

BP 神经网络在图像字符识别中的改进和应用

张永宏¹ 吴鑫¹

摘要

传统 BP 神经网络算法虽然具有良好的学习能力和容错能力,但是收敛速度慢,易陷入局部极小点等缺点制约了它的进一步发展和应用.针对这些不足,采用自适应学习率结合附加动量因子的方法可以有效缩短训练时间,加快收敛速度,同时抑制寻优算法陷入局部极小点.将该算法应用于图像字符识别系统中,通过一系列实验优化系统参数之后给出系统识别结果,表明该系统识别具有较高的准确性和鲁棒性.

关键词

字符识别;BP 神经网络;动量因子;自适应学习率

中图分类号 TP391.1

文献标志码 A

0 引言

随着计算机技术的飞速发展,字符识别技术经过多年的不断改进和完善,现已广泛应用于符号信息自动处理的各个领域^[1].目前,字符识别常用的方法有模板匹配法、特征匹配法以及神经网络识别等.其中,模板匹配法应用最为广泛,其基本原理是对每个模式类都定义一个标准的模式并将它作为模板^[2],特点是识别速度快,但是易受噪声干扰.特征匹配法主要是按照字符特征信息计算相似度进行匹配,能够很容易地区分形近字符,但抗干扰能力弱.基于人工神经网络的字符识别方法具有一些传统技术所没有的优点,如速度较快、分类能力强,具有良好的容错能力、并行处理能力和自学习能力.其中,在字符识别中使用最为广泛的是 BP 神经网络,特点是利用样本字符训练神经网络来获得字符输入和输出的复杂分类规律,最终实现字符的分类识别.在实际应用中,BP 神经网络算法存在收敛速度慢、易陷入局部极小值等缺陷,本文采用附加动量法和自适应调整学习率相结合的策略,改进 BP 神经网络字符识别的效率和准确性.

1 BP 神经网络原理及其改进方法

1.1 BP 神经网络原理

BP 神经网络是一种按误差反向传播算法训练的多层前馈网络,其拓扑结构由输入层、输出层、隐含层构成.它的学习规则是采用最速下降法,通过反向传播不断调整网络的权值和阈值,使网络的误差平方和最小,一个典型的 3 层 BP 神经网络模型如图 1 所示^[3].

BP 神经网络执行的是有导师监督的训练模式,在训练之前必须准备好所有训练样本,在整个训练过程中不可中断,一经训练完成,网络连接权值不可以再更改,否则信息将会全部丢失.BP 神经网络的训练过程主要分为信息的正向传播和误差的反向传播 2 个阶段.在正向传播过程中,输入样本从输入层经过隐含层处理后传送到输出层,再将输出层的输出结果与实际期望输出数据的比较,如果训练未达标,则计算网络误差,并将误差逐级反方向传递到各层并分摊给各层的节点,依据各单元误差动态修正其相应连接权值,不断重复该过程,直到满足训练结束条件^[4-8].

1.2 BP 神经网络的改进

BP 神经网络在实际应用中极为广泛,具有可靠、精度高、通用性

收稿日期 2011-11-03

资助项目 教育部留学回国人员启动基金(2010609);江苏省“六大人才”高峰资助项目(2010-JXQC-132)

作者简介

张永宏,男,教授,主要从事图像处理、模式识别研究.nuis_t_zyh@yahoo.com.cn

¹ 南京信息工程大学 信息与控制学院,南京,210044

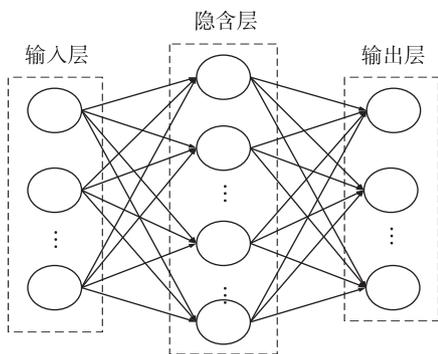


图1 3层BP神经网络结构示意图

Fig. 1 Structure of 3-level BP neural network

好等优点,但标准BP神经网络也存在以下问题:由于采用非线性梯度优化算法,易陷入局部极小点而得不到全局最优;训练次数多,学习效率低,收敛速度慢;难以确定隐层数和隐层节点个数,缺乏相应的理论指导.为了提高BP神经网络训练速度,避免陷入局部极小值和改善其他能力,目前改进算法主要有2种.

1) 附加动量法. 为了提高网络的训练速度,降低网络误差曲面局部细节的敏感性,通过在权值调整中增加动量项,可以带动梯度下降搜索冲过狭窄的局部极小值而不是陷入其中.同时,动量项还能在梯度不变的区域逐渐增大搜索步长,从而加快收敛的作用.带有附加动量因子的权值更新为

$$\Delta w_{ji}(n+1) = (1 - \alpha)\eta\delta_j x_{ji} + \alpha\Delta w_{ji}(n), \quad (1)$$

其中, $w_{ji}(n)$ 为第 n 次迭代时的权值更新, $\alpha(0 \leq \alpha \leq 1)$ 为动量系数.附加动量法的实质是将最后一次权值变化的影响,通过一个动量因子来传递.以此方式,当增加动量项后,促使权值的调节向着误差曲面底部的平均方向变化,当网络权值进入误差曲面底部的平坦区域时 δ_j 将变得很小,于是 $\Delta w_{ji}(n+1) \approx \Delta w_{ji}(n)$,从而防止了 $\Delta w_{ji}(n) = 0$ 的出现,有助于使网络从误差曲面的局部极小值中跳出.

2) 自适应调整学习率. BP算法的收敛特性和收敛速度很大程度上取决于学习率.在实际应用中,很难确定一个从始至终都合适的最佳学习率.通过自适应调整学习率,改善学习率选择不当所带来的震荡与收敛速度慢等缺陷,使得在误差曲面较为平坦的区域增加学习率,避免训练陷入震荡,提高收敛速度.自适应学习率调整公式为

$$\eta(n+1) = \begin{cases} \theta_1 \eta(n), & E(n) < E(n-1), \\ \theta_2 \eta(n), & E(n) > \xi E(n-1), \\ \eta(n), & \text{其他.} \end{cases} \quad (2)$$

其中 η 为学习率, θ_1 、 θ_2 为学习速率调整比例因子, ξ 为允许反弹误差系数, $E(n)$ 为网络误差, $\theta_1 > 1, 0 < \theta_2 < 1, \xi$ 通常略大于1.初始学习率 $\eta(0)$ 的选取有较大随意性,在试验部分给出不同的初始值.

为了充分发挥附加动量法和自适应学习率在提高收敛速度和缩短训练时间的特点,本文采用动量-自适应学习率调整算法.带附加动量的自适应学习率权值更新为

$$\Delta w_{ji}(n) = \begin{cases} (1 - \alpha)\theta_1 \eta(n)\delta_j x_{ji} + \alpha\Delta w_{ji}(n-1), & E(n) < E(n-1), \\ (1 - \alpha)\theta_2 \eta(n)\delta_j x_{ji} + \alpha\Delta w_{ji}(n-1), & E(n) > \xi E(n-1), \\ (1 - \alpha)\eta(n)\delta_j x_{ji} + \alpha\Delta w_{ji}(n-1), & \text{其他.} \end{cases} \quad (3)$$

在权值更新过程中若某次更新后误差减小,则说明搜索继续向全局最小收敛,此时通过 θ_1 系数加大搜索步长以提高收敛速度;若某次更新后误差增大,说明搜索可能越过了误差曲面的极小值,此时通过 θ_2 系数缩小搜索步长以提高搜索精度.同时,添加动量项可以帮助网络冲出高频震荡区域,加速网络收敛.总之,基于附加动量项和学习率自适应调整策略可以使BP神经网络在较短的时间内得到更优解.

2 图像字符识别实验

本部分利用动量-自适应学习率调整算法,通过对图像字符识别的实验分析,研究字符的识别效率和准确性.

2.1 实验设计

BP神经网络输入层的节点个数即为图像字符预处理后所提取特征的维数.本文从训练样本图像中分割出单个字符样本,将分割出的每个字符样本归一化到 8×16 的矩形区域中,特征提取直接利用每个字符样本的像素值作为特征.对于每一个输入字符样本,就有 $128(8 \times 16)$ 个特征,即输入层的节点数为128.

神经网络隐层神经元的数目越多,训练时间也越长.神经元数目计算公式^[6]为

$$m \leq \sqrt{N \times (l+3)} + 1, \quad (4)$$

其中, N 为输入神经元数, l 为输出结点数.

神经网络输出层结点数的确定,取决于如何设定标准输出.本文采用6位二进制输出编码,即需要

6 个输出单元,如数字 3 被编码为 000011. 由式(4)可知,BP 神经网络隐层节点数取最大值 33.

2.2 实验结果及分析

根据具体的待测目标,实验选取 1 幅含有 196 个字符的图像,每组包含 10 个数字字符,26 个大写英文字符和 13 个符号字符,共 4 组作为训练样本,训练样本抽取了部分测试样本中的字符,组成字符序列如图 2 所示. 图 3 所示的扫描字符图像作为待测样本.

```
0123456789ABCDEFGHIJKLMN O P Q R S T U V W X Y Z # $ % & ' ( ) * + , - . /
0123456789ABCDEFGHIJKLMN O P Q R S T U V W X Y Z # $ % & ' ( ) * + , - . /
0123456789ABCDEFGHIJKLMN O P Q R S T U V W X Y Z # $ % & ' ( ) * + , - . /
0123456789ABCDEFGHIJKLMN O P Q R S T U V W X Y Z # $ % & ' ( ) * + , - . /
```

图 2 训练样本

Fig. 2 Samples for training

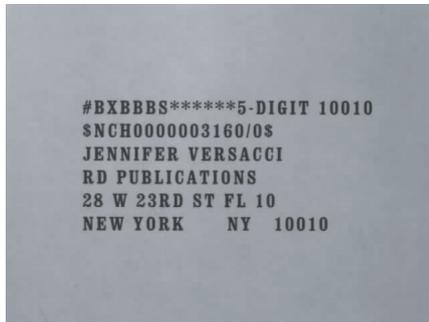


图 3 测试样本

Fig. 3 Samples for testing

实验过程中,分析了学习率、附加动量因子等因素对 BP 神经网络性能的影响,结果如表 1 所示. 传统固定学习率法对初始学习率的依赖度较高,较小的学习率会使训练时间较长,但可以获得较高的字符识别率;而学习率过高时,则存在震荡特性,对网络性能带来不利影响. 自适应学习率法通过在搜索过程中根据误差自动调节学习率,从而增大收敛方向的搜索步长来加速收敛或者减小步长来提高搜索精度,但是依然可能陷入局部极小、高频震荡等. 根据表 1 可知,通过加入动量因子有助于加快网络的训练速度,避免陷入局部极小,可以进一步提高系统收敛速度和精度.

为了对比传统 BP 神经网络和本文采用的动量-自适应学习率调整算法 BP 神经网络的性能,将 BP 神经网络参数统一设置为初始学习率 0.2,误差阈值 0.002,隐层节点数 33,动量因子 0.5. 网络性能比较结果如表 2 所示,可知本文方法具有更高的识别效率和准确率. 图 4 给出了测试样本图像的识别结果.

表 1 各参数对 BP 神经网络性能的影响

Table 1 Influence of different parameters on performance of BP neural network

初始学习率	训练误差	附加动量因子	固定学习率		自适应学习率	
			迭代次数	识别率/%	迭代次数	识别率/%
0.05	0.001	0	4 881	90	2 763	90
		0.3	3 823	87	2 032	90
		0.5	2 934	90	1 960	92
0.10	0.001	0	1 136	82	1 021	85
		0.3	1 080	85	1 017	83
		0.5	916	87	964	85
0.15	0.001	0	1 001	91	989	90
		0.3	992	90	1 031	88
		0.5	873	90	827	92
0.20	0.002	0	1 796	93	1 202	95
		0.3	1 531	90	811	92
		0.5	1 370	94	793	94
0.25	0.002	0	948	90	986	93
		0.3	825	92	1 025	90
		0.5	704	90	882	92
0.30	0.003	0	774	45	823	65
		0.3	793	43	704	65
		0.5	615	40	595	55

表 2 BP 神经网络图像字符识别性能比较

Table 2 Performance comparison of image character identification methods based on BP neural network

BP 神经网络模型	迭代次数	识别率/%
BP 网络	2 422	83
BP 网络 + 附加动量因子	1 955	90
BP 网络 + 自适应学习率	1 573	92
BP 网络 + 自适应学习率 + 附加动量因子	780	94

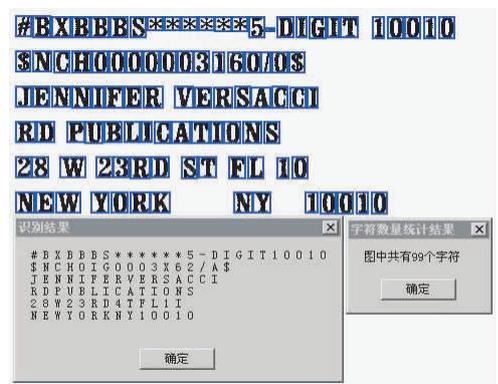


图 4 参数调整后的识别结果

Fig. 4 Identification results after parameter modification

3 结语

传统的 BP 算法存在的不足会影响网络的实际性能,通过引入附加动量和自适应调节学习率的 BP 神经网络保留了传统 BP 算法的优点,如非线性处理能力和较强的容错能力. 自适应学习率的引入有效缩短了训练时间,能够加快收敛速度,而附加动量因子则有效降低了收敛算法陷入局部极小点的危险,两者结合可以明显改善网络实际性能. 该算法应用于图像字符识别系统中,通过一系列实验可知,改进之后的 BP 算法的识别速度更快,准确性更高,鲁棒性更好. 今后的工作重点是减小网络规模,进一步优化网络参数,改善收敛算法,进一步提高识别的精度和速度,广泛应用于车牌识别、身份认证等领域.

参考文献

References

- [1] 马锐. 人工神经网络原理[M]. 北京:机械工业出版社,2010
MA Rui. Principle of artificial neural networks[M]. Beijing: China Machine Press, 2010
- [2] 张彤,肖南峰. 基于 BP 网络的数字识别方法[J]. 重庆理工大学学报,2010,24(3):44-47
ZHANG Tong, XIAO Nanfeng. Character identification based on a BP neural network[J]. Journal of Chongqing

- University of Technology, 2010, 24(3):44-47
- [3] 张煜东,吴乐南,王水花,等. Li-Hopfield 神经网络用于汉字字符识别[J]. 南京信息工程大学学报:自然科学版,2010,2(1):6-12
ZHANG Yudong, WU Lenan, WANG Shuihua, et al. Li-Hopfield neural network used for Chinese character recognition[J]. Journal of Nanjing University of Information Science & Technology: Natural Science Edition, 2010, 2(1):6-12
- [4] 王燕. 一种改进的 BP 神经网络手写体数字识别方法[J]. 计算机工程与科学,2008,30(4):50-52
WANG Yan. An improved handwritten numeral recognition method based on the BP neural network[J]. Computer Engineering & Science, 2008, 30(4):50-52
- [5] 许宜申,顾济华,陶智,等. 基于改进 BP 神经网络的手写字符识别[J]. 通信技术,2011,44(5):106-109
XU Yishen, GU Jihua, TAO Zhi, et al. Handwritten character recognition based on improved BP neural network[J]. Communications Technology, 2011, 44(5):106-109
- [6] Gan D Q. The optimal number of hide nodes in multi-layered feedforward neural networks[C]//IEEE SMC CE-SA, LiLLe, France, 1996
- [7] LU Y. Machine printed character segmentation: An overview[J]. Pattern Recognition, 1995, 28(1):67-80
- [8] 沙妮,李宏,李平,等. 基于 BP 神经网络的印刷字符识别系统[J]. 电脑知识与技术,2009,5(19):5238-5241
Jara S C, LI Hong, LI Ping, et al. Machine printed character recognition system using backpropagation neural network[J]. Computer Knowledge and Technology Journal, 2009, 5(19):5238-5241

Improvement and application of BP neural network in image character recognition

ZHANG Yonghong¹ WU Xin¹

¹ School of Information and Control, Nanjing University of Information Science & Technology, Nanjing 210044

Abstract The traditional BP neural network algorithm is good in learning ability and fault tolerance, while its disadvantages such as slow convergence rate and easily falling into local minimum restrict its further development and application. An improved BP algorithm with self-adaptive learning rate and additional momentum factors can effectively reduce the training time, speed up the convergence rate and inhibit the possibility of falling into a local minimum. The improved algorithm is applied to the image character recognition system. The influences of model parameters on performance of BP neural network are analyzed, and the recognition results are given after a series of parameter optimization. The experimental results show that the improved BP neural network can recognize image characters with high accuracy and robustness.

Key words character recognition; BP neural network; additional momentum; self-adaptive learning rate