

• 计算机工程与应用 •

# 基于投影阈值分割和数字序列校正的高噪声 数字仪表图像识别方法



郝琨<sup>1</sup>, 韩冰<sup>1</sup>, 李志圣<sup>1\*</sup>, 王传启<sup>2</sup>

(1. 天津城建大学计算机与信息工程学院 天津 西青区 300384; 2. 天津凯发电气股份有限公司 天津 西青区 300384)

**【摘要】**针对数字仪表图像噪声大、图像特征信息不足导致图像识别准确率低的问题,提出了一种基于卷积递归神经网络结合投影阈值分割和数字序列校正的高噪数字仪表图像识别方法。首先,用投影阈值分割二值化算法对图像进行预处理:使用垂直投影法将图像划分为不同区域,根据不同区域的噪声强度自适应设定二值化阈值,对图像进行二值化处理,降低噪声;其次,根据图像之间数字规律变化特点,利用数字序列校正算法将单个数字识别转换为数字序列识别,通过对比不同数字序列的识别概率得出识别结果,解决单张图像特征信息不足导致识别准确率低等问题。实验结果表明,在高噪声数据集上,相较于卷积递归神经网络模型,提出的高噪声数字仪表识别模型在准确率方面提高了约61.95%,达到93.58%。

**关键词** 卷积递归神经网络; 数字仪表识别; 图像去噪; 数字序列校正; 投影阈值分割  
中图分类号 TP391.4; TP183 文献标志码 A doi:10.12178/1001-0548.2022217

## High Noise Digital Instrument Image Recognition Method Based on Projection Threshold Segmentation and Number Sequence Correction

HAO Kun<sup>1</sup>, HAN Bing<sup>1</sup>, LI Zhisheng<sup>1\*</sup>, and WANG Chuanqi<sup>2</sup>

(1. School of Computer and Information Engineering, Tianjin Chengjian University Xiqing Tianjin 300384;

2. Tianjin Keyvia Electric Co., Ltd Xiqing Tianjin 300384)

**Abstract** To solve the problem of low recognition accuracy due to high noise and insufficient feature information of digital instrument images, this paper proposes an image recognition method of high noise digital instrument based on convolution recursive neural network combining projection threshold segmentation and number sequence correction. Firstly, the projection threshold segmentation binarization algorithm is proposed to preprocess the image. The vertical projection method is used to divide the image into different regions, and the binarization threshold is set adaptively according to the noise intensity of different regions to binarization the image and reduce the noise. Secondly, according to the changing characteristics of the number rules between images, the number sequence correction algorithm is used to transform a single number recognition into a number sequence recognition, and the recognition result is obtained by comparing the recognition probability of different number sequences, so as to solve the problem of low recognition accuracy caused by insufficient feature information of a single image. The experimental results show that, compared with the convolutional recursive neural network model, the accuracy of the high-noise digital instrument recognition model proposed in this paper is improved on the high-noise data set by about 61.95%, reaching 93.58%.

**Key words** convolutional recursive neural network; digital instrument recognition; image denoising; number sequence correction; projection threshold segmentation

数字仪表在生产中发挥着重要作用,仪表读数的精准度是决定生产质量的关键因素。在工业领域,数字仪表的数据读取一般通过内置装置采集,同时配合人工定期巡检。但由于数字仪表的工作环

境噪声复杂,采用人工读数易受巡检人员精力和经验影响,准确率较低,且成本较高。因此,大部分企业为了提高仪表识别准确率、节约生产成本,常采用计算机图像识别技术代替人工巡检。

收稿日期: 2022-07-04; 修回日期: 2022-09-12

基金项目: 国家自然科学基金(61902273); 天津自然科学基金面上项目(18JCYBJC85600)

作者简介: 郝琨(1979-),女,博士,教授,主要从事水下传感器网络、计算机视觉方面的研究。

\*通信作者: 李志圣, E-mail: lzs@tcu.edu.cn

## 1 相关工作

数字仪表的数据读取分为数字区域检测阶段和数字识别阶段, 检测阶段将图像中可能包含数字的区域使用锚框标记, 并将标记区域送入识别阶段进行识别。常见的数字区域检测模型有 EAST<sup>[1]</sup>、TextBox<sup>[2]</sup>、DBnet<sup>[3]</sup> 等。DBnet 模型使用可微二值化方法, 通过网络模型训练得出图像二值化阈值, 相较于其他检测模型具有较快的检测速度和较高的检测精度。因此本文使用 DBnet 模型对数字区域进行检测, 由于仪表数据集中数字区域和背景区域对比度较大, 因此检测精度可达 100%。数字识别阶段需提取图像中的数字特征并进行解码识别。传统的数字式仪表识别方法有特征检测<sup>[4]</sup>和模板匹配算法等<sup>[5]</sup>, 但传统方法模板制作复杂, 在识别字符种类较多情况下识别速度较慢且对于变形字符匹配效果较差, 适用范围较窄, 对于噪声干扰的鲁棒性较低。

随着深度学习技术的发展, 卷积递归神经网络 (convolutional recurrent neural network, CRNN)<sup>[6]</sup> 被广泛应用于数字式仪表的识别。文献 [7] 对数字仪表进行像素级的语义分割, 提高了数字仪表的识别准确率, 但此方法受脏污噪声影响较大, 在脏污遮挡环境下易出现误识现象; 文献 [8] 将可变形卷积应用到识别网络中, 在传统卷积和池化操作中添加二维偏置值, 使得神经网络可以适应不规则物体的识别, 该方法提高了在少量脏污遮挡情况下的数字识别准确率, 但在光线不均匀情况下该方法识别率较低; 文献 [9] 对现有的 CRNN 结构做出改进, 增加注意力机制, 在识别算法中使用正反两个解码器, 通过结合正序和逆序两种识别结果得出读数, 提高了字符粘连情况下的识别准确率, 但该方法会放大图像中污渍噪声的干扰; 文献 [10] 提出一种图像增强算法, 该算法通过计算图像不同区域内的动态截断值, 最终得到局部细节最优图像, 解决了图像中噪声分布不均匀等问题。但该方法需要人工选定参数且会带来局部失真问题, 在高曝光噪声图像上效果不佳; 文献 [11] 使用高斯滤波平滑仪表图像, 再通过自适应伽马增强算法去除复杂光线影响, 解决了图像高曝光问题, 但该方法无法去除密集噪点干扰; 文献 [12] 提出一种改进的最大熵阈值分割预处理算法, 能较好地去除密集噪点干扰, 配合卷积神经网络, 提高了高噪声环境下的仪表识别准确率, 但通过该方法计算的二值化阈值偏大, 在数字部分缺失情况下容易忽略特征信息; 文献 [13] 提出一种互补序列对的运动模糊图像复原方法, 利

用不同图像的特征信息, 使用互补帧图像修复模糊图像损失部分, 但该方法使用范围较窄且无法解决多幅图像出现特征信息不足的情况。

综上, 在噪声情况复杂、图像特征信息较少的情况下, 数字仪表识别技术存在识别准确率较低的问题。且由于仪表图像中缺少关键特征信息, 现有的去噪声算法和神经网络优化算法对于识别准确率提升有限。因此, 本文提出了结合投影阈值分割和数字序列校正的高噪声数字仪表图像识别方法 PN-CRNN。该方法使用投影阈值分割二值化算法 (projection threshold segmentation, PTS), 将图像按噪声强度划分不同区域, 并自适应设定阈值进行二值化处理, 有效降低噪声干扰; 接着提出数字序列校正算法 (number sequence correction algorithm, NSC), 通过数字序列概率代替单个数字概率, 利用前后帧图像之间的数字规律, 弥补当前图像缺失的特征信息, 从而提高数字图像识别准确率。实验结果表明, 本文方法在高噪声环境下具有较高的识别准确率。

## 2 高噪声数字仪表图像识别方法 PN-CRNN

### 2.1 PN-CRNN 网络模型

目前主流的文本识别网络为 CRNN。该模型包括卷积神经网络 (convolutional neural networks, CNN)、循环神经网络 (recurrent neural network, RNN) 和时间关联性序列分类模块 (connectionist temporal classification, CTC)<sup>[14]</sup>, 分别对应卷积层, 循环层和翻译层。CRNN 模型通过 CNN 网络和 RNN 网络充分提取图像特征信息并通过 CTC 算法解码翻译进行对比识别, 在无噪声干扰或低噪声干扰情况下识别准确率较高, 但是在高噪声干扰情况下, 由于数字特征信息较少, 导致识别准确率较低, 如图 1 所示, 数字“5”由于脏污遮挡影响, 数字特征信息不足, CRNN 网络将其误识为“6”。

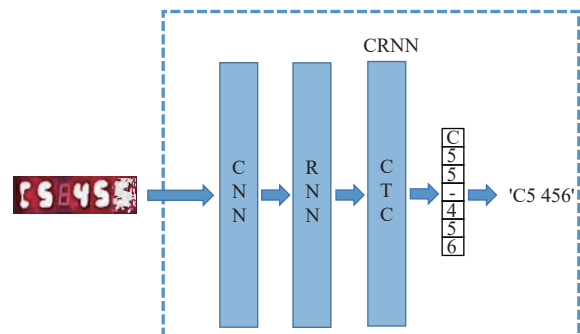


图1 高噪声情况下 CRNN 识别结果

因此，为提高识别准确率，解决图像中噪声复杂以及数字特征不足导致的误识等问题，本文基于 CRNN 网络提出 PN-CRNN 模型。模型结构如图 2 所示。PN-CRNN 首先对检测阶段得到的数字区域图像进行预处理，通过投影阈值分割算法 PTS 去除图像中的噪声影响；由 CRNN 网络提取图像中的特征信息，并生成预测概率矩阵；最后，

利用数字序列校正算法 NSC 得到精准的图像识别结果。相比于 CRNN，PN-CRNN 基于数字变化规律，利用不同仪表图像之间的相关信息，结合投影阈值分割和数字序列校正算法，有效地提高了图像特征信息利用率，解决了在高噪声环境下数字仪表识别率低的问题。

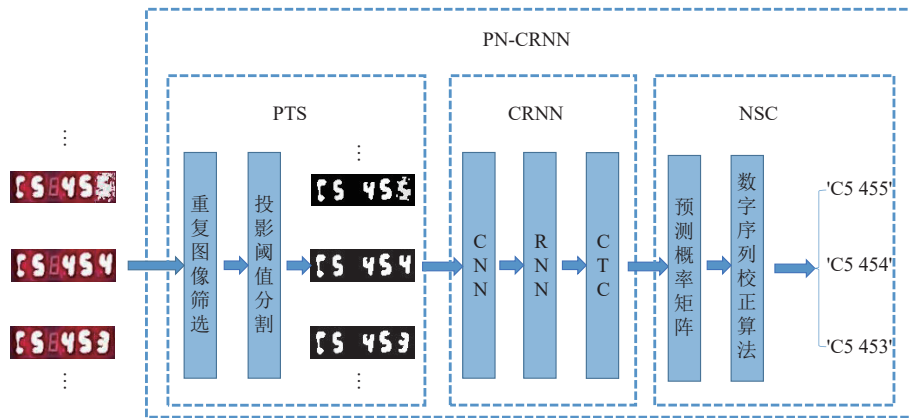


图 2 PN-CRNN 模型结构

## 2.2 PTS 图像预处理

### 2.2.1 重复图像筛选

在实际的生产过程中，当图像采集设备的捕捉频率高于仪表读数变换频率时，会出现图像重复采集情况，导致识别过程时间损耗增加，因此本文通过对比图像中关键区域的像素密度进行重复图像筛选。

由于数字部分在图像中占较大比例且与背景有较大对比度，因此图像中的关键区域可认为是数字区域。本文通过垂直投影法对数字区域进行划分。将图像中数字边界像素值设为阈值进行垂直投影，选取投影图中波谷坐标作为划分边界将图像划分为不同数字区域。在无噪声情况下，图 3 所示图片可划分出 5 个数字区域；当图像噪声较大时，数字出现粘连现象，导致投影图中波谷相互连接，此时将连接区域作为整体处理，则图片可划分出 3 个数字区域，如图 4 所示。

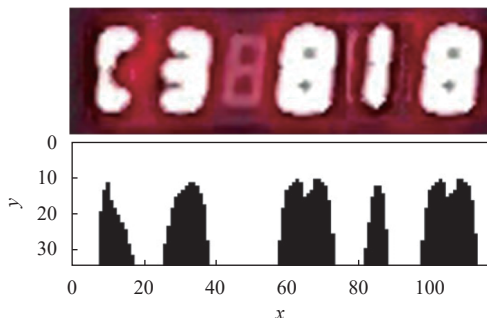


图 3 垂直投影图

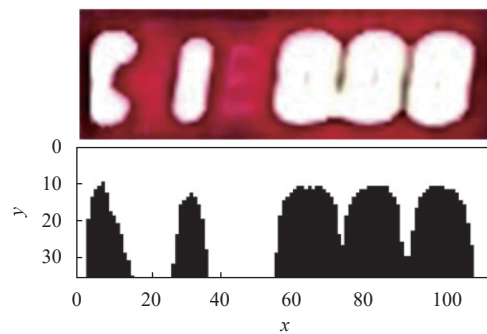


图 4 高噪声情况下垂直投影图

设数字边界像素值为  $C_0$ ，数字区域中灰度级  $i$  的像素个数为  $m_i$ ，灰度范围为  $[C_0, 255]$ ，则区域内总像素数为：

$$N = \sum_{i=C_0}^{255} m_i \tag{1}$$

数字区域像素点密度为：

$$p = \frac{N}{lh} \tag{2}$$

式中， $l$ 和 $h$ 分别表示数字区域的长度和高度。当两幅图像中数字区域个数相同且对应数字区域内像素点密度相似时，则判断为重复图像。

### 2.2.2 投影阈值分割算法

由于数字仪表工作环境复杂，噪声影响较大，常见的图像增强算法和图像二值化算法均不能完全满足实际生产需求<sup>[15]</sup>。因此本文结合全局阈值分割

和局部阈值分割优点, 根据图像数字区域内噪声强度大小自适应设定二值化阈值。通过实验验证, 本文提出的投影阈值分割算法效果优于常见的图像增强算法和二值化算法。

数字区域内二值化阈值表示为:

$$T = \frac{p}{\bar{p}} C_0 \quad (3)$$

式中,  $p$  和  $\bar{p}$  分别表示该数字区域像素点密度和无噪声情况下数字区域平均像素点密度。

图像二值化操作后需要通过形态学算法去除剩余噪声, 当图像受到水滴、雾气等噪声影响, 容易发生图像特征丢失现象, 如图 5a 所示。此时使用腐蚀等算法会造成部分特征丢失, 受噪声影响的数字区域经二值化处理后出现断裂现象, 数字被分成多个连通区域, 如图 5b 所示。因此需要对二值化处理后的图像进行修补。

本文提出的修补方法是以数字“1”的横向宽度作为修补标准, 对数字区域内像素点进行筛选。当以该像素点为中心的矩形区域内纵向像素点数量过少时, 则认为是噪声点或是边界点, 不做处理; 若竖向像素点形成连通区域且横向像素点数量小于标准值则进行修补。

修补过程将图像中断裂数字进行补全, 多个特征信息融合为所需特征信息。如图 5c 所示, 数字“7”经修补后仅存在一个连通区域, 且数字特征十分明显, 避免了特征信息不足导致的误识问题。

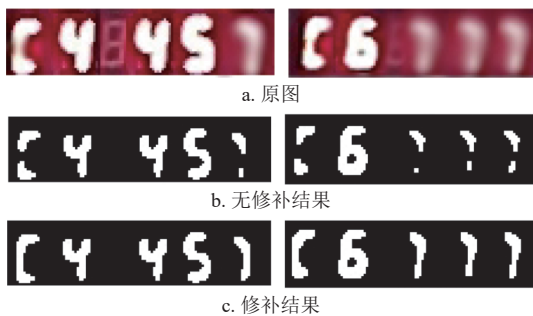


图 5 数字修补对比图

### 2.3 NSC 数字序列校正算法

#### 2.3.1 NSC 算法概念

实际生产中, 大部分仪表的读数是基于一定规律变化的, 如水表、电表。水表的读数表示用水量, 数字按规律逐渐变大, 如“1, 2, 3”或“1, 3, 5”。因此当规律变化的数字中有无法识别的情况, 可以借助数字变化规律进行推测, 如图 6 所示, 已知前两帧图像中最后一位数字分别为“3”

和“4”, 第三帧图像中最后一位数字受到噪声影响无法识别, 但依照前两帧图像数字变化规律, 该数字有较大的概率为“5”。



图 6 数字变化规律图

NSC 算法基于数字规律变化的前提, 将单个数字的识别变为对包含该数字的数字序列识别。原始的 CRNN 模型选择识别概率最大的字符作为识别结果, 当遇到高噪声导致的图像特征信息不足情况, 会产生误识的情况。NSC 算法选择识别概率最大的数字序列作为序列识别结果, 数字序列的识别概率由组成该数列的多个数字识别概率相乘所得, 因此识别准确率受数字序列整体影响, 降低了噪声对单个数字识别准确率的干扰权重。

#### 2.3.2 NSC 算法流程

假设数字跳动间隔为  $k$ , 数字序列长度为  $n$ , 上限为  $N$ 。STR<sub>1</sub>, STR<sub>2</sub>, ..., STR<sub>m</sub> 用于存放可行数字序列,  $m$  表示可行数字序列的数量。设  $P_1, P_2, \dots, P_m$  为可行数字序列概率。P<sub>out</sub> 表示输出概率, P<sub>idx</sub> 表示输出的数字序列。则 NSC 算法流程如图 7 所示。

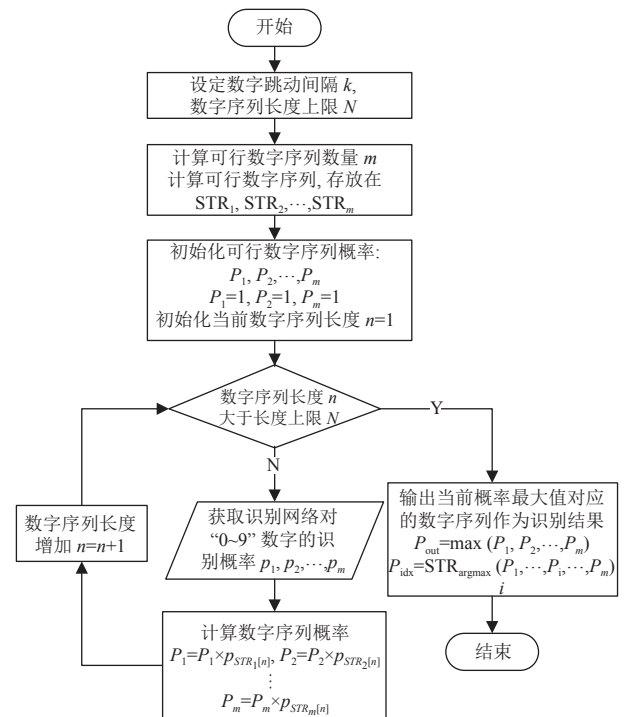


图 7 NSC 算法流程图

1) 设定数字跳动间隔  $k$  和数字序列长度上限  $N$ , 计算可行数字序列数量  $m$  和可行数字序列, 并将可行数字序列存放在 STR<sub>1</sub>, STR<sub>2</sub>, ..., STR<sub>m</sub> 中;

2) 对数字序列概率  $P_1, P_2, \dots, P_m$  和序列长度  $n$  赋初值, 令  $P_1 = 1, P_2 = 1, 2, \dots, P_m = 1, n = 1$ ;

3) 判断数字序列长度  $n$  是否大于上限  $N$ , 如果不是, 转入步骤 4), 如果是, 转入步骤 5);

4) 获取识别网络对数字“0~9”的识别概率  $P_1, P_2, \dots, P_9$ , 计算数字序列概率, 序列长度  $n = n + 1$ , 转入步骤 3);

5) 输出最大识别概率  $P_{\text{out}}$  和对应的数字序列  $P_{\text{idx}}$  作为识别结果, 流程结束。

### 2.3.3 NSC 算法有效性证明

为说明 NSC 算法的原理和有效性, 本文做如下数学证明。

将图片中某个位置数字的真实值记为  $i$ , 识别为正确数字  $i$  的概率记为  $p_{i,i}$ , 识别为错误数字  $i'$  的概率记为  $p_{i,i'}$ , 其中  $i, i' \in [0, 9], i \neq i'$ 。假设该位置数字的跳动间隔为 1, 将该位置上长度为  $n$  的数字序列记为  $S_{[i][i+1][i+2] \dots [i+j] \dots [i+(n-1)]}$ , 其中  $i+j$  为  $(i+j) \bmod 10$  的简记, 下文中的  $i+j, i+j+k$  等均为简记。数字序列的概率为组成序列中数字的概率乘积, 假设数字序列的初始数字为  $i$ , 序列长度为  $n$ , 识别为正确数字序列的概率记为:

$$P_{\text{correct}} = \prod_{j=0}^{n-1} p_{i+j,i+j} \quad 0 \leq i \leq 9 \quad (4)$$

识别为错误数字序列的概率记为:

$$P_{\text{wrong}} = \prod_{j=0}^{n-1} p_{i+j,i+j+k} \quad 0 \leq i \leq 9, 1 \leq k \leq 9 \quad (5)$$

经观察发现, 存在以下启发规则。

在高噪声条件下, 识别为正确数字的概率  $p_{i,i}$  会发生下降, 但从整体数据分析, 识别为正确数字平均概率仍大于识别为错误数字的平均概率。

假设  $i \in [0, 9]$ , 若序列  $S$  长度为  $n$ , 识别为正确数字的平均概率  $\bar{\mu}_{\text{correct}} = \sum_{j=0}^{n-1} p_{i+j,i+j} / n$ , 识别为错误数字的平均概率  $\bar{\mu}_{\text{wrong}} = \sum_{k=1}^9 \left( \sum_{j=0}^{n-1} p_{i+j,i+j+k} \right) / 9n$ , 则:

$$\bar{\mu}_{\text{correct}} > \bar{\mu}_{\text{wrong}} \quad (6)$$

数字概率分析实验说明了该启发规则的正确性。

已知数字序列中, 单个数字的识别概率分布相互独立<sup>[16]</sup>, 假设识别为正确数字的概率服从以  $\bar{\mu}_{\text{correct}}$  为均值的分布; 识别为错误数字的概率服从

以  $\bar{\mu}_{\text{wrong}}$  为均值的分布<sup>[17]</sup>。由启发规则和假设条件提出如下命题:

**命题 1** 在启发规则  $\bar{\mu}_{\text{correct}} > \bar{\mu}_{\text{wrong}}$  成立的条件下, 记数字序列  $S$  的长度为  $n$ , 则:  $\lim_{n \rightarrow \infty} \frac{E_S(P_{\text{correct}})}{E_S(P_{\text{wrong}})}$  依概率趋向于  $\infty$ 。

证明: 正确数字序列概率乘积的期望  $E_S(P_{\text{correct}})$  和错误数字序列概率乘积的期望  $E_S(P_{\text{wrong}})$  分别为:

$$E_S(P_{\text{correct}}) = E_S \left( \prod_{j=0}^{n-1} p_{i+j,i+j} \right) = \prod_{j=0}^{n-1} E(p_{i+j,i+j}) = (\bar{\mu}_{\text{correct}})^n \quad 0 \leq i \leq 9 \quad (7)$$

$$E_S(P_{\text{wrong}}) = E_S \left( \prod_{j=0}^{n-1} p_{i+j,i+j+k} \right) = \prod_{j=0}^{n-1} E(p_{i+j,i+j+k}) = (\bar{\mu}_{\text{wrong}})^n \quad 0 \leq i \leq 9, 1 \leq k \leq 9 \quad (8)$$

则可得:

$$\frac{E_S(P_{\text{correct}})}{E_S(P_{\text{wrong}})} = \left( \frac{\bar{\mu}_{\text{correct}}}{\bar{\mu}_{\text{wrong}}} \right)^n \quad (9)$$

由式 (6) 可知  $\frac{E_S(P_{\text{correct}})}{E_S(P_{\text{wrong}})} > 1$ , 即  $P_{\text{correct}}$  依概率<sup>[18]</sup>

大于  $P_{\text{wrong}}$ , 命题 1 成立。

由命题 1 可知在数字序列长度  $n$  达到一定程度时, 正确数字序列依概率大于错误数字序列, 算法的有效性得到保证。在实际测试中, 当数字序列长度达到 8 时, PN-CRNN 模型识别准确率趋于稳定。

### 2.3.4 NSC 算法的影响因素

#### 2.3.4.1 数字序列长度

由命题 1 可知 NSC 算法受数字序列长度影响, 当  $n \rightarrow \infty$ , 有  $\lim_{n \rightarrow \infty} \frac{E_S(P_{\text{correct}})}{E_S(P_{\text{wrong}})} = \infty$ 。当序列长度

接近无穷大时, 正确数字序列概率远大于错误数字序列概率, 即数字序列长度越长, 算法效果越好。

#### 2.3.4.2 可行数字序列数量

数字序列有多种可能, 将满足已知规律的数字序列定义为可行数字序列, 其余为干扰数字序列。

在固定数字序列长度  $n$  的前提下, NSC 算法受到可行数字序列数量的影响。可行数字序列数量决定算法的时间复杂度, 当可行数字序列数量过于庞大时, 算法难以满足实时识别。

可行数字序列数量受数字跳动间隔影响, 数字跳动间隔可认为是数字序列的规律。如数字跳动间

隔取值为 1 时, 数字序列 “1, 2, 3” 为可行数字序列, 数字序列 “1, 5, 8” 为干扰数字序列。当数字跳动间隔未知时, 数字序列没有规律约束, 所有数字序列都是可行数字序列。

### 2.3.5 NSC 算法参数的优化

#### 2.3.5.1 重复图像筛选

重复图像会极大增加可行数字序列的数量。以数字  $i \in [0, 9]$ , 序列长度  $n$  为例, 此时可行数字序列共有  $10^n$  种可能, 如图 8a 所示。通过 PTS 预处理算法剔除重复图像后, 可行数字序列数量降低为  $10 \times 9^{n-1}$  种可能, 如图 8b 所示。

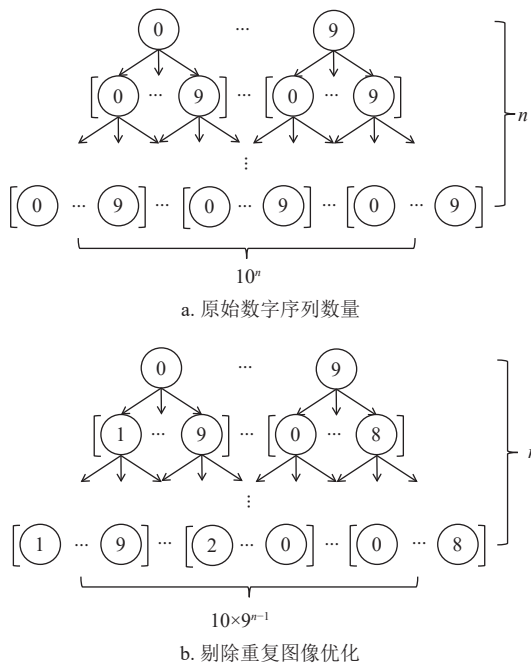


图 8 可行数字序列数量

#### 2.3.5.2 数字跳动间隔约束

在实际生产中, 由于仪表数字变化情况不固定, 因此数字跳动间隔取值情况存在多种可能。

假设数字  $i \in [0, 9]$ , 数字序列长度为  $n$ , 当数字跳动间隔有两种取值时, 以 1, 2 混合情况为例, 如图 9a 所示, 可行数字序列数量为  $10 \times 2^{n-1}$  种; 当数字跳动间隔有 3 种取值时, 以 1, 2, 3 混合情况为例, 如图 9b 所示, 可行数字序列数量为  $10 \times 3^{n-1}$  种。由此可得, 可行数字序列数量随数字跳动间隔的取值数量增大而增大, 因此减少数字跳动间隔的取值数量可以有效减少可行数字序列数量。

数字跳动间隔在影响可行数字序列数量的同时也影响了 NSC 算法的准确率。假设数字仪表规律变化, 数字跳动间隔为 1 所得数字序列如图 10a 所示, 数字跳动间隔为 2 所得数字序列如图 10b 所

示, 数字跳动间隔由 1 加大为 2 时, 数字序列的长度对应由 10 减少为 5。当数字跳动间隔过大时, 数字之间呈现无序跳动, 导致 NSC 算法失效。因此减小数字跳动间隔取值大小可以有效增加 NSC 算法的识别准确率。

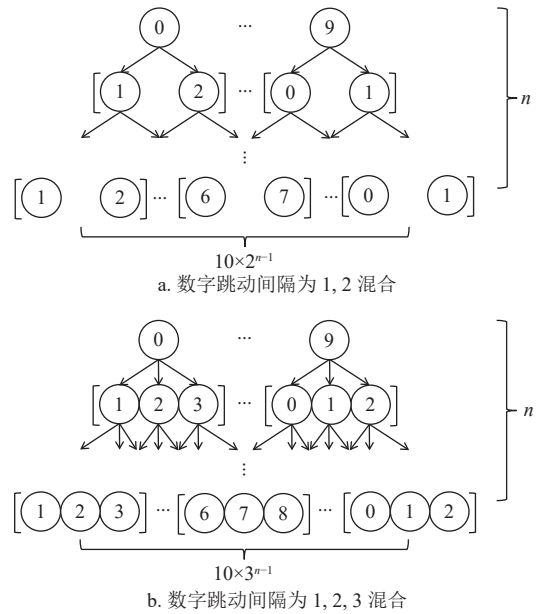


图 9 控制数字跳动间隔优化

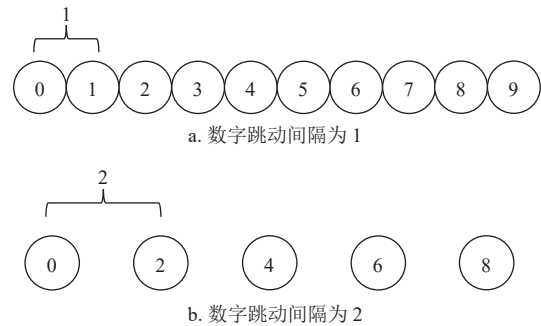


图 10 跳动间隔对数字序列影响

实际生产中, 数字跳动间隔可通过改变采集设备的图像采集频率约束, 当仪表图像变化规律时, 加快采集频率可以减小数字跳动间隔取值数量和取值大小。

## 3 实验与结果分析

### 3.1 数据集

本文采用的数据集 1 由电力企业通过监控摄像头对仪表设备进行实时采集制成, 共 3 510 张。为了验证 NSC 算法在高噪声情况下的效果, 本文在高噪声环境下对仪表设备进行采集, 得到噪声图像 500 张; 同时数据集 1 中随机抽取部分图像并通

过数据增强方法添加噪声, 得到噪声图像 404 张, 将其与采集所得噪声图像混合, 制作了高噪声情况下的数据集 2, 共 904 张, 如图 11 所示, 其中包括仪表镜头被脏污遮挡、光线过强导致的过曝、仪表电子元件损坏导致数字显示不完整等情况。数据集共有 4 414 张图片, 2 822 张用于网络训练, 1 592 张用于测试。

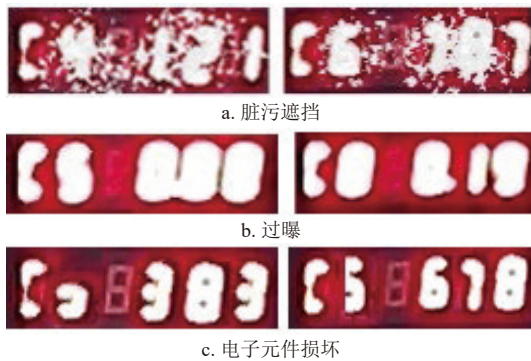


图 11 不同图像噪声

### 3.2 实验平台

本文实验使用计算机硬件设备为 Intel(R) Core(TM) i7-7800X CPU@3.50 GHz, 模型训练使用 2 个显存为 11 GB 的 GeForce GTX 1080Ti 上训练。计算机系统为 Ubuntu18.04LTS。

### 3.3 评价指标

本文采用准确率 (accuracy, Acc) 和每秒传输帧数 (frames per second, FPS) 作为评价指标。correct\_num 是正确的识别图像数量, all\_num 是图像总数, 准确率计算公式为:

$$\text{Acc} = \text{correct\_num} / \text{all\_num} \quad (10)$$

### 3.4 结果分析

#### 3.4.1 PTS 算法实验结果分析

预处理算法主要用于去除图像噪声, 在保留特征信息的同时减少噪声影响。图像的预处理方法分为非物理模型的增强方法和基于物理模型的复原方法。

常用的增强方法有带色彩恢复的多尺度视网膜增强算法 (multi-scale retinex with color restoration, MSRRCR)<sup>[19]</sup>, 自动彩色均衡算法 (automatic color equalization, ACE)<sup>[20]</sup>, 限制对比度自适应直方图均衡算法 (contrast limited adaptive histogram equalization, CLAHE)<sup>[21]</sup>。

除图像增强算法之外, 图像二值化处理也是常用的降噪方法<sup>[22]</sup>。常见的图像二值化算法有平均灰

度阈值分割和自适应的最大类间方差法 (maximum inter class variance, OTSU)<sup>[23]</sup>。

本文选择 ACE 算法、MSRRCR 算法、CLAHE 算法、OTSU 算法、全局阈值分割算法以及 PTS 算法进行对比实验。

图像增强算法可以增强图像中前景和背景的对比度, 改善过曝情况, 但同时会提高噪声影响。如图 12a~图 12d 所示, 图像增强算法无法去除图像密集噪点。

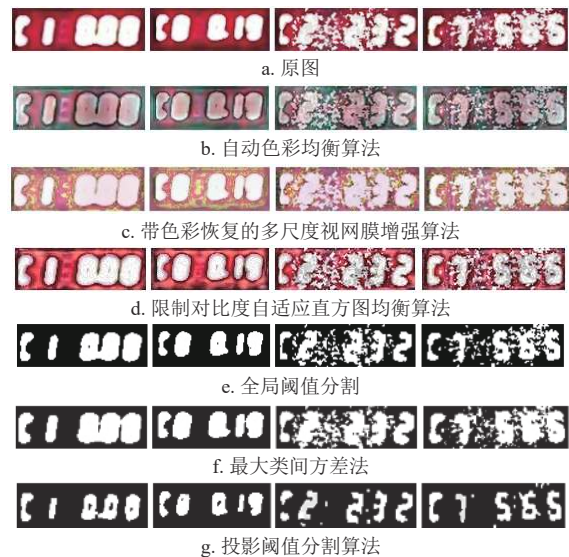


图 12 预处理算法对比

全局阈值二值化分割时间复杂度较低, 通过选取图像中字体区域的平均像素值作为阈值对图像进行简单的分割。阈值选大会导致图像丢失部分特征, 阈值过小则无法去除噪声干扰, 如图 12e 所示。自适应的阈值分割算法, 通过自适应选取阈值对图像进行分割, 当图片中前景和背景对比度较大时, Otsu 算法计算所得阈值偏小, 去噪效果较差, 如图 12f 所示。PTS 算法通过垂直方向投影将图像划分为不同区域, 针对区域内噪声强度自适应选取分割阈值, 并使用形态学方法减少噪声干扰。相比于全局阈值分割算法和 Otsu 算法有更好的去噪效果, 如图 12g 所示。

预处理算法对比实验使用 CRNN 模型, 主干网络采用 MobileNetV3(slim), 首先在低噪声数据集 1 上测试噪声干扰较小情况下预处理算法效果。实验结果如表 1 所示。从表 1 中可以看出, 图像增强算法对识别率提升效果有限; 全局阈值分割算法损失了图像部分特征, 导致识别准确率下降; Otsu

算法对噪声比较敏感, 相比于原图识别准确率下降约 14%; PTS 算法使用二值化处理消除部分噪声, 并通过投影分割方法解决阈值选取过高造成特征信息丢失的问题, 相比于原图识别准确率提高约 1.3%, 相比全局阈值算法识别准确率提高约 2.7%。

表 1 在数据集 1 上的预处理算法对比

算法名称	准确率(Acc)	FPS
无预处理	0.918 6	387
MSRCR	0.925 9	395
ACE	0.928 8	392
CLAHE	0.925 9	394
全局阈值分割	0.904 1	389
Otsu	0.776 2	387
PTS	0.931 7	396

通过在低噪声数据集 1 上的预处理对比实验可知, 图像增强算法提升效果有限, 且相比于图像二值化算法, 图像增强算法复杂度较高<sup>[24]</sup>。为进一步对比几种预处理算法的效果, 使用高噪声数据集 2 进行测试, 实验结果如表 2 所示, 使用 PTS 算法处理后, 图像识别准确率对比原图提高约 15%, 对比使用全局阈值分割算法高约 10%, 对比使用图像增强算法提高约 6%。实验结果表明, PTS 算法对比其他算法, 有更强的去噪效果。

表 2 在高噪声数据集 2 上的预处理算法对比

算法名称	准确率(Acc)	FPS
无预处理	0.316 3	394
MSRCR	0.398 4	353
ACE	0.400 7	364
CLAHE	0.389 5	362
全局阈值分割	0.358 4	375
PTS	0.462 1	363

### 3.4.2 数字序列概率分析

为验证 NSC 算法启发规则  $\bar{\mu}_{\text{correct}} > \bar{\mu}_{\text{wrong}}$  的正确性, 使用低噪声数据集 1 和高噪声数据集 2 进行实验。

高噪声数据集 2 分析结果见表 3, 正确数字平均识别率为 0.475, 错误数字平均识别率为 0.057, 错误数字平均识别率小于正确数字平均识别率。低噪声数据集 1 分析结果见表 4, 正确数字平均识别率为 0.962, 错误数字平均识别率为 0.014, 错误数字平均识别率远小于正确数字平均识别率。

表 3 在高噪声数据集 2 上的数字平均识别率乘积

类别	识别率乘积
正确识别	0.475
错误识别	0.057

表 4 在数据集 1 上的数字平均识别率乘积

类别	识别率乘积
正确识别	0.962
错误识别	0.014

从两组实验数据对比可知, NSC 算法的启发规则成立。

为证明假设中序列长度  $n$  增大,  $\frac{E_S(P_{\text{correct}})}{E_S(P_{\text{wrong}})} > 1$

依概率趋于无穷, 在高噪声数据集 2 上进行实验, 对数字识别率进行乘积对比。实验结果如表 5 和表 6 所示, 正确识别率乘积与错误识别率乘积比值如表 7 所示, 当序列长度为 2 时, 乘积比值为 1.005, 可知错误识别率乘积近似于正确识别率乘积, 即  $n$  当较小时, 有一定概率发生  $\frac{E_S(P_{\text{correct}})}{E_S(P_{\text{wrong}})} \leq 1$ , 但随  $n$  增大, 正确识别率乘积稳定大于错误识别率乘积。

表 5 在高噪声数据集 2 上的正确识别率乘积

序列长度 $n$	识别率乘积
2	$9.997 \times 10^{-3}$
3	$9.991 \times 10^{-3}$
7	$3.516 \times 10^{-5}$
50	$5.728 \times 10^{-30}$
100	$2.355 \times 10^{-60}$

表 6 在高噪声数据集 2 上的错误识别率乘积

序列长度 $n$	识别率乘积
2	$9.994 \times 10^{-3}$
3	$9.943 \times 10^{-5}$
7	$9.943 \times 10^{-13}$
50	$6.189 \times 10^{-93}$
100	$5.213 \times 10^{-186}$

表 7 在高噪声数据集 2 上的正确识别率与最大错误识别率乘积比值

序列长度 $n$	识别率乘积的比值
2	1.005
3	100.488
7	$3.536 \times 10^7$
50	$9.254 \times 10^{62}$
100	$4.571 \times 10^{125}$



### 3.4.3 NSC 算法效果测试

为验证 NSC 算法的效果, 使用高噪声数据集 2 测试。原始 CRNN 模型选用 PaddlePaddle 平台数字识别模型<sup>[25]</sup>, 并对其添加 PTS 算法和 NSC 算法。如表 8 所示, 仅使用 NSC 算法相比于原模型识别准确率提高约 37%, 同时使用 PTS 算法和 NSC 算法相比于原模型提高约 62%。

表 8 在高噪声数据集 2 上的 NSC 算法效果

算法名称	准确率(Acc)	FPS
原模型	0.316 3	371
原模型+NSC算法	0.689 1	389
原模型+PTS算法+NSC算法	0.935 8	390

### 3.4.4 数字序列长度对 NSC 算法影响

NSC 算法的序列长度由参与运算的数字个数决定, 序列长度为 1 的情况即原始模型。当序列长度增大时, 模型识别准确率逐渐增大, 如表 9 所示, 因此 NSC 算法效果随数字序列长度  $n$  增加而增强。随序列长度增大, 识别准确率上升趋势逐渐变缓, 如图 13 所示。在实际测试中, 当序列长度增大到某一阈值, 识别准确率达到最大值并不再增长。

表 9 序列长度对 NSC 算法影响

序列长度 $n$	准确率(Acc)	FPS
1	0.462 1	363
2	0.595 1	390
3	0.685 8	365
4	0.721 2	370
5	0.794 2	373
6	0.869 4	391
7	0.915 9	372
8	0.935 8	390

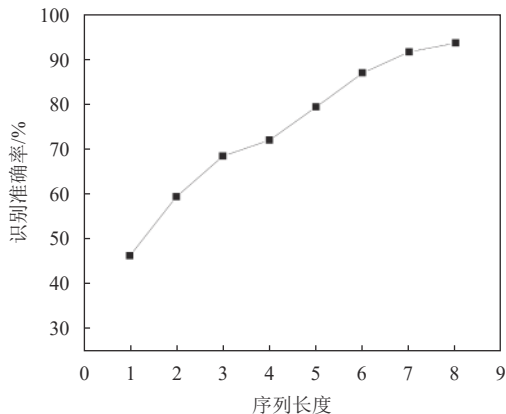


图 13 序列长度与识别准确率关系

### 3.4.5 数字跳动间隔对 NSC 算法的影响

NSC 算法受到数字序列数量影响, 数字序列数量由数字跳动间隔和起始位置决定。数字跳动间隔由 1 增加到 2 时, 如表 10 所示, 准确率下降约 5%; 当改变起始位置时, 数字跳动间隔变为 1 和 2 混合情况, 准确率下降约 9%。同时改变数字跳动间隔和起始位置, 准确率下降约 12%。由实验结果可知, 数字的跳动间隔是影响算法识别准确率的重要因素。在实际测试中, 算法使用的序列长度较短, 当数字跳动间隔过大时, 数字无法形成序列。针对此问题可通过调校设备采集时间间隔, 增加采样次数, 避免数字无序跳动, 保证算法效果。

表 10 数字间隔对 NSC 算法影响

数字间隔	准确率(Acc)	FPS
1	0.935 8	390
2	0.885 0	387
(1,2)	0.837 1	386
(2,3)	0.813 6	385

### 3.4.6 模型对比实验

在数据集 1 上, 对比 PN-CRNN 模型与 CRNN 模型在正常情况下识别准确率, 主干网络选择 MobileNetEnhance、MobileNetV3<sup>[26]</sup>, ResNet<sup>[27]</sup> 和 MobileNetV3(slim)。如表 11 所示, 添加 PTS 算法和 NSC 算法后模型准确率均有不同程度的提高, 其中 MobileNetV3(slim) 的识别准确率可达到 96.32%。

表 11 在数据集 1 上的模型对比

模型名称	主干网络	准确率	模型大小	FPS
CRNN	MobileNetV1Enhance	0.792 1	8.5 M	450
	MobileNetV3	0.921 5	6 M	386
	ResNet34	0.930 2	94.8 M	229
	MobileNetV3(slim)	0.918 6	2.6 M	387
PN-CRNN	MobileNetV1Enhance	0.944 1	8.5 M	376
	MobileNetV3	0.960 3	6 M	304
	ResNet34	0.947 1	94.8 M	165
	MobileNetV3(slim)	0.963 2	2.6 M	311

为进一步验证模型效果, 本文选择基于 SVTR<sup>[28]</sup> 和 PP-LCNet<sup>[29]</sup> 的 PP-OCRv3 模型进行对比实验。如表 12 所示, PN-CRNN 模型相比于 PP-OCRv3 模型在数据集 1 上准确率提高约 2%, 在高噪声数据集 2 上提高约 22%。

表 12 PN-CRNN 与 PP-OCRv3 模型对比

模型名称	主干网络	数据集名称	准确率	模型大小	FPS
PP-OCRv3	MobileNetV1Enhance	数据集1	0.946 2	2.6 M	246
		高噪声数据集2	0.716 5	167.5 M	231
PN-CRNN	MobileNetV3(slim)	数据集1	0.963 2	2.6 M	304
		高噪声数据集2	0.935 8	167.5 M	311

## 4 结束语

本文针对仪表图像噪声复杂问题, 提出了一种基于投影阈值分割和数字序列校正的高噪数字仪表图像识别方法。首先使用 PTS 投影阈值分割算法去除仪表图像噪声, 以便于模型提取数字特征。然后通过 NSC 数字序列校正算法将单个数字概率转变为数字序列概率, 避免了因数字特征信息不足导致的误识、漏识问题。实验结果表明, 在高噪声数据集上, 本文算法对数字的识别准确率达到 93.64%, 远远高于其他算法, 具有很好的实用性。然而, 本文算法存在对于数字序列条件要求较高、忽略了图像中数位之间存在的规律等不足, 未来将会针对这些不足继续开展后续研究工作。

## 参 考 文 献

- [1] ZHOU X Y, YAO C, WEN H, et al. East: An efficient and accurate scene text detector[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2642-2651.
- [2] LIAO M H, SHI B, BAI X, et al. Textboxes: A fast text detector with a single deep neural network[C]//Thirty-first AAAI Conference on Artificial Intelligence. California: AAAI, 2017: 4161-4167.
- [3] LIAO M H, WAN Z Y, YAO C, et al. Real-time scene text detection with differentiable binarization[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020, 34(7): 11474-11481.
- [4] 陈刚, 胡子峰, 郑超. 基于特征检测的数字仪表数码快速识别算法[J]. 中国测试, 2019, 45(4): 146-150.  
CHEN G, HU Z F, ZHENG C. Digital fast recognition algorithm for digital instrument based on feature detection[J]. China Measurement & Testing, 2019, 45(4): 146-150.
- [5] 高菊, 叶桦. 一种有效的水表数字图像二次识别算法[J]. 东南大学学报(自然科学版), 2013, 43(S1): 153-157.  
GAO J, YE H. An effective secondary recognition algorithm for water meter digital image[J]. Journal of Southeast University (Natural Science Edition), 2013, 43(S1): 153-157.
- [6] SHI B G, XIANG B, CONG Y. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39(11): 2298-2304.
- [7] 汤鹏, 刘毅, 魏宏光, 等. 基于 Mask-RCNN 海上升压站数字式仪表读数的自动识别算法[J]. 红外与激光工程, 2021, 50(S2): 163-170.  
TANG P, LIU Y, WEI H G, et al. Automatic recognition algorithm of digital instrument readings based on Mask-RCNN offshore booster station[J]. Infrared and Laser Engineering, 2021, 50(S2): 163-170.
- [8] 郭兰英, 韩睿之, 程鑫. 基于可变形卷积神经网络的数字仪表识别方法[J]. 计算机科学, 2020, 47(10): 187-193.  
GUO L Y, HAN R Z, CHENG X. Digital instrument recognition method based on deformable convolutional neural network[J]. Computer Science, 2020, 47(10): 187-193.
- [9] 曲超然, 陈立伟, 王建生, 等. 基于深度学习的工业数字仪表识别算法研究[J]. 应用科技, 2022, 49(2): 100-105.  
QU C R, CHEN L W, WANG J S, et al. Research on industrial digital instrument recognition algorithm based on deep learning[J]. Applied Science and Technology, 2022, 49(2): 100-105.
- [10] 梅英杰, 宁媛, 陈进军. 融合暗通道先验和 MSRRCR 的分块调节图像增强算法[J]. 光子学报, 2019, 48(7): 124-135.  
MEI Y J, NING Y, CHEN J J. A block-adjusted image enhancement algorithm based on dark channel prior and MSRRCR[J]. Acta Photonica Sinica, 2019, 48(7): 124-135.
- [11] WU X, GAO X T, GONG J. Intelligent instrument recognition scheme based on unattended substation inspection[C]//2020 39th Chinese Control Conference (CCC). Shenyang: IEEE, 2020: 6550-6555.
- [12] SHEN W D, TIAN G Y, LIU L, et al. Research on digital instrument segmentation based on improved maximum entropy algorithm[C]//2020 International Conference on Internet of Things and Intelligent Applications(ITIA). [S.l.]: IEEE, 2020: 1-5.
- [13] 叶晓杰, 崔光茫, 赵巨峰, 等. 基于闪动快门的互补序列对的运动模糊图像复原[J]. 光子学报, 2020, 49(8): 161-175.  
YE X J, CUI G M, ZHAO J F, et al. Motion blur image restoration based on complementary sequence pairs of flash shutter[J]. Acta Photonica Sinica, 2020, 49(8): 161-175.
- [14] GRAVES A, SANTIAGO F, GOMEZ F. Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks[C]//International Conference on Machine Learning. Pittsburgh: ACM, 2006: 369-376.
- [15] 张娜娜, 张媛媛, 丁维奇. 经典图像去噪方法研究综

- 述[J]. 化工自动化及仪表, 2021, 48(5): 409-412.  
ZHANG N N, ZHANG Y Y, DING W Q. An overview of denoising methods for classical images[J]. *Chemical Automation & Instrumentation*, 201, 48(5): 409-412.
- [16] LEE S, LEI M K, BRODY G H. Confidence intervals for distinguishing ordinal and disordinal interactions in multiple regression[J]. *Psychological Methods*, 2015, 20(2): 245-258.
- [17] CAGINALP C, CAGINALP G. The quotient of normal random variables and application to asset price fat tails[J]. *Physica A: Statistical Mechanics and Its Applications*, 2018, 499: 457-471.
- [18] SCHEICHL R, STUART A M, TECKENTRUP A L. Quasi-Monte carlo and multilevel monte carlo methods for computing posterior expectations in elliptic inverse problems[J]. *SIAM/ASA Journal on Uncertainty Quantification*, 2017, 5(1): 493-518.
- [19] FU X Y, ZENG D L, HUANG Y, et al. A fusion-based enhancing method for weakly illuminated images[J]. *Signal Processing*, 2016, 129: 82-96.
- [20] FU X Y, LIAO Y H, ZENG D L, et al. A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation[J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 4965-4977.
- [21] LI Y C, GUO J C, CONG R M, et al. Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior[J]. *IEEE Transactions on Image Processing*, 2016, 25(12): 5664-5677.
- [22] 刘硕. 阈值分割技术发展现状综述[J]. *科技创新与应用*, 2020(24): 129-130.  
LIU S. A review of threshold segmentation technology[J]. *Science and Technology Innovation and Application*, 2020(24): 129-130.
- [23] 黄梦涛, 连一鑫. 基于改进 Canny 算子的锂电池极片表面缺陷检测[J]. *仪器仪表学报*, 2021, 42(10): 199-209.  
HUANG M T, LIAN Y X. Based on the improved Canny operator of lithium battery pole piece surface defect detection[J]. *Journal of Instruments and Meters*, 2021, 42(10): 199-209.
- [24] 王一竹, 李渊, 杨宇. 基于人眼亮度感知的 S 型函数图像对比度增强算法[J]. *电子科技大学学报*, 2022, 51(4): 600-607.  
WANG Y Z, LI Y, YANG Y. Sigmoid function image contrast enhancement algorithm based on human eye brightness perception[J]. *Journal of University of Electronic Science and Technology of China*, 2022, 51(4): 600-607.
- [25] DU Y, LI C X, GUO R Y, et al. Pp-ocr: A practical ultra lightweight ocr system[EB/OL]. (2020-09-21). [2022-06-05]. <https://arxiv.org/abs/2009.09941>.
- [26] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul Korea: IEEE/CVF, 2019: 1314-1324.
- [27] VEIT A, WILBER M J, BELONGIE S. Residual networks behave like ensembles of relatively shallow networks[J]. *Advances in Neural Information Processing Systems*, 2016, 29: 550-558.
- [28] DU Y, CHEN Z, JIA C, et al. SVTR: Scene text recognition with a single visual model[EB/OL]. (2022-04-30). [2022-09-11]. <https://arxiv.org/abs/2205.00159>.
- [29] CUI C, GAO T, WEI S, et al. PP-LCNet: A Lightweight CPU Convolutional Neural Network[EB/OL]. (2021-09-17). [2022-09-11]. <https://arxiv.org/abs/2109.15099>.

编辑 税红