

## 基于自注意力机制与 1D-CNN 的变压器故障诊断方法

刘国柱

安徽理工大学 电气与信息工程学院, 安徽 淮南 232001

**摘要:**目的 变压器是电力系统中重要的设备,其发生故障时能够被有效地判别出故障类别,使得电力检修效率提升,这对电网的安全运行具有重要意义。针对电网电力检修中出现的变压器故障判别精度不足这一问题,提出了基于自注意力机制与 1D-CNN 的变压器故障诊断方法。常规卷积在处理 DGA 气体样本数据时容易损失特征信息,导致故障诊断的准确率偏低,论文将自注意力机制与 1D-CNN 结合,有效改善了上述问题,提高了变压器故障诊断的准确率和可靠性。**方法** 为减少卷积网络提取到的特征提取信息在模型层间传播时造成的损失,论文在 1D-CNN 的基础上使用 LeakyReLU 函数替代原模型中的 ReLU 激活函数,相比于 ReLU 激活方式下很多神经元都没有被激活,LeakyReLU 可以降低模型的稀疏性,使得网络特征信息多样性增加。自注意力机制可实现对变压器油中溶解气体数据的特征信息加权处理,实现了有效特征信息增强作用,采用动态衰减学习率策略对优化器进行优化。**结果** 所提的方法损失率可降低至 0.078,相比于无动态衰减学习率和 ReLU 激活方式,损失率分别降低了 44.7% 和 38.6%;诊断准确率可达到 93.79%,较 1D-CNN 和 GOA-BP 方法诊断准确率提高了 0.36% 和 2.12%。**结论** 算例仿真验证了所提方法的有效性和优越性,表明基于自注意力机制与 1D-CNN 的变压器故障诊断方法能有效提高诊断的准确率,降低模型的损失率。

**关键词:** 变压器;故障诊断;自注意力机制;卷积神经网络

**中图分类号:** TM407 **文献标识码:** A **doi:** 10.16055/j.issn.1672-058X.2025.0001.010

### Transformer Fault Diagnosis Method Based on Self-attention Mechanism and 1D-CNN

LIU Guozhu

School of Electrical and Information Engineering, Anhui University of Science and Technology, Anhui Huainan 232001, China

**Abstract: Objective** Transformers are important equipment in the power system. Effective identification of fault categories when transformers fail can improve the efficiency of power maintenance, which is of great significance for the safe operation of the power grid. In response to low accuracy in transformer fault identification in power grid maintenance, this paper proposed a transformer fault diagnosis method based on a self-attention mechanism and 1D-CNN. Conventional convolution often loses feature information when processing DGA gas sample data, resulting in low accuracy in fault diagnosis. By combining the self-attention mechanism with 1D-CNN, this paper effectively addressed the above issues, improving the accuracy and reliability of transformer fault diagnosis. **Methods** To reduce the loss of feature extraction information during inter-layer propagation in convolutional networks, this paper replaced the ReLU activation function in the original model with the LeakyReLU function. Compared with ReLU activation, where many neurons are not activated, LeakyReLU can reduce the sparsity of the model and increase the diversity of network feature information. The self-attention mechanism can weight the feature information of dissolved gas data in transformer oil, realizing effective feature information enhancement. A dynamic decay learning rate strategy was used to optimize the optimizer. **Results** The proposed method reduced the loss rate to 0.078, a decrease of 44.7% and 38.6% compared with models without dynamic

**收稿日期:** 2023-04-11 **修回日期:** 2023-06-21 **文章编号:** 1672-058X(2025)01-0072-07

**作者简介:** 刘国柱(1997—),男,安徽合肥人,硕士,从事电力设备状态感知及人工智能应用研究。

**引用格式:** 刘国柱. 基于自注意力机制与 1D-CNN 的变压器故障诊断方法[J]. 重庆工商大学学报(自然科学版), 2025, 42(1): 72-78.

LIU Guozhu. Transformer fault diagnosis method based on self-attention mechanism and 1D-CNN [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2025, 42(1): 72-78.

decay learning rate and ReLU activation, respectively. The diagnostic accuracy can reach 93.79%, an increase of 0.36% and 2.12% compared with 1D-CNN and GOA-BP methods, respectively. **Conclusion** Case study simulations validate the effectiveness and superiority of the proposed method, demonstrating that the transformer fault diagnosis method based on the self-attention mechanism and 1D-CNN can effectively improve diagnostic accuracy and reduce model loss.

**Keywords:** transformer; fault diagnosis; self-attention mechanism; convolutional neural network

## 1 引言

电力变压器在电力系统中起着电能变换的重要作用,其不正常运行状态将会给电力系统的安全性和稳定性带来威胁<sup>[1]</sup>。因此,建立一套有效的故障诊断流程尤为重要,这能使基层电力巡检维护人员及时有效地诊断出故障并进行处理,保证供电的可靠性<sup>[2-3]</sup>。

变压器内部发生绝缘老化时,故障一般可分为两类:即过热故障和放电故障,过热故障按温度高低,可分为低温过热,中温过热与高温过热 3 种;放电故障按照能量密度的不同,可分为局部放电,低能放电和高能放电 3 种。发生上述故障时,变压器中产生的  $H_2$ 、 $CH_4$ 、 $C_2H_6$ 、 $C_2H_4$ 、 $C_2H_2$  等气体会溶解在绝缘油中,通过对溶于油中气体成分进行分析就能够反映变压器当前运行状态,这种方法被称为油中溶解气体分析法(Dissolved Gas Analysis, DGA)<sup>[4-6]</sup>。

变压器故障诊断的方法主要有直观检查方法、电气预防性试验方法、专家系统(TFDES)<sup>[7]</sup>及人工神经网络法(TFDANN)等。其中,油中溶解气体分析法是最常用的一种方法,它根据变压器油中溶解气体的浓度来判断变压器内部是否发生故障以及故障类型,然而这种方法存在一些缺点,例如,诊断准确率不高、受外界因素影响大、无法反映故障的实时状态等<sup>[8]</sup>。

近年来随着人工智能领域的发展,将机器学习算法和实际采集到的数据进行结合使得故障诊断取得了更好的效果。目前,国内外学者在进行多分类故障诊断研究时常用的有支持向量机(SVM)<sup>[9]</sup>、概率神经网络(PNN)<sup>[10]</sup>及 BP 神经网络<sup>[11]</sup>等机器学习算法,SVM 是一种二分类模型,对于多分类问题需要将某转化为多个二分类问题会增加计算复杂度和运算时间;PNN 是一种基于概率密度函数的神经网络,它对数据的分布假设较强,若数据不符合高斯分布或者类别之间的方差不相等,那么 PNN 的分类效果会下降,受到样本分布影响较大;BP 神经网络是一种基于误差反向传播的神经网络,容易陷入局部最优解,且容易受到学习率的影响导致收敛性不足误差较大。

深度算法作为机器学习的细分领域,有卷积神经网络(CNN)、循环神经网络(RNN)、深度置信网络(DBN)等,RNN 主要用于处理文本、语音等时间序列数据,对于空间数据的处理能力较弱,不能很好地提取数

据的局部特征;DBN 是一种基于无监督预训练的深度学习学习方法,其预训练过程比较耗时,且对参数的选择较为敏感;CNN 因为其自身卷积核具有稀疏性,所以在参数提取过程中更加注重对于有用特征数据的处理,基于此,本文在 1D-CNN 的基础上加入自注意力机制(Self-Attention),并对激活函数与学习率进行优化,改进后的模型在对特征信息进行提取时,通过保留一定的负值权重实现对于负值有效信息的利用,并通过自注意力机制对于所提取到的特征数据按重要程度进行权重分配,增强了对有效信息的利用,实现模型损失率的降低,实现了对于变压器故障多分类问题的有效判别。

## 2 相关算法介绍

### 2.1 一维卷积神经网络模型

CNN 是深度学习理论中最成功的一个分支,通过多网络结构、卷积运算和降采样实现对信号深层信息提取,具有稀疏连接和权重共享两大突出特点,有效地降低了网络复杂度和过拟合风险<sup>[12]</sup>。在特征提取和运算处理速度方面相较于其他神经网络有较大优势。

国内外学者常将一维数据转化为二维图像,通过间接对图像的处理实现故障分类,Waibel 等<sup>[13]</sup>于 1987 年提出时间延迟网络,将 1D-CNN 应用于一维数据方面的研究,从此开启了 1D-CNN 的发展之路。相比于普通卷积神经网络,1D-CNN 的输入通常为二维数组或包含多个通道的二维数组,其卷积核只朝着一个方向进行移动,卷积的填充方式包含有效填充(Valid padding)和相同填充(Same padding)两种。Valid padding 时,卷积核只对有效位置进行扫描;Same padding 时,填充数值由算法依据卷积核大小计算,若步长(stride,  $s$ )为 1,则特征输出形状和输入形状一致,若  $s > n$  且  $n$  为正整数,则卷积核进行降采样填充,输出缩小为原来的  $1/n$  倍。

如图 1 所示,  $X^i$  和  $X^o$  分别为输入特征和输出特征,  $K$  为卷积核组;  $w_1$  和  $h_1$  分别为输入特征的特征通道数和序列长度;  $\otimes$  代表进行一维卷积操作,  $q$  为卷积核大小,卷积核的大小决定了特征提取时所对应的局部时间窗的长度;  $h_2$  为输出特征的序列长度,  $A$  为非线性激活函数;  $w_2$  为卷积核个数,其决定输出特征的通道数;此外,卷积核参数还包括卷及步长(stride,  $s$ )和填充方式(padding,  $p$ ),图中  $s$  为 2,  $s$  的参数设置若大于 1 则是希望减小输入参数的数目,减少计算量。图中的

1D-CNN 卷积操作过程用公式表达为

$$A(x)_{\text{Relu}} = \text{Max}\{0, x\}$$

$$X^o = A(K \otimes X^i + B)$$

$$X^o = X_1^o \oplus X_2^o \oplus X_3^o \oplus \dots \oplus X_{w_2}^o$$

其中,  $\oplus$  为矩阵拼配运算符号,  $B$  为卷积偏置, 卷积层激活函数  $A(x)$  常选用 Relu。

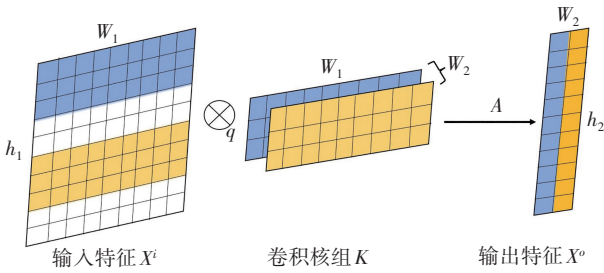


图 1 1D-CNN 处理过程

Fig. 1 The process of 1D-CNN processing

常规一维卷积的填充方式分为有效填充和相同填充。有效填充时,卷积核只扫描输入的特征图有效窗口位置,其特征输出如下:

$$h_2 = 1 + (h_1 - q) / s$$

当进行相同填充,填充的值由算法内部依据卷积核大小计算,当步长  $s = 1$  时,能够保证特征输出尺寸和输入尺寸相等,若  $s > n$  时且  $n$  为正整数,则卷积核进行降采样填充,输出特征为输入特征的  $1/n$  倍,其特征输出如下:

$$h_2 = 1 + (h_1 + 2p - q) / s$$

### 2.2 自注意力机制

自注意力机制可将数据的不同部分赋予权重,依据各部分所占权重获得相应的关注,并对其进行有效的运算及判断<sup>[14-15]</sup>,这提高了计算机对有效数据的运算效率,其原理如图 2 所示。

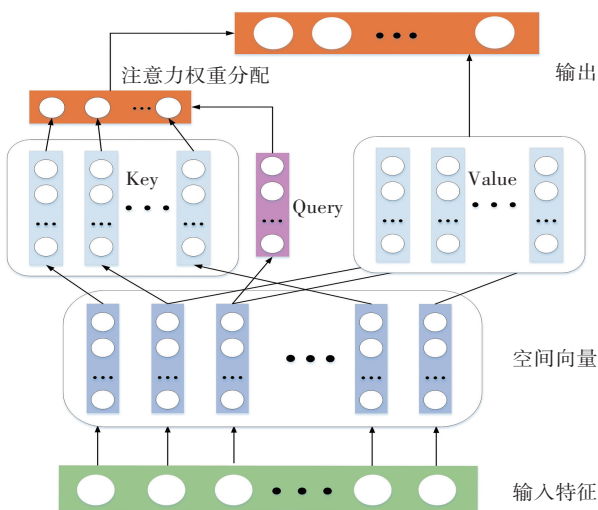


图 2 自注意力机制原理图

Fig. 2 Schematic diagram of the self-attention mechanism

图 2 中,Query 为查询向量,Key 为键向量,Value 为

值向量,  $d$  是维度。经 Embedding 线性变换后的输入的特征变换为上述空间向量,Query 通过在 Key 上投影获得数值实现对向量之间相关度的判断,两者的矩阵内积后通过 softmax 函数将相关度进行归一化得到注意力权重分配,将其与 Value 进行加权求和,最终实现对于高权重值的特征信息量的保留。上述原理的数学表达式为

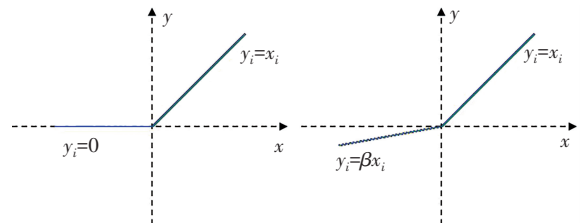
$$\text{Self\_Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_K}}\right)V$$

### 3 自注意力机制与 1D-CNN 模型

论文模型的搭建均基于 Python 语言 TensorFlow2.5 框架,并通过 Jupyter notebook 平台上进行仿真计算,计算机的硬件配置为 Intel(R) Core(TM) i5-7300HQ CPU@2.50 GHz, GTX1050 GPU。

#### 3.1 模型的改进

卷积神经网络中常采用激活函数 ReLU 来解决梯度消失问题,即在运算过程中输入的数值大于 0,则保持数值输出;若输入小于 0 时,将原数值置零输出,这会使对应的权重和偏置参数此次无法得到更新,同时也会损失所提取到的负值有用信息。为此,论文提出以 LeakyReLU 替代 ReLU 来解决上述问题。LeakyReLU 可以有效地解决神经元输入特征信息为负值时梯度不更新的问题,信息输入数值为正数时进行完全输出,在反向传播过程中出现负值输入时,通过设置较小的权重  $\beta$  来保留一定的负值信息,在进行计算后得到梯度更新信息,能有效消除梯度方向锯齿现象,原理如图 3 所示。



(a) ReLU 激活 (b) LeakyReLU 激活

图 3 激活函数

Fig. 3 Activation functions

在训练过程中,固定的学习率会使得模型训练速度缓慢,容易陷入局部最优点或鞍点从而导致稳定性不足问题,本文在 Adam 优化器的基础上采用动态衰减学习率策略对优化器进行优化,公式如下:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \delta}} \hat{m}_t$$

$$lr_{t,i} = \frac{\eta}{\sqrt{\hat{v}_{t,i} + \delta}}$$

$$lr_{t,i} = \frac{\eta}{\sqrt{\hat{v}_{t,i} + \delta}} * decay\_rate^{\left(\frac{global\_step}{decay\_steps}\right)}$$

其中,  $g_t$  是第  $t$  轮的梯度,  $m_t$  和  $v_t$  是梯度的一阶矩和二阶矩的指数移动平均,  $\hat{m}_t$  和  $\hat{v}_t$  是偏差修正后的一阶矩和二阶矩,  $\theta_t$  是第  $t$  轮的参数,  $\eta$  是初始学习率,  $\beta_1$  和  $\beta_2$  是移动平均的衰减率, 这里将  $\eta$  和  $\beta$  参数分别设为 0.04 和 0.92,  $\epsilon$  是一个很小的数, 用来防止除以零。

Adam 优化器的学习率不是固定的, 而是根据每个参数的梯度变化而调整。学习率等于初始学习率除以梯度平方的平均值的平方根, 其中  $i$  表示第  $i$  个参数,  $lr_{t,i}$  表示第  $t$  轮第  $i$  个参数的学习率,  $lr_{0,i}$  是第  $i$  个参数的初始学习率,  $decay$  是一个衰减函数, 可以根据  $t$  或其他因素来计算新的学习率, 其中  $decay\_rate$  和  $decay\_steps$  是超参数, 用于控制衰减速度和周期。改进前后模型训练的损失率变化对比图如图 4 所示。

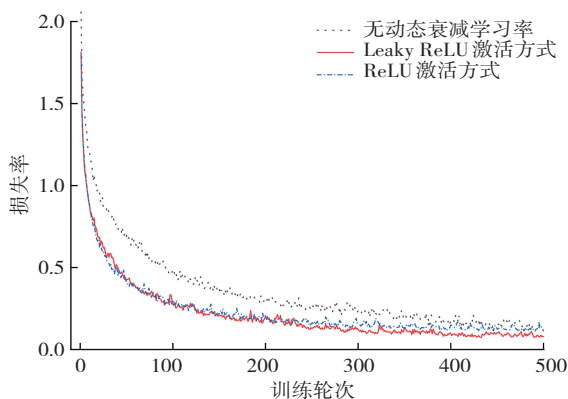


图 4 不同激活函数下模型训练的损失率

Fig. 4 Loss rate of model training with different activation functions

由图 4 曲线可知, 在无动态衰减学习率、有动态学习率的 LeakyReLU 激活方式和 ReLU 激活方式下, 模型的训练损失率分别为 0.141、0.078 和 0.127, 未添加动态衰减学习率的模型在训练时效率较低, 无法有效地加速模型的迭代过程; 而在 ReLU 激活方式下, 会将所提取到的负值有效信息全部置零从而增大损失率, 从而导致模型诊断精度的降低。优化器采用动态衰减学习率策略和 Leaky ReLU 激活方式可以明显地降低模型的损失, 相比于无动态衰减学习率和 ReLU 激活方式, 损失率分别降低了 44.7% 和 38.6%。

为了分析不同优化器的选取对本文所搭建的模型诊断性能影响, 将上文所设置的参数与不同优化器进行对比, 这里的批尺寸参数均设置为 32, 对比仿真结果

如图 5 所示。

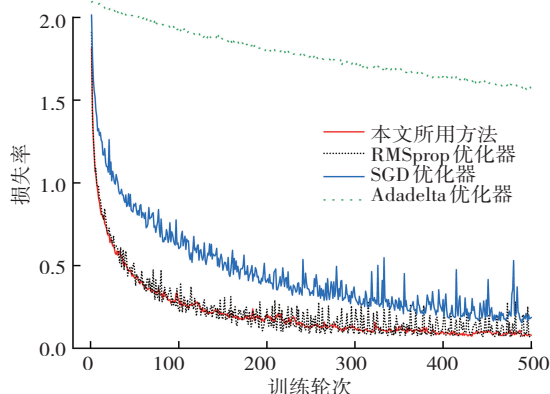


图 5 不同优化器训练效果对比

Fig. 5 Comparison of training effects of different optimizers

由图 5 中可以看出, 优化器的选取对于模型的训练效果有较大影响, 当选用 Adadelta 优化器时, 模型训练中的梯度下降缓慢且最终无法收敛, 难以用于故障诊断; RMSprop 和 SGD 优化器在训练过程中无法动态地调整学习率的下降速度, 从而导致梯度更新不稳定, 相较于 Adam 优化器, 两者收敛稳定性不足。

### 3.2 模型结构与参数设置

一维卷积神经网络模型常用于提取输入样本数据的局部特征, 并且通过堆叠层数来增加感受野和抽象能力。然而, 一维卷积神经网络模型也有一些局限性, 随着层数的增加, 梯度消失和信息损失的问题也会加剧。为了解决上述问题, 论文将自注意力机制与 1D-CNN 模型结合, 以增强一维卷积神经网络模型的表达与泛化能力。在卷积层之后添加自注意力机制, 这样可以让卷积层提取局部特征, 而自注意力层提取全局特征, 可以让自注意力层在每个卷积层之后对特征进行重新组合, 并增强特征信息的权重。

诊断模型结构图如图 6 所示。图 6 中, Normalization 层的作用是对神经网络的输入或输出进行规范化, 使分布更加接近标准正态分布, 有助于加速训练过程和解决梯度消失问题, 在 Normalization 层后加入自注意力机制层, 在特征信息经过规范化处理后, 将数据的各部分赋予相应的权重, 再根据所占权重获得相应的关注。Flatten 层能将多维的输入一维化, 常用于卷积层到全连接层的过渡, 可以简化数据的表示, 减少计算量, 方便后续的全连接层进行处理。

两个 Conv1D 的卷积核均为 5, 卷积核核数分别为 64 和 128, MaxPool1D 中 pool\_size 为 2, 两层 Dense 神经元个数为分别为 256 和 7, Activation 层中激活函数为 LeakyReLU, 其权重参数  $\beta$  为 0.01, 两个 Self-Attention 模块的注意力头维度分别为 64 和 128。

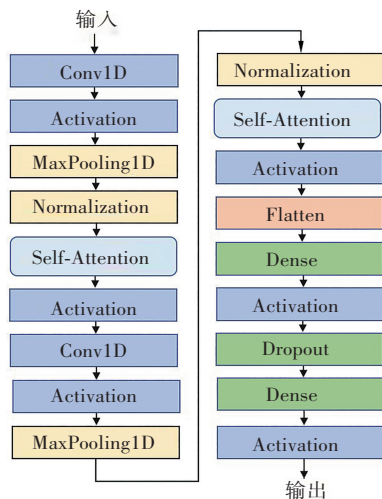


图 6 故障诊断模型结构图

Fig. 6 Structure of the fault diagnosis model

#### 4 仿真实验与结果分析

论文数据集由某市电网供电公司提供,数据集中包含氢气( $H_2$ )、甲烷( $CH_4$ )、乙烷( $C_2H_6$ )、乙烯( $C_2H_4$ )、乙炔( $C_2H_2$ )五种气体,共 7 种样本类型,分别为 6 种故障状态和正常状态,6 种故障状态分别为:低温过热、中温过热、高温过热、低能放电、局部放电、高能放电。实际测量共计 600 例样本,对实际样本中进行人工筛选,选出 550 条有效样本,DGA 数据样本图如图 7 所示。

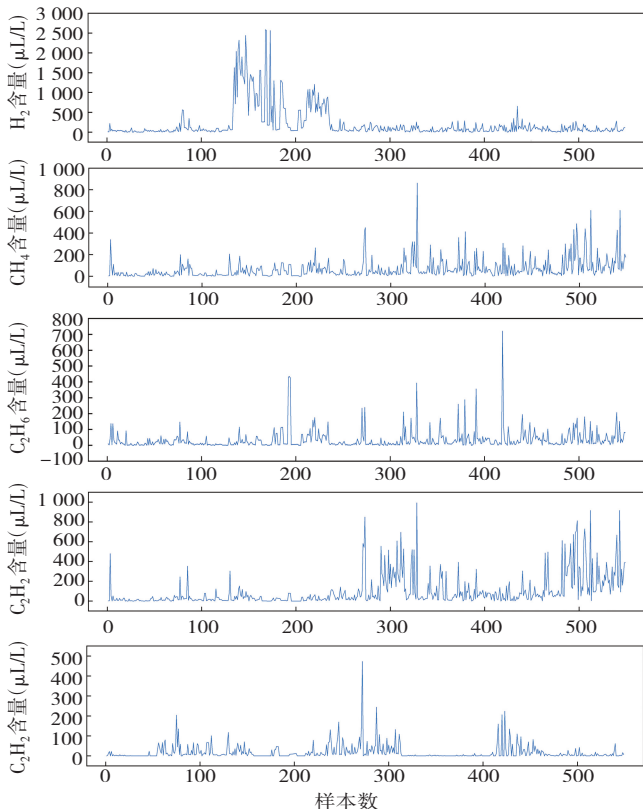


图 7 DGA 数据样本图

Fig. 7 Sample plot of DGA data

#### 4.1 模型的训练与测试

将样本按照 4 : 1 划分为训练集和测试集,样本数据划分及编码如表 1 所示。由于样本数据集有限,难以在统计学上代表数据,故采取 4 折验证法,最终的模型参数为 4 个模型参数取值的平均数,其原理如图 8 所示。

表 1 样本数据划分及编码

Table 1 Division and coding of sample data

状态类别	训练样本	测试样本	编 码
高能放电	60	14	1
低能放电	60	14	2
局部放电	66	16	3
高温过热	66	18	4
中温过热	64	17	5
低温过热	62	15	6
正常状态	62	16	7



图 8 四折验证法

Fig. 8 Four-fold verification method

对模型进行训练和测试后,得到图 9。图 9 中可看出在训练轮次达到 100 时,基于自注意力机制与 1D-CNN 的诊断模型的训练集准确率已达到 88.7%,测试集准确率达到 86.51%,在训练轮次达到 400 次时,模型的准确率趋近于恒定值,说明模型已经收敛。多次仿真后,模型的平均训练准确率和平均测试准确率分别为 95.94% 和 93.68%。

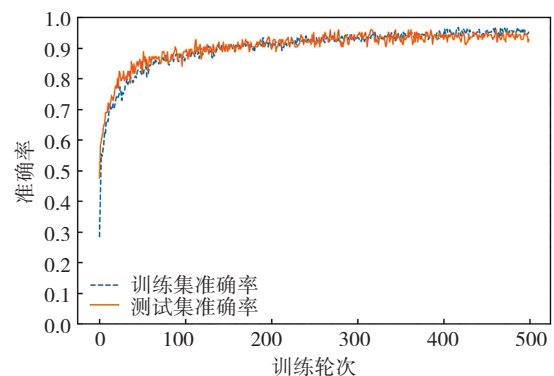


图 9 模型的训练与测试

Fig. 9 Training and testing of the model

以所搭建的模型进行实际故障诊断,将 DGA 特征气体含量数据进行诊断测试,部分实例如表 2 所示。

表 2 变压器故障诊断方法部分实例

Table 2 Selected examples of transformer fault diagnosis methods

DGA 特征气体含量 ( $\mu\text{L} \cdot \text{L}^{-1}$ )					实际故障	诊断故障
H <sub>2</sub>	CH <sub>4</sub>	C <sub>2</sub> H <sub>6</sub>	C <sub>2</sub> H <sub>4</sub>	C <sub>2</sub> H <sub>2</sub>	类型	类型
255.7	104.4	21.3	117.3	169.4	高能放电	高能放电
150	11.6	29	50.6	1.1	局部放电	局部放电
91	307	78	313	1.4	高温过热	高温过热
66.7	97.6	55.8	92.8	0.1	低温过热	低温过热
86	112	67	46	85.8	低能放电	低能放电
193	29	6	13	14	高能放电	高能放电
6.9	80.6	47.4	351	4.9	高温过热	高温过热
9.1	24	5.2	8.8	90.8	低能放电	正常状态
13.6	5.3	8.2	29	2.1	正常状态	正常状态
193	29	6	13	14	高能放电	高能放电
242	24.4	5.5	5.1	0.1	局部放电	局部放电
31.2	150	50	208	1.4	高温过热	高温过热
3.7	153.8	99	195.2	0.1	中温过热	中温过热
18	137	50	152	0.4	中温过热	中温过热
46	12	33.3	9	0.8	正常状态	正常状态

为直观分析模型故障诊断的分类性能,将实际样本放入模型中进行预测,重复多次试验后,结果如图 10 所示。图 10 中,低能放电和局部放电状态全部样本均预测正确,诊断正确率可达 100%,高能放电、中温过热和高温过热样本中均只有一例样本误判,实际 110 个样本中,103 例样本诊断后可获得正确的诊断结果,仿真实验有效验证了基于自注意力机制与 1D-CNN 的变压器故障诊断方法的性能。

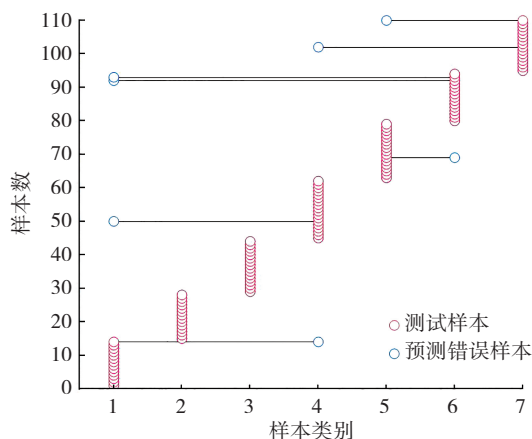


图 10 模型的实际样本预测

Fig. 10 Actual sample predictions from the model

#### 4.2 算例对比

为了验证基于自注意力机制与 1D-CNN 的变压器故障诊断方法的优越性,本文搭建 1D-CNN 和 LSTM 神经模型并与文献[16]中的蝗虫优化神经网络(GOA-BP)作对比,用以上算法做故障诊断模型,各优化算法模型的相关参数如表 3 所示。

表 3 各算法相关参数对比

Table 3 Comparison of relevant parameters for each algorithm

模型	$P_{accuracy}/\%$	$P_{recall}/\%$	$Mean\_accuracy/\%$
论文方法	93.79	89.98	93.68
1D-CNN	93.43	88.61	93.09
LSTM	85.57	84.58	85.02
GOA-BP	91.67	—	91.67

在此采用 Accuracy、Recall、Mean\_accuracy 三个参数作为评估模型稳定性的标准,Accuracy 是准确率,是指预测正确的样本数占测试所有的样本数的比,Mean\_Accuracy 为平均准确率,用以衡量模型的平均诊断性能,Recall 为召回率,即正确预测正分类的样本数占测试的所有正样本的比例比,False Positive (FP,  $N_{FP}$ ) 为被判定成正确样本的错误样本,False Negative (FN,  $N_{FN}$ ) 为被判定成错误样本的正确样本,True Positive (TP,  $N_{TP}$ ) 为判定成正确样本的正确样本,True Negative (TN,  $N_{TN}$ ) 为被判定为错误样本的错误样本,相关计算公式如下:

$$P_{accuracy} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}}$$

$$P_{recall} = \frac{N_{TP}}{N_{TP} + N_{FN}}$$

表 3 中所使用的方法对于变压器故障多分类诊断的平均准确率为 93.68%,测试准确率可达 93.79%,比 1D-CNN、LSTM 和 GOA-BP 的预测准确率分别高出 0.36%、8.22% 和 2.12%,算例对比表明论文所提方法的有效性和优越性。

### 5 结论

为有效提高变压器的故障诊断精度,本文提出将 1D-CNN 与自注意力机制结合实现对变压器的故障诊断,通过自注意力机制对所提取出的 DGA 气体数据进行合理的权重分配,激活函数方式及学习率进行改进优化后,有效地降低了模型训练和测试中的损失率,提升了模型诊断性能,所提故障诊断方法诊断准确率可达到 93.79%,损失率降低至 0.078。相较于 1D-CNN、LSTM 和 GOA-BP 模型,本文所提模型预测准确率分别高出 0.36%、8.22% 和 2.12%,平均准确率分别高出 0.59%、8.66% 和 2.01%。本文所提方法可为解决变压器故障诊断精度不足的问题提供一个新思路。

本文的研究仅聚焦于对于电力变压器的故障诊断,在后续的工作中,针对其他相关电气设备的相关特性,把本文所用方法改进后与其结合,可探索对于多类型电气设备的故障诊断。

#### 参考文献(References):

[1] 代杰杰, 宋辉, 杨伟, 等. 基于油中气体分析的变压器故障诊断 ReLU-DBN 方法[J]. 电网技术, 2018, 42(2): 658-664.

- DAI Jie-jie, SONG Hui, YANG Yi, et al. Dissolved gas analysis of insulating oil for power transformer fault diagnosis based on ReLU-DBN[J]. Power System Technology, 2018, 42(2): 658-664.
- [2] 李宗博, 焦在滨, 何安阳. 基于卷积神经网络特征迁移策略的变压器智能保护方法[J]. 中国电机工程学报, 2021, 41(15): 5201-5212.  
LI Zong-bo, JIAO Zai-bin, HE An-yang. Intelligent protection for power transformer using convolutional neural network integrated into features transferring strategy[J]. Proceedings of the CSEE, 2021, 41(15): 5201-5212.
- [3] 刘志坚, 何蔚, 刘航, 等. 基于格拉姆角场变换和深度压缩模型的变压器故障识别方法[J]. 电网技术, 2023, 47(4): 1478-1490.  
LIU Zhi-jian, HE Wei, LIU Hang, et al. Fault identification method for power transformer based on Gramian angular field transformation and deep compression model[J]. Power System Technology, 2023, 47(4): 1478-1490.
- [4] 李辉, 张志攀, 张中卫. 基于卷积神经网络的变压器故障诊断[J]. 河南理工大学学报(自然科学版), 2018, 37(6): 118-123, 154.  
LI Hui, ZHANG Zhi-pan, ZHANG Zhong-wei. Fault diagnosis of transformer based on convolutional neural network[J]. Journal of Henan Polytechnic University (Natural Science), 2018, 37(6): 118-123, 154.
- [5] 刘可真, 苟家莫, 骆钊, 等. 基于粒子群优化-长短期记忆网络模型的变压器油中溶解气体浓度预测方法[J]. 电网技术, 2020, 44(7): 2778-2785.  
LIU Ke-zhen, GOU Jia-qi, LUO Zhao, et al. Prediction of dissolved gas concentration in transformer oil based on PSO-LSTM model[J]. Power System Technology, 2020, 44(7): 2778-2785.
- [6] 张育杰, 冯健, 李典阳, 等. 基于油色谱数据的变压器故障征兆新优选策略[J]. 电网技术, 2021, 45(8): 3324-3332.  
ZHANG Yu-jie, FENG Jian, LI Dian-yang, et al. New feature selection method for transformer fault diagnosis based on DGAdata[J]. Power System Technology, 2021, 45(8): 3324-3332.
- [7] 徐阳, 谢天喜, 周志成, 等. 基于多维度信息融合的实用型变压器故障诊断专家系统[J]. 中国电力, 2017, 50(1): 85-91.  
XU Yang, XIE Tian-xi, ZHOU Zhi-cheng, et al. Practical expert system for transformer fault diagnosis based on multi-dimensional information fusion technology[J]. Electric Power, 2017, 50(1): 85-91.
- [8] 沙伟燕, 李秀广, 何宁辉, 等. 基于油色谱时频域信息和残差注意网络的变压器故障诊断方法[J]. 电网与清洁能源, 2022, 38(1): 66-75.  
SHA Wei-yan, LI Xiu-guang, HE Ning-hui, et al. Transformer fault diagnosis method based on oil chromatogram time frequency domain information and residual attention network[J]. Power System and Clean Energy, 2022, 38(1): 66-75.
- [9] 邵凯旋, 何怡刚, 汪磊. 基于多尺度熵分析与改进 SVM 的变压器故障识别[J]. 电子测量与仪器学报, 2022, 36(6): 161-168.  
SHAO Kai-xuan, HE Yi-gang, WANG Lei. Fault identification of transformer based on multiscale entropy analysis and improved SVM[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(6): 161-168.
- [10] 安源, 张智恒. 基于 PNN 的电力变压器故障诊断[J]. 电气应用, 2020, 39(11): 12-17.  
AN Yuan, ZHANG Zhi-heng. Fault diagnosis of power transformer based on PNN[J]. Electrotechnical Application, 2020, 39(11): 12-17.
- [11] 郭林, 唐晶, 唐黎哲, 等. 一种基于改进 BP 神经网络的变压器故障诊断方法[J]. 控制与信息技术, 2021(5): 71-77.  
GUO Lin, TANG Jing, TANG Li-zhe, et al. A method of transformer fault diagnosis based on improved BP neural network[J]. Control and Information Technology, 2021(5): 71-77.
- [12] 何成兵, 王润泽, 张霄翔. 基于改进一维卷积神经网络的汽轮发电机组轴系扭振模态参数辨识[J]. 中国电机工程学报, 2020, 40(S1): 195-203.  
HE Cheng-bing, WANG Run-ze, ZHANG Xiao-xiang. Modal parameters identification of torsional vibration of turbogenerator shafting based on improved one-dimensional convolution neural network[J]. Proceedings of the CSEE, 2020, 40(S1): 195-203.
- [13] WAIBEL A, HANAZAWA T, HINTON G, et al. Phoneme recognition using time-delay neural networks[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989, 37(3): 328-339.
- [14] 万黎明, 张小乾, 刘知贵, 等. 基于空洞空间金字塔池化和多头自注意力的特征提取网络[J]. 计算机应用, 2022, 42(S2): 79-85.  
WAN Li-ming, ZHANG Xiao-qian, LIU Zhi-gui, et al. Feature extraction network based on pyramid pooling in empty space and multi-head self-attention[J]. Journal of Computer Applications, 2022, 42(S2): 79-85.
- [15] 张辉宜, 张进, 黄俊. 基于图注意力网络的多标签图像分类模型[J]. 重庆工商大学学报(自然科学版), 2022, 39(1): 34-41.  
ZHANG Hui-yi, ZHANG Jin, HUANG Jun. Multi-label image classification model based on graph attention network[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2022, 39(1): 34-41.
- [16] 徐新, 蒋波涛, 曹雯. 蝗虫优化神经网络在变压器故障诊断中的应用[J]. 电网与清洁能源, 2021, 37(5): 17-23.  
XU Xin, JIANG B-otao, CAO Wen. Application of grasshopper optimization neural network in power transformer fault diagnosis[J]. Power System and Clean Energy, 2021, 37(5): 17-23.

责任编辑:陈芳