

基于神经网络的字母识别系统的设计与实现

Design and Implemen of Letter Recognition System Based on Neural Network

贺 杰

He Jie

(广西大学梧州分校计算机科学系, 广西梧州 543002)

(Dept. of Comp. Sci., Guangxi Univ. Wuzhou Branch, Wuzhou, Guangxi, 543002, China)

摘要:基于神经网络,采用Matlab 6.5和Visual C,设计一个字母识别系统。该系统通过对BMP图片的二值化处理,在VC环境下调用Matlab,并将把二值化后的数据进行网络训练,从而实现26个英文字母的识别。系统性能的测试表明,系统所训练的神经网络有很好的抗干扰能力。

关键词:字母 模式识别 神经网络 Matlab Visual C

中图分类号:TP315 文献标识码:A 文章编号:1002-7378(2005)04-0278-04

Abstract:A letter recognition system based on neural network is developed by using Matlab 6.5 and Visual C. In this system, the BMP image is binarized. The binarized data are used in the neural network for the recognition of 26 letters. The system test reveals that the neural network has a good anti-jamming.

Key words:letter, pattern recognition, neural network, Matlab, Visual C

手写字母识别一直是热门的研究课题,字母识别要解决数据的采集、处理及选择、输入样本表达的选择、模式识别分类器的选择以及用样本集对识别器的有指导的训练等典型问题。随着人工神经网络研究的深入,为解决计算机一些不易处理难题,特别是语音和图像的识别、理解、知识的处理,优化组合计算机和智能控制等难题提供强有力的工具。本文基于神经网络,采用Matlab 6.5和Visual C设计一个字母识别系统,该系统通过对BMP图片的二值化,在VC环境下调用Matlab并且对无噪声和有噪声训练进行比较,实现了26个英文字母的识别。

1 系统的设计思路

要实现字母识别,先把要识别的字母进行二值化处理,通过神经网络训练后,实现字母识别,其主要思路如下:

- (1)对字母图像二值化,转化成 5×7 的矩阵。
- (2)在VC环境下把二值化后的数据,调入到Matlab中,保存成.mat数据文件^[1]。
- (3)初始化前后神经网络,用改进的快速算法和

样本向量训练前后神经网络,绘出网络平方误差和训练步长变化曲线。

(4)采用有噪声样本对神经网络训练。

(5)将待识向量输入,对神经网络进行仿真。得到结果。

2 神经网络的创建和训练

2.1 网络结构的创建

该系统的神经网络需要具有35个输入节点和26个输出神经元,采用输入在(0,1)范围的对数S型激活函数两层logsig/logsig网络^[2],这种网络对0-1型布尔值是相当完美的,网络取35-10-26的结构。隐含层取10个神经元。如果需要网络有更高的识别精度,可以再增加一些隐含层神经元。

网络设计时,为了使其输出矢量在正确的位置上输出为1,而在其他位置上输出为0,网络的结构如图1所示。

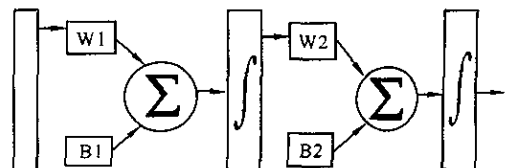


图1 网络结构

传递函数一般为(0,1)S型函数

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

2.2 网络初始化和网络训练

在训练神经网络之前,必须设置权值和阈值的初始值。当神经网络的权值和阈值初始化后,就可以对网络进行训练。其中的网络训练算法使用自适应修改学习率加动量因子算法 traingdx^[3]来进行。

误差函数:对第 p 个样本误差计算公式^[4]为

$$E_p = \frac{\sum_i (t_{pi} - o_{pi})^2}{2}$$

其中, t_{pi} 、 o_{pi} 分别为期望输出和网络的计算输出。

在负梯度算法中,学习率是一个固定的常数,且它的值将直接影响到网络的训练性能。如果选择得太大,会降低网络的稳定性;如果选择得过小,会导致过长的训练时间。

在每次训练结束后,利用此时的学习率计算出网络的权值和阈值,并且计算出网络此时的输出误差。如果此时的输出误差与前一时刻的输出误差的比值大于预先定义的参数max-perf-inc,那么就减少学习率,反之,就增加学习率。再重新计算网络的权值和阈值以及输出误差,直到前后输出误差的比值小于参数 max_perf_inc 为止。

3 字母识别系统的实现

对字母进行识别前,首先对字母的BMP 图片二值化,在 VC 环境下调用 Matlab,并通过对无噪声和有噪声训练进行比较,从而实现字母的识别。

3.1 字母图像的二值化实现

选取从 A 到 Z 26 个英文字母的 256 色图片,例如 A 的 bmp 图片如图2 所示。字母识别时,首先将印在纸上的字母经过光电扫描产生模拟信号,再通过模数转换变成表示灰度值的数字信号输入计算机。二值化的关键在于阈值 T 的选择。考虑到各个样本数字的灰度范围的差别,可以采用不同的二值化阈值进行处理。



图2 字母 A 的 bmp 图片

设像素点 (i, j) 的灰度为 $g(i, j)$ 。有笔划的部分的 $g(i, j)$ 小。白背景部分的 $g(i, j)$ 大。二值化可以通过设定灰度阈值来完成,即当 $g(i, j)$ 值比预定阈值大,看成背景,以 0 表示;否则看成笔划,以 1 表示。这样,经过二值化,各点重新表示为:

$$c(i, j) = \begin{cases} 1, & g(i, j) < T(\text{笔划}) \\ 0, & g(i, j) \geq T(\text{空白}) \end{cases} \quad i = 1, 2 \dots N; j = 1, 2 \dots N.$$

比较阈值,大于阈值,则这个点取 0,为白色。小于阈值,则这个点取 1,为黑色,对图像进行二值化。

输入一个 bmp 图片,显示出 bmp 经过二值化后的 0,1 矩阵。读入一个 bmp 图片如图3 所示,显示的矩阵如图 4 所示。



图3 被识别的 bmp 图片

```

Input Bmp File Name
01.bmp

1 1 1 1 1
1 0 0 0 0
1 0 0 0 0
1 1 1 1 0
1 0 0 0 0
1 0 0 0 0
1 0 0 0 0
    
```

图4 二值化后的矩阵

3.2 VC 调用 Matlab 的实现

经过图片二值化后,把二值化后的数组数据调入到 Matlab 中作为待识向量,以 recgNow 调用,将其识别出来。

如果在 Matlab 保存的是一个列向量,就要把一个高 x 宽的矩阵换成一个高 * 宽 x 1 的矩阵,所以在 VC 对 Matlab b 的调用上创建一个 mxArray^[5],宽为 InfoHeader.biWidth * InfoHeader.biHeight,长为 1,数据类型为 DOUBLE 的数组。

```

在 matlab command Windows 显示:
recgLetter = recgLetter: [35x1 uint8]
    
```

3.3 网络训练的实现

设计一个能够识别 26 个英文字母的神经网络,意味着每当给训练过的网络一个表示某一个字母的输入时,网络能够正确的在输出端指出该字母。神经网络训练应当是有监督的,训练出对输入端的 26 组分别表示字母 A 与 Z 的数组,能够对应输出端 1 到 26 的具体位置^[6]。首先必须将每个字母进行数字化处理以便构造输入样本。用 5x7 矩阵的布尔值可以清楚地表示出每个字母,例如字母可以用 0,1 矩阵表示为:

```

letter A = [0 0 1 0 0 ...
            0 1 0 1 0 ...
    
```

```

0 1 0 1 0 ...
1 0 0 0 1 ...
1 1 1 1 1 ...
1 0 0 0 1 ...
1 0 0 0 1 ];

```

每个字母都可以用 $5 \times 7 = 35$ 的元素组成相应的列矩阵,那么26个字母则分别由表示26个字母的列矩阵组成 26×35 的输入矩阵,然后把这26个字母送入数组:

```
Alphabet = [letter A, letter B, ..., letter Z]
```

3.3.1 对无噪声的网络进行训练

通过网络对样本的训练,得到一组稳定的权值,存入一个适当数据文件.mat中,以备网络运行使用。画出步长与误差的关系如图5所示。

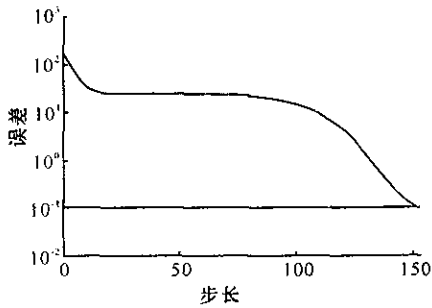


图5 无噪声环境下步长与误差的关系

训练结束后应当将训练好的网络权矢量存入一个适当数据文件 recg-net.mat 中,以备网络运行时用。

3.3.2 对有噪声的网络进行训练

首先训练一个具有两组理想输入矢量 alphabet 加上两组带有噪声的输入矢量,目标矢量为4组期望矢量 targets,噪声矢量为均值为0.1到0.2的随机噪声。这迫使网络学习适当地对含有噪声的字母进行正确的识别。

然后,对无噪声输入再进行训练,以保证网络能够准确无误地识别出理想的字母。最终训练结果应存放在另外一个后缀名为.mat的数据文件中,作为将要被使用的网络权值。

经上述神经网络训练后,虽能够辨识出带有噪声的字母的正确位置,但这却是以牺牲理想无噪声输入辨识的正确性作为代价的,即此时网络可能出现对理想无噪声输入辨识错误的可能性。故应将网络对训练后的权矢量做为初始权值,对无噪声输入再进行训练,以此来保证网络对理想输入输出的正确性。本文用十组“理想加随机噪声”的输入样本对有噪声的网络进行训练后得到步长与误差的关系图

如图6所示。

3.4 字母的识别

网络训练后,把从二值化后调用的 recg-Dat.mat 数据文件中的数据作为待识向量。读出数据,识别算法中调用,在这里不能用 load,因为 load 的是调用一个 1×1 的结构数组,而在这里是调用的一个数据数组^[7]。

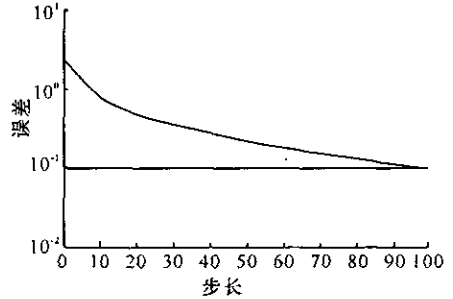


图6 有噪声环境下步长与误差的关系

以一个待识别的图像为例(图7),二值化后的矩阵如图8。总计识别2600次,所有识别结果均为字符B,正确率达100%。



图7 待识字母图片

1	1	1	1	0
1	0	0	0	1
1	0	0	0	1
1	1	1	1	0
1	0	0	0	1
1	0	0	0	1
1	1	1	1	0

图8 二值化后的矩阵

4 系统性能测试

用神经网络进行模式识别的可靠性测试,首先加入噪声,使训练样本发生畸变。然后,对发生畸变的样本进行识别。通过大量的实验发现,神经网络对发生较大畸变的训练样本也有很好的识别能力,系统中将干扰噪声进行数字化处理后即变成平均值为0到0.2之间变化的随机值,由此随机产生100个输入矢量,通过网络识别后输出。为了增强网络的抗干扰能力,网络在使用时,应加上一个竞争网络,以使网络对没个字母的输出只有一个位置为1,其他位置为0。步骤如下:

步骤1: 读出加噪声的网络训练后的网络权矢量。

步骤2: 对样本随机制造噪声,然后进行训练。

步骤3:对无噪声的网络进行仿真和识别。

步骤4:对有噪声的网络进行仿真和识别。

步骤5:画出字母辨识的出错率的图表。

测试结束后,绘制辨识的出错百分比。为比较两者的性能,在测试的同时,加进对无噪声的理想字母辨识神经网络的测试,并作出辨识字母的出错率。对有噪声和无噪声的网络进行训练和识别,出错率如图9所示。平均值0~0.1的噪声的影响几乎100%地被辨识出来,而在较大的噪声出现后,0.1~0.2,理想网络比带噪声训练的出错率高许多。两者的比较如表1所示。表1说明,有噪声训练的网络有更好的识别率。

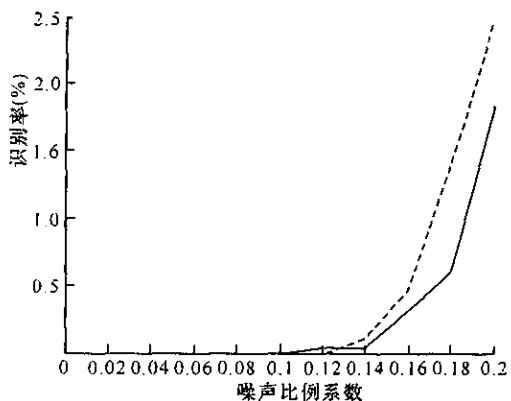


图9 字母表示的出错率

——有噪声训练;---无噪声训练。

表1 无噪声与有噪声训练网络的识别率比较

样本	无噪声训练的网络(%)	有噪声训练的网络(%)
标准样本	100	100
80 以上	噪声样本	70 左右

5 结束语

字母识别系统通过对 bmp 图片的二值化处理,在 VC 环境下调用 Matlab,并且对无噪声和有噪声训练进行比较,从而实现字母识别。对系统性能的测试说明,所训练的神经网络有很好的抗干扰能力。由于系统采用的编程工具是 Matlab6.5 和 Visual C,因而在 Windows 9X、Windows 2000/NT、Windows XP 平台均可实现 26 个英文字母的识别,并有很高的识别率。

参考文献:

- [1] 徐 慧. Visual C++ 数字图像使用工程案例精选 [M]. 北京:人民邮电出版社,2004
- [2] 闻 新,周 露,李 翔,等. Matlab 神经网络仿真与应用 [M]. 北京:科学出版社,2003.
- [3] 张 铮,杨文平,石博强,等. MATLAB 程序设计与实例应用 [M]. 北京:中国铁道出版社,2003.
- [4] 丛 爽. 面向 Matlab 工具箱的神经网络理论与应用 [M]. 合肥:中国科学技术大学出版社,2000.
- [5] 张宏林. Visual C++ 数字图像模式识别技术及工程实践 [M]. 北京:人民邮电出版社,2003.
- [6] J P Marques de Sa. 模式识别——原理、方法及应用 [M]. 北京:清华大学出版社,2002.
- [7] Abhijit S Pandya. 神经网络模式识别及其实现 [M]. 北京:电子工业出版社,1999.

(责任编辑:黎贞崇)

(上接第 277 页)

91%。广西家校通信息发布系统主要面向广西各市、县教育局及全区中、小学校而设计开发,将先进的教育理念和现代信息技术有机结合,代表了教育发展的趋势。

参考文献:

- [1] 教育部. 教育管理信息化标准:学校管理信息标准[S]. 2002.
- [2] 陈庆章. 基于用户角色和阶段性控制的网上公文传送

的安全机制[J]. 信息和通讯安全,1999,(12):1-6.

- [3] Solms S He Von, Merwe Lsak Van Der. The management of computer security profiles using a role-oriented approach [J]. Computers Security, 1994, 13 (9):673-680.
- [4] Bruce Schneier. 应用密码学:协议、算法与 C 源程序 [M]. 吴世忠,等译. 北京:机械工业出版社,2000,185-207.

(责任编辑:黎贞崇)