

基于人工神经网络的数字识别系统的研究

王建雄, 刘应龙

(中南大学 信息科学与工程学院, 湖南 长沙 410083)

摘 要:在社会生活各领域中,存在着含有大量数字信息的文件,实现计算机对数字的自动识别已经成为当今社会的一种需求。基于此,文中提出了一种神经网络的方法来实现数字识别。经实验结果表明,该系统取得了较好的识别精度。

关键词:人工神经网络;特征提取;BP网络

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2006)05-0026-02

Digital Recognition System Based on Artificial Neural Network

WANG Jian-xiong, LIU Ying-long

(Information Science and Engineering College, Central South University, Changsha 410083, China)

Abstract: There is much information which includes the massive digit in the world. It is essential to recognize digit automatically by computer in current society. So, the paper presents a system based on artificial neural network to realize digital diagnosis. The experiment shows that the system has obtained the good precision.

Key words: artificial neural network; characteristics extraction; BP network

0 引言

为了解决计算机字符的自动识别、输入问题,并进行高速加工处理,使计算机达到真正智能化,人们对计算机的自动识别进行了多年的研究,并取得了很大的进步。数字识别一般通过特征匹配及特征判别的方法来进行处理,前者一般适用于规范化的印刷体字符识别;后者多用于手写字符识别,其研究还处于探索阶段,识别率还比较低。因此,为了提高识别率,就必须寻求新的方法和途径。

近年来,人工神经网络技术取得了巨大的发展,它具有分布式存储信息、并行处理信息、自组织、自学习信息等显著优点,为模式识别开辟了新的途径。文中所介绍的系统正是借助人工神经网络基本原理并结合实际实现的。

1 输入及预处理

1.1 输入环节

通过扫描仪将数字扫描输入到计算机内部,并以点阵形式保存为数据文件,供预处理程序随时调用和处理。该输入环节对扫描所得的点阵大小并无限制,因它要由预处理环节进行归一化处理,最终变成一种规格化的点阵。

1.2 预处理环节

采用了许多图像处理技术,最后把每个数字的特征提

取出来。这些技术包括图像二值化、噪声点的去除、数字分离、图像平滑、归一化处理、特征提取。图像二值化、噪声点的去除、数字分离、图像平滑^[1,2]都是采用数字图像处理中一般常规方法,即可实现。

归一化处理部分是把字符压缩成固定的 16×16 点阵形式。因为扫描进来的图像中字符大小存在较大的差异,而相对来说,统一尺寸的字符识别的标准性更强,准确率也更高。具体算法如下:先得到原来字符的高度,并与系统要求的高度做比较,得到变换系数,然后根据得到的系数求得变换后应有的宽度。在得到宽度、高度后,把新图像中的点按插值的方法映射到原图像中。

最后一部分是特征提取。经过上述处理,原来大小不同,分布不规则的各个字符变成了大小相同,排列整齐的字符。下面就要从被分割归一化处理完毕的字符中,提取最能体现其特点的特征向量。将从训练样本中提取出来的特征向量代入BP(Back Propagation)网络中就可以对网络进行训练,提取出待识别的样本中的特征向量代入BP训练好的BP网络中,就可以对字符进行识别。具体算法如下:对图像进行逐行逐列的扫描,当遇到黑色像素时取其特征值为1;遇到白色像素时取其特征值为0,这样当扫描结束时就形成了一个维数与图像中像素点个数相同的特征向量矩阵。

2 数字识别

该环节为系统的核心环节,采用误差反向传播神经网络模型,简称BP神经网络^[3~6]。它使用的是梯度下降

收稿日期:2005-09-08

作者简介:王建雄(1979-),男,江西萍乡人,硕士研究生,研究方向为计算机工业视觉监测;刘应龙,副教授,研究方向为计算机工业视觉监测。

法,作用函数采用可微的S型函数。

2.1 网络结构

本系统采用的是三层BP网络,网络由输入层、隐层和输出层组成。其结构如图1所示。

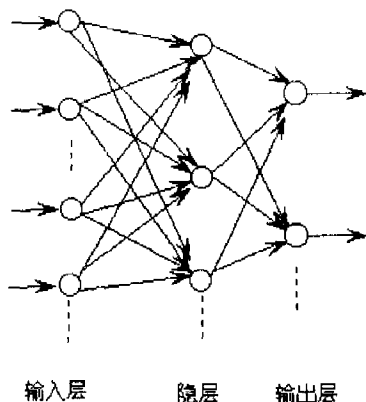


图1 BP神经网络

2.2 BP学习算法

此算法由正向传播和反向传播组成。正向传播是输入信号从输入层经隐层传向输出层,若输出层得到了期望的输出,则学习算法结束;否则,转向反向传播。反向传播就是将误差信号(样本输出与网络输出之差)按原联结通路反向计算,由梯度下降法调整各神经元的权值和阈值,使误差信号减小。

算法步骤:

- 1) 设置初始权系 $W(0)$ 为较小的随机非零值。
- 2) 给定输入/输出样本对,计算网络的输出。

设第 p 组样本输入、输出分别为 $u_p = (u_{1p}, u_{2p}, \dots, u_{np})$; $d_p = (d_{1p}, d_{2p}, \dots, d_{kp})$, $p = 1, 2, \dots, L$ 。节点 i 在第 p 组样本输入时,输出为:

$$y_{ip}(t) = f[x_{ip}(t)] = f[\sum_j w_{ij}(t) I_{jp}]$$

式中 I_{jp} 为在第 p 组样本输入时,节点 i 的第 j 个输入。 $f(\cdot)$ 为S型函数,可由输入层经隐层至输出层,求得网络输出层节点的输出。

- 3) 计算网络的目标函数 J 。

$$E_p(t) = \frac{1}{2} \|d_p - y_p(t)\|_2^2 = \frac{1}{2} \sum_k [d_{kp} - y_{kp}(t)]^2 = \frac{1}{2} \sum_k e_{kp}^2(t)$$

式中, $y_{kp}(t)$ 为在第 p 组样本输入时,经 t 次权值调整,网络的输出, k 是输出层第 k 个节点。

网络的总目标函数为: $J(t) = \sum_p E_p(t)$, 作为对网络学习状况的评价。

- 4) 判别。

若 $J(t) \leq \epsilon$ (ϵ 为预先确定的, $\epsilon > 0$) 则算法结束;否则转至步骤5。

- 5) 反向传播计算。

由输出层,依据 J 按“梯度下降法”反向计算,逐层调整权值。

$$w_{ij}(t+1) = w_{ij}(t) - \eta \frac{\partial J(t)}{\partial w_{ij}(t)} = w_{ij}(t) - \eta \sum_p \frac{\partial E_p(t)}{\partial w_{ij}(t)}$$

$$\frac{\partial E_p(t)}{\partial w_{ij}(t)} = w_{ij}(t) + \Delta w_{ij}(t)$$

其中 η 为步长。具体算法如下:

$$\frac{\partial E_p}{\partial w_{ij}} = \frac{\partial E_p}{\partial x_{ip}} \frac{\partial x_{ip}}{\partial w_{ij}}$$

设 $\delta_{ip} = \frac{\partial E_p}{\partial x_{ip}}$ 为第 i 个节点的状态 x_{ip} 对 E_p 的灵敏度,

由此可得 $\frac{\partial E_p}{\partial w_{ij}} = \delta_{ip} I_{jp}$ 。

当 i 为输出节点,即 $i = k$ 时,

$$\delta_{ip} = \delta_{kp} = \frac{\partial E_p}{\partial x_{kp}} = \frac{\partial E_p}{\partial y_{kp}} \frac{\partial y_{kp}}{\partial x_{kp}} = -e_{kp} f'(x_{kp});$$

当 i 不是输出节点,即 $i \neq k$ 时,

$$\delta_{ip} = \frac{\partial E_p}{\partial x_{ip}} = \frac{\partial E_p}{\partial y_{ip}} \frac{\partial y_{ip}}{\partial x_{ip}} = \frac{\partial E_p}{\partial y_{ip}} f'(x_{ip})$$

其中 $\frac{\partial E_p}{\partial y_{ip}} = \sum_m \frac{\partial E_p}{\partial x_{mp}} \frac{\partial x_{mp}}{\partial y_{ip}} = \sum_m \frac{\partial E_p}{\partial x_{mp}} \frac{\partial}{\partial y_{ip}} \sum_j w_{mj} I_{jp}^* =$

$$\sum_m \frac{\partial E_p}{\partial x_{mp}} w_{mi} = \sum_m \delta_{mp} w_{mi}$$

为使学习速率足够大,又不易产生振荡,在权值调整算法中加入平滑因子 β :

$$w_{ij}(t+1) = w_{ij}(t) - \eta \frac{\partial J(t)}{\partial w_{ij}(t)} + \beta [w_{ij}(t) - w_{ij}(t-1)]$$

本系统中采用步长 $\eta = 0.8$, 平滑因子 $\beta = 0.4$ 。

2.3 BP网络设计

BP网络应用过程如图2所示。

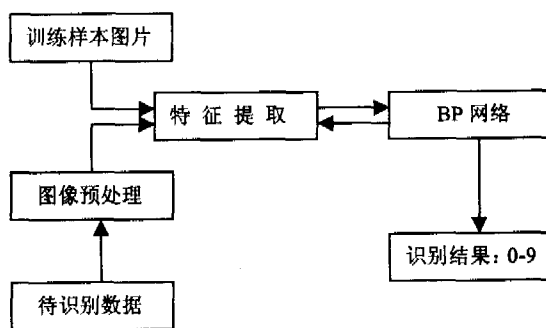


图2 BP网络应用流程图

BP网络应用的第一步就是要用已知训练样本来训练BP网络。这里BP网络的输入层的节点个数为图像预处理后所输出的特征维数。本系统采用 16×16 归一化,特征提取采用的是直接利用每个点的像素值来作为特征,这样,对于每一个输入样本就有 $16 \times 16 = 256$ 个特征。因此输入层的节点数为256。

对于隐层的节点数,一般来说,隐层神经元的数目越多,那么BP网络也就越精确,训练时间也越长。本系统采用了20个隐层神经元。

对于输出层节点数的确定,这取决于如何设定标准输出。这里采用了8421BCD码来对0到9进行编码。这样

(下转第30页)

④保留最优分类器的结构,利用整个的数据集进行参数学习(条件概率表)。

⑤输出这个新的分类器。

数据集规模不大的情况下,选择用 LOO 验证法。LOO 充分利用训练样本,是最严格、最精确的一种评估方法之一,但是它适用于小规模数据集,否则增加机器运行时间。数据集规模很大的情况下,选择保留(Holdout)方法来测试分类器的性能。保留方法中一般采用数据集的 2/3 作为训练集,数据集的 1/3 作为测试集。

利用保留方法测试小数据集的分类器性能会降低分类器的分类精度;而利用 LOO 方法测试大数据集的分类器,会增大机器的运行时间。因此不同的数据规模下选择合理的检验方法既保证分类器分类精度,又可以减少机器运行时间。

4 实验结果

标准数据集是从 UCI^[5]上下载的,选取了 8 个数据集。有关数据集的概况见表 1。所选择的数据集都是小规模的,所以本实验只针对评估方法是 LOO 情况进行。在实验中不同的数据集是在相同的环境下运行的。

表 1 算法的实验结果

Dataset	Attributes	Classes	Instances	HCS	SP	HCS&SP
Vehicle	18	4	846	70.1	70.3	70.3
Post-op	9	3	90	72.7	72.1	72.7
Australia	14	2	690	84.7	85.3	85.3
Hepatitis	19	2	270	84.8	84.3	84.8
Vote	16	2	435	95.6	95.7	95.7
Heart	13	2	270	78.7	76.1	78.7
Soybean - Large	35	19	562	88.6	88.4	88.6
Pima	8	2	768	78.0	78.2	78.2

表 1 列出了实验结果,实验结果基本和文献[4]中的数据差不多,第七列是改进算法的实验结果,每次输出的分类器确实是 HCS 和 SP 中性能最好的,验证了所建立的

分类器的有效性和正确性。同时也发现当属性变量之间的依赖关系不是很复杂的情况下,分类器的分类精度会很高。例如数据集 Vote,说明文中讨论的分类器在对 Vote 操作时几乎把影响分类准确率的弧都加上了,丢失有价值的弧的个数相应减少。当属性间的关系很复杂时,由于分类器属性之间构成的是有向森林,有可能丢失一些对分类有价值的弧,如 Vehicle。再一次表明了属性结点之间的条件依赖关系不是很复杂的情况下,HCS 和 SP 算法的分类准确性非常高,改进的算法的分类准确性更高。

5 结束语

通过对 HCS 算法和 SP 算法的分析,又提出了一个改进的算法,明显提高了分类器的分类精度。扩展朴素贝叶斯分类器既有朴素贝叶斯分类器的简单性,又有很好的分类能力,应努力把它应用到更多的现实领域中。当属性结点之间的相关性很复杂的情况,即有简单性又有很好的分类性能的分类器待于进一步的研究。

参考文献:

- [1] 林士敏,田凤占,陆玉昌.用于数据挖掘的贝叶斯分类器研究[J].计算机科学,2000,27(10):73-76.
- [2] Friedman N, Geiger D, Goldszmidt M. Bayesian Network Classifiers[J]. Machine Learning, 1997(29):131-163.
- [3] Cheng J, Greiner R. Comparing Bayesian network classifiers [A]. Proceedings of the fifteenth conference on uncertainty in artificial intelligence [C]. San Francisco: Morgan Kaufmann, 1999. 101-108.
- [4] Eamonn J, Pazzani J. Learning augmented bayesian classifiers: a comparison of distribution-based and classification-based approaches [A]. Proceedings of the Seventh International Workshop on Artificial Intelligence and Statistics [C]. Lauderdale: [s. n.], 1999. 225-230.
- [5] Blake C, Keogh E, Merz C. UCI repository of machine learning database [EB/OL]. <http://www.ics.uci.edu/mllearn/MLRepository.html>, 1998.

(上接第 27 页)

就可以确定输出层的神经元数目为 4。但是通过其后的实践发现,当采用(0,0,0,0)这样的目标输出向量时,BP 网络无法收敛,那是因为采用的激发函数永远不可能达到 0 或 1,而只能是接近。所以还要对其重新编码,最后确定其编码为:0(0.1,0.1,0.1,0.1),1(0.1,0.1,0.1,0.9),...,9(0.9,0.1,0.1,0.9)。

3 实验结果

本系统采用了 100 个训练样本用来测试算法的性能,这 100 幅图像包括 Arial, Batang, Gautami 字体,单个数字的平均识别率达到 94%,网络训练时间大约 3~6 秒,当然如果采用更多的训练样本,则识别率将会进一步提高。

参考文献:

- [1] 张宏林,蔡锐. Visual C++ 数字图像识别技术及工程实践[M]. 北京:人民邮电出版社,2003. 422-442.
- [2] 冈萨雷斯. 数字图像处理(第 2 版)[M]. 北京:科学出版社,2003. 60-127.
- [3] Hagan M T. Neural Network Design[M]. 北京:机械工业出版社,2002. 16-64.
- [4] Bruck J, Sanz J. A Study on Neural Networks[J]. Int J Intelligent System, 1988(3):4-15.
- [5] 徐丽娜. 神经网络控制[M]. 哈尔滨:哈尔滨工业大学出版社,1999. 7-16.
- [6] 袁曾任. 人工神经网络及其应用[M]. 北京:清华大学出版社,1999. 16-29.