

DOI: 10.7652/xjtuxb201810007

堆叠自编码网络性能优化及其在滚动轴承故障诊断中的应用

张西宁, 向宙, 夏心锐, 李立帆

(西安交通大学机械制造系统工程国家重点实验室, 710049, 西安)

摘要: 为了解决堆叠自编码网络在参数较多时的梯度弥散问题,对网络每层的编码值进行了统计分析,发现大部分分布于激活函数的饱和区,这直接导致了神经元权值梯度的消失。为此,引入了一种标准化策略,将神经元按照样本进行归一化,然后引入两个待学习参数进行缩放和平移,最后通过激活函数输出到下一级神经元。运用带标准化的堆叠自编码网络进行滚动轴承故障诊断,将振动信号的频谱输入到网络中。与普通堆叠自编码网络相比,该标准化策略可有效地使网络编码值均匀分布,如将第一层编码值的熵从 0.88 bit 提高到了 16.29 bit。带标准化的堆叠自编码网络可有效提高网络的抗噪能力和训练速度:在凯斯西储大学滚动轴承数据集上,当人为添加噪声信号的信噪比为 0 dB 时,识别正确率从 16.18% 提高到了 100%;在实验室实测数据集上,不仅训练时间下降了 37.22%,而且识别正确率从 97.93% 提高到了 99.95%。对网络的编码值进行分析以及引入的标准化策略,可为科研技术人员构建堆叠自编码网络时提供参考,也为滚动轴承故障诊断提供了一种策略。

关键词: 故障诊断;堆叠自编码网络;标准化;滚动轴承

中图分类号: TH17 **文献标志码:** A **文章编号:** 0253-987X(2018)10-0049-08

Optimization of Stacking Auto-Encoder with Applications in Bearing Fault Diagnosis

ZHANG Xining, XIANG Zhou, XIA Xinrui, LI Lifan

(State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: To solve the problem of gradient dispersion in stacking auto-encoder (SAE) with large number of parameters, we analyze the distribution of encoding values in each hidden-layer of network. It is found that most of them are distributed in the saturation area of the activation function, which directly leads to a weight gradient loss, thus a normalizing strategy is introduced. The node is normalized according to the sample, then two parameters are introduced to scale and move the encoding values. Then the modified values are passed to the activation function to next layer. The fault diagnosis of rolling bearing is carried out by the normalized SAE, and the spectrum of vibration signal is input into the network. Compared with ordinary SAE, the encoding values of normalized SAE are more well-distributed. For example, the entropy of encoding values in the first level is increased from 0.88 bit to 16.29 bit. The normalized SAE has higher anti-noise ability and faster training rate. When the signal to noise

收稿日期: 2018-04-19。 作者简介: 张西宁(1965—),男,教授,博士生导师。 基金项目: 国家自然科学基金资助项目(51275379);国家自然科学基金创新研究群体资助项目(51421004)。

网络出版时间: 2018-06-15 网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20180615.1626.002.html>

ratio (SNR) is 0 dB, the recognition accuracy is increased from 16.18% to 100% on the rolling bearing data sets of Case Western Reserve University. On laboratory data sets, the training time is decreased by 37.22%, and the recognition accuracy is increased from 97.93% to 99.95%. The introduced normalizing strategy provides a reference for subsequent research on the construction of SAE, and also provides a strategy for fault diagnosis of rolling bearings.

Keywords: fault diagnosis; stacking auto-encoder; normalization; rolling bearing

随着物联网、计算机的发展与普及,各种信息数据呈现出了井喷式的增长,如何从浩瀚的数据流中挖掘出所需要的信息,已经成为了当前急需解决的问题之一。在机电大数据时代,反映装备状态的数据往往具有总量大、形式多、价值密度低等特点,如再依靠诊断专家手工处理信号获取信息,犹如大海捞针,因而传统机械故障诊断“特征提取+模式识别”的模式已然不能满足从机电大数据中自动挖掘信息的需求^[1]。

深度学习作为大数据时代的一颗明星,在学术界和工业界都引起了足够的重视,在很多传统的识别任务上都取得了显著成果。依靠多个隐藏层,深度学习可以很好地实现复杂高维函数的表示,因而具有强大的表征能力,在特征自动提取上具有很多其他算法不可比拟的优势^[2-4]。堆叠自编码网络(SAE)是深度学习中最重要模型之一,由 Hinton 于 2006 年对 Rumelhart 提出的自编码器进行改进而来^[5-6]。SAE 采用逐层无监督学习的方式初始化深度神经网络的权值,然后使用少量标签样本对网络进行微调,逐层贪心的训练方式可有效避免网络陷入局部最优,从而改善网络性能,使深度神经网络的训练成为了可能。2007 年,Benjio 模拟人脑神经元被稀疏激活的特性,在损失函数中加入稀疏项,构成了稀疏 SAE^[7];2008 年,Vincent 向训练数据中添加随机噪声,有效地避免了网络陷入局部最优,构成了降噪 SAE^[8];2013 年,Telmo 研究了不同损失函数下 SAE 的性能,对损失函数的选用和改进做出了指导^[9]。

在机械故障诊断领域,应用 SAE 完成了很多故障识别、分类的任务。2015 年,雷亚国等人利用普通 SAE 和降噪 SAE 实现了对滚动轴承以及行星齿轮箱的故障诊断,并对比了 SAE 和传统模式识别方法之间的识别正确率^[3,10];2016 年,Shao Haidong 等人对 SAE 的损失函数进行了改进,并用人工鱼群算法优化了网络的关键结构参数,有效地提高了对齿轮箱以及列车滚动轴承故障的识别率^[11];2017 年,崔江等人用经过灰色关联度优化后的 SAE 提取

特征,支持向量机分类,实现了航空发电机旋转整流器二极管的故障诊断^[12]。此外,SAE 在变压器^[13]、核电站^[14]、风力发电机组设备^[15]等复杂机电系统的故障诊断领域均有应用,并取得了良好的效果。但是,目前机械故障诊断对 SAE 的研究大多还处于从人工智能领域“引进+应用”的阶段,对网络性能退化的原因分析及改进还有所欠缺。

2015 年,Sergey Ioffe 提出了批标准化(BN)的思想,并成功运用在卷积神经网络(CNN)上,不仅减少了模型的训练时间,而且提高了图像识别的正确率,但是由于主要运用在大型图像 CNN 上,对 BN 改善网络性能的原因分析还有所欠缺^[16]。

本文从神经元激活值分布的角度,分析了自动编码器在参数较多时性能退化的原因,将 BN 中的标准化策略应用到 SAE 上,把神经元按照样本进行归一化,然后引入两个待训练参数进行缩放和平移,最后通过激活函数输出到下一级神经元。在凯斯西储大学滚动轴承数据集以及实验室实测数据集上对比了带标准化的 SAE 和普通 SAE 在抗噪性、训练速度、模型精度等方面的优势。

1 基础理论

1.1 自编码器

自编码器(AE)是由 Rumelhart 在 1986 年提出的一种典型单隐层神经网络^[5]。由于这种网络是在保持输入与输出尽可能一致的情形下实现无监督方式下的隐层特征提取与参数学习,因而被称作自编码器,从输入到隐藏层、隐藏层到输出的过程分别被称为编码和解码。

如图 1 所示,自编码器采用全连接的方式将输入加权求和,加上偏置后通过激活函数,得到隐藏层;隐藏层通过同样的方式得到输出层;然后以输入、输出层的差异最小为目标对权值和偏置进行优化求解。自编码器的数学表述如下

$$Y = \sigma_a(W_a \cdot X + b_a) \quad (1)$$

$$\hat{X} = \sigma_s(W_s \cdot Y + b_s) \quad (2)$$

$$\min_{\theta} L(\theta) = \|\hat{X} - X\|_2^2; \theta = [W_a, b_a; W_s, b_s] \quad (3)$$

式中: $\mathbf{X}=(x_1,x_2,\cdots,x_n)$ 、 $\hat{\mathbf{X}}=(\hat{x}_1,\hat{x}_2,\cdots,\hat{x}_n)$ 、 $\mathbf{Y}=(y_1,y_2,\cdots,y_m)$ 分别为输入、输出和隐藏层神经元的值; $\mathbf{W}_a\in\mathbf{R}^{n\times m}$ 、 $\mathbf{W}_s\in\mathbf{R}^{m\times n}$ 、 $\mathbf{b}_a\in\mathbf{R}^m$ 、 $\mathbf{b}_s\in\mathbf{R}^n$ 为需要优化求解的权值和偏置; $\sigma_a(\cdot)$ 、 $\sigma_s(\cdot)$ 为激活函数。

自编码器只有简单的矩阵乘法 and 矩阵加法运算,激活函数通常为可导的简单函数,因而利用链式法则,容易得到网络的梯度,所以网络的权值和偏置可以采用反向传播算法进行优化求解。

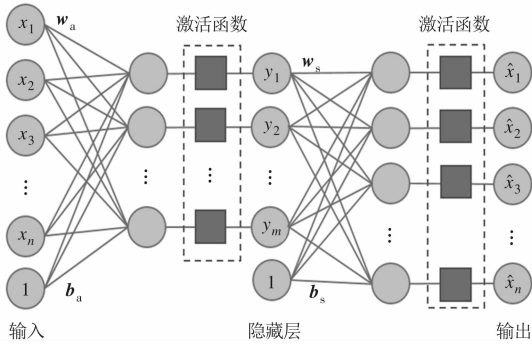


图 1 自编码器结构示意图

1.2 堆叠自编码网络(SAE)

如图 2 所示,将自编码器的编码部分进行堆叠,即依次将上一层自编码器的隐藏层作为下一层自编码器的输入,最后给网络加上分类器,即构成 SAE。网络使用逐层自编码的方式进行预训练,由于这种预训练不需要用到样本标签,因而是一种无监督的训练方式。预训练后,网络只需要用少量带标签的样本进行微调,即可达到较好的性能。由于每一层自编码器为浅层神经网络,因而可以很好地避免网络陷入局部最优^[17]。

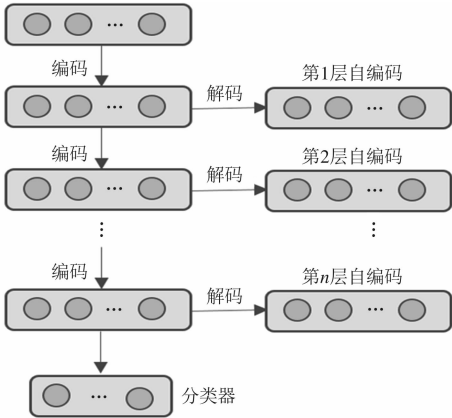


图 2 SAE 结构示意图

自编码器是在尽量保证信息不丢失的前提下,对输入进行编码,以期编码特征可以以足够小的误差重构回输入。这种特征编码可以视为一种特征提取,随着层级加深,编码特征便越加抽象,具有全局

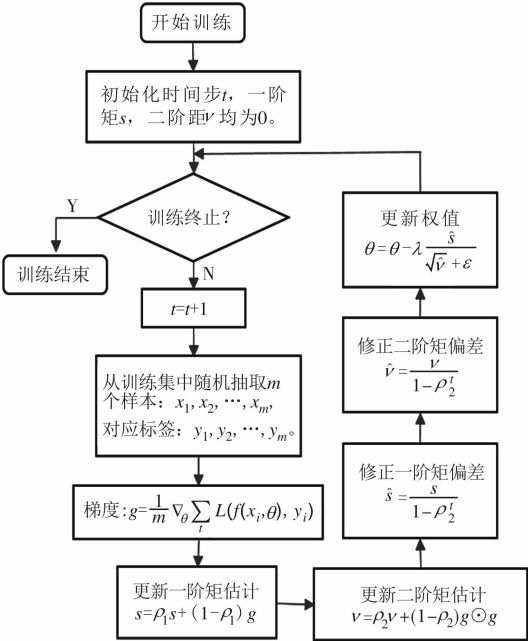
整体特性。

由于 SAE 由自编码器的编码部分堆叠而成,因而整个网络也只有简单的矩阵乘法和矩阵加法运算,激活函数简单可导,所以网络也可以采用梯度相关的算法进行优化求解。

1.3 自适应动量项与置零比例

利用反向传播算法求得梯度后,本文利用了自适应动量项 (Adam) 算法更新需要训练的参数。Adam 是 Kingma 和 Ba 在 2014 年提出的一种学习率自适应的优化算法,具有良好的超参数鲁棒性。Adam 包括偏置矫正环节,可以修正从原点的初始化一阶矩(动量项)以及(非中心的)二阶矩估计,算法流程如图 3 所示^[18]。

为了减少网络训练过程中数据的过拟合现象,本文在模型训练时使用了置零比例 (Dropout) 训练技巧。Dropout 在每一轮的训练时,随机将隐层神经元的输出按照一定的比例归 0,使得部分神经元不参与正向传播,这部分神经元在反向传播阶段也没有任何贡献,但是其权值却保存以供下一轮训练和测试阶段使用^[19]。



λ 为学习率; ρ_1 和 ρ_2 为矩估计衰减率,推荐 0.9 和 0.999;
 ϵ 为数值稳定常数,推荐 10^{-8}

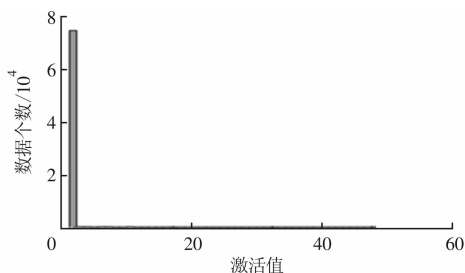
图 3 Adam 算法流程图

2 网络性能分析与优化

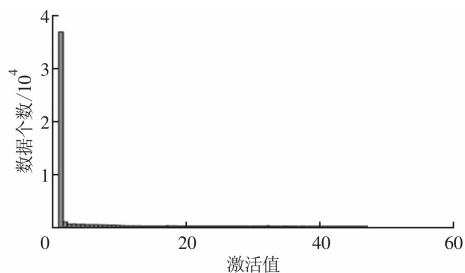
2.1 SAE 性能退化分析

为了分析 SAE 性能退化的原因,在凯斯西储大学轴承数据集上构建了一个具有 3 层编码层的

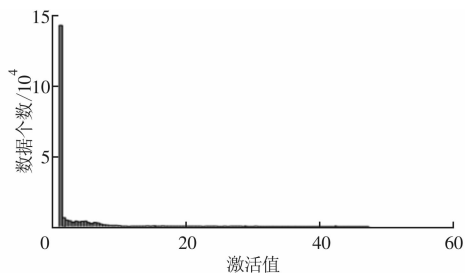
SAE。该 SAE 网络拥有 200 个训练样本,各层神经元的个数分别为 600、400、200、100、10,所以中间 3 个隐藏层分别拥有 400×200 (8 万)、 200×200 (4 万)、 100×200 (2 万)个激活值,对这些激活值的分布进行统计如图 4 所示。从图中可以看出,各隐藏层神经元的激活值大多集中在 -1 左右,结合所用激活函数 Exponential Linear Unit(ELU)的图像(见图 5)可知,网络大部分神经元已经处于了饱和阶段。



(a) 第 1 隐藏层



(b) 第 2 隐藏层



(c) 第 3 隐藏层

图 4 各隐藏层神经元激活值分布情况

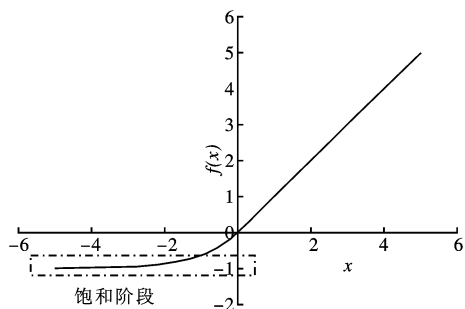


图 5 ELU 激活函数图像

了学习能力,严重限制了网络更好地拟合需要学习的拓扑映射关系。

2.2 带标准化 SAE

为了解决神经元处于饱和阶段学习能力丧失的问题,引入了一种标准化策略。将通过激活函数前的神经元按照样本分别进行标准化,然后引入两个参数进行缩放和平移,最后再通过激活函数输出到下一级神经元,如图 6 所示。

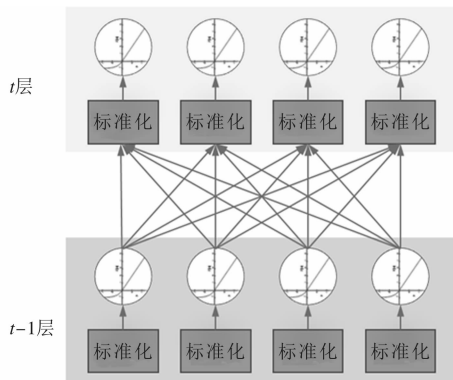


图 6 带标准化 SAE

标准化的具体流程为每个神经元都减去均值后除以标准差化为标准分布,为了抵消这种标准化对网络学习效果的弱化,引入两个需要学习的参数对其进行了缩放和平移,计算公式如下

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (4)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (5)$$

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (6)$$

$$y_i = \gamma \hat{x}_i + \beta \quad (7)$$

式中: x_i 、 y_i 为标准化前、后神经元激活值; m 为训练样本数量; γ 和 β 为两个待学习的参数; ϵ 为一个很小的稳定常数。

为了使反向传播算法能够正常进行,对梯度在该步骤的传播情况进行了分析,计算公式如下

$$\frac{\partial L}{\partial x_i} = \frac{\partial L}{\partial \hat{x}_i} \frac{1}{\sqrt{\sigma_B^2 + \epsilon}} + \frac{\partial L}{\partial \sigma_B^2} \frac{2(x_i - \mu_B)}{m} + \frac{\partial L}{\partial \mu_B} \frac{1}{m} \quad (8)$$

$$\frac{\partial L}{\partial \gamma} = \sum_{i=1}^m \frac{\partial L}{\partial y_i} \hat{x}_i \quad (9)$$

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^m \frac{\partial L}{\partial y_i} \quad (10)$$

其中

$$\frac{\partial L}{\partial \hat{x}_i} = \frac{\partial L}{\partial y_i} \gamma \quad (11)$$

$$\frac{\partial L}{\partial \sigma_B^2} = \sum_{i=1}^m \frac{\partial L}{\partial \hat{x}_i} (x_i - \mu_B) \frac{-1}{2} (\sigma_B^2 + \epsilon)^{-3/2} \quad (12)$$

$$\frac{\partial L}{\partial \mu_B} = \sum_{i=1}^m \frac{\partial L}{\partial \hat{x}_i} \frac{1}{\sqrt{\sigma_B^2 + \epsilon}} \quad (13)$$

训练阶段,样本通常是成批输入到 SAE 中的, μ_B 、 σ_B^2 由式(4)、式(5)计算得到。但是,到了测试阶段,往往只有单个样本输入到网络,此时再利用式(4)、式(5)计算 μ_B 、 σ_B^2 ,网络的测试精度就会大打折扣。针对此,采用了如图 7 所示的策略,将 SAE 训练时最后一次迭代的 μ_B 、 σ_B^2 带入式(6),对测试样本进行标准化。

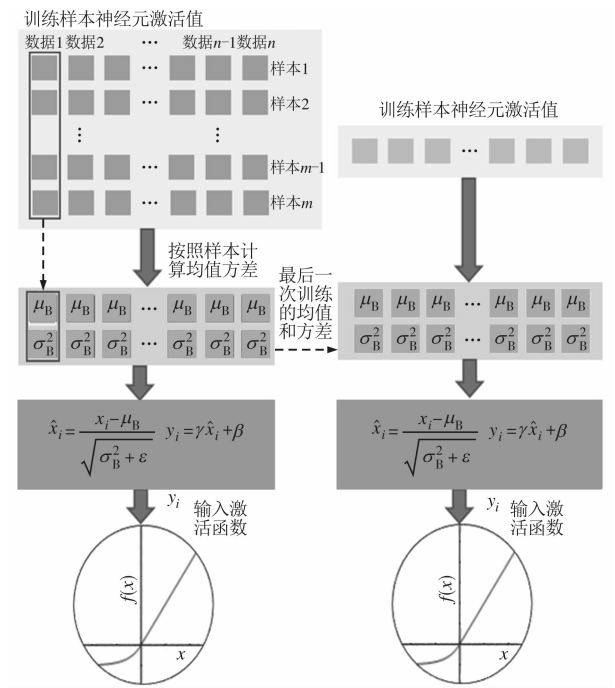


图 7 训练和测试样本的标准化策略

对 2.1 节中的 SAE 进行改进,在每次通过激活函数前都加入标准化层,按照图 4 同样的方式对隐藏层的激活值进行了统计,如图 8 所示。利用下式计算图 4、图 8 中激活值的分布熵

$$H = - \sum_{i=1}^q P(x_i) \log P(x_i) \quad (14)$$

式中: $P(x_i)$ 为激活值 x_i 出现的概率; q 为不重复激活值的数量。计算结果如表 1 所示。

表 1 两种网络隐藏层激活值的分布熵

网络	H/bit		
	第 1 隐藏层	第 2 隐藏层	第 3 隐藏层
SAE	0.88	6.18	11.33
标准化 SAE	16.29	15.29	14.29

从分析可以看出,加入标准化后,神经元激活值

的分布明显趋于均匀,表明这些激活值处于 ELU 激活函数的不同阶段,不像图 4 那样大部分处于饱和区,这种特性有利于缓解网络梯度消失的问题。所以,在网络每次通过激活函数前加入标准化策略,可有效地避免网络由于过早陷入激活函数饱和区而导致的学习能力退化问题。

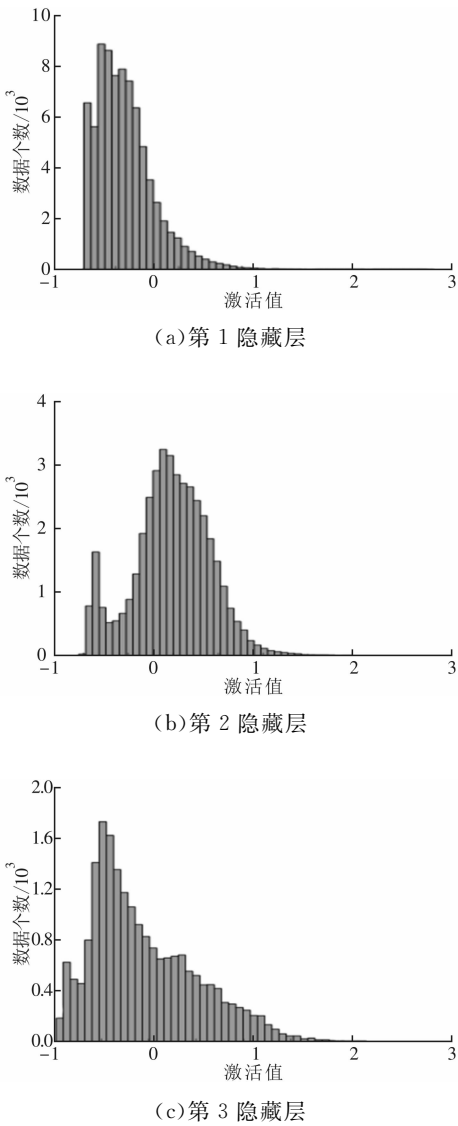


图 8 标准化后各隐藏层神经元激活值分布情况

3 基于标准化 SAE 的轴承故障诊断

3.1 凯斯西储大学轴承数据集

凯斯西储大学(CWRU)滚动轴承数据集^[20]数据质量高、故障特征明显,因而是学术界常用的滚动轴承故障诊断标准数据集。本文利该数据集对比标准化前后 SAE 对故障识别的效果。诊断对象为深沟球轴承 SKF6205,分别在其外圈滚道、内圈滚道、滚动体上通过电火花加工模拟不同故障,故障直径

分别为 0.18、0.36、0.54 mm 3 类不同程度,加上无损伤轴承,实验共计需要识别 10 类不同状态的轴承。数据集为用加速度传感器采集的轴承安装座上的振动,采集频率为 12 kHz,采样长度为 1 200 个点。每类状态有 100 个样本,共计 1 000 个,普通 SAE 和标准化后的 SAE 均选择 200 个样本训练,其余 800 个为样本测试。为了克服各样本采集起始时间不一致的问题,本文对数据进行了快速傅里叶变换(FFT),将信号的频谱输入到网络中进行训练和测试。

本文在 Python3.5.2 环境下利用 Google 公司的深度学习框架 TensorFlow,搭建了普通 SAE 和带标准化 SAE 的数据流图;将训练数据频谱输入到图中,启动数据流图,利用权值更新模块自动更新权值;然后网络训练完毕后,将测试数据频谱输入数据流图进行诊断。网络结构及训练超参数如表 2 所示。

普通 SAE 和标准化 SAE 都分别重复实验 20 次,计算 20 次实验的平均测试正确率,结果显示,普通 SAE 和标准化 SAE 的识别正确率分别达到了 99.99%和 100%。为了更加明显地区分两种网络在识别正确率上的差异,按照下式对原始信号添加了不同信噪比的高斯噪声

$$\sigma_{\text{SNR}}=-10\log\sum_{i=1}^nx_i^2/\sum_{i=1}^nz_i^2\tag{15}$$

式中： $\sum_{i=1}^nx_i^2$ 为原始信号能量； $\sum_{i=1}^nz_i^2$ 为所加噪声能量。同样,对每组实验重复 20 次,平均识别正确率如图 9 所示。

表 2 凯斯西储大学轴承数据网络结构及训练超参数

编号	普通 SAE		标准化 SAE		学习率	迭代步数	置零比例
	操作	数据个数	操作	数据个数			
1	输入	1 200	输入	1 200			
2	FFT	600	FFT	600			
3	自编码	400	自编码	400	0.01	100	0.5
4			标准化	400			
5	自编码	200	自编码	200	0.01	100	0.5
6			标准化	200			
7	自编码	100	自编码	100	0.01	100	0.5
8			标准化	100			
9	Softmax	10	Softmax	10	0.001	100	0.2
10	微调	10	微调	10			

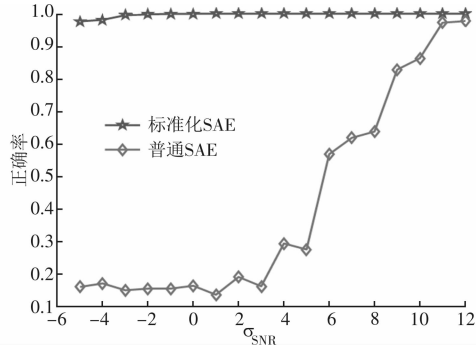


图 9 不同信噪比的测试结果

从图 9 可以看出:在信噪比较大的情形下,普通 SAE 与标准化 SAE 对故障的识别能力相当,都接近 100%;但是在信噪比较小时,标准化 SAE 对故障的识别能力明显强于普通 SAE。例如,当信噪比为 0 时,原信号和添加的噪声信号的能量相等,标准化 SAE 识别正确率为 100%,而普通 SAE 仅为 16.18%。这说明标准化过后的 SAE 极大地提高了在噪声环境下网络对故障的识别能力。

当信噪比为 0 时,第一次实验第一层编码的误差以及所有神经元梯度总和随训练步数的变化如图 10 所示。从中可以看出,加入标准化后的 SAE 可使网络误差收敛到更小的区间,避免神经元权值梯度过早消失。

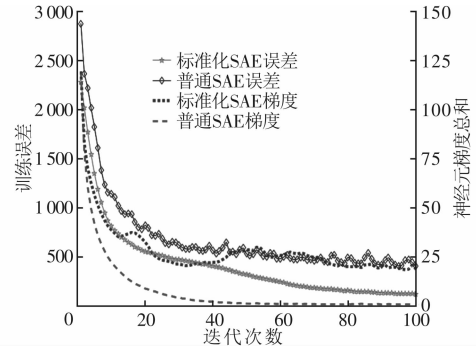


图 10 误差与神经元梯度总和的变化对比

3.2 实验室实测数据集

在实验室实测数据上对标准化前后 SAE 的性能做了进一步的测试。如图 11 所示,实验将滚动轴承安装在安装座内,通过预紧装置对滚动轴承进行轴向预紧,轴承安装座上的绳索通过滑轮与重物相连,进而模拟滚动轴承的载荷。加速度传感器 IMI601A11 采集安装座上的振动信号,然后传送到数采卡 UA300,进行采集后存储到计算机上。实验所用的 6308 滚动轴承共有 5 种状态,分别为正常、外圈剥落、内圈剥落、滚动体剥落和保持架故障。每种状态都分别采集 400 个样本,共计 2 000 个,采样

频率为 10 kHz,采样长度为 2 400 个点。

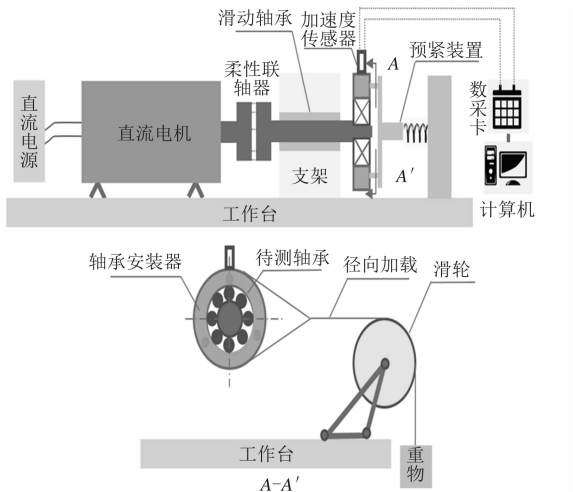


图 11 滚动轴承实验台示意图

数据样本被分为训练集、验证集、测试集 3 部分,每部分样本个数分别为 1 000、500、500。验证集的目的主要是为了防止网络过拟合,如图 12 所示,当训练集上训练误差稳定后,选择在验证集上验证正确率连续 5 次下降,或者训练误差连续 5 次上升,作为网络训练终止的标志。

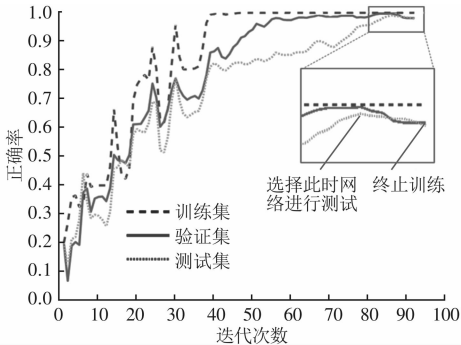


图 12 网络训练终止条件

在 TensorFlow 框架下搭建如表 3 所示的普通 SAE 和标准化 SAE 的数据流图,将训练集输入网络进行训练;每一步训练完成后,在验证集上进行验证,以判断网络是否终止;网络训练终止后将测试集输入进行识别。

对普通 SAE 和标准化 SAE 均独立重复训练测试 20 次,所用计算机配置为 Core i5-6500 CPU @ 3.20 GHz,内存为 16 GB。20 次实验的测试正确率如图 13 所示,各层编码以及微调环节的训练时间如图 14 所示。

由图 13 可以看出,标准化 SAE 的测试正确率均在 99% 以上,而普通 SAE 的正确率在 98% 上下波动,统计 20 次实验的平均正确率分别为 99.95%、

表 3 实测数据网络结构及训练超参数

编号	普通 SAE		标准化 SAE		学习率	置零比例
	操作	数据个数	操作	数据个数		
1	输入	2 400	输入	2 400		
2	FFT	1 200	FFT	1 200		
3	自编码	600	自编码	600	0.01	0.5
4			标准化	600		
5	自编码	400	自编码	400	0.01	0.5
6			标准化	400		
7	自编码	200	自编码	200	0.01	0.5
8			标准化	200		
9	自编码	100	自编码	100	0.01	0.5
10			标准化	100		
11	Softmax	5	Softmax	5	0.001	0.2
12	微调	5	微调	5		

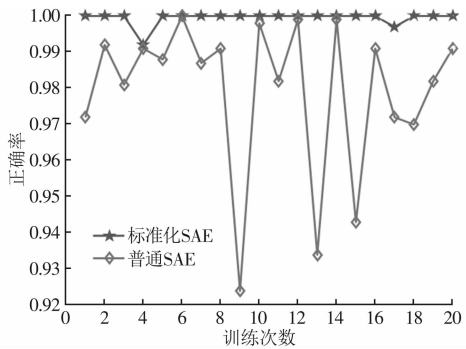


图 13 20 次实验测试结果对比

97.93%。由图 14 可以看出,标准化后 SAE 各次实验的训练时间都比普通 SAE 低。对这 20 次训练时间进行统计,标准化 SAE 的平均训练时间为 88.66 s,相比普通 SAE 平均训练时间 141.22 s,下降约 37.22%。

4 结 论

为了提高 SAE 的性能,本文从网络每层编码值分布的角度对 SAE 进行了分析。引入了一种标准化策略,在神经元每次通过激活函数前进行标准化,将神经元按照样本进行归一化,然后引入两个待学习参数进行缩放和平移,最后通过激活函数输出到下一级神经元。为了验证带标准化 SAE 的性能,在凯斯西储大学轴承数据集以及实验室实测数据集上进行了实验,得出了如下结论。

(1)对网络每层编码值的统计分析表明:普通 SAE 各层编码值大部分分布于激活函数的饱和区,

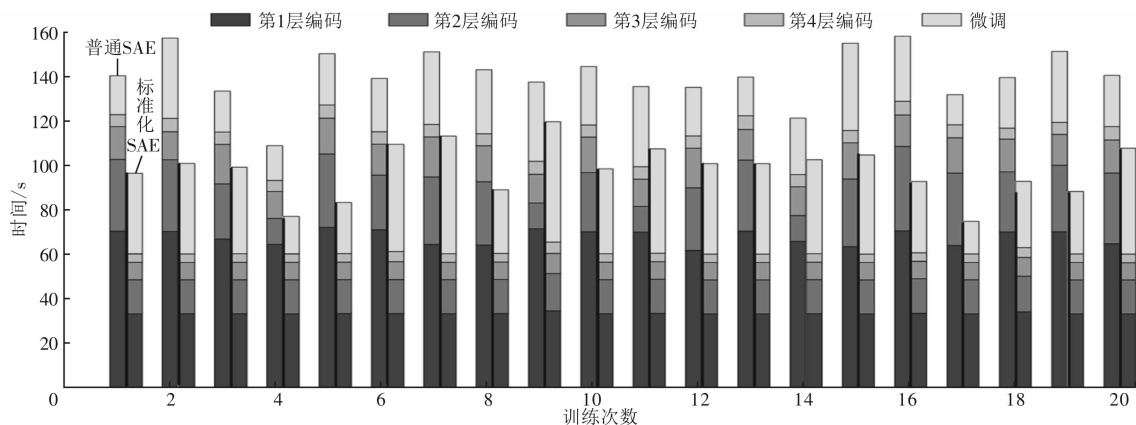


图14 各层编码及微调训练时间对比

这直接导致了神经元权值梯度的消失,使网络学习能力退化。

(2)引入的标准化策略可有效地使网络编码值均匀分布,例如在凯斯西储大学轴承数据集上将第一层编码值的熵从 0.88 bit 提高到了 16.29 bit。神经元激活值的均匀分布表明,这些激活值处于激活函数的不同阶段,不是大部分处于饱和区,这种特性有利于减少 SAE 权值梯度的消失,避免 SAE 由于神经元激活值过早陷入激活函数饱和区而导致的学习能力退化问题。

(3)利用普通 SAE 和标准化 SAE 进行滚动轴承故障诊断,结果表明,加入标准化步骤后的 SAE 不仅可以提高网络在噪声环境下的识别能力,而且能够有效地减少网络的训练时间。例如在凯斯西储大学数据集上,当信噪比为 0 dB 时,将识别正确率从 16.18% 提高到了 100%;在实验室实测数据集上,不仅训练时间下降了 37.22%,而且识别正确率从 97.93% 提高到了 99.95%。

本文对 SAE 的分析为科研技术人员寻找 SAE 性能退化原因提供了一种参考,引入的标准化策略为提高 SAE 性能提供了一种思路。

参考文献:

- [1] 雷亚国,贾峰,孔德同,等. 大数据下机械智能故障诊断的机遇与挑战 [J]. 机械工程学报, 2018, 54(5): 94-104.
LEI Yaguo, JIA Feng, KONG Detong, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era [J]. Chinese Journal of Mechanical Engineering, 2018, 54(5): 94-104.
- [2] 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 and Decision, 2017(8): 1345-1358.
- [3] 雷亚国,贾峰,周昕,等. 基于深度学习理论的机械装备大数据健康监测方法 [J]. 机械工程学报, 2015, 51(21): 49-56.
LEI Yaguo, JIA Feng, ZHOU Xin, et al. A deep learning-based method for machinery health monitoring with big data [J]. Chinese Journal of Mechanical Engineering, 2015, 51(21): 49-56.
- [4] ZHANG Qingchen, YANG L T, CHEN Zhikui. Deep computation model for unsupervised feature learning on big data [J]. IEEE Transactions on Services Computing, 2016, 9(1): 161-171.
- [5] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors [J]. Nature, 1988, 323(6088): 399-421.
- [6] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504-507.
- [7] SCHÖLKOPF B, PLATT J, HOFMANN T. Greedy layer-wise training of deep networks [C]// International Conference on Neural Information Processing Systems. Cambridge, MA, USA: MIT Press, 2006: 153-160.
- [8] VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoders [Z]. New York, USA: ACM, 2008: 1096-1103.
- [9] AMARAL T, KANDASWAMY C, SANTOS J M. Using different cost functions to train stacked autoencoders [C]// Mexican International Conference on Artificial Intelligence. Piscataway, NJ, USA: IEEE, 2014: 114-120.

(下转第 87 页)