基于 BP 神经网络的脱机手写数字识别

Off - line Handwritten Numeral Recognition Using Topological Feature Structures

杨金伟* 段会川

YANG Jin - wei DUAN Hui - chuan

摘 要 脱机手写体数字识别有着重大的使用价值,特征提取占据了重要的位置,本文针对手写体数字识别中单一识别方法的局限性,提出采用 BP 神经网络进行识别,并且提出了一种特征提取方法。采用 BP 神经网络,利用其良好的监督学习功能进行识别,结合提取的降维数字符号的特征,能较好的识别出手写数学符号。BP 神经网络(Back - Propagation),又称误差反向传递神经网络,是一种依靠反馈值来不断调整节点之间的连接权值而构建的一种网络模型。最后,在 Mnist 手写数据库上的试验结果表明,该方法具有较好的识别率和较高的可靠性。

关键词 预处理 BP 神经网络 手写数字识别

Abstract Off - line handwritten numeral recognition has great value, in which feature extraction occupy an important position. This paper uses BP neural network to recognize handwritten numerals against the limitations of a single recognition method, and put forward a method of feature extraction. BP network has good learning supervision power. With the symbolic features of the extracted dimension - fallen figures, the handwriting numerals can be well recognized. BP neural network (Back - Propagation), also known as the reverse error transferred neural network, is a network model of a feedback on the value of continuous adjustment to connected value among the nodes. Finally, the test on the Mnist handwriting database shows that the method has a good recognition rate and high reliability.

Keywords Preprocessing BP neural network Handwritten numeral recognition

引言

手写体数字识别在邮政编码自动识别、银行业务方面有重要的应用,但是由于字体变化大,对识别率要求高,因此有较大的困难。字符识别一般分为两类[1]:联机手写数字识别和脱机手写数字识别。目前,在脱机手写字符识别研究中使用最广泛的是光学字符识别,即 OCR 方法。其统计模式识别方法注重数量特征,便于特征提取、分析和计算。但是,它将字符看成一种随机的二维点阵,没有考虑字符的结构特征和结构信息。因此,这种方法对单一字符比较有效,而对不同字体的字符识别效果则较差。

结构模式识别[2]的方法是把待识别的模式看成由若干个比较简单的子模式构成的集合,任何模式都可以用一组基元及一定的组合关系来描述。由于字符含有丰富的结构信息,可以设法提取含有这种信息的结构特征,作为字符识别的依据。但是,由于字符结构比较复杂,实际应用中仍有较大困难。近年来,出现了将统计和结构识别结合起来的途径,既吸取了统计识别的优点,又利用了字符的结构信息。手写数字识别是字符识别的一个特定方向。由于问题本身的特殊性,传统的 OCR 方法不能有效地解决这一问题。因此,手写数字的识别方法应该是一种具有自适应、抗干扰、能够有效地解决手写数字分割、统计模式识别与结构模式识别相结合的方法。神经网络技术的运用能够部分解决上述问题。神经网络具有以下几方面的优点:神经网络是自适应的,它能从数据中自动地学习到解决问题的知识。本文利用了 Mnist 手写体数字数据库,识别率达到 93% 以上。

* 山东师范大学信息科学与工程学院 济南 250014

1 预处理

图像预处理是字符识别重要的一环,它把原始的图像转换成识别器所能接受的二进制形式。要识别手写体数字首先要对其字符图像进行预处理。预处理的目的是去除字符图像中的噪声、压缩冗余信息,得到规范化的点阵,为特征提取做好准备。由于数据量比较大,需要批量读入字符图像。求取其能包络最大像素区域的 BoundingBox,即去掉外边缘和噪声等无用信息。对 BoundingBox 区域进行细化,进行断笔填充。使得预处理后的图像成为一个无间断点的数字。

2 特征提取

手写数字特征的提取和选择是一项极为重要的工作。特征的选择是否恰当,提取的方法是否有效直接关系到最终的识别结果。研究者们提出了许许多多的识别方法,按使用特征的不同,这些方法可以分为两类[3]:基于结构特征的方法和基于统计特征的方法。统计特征通常包括点密度的测量、矩、特征区域等等;结构特征通常包括圈、端点、交叉点、笔画、轮廓等等。一般来说,两类特征各有优势。例如,使用统计特征的分类器易于训练,而且

对于使用统计特征的分类器，在给定的训练集上能得到相对较高的识别率；而结构特征的主要优点之一是描述字符的结构，在识别过程中能有效地结合几何和结构的知识，因此能得到可靠性较高的识别结果。

本文采用预处理后的图像，将字符图像内的像素分成 $7\times7=49$ 个小矩形，对每一个矩形里的像素求和。即将原来的 $28\times28=784$ 维降到 $7\times7=49$ 的点阵。由于 49 维的特征数量较大，会大大增加运算时间，因而采用垂直投影法将其压缩为 7 维特征再输入 BP 神经网络。试验表明，该方法简单而且识别率高。

3 BP 神经网络结构设计

本文采用的 BP 神经网络结构有三层：输入层、隐含层、输出层。将对字符图像提取每一个数字的 7 维的特征值作为神经网络的输入，因此输入节点为 7 个。由于隐层神经元的数目很难确定，通过反复实验，采用 10 个神经元的学习速度和准确度比较高。输出层有 10 个节点，最大匹配的数字序号即为得到的输出值。

4 BP 网络结构及学习规则

4.1 BP 网络结构

BP 神经网络(Back - Propagation)[2]，又称误差反向传递神经网络。它是人工神经网络(ANN)中的一种模型，是利用率很高的一种神经网络，有 80% ~90% 的神经网络采用了 BP 神经网络或者它的变化形式。BP 网络是前向网络的核心部分，体现了神经网络中最精华的内容。BP 神经网络是一种依靠反馈值来不断调整节点之间的连接权值而构建的一种网络模型。它的整个体系结构(如下图 1)所示，分为输入层、隐藏层和输出层，其中隐藏层根据具体情况的需要，可以是一层结构也可为多层结构。上下层之间实现全连接，而每层神经元之间无连接。当一对学习样本提供给网络后，神经元的激活值从输入层经各中间层向输出层传播，在输出层的各神经元获得网络的输入响应。按照减少目标输出与实际误差的方向，从输出层经过各中间层逐层修正各连接权值，最后回到输入层，这种算法称为“误差逆传播算法”，即 BP 算法。随着这种误差的传播修正不断进行，网络对输入模式响应的正确率也不断上升。

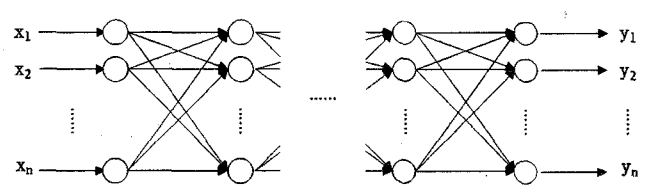

图 1 BP 网络结构

4.2 BP 网络学习规则

为方便阐述，首先对各符号的形式及意义进行说明。

网络输入向量 $P_k=(a_1,a_2,\cdots,a_n)$；

网络目标向量 $T_k=(y_1,y_2,\cdots,y_q)$；

中间层单元输入向量 $S_k=(s_1,s_2,\cdots,s_p)$；输出向量 $B_k=(b_1,b_2,\cdots,b_p)$；

输出层单元输入向量 $L_k=(l_1,l_2,\cdots,l_q)$；输出向量 $C_k=(c_1,c_2,\cdots,c_q)$；

输入层至中间层的连接权值 w_{ij}，$i=1,2,\cdots,n$，$j=1,2,\cdots,p$；

中间层至输出层的连接权值 v_{jt}，$j=1,2,\cdots,p$，$t=1,2,\cdots,p$；

中间层各单元的输出阈值 θ_j，$j=1,2,\cdots,p$；

输出层各单元的输出阈值 γ_t，$t=1,2,\cdots,p$；

参数 $k=1,2,\cdots,m$。

(1)初始化。给每个连接权值 w_{ij}、v_{jt} 阈值，θ_j 与 γ_t 赋予区间 $(-1,1)$ 内的随机值。

(2)随机选取一组输入和目标样本 $P_k=(a_1^k,a_2^k,\cdots,a_n^k)$、$T_k=(S_1^k,S_2^k,\cdots,S_p^k)$ 提供给网络。

(3)用输入样本 $P_k=(a_1^k,a_2^k,\cdots,a_n^k)$、连接权值 w_{ij} 和阈值 θ_j 计算中间层各单元的输入 s_j，然后用 s_j 通过传递函数计算中间层各单元的输出 b_j。

$$s_j=\sum_{i=1}^{n}w_{ij}\alpha_i-\theta_j \quad j=1,2,\cdots,p$$

$$b_j=f(s_j)\quad j=1,2,\cdots,p$$

(4)利用中间层的输出 b_j、连接权 v_{jt} 和阈值 γ_t 计算输出层各单元的输出 L_t，然后利用 L_t 通过传递函数计算输出层各单元的响应 C_t。

$$L_t=\sum_{j=1}^{p}v_{jt}b_j-\gamma_t \quad j=1,2,\cdots,q$$

$$C_t=f(L_t)\quad t=1,2,\cdots,q$$

(5)利用网络目标向量 $T_k=(y_1^k,y_2^k,\cdots,y_q^k)$，网络的实际输出 C_t，计算输出层的各单元一般化误差 d_k^t。

$$d_k^t=(y_t^k-C_t)\cdot C_t(1-C_t)\quad t=1,2,\cdots,q$$

(6)利用连接权 v_{jt}、输出层的一般化误差 d_t 和中间层的输出 b_j 计算中间层各单元的一般化误差 e_j^k。

$$e_j^k=[\sum_{t=1}^{q}d_tv_{jt}]b_j(1-b_j)$$

(7)利用输出层各单元的一般化误差 d_k^t 与中间层各单元的输出 b_j 来修正连接权 v_{jt} 和 γ_t。

$$v_{jt}(N+1)=v_{jt}(N)+\alpha d_t^k b_j$$

$$\gamma_t(N+1)=\gamma_t(N)+\alpha d_t^k\quad t=1,2,\cdots,q,0<\alpha<1$$

(8)利用中间层各单元的一般化误差 e_j^k，输入层各单元的输入 $P_k=(a_1,a_2,\cdots,a_n)$ 来修正连接权 w_{ij} 和阈值 θ_j。

$$w_{ij}(N+1)=w_{ij}(N)+\beta e_j^k a_i^k$$

$$\theta_j(N+1)=\theta_j(N)+\beta e_j^k\quad i=1,2,\cdots,n,j=1,2,\cdots,p,0<\beta<1$$

(9)随机选取下一个学习样本向量提供给网络，返回到步骤(3)，直到 m 个训练样本训练完毕。

(10)重新从 m 个学习样本中随机选取一组输入和目标样本，返回步骤(3)，直到网络全局误差 E 小于预先设定的一个极小值，即网络收敛。如果学习次数大于预先设定的值，网络就无法收敛。

(11)学习结束。

其中，(7) ~ (8)步为网络误差的“逆传播过程”，(9) ~ (10)步用于完成训练和收敛过程。

(下转第 54 页)

容易通过阴性选择,其中可以发现作为基因提取对象的记忆检测元集对 MosCS 算法有一定的影响,集合越大,效果越好。

4.2 检测异常模式的效果

计算一定时间内两种方法产生的成熟检测元检测到异常模式数量的均值,设:

MosCS 算法均值 Vcs = 成熟检测元检测到的 NP(NPcs)总数/成熟检测元(cs)总数

随机算法均值 Vr = 成熟检测元检测到的 NP(NPr)总数/成熟检测元(r)总数

参数:检测元数量上限 = 500,变异度 = 0.50,时间 Aug 25,2005—Aug 29,2005

MosCS 算法	202.113.76.82	202.113.76.83	202.113.76.99
NPcs	2633	2880	2675
cs	12502	13783	13196
Vcs %	21.06	20.90	20.27

参数:检测元数量上限 = 500,时间 Aug 25,2005—Aug 29,2005

MosCS 算法	202.113.76.82	202.113.76.83	202.113.76.99
NPr	960	1573	1237
r	9887	10675	10869
Vr %	9.71	14.74	11.38

由表中数据可知 Vcs > Vr,说明 MosCS 算法产生的检测元绑定异常行为的效果更好。

5 结论

通过 MosCS 算法的分析和实验验证,MosCS 相对现有的检测元生成算法在基因库的基因最优化方面有一定优势,MosCS 产生的检测元更容易通过阴性选择,生成有效检测元的效率更高,同时可以显著提高检测覆盖率。

参考文献:

[1] 大众医药网. 医学免疫学[EB/OL]. http://www.windrug.com/book/book22.php.

[2] HOFMEYR S A, FORREST S. Architecture for an artificial immune system [J]. Evolutionary Computation, 2000, 7(1): 45 - 68.

[3] HOFMEYR, S. An Immunological Model of Distributed Detection and Its Application to Computer Security [D], PhD Thesis, Dept of Computer Science, University of New Mexico, 1999.

[4] KIM, J., BENTLEY, P. J. Towards an Artificial Immune System for Network Intrusion Detection: An Investigation of Dynamic Clonal Selection [J], the Congress on Evolutionary Computation (CEC - 2002), Honolulu, pp. 1015 - 1020, May 12 - 17, 2002.

[5] KIM, J., BENTLEY, P. J. A Model of Gene Library Evolution in the Dynamic Clonal Selection[C]. Proceedings of the First International Conference on Artificial Immune Systems (ICARIS) Canterbury. 2002. 175 - 182.

[6] DE CASTRO, L. N. VON ZUBEN, F. J. Learning and Optimization Using the Clonal Selection Principle [A]. IEEE Transactions on Evolutionary Computation, Special Issue on Artificial Immune System. 2001.6(3): 239 - 251.

[作者简介] 王斌(1979 ~),天津人,助教,硕士,主要研究方向:网络安全、人工智能。

(收稿日期:2008 - 01 - 09)

(上接第 50 页)

5 网络训练及识别结果

本文实验采用 Mnist 手写数据库,随机挑选了 2000 个训练样本,500 个测试样本。实验使用 matlab7.4 标准完成。值得说明的是,在训练集上做到的识别率不是 100%,这在大样本库上是允许的。测试结果如表 1 所示:

表 1

	识别率	误识率	拒识率
训练集	99.6%	0.3%	0.1%
测试集	93.4%	6.1%	0.50%

6 总结

BP 神经网络模型在模式识别方面得到了较广泛的发展。本文在经典 BP 网络结构的基础上,根据问题的解决需要,通过改变输入层、隐含层及输出层的网络结点数,用已知样本库进行训练,修改输入层、隐含层及输出层间的权值,从而使整个网络达到稳定。实验表明,能较好地识别手写的数字符号。由于 BP 网络本身固有的缺点,难以找到全局收敛点,所以还需进行改进,如:结合模拟退火算法或遗传算法,可以适当改变其收敛特性,这将在以后继续加以研究。

参考文献:

[1] Cao J, Shridhar M. A Hierarchical Neural Network Architecture for Handwritten Numerral Recognition [J]. Pattern Recognition, 1997,30(2): 289 - 294.

[2] 边肇祺, 张学工. 模式识别[M]. 北京:清华大学出版社, 1999:250 - 257.

[3] 韩宏,杨静宇. 神经网络分类器的组合(J). 计算机研究与发展. 2000, (12):1488 - 1492.

[4] 孙志强,葛哲学. 神经网络理论与 MATLAB7 实现[M] 北京:电子工业出版社,2005:99 - 102.

[作者简介] 杨金伟(1982 ~),女,在读硕士研究生,主要研究方向为数字图像处理;段会川(1967 ~),男,教授,硕士生导师,主要研究方向是数字图像处理、模式识别。

(收稿日期:2008 - 03 - 14)