

DOI:10.15923/j.cnki.cn22-1382/t.2019.1.06

基于人工神经网络的字母识别

陈戈珩, 程青青*

(长春工业大学 计算机科学与工程学院, 吉林 长春 130012)

摘要: 在识别过程中加入噪声非线性干扰, 利用 MATLAB 工具箱对字母进行模拟识别, 搭建神经网络并且确定隐含层的个数, 利用两类不同的数据对神经网络进行训练, 最终实现字符识别的功能。

关键词: BP 神经网络; MATLAB 仿真; 模式识别

中图分类号: TP 393 **文献标志码:** A **文章编号:** 1674-1374(2019)01-0031-07

Letter identification based on artificial neural network

CHEN Geheng, CHENG Qingqing*

(School of Computer Science & Engineering, Changchun University of Technology, Changchun 130012, China)

Abstract: Nonlinear noise interference is added in the process of identification, and MATLAB toolbox is used simulate the letters. We build up a neural network and determine the number of hidden layers. Two kinds of data are used to train the neural network to realize the function of character recognition.

Key words: BP neural network; MATLAB simulation; pattern recognition.

0 引言

21 世纪以来,世界进入信息化时代。计算机技术的快速发展,使人们的工作生活变得越来越方便、越来越智能化。它让人类从很多枯燥乏味的工作中解放出来,在人工智能领域中字符识别(Optical Character Recognise, OCR)技术成为其中一个非常重要的应用^[1]。

在这个信息爆炸的时代,人们的日常学习、工作和生活中几乎每时每刻都充斥着与字符相关的各种信息,字符信息是人类从外界获取信息最重

要的方式之一,这些字符信息一般可以分为两类:手写体和印刷体。而我们所说的 OCR 技术一般指的是对印刷体字符的识别技术,原因在于手写体因人而异,一般没有统一的字符模板可供计算机参考,因此,计算机对这类字符的识别通常比较困难,不仅识别率比较低,而且对手写体字符的识别速度相对印刷体字符来说也比较慢。因此,接下来所提到的字符识别技术都是针对印刷体字符的。印刷体字符通常具有一定的规则性,而且不论对于哪一种字体来说,其对应的印刷体字符都有相应的字符模板可供参考,这也是印刷体字符

收稿日期: 2018-11-05

基金项目: 国家科技支撑计划基金资助项目(2007BAQ00097); 吉林省自然科学基金资助项目(20101523)

作者简介: 陈戈珩(1961—),女,汉族,吉林长春人,长春工业大学教授,主要从事嵌入式系统设计方向研究,E-mail:1026459117@qq.com. * 通讯作者:程青青(1993—),女,汉族,安徽亳州人,长春工业大学硕士研究生,主要从事车联网、无线网络通信方向研究,E-mail:1364619658@qq.com.

的识别率远高于手写体字符的重要原因之一。

人类对印刷体字符的识别技术早在上个世纪初就开始了。1929 年,德国科学家 Tausheck 提出了光学文字识别的概念,这是最早的一种 OCR 技术。之后的几十年,世界各国也相继开始 OCR 技术的研究,但是由于硬件技术等条件的限制,直到上个世纪六七十年代,才相继有少数学术成果以及基于 OCR 技术实现的产品出现^[2]。1966 年,IBM 公司的 Casey 和 Nagy 两人联合发表了第一篇基于汉字识别的文章,填补了人类发展历史上这一页的空白。与此同时,日本在 OCR 技术上的研究也取得了一定的进展,研制出了一款基于印刷文字识别的邮政编码识别系统。我国在 OCR 技术方面的研究晚于这些国家,20 世纪 70 年代才开始对数字和符号的识别进行研究,而且由于条件的限制,研究进度相对较慢,直到 90 年代,我国对字符识别技术的研究才开始步入正轨,相应的技术和产品才逐渐得到应用和普及。

1 BP 神经网络

1.1 基本原理

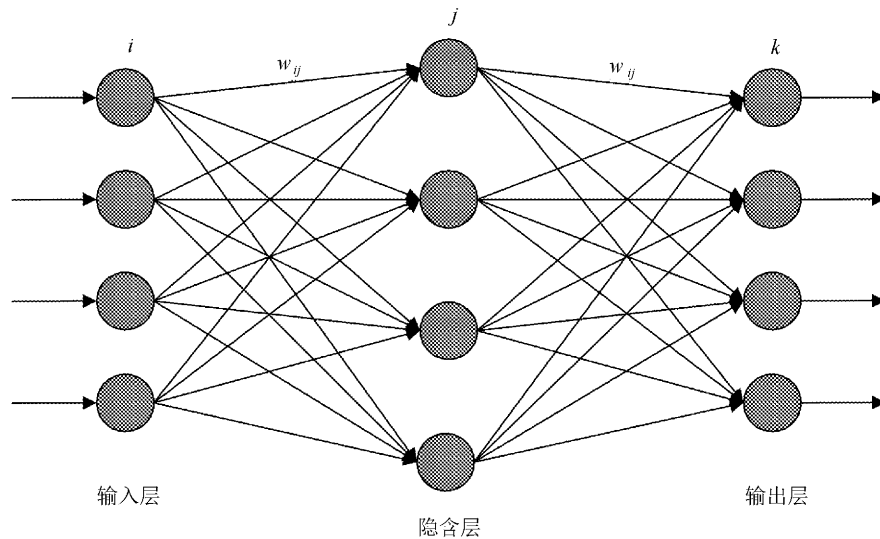


图 1 BP 神经网络模型

前面已经提到, BP 神经网络是一种具有多层结构的前馈神经网络,一个完整的 BP 神经网络由输入层、隐含层和输出层组成,其中隐含层可以具有多层^[5]。

1.3 附加动量法

附加动量法相比反向传播法在进行权值修正时,考虑的因素更多,在反向传播法的基础上引入

BP(Back Propagation)神经网络是人工神经网络的一种,是由 Rumelhart 和 McClelland 的团队 1986 年提出的,也是目前应用最广泛的人工神经网络之一^[3]。总的来说, BP 神经网络属于前馈多层神经网络的一种,它在函数逼近与预测分析、模式识别、分类以及数据压缩等方面具有非常广泛的应用^[4]。基本工作原理是依靠误差的反向传播进行网络训练,以期望网络的实际输出值与期望输出值之间的误差均方差达到最小。

BP 神经网络的核心部分是 BP 算法, BP 算法的核心思想是以网络误差平方为目标函数、采用梯度下降法来计算目标函数的最小值^[5]。算法的执行过程包含两个部分,信号的正向传播以及误差的反向传播。信息首先沿正向传播,当网络输出值与期望输出值的误差不满足要求时,转入误差的反向传播阶段,并修改各个神经元权值,通过这样的反复迭代过程,最终可以得到期望的网络输出值。

1.2 网络模型

BP 神经网络模型如图 1 所示。

误差在曲面上的变化趋势的影响,这样做的好处就是网络在进行每一次权值修正时,都要加上相比于前一次的权值变化量的值,新的权值变化量的产生方式依据反向传播方法的原理^[6]。

不仅如此,附加动量法忽略了网络上一些无关紧要的微小变化特性,这样做不仅能够使网络的训练速度加快,而且可以防止网络陷入局部极

小值。

下面是带有附加动量因子修正公式:

$$\Delta w_{ij}(k+1) = (1 - mc)\eta\delta_i P_j + mc\Delta w_{ij}(k)$$

$$\Delta b_i(k+1) = (1 - mc)\eta\delta_i + mc\Delta b_i(k)$$

式中: mc —— 动量因子, 取值 0 或 1;

k —— 训练次数。

附加动量法就是利用动量因子将最新的权值修正情况传递下去。从上式可以看出, 动量因子取值不同, 权值的变化情况是截然不同的。当 $mc = 0$ 时, 权值依据梯度法的规则变化; 当 $mc = 1$ 时, 自动忽略梯度法产生的权值变化, 并将最新的权值变化设定为当前的权值变化。

另外, 附加动量法规定了两种特殊情况, 一是修正后的权值导致误差产生较大的变化, 二是误差变化率超过当前设定的最大误差变化率, 最大误差变化率应当大于等于 1, 通常取 1.04。这两种情况看似在说明同一个问题, 其实不然, 因为误差的产生不仅仅是因为权值的改变, 第一种情况说明的仅仅是因为权值改变导致的误差, 属于局

部误差的范围, 而第二种情况介绍的是全局误差。当这两种情况发生时, 应当舍弃当前计算得出的权值变化^[7]。这就要求在进行程序设计时注意选用正确的动量因子的值, 具体值及判断公式如下:

$$mc = \begin{cases} 0, & SSE(K) > SSE(K-1) * 1.04 \\ 0.95, & SSE(K) < SSE(K-1) \\ mc, & \text{其他} \end{cases}$$

式中: $SSE(K)$ —— 第 K 次的误差变化率。

1.4 自适应学习速率

考虑到一个固定的学习效率可能不能完全适应训练前后期的要求, 文中采用可变的学习效率。首先, 检查修正后的权值是否能够真正有效降低误差, 如果能, 说明当前选取的学习效率较低, 此时适当提高学习效率; 如果不能有效降低误差, 说明当前学习效率可能过高, 适当降低学习效率, 直到网络的学习过程趋于稳定为止。这样做的好处是, 可以保证网络时刻都是以当前可以接受的最大效率进行学习, 可以大大缩短网络训练的时间。

具体的自适应学习效率公式及条件如下:

$$\eta(k+1) = \begin{cases} 1.05\eta(k), & SSE(K+1) < SSE(K) \\ 0.7\eta(k), & SSE(K+1) > 1.04 * SSE(K) \\ \eta(k), & \text{其他} \end{cases}$$

式中: $\eta(k)$ —— 第 k 次的学习效率。

2 神经网络设计

用神经网络对字母进行识别之前需要对字母进行归一化处理, 用一个 7 行 5 列的矩阵来对字

母进行数字化的处理, 其中有数据的地方用 1 表示, 反之则用 0 表示。字母 A、B、C 的数字化过程如图 2 所示。

也可以用 $1 * 35$ 的向量来表示数字化处理后的字母。则可以表示为:

$$\text{LetterA} = [00100010100101010001111111000110001]$$

$$\text{LetterB} = [11110100011000111110100011000111110]$$

$$\text{LetterC} = [01110100011000010000100001000101110]$$

0	0	1	0	0	1	1	1	1	0	0	1	1	1	0
0	1	0	1	0	1	0	0	0	1	1	0	0	0	1
0	1	0	1	0	1	0	0	0	1	1	0	0	0	0
1	0	0	0	1	1	1	1	1	0	1	0	0	0	0
1	1	1	1	1	1	0	0	0	1	1	0	0	0	0
1	0	0	0	1	1	0	0	0	1	1	0	0	0	1
1	0	0	0	1	1	1	1	1	0	1	1	1	1	0
A					B					C				

图 2 数字化字母

由此可以得到 26 个标准字母的神经网络输入样本,将输入样本定义为 alphabet 向量矩,其中 $\text{alphabet} = [\text{letterA}, \text{letterB}, \dots, \text{letterZ}]$ 。由于每个字母都由 35 个元素组成一个向量,故 alphabet 是一个 35×26 的矩阵,故将网络输出矩阵定义为 26×26 的单位阵,记为: $\text{target} = \text{eye}(26)$ 。在实验时,将理想标准信号和加入噪声的标准输入信号这两类数据分别作为数据的输入信号。

在对 BP 网络设计时,隐层节点个数的选择是 BP 网络的重要环节,它也是网络实现预测成败的关键。在实际的预测实验中,隐层节点个数

过多会导致网络学习过度,学习时间剧增。隐层节点数过少,则不能达到所需要的实验预测要求。由此可见,隐层节点对于神经网络的学习和计算特性起着非常大的作用^[8]。

实验中,根据 Kolmogorov 定理和单隐层的设计经验公式确定隐层节点数范围介于 8~17 之间。

通过实验的验证和误差值的分析,将最终值确定为 14。此时,BP 网络的输出显示以及网络训练速度和精度因素都是最好的。实验仿真如图 3~图 5 所示。

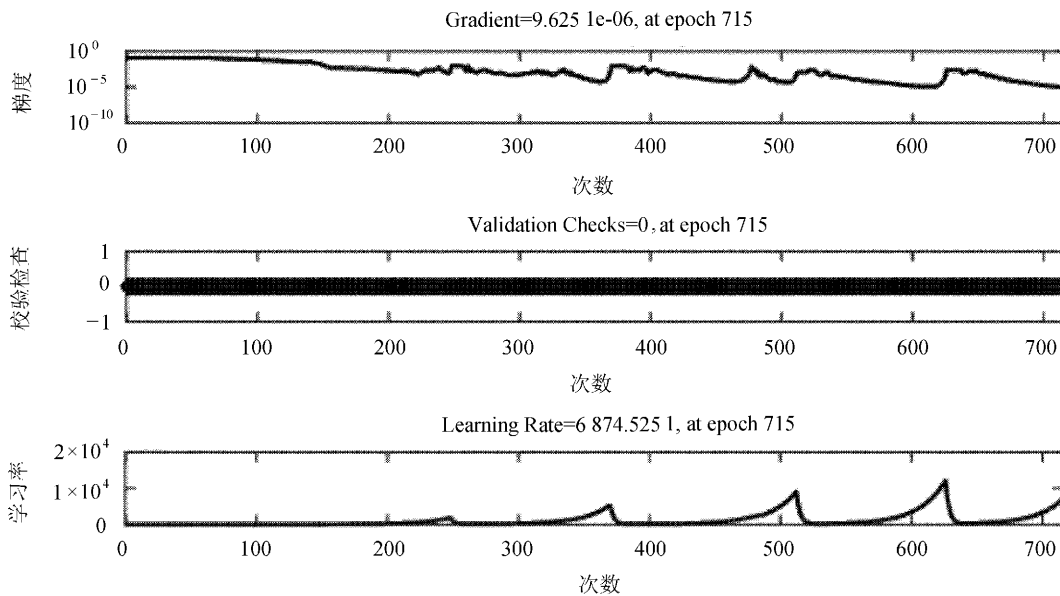


图 3 理想标准信号训练状态图

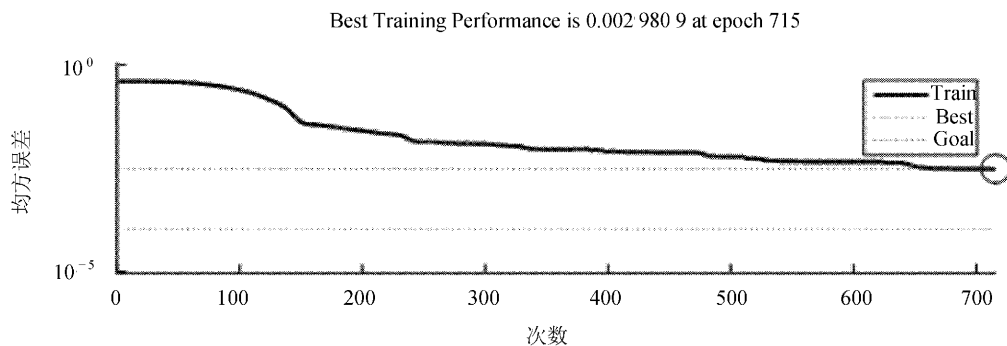


图 4 理想标准信号训练误差性能曲线

3 仿真实验

文中选择 MATLAB 开发环境,基于

MATLAB 对于图像、神经网络等提供的工具箱函数建立了 BP 神经网络,并对获取的样本进行识别。

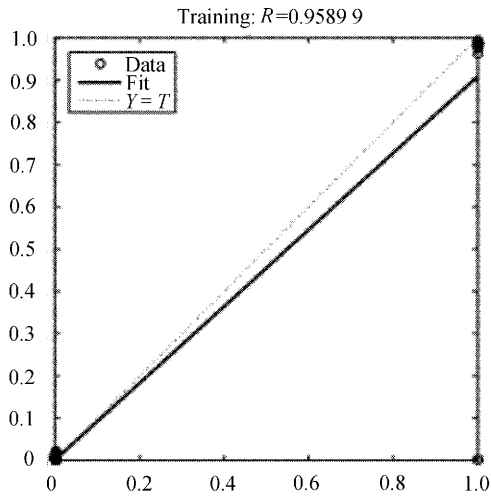


图 5 训练回归曲线

实验过程中,为了使产生的网络具有一定的容错能力,需要采用 BP 网络对不同的信号进行训练,再使用 `trainlm` 函数达到快速训练的目的,在这个过程中,网络的学习速率和冲量参数的设定均为自适应改变。先将理想信号输入,直到平方和误差足够小再停止训练。文献[9]在实验时,将训练结束的条件规定为:最大次数 1 000,误差平方和为 0.000 01。通过 MATLAB 训练后得到实验结果: TRAINLM, Epoch 368/1000, SEE 3.68e-05/1e-05, Gradient 9.972e-06 TRAINLM, Performance goal met。由此可见,经过 368 次训练后误差达到要求。实验仿真如图 6~图 8 所示。

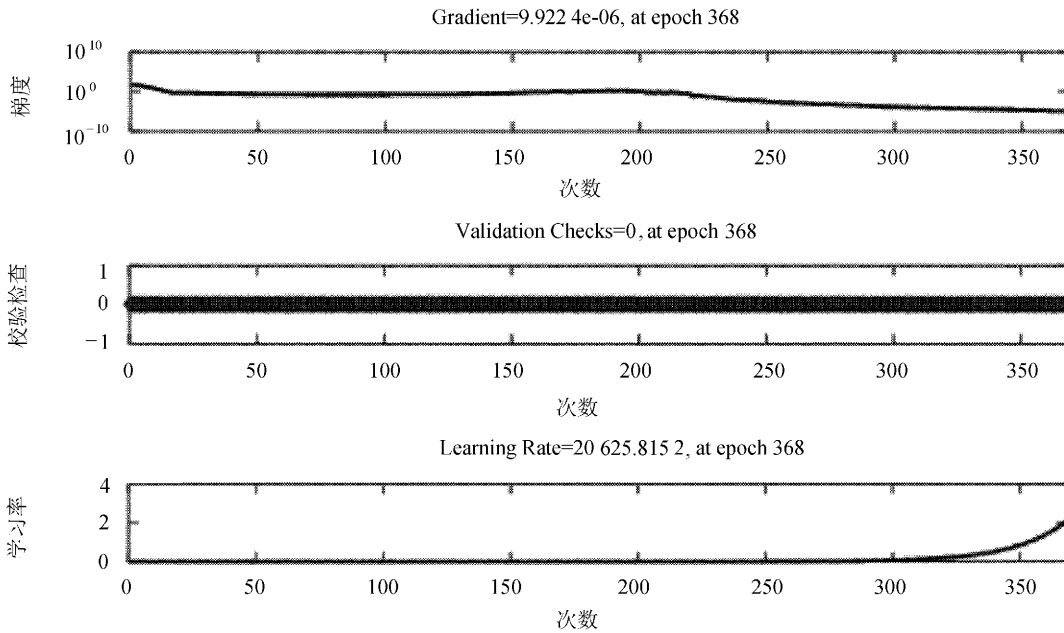


图 6 加入噪声信号的训练状态图

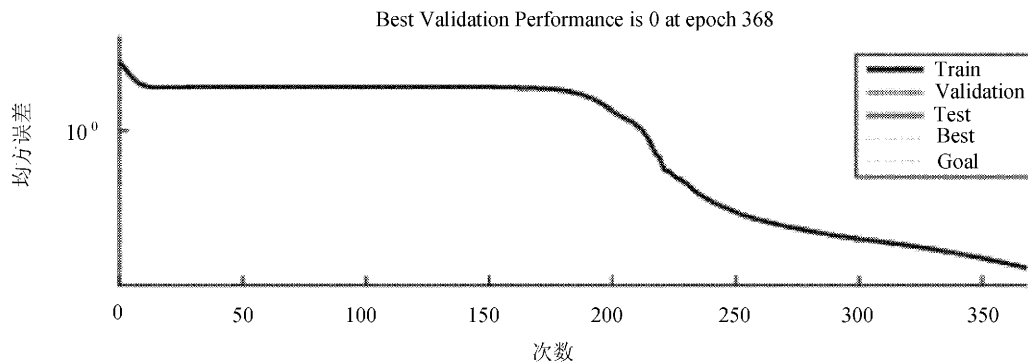


图 7 加入噪声信号的训练误差性能曲线

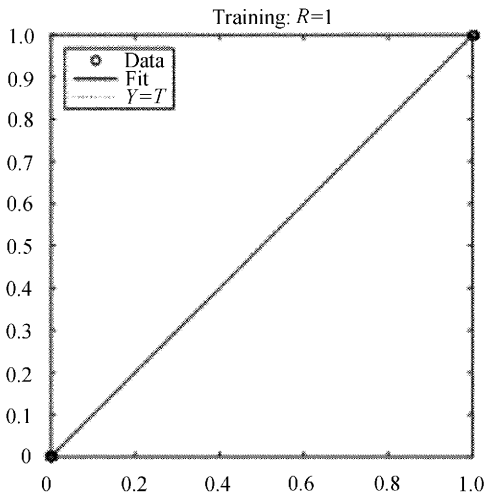


图 8 训练回归曲线图

接下来进行的是含有噪声的信号实验,选用 15 组带有噪声的信号,将加入噪声的均值设定为 0.1 和 0.2。这是为了保证神经网络既能够对含噪声信号识别,又能够对于理想情况下的字母进行识别。为了保证网络在识别理想输入向量的稳定性,要在输入时满足同时输入有误差向量和两倍重复的无误差信号。运行程序,查看识别错误与噪声信号的关系,实验结果如图 9~图 11 所示。

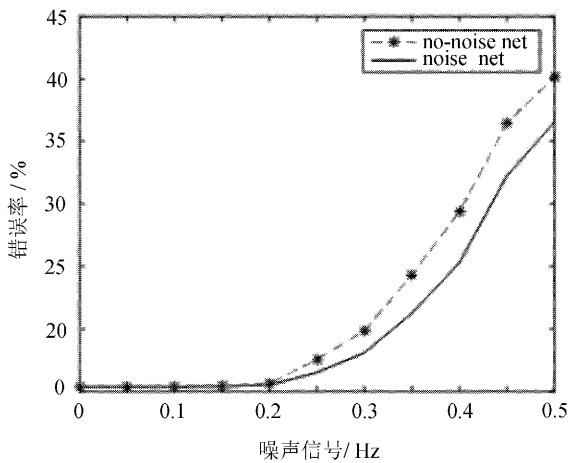


图 9 第一次运行结果

通过三次结果的对比分析,可以看出每一次的运行结果都存在差异,这是因为在实验的过程中,每次运行时产生的随机数是不同的。

由实验数据可以得出,当噪声小于 0.15 的时候,两个网络的识别错误率基本上是不同的,而当噪声不断增加时,经过训练的网络会明显比没有

经过训练的网络识别错误率要低,噪声越大,差距也会越大。

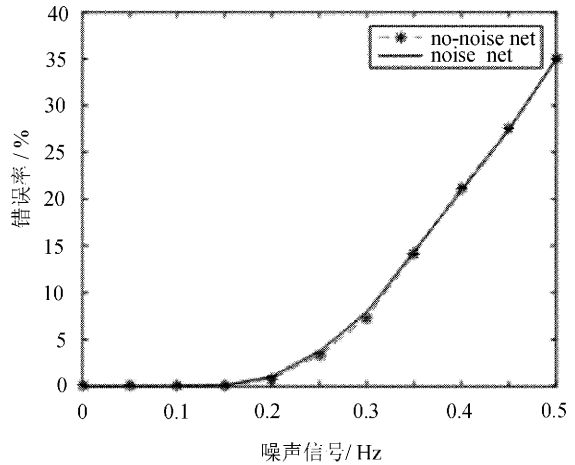


图 10 第二次运行结果

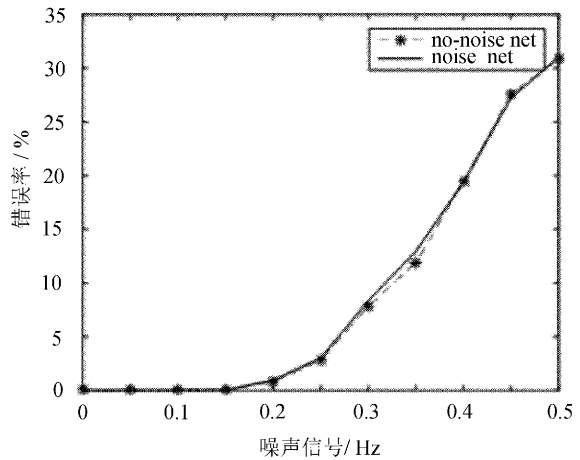


图 11 第三次运行结果

编辑 alphabet_recognise.m 代码,运行程序,结果如图 12~图 15 所示。

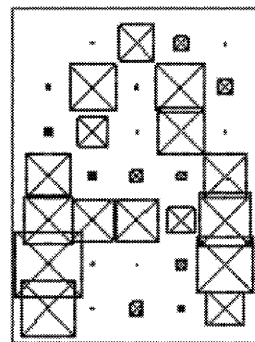


图 12 含噪声的字母 A 显示图

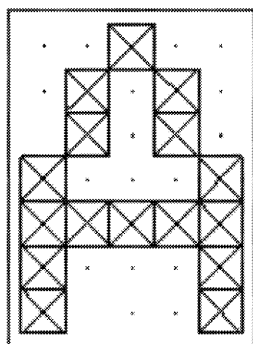


图 13 经过训练后的网络显示的 A 图

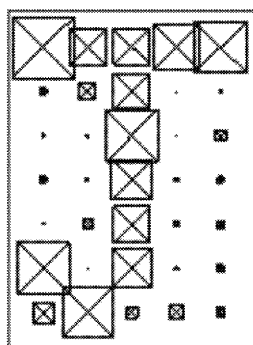


图 14 含噪声的字母 J 显示图

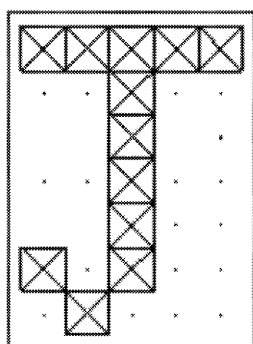


图 15 经过训练后的网络显示的 J 图

4 结 语

利用 BP 网络对有噪声的字母进行识别和仿真,结果表明,此网络具有联想记忆和抗干扰功能,对于字母具有一定的辨别能力,可以作为一种

对字母识别的有效方法。改进后的 BP 传播算法在加快训练速度的基础上还解决了在实验时网络陷入局部极小值的问题,在训练过程中可以根据当时的具体情况来自动调整学习速率,从而将学习速率和误差值达到最佳的平衡状态。在实验过程中发现,均值为 0~0.15 之间的噪声环境下,两个网络都能够准确的进行识别。但是当所加的噪声均值超过 0.15 时,待识别字符在噪声作用下不再接近于理想字符,无噪声训练网络的出错率急剧上升,此时有噪声训练网络的性能较好。为了获得更高的识别准确率,可以训练更多的样本,尽量选择一些带有噪声的字母来多次训练网络,使得识别的准确率更高。

参考文献:

- [1] 高振斌,赵盼,王霞,等.印刷体数字识别系统的 FPGA 实现[J].重庆邮电大学学报:自然科学版,2015,27(2):213-218.
- [2] 陈明华.印刷体数字识别算法研究[D].武汉:华中科技大学,2012:12-14.
- [3] D E Rumelhart, G E Hinton, R J Williams. Learning representations of back-propagation errors [J]. Nature(London),1986,323:533-536.
- [4] 杜慧茜,梅文波,李德生.一种改进的 BP 神经网络在遥感图像分类中的应用[J].北京理工大学学报,1998,18(4):485-488.
- [5] 张德江.遗传神经网络能耗预测模型在钢铁企业中的应用[J].长春工业大学学报:自然科学版,2007,28(2):186-189.
- [6] 钟淑瑛,李陶深.基于 MATLAB 的 BP-LVQ 神经网络组合分类器模型[J].计算机技术与发展,2006,16(2):114-116.
- [7] 荣秋生,潘海森,颜君彪.基于 Hopfield 神经网络的图像矢量量化[J].微计算机信息,2007(6):301-302,204.
- [8] 王静娇,孙晶,周玉冰,等.基于 TMS320DM642 的人民币图像特征识别系统[J].数据采集与处理,2012,27(S2):206-211.
- [9] 丛爽.面向 MATLAB 工具箱的神经网络理论和应用[M].合肥:中国科学技术大学出版社,2009:64-99.