

基于改进 BP 神经网络的手写体数字识别

何松¹ 戚建宇²

(常州工学院计算机信息工程学院 常州 213002)¹ (常州工学院延陵学院 常州 213002)²

摘要 数字识别在许多领域有广泛的应用。通过对人工神经网络的研究与学习,运用改进的 BP 神经网络对无约束手写体数字识别过程中的数字样本进行识别。实验证明,该方法具有很强的抗干扰性,克服了传统 BP 算法的局限性,其识别率和准确率都有很大提高。

关键词 神经网络,数字识别,改进 BP 网络

Handwritten Numeral Recognition Based on the Improved BP Neural Network

HE Song¹ QI Jian-yu²

(School of Computer Information & Engineering, Changzhou Institute of Technology, Changzhou 213002, China)¹

(School of Yanling, Changzhou Institute of Technology, Changzhou 213002, China)²

Abstract The digital recognitions have extensive application in a lot of important fields. The artificial network was studied in the paper. The figure specimens were filtered through the process of unconstrained handwritten numeral, that it was recognized by the improved BP neural network. The result of experiment demonstrates that the method has very high noise immunity capacity and overcomes the limitation of traditional BP algorithm. The recognition rate and precision rate are greatly improved at the same time.

Keywords Neural network, Digital recognition, Improve BP network

数字识别前景广阔,广泛应用于汽车牌照的数字识别、表格中数字的自动识别和个人成绩单的识别等多个项目,涉及到教育、交通、商业和邮政、银行等多个领域。数字的自动识别系统的实现和应用给人们的生活和工作提供了很大的方便。相对于印刷体的数字识别,无约束手写体数字识别是模式识别领域的难点,也是现在的一个研究热点。由于手写数字的样本库较少也较为复杂,因此它成为检测各种识别算法优劣的重要手段。近几年众多学者对手写体的识别进行了较多研究,提出了多种的算法^[1-4]。不过当前运用较好的主流算法还是以统计、神经网络和聚类分析的识别算法为主,如支持向量机算法(Support Vector Machine, SVM)、误差反向传播(Back Propagation, BP)算法、Bagging 算法等。

神经网络是近几年发展起来的一门新兴学科,具有很强的学习性、自适应性。神经网络具有并行处理、容错性、自学习功能等多种独有特性,有别于传统方法。目前神经网络已在模式识别、自动化控制等领域实现应用并取得了较好的效果^[5]。BP 神经网络是目前应用最为广泛的一种网络模型。由于 BP 神经网络对输入与输出元素之间存在复制的多元非线性关系,具有很强的处理能力,因此对手写体数字识别有较大的优势。但是在该网络运用的同时,我们也发现其仍存在很多不足之处,如学习速度慢,网络训练失败可能性大,训练结果局部收敛。本文通过对 BP 神经网络进行改进,并用改进的 BP 神经网络完成对手写体数字进行识别。结果显示,其识别效果相当理想。

1 BP 神经网络的改进^[6-8]

BP 神经网络一般在逻辑上分为 3 层,分别为输入层、隐含层和输出层,具体结构如图 1 所示。设网络有 M 个输入节点—— M 个输入特征量分别用 Net_i 表示,有 n 个节点隐含层, W_{ij} 是输入层和隐含层节点之间的连接权值,输出层有 L 个节点。隐含层及输出层节点的输入值为前一层节点输出的加权和,每一个节点的激励程度由它的激发函数来决定。

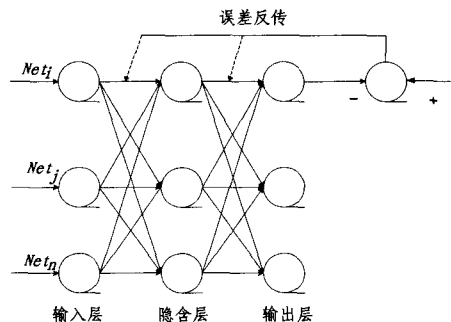


图 1 BP 网络

本文为了表达方便,设如下规定。

- (1) m : 处理层号,且 $m=0, 1, \dots, M-1$, 其中 $m=0$ 表示输入层, $m=M-1$ 表示输出层。
- (2) Nm : 第 m 层所有节点的数目。
- (3) y_i^m : 第 m 层第 i 个处理节点的输出值。
- (4) net_i^m : 第 m 层第 i 个处理节点的输入值。其中

$$y_i^m = \sum_{j=1}^M W_{ij} net_i^m \quad (1)$$

$$net_i^m = f(y_i^m) \quad (2)$$

在训练网络学习阶段,设有 P 个训练样本 X_p , 其中 X_p 对应一个 N 维输入特征矢量, 即 $X_p = \{x_1, x_2, \dots, x_p\}$ 。其相对应的导师信号为这 p 个样本的主观输出值 $D_p = \{d_1, d_2, \dots\}$, D_p 对应一个 N 维的主观评价分。其对应的网络客观评价输出 $Y_p = \{y_1, y_2, \dots, Y_{yp}\}$, Y_p 对应的是一个 N 维的客观评价分。

定义系统的误差函数:

$$E = \sum_{p=1}^P E_p(w, f, D_p, X_p) = \frac{1}{2} \sum_{p=1}^P \sum_{j=0}^{N(m-1)-1} (d_{pk} - Y_{pk})^2 \quad (3)$$

第 k 层中第 i 个神经元的输入为 net_i^k , 输出 Y_i^k , 则输出与输入的关系为:

$$Y_j^k = f\left[\sum_{i=1}^{N_{k-1}} w_{ij}^{(k-1)} \cdot net_i^k + \theta_j^k\right], \quad (4)$$

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

式中, w_{ij}^{k-1} 为第 $k-1$ 层第 i 个神经节点到第 k 层第 j 个神经节点的权值, 其中 $f(x)$ 为神经元节点的传递函数。

传统的 BP 网络的优点明显, 但它存在着容易陷入局部最小, 学习速度慢, 网络隐含层及其单元个数难于确定以及会出现过拟合现象等缺陷, 本文将从以下 3 方面对传统的神经网络进行改进。

1.1 神经元函数输出的限制

在传统的 BP 网络中, 权值校正量和误差项都与隐含层的实际输出有关。当神经元传递函数 $f(x)$ 输出为“0”值或“1”值时, 实际连接校正量不起作用, 因而在其后的较多次的学习过程中, 真正的校正量很小。因此, 为了提高网络的收敛速度, 对 $f(x)$ 函数做如下限制: 若激励函数 $f(x) < 0.01$, 令 $f(x) = 0.01$; 若 $f(x) > 0.99$, 令 $f(x) = 0.99$ 。但是这种方法容易导致激活函数的 $\exp$ 的值溢出, 为此, 可以这样定义: 若 $x > 100$, 令 $x = 100$; 若 $x < -100$, 令 $x = -100$ 。

1.2 添加“惯性项”

传统 BP 网络中通过样本训练调整网络参数, 使得网络实际评测输出与理想教师信号输出的误差小于允许误差, 从而逼近客观网络输出结果到输入的理想主观评测映射。BP 网络的权值的和阈值的调节为:

$$\begin{aligned} \Delta v_{ji}(n+1) &= \alpha_1 d_i^k b_j \\ \Delta \gamma_i(n+1) &= \alpha_1 d_i^k \\ \Delta w_{ij}(n+1) &= \alpha_2 e_j^k x_i^k \\ \Delta \theta_j(n+1) &= \alpha_2 e_j^k \end{aligned}$$

上式中, 学习速率 α_1 和 α_2 在实际的学习过程中有很大作用。 α_1 和 α_2 越大, 权值和阈值的变化率就越为强烈, 网络实际输出就越容易收敛。一般是在不导致振荡的前提下取尽量大的 α_1 和 α_2 值。本文中为了使 BP 网络学习速度足够快而又不易产生振荡, 给它加上一个“惯性项”, 即

$$\begin{aligned} \Delta v_{ji}(n+1) &= \alpha_1 d_i^k b_j + \eta \Delta v_{ji}(n) \\ \Delta \gamma_i(n+1) &= \alpha_1 d_i^k + \eta \Delta \gamma_i(n) \\ \Delta w_{ij}(n+1) &= \alpha_2 e_j^k x_i^k + \eta \Delta w_{ij}(n) \\ \Delta \theta_j(n+1) &= \alpha_2 e_j^k + \eta \Delta \theta_j(n) \\ \Delta E(N) &= E(N) - E(N-1), \phi > 1, \beta < 1 \end{aligned}$$

如果 $\Delta E(N) = [E(N) - E(N-1)] > 0$, 表示学习误差增大, 当前的网络客观输出值正向偏离理想的期望值, 这时需要减小权值的调整量, 从而减小学习率, 甩掉惯性项。如果 $\Delta E < 0$, 表示梯度修改方向正确, 这时加入惯性项, 增大学习率, 大大提高 BP 网络学习效率从而网络快速收敛。

1.3 能动的改变权值和阈值

为使网络学习过程更快地跳出局部最小, 可做如下规定: 如果 $\Delta E(N)$ 连续多次小于一个很小的数, 这表示按当前网络的权值和阈值的调节规律, 对于全局误差的减小不会有很多的帮助, 于是令权值跟阈值加上一个很小的随机数从而避开局部最小。

2 基于改进 BP 神经网络的识别方法

BP 神经网络具有高度的自适应性, 能从样本的训练中学习问题的规律, 对相互交叉的识别因素起到较好的筛选作用, 从而为后续的识别提供有效的帮助。本文提出了将改进 BP 神经网络运用于数字识别中, 其基本的思想是通过改进 BP 网络对数字样本训练, 完成数字识别。如图 2 所示。

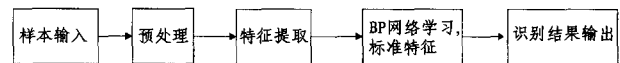


图2 识别系统

2.1 预处理

预处理过程一般包括二值化、噪声点去除、数字特征分离、图像平滑、归一化处理和细化等过程, 不同的识别算法的预处理项目和要求有所不同。在该识别系统的设计中, 预理由二值化 (binarization)、归一化 (normalization) 和平滑 (smooth) 3 大过程组成。首先对图像字符分割处理后得到单个数字图像, 并进行归一化处理, 因为扫描出来的图像中字符的大小存在较大的差异, 而相对来说, 统一尺寸的字符识别的标准性更强, 准确率也更高, 处理起来相对简单。通过归一化以消除各数字在位置和大小上的差异, 从而提高数字识别的准确率。归一化处理后的字符便可进行特征向量提取。

简单地说, 对其字形有笔划的部分, 其位点为黑像素, 设其位点值为“1”, 字形的空白背景部分, 其位点为白像素, 设其位点值为“0”。因此, 一幅单个数字的图像就变成由“0”和“1”组成的二值数据矩阵。本文将以预处理后的 16×16 二值点阵为例进行讨论。

2.2 特征提取和分析^[9]

特征提取方法的选择是影响识别率的一个十分重要的因素。必须指出的是, 对于不同的识别问题和不同的样本数据, 使用不同的特征提取方法的适应性差别很大。在一定意义上, 特征提取和特征选择的主要目的都是要达到特征降维。特征提取就是要通过某种变换的方法重新组合原始高维特征, 获得一组较为低维的新特征, 而特征选择是根据某种评价准则或根据专家的经验知识来挑选出那些对分类最具影响力的特征, 并生成新的较易处理的特征。

不同的字符由于笔划不同, 其像素点的分布也不同, 因此造成不同字符的重心位置不同, 与其有关的一些离散量包含了字符几何特征的信息。本文采用重心及重心矩特征。令 f_{mn} 表示点阵第 m 行, 第 n 列的像素, 定义:

$$\bar{m} = \frac{\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f_{mn} \frac{m}{N-1} \frac{m}{M-1}}{\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f_{mn}} \quad (5)$$

(下转第 248 页)

了研究。相关结论揭示了软件系统的动态执行特性,对于软件运行维护及质量保障具有重要的指导意义。

参考文献

[1] 岳昆,王晓玲,周傲英. Web 服务核心支撑技术,研究综述[J]. 软件学报,2004,15(3):428-442

[2] Barabasi A L, Albert R. Emergence of scaling in random networks. Science,1999,286(5439):509-512

[3] Newman M E J. The structure and function of complex networks[J]. SIAM Review,2003,45:167-256

[4] Valverde S,Ferrer-Cancho R,Sole R. Hierarchical Small Worlds in Software Architecture, Santa Fe Inst[J]. Working Paper, 2003:SFI/03-07-044

[5] 马于涛,何克清,李兵,等. 网络化软件的复杂网络特性实证[J]. 软件学报,2011,22(3):381-407

[6] Myers C. Software system as complex networks:structure, function, and evolvability of software collaboration graphs[J]. Physi-

cal Review E 68(2003) 046116,1-046116. 15

[7] de Moura A,Lai Y-C,Motter A, Signature of small-world and scale-free properties in large computer programs[J]. Physical Review E 68(2)(2003) 017102,1-017102. 4

[8] Concas G, Marchesi M, Pinna S, et al. Power-laws in a large object-oriented software system[J]. IEEE Transactions on Software Engineering,2007,33(10):687-707

[9] 闫栋,祁国宁. 大规模软件系统的无标度特性与演化模型[J]. 物理学报,2006,55:3799-3804

[10] 韩言妮,李德毅,陈桂生. 软件网络的多粒度拓扑特性分析及其应用[J]. 计算机学报,2009,32(9):1711-1721

[11] Cai K Y, Yin B B. Software Execution Processes as an Evolving Complex Network[J]. Information Sciences, 2009, 179 (12): 1903-1928

[12] Potanin A,Nobbe J,Frean M, et al. Scale-free geometry in object-oriented programs[J]. Communication of the ACM, 2005 (48):99-103

(上接第 205 页)

$$\bar{n} = \frac{\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f_{nm} \cdot n}{\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f_{nm}} \quad (6)$$

式中, $m=1,2,\dots,M-1, n=1,2,\dots,N-1$; $(\bar{m}, \bar{n})$ 就定义为整个字符的中心位置。定义中心矩如下:

$$L_{i,j} = \sum_{n=(j-1)\frac{N}{2}}^{\frac{N}{2}} \sum_{m=(i-1)\frac{M}{2}}^{\frac{M}{2}} (m-\bar{m})(n-\bar{n}) \quad (7)$$

式中, $i=1,2; j=1,2$, 这样可以得到 4 个重心矩特征: $L_{ij} (i=1,2; j=1,2)$ 。它们分别反映 4 个象限中像素偏离重心位置的整体特征。本文中选用重心及重心矩特征计算后点阵的重心位置,以及 4 个象限的重心矩,共 5 个特征,包括一个二维特征,4 个一维特征。

2.3 BP 神经网络的学习和训练

从特征的选取过程可以发现,被选取的特征因素很可能存在相互交叉现象。由于多层 BP 网具有自学习、容错性、分类能力强和并行处理等特点,因此用它来训练与识别手写体数字比较合适。文中使用的 BP 网络分为 3 层,其输入层有 6 个点,分别对应特征矢量的 6 个分量,其输出层有 10 个输出点。通过对数字样本进行训练,将学习的结果纳入到聚类源。

3 结果分析

表 1 实验结果

样本输入	正确率	误识率	拒识率
测试样本 1(印刷体)	95.50%	2.50%	2.0%
测试样本 2(手写体)	90.10%	7.20%	2.70%
测试样本 3(手写体)	93.10%	5.30%	1.60%
测试样本 4(手写体)	95.20%	4.10%	0.70%
测试样本 5(手写体)	92.60%	2.50%	4.90%

提取测试片断,按上述方法对每个手写体数字提取特征矢量,得到 5 个测试样本,每个样本有 115 个待识别数字,每个数字的特征矢量为 6 位。对于这个输入样本,经过上述 BP 网络学习以后,将得到一个十维的输出向量。在该 BP 网络中,选择 η 的初始值为 0.15,然后用时间退火函数让 η 值慢慢减小。最后,选择定势态因子 α 为 0.175,学习结束的条件为

网络的均方根 $\Delta E=0.001$ 。实验结果见表 1 及图 3 所示。

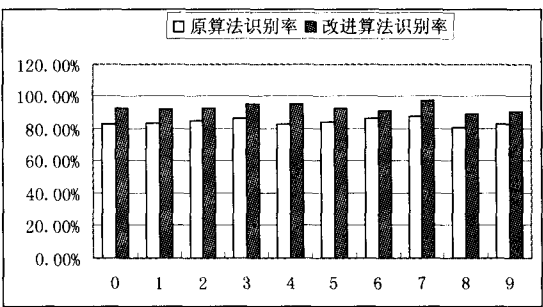


图 3 两种 BP 算法分别对样本中 0~9 的识别率

参考文献

[1] 刘勇,赵斌,夏绍玮. 模糊超椭圆分类算法及其在无约束手写体数字识别中的应用[J]. 清华大学学报:自然科学版,2000,40(9):120,124

[2] Demirekler M, Altincy H. Plurality Voting-based Multiple Classifier Systems; Statistically Independent with Respect to Depend Classifier Sets[J]. Pattern Recognition,2002,35:2365-2379

[3] Altineay H, Demirelder M. Undesirable Effects of Out-put Normalization in Multiple Classifier Systems[J]. Pattern Recognition Letters,2003,24:1163. 1170

[4] Bezdek J. Pattern Recognition with Fuzzy Objective Function Algorithms[M]. New York: Plenum Press,1981

[5] 胡瑞敏. 广义神经网络理论和信号识别系统的研究[D]. 武汉:华中理工大学,1994,9

[6] 王爱国,肖定中. 一种改进的 BP 神经网络模型及其在语音识别中的应用[J]. 光通信研究,1996

[7] 陈敏,刘君. BP 网络的改进及其应用[J]. 湖南文理学院学报, 2005

[8] 戚建宇,何松. 改进 BP 神经网络的普通话单字发音标准度研究[J]. 福建电脑,2009

[9] 董慧. 手写体数字识别中的特征提取和特征选择研究[D]. 北京:北京邮电大学,2007