

BP神经网络改进及其在手写数字识别中的应用

刘接胜

(华南理工大学 自动化科学与工程学院 广东 广州 510640)

【摘要】对BP网络的激励函数进行了改进,并用matlab7对改进后的BP网络性能进行了仿真;对手写数字所提取的25维特征进行水平投影降为5维特征,在基于改进后的BP网络对手写数字进行了识别,并用VC++6.0使之得以实现。实验结果表明,BP网络完全可以对手写数字进行快速、准确地识别,具有广泛的应用前景。

【关键词】BP网络;激励函数;数字识别;降维;识别率

在20世纪80年中期,David Rumelhart、Geoffrey Hinton和Ronald Williams等完整地提出了误差反向传播的学习算法,即我们熟知的BP算法,其系统地解决了多层感知机的学习问题。当前,该网络在很多行业都得到了广泛的应用,如制造业、机器人、电子领域等,其中用的最多的是模式识别方面。但是在该网络得到广泛应用的同时,我们也发现其仍存在很多不足之处,如学习速度慢、网络训练失败的可能性较大等,针对这些缺陷,有不少改进方案先后被提了出来。

数字、英文字母及文字的识别已广泛地应用于交通、边检、银行、邮政等部门,其重要应用价值得到日益体现。近些年来,很多学者采用了多种识别技术对数字、英文字母及文字进行了识别,但是其识别效果都不甚理想。由于神经网络对输入与输出元素间存在复杂的多元非线性关系的问题处理能力很强,本文采用BP网络对手写数字进行了识别。结果显示,其识别效果是相当理想的。

1、BP算法激励函数的改进

理论上,任意阶可导且为非常数的函数都可以作为BP网络的激励函数。通常BP算法都采用sigmoid函数作为激励函数,但是研究发现,sigmoid函数往往会造成网络的收敛速度慢,学习效率低,并且网络容易陷入局部极小值而不是全局最小值等不足,这都表明有必要对现有的BP网络模型进行改造,对激励函数改进是BP模型改进的思路之一。

(1) 对三角函数的改进:

有学者提出采用下面的三角函数代替sigmoid函数^[1]:

$$f(x)=0.5\sin(\lambda x)+0.5 \quad (1)$$

其中 λ 的取值由经验在1.2~1.8之间选取。潘涛的仿真结果显示:采用该三角函数作为BP网络的激励函数在全局最优化上有显著效果,但是对网络的学习速度的提高效果不明显。通过进一步分析和实验发现学习速度的提高效果不明显的原因是式(1)中的两个参数:0.5与 λ 无联系,如果将式(1)改为:

$$f(x)=(0.5/\lambda)\sin(\lambda*x)+0.5/\lambda \quad (2)$$

这样,函数 $f(x)$ 的周期改变的同时,其幅值也能随函数周期的改变而改变,从而能够改变层间两个相连的神经元之间的兴奋程度,也即改变相应的连接权值。其中 λ 仍在1.2~1.8之间选取值,则网络能保证全局最优化的同时,学习速度也明显提高。

(2) 对双极性压缩函数的改进:

传统BP网络采用的Sigmoid函数,其形状固定不变,从而影响了网络收敛速度,可以通过加大其陡峭度同时改变其幅度来加快网络收敛速度,即将其改进为:

$$f(x)=-1+\lambda/[1+\exp(-\lambda*x)] \quad (3-1)$$

$$f(u)=-1+\lambda/[1+\exp\{\lambda*\sum_i(w_{ij}x_i-\theta_j)\}] \quad (3-2)$$

式中的 λ 是所加入的陡峭度因子,其值由实验得出在0.28~1.32中取值。 u_j 为第 j 个神经元的状态值, w_{ij} 代表前级输入 x_i 至第 j 个神经元的权重, θ_j 则为该神经元的阈值。由于 λ (形态因子)的引入,使得函数对于输入而言可以自由地伸缩和平移变换。网络的收敛速度可以非常大的提高,通常可以加快34%~

55%。

图1是利用matlab7对改进激励函数的BP网络在数字识别训练过程中的识别率变化情况的仿真,从图中可看到,尽管误差定得很小,仅为0.1%,但是改进BP网络的收敛速度还是很快,所需的步数很少,缺点是收敛的过程存在小幅度的波动,不是非常平稳。

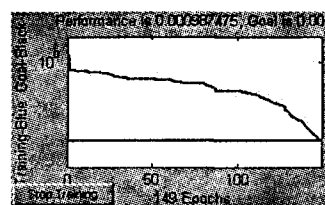


图1改进BP模型数字识别训练过程中识别率变化情况

2、应用改进的BP网络对手写数字的识别

利用BP网络来对对象进行识别之前,应先对待识别对象样本进行预处理,即将手写数字二值化,然后对其进行特征提取,将所提取的特征输入到神经网络进行识别,最后输出并显示识别结果,过程如下图2所示

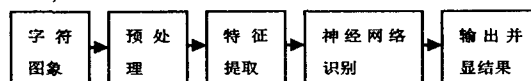


图2 数字识别原理

2.1 BP网络模式的建立:

尽管数字手写写法多种多样,但是数字只有0到9共10个,故采用三层BP网络:输入层、隐含层及输出层,就能解决手写数字识别的问题。其网络结构如图3所示。

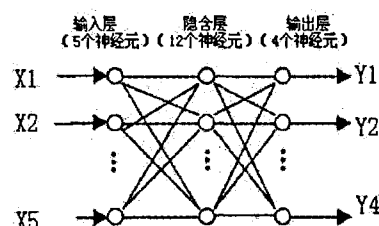


图3 BP网络模型结构

其中 X_i ($i=1,2,3,4,5$)是5个所提取的降维后的数字特征作为输入, Y_i ($i=1,2,3,4$)是输出,其值只能是0或1,用来表示0~9中的一个二进制数。

网络模型各参数的确定:

网络参数的选择直接决定了网络的识别效果,综合考虑各因素,对网络的各参数选择如下:

(1)输入神经元个数确定:一般来说,用5X5的模板对手写数字提取5x5=25个特征就足以对数字进行识别,但是若直接采用这25个特征,网络会比较庞大,为了简化网络,将这25个特征经过水平投影只剩下5个^[2],并将此5个特征作为网络的输入,因此输入神经元的个数定为5个。

(2)隐层神经元个数确定:隐层神经元个数选取要考虑到网络的收敛快慢,因此可以根据式^[9]:

$$S = \sqrt{n + m} + a \quad (4)$$

式(4)中 m 为输出神经元数, n 为输入数, a 为[1, 10]之间常数,因此可选取 $S=12$

(3)输出层神经元个数确定:因为网络的输出只有 0~9 共十个数字,可以采用 4 位二进制数来表示,因此输出层的神经元个数为 4。

(4)激励函数的确定:

根据上面的分析,选取式(3)所表示的函数作为输入层与隐含层之间的激励函数,并采用式(2)所表示的函数作为隐含层与输出层的之间的激励函数。

2.2BP 网络模型的训练

(1)训练数据库

本文使用的数据库是作者自己据常见的数字手写体所写的 110 个数字(0~9 共十个数字各有 11 个样本)其中 80 个作为训练样本,30 个作为测试样本,其中每个数字各为 3 个测试样本。

(2)训练方法

采用“留一法”训练网络^[10],具体做法如下:数据库中有训练 80 个样本,第一次,从样本集中取出第一个样本,再用其余 80-1=79 个样本训练网络,然后用训练得到的 BP 模型对取出的那个样本进行预测;第二次:将第一次取出的样本放回样本库,取出第二个样本,同样用剩余 79 个样本训练网络,然后用训练得到的 BP 模型对此次取出的样本进行预测……,如此依次进行,用预测结果来评价 BP 模型的预测性能。

训练过程曲线如图 1 所示。

2.3 实验结果及分析

基于改进 BP 神经网络的手写数字识别方法,将提取的 5 个特征作为神经网络的输入。选取了 110 个手写数字字符,其中 80 个作为训练样本,剩余的样本作为测试数据。使用 VC++6.0+SP6 补丁作为工具,对所搭建的 BP 网络进行了仿真,实验结果如表 1 所示(表中网络的各项输出指标都是三个测试样本输出的平均值),该 BP 网络模型完全能够识别 30 个样本,网络训练时间短,识别率高,收敛速度快,跟其它识别方法相比,性能有显

著提高。

期望值	网络输出	误差	期望值	网络输出	误差
0	0.008377	-0.008377	5	5.001736	-0.001736
1	1.000591	-0.000591	6	6.0035	-0.009735
2	1.999008	0.000992	7	6.989958	0.010042
3	3.004652	-0.004652	8	8.000344	-0.000344
4	3.991343	0.008657	9	9.001128	-0.001128

表 1 样本期望输出与 BP 网络的实际输出比较表



图 4 部分改进 BP 网络识别效果图

3. 结束语

笔者对 BP 网络的激励函数进行了改进,并基于改进后的 BP 网络模型,对手写数字所提取的特征首先降维的后再进行了识别。仿真结果表明,改进后的 BP 网络不但训练时间短而且识别率高。

参考文献:

1. 陆琼瑜, 童学锋, BP 算法改进的研究, 计算机工程与设计, 2007 年 2 月
2. 杨淑莹, 图像模式识别-VC++技术实现, 清华大学出版社, 2005 年 7 月第一版, 115-134.
3. 飞思科技研发中心, 神经网络理论与 MATLAB7 实现, 电子工业出版社, 2005, 44-51
4. Sung-Sau So, Martin Karplus. Evolutionary Optimization in Quantitative Structure-Activity Relationship: An Application of Genetic Neural Networks, Journal of Medicinal Chemistry, 1996, 39(7):1524.
5. 瞿金海, 手写体数字识别方法的研究与实现, 南京理工大学硕士学位论文, 2005, 9
6. 刘维群, 李元臣, BP 网络的隐节点可调整改进, 洛阳师范学院学报, 2005, (2):68-70

(上接第 122 页)

首先,明确管理流程中各角色及其操作范围。如表 1 所示:

角色	操作范围
学生(申请人)	创建新申请;
任课教师(申请人, 证明人)	创建新申请; 提供成绩资料; 证明; 给出参考意见;
成绩管理员(证明人, 操作人)	提供相关材料; 证明; 执行相关操作;
课程管理员(证明人, 操作人)	提供相关材料; 证明; 执行相关操作;
学籍管理员(证明人, 操作人)	提供相关材料; 证明; 执行相关操作;
教务办主任(审批人)	给出审批意见(初审、终审);
分院教学院长(审批人)	给出审批意见(初审、终审);
院教学院长(审批人)	给出审批意见(终审);

表 1 工作流中角色操作对应表

其次,明确流程中出现的各种状态及对应的操作范围。如表 2 所示:

状态	操作范围
新申请	填写申请内容; 设置相关管理人员;
准备审批中	判断需要出具的证明材料; 提交相关人员;
证明中	给出证明; 设置审批人; 拒绝; 给出原因;
审批中	签署意见; 设置后继审批人; 审批完成;
被拒绝的申请	返回申请人;
被同意的申请	执行相关申请; 归档;

表 2 工作流中状态操作对应表

上面两个表格对任务分解后,将系统中角色和状态及相应操作对应起来。使管理流转的过程清晰,透明,易于完成,充分体现了各部门协同工作的特点。例如,学生提出更改成绩的申请后,就分别由任课老师,成绩管理员和课程管理员给出相关材料,然后进行领导审批。一旦进入设定的工作流后,申请就自动

流转。不再需要手工转发,提高了工作效率。

最后,给出成绩管理工作流的流程图。如图 2 所示:

图中分别标识出了:A 审批准备中, B 证明中, C 审批中三个过程。在申请证明和审批过程中,流转都是循环进行的,通过对下一经手人的设置来实现自动流转。

4. 总结与展望

通过定义系统自动的流转模式,使得教学管理更加规范。并且在一定的规范下灵活变更事务处理过程。很适合高校的教学管理模式。理想的工作流模型能够清楚地定义任意情况下的工作流,能够适应用户在建模过程中所提出的各种要求。然而,从目前模型的能力上看,距这一理想情况尚有一定的距离,还不能处理复杂的过程逻辑,缺乏柔性。

参考文献:

1. 冯峰, 王小刚, OAS 中的工作流设计方法, 计算机工程与应用 2000.3 p91
2. 姬风波, 云红艳, 高岭, 工作流技术在高校 OA 系统中的应用, 西北大学学报(自然科学版)2001.8 vol31.no4 p282
3. 罗海滨, 范玉顺, 吴澄, 工作流技术综述, 软件学报 2000(11):899-907
4. 杨旭东, 我校办公自动化建设工作中的若干问题探讨, 宁波大学学报(教育科学版)2002.12 vol24.no6 p70