

文章编号:1671-6833(2020)04-0028-06

BP 网络预测阈值的仪表重影字符识别方法研究

孙国栋, 江亚杰, 徐 亮, 胡 也, 席志远

(湖北工业大学 机械工程学院,湖北 武汉 430068)

**摘 要:** 针对仪表数字获取过程中多出现光照不均匀和字符重影现象,导致二值化困难、识别率低等问题,提出了一种新的二值化方法。在对图像二值化之前,由于图像质量不佳,首先需要对图像进行预处理。针对光照不均现象,使用了非线性函数彩色图像校正方法。针对重影现象,以图像的灰度级分布统计量作为输入,自适应二值化全局阈值作为标签训练 BP 神经网络预测模型,使用训练好的 BP 网络对图像全局阈值进行预测并二值化,达到分离重影的目的。同时,采用改进 LeNet-5 网络对分割后的单个字符进行识别。结果表明,提出的二值化方法效果优于经典方法,改进的 LeNet-5 能够满足分割后的仪表字符识别,其识别率能达到 98.94%,分割后单个字符识别时间只需要 0.001 4 s。

**关键词:** 光照不均;重影;字符识别;预测阈值;LeNet-5;BP 神经网络

**中图分类号:** TP391.41      **文献标志码:** A      **doi:** 10.13705/j.issn.1671-6833.2020.04.011

0 引言

电力计量仪器的精度对于测量至关重要,需定期检测其精确度。而传统的数显式仪表不全都具备通信接口,无法直接获取标准测量仪表和被测仪表的数值,只能通过双相机分别采集其在快速升压和降压过程中的图像,自动识别仪表读数,最后对比标准表和被测表的结果,以判定被测表的精确度<sup>[1]</sup>。然而,在仪表数字变化时,部分仪表会出现严重重影,光照不均的影响也较为突出,使得仪表字符无法正常二值化。

针对光照不均的影响,常用的图像增强算法包括灰度变换方法和直方图均衡化<sup>[3-4]</sup>。Kim 等<sup>[5]</sup>使用快速迭代直方图方法增强图像的对比度,同时使用移动模板对图像的各个部分进行部分重叠的字块直方图均衡化。近年来,Retinex 理论<sup>[6]</sup>研究逐渐成为图像增强、去雾的热点,并衍生出多种改进算法,如单尺度 Retinex (SSR)<sup>[7]</sup>、同态滤波的 Retinex<sup>[8]</sup>、双边滤波的 Retinex<sup>[9]</sup>。刘健等<sup>[10]</sup>提出了基于 Retinex 理论与 LIP 模型的低照度图像增强方法,在 HSV 颜色空间,结合引导滤波与高斯平滑估计图像 V 分量的照度分量,利用 LIP 模型替代传统对数运算。这些方法机制

简单并且具有较低复杂度,然而都容易丢失图像细节和增加过多的噪声。

常用的二值化方法主要分为两大类:全局阈值算法和局部阈值算法。全局阈值算法主要包括大律法 (Otsu)、最大熵法、迭代法等,主要适用于光照均匀、灰度直方图有明显双峰的图像。局部阈值算法主要包括 Sauvola 算法、Niblack 算法、Bernsen 算法等,主要适用于光照不均的图像。以上算法都只能分离仪表图像的背景,无法分离数字重影。

为克服图像光照不均而导致的二值化困难,笔者根据文献[11]的方法,分别对具有强光照、弱光照、无底光的图像进行增强。通过对 RGB 彩色图进行灰度化,以灰度图的灰度级分布统计量作为输入,利用误差反向传播神经网络 (back propagation neural network, BPNN) 建立仪表字符图像全局阈值预测模型,预测理想二值化全局阈值,以分离重影。

1 仪表字符识别流程

笔者结合提出的 BP 神经网络<sup>[12]</sup> 阈值预测算法和改进 LeNet-5 的字符识别模型,提出了一种新的仪表字符识别方法,具体流程如图 1 所示,

其中 BP 神经网络模型和改进 LeNet-5 模型的训练过程提前完成,未在流程图中画出,主要包括以下步骤:

- (1)摄像头采集图像,图像增强并灰度化。
- (2)计算灰度图像的灰度级分布统计量,作为已经训练好的 BP 神经网络阈值预测模型的输入,预测理想的全局阈值,对图像进行二值化。
- (3)去除二值图的小连通域,降低噪声对图像的影响。
- (4)进行倾斜矫正,并将二值图分割成单个字符。
- (5)对单个字符图像进行尺寸归一化,使其达到改进 LeNet-5 模型输入图像尺寸统一的要求。
- (6)将尺寸归一化的图像输入到训练好的改进 LeNet-5 模型进行识别,得到仪表字符的识别结果。

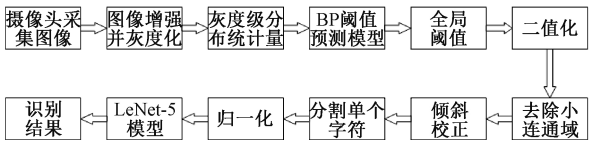


图 1 仪表字符识别流程图

Figure 1 Flow chart of instrument character recognition

其中,图像增强采用基于非线性函数变换的彩色图像校正方法,先将原始 RGB 图像转换到 HSV 颜色空间,对 V 分量采用多尺度加权高斯滤波的方式估计照射分量,然后根据估计光照分量的分布调整自适应增强函数的参数,得到两幅图像。利用图像融合从图像中提取重要信息来增强 V 分量。最后,将图像从 HSV 空间转换回 RGB 空间<sup>[11]</sup>。增强后的效果如图 2 所示。



图 2 光照不均实验对比

Figure 2 Uneven illumination experiment comparison

从图 2 可以看出,增强后的图像相比原图亮度得以提高,整体对比度有明显提升,极大地提高了图像质量。增强后彩色图像的灰度化使用 MATLAB 的函数模型 `rgb2gray` 完成。

2 BP 神经网络二值化阈值预测模型设计及优化

2.1 数据来源

为了验证本文算法的有效性,制作了一个含

有复杂重影的仪表数字库。通过大恒 MER-131-210U3C 相机与 50 mm 镜头采集 5 000 张多种旋转角度、不同污渍同光照下电压快速变化时 FLUKE289C 万用表的图像,不同的污渍状态是通过在数字显示处加灰尘和水渍得到,采集过程中固定图像为 1 141 像素×317 像素,并通过手动调整相机位置使得仪表数字全部出现在采集图像中,部分样本如图 3 所示。



图 3 部分仪表数字样本

Figure 3 Digital sample of some instruments

2.2 样本预处理和标签制作

采集的样本需进行灰度化处理,再计算出样本的灰度级分布统计量。经分析发现,仪表字符中重影的灰度明显比真实数字低。图 4(a)中显示的实际数字是 31.936,而数字“1”、“9”和“3”有明显拖影,实际数字灰度偏暗,拖影数字偏亮,两者灰度存在明显区别,图 4(a)对应的灰度直方图如图 4(c)所示,根据灰度直方图手动调整二值化全局阈值,将灰度直方图分为前景和背景,前景则是最理想二值化结果,如图 4(b)所示,二值化结果中无拖影,且数字完整,基本无噪声。

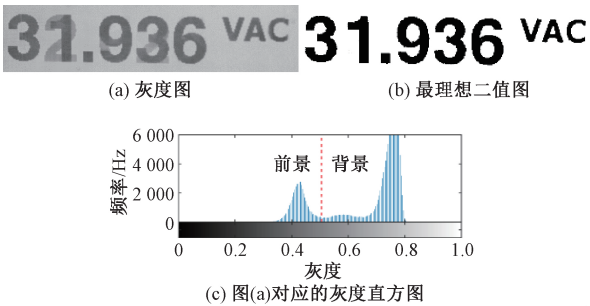


图 4 重影字符的理想二值化过程

Figure 4 The ideal binarization process of ghost characters

灰度级分布统计量是灰度直方图数据的来源,直接反映了图像的灰度信息,包含了图片的拖影和真实数字以及背景的灰度信息,且与二值化全局阈值密切相关。笔者以灰度级分布统计量作为 BP 神经网络预测模型的输入,理想全局阈值作为网络的输出,对模型进行训练,并将预测值与实际值进行比较,最终实现 BP 神经网络<sup>[12]</sup>自动预测仪表数字图像的最佳二值化全局阈值。

在网络训练前,由于各灰度级统计量相差过

大,导致网络的训练时间增加,甚至无法收敛。为保证网络的收敛速度和可靠性,必须对样本数据进行归一化处理。笔者采用最大值归一化处理:

$$x'_n = \frac{x_n}{x_{\max}}, \tag{1}$$

式中: $x_n$ 为样本的第  $n$  级灰度统计量; $x_{\max}$  为 256 维灰度级中最大灰度统计量; $x'_n$  为  $x_n$  归一化之后的输出值。归一化后 BP 阈值预测网络的输入量归一化到  $[0,1]$ 。

通过 BP 神经网络训练输出全局阈值的训练标签根据经验选取。从拍摄的图像库中选取 4 000 张样本,选取每个样本合理的全局阈值的最小值和最大值,并取平均值作为参考标准,手动调整直到图像清晰,此时对应的阈值即是理想二值化阈值,训练样本的数量为留下的 4 000 张样本的 60%,测试样本为剩余的 40%。

对训练样本进行预处理并制作阈值标签后,设计 BP 阈值预测网络的参数,使其能够通过样本的灰度级分布统计量预测出理想的二值化全局阈值。需要设计的参数包括隐含层的层数和节点数、节点转移函数和训练函数。

2.3 隐含层的层数和节点设计

在实际应用中,隐含层一般为 1~2 层,节点数大多根据以下经验公式<sup>[13]</sup> 选择:

$$l = \sqrt{n + m} + a, \tag{2}$$

式中: $l$  为隐含层节点数; $n$  为输入节点数; $m$  为输出节点数; $a$  为 1~10 的整数。

BP 阈值预测网络输入为 256 维的灰度级分布统计量,输出为单输出,则  $n$  和  $m$  分别设置为 255 和 1,即隐含层节点数的范围为 17~26。笔者选取不同的节点数和网络层数对网络进行训练,比较不同情况下的均方根误差  $E_{\text{RM}}$  和相关系数  $R^2$ <sup>[14]</sup> 来评价训练网络的好坏。均方根误差衡量实际值和预测值之间的偏差,其值越小,模型越好。相关系数衡量预测值与实际结果的精确度,其值越接近于 1,越精确。均方根误差和相关系数计算公式如下:

$$E_{\text{RM}} = \left[ \sum_{x=1}^N (a_x - g_x)^2 / N \right]^{\frac{1}{2}}; \tag{3}$$

$$R^2 = 1 - \frac{\sum_{x=1}^N (a_x - g_x)^2}{\sum_{x=1}^N g_x^2}, \tag{4}$$

式中: $N$  表示测试样本总数; $a_x$  为第  $x$  个样本的理

想阈值; $g_x$  为第  $x$  个样本的预测阈值。

图 5 给出了不同隐含层数和节点数下的均方根误差值和相关系数,其中第 2 层隐含层的节点数为 0 时,代表只有 1 个隐含层。由实验结果可以看出,当有两个隐含层且节点数都为 25 时,均方根误差最小,相关系数最接近于 1,此时隐含层的层数和节点数最优。因此,选取最优的隐含层结构为两隐含层,各层的节点数均为 25。

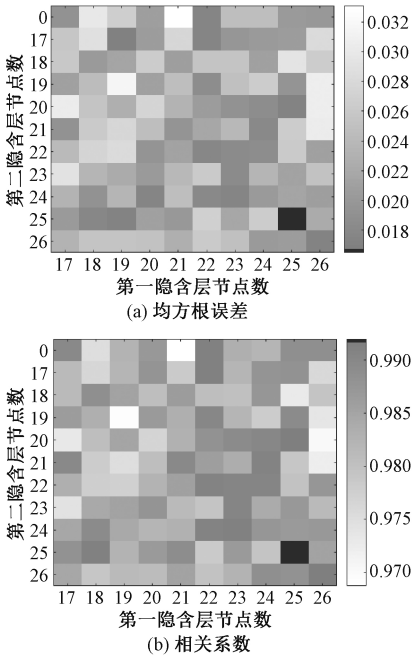


图 5 不同隐含层数和节点数下的均方根误差  
和相关系数

Figure 5 Root-mean-square error and correlation coefficient under different hidden layers and nodes

2.4 节点传递函数的选择

节点传递函数对于预测的精度有着重要的影响<sup>[15]</sup>,笔者在 Logsig、Purelin 和 Tansig 3 种节点传递函数中进行选择。不同节点传递函数的训练结果如表 1 所示,由表 1 可以看出,当隐含层和输出层函数都为 Tansig 时,相关系数最接近于 1,均方根误差也最小。因此,笔者选择 Tansig 作为隐含层和输出层的节点传递函数。

2.5 训练函数的选择

BP 训练函数根据网络的输入和目标期望输出,修正 BP 网络的权值和阈值,最终达到设定的网络性能指标,其直接影响 BP 模型的预测精度<sup>[16]</sup>。笔者分别对 Traingdx、Traingdm、Traingd、Trainrp、Trainlm、Traincgb 和 Trainscg 等 7 种训练函数进行了测试,对比实验结果见表 2。由表 2 可以看出,Traingdm 训练函数对应模型的相关系数最接近于 1,均方根误差也最小,优于其他 6 种

表 1 不同节点传递函数的训练结果

Table 1 Training results of transfer functions of different nodes			
隐含层函数	输出层函数	相关系数	均方根误差
Logsig	Logsig	0.990 97	0.017 616
Logsig	Purelin	0.985 34	0.022 557
Logsig	Tansig	0.985 08	0.022 537
Purelin	Logsig	0.976 16	0.028 609
Purelin	Purelin	0.960 05	0.036 811
Purelin	Tansig	0.933 81	0.045 791
Tansig	Logsig	0.985 94	0.021 659
Tansig	Purelin	0.982 26	0.024 199
Tansig	Tansig	0.992 04	0.016 835

表 2 不同训练函数的训练结果

Table 2 Training results of different training functions		
训练函数	相关系数	均方根误差
Traingdx	0.990 97	0.017 616
Traingdm	0.992 23	0.016 643
Traingd	0.985 62	0.022 197
Trainrp	0.977 33	0.028 295
Trainlm	0.990 35	0.017 941
Traincgb	0.978 61	0.026 761
Trainscg	0.971 77	0.030 962

训练函数。因此,笔者选择 Traingdm 作为 BP 训练函数。

3 阈值分割实验及字符识别

实验环境如下:Windows 10 系统,Intel Core i5 处理器,8 GB 内存,编程软件为 MATLAB2017a。

3.1 阈值分割实验

笔者选取大律法、最大熵法、迭代法、Sauvola 算法、Niblack 算法和 Bernsen 算法等 6 种经典的全局阈值和局部阈值算法作对比。实验所用的样本均是增强后的实验样本,结果如图 6 所示,各二值化方法的时间比较如表 3 所示。不同二值化算法的单张样本处理时间由相同条件下处理 1 600 张测试样本的总时间取平均求得。

由实验结果可以看出,笔者提出的 BP 预测阈值的二值化方法效果最好,基本能分离重影;而其他 6 种方法二值化效果较差,无法分离重影。而且增强后二值化能分割出完整的字符。说明了本文算法的有效性。单个样本的处理时间仅 0.036 8 s,仅比大律法和迭代法慢,说明本文算法的实时性较好,满足仪表字符的快速识别要求。而本文算法的网络结构简单,由于输入层和最终输出

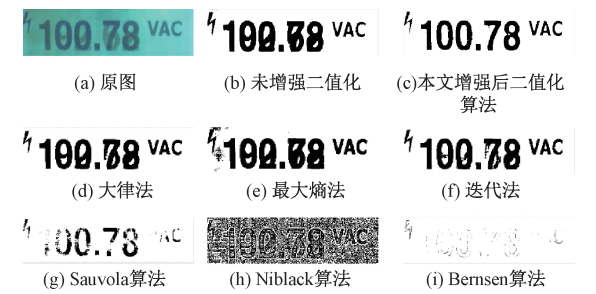


图 6 不同算法的二值化结果

Figure 6 Binary results of different algorithms

表 3 不同二值化算法的时间比较

Table 3 Time comparison of different binarization algorithms	
算法	单张图片处理时间/s
大律法	0.012 5
最大熵法	0.120 7
迭代法	0.031 3
Sauvola 算法	0.060 4
Niblack 算法	0.059 4
Bernsen 算法	0.133 8
本文算法	0.036 8

层结点数量是确定的,可以视为常量,中间的隐含层节点数均设定为 25,反向传播的时间复杂度和前馈计算相同,所以对一个样本的前馈计算时间复杂度应该是  $O(N)$ 。

3.2 倾斜矫正和字符分割

倾斜矫正算法<sup>[17]</sup>采用先创建最小外接矩形并计算出倾斜角度,后仿射变换,具体过程如图 7 所示。



图 7 校正算法流程图

Figure 7 Flow chart of calibration algorithm

字符分割算法采用投影法,包括垂直投影和水平投影两步,可以很好地对经过倾斜矫正的二值化仪表字符进行分割。分割过程如图 8 所示,由图 8 可以看出,数字、小数点和“VAC”字符水平投影的区间不同。

笔者主要考虑电压变化下仪表数字识别,而“VAC”字符不会随电压改变而改变,小数点可以直接通过水平投影区间判断,只需要识别重影数字,因此,在分割时只保留数字。

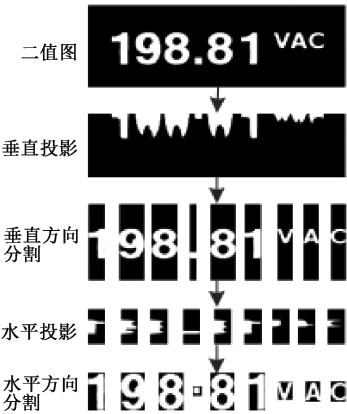


图 8 投影法分割流程图

Figure 8 Flow chart of projection method

3.3 重影字符识别实验

LeNet-5 是适用于手写字符识别的经典卷积神经网络,主要包括 1 个输入层、1 个输出层、2 个卷积层、2 个池化层和 3 个全连接层。笔者对 LeNet-5 进行了改进,并应用于仪表字符识别,主要改进如下:

(1) LeNet-5 使用的 sigmoid 激活函数会出现梯度消失问题且泛化能力差,而 ReLU 可以很好地解决此问题,并且收敛速度更快,故采用 ReLU 替换 sigmoid 函数。

(2) 使用 RMSprop 优化算法可以解决学习率选择困难和避免陷入局部最小值的问题,且收敛速度更快,因此采用 RMSprop 优化算法更新权重和偏置。

对于采集的重影字符库,传统大律法二值化后单个字符图像如图 9 所示,图像包含大量重影,难以辨别显示的数字。然而,本文算法分割得到单个字符如图 10 所示,图像基本不包含重影,可以达到识别要求。

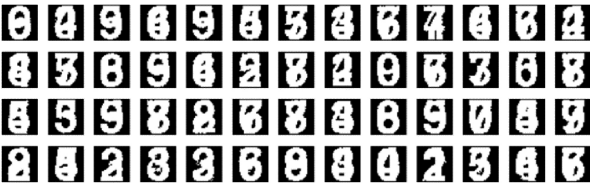


图 9 大律法二值化单个字符部分样本图

Figure 9 A part of sample image of the binarization of a single character in the Great Law

将训练样本和测试样本分割后的单个字符进行数字归类,制作数字标签,使用训练样本和测试样本分别对改进 LeNet-5 模型进行训练和测试。

改进 LeNet-5 和原始的 LeNet-5 参数设置如下:输入图像为 32 像素×32 像素,训练周期为 3 000, batch size 为 64,学习率为 0.000 1。SVM



图 10 本文算法分割后单字符部分样本图

Figure 10 Part of Sample image of single-character after segmentation algorithm in this paper

使用的 LIBSVM 工具包,采用线性核函数。HOG 特征提取使用的是 MATLAB 自带函数,其 cell size 设置为 3×3。各参数由多次实验取平均值得到。

实验结果如表 4 所示,由表 4 可以看出,改进 LeNet-5 和其他方法都能达到很高的识别率,表明本文二值化算法性能好,二值化后样本区分度高,但是改进 LeNet-5 不需要手动提取特征,在单个样本识别速度上比其他方法更快,虽然改进 LeNet-5 训练时间更长,但是训练过程可以预先完成,不影响实时识别。因此,改进 LeNet-5 更适合仪表数字的实时识别。

表 4 字符识别结果

Table 4 Character recognition results

方法	训练时间/s	单个样本测试时间/s	测试识别率/%
HOG+SVM	32.09	0.005 6	98.93
LeNet-5	94.76	0.002 7	98.65
改进 LeNet-5	93.87	0.001 4	98.94
KNN	45.26	0.004 7	96.77

根据本文提出的仪表重影字符识别方法,并通过 MATLAB 设计 GUI 交互界面,可以准确识别出仪表重影字符上显示的数字,如图 11 所示。



图 11 仪表重影识别交互界面

Figure 11 Instrument ghost recognition interactive interface

4 结论

笔者提出了一种基于 BP 神经网络的二值化方法,根据重影数字图像的灰度级分布统计量预



测理想二值化全局阈值,在不同光照情况下该二值化方法能够有效地消除仪表数字刷新时形成的重影。并采用改进 LeNet-5 对分割后的二值化单个字符进行识别。实验结果表明,提出的二值化方法优于现有的大津法、最大熵法、迭代法、Sauvola 算法、Niblack 算法和 Bernsen 算法,改进 LeNet-5 网络优于传统 HOG+SVM 算法和原始的 LeNet-5 网络,具有很高的实用价值。

参考文献:

[1] 孙国栋,梅术正,汤汉兵,等.基于密度特征与 KNN 算法的最优特征维数选择[J].现代电子技术,2018,41(16):80-83.

[2] 李智成,李文婷,梅术正,等.基于机器视觉的高压计量数显表自动校对系统[J].电测与仪表,2017,54(20):105-109,121.

[3] 李娜.利用灰度变换法增强数字图像[J].北京工业职业技术学院学报,2009,8(3):36-39.

[4] KANDHWAY P,BHANDARI A K,SINGH A.A novel reformed histogram equalization based medical image contrast enhancement using krill herd optimization[J].Biomedical signal processing and control, 2020, 56:101677.

[5] KIM J Y,KIM L S,HWANG S H.An advanced contrast enhancement using partially overlapped sub-block histogram equalization[J].IEEE transactions on circuits and systems for video technology, 2001, 11(4):475-484.

[6] 梁琳,何卫平,雷蕾,等.光照不均图像增强方法综述[J].计算机应用研究,2010,27(5):1625-1628.

[7] JIN H,TU L,DENG X.Night image enhancement algorithm based on retinex theory[J].International journal of advancements in computing technology, 2011, 3(10):291-298.

[8] 董静薇,徐博,马晓峰,等.基于同态滤波及多尺度 Retinex 的低照度图像增强算法[J].科学技术与工程,2018,18(22):238-242.

[9] 任崇巍.基于双边滤波和 Retinex 算法的货车图像预处理方法[J].现代城市轨道交通,2019(6):39-43.

[10] 刘健,郭潇,徐鑫龙,等.基于 Retinex 理论的低照度图像增强技术[J].火力与指挥控制,2019,44(9):139-143.

[11] WANG W C,CHEN Z X,YUAN X H,et al.Adaptive image enhancement method for correcting low-illumination images[J].Information sciences, 2019,496:25-41.

[12] 查刘根,谢春萍.应用四层 BP 神经网络的棉纱成纱质量预测[J].纺织学报,2019,40(1):52-56,61.

[13] 夏克文,李昌彪,沈钧毅.前向神经网络隐含层节点数的一种优化算法[J].计算机科学,2005,32(10):143-145.

[14] 冉成科,夏向阳,杨明圣,等.基于日类型及融合理论的 BP 网络光伏功率预测[J].中南大学学报(自然科学版),2018,49(9):2232-2239.

[15] 范伟,林瑜阳,李钟慎.基于 BP 神经网络的压电陶瓷蠕变预测[J].计量学报,2017,38(4):429-434.

[16] 李标,王磊,朱金营,等.BP 网络训练函数选取及其岩土工程应用分析[J].西部探矿工程,2008,20(10):58-60.

[17] 李少辉,周军,刘波,等.基于机器视觉和神经网络的低质量文本识别研究[J].机电工程,2018,35(9):1006-1010.

Study on Instrument Ghosting Character Recognition Method for Predicting Binarization Threshold by BP Network

SUN Guodong, JIANG Yajie, XU Liang, HU Ye, XI Zhiyuan

(School of Mechanical Engineering, Hubei University of Technology, Wuhan 430068, China)

**Abstract:** During instrument digital image acquisition, there were many phenomena of uneven illumination and character double shadow, which led to the difficulty of binarization and low recognition rate. A new binarization method is proposed. Before image binarization, the image would be preprocessed because of the poor image quality. Due to uneven illumination, the color image correction method based on nonlinear function was used. In view of the imaging ghosting, the image gray scale distribution statistics were taken as the input, and the adaptive binarization global threshold is used as the label of prediction model to train BP neural network. The trained BP network was used to predict the global threshold and binarize the image, in order to achieve the separating the ghosting. At the same time, the improved LeNet-5 network was adopted to recognize the single character after segmentation. The experimental results showed that the proposed binarization method was better than the classical methods, and the improved LeNet-5 could satisfy the instrument character recognition after segmentation, with the recognition rate of 98.94%, and the single character recognition time of only 0.0014s.

**Key words:** uneven illumination; ghost image; character recognition; prediction threshold; LeNet-5; BP neural network