

基于神经网络的带噪声英文字母和数字识别

牛慧娟, 汪森霖

(广州涉外经济职业技术学院, 广州 510540)

摘要: 应用神经网络的方法, 设计了一个 432-26-36 结构的三层 BP 训练网络, 来识别具有一定噪声的 26 个大写英文字母及 10 个阿拉伯数字, 并考察了网络初始参数对于收敛速度的影响, 通过多次对试验结果的比较选取了最佳隐层节点个数。系统用 MatLab 进行仿真实现, 试验结果表明该系统能够对具有一定噪声的大写英文字母和数字进行非常好的识别。

关键词: 英文字母识别; 数字识别; BP 网络; MatLab; 噪声

0 引言

英文字母和数字识别技术是车牌识别过程的一个关键环节。由于实际中获得的图片往往带有一些噪声, 要求所设计的识别系统具有一定抗干扰能力, 使其具有在一定不规范书写的情况下的仍然能够识别字母或数字能力, 这给系统设计实现带来很大困难, 人工神经网络技术的出现和发展为解决这一问题提供了很好的方法和技术。人工神经网络具有一些传统技术所没有的优点, 例如良好的容错能力、较强的分类能力、并行处理能力和自学习能力。因而, 采用神经网络识别方式是一种很好的选择。而其中的 BP 网络 (Back-Propagation Network, 反向传播网络) 是用在模式识别的典型方法。

1 BP 网络简介

BP 网络是一种多层前馈网络, 包括输入层、隐层 (中间层) 和输出层, 隐层可以有一个, 也可以有多个。前、后两层之间实现全连接, 而每层神经元之间无连接。以三层 BP 网络为例, 其网络拓扑结构如图 1 所示。

BP 网络因采用 BP 学习算法而得名。BP 算法工作过程如下:

(1) 正向传播: 输入的样本从输入层经过隐层处理后传向输出层; 在逐层处理的过程中, 每一层神经元的状态只对下一层神经元的状态产生影响。在输出层把输出值和期望值进行比较, 如果输出值不等于期

望值, 则进入反向传播过程。在正向传播过程中, 网络的权值不变;

(2) 反向传播: 根据输出值与期望值的误差, 把误差信号按原来正向传播的通路反向传回, 并对每个隐层的各个神经元的权值进行修改, 以使误差信号趋于最小。

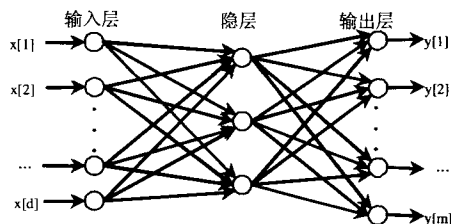


图 1 三层 BP 网络拓扑结构图

2 系统所用 BP 网络结构的设计

BP 网络的设计一般要考虑这几个问题: ①网络的层数; ②每层的神经元数; ③每层的激活函数; ④初值的选择。

2.1 网络层数的选择

理论上已经证明: 只要在输入输出线性神经元之间, 增加采用 S 型函数的隐层神经元, 所构成的多层前馈网络可以实现任意连续函数的逼近。也就是说 BP 网络解决非线性问题首选的模型为具有一个隐含层的三层网络结构。其拓扑结构如前面的图 1 所示。如果试验中发现网络的训练精度无法达到预期要求,

收稿日期: 2008-06-25 修稿日期: 2008-10-03

作者简介: 牛慧娟 (1982-), 女, 河北任丘人, 助教, 本科, 研究方向为多媒体技术

可以考虑适当增加隐层神经元的个数,这样修改之后的网络的训练效果要比增加层数更容易观察和调整。

2.2 每层神经元数

(1) 输入层节点个数

由于 BP 网络不能直接处理图片信息,需要将要识别的图像进行预处理。将含有英文字母和数字的部分用点阵画出其图形,并将其数字化为 24×18 的 0、1 矩阵,例如字母 A 的点阵如图 2。

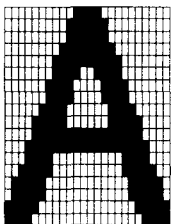


图 2 大写字母 A 的 24×18 点阵表示

由此 A 的 0、1 矩阵再 MatLab 中可以表示为一个 432×1 的矩阵,用同样的方法来定义其余字母以及数字,最终得到样本矩阵为:

alphabetnum = [letterA, letterB, ..., letterZ, Number0, Number1, ..., Number9];

输入端为其中的每一个列向量,因此输入端的节点个数应为 432 个。

(2) 输出层节点个数

由于最终希望系统可以识别出 26 个大写英文字母及 10 个数字,网络的输出端可以采用一个 36×36 的单位矩阵,当输入一个字母或数字时,矩阵按 alphabetnum 中给定顺序在其相应位置上输出为 1,其他位置输出 0。所以输出层的节点选定为 36 个。

(3) 隐层节点个数

关于隐层神经元个数的选择至今没有很好的理论指导,需要在试验中反复比较,对比不同神经元个数的训练结果,选择最优的节点个数。以下公式(1)可以作为选择隐层节点个数的参考:

$$n_1 = \sqrt{n+m} + a \quad (1)$$

其中, m 为输出层节点个数, n 为输入层节点个数, a 为 $[1, 10]$ 之间的常数。

根据以上公式和输入、输出层节点个数可知隐层节点大致可以设置为 21~31 之间的数,由于隐层的节点个数并不是越多越好,隐层节点数目太多会导致学习时间过长,误差也不一定最佳,还可能会导致网络容错性差、不能识别以前没有看到的样本等问题;而

且考虑到要防止有不发生作用的节点影响网络训练效果,所以通过多次修改网络隐层节点个数进行试验,并记录其训练循环次数,结果如表 1,其中 $S1$ 为隐层节点个数, Epochs 为训练循环次数。

表 1 修改 $S1=15, 16, \dots, 34$ 时的训练结果

S1	Epochs	S1	Epochs	S1	Epochs	S1	Epochs
15	438	20	258	25	272	30	257
16	276	21	274	26	249	31	276
17	272	22	276	27	274	32	272
18	271	23	274	28	269	33	271
19	260	24	262	29	266	34	265

通过表 1 可以观察出,除了当 $S1$ 为 15 时,网络训练需要循环次数比较多之外,当隐层节点为表中其他值时网络的训练循环次数并没有较大的变化。

为了选择最优节点个数,试验同时还记录下了 $S1$ 在 16~34 之间的误差曲线图,由于篇幅限制,只列出其中 4 个结果。

如图 3 所示,图中实线为有噪声的网络的训练结果,虚线是没有噪声的理想网络训练结果。当 $S1=20$ 时,比 $S1=16$ 的网络识别错误率有明显的降低;当 $S1$ 加大到 26 时网络性能能进一步提高;而当 S 取 30 时,网络的性能反倒略有下降,这也很好地说明了,对于多层 BP 网络来说,并不是隐层节点越多越好的问题。根据以上试验结果,本系统的隐层节点数最终确定为 26 个。

2.3 每层的激活函数

选择采用两层对数 S 型函数作为网络激活函数。在 MatLab 中已经将其编写成函数的形式,可以调用函数 logsig.m。这里它的输出计算公式如公式(2)。

$$A = \text{logsig}(W * P, B) \quad (2)$$

2.4 系统最终结构

根据以上分析,系统的结构设计为:只有一个隐含层的三层 BP 网络,432 个输入节点、26 个隐层节点及 36 个输出节点,为 432-26-36 结构。采用 (0,1) 范围的对数 S 型激活函数两层 logsig/logsig 网络。系统最终实现所用 BP 网络结构模型如图 4 所示。

2.5 BP 网络的初值设定

(1) 初始权值

由于网络是非线性的,初始权值对于学习是否达到局部最小、是否能够收敛以及训练时间长短有很大关系。初始值过大,可能会使得加权后的输入落在激活函数的饱和区,误差变化率将会非常小,进而使得权值几乎不能调整,网络陷入瘫痪;而初值过小,又会

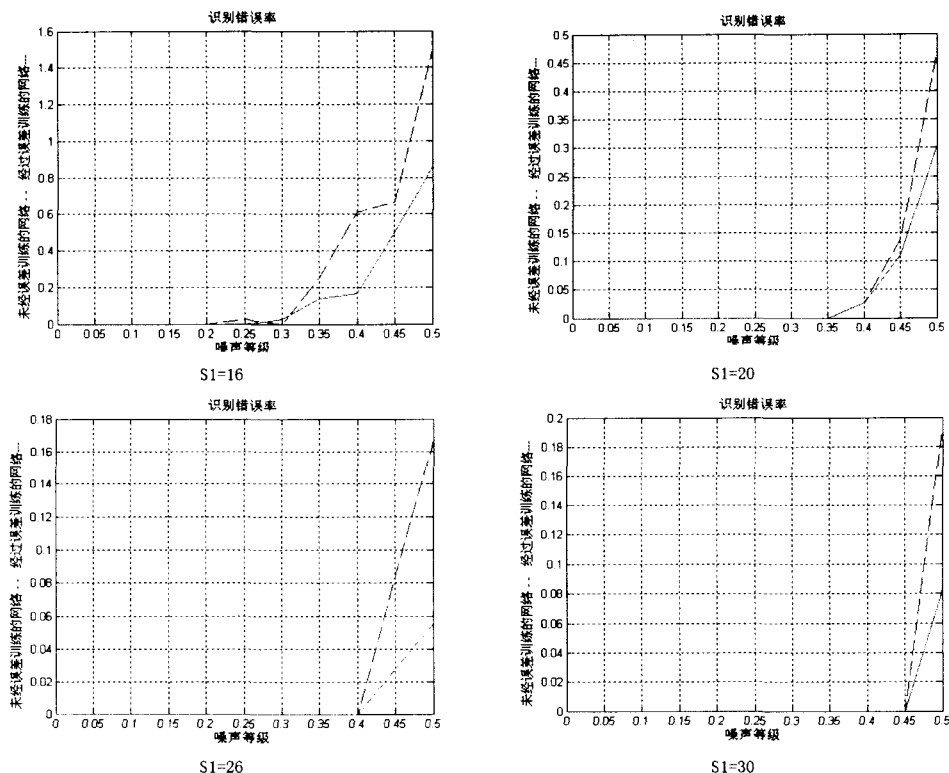


图3 S1=16,20,26,30时的误差曲线

影响学习速度。所以,一般取初始权值在 $(-1,1)$ 之间的随机数,希望经过初始加权后的每个神经元的输出值都接近于零,这样可以保证每个神经元的权值都能够在它们的S型激活函数变化最大之处进行调节。

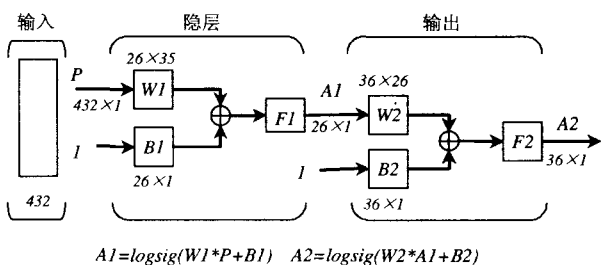


图4 系统实现所采用BP网络结构模型

(2)学习速率

学习速率决定每一次循环训练中所产生的权值的调整量,学习速率过大会导致权值修改过大,从而使系统不稳定;过小可能会导致较长的训练时间,收敛很慢,但是能保证网络的误差值不跳出误差表面的低谷而最终趋于最小误差值,所以一般倾向于选择较小的学习速率以保证系统的稳定性。

一种对于BP网络的改进方法是在网络训练的过程中使学习速率自适应;检查权值的修正值是否真正降低了误差函数,如果确实如此,则说明所选取的学习速率值小了,可以对其增加一个量;如果不是,而产生了过调,那么就应减小学习速率的值,公式(3)给出了一种自适应学习速率的调整公式。

$$\eta(k+1)=\begin{cases} 1.05\eta(k) & SSE(k+1)<SSE(k) \\ 0.7\eta(k) & SSE(k+1)>SSE(k) \\ \eta(k) & \text{其他} \end{cases} \quad (3)$$

学习速率可以自动调整后使得初始学习速率(0)的选取范围可以有很大的随意性。

(3)动量因子

BP算法采用的是梯度下降法,训练是从某一起点沿着误差函数的斜面逐渐达到误差的最小值,但由于很多误差函数为多维空间曲面,如同连绵起伏的山脉,使得在训练过程中,可能使网络陷入某一小谷底,即某一局部极小值,训练可能无法逃出这一局部极小值。解决这一问题的方法是训练网络时采用附加动量法,这种方法使网络在修正权值时,不仅考虑误差在梯度上的作用,而且考虑在误差曲面上变化趋势的影

响,如同一个低通滤波器,它允许网络忽略其上的微小变化,使其有可能滑过一些局部极小值。

MatLab 神经网络工具箱里提供了训练函数 `traindx`, 它的学习速率是自适应的,而且所采用的学习算法就是附加动量法,所以系统的网络训练函数选择 `traindx`。

由此,生成系统 BP 网络的 MatLab 语句定义为:
`net = newff (minmax (alphabetnum), [S1 S2], {'logsig' 'logsig'}, 'traindx');` 其中 S1、S2 分别为隐层节点个数和输出层节点个数。

(4) 期望误差

网络的目标误差 `err_goal` 设定为 0.01。

3 系统流程图

为了使网络具有抗噪声的能力,在训练时除了训练理想样本之外,还人为添加了一些随机的噪声作为训练样本。训练具有噪声的输入矢量是通过在对两组无噪声矢量训练的同时,加上两组带有随机噪声输入矢量来实现的,这主要是为了保持网络同时具有对理想样本的识别能力。网络经过上述训练后,可能会出现对理想无噪声输入辨识错误的情况。为此,需要将上述训练得到的权值作为初始权值,对无噪声理想输入再进行训练。

训练结束后可以通过使用若干个具有 0.1~0.5 随机噪声的输入矢量对网络性能进行测试。网络训练如果达到预期设定的误差要求,则绘制出识别错误率曲线。

由此,系统流程设置如图 5 所示。

4 网络训练的结果

从图 6 中可以看出,网络具有较好的收敛速度。图 7 则得到了网络识别字母和数字的出错百分比。由

图可以看出,理想和有噪声两种网络都有一定的抗干扰能力,在平均值为 0~0.4 之间的噪声影响下,两个网络都能 100% 准确无误地进行识别,而当噪声加大到 0.4~0.5 时,理想网络的出错率明显比带噪声训练网络的出错率要高,在平均噪声为 0.5 的情况下,理想网络的出错率为 0.17%,而带噪声训练的网络出错率仅为不到 0.06%。

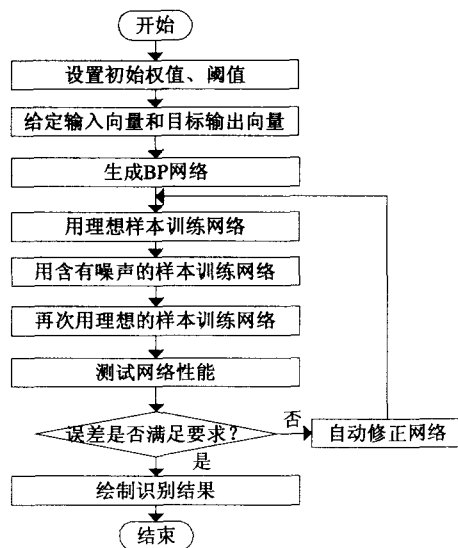


图 5 系统流程图

5 结语

本文设计实现了一个带动量因子及学习率自适应的 BP 网络来识别带有一定噪声的 26 个大写英文字母和 10 个阿拉伯数字,试验结果表明,网络的识别率比较高,系统可以作为实际车牌等识别过程中的一个关键环节使用。

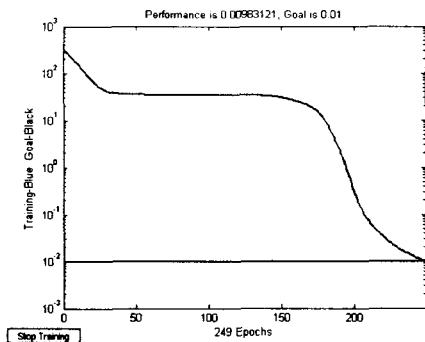


图 6 网络的训练记录

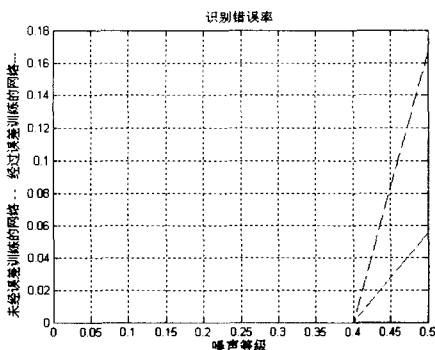


图 7 网络的识别错误率

参考文献

- [1] 飞思科技产品研发中心. 神经网络理论与 MatLab 7 实现[M]. 电子工业出版社, 2005: 99~105
[2] 从爽. 面向 MatLab 工具箱的神经网络理论与应用[M]. 中国科学技术大学出版社, 1998: 53~70
[3] 四维科技, 胡小锋, 赵辉. Visual C++/MatLab 图像处理

与识别实用案例精选[M]. 人民邮电出版社, 2004: 196~227

- [4] 陈先锋, 舒志兵, 赵英凯. “弹性”BP 神经网络在识别带有噪声字母中的应用[N]. 计算机仿真, 2005, 22(9): 153~155

Recognition of English Letters and Numbers with Noise Based on Neural Network

NIU Hui-juan , WANG Sen-lin

(Guangzhou International Economics College, Guangzhou 510540)

Abstract: Applies the neural network to design a tri-level Back-Propagation network with a 432-26-36 structure, and uses this network to recognize 26 upper case letters in English and 10 Arabic numerals with some noise, at the same time, tests the influence of the network's initial parameters to the speed of convergence, through comparing the results of many trials, chooses the best number of hidden nodes. This system uses MatLab to simulate, and the results of trials shows it can recognize the upper case letters in English and the numbers efficiently.

Keywords: English Letters Recognition; Numbers Recognition; Back-Propagation Network; MatLab; Noise

(上接第 39 页)

Research on Maximal Frequent Item FP-MFI Algorithm Based on FP-tree

ZHENG Hai-ming

(Computer Science and Technology Institute, Jilin University, Jilin 135000)

Abstract: Because of generating the candidate ones in the maximal frequent item-sets and it will bring on a batch of the candidate sets, through improving the traditional FP-tree structure and proposes the maximal frequent item-sets mining algorithm based on the improved FP-tree. It needn't to generate the candidate maximal frequent item-sets. The improved FP-tree is unilateralist, and each point saves the pointers of the parents', which will economize memory. It is shown in our experimental results that the FP-MFI algorithm is more effectively than the DMFIA also based on FP-tree in the mining maximal frequent item-sets when the frequent items are very much in database and they are so few in each transaction.

Keywords: Data Mining; Association Rules; Maximal Frequent Item; FP-tree