

改进的 BP 网络在纸病识别中的应用

魏爱娟¹, 李 茜¹, 汤 伟²

(1. 陕西科技大学 电气与信息工程学院, 西安 710021; 2. 陕西科技大学 造纸工程学院, 咸阳 710021)

摘要:针对BP网络收敛速度慢、计算量大等缺点,采用学习样本产生的总误差调整权值,以提高BP算法的速度。文中通过仿真说明其算法的有效性,对改进BP算法用于纸病检测有一定的参考价值。实践表明,改进的BP网络能够满足在线纸病检测要求,平均识别率达94%以上。

关键词:纸病;BP网络;模式识别;在线检测系统

中图分类号:TP183 **文献标志码:**B

Application of Improved BP Network in the Paper Defect Recognition

WEI Ai-juan¹, LI Qian¹, TANG Wei²

(1.College of Electric and Information Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China; 2.College of Papermaking Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China)

Abstract: According to the BP network was slow convergence and large amounts of calculation, this paper proposed the line detection method of eliminate the influence of sample order and adaptive correction weights of BP network method. To improve speed and efficiency of the training at the same time, the recognition process was more stable and reliable, more accurate identification results. The experiments show that improved BP network can meet the requirements of online testing, and recognition rate can reach more than 94%.

Key words: paper defect; BP Network; pattern recognition; on-line detection system

纸是现代工业不可缺少的原材料,常见的脏斑、孔洞、褶皱和裂口等缺陷严重影响着纸的品质,所以要对纸病进行检测识别。纸张表面缺陷检测经历了人工目测、传统无损检测和机器视觉检测3个发展阶段^[1-2]。人工目测纸病的方法效率低、容易漏检、劳动强度大和实时性差。传统无损检测方法识别的缺陷种类少、检测实时性不强^[3-4]。本文采用机器视觉技术检测纸病表面缺陷,将改进的BP网络

算法^[5]应用到纸病的分类与识别,达到较好的识别和分类效果。

1 神经网络的纸病识别原理

在各种神经网络模型中,BP网络用于模式识别的效果较好。BP网络是前馈多层网络和误差反向传播算法相结合而成的一种神经网络。误差通过输出层按梯度下降的方式修正各层权值,向隐层、输入层

收稿日期:2012-08-14;修订日期:2012-10-11

基金项目:国家自然科学基金项目(30972322)

作者简介:魏爱娟(1986—),女,硕士研究生,研究方向为工业智能控制及工业高级过程控制;李茜(1957—),女,硕士,教授,学术带头人,研究方向为自动控制等;汤伟(1971—),男,教授,博士生导师,学术带头人,研究方向为工业智能控制、工业过程高级控制等。

逐层反传,此过程一直进行到网络输出的误差最小,经过多次迭代直至满足精度所要求的网络连接权和阈值后,便可对样本进行识别。神经网络模式识别功能的系统结构如图1所示。

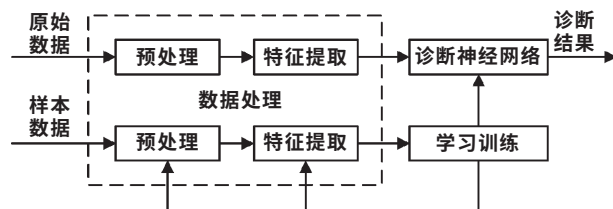


图1 神经网络模式识别功能的系统结构

Fig.1 Neural network pattern recognition system structure

2 改进的BP网络的原理及实现

2.1 BP网络模型的优缺点分析

神经网络能够充分逼近任意复杂的非线性关系,所有定量和定性的信息都均等分布存储在网络内的各神经元,有很强的鲁棒性和容错性。但是,BP网络算法的缺点也是不可忽视的。首先,在每输入一个学习样本后,网络就会根据误差对权值进行调整,因此网络受到后面输入样本的影响较大,会影响用户要求的训练精度。其次,学习速率是固定的,网络收敛速度慢,需要较长的训练时间。

2.2 BP网络的改进算法

1)为了消除样本顺序对结果的影响,可采用一批学习样本产生的总误差来调整权值,表达式如下:

$$\Delta w_{ij} = \sum \Delta_p w_{ij} \quad (1)$$

式中, Δw_{ij} 为连接权 w_{ij} 关于一批学习样本的调整量,即所有学习样本产生的总误差。由相应的计算公式求得每一个样本产生的权值调整量,从而有效地消除样本顺序对学习算法的影响。

将网络初始化,给 w_{ih} 、 w_{ho} 分别赋一个 $(-1, 1)$ 内的随机函数,设定误差函数:

$$e = \frac{1}{2} \sum_{o=1}^q (d_o(k) - y_o(k))^2, \text{ 给定计算精度值 } \varepsilon \text{ 和}$$

最大学习次数 M 。

输出层的权值调整值为

$$w_{ho}^{N+1} = w_{ho}^N + \sum \Delta_p w_{ho} \quad (2)$$

隐含层的权值调整值为

$$w_{ih}^{N+1} = w_{ih}^N + \sum \Delta_p w_{ih} \quad (3)$$

全局误差为

$$E = \frac{1}{2m} \sum_{k=1}^m \sum_{o=1}^q (d_o(k) - y_o(k))^2 \quad (4)$$

当 E 或学习次数大于设定的最大次数 M 时,则结束算法;否则,返回到输出层权值的调整,进行下一轮学习,以上改进的算法较好地解决了因样本输入顺序引起的精度和训练问题,但是收敛速度比较慢。

2)采用自适应调整参数是通过学习率 η 应根据误差的变化而自适应的调整,以使权系数调整向误差减小的方向变化,其迭代过程可表示为

$$w(k+1) = w(k) - \eta \nabla f(w(k)) \quad (5)$$

在训练过程中,不但消除样本顺序影响,而且也在学习速率进行了调整。采用消除样本输入顺序影响和自适应调整参数相结合算法的优点,能够有效地提高BP算法的训练速度和识别准确度。

3 改进的BP网络的结构设计

1)输入层和输出层变量的设计

三层的BP网络能够逼近任意的 n 维输入空间到 m 维输出的映射。输入量为表征纸病图像特征的特征向量,在特征提取过程中,提取了纸病图像的特征向量,网络输入层神经元个数为3个,输出层神经元个数为3个;隐含层的神经元个数为7个,隐含层的神经元个数是经过实际训练检验来不断调整的。选取常见的脏斑、孔洞、褶皱和裂口纸病的特征量作为网络的输入向量,需要网络区分优质(0,0,1)、合格状态(0,1,0)、纸病(1,0,0)作为网络的输出量,如图2所示。分别测试了节点为2、3、4、5,比较误差和收敛速度,确定隐层节点数为3。BP网络的输入层与隐层之间的传递函数,采用S型传递函数 tansig ,输入值取任意值,输出值为 $(-1, 1)$;隐层与输出层之间的传递函数选用 purelin ,输入值与输出值均可取任意值。

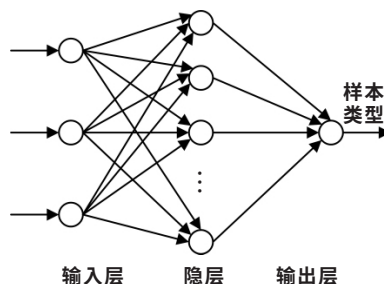


图2 BP网络结构

Fig.2 BP network structure

2)数据归一化

实验数据来自于湖南浏阳某造纸厂纸病在线检测过程所得数据。采样 240 组样本数据进行训练,6 组测试输入输出数据如表 1 所示。

表 1 输入样本

Tab.1 Input samples

样本 序号	输入样本特征数据								目标 向量
1	0.172	0.144	0.177	0.183	0.192	1.834	0.285	(0 0 1)	
2	0.170	0.157	0.158	0.200	0.188	1.786	0.280	(0 0 1)	
3	0.099	0.139	0.167	0.282	0.169	1.602	0.199	(0 1 0)	
4	0.024	0.039	0.076	0.166	0.165	1.285	0.266	(0 1 0)	
5	-0.234	-0.206	-0.071	-0.040	-0.049	-0.204	-0.108	(1 0 0)	
6	-0.250	-0.263	-0.082	-0.041	-0.032	-0.245	-0.122	(1 0 0)	

3)学习速率、初始权值及允许误差的设置

学习速率的选取是影响收敛速度以及训练结果的重要因素。每一个具体网络都存在一个适合的学习速率。训练次数为 500,训练目标为 0.01,学习速率为 0.1,取一组(-1 1)均匀分布的随机数作为初始权值,允许误差设为 0.002。

4 测试与结果分析

采集 4 种纸病样本共 200 组,130 组样本作为训练集,其余的作为测试集。利用 BP 网络工具箱训练样本数据,再用表 1 中的 6 组数据进行测试,阈值取为 0.8,仿真误差结果如图 3 和图 4 所示。在 MATLAB 的 M-File 中编写前面描述的仿真语句:

```
P=[0.172 0.144 0.177 0.183 0.192 1.834 0.285;0.170 0.157
0.158 0.200 0.188 1.786 0.280;
0.099 0.139 0.167 0.282 0.169 1.602 0.199;0.024 0.039 0.076
0.166 0.165 1.285 0.266;
-0.234 -0.206 -0.071 -0.040 -0.049 -0.204 -0.108;-0.250
-0.263 -0.082 -0.041 -0.032 -0.245 -0.122;];
P=P';
T=[0 0 1;0 0 1;0 1 0;0 1 0;1 0 0;1 0 0;];
T=T';
threshold=[0 0.8;0 0.8;0 0.8;0 0.8;0 0.8;0 0.8;0 0.8;];
net=newff(threshold,[7,3],{'tansig','logsig','trainlm'});
net.trainParam.epochs=500;
net.trainParam.goal=0.01;
LP.lr=0.1;
net=train(net,P,T)
```

由图 3 可知,网络训练误差在 56 次才达到训练目标,收敛性较差,而改进 BP 网络训练误差在训练仅为 18 次时就达到了设定的最小值且曲线较为

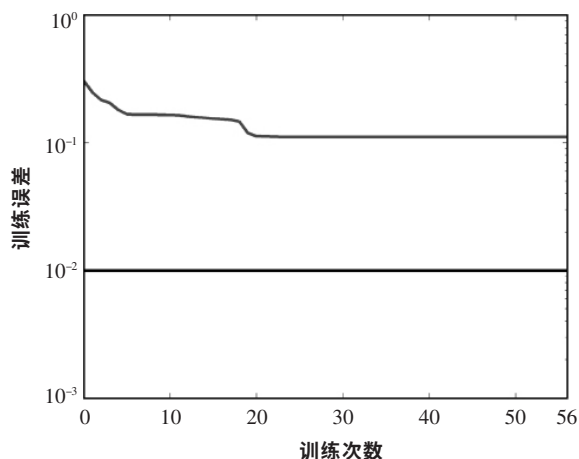


图 3 BP 网络样本训练

Fig.3 BP network sample training

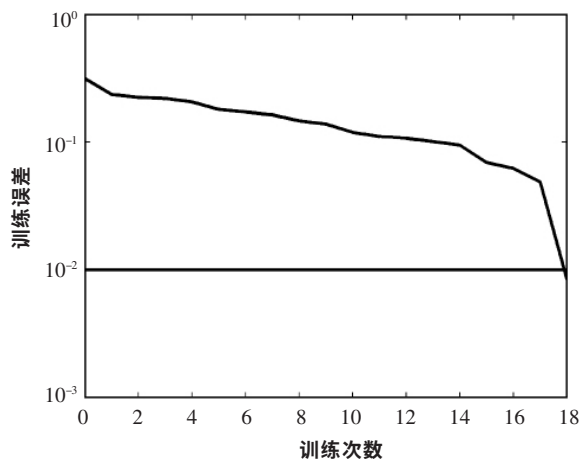


图 4 改进的 BP 网络样本训练

Fig.4 Improved BP network sample training

平滑,收敛性较好,如图 4 所示。对训练好的网络进行测试,抽取 3 组新的数据作为网络的测试数据,如表 2 所示。

表 2 测试数据

Tab.2 Test data

样本序号	测试样本特征数据								样本类型
1	0.186	0.146	0.167	0.203	0.192	1.881	0.267		优质
2	0.170	0.158	0.156	0.234	0.188	1.686	0.201		合格
3	0.035	0.039	0.016	0.082	0.069	1.002	0.151		纸病

测试代码为:Y=sim(net,P_test)

测试结果:Y=0.0195 0.9881 0.9763 0.9420
0.0514 0.8602

据欧式范数理论得 3 组测试样本的误差分别为 0.06697,0.02228,0.01322,这些误差是非常小的。因

此判定,采用改进的 BP 网络训练后,网络完全可以满足质量检测要求,识别结果比普通 BP 网络更准确。

在工程项目中,前端利用高速相机镜头收集来自纸张表面的光信号并保存到 CCD 器件上,由 CCD 元器件将光信号转换为像素对应的数字信号,由采集卡传输给计算机进行处理。通过对图像数据进行 BP 网络算法的分析与处理,辨识出纸病的大小、种类和位置信息,最后将这些信息显示出来并保存到数据库中。在某造纸厂的项目中得到测试运行,如图 5 所示。

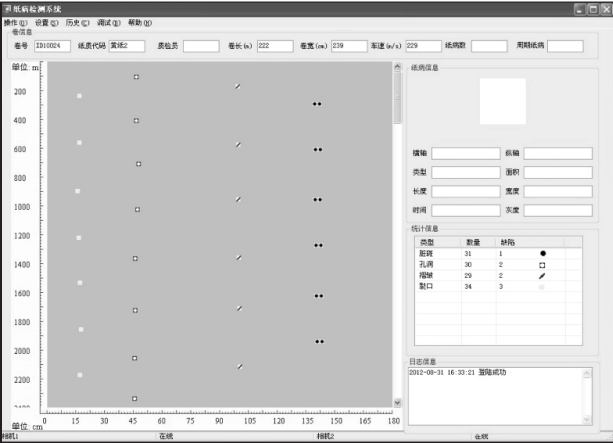


图 5 在线检测纸病监控画面

Fig.5 Monitoring screen of disease paper online detection

使用独立于训练集的纸病测试样本进行分类器验证,根据网络实际输出和目标样本对照,对测试样本进行检测,得到了纸病识别结果如表 3 所示。

从表 3 中可以看出,改进的 BP 网路对纸病缺陷图像进行分类识别的平均识别率为 94%,其正确识别率整体明显高于普通 BP 网络,而且耗时少,稳定性好。

表 3 BP 网络在线检测识别结果统计信息

Tab.3 BP network online detection and recognition of the results of statistical information

缺陷类型	裂口	脏斑	褶皱	孔洞	平均识别率
测试样本	37	32	31	32	
普通 BP 网路	29	28	26	27	83.5%
	78.4%	87.5%	83.9%	84.3%	
改进 BP 网络	34	31	29	30	94%
	91.9%	96.8%	93.5%	93.8%	

5 结语

本文利用 BP 网络解决了纸病识别的难题,通过两种方法的仿真对比可以看出,较 BP 网络而言,改进的 BP 网络检测效果较好,网络受到后面输入样本影响较小且在一定的范围内提高了学习效率,得到比普通 BP 网络更快更好的收敛性。同时,识别率具有更精确、更稳定的效果,在很大程度上增强了检测的效率和品质。将该算法应用到某造纸厂在线纸病检测过程且对识别率进行测试,正确识别率达 91% 以上,满足在线检测要求。

参考文献:

[1] 孙建镇,邱书波,刘芹.纸病检测方法的研究进展[J].中华纸业,2005,26(9):56-58.

[2] 倪洁,徐杰,胡幕伊.基于 BP 神经网络的纸病分类器设计[J].中国造纸学报,2010,25(2):76-78.

[3] Gallant Stephen I. Neural network learning and expert systems [M].London:The MIT Press,1993.

[4] 朱正为,郭玉英.基于 BP 网络的香烟包装质量在线检测与诊断[J].微计算机信息,2006,22(13):98-101.

[5] 张德丰.MATLAB 神经网络应用设计[M].北京:机械工业出版社,2009.

国务院五举措扶持光伏业

国务院总理温家宝 2012 年 12 月 19 日主持召开国务院常务会议,研究确定促进光伏产业健康发展的政策措施。分析人士表示,这说明光伏行业受支持的大方向已经明确,国内市场有望快速启动。

会议确定 5 大政策措施:加快产业结构调整和技术进步,善加利用市场“倒逼机制”,鼓励企业兼并重组,淘汰落后产能,提高技术和装备水平。严格控制新上单纯扩大产能的多晶硅、光伏电池及组件项目;规范产业发展秩序,积极开拓国内光伏应用市场;完善支持政策,根据资源条件制定光伏电站分区域上网标杆电价,对分布式光伏发电实行按照电量补贴的政策,根据成本变化合理调减上网电价和补贴标准。完善中央财政资金支持光伏发展的机制,光伏电站项目执行与风电相同的增值税优惠政策;充分发挥市场机制作用,减少政府干预,禁止地方保护。

转自证券日报