

14

# 多字体字符识别的模糊神经网络模型\*

王磊 戚飞虎

(上海交通大学计算机科学与工程系, 上海, 200030)

TP391.41  
TP18

**摘要** 提出了一种用于多字体字符识别的模糊神经网络模型. 通过对一个 3 层 MLP 的输入层、输出层以及学习算法的模糊化, 构造出能有效处理具有模糊边界的模式分类问题的模糊神经网络. 经过大量实际采样多字体字符样本的测试表明, 该模型能对字体字符识别取得很高的识别率. 对加噪字符的识别试验还表明该模型具有较好的鲁棒性.

**关键词** 模糊神经网络, 多字体字符识别, 识别率.

## 引言

近年来, 模糊神经网络(Fuzzy Neural Network)正逐渐引起人们越来越多的注意<sup>[1]</sup>, 该技术试图通过模糊逻辑和神经网络的结合来获得单一技术所无法达到的良好性能.

尽管已有很多文献讨论字符识别, 但对手写体和多字体识别问题至今仍未有很好的解决方案. 这是由于它们都面临比较复杂的模式分类情况: 一是需要将一些差异较大的模式归为同一类, 二是需要将一些比较接近的模式归为不同类. 这一对看似矛盾的条件在识别边界模糊的模式中往往需要同时得到满足, 这也正是传统识别方法在这一领域难以奏效的症结所在. 目前正在进行的许多研究都在试图为具有模糊边界的模式识别寻找新的模型. 本文就多种字体印刷字符的识别提出了一种模糊神经网络模型, 并对该模型的识别性能进行了一系列测试, 得到令人满意的结果. 模糊神经网络的实质是对神经网络的模糊化, 模糊化既可以在神经网络的输入、输出和学习算法上进行, 也可以在神经元上进行<sup>[2]</sup>. 目前所出现的一些模糊神经网络类型如模糊 MLP<sup>[3]</sup>、模糊 Kohonen<sup>[4]</sup>和模糊 ART<sup>[5]</sup>等等, 都是基于这一原理构造的. 本文提出的 FNN 模型属于模糊 MLP 类型.

## 1 模糊 MLP 模型

尽管目前有许多神经网络结构可与模糊逻辑相结合, 但考虑到 MLP 仍是目前运用最广的一种神经网络, 故选择它作为构造 FNN 模型的基础结构. 本文的 FNN 模型是通过对一个普通 3 层 MLP 的输入层、输出层以及学习算法的模糊化实现的, 其基本结构如图 1.

### 1.1 输入层的模糊化

为了使 FNN 模型同时适用于二值图像和灰度图象输入, 需要对输入矢量 X 进行归一化操作, 即

$$y_{ij} = x_{ij}/x_{\max}, \quad (i = 1, \dots, M, j = 1, \dots, N) \quad (1)$$

\* 国防预研基金(编号 96J2. 4. 3)和国家自然科学基金(编号 69772002)资助项目  
稿件收到日期 1998-06-12, 修改稿收到日期 1998-09-16

其中  $x_{max}$  为输入矢量  $X$  中各元素的最大值。对归一化以后的输入矢量再进行模糊化，二维模糊化运算可由定义在模糊集合上的一维模糊化运算<sup>[7]</sup>推广而得：

$$z_{ij} = \max_s \{ \max_l [y_M f_{ij,M} - l] \} \quad (2)$$

其中的  $f$  为模糊化函数，模糊化结果取决于模糊化函数的选取。本文选用二维高斯函数来定义模糊函数，即

$$f(i-k, j-l) = \begin{cases} e^{-\beta^2[(i-k)^2 + (j-l)^2]}, & (|i-k| \leq \sigma, |j-l| \leq \sigma) \\ 0, & \text{其他} \end{cases} \quad (3)$$

其中  $\sigma \geq 0$  决定模糊化的区域大小， $\beta$  决定模糊程度，含有不同  $\beta$  参数的  $f$  如图 2 所示。

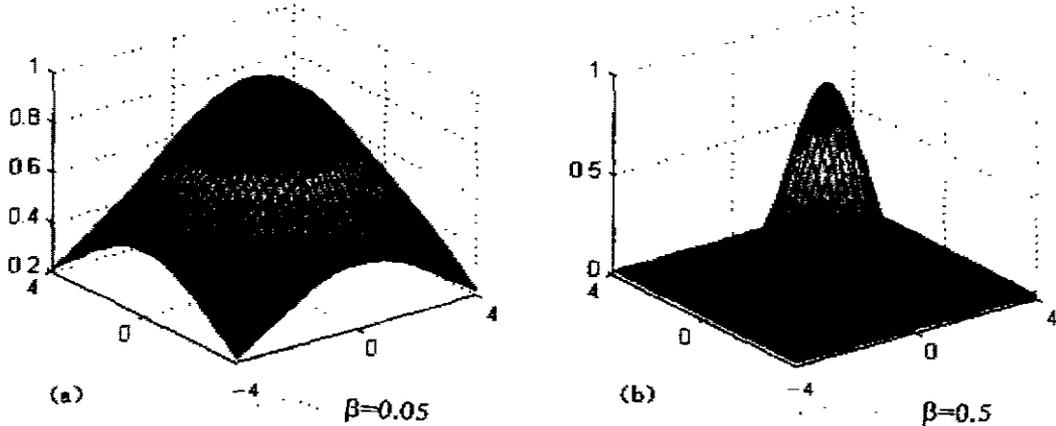


图 2 含有不同  $\beta$  参数的模糊函数  $f$   
Fig. 2 Fuzzy function  $f$  with different  $\beta$

### 1.2 BP 算法和输出层的模糊化

MLP 广泛采用的 BP 算法是一种有教师学习算法，它在训练时需要指定输入矢量与目标输出矢量。设输入矢量为  $X$ ，相应的目标输出矢量为  $O = \{O_i | i=1, \dots, P\}$ ，则有

$$O_i = \begin{cases} 1, & X \in C_i \\ 0, & \text{其他} \end{cases} \quad (4)$$

其中  $C_i$  为第  $i$  类。式(4)说明 MLP 训练时，其输出层只允许一个神经元的输出为 1，其它神经元的输出为 0，显然，这种训练方式是基于分类模式之间具有清晰边界这一假设的。而在实际问题中有很多模式之间是相互重叠的，模式的边界是模糊的，要正确分类这些模式需要借助模糊逻辑这一有力工具。在模糊集合理论中，一个元素可按不同的程度隶属于不同的模糊集合，其程度是由一个取值为  $[0, 1]$  的隶属函数来描述的。这一原则体现在模式识别上就

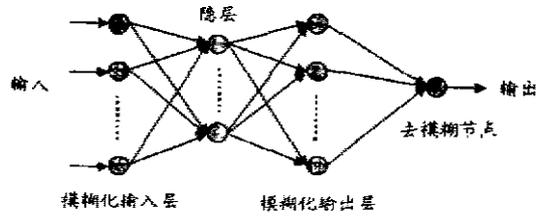


图 1 FNN 模型结构  
Fig. 1 FNN structure

是一个输入模式在输出层将允许有多个非零输出,即

$$O_k = \mu_k(X) \in [0,1]. \quad (5)$$

式(5)中  $\mu_k(X)$  为输入矢量  $X$  对于第  $k$  类的隶属度,在网络结构中引入隶属函数是对神经网络模糊化的关键步骤之一.这样,通过在网络的输出层建立适当的隶属函数来刻划模式间边界的模糊性,从而使得训练后的 FNN 能有效地区分边界模糊的模式.对于

$$D(X,Y) = \sqrt{\frac{1}{MN} \sum_i^M \sum_j^N (x_{ij} - y_{ij})^2}. \quad (6)$$

若  $X$  矢量代表某一模式类  $A$ ,则输入矢量  $Y$  对于这个模式类的隶属函数定义为

$$\mu_A(Y) = 1 - D(X,Y). \quad (7)$$

若  $X=Y$ ,则  $\mu_A(Y)=1$ .一般情况下  $\mu_A(Y)$  在 0 到 1 之间取值.

为了增强同一输入矢量对不同模式类的隶属度之间的差别,可用模糊集中(Fuzzy Concentration)和模糊强化(Fuzzy Intensification)两种运算<sup>[6]</sup>来修正式(7)中定义的隶属函数,即

$$\mu'_\alpha(Y) = [\mu_A(Y)]^\alpha, \quad (8)$$

其中  $\alpha$  为一正整数,用来控制模糊集中的程度;并有

$$\mu'^\lambda_\lambda(Y) = \begin{cases} 2^{\lambda-1} \mu'_\alpha(Y)^\lambda, \\ 1 - 2^{\lambda-1} (1 - \mu'_\alpha(Y))^\lambda, \end{cases} \quad (9)$$

其中  $\lambda$  为一正整数,用来控制模糊强化的程度.

本文提出的 FNN 模型,是通过每个训练样本使用式(6)~(9)以得到相应的模糊输出,从而完成整个训练过程的模糊化.

### 1.3 去模糊

图 1 中的去模糊节点用于将 FNN 的模糊化输出转换成最终的非模糊识别结果.若模糊化输出为  $\{p_i | i=1, \dots, P\}$ ,则最终识别结果  $o$  为  $o = j | p_j = \max(p_i)$ .

## 2 实验结果分析

本文用 3 种字体的 26 个小写字母共 1560 个样本对上述 FNN 模型进行测试,这 3 种字体分别是 Times New Roman, Arial 和 Courier New,字符大小为  $22 \times 28$ .图 3 列出了部分测试样本及其模糊化结果.识别结果见表 1,为对比起见,表中同时给出同样规模的 MLP 识别结果.为了检验本文提出的 FNN 模型的鲁棒性,还使用叠加了随机高斯噪声的字符样本(见图 3(c))对其测试,对 1560 个字符的总识别率高于 96%.

由表 1 可见,FNN 对 3 种字体的识别率均为 100%,性能明显高于相同规模的 MLP.由于 3 种字体的 g 和 t 在字形和大小上差异较大,MLP 因无法容忍这些差异而将它们错分成不同的类,故导致对它们的识别率均低于 80%.字符 o 和 r 因与字符 c 和 l 相似,而 MLP 又不能正确地地区分它们之间的细微差别,因此将它们相互混淆,同样导致识别率的降低.这两种情况正好体现了本文前面所述的一对矛盾,即有些差异大的模式需归于同一类,而有些



图3 部分测试样本及模糊化结果。(a)部分无噪声测试样本,  
(b) (a)的模糊化结果,(c)部分叠加了高斯噪声的测试样本

Fig. 3 Part of examples and their fuzzified results, (a) part of examples with no noise,  
(b) fuzzified result of (a), (c) part of examples with Gauss noise

差异小的模式又需分开成不同类,实验证明 MLP 不具备解决这一矛盾的能力,而 FNN 却很好地解决了这一矛盾,既能容忍同一字符在不同字体间的差异,又能准确地区分不同字符在不同字体上出现的相似性,因此有效地提高了识别率。正因为这些优点,FNN 非常适合模糊边界模式识别,另外,对加噪字符的识别结果也证明了 FNN 具有良好的鲁棒性。

表1 FNN 与 MLP 的多字体字符识别性能对比

Table 1 Performance comparison between FNN and MLP for multi-font character recognition

网络	字体	b	e	g	h	n	o	r	t	u	其他
FNN	T. N. R Arial Courier					100%					100%
MLP	T. N. R Arial Courier	85%	95%	70%	100%	90%	100%	45%	80%	95%	100%
		95%	100%	75%	100%	100%	70%	100%	70%	100%	
		100%	90%	80%	90%	100%	80%	100%	80%	100%	

### 3 结论

本文讨论了一种用于多字体字符识别的模糊神经网络模型,旨在能够改善传统神经网络在处理模糊边界模式分类时的性能,初步实验表明该模型不仅能很好的识别多字体字符,而且对于加噪字符也有着很好的适应性。可见,与传统神经网络相比,FNN 在性能上表现出明显的优势,但仍有很多问题待进一步研究,以期进一步提高性能。例如,基于矢量间的欧氏距离构造出的输出层隶属函数是否最有利于模糊边界模式的分类,以及能否利用多个非零模糊输出之间的内在联系来进一步提高识别率,这些问题都值得进行更深入的研究和探讨,相信随着对模糊神经网络研究的不断深入,其应用领域也会不断得到扩展,应用效果也会不断提高。

## REFERENCES

- 1 Buckley J J, Hayashi Y. Fuzzy neural networks, A survey, *Fuzzy Sets and Systems*, 1994, **66**(1):1-13
- 2 Pal S K, Srimani P K. Neurocomputing Motivation, Models, and Hybridization, *Computer*, 1996, **29**(3):24-28
- 3 Pal S K, Mitra S. On Fuzzy Neuron Models, *IEEE Transactions on Neural Networks*, 1992, **3**(5):683-697
- 4 Mitra S, Pal S K. Multilayer Perceptron, Fuzzy Sets, and Classification, *IEEE Transactions on Systems, Man and Cybernetics*, 1994, **24**(3):385-399
- 5 Carpenter G. Self-organizing neural networks as a fuzzy classifier, *IEEE Transactions on Neural Networks*, 1992, **3**(5):698-713
- 6 Klir G J, Folger T A. Fuzzy ARTMAP: A Network Architecture for Incremental Supervised Learning of Analogue Multi-dimensional Maps, *Fuzzy Sets, Uncertainty, and Information*, Englewood Cliffs: Prentice-Hall, 1988:5-17
- 7 Klir G J, Folger T A. *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

## FNN MODEL FOR MULTI-FONT CHARACTER RECOGNITION\*

WANG Lei QI Fei-Hu

(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

**Abstract** A novel fuzzy neural network (FNN) model for multi-font character recognition was presented, which can efficiently process the fuzzy pattern classification problem. This FNN model is built by fuzzifying the input layer, output layer and the training algorithm of a conventional multilayer perceptron (MLP). The simulation with a lot of multi-font character samples shows that the FNN presented here can get a high recognition rate, and has low sensitivity for different character fonts in comparison with classical MLP. Also, this FNN is proved to have a good robustness.

**Key words** fuzzy neural network (FNN), multi-font character recognition, recognition rate.

---

\* The project supported by the Foundation of Preliminary Research in National Defense and the National Natural Science Foundation of China  
Received 1998-06-12, revised 1998-09-15