

基于多层神经网络的非线性图像分割

郭 平^a 刘大禾^b 崔建生^a

a, 北京师范大学分析测试中心, 北京 100875
b, 北京师范大学物理系, 北京 100875

摘 要 提出了一种用多层神经网络对图像进行非线性分割的方法, 讨论了所用多层神经网络的学习速度的改进与训练样本的选择方法。实验表明, 该多层神经网络系统可用于实时图像分割, 并能获得很好的结果。

关键词 多层前馈神经网络, 实时学习, 非线性图像分割。

1 引 言

在数字图像处理中, 常常要对一幅图像中一些较为复杂封闭区域进行分割, 以提取所需要的部分。常用的分割方法是门限技术。但对于复杂的图像, 其灰度分布极为复杂。采用门限技术常常会使被分割区域内出现“孔”, 而区域外出现噪声。另外, 不同的图形可以有相同的灰度, 不同的图形也可能重迭。这时, 门限技术都显得无能为力。例如在染色体样品制作过程中, 常常因为散开性不好, 造成两条或多条染色体的重迭, 给自动图像处理及染色体识别带来极大的困难。如果把重迭的染色体部分抛弃不用, 则平均只有 83% (38.2/46) 的可用率, 很可能造成重要信息的丢失。因此, 必须寻找其他的分割方法。

多层前馈神经网络^[1-3]是目前应用最为广泛的一种人工神经网络, 可用于语言综合、语音识别、图形分类、自适应控制等方面。这种网络的优点之一是对任意复杂图形的识别, 只要经过对给定样本的学习后, 即可进行正确率相当高的分类。若把图像分割问题看成是一种对空间坐标点的分类, 则可用多层前馈神经网络解决门限技术无法解决的困难。本文讨论了用多层前馈神经网络实现重迭图形的分割。

2 多层前馈神经网络

一般来说, 通常的图像分割可对应于把二维空间坐标点分成两类。一类是所需要的区域, 另一类是其他区域。定义图像分割为空间坐标点 A 类与 B 类的分离, 则对图像分割所用的多层神经网络问题的数学描述为:

$$Net_{output}(x, y) = \begin{cases} 1 & (x, y) \in A \\ 0 & (x, y) \in B \end{cases} \quad (1)$$

式中 (x, y) 为空间坐标点，网络输出是与坐标点相关的。只要把图像像素的坐标输入到经过训练的神经网络，根据其输出则可判定该点属于哪一类区域，从而分离 A 、 B 两类区域。

多层前馈神经网络一般有一个输入层、一个输出层和几个隐含层。Lippmann 指出：具有两个隐层的神经网络就可对任意复杂的区域进行正确分类^[2]。实际应用中，多层神经网络结构中的隐层数目不会超过两个。对一幅图像所关心的区域不是非常复杂的情况，如对“X”型交叉染色体图像，有一个隐层的网络结构就可以解决分割问题。

对多层前馈神经网络的训练，通常采用误差背向传播算法，但该算法是一种梯度下降算法，收敛速度慢，有时会陷于局部极小。这种算法还有一个缺点，即需要使用者选择如步长、动量项等学习参数。然而，到目前为止，尚没有理论指导学习参数的选择，只有通过多次调试才能获取较迁当的参数。

本文采用了一种新的三层神经网络的学习算法，不需要使用者调试学习参数，而且在样本数目不是很多的情况下，能实现快速、精确的学习^[4]。下面对本文所采用的神经网络结构与算法作一简介。

对于只有一个输入层、一个隐层和一个输出层的三层前馈神经网络，假定矩阵 V 连结输入层与隐层的神经元，矩阵 W 连结隐层与输出层的神经元。隐层的激活函数采用 Sigmoid 型，即：

$$f(Net) = 1/[1 + \exp(-Net)] \tag{2}$$

式中 (Net) 表示函数变元。输入层上有两个输入神经元和一个常数偏置为 1.0 的神经元，输出层只有一个神经元。隐层神经元数目选择与学习样本数目一致，假定为 N ，图 1 为此种网络结构。学习样本中的所有输入信号组成矩阵 X 。矩阵 X 中最后一列是常数 1.0，每行表示一个输入样本，每行的形式是 $(x, y, 1.0)$ 。而所有期望输出信号组成目标矩阵 T 。矩阵 T 中每行表示一个期望输出样本，与输入样本一一对应。则网络结构可用矩阵形式表示为：

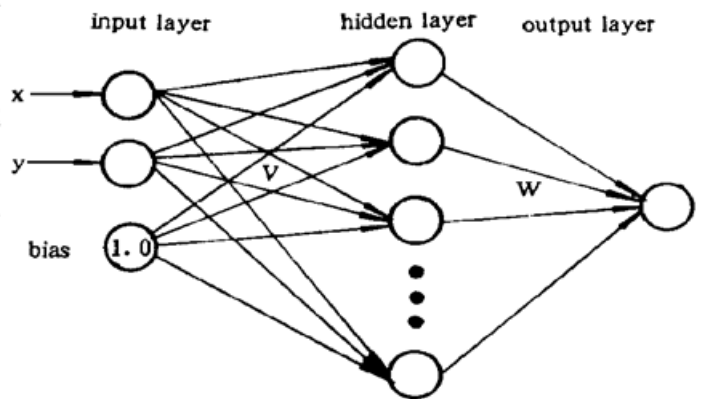


Fig. 1 The network used for image segmentation architecture sketch map

$$T = f(XV)W \tag{3}$$

在本文中， X 为 $N \times 3$ 矩阵， V 为 $3 \times N$ 矩阵， W 、 T 均为 $N \times 1$ 矩阵。

在数学上，训练网络的目的就是在给定的输入输出样本组 X 、 T 下，找到互联权重矩阵 V 、 W ，使得误差：

$$E = \|T - O\| \tag{4}$$

最小。这里 $O = f(XV)W$ 是网络的实际输出，理想情况下，(3) 式成立。本文提出的算法采用批处理方式，其过程如下：

- 1) 计算 $X^+ = \text{Pseudoinverse}(X)$ ， X^+ 为 X 的广泛逆矩阵。
- 2) 计算 $Y = f(XX^+)$ ， $f()$ 为(2) 式所表示的激活函数。
- 3) 计算 $Y^+ = \text{Pseudoinverse}(Y)$ ， Y^+ 为 Y 的广义逆矩阵。
- 4) 计算 $W = Y^+ T$ ， $V = X^+$ 。

5) 实际输出 $O = f(XV)W$, 结束训练。

由于采用了一次计算权重矩阵, 而不是误差背向传输算法中的梯度下降法, 且无需选择学习常数, 故学习过程较误差背向传输算法要快, 可实时学习, 而且效果好。

3 学习样本提取分析

由于多层前馈神经网络采用的是有教师指导学习法, 对要解决的实际问题, 首先要选择适当的学习样本对网络进行训练。学习样本选择的好坏对网络的归纳能力有很大的影响。为了测试所提出以及训练后的网络对空间分类的能力, 首先用下例对网络进行训练, 然后再用训练好的网络进行分类。

多层神经网络与感知器的最大区别是在于它可进行非线性分类。通常用异或(XOR)问题来测试学习算法, 本文选一较异或更为复杂的非线性分割区域问题来测试学习算法与网络结构, 即将异或问题扩展为:

在正方形 $0 \leq x \leq 1, 0 \leq y \leq 1$ 的区域内, 选择 12 组点作为学习样本, 对网络进行训练。其中 6 组样本为 A 类, 输出期望值 1.0, 6 组样本为 B 类, 输出期望值 0。图 2(a) 是样本点的坐标。相应的输入、输出矩阵为

$$X = \begin{bmatrix} 1.0 & 1.0 & 0.0 & 0.0 & 0.2 & 0.8 & 0.2 & 0.8 & 0.6 & 0.4 & 0.4 & 0.6 \\ 1.0 & 0.0 & 1.0 & 0.0 & 0.8 & 0.2 & 0.2 & 0.8 & 0.6 & 0.4 & 0.6 & 0.4 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \end{bmatrix}^t,$$

$$T = (0.0, 1.0, 1.0, 0.0, 0.0, 0.0, 1.0, 1.0, 0.0, 0.0, 1.0, 1.0)^t。$$

采用上节的学习算法, 计算得到的权重互联矩阵为:

$$V = \begin{bmatrix} a_1, & a_1, & -a_1, & -a_1, & -a_2, & a_2, & -a_2, & a_2 & a_3 & -a_3, & -a_3, & a_3 \\ a_1, & -a_1, & a_1, & -a_1, & a_2, & -a_2, & -a_2, & a_2, & a_3 & -a_3, & a_3, & -a_3 \\ -a_4, & a_5, & a_5, & a_6, & a_5, & a_5, & a_7, & a_8, & a_9, & a_{10}, & a_5, & a_5 \end{bmatrix}$$

其中 $a_1 = 0.357143, a_2 = 0.214286, a_3 = 0.0714286, a_4 = 0.27381, a_5 = 0.083333, a_6 = 0.440476, a_7 = 0.297619, a_8 = 0.130952, a_9 = 0.0119048, a_{10} = 0.154762$ 。

$W = (w_1, -w_1, -w_1, w_1, w_2, w_2, -w_2, -w_2, w_3, w_3, -w_3, -w_3)^t$, 其中 $w_1 = 1.58891 \times 10^8, w_2 = 1.30458 \times 10^9, w_3 = 7.76924 \times 10^9$ 。

上面各式中的上标符号 t 表示矩阵的转置。该例的学习误差 E 为 6×10^{-11} 量级, 可以说达到了完备学习。

利用训练好的神经网络对 $0 \leq x \leq 1, 0 \leq y \leq 1$ 的正方形区域进行实验分割时, 还需说明一点。由于网络对不在训练集中的点的响应与在训练集中的样本点有所差别, 输出值将不会正好是 1.0 或 0, 在网络输出端加一阈值器。考虑 A 类样本输出为 1.0, B 类样本输出为 0, 其它不在训练集中的点输出介于 0 与 1.0 之间, 取中间值 0.5 作为阈值。若输出 $O_i > 0.5$, 则输入点属于 A 类, 反之则属于 B 类。显然阈值为 0.8 时 B 类比较占优, 而 0.3 时 A 类占优。阈值为 0.5 时分割结果如图 2(b) 所示。阈值为 0.8 时, 分割结果如图 2(c) 所示。

可以看出, 网络能对并不相邻或封闭的区域进行分割, 而这些区域是不能线性分离的。反过来说, 只要选择适当的样本及阈值, 则可把区域分割为较复杂的子区域。这表明, 三层神经网络可以解决门限技术无法实现的非线性图像分割问题, 关键在于学习样本与阈值的选

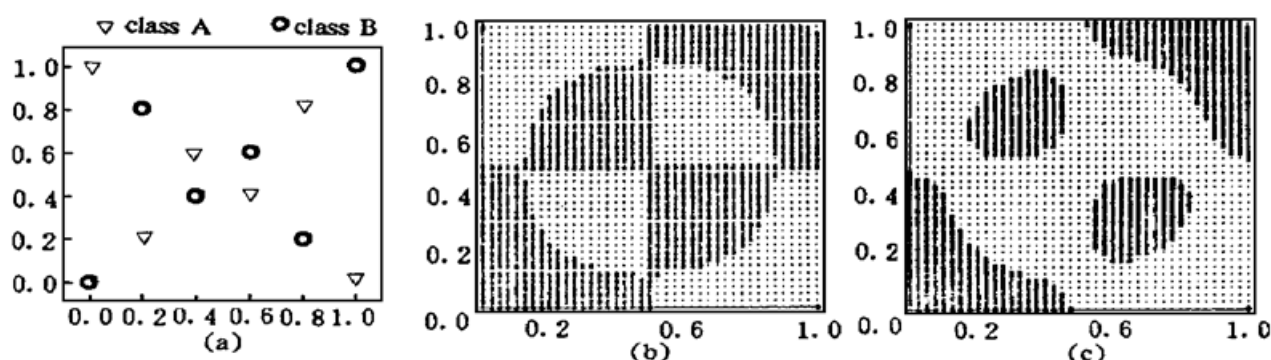


Fig. 2 The learning examples and segmentation results. (a) The examples used to train the network. (b) Segmentation result with threshold 0.5. (c) Segmentation result with threshold 0.8. Thin dot area is for class A, while thick dot area for class B.

取。

需要指出的是, 只有一个隐层的神经网络对区域的分割能力与 Lippmann 的结论^[2]并不一致。上面例子表明, 可以把两个不相邻的封闭区域划分为一类, 而 Lippmann 认为有两个隐层的神经网络才能做到这一点。这个问题还需要作进一步的理论研究。

对一幅实际图像进行基于多层神经网络的分割时, 只有适当选取某些像素点的坐标作为学习样本来训练神经网络, 网络才能对图像进行正确的分割。通过分析上例可知, 神经网络的分区特点是比较平滑的。若样本点选择得比较对称, 则分区结果也比较对称。在 x 或 y 方向, 两类点都扩展到均分区域处。而在对角线方向, 区域中每个点, 都有向外扩展的趋势, 直至遇到另一类点作为边界。因此, 选择区域边界点作为学习样本, 会有较好的分割效果。而且只要样本选择适当, 其数目就可大大减少。这对提高网络的速度极为有利。

对于不同的图像, 要求分割的区域不同, 而学习样本的选择取决于被分割区域的边界, 故学习样本的选择需视具体图像而定。对于染色体交叉图像的分割, 所采用的学习样本选择方法为:

- 1) 用边缘检测算法找出整个区域边缘;
- 2) 将区域边缘二值化, 得到描述区域边缘的函数 $g(x_i, y_i)$, (x_i, y_i) 为边缘像素点的位置坐标;
- 3) 对 $g(x_i, y_i)$ 求一阶和二阶导数, 分析得到四个交叉边缘点的信息;
- 4) 沿交叉处四个点相对应的方向沿边缘向外延伸, 得到一条染色体的边缘信息, 这些点构成 A 类学习样本;
- 5) 将上述坐标适当向外扩展, 得到 B 类学习样本。

4 实验结果

图像分割实际上是利用神经网络的拓扑性质, 对图像进行具有最小方差意义上的分割。图 3(a) 是由显微镜连接 CCD 摄像机得到的两条交叉染色体的数字图像。图 3(b) 是采用上述三层神经网络对其中一条染色体分割的结果。图 3(c) 是重新选择样本点训练网络后, 对另一条染色体分割的结果。可以看到, 分割是相当成功的。

最后强调指出, 本文所讨论的基于多层前馈神经网络的图像分割是对空间坐标, 而不是对灰度的分类。它可以解决门限技术无法解决的问题, 因而具有广泛的应用前景。

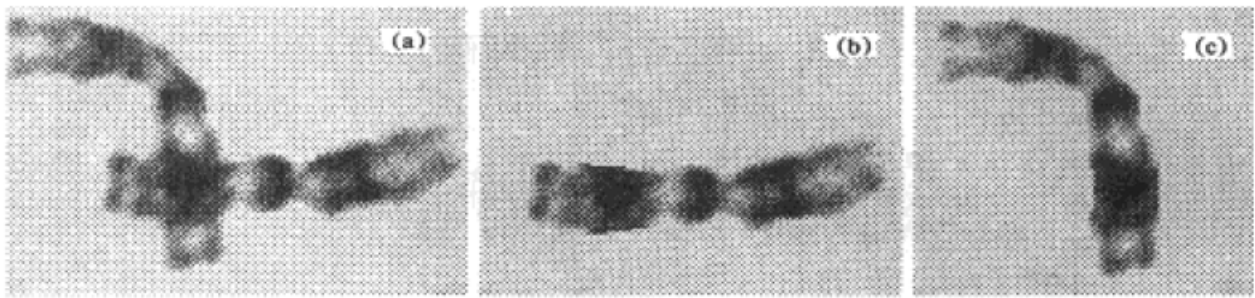


Fig. 3 The crossing chromosome and segmented results. (a) The crossing chromosome digital image. (b) and (c) segmentation results using trained networks with different learning examples

参 考 文 献

- [1] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning internal representations by error propagating, in D. E. Rumelhart, J. L. Clelland Eds., *Parallel Distributed Processing*, Cambridge, MA: MIT Press, 1986, **1**: 318~ 362
- [2] R. P. Lippmann, An introduction to computing with neural nets, *IEEE ASSP Magazine*, 1987, **4**(1): 4~ 22
- [3] Kelvin Wagner, Demetri Psaltis, Multilayer optical learning networks, *Appl. Opt.*, 1987, **26**(23): 5061~ 5076
- [4] Ping Guo, C. L. Philip Chen, Yinguan Sun, A exact supervised learning for a three-layer supervised neural network, in Yixin Zhong, Yixian Yang, Minghui Wang Eds., *Proceedings of International Conference On Neural Information Processing*, Beijing: Publishing House of Electronics Industry, 1995, 1041~ 1044

Image Nonlinear Segmentation Based on Multilayer Neural Networks

Guo Ping^a Liu Dahe^b Cui Jiansheng^a

[^a, *Analysis and Testing Center, Beijing Normal University, Beijing 100875*
^b, *Department of Physics, Beijing Normal University, Beijing 100875*]

(Received 10 September 1995; revised 8 January 1996)

Abstract A new image nonlinear segmentation method based on multilayer neural network is proposed in this paper. The technique of improving the learning speed and the method of selecting learning patterns are discussed. The experiment shows that proposed multilayer neural network can be used for real time image segmentation and the good results can be obtained.

Key words multilayer feedforward neural network, real time learning, image nonlinear segmentation.