

DOI: 10.3901/JME.2019.07.009

基于经验模态分解和深度卷积神经网络的行星齿轮箱故障诊断方法*

胡芑庆^{1,2} 陈徽鹏^{1,2} 程哲^{1,2} 张伦^{1,2} 张宇^{1,2}

(1. 国防科技大学智能科学学院 长沙 410072;

2. 国防科技大学装备综合保障技术重点实验室 长沙 410072)

摘要: 行星齿轮箱振动信号具有非平稳特性, 需要一定的先验知识和诊断专业知识设计和解释特征从而实现故障诊断。为了实现行星齿轮箱的智能诊断, 提出一种基于经验模态分解(Empirical mode decomposition, EMD)和深度卷积神经网络(Deep convolutional neural network, DCNN)的智能故障诊断方法。首先对振动信号进行经验模态分解得到内禀模式函数(Intrinsic mode function, IMF); 然后利用 DCNN 融合特征信息明显的 IMF 分量, 并自动提取特征; 最后, 将特征用于分类器分类识别, 从而实现行星齿轮箱故障诊断的自动化。试验结果表明: 该方法能准确、有效地对行星齿轮箱的工作状态和故障类型进行分类。

关键词: 故障诊断; 经验模态分解; 深度卷积神经网络; 行星齿轮箱

中图分类号: TH17

Fault Diagnosis for Planetary Gearbox Based on EMD and Deep Convolutional Neural Networks

HU Niaoqing^{1,2} CHEN Huipeng^{1,2} CHENG Zhe^{1,2} ZHANG Lun^{1,2} ZHANG Yu^{1,2}

(1. College of Mechatronics and Automation, National University of Defense Technology, Changsha 410072;

2. Laboratory of Science and Technology on Integrated Logistics Support,
National University of Defense Technology, Changsha 410072)

Abstract: As the vibration signal of the planetary gearbox is usually nonstationary, a significant level of prior knowledge and diagnostic expertise is required to engineer and interpret features for fault diagnosis. In order to achieve intelligent diagnosis of the planetary gearbox, an intelligent fault diagnosis method based on empirical mode decomposition (EMD) and deep convolutional neural networks (DCNN) is proposed. Firstly, EMD is used to decompose the vibration signal to obtain intrinsic mode function (IMF) components. Then the IMFs with obvious fault character are fused through DCNN and features are extracted automatically. Finally, the learned features serve as the input parameters of classifier to classify working condition, and the atomization of the planetary gearbox fault diagnosis can be implemented. The experimental results show that the method can classify the working state and fault type of the planetary gearbox accurately and effectively.

Key words: fault diagnosis; empirical mode decomposition; deep convolutional neural networks; planetary gearbox

0 前言

行星齿轮箱是风机、直升机、舰船等重大装备和设备用于改变转速和转矩以匹配动力源和负载的关键传动部件^[1], 具有质量小、体积小、减速比大及传动效率高等优点^[2]。其低速重载的工作环

境常常导致太阳轮、行星轮和齿圈等关键部件发生严重磨损和疲劳裂纹等故障^[3]。装备设备旋转部件的故障如果置之不理, 不及时采取恰当的维修作业, 将很可能导致整个传动系统停机, 甚至引发灾难性后果。因此, 研究行星齿轮箱故障诊断方法对保持装备和设备的技术状态, 延长服役时间具有重要意义。

然而, 行星齿轮箱的故障诊断问题面临诸多棘手问题。一方面, 行星齿轮箱复杂的内部结构造成

* 国家自然科学基金(51475463, 51775550)和湖南省自然科学基金(2018JJ3604)资助项目。20180709 收到初稿, 20181212 收到修改稿

振动传递路径复杂,导致故障响应微弱,而且常常受到外界环境噪声的干扰,使得信号的信噪比低。另一方面,行星齿轮箱中零部件种类数目较多,齿轮的运动是自转与公转的复合运动,振动响应复杂,各部件振动相互耦合,测取的振动信号呈现出明显的非线性,而且载荷的大范围瞬时波动引起振动具有强烈的非平稳性^[3-4]。因此有必要对采集得到的原始信号进行预处理来增强特征信息,以获得更好的诊断结果。

经验模态分解(Empirical mode decomposition, EMD)是故障诊断中广泛使用的时频域信号分析方法,能把复杂的信号分解为一系列内禀模式函数(Intrinsic mode function, IMF)之和^[5-6],各个 IMF 分量代表了信号一个固有的振动模态,包含了不同的特征时间尺度,使信号特征在不同的分辨率下显示出来。由于每一个 IMF 所包含的频率成分不仅仅与采样频率有关,而且还随着信号本身的变化而变化,因此 EMD 方法是一种自适应的信号分解方法,具有很高的信噪比,非常适合处理时变、非平稳信号。基于 EMD 方法,学者们在故障诊断领域已经取得了一定的成果。于德介等^[4]针对齿轮振动信号的非平稳特征的实际情况,提出了一种基于支持向量机(Support vector machine, SVM)和 EMD 的齿轮故障诊断方法;付大鹏等^[7]在应用 EMD 进行轴承故障特征提取的基础上,结合 SVM 良好的分类性能进行模式识别,实现了故障诊断;杨宇等^[8]提出以 EMD 为预处理器提取各频带能量作为特征参数的神经网络诊断方法,成功应用于滚动轴承工作状态和故障类型识别。这些研究采用的模式识别方法主要是基于“浅层学习”的算法,对信号采集和处理的要求较高,对复杂函数的表达能力有限,模型的泛化能力较差^[9]。

深度学习起源于人工神经网络,它模仿人脑的机制对多源异类数据进行解释,在计算机视觉^[10-11]、语音识别^[12-13]等领域已经取得了辉煌的成果。深度学习强大的特征学习能力,满足机械故障诊断的自适应特征提取要求,将其应用于故障诊断领域能减少对专家故障诊断经验和信号处理技术的依赖,降低传统故障诊断方法由于人工设计并提取特征引入的不确定性。另外,深度学习通过构建含有多个隐含层的深层模型,以非线性映射的方式多级抽象来学习得到更深层次的特征表征监测数据与故障状态之间复杂的映射关系,非常适合于复杂数据背景下故障诊断需求。因此,研究基于深度神经网络(Deep neural network, DNN)从数据中自适应学习特征进行故障诊断具有重要意义。

卷积神经网络(Convolutional neural network, CNN)是 DNN 的典型模型之一,其实质是构建多个滤波器对输入数据进行逐层卷积及池化,逐层提取输入数据的特征。其独特的网络结构有效减少训练参数数量,降低了网络的复杂度,同时保证对一定程度上的平移、旋转及缩放具有不变性。因此, CNN 能高效地提取出海量数据中蕴含的特征信息,是处理大数据非常好的选择。CNN 自 LECUN 等^[14]首次提出至今,网络深度不断加深, CNN 的概念也转变为深度卷积神经网络(Deep convolutional neural network, DCNN),相关理论与结构得到快速发展。截止目前,发展得到了许多 DCNN 结构,如 VGG-net^[15], Res-net^[16]和 inception v4^[17]。DCNN 在图像识别领域取得了相当成功的应用,在机械故障诊断领域也有一定的成果。JANSSENS 等^[18]探索研究了用原始频域数据作为模型的输入来进行轴承的故障诊断。采集到的轴承座振动数据先通过 FFT 进行预处理,然后输入模型中以检测轴承状态,实现了无需专家经验的故障检测与识别算法;GUO 等^[19]构建了由两个 CNN 组合而成的分层深度卷积神经网络来监测滚动轴承的健康状况。一个 CNN 用于从原始振动数据中学习特征并识别故障模式,另一个用于评估每个故障模式的故障大小,将原始振动信号作为 CNN 故障诊断模型的输入,无需手工特征提取过程;文献[20]研究轴承的故障程度与故障位置的诊断与识别,达到了远优于传统神经网络、SVM 等方法的准确率;文献[21]中提出了 WDCNN 模型进行轴承故障诊断,即使在嘈杂的环境中也可以实现高诊断精度。

本文借鉴深度学习相关的研究进展,在前人研究的基础上,重点研究行星齿轮箱的智能故障诊断。针对行星齿轮振动信号的非平稳特征,将 EMD 方法与 DCNN 结合,以 EMD 作为振动信号的预处理器,根据各分量的峭度值选择故障特征信息明显的 IMF 分量,然后通过 DCNN 自适应地融合 IMF 分量的信息并从中提取特征,完成对行星齿轮箱状态的智能分类识别。最后通过试验数据将该方法与其他故障诊断方法进行了对比与分析,结果显示所提出的故障诊断方法性能远优于 SVM 和反向传播神经网络(Back propagation neural network, BPNN)这两种传统诊断方法,与 WDCNN 这目前先进的诊断方法相比,准确率更高,且诊断结果稳定性更好。

1 经验模态分解方法

经验模态分解方法由 HUANG 等^[5]于 1998 年提

出, 该方法适用于分析和处理非线性、非平稳随机信号。EMD 方法基于信号的局部特征时间尺度, 将信号分解为一系列内禀模式函数 IMF 之和。IMF 分量必须满足下面两个条件。

(1) 具有相同数量的极值点和过零点(或至多相差一个)。

(2) 任意时刻, 极大值和极小值的包络关于时间轴局部对称, 平均值为零。

在满足上述条件的基础上, 对任一实信号 $x(t)$ 进行 EMD 的步骤如下。

(1) 确定信号 $x(t)$ 所有局部极值点, 然后用三次样条曲线将所有的局部极大值点连接起来形成上包络线; 同样, 用三次样条曲线将所有的局部极小值点连接起来形成下包络线。

(2) 计算上下包络线的平均值 m_1 , 然后求出

$$h_1 = x(t) - m_1 \quad (1)$$

判断 h_1 是否满足 IMF 的条件, 若满足, 则 h_1 为 $x(t)$ 的第 1 个 IMF 分量。

(3) 如 h_1 不满足 IMF 的条件, 把 h_1 作为原始数据, 重复步骤(1)和(2) k ($k=1, 2, \dots$) 次, 直到 $h_{1k} = h_{1(k-1)} - m_{1k}$ 满足 IMF 的条件(其中 m_{1k} 为第 k 次重复过程上下包络线的平均值)。记 $c_1 = h_{1k}$, 则 c_1 为信号 $x(t)$ 的第 1 个 IMF 分量。

(4) 从 $x(t)$ 中将 c_1 分离出来, 得到

$$r_1 = x(t) - c_1 \quad (2)$$

将 r_1 作为原始数据重复步骤(1)~(3), 得到 $x(t)$ 的第 2 个 IMF 分量 c_2 , 重复循环 n 次, 得到 n 个 IMF 分量。这样就有

$$r_1 - c_2 = r_2, \dots, r_{n-1} - c_n = r_n \quad (3)$$

当满足给定的终止条件(常设定的终止条件为 r_n 成为一个单调函数不能再从中提取满足 IMF 条件的分量)时, 循环结束。这样由式(2)和式(3)得到

$$x(t) = \sum_{i=1}^n c_i + r_n \quad (4)$$

式中, r_n 称为残余函数, 代表信号的平均趋势。

EMD 方法通过“筛分”过程将信号分解为一系列平稳的 IMF 分量。在 EMD 方法中, 相邻两个极大值点或两个极小值点之间的时间跨度被称为时间特征尺度, 反映了信号不同模态的局部振动特征。

“筛分”过程按照特征时间尺度从小到大的顺序依次分离各 IMF, 分别包含信号从高到低不同频率段的成分, 因此可以把 EMD 方法看成是一组高通滤波器。“筛分”过程如图 1 所示。

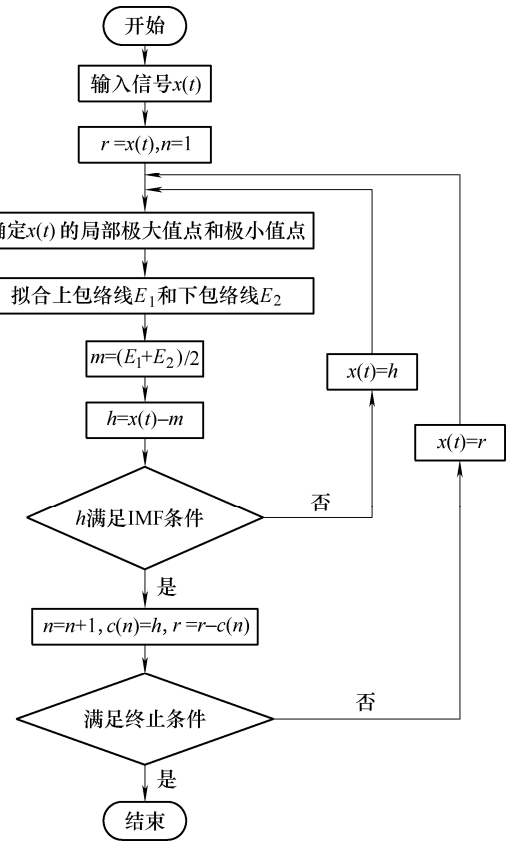


图 1 EMD 算法的流程图

2 DCNN 简介

2.1 DCNN 结构

DCNN 是一个典型的前馈神经网络, 其实质是构建多个能够提取输入数据特征的滤波器。通过局部感受野, 共享权重和空间降采样来保持输入的初始信息, 最终获得输入数据的平移、旋转及缩放不变的特征表示。

典型的 DCNN 由特征提取阶段和分类阶段构成, 其中特征提取阶段由卷积和池化交替进行, 分类阶段则通常由多层神经网络与分类器组成。图 2 给出了典型的 DCNN 架构图, 本文中特征提取阶段组成单元结构主要包括卷积层、批归一化层、激活层和池化层。在接下来的部分, 将详细介绍 DCNN 各层实现原理及作用。

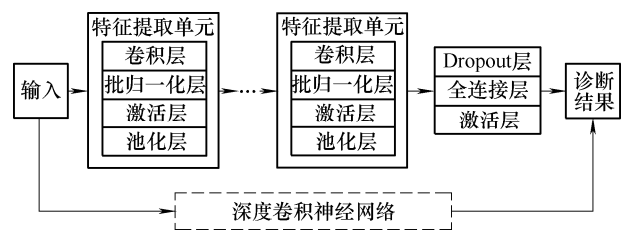


图 2 DCNN 典型结构

卷积层是 DCNN 的核心组成部分, 由多个特征

图组成,每个特征图上的神经元节点又通过滤波器(卷积核)与上一层特征图的局部区域相连,即稀疏连接。每一个滤波器作用于局部感受野,通过设定的移动步长作用于全部感受野,完成对各输入特征图的卷积滤波,提取局部特征,输出特征图,而输出特征图的数量就是该卷积层的深度。另外,同一输入特征图与同一输出特征图之间,滤波器的参数共享,包括相同的权重矩阵和偏置项^[22]。

卷积层的数学模型可以简述为

$$\mathbf{X}_k^{(m)} = \sum_{c=1}^C \mathbf{W}_k^{(c,m)} * \mathbf{X}_{k-1}^{(c)} + \mathbf{B}_k^{(m)} \quad (5)$$

式中, k 表示第 k 层卷积层,运算符 $*$ 表示对各层(c 为层数, $c=1, \dots, C$)输入特征图 $\mathbf{X}_{k-1}^{(c)}$ 与权重矩阵 $\mathbf{W}_k^{(c,m)}$ 做卷积运算,对应得到第 m 个输出特征映射 $\mathbf{X}_k^{(m)}$ ($m=1, \dots, M$)。 $\mathbf{B}_k^{(m)}$ 表示偏置项。

在每个卷积层之后,通常会应用一个非线性层(或激活层)。其目的是给一个在卷积层中刚经过线性计算操作的系统引入非线性特征。在本论文中,选择整流线性单元^[23](Rectified linear units, ReLU)来加速卷积神经网络的收敛,ReLU 的函数表达如下

$$f(x) = \max\{0, x\} \quad (6)$$

池化层使用某一位置相邻输出的总体统计特征来代替网络在该位置的输出,是对上一层数据图的缩放映射,减少了数据维度与网络参数。使用池化层既可以加快计算速度,同时还可以防止过拟合。其中最大池化函数给出相邻矩形区域内的最大值,是最常用的池化函数。本文的网络结构中就使用该池化结构,它的数学表达如下

$$P_i^{l+1}(j) = \max_{(j-1)W+1 \leq t \leq jW} \{q_i^l(t)\} \quad (7)$$

批归一化(Batch normalization, BN)层^[24]是对每一批数据进行归一化的操作,不仅解决中间层数据分布发生改变的情况,而且在一定程度上解决了深层网络中“梯度弥散/爆炸”的不稳定问题,加快模型的收敛速度,从而使得训练深层网络模型更加容易和稳定。在 DCNN 中, BN 层作用在非线性映射前。

DCNN 的分类阶段一般由全连接层和分类器组成,主要对来自先前层的信息进行整合、分类。在 DCNN 中,常采用的分类器为 Softmax 分类器,采用 Softmax 逻辑回归进行分类,估计得到样例属于不同种类的概率分布情况。Softmax 回归的输出表达如下

$$f(\theta^{(i)} x) = P(y=i|x;\theta) = \frac{\exp(\theta^{(i)} x)}{\sum_{j=1}^K \exp(\theta^{(j)} x)} \quad (8)$$

式中, $\theta^{(i)}$ ($1 \leq i \leq K$) 是模型的参数,得到的 $f(\theta^{(i)} x)$ 即为 DCNN 的输出。

2.2 训练方法

选择交叉熵作为代价函数来表达真实值与估计值的误差

$$L = -\sum_x p(x) \ln q(x) \quad (9)$$

式中, $p(x)$ 表示真实的分类结果, $q(x)$ 是 Softmax 输出的分类结果。

使用 ADAM^[25]优化算法来训练 DCNN,确定网络参数,即权重和偏差。ADAM 能够利用梯度的一阶矩估计(平均值)和二阶矩估计(方差)动态调整每个参数的学习率,在优化 DCNN 的学习速率方面已经取得了成功。

2.3 DCNN 设计

DCNN 在图像识别领域已经被广泛运用,图像具有平移不变性,而具有二维卷积结构的 DCNN 能够有效提取出输入中具有空间相关性的特征信息。虽然振动信号是一维时间序列,但同样具有一定程度的平移不变性,一维卷积结构的 DCNN 同样适合处理振动信号。

DCNN 的性能与卷积层的数量和卷积核的大小密切相关。首先,DCNN 的学习能力与层数正相关。在一定程度上,网络结构越深,表达能力就越强。如果 DCNN 结构太简单,学习能力差,不能有效提取输入中蕴含的信息。而网络的深层结构意味着需要更多的训练数据来训练大量的参数。否则可能产生过拟合,局部最优等问题。同时,由于网络复杂程度上升,对模型进行训练非常耗时,增加了建立合适、高效模型的难度^[26]。其次,卷积核尺寸大小需要根据一些基本的设计原则^[21, 27]进行测试和确定。与小卷积核相比,第一卷积层的卷积核尺寸越大,感受野区域越大,可以获取更多的数据并为更深层提供更多信息。更重要的是,对于机械故障诊断而言,大卷积核可以更好地抑制高频噪声。另外,考虑将第一层卷积层之后的卷积核设置为小尺寸以加深网络,提高网络的表达能力,从而学习获得输入信号良好的特征表示。

尽管上面提到了设置超参数的一些基本原则,但在实际过程仍需要进行大量调试,得到合适的参数值。

3 故障诊断方法流程

基于 EMD 和 DCNN 进行行星齿轮箱故障定

性诊断时, 分为五个步骤: 信号采集与划分、样本信号 EMD 分解、数据集创建、DCNN 设计与

训练以及行星齿轮箱故障定性诊断, 流程如图 3 所示。

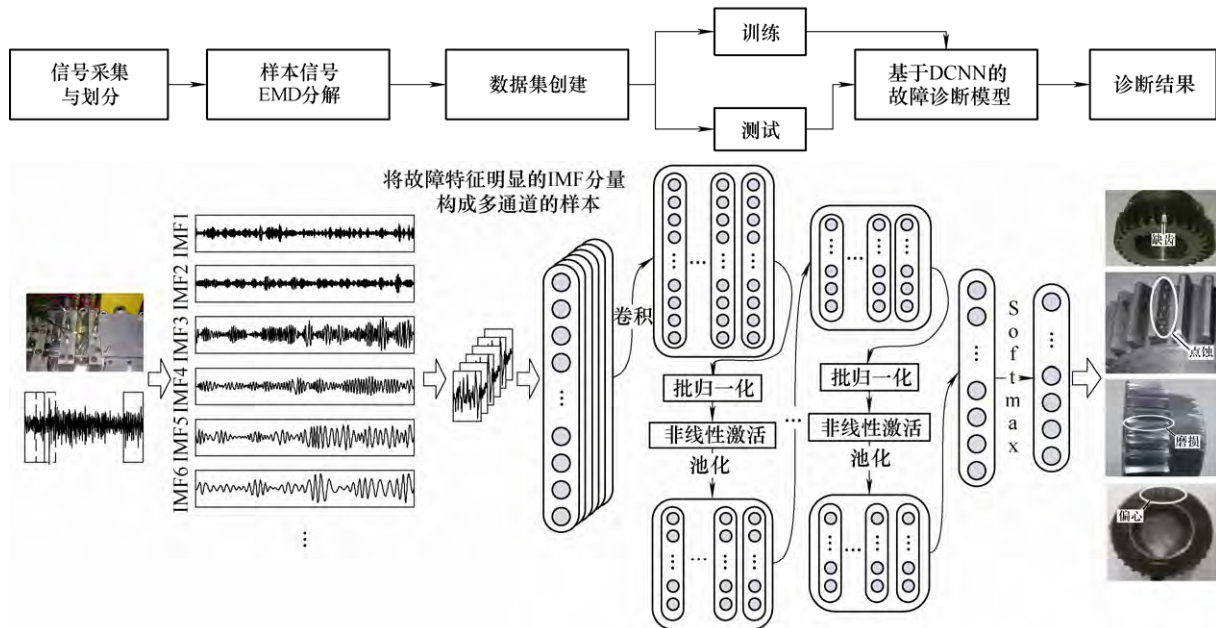


图 3 基于 EMD 和 DCNN 的行星齿轮箱故障诊断方法流程

(1) 信号采集与划分: 通过加速度传感器获取齿轮箱振动信号, 用相等长度的窗口进行划分, 获取样本信号。当窗口移动步长小于单个样本的信号长度时, 样本之间存在重叠部分, 固定长度的信号能够提取更多的样本, 将此数据增强处理称为重叠切片。使用该方法, 长度为 l_{signal} 的时域信号可以划分成 $N_{\text{samples}} = (l_{\text{signal}} - N_{\text{in}}) / s$ 个样本, 其中样本的长度为 N_{in} , 切片窗口的移动距离为 s 。

(2) 样本信号 EMD 分解: 对振动信号重叠切片获得的样本信号进行 EMD 分解, 把非平稳信号 $x(t)$ 分解为一系列具有不同特征尺度的平稳信号 $c^{(1)}, c^{(2)}, \dots, c^{(n)}$, 然后通过分析 IMF 分量的峭度, 选择故障特征信息明显的 6 个 IMF 分量。

(3) 数据集创建: 将故障特征明显的 6 个 IMF 分量按规定的顺序堆叠成一个多通道样本(同一状态下顺序相同即可), 样本数据维度增加一维。将所有的样本信号进行上述操作, 创建数据集, 并将数据集划分为训练集与测试集。

(4) DCNN 设计与训练: 按照第 2.3 节所述设计原则设计可行的 DCNN 并使用训练集进行训练, 调试超参数, 获得性能较好的 DCNN。

(5) 行星齿轮箱故障定性诊断: 通过测试集验证基于 EMD 和 DCNN 的行星齿轮箱故障诊断模型的有效性。

该方法通过 DCNN 将故障特征信息明显的 IMF 分量进行自适应的融合, 通过训练得到的卷积核参

数就是各分量融合的权重, IMF 分量自适应融合的示意如图 4 所示, 其中同一输入特征图与同一输出特征图之间参数共享, 输出特征图中的神经元由各分量同一位置的元素共同决定。

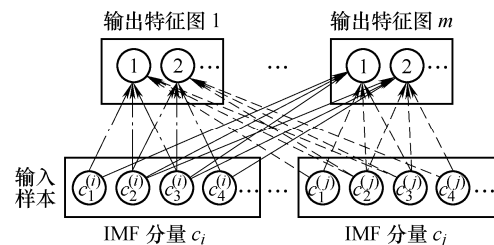


图 4 IMF 分量的自适应融合示意图

4 试验

4.1 试验说明

进行故障植入试验, 采集振动数据并进行处理来验证本文所提出的故障诊断方法。

图 5 所示为行星齿轮箱试验装置及其原理图, 它由驱动电动机、斜齿轮箱、两级的行星齿轮箱、扭矩传感器、两个两级的增速齿轮箱、加载电动机、润滑系统和控制系统等组成。

试验在第一级行星齿轮箱进行, 在太阳轮上人工植入磨损、点蚀、缺齿、偏心缺陷用作故障植入部件, 如图 6 所示。共安装 4 个加速度传感器用于同步采集振动数据, 每一级行星轮上各有

两个, 分别测量水平方向和竖直方向的加速度, 第一级行星齿轮的垂直方向振动数据用于本文研究。转速传感器安装于第二级行星齿轮输出轴端。传感器的安装如图 7 所示, 试验过程中参数设置如表 1 所示。

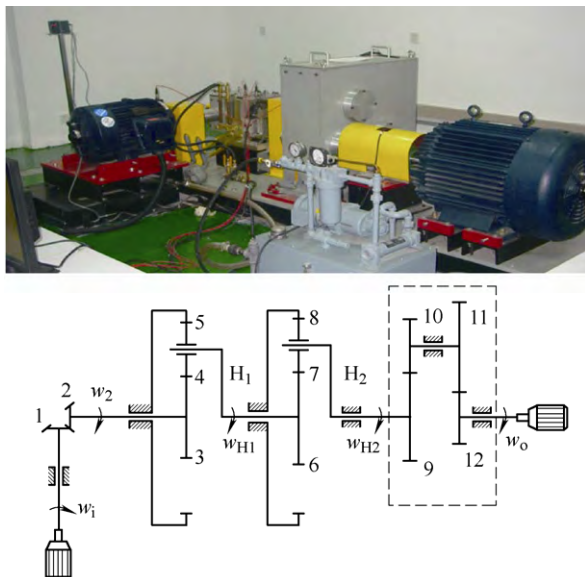


图 5 行星齿轮箱试验装置及其原理图

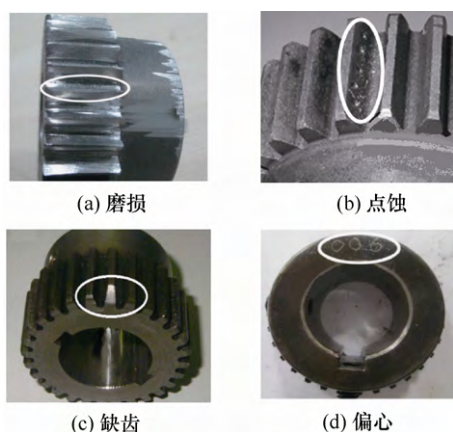


图 6 太阳轮故障类型

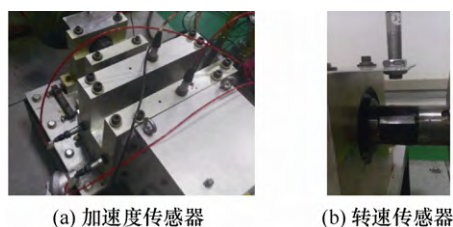


图 7 传感器安装

表 1 试验参数设置表

参数	数值
电动机转速/(r/min)	1 200
采样频率/kHz	5
负载/(N·m)	41.2

4.2 数据说明

将各状态下采集得到的第一级行星齿轮箱的振动信号按照第 3 节所述的方法进行处理。其中每段样本信号包含 2 048 个数据点, 各状态下得到 660 个训练样本信号和 25 个测试样本信号。其中某一太阳轮缺齿故障的行星齿轮箱振动加速度样本信号如图 8 所示。将其进行 EMD 分解后, 取前 8 个 IMF 分量如图 9 所示, 不同的 IMF 分量包含了不同的时间尺度, 可以使信号的特征在不同的分辨率下显示出来。而行星齿轮产生的故障特征信息仅包含在部分 IMF 中。在行星齿轮啮合过程中会产生冲击分量, 冲击分量通常包含故障特征信息。因此, 选择峭度为有效 IMF 的判断标准, 各 IMF 分量的峭度如图 10 所示。

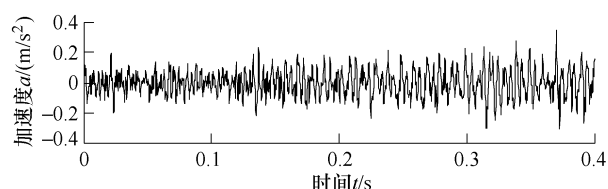


图 8 太阳轮缺齿故障振动加速度信号

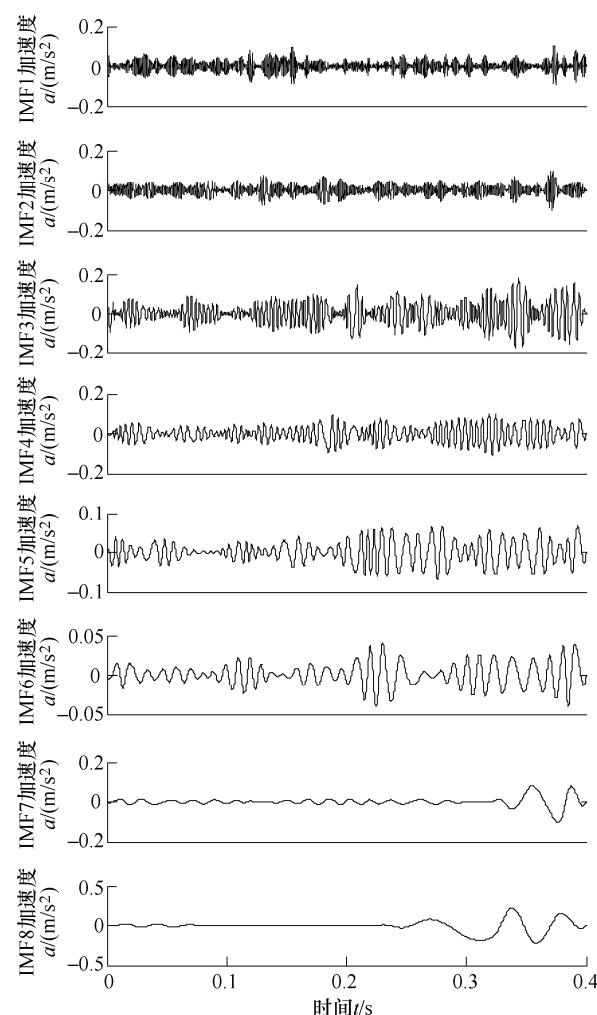


图 9 太阳轮缺齿故障振动信号前 8 个 IMF 分量

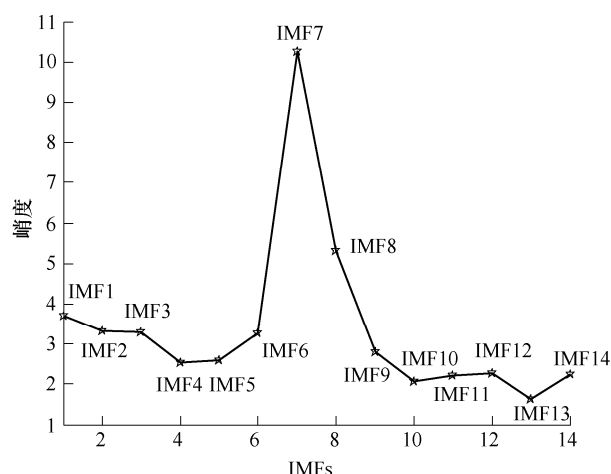


图 10 太阳轮缺齿故障信号各 IMF 分量峭度

通常, 峰度越大, 信号中包含更多的冲击成分。因此, 选择峭度最大的六个分量 IMF1, IMF2, IMF3, IMF6, IMF7 和 IMF8 作为有效 IMF, 其包含较多由行星齿轮故障产生的振动特征信息。然后将故障特征明显的 6 个 IMF 分量按峭度从大到小的顺序堆叠成一个多通道样本, 样本数据维度由 1×2048 变为 $1 \times 2048 \times 6$ 。将所有的样本信号进行上述操作, 创建数据集, 并将数据集划分为训练集与测试集, 最终得到的数据集如表 2 所示。

表 2 行星齿轮箱数据集的描述

齿轮状态	训练集	测试集	类别标签
健康	660	25	1
磨损	660	25	2
点蚀	660	25	3
缺齿	660	25	4
偏心	660	25	5

4.3 DCNN 模型参数

根据第 2.3 节所述的设计原则设计 DCNN, 并通过反复试验调整相关参数, 最终得到的模型结构及参数如表 3 所示。其中 CW 表示卷积核宽度; CH 表示卷积核高度; CC 表示当前层输入特征图深度, 即输入特征图数量; CN 表示卷积核深度, 即输出特征图数量; S 指池化层池化带宽度; Strides 表示卷积核的移动步长; Dropout 表示以一定概率 p 将网络中的神经元置零, 可以在一定程度上抑制过拟合问题, 提高 DCNN 的泛化性能。另外, 所有的卷积层都进行了称为 SAME 的边缘处理, 在输入矩阵的边界上加入全 0 填充, 避免尺寸的变化, 而卷积层与池化层之间都存在批归一化层与激活层。

设计的 DCNN 由 5 个特征提取单元进行特征提取, 一个全连接层进行特征整合并通过 Softmax 层进行分类识别。第一级特征提取单元中卷积层的卷积核尺寸为 64×1 , 增大了模型输入的感受野, 增加

了输入信息量, 并且同短时傅里叶变换类似, 其提取数据中的短时特征。之后共有 4 个特征提取单元, 其中每层卷积层中卷积核大小都为 3×1 , 实现了网络的深层结构, 提高网络的表达能力。

表 3 DCNN 模型主要结构参数

结构	参数
卷积层	CW = 64; CH = 1; CC = 6; CN = 16; Strides: 2
池化层	S = 2
卷积层	CW = 3; CH = 1; CC = 16; CN = 32; Strides: 1
池化层	S = 2
卷积层	CW = 3; CH = 1; CC = 32; CN = 32; Strides: 1
池化层	S = 2
卷积层	CW = 3; CH = 1; CC = 32; CN = 64; Strides: 1
池化层	S = 2
卷积层	CW = 3; CH = 1; CC = 64; CN = 64; Strides: 1
池化层	S = 2
全连接层	节点数: 100, 激活函数: ReLU
Dropout	比率: 0.5
Softmax	输出节点数: 5, 激活函数: Softmax

4.4 结果及分析

4.4.1 评估指标

为了量化模型分类诊断的性能, 计算四个指标: 准确率(Accuracy), 精确率(Precision), 召回率(Recall)和 F_1 -score, 各指标的计算公式分别见式(10)~(13)

$$A_{cc} = \frac{A + D}{A + B + C + D} \quad (10)$$

$$P = \frac{A}{A + B} \quad (11)$$

$$R = \frac{A}{A + C} \quad (12)$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} \quad (13)$$

式中, A 为某一类别样本被正确识别的数目(假设样本标签为 y); D 为其他类别(不属于类别 y)的样本被正确识别的数目; B 为其他类别的样本被错误识别为类别 y 的数目; C 为类别 y 的样本识别错误的数目。选择这些指标是因为它们直接反映了状态监测需求。如果状态监测系统在分类器检测到故障时能够触发警报, 那么即使故障中存在一些错误警报, 也能对所有故障进行警报。然而, 维修人员并不希望误警太多, 因为这会导致不必要的停机而增加操作成本。换句话说, 如果触发的警报越多, 越多的故障会被维修人员意识到, 这时召回率越高。但是也会有更多的错误警报, 这就导致精确率降低。另一方面, 如果只标记了真正的故障, 没有误警, 但是有些故障漏警了, 那么精确率将会较高, 但是召回率较低。一个好的分类器能够使两者最大化, 这

样才能保证在确实发生故障时才触发警报,不会错过任何故障,并且不会出现错误警报情况。而 F_1 -score 是精确率和召回率的加权调和平均,综合了两者的结果。此外,准确率表示对于给定的测试数据集,分类器正确分类的样本数与总样本数之比,也常用于评估模型性能。

4.4.2 试验结果

表 4 总结了基于同一数据集,使用本文所提出的方法、传统诊断方法 SVM 和 BPNN,及文献[21]所提出的 WDCNN 所实现的行星齿轮箱状态监测结果。其中 SVM 使用高斯径向基函数(RBF)作为核函数, BPNN 结构为 2048-200-100-50-5,其激活函数为 ReLU, WDCNN 将原始数据作为模型输入。另

外,为了消除偶然误差的影响,每种方法各进行了 10 次试验,将各类评估指标的平均值作为每次试验结果,10 次试验结果的平均值作为该方法的分类诊断性能评估指标。可以看出,本文提出的基于 EMD 和 DCNN 的诊断方法明显优于 SVM 和 BPNN 两种传统诊断方法。而与目前先进的诊断方法 WDCNN 相比,本文提出的方法不仅识别各指标都达到了 100%,优于 WDCNN,而且诊断结果稳定,表现为多次试验得到的评估指标的标准差为 0。原因在于经过 EMD 处理后提高了信号的信噪比,而且将峭度较大的有效 IMF 分量构造为 DCNN 的多通道输入,通过 DCNN 进行信息融合,自适应地得到各 IMF 分量对输出的权重。

表 4 本文所提出的方法与 SVM、BPNN 性能比较(%)

评估指标	EMD+DCNN	SVM	BPNN	WDCNN
准确率	100($\sigma=0\%$)	72.16($\sigma=2.06\%$)	68.73($\sigma=3.51\%$)	98.40($\sigma=0.65\%$)
精确率	100($\sigma=0\%$)	74.56($\sigma=2.06\%$)	69.66($\sigma=3.71\%$)	98.45($\sigma=0.64\%$)
召回率	100($\sigma=0\%$)	72.16($\sigma=2.06\%$)	68.73($\sigma=3.51\%$)	98.40($\sigma=0.65\%$)
F1-score	100($\sigma=0\%$)	72.85($\sigma=2.04\%$)	67.13($\sigma=3.07\%$)	98.39($\sigma=0.66\%$)

4.4.3 t-SNE 可视化分析

t-SNE^[28]是一种非线性降维算法,用于可视化高维数据。通过 t-SNE 来研究 DCNN 各层学习的特征分布,降维并进行归一化后的二维散点图如图 11

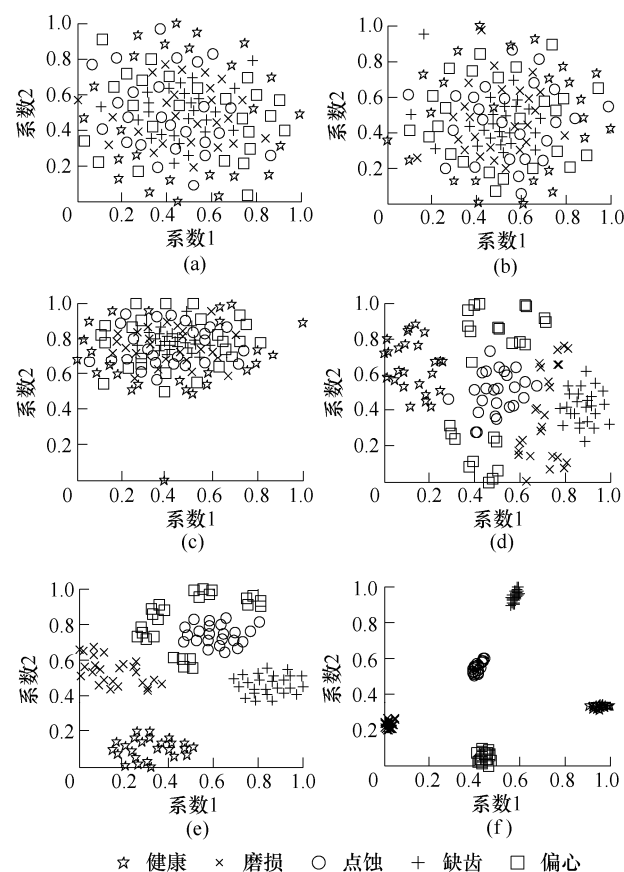


图 11 测试样本输入 DCNN 后通过 t-SNE 特征可视化

所示。其中,图 11a~11e 为各激活层输出;图 11f 为 DCNN 全连接层输出。由于散点图的坐标轴无实际意义,因此用系数 1、2 分别作为两个坐标的名称。可以发现第三层开始各状态的特征已经开始慢慢各自聚拢,而到了第四层各状态特征之间已经基本分散。再经过一层卷积运算,类间距变大而类内距变小,各类明显可分。全连接层对特征进行整合处理,进一步缩小同一状态下特征分布的距离,增大不同状态下特征分布的距离,使得特征的可分性更好。整个过程显示出了 DCNN 强大的非线性映射能力。

5 结论

(1) 本文针对振动信号通常是非平稳并噪声污染严重的问题,提出一种基于经验模态分解和深度卷积神经网络的故障诊断方法,实现数据的自适应处理,简化了故障诊断的操作。

(2) 结合 EMD 分析和处理非线性、非平稳随机信号的优势及 DCNN 强大的自动特征提取能力,实现行星齿轮箱的故障诊断的智能化。

(3) 通过峭度确定有效 IMF 分量,构建含振动信息更集中的数据集,提升了 DCNN 的分类性能。

(4) 构建第一层卷积层卷积核尺寸大,深层卷积核尺寸小的 DCNN,进一步提升网络融合 IMF 分量信息并自动提取状态特征的能力。

(5) 通过试验数据测试,所提出的故障诊断方

法能达到 100% 的准确性, 远优于 SVM 和 BPNN 这两种传统诊断方法。与目前先进的诊断方法 WDCNN 相比, 准确率更高, 且诊断结果稳定性更好。

参 考 文 献

- [1] LEI Yaguo, LIN Jing, ZUO Mingjian, et al. Condition monitoring and fault diagnosis of planetary gearboxes: A review[J]. Measurement, 2014, 48(1): 292-305.
- [2] KHAZAEE M, AHMADI H, OMID M, et al. Feature-level fusion based on wavelet transform and artificial neural network for fault diagnosis of planetary gearbox using acoustic and vibration signals[J]. Insight: Non-Destructive Testing and Condition Monitoring, 2013, 55(6): 323-329.
- [3] 雷亚国, 何正嘉, 林京, 等. 行星齿轮箱故障诊断技术的研究进展[J]. 机械工程学报, 2011, 47(19): 59-67.
LEI Yaguo, HE Zhengjia, LIN Jing, et al. Research advances of fault diagnosis technique for planetary gearboxes[J]. Journal of Mechanical Engineering, 2011, 47(19): 59-67.
- [4] 于德介, 杨宇, 程军圣. 一种基于 SVM 和 EMD 的齿轮故障诊断方法[J]. 机械工程学报, 2005, 41(1): 140-144.
YU Dejie, YANG Yu, CHENG Junsheng. Fault diagnosis approach for gears based on EMD and SVM[J]. Chinese Journal of Mechanical Engineering, 2005, 41(1): 140-144.
- [5] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 1998, 454(1971): 903-995.
- [6] HUANG N E, SHEN Z, LONG S R. A new view of nonlinear water waves: The Hilbert spectrum[J]. Annual Review of Fluid Mechanics, 1999, 31(1): 417-457.
- [7] 付大鹏, 翟勇, 于青民. 基于 EMD 和支持向量机的滚动轴承故障诊断研究[J]. 机床与液压, 2017, 45(11): 184-187.
FU Dapeng, ZHAI Yong, YU Qingmin. Study on fault diagnosis of rolling bearing based on EMD and support vector machine[J]. Machine Tool & Hydraulics, 2017, 45(11): 184-187.
- [8] 杨宇, 于德介, 程军圣. 基于 EMD 与神经网络的滚动轴承故障诊断方法[J]. 振动与冲击, 2005, 24(1): 85-88.
YANG Yu, YU Dejie, CHENG Junsheng. Roller bearing fault diagnosis method based on EMD and neural network[J]. Journal of Vibration and Shock, 2005, 24(1): 85-88.
- [9] 任浩, 屈剑锋, 柴毅, 等. 深度学习在故障诊断领域中的研究现状与挑战[J]. 控制与决策, 2017, 32(8): 1345-1358.
REN Hao, QU Jianfeng, CHAI Yi, et al. Deep learning for fault diagnosis: The state of the art and challenge[J]. Control & Decision, 2017, 32(8): 1345-1358.
- [10] SOCHER R, HUVAL B, BATH B, et al. Convolutional-recursive deep learning for 3d object classification[C]// Advances in Neural Information Processing Systems, December 3-6, 2012, Lake Tahoe, Nevada. New York: ACM, 2012: 656-664.
- [11] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436.
- [12] DENG L, LI J, HUANG J T, et al. Recent advances in deep learning for speech research at Microsoft[C]// International Conference on Acoustics, Speech and Signal Processing (ICASSP). May 26-31, 2013, Vancouver, BC, Canada. Piscataway: IEEE, 2013: 64.
- [13] HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups[J]. IEEE Signal Processing Magazine, 2012, 29(6): 82-97.
- [14] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv Preprint arXiv: 2014(1): 1409.1556.
- [16] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]// Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016. Las Vegas, NV, USA. Piscataway: IEEE, 2016: 770-778.
- [17] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]// Association for the Advancement of Artificial Intelligence (AAAI), February 04-09, 2017, San Francisco, San Francisco, California. Menlo Park: AAAI, 2017: 12.
- [18] JANSSENS O, SLAVKOVIKJ V, VERVISCH B, et al. Convolutional neural network based fault detection for rotating machinery[J]. Journal of Sound and Vibration, 2016, 377: 331-345.
- [19] GUO Xiaojie, CHEN Liang, SHEN Changqing. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis[J]. Measurement, 2016, 93: 490-502.
- [20] ZHANG Wei, PENG Gaoliang, LI Chuanhao. Rolling

- element bearings fault intelligent diagnosis based on convolutional neural networks using raw sensing signal[M]. Cham: Springer, 2017.
- [21] ZHANG Wei, PENG Gaoliang, LI Chuanhao, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals[J]. *Sensors*, 2017, 17(2): 425.
- [22] WON Y, GADER P D, COFFIELD P C. Morphological shared-weight networks with applications to automatic target recognition[J]. *IEEE Transactions on Neural Networks*, 1997, 8(5): 1195-203.
- [23] NAIR V, HINTON G. Rectified linear units improve restricted boltzmann machines[C]// *International Conference on Machine Learning (ICML)*, June 21-24, 2010, Haifa, Israel. Madison: Omnipress, 2010: 807-814.
- [24] IOFFE S, SZEGEDY C. Batch normalization : Accelerating deep network training by reducing internal covariate shift[J]. *arXiv Preprint arXiv*, 2015(1): 1502.03167.
- [25] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. *arXiv Preprint arXiv*, 2014(1): 1412.6980.
- [26] JING Luyang, WANG Taiyong, ZHAO Ming, et al. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox[J]. *Sensors*, 2017, 17(2): 414.
- [27] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]// *Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. Piscataway: IEEE, 2016: 2818-2826.
- [28] MAATEN L, HINTON G. Visualizing data using t-SNE[J]. *Journal of Machine Learning Research*, 2008(9): 2579-2605.

作者简介: 胡芑庆, 男, 1967 年出生, 博士, 教授, 博士研究生导师。主要研究方向为状态监控与故障诊断。

E-mail: hnq@nudt.edu.cn

陈徽鹏(通信作者), 男, 1994 年出生, 硕士研究生。主要研究方向为状态监控与故障诊断。

E-mail: chenhuipeng12@nudt.edu.cn