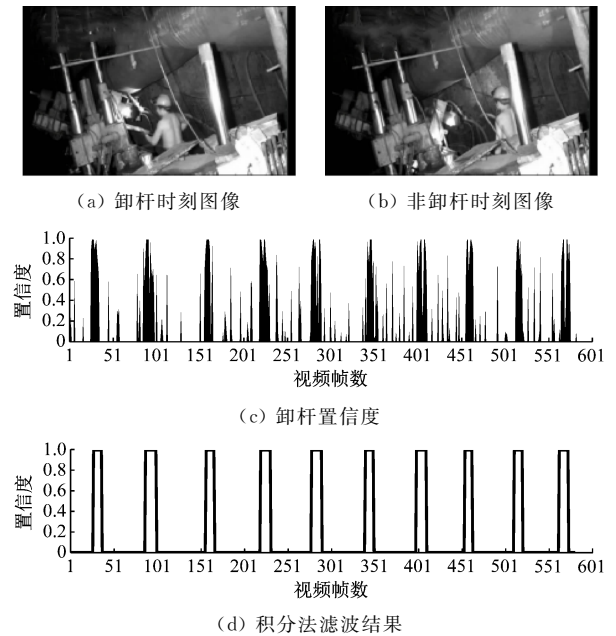


图 7 图像可视化分类过程

Fig. 7 Image visualization classification process
选取打钻角度为 30°,60°,90°,对 3 种情况下的钻杆检测结果进行统计,结果见表 4。人工检测时,将原视频以 16~32 倍速度播放,统计钻杆数量。

表 4 钻杆统计结果

Table 4 Statistical results of drill pipe			
打钻角度/(°)	实际钻杆数量	人工检测结果	图像检测结果
30	100	95	99
60	100	102	103
90	100	100	102

分析表 4 可知,人工检测经常受人员体力影响,识别率可能随时间的延长而降低,而图像检测不受体力等因素影响,识别更稳定,但也存在误识别的现象。现阶段,图像检测识别率并不完全高于人工检测,还需不断对识别过程进行优化,以降低误检率。

最后在河南平煤集团八矿搭建的钻杆统计平台对基于改进 ResNet 网络的井下钻杆计数方法进行验证,累计卸钻杆 1 900 次,平均钻杆统计精度为 97%。

4 结论

(1) 提出了一种基于改进 ResNet 网络的井下钻杆计数方法,采用 ResNet-50 网络建立图像分类模型,并结合线性学习率预热和基于 Logistic 曲线的学习率衰减策略进行学习率更新。对比分析结果表明,与传统学习率更新策略相比,预热+衰减的更新策略效果最好,模型分类准确率为 89%。

(2) 利用积分法对图像分类置信度进行滤波,有效降低了神经网络模型训练结果的误差。实际应用结果表明,基于改进 ResNet 网络的井下钻杆计数方法的平均精度为 97%。

(3) 通过图像分类方法检测钻杆数量,实质是通过图像识别工人卸杆动作来统计钻杆数量,若图像中包含卸杆动作而实际并未卸杆,则会造成误统计的情况。对于如何避免此类误识别情况,今后还需要深入研究。

参考文献(References):

[1] BONDARENKO V,KOVALEVSKA I, ASTAFIEV D, et al. Examination of phase transition of mine methane to gas hydrates and their sudden failure-percy bridgman's effect[J]. Solid State Phenomena,2018, 277:137-146.

[2] 宁小亮.煤与瓦斯突出预警技术研究现状及发展趋势[J].工矿自动化,2019,45(8):25-31.

NING Xiaoliang. Research status of early warning technology of coal and gas outburst and its development trend [J]. Industry and Mine Automation,2019,45(8):25-31.

[3] SUI D, SUKHOBOKA O, AADNØY B S. Improvement of wired drill pipe data quality via data validation and reconciliation[J]. International Journal of Automation and Computing,2018,15(5):625-636.

[4] 刘勇,何岸,魏建平,等.水射流卸压增透堵孔诱因及解堵新方法[J].煤炭学报,2016, 41(8):1963-1967.

LIU Yong, HE An, WEI Jianping, et al. Plugging factor and new plugging method to hydraulic relieving stress[J]. Journal of China Coal Society,2016,41(8):

1963-1967.

[5] 高琨. 煤矿井下钻孔深度检测技术研究[J]. 煤炭科学技术, 2016, 44(4):106-109.

GAO Jun. Study on drilling depth detection technology of underground coal mine[J]. Coal Science and Technology, 2016, 44(4):106-109.

[6] 仇海生, 孙波, 杨春丽. 随钻钻孔深度的声学测量技术实验研究[J]. 世界科技研究与发展, 2015, 37(4):338-341.

QIU Haisheng, SUN Bo, YANG Chunli. Experimental study on acoustic measurement technique of drilling depth[J]. World Sci-Tech R & D, 2015, 37(4):338-341.

[7] 路拓, 刘盛东, 王勃, 等. 深孔钻孔深度的弹性波测量方法[J]. 地球物理学进展, 2015, 30(5):2176-2180.

LU Tuo, LIU Shengdong, WANG Bo, et al. Measurement of deep drilling depth using elastic wave [J]. Progress in Geophysics, 2015, 30(5):2176-2180.

[8] 董立红, 王杰, 匡向阳. 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] 彭业勋. 煤矿井下钻杆计数方法研究[D]. 西安: 西安科技大学, 2019.

PENG Yexun. Research on the counting method of underground drill pipe in coal mine [D]. Xi'an: Xi'an University of Science and Technology, 2019.

[10] 黄一鸣, 雷航, 李晓瑜. 量子机器学习算法综述[J]. 计算机学报, 2018, 41(1):145-163.

HUANG Yiming, LEI Hang, LI Xiaoyu. A survey on quantum machine learning [J]. Chinese Journal of Computers, 2018, 41(1):145-163.

[11] 张钢, 田福庆, 梁伟阁, 等. 基于多尺度 AlexNet 网络的健康因子构建方法[J]. 系统工程与电子技术, 2020, 42(1):245-252.

ZHANG Gang, TIAN Fuqing, LIANG Weige, et al. Construction method of bearing health indicator based on multi-scale AlexNet network [J]. Systems Engineering and Electronics, 2020, 42(1):245-252.

[12] 庄连生, 吕扬, 杨健, 等. 时频联合长时循环神经网络[J]. 计算机研究与发展, 2019, 56(12):2641-2648.

ZHUANG Liansheng, LYU Yang, YANG Jian, et al. Long term recurrent neural network with state-frequency memory[J]. Journal of Computer Research and Development, 2019, 56(12):2641-2648.

[13] WAN J, YILMAZ A, YAN L. DCF-BoW: Build match graph using bag of deep convolutional features for structure from motion[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(12):1847-1851.

[14] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 2016.

[15] HOSSEINI E A, NGUYEN K P, JOINER W M. The decay of motor adaptation to novel movement dynamics reveals an asymmetry in the stability of motion state-dependent learning [J]. PLoS Computational Biology, 2017, 13(5):e1005492.