

文章编号: 10062475(2008)052017203

基于 BP 神经网络的图像识别研究

张海波, 董槐林, 龙 飞, 郭世可

(厦门大学软件学院, 福建 厦门 361005)

摘要: BP 神经网络具有较强的容错性和自适应学习能力, 因而在数字图像识别领域有着广泛的应用。本文在经典 BP 神经网络的基本算法的基础上, 对 BP 算法的参数设置进行了优化, 实现了一种基于分类的改进 BP 神经网络算法。通过探讨 BP 神经网络在数字图像分类识别中的应用, 详细考察了各种参数对识别效果的影响。实验结果证明改进后的算法有很好的实用价值。

关键词: BP 神经网络; 动量因子; 图像识别

中图分类号: TP391.41 文献标识码: A

Research on Image Recognition Based on BP Neural Network

ZHANG Haibo, DONG Huailin, LONG Fei, GUO Shike

(Software School, Xiamen University, Xiamen 361005, China)

Abstract: Backpropagation neural network with high fault tolerance and good adaptive learning ability is found great applications in digital image recognition. This paper presents an improved algorithm through setting new parameters based on traditional BP algorithm, applies it to digital image recognition and analyzes the influence produced by improved parameters. The experimental result shows that the improved algorithm has a certain practical value.

Key words: BP neural networks; momentum factor; image recognition

0 引 言

数字图像识别前景广阔, 广泛应用于交通、银行、教育和邮政等多个领域, 在这些应用中通常的方法是把待识别的数字区域从原始的图像中提取出来, 然后把其中的数字串分割成单个字符, 再用分类器进行识别。

由于神经网络具有自学习、容错性、分类能力强和并行处理等特点, 对于数字图像来说, 采用神经网络是一种有效的手段。然而, BP 算法还存在许多值得探讨的问题, 例如学习算法的收敛速度慢、局部极小问题、网络的初始权值、阈值及隐层单元数根据经验选取等。针对 BP 算法存在的问题, 研究者提出一些改进方案, 主要包括改变误差函数^[9], 网络拓扑结构^[10]、学习率和参数的动态调整^[4, 11]以及初值的确定方法^[12]等。已有算法通过修改步长在收敛速率方面有了很大改善, 本文通过优化参数对原算法进行了改进, 对数字图像识别的实验证明改进后的 BP 神经网络的识别效果更好。

1 BP 算法

1.1 数学模型

BP 神经网络是一种多层前向网络, 由输入层、输出层、隐含层 (可以是一层或多层) 构成, 一种典型的三层 BP 神经网络模型如图 1 所示。

反向传播算法的主要思想是把学习过程分为 2 个阶段: 第 1 阶段 (正向传播过程), 输入信息从输入层开始逐层计算各单元的实际输出值, 每一层神经元的状态只对下一层神经元的状态产生影响; 第 2 阶段 (反向传播过程), 若在输出层未能得到期望的输出

收稿日期: 2007205218

基金项目: 国家 985 二期信息创新平台项目 (0000-X07204)

作者简介: 张海波 (19832), 男, 山东金乡人, 厦门大学软件学院硕士研究生, 研究方向: 模式识别, 图像处理; 董槐林 (19572), 男, 福建龙岩人, 教授, 研究方向: 数值分析, 软件工程, 图像处理; 龙飞 (19772), 男, 讲师, 博士, 研究方向: 生物特征识别技术, 模式识别与机器学习, 智能信息处理; 郭世可 (19832), 男, 福建厦门人, 硕士研究生, 研究方向: 图像处理, 模式识别。

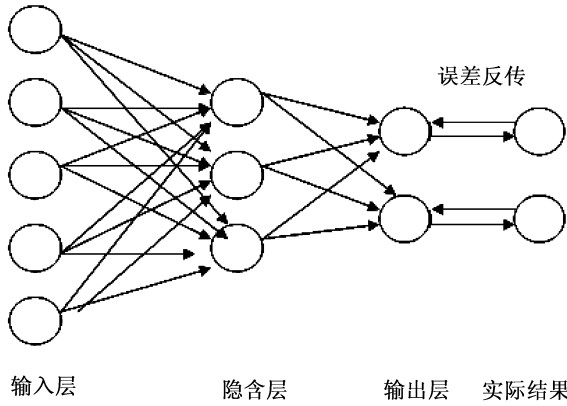


图 1 一个典型的 BP神经网络模型

值, 则逐层递归计算实际输出与期望输出之间的差值, 根据此误差修正前层权值使误差信号趋向最小。它通过连续不断地在相对于误差函数斜率下降的方向上计算网络权值和偏差变化而逐渐逼近目标。每一次权值和误差的变化都与网络误差的影响成正比。

假设输入层、中间层、输出层的单元数分别是 N 、 L 和 M 。 $X = (x_0, x_1, \dots, x_{N-1})$ 是加到网络的输入矢量, $H = (h_0, h_1, \dots, h_{L-1})$ 是中间层输出矢量, $Y = (y_0, y_1, \dots, y_{M-1})$ 是网络的实际输出矢量, 并且用 $D = (d_0, d_1, \dots, d_{M-1})$ 来表示训练组中各模式的目标输出矢量。输出单元 j 到隐单元 i 的权值是 V_{ij} 而隐单元 j 到输出单元 k 的权值是 W_{jk} 。另外用 H_k 和 θ_j 来分别表示输出单元和隐单元的阈值。

传递函数是反映下层输入对上层节点刺激脉冲强度的函数又称刺激函数, 一般取为 $(0, 1)$ 内连续取值 Sigmoid 函数, 即

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

误差函数为:

$$E = \frac{1}{2} \sum_{k=1}^m (d_k - y_k)^2 \tag{2}$$

中间层和输出层各单元的 outputs 为:

$$h_j = f\left(\sum_{i=0}^{N-1} V_{ij} x_i + \theta_j\right) \tag{3}$$

$$y_k = f\left(\sum_{j=0}^{L-1} W_{jk} h_j + H_k\right) \tag{4}$$

BP 算法采用梯度下降法调整权值:

$$W_{ij}(n+1) = W_{ij}(n) + G D_j x_i^c \tag{5}$$

式中, j 为本结点的序号, i 则是隐含层或输入层结点的序号。 x_i^c 是结点 i 的输出或者是外部输入; 式中 G 称为学习速率, D_j 为误差项。

1.2 构建 BP 神经网络步骤

第 1 步, 用均匀分布随机数将权值和阈值初始化。

第 2 步, 对样本数据进行训练:

- (1) 加载输入向量 X 和目标向量 D
- (2) 计算隐含层和输出层的实际输出 H 和 Y ;
- (3) 计算输出层的误差;
- (4) 计算隐含层单元误差;
- (5) 用式 (5) 对权值和阈值进行修改, 将误差信号沿原来的连续返回, 通过修改各层神经元的权值, 逐次地向输入层传播去进行计算。再经过正向传播过程, 这两个过程反复进行使得误差信号最小;
- (6) 计算全部误差是否满足要求, 若满足则学习结束, 否则回到 (3)。

第 3 步, 输入信息向量 X 经仿真计算预测输出向量 Y 。

2 改进后的算法

由于在 BP 网络中采用不同的传递函数时, 产生的激励值不同。在这里选择反正切函数和 S 型函数结合使用。反正切公式如下:

$$g(x) = \frac{3}{\pi} \arctan x \tag{6}$$

反正切函数曲线比 Sigmoid 函数曲线陡峭。根据生物神经元的基本特点, 在隐含层采用反正切函数, 输出层采用 Sigmoid 函数。若采用反正切函数作为输出函数, 则由于其曲线斜率大, 激励程度比较大, 会产生过激的响应, 易引起振荡或不收敛, 而 Sigmoid 函数具有饱和非线性特性, 可增强网络的非线性映射能力。

针对 BP 算法收敛速度慢的缺点, 结合已有的对学习速率 G 的改进方法^[4], 同时引入动量因子, 对权值的学习过程进行改进。参数取值如下:

$$v W_{jk}(n) = (A/(1+L)) * (v W_{jk}(n-1) + 1) * D_k * h_j + B v W_{jk}(n-1) \tag{7}$$

$$v V_{ij}(n) = (A/(1+N)) * (v V_{ij}(n-1) + 1) * D_j^* x_i + B v V_{ij}(n-1) \tag{8}$$

$v W_{jk}(n)$, $v V_{ij}(n)$ 分别为中间层到输出层的权值和输入层到中间层的权值。其中 B 为动量因子, 取值范围为 $(0, 1)$, 它的存在可以平滑学习过程, 在一定程度上避免振荡, 减小陷入局部极小的概率。

3 实验结果与分析

3.1 实验环境

将本文方法应用到手写体数字识别问题, 选用 MNIST 数据库作为训练和测试用字符集, 有关该数据库更详细的信息请见 <http://yann.lecun.com>。本文从中抽取 5000 个样本作为训练样本, 另外抽取 200 个样本作为识别样本。样本图像如图 1 所示。

BP 网络的输入层神经元个数为 784 输出层神经元个数为 10 个。

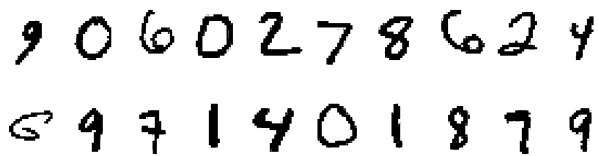


图1 样本实例

3.2 不同隐结点数对识别率的影响

图2给出隐层结点数对识别率的影响。当隐层结点数取10到30之间时,识别率随着隐层结点数的增多呈上升趋势;当隐层结点数增加到大于40后,识别率不再显著增加,而是围绕0.95振荡,由于隐层结点数的增加将伴随着网络计算规模的扩大,从而影响计算效率,这里取隐层结点数为40。

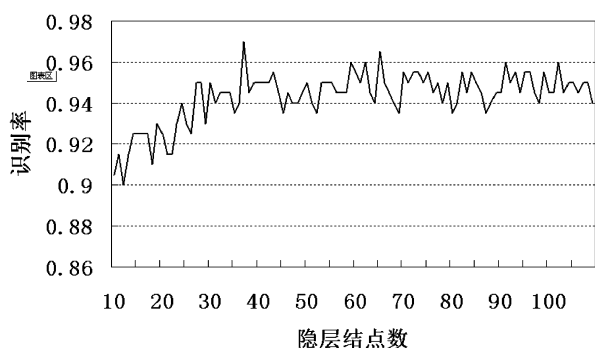


图2 不同隐结点数对识别率的影响

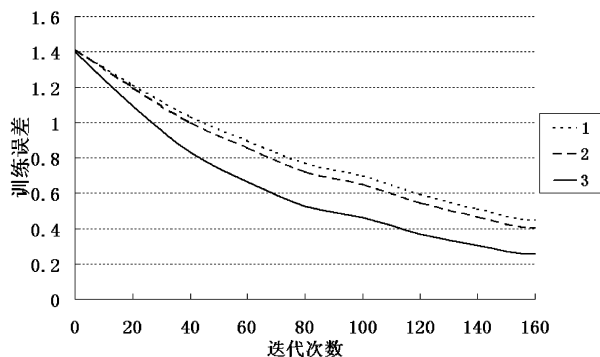


图3 不同动量因子对训练误差的影响

3.3 不同动量因子对训练过程的影响

图3给出隐含层节点数和学习速率在合适的取值时,动量因子对学习过程的影响。其中,隐含层节点数为40,线1为未加入动量因子的情况,线2的动量因子为0.05,线3的动量因子为0.15。从中可以看出图中动量因子为0.15的误差曲线,明显地能较好地避开局部最优问题,在训练的后期,能够降到误差更小的极小点。

3.4 不同参数对识别结果的影响

表1给出算法改进前后图像识别率的比较,结果显示通过优化参数改进的算法的识别效果优于原有

算法。

表1 改进前后数字识别率比较

数字	经典BP算法	文献[4]中算法	本文算法
0	88.2%	94.1%	100%
1	100%	100%	100%
2	75.0%	87.5%	93.8%
3	87.3%	93.8%	93.8%
4	92.9%	96.4%	96.4%
5	100%	95.0%	95.0%
6	95%	95.0%	100%
7	95.8%	95.8%	95.8%
8	100%	100%	100%
9	80.9%	81.0%	85.7%
平均	92%	93.9%	96.1%

4 结束语

BP算法始终是一个非线性梯度优化问题,存在局部最小误差。文章通过优化网络参数和附加动量等方法改进了现有的BP神经网络,改进后的网络能够有效地避免局部极小值,提高网络的识别效果。实验结果也证明了此改进算法的实用价值。

参考文献:

- [1] 边肇祺,等. 模式识别[M]. 北京:清华大学出版社, 1988.
- [2] 王永乾,吕蓉. 基于BP网络的手写体数字识别方法[J]. 山东电子, 2004(3): 24225.
- [3] 柳回春,马树元,吴平东,李晓梅. 手写体数字识别技术的研究[J]. 计算机工程, 2003, 29(4): 24225.
- [4] 周敬利,等. 基于BP神经网络的人脸检测算法[J]. 计算机工程, 2004, 30(11): 34236.
- [5] Henry A Rowley, Shumeet Baluja, Takeo Kanade. Neural network based face detection[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1998, 20(1): 23238.
- [6] 张伟,王克俭,秦臻. 基于神经网络的数字识别的研究[J]. 微电子学与计算机, 2006, 23(8): 2062208.
- [7] 樊亚军,曲仕茹. 利用神经网络实现三维飞机目标识别[J]. 郑州大学学报, 2004, 25(4): 56259.
- [8] 徐家树,覃征,杨盾. 基于BP神经网络的Web页面分类算法[J]. 微电子学与计算机, 2006, 23(5): 83285.
- [9] 马义德,齐春亮,杜鸿飞. 一种基于分类的改进BP神经网络图像压缩方法[J]. 兰州大学学报(自然科学版), 2005, 41(4): 70272.
- [10] 田旭光,宋彤,刘宇新. 结合遗传算法优化BP神经网络的结构和参数[J]. 计算机应用与软件, 2004, 21(6): 69271.
- [11] 陈为龙,李晓宁. 一种基于总量风险函数的改进BP算法[J]. 四川大学学报(自然科学版), 2006, 43(5): 102321026.
- [12] 乔双,董智红. BP网络初始权值的选取方法[J]. 东北师大学报(自然科学版), 2004, 36(3): 25230.