

文章编号: 2095-2163(2019)03-0045-05

中图分类号: TN919.5

文献标志码: A

油井动液面位置智能识别算法研究

仲志丹, 吴进峰, 任金梅

(河南科技大学 机电工程学院, 河南 洛阳 471003)

摘要: 针对采用声波探测法测得油井动液面数据时,采集到的信号由于受到长距离衰减和复杂背景噪声的影响,动液面回波位置容易淹没在复杂噪声之中不易识别的问题,本文采用一种新的带有宽第一层核的深度卷积神经网络(WDCNN)的方法。即使用采集到的原始声波信号作为输入,并使用第一卷积层中的宽内核来提取特征和抑制高频噪声;卷积层中的小卷积核用于多层非线性映射,池化层用来减少特征的空间大小和网络的参数;在输出层使用 softmax 函数转化识别的不同液面深度值。现场试验结果表明,构建的 WDCNN 模型提高了动液面位置识别的准确性与识别效率,智能识别技术取代了传统的耗时且不可靠的人工分析,降低了油田开采生产成本,提高了经济效益。

关键词: 声波探测法; 液面回波; 抗噪; 深度卷积神经网络; 智能识别

Research on intelligent recognition algorithms for dynamic fluid level position of oil wells

ZHONG Zhidan, WU Jinfeng, REN Jinmei

(School of Mechatronics Engineering, Henan University of Science and Technology, Luoyang Henan 471003, China)

[Abstract] Aiming at the problem that the echo position of the dynamic liquid level is easily submerged in the complex noise due to the influence of long-distance attenuation and complex background noise when the acoustic detection method is used to measure the dynamic liquid level data of oil wells, this paper uses a new method of Deep Convolutional Neural Network (WDCNN) with a wide first-layer kernel. That is, the original acquired acoustic signal is used as an input, and a wide kernel in the first convolutional layer is used to extract features and suppress high frequency noise. The small convolution kernel in the front layer is used for multi-layer nonlinear mapping, the pooling layer is used to reduce the spatial size of the feature and the parameters of the network; the softmax function is used in the output layer to convert the identified liquid surface depth value. Field test results show that the WDCNN model improves the accuracy and efficiency of dynamic liquid level position recognition. Intelligent recognition technology replaces the traditional time-consuming and unreliable manual analysis, reduces the cost of oil field production and improves economic benefits.

[Key words] sound wave detection method; liquid surface echo; noise immunity; Deep Convolutional Neural Network; intelligent identification

0 引言

中国现有油田大多数都处于开发的中后期,由于采油井自身的原油渗出率低和长时间开采造成地层自身压力降低等因素的影响,即使得现有的油田开采必须依靠人工举升的方式来保障任务目标的实现。目前,国内大多数油田矿井使用的是杆泵式的抽油设备进行生产作业,这种抽采油技术具有设备体积小、生产效率高、运行功耗小等诸多优点,因而迄今为止的油田开发中也一直都在使用该项技术^[1]。为了更精细地掌握抽油井生产状态和更全面地判断抽油井的工作性能,需要测量多个油井的

相关参数,其中,最重要的参数就是抽油井动液面深度。而动液面是指产油井中,抽油杆和套气管之间环形空间内的油液形成的油面,而从油面到井口之间的距离深度就是指该井的油液面深度。抽油井的内部结构如图 1 所示。

油井动液面深度值不仅能够反映出油井内动态液面的高度变化和油井内部流动压力大小,而且还可以通过动液面深度值确定泵的沉没度^[2]。确定合理的沉没度是抽油机进行安全生产作业的前提。实时监测动液面深度值,有利于合理安排抽油机的采油速率和抽油机生产作业时间,从而节约能源、降低成本,获得最大的经济效益。

基金项目: 国家自然科学基金(50906022); 河南省高等学校重点科研项目(15A460023)。

作者简介: 仲志丹(1975-),男,博士,副教授,硕士生导师,主要研究方向:工业自动化控制系统、先进能源技术研究;吴进峰(1992-),男,硕士研究生,主要研究方向:人工智能、信号处理。

收稿日期: 2019-03-05

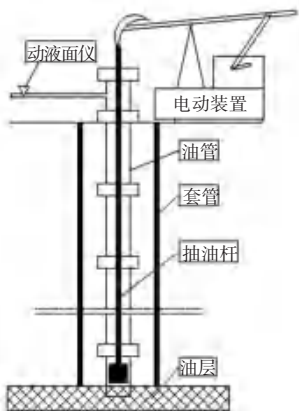


图1 抽油井内部结构图

Fig. 1 Internal structure chart of pumping well

目前,井下压力测量法、示功图法、光纤测量法、声波探测法等都可以用于抽油井动液面深度的测量分析,但声波探测法是应用最广泛、最高效的,因为这种测量方法具有容易操作、低成本和在线实时监测等诸多优点。但也存在一定的不足,较为突出的就是采集的回波信号由于受长距离衰减和复杂背景噪声影响,动液面回波位置容易淹没在噪声之中不易识别。

传统的信号处理方法是通过对傅里叶变换^[3]、均值滤波^[4]、维纳滤波^[5]等来做消噪处理,再进行液面位置的识别,但是这类滤波方法在处理非平稳信号时,具有局限性;还有一些采用双频段带通滤波法和谱减法,但却存在识别的动液面深度值成功率有限和关键信号容易消失等各种问题。

近年来,深度学习^[6]相关理论迅速发展,在机器视觉、图像处理、语音识别等领域都得到了广泛的应用,深度学习具有从数据中自主学习的能力,也可用于油井动液面信号的识别。卷积神经网络^[7-8]算法即是其中的一种代表性算法,其输入可以是一维的时间序列信号,也可以是二维、三维的图片信息。研究可知,通常采用卷积神经网络进行动液面识别时,都需要对采集信号进行预处理,本文则采用一种宽第一层核的深度卷积神经网络的方法,可以使用采集到的原始声波动液面信号进行输入,经过训练得到准确的动液面测量值,提高动液面位置识别率。

1 卷积神经网络研究

卷积神经网络的结构主要包括由卷积层、池化层构成的过滤器阶段和以全连接层构成多层感知器的分类阶段。对此可做阐释分析如下。

1.1 过滤器阶段

卷积层主要用来进行权重分配,卷积层内的过

滤器都使用相同数量的卷积核,把输入原始的动液面信号局部区域与卷积核进行卷积运算,然后通过相应的激活函数生成输出特征。卷积运算过程如式(1):

$$Y_i^{l+1}(j) = K_i^l * x^l(j) + b_i^l, \quad (1)$$

其中,符号 $*$ 表示进行的卷积运算; K_i^l 和 b_i^l 表示第 l 层中第 i 个滤波器的权重和偏差; $x^l(j)$ 表示第 l 层的第 j 个局部区域; $Y_i^{l+1}(j)$ 表示第 $l+1$ 层中第 i 个滤波器中的第 j 个局部区域。

在卷积层的设计运算后,还需要添加池化层,重点是用于减少计算空间的大小和神经网络参数的数量。池化层中,最常用的是max-pooling层。对经过卷积运算并通过相应激活函数获得非线性表达的信号进行最大池化操作,从而使得网络相关运算参数减少,同时将保持截取的局部位置区域不变的特征。研究推得最大池化过程的数学方法公式如下:

$$M_i^{l+1}(j) = \max_{(j-1)W+1 \leq t \leq jW} \{q_i^l(t)\}, \quad (2)$$

其中, $q_i^l(t)$ 表示第 l 层的第 i 个局部区域中第 t 个神经元的值; $t \in [(j-1)W+1, jW]$; W 是截取的局部区域的宽度; $M_i^{l+1}(j)$ 表示进行池化操作后第 $l+1$ 层中对应神经元的值。

研究发现,在经过了卷积层、池化层的设计开发后,为加速神经网络的训练过程,还需要施以批量标准化的操作。设定进行批量标准化层操作的 p 维源信号输入为 $x = (x^{(1)}, \dots, x^{(p)})$,其转换过程如式(3)所示:

$$\hat{x}^{(i)} = \frac{x^{(i)} - E[x^{(i)}]}{\sqrt{\text{Var}[x^{(i)}]}},$$

$$y^{(i)} = \gamma^{(i)} \hat{x}^{(i)} + \beta^{(i)}. \quad (3)$$

其中, $y^{(i)}$ 表示一个神经元响应的输出, $\gamma^{(i)}$ 和 $\beta^{(i)}$ 分别表示学习率和偏差,主要用于恢复神经网络的表示能力,增强网络识别的准确率。

1.2 分类阶段

在分类阶段一般由完全连接层组成,旨在用于分类识别不同的动液面位置。主要有使用Softmax函数用于转换多个神经元的logits以符合多种不同动液面深度值的概率分布形式。

2 宽第一层核的深度卷积神经网络(WDCNN)

2.1 WDCNN网络结构的构建

抽油机采集到的动液面信号是基于时间序列的一维信号,与神经网络常用于识别的二维图像信号有所不同^[9]。常用的神经网络的第一层都是非常小的卷积核,但在实际的动液面信号采集过程中,由

于抽油杆上下滑动,电机运转,井内存气结蜡等因素影响,故而经常会受到高频噪声的干扰。为了捕获所采集到的信号中包含的中频、低频的关键信息,在原来传统的小内核层的神经网络的结构基础上添加宽内核用于提取采集到信号中更多的有用信息,此

后再继续使用小内核进行卷积运算,为此就将其称为宽第一层核深度卷积神经网络(WDCNN)。本文模型构建中使用的 WDCNN 的体系结构包括 5 个卷积和池化层,然后是完全连接的隐藏层,最后是 Softmax 层。本文模型的设计架构如图 2 所示。

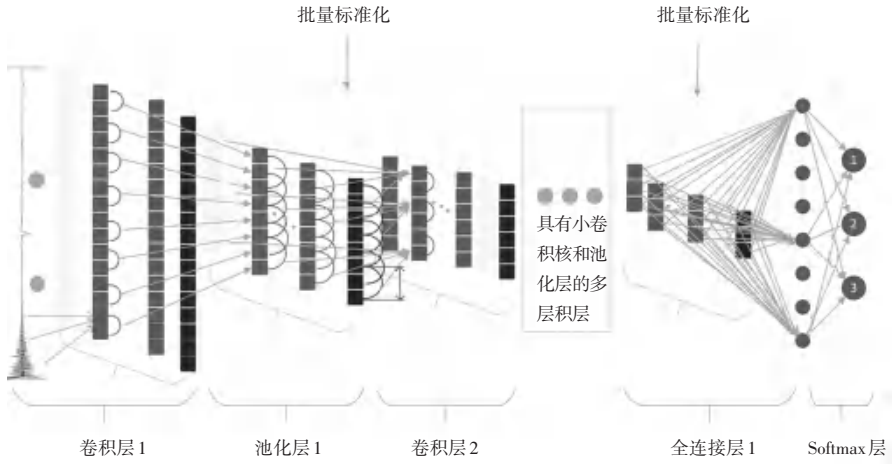


图 2 构建的 WDCNN 模型
Fig. 2 WDCNN model diagram

2.2 WDCNN 网络的训练

动液面液面位置识别的声波信号输入的是一维时间序列信号,与传统的二维图像识别不同,使用了一维的卷积核和反向传播算法。

神经网络结构的代价函数估计的 Softmax 输出概率分布与目标类概率分布之间的交叉熵。考虑到在模型构建中,采用的批量标准化过程已经包含正则化处理的效果,所以,在代价函数中就没有加入正则化项。设 $p(x)$ 是目标分布, $q(x)$ 是估计分布,则代价函数的数学公式可表述为:

$$J(x) = H(p, q) = - \sum_x p(x) \log q(x), \quad (4)$$

反向传播算法中使用梯度下降进行参数的迭代更新,其更新的过程如下:

$$W_{ij}^l = W_{ij}^l - \alpha \frac{\partial H}{\partial W_{ij}^l},$$

$$b_i^l = b_i^l - \alpha \frac{\partial H}{\partial b_i^l}. \quad (5)$$

其中, α 是学习率。

研究中构建的 WDCNN 模型参数有多个,为了防止过度拟合并提高模型的泛化能力,需要大量的训练样本。动液面训练样本的实验数据分别从延长油田靖边采油厂井号为 586-11(动液面深度在 800 m)、靖边采油厂井号 51-1(动液面深度在 1 500 m)和吴起采油厂井号 38-15(动液面深度在 2 100 m)

三口采油井中获得。为了研究获知应选用何种规模的训练样本数据,分别使用 100、300、900、1 500、3 000、6 000 和 10 000 个训练样本来训练构建网络模型的性能。多组训练样本对动液面识别的准确率如图 3 所示。

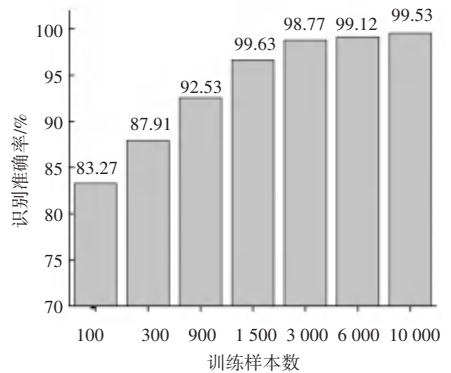


图 3 训练样本数对识别准确率的影响

Fig. 3 The influence of training sample number on recognition accuracy

不断调节学习率 α 大小和梯度下降过程中的迭代次数,寻找代价函数损失最小的学习率 α 和迭代次数。在调节学习率 α 时,分别使用 0.5、0.8、1 等数值进行实验仿真;在调节迭代次数时,分别迭代 300、500、800、900、1 000 次,可得效果绘制曲线如图 4 所示。分析图 4 可知,当进行 300、500 次迭代时,代价函数的损失一直在减小,但当迭代 1 000 次时,在 800 与 900 次之间,出现了代价函数损失的上升。

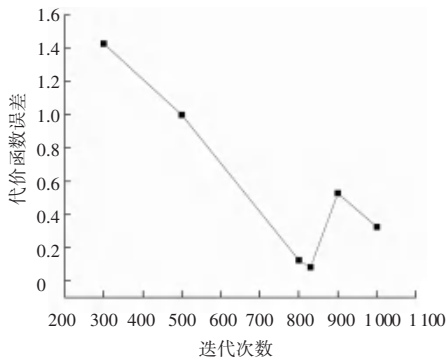


图4 迭代次数对代价函数误差的影响

Fig. 4 The influence of iteration times on the error of cost function

综合前文论述可知,经多次调参和实验,最终确定得到了学习率 $\alpha = 0.8$,迭代次数为 830 次,训练样本为 10 000 的最优训练模型。这里,进一步给出迭代 830 次时代价函数损失值变化如图 5 所示。

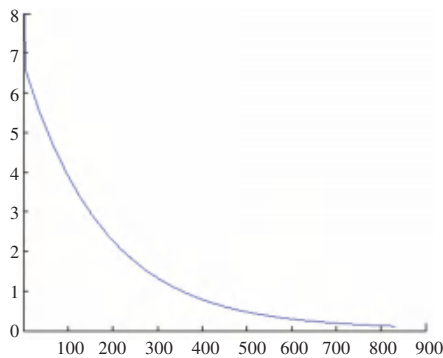


图5 迭代 830 次时代价函数损失值变化

Fig. 5 Loss value change of price function at 830 iterations

另有研究表明,域适应是研究训练的 WDCNN 网络结构在不同井况动液面识别中的挑战性问题^[10]。当所构建的 WDCNN 网络结构中的分类器由在另一个不同的抽油井中收集的样本训练时,很难对这个抽油井采选收集样本再进行准确的分类。不同的抽油机井况可以被视为一个域,因此研究中即将获取标记数据并初始训练本文模型的域称为源域,而研究中只是获取未标记数据并测试本文模型的域名称为目标域。研究至此得到的文中使用的 WDCNN 网络结构的域适应总体框架如图 6 所示。

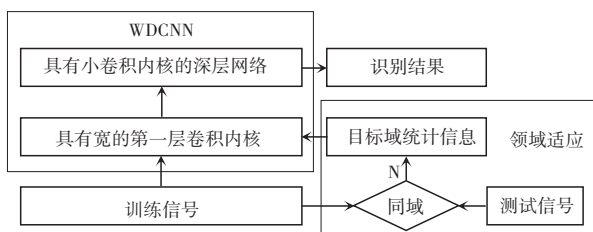


图6 WDCNN 网络结构的域适应总体框架

Fig. 6 Domain adaptation overall framework for WDCNN network architecture

3 实验验证

用于动液面识别实验构建的 WDCNN 模型经过调参训练结束之后,为验证所构建的模型对油井液面识别准确率的性能优劣,从延长油田吴起和靖边采油厂三口井中各取 100 组、共 300 组原始数据进行验证,同时也采用其它传统的信号处理方法,例如带通滤波法、谱减法、稀疏分解法、FFT-SVM 等方法开展对比性的识别验证。仿真实验运行后可得出,各种方法的识别准确率如图 7 所示。

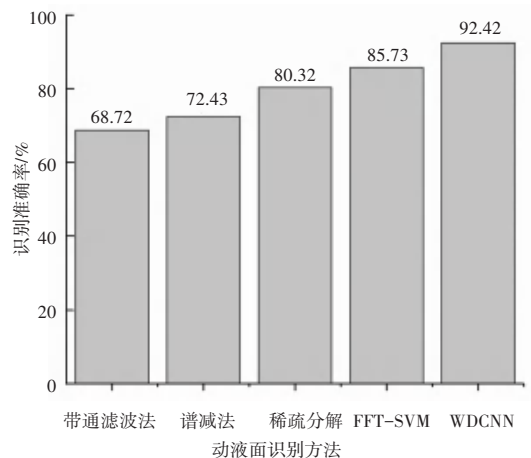


图7 不同方法液面识别准确率

Fig. 7 Accuracy of liquid level recognition with different methods

由图 7 看到,所使用的宽第一核深度卷积神经网络的动液面位置识别方法相较传统的信号处理方法,识别准确率得到了较大的提升,满足了油田生产作业的要求。

4 结束语

(1) 本文使用一种新的宽第一层核卷积神经网络(WDCNN)结构来解决油井动液面位置识别问题。所构建的 WDCNN 可以直接作用于采集到的原始动液面信号,无需其他任何耗时费力的前置人工预处理过程。

(2) 建立使用不同抽油机井域适应框架,更好地适用不同油田、不同井况、不同工作环境下动液面位置信号的准确识别。

(3) 本文使用智能算法用于抽油机井动液面位置识别,节约人力成本,提高识别准确率与识别效率。

参考文献

- [1] 林立星. 声波法测油井动液面信号辨识技术研究[D]. 北京:中国石油大学, 2011.