

# 基于神经网络信息融合的印刷体字符识别研究

张 宇

(华北计算机系统工程研究所,北京 100083)

**摘 要:** 针对印刷体字符识别,提出一种基于神经网络信息融合的方法。在对待识别目标提取特征后,分别采用 2 种反向传播算法的改进算法和遗传算法构造神经网络分类器模型,并进行网络的训练和识别工作。通过实验数据着重分析和比较了 3 种算法的特点,将此 3 种分类器得出的分类结果进行决策级的信息融合,最终得出识别结果。实验结果表明,此方法简单可行,具有较高的鲁棒性和识别率。

**关键词:** 神经网络;反向传播算法;遗传算法;信息融合

中图分类号: TP391.43

文献标识码: A

## Research of printed character recognition based on neural network information fusion

ZHANG Yu

(National Computer System Engineering Research Institute of China, Beijing 100083, China)

**Abstract:** Considering the status of the printed character recognition, this paper proposed a new approach which based on the neural network information fusion. After extracting the features of the target, using two back-propagation algorithm and genetic algorithms to build the neural network models to train and recognize the targets. Analyzes and compares the characteristic of the three algorithms through experiment data, take identification results from the three classifiers to the decision-making level information fusion and get the final results. The experiments results show that this method is simple and feasible with high robustness and recognition rate.

**Key words:** neural network; back-propagation algorithm; genetic algorithm; information fusion

印刷体字符识别作为字符识别的一个特定领域,在文摘扫描、名片识别和车牌识别等方面都有重要的应用。基于神经网络的字符识别技术具有一些传统识别技术没有的优点,如识别速度较快、分类能力强,且具有较好的容错性能和自学习能力,因此,基于神经网络的识别技术得到众多研究人员的青睐并已成为当今模式识别应用领域的技术热点。

本文采用了一种基于神经网络信息融合的方法,首先设计了 3 种神经网络分类器模型,分别采用了反向传播算法<sup>[1-2]</sup>(BP 算法)和遗传算法<sup>[1,3]</sup>。通过对 3 种分类器对印刷体字符分类结果的分析,结合各自分类器的特点,设计出相应的信息融合策略<sup>[4]</sup>,从而将 3 种神经网络分类器和信息融合策略结合,形成一种有效的印刷体字符识别的解决方法。由于采用了决策级的信息融合技

术,本方法除具有一般神经网络模型进行字符识别的优点外,且提高了识别效率,更加稳定、高效、合理。

### 1 算法

本文通过对区域面积特征提取方法的改进来对待识别目标进行特征提取,并将提取的特征向量应用于神经网络的训练和识别。在分类器的选取方面,本文采用了 2 种 BP 算法的改进算法和基于遗传算法的(0, 1, \*)-矩阵法<sup>[5]</sup>构造分类器模型。

#### 1.1 基于区域面积的特征提取

基于区域面积的特征提取方法的改进力求在保持区域面积方法特征提取的优点下,简化计算,易于实现。本文采用的改进方法的流程描述如下:

- (1)将 RGB 图转换为灰度图,进而转换为二值图。
- (2)寻找二值图像的上下左右 4 个方向的边界,将

由边界围成的图像划分成  $M \times N$  个小区域, 计算图像中各个小区域像素值为 0 的点的个数  $\text{pixelBlack}(i)$ 。

(3) 计算出单个小区域中总的像素  $\text{pixelTotal}$ , 用每个小区域中像素为 0 的像素数除以单位小区域中总的像素数  $(\text{pixelBlack}(i)/\text{pixelTotal})$ , 将得到的数值作为该图像特征的一部分。

基于简化神经网络计算和尽量减少输入特征维数的原则, 经过实验验证, 本文采用了基于  $4 \times 4$  的区域面积特征提取方法。

## 1.2 BP 算法

BP 算法即反向传播算法 (Back-Propagation Algorithm) 通过网络将误差信号沿原来的连接通路反转回来修改神经元的权值, 直到达到期望目标。采用 BP 算法训练的网络称为 BP 网络, BP 网络是一种由输入层、输出层及若干隐层的节点互连而成的多层前馈型网络。其步骤如下:

- (1) 通过网络将输入向前传播。
- (2) 通过网络将敏感性反向传播。
- (3) 使用近似的速下降法更新权值和偏置值。

本文采用了 2 种 BP 算法的改进算法来设计分类器: (1) BP 网络模型在隐层中选用 sigmoid 函数, 在输出层中选用线性函数, 此时几乎可任意精度逼近任何感兴趣的函数; (2) BP 网络模型在隐层和输出层中均选用 sigmoid 函数, 确保网络的输出在  $[0, 1]$  之间, 利于网络的识别。另外, 此类型的 BP 网络模型中将网络的权值调整从总体考虑, 将组成 1 个训练周期的全部样本都依次输入后计算总的平方误差, 进而从总体上对网络的权值进行调整。

## 1.3 基于遗传算法的 $(0, 1, *)$ -矩阵法

基于遗传算法的  $(0, 1, *)$ -矩阵法不需要提取特征, 也不需要大量样本进行反复训练, 而是考虑到印刷体字符相对稳定的特点, 直接基于像素进行整数运算, 从而简化了计算复杂度。本方法全面地考虑了代表样本间相似性与相异性的因素, 并根据这些因素的不同重要性分别赋以不同的权系数。遗传算法是一种广泛应用的高效的随机化搜索与优化方法。本文的重要参数设定均采用遗传算法。

### 1.3.1 $(0, 1, *)$ -矩阵法算法

(1) 将所有输入的样本图像归一化为  $16 \times 16$ , 构成  $256 \times 1$  的列向量。假设样本类别数为  $n$ , 分别计算训练样本中  $n$  类样本向量的像素算术平均值, 得到  $n$  个实向量  $\rho_1, \rho_2, \dots, \rho_n$ 。

(2) 通过设定 2 个阈值  $t_1, t_2 (1 > t_1 > 0.5 > t_2 > 0)$ , 经实向量  $\rho_1, \rho_2, \dots, \rho_n$  转化为  $(0, 1, *)$  向量  $\alpha_1, \alpha_2, \dots, \alpha_n$ 。当像素值大于  $t_1$  时取“1”, 小于  $t_2$  时取“0”, 否则取“\*”。

(3) 建立模板矩阵  $A$  ( $A$  为  $n \times 256$  的矩阵), 其每 1 行

依次为  $\alpha_1^T, \alpha_2^T, \dots, \alpha_n^T$ 。记待识别的印刷体字符所构成的  $256 \times 1$  的列向量为  $\beta$ 。

(4) 定义矩阵与向量的运算  $\otimes$ 。令矩阵  $D_{p \times q}$  的元素为  $d_{ij} (i=1, 2, \dots, p; j=1, 2, \dots, q)$ , 向量  $\varepsilon_{q \times 1}$  的元素为  $b_k$ , 则定义:

$$D \otimes \varepsilon = \left[ \sum_{j=1}^q d_{1j} \otimes b_j, \sum_{j=1}^q d_{2j} \otimes b_j, \dots, \sum_{j=1}^q d_{pj} \otimes b_j \right]^T \quad (1)$$

式中  $\otimes$  的运算规则为:

$$0 \otimes 1 = 1 \otimes 0 = -1 \quad (2)$$

$$a \otimes * = * \otimes a = 0, a = 0, 1 \text{ 或 } * \quad (3)$$

$$b \otimes b = 1, b = 1 \text{ 或 } 0 \quad (4)$$

取反规则为:

$$\bar{1} = 0, \bar{0} = 1, \bar{*} = * \quad (5)$$

矩阵  $A$  取反为  $\bar{A}$ , 即将矩阵的每一个元素按规则 (5) 取反。同样, 向量  $\beta$  取反为  $\bar{\beta}$  亦如此计算。

定义  $\gamma^{(0,0)}, \gamma^{(0,1)}, \gamma^{(1,0)}, \gamma^{(1,1)}$  如下:

$$\begin{aligned} \gamma^{(0,0)} &= A \otimes \beta & \gamma^{(1,1)} &= A \otimes \beta \\ \gamma^{(1,0)} &= A \otimes \bar{\beta} & \gamma^{(0,1)} &= \bar{A} \otimes \beta \end{aligned} \quad (6)$$

其中,  $A$  是模板矩阵,  $\beta$  是待识别样本向量,  $\gamma^{(i,j)} (i, j = 0, 1)$  为  $n \times 1$  列向量。

$\gamma^{(0,0)}$  是模板与待测样本间相似性的度量。表明模板向量  $\alpha_i (i=1, 2, \dots, n)$  与待测样本向量  $\beta$  对应位都含有“0”的个数。

$\gamma^{(0,1)}$  是模板与待测样本间相似性的度量。表明模板向量  $\alpha_i (i=1, 2, \dots, n)$  与待测样本向量  $\beta$  对应位都含有“1”的个数。

$\gamma^{(1,1)}$  是模板与待测样本间相异性的度量。其元素绝对值表明模板向量  $\alpha_i (i=1, 2, \dots, n)$  中元素“1”, 而  $\beta$  中对应元素为“0”的元素个数。

$\gamma^{(1,0)}$  是模板与待测样本间相异性的度量。其元素绝对值表明模板向量  $\alpha_i (i=1, 2, \dots, n)$  中元素为“0”, 而  $\beta$  中对应元素为“1”的元素个数。

下面依据相关度向量:

$$\sigma = [\gamma^{(1,1)}, \gamma^{(0,0)}, \gamma^{(0,1)}, \gamma^{(1,0)}] \times [\omega^{(1,1)}, \omega^{(0,0)}, \omega^{(0,1)}, \omega^{(1,0)}]^T \quad (7)$$

对样本进行分类。式中,  $\omega^{(1,1)}, \omega^{(0,0)}, \omega^{(0,1)}, \omega^{(1,0)}$  分别对应于  $\gamma^{(1,1)}, \gamma^{(0,0)}, \gamma^{(0,1)}, \gamma^{(1,0)}$  的权系数。 $\sigma$  为  $n \times 1$  列向量, 其元素  $s_i (i=1, 2, \dots, n)$  为待测样本  $\beta$  与第  $i$  类模板向量  $\alpha_i$  的相关度。根据本文算法,  $s_i$  越大表示待测样本  $\beta$  与第  $i$  类模板向量  $\alpha_i$  越相关, 也就越有可能被归为第  $i$  类。

### 1.3.2 通过遗传算法确定权系数和阈值

由上述算法可知, 阈值  $t_1, t_2$  以及权系数  $\omega^{(1,1)}, \omega^{(0,0)}, \omega^{(0,1)}, \omega^{(1,0)}$  是本文算法的待求参数, 对最终识别率有至关重要的影响。下面利用遗传算法求解这 6 个参数。

(1) 编码

采用二值编码。设定权系数  $[0, 15.75]$  之间的实数,

阈值  $t_1, t_2$  分别为  $[0.6, 0.95]$  和  $[0.05, 0.4]$  之间的实数。则权系数  $\omega^{(1,1)}, \omega^{(0,0)}, \omega^{(0,1)}, \omega^{(1,0)}$  可用 6 位二进制串来表示。其映射精度  $\Pi_\omega$  为:

$$\Pi_\omega = \frac{15.75-0}{63-0} = 0.25$$

而阈值可用 3 位二进制串来表示,其映射精度  $\Pi_1, \Pi_2$  分别为:

$$\Pi_1 = \frac{0.95-0.6}{7-0} = 0.05 \quad \Pi_2 = \frac{0.4-0.05}{7-0} = 0.05$$

### (2) 适应度的选取与定标

以识别率作为衡量适应度的唯一指标。在每个个体的适应度正比于该个体对应的阈值和权值分布下测试样本集的识别率。为了加大选择力度,必须对适应度进行定标。记识别率为  $r$ , 则适应度函数<sup>[5]</sup>为:

$$f(r) = \frac{1}{0.21 - (r-0.8)u(r-0.8)} \quad (8)$$

式中  $u(x)$  为阶跃函数。

### (3) 遗传算子

采用常用的选择、交叉和变异算子。选择采用赌轮选择机制,交叉采用单点交叉,变异采用小概率突变。具体步骤如下:

① 随机产生一定数目的染色体串长为 30 位的初始群体。

② 判断是否满足终止准则,若满足则停止。

③ 解码个体,产生每一个体对应的阈值和权系数;根据阈值可求得该个体对应的模板矩阵,结合权系数按上述算法将所有训练样本归类,计算出该个体对应下的识别率,进而求出适应度。

④ 进行遗传算法迭代:选择-交叉-变异。

⑤ 返回步骤②。

## 2 实验结果分析和信息融合

本文以印刷体数字字符识别为例对上述模型构造的 3 个分类器进行了实验。本实验选取印刷体数字字符样本库中的 540 个数字字符图像作为实验样本。实验样本如图 1 所示。

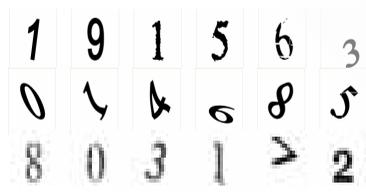


图 1 样本库图例

对于遗传算法,最大迭代次数为 20 代,交叉概率为 85%,变异概率为 5%,群体大小为 20 个个体,染色体串长为 30 位。实验得到的 6 个参数值如下:

$$t_1 = 0.8, t_2 = 0.2$$

$$[\omega^{(1,1)}, \omega^{(0,0)}, \omega^{(0,1)}, \omega^{(1,0)}] = [3, 15.5, 8.5, 1.75]$$

表 1 列出了 3 种模型构造的分类器的识别率比较。

另外,本实验在基于遗传算法的  $(0,1,*)$ -矩阵法中

表 1 3 种模型构造的分类器的识别率比较

分类器所采用的识别方法	识别率/%
一般 BP 网络	91.7
基于总体权值修改的 BP 网络	95.0
基于遗传算法的 $(0,1,*)$ -矩阵法	93.1

对数字字符 0~9 的识别效果进行了对比,结果如表 2 所示。

表 2 对数字字符 0~9 的识别错误数的对比

字符类型	0	1	2	3	4	5	6	7	8	9
识别错误数	1	1	0	0	0	4	0	2	17	0

由表 2 可以看到,基于遗传算法的  $(0,1,*)$ -矩阵法对除数字字符“8”外的其他数字字符均有较高的识别率。因此,本文依据实验结果和分析,针对上述 3 类分类器给出的识别结果,设计出一种简单易行的信息融合策略。由于此方法是在分类器得出识别结果后进行的信息融合,因而称作决策级信息融合<sup>[4]</sup>。该信息融合策略主要包含 3 个步骤。

(1) 应用 3 种分类器对待测样本进行识别,若存在 2 种分类器得出相同的结果,则采用‘少数服从多数’的原则确定最终的识别结果。

(2) 对于 3 种分类器各自给出了不同的识别结果,若基于 BP 算法的 2 种识别方法给出的结果都不是字符“8”,则最终结果为遗传算法的识别结果。因为遗传算法对于除字符“8”外的其他字符的识别效果相对最好,此时应给予该分类器较高的置信度。

(3) 对于 3 种分类器各自给出了不同的识别结果,若基于 BP 算法的 2 种识别方法给出的结果中含有字符“8”,则最终结果为基于总体修改权值的 BP 算法的识别结果。因为此时基于遗传算法的识别结果可信度较低,而基于总体修改权值的 BP 算法的识别率是 2 种 BP 算法中识别率最优的,此时应给予该分类器较高的置信度。

依据上述决策级信息融合策略对同样的样本库进行实验得出的识别率为 98.9%,高于本文 3 种分类器中的任何 1 种,基本符合实际应用的需求,从而达到了本文研究的目的。

字符识别技术是当今模式识别技术研究的热点,字符的种类繁多,对于字符的识别很难找到一种通用的、达到满意效果的方法,所以应根据实际应用具体分析,找到一种解决特定问题的高效方法。基于神经网络的字符识别技术具有一些传统技术没有的优点,如识别速度较快,分类能力强,且具有较好的容错性能和自学能力。

本文提出的基于神经网络信息融合的字符识别方法,在吸取了基于神经网络的字符识别技术的优点外,依据印刷体字符相对稳定的特点,本着简单易行、减少

算法的计算复杂度和识别时间的思路设计分类器,并依据分类器的特点对各分类器给出的识别结果进行决策级的信息融合,最终取得了预期的效果。通过实验表明,此方法对于印刷体字符识别是一种快速、高效、简单易行的方法。本方法对于种类较少、且特征相对稳定的字符进行识别也具有一定的参考价值。

#### 参考文献

- [1] THEODORIDIS S, KOUTROUMBAS K. Pattern re-cognition, third edition[M]. Academic Press, 2006.
- [2] 陈蕾, 黄贤武, 仲兴荣, 等. 基于改进 BP 算法的数字字符识别[J]. 微电子学与计算机, 2004, 21(12): 127-130.
- [3] 温泉彻, 彭宏, 黎琼. 一种自适应遗传 BP 神经网络模型研究及应用[J]. 计算机仿真, 2006, 23(12): 160-162, 166.
- [4] 葛红, 田联房. 信息融合技术在模式识别中的应用[J]. 计算机应用研究, 2009, 26(1): 19-24.
- [5] 郑朝晖, 裘聿皇, 陈峻峰. 一种印刷体字符识别的新方法: 基于遗传算法的(0, 1, \*)-矩阵法[J]. 控制与决策, 2001, 16(3): 296-298, 302.

(收稿日期: 2009-06-15)

电子技术应用  
APPLICATION OF ELECTRONIC TECHNIQUE  
www.chinaAET.com