**基于变压器振动信号故障分析方案**

摘要：针对目前变压器振动信号故障诊断领域多是对基于振动信号的特征提取方法进行研究，而缺乏对于故障诊断模型训练方法的研究以及对于模型的有效评价的问题，提出了基于主成分分析方法与K折交叉验证方法优化的BP神经网络故障诊断模型训练方法。对采用经验小波+希尔伯特变换提取的时频特征进行主成分降维，通过K折交叉验证方法对样本数据进行划分，然后采用BP神经网络进行模型训练，并对训练结果采用监督学习评价指标——ROC曲线与PRC曲线——进行评价。使用实测数据进行实验表明，采用该方法能够有效降低模型训练的运算复杂度并提高故障诊断的准确性。与常用机器学习方法相比，BP神经网络故障诊断模型能够更为准确的对变压器故障进行诊断，是变压器故障诊断的有效工具。

## 0引言

电力变压器在电网中实现电压转换与电力分配，电力变压器的正常运行对整个电网的安全稳定运行起着至关重要的作用。据国内相关部门统计，在2001年度，500KV变压器的故障率为5.1%，220KV变压器的故障率为0.4%，110KV变压器的故障率为0.49%[1] [2]。及时发现变压器存在的潜在故障，准确识别故障类型，及时进行维护，能够有效避免变压器进一步恶化为严重故障导致生命财产的重大损失。因此，对变压器潜在故障的预警及诊断，降低变压器故障率，提高安全可靠型具有重要研究意义。

从国内外关于电力变压器故障的统计分析得出：绕组所引起的机械故障是变压器故障的主要原因,绕组故障占比总故障的60%以上，其次是铁心故障[3] [4]。研究发现，变压器内部绕组与铁芯的位移和结构等机械状态发生变化时可通过箱体表面的振动信息反馈出来[5]，因此通过监测变压器箱体表面振动信号能够有效监测变压器绕组、铁芯等主要部件的机械结构变化。振动分析法相较于常用变压器故障诊断方法——油色谱分析法[6]和三比值法[7]——优点在于无需把采集振动信号的加速度传感器放置进变压器内部，直接将传感器贴在变压器箱体表面即可对振动信号进行采集[8]。传感器可安装在油箱前侧，此处具有能量传播路短，距离绕组较近幅值明显的特点，可以更容易采集到箱体振动信号[9]。信号采集测试发现变压器绕组与铁心产生的振动信号的特征频率主要位于100Hz倍频处，而变压器故障状态下的振动特征主要包括在100Hz倍频处能量会出现变化[10]。为进一步提取变压器振动信号特征，文献[11]分析了傅里叶变换、小波变换、希尔伯特黄变换等常用变压器振动信号特征提取方法的局限性，提出了基于经验小波的变压器振动信号特征提取方法。经验小波变换通过对信号频谱进行自适应划分，将单一信号分解为多个含有不同频率特征信息的模态分量，实现对信号的特征提取。实验表明该方法对变压器振动信号特征能够有效提取。

目前机器学习在故障诊断领域中使用已经十分广泛，为故障诊断提供了有力的支持。机器学习算法分为无监督学习和有监督学习，在其他领域中针对无监督学习算法在故障诊断中的应用进行了研究，发现无监督特征提取往往会出现大量无关特征和冗余特征，从而影响故障诊断的准确率[12]。因此针对有监督学习法特征提取以及模型训练在变压器振动信号故障诊断中的研究变得十分重要。常用机器学习算法如BP神经网络[13]、RBF神经网络[14]、极限学习机[15]、广义回归神经网络[16]、概率神经网络[17]等相继被应用于故障诊断领域，取得了较好的效果。如文献[18]提出了一种基于广义回归神经网络的变压器振动基频幅值计算方法，为变压器振动在线监测提供了重要参考。文献[19]通过对变压器振动信号小波变换后的能谱熵作为特征输入向量，提出了基于多分类支持向量机的故障诊断方法，提高了绕组故障诊断的有效性和准确性。这些方法均取得了一定的效果，但这些文献中往往只是针对某一种机器学习方法在变压器振动信号故障模型训练方面进行了讨论，缺乏系统的评价与对比。

本文在文献[11]采用经验小波对变压器振动信号进行分析的基础上，提出了一种基于变压器振动信号时频能量分布为特征，以BP神经网络为模型进行故障诊断训练的变压器振动信号故障诊断方法。并对故障诊断模型训练中的样本特征矩阵采用PCA降维减小特征向量的维数，压缩数据减小运算量并提升了模型训练效率。采用K折交叉验证法对BP神经网络进行诊断模型训练，提高了模型的准确率与稳定性。同时也对常用机器学习算法进行了对比训练，采用监督学习算法评价指标——ROC曲线与PRC曲线——对BP神经网络训练模型以及常用机器学习算法训练模型进行了综合评价。

## 1算法理论

### 1.1主成分分析法降维

在信号特征提取的过程中，原变量矩阵中往往包含冗余信息，增加了计算的复杂度。主成分分析算法[20]（Principal Component Analysis，简称PCA）是一种常见的降维算法，其基本思想是通过消除原变量矩阵中冗余的信息从而减少变量矩阵的维度，使得数据变得更加简单高效。该算法首先通过构造一个降纬矩阵，对原始的变量矩阵进行转换，使得转换后的新变量矩阵的维度更低，同时要求原始变量矩阵所携带的信息可以在新变量矩阵中最大限度的体现。此外为了消除原变量矩阵中冗余的信息，要求新变量矩阵的不同维度之间相互正交。

设有n个样本的特征向量分别为，每个样本含有k个特征维度，则其构成的n\*k阶数据矩阵可表示为：

 （1）

将样本矩阵通过主成分分析方法降维步骤如下：

1. 对每个特征求平均值

 （j=1,2,…,k） （2）

1. 中心化数据：将原数据矩阵中的特征数据减去平均值得到新的中心化后的数据；

 （3）

1. 计算协方差矩阵；

 （4）

1. 采用特征值分解方法求得协方差矩阵的特征值λi(i=1,2,…,k)与对应特征向量ai (i=1,2,…,k)，并要求 ;
2. 对特征值进行降序排序，通过式（5）计算贡献率和式（6）计算累计贡献率，一般选择累计贡献率>85%的前m个主成分。这m个特征值所对应的特征向量分别作为行向量组成特征向量P；

 （5）

 （6）

1. 将数据映射到m个特征向量构建的新空间，即Y=PX，其中Y为降维到m维后的数据。

### 1.2 K折交叉验证法参数寻优

在机器学习中，如果对同一数据集既进行训练又进行模型误差估计，会导致获得的模型对误差估计不准确。交叉验证的基本思想是将数据分为两部分，其中一部分做训练集，进行模型训练。另一部分数据即验证集用来测试模型的误差。因为两部分数据不重合，因此采用该方式所得出的泛化误差更接近该模型对真实数据评估的效果。常用的交叉验证方法有留一交叉验证（leave-one-out）和留p交叉验证（leave-P-out），其基本思想是将样本数据集划分为N个样本，留取一个或者P个样本作为验证集，剩下的N-1个或者N-P个作为训练集，重复进行N次，最后平均这N次的结果作为泛化误差估计。这两种方法都需要多次对数据进行切分，重复次数多，也会导致计算量过大。

K折交叉验证法是将数据的随机等分成K份，将每个子数据集都分别做一次验证集，其余的K-1个子数据集作为训练集对模型进行训练，重复K次。取这K个模型最终在验证集的分类准确率的平均结果作为该模型的性能指标。与留一交叉验证法和留P交叉验证法相比，该方法只需要计算K次，大大减小了算法复杂度。同时，该方法也能有效降低模型的过拟合，从有限的数据中过得更多信息。

实验表明，K取10能取得较为理想的误差估计结果，以10折交叉为例，模型结构如图所示。



图 1 10折交叉验证法模型

Fig. 1 The model of 10 fold cross validation

### 1.3变压器振动信号故障诊断模型训练

BP神经网络（back propagation network）是一种以误差为依据进行反向传播算法训练的多层前馈网络，于1986年由以Rumelhart和McClelland为首的科学家提出，其后逐渐发展为应用最为广泛的神经网络。包括输入层、隐含层和输出层，其中隐含层又叫中间层，其往往包含多层神经元，拓扑结构如图 2所示。其通过输入样本正向传播进行计算输出与输出误差反向传播进行隐含层权值调整两个过程完成模型训练。在网络理论以及性能方面已经比较成熟，具有非线性映射能力强、具有柔性的网络结构，自适应和自学习能力强等优点，且网络的中间层数以及各层神经元的个数可根据情况任意设定。但BP神经网络同时也具有学习速度慢，容易陷入局部极小化等缺点。



图 2 BP网络拓扑结构示意图

Fig. 2 Diagram of BP network topology

优化后的BP神经网络模型训练步骤如图所示，通过PCA降维，降低训练集数据冗余度，降低BP神经网络的训练复杂度。同时通过10折交叉验证法对样本数据分别进行训练模型，然后使用各模型分别进行诊断，取诊断平均结果作为最终故障诊断结构，该方法能够有效提高模型的诊断准确率，提高模型故障诊断的泛化性。其训练模型结构框图如图 3。



图 3改进BP神经网络训练模型

Fig. 3 Improved BP neural network training model

### 1.4 监督学习算法性能评价指标

对于训练好的模型，传统的模型评价方法往往是通过在验证集或者测试集上的准确率来评价模型的好坏，这种方法虽然简单直接，但不能综合的评价一个模型的性能。模型查全能力与查准能力可以用受试者工作特征曲线（Receiver Operating Characteristic Curve，简称ROC）与准确召回率曲线（Precision Recall Curve，简称PRC）来进行衡量。在分类任务中会出现4种情况，如表 1所示，分别是：

真阳性（True Positive，简称TP）：实际为正样本且模型的预测也为正样本的数目。也就是模型正确分类的正样本数。

假阳性（False Positive，简称FP）：实际为负样本且模型的预测为正样本的数目。也就是被模型错误的标记为正样本的负样本数。

真阴性（True Negative，简称TN）：实际为负样本且模型的预测也为负样本的数目。也就是模型正确分类的负样本数。

假阴性（False Negative，简称FN）：实际为正样本且模型的预测为负样本的数目。也就是被模型错误的标记为负样本的正样本数。

表 1预测的四种可能情况

Table 1 The result of prediction

|  |  |  |
| --- | --- | --- |
|  | 真实正例 | 真实反例 |
| 预测正例 | TP（True Positive） | FP（False Positive） |
| 预测反例 | FN（False Negative） | TN（True Negative） |

根据表 1可以得到如下几个公式：

查准率：

 (7)

查全率：

 (8)

真阳性率：

 （9)

伪阳性率：

 (10)

把分类阈值在0到1之间进行调整时可以得到不同的分类结果，把所有阈值情况下的精准率、查全率、真阳性率、伪阳性率记录下来，就可以以查全率为横轴，精准率为纵轴绘制ROC曲线，以伪阳性率为横轴，真阳性率为纵轴绘制PRC曲线。其中ROC描绘了模型查全的能力，PRC描绘了模型查准的能力。如果一个模型的ROC曲线完全覆盖另外一个模型的ROC曲线，则可以断言这个模型的查全能力更加优秀，如果2个模型的ROC曲线存在交叉情况，则往往使用ROC曲线与查全率轴所围成的面积AUC（Area Under Curve）来衡量，AUC越大，则模型越优。同理，如果一个模型的PRC曲线完全覆盖另外一个模型的PRC曲线，那么可以断言这个模型的查准能力更加优秀，如果2个模型的PRC曲线存在交叉情况，则往往使用PRC曲线与伪阳性率轴所围成的面积AP(Average precision)来衡量，AP越大，则模型越优。如图 4所示，模型B的AP值与AUC值都高于模型A，所以可以认为模型B比模型A拥有更好的查全能力与查准能力。

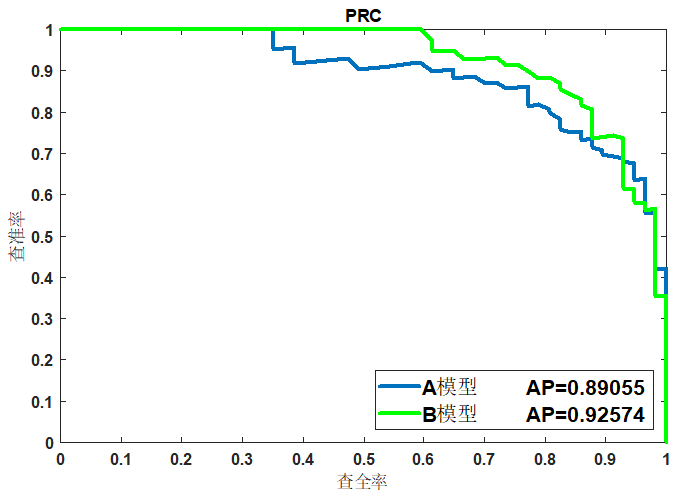
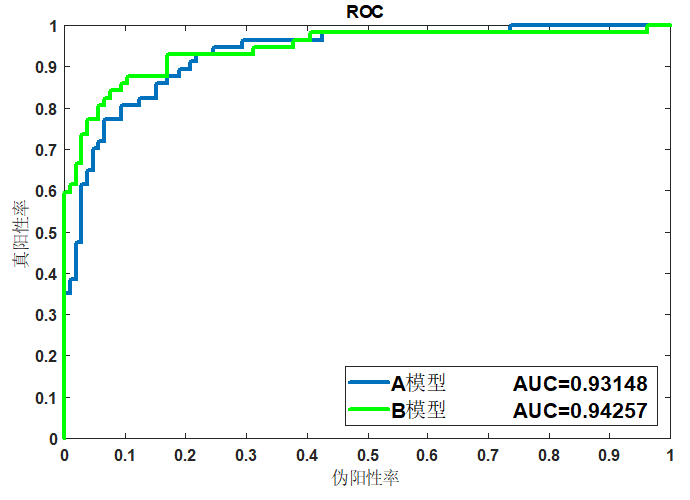


图 4 ROC与PRC

Fig. 4 ROC and PRC

## 3 实验与结果分析

### 3.1数据采集与特征提取

首先需要对变压器器身振动信号进行采集，本文采用美国PCB PIEZOTRONICS公司生产的压电式传感器分别固定在正常工况、绕组变形、铁芯松动三种不同工况下的额定容量为500kV·A,绕组额定电压为10.5/0.4kV，接线组别为Dyn11的变压器油箱两侧。传感器输出信号经信号解调仪做信号调理后直接接在数据采集仪上对振动信号进行采集。因为变压器箱体振动信号特征频率小于1000Hz，因此在实验室中，分别对三种不同工作状态下的变压器振动状态进行了采集，采样频率为2048Hz，采样时长1s，采样间隔1s。

对采集到的箱体振动信号数据采用经验小波分析方法对三种工况下的振动信号进行分析，步骤如下：

1. 对振动信号进行FFT变换；
2. 对FFT频谱进行EWT分解，获得EWF分量；
3. 对EWF特征分量求希尔伯特变换，并绘制希尔伯特谱；
4. 以50Hz频率段划分希尔伯特谱，并求划分后的20个频段内信号能量分布大小；
5. 因为在0到50Hz范围内的振动信号均为冷却系统等的噪声信号，因此在这里对0到50Hz范围内的能量清零；
6. 将每个频段能量大小作为信号特征矩阵。

对三种工况下的信号通过经验小波分析后进行希尔伯特变换绘制的希尔伯特谱如图 5（a）图 6（a）图 7（a）所示，对希尔伯特谱进行分段能量统计柱状图如图 5（b）图 6（b）图 7（b）所示。

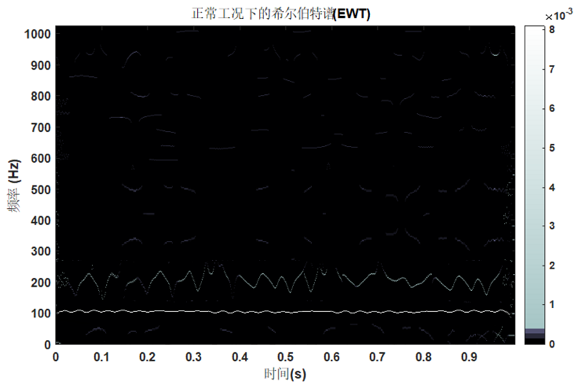
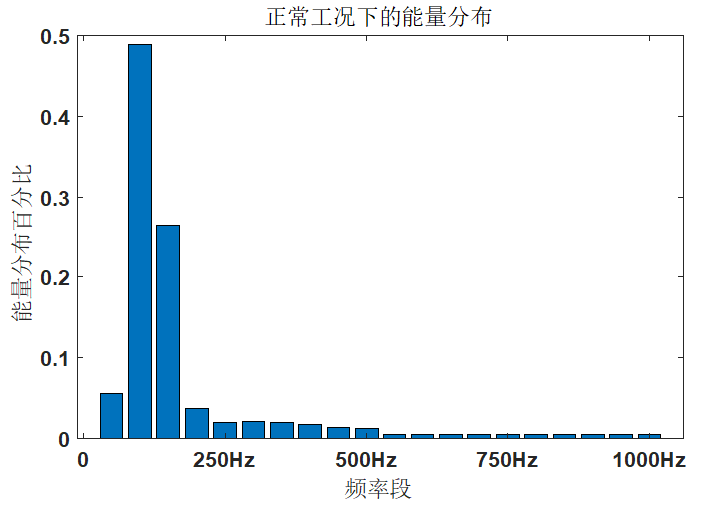
 

图 5正常工况下的希尔伯特谱与能量分布

Fig. 5 The Hilbert spectrum and energy distribution under normal conditions

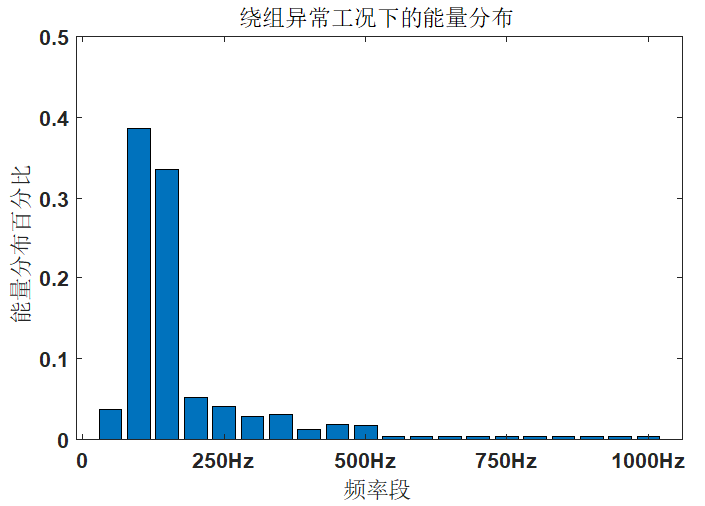
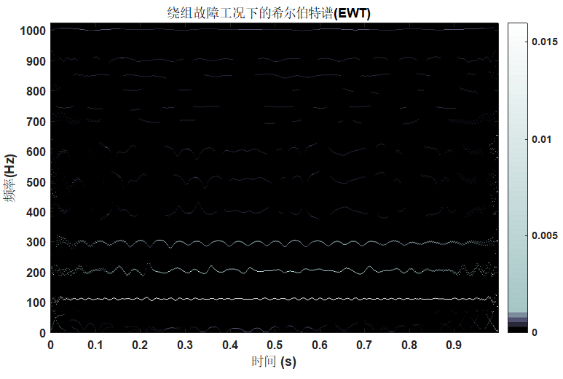


图 6绕组故障工况下的希尔伯特谱与能量分布

Fig. 6 The Hilbert spectrum and energy distribution under winding fault conditions

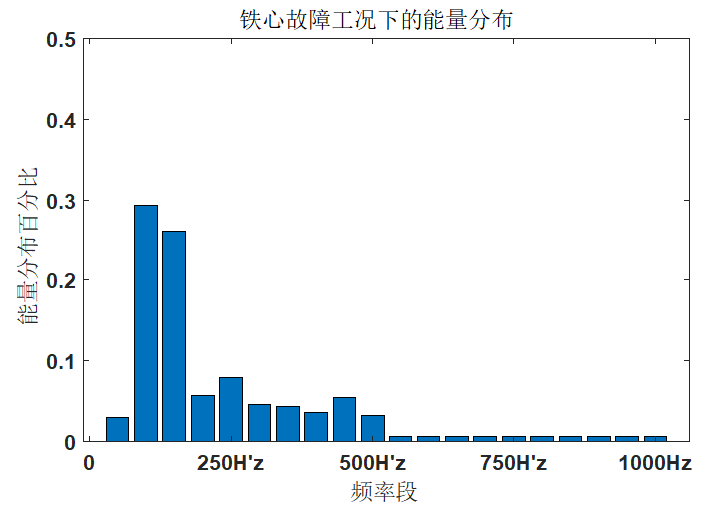
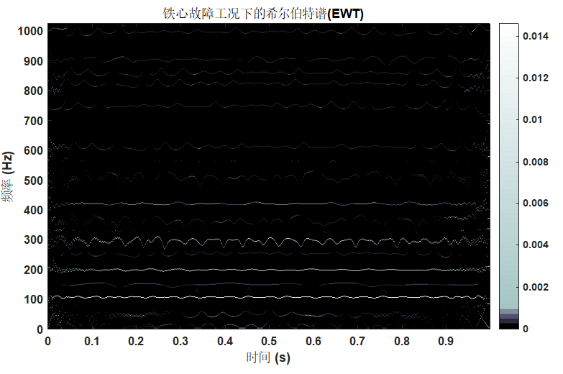


图 7铁心故障工况下的希尔伯特谱与能量分布

Fig. 7 The Hilbert spectrum and energy distribution under core fault conditions

通过希尔伯特谱与能量分布可以看出在正常工况下，振动信号的频率特征主要集中在100Hz，在200Hz处有微弱的频率特征，能量分布主要集中在50Hz到150Hz这个段。绕组故障工况下，振动信号的频率特征仍主要集中在100Hz，在200Hz到500Hz的100Hz倍频处有微弱的频率特征，能量分布仍然集中在50Hz到150Hz这个段，但是与正常工况相比，50Hz到100Hz这个段的能量分布有所下降，而100Hz到150Hz这个段的能量分布有所上升。铁心故障工况下，振动信号的频率特征主要集中在100Hz与200Hz，在高频处存在微弱特征频率，能量分布主要集中在50Hz到150Hz这个段，同时在150Hz到500Hz频段的能量有所增加。

### 3.2 PCA降维有效性验证

对比3种不同工况下的能量分布，发现不论在何种工况下，在500Hz以上的能量分布均无明显差异，这部分的能量对变压器的状态判断影响不大。因此通过降维算法对特征进行降维处理。选择810条数据进行PCA分析，其中变压器正常工况、绕组异常工况、铁心异常工况的数据比例为1:1:1，把这810条数据按照8:2随机划分为训练集与测试集。对训练集采用PCA主成分进行分析得到的主成分贡献率如图 8所示。从图中可以发现前4个最大特征值所对应的特征向量组成的降维矩阵对训练集进行降维后，把训练集从以前的20维降低到了4维，新的训练集携带了原始训练集95%以上的信息。达到了降维的目的。

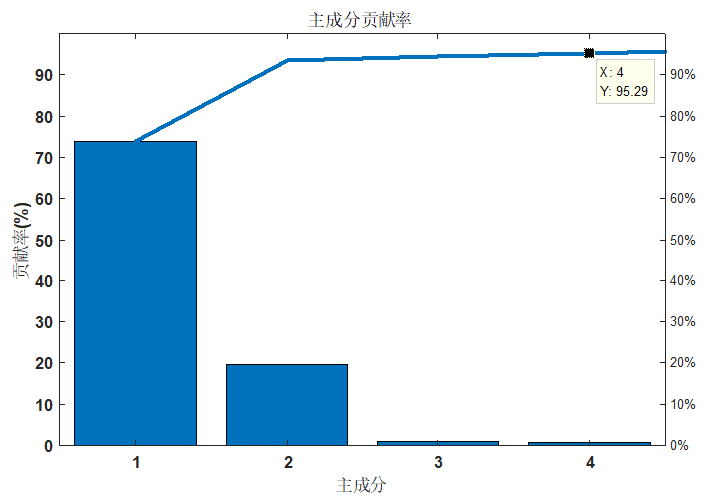


图 8主成分贡献率

Fig. 8 Principal component contribution rate

对训练集降维后，使用该降维后的训练集来训练BP神经网络，使用从训练集上获得的降维矩阵把测试集也降低到4维，然后输入到之前训练好的BP神经网络中进行测试。记录降维后的测试集在BP神经网络上的ROC与PRC。考虑到变压器的安全运行关系到人民的正常生活与生命财产安全，所有希望模型能尽可能的把存在安全隐患的变压器识别出来，所以将分类阈值设置为0.4。记录分类阈值为0.4时的查全率、查准率、综合准确率。其中综合准确率是在测试集上被正确分类的样本所占的比例。同时为了验证PCA降维的有效性，这里将未降维处理的训练集与测试集也用来训练BP神经网络，将其作为实验的对照组，并且测试没有降维所训练出来的模型在测试集上的ROC与PRC，阈值设置为0.4时的查全率、查准率、综合准确率。实验结果如图 9、图 10与表 2所示。

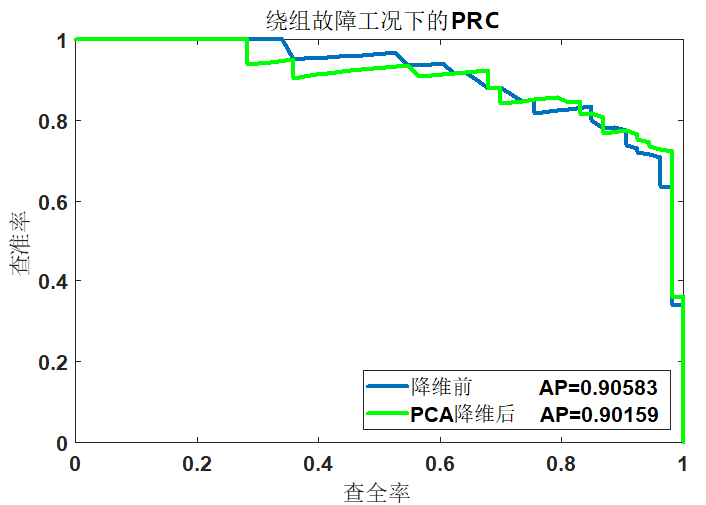
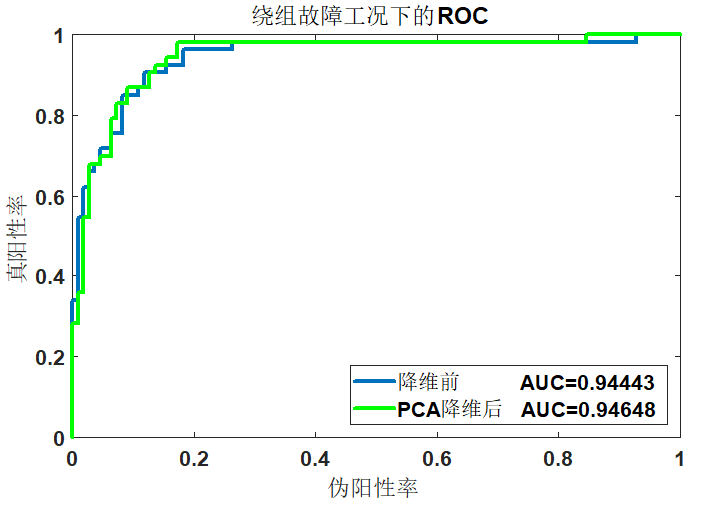


图 9 PCA前后训练出来的模型对于绕组故障的ROC与PRC

Fig. 9 ROC and PRC of the model trained before and after PCA for winding failure

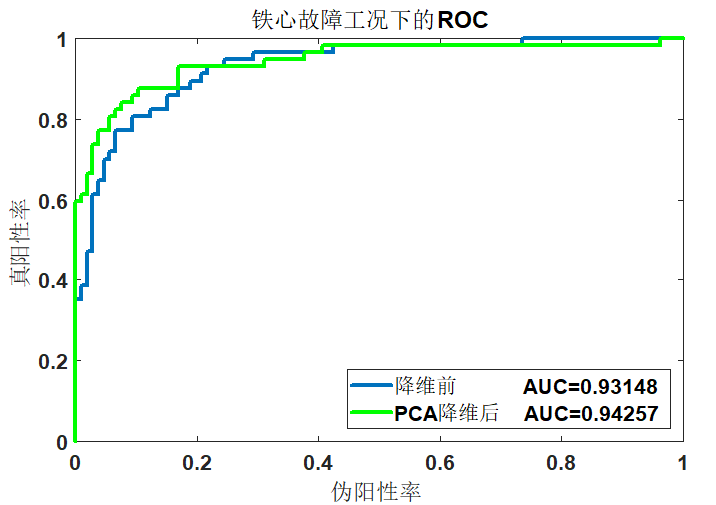
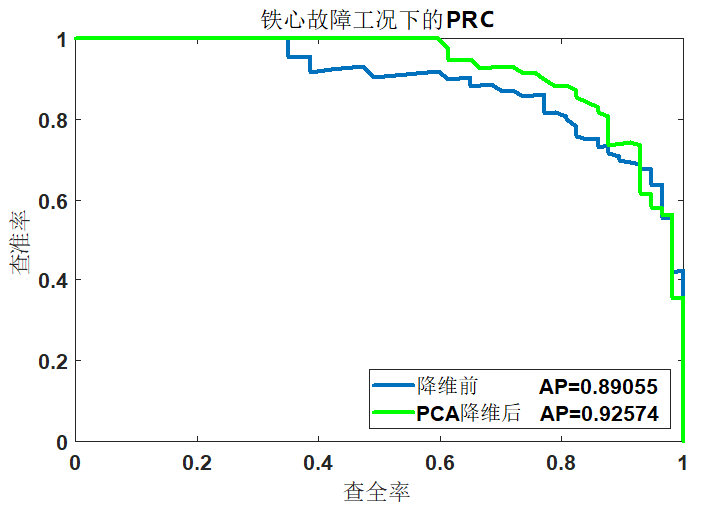


图 10 PCA前后训练出来的模型对于铁心故障的ROC与PRC

Fig. 10 ROC and PRC of the model trained before and after PCA for core failure

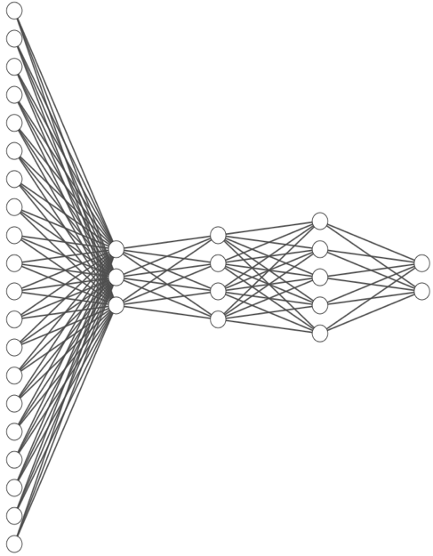
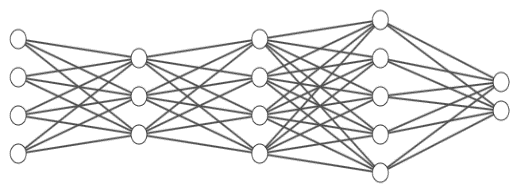
表 2 PCA前后训练出来的模型查全率与查准率

Table 2 The Recall and Precision of the model trained before and after PCA

|  |  |  |
| --- | --- | --- |
| 阈值：0.4 | 降维前 | PCA降维后 |
| 测试集上综合准确率 | 0.8160 | 0.8589 |
| 测试集上绕组故障工况查全率 | 0.8868 | 0.8679 |
| 测试集上铁心故障工况查全率 | 0.8426 | 0.8596 |
| 测试集上绕组故障工况查准率 | 0.7833 | 0.8070 |
| 测试集上铁心故障工况查准率 | 0.7460 | 0.8305 |
| 测试集上绕组故障工况AUC | 0.9444 | 0.9564 |
| 测试集上铁心故障工况AUC | 0.9314 | 0.9425 |
| 测试集上绕组故障工况AP | 0.9058 | 0.9015 |
| 测试集上铁心故障工况AP | 0.8905 | 0.9257 |

从表 2中可以看出降维后的故障诊断模型仍然可以有效的对变压器的绕组故障与铁心故障进行识别，此外从铁心故障工况下的PRC可以看出降维后的模型与降维前相比拥有更好的查准能力。这是因为在原始的振动信号中含有噪声成分，噪声与原始振动信号的相关性很小，这也使得因为噪声所导致的主成分贡献率很小，从而使得在选择特征向量构成降维矩阵时不会把噪声所对应的特征给选择进去，最终可以实现间接去噪的效果，从而在一定程度上提升模型的查全能力与查准能力。

此外，降维后也大大简化了模型的复杂程度。图 11(a)表示的是降维前的BP神经网络模型，我们可以看到从输入层到第一隐藏层有60条线段，每条线段都对应了神经网络需要在反向传播时更新的参数。未做降维处理的训练方法耗时为0.6680s（这里标注上实验条件）。图 11(b)表示降维后的BP神经网络模型。从该图可以看见降维后的模型与降维前相比得到了大幅简化，从输入层到第一隐藏层只有12条线段，这也就意味着在每次反向传播时需要更新的参数相比降维前会更少。降维后测试的训练耗时为0.1150s，与降维前相比，训练时间下降了82.78%。综上所述，对训练集经过PCA降维后，能在保证准确率的情况下，能有效的降低模型计算的复杂度，提升模型的训练与预测效率。

1. (b)

图 11 (a)降维前训练出来的BP神经网络;(b) 降维后训练出来的BP神经网络

Fig. 11 (a)[BP-neural network](http://www.baidu.com/link?url=xWnRVuI-i-2m0xeqwUzujxdIwH0OZoPDcD-YtHpSFeAyzWRxeq6Y5H8L085mSSdNIIy-C6RXWCcXRI6T8sLD1zWT9Qa-kNpiBj0BDTlnGDKZt1zq78rumcpbmczUpEjU) befor dimension reduction;(b) [BP-neural network](http://www.baidu.com/link?url=xWnRVuI-i-2m0xeqwUzujxdIwH0OZoPDcD-YtHpSFeAyzWRxeq6Y5H8L085mSSdNIIy-C6RXWCcXRI6T8sLD1zWT9Qa-kNpiBj0BDTlnGDKZt1zq78rumcpbmczUpEjU) after dimension reduction

### 3.4 BP神经网络与常用故障诊断模型训练及诊断结果对比

本文为对比BP神经网络、RBF神经网络、极限学习机、广义回归神经网络、概率神经网络几种常见监督学习方法对于变压器箱体振动信号故障诊断模型的训练效果，分别对这5种监督学习方法训练了故障诊断模型。具体的训练方法如下：

1. 对810条样本数据（正常工况、绕组故障工况、铁心故障工况的数据比例为1:1:1）按照9:1:1的比例随机划分为训练集、测试集与验证集；
2. 对训练集采用PCA进行分析，选择前4个特征向量作为映射矩阵，对训练集、测试集与验证集进行降维；
3. 对降维后的训练集与测试集采用10折交叉验证的方法对模型进行训练，训练完成后得到10个模型；
4. 取这10个模型的输出结果的平均值作为诊断结果，记录在验证集上的ROC，PRC，AUC，AP，计算阈值为0.4时的模型查全率、查准率以及综合准确率。

采用上述方法对BP神经网络、RBF神经网络、极限学习机、广义回归神经网络、概率神经网络进行训练与预测得到的结果如图 12，图 13与表 3所示。

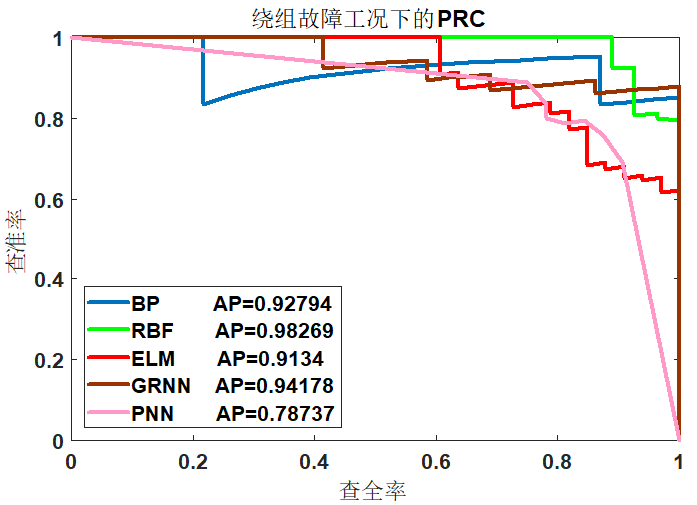
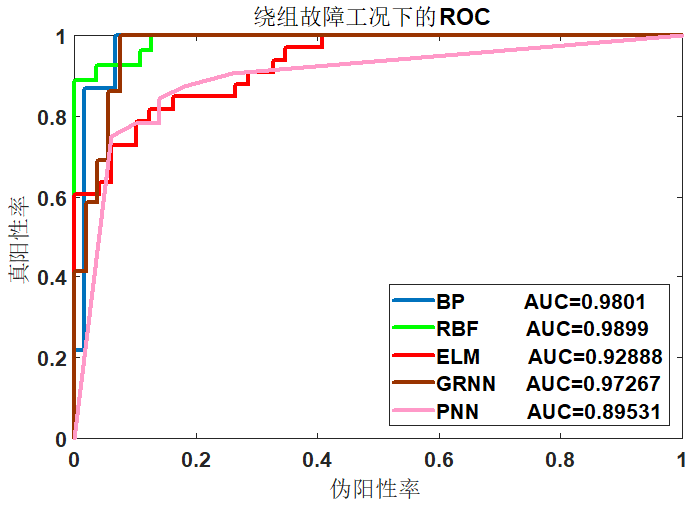


图 12 五种监督学习算法所训练出来的模型对绕组故障工况的ROC与PRC

Fig. 12 ROC and PRC of models trained by five supervised learning algorithms for winding fault conditions

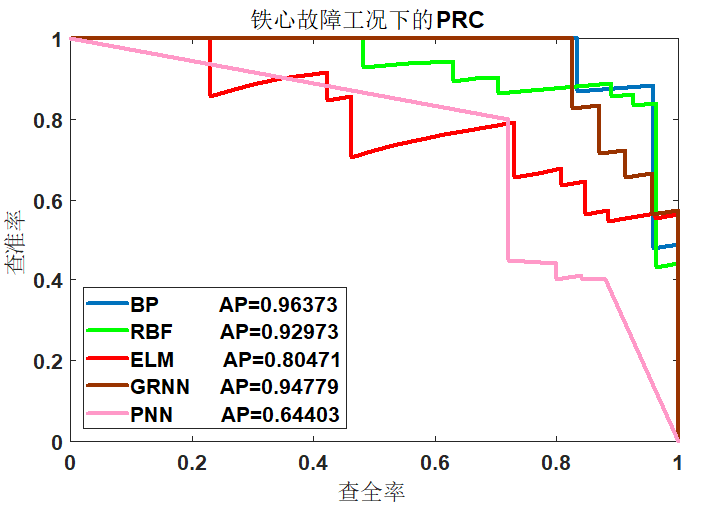
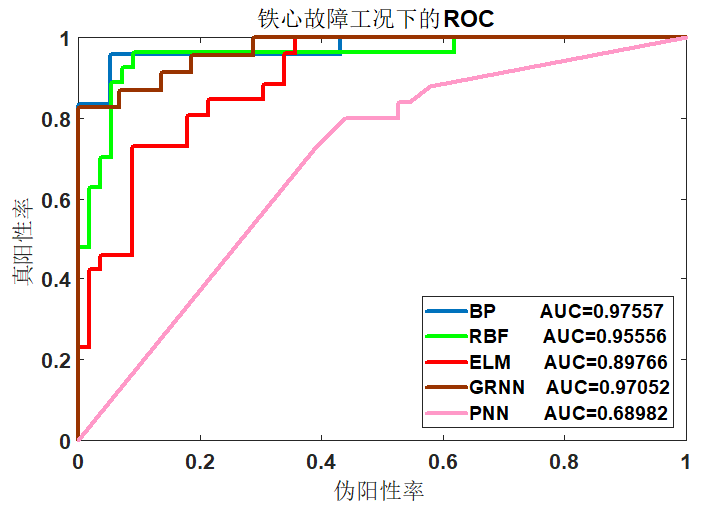


图 13五种监督学习算法所训练出来的模型对铁心故障工况的ROC与PRC

Fig. 13 ROC and PRC of models trained by five supervised learning algorithms for core fault conditions

从图 12可以看出，在绕组故障的查全能力也即AUC的值来看，RBF（AUC=0.9899）略微优于BP（ AUC=0.9801），二者均远优于其他三种模型的AUC。在绕组故障的查准能力也即AP的值来看，RBF（AP=0.9826）最为优秀，其次是是广义回归神经网络（AP =0.9417）与BP神经网络（AP =0.9279）。从图 13可以看出，在铁芯故障的查全能力也即AUC的值来看，BP神经网络（AUC=0.9755）优于广义回归神经网络（AUC=0.9705）与RBF神经网络（AUC=0.9555）。在铁芯故障的查准能力也即AP的值来看，BP神经网络（AP=0.9637）明显优于排在第二的广义回归神经网络（AP =0.9477）与排在第三的RBF神经网络（AP =0.9297）。综合考虑到变压器的故障会对民生安全以及经济财产带来巨大的损失，查全能力相较于查准能力更为重要，BP神经网络对铁心故障的查全能力（AUC=0.9755）是最好的，在对绕组故障的查全能力（AUC=0.9801）上略逊于RBF神经网络（AUC=0.9899）。同时在阈值设置为0.4时，BP神经网络在测试集与验证集上的综合准确率均高于90%，远优于RBF神经网络。可以看出，改进后的BP神经网络作为变压器故障诊断的模型训练方法能够得到较好的查全率与查准率。

表 3五种监督学习算法所训练出来的模型的性能

Table 3 Performance of models trained by five supervised learning algorithms

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | BP | ELM | RBF | PNN | GRNN |
| 测试集上综合准确率 | 0.9026 | 0.6627 | 0.8599 | 0.4884 | 0.8916 |
| 测试集上绕组故障工况查全率 | 0.9356 | 0.9332 | 0.8572 | 0.6916 | 0.9235 |
| 测试集上铁心故障工况查全率 | 0.9030 | 0.8077 | 0.8657 | 0.7849 | 0.8449 |
| 测试集上绕组故障工况查准率 | 0.8501 | 0.6758 | 0.6497 | 0.6866 | 0.8032 |
| 测试集上铁心故障工况查准率 | 0.8877 | 0.5897 | 0.6318 | 0.3957 | 0.9107 |
| 验证集上综合准确率 | 0.9146 | 0.6707 | 0.8415 | 0.5488 | 0.9146 |
| 验证集上绕组故障工况查全率 | 0.9130 | 0.8182 | 0.9899 | 0.8125 | 0.9899 |
| 验证集上铁心故障工况查全率 | 0.9583 | 0.9899 | 0.9300 | 0.8000 | 0.8261 |
| 验证集上绕组故障工况查准率 | 0.8400 | 0.7941 | 0.7714 | 0.7879 | 0.8788 |
| 验证集上铁心故障工况查准率 | 0.8846 | 0.5652 | 0.8317 | 0.4000 | 0.8636 |
| 验证集上绕组故障工况AUC | 0.9801 | 0.9288 | 0.9899 | 0.8953 | 0.9726 |
| 验证集上铁心故障工况AUC | 0.9755 | 0.8976 | 0.9555 | 0.6898 | 0.9705 |
| 验证集上绕组故障工况AP | 0.9279 | 0.9134 | 0.9826 | 0.7873 | 0.9417 |
| 验证集上铁心故障工况AP | 0.9637 | 0.8047 | 0.9297 | 0.6443 | 0.9477 |

同时，该实验中的BP网络模型诊断结果对比前一节PCA有效性严验证实验的结果，可以发现在均采用PCA降维的情况下，本节采用10折交叉验证法对BP神经网络进行模型训练所得到的故障诊断模型对于绕组变形工况下的AUC=0.9801、AP=0.9279,对于铁芯松动工况下的AUC=0.9755、AP=0.9637均优于未采用10折交叉验证法下模型对于绕组变形工况的AUC=0.94648、AP=0.90159,模型对于铁芯松动工况的AUC=0.92574、AP=0.94257。采用10折交叉验证方法所改进后训练出的BP神经网络模型的查全率与查准率均有了明显上升。

## 结束语

在采用经验小波分析+希尔伯特变换方法对变压器箱体振动信号进行特征提取的基础上，本文提出了通过主成分分析与K折交叉验证方法进行数据优化的BP神经网络变压器故障诊断方法，并采用监督学习算法评价指标ROC曲线与PRC曲线对诊断模型进行了综合评价。实验结果表明，采用主成分分析方法与K折交叉验证方法优化后的BP神经网络相较于常用机器学习算法能够更有效的诊断出变压器故障。

# 参考文献

1. [1]【董其国,2001电力变压器故障与诊断[M],北京:中国电力出版社.】
2. [2]【王晓莺,2004变压器故障与监测,.[M],北京:机械工业出版社.】
3. [3]【 李晓兰,黄海,陈祥献等.基于振动法的电为变压器在线状态监测系统设计[J],变压器Transformer.2008,45(口):60-64.】
4. [4]【虞海强.基于振动分析法的变压器状态检测研究[D].西华大学,2012.】
5. 【孙翔，何文林，詹江杨，等 . 电力变压器绕组变形检测与诊断技术的现状与发展 [J]. 高电压技术，2016，42(4) ：1207-1220.】
6. [6]【基于油色谱分析的变压器复合绝缘缺陷多指标综合权重评估方法\_王学磊[J].】
7. [7]【基于BP人工神经网络的电力变压器故障诊断研究\_谭子兵[J].加速度传感器位置】
8. 【振动法在线监测变压器的研究现状与发展】
9. [9]【马宏忠，耿志慧，陈楷，等 . 基于振动的电力变压器绕组变形故障诊断新方法 [J]. 电力系统自动化，2013，37(8) ：89-95.】
10. [10]【李琪菡，雷勇，闫志强，等 . 基于快速傅里叶分析法的油浸式变压器绕组振动特性分析 [J]. 科学技术与工程，2017，17(28) ：211-218.】
11. [11]【赵妙颖,许刚. 基于经验小波变换的变压器振动信号特征提取\_赵妙颖[J]. 电力系统自动化, 2017, 41(20): 63-69, 91.】
12. [12]【基于多故障流形的旋转机械故障诊断\_苏祖强[J].】
13. [13]【基于 BP 人工神经网络的电力变压器故障诊断研究】
14. [14]【张奎,王建南,王肖峰.基于神经网络的变压器故障诊断[J].电子测量技术,2017,40(12):98-101.】
15. [15]【徐睿,梁循,齐金山,李志宇,张树森.极限学习机前沿进展与趋势[J].计算机学报,2019,42(07):1640-1670.】
16. [16]【张宇航,兰生.基于广义神经网络与模糊聚类的变压器故障诊断[J].高压电器,2016,52(05):116-120+125.】
17. [17]【王淼.基于概率神经网络的变压器故障诊断研究[J].中国设备工程,2019(02):99-100.】
18. [18]【基于广义回归神经网络的变压器表面振动基频幅值计算\_李中[J].】
19. [19]【基于电力变压器振动信息的绕组形变诊断方法\_张彬[J].】
20. [20]【熊强.基于PCA和预测误差法的轴承故障预测方法研究[J].现代信息科技,2019,3(16):165-167.】