

# 面向电力变压器油中溶解气体的卷积神经网络 诊断方法

裴小邓, 罗 林, 陈 帅, 王 乔

(辽宁石油化工大学 信息与控制工程学院, 辽宁 抚顺 113001)

**摘 要:** 油中溶解气体分析法(Dissolved Gas Analysis, DGA)是判断变压器内部故障的重要方法之一。针对传统基于浅层的机器学习方法在变压器故障诊断中存在的特征提取和泛化能力方面的不足,提出了一种基于卷积神经网络的变压器故障诊断方法。利用网络中的卷积层对油中溶解气体进行特征转换,结合池化层强化重要特征的能力,对故障敏感特征进行提取。通过实验研究了卷积核数目、卷积核大小、池化层、网络深度对模型诊断性能的影响。通过混淆矩阵、ROC曲线和PR曲线对比分析了卷积神经网络模型、支持向量机(Support Vector Machine, SVM)模型、BP神经网络(Back Propagation Neural Network, BPNN)模型。实验结果表明,卷积神经网络模型的诊断性能更为优秀。

**关键词:** 油中溶解气体; 变压器; 故障诊断; 卷积神经网络

**中图分类号:** TM41

**文献标志码:** A

doi:10.3969/j.issn.1672-6952.2020.05.014

## A Convolutional Neural Network Diagnosis Method for Dissolved Gas in Power Transformer Oil

Pei Xiaodeng, Luo Lin, Chen Shuai, Wang Qiao

(School of Information and Control Engineering, Liaoning Shihua University, Fushun Liaoning 113001, China)

**Abstract:** Dissolved Gas Analysis (DGA) is one of the important methods for determining transformer internal faults. Aiming at the shortcomings of the traditional shallow-based machine learning method in transformer fault diagnosis in feature extraction and generalization ability, a transformer fault diagnosis method based on Convolutional Neural Network (CNN) was proposed. The convolution layer in the network was used for feature conversion of dissolved gas in oil, and fault sensitive features were extracted by combining with the ability of pooling layer to strengthen important features. The effects of the number of convolution kernels, the size of convolution kernels, pooling layer and network depth on the diagnostic performance of the model were studied experimentally. The models of Convolutional Neural Network, Support Vector Machine (SVM) and Back Propagation Neural Network (BPNN) were compared and analyzed by confounding matrix, ROC curve and PR curve. The experimental results show that the Convolutional Neural Network model has better diagnostic performance.

**Keywords:** Dissolved gas in oil; Transformer; Fault diagnosis; Convolutional neural network

电力变压器是电力系统中关键电气设备之一,一旦变压器因故障而停止工作,将会对整个区域的生产生活造成不可估量的损失。因此,实时识别电力变压器的运行状态是非常必要的<sup>[1]</sup>。电力变压器在实际运行条件下受电气应力、机械应力和热压力等因素

的影响,其绝缘材料与变压器油会发生缓慢的化学变化,进而放出CO<sub>2</sub>、CO气体以及生成某些碳氢化合物(CH<sub>4</sub>、C<sub>2</sub>H<sub>6</sub>等)溶解在油中。当变压器发生故障时,将加速这些化学反应的进行,所生成的化合物种类及比例与放电、放热等故障有着密切的联系<sup>[2]</sup>。

收稿日期:2020-05-05 修回日期:2020-05-21

基金项目:国家自然科学基金项目(61703191); 辽宁省教育厅青年基金项目(L2017LQN028); 辽宁石油化工大学科研启动基金项目(2017XJJ-012)。

作者简介:裴小邓(1996-),男,硕士研究生,从事变压器智能故障诊断方向研究;E-mail:peixiaodeng@163.com。

通信联系人:罗林(1984-),男,博士,讲师,从事输变电装备在线监测与智能故障诊断的研究;E-mail:luolin@lnpu.edu.cn。

基于DGA的电力变压器故障诊断方法主要分为传统诊断方法和人工智能诊断方法。传统诊断方法主要有特征气体法和比值法(包括IEC三比值法、改进的IEC三比值法)。虽然这些方法便于工程应用,但是编码并不完备,对于某些故障很难找到与之对应的编码,因此诊断的准确率较低。人工智能诊断方法主要包括模糊理论<sup>[3]</sup>、专家系统<sup>[4]</sup>、神经网络<sup>[5]</sup>、支持向量机<sup>[6]</sup>等。虽然这些方法取得了较好的应用效果,提高了故障识别率,但同时存在着一定不足。模糊理论诊断法隶属函数难以确定,需要大量的实际故障数据;专家系统诊断法需要大量的专家经验,获取到的经验难以表达,推理能力较弱;神经网络诊断法存在易陷入局部最优、收敛速度慢、泛化能力差等缺点;支持向量机诊断法在本质上属于二分类问题,处理多分类问题时需要经历“一对一”或“一对多”等复杂的过程,核函数和正则化参数选择困难。

随着机器学习的发展,深度学习的发展推动了人工智能技术在变压器DGA故障诊断方面的诸多应用,如深度置信网络(Deep Belief Network, DBN)<sup>[7]</sup>、深度自编码网络(Deep Auto-Encoder Network, DAEN)<sup>[8]</sup>、卷积神经网络(Convolutional Neural Network, CNN)<sup>[9]</sup>。文献[10]提出了一种基于深度置信网络构建变压器故障诊断模型,利用快速歧化算法加快了网络收敛速度,在诊断的准确率上较BPNN和SVM有了提升。文献[11]提出了一种基于贝叶斯正则化深度信念网络的电力变压器故障诊断方法,克服了传统的DBN方法在大量无标签样本的情况下才能得到较高准确率的缺点。文献[12]提出了一种基于深度自编码网络的电力变压器故障诊断方法,解决了一些样本数据缺少标签的问题,但是该方法需要大量的预训练样本,在预训练样本数据较少时,诊断的效果并不理想。文献[13]提出了一种基于卷积神经网络的变压器故障诊断方法,虽然克服了一些浅层机器学习方法存在的问题,但是池化类型和网络深度的选择并未说明,卷积网络强大的特征提取能力并未得到展示。

本文提出了一种基于卷积神经网络的变压器故障诊断方法,利用卷积层提取特征和池化层强化重要特征的能力来解决特征提取和过拟合的问题,达到提高诊断准确率的目的。

## 1 卷积神经网络的基本原理

一个完整的卷积神经网络一般是由输入层、卷积层、池化层、展开层、全连接层、输出层组成,其中卷积层和池化层是整个网络的核心。

### 1.1 卷积层

卷积层是通过使用卷积核对上一层的部分区域进行卷积运算,提取出相应的特征信息。卷积层最重要的两个特点是稀疏连接和权值共享,这两点也是卷积神经网络和普通神经网络本质上的区别。稀疏连接和全连接的不同之处在于卷积层中的神经元只与上一层的部分神经元相连,权值共享的优点在于特征提取时可以不考虑局部特征的位置,同时可以有效地减少卷积层网络参数的数量,从而降低因参数过多而发生拟合的几率。全连接、稀疏连接、权值共享示意图如图1所示。

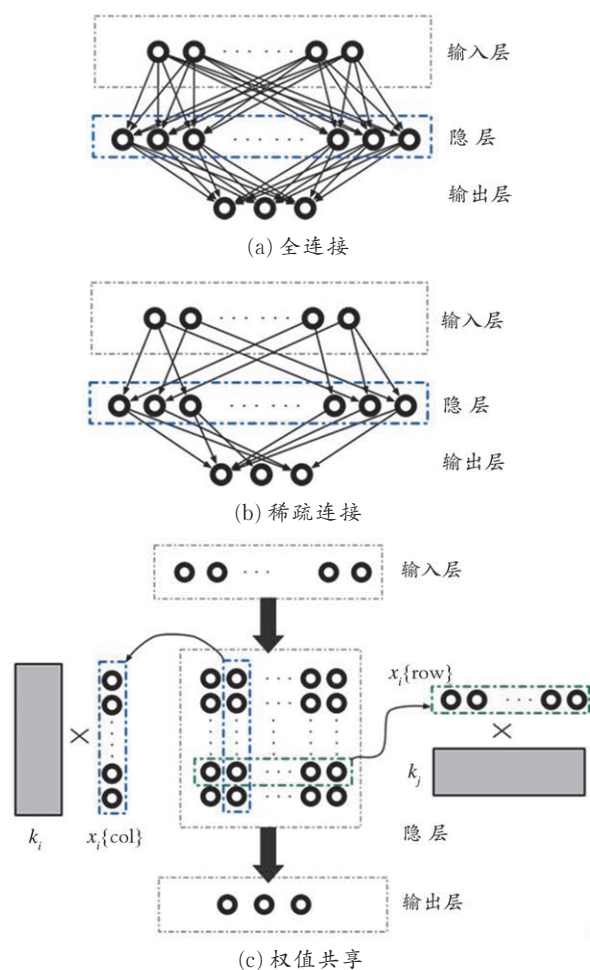


图1 全连接、稀疏连接、权值共享示意图

卷积层的输出结果为:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_i^l\right) \quad (1)$$

式中,  $x_j^l$  为第  $l$  层的  $j$  个特征图;  $f$  为激活函数;  $M_j$  为输入特征图的集合;  $*$  为卷积算;  $k_{ij}^l$  为卷积核;  $b_i^l$  为偏置项。

### 1.2 池化层

卷积层提取到的特征在传入下一层之前,可以对其进行处理,池化是最常用的处理方式。通过池化层的处理,卷积层提取到的特征维度得到有效的

压缩,可以减少网络中参数的数量,网络的计算量得到降低,同时也降低网络发生过拟合的几率。池化有最大池化和平均池化两种方式,最大池化是对某个区域的特征取最大化的操作,可以滤去一些不重要的特征信息。平均池化是对某个区域的特征进行平均化的操作,对一些不重要的特征信息并不像最大池化那样完全滤去,而是淡化,以此来强化那些重要的特征信息。由此可见,模型经过池化处理后,其抗干扰能力和稳定性都得到加强。

池化层的输出结果为:

$$x_j^l = f(\beta_j^l \text{down}(x_i^{l-1}) + b_j^l) \quad (2)$$

式中,down(•)为降采样函数; $x_i^{l-1}$ 为上一层的特征图(卷积层); $\beta_j^l$ 为特征图的乘性偏置; $b_j^l$ 为加性偏置。

### 1.3 Softmax 回归模型

本文采用 Softmax 回归模型。Softmax 回归模型是 Logistic 回归模型的推广,常用于求解多分类问题。假设训练集为 $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$ ,其中 $x^{(i)} \in \mathbb{R}^{n+1}$ , $y^{(i)} \in \{1, 2, 3, 4, \dots, k\}$ ,当输入样本为 $x$ 时,可由激励函数 $h_\theta(x)$ 求出样本属于任意类别时的概率 $p(y=j|x)$ 。假设函数将要输出一个 $k$ 维的向量来表示 $k$ 个估计的概率值,这些向量元素之和为1。激励函数可表示为:

$$h_\theta(x^{(i)}) = \begin{bmatrix} p(y^{(i)}=1|x^{(i)}; \theta) \\ p(y^{(i)}=2|x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)}=k|x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \quad (3)$$

式中, $\theta_1, \theta_2, \dots, \theta_k \in \mathbb{R}^{n+1}$ 为模型参数; $\frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}}$ 项是对概率分布进行归一化处理,使所有概率的和等于1。将 $\theta$ 用一个 $K \times (n+1)$ 的矩阵可表示为:

$$\theta = \begin{bmatrix} -\theta_1^T \\ -\theta_2^T \\ \vdots \\ -\theta_k^T \end{bmatrix} \quad (4)$$

接下来对 Softmax 回归代价函数进行分析。代价函数 $J(\theta)$ 见式(5)。在式(5)中, $1\{\cdot\}$ 是指示性函数,运算规则为: $1\{\text{表达式的值为真}\}=1$ , $1\{\text{表达式的值为假}\}=0$ 。

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)}=j\} \lg \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right] \quad (5)$$

将样本 $x$ 分类,分为第 $j$ 类的概率为:

$$p(y^{(i)}=j|x; \theta) = \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \quad (6)$$

在实现 softmax 回归算法时,通过在式中添加

一个权重衰减项 $\frac{\lambda}{2} \sum_{i=1}^m \sum_{j=1}^k \theta_{ij}^2$ ,对过大的参数值进行惩罚。回归代价函数公式将表示为:

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)}=j\} \lg \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{i=1}^m \sum_{j=1}^k \theta_{ij}^2 \quad (7)$$

增加权重衰减项后,代价函数成为一个凸函数,这样能防止优化过程中陷入局部收敛,得到最优解。为对其进行优化,需要计算 $J(\theta)$ 的导数,其梯度公式为:

$$\nabla_{\theta_l} J(\theta) = -\frac{1}{m} \sum_{i=1}^m [x^{(i)} (1\{y^{(i)}=j\} - p(y^{(i)}=j|x^{(i)}; \theta))] + \lambda \theta_j \quad (8)$$

式中, $\nabla_{\theta_l} J(\theta)$ 为向量,第 $l$ 个元素 $\frac{\partial J(\theta)}{\partial \theta_{jl}}$ 是 $J(\theta)$ 对 $\theta_j$ 第 $l$ 个分量的偏导数。利用梯度下降法对代价函数 $J(\theta)$ 进行最小化。每次迭代过程中都需要对参数进行更新。

$$\theta_j = \theta_j - \alpha \nabla_{\theta_l} J(\theta) \quad j=1, 2, \dots, k \quad (9)$$

式中, $\alpha$ 为学习率。

## 2 卷积网络诊断模型的搭建

### 2.1 输入和输出向量的选择

根据《变压器油中溶解气体分析和判断导则》可知,变压器发生故障状态时共有氢气( $H_2$ )、甲烷( $CH_4$ )、乙烷( $C_2H_6$ )、乙烯( $C_2H_4$ )、乙炔( $C_2H_2$ )、一氧化碳( $CO$ )、二氧化碳( $CO_2$ )七种气体存在,本文选取其中五种气体( $H_2$ 、 $CH_4$ 、 $C_2H_6$ 、 $C_2H_4$ 、 $C_2H_2$ )作为变压器故障诊断的输入特征向量。选取正常N、局部放电PD、低能放电D1、高能放电D2、中低温过热T1、高温过热T2作为卷积神经网络的输出向量。

### 2.2 数据预处理

本文将每种气体(共五种)体积与气体总体积之比作为卷积网络诊断模型的输入,按照式(10)做规范化处理。

$$x_{\text{new}}^{(i)} = \frac{x^{(i)}}{x^{(1)} + x^{(2)} + x^{(3)} + x^{(4)} + x^{(5)}} \quad (10)$$

式中, $i=\{1, 2, 3, 4, 5\}$ ;  $x_{\text{new}}^{(i)}$ 为规范化后的气体数据; $x^{(i)}$ 为规范前的气体数据; $x^{(1)}-x^{(5)}$ 分别为 $H_2$ 、 $CH_4$ 、 $C_2H_2$ 、 $C_2H_4$ 、 $C_2H_6$ 五种气体的体积分数, $\mu\text{L/L}$ 。

### 2.3 变压器故障编码

对六种变压器故障类型进行二进制编码,结果如表1所示。

### 2.4 CNN模型诊断结构图

卷积神经网络处理的数据一般是 $m \times n$ 的二维



格式,但是本文用于诊断变压器故障的数据是  $m \times 1$  的一维格式,所以采用一维卷积神经网络搭建诊断模型。由于数据维度小,只搭建了单层和双层两种诊断模型。

表 1 变压器故障类型编码

编号	故障类型	二进制编码
1	正常	100000
2	局部放电	010000
3	低能放电	001000
4	高能放电	000100
5	中低温过热	000010
6	高温过热	000001

CNN 模型诊断结构如图 2 所示。输入层的大小为  $5 \times 1$ ,因此卷积层  $C_1$  卷积核的大小有  $1 \times 1$ 、 $2 \times 1$ 、 $3 \times 1$ 、 $4 \times 1$  四种选择。为了体现卷积层稀疏连接的优势, $C_1$  层只选择卷积核大小为  $2 \times 1$ 、 $3 \times 1$ 、 $4 \times 1$  三种进行实验。

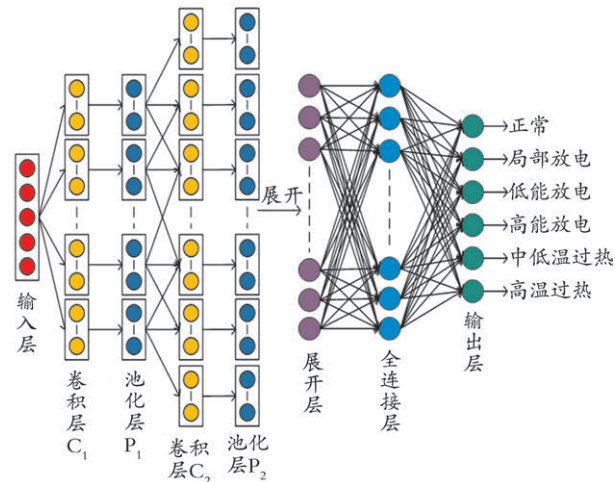


图 2 CNN 模型诊断结构

2.4.1 单层卷积诊断模型 当卷积层有  $K$  个大小为  $m \times 1$  ( $m = 2, 3, 4$ ) 的卷积核,可以提取到  $K$  个大小为  $(6 - m) \times 1$  的特征图。当池化层的大小为  $2 \times 1$ ,可以得到  $K$  个大小为  $((6 - m)/2) \times 1$  的特征图,经展开后共获得  $K(6 - m)/2$  个特征。

2.4.2 双层卷积诊断模型

(a)当卷积层  $C_1$  有  $K_1$  个大小为  $2 \times 1$  的卷积核,可以提取到  $K_1$  个大小为  $4 \times 1$  的特征图。当池化层  $P_1$  的大小为  $2 \times 1$ ,可以得到  $K_1$  个大小为  $2 \times 1$  的特征图。当卷积层  $C_2$  有  $K_2$  个大小为  $2 \times 1$  的卷积核,可以提取到  $K_2$  个大小为  $1 \times 1$  的特征图,当池化层  $P_2$  的大小为  $1 \times 1$ ,可以得到  $K_2$  个大小为  $1 \times 1$  的特征图,经展开后共获得  $K_2$  个特征。(b)当卷积层  $C_1$  有  $K_1$  个大小为  $3 \times 1$  的卷积核,可以提取到  $K_1$  个大小为  $3 \times 1$

的特征图。当池化层  $P_1$  的大小为  $2 \times 1$ ,可以得到  $K_1$  个大小为  $2 \times 1$  的特征图。当卷积层  $C_2$  有  $K_2$  个大小为  $2 \times 1$  的卷积核,可以提取到  $K_2$  个大小为  $1 \times 1$  的特征图。当池化层  $P_2$  的大小为  $1 \times 1$ ,可以得到  $K_2$  个大小为  $1 \times 1$  的特征图,经展开后共获得  $K_2$  个特征。

全连接层的神经元个数设为 32,卷积层和全连接层的激活函数为 Relu。为防止过拟合,在全连接层后面使用 Dropout 技术,比率设为 0.2,学习率设为 0.000 1,迭代次数设为 2 500。

3 实验分析

本文变压器故障实验基于 python3.7 语言环境,在操作系统为 Windows10 家庭中文版(64 位)、CPU 为 Intel(R) Core(TM) i5-5200U (2 201 MHz)、内存为 8 G 联想 80FA 电脑上完成。

3.1 样本数据的选取

本文从相关文献中一共获取了 393 条完整的变压器 DGA 数据进行实验。将故障样本按照 7:3 的比例划分为训练集和测试集。训练集和测试集的分布状况如表 2 所示。

表 2 训练集和测试集的分布情况

变压器状态	训练集样本数	测试集样本数
正常	18	8
局部放电	36	16
低能放电	20	8
高能放电	64	28
中低温过热	41	17
高温过热	96	41

3.2 卷积核数目对模型性能的影响

为了保证实验的准确性,将卷积网络其他参数的设置保持一致(包括卷积核大小、池化类型的选择、全连接层单元数目、学习率、Dropout 比率等参数)。卷积核数目对训练集准确率、测试集准确率以及训练损失的影响如图 3 所示。

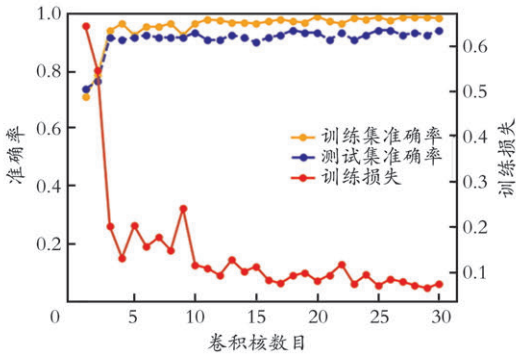


图 3 卷积核数目对训练集准确率、测试集准确率以及训练损失的影响

从图 3 可以看出,当卷积核数目从 1 增加到 3 时,训练集准确率和测试准确率不断提高,卷积核数目从 3 增加到 30 时训练集准确率和测试准确率趋于平稳。当卷积核数目从 1 增加到 4 时,训练损失大幅减少,卷积核数目从 4 增加到 17 时总体上为下降趋势,卷积核数目从 17 增加到 30 时有小幅度变化,总体是趋于平稳的。由此可以得出,随着卷积核数目的增加,训练集和测试集的准确率并不会一直增大,训练损失也不会一直减小。为了使训练集和测试集准确率高,同时使训练损失小,卷积核数目为 17 时最为合适。

3.3 卷积核大小对模型性能的影响

在卷积核数目为 17、其他参数保持相同的情况下,选择卷积核大小为  $2 \times 1$ 、 $3 \times 1$ 、 $4 \times 1$  来测试卷积核对模型性能的影响,从训练损失和训练准确率两方面来比较说明。不同卷积核的训练损失及训练集准确率迭代图如图 4 所示。

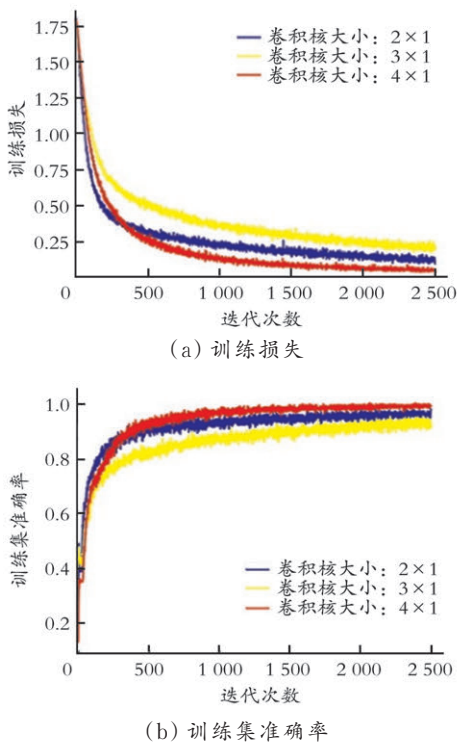


图 4 不同卷积核的训练损失及训练集准确率迭代图

从图 4 可以看出,当卷积核大小为  $3 \times 1$  时,训练损失最大,训练集准确率最低;当卷积核大小为  $4 \times 1$  时,训练损失最小,训练准确率最高。该实验证明,并不是卷积核越小训练集准确率越高,训练损失越小。为了使训练集准确率高同时训练损失小,卷积核大小为  $4 \times 1$  最为合适。

3.4 池化层对模型性能的影响

分别测试最大池化和平均池化对模型性能的

影响,选择卷积核数目为 17,卷积核大小为  $4 \times 1$ ,池化层大小为  $2 \times 1$ ,其他的参数保持相同。不同池化层下的训练损失、训练集准确率、测试集准确率如表 3 所示。从表 3 可以看出,最大池化的训练集和测试集准确率更高,训练损失更小,证明最大池化使模型的诊断性能更为优秀。

表 3 不同池化层的训练损失、训练集准确率、测试集准确率

池化类型	训练损失	训练集准确率/%	测试集准确率/%
最大池化	0.086	97.818	91.525
平均池化	0.162	94.545	87.288

3.5 学习率和 Dropout 比率对诊断结果的影响

为了优化模型性能,测试学习率和 Dropout 比率对变压器诊断结果的影响,结果如图 5 所示。

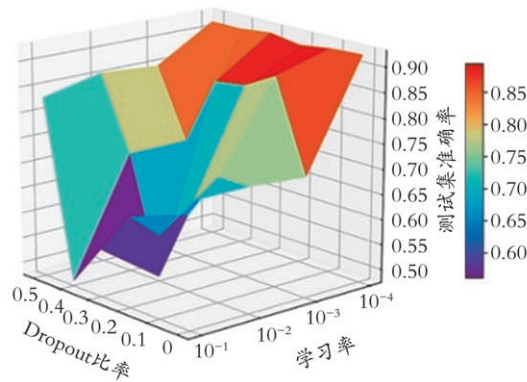


图 5 学习率和 Dropout 比率对变压器诊断结果的影响

从图 5 可以看出,当 Dropout 比率一定时,诊断的准确率会随着学习率的减小而增加。当学习率为 0.100 0、0.010 0、0.001 0 时,Dropout 比率对诊断结果影响较大,当学习率为 0.000 1 时,Dropout 比率对于诊断结果没有影响。实验结果表明,当学习率选择合适时,模型并不会发生过拟合现象,证明了卷积层的稀疏连接和权值共享能够有效地防止过拟合,同时也证明了池化层增强了模型的稳定性。

3.6 网络深度对模型性能的影响

为了测试网络深度对模型诊断性能的影响,本实验将对单层和双层(单个卷积层加上单个池化层为一层网络)卷积诊断模型的性能进行对比。单层模型参数选择实验最佳参数,即卷积核数目为 17,卷积核大小为  $4 \times 1$ ,池化层选择最大池化。双层模型有 2 种方式,单层卷积和双层卷积诊断模型性能如表 4 所示。卷积层中 17- $4 \times 1$  表示卷积核数目为 17,卷积核大小为  $4 \times 1$ ,池化层中  $2 \times 1$  表示池化层大小为  $2 \times 1$ ,池化类型同样选择最大池化。其余的参数单层和双层保持一致(从结构上讲,双层是单层网络的复制,仅设定

单层网络初始参数)。从表 4 可知,单层和双层的训练损失、训练集准确率、测试集准确率几乎相同,证明

了单层和双层诊断模型的性能相同,这是由于 DGA 数据结构简单,受网络深度影响较小。

表 4 单层卷积和双层卷积诊断模型性能

卷积层 C <sub>1</sub>	池化层 P <sub>1</sub>	卷积层 C <sub>2</sub>	池化层 P <sub>2</sub>	训练损失	训练集准确率/%	测试集准确率/%
17-4×1	2×1	无	无	0.086	97.818	91.525
17-2×1	2×1	17-2×1	1×1	0.088	98.182	92.373
17-3×1	2×1	17-2×1	1×1	0.101	97.454	91.525

3.7 实验对比

从混淆矩阵、ROC 曲线、PR 曲线等多个方面来比较 CNN、SVM、BPNN 模型诊断性能。CNN、BPNN、SVM 的混淆矩阵如图 6 所示。从图 6 可以看出,CNN 模型对于正常、低能放电、高能放电、中低温过热、高温过热状态预测上效果较好,局部放电状态的预测效果一般;BPNN 模型在正常、中低温过热、高温过热状态的预测效果较好,在正

常、局部放电、低能放电状态的预测效果较差;SVM 模型在正常、中低温过热、高温过热状态的预测效果较好,在局部放电状态的预测效果一般,在正常、低能放电状态预测效果非常差。尤其是低能放电状态,SVM 模型预测准确率为 0,BPNN 模型预测准确率为 50%,而 CNN 模型预测准确率为 100%。实验证明了 CNN 模型的分类性能要明显优于 BPNN 模型和 SVM 模型。

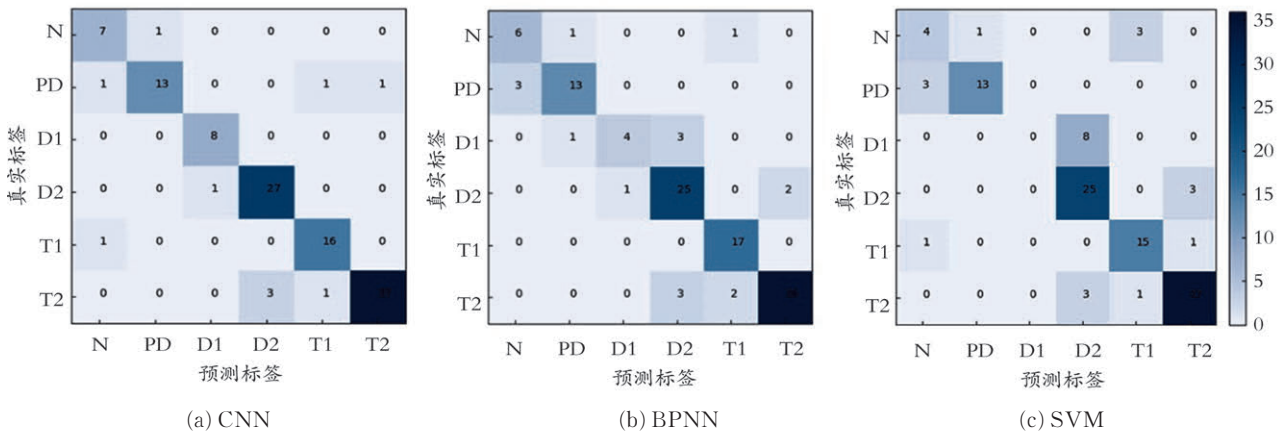


图 6 CNN、BPNN、SVM 的混淆矩阵

不同模型的训练集和测试集准确率如表 5 所示。

表 5 不同模型的训练集和测试集准确率 %		
模型	训练集准确率	测试集准确率
CNN	97.818	91.525
SVM	85.455	79.661
BPNN	95.636	85.593

从表 5 可以看出,CNN 模型的训练集准确率和测试集准确率都超过了 90%,明显高于 SVM 模型和 BPNN 模型,这是由于 CNN 模型特征提取能力要比 SVM 和 BPNN 强。SVM 模型训练集和测试集的准确率都不高,说明 SVM 模型在多分类问题上分类能力一般。BPNN 模型虽然训练集准确率高达 95.636%,但是训练集的准确率只有 85.593%,说明 BPNN 模型的泛化能力较差,容易发生过拟合。

不同模型的 ROC 曲线如图 7 所示。从图 7 可以

看出,CNN 模型 ROC 曲线 AUC 值高达 0.99,比 SVM 模型 AUC 值高 0.02,比 BPNN 模型 AUC 值高 0.01,AUC 值接近的原因是由于样本不平衡造成的,AUC 值受样本数量影响较大,从混淆矩阵中可以看出 BPNN、SVM 模型在正常、高温过热状态预测结果非常出色,最终导致 AUC 值接近。但是,CNN 模型的分类和泛化能力比 SVM 和 BPNN 模型要优秀。

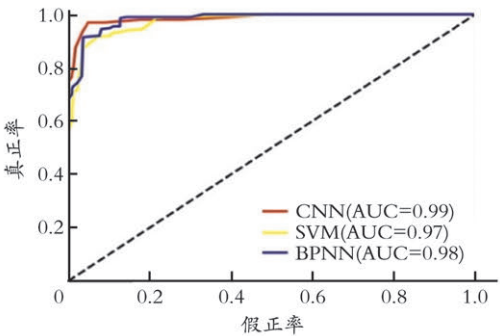


图 7 不同模型的 ROC 曲线



PR曲线跟ROC曲线相比,AUC的值受样本不平衡影响较小。不同模型的PR曲线如图8所示。

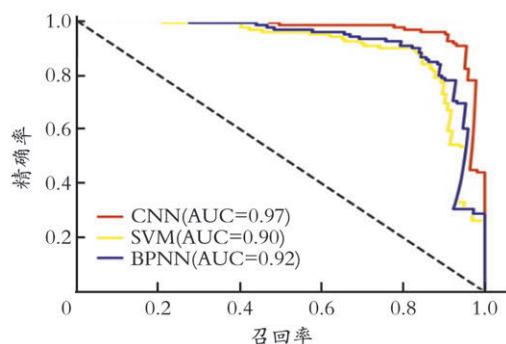


图8 不同模型的PR曲线

从图8可以看出,CNN模型PR曲线AUC值高达0.97,比SVM模型AUC值高出0.07,比BPNN模型AUC值高出0.05,证明CNN模型的泛化和分类能力更为优秀。

### 3.8 特征可视化

CNN模型诊断准确率高的原因之一在于它强大的特征提取能力。为了验证CNN模型强大的特征提取能力,采用 $t$ -SNE( $t$ -Distributed Stochastic Neighbor Embedding)技术,把CNN模型提取到的高维特征映射成二维特征,并以散点图的形式可视化出来。首先把CNN模型的全连接层提取到的393个样本的三十二维特征提取出来,然后利用 $t$ -SNE技术将393个样本的三十二维特征降为二维并展示出来。CNN模型特征可视化如图9所示。从图9可以看出,相同的类聚集在一起,不同的类有一定的距离,聚类的准确率在98%以上,说明CNN模型提

取到的特征可分性好,证明CNN模型有着优秀的特征提取能力。

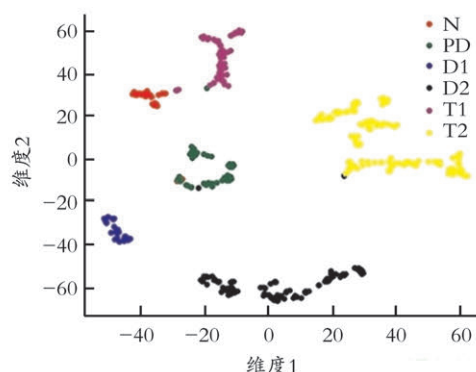


图9 CNN模型特征可视化

## 4 结 论

在变压器诊断领域,浅层机器学习诊断方法虽然在诊断的精度上较传统的方法有所提升,但是还存在着一些缺点,为克服这些缺点,构建基于卷积神经网络的变压器故障诊断模型。通过混淆矩阵、ROC曲线、PR曲线多方面实验验证,基于卷积神经网络的变压器故障诊断方法跟支持向量机、BP神经网络相比,特征提取和泛化能力更强,诊断的准确率更高。实验结果表明,双层卷积网络模型诊断准确率不一定比单层的高,卷积核的数量并不是越多越好,卷积核的尺寸并不是越小越好,因此卷积网络的深度、核数目、核大小要根据数据的实际情况来选择。对于池化层的大小和全连接层神经元数目的选择并未讨论,这也是下一步研究的重点。

## 参 考 文 献

- [1] 代杰杰. 基于深度学习的变压器状态评估技术研究[D]. 上海:上海交通大学,2018.
- [2] 贾亦敏. 变压器在线监测与智能故障诊断系统研究[D]. 徐州:中国矿业大学,2009.
- [3] Huang Y C, Sun H C. Dissolved gas analysis of mineral oil for power transformer fault diagnosis using fuzzy logic[J]. IEEE Transactions on Dielectrics & Electrical Insulation, 2013, 20(3):974-981.
- [4] 刘灿萍. 基于DGA的变压器故障诊断专家系统的研究[D]. 南宁:广西大学,2007.
- [5] Zhang Y, Chen E, Guo P, et al. Application of improved particle swarm optimization BP neural network in transformer fault diagnosis[C] // 2017 Chinese Automation Congress (CAC), Jinan: IEEE, 2017: 6971-6975.
- [6] Li J, Zhang Q, Wang K, et al. Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2016, 23(2): 1198-1206.
- [7] Dai J, Song H, Sheng G, et al. Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2017, 24(5): 2828-2835.
- [8] 许倩文, 吉兴全, 张玉振, 等. 栈式降噪自编码网络在变压器故障诊断中的应用[J]. 电测与仪表, 2018, 55(17): 62-67.
- [9] 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.
- [10] 姜有泉, 黄良, 王波, 等. 基于DGA和深度置信网络的变压器内部故障诊断[J]. 武汉大学学报(工学版), 2017, 50(5): 749-753.
- [11] 王德文, 雷倩. 基于贝叶斯正则化深度信念网络的电力变压器故障诊断方法[J]. 电力自动化设备, 2018, 38(5): 129-135.
- [12] 石鑫, 朱永利, 宁晓光, 等. 基于深度自编码网络的电力变压器故障诊断[J]. 电力自动化设备, 2016, 36(5): 122-126.
- [13] 贾京龙, 余涛, 吴子杰, 等. 基于卷积神经网络的变压器故障诊断方法[J]. 电测与仪表, 2017, 54(13): 62-67.

(编辑 陈 雷)