



# 智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

## CBiA-PSL抽油井异常工况预警模型

李克文, 杜苾聪, 黄宗超, 李潇, 柯翠虹

引用本文:

李克文,杜苾聪,黄宗超,李潇,柯翠虹. CBiA-PSL抽油井异常工况预警模型[J]. 智能系统学报, 2022, 17(2): 295–302.

LI Kewen, DU Congcong, HUANG Zongchao, LI Xiao, KE Cuihong. Early warning model for abnormal working conditions of CBiA-PSL pumping wells[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(2): 295–302.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202106007>

## 您可能感兴趣的其他文章

### 层次化双注意力神经网络模型的情感分析研究

Hierarchical double-attention neural networks for sentiment classification

智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

### 基于双向消息链路卷积网络的显著性物体检测

Salient object detection based on bidirectional message link convolution neural network

智能系统学报. 2019, 14(6): 1152–1162 <https://dx.doi.org/10.11992/tis.201812003>

### 基于注意力融合的图像描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

### 应用于不平衡多分类问题的损失平衡函数

Application of the loss balance function to the imbalanced multi-classification problems

智能系统学报. 2019, 14(5): 953–958 <https://dx.doi.org/10.11992/tis.201808004>

### 区域损失函数的孪生网络目标跟踪

Regional loss function based siamese network for object tracking

智能系统学报. 2020, 15(4): 722–731 <https://dx.doi.org/10.11992/tis.201910005>

### 基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

微信公众平台



关注微信公众号，获取更多资讯信息

DOI: 10.11992/tis.202106007

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20211018.1854.002.html>

# CBiA-PSL 抽油井异常工况预警模型

李克文, 杜苾聪, 黄宗超, 李潇, 柯翠虹

( 中国石油大学(华东) 计算机科学与技术学院, 山东 青岛 266580 )

**摘要:** 油田生产过程中, 油井受各种因素的影响容易发生泵漏、管漏等异常工况, 会降低油井产出甚至导致躺井, 对异常工况预警是油田智能化管理的重要任务。基于 CNN-BiGRU 联合网络, 提出一种改进的网络结构 CBiA-PSL 模型 (CNN BiGRU attention-positive sharing loss), 用于油井异常工况预警。模型利用 CNN 学习工况样本灰度图像的深度特征, BiGRU 有效避免信息损失并加强 CNN 池化层特征的联系, 注意力机制对隐藏状态加权计算以完成有效特征筛选。针对工况数据集不平衡的问题, 提出正共享损失函数 PSL, 将异常数据 (正类) 划分为子类, 每个子类都共享整个正类的损失, 且给样本少的正类更高的权重。实验结果表明, CBiA-PSL 模型预测效果更佳, 对于异常类和整体的预测都有较高的精度。

**关键词:** 卷积神经网络; 双向门控循环单元; 注意力机制; 正共享损失; 损失函数; 异常工况识别; 工况诊断与预警; 数据不平衡

中图分类号: TP18 文献标志码: A 文章编号: 1673-4785(2022)02-0295-08

中文引用格式: 李克文, 杜苾聪, 黄宗超, 等. CBiA-PSL 抽油井异常工况预警模型 [J]. 智能系统学报, 2022, 17(2): 295-302.

英文引用格式: LI Kewen, DU Congcong, HUANG Zongchao, et al. Early warning model for abnormal working conditions of CBiA-PSL pumping wells[J]. CAAI transactions on intelligent systems, 2022, 17(2): 295-302.

## Early warning model for abnormal working conditions of CBiA-PSL pumping wells

LI Kewen, DU Congcong, HUANG Zongchao, LI Xiao, KE Cuihong

(College of Computer Science and Technology, China University of Petroleum, Qingdao 266580, China)

**Abstract:** In the process of oilfield production, which is affected by various factors, oil wells are prone to abnormal working conditions, such as pump and pipe leakages, which will reduce the output of oil wells and even lead to lying in wells. Early warning of abnormal working conditions is an important task of intelligent oilfield management. Based on the convolutional neural network-bidirectional gated recurrent unit (CNN-BiGRU) joint network, an improved network structure CNN-BiGRU attention-positive sharing loss (CBiA-PSL) model is proposed for early warning of abnormal oil well conditions. The model uses the CNN to learn the depth features of the gray image of the sample, BiGRU to effectively prevent information loss and strengthen the connection between the features of the CNN pool layer, and attention mechanism to weigh the hidden state to complete effective feature screening. To address the imbalance of the working condition dataset, a PSL function, which divides the abnormal data (positive class) into subclasses, is proposed. Each subclass shares the loss of the entire positive class and provides a high weight to the positive class with only a few samples. The experimental results show that the CBiA-PSL model has a better prediction effect and a higher accuracy for anomaly and overall prediction than other models.

**Keywords:** convolutional neural network; bidirectional gated recurrent unit; attention; positive sharing loss; loss function; identification of abnormal working conditions; working condition diagnosis and early warning; data imbalance

油田生产信息化建设, 基本实现了油水井、站库数据的实时采集, 油井的智能化管理需求日益突出。

在油田生产过程中, 油井受各种因素的影响出故障的概率很高, 经常造成油井产量降低甚至躺井。目前, 识别异常工况主要依靠简单的参数超阈值报警, 导致报警频繁且有效率不高, 问题原因仍主要采取人工分析, 数据变化趋势难

收稿日期: 2021-06-03. 网络出版日期: 2021-10-19.

基金项目: 国家自然科学基金重大项目 (51991365); 国家自然科学基金面上项目 (61673396).

通信作者: 李克文. E-mail: [likw@upc.edu.cn](mailto:likw@upc.edu.cn).

以自动跟踪、问题隐患难以超前发现、预防性治理和优化措施难以精准实施,难以满足信息化条件下“预警式管理”的要求。所以,预警油井异常工况,指导技术人员超前采取治理与优化措施,对于提高油井开采效率,降低躺井率与维护成本,实现智能化精细化管理具有重要意义<sup>[1]</sup>。

为了提高油井异常工况的识别精度,世界各地的学者进行了大量的研究。考虑到油井生产数据的特征,即异常工况样本少、维数大等特点,近年来识别异常工况的方法可分为两类:基于统计<sup>[2]</sup>和传统机器学习的方法<sup>[3]</sup>以及基于深度学习的方法<sup>[4-6]</sup>。对于第一类,Zhang 等<sup>[7]</sup>将支持向量机与遗传算法相结合,提出基于 GA-SVM 与频域光谱法的变压器油浸绝缘体的湿度预测;周斌等<sup>[8]</sup>提出一种基于 Hessian 正则化支持向量机(Hessian 正则化 SVM)的多视角协同识别抽油机井工况方法。这些研究提高了异常工况检测的准确性,但它们都忽略了工况数据类别不平衡的问题。王利君等<sup>[9]</sup>使用集成 SMOTE、CLUSTER 与随机森林的集成学习方法 SCRF 进行结蜡预测,考虑了类别不平衡,但忽略了高维数据的问题且目标数据集相对较小,模型的泛化性能和自学习能力不足。

近年来,由于深度学习模型具有强大的特征提取能力及拟合海量数据的能力,因此被广泛应用于故障检测领域<sup>[10]</sup>。例如,Wei 等<sup>[11]</sup>提出基于电机数据的深度学习的抽油杆泵故障诊断,利用 CNN 作特征提取器弥补了专家经验的不足;Chen 等<sup>[12]</sup>提出基于多尺度 CNN 和 LSTM 的轴承故障诊断,使用两个不同内核大小的 CNN 从原始数据中提取不同的频率信号特征,然后利用 LSTM 识别故障类型;Cabrera 等<sup>[13]</sup>提出基于 LSTM 的往复压缩机故障诊断模型,其中超参数搜索由在每次迭代中限制搜索空间的贝叶斯方法完成;魏晓良等<sup>[14]</sup>提出了一种基于长短时记忆(LSTM)和一维卷积神经网络(1D-CNN)相结合的空化故障诊断方法,用于高速柱塞泵故障诊断;Liang 等<sup>[15]</sup>提出一种使用自适应矩估计最大值(adamax)优化算法的双向门控递归神经网络(adamax-BiGRU)的瓦斯浓度预测模型,预警煤矿瓦斯事故。这些研究初步证明了深度学习模型在故障检测领域的有效性。但是,LSTM 具有许多参数和复杂的结构,容易出现过拟合的问题。作为 LSTM 的变体,GRU 具有简单的结构、较少的参数和较短的训练时间,这比 LSTM 更具优势。但是,GRU 仅按顺序考虑前向信息,不考虑反向信息。由两个 GRU 组成的双向 GRU(BiGRU)可以利用附加的后向生产数据,从而进一步提高模型预警异常的准确性。

在上述方法中,存在一个主要问题:实际生产数据中抽油井异常记录远小于正常生产记录,即抽油井异常工况数据集往往存在类别不平衡的问题。上述方法仅考虑总体准确率,而忽略了类不平衡问题,导致少数类分类出错率较高,实际上的异常工况检测效果不理想。因此,本文提出了一种正共享损失函数<sup>[16]</sup>,以增强少数类,即异常工况的检出效率。

本文的主要工作如下:

1)针对工况样本的灰度图像,在 CNN-BiGRU 联合网络模型上引入注意力机制,从而提高模型对特征的学习能力,进一步提高识别异常工况的准确率。

2)除了将 CNN-BiGRU-Attention 网络用作特征提取器外,提出了正共享损失函数 PSL。PSL 引入一个额外的正则化项,以强调正负类的损失,且给样本少的正类更高的权重,旨在减弱不平衡,有助于更好地识别异常工况。

## 1 CNN-BiGRU 联合网络特征学习

CNN 网络主要有两个算子<sup>[17]</sup>,一个是卷积层,另一个是池化层。卷积层作为特征提取器将数据集的工况样本的灰度特征矩阵分割成若干子矩阵,每个卷积层中所有的特征子矩阵与同一个权值矩阵(卷积核)做卷积运算,通过卷积核刻画图片的局部模式来提取图像的局部特征<sup>[18]</sup>。卷积运算可以提取数据集中人类无法理解的异常工况的局部抽象特征,起到过滤作用。池化层在卷积层之后,对卷积得来的特征进行筛查,减少特征数量来降低计算量,同时可以起到保留异常工况特征以及防止过拟合的作用。

门控循环单元(gate recurrent unit, GRU)是一种特殊的循环神经网络(recurrent neural network, RNN),其与长短期记忆网络(long-short term memory, LSTM)相似,是为了解决长期记忆和反向传播中的梯度爆炸、梯度消失和长距离依赖等问题而提出的。

GRU 适宜于处理时间序列数据,对比 LSTM,GRU 在性能相当的同时参数量更少,结构简单,更易收敛,且计算速度比 LSTM 更快。双向 GRU<sup>[19]</sup>在输入序列上有两个 GRU 互相连接,每一个输入的异常工况特征图都会从正向和反向经过 GRU,为神经网络提供上下文的全局特征。在图 1 中,每个 GRU 单元都在两个方向上进行处理:GRU<sub>1</sub>是正向 GRU,其内部结构如图 2 所示;GRU<sub>2</sub>是反向 GRU,其内部结构如图 3 所示。

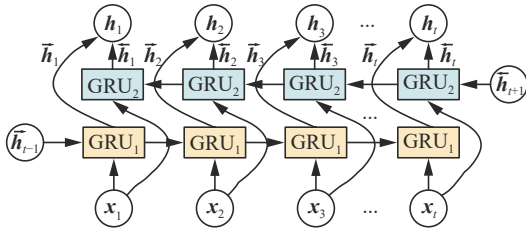


图1 BiGRU 的网络结构

Fig. 1 Network structure of BiGRU

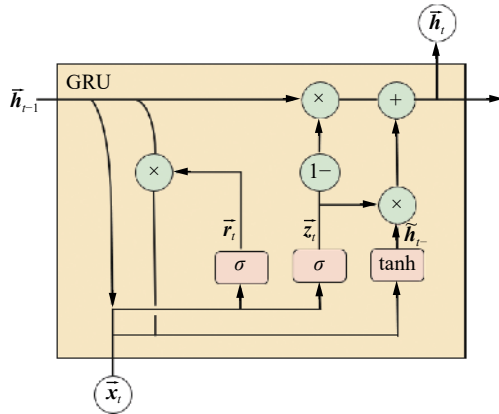


图2 正向 GRU 单元的内部结构

Fig. 2 Internal structure of forward GRU unit

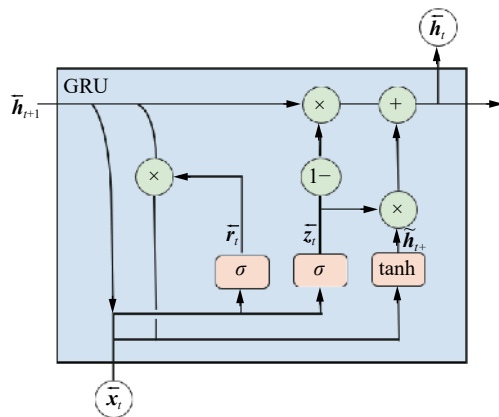


图3 反向 GRU 单元的内部结构

Fig. 3 Internal structure of reverse GRU unit

图2中的正向计算过程如下:

定义 $r_t$ 是正向 GRU 在 $t$ 时刻的复位门。公式如下:

$$\vec{r}_t = \sigma(\vec{W}_r \vec{x}_t + \vec{U}_r \vec{h}_{t-1})$$

式中: $\sigma$ 为 sigmoid 函数; $x_t$ 和 $h_{t-1}$ 分别是当前输入值和上一个激活值; $W_r$ 是输入权重矩阵; $U_r$ 是循环连接的权重矩阵。

类似地,定义 $z_t$ 是正向 GRU 在 $t$ 时刻的更新门。公式如下:

$$\vec{z}_t = \sigma(\vec{W}_z \vec{x}_t + \vec{U}_z \vec{h}_{t-1})$$

定义 $h_t$ 是正向 GRU 在 $t$ 时刻的激活值,即上一个激活值 $h_{t-1}$ 和候选激活值 $h_{t+}$ 之间的折中。

$$\vec{h}_t = (1 - \vec{z}_t) \cdot \vec{h}_{t-1} + \vec{z}_t \cdot \vec{h}_{t+}$$

$h_{t-}$ 公式如下:

$$h_{t-} = \tanh(\vec{W}_h \vec{x}_t + \vec{U}_h \vec{h}_{t-1})$$

式中:“ $\cdot$ ”代表哈达玛乘积(Hadamard)。

当重置门 $r_t$ 关闭时,即其值接近于0,GRU 忽略先前的激活值 $h_{t-1}$ ,仅由当前输入 $x_t$ 决定。这允许 $h_t$ 丢弃不相关的信息,从而更有效地表达有用的信息。

另一方面,更新门 $z_t$ 控制将 $h_{t-1}$ 中的多少信息传递给当前 $h_t$ 。

同样,图3中的反向计算过程如下:

$$\overleftarrow{r}_t = \sigma(\overleftarrow{W}_r \overleftarrow{x}_t + \overleftarrow{U}_r \overleftarrow{h}_{t+1})$$

$$\overleftarrow{z}_t = \sigma(\overleftarrow{W}_z \overleftarrow{x}_t + \overleftarrow{U}_z \overleftarrow{h}_{t+1})$$

$$\overleftarrow{h}_t = (1 - \overleftarrow{z}_t) \cdot \overleftarrow{h}_{t+1} + \overleftarrow{z}_t \cdot \overleftarrow{h}_{t+}$$

$$h_{t+} = \tanh(\overleftarrow{W}_h \overleftarrow{x}_t + \overleftarrow{U}_h \overleftarrow{h}_{t+1})$$

将两个方向的结果求平均,以获得最终输出 $h_t$ 。

$$h_t = (\vec{h}_t + \overleftarrow{h}_t) / 2$$

## 2 CBiA-PSL 工况预警模型

为了更好地解决深度学习预测模型中工况属性的选择问题,且所提出的方法对图像处理更有效,将油井数据转化为灰度图像。数据集中共有41维特征,为了保留数据集中每个样本的所有特征,用8个0填充,将其转化为 $7 \times 7$ 灰度图像。转化后的灰度图像如图4所示,从左到右依次为无异常、泵漏、管漏,同一类别的图片几乎相同,但是不同类别的图片之间存在很大差异。

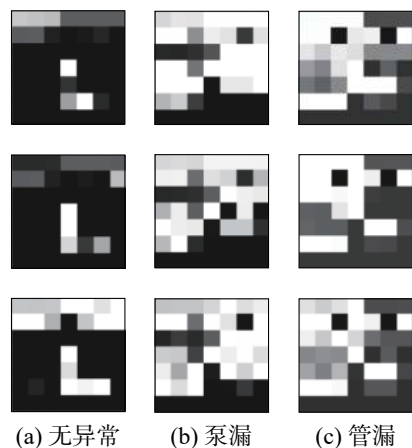


图4 样本灰度图像

Fig. 4 Sample gray image

### 2.1 CNN-BiGRU-Attention 特征筛选

本文提出的 CBiA-PSL 模型是基于 CNN-BiGRU-Attention 网络和正共享损失函数 PSL。CNN-BiGRU-Attention 在 CNN-BiGRU 联合网络的基础上引入注意力机制,其结构如图5所示。首先,



7×7 工况样本灰度图像作为 CNN 的输入, 利用 CNN、BiGRU 提取前后向相关特征, 克服了 CNN 缺乏对上下文的全局关注与 BiGRU 缺乏对局部关注的不足, 结合两者的优点从全局和局部对异常工况特征学习训练, 并通过 Attention 层增强相关特征表示<sup>[20-21]</sup>, 最后, 通过 softmax 层输出分类结果。

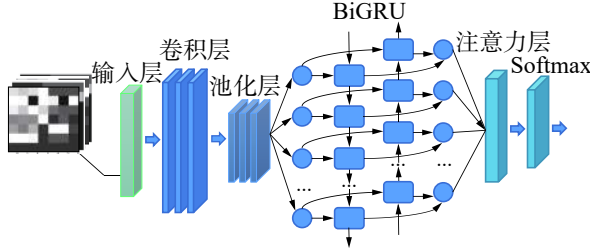


图 5 CNN-BiGRU-Attention 模型结构  
Fig. 5 Model structure of CNN-BiGRU-Attention

在 CNN-BiGRU 联合网络架构的基础上增加注意力机制对隐藏状态加权计算以完成有效特征筛选。其核心是权重系数, 首先学习每个特征的重要程度, 而后根据重要程度为每个特征分配相应的权重, 以区分各特征的重要性大小, 提高工况识别的准确率。本文使用前馈注意力机制, 其结构如图 6 所示。

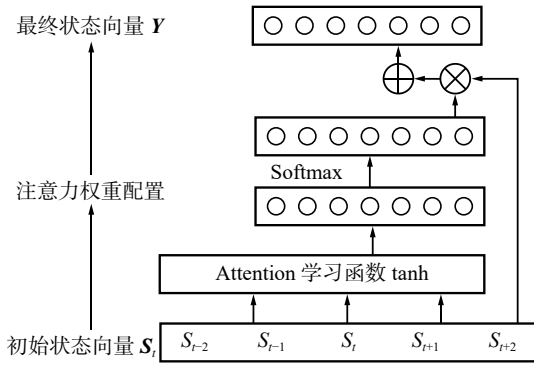


图 6 注意力机制结构  
Fig. 6 Structure of attention mechanism

首先生成目标注意力权重  $e_t$ , 公式如下:

$$e_t = \sigma(w_t S_t + b_t)$$

式中:  $\sigma$  是注意力学习函数  $\tanh$ ;  $S_t$  是第  $t$  个特征向量的初始状态向量;  $w_t$  表示第  $t$  个特征向量的权重系数矩阵;  $b_t$  表示第  $t$  个特征向量相对应的偏移量。

然后将注意力权重概率化, 通过 softmax 函数生成概率向量  $\alpha_t$ , 公式如下:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{i=1}^t \exp(e_i)}$$

最后, 注意力权重配置。将生成的注意力权重配置给对应的隐层状态  $S_t$ , 使模型生成的注意力权重发挥作用,  $Y$  是最终输出的状态向量, 为

$S_t$  的加权平均值, 权值是  $\alpha_t$ , 公式如下:

$$Y = \sum_{i=1}^n \alpha_i \cdot S_i$$

## 2.2 正共享损失函数

在交叉熵损失函数中通过引入一个额外的正则化因子强调正类、负类的区别, 构造正共享损失函数 PSL, 且给样本数少的正类更高的权重, 以减弱不平衡, 更好地区分各类异常与正常的生产状态, 提高识别异常工况的准确率。

训练 CNN-BiGRU-Attention 网络的目标是对异常工况特征进行提取, 最大化识别异常工况的概率, 这是通过最小化交叉熵损失函数实现的。给定一个包含  $m$  个样本的训练集:  $\{x^{(i)}, y^{(i)}\}_{i=1}^m$ ,  $x^{(i)}$  是第  $i$  个样本,  $y^{(i)} \in \{0, 1, 2, \dots, K\}$  是它的标签。  $y^{(i)} = 0$  表示  $x^{(i)}$  是负样本,  $y^{(i)} = k > 0$  表示  $x^{(i)}$  是正样本且  $x^{(i)}$  属于第  $k$  种工况。  $(\alpha_j^{(i)}; j = 1, 2, \dots, K)$  表示 Softmax 层的输出,  $x^{(i)}$  分类为  $j$  的概率为

$$p_j^{(i)} = \frac{\exp(\alpha_j^{(i)})}{\sum_{l=0}^K \exp(\alpha_l^{(i)})}$$

损失函数如下:

$$\text{Loss}_0 = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=0}^K 1(y^{(i)} = j) \log p_j^{(i)} \right]$$

标准交叉熵损失函数会平均惩罚每个类的错分类误差。但在实际情况下, 将正样本分类为错误的非零标签并不是重大错误, 因为仍将其识别为异常。也就是说, 异常类间的分类出错一般可以忽略, 但异常和无异常间分类出错是不可容忍的, 即应更关注零标签和非零标签之间的错误分类导致的损失。为此, 引入了额外的正则化因子, 增加了异常工况错分类到正常和正常错分类到异常类的损失。改进后的损失函数如下:

$$\text{Loss} = \text{Loss}_0 - \frac{1}{m} \cdot$$

$$\left\{ \sum_{i=1}^m \lambda \left[ (1(y^{(i)} = 0)) \log p_0^{(i)} + \omega \sum_{j=1}^K (1(y^{(i)} = j)) \log(1 - p_0^{(i)}) \right] \right\} \quad (1)$$

式中:  $\lambda$  为控制参数。当  $\lambda$  趋向于 0 时, 式(1)为标准交叉熵损失函数; 当  $\lambda$  足够大时, 区分不同异常工况的效果变弱, 式(1)变为解决二分类问题的损失函数, 旨在识别异常工况和正常生产数据。实验中, 默认设置  $\lambda = 1$ 。在实际情况下, 我们更关注能不能识别出异常工况, 而不是异常工况间被错分类的概率, 因此引入的正则化项中各正类别的损失在其他正类间共享, 称式(1)为正共享损失函数。

异常工况样本数少, 数据集中正样本数远小于负样本数, 数据不平衡, 因此为正类项引入参数  $\omega$ ,  $\omega$  为负样本数与正样本数的比值, 旨在减弱不平衡, 公式如下:

$$\omega = \frac{\sum_{i=1}^m 1(y^{(i)} = 0)}{\sum_{i=1}^m \sum_{j=1}^K 1(y^{(i)} = j)}$$

通过以上这两个措施, 从而让 CNN-BiGRU-Attention 模型学习到更加具有辨别力的特征。通过标准反向传播来更新网络参数,  $\text{Loss}_0$  为标准交

$$\frac{\partial \text{Loss}}{\partial a_0^{(i)}} = \frac{1}{m} \left\{ (\lambda + 1) [1(y^{(i)} = 0)] (p_0^{(i)} - 1) + (\lambda + 1) \sum_{j=1}^K [1(y^{(i)} = j)] p_0^{(i)} \right\}$$

$$\frac{\partial \text{Loss}}{\partial a_l^{(i)}} = \frac{1}{m} \left\{ (\lambda [1(y^{(i)} = 0)] + 1) p_l^{(i)} - [1(y^{(i)} = l)] - \lambda \sum_{j=1}^K [1(y^{(i)} = j)] \left( \frac{p_0^{(i)} p_l^{(i)}}{1 - p_0^{(i)}} \right) \right\}$$

### 3 实验与结果分析

#### 3.1 数据集描述

本文使用的数据集来源于胜利油田若干采油厂上百万条抽油井生产数据。原始的数据包括井组相关资料、日常管理资料、实时生产数据、示功图采集数据、功图分析数据、工况日志数据和管柱等资料, 时间范围是 2019~2020 年。井组相关资料主要存储单井基础信息, 包括井号、井别、地质储量、对应的注水井配注量等数据资料; 日常管理资料、实时生产数据、示功图采集数据和功图分析数据主要存储动态示功图及动态参数, 如上下行电流、热洗周期、日产液量、日产油量、含水等; 工况日志数据主要存储异常发生的井号、时间、工况类型、治理措施等; 管柱资料主要存储管柱图、泵效、泵挂深等资料。以上数据在实际使用时往往存在问题: 1) 数据存在缺失、无效波动、重复值; 2) 数据时间段不连续; 3) 工况标签缺失等。因此, 为提高工况预警的准确性和精度, 在建立异常预测模型前需进行预处理, 得到数据质量较高的数据集。

#### 3.2 数据预处理

初始油井实时数据通过缺失值填写、光滑噪声数据、删除重复数据、去除无效波动、数值规约、切片等手段预处理, 共保留 135 个特征字段。在数百个工况属性中, 由于许多工况属性具有相似的公式和表达式, 这些属性之间相关性很高, 去冗余后保留 38 个特征字段。考虑到油井异常会导致示功图载荷和面积有较大变化, 对示功图进行特征量分解, 获取功图面积、载荷差、载荷比 3 个新的特征, 同时构造工况标识字段, 泵漏标

叉熵损失函数, 其导数已在文献 [16] 中提供, 损失函数中第二部分  $\alpha_0^{(i)}$  和  $\alpha_{l(l=1,2,\dots,K)}^{(i)}$  的偏导数计算如下:

$$\frac{\partial \log p_0^{(i)}}{\partial a_0^{(i)}} = 1 - p_0^{(i)}$$

$$\frac{\partial \log(1 - p_0^{(i)})}{\partial a_0^{(i)}} = -p_0^{(i)}$$

$$\frac{\partial \log p_l^{(i)}}{\partial a_l^{(i)}} = -p_l^{(i)}$$

$$\frac{\partial \log(1 - p_0^{(i)})}{\partial a_l^{(i)}} = \frac{p_l^{(i)} p_0^{(i)}}{1 - p_0^{(i)}}$$

正共享损失函数的偏导数计算如下:

记为 1, 管漏标记为 2, 正常标记为 0。原始数据经过预处理后, 得到样本数据集, 共包含 41 个特征参数和 1 个目标参数。将异常工况类(含泵漏、管漏)统称为正类, 无异常统称为负类。预处理后的样本情况如表 1 所示, 正负类样本极不平衡。划分数据集中 70% 作为训练集, 30% 作为测试集。

表 1 预处理后的样本情况  
Table 1 Sample condition after pretreatment

工况	样本数量	占总样本比例
正常	2 288 074	0.93
泵漏	60 158	0.02
管漏	114 720	0.05

#### 3.3 实验设置

本文实验在服务器下进行, 硬件设备 CPU 型号为 Intel(R) Xeon(R) CPU E5-2630L v3 @ 1.80 GHz, 内存大小 48 GB。实验环境为 Python3.8, 借助 Keras2.4.3 搭建神经网络, 后端使用 TensorFlow 2.4.0。

实验参数设置如下: 卷积核大小设置为  $5 \times 5$ , 步长为 1, 隐藏层的数量为 1, 隐藏层单元数为 128, 隐藏层丢包率为 0.5。学习率为 0.01, BiGRU 的时间步长为 8, 隐层单元数为 64, 批次大小为 128, 迭代次数为 50。

#### 3.4 评价指标

##### 3.4.1 对比度得分

可以通过测试集上每个样本的分类准确率来验证模型的有效性, 但是对于检测问题, 正样本和负样本之间的对比度更能反映模型性能。因此, 定义对比度得分作为度量,  $\{x_v^{(i)}, y_v^{(i)}\}_{i=1}^m$  为测试集,  $p_0^{(i)}$  是模型把  $x_v^{(i)}$  分为负类的概率。对比度得分

定义如下:

$$\gamma = \frac{1}{m} \sum_{i=1}^m \left[ (1(y_v^{(i)} = 0) - 1(y_v^{(i)} > 0))(p_0^{(i)} - (1 - p_0^{(i)})) \right]$$

$\gamma$  的取值范围是  $-1 \sim 1$ , 表示模型识别正样本、负样本的能力。

### 3.4.2 混淆矩阵

混淆矩阵如表 2 所示。 $TP_m$  表示正确预测的  $m$  类阳性样本数,  $E_{mn}$  表示  $m$  被归类为  $n$  的错误分类样本数。 $FN_m$  表示  $m$  被归类为其他类的错分类样本数,  $FN_A = E_{AB} + E_{AC}$ ,  $FN_{BC}$  可同样计算得到;  $FP_m$  表示其他类被归类为  $m$  的错分类样本数,  $FP_A = E_{BA} + E_{CA}$ ;  $TN_m$  表示正确预测的  $m$  类负样本数,  $TN_A = S - FN_A - FP_A - TP_A$ 。真阳性率  $TPR_m$  表示所有实际为  $m$  类的样本被正确判断为  $m$  类的比率, 公式如式 (1) 所示; 假阳性率  $FPR_m$  表示所有实际为其他类的样本被错误判断为  $m$  类的比率, 公式如式 (2) 所示。

$$TPR_m = \frac{TP_m}{TP_m + FN_m} \quad (2)$$

$$FPR_m = \frac{FP_m}{FP_m + TN_m} \quad (3)$$

表 2 多分类的混淆矩阵

Table 2 Confusion matrix of multi classification

真实标签	预测标签		
	A	B	C
A	$TP_A$	$E_{AB}$	$E_{AC}$
B	$E_{BA}$	$TP_B$	$E_{BC}$
C	$E_{CA}$	$E_{CB}$	$TP_C$

为了评估 CBiA-PSL 模型, 本文选择了 3 个指标: AC(准确率)、DR(检出率)和 FR(错误报警率):

$$AC = \sum_m \frac{TP_m}{S} \quad (4)$$

$$DR_m = TPR_m \quad (5)$$

$$FR_m = FPR_m \quad (6)$$

### 3.5 结果分析

将学习速率设置为 0.01, 对损失函数执行随机梯度下降 (SGD), 50 次迭代后, 采用标准交叉熵损失函数 ( $Loss_0$ ) 的 CNN-BiGRU-Attention 模型的对比度得分为 0.58, 采用正共享损失函数 PSL ( $Loss$ ) 的 CNN-BiGRU-Attention 模型的对比度得分为 0.63, 有 0.05 的提高, 如图 7 所示。这表明引入正共享损失函数可以提高异常类的检出效率。

第 2 个实验将 CBiA-PSL 与其他 5 种用于异常工况检测的机器学习/深度学习方法进行比较, 包括: CNN-BiGRU-Attention、卷积神经网络 CNN、双向长短时记忆网络 BiLSTM、双向门控循

环单元 BiGRU、随机森林 RF。各模型的 AC 值、DR 值、FR 值如图 8、表 3、表 4 所示, CBiA-PSL 模型训练过程的损失如图 9 所示。从图 8 可以看到, CBiA-PSL 的 AC 值为 88.2%, 高于其他 5 种方法。此外, 分析表 3 可以得出结论, CBiA-PSL 可以提高少数类 (泵漏、管漏) 的 DR 值, 且正常类的 DR 值保持不变, 即提高了异常类的检出率。分析表 4, 少数类的 FR 值有时为 0, 式 (1)~(5) 表明该方法未检测到异常类别时, DR 和 FR 均为 0。表 4 表明本文提出的方法可以将异常类的错误报警率维持在较低水平。因此, 本文所提出的方法 CBiA-PSL 可以在总体准确率较高的情况下, 提高异常类的检出率, 并降低错误报警率。

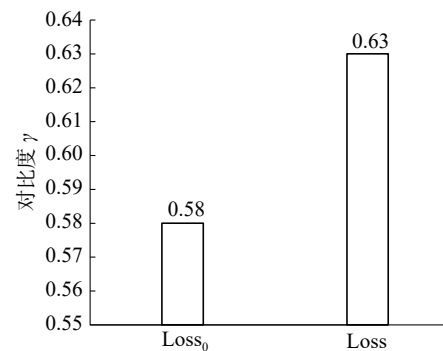


图 7 对比度得分

Fig. 7 Contrast score

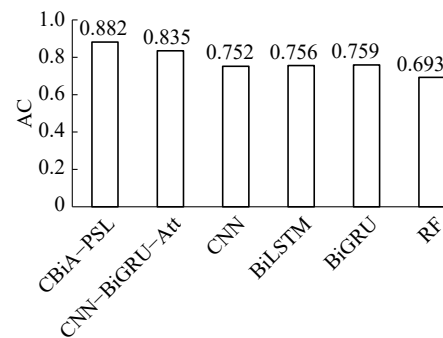


图 8 各模型的 AC 值

Fig. 8 AC value of each model

表 3 各模型的 DR 值

Table 3 DR value of each model

算法	工况		
	正常	泵漏	管漏
CBiA-PSL	<b>0.921</b>	<b>0.712</b>	<b>0.781</b>
CNN-BiGRU-Att	0.921	0.601	0.710
CNN	0.921	0.499	0.645
BiLSTM	0.921	0.531	0.653
BiGRU	0.920	0.529	0.657
RF	0.921	0.503	0.594

表 4 各模型的 FR 值  
Table 4 FR value of each model

算法	工况		
	正常	泵漏	管漏
CBiA-PSL	0.045	0.003	0.011
CNN-BiGRU-Att	0.113	0.006	0.011
CNN	0.120	0.003	0.010
BiLSTM	0.127	0.007	0.010
BiGRU	0.129	0.007	0.010
RF	0.097	0	0.013

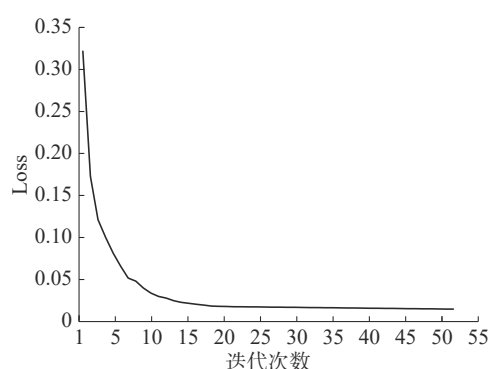


图 9 模型训练集的损失图  
Fig. 9 Loss graph of model training set

## 4 结束语

本文基于 CNN-BiGRU 联合网络特征学习模型, 使用改进的 CBiA-PSL 网络进行抽油井异常工况预警, 利用 CNN 局部特征的强学习能力提取工况样本灰度图像的局部特征, BiGRU 提取全局前后向特征并加强 CNN 池化层特征的联系, 加入注意力机制获取样本中的重点特征, 降低噪声特征的干扰, 从而完成有效特征筛选, 提高模型对异常工况特征的学习能力。除了将 CNN-BiGRU-Attention 网络用作特征提取器外, 针对工况样本数据集不平衡的问题, 本文提出正共享损失函数 PSL, 该函数强调异常类和非异常的损失, 而不是每个子类的损失, 有助于学习比 Softmax 损失函数更多的判别特征, 且给样本少的正类更高的权重, 以学习参数, 减弱了不平衡, 有助于更好地识别异常工况。

实验结果表明, 用本文提出的 CBiA-PSL 网络模型进行异常工况预警, 可以取得较高的 AC 值、DR 值, 较低的 FR 值, 即本文提出的 CBiA-PSL 方法能有效处理不平衡数据集, 并且对于异常类和整体的预测都有较高的精度。未来的工作如下, 由于预警时间也是异常预警的关键, 因此

在保证准确率的同时必须确保模型能满足异常预警的时间要求, 同时优化网络模型参数, 进一步提升模型对异常工况预警的精度。

## 参考文献:

- [1] ZHANG Bin, GAO Xianwen, LI Xiangyu. Complete simulation and fault diagnosis of sucker-rod pumping (includes associated comment)[J]. *SPE production & operations*, 2021, 36(2): 277–290.
- [2] DUAN Jimiao, LIU Huishu, JIANG Junzhe, et al. Numerical prediction of wax deposition in oil-gas stratified pipe flow[J]. *International journal of heat and mass transfer*, 2017, 105: 279–289.
- [3] ALSAIHATI A, ELKATATNY S, MAHMOUD A A, et al. Use of machine learning and data analytics to detect downhole abnormalities while drilling horizontal wells, with real case study[J]. *Journal of energy resources technology*, 2021, 143(4): 043201.
- [4] ÖZMEN Ö, SINANOĞLU C, CALISKAN A, et al. Prediction of leakage from an axial piston pump slipper with circular dimples using deep neural networks[J]. *Chinese journal of mechanical engineering*, 2020, 33(1): 28.
- [5] RAHMANIFARD H, PLAKSINA T. Application of artificial intelligence techniques in the petroleum industry: a review[J]. *Artificial intelligence review*, 2019, 52(4): 2295–2318.
- [6] ZHAO Minghang, KANG M, TANG Baoping, et al. Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes [J]. *IEEE transactions on industrial electronics*, 2018, 65(5): 4290–4300.
- [7] ZHANG Yiyi, LI Jiaxi, FAN Xianhao, et al. Moisture prediction of transformer oil-immersed polymer insulation by applying a support vector machine combined with a genetic algorithm[J]. *Polymers*, 2020, 12(7): 1579.
- [8] 周斌, 王延江, 刘伟锋, 等. 基于 Hessian 正则化支持向量机的多视角协同识别抽油机井工况方法 [J]. *石油学报*, 2018, 39(12): 1429–1436.  
ZHOU Bin, WANG Yanjiang, LIU Weifeng, et al. A working condition recognition method of sucker-rod pumping wells based on Hessian-regularized SVM and multi-view co-training algorithm[J]. *Acta petrolei sinica*, 2018, 39(12): 1429–1436.
- [9] 王利君, 支志英, 贾鹿, 等. 基于 SCRF 的抽油井结蜡预测方法优化研究 [J]. *计算机科学*, 2019, 46(S2): 599–603.  
WANG Lijun, ZHI Zhiying, JIA Lu, et al. Study on optimized method for predicting paraffin deposition of pumping wells based on SCRF[J]. *Computer science*,



- 2019, 46(S2): 599–603.
- [10] LEI Yaguo, YANG Bin, JIANG Xinwei, et al. Applications of machine learning to machine fault diagnosis: a review and roadmap[J]. *Mechanical systems and signal processing*, 2020, 138: 106587.
- [11] WEI Jingliang, GAO Xianwen. Fault diagnosis of sucker rod pump based on deep-broad learning using motor data[J]. *IEEE access*, 2020, 8: 222562–222571.
- [12] CHEN Xiaohan, ZHANG Beike, GAO Dong. Bearing fault diagnosis base on multi-scale CNN and LSTM model[J]. *Journal of intelligent manufacturing*, 2021, 32(4): 971–987.
- [13] CABRERA D, GUAMÁN A, ZHANG Shaohui, et al. Bayesian approach and time series dimensionality reduction to LSTM-based model-building for fault diagnosis of a reciprocating compressor[J]. *Neurocomputing*, 2020, 380: 51–66.
- [14] 魏晓良, 潮群, 陶建峰, 等. 基于 LSTM 和 CNN 的高速柱塞泵故障诊断 [J]. *航空学报*, 2021, 42(3): 423876.  
WEI Xiaoliang, CHAO Qun, TAO Jianfeng, et al. Cavitation fault diagnosis method for high-speed plunger pumps based on LSTM and CNN[J]. *Acta aeronautica et astronautica sinica*, 2021, 42(3): 423876.
- [15] LIANG Rong, CHANG Xintan, JIA Pengtao, et al. Mine gas concentration forecasting model based on an optimized BiGRU network[J]. *ACS omega*, 2020, 5(44): 28579–28586.
- [16] SHEN Wei, WANG Xinggang, WANG Yan, et al. DeepContour: a deep convolutional feature learned by positive-sharing loss for contour detection[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA, 2015: 3982–3991.
- [17] CHEN Yushi, JIANG Hanlu, LI Chunyang, et al. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks[J]. *IEEE transactions on geoscience and remote sensing*, 2016, 54(10): 6232–6251.
- [18] LU Wenjie, LI Jiazheng, WANG Jingyang, et al. A CNN-BiLSTM-AM method for stock price prediction[J]. *Neural computing and applications*, 2021, 33(10): 4741–4753.
- [19] DIAZ M, MOETESUM M, SIDDIQI I, et al. Sequence-based dynamic handwriting analysis for Parkinson's disease detection with one-dimensional convolutions and BiGRUs[J]. *Expert systems with applications*, 2021, 168: 114405.
- [20] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[EB/OL]. (2017-05-12)[2021-05-01]. <https://arxiv.org/abs/1706.03762>.
- [21] JAGVARAL B, LEE W K, ROH J S, et al. Path-based reasoning approach for knowledge graph completion using CNN-BiLSTM with attention mechanism[J]. *Expert systems with applications*, 2020, 142: 112960.

#### 作者简介:



李克文, 教授, 博士生导师, 博士, CCF 会员, 主要研究方向为人工智能、软件工程、数据挖掘、深度学习。发表学术论文 100 余篇。



杜苾聪, 硕士研究生, 主要研究方向为深度学习、故障检测。