

进程择优法及在心音深度信任网络中的应用

成谢锋^{1,2)} 杨 贺¹⁾ 马 勇³⁾ 张学军^{1,2)} 张少白¹⁾ 王 悦¹⁾

¹⁾(南京邮电大学电子科学与工程学院 南京 210003)

²⁾(南京邮电大学射频集成与微组装技术国家地方联合工程实验室 南京 210003)

³⁾(南京理工大学计算机学院 南京 210094)

摘 要 深度学习算法因其在自然环境下对大数据处理的优良特性已成为图像、语音识别方面的主流算法。为解决深度学习网络结构选择困难的问题,文中深入探究了深度学习网络的结构特性,提出了一种进程择优法来帮助深度学习网络结构的选择,可方便、快速地给出深度学习网络的优选范围。经实验验证,此方法在多种数据库下都有良好效果,方法具有一定的普适性。而心音作为一种生理信号,反映了人体心脏的跳动情况,与人体心脏的健康息息相关,在心音分类识别、健康鉴定中得到广泛的应用。文中首先使用进程择优法来优选、构建出一种心音深度学习网络,再以心音深度学习网络为核心,加入 BP 神经网络作为分类器,设计出了一种心音深度信任网络。该网络相比同类其它层次结构的深度信任网络拥有更低的误识别率,平均误识别率在 10% 左右。特别是将原系统优化为融合心音能量特征输入的心音深度信任网,其平均误识别率可下降到 3%。文中的研究对于提高心音识别算法在自然环境下处理数据的能力具有积极的意义。

关键词 深度学习算法;进程择优法;心音;心音深度学习网络;心音深度信任网络

中图法分类号 TP18 DOI号 10.11897/SP.J.1016.2018.00208

A Preferred Method of Process and Application in Heart Sounds Deep Belief Networks

CHENG Xie-Feng^{1,2)} YANG He¹⁾ MA Yong³⁾ ZHANG Xue-Jun^{1,2)}

ZHANG Shao-Bai¹⁾ WANG Yue¹⁾

¹⁾(College of Electronic Science and Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003)

²⁾(Nation-Local Joint Project Engineering Lab of RF Integration & Micropackage,

Nanjing University of Posts and Telecommunications, Nanjing 210003)

³⁾(School of Computer Science, Nanjing University of Science and Technology, Nanjing 210094)

Abstract Deep learning algorithm has become a main method in the image and voice recognition field because of its excellent characteristics under the process of big data in a natural environment. For the purpose of solving the problems that it is difficult to select an appropriate structure of the deep learning network. In this paper, we get an in-depth explore at the structure characteristics of deep learning network and propose a method the preferred method of process to help us to select the structure of the deep learning network. This method uses the variation of the deep reconstruction error of single layer in the deep learning network as a criterion. By introducing the pare to principle in economics into our network, the threshold of the deep reconstruction error of single layer can be determined conveniently and quickly, and deep learning network optimization

收稿日期:2016-05-14;在线出版日期:2017-02-24。本课题得到国家自然科学基金(61271334,61373065)资助。成谢锋,男,1956年生,教授,中国计算机学会(CCF)会员,主要研究领域为智能信息处理、智能仪器。E-mail: chengxf@njupt.edu.cn。杨 贺,女,1990年生,硕士,主要研究方向为智能信息处理、深度学习算法。马 勇(通信作者),男,1983年生,博士,讲师,主要研究方向为计算机应用技术、机器人。E-mail: mayong@163.com。张学军,男,1965年生,博士,教授,主要研究领域为无线通信网络、脑电(EEG)和脑磁(MEG)信号处理。张少白,男,1954年生,博士,教授,主要研究领域为计算机技术、脑机接口。王 悦,女,1992年生,硕士研究生,主要研究方向为智能信息处理。

is given. The preferred method of process can effectively solve the uncertainty of the layers' number and the nodes' number per layer when set up the network. It is validated by experiment that this method has large amount of good effectiveness under a variety of heart sound databases and this method has a certain universal applicability to be used in other studies. As a kind of valuable biological signals, heart sounds can reflect the beating of the heart and are closely related to the health of the heart or the human's body. Heart sounds has been widely used in health monitoring and identity recognition and it has attracted the attention of researchers both domestic and abroad. In this paper, we use the preferred method of process to optimize and build a heart sound deep learning network firstly. Then we design a heart sounds deep belief network by combining the classifier BP neural network with the heart sound deep learning network, the core of it. Compared to other belief network in the same category, this heart sounds deep belief network has a lower error recognition rate and the average error recognition rate is only about 10%. The deep belief network has little limit for input data. So referencing to the previous method of heart sound recognition, an effective method to greatly improve the recognition rate is to extract the feature vector of input data, and then used it as input of the deep belief network. In this paper the average error recognition rate can be reduced to 3% when the network is optimized for a new deep belief network. The new deep belief network is based on the original system data which optimized by the orthogonal wavelet transform and then extract heart sound energy characteristics. The research in this paper has a positive significance to improve the ability to process data of heart sound recognition algorithm under normal natural circumstances. Building better depth of heart sound recognition system, is the next step research work.

Keywords deep learning algorithm; the preferred method of process; heart sounds; heart sounds deep learning networks; deep belief networks

1 引言

在大数据分析中, 如何对数据进行有效表达尤其重要. 而在以往的研究中, 模型多为简单、浅层的网络, 模型的表达能力有限, 在复杂的自然环境中处理大数据时很难获得优良的输出结果. 深度学习是模拟人脑的功能, 利用多层次的架构, 获得对象在不同层次上的表达, 以解决一些浅层结构无法很好解决的复杂、抽象的问题^[1]. 深度学习算法无需人工设计特征, 由计算机自主地对数据进行学习, 可有效地对复杂数据进行降维, 极大地提高了系统的自组织性和系统的效率. 但其强大的自主性需要大量的数据进行学习才能获得良好的效果, 而复杂的大数据条件恰好满足了深度学习算法对于数据量的要求^[1-3].

最早的深度学习系统产生于 2006 年, 多伦多大学的 Hinton 等人成功解决了多层深度模型优化困难的问题, 从而掀起了深度学习的浪潮^[4-6]. 在深度

学习算法的相关应用中, 最引人瞩目的是将深度学习算法与前馈神经网络结合. 结合之后的系统不仅在模式识别中获得了良好的效果, 在图片的分割、实物的探测上也取得了不俗的效果. 深度信任网络、卷积神经网络以及编码器是这其中最常用的三种深度学习模型^[5-7].

深度学习网络中有很多可调参数, 这一点保证了深度学习网络拥有很高的灵活性, 可以按照个人要求进行定制. 但是节点数、层数、学习率、训练时每组数据的个数等等参数都没有现行的标准进行参考, 以往的研究中大多按照经验数据进行设定, 这也限制了深度学习网络的性能.

心音是人体重要的生物信号, 现已广泛应用到心音识别、人体健康情况评价中^[8]. 心音反映了人体心脏的跳动情况, 与心脏的健康情况息息相关. 心音采集方便, 通过对采集出的心音进行分析可方便检测出心脏的健康情况. 心音信号一般具有以下特点: (1) 心音信号具有周期性、非平稳性与随机性; (2) 心音信号的共性比较明显、个性比较微弱; (3) 心音主

要信息一般集中在 0~600 Hz 频率段内. 由于心音信号从总体上来讲共性比较明显, 在分析时难免会遇到很多困难, 导致分析结果不尽人意. 于是能否寻找到一个适合于对心音信号进行识别的优良算法成为能否将心音真正应用到为人类造福这项事业的核心问题.

对心音进行分类识别的方法较多, 主要有以 FFT 方法为代表的功率谱特征表征、以小波变换法为代表的时频特征表征及采用欧氏距离作为心音识别分类器的方法^[8]. Patangay 等人^[9]发表的论文总结出了也许是自动听诊最成功的方法. 他结合了以前建立的一些方法, 选择性提取特征, 为每个类型的心音创建了一个标志性数据集, 使用动态规划选择主要的特征, 并发展和学习 (GAL) 神经网络来优化以关键要素为基础的神经结构. 这种方法对七种不同的心脏病变有非常高的预测性能. 但是, 这种方法并没有预测心音含有无关杂音的情况, 准确度、适应性还需提高. 因此在分析小波构造理论的基础上, Varghees 等人在 2014 年提出了针对心音的声学边界自动分析的方法^[10], 参考文献^[11]构造了一种专门用于心音信号处理的小波基——心音小波. 实验结果表明, 相比常用的 db、bior 系列小波, 运用心音小波对心音信号进行处理, 能够获得更好的去噪效果、更精确的心音分类信息以及更小的重构误差率, 在表征心音个体特征的细节方面具有积极的意义. 但是, 在这些心音分类识别中, 所使用的算法多面向实验室环境, 对采集到的心音信号需要进行去噪、分段、特征提取等预处理, 在自然环境下的应用效果往往并不良好.

目前在对心音识别算法进行研究时, 研究者的重点多集中在特征提取方面. 而且同一个人、不同采集设备获取的心音会有一些的区别. 这意味着即使一种算法已经在数据库中得出十分优秀的识别效果, 也很难保证其在广泛应用时的可靠性. 本文研究的方法面向大数据、自然环境下的心音, 对心音信号的获取要求极低, 相比传统方法具有更好的实际应用性.

为解决深度学习网络结构难以确定的问题, 本文深入探究了深度学习网络的结构特性, 提出了一种进程择优法来帮助深度学习网络结构的选择, 可方便、快速地给出深度学习网络的优选范围.

为了提高心音识别算法在自然环境下处理大数据的能力, 本文使用进程择优法辅助构建出了一种心音深度学习网络, 经验证使用心音深度学习网络构建出的心音深度信任网络相比其它层次结构的深

度信任网络拥有最低的误识别率, 平均误识别率在 10% 左右. 特别是对心音进行简单的能量特征提取, 将原系统优化为融合心音能量特征输入的心音深度信任网络, 实验表明该系统的误识别率显著下降, 误识别率仅为 3% 左右, 可达到传统识别方法下所能达到的效果.

2 深度学习网络与深度信任网络

2.1 深度学习网络

深度学习系统可以由多层的限制玻耳兹曼机 (Restricted Boltzmann Machine, RBM) 叠加而成. 任意一个 RBM 的概率模型为^[12]

$$P(x, h) = \frac{e^{-E(x, h)}}{Z} \quad (1)$$

$$E(x, h) = -b'x - c'h - h'wx \quad (2)$$

$$Z(x, h) = \sum_{x, h} e^{-E(x, h)} \quad (3)$$

其中 x 为较低的可见层, h 为较高的隐层, $'$ 表示转置. 深度学习即是训练数据由底层的 RBM 训练并输出, 输出数据作为高一层的输入逐层传递. 可见层和隐层之间的连接权重用 w 表示. E 为 RBM 的能量函数, Z 为 RBM 的归一化因子. b, c 分别是可见层和隐层的偏置.

对于一个 RBM 网络, 若各隐层的状态已确定, 则各显层的条件概率即可确定. 有

$$P(h|x) = \frac{\exp(b'x + c'h + h'wx)}{\sum \exp(b'x + c'h + h'wx)} \quad (4)$$

其中 $'$ 表示转置. 上式可化简为

$$P(h|x) = \prod_i P(h_i|x) \quad (5)$$

由式(5)可以看出, 在已知输入数据层的情况下, 所有的隐藏层节点之间是条件独立的, 同理, 在已知隐藏层的情况下, 所有的可视节点都是条件独立的. 由于神经元是二进制的, 所以当神经元激活时 (值为 1) 其概率分布为

$$P(h_i=1|x) = \frac{\exp(c_i + w_i x)}{1 + \exp(c_i + w_i x)} \quad (6)$$

或者表示为

$$P(h_i=1|x) = \text{sigm}(c_i + w_i x) \quad (7)$$

同理, 可见层的概率分布为

$$P(x_i=1|h) = \text{sigm}(b_i + w_i h) \quad (8)$$

为了训练这个模型, 需使此模型转化的数据与测试数据尽可能相似. 从数学角度上讲, 我们希望最大化训练数据的对数概率, 最小化训练数据的负对数概率. 对数概率比较容易计算, 而负对数概率很难

计算. 为简化训练步骤, 避免计算负相, 我们将对模型进行抽样.

在进行抽样时我们需要给定可见层状态, 更新隐层, 将隐层得到的数据返还给可见层从而更新可见层, 依次循环. 即

$$\begin{cases} h^{(0)} = P(h | x^{(0)}) \\ x^{(1)} = P(x | h^{(0)}) \\ h^{(1)} = P(h | x^{(1)}) \\ x^{(2)} = P(x | h^{(1)}) \\ \vdots \\ x^{(n)} = P(x | h^{(n-1)}) \end{cases} \quad (9)$$

在每次迭代中, 整个层得以更新. 而为了得到恰当的抽样, 初始化应该随机, 抽样次数也应该越多越好. 为了简化, 我们将训练样本作为初始化数据, 迭代一步之后, 将更新的数据作为负样本. 这就是对比分歧算法^[12-13].

但对比分歧算法也有一定的缺陷, 即由于深度学习在训练过程中使数据收敛的方法是层与层数据之间进行对比, 从而对网络权重进行调整, 而非普通训练方法将训练后的数据与原数据进行对比, 所以算法容易导致训练误差的逐层增大. 所以深度学习系统存在一个恰当的层数, 并非层数越多越好. 而每一层 RBM 的节点数意味着其对数据的降维程度, 若网络收敛速度过快则意味着每一层 RBM 所提取的信息量过少.

深度学习网络具有如下主要特点:

(1) 深度学习网络是由多个 RBM 叠加而成的, 每一个 RBM 的隐层作为更高层 RBM 的可见层使用, 换句话说, 每一个 RBM 的输出为其更高层 RBM 的输入. 任一个 RBM 输出的数据也仅与输入有关, 其权值的调整情况也仅与输入有关, 与此 RBM 之后的状态、结构均无关.

(2) 深度学习网络的训练方式是无监督的. 在进行深度学习网络的训练过程中需要的仅是一个无标签的数据库, 对数据的要求很低. 但深度学习网络仅仅是一个用于数据学习的网络, 可认为是个对数据进行自适应特征提取的网络, 并不能进行数据识别分类.

(3) 深度学习网络可对数据进行有效地降维. 理论上, 经过深度学习网络学习后的数据会比原数据更具有组织性、特征更明显. 深度学习网络中每层的节点数决定了此层输出数据的维度. 所以一个合适的节点数就可以将数据进行有效地降维, 降低数据的运算代价.

(4) 深度学习网络中底层的误差最大且训练速度最慢. 由于使用随机初始化的方法, 最初的网络权

值随机分布, 初始的训练误差会很大且训练速度也会比较慢. 而由于使用了反向传播, 在计算梯度即误差导数时, 随着网络一层层的增加, 梯度的幅值会逐渐减小, 整体上权值变化的导数会越来越小即趋势会越来越缓.

(5) 深度学习网络的训练效率较高. 由于使用 CD(Contrastive Divergence)算法, 深度学习网络在进行训练时, 往往迭代一次就可以获得一套比较好的网络权值, 不需要很多次迭代, 网络训练的效率较高.

(6) 深度学习网络可以与任意一种分类器结合. 目前效果较好的是与 BP(Back Propagation)神经网络结合, 组成深度信任网络. BP 神经网络采用的是有监督的训练方式, 可以帮助深度学习网络将学习到的数据更好地拟合到所期望的输出.

(7) 深度学习网络中很多参数没有一个现行标准可以进行参考. 深度学习网络中有很多可调参数, 这一点保证了深度学习网络拥有很高的灵活性, 可以按照个人要求进行定制. 但是节点数、层数、学习率、训练时每组数据的个数等等参数并没有现行的标准可以参考, 以往的研究中大多按照经验数据进行设定, 这也限制了深度学习网络的性能.

2.2 深度信任网络

在利用深度学习网络进行分类时, 最简便的模型即为深度信任网络. 深度信任网络是将深度学习算法与 BP 算法结合, 使用 BP 神经网络作为深度学习网络的最后一层, 使等级最高的隐层作为分类器的一部分, 并将整个网络作为一个反向神经网络进行训练, 是无监督学习与有监督学习的结合^[12-16]. 整个网络将深度学习网络通过学习训练数据所得到的网络权值赋予 BP 神经网络的隐层, 使得 BP 神经网络可以得到深度学习网络学习到的先验知识. 这种方法可以使得 BP 神经网络训练成功率增加, 收敛速度提高. 图 1 为深度信任网络结构示意图.

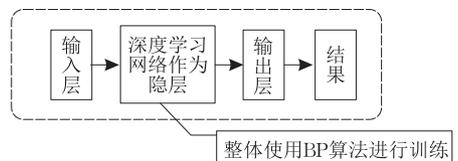


图 1 深度信任网络结构示意图

深度信任网络的整体训练方法如下:

步骤 1. 对一个多层的深度学习网络进行初始化, 使用 CD 算法对训练数据进行训练, 并得到网络的所有权值;

步骤 2. 构建一个隐层层数、节点数与深度学习网络一致的人工神经网络, 将深度学习网络学习后

的权值赋予人工神经网络的各隐层；

步骤 3. 使用 BP 算法对神经网络进行微调, 最终获得一个可用的深度信任网络.

深度信任网络拥有 BP 神经网络的自组织性, 对输入数据没有任何要求. 由于增加了隐层层数, 理论上可更好地学习到数据深层次的特征, 提高系统的整体性能. 另外深度信任网络克服了 BP 神经网络收敛速度慢的问题, 使得整个网络效率得到大大提高. 但深度信任网络仍未克服 BP 神经网络容易陷入局部最优解的问题, 而且由于隐层层数的增加和训练时使用层与层数据进行比较的问题, 使得误差层层传递. 所以隐层层数以及隐层节点数的确定问题成为了改善深度信任网络性能的难题.

3 进程择优法和深度学习网络的快速设计方法

3.1 进程择优法的理论基础

在深度信任网络中, 确定网络的隐层层数以及每层的节点数是进行系统架构的大前提. 在以往的研究中, 往往采用经验值来确定网络的隐层层数以及每层的节点数. 本文特提出一种在进程中进行网络架构择优的方法——进程择优法. 通过本方法可以给出深度学习网络的深度与节点数的优选范围, 最后利用分类器得出的训练集误差进行网络的适当微调即可方便地给出拥有最优识别率的深度学习网络. 根据文献 [15, 17] 我们认为:

定理 1. 重构误差与网络能量正相关.

证明. 因为

$$Err = \frac{\sum_i^n \sum_j^m (F_{i,j} - T_{i,j})}{nmN} \quad (10)$$

式(10)为重构误差的定义公式. 其中 n, m 分别为数据的维度, N 为数据的总个数, F 为经过网络处理后的值, T 为原始值. 则

$$\begin{aligned} F &= P(x_0)P(h|x_0)P(x|h) \\ &= P(x_0) \frac{P(x_0, h)}{P(x_0)} \frac{P(x, h)}{P(h)} \\ &= P(x_0, h) \frac{P(x, h)}{P(h)} \\ &= P(x_0|h)P(h) \frac{P(x, h)}{P(h)} \\ &= P(x_0|h)P(x, h) \end{aligned} \quad (11)$$

$$T = P(x_0) \quad (12)$$

将重构误差的公式代入, 可得

$$Err = \frac{\sum_i^n \sum_j^m (p_{i,j} - T_{i,j})}{nmN} = P - T \quad (13)$$

将式(11)、(12)代入有

$$\begin{aligned} Err &= P(x_0|h)P(x, h) - P(x_0) \\ &= P(x_0)[P(x, h) - 1] \end{aligned} \quad (14)$$

由式(1)的 RBM 的概率模型可以获得

$$Err \propto E(x, h) \quad (15)$$

证毕.

即网络的重构误差可体现网络的整体特性. 在使用网络重构误差对网络层数进行选择时, 其规则如下: 当重构误差下降至所设定的阈值时, 保留原有的网络结构, 否则网络层数增加一层.

定义 1. 深度单层重构误差 (Deep Learning Layer-layer Reconstruction Error), 是深度学习网络中每一个层 RBM 输入与输出数据的重构误差. 它可表示为

$$Err_{\text{单}} = \frac{\sum \sum (x_i - x_{i-1})^2}{N} \quad (16)$$

其中, x_i 为第 i 个 RBM 的输出, 即第 $i+1$ 个 RBM 的输入, x_{i-1} 为第 i 个 RBM 的输入即第 $i-1$ 个 RBM 的输出, N 为测试数据集的数据量.

为避免出现负值导致计算误差, 本文对式(10)的重构误差进行改进, 使用二阶范式, 不仅可以减少计算量, 而且改进后的重构误差仍可表现每层 RBM 对数据的转移情况. 深度单层重构误差是进程择优法中十分重要的参考量. 由于深度学习网络每一层独立, 每一个 RBM 是单独训练的, 则如果将每一层看成一个网络, 由定理 1 可知, 每一层的重构误差与每一层网络的能量相关, 每一层重构误差所体现出的趋势就与网络的整体特性相关.

所以深度单层重构误差与网络的总体特性也是相关的.

引理 1. 通过无监督的学习, 深度学习网络的权值已经处于一个比较好的位置. 后续的训练对其权值的调整十分微小.

引理 2. 网络的训练精度随层数的增加而提高.

引理 3. 由于使用反向梯度算法, 系统的训练误差会层层累加.

引理 4. 对于每层 RBM 节点数相同的系统, 节点数与系统提取的数据量正相关, 随着层数的增加每层的收敛速度应慢慢下降.

引理 5. 深度学习网络迭代一次所得出的重构误差与迭代多次稳定后的误差趋势大致相同.

引理 1~4 由深度学习网络的结构特点引出, 引理 5 由深度学习算法使用的对比分歧算法的原理引出.

推理 1. 在深度学习网络中, 某层深度单层重构误差下降到最底层深度单层重构误差的 20% 时, 系统达到最优.

此推理是将经济学中的二八定律转移到系统学中. 二八定律又名帕累托定律, 是 19 世纪末 20 世纪初意大利经济学家帕累托发明的. 他认为, 在任何一组东西中, 最重要的只占其中一小部分, 约 20%, 其余 80% 的尽管是多数, 却是次要的. 而现在二八定律已不仅仅只是运用在社会学以及经济学中. 经研究者研究发现, 在计算机系统也存在二八定律^[18]. 如 80% 的错误存在于 20% 的代码中、80% 的读写操作应用在 20% 的硬盘空间中等等. 这里, 本文将二八定律应用在深度学习网络中. 推理 1 表明, 当深度单层重构误差下降到最底层深度单层重构误差的 20% 时, 该网络达到最优. 大量的统计结果已经证明了二八定律的正确性, 因此本推理也是成立的.

3.2 深度学习网络的快速设计方法

基于以上分析, 我们可以获得进程择优法及深度学习网络的快速设计方法如下:

步骤 1. 构建一个层数、节点随机的深度学习网络, 并将训练数据带入网络中进行训练, 得出每层的深度单层重构误差.

步骤 2. 对深度单层重构误差进行分析:
 (1) 若该深度学习网络的重构误差随着层数增加逐渐下降, 且在某层已到达所设定的误差阈值, 则保留此网络结构;

(2) 若在达到阈值之前出现重构误差加大, 则说明该网络的节点过多, 所提取数据量过大, 应减少节点, 重新构建一个深度学习网络;

(3) 若最后一层重构误差都未达到阈值, 则层数过少, 应增加层数, 重新构建一个深度学习网络;

(4) 若该网络深度单层重构误差下降过快, 当出现某层重构误差快速下降到上层重构误差所设定的阈值时(该阈值通常设定为某层重构误差快速下降到上层重构误差的 10%~15%), 则代表网络节点数过低, 该网络降维剧烈, 所提取的信息量过少, 应适当增加节点数, 需重新构建一个深度学习网络.

步骤 1、2 所描述的进程择优法的应用流程图如图 2 所示.

步骤 3. 挑选出符合上两条规则的网络, 然后加入分类器组成一种深度信任网络, 利用训练数据对

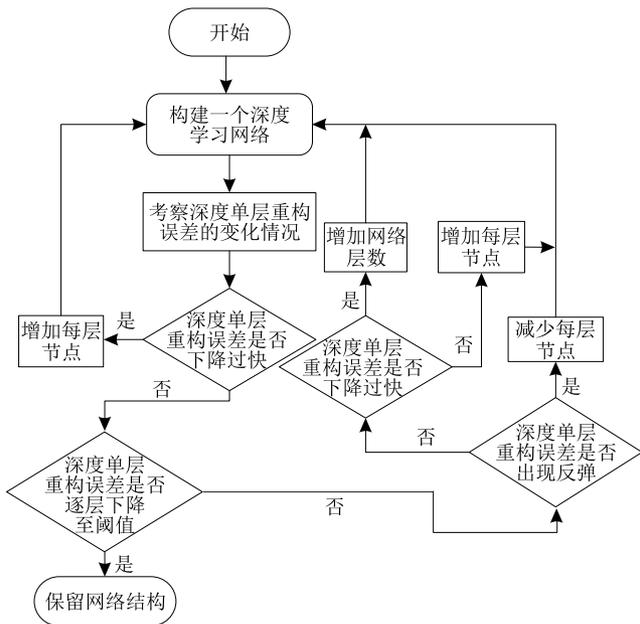


图 2 进程择优法应用流程

它进行训练, 得出相关的误识别率, 在对全部误识别率进行对比分析的基础上, 最终构建出最优的深度学习网络.

使用本进程择优法, 可以实现: (1) 解决了目前深度学习网络的节点数、层数、学习率等参数没有选择标准的问题; (2) 可方便、快速地给出深度学习网络结构的优选范围, 达成深度学习网络的快速构建.

4 心音深度学习网络的构建

按照上述深度学习网络的快速设计方法, 本文构建一种用于对自然环境下的心音信号进行分类处理的深度学习网络, 其系统示意图如图 3 所示.

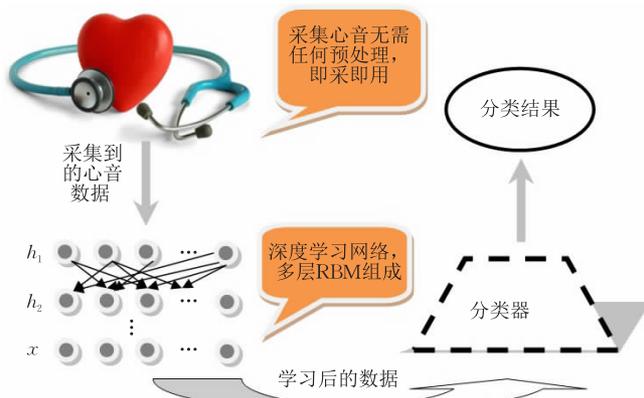


图 3 心音深度学习网络示意图

要处理的心音数据库由 1600 个正常与非正常心音组成, 每条心音样本时长不少于 3 s, 正常心音

样本有 800 条,非正常心音 800 条(包括早搏心音),采样频率统一为 4000 Hz. 该数据库主要由本实验室自行采集取得,并且参考使用了部分其它心音信号^{①②}. 使用该数据库的心音时不需要做任何的预处理. 图 4 是 4 种不同心音的波形图.

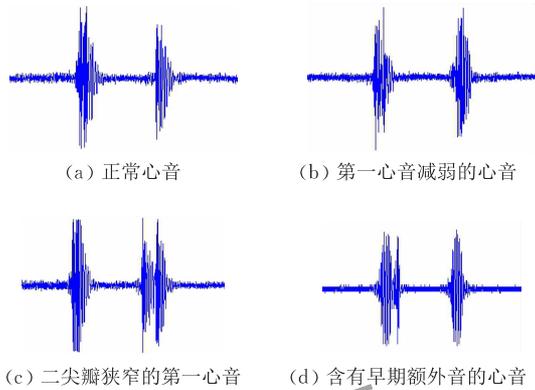


图 4 不同种类心音的波形图

表 1 使用训练库数据所获得层与层之间的重构误差表

$b \backslash a \backslash c$	20	40	60	80	100	120	150	200
1	59.1436	54.6718	53.5645	52.1117	51.1032	49.8868	49.8606	47.8010
2	5.4684	13.1525	20.1444	28.4842	37.2516	44.4062	55.7666	75.2152
3	2.1638	4.5349	6.9003	7.8344	8.8434	11.1311	9.2948	10.2240
4	2.3663	3.1348	3.8904	5.2701	6.4173	9.1442	9.9193	12.4477
5	1.5783	2.6103	2.5187	3.0984	3.6100	4.3294	4.1229	4.6726
6	0.98919	1.7434	2.9409	4.0001	4.0132	4.2219	5.9614	7.7641

步骤 2. 按照上述步骤,我们在实验中发现,当每层的节点数超过 150 个时,从第二层开始重构误差就逐步增加,这说明,网络节点数过多,第一层即底层累积了过多的误差. 适当降低节点数后,网络每层的重构误差开始下降,当节点数下降到 20 节点之后,第二层的重构误差即下降到底层的 10% 以下,这说明网络降维太过剧烈. 应适当增加节点数. 此时我们画出每层节点数在 40 到 120 之间时,深度单层重构误差随层数变化的折线图如图 5 所示.

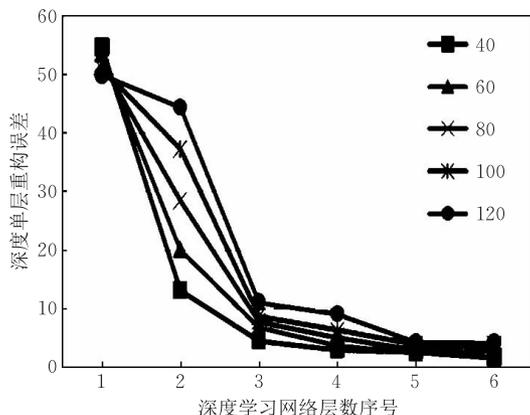


图 5 每层节点数 40 到 120 深度单层重构误差随层数变化的折线图

在以往的心音识别中,心音往往要进行一系列复杂的预处理,比如去噪、分段,并且一次性输入系统进行识别的心音个数也非常有限. 而本文构造的心音深度学习网络针对数据大、自然环境下的心音信号,相比传统的实验室方法更加实用,并有一定的工程应用性. 先在心音库中随机提取 800 个心音作为训练库使用,另外 800 个心音分为两个测试库(测试库 1, 测试库 2),训练库与测试库没有任何交集,排除相同数据对训练与测试的干扰.

步骤 1. 随机构造一个深度学习网络,层数是 2 到 6 层,每层节点数是 20 到 300,表 1 是使用训练库数据所获得的层与层之间的重构误差. a 为每层的节点数, b 为深度单层重构误差, c 为深度学习网络的层数序号,下文其他表格中 a 、 b 、 c 与此表意义相同.

根据深度学习网络的特点和进程择优法的要求,深度单层重构误差下降的速度应慢慢变缓. 从图 5 中可以看出,对于大于 80 节点的网络层,深度单层重构误差的下降速度有升高的过程,这不符合进程择优法中深度单层重构误差下降速度应慢慢减缓的原则. 通过上述分析,我们得到一个网络架构范围如表 2 所示(下面各表中加粗部分为网络参数备选范围). 最后根据进程择优法的要求,当深度单层重构误差逐层下降到所设定的最底层重构误差的 20% 时,选定该网络. 从表 2 中可得,40 节点、60 节点最底层(第 1 层)的重构误差分别为 54.6718 和 53.5645,它们的 20% 为 10.8 左右. 因此,按照该深度单层重构误差阈值,最终选定的心音深度学习网络的大致网络结构为 40-40-40 或 60-60-60 的结构.

① Heart Sound & Murmur Library. http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html

② FRONTIERS IN BIOSCIENCE; NORMAL AND ABNORMAL KEFS. http://int-prop.lf2.cuni.cz/heart_sounds/h14/sound.htm

表 2 初步选择出的网络结构表

$b \backslash a \backslash c$	20	40	60	80	100	120	150	200
1	59.1436	54.6718	53.5645	52.1117	51.1032	49.8868	49.8606	47.8010
2	5.4684	13.1525	20.1444	28.4842	37.2516	44.4062	55.7666	75.2152
3	2.1638	4.5349	6.9003	7.8344	8.8434	11.1311	9.2948	10.2240
4	2.3663	3.1348	3.8904	5.2701	6.4173	9.1442	9.9193	12.4477
5	1.5783	2.6103	2.5187	3.0984	3.6100	4.3294	4.1229	4.6726
6	0.98919	1.7434	2.9409	4.0001	4.0132	4.2219	5.9614	7.7641

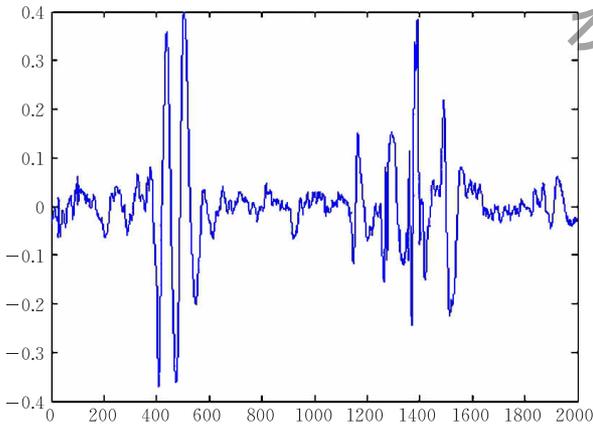
步骤 3. 我们使用 BP 神经网络作为分类器组成心音深度信任网络,使用心音训练数据库对它进行训练. BP 神经网络对数据的种类、数据量没有要求,也不需要参考的标准心音,训练后的训练误差以及识别误差可作为衡量标准.表 3 为不同节点数、不同层数的心音深度信任网络的训练误差表.其中 a

是深度信任网络每层的节点数, b 是训练误差, c 是深度信任网络的层数.

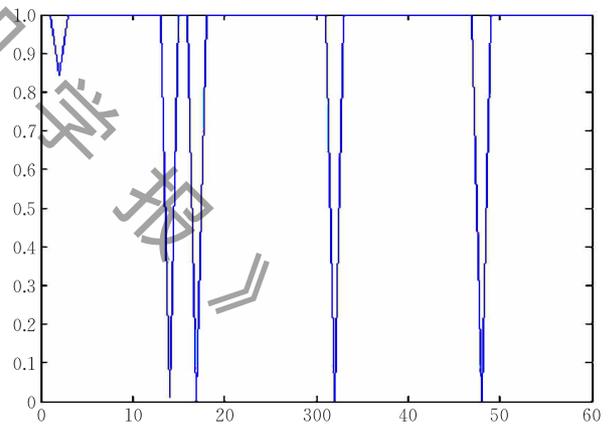
与训练集误差结合分析,按照深度学习网络的快速设计方法,可得出 60-60-60 的网络结构最优,它的训练误差是 0.0225,最低.即每层 60 节点共三层的网络为最优网络结构.图 6 为 60-60-60 结构的

表 3 不同节点数、不同层数的心音深度信任网络的训练误差表

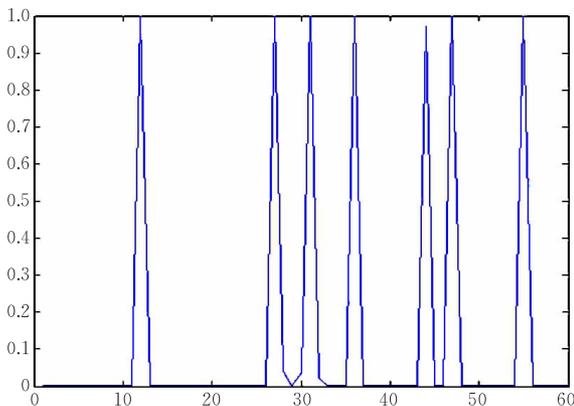
$b \backslash a \backslash c$	20	40	60	80	100	120	150	200
2	0.0512	0.0325	0.0225	0.0225	0.0225	0.0313	0.0325	0.0775
3	0.0750	0.0250	0.0225	0.0338	0.0525	0.0288	0.0425	0.0462
4	0.2062	0.0525	0.0462	0.0537	0.0638	0.0537	0.1250	0.1125
5	0.2062	0.0950	0.0900	0.0612	0.1588	0.1038	0.1575	0.1425
6	0.2062	0.1250	0.1850	0.0688	0.1762	0.1437	0.2062	0.1550



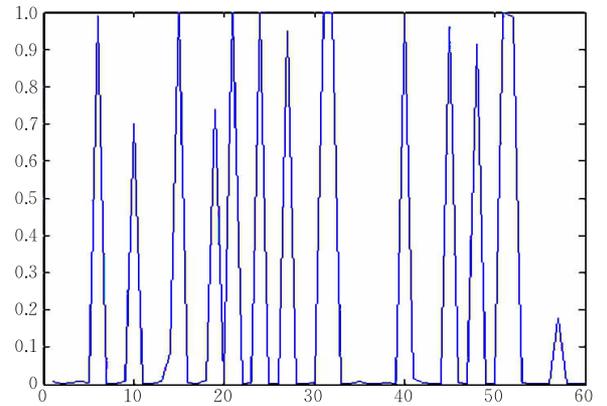
(a) 原始心音数据的波形



(b) 经过第一层网络学习后的心音数据



(c) 经过第二层网络学习后的心音数据



(d) 经过第三层网络学习后的心音数据

图 6 60-60-60 结构的深度学习网络对正常心音进行学习后,每一层输出数据的示意图

深度学习网络对正常心音进行学习后,每一层输出数据的波形图.其中图 6(a)为原始心音数据的波形;图 6(b)是经过输入层网络学习后获得的心音特征数据;图 6(c)是经过中间层网络学习获得后的心音特征数据;图 6(d)是经过输出层网络学习获得后的心音特征数据;可见经过 60-60-60 结构的深度学习网络后,心音特征数据变得比较丰富了.

然后采用心音测试库进行验证,60-60-60 的最优网络结构对于心音库中两测试库(测试库 1,测试库 2)的测试误差仅为 0.1075 和 0.0975,而 60-60-60-60(60 节点四层)的网络结构误差为 0.1125 和 0.1175,40-40-40(40 节点三层)的网络结构误差为 0.1275 和 0.1150,100-100-100(100 节点三层)层数结构的误差为 0.1700 和 0.1400.所述结果如表 4 所示.显然,60-60-60 最优网络结构相比其它结构的深度信任网络拥有更低的误识别率,平均误识别率在 10%左右.从而证实我们提出的进程择优法行之有效.

表 4 多组网络结构的误识别率比较

误识别率 \ 测试库	测试库 1	测试库 2
网络结构		
60-60-60	0.1075	0.0975
60-60-60-60	0.1125	0.1175
40-40-40	0.1275	0.1150
100-100-100	0.1700	0.1400

5 心音深度信任网络的识别实验

5.1 心音能量特征的提取

小波分析可将原始信号从时域转移到频域内,将原信号转为一系列小波的组合,从而将原信号进行频率上的分解.分解后,将各频率段上的小波归一化能量计算出来即可得到信号的一组能量特征.

在对信号进行小波分解前,一般要先选定需要分解的层数.如原信号的频率为 f ,使用小波对信号进行 N 层分解,则可将原始信号分解到 2^N 个频率段上,每一个频率段的带宽为 $f/2^N$.基于小波分解的这种特性,可利用原信号在不同频带上的特性选定小波分解的层数.能量为信号每点幅值的平方和.设 E_i 为心音信号经过 N 层分解后第 i 个频率段的能量值,则相应的归一化能量值为 $E'_i = E_i / \sum_{k=1}^{2^N} E_k$.

一般来说心音信号的主要成分位于 $0 \sim 600$ Hz 内,第一心音主要集中在 $0 \sim 150$ Hz 内,第二心音主

要集中在 $0 \sim 200$ Hz 内,在 $250 \sim 350$ Hz 内第二心音会出现第二个峰值.本文所使用的心音数据频率为 4000 Hz,为更好地体现心音的频率特性,将数据库中心音进行 5 层分解,经分解后每个数据段的带宽为 125 Hz,由于频率在 $1000 \sim 2000$ Hz 以上,心音的有效成分已经非常少, 2000 Hz 以上基本都是无用的噪声.所以为提高特征提取的效率, $1000 \sim 2000$ Hz 以及 2000 Hz 以上的频率段不再分解,由此可得以下 10 个频段.

$0 \sim 125$ Hz、 $126 \sim 250$ Hz、 $251 \sim 375$ Hz、 $376 \sim 500$ Hz、 $501 \sim 625$ Hz、 $626 \sim 750$ Hz、 $751 \sim 875$ Hz、 $876 \sim 1000$ Hz、 $1001 \sim 2000$ Hz、 $2001 \sim 4000$ Hz.由归一化能量公式可得到心音数据的能量特征向量 $T = \{E'_1, E'_2, E'_3, E'_4, E'_5, E'_6, E'_7, E'_8, E'_9, E'_{10}\}$.

5.2 基于心音能量特征的心音深度信任网络的仿真与分析

本文使用 db5 小波对原始心音数据库进行特征提取,并输入到上述心音深度信任网络中.

图 7 是图 4(a)正常心音所提取出的能量特征向量柱状图.

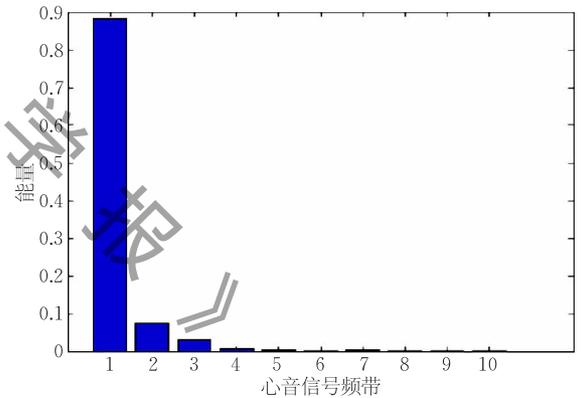


图 7 一组正常心音提取出的能量特征向量柱状图

表 5 是基于心音能量特征的心音深度信任网络识别率统计表.

表 5 基于心音能量特征的心音深度信任网络误识别率

	测试库 1	测试库 2
误识别率	0.0300	0.0350

从表 5 中可以明显看出,基于心音能量特征的心音深度信任网络拥有良好的识别效果,误识别率比之原先不采用心音能量特征的心音深度信任网络明显降低,误识别率仅为 3%,已接近传统实验室小数据心音识别方法的效果.

由于心音深度信任网络对输入数据的限制很小,所以可以提取出多种心音的特征向量,并将其简

单结合为一个融合的特征向量输入到心音深度信任网络中进行识别. 因此提取心音特征是大幅度提高识别率的一种有效方法.

5.3 与 KNN 和 SVM 算法识别效果的比较

我们测试单纯使用 KNN 最近邻算法以及 SVM 算法对心音数据库的数据进行识别的效果.

KNN 最近邻算法中的 K 值使用最大类间方差法. 经计算当 $K=20$ 时, 类间方差最大.

表 6 为 KNN、SVM 分类器对原始心音数据的误识别率表. 从识别效果上看, 这两种分类器并不适用于直接应用于心音数据的识别, 错误率相当高, SVM 甚至无法在有限的迭代次数中收敛. 通常需要对心音进行去噪、分段、特征提取后再使用 KNN、SVM 分类器效果才好^[8]. 当基于心音能量特征进行识别实验, KNN、SVM 分类器的识别效果明显上升, 误识别率分别降为 0.05250、0.04500 和 0.0750、0.0600.

表 6 KNN、SVM 分类器对心音数据的误识别率表

	KNN		SVM	
	测试库 1	测试库 2	测试库 1	测试库 2
原始心音数据误识别率	0.5850	0.5925	在有限迭代次数内不收敛	在有限迭代次数内不收敛
心音能量特征误识别率	0.0525	0.0450	0.0750	0.0600

表 7 经过第一次筛选的不同节点数、不同层数的 MNIST 深度学习网络的深度单层重构误差表

b/a c	60	80	100	150	200	300	400
1	71.1275	70.6322	65.6136	64.3139	63.4662	63.2673	66.2706
2	7.4401	8.4717	10.4164	11.9364	13.0133	14.4253	13.3402
3	6.9878	8.1653	9.0310	10.9085	12.7530	13.1537	15.4317
4	4.4323	5.3721	5.9832	6.7614	7.3692	8.9805	8.1867
5	4.4969	5.3822	7.3353	7.6916	8.2884	8.2198	7.6088
6	3.7849	4.1903	4.8042	5.6273	5.0133	4.3971	6.0024

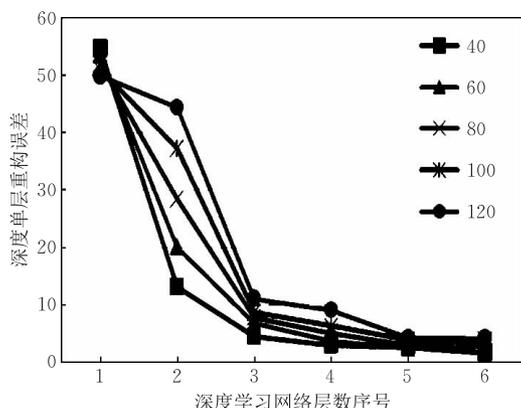


图 8 每层节点数 80 到 300 深度单层重构误差随网络层数变化的折线图

6 进程择优法在手写数据分析中的应用

本文提出的进程择优法不仅可以用于心音深度学习网络的快速构造, 而且还可用于其它数据的分类识别. 为检测进程择优法的普适性, 我们特选取 MNIST 手写数据库^[3]进行验证. 按照深度学习网络的快速设计方法, 构建一种用于对手写数据进行分类处理的深度学习网络. MNIST 手写库中拥有七万张手写数字图片, 图片像素为 28×28 , 其中六万张作为训练数据, 一万张为测试数据.

对手写数据库所建立的深度学习网络, 根据进程择优法, 当每层节点数不超过 400 时, 深度单层重构误差会逐层下降至阈值. 当每层节点数少于 60 时, 第二层的深度单层重构误差就会下降到最底层的 10% 以下. 则可确定网络每层节点的大致范围为 80~300 节点. 表 7 是经过第一次筛选的不同节点数、不同层数的 MNIST 深度学习网络的深度单层重构误差表. 因为对 MNIST 数据库来说, 由于其种类较多, 有 0~9 十种手写数字, 需要节点数较多、层数也比较多的深度学习网络进行分类处理. 从表 7 中可以看出这个结论. 画出可选节点范围内深度单层重构误差的折线图, 图 8 是该网络每层节点数在 80 到 300 之间时, 深度单层重构误差随网络层数变化的折线图.

可以看出对于 MNIST 数据库, 其深度单层重构误差下降速度普遍较快, 下降速度较缓是节点数为 200 和 300. 参考最底层重构误差的阈值, 给出的优选结构范围如表 8 所示.

采用进程择优法可得, 在节点数 80 到 300 的五种结构中, 当网络结构为 200-200-200-200 (200 节点四层) 时训练集的误差最小, 从表 8 可得, 其测试误差仅为 0.073%, 相比较其他几种层数结构的测试误差是最小的. 当选用网络结构为 200-200-200-200 (200 节点四层) 对测试集进行处理时, 其测试误差仅为 0.022%. 综上所述, 我们可以获得, 200-200-

表 8 经过最终筛选的不同节点数的深度学习网络各层深度单层重构误差表

$\begin{matrix} a \\ b \\ c \end{matrix}$	60	80	100	150	200	300	400
1	71.1275	70.6322	65.6136	64.3139	63.4662	63.2673	66.2706
2	7.4401	8.4717	10.4164	11.9364	13.0133	14.4253	13.3402
3	6.9878	8.1653	9.0310	10.9085	12.7530	13.1537	15.4317
4	4.4323	5.3721	5.9832	6.7614	7.3692	8.9805	8.1867
5	4.4969	5.3822	7.3353	7.6916	8.2884	8.2198	7.6088
6	3.7849	4.1903	4.8042	5.6273	5.0133	4.3971	6.0024

200-200 较其他结构深度单层重构误差下降较和缓,是一种针对手写数据库的最优深度学习网络结构.显然,手写库中数据类别较多,理论上也的确需要更多的网络层数进行识别.从而表明我们提出的进程择优法是行之有效的.

7 结 论

本文对心音分类识别技术在大数据、云计算领域的应用进行了先导性的研究,归纳出深度学习网络的一系列特点,并且将深度学习算法与心音识别技术进行结合,构建出一个适于在自然环境下,对大量心音数据进行学习的深度学习网络,并辅以适当的分类器构建出一种心音深度信任网络.

(1) 基于深度学习网络的特点,本文提出了一种深度学习网络结构的快速确定方法——进程择优法.进程择优法使用深度学习网络中深度单层重构误差的变化情况作为考量标准,并将经济学中的二八原则引入到网络架构分析中帮助确定深度单层重构误差的阈值,可方便快捷地给出深度学习网络中每层节点数以及网络层数的优选范围.进程择优法可有效解决深度学习网络在进行网络架构时每层节点数与网络层数无法确定的问题.

(2) 针对自然条件下的心音信号,应用进程择优法快速构建出一种心音深度学习网络.经实验验证,由心音深度学习网络组成的深度信任网络比其它层数结构组成的深度信任网络具有更低的误识别率,最适合应用于心音数据的分类识别.

(3) 由于深度信任网络对输入数据的限制很小,所以先提取输入数据的特征向量,然后将其作为深度信任网络的输入,是大幅度提高识别率的一种有效方法.

(4) 本文提出的进程择优法不仅可以用于心音深度学习网络的快速构造,而且还可推广应用于手写数据等深度学习网络的构造,具有一定的应用

前景.

本文的研究对于快速构建一种深度学习网络,特别是有针对性地设计心音深度学习网络,对提高心音识别算法在自然环境下处理大数据的能力,具有积极的意义.构建更完善的心音深度识别系统,则是下一步的研究工作.

参 考 文 献

- [1] Cheng Xue-Qi, Jin Xiao-Long, Wang Yuan-Zhuo, et al. Survey on big data system and analytic technology. *Journal of Software*, 2014, 25(9): 1889-1908(in Chinese)
(程学旗, 靳小龙, 王元卓等. 大数据系统和分析技术综述. *软件学报*, 2014, 25(9): 1889-1908)
- [2] Zhou S, Chen Q, Wang X. Active semi-supervised learning method with hybrid deep belief networks. *PLoS One*, 2014, 9(9): e107122-e107122
- [3] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, 521(7553): 436-44
- [4] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, 18(7): 1527-1554
- [5] Abdel-Zaher A M, Eldeib A M. Breast cancer classification using deep belief networks. *Expert Systems with Applications*, 2016, 46(1): 139-144
- [6] Bengio Y. Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, 2009, 2(1): 1-127
- [7] Seyyedsalehi S Z, Seyyedsalehi S A. A fast and efficient pre-training method based on layer-by-layer maximum discrimination for deep neural networks. *Neurocomputing*, 2015, 168(C): 669-680
- [8] Cheng Xie-Feng, Fu Nv-Ting. Heart sound identification recognition reviews. *Journal of Shanghai Jiaotong University*, 2014, 48(12): 1745-1750(in Chinese)
(成谢锋, 傅女婷. 心音身份识别综述. *上海交通大学学报*, 2014, 48(12): 1745-1750)
- [9] Patangay A, Tewfik A. Low complexity tracking for long term monitoring of heart sounds//*Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2008)*. Vancouver, Canada, 2008: 17-20

- [10] Varghees V N, Ramachandran K I. A novel heart sound activity detection framework for automated heart sound analysis. *Biomedical Signal Processing & Control*, 2014, 13(1): 174-188
- [11] Cheng Xie-Feng, Zhang Zheng. A construction method of biorthogonal heart sound wavelet. *Acta Physica Sinica*, 2013, 62(16): 168701(in Chinese)
(成谢锋, 张正. 一种双正交心音小波的构造方法. *物理学报*, 2013, 62(16): 168701)
- [12] Dahl G E. *Deep Learning Approaches to Problems in Speech Recognition, Computational Chemistry, and Natural Language Text Processing*[Ph. D. dissertation]. University of Toronto, Toronto, Canada, 2015
- [13] Mohamed A R, Sainath T N, Dahl G, et al. Deep belief networks using discriminative features for phone recognition// *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2011)*. Prague, Czech Republic, 2011; 5060-5063
- [14] Fasel I, Berry J. Deep belief networks for real-time extraction of tongue contours from ultra-sound during speech// *Proceedings of the 20th International Conference on Pattern Recognition*. Stroudsburg, USA, 2010; 1493-1496
- [15] Pan Guang-Yuan, Chai Wei, Qiao Jun-Fei. Calculation for depth of deep belief network. *Control and Decision*, 2014, 30(2): 256-260(in Chinese)
(潘广源, 柴伟, 乔俊飞. DBN 网络的深度确定方法. *控制与决策*, 2014, 30(2): 256-260)
- [16] Jiao Li-Cheng, Zhao Jin, Yang Shu-Yuan, et al. Research advances on sparse cognitive learning, computing and recognition. *Chinese Journal of Computers*, 2016, 39(4): 835-852(in Chinese)
(焦李成, 赵进, 杨淑媛等. 稀疏认知学习、计算与识别的研究进展. *计算机学报*, 2016, 39(4): 835-852)
- [17] Zhou Ming-Ke, Zhang Xu-Yao, Yin Fei, et al. Discriminative quadratic feature learning for handwritten Chinese character recognition. *Pattern Recognition*, 2016, 49(6): 7-18
- [18] Han Wei, Wang Yong. *National Computer Rank Examination Level 4 Software Testing Engineer*. Beijing: Electronic Industry Press, 2011; 25(in Chinese)
(韩为, 王勇. *全国计算机等级考试四级软件测试工程师*. 北京: 电子工业出版社, 2011; 25)



CHENG Xie-Feng, born in 1956, professor, Ph. D. supervisor. His research interests include intelligent information processing and intelligent instruments.

YANG He, born in 1990, M. S. Her research interests include intelligent information processing and deep learning algorithm.

MA Yong, born in 1983, Ph. D., lecturer. His research

interests include computer application technology and robot.

ZHANG Xue-Jun, born in 1965, Ph. D., professor. His research interests include wireless communication network, electroencephalogram (EEG) and magnetoencephalogram (MEG) signal processing.

ZHANG Shao-Bai, born in 1954, Ph. D., professor. His research interests include computer technology and brain-computer interface.

WANG Yue, born in 1992, M. S. candidate. Her research interests include intelligent information processing and intelligent instruments.

Background

Heart sounds can reflect the beating of the heart and are closely related to the health of the heart or the human's body. As a kind of valuable biological signal, it has attracted the attention of researchers both domestic and abroad. In the sophisticated natural environment, deep learning algorithm has become the mainstream algorithm in image and voice recognition field because of its excellent characteristics in big data processing. In the traditional recognition of the heart sounds, the heart sound signals are usually laboratory collected and need complex pretreated.

This paper made a pilot study of heart sound classification and recognition technology in big data and cloud computing field. It summed up a series features of deep learning network,

by combining deep learning algorithm and heart sound recognition technology together to construct a deep learning network, suitable for natural environment and large number of heart sound data for learning. Then assisted by appropriate classifiers, we put forward a heart sound deep belief network.

(1) This paper presents a method for fast determination of deep learning network structure, the preferred method of process, based on the characteristics of the deep learning network. This method uses the variation of the deep reconstruction error of single layer in the deep learning network as a criterion. By introducing the pareto principle in economics into the preferred method of process, to help determine the threshold for single layer reconfiguration errors in deep

learning network, it can be convenient, quickly to give optimization range of each layer node number and the number of network layer. The preferred method of process can effectively solve the uncertainty of the layers' number and the nodes' number per layer when set up the network. The data dealt by the heart sound deep learning network can be divided correctly is obvious higher than the original data.

(2) In order to improve the practical applications of heart sound recognition algorithm in natural environment, this paper used the preferred method to build a kind of heart sound deep learning network. Verified by experiment, using the heart sound deep learning network to build the heart sound deep belief system, the system's error recognition rate was declined, and it will achieve best effect when it is used for heart sound classification and recognition technology.

(3) The deep belief network has little limit for input

data. So an effective method to greatly improve the recognition rate is to extract the feature vector of input data, and then used it as input of the deep belief network.

This work was supported by the National Fund Project, "New Method Research and Application of Heart Sounds Feature Extraction and Identification". Our team had done a thorough research on the extraction, recognition and application of heart sound signals. At present, we have mature theoretical research and practical application achievements. And many papers have been Published on domestic journals and international conferences, such as *Chinese Science (F) (Chinese, English version)*, *Physics Journal*, *Electronic Journal* and *Journal of Computational Information Systems*. However, this paper is another achievement in heart sound identification technology research.

《计算机学报》