

# 基于整体特征神经网络数字二次识别方法<sup>①</sup>

张翼成, 陈欣, 嵇小波, 张晓荣

(上海宝信软件股份有限公司, 上海 201900)

**摘 要:** 根据数字字符整体特征, 提出一种基于字符整体特征的 Bp 神经网络数字二次识别方法. 该方法首先根据 Bp 神经网络原理对数字字符进行预识别; 然后对预识别结果中存在混淆的字符按照字符整体特征进行二次识别, 从而准确获得识别结果. 该方法结合了神经网络非线性、自主学习特点和字符整体特征形状结构不变性特点, 有效的在低样本量情况下, 获得较高的字符识别精度.

**关键词:** 数字识别; 神经网络; 字符特征; 二次识别

## Secondary Digital Character Recognition Based on the Global Feature and BP Neural Network

ZHANG Yi-Cheng, CHEN Xin, JI Xiao-Bo, ZHANG Xiao-Rong

(Shanghai Baosight Software Co., Ltd, Shanghai 201900, China)

**Abstract:** The secondary digital character recognition based on the character global feature and BP neural network is proposed in this paper. First of all, digital characters should be pre-recognized by the method of BP neural network, and a preliminary result can be obtained. After that, the secondary recognition to those characters which may be very close to confusing in the preliminary result should be arranged, so that, those digital characters confused can be distinguished and the better result can be reached. As a result, the method, combined with nonlinear and self-learning of the neural network, and invariance of the overall character feature of the shape and structure, can reach the high recognition accuracy on the condition of the low sample size.

**Key words:** Digital characters recognition; Bp neural network; character feature; secondary recognition

### 1 前言

数字识别是指对包含数字的数字图像特征信息进行分析 and 处理, 以对数字图像中的数字字符进行辨认和分类的过程. 数字识别方法根据分类器设计的不同通常可以分为基于统计特征的数字识别、基于结构特征的数字识别、基于神经网络的数字识别和基于支持向量机的数字识别四类. 其中基于神经网络和基于结构特征的数字识别方法是该领域的研究热点, 如文献[1]使用逐像素特征提取方法进行特征向量提取, 构建 Bp 神经网络分类器实现数字识别; 文献[2]研究一种 13 维新特征提取方法与量子神经网络结合的方法构建神经网络分类器实现数字识别; 文献[3]则针对神经网络易陷入局部极小、收敛速度慢的缺点, 提出一种 Bp

神经网络改进算法实现数字快速识别. 上述神经网络数字识别方法都是依赖于大量数字字符样本, 对这些数字字符样本进行训练从而实现数字识别. 而文献[4,5]则根据字符的结构特征提出相应的数字识别方法, 这种方法不需要大量的数字字符样本进行训练, 仅仅通过提取数字字符的结构特征来实现识别, 但是这种方法的鲁棒性和自主学习能力较差, 一般识别率较低. 本文针对上述特点, 提出了一种基于字符整体特征的 Bp 神经网络数字二次识别方法. 该方法结合了 Bp 神经网络的鲁棒性高、自主学习能力强和字符整体特征形状结构不变性特点, 在数字字符样本较少的情况下仍然能够准确的识别数字字符, 而且误识率很低, 准确率较高.

<sup>①</sup> 收稿时间:2012-12-07;收到修改稿时间:2013-01-24

## 2 Bp神经网络数字识别原理

### 2.1 Bp神经网络模型

Bp神经网络是基于误差反向传播算法的人工神经网络,也是目前应用最广泛的神经网络模型之一。Bp神经网络不需要给出有关模式的经验知识和判别函数,只需要通过自身的学习机制揭示输入输出模型的映射关系,从而形成决策区域。由于Bp神经网络具有极强的非线性数据处理能力、容错能力和自主学习能力,因而在大多数的神经网络模型都是采用Bp网络和它的变化形式。Bp神经网络为拓扑类型结构模型,通常包括输入层、隐层和输出层。如图1所示。

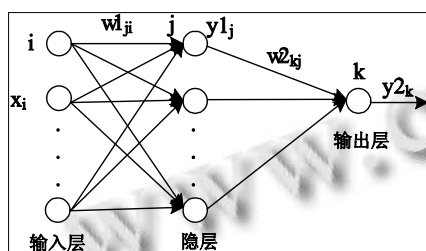


图1 Bp神经网络拓扑结构模型

当Bp神经网络正向传播,模式作用于输入层,输入层各单元对应于输入模式向量的各个元素,经隐层处理后传向输出层。若输出层未能得到期望的输出,则转入误差的逆向传播阶段,将输出误差按照某种形式通过隐层向输入层逐层返回,并分配给各层的所有单元,从而获得各层单元的参考误差或称误差信号,以作为修改各单元权值的依据。这种信号正向传播与逆向传播的各层权矩阵的修改过程是周而复始的进行。周而复始的信息正向传播和误差反向传播过程是各层权值不断调整的过程,也是神经网络的学习训练的过程,此过程一直进行直到网络输出的误差减少到可以接受的程度,或者进行到预先设定的学习次数为止。

### 2.2 Bp神经网络设计

#### 2.2.1 输入层节点确定

Bp神经网络输入层节点的个数为图像特征提取后的输出维数。输入层节点的值为图像提取的特征向量元素值。本文采用逐像素特征提取方法,对于每一个输入的样本,归一化大小为 $16 \times 8$ (高 $\times$ 宽),那么输入层结点数位128,各结点值为归一化之后的图像像素灰度值。

#### 2.2.2 输出层节点确定

神经网络分类器主要根据输出层节点来进行分类

和决策。对于0~9数字识别系统,所需分类数目为10,可以采用10结点型输出层和4结点型输出层。对于4结点型输出层,它采用8421码来识别。如输出结点向量为(0.1,0.1,0.1,0.1)表示识别数字0,(0.9,0.1,0.1,0.1)表示识别数字8,等等。对于10结点型输出层则比较直观,结点向量元素的序号表示识别的数字,如输出层结点向量(0.1,0.1,0.1,0.1,0.1,0.1,0.9,0.1,0.1,0.1)表示识别的数字为3。本文选择10结点型的输出层。

#### 2.2.3 隐层节点及相关参数确定

##### 1) 确定隐层节点数

确定隐层节点数的最基本原则是:满足精度要求的前提下取尽可能紧凑的结构,即取尽可能少的隐层节点数目。研究表明,隐层节点数目不仅与输入和输出层的节点数有关,而且也与需解决的问题的复杂程度和转换函数的形式以及样本数据特性等因素有关。确定隐层节点数一般满足两个条件:1)隐层节点数目小于 $N-1$ (其中 $N$ 为训练样本数目),否则,网络模型的系统误差与训练样本的特征无关而趋于零,即建立的网络模型没有泛化能力,也没有任何的使用价值;2)训练样本数必须多于网络模型的连接权数,一般为2-10倍,否则,样本必须分成几部分并采用“轮流训练”的方法才能得到可靠的神经网络模型。

总之,隐层结点数不能太少,太少则网络可能根本不能训练或网络性能差;同样隐层节点也不能太多,太多虽然可使网络的系统误差减少,但可能会使训练容易陷入局部极小点而得不到最优点。本文根据上述的原则将隐层结点数目设定为13。

##### 2) 确定初始权值

权值的初始值选择遵循两个原则:第一,所选权值不能相同;第二,在定义的值域空间均匀分布。权值的初始值选得好则网络收敛快,否则网络处于饱和区难于收敛或陷入局部极小,导致输出节点误差难以满足,训练失败。本文选取-1到+1范围内的随机数作为神经网络的初始权值。

##### 3) 确定权值学习算法

Bp神经网络权值调整过程实际是误差反向传播调整过程。根据文献[6],误差反向传播过程实际上通过计算输出层的误差 $e_k$ ,然后将其与输出层的激活函数的一阶导数 $f_2'$ 相乘来求得 $\delta_{kj}$ 。由于隐层中没有直接给出目标矢量,所以利用输出层 $\delta_{kj}$ 的进行误差反向传递求出隐层权值的变化量 $\Delta w_{2kj}$ 。然后计算 $e_k = \sum_{k=1}^m \delta_{kj} w_{2kj}$ 并同

时通过  $e_j$  将与该层激活函数的一阶导数  $f'_j$  相乘而求得, 以此求出前层权值的变化量  $\Delta w_{ij}$ . 在这个过程中权值的调整与激活函数  $f(x)$  和步长  $\eta$  有关, 在本文中, 激活函数为  $f(x) = \frac{1}{1+e^{-x}}$ , 步长  $\eta = 0.015$ , 误差阈值  $e = 0.0010$ .

当输入层结点、隐层结点和参数、以及输出层结点确定之后, 输入层到隐含层权值矩阵和隐含层结点到输出层结点权值矩阵即可确定, 从而建立了输入到输出的对应关系, Bp 神经网络分类器确定. 在获得输入层结点向量后, 通过神经网络分类器, 可以实现数字字符的分类识别.

### 2.3 Bp 神经网络的数字预识别

利用 Bp 神经网络对阿拉伯数字进行识别前, 应先对阿拉伯数字图像样本进行预处理<sup>[7]</sup>(图像预处理组成如图 2 所示), 以消除各数字字符在位置和大小上的差异, 从而提高识别的准确率.

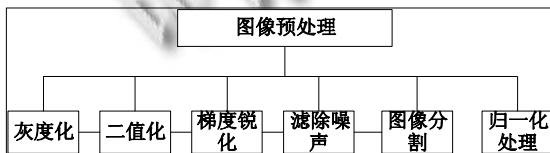


图 2 数字图像预处理过程图

预处理后的数字字符不能直接作为神经网络输入层结点向量, 需要进行特征向量提取, 按照 2.2 部分神经网络设计可知, 数字字符特征向量由 128 个元素组成, 每个元素取值由归一化后的字符像素灰度确定. 特征向量提取之后即可将其送到神经网络分类器进行分类识别. 分类器分类识别主要根据训练确定的输入层结点到隐层结点权值矩阵和隐层结点到输出层结点权值矩阵确立的输入到输出层结点的映射关系来进行分类识别, 最后获得识别结果. 整个 Bp 神经网络数字预识别过程如下图 3:

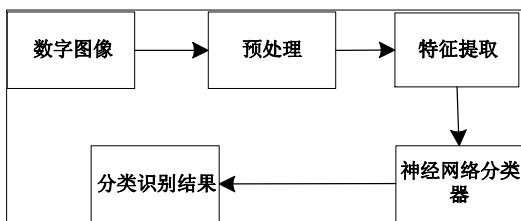


图 3 Bp 神经网络数字识别流程

根据上述步骤, 能够初步将数字字符识别出来,

当样本空间趋近无穷大的时候, 数字字符的识别率很高, 但是在现实生产生活中, 样本空间往往是不够的, 甚至在某些条件下, 样本空间比较少, 那么当数字字符特征作为输入层结点输入到神经网络分类器后, 输出层 10 个结点的输出值就可能存在比较接近的几个值, 这样就会存在对混淆数字字符的误识别, 如字符 5 和 6、1 和 7 等. 这样将会造成数字识别的准确率不高. 如果在输出层结点判定环节根据数字字符结构特征进行二次识别, 那么将可以大大提高数字识别的准确率.

### 3 数字字符整体结构特征

根据文献[4], 数字字符整体特征(凹凸特征)可以实现数字字符的分类. 分类的过程比较简单, 如下图 4 所示.

如图 4, 圆圈是第一个重要的数字字符整体结构特征. 圆圈的判定可以根据数字字符图像的欧拉数进行判定, 如果数字图像的欧拉数为 1 表示字符图像中不存在圆圈, 欧拉数为 0 表示字符图像中存在一个圆圈, 欧拉数为 -1 表示字符图像中存在两个圆圈, 这样可以比较容易的将十个字符分为三类, 一类是唯一数字 8, 一类是一个圆圈的 0,4,6,9, 一类是没有圆圈的 1,2,3,5,7.

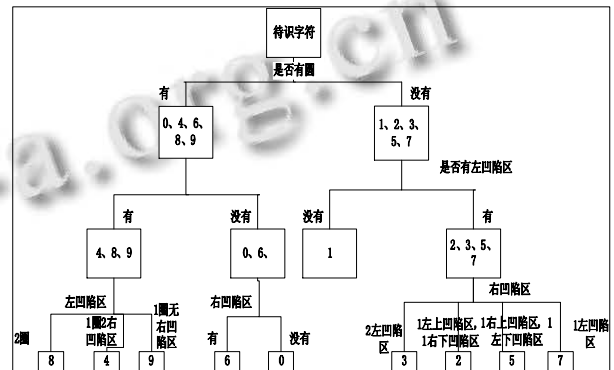


图 4 数字字符整体特征识别

凹陷区域是第二个重要的数字字符整体结构特征. 从上图 4 可以看出, 在根据圆圈进行一次分类后, 凹陷区域位置将就可以唯一确定圆圈分类后的各个字符数字. 如 6 欧拉数为 0, 且不存在左凹陷区, 存在右凹陷区, 就很容易将其与 5 进行区分; 而 7 存在左凹陷区, 于是容易混淆的 1 则不存在左凹陷区, 整体结构特征明显, 比较容易区分.

### 4 二次识别原理

根据本文第 2 部分的内容, 当输入层结点向量确定之后, 根据输入层到隐层权值矩阵和隐层到输出层权值矩阵, 将会获得输出层结点矩阵. 设定一个输出层结点阈值  $T$ , 如果输出层结点向量中存在大于阈值  $T$  的元素, 那么将该元素所在的结点序号作为识别的数字字符. 这样做存在一个问题, 当输出层结点向量中存在若干个(大于 1)满足条件的元素, 且这些元素的值都比较接近, 如果直接取其中一个元素对应的结点序号作为识别的数字字符, 这将是合理的, 会产生错误识别. 因为这些元素对应的数字字符一般都是比较相似的, 如文中提到的 5 和 6, 1 和 7 等, 如果在这个阶段再进行二次识别, 即根据本文第 3 部分内容, 对满足阈值  $T_1$  条件的元素结点序号数字字符进行字符整体特征识别, 那将比较容易的解决错误误判的问题.

Bp 神经网络二次识别将分为两步. 第一步为根据输出层结点向量, 找出可能的识别数字; 第二步是在可能的识别数字中确定唯一的识别数字. 其中第一步的流程图如下图 5 所示.

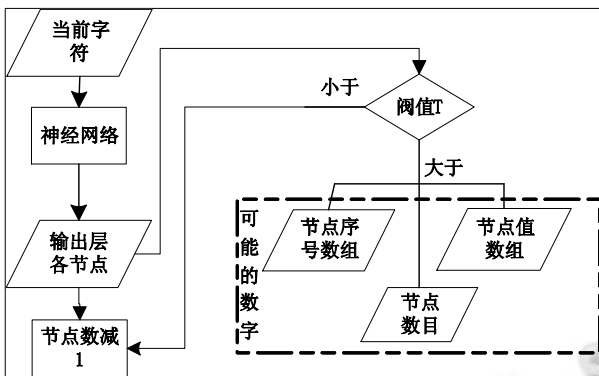


图 5 二次识别第一步流程图

如图 5 所示, 当图像预处理完成后, 获得分割后的归一化数字字符图像. 设定某一个归一化的数字字符为当前字符, 经过神经网络分类器, 可以获得输出层结点向量各个节点值, 这些输出层结点向量节点值能够反映与真实数字的相似程度. 将输出层结点向量各个节点值与设定好的阈值  $T$  比较, 将大于阈值的节点值、节点序号分别以数组形式保存, 并统计满足阈值  $T$  条件的节点数目.

神经网络二次识别的第二步流程图如下图 6 所示.

如图 6 所示, 根据第一步的结果可以获得满足阈值  $T$  节点值数组、节点对应的序号数组和节点的数量,

如果满足阈值  $T$  节点的数量不大于 1, 那么很明显存在两种可能, 第一该数字字符拒识, 第二该数字字符唯一确定, 对应的唯一满足要求节点的序号即为识别的数字. 如果满足条件的节点数量大于 1, 那么满足条件的节点数组中一定存在一个数值最大值, 该数值最大值的节点序号一般情况下能作为唯一识别的数字, 可以将其对应的序号设定为标准数字(相似度最高的数字), 但是如果节点值数组中存在和最大值极为接近的节点值, 那么将不能准确判定. 可以通过将节点值数字中各个元素与节点最大值作差, 并再设定一个阈值  $T_1$ , 如果差值大于  $T_1$ , 说明该节点值与标准数字差别较大, 将从节点值数组中剔除该值和序号数组中剔除该值对应的序号; 如果差值小于阈值  $T_1$ , 那说明该节点对应的序号值与标准数字极为相似, 将这样满足条件的节点再重新组成一个新的节点值数组. 通常情况下, 存在与标准数字即为相似的输出节点数量不会大于 3 个, 可以很容易的将这些极为像素的数字字符通过字符的整体结构特征进行识别, 从而唯一确定数字字符的对应数字, 获得识别结果.

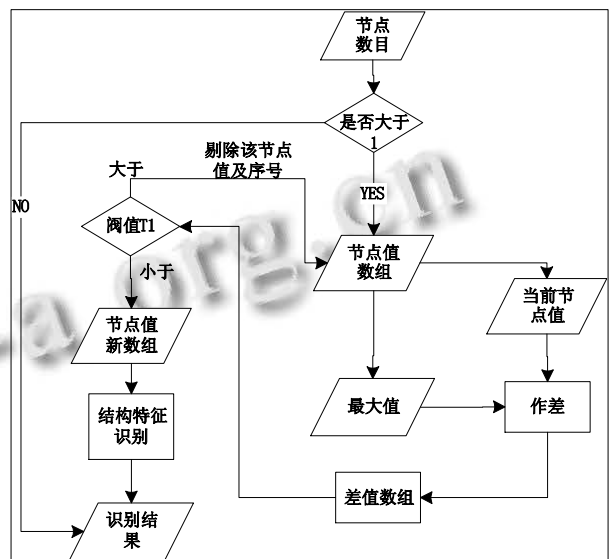


图 6 二次识别第二步流程图

### 5 实验及结论

实验从现场采集图片, 分为两个部分, 一部分为训练的样本库, 一部分为测试的样本库. 训练的样本库是采集图片经过图像预处理阶段获得数字 0~9 各个字符图像各 100 幅, 总计 1000 个样本. 测试样本库是 700 幅采集图片, 每幅图片数字字符数量相等, 总计有

7000 个数字字符. 对这些样本库, 采用两种方法进行测试, 一种普通的 Bp 神经网络进行数字识别, 一种是

按照本文提供的神经网络二次识别方法进行识别, 实验结果如表 1 所示.

表 1 普通 Bp 神经网络数字识别与 Bp 神经网络二次识别方法识别结果对比

	普通 Bp 神经网络识别方法			Bp 神经网络二次识别方法		
	正确率	拒识率	误识率	正确率	拒识率	误识率
数字 0	93.54%	1.03%	5.43%	96.48%	1.26%	2.26%
数字 1	92.72%	1.57%	5.71%	95.71%	1.44%	2.85%
数字 2	91.86%	1.43%	6.71%	96.23%	1.37%	2.40%
数字 3	92.28%	0.89%	6.83%	96.56%	1.01%	2.43%
数字 4	92.37%	1.97%	5.66%	94.66%	1.78%	3.54%
数字 5	91.16%	1.17%	7.67%	96.36%	1.15%	2.49%
数字 6	89.77%	1.68%	8.45%	94.95%	1.72%	3.33%
数字 7	92.68%	2.14%	5.18%	95.82%	1.98%	2.20%
数字 8	90.93%	1.37%	7.70%	97.21%	1.35%	1.44%
数字 9	91.64%	1.55%	6.81%	96.47%	1.29%	2.24%
平均	91.89%	1.48%	6.63%	96.05%	1.43%	2.52%

表 1 针对所有的测试样本库字符图片, 分别统计了数字 0~9 这 10 中不同字符在两种不同算法下识别的情况, 字符识别率通常是指字符识别的正确率, 但是一个字符未识别正确存在两种状态, 一种是不符合自己在算法中设定的状态, 直接过滤掉了, 这种称为拒识别; 一种则是同正确识别的字符一样, 所有的识别流程全部完成, 只是结果错误, 这种称为误识别. 本文提供的算法能够很好处理这两种情况, 将未能识别准确的字符进行二次识别, 进一步从错误的结果中识别出准确的字符, 从而降低了拒识率和误识率, 提高了识别的准确率. 从实验数据表 1 可以看出在相同的样本空间情况下, 采用 Bp 神经网络二次识别方法可以有效降低误识率, 提高数字字符识别的准确率.

#### 参考文献

- 1 施少敏, 马彦恒, 陈建泗. 基于 BP 神经网络的数字识别方法. 兵工自动化, 2006, 25(10): 40-41.
- 2 肖婧, 谭阳红. 基于新特征提取法和量子神经网络的手写数字识别. 电子测量技术, 2009, 32(6): 85-87.
- 3 王婷, 江文辉, 肖南峰. 基于改进 BP 神经网络的数字识别. 电子设计工程, 2011, 29(3): 108-111.
- 4 叶飞, 黎峰. 基于整体特征的快速手写体数字字符识别. 计算机工程与设计, 2006, 27(22): 4347-4352.
- 5 原玉磊, 蒋理兴, 钦桂勤. 基于字符特征的数字字符识别算法. 海洋测绘, 2009, 29(1): 56-58.
- 6 Matrin T, Hagan BD, Howard, Hbeal M. 戴奎等译. Neural Network Design. 北京: 机械工业出版社, 2006.
- 7 徐大宏, 王润生. 一种实时精确的数字识别方法. 计算机工程与应用, 2005, 2: 87-89.