



华南理工大学  
South China University of Technology

# 专业学位硕士学位论文

基于深度学习的指针式仪表

全过程自动读数识别算法研究

作者姓名 安毅  
学位类别 电子信息硕士  
指导教师 杜启亮 副研究员  
董超 研究员  
所在学院 自动化科学与工程学院  
论文提交日期 2023年4月23日

# **Research on Automatic Reading Recognition Algorithm of pointer instrument based on deep learning**

A Dissertation Submitted for the Degree of Master

**Candidate: AnYi**

**Supervisor: Associate Researcher. Du Qiliang**

**Researcher. Dong Chao**

South China University of Technology

Guangzhou, China

分类号：TP391.4  
学 号：202021017344

学校代号：10561

### 华南理工大学硕士学位论文

# 基于深度学习的指针式仪表 全过程自动读数识别算法研究

作者姓名：安毅 指导教师姓名、职称：杜启亮 副研究员

申请学位级别：专业硕士 学科专业名称：电子信息

研究方向：计算机视觉

论文提交日期：2023 年 4 月 23 日 论文答辩日期：2023 年 6 月 7 日

学位授予单位：华南理工大学 学位授予日期：年 月 日

答辩委员会成员：

主席：董超

委员：田联房、莫鸿强、李彬、吴玉香

## 摘要

指针式仪表由于具有成本低、抗电磁干扰性能好等优点，在电力、石油、化工等领域得到了广泛应用。目前，获取指针式仪表读数主要依赖人工巡检，效率低且工作枯燥。因此，应用机器人替代人工巡检相关课题备受研究者关注，然而相关方法存在依赖先验信息、泛化能力不足等问题，成为限制其推广的主要制约条件。鉴于此，本文旨在研究一种仪表全过程自动读数识别算法，面向未知仪表也能得到可靠读数结果。

为实现上述目标，本文提出了一种仪表读数识别框架，首先从巡检图像中提取表盘图像并对其进行图像增强，然后基于表盘图像分别提取仪表的标度信息和指针信息，最后基于标度和指针信息使用仪表读数计算公式得到仪表读数识别结果。

在表盘图像提取环节，针对现有表盘检测方法计算量大，占用计算资源多等问题，本文基于 Fully Convolutional One-Stage Object Detection(FCOS)模型改进提出一种轻量级目标检测模型 Lite-FCOS，应用了本文提出的 Fast Global Context Network (FGCNet) 轻量级主干网络，并对 FCOS 模型的特征融合和检测头网络进行了轻量化改进。同时，针对检测得到的表盘图像可能存在的模糊和分辨率低等问题，本文提出了一种应用图像超分辨率重建技术提高表盘清晰度的方法，提升了算法泛化能力。

标度信息为表盘中所有标度数字的坐标及对应标度数值，针对当前同类算法普遍依赖仪表先验信息难以自主获取标度信息的问题，本文提出两种直接从表盘图像获取标度信息的方法。方法 1 基于光学字符识别 Optical Character Recognition(OCR)技术，使用 OCR 模型识别表盘全部字符，然后提出一种字符筛选方法从全部字符中筛选得到各个标度数字信息；方法 2 基于目标检测技术，由目标检测模型提取表盘中单个数字字符，然后提出跨类别非极大值抑制 Non-Maximum Suppression(NMS)，待融合字符组合和数字融合等算法，从单个数字字符中还原各个标度数字信息。

指针信息为指针中心点和尖端点坐标，针对当前同类算法对参数敏感导致鲁棒性差和抗干扰能力差等问题，本文提出两种方法以解决上述问题。首先两种方法都使用目标检测模型提取指针图像，然后方法 1 提出基于关键点检测模型从指针图像直接预测两点坐标，从而得到指针信息；方法 2 提出一种能够依据标度信息进行计算的指针中心点拟合算法，改进了 Hough 直线检测算法用于指针中心线提取，然后从指针中心线和指针图像边界框的两个交点中选取距指针中心点较远的作为指针尖端点并求其坐标。

针对传统角度法读数计算公式有效读数范围小，高度依赖仪表先验信息等问题，本

文提出了一种基于读数基准标度区间的仪表读数计算公式，大幅拓展了公式的有效范围，并且基于少量标度数字就可以实现可靠读数，同时提出了一种标度区间匹配方法以减少错误匹配标度区间带来的错误读数结果。

基于测试数据集对本文提出的仪表读数识别算法及各项改进实施了充分的消融实验和对比实验，验证了本文各项改进的有效性，实验结果表明本文算法相对同类算法具备更好的泛化能力。最后，基于已实现的算法设计和实现了一种智能巡检系统界面。

**关键字：**指针式仪表；仪表读数识别；目标检测；表盘信息提取；深度学习

## Abstract

Pointer-type meters, due to their advantages of low cost and good anti-electromagnetic interference performance, have been widely used in fields such as electric power, petroleum, and chemical industry. Nowadays, obtaining results from pointer-type meters mainly relies on manual inspection, which is inefficient and tedious. Therefore, the use of robots to replace human has been extensively studied. However, existing methods have problems such as dependence on prior information and lack of a strong generalization ability, which are the main limiting factors to commercialize it extensively. This paper aims to study an automatic reading recognition algorithm for the entire process of the meter, which can get reading results for unknown meters.

To achieve this goal, a framework for meter reading recognition is proposed in this paper. Firstly, the dial image is extracted from the inspection image and enhanced. Then, the scale and pointer information of the meter are extracted based on the dial image, and finally, we can calculate to get the result of the meter using the meter reading calculation formula.

To extract dial image from inspection image, we propose a lightweight object detection model called Lite-FCOS to address the problems of large computational overhead and resource consumption in other dial detection models. This model is based on FCOS model and uses the FGCNet lightweight backbone network we proposed in this paper. The feature fusion and detection head network of the FCOS model are also optimized for lightweight performance. Furthermore, to address the potential issues of blurriness and low resolution in the obtained dial images, this paper proposes a method to improve the clarity of blurry dials using image super-resolution reconstruction technology, thereby enhancing the algorithm's generalization ability.

The scale information refers to the coordinates and corresponding values of all scale number characters on the dial. To solve the problem of the current state-of-the-art algorithms relying on prior knowledge and facing difficulties in obtaining scale information, this paper proposes two methods for directly obtaining scale information from the dial image. Method 1 is based on OCR technology, which recognizes all characters on the dial using an OCR model and proposes a character filtering method to extract the information of each scale number from all characters. Method 2 is based on object detection technology, which extracts single-digit characters from the dial image using an object detection model and proposes cross-class NMS, character grouping, and number merging algorithms to restore the information of each scale number from the single-digit characters.

The pointer information includes the coordinates of the pointer's center point and tip point. To address the issues of poor robustness and anti-interference ability caused by the sensitivity of current comparable algorithms to parameters, this paper proposes two methods to solve these problems. Both methods use an object detection model to extract the pointer image at first. After that, method 1 proposes a keypoint detection model to predict the coordinates of the two points directly from the pointer image. Method 2 proposes a pointer center point fitting algorithm that can be calculated based on the scale information, which improves the Hough line detection algorithm for pointer center line extraction. Then, the coordinates of the pointer tip are obtained by selecting the intersection point farther from the pointer center point from the pointer center line and the two intersection points of the pointer image bounding box.

To optimize the limitations of traditional angle-based calculation methods for meter reading, which include a small effective reading range and a high dependence on prior knowledge of the meter, this thesis proposes a method involves a reading calculation formula based on a reference scale interval, which greatly expands the effective range of the formula and enables reliable readings with only a small number of scale digits. Additionally, we propose a scale interval matching method to reduce errors in reading caused by incorrect scale matching.

To validate the proposed meter reading recognition algorithm and its improvements, extensive ablation and comparative experiments are conducted on a test dataset. The results demonstrate the effectiveness of the proposed improvements and show that the algorithm has better generalization ability compared to similar algorithms. Finally, an intelligent inspection system interface is designed and implemented based on the implemented algorithm.

**Keywords:** pointer-type meter; meter reading recognition; object detection; dial information extraction; deep learning.

# 目 录

第一章 绪论.....	1
1.1 研究背景和意义.....	1
1.2 国内外研究现状.....	2
1.2.1 仪表读数计算研究现状.....	2
1.2.2 表盘图像提取研究现状.....	4
1.2.3 指针信息提取研究现状.....	5
1.2.4 标度信息提取研究现状.....	6
1.3 研究内容和结构安排.....	7
1.3.1 研究内容.....	7
1.3.2 论文章节安排.....	8
第二章 表盘图像提取 .....	10
2.1 FCOS 目标检测模型.....	10
2.2 Lite-FCOS 模型.....	12
2.2.1 FGCNet 轻量级主干网络 .....	12
2.2.2 FCOS 模型轻量化.....	15
2.3 Lite-FCOS 模型实验.....	16
2.3.1 FGCNet 主干网络实验 .....	16
2.3.2 Lite-FCOS 模型消融实验.....	17
2.3.3 Lite-FCOS 模型对比实验 .....	17
2.4 表盘图像增强.....	18
2.4.1 图像去噪.....	19
2.4.2 基于超分辨率重建的表盘图像清晰度增强 .....	21
2.4.3 大角度倾斜表盘检测.....	22
2.5 本章小结 .....	23
第三章 标度信息提取 .....	24
3.1 OCR 提取标度信息 .....	24
3.1.1 DBNet 文本检测模型 .....	25
3.1.2 CRNN 文本识别模型.....	27
3.1.3 OCR 模型的训练和优化 .....	28
3.1.4 OCR 识别结果筛选 .....	29
3.2 OCR 提取仪表类型 .....	30
3.3 目标检测提取标度信息.....	31
3.3.1 数字字符检测和跨类别 NMS .....	32
3.3.2 待融合字符组合.....	34

3.3.3 数字字符融合 .....	36
3.3.4 小数标度数字的处理机制 .....	38
3.4 本章小结 .....	40
第四章 指针信息提取 .....	41
4.1 指针区域提取 .....	41
4.2 基于 Hough 直线检测提取指针信息 .....	42
4.2.1 拟合指针中心点 .....	42
4.2.2 拟合指针中心线 .....	44
4.2.3 拟合指针尖端点 .....	48
4.3 关键点检测提取指针信息 .....	49
4.3.1 关键点检测技术简介 .....	50
4.3.2 HRNet 关键点检测网络 .....	51
4.3.3 关键点检测提取指针信息 .....	52
4.4 基于边界框的指针尖端点近似 .....	53
4.5 指针尖端点坐标标准化 .....	54
4.6 本章小结 .....	56
第五章 读数计算与总体框架 .....	57
5.1 仪表读数计算公式 .....	57
5.2 读数基准标度区间匹配 .....	60
5.3 仪表读数计算总体框架 .....	63
5.4 本章小结 .....	67
第六章 仪表读数实验及分析 .....	68
6.1 实验说明 .....	68
6.1.1 数据集 .....	68
6.1.2 评价指标 .....	69
6.1.3 硬件环境与模型训练配置 .....	69
6.2 读数识别性能验证 .....	70
6.3 消融实验 .....	71
6.4 读数对比实验 .....	75
6.5 抗扰能力分析 .....	77
6.6 智能巡检系统界面设计 .....	81
6.6.1 界面设计 .....	82
6.6.2 案例展示 .....	83
6.7 本章小结 .....	84
结论与展望 .....	85
参考文献 .....	87

# 第一章 绪论

## 1.1 研究背景和意义

仪表是显示数值的仪器的总称，是用于获取生产生活中各种数据的重要工具，对人类了解生产和生活中的环境状态都起着重要作用。虽然目前科技发展迅速，人类社会已经步入了数字化时代，一些基于互联网的数据采集系统已经可以直接获取实时生产数据传输到计算机中，可是在电力、石化等很多传统生产场景中，传统的指针式仪表仍然发挥着不可替代的重要作用。

指针式仪表具有以下几个特点：1) 结构简单，维护方便，价格便宜，具有较高的性价比；2) 抗电磁干扰能力强，能够在复杂电磁环境下正常稳定的工作；3) 发展历史长，已经部署的场景十分广泛，绝大多数现有的变电站，油气生产单位目前使用的系统都在应用指针式仪表，这些系统本身极为复杂且更新换代周期较长，因此不会在短时间内替换这些现有的仪表。由此可知，指针式仪表仍然将会在较长一段时间内继续在人类的生产生活中扮演重要角色。

目前获取指针式仪表数据主要依赖人工肉眼读数的手段。绝大部分生产场景中都还需要安排专门的巡检工作人员，每天前往各个仪表的部署地点进行读表和记录。这一过程耗时长，劳动强度大且高度重复，是一种低质量的体力劳动，难以使劳动者获得充分的价值感和发挥其创造性。同时，巡检人员的读数可靠性受到诸多因素的影响，长时间重复性劳动带来的疲劳、工作态度不端正带来的敷衍了事，专业能力不足带来的无法胜任，种种原因都可能导致的人工巡检读数出错，而这些错误很有可能引发更大的生产甚至安全问题。因此，不论是从劳动者的工作体验还是从安全生产对仪表读数可靠性的高要求来说，人工巡检都有着诸多的弊端。

随着人类科技的发展不断进步，研究者们一直都在关注如何使用机器人和计算机等技术替代人类完成上述这种低价值同时又对可靠性有较高要求的工作，仪表自动读数识别也一直受到研究者们的关注。早在上世纪 90 年代就有研究者提出了利用计算机视觉技术完成仪表图像的读数任务，但是当时的计算机视觉算法还比较初级，因此无法将表盘从复杂的背景中分割出来，因此研究的重点主要在于如何获取仪表中的指针的偏转角度。基于这种早期的仪表识别算法，一些研究者探索了利用固定摄像头来获取各个仪表的表盘图像，然后进行读数识别的机器视觉系统，然而这类方法的应用成本较高，因此没有得到广泛的使用。随着计算机视觉技术不断发展，研究者接连探索了基于 Hough

圆检测和基于图像特征的模板匹配等方法来对自然场景中的表盘图像进行定位和提取。2012年以来，深度学习技术赢来了一波大的爆发，推动计算机视觉领域取得了快速的发展，此后将深度学习模型应用在仪表识别算法中成为了研究的主流。这一阶段的早期主要是应用了基于卷积神经网络实现的目标检测和实例分割模型完成表盘检测和定位的任务，随着技术不断发展，也有越来越多的研究者开始探索应用其他计算机视觉任务的深度学习模型解决仪表读数问题以获得更好的泛化性能。

然而现有研究技术方案普遍对表盘先验信息高度依赖，不具备全过程读数识别的能力，因此面对未知的仪表无法进行读数识别。这就很大程度上制约了此类方法的应用和推广，因为每到一个新的场景都需要完成数据集采集，模型训练，参数设定等一系列工作，因此不利于方法的普及和真正下沉到实际工业场景中。

近年来国家大力提倡智能电网的发展，在此背景之下，进一步研究仪表的全过程自动读数识别算法具有很强的现实意义。通过进一步推动仪表读数识别算法的智能化和通用性使算法可以更加方便可靠的部署在不同场景中，从而让机器人巡检智能自动读表真正成为了可能。由此，本文选择单指针单标度的指针式仪表作为研究对象，以仪表的全过程自动读数识别为研究目标，探索提出了一种基于表盘图像提取、表盘信息识别和读数计算三大部分的仪表读数框架，并在此基础上进一步探索了目标检测、实例分割、OCR 和关键点检测模型在仪表读数识别算法中的应用，实现了基于多深度学习模型和基于目标检测模型的两种具体算法，并自制数据集完成了相关实验，并设计了智能巡检系统界面。

## 1.2 国内外研究现状

### 1.2.1 仪表读数计算研究现状

早在上世纪 90 年代，国外的研究团队就率先尝试利用计算机视觉技术对指针式仪表进行读数识别，Robert Sablatnig<sup>[1]</sup>团队在首先提出了一种针对水表进行自动读数识别的算法，这是计算机视觉技术第一次在仪表读数识别任务上的应用。国内研究者黎明等人<sup>[2]</sup>和郭子海等人<sup>[3]</sup>也分别在国内提出了应用 Hough 变换来求解仪表指针偏转角度然后基于仪表读数和指针偏转角度的对照表获取仪表的读数的算法，奠定了后续多年仪表读数识别算法的基本框架。由此可见，指针式仪表的自动读数识别其实有着较长的研究历史。在多年以来的不断发展中，仪表自动读数识别技术大致发展出了角度法和距离法两种不同技术路线。

角度法读数和指针式仪表的设计逻辑是一致的，仪表的指针偏转角度和仪表的读数成比例关系，因此如果能够求得指针相对某一个刻度的偏转角度，就可以通过这种比例关系计算仪表的最终读数。

早期的角度法一般利用指针相对于零标度线的偏转角度作为读数计算依据。国内的王涛等<sup>[4]</sup>提出了通过对指针中心线进行 Hough 直线检测得到指针中心线的偏转角度，然后基于已知的仪表量程等相关信息计算读数的方法。陶冰洁等<sup>[5]</sup>利用一种改进的自适应中值滤波方法改善图像效果，然后用双阈值 Hough 变换提取指针相对于零标度线的偏转角度，最后依据角度和仪表读数的线性关系计算读数结果。国外的 Ni 等人<sup>[6]</sup>直接绕过复杂的指针中心线提取环节，使用关键点检测网络预测零刻度、最大量程刻度、指针尖端点和指针中心点，然后基于同样的思路实现读数计算。

随着研究者获取表盘信息能力的增强，一些研究者探索了缩小标度区间的范围，不再使用从零标度到最大标度数字的量程区间，而是基于两个标度数字构成的标度区间作为读数计算的参照。徐发兵等<sup>[7]</sup>利用指针轮廓和相邻标度数字连线相交关系，将指针定位到两个标度数字所构成的标度区间内部，并在此基础之上计算读数。郭月阳等<sup>[8]</sup>综合角度法和距离法的优势设计了一种基于精定位之后指针所在标度区间进行计算读数识别方法。

距离法读数的底层逻辑其实和角度法是一致的，也是利用了指针偏转角度和仪表读数呈比例关系这一特点，只不过其并不直接计算指针的偏转角度，而是先对表盘进行极坐标变换处理，使得原本为圆弧形状的标度盘刻度部分在变换之后呈现条带形状，仿佛被“拉直”了一样，在此基础上再进一步寻找指针在这一条带状标度盘区域的位置，通过指针和某一个刻度的距离之比计算仪表的最终读数。F.Corrà Alegria 等人<sup>[9]</sup>最早在 2000 年提出了这一技术构想。Hung 等人<sup>[10]</sup>提出了一种基于刻度线提取特征点并进一步利用特征点拟合椭圆弧线，并在该椭圆弧线上应用距离法实现读数；在近年的相关文献中，Fan 等人<sup>[11]</sup>也通过一种基于 U-Net 模型<sup>[12]</sup>改进的语义分割和关键点检测模型分别获取指针掩模和刻度线关键点，然后利用刻度线关键点定位刻度区域并基于距离法实现了读数。刘葵等人<sup>[13]</sup>通过检测得到的指针端点确定表盘的半径，使用刻度线的延长线交点确定表盘圆心，利用表盘的圆心和半径截取一个扇形圆环区域。

由于距离法读数包含将标度盘刻度线部分转化为条带状的极坐标变换这一环节，因此对于仪表的拍摄条件要求更高，会放大表盘倾斜等不利拍摄条件导致的读数误差，同时刻度线本身的精确提取较为困难，对于本文的目标仪表全过程自动读数识别挑战较大。

因此本文选择的基本读数识别方法是角度法，下文的表盘图像提取和表盘信息提取部分也都是相对于角度法中的对应研究环节进行的文献综述。

### 1.2.2 表盘图像提取研究现状

早期的仪表自动读数识别算法并没有讨论过表盘图像提取的问题。受限于当时的计算机视觉技术能力，当时的仪表自动读数识别算法研究针对的是质量比较好的表盘图像，这些图像均是在理想条件下拍摄的高质量表盘图像，具体来说指表盘正对摄像头进行垂直方向的拍摄，且表盘部分占据了图像的绝大部分面积，基本没有干扰的背景部分。

为了在应用场景中也能够满足上述高质量表盘图像的条件，一些研究者提出了为仪表读数识别算法设计相应的机器视觉系统，即从工业相机选型安置、光源设计等方面构建专门的系统来获取高质量表盘图像。如 Wang 等人<sup>[4]</sup> 和 Zhang 等人<sup>[14]</sup> 的文献中就提及了类似机器视觉系统的搭建方案。由工业生产场景下的每个仪表都建立一个独立的机器视觉系统是一件成本极为高昂的事情，因此，早期基于机器视觉系统的仪表自动读数算法并没有得到普及。

随着机器人技术的进步和发展，使用巡检机器人代替机器视觉系统作为图像的获取平台来实现仪表自动读数识别算法的部署成为了主流的技术方案。巡检机器人获取的图像为巡检中的自然场景图像，以下简称“场景图像”，包含复杂的背景部分，对仪表读数识别干扰很大，因此如何从复杂背景中获取表盘部分成为了研究者们的重点关注方向。在深度学习热潮到来之前，主流的表盘检测技术主要包括 Hough 圆检测、模板匹配和基于特征点的模板匹配等几类。Hough 圆检测技术针对圆形的仪表提出，利用圆形仪表的表盘轮廓为圆形而自然场景中较少出现正圆形的物体这一特点，利用 Hough 变换进行圆形轮廓的检测，并将圆形轮廓内部的部分视为表盘部分提取出来。Gao 等人<sup>[15]</sup> 就通过首先使用 ORB(Oriented FAST and Rotated BRIEF) 特征进行仪表粗定位，然后使用云台姿态控制将表盘调整到图像中心位置，然后使用 Hough 圆检测的方法提取表盘部分，Zhang<sup>[16]</sup> 等研究者同样采用了类似的方法。模板匹配原指利用现有的模板和场景图像中的仪表进行匹配，通过滑窗法将模板在场景图象中进行遍历，选取和模板图像相似度最高的区域作为表盘部分提取出来。实践中更常用的是基于特征点匹配的模板匹配方法，通常利用 SIFT<sup>[17]</sup>( Scale Invariant Feature Transform) 或者 SURF<sup>[18]</sup>( Speeded-Up Robust Features) 等图像特征点提取算法提取模板图像和场景图象中的特征点，然后使用滑窗法遍历场景图像利用图像特征和模板进行匹配。Mai 等人<sup>[19]</sup> 提出利用 SIFT 模板匹配的方法，基于模板对场景图像中的仪表进行检测和提取，并且使用 RANSAC(Random Sample Consensus)

方法过滤错误的匹配。

然而上述方法的缺点都十分明显，泛化能力不足，高度依赖具体模板，不具备从检测一种仪表泛化到其他仪表的能力。

随着深度学习技术的发展，深度神经网络模型具有泛化能力上的强大优势，目标检测和实例分割模型解决的就是从包含复杂背景的自然场景中提取特定类型目标的问题，因此这两种模型也就成为了主流解决方案。经过仪表表盘数据集训练过的目标检测模型能够直接从场景图象中预测表盘目标，然后用边界框（bounding-box）描述目标的具体位置和尺寸，实例分割模型则再此基础上还能得到目标的二值掩膜。邢浩强等<sup>[20]</sup>率先在国内提出了一种结合卷积神经网络、仪表目标定位和 Hough 圆检测的表盘检测算法。Zhang 等人<sup>[21]</sup>提出了使用 Faster R-CNN 目标检测模型提取表盘图像并同时用于仪表分类工作。王俊凯等人<sup>[22]</sup>使用 FacesBoxes 模型提取指针图像，Zuo 等人<sup>[23]</sup>使用 Mask R-CNN 实例分割模型提取表盘和指针的掩模，然后基于掩模分割得到的表盘图像继续进行后续计算读数识别计算。

由于巡检拍摄的图像质量有限，存在因表盘图像质量不佳导致识别算法性能下降的问题，引起了研究者的关注和重视。孙绍祖等<sup>[24]</sup>关注了相机拍摄过程中的相对运动对仪表读数造成的不利影响，利用维纳滤波恢复了模糊的图像，改进了 PSF（Point Spread Function）参数的估计方法，通过提高参数估计效果实现更好的图像复原效果。胡东华等<sup>[25]</sup>关注了表盘倾斜校正问题，探索了一种基于轮廓的圆形仪表自动透视校正算法。胡鑫等<sup>[26]</sup>利用 CSS 角点检测设计了一种表盘校正算法。

### 1.2.3 指针信息提取研究现状

指针信息提取一直是仪表读数识别算法相关研究者关注的核心问题，因此发展出了较多方法和求解策略。指针信息都代表着指针的偏转角度，在多数基于角度法的相关文献中，指针的偏转角度都是通过指针中心线求取的，因而可以认为指针信息的提取关键在于指针中心线的提取。

早期的指针信息提取一般是直接在表盘图像中提取的，主要步骤一般包括：图像二值化和指针中心线检测两个主要环节，得到指针中心线表达式之后就可以计算出指针的偏转角度。图像二值化主要包括减影法<sup>[27]</sup>，首先需要获取一个指针处于调零状态下的表盘图像，作为模板和待检测图像做差分运算，从而得到指针部分的二值图像，然后再进一步检测其直线；

Hough 变换是最常用的指针直线检测算法，从早期到近年的文献都经常出现，Zhang

等人<sup>[28]</sup>使用利用仪表的刻度线延长线拟合指针中心点位置坐标用于将表盘图像调整到图像中心位置，然后使用 Hough 直线检测器检测指针中心线。其他常用的指针检测算法还包括 PCA(Principal Component Analysis) 算法<sup>[23]</sup>，最小二乘法<sup>[29, 30]</sup>和 RANSAC 随机采样一致算法<sup>[31]</sup>等。

随着基于深度学习的目标检测模型受到关注，一些研究者使用该技术以提高指针信息提取的效果。利用目标检测模型对指针区域进行提取，从而将图像二值化和指针中心线检测局限于指针区域内进行，有效降低了表盘内大量字符对这一过程的干扰，能够更好的提升检测效果。胡鑫等人<sup>[26]</sup>使用 yolov5 算法从表盘图像中获取指针图像然后进一步通过二值化和图像细化算法提取指针中心线。He 等人<sup>[32]</sup>使用实例分割模型从场景图像中直接提取指针掩模，然后在指针掩模的基础上使用 PCA 主成分分析方法拟合指针中心线，Liu 等人<sup>[33]</sup>使用实例分割 Mask R-CNN 模型同时检测表盘和指针掩模，然后对指针掩模应用图像细化算法以获取指针中心线。

#### 1.2.4 标度信息提取研究现状

和指针信息相反，标度信息提取一直是相关研究中的短板，没有受到研究者们重视，相关研究高度依赖先验信息，本文认为标度信息提取能力不足是导致仪表读数识别算法无法具备很强泛化能力的主要制约因素。

早期的仪表自动读数识别算法基本都只针对一种具体的仪表开展技术研究，因此文献内不涉及仪表标度信息获取的讨论，而是默认其为已知信息。这种做法本质上是预先构建了一种指针偏转角度到读数结构的映射表，从而将仪表读数问题简化为指针偏转角度的求解问题，仅需要求得指针偏转角度就可以直接通过查表的方式获得最终的仪表读数。张雪飞<sup>[34]</sup>和莫文雄<sup>[35]</sup>等都在其文献中提出了通过查表直接从得到的指针偏转角度得到读数结果的方法。类似的研究还有孙慧媛等<sup>[36]</sup>提出了一种基于二维码匹配的方法，通过扫描仪表下方的二维码获取仪表的完整先验信息。

模板是标度信息的一种呈现形式，随着仪表自动读数识别算法从面向单一固定仪表到面向场景中不同仪表，模板匹配最先被研究者用来提供标度信息。此类方法和 1.2.2 小节中基于模板匹配的表盘检测算法本质是上完全相同，这里不再赘述。

近年来，随着深度学习技术的不断发展，一些比较前沿的研究者开始针对标度信息的自动获取过程展开了研究。Gao 等研究者<sup>[37]</sup>使用轮廓分析提取标度数字区域，然后利用提出了一种结合了支持向量机二元分类器、字符滤波器以及 HOG(Histogram of Oriented Gradients)/多类支持向量机数字分类器的鲁棒字符分割分类器，用于识别仪表

板上的数字字符。靳一丹等人<sup>[38]</sup>使用轮廓检测的方法获取字符区域图像，然后基于一个小型 CNN(convolutional neural network) 实现标度数字识别。徐发兵等人<sup>[7]</sup>通过 EAST(Efficient and Accurate Scene Text detection) 文本检测模型进行标度数字检测，然后利用一个和靳一丹类似的 8 层的小型卷积神经网络进行印刷字体的识别得到标度信息。

一些研究者利用关键点检测网络预测指针的关键点，绕过了表盘信息提取步骤。宫倩<sup>[20]</sup>和 Meng 等<sup>[39]</sup>研究者都提出了利用关键点检测网络直接预测表盘的中心点、指针顶点和零标度点的坐标，然后结合仪表量程信息和角度法计算仪表读数方法，其中，宫倩等提出了一种以 ResNet18<sup>[40]</sup>为主干网络的轻量化仪表关键点检测网络，Meng 等提出了一种基于 Hourglass 模型<sup>[41]</sup>的改进实现了仪表关键点检测模型。

## 1.3 研究内容和结构安排

### 1.3.1 研究内容

指针式仪表巡检主要依靠巡检人员肉眼识别和记录，这是一项重复而且枯燥的工作，且容易因疲劳等原因导致读数或记录出错。自从国家提出智能电网概念以来，无人化和智能化一直是智能电网建设的重要方向，因此仪表自动读数识别也成为了研究热点。现有技术对仪表先验信息依赖度高，需要建立高质量模板或者数据集才能有效读数，然而巡检机器人应用的很多场景都属于电力部门等重要生产企业，厂区机密信息多，难以快速构建大规模训练用数据集，因此限制了此类技术的推广。为满足越来越迫切的仪表巡检无人化需求，同时解决现有相关算法泛化能力不足的问题，本文以基于深度学习的计算机视觉技术为主要方法，紧扣仪表自动读数识别这一具体应用场景，对仪表自动读数识别问题进行了细致而深度的研究。

结合上文的阐述，为进一步增强仪表自动读数识别算法的泛化能力，尤其是实现面向未知仪表的读数识别能力，本文提出两种技术路线作为后续研究方向：技术路线 1 充分利用深度学习算法泛化能力强的优势，将仪表读数识别问题拆分成多个子问题并针对性地采用不同领域的深度学习模型来实现功能；技术路线 2 专注于目标检测模型，通过分层次提取仪表相关信息并设计方法从目标信息还原表盘信息，同时注重网络的轻量化，追求以尽可能小的参数量和尽可能高的帧率实现功能。

针对现有算法存在的不足，综合上文的任务分析和实际生产的应用需求，提出本文的主要研究内容包括：

- 1) 表盘图像提取算法：主要解决在场景图像中对仪表表盘部分进行定位和分割的

问题，并且设计表盘提取之后进行的图像增强算法以得到高质量表盘图像。

2) 表盘标度信息提取算法：主要解决从表盘图像中获取标度信息的问题，标度信息由标度数字及其在表盘图像中的位置坐标描述。本阶段能否脱离对先验信息的依赖可以作为仪表自动识别算法能否泛化到未知仪表的主要判断依据。

3) 表盘指针信息提取算法：主要解决从表盘图像中获取指针信息的问题，尽可能减少算法对参数敏感性，指针信息包括指针旋转中心点和指针尖端点及其对应坐标。

4) 仪表读数计算方法：主要解决仪表读数计算公式设计的问题，并依据读数方法确定如何从表盘信息中获取读数计算所需的有效信息。

本文的技术路线总结如图 1-1 所示：

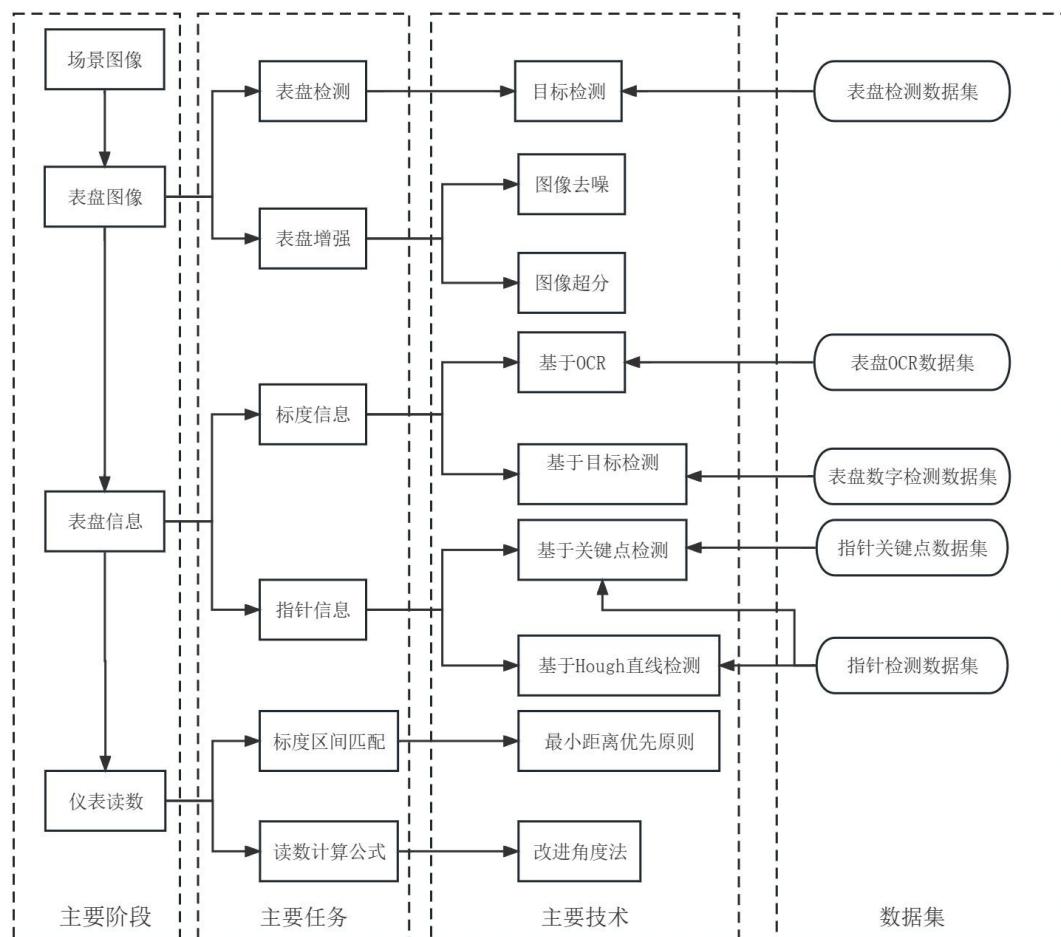


图 1-1 技术路线图

### 1.3.2 论文章节安排

本论文由七个章节构成，具体章节内容安排如下：

第一章：绪论。通过文献阅读和网上调研，对我国巡检机器人仪表自动读数识别算法进行了背景调研，从仪表读数计算方法、表盘图像提取、表盘指针信息提取和表盘标

度信息提取四个维度进行了文献综述，了解了该领域的发展状况和主流方法，深入分析了该领域当前面临的困难和挑战，最终确定本文的研究内容和章节安排。

第二章：表盘图像提取。研究了基于目标检测的表盘检测算法，基于对 FCOS 的改进提出 Lite-FCOS 轻量级目标检测模型作为本文目标检测相关任务的主要模型，并进行实验验证模型效果。考虑到直接提取的表盘图像可能质量不高，进一步研究了应用超分辨率重建模型对表盘图像进行清晰度增强的方法。

第三章：标度信息提取。研究了可面向未知仪表的标度信息提取方法，提出了一种基于 OCR 模型提取表盘全部字符然后进行筛选的标度信息提取算法和一种基于目标检测模型提取表盘单个数字字符然后进行数字融合的标度信息提取算法，并对基于 OCR 模型的方法进一步研究了筛选仪表类型相关字符和确定仪表类型信息的相关算法。

第四章：指针信息提取。研究了更加鲁棒和面向未知仪表的指针信息提取方法，首先使用目标检测模型提取表盘图像的指针区域，然后针对指针图像分别提出了一种改进 Hough 直线检测的指针信息提取算法和一种基于关键点检测模型的指针信息提取算法。同时提出了指针直线检测缺失的情况下利用指针图像边界框进行近似估计的备选策略。最后提出了指针尖端点标准化的问题和方法。

第五章：仪表读数计算。提出了本文选用的仪表读数计算公式和读数基准标度区间的匹配方法及其优化策略。综合第二、三、四章的内容提出了完整的仪表读数识别总体框架，并将全文提出的算法整合为了 MRM 和 MROD 两种具体仪表读数识别算法作为整篇论文的最终成果。

第六章：仪表读数实验及分析。基于自制的测试数据集对本文提出的两种仪表读数识别算法进行测试，从读数识别性能验证、消融实验、指针信息提取方法对比和抗干扰性能等方面设计并完成了实验，得到了充足的实验结果并进行了分析。最后，使用 PyQt5 框架设计并实现了一种智能仪表巡检系统界面，可以完成算法的人机交互，便于进一步推广和应用。

第七章：结论与展望。总结了本文的研究内容，分析当前提出方法仍然存在的不足之处，并基于现存不足和实际应用的需求提出对未来进一步研究工作的展望。

## 第二章 表盘图像提取

实际应用中的指针式仪表一般被部署在变电站或者机房等工业场景中，巡检机器人行进至预设拍摄位点后，调整摄像头面向仪表拍照获取场景图像。场景图像常包括复杂背景，对表盘信息提取不利，因此有必要将表盘区域分割后再进行后续读数识别步骤。现有表盘检测算法存在的计算量大和资源消耗多等问题，同时难以保证表盘图像质量。为解决上述问题，本文提出了一种基于 FCOS 模型改进的 Lite-FCOS 表盘检测模型，使用了本文提出的 FGCNet 主干网络并对特征融合和检测头网络进行了轻量化改造，另外提出了一种应用图像超分辨率重建模型进行表盘图像增强的方法，可以大幅改善低分辨率表盘图像效果。

### 2.1 FCOS 目标检测模型

FCOS 模型<sup>[42]</sup>是一种基于全卷积神经网络 FCN<sup>[43]</sup>思想的目标检测算法，具有不依赖 anchor 锚点框和 proposal 建议框的特点，整个算法仅依靠后处理过程的非极大值抑制 Non-Maximum Suppression（NMS）就实现了较高检测水平。

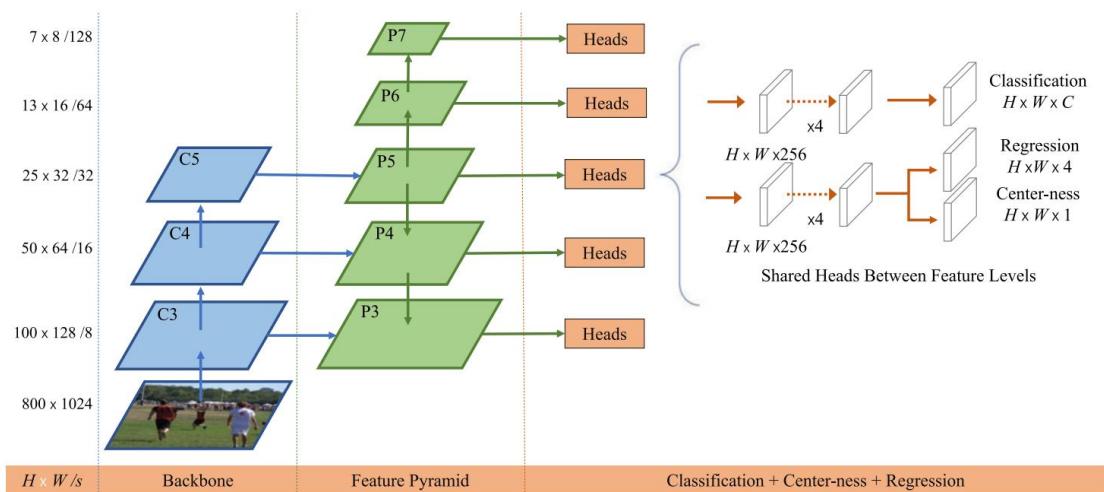


图 2-1 FCOS 模型结构图<sup>[42]</sup>

如图 2-1 所示，FCOS 模型主要包括主干网络，特征融合网络和检测头网络三大部分。主干网络使用 ResNet 网络<sup>[40]</sup>或其改进型 ResNeXt 网络<sup>[44]</sup>实现，用于对输入图像进行初步特征提取。ResNet 网络应用了一种残差网络结构，成功地解决了层数较多的神经网络在训练过程中常常出现的梯度消失和模型训练难以拟合的问题，能够让模型在训练过程中更好的收敛，因此能够增加模型的深度进而获得更好的特征提取效果，使得模型成为目前最常用主干网络之一。特征提取网络使用了特征金字塔网络即 FPN 网络<sup>[45]</sup>，用于对将不同层次的特征信息综合在一起，增强表征能力，提高特征的显著性、多样性，

同时也是实现多尺度特征提取的基础。检测头网络在设计过程中参考了 RetinaNet 模型<sup>[46]</sup>的检测头设计，用两个分支分别预测每个特征所对应的类别和边界框信息，同时还单独确定了一个 center-ness 的分支，用来预测当前特征相对于目标中心点位置的偏移程度。

FCOS 模型的最大创新点在于其结合了 FCN 的思想，对特征层逐像素点生成了边界框的预测结果，使得不依赖 anchor 机制和 proposal 机制也能对目标边界框进行预测，由此，所有相关计算和超参数也都被省略，从而节省了模型的计算量，同时避免了人工设计超参数不合理导致错检或漏检。边界框的具体生成方法为：首先将特征层中的点反向映射回其在原始图像中的坐标，然后根据其在原始图像中坐标和标注框的位置关系来确定其是否对应某一目标，若该原始图像中坐标在标注框的内部，则将其判定为正样本，然后以该原始图像中的坐标为参照点，基于上下左右四个方向边界框边界距离该参照点的距离  $t^*, b^*, l^*, r^*$  生成一个预测框，生成效果如图 2-2 所示。

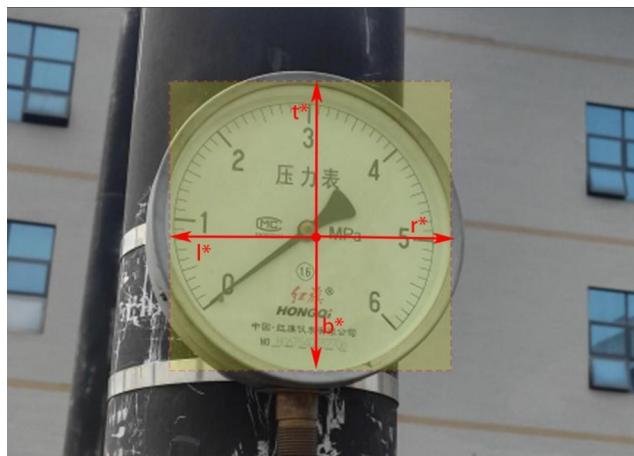


图 2-2 FCOS 模型边界框生成示意图

通过使用 FPN 网络，FCOS 模型实现了多级预测，提高了模型的召回率 recall，同时解决了边界框重叠带来的目标歧义问题。FPN 网络输入由主干网络输出的  $C_3, C_4, C_5$  三个公共特征层，对这些特征层进一步融合生成了特征层  $P_3, P_4, P_5$ ，并对  $P_5$  特征层继续用卷积层提取尺度更小的特征得到了  $P_6$  和  $P_7$ ，最终构成 5 层的特征金字塔。在边界框回归的环节，FCOS 模型限制了不同特征层的可预测范围，具体是通过限定了每个特征层的边界框回归范围  $(m_{i-1}, m_i)$  实现的，其中  $m_i$  表示特征金字塔的第  $i$  层特征。若一个特征生成的预测框最大方向距离值  $\max(t^*, b^*, l^*, r^*)$  在回归范围之外，则将这个特征点判定为一个负样本，从而实现在不同特征层预测不同尺度目标的多级预测。相比基于 anchor 框的模型，FCOS 能够更好应对目标遮挡等情况，提高检测性能。

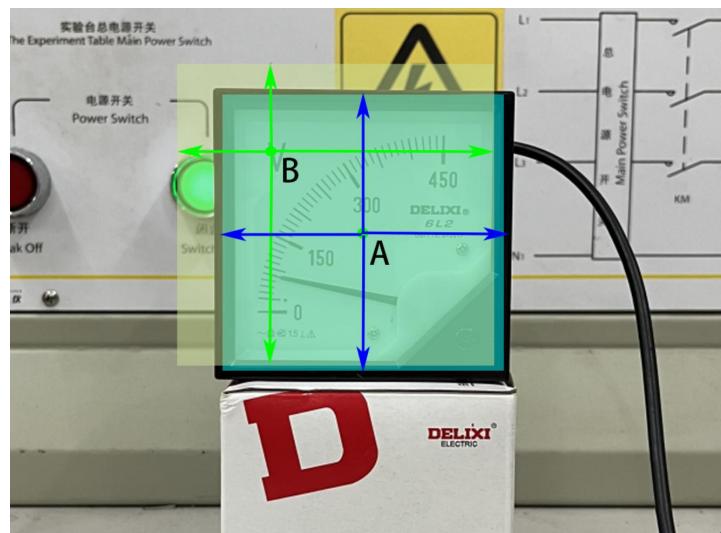


图 2-3 center-ness 效果示意图

通过在检测头新增 center-ness 分支, FCOS 模型进一步抑制了距目标较远的边界框, 降低了低质量边界框对检测效果的不利影响。center-ness 是作者定义的一种用于度量特征点和对应边界框位置关系的指标, 计算公式如(2-1)所示。当特征点距边界框中心越近, 其值越接近于 1, 否则其值越接近于 0。作者将 center-ness 作为权重引入边界框的置信度计算过程中, 降低特征点距离中心较远的边界框置信度, 进而在 NMS 阶段就能将其剔除。如图 2-3 所示, 图中 AB 两点图像的特征点生成了两个非常相近的边界框, 但是 A 点相比 B 更靠近边界框中心, center-ness 也就更高, 因而会降低 NMS 过程中 B 点对应边界框的置信度权重, 增加 A 点对应边界框的权重。

$$\text{centerness} = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)} \times \frac{\min(t^*, b^*)}{\max(t^*, b^*)}} \quad (2-1)$$

## 2.2 Lite-FCOS 模型

FCOS 模型是一种非常优秀的目标检测模型, 作为一种 anchor-free 的单阶段算法, 其性能能够达到同时代主流的 FasterR-CNN<sup>[47]</sup>等双阶段算法的水平。不过 FCOS 模型本身是面向 COCO 等大型通用目标检测数据集而设计的, 处理的目标种类多达 80 种, 且模型本身参数量和计算量都比较大, 因而不利于在巡检机器人这样的移动平台进行部署。因此, 本文在 FCOS 模型的基础上对其进行了优化改进, 在维持模型性能的前提下进行轻量化设计, 以降低模型参数量和加快检测速度为目标, 提出了 Lite-FCOS 模型。

### 2.2.1 FGCNet 轻量级主干网络

在文献中 FCOS 模型使用的主干网络是 ResNet 等相对复杂的模型, 其优点是特征提取能力强大, 但是缺点是部署计算资源要求较高, 同时推理过程耗时较长, 因而不利

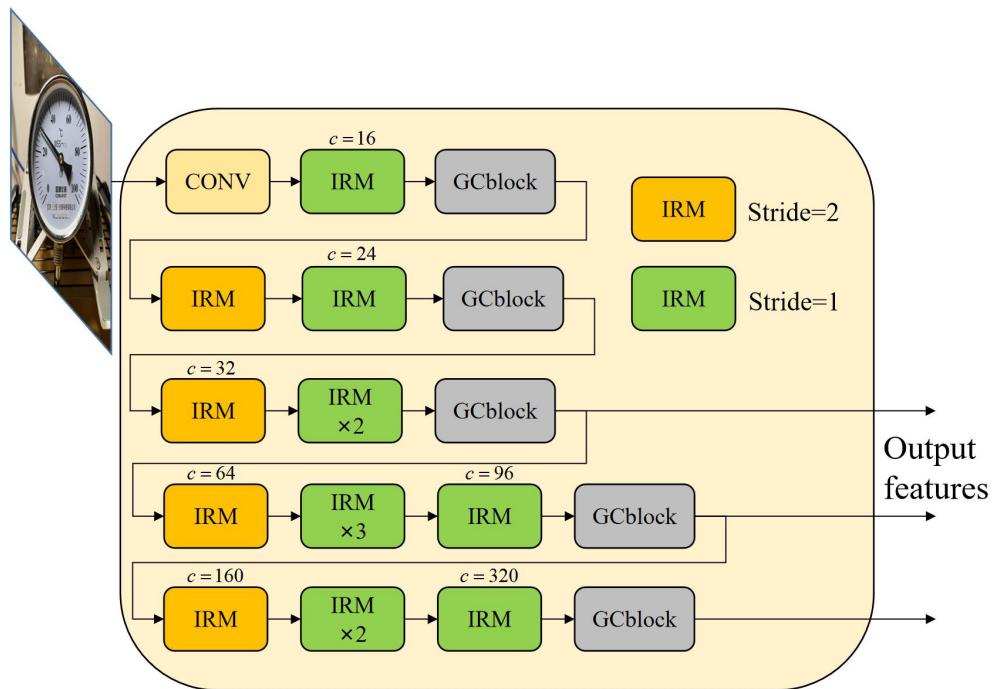


图 2-4 FGCNet 网络结构图

于实际部署。为使表盘检测模型足够轻量和快速，本文提出 FGCNet 网络（Fast Global Context Network）作为主干网络以替代 ResNet 网络，其网络结构如图 2-4 所示，该网络主要包括 CONV 卷积层，IRM 层和 GCblock 层三种基本组件。

CONV 卷积层是卷积神经网络中最基础的组件之一，用于提取输入图像的特征信息。

深度可分离卷积<sup>[48]</sup>是对卷积层的一种轻量化改进，包括两个步骤：深度卷积和逐点卷积，深度卷积用一个单通道的卷积核对输入特征层的每个 channel 单独进行卷积计算，因此通道数和输入保持一致，特征层的尺寸根据深度卷积的卷积 kernel-size 和 stride 大小而改变；逐点卷积对每个特征点进行  $1 \times 1$  卷积操作，仅改变通道数进行不同通道间的特征融合，不改变特征层的尺寸。用深度可分离卷积替代原始的卷积层可以实现相同特征变换的同时减少参数量和计算量。

IRM 层全称为 Inverted Residual Modules，即倒残差模块，其结构如图 2-5 所示，图 2-5 a) 表示 stride 为 1 的情形，图 2-5 b) 表示 stride 为 2 的情形，是一种对深度可分离卷积进一步改进后的网络结构。首先使用一个逐点卷积层扩展特征层的通道数，然后使用卷积核尺寸为  $3 \times 3$  的深度卷积进行特征提取，然后再使用一个逐点卷积对通道数进行压缩。这种设计相比原始的残差模块，内存效率更好，同时实验效果也有提升。

GCblock<sup>[49]</sup>是一种能够获取全局上下文信息的网络模块，应用了注意力机制以提高深度神经网络模型的识别能力。由于卷积层是一种局部连接的计算模块，每个卷积层处

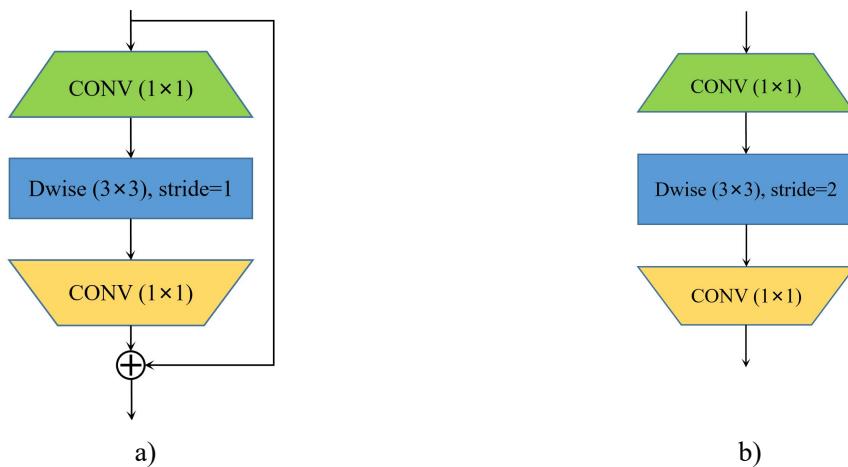


图 2-5 IRM 层结构图

理后的特征点由上一层特征部分区域计算得到，因此仅能够反应部分区域的信息，该区域也被称为感受野。特征点所在的特征层越深对应原图的感受野面积越大。但是感受野总是有限的，导致卷积神经网络难以反应全局信息。为解决上述问题，研究者基于注意力机制提出了 GCblock 模块，该模块包括上下文建模、特征转换和特征融合三大部分，兼具 Non-local block<sup>[50]</sup>全局上下文建模能力强和 SENet 模型<sup>[51]</sup>计算量小的优势，将其引入网络的特征层层尺度压缩的阶段，通过帮助特征层获取全局信息显著改善了模型的特征提取能力。

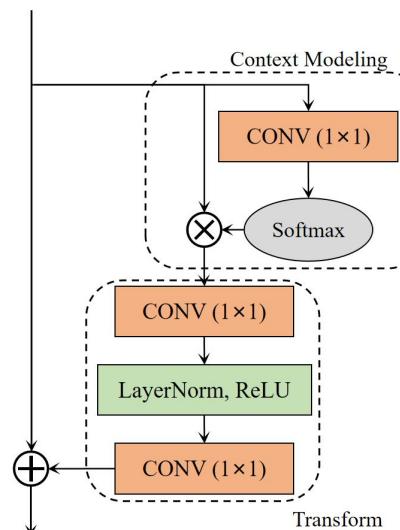


图 2-6 GCblock 结构图

经实验验证，使用 FGCNet 代替原始的 ResNet 主干网络，能够在保持模型检测性能稳定的前提下大幅降低模型的参数量和计算量，使得该模型能够更好地适应巡检任务中快速部署的需求。

## 2.2.2 FCOS 模型轻量化

由于表盘检测任务目标种类有限，仅区分单指针圆形表盘、单指针方形表盘和多指针表盘三类目标，相比 COCO 等大型目标检测数据集中 80 类目标这样的通用检测任务难度较低，因此对特征层数量需求也相对较低。基于上述认知，本节从精简特征层的角度继续研究了 FCOS 模型的轻量化问题，最终提出 Lite-FCOS 模型，网络结构图如图 2-7 所示。

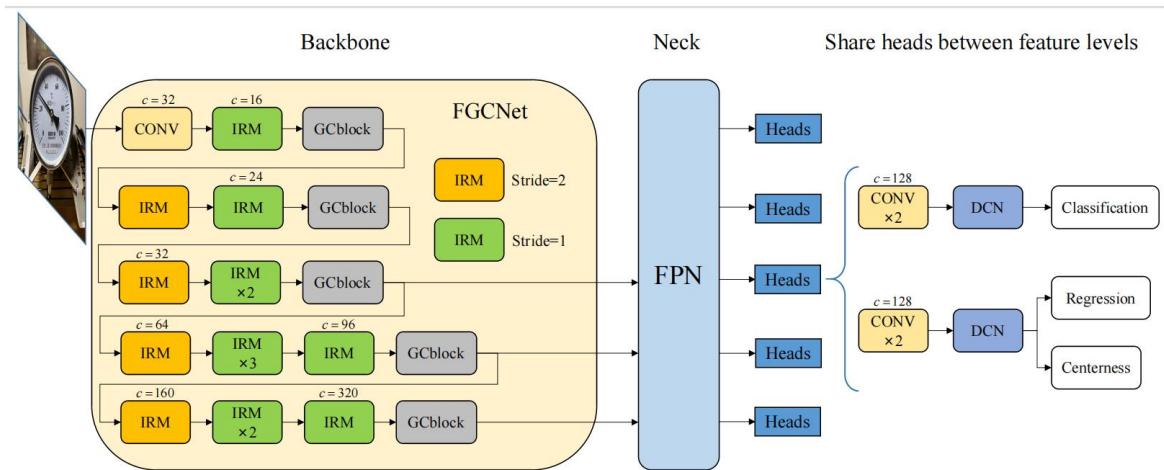


图 2-7 Lite-FCOS 模型网络结构图

Lite-FCOS 模型对 FPN 网络构建的特征金字塔进行了精简。标准 FPN 网络生成的特征金字塔中的每一层特征都设定为 256 个通道数，为对应 ResNet 主干网络的输出，其输入的三个特征层尺寸分别为 512, 1024 和 2048。根据卷积神经网络的特点，特征层越多模型的特征提取能力越强，但也会带来计算量、参数量增大和训练收敛难度增大的问题。由于本文已将主干网络替换为 FGCNet 网络，因此 FPN 输入特征层也要对应调整为 32, 96 和 320 以适应主干网络的改变。同时本文进一步压缩了 FPN 特征金字塔中每一层特征层的通道数为 128 层。由于 FGCNet 带来的特征层质量提高，上述改进使 Lite-FCOS 模型在显著减少了网络计算量和参数量的同时仍然保持了较好的检测精度和效果。

Lite-FCOS 模型对 FCOS 模型的检测头网络进行了精简和优化。原始 FCOS 模型参考 ReinaNet 模型在检测头网络分别使用 4 组重复的卷积对特征金字塔中的公共特征层进行特征提取然后进行类别预测和边界框的回归。由于在 FPN 网络中已经大幅降低了特征层的规模，因此在检测头环节也做出相应调整，将 4 组重复卷积调整为 3 组，包括两组卷积和一组 DCN 可变形卷积层<sup>[52, 53]</sup>。DCN 在正常卷积采样的基础上增加了一个偏移量，因此能够获取非固定的上文信息来进行卷积，能够显著提高模型的检测和感知能力，

DCN 可变性卷积示意图如 2-8 所示。

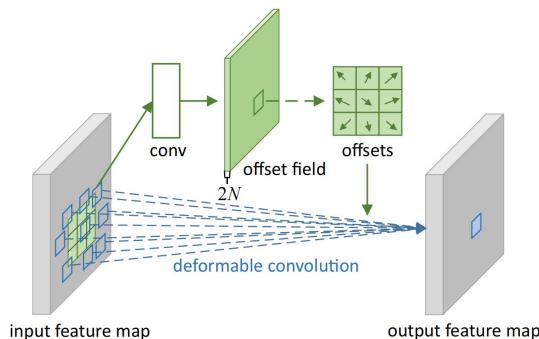


图 2-8 可变性卷积示意图<sup>[52]</sup>

## 2.3 Lite-FCOS 模型实验

为验证本文提出的 Lite-FCOS 模型的性能能够达到预期水平且相关改进都是有效的，本文设计并完成了两组对比实验和一组消融实验。

### 2.3.1 FGCNet 主干网络实验

为完成表盘检测模型的训练和测试，本文收集图像并制作了一个仪表巡检场景数据集，命名为 PMDD，为英文名 Pointer Meter Detection Dateset 的简写。其中所有图片均来自网络开源和实际拍摄的图像，共收集图像 6058 张，包括单指针圆形仪表、单指针方形仪表和多指针仪表三类。

使用了领域内通用的 COCO 评价标准作为目标检测模型的评价指标，具体包括：

- 1) 边界框平均预测精度 mAP：预测边界框和标签边界框 IOU0.5~0.95 每隔 0.05 测一个 AP（单目标平均预测精度）之后的均值，用来度量模型的检测精度。其衍生指标 mAP50 和 mAP75，分别表示 IoU 设置为 0.5 和 0.75 的 AP 平均值。
- 2) 模型推理帧率 FPS：为模型推理过程中单张每秒能够预测的图像数量，用来度量模型的检测速度和实时性。
- 3) 模型参数量 Params：为模型所占用的显存大小，用来度量模型对显存的要求。

为了验证 FGCNet 主干网络的性能，本文选用多种主流 SOTA 主干网络在 FCOS 模型中完成了主干网络对比实验，实验结果如表 2-1 所示。

由实验结合可知，本文提出的 FGCNet 网络在测试集上的性能指标在所有测试主干网络中达到了最优，其 mAP 甚至略高于参数量最大的 ResNet101 网络，验证其具备可靠的检测能力。与此同时，FGCNet 网络的参数量和检测速度也是所有测试主干网路中较优的，其参数量和检测时间仅略大于 MobileNetV2 网络，证明了 FGCNet 网络具备良

表 2-1 主干网络对比试验

方法	检测精度			检测速度	参数量
	mAP	mAP50	mAP75		
ResNet50 <sup>[40]</sup>	74.3	93.5	85.2	29.8	31.86 M
ResNet34	65.1	86.9	75.0	40.9	28.98 M
ResNet18	64.0	88.4	73.2	47.8	18.95 M
ResNet101	74.4	94.0	84.6	23.5	50.80 M
ResNest50	72.5	91.6	82.0	25.8	33.74 M
pvt <sup>[54]</sup>	72.9	93.1	83.6	26.5	20.67 M
HRNet <sup>[55]</sup>	68.4	90.9	76.0	18.7	37.14 M
MobileNetV2 <sup>[56]</sup>	74.2	94.3	85.5	46.4	10.05 M
FGCNet (ours)	75.0	94.7	85.8	43.9	10.69 M

好的轻量化性能，适宜应用在实际部署中。

### 2.3.2 Lite-FCOS 模型消融实验

为验证 Lite-FCOS 模型中 FPN 网络和 DCN 网络的改进有效性，本文对 FGCNet 网络进行了消融实验。由表格可以看出，使用了 FPN 网络通道数压缩后，模型的精度指标略有下降，但是速度和参数量指标得到了大幅改善。使用了 DCN 之后，速度和参数量指标略有下降，但是精度指标得到改善。实验结果表明，本文提出的两个网络轻量化的改进都是有效的，综合使用后，精度指标在维持稳定的基础上略有提高，而速度指标和参数量指标均得到了显著改善。

表 2-2 Lite-FCOS 消融实验

通道数压缩	DCN	mAP	mAP50	mAP75	FPS	Params
		75.0	94.7	85.8	43.9	10.69M
✓		74.1	93.8	84.8	59.0	4.14M
✓	✓	75.2	94.4	86.8	56.2	4.20M

### 2.3.3 Lite-FCOS 模型对比实验

为考察 Lite-FCOS 和其他同类目标检测模型之间的性能对比，本文使用相同数据集对若干典型目标检测 SOTA 模型进行了训练并完成了对比试验，实验结果如表 2-3 所示。由实验结果可知，Lite-FCOS 模型在检测精度方面处于中游水平，略低于 Faster R-CNN 等二阶段模型，但是优于 FCOS 等一阶段模型。Lite-FCOS 模型的优势在于其在具有较

表 2-3 Lite-FCOS 模型对比实验

方法	检测精度			检测速度	参数量
	mAP	mAP50	mAP75		
Faster R-CNN <sup>[47]</sup>	76.2	96.3	87.0	28.4	41.17M
Cascade R-CNN <sup>[57]</sup>	77.2	95.1	89.0	24.3	68.93M
Libra R-CNN <sup>[58]</sup>	77.4	95.8	88.7	27.4	41.40M
RetinaNet <sup>[46]</sup>	75.0	94.9	85.5	29.7	36.31M
Fovea Box <sup>[59]</sup>	74.0	93.7	86.3	30.7	36.02M
ATSS <sup>[60]</sup>	75.3	95.1	86.1	30.2	31.89M
FCOS <sup>[42]</sup>	74.3	93.5	85.2	29.8	31.86M
Lite-FCOS	75.2	94.4	86.8	56.2	4.20M

好的检测精度的前提下，大幅改善了检测速度和参数量，其检测帧率达到了 56.2FPS，相较其他所有同类模型快一倍左右，模型参数量仅为 4.2M 相比对比模型中参数量最小的 FCOS 模型也有 7.6 倍的提升，相比参数量较大的 Cascade R-CNN 模型提升了 16 倍。

总体来看，Lite-FCOS 模型在保证了检测精度的同时，大大改善了模型在检测速度和参数量方面的表现，这无疑将给模型的实际应用带来很大优势。

## 2.4 表盘图像增强

使用 Lite-FCOS 可以从场景图像中分割下仪表的表盘区域得到表盘图像。表盘图像将被用于表盘信息提取，并基于这些提取得到的表盘信息计算仪表的读数结果。表盘检测效果如图 2-9 所示。然而，表盘信息提取对图像质量要求较高，因此有必要对获取的表盘进行图像增强处理，以得到更高质量的表盘图像。



图 2-9 表盘检测效果图

图像常见的质量问题包括：噪声、清晰度不够以及遮挡。由于仪表读数识别需要算法正确地获取标度数字、指针的有效信息，基于角度法的仪表读数识别算法需要实际拍

摄的仪表图像正确反映指针和标度数字之间的空间位置关系。因此，对表盘图像进行增强十分有必要，对保证获取可靠的仪表读数信息有着重要意义。

### 2.4.1 图像去噪

图像噪声是数字图像处理领域最常见的图像干扰因素之一，对待处理图像进行图像去噪也是几乎所有计算机视觉任务中图像预处理阶段的重要环节。图像噪声产生的常见原因包括：图像传感器在拍摄过程中视场不够明亮、电路传感器自身噪声和相互作用的影响、图像传感器长时间工作导致的温度变化以及传输信道和解码过程受到干扰等。图像噪声可以通过优化图像获取和传输设备来尽量降低，但是难以完全避免，因此在处理图像之前先进行去噪是很有必要的。目前图像处理技术已经发展出了多种图像去噪方法，主要包括：高斯滤波、中值滤波和双边滤波等。

高斯滤波是最常见的一种图像平滑滤波器，其作用在于消除高斯噪声的影响，能够起到平滑图像的作用，基本原理和均值滤波器比较类似。二维高斯函数如下：

$$h(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2-2)$$

式中  $(x, y)$  表示待计算一个点的坐标， $\sigma^2$  表示图像的标准差。

利用二维的高斯函数进行离散化处理可以得到高斯滤波核模板，相当于以模板的中心位置为坐标圆点进行采样，用得到的高斯函数值作为模板的系数。在实际计算过程中，相当于利用高斯滤波核模板对一个像素点及其周围像素值求加权后的均值，然后用这个新的加权平均值作为该像素点滤波后的像素值。

对于高斯滤波这样的常规低通滤波器，默认假设是一个像素点和其周围的像素点像素数值总是比较接近的，所以可以在其邻域内通过高斯模板等滤波核计算出一个新的像素值作为滤波结果。这种假设在图像边缘位置则会带来问题，因为图像边缘本身就是像素值剧变的位置，通常来说像素值差异越明显，图像的边缘信息也就越强烈，对边缘相关的检测任务越有利，然而高斯滤波的计算过程会将这一部分也变得平滑，这就导致边缘高频信息的模糊甚至损失。

双边滤波<sup>[61]</sup>是一种同时考虑了被滤波像素点的空间域和频率域的综合滤波方式，基本思想是在高斯滤波的基础上增加了像素值权重项，使得滤波过程既能够考虑到距离因素，也能够兼顾像素值差异的影响。在计算过程中分别用两个高斯函数值表示空间距离和像素值的权重，分别记为  $G_s$  和  $G_r$  如公式(2-3)和公式(2-4)所示：

$$G_s = \exp\left(-\frac{\|p-q\|^2}{2\sigma_s^2}\right) \quad (2-3)$$

$$G_r = \exp\left(-\frac{\|I_p - I_q\|^2}{2\sigma_r^2}\right) \quad (2-4)$$

式中 q 和 p 分别代表了两点的坐标,  $I_q$  和  $I_p$  分别代表了两点的像素值,  $\sigma_s$  和  $\sigma_r$  分别代表了邻域内坐标和像素值的标准差。

该公式表示计算以 q 为中心的窗口中, 某点 p 的空间距离和像素值权重, 由上述两个模板相乘就得到了双边滤波的模板, 记双边滤波的计算结果为 BF, 其计算结果如下方公式(2-5)和公式(2-6)所示:

$$BF = \frac{1}{W_q} \sum_{p \in S} G_s(p) G_r(p) * I_p \quad (2-5)$$

$$BF = \frac{1}{W_q} \sum_{p \in S} \left( \exp\left(-\frac{\|p-q\|^2}{2\sigma_s^2}\right) \exp\left(-\frac{\|I_p - I_q\|^2}{2\sigma_r^2}\right) * I_p \right) \quad (2-6)$$

式中,  $W_q$  表示整个窗口内的所有像素值权重之和, 用于进行归一化处理:

$$W_q = \sum_{p \in S} G_s(p) G_r(p) = \sum_{p \in S} \exp\left(-\frac{\|p-q\|^2}{2\sigma_s^2}\right) \exp\left(-\frac{\|I_p - I_q\|^2}{2\sigma_r^2}\right) \quad (2-7)$$

由于后续需要识别表盘图像中的指针和字符目标, 而这两类目标的边缘特征都十分明显, 能够保留清晰和完整的边缘信息对于后续识别也非常重要, 因此使用双边滤波能够更好实现表盘图像去噪, 在对表盘图像中的连续区域进行平滑的同时, 也能最大程度的保留图像的边缘信息。

中值滤波是一种非线性统计排序滤波器, 其作用在于消除椒盐噪声影响和平滑图像。该滤波器把图像中的一个像素点用其一定邻域内所有像素的中值点像素值来替换, 只需要确定该邻域的窗口大小, 不需要应用其他的人工设计参数。中值滤波尤其适合应对椒盐噪声的问题, 因为椒盐噪声通常以大量孤立脉冲点的形式存在于图像中, 在窗口内排序的过程中一般都在数值的最大或者最小一端, 因此不会被保留下。

综上所述, 本文结合双边滤波和中值滤波作为图像去噪的滤波器, 滤波效果如图 2-10 所示。

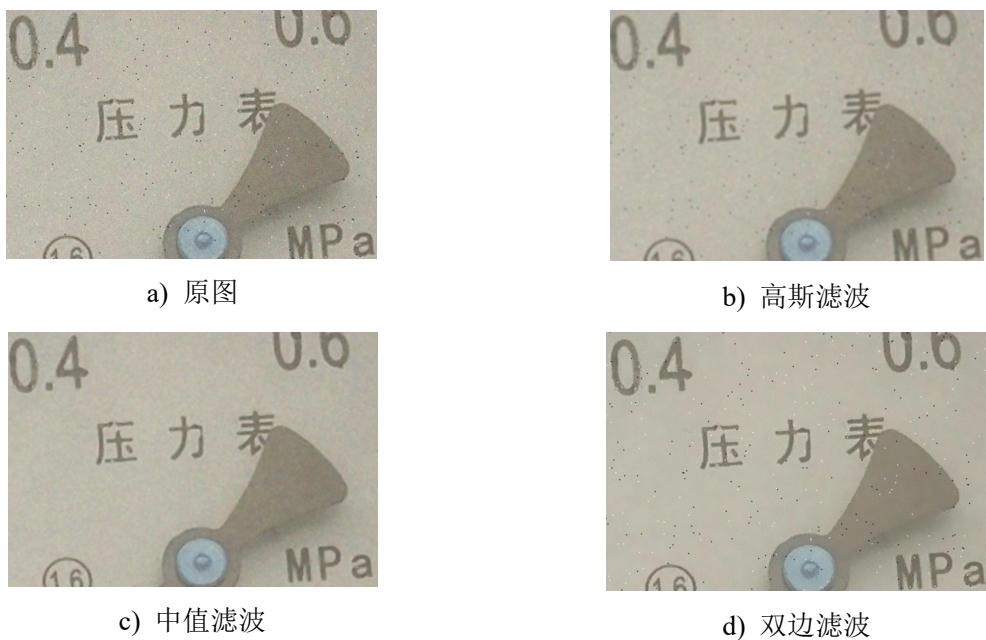


图 2-10 不同图像去噪方法对比效果图

#### 2.4.2 基于超分辨率重建的表盘图像清晰度增强

真实工业场景中的仪表常因距巡检机器人过远，表盘表面存在积灰或雾气等原因导致表盘清晰度不足，在本文中统称这种情况为模糊表盘，如图 2-11 中的例子所示。模糊表盘对表盘信息提取造成干扰，因此有必要设计相应处理环节以提高模糊表盘图像的清晰度。为最大限度提高表盘图像的质量，本文采用超分辨率重建技术以提高表盘图像清晰度，相比常规的去模糊、去噪等图像处理方法拥有更好的图像增强效果。

图像超分辨率重建，简称 SR，是由一张或多张低分辨率图像得到高分辨率图像的过程<sup>[62]</sup>。随着深度学习技术兴起，基于深度神经网络模型的 SR 方法逐渐占据了主流。2014 年，Dong 等提出一个包含 3 层卷积的神经网络，命名为 SR-CNN<sup>[63]</sup>，即用于超分辨率重建的卷积神经网络，该网络本质是形成了一个从低分辨率图像到高分辨率图像的



图 2-11 清晰度不佳表盘示意图

映射，从而实现了端到端的图像超分辨率重建，其效果比传统的图像超分算法更好。后续，SR 技术经历了深度学习方法的研究热潮，不断有新的模型<sup>[64]</sup>被不断提出。

本文选用近年效果较好的 SR 模型 RealSR<sup>[65]</sup>实现表盘清晰度增强任务，该模型利用一种无监督超分辨率重建算法，基于 ESRGAN 模型<sup>[66]</sup>改进而实现，能够直接对真实世界中的图像进行重建。文献中提出了一种退化框架，使数据集更好的保留了与频率相关的细节问题，从而使训练得到的模型在真实世界图像中的 SR 效果更好。由于缺乏足量高质量的表盘图像数据支持 SR 模型的微调训练，本文直接应用 PPGAN 框架中的开源 RealSR 超分辨率重建模型实现此部分，并使用 OCR 模型的识别结果对比展示图像清晰度增强带来的改善效果。

实验效果如图 2-12 所示，图中上方图像为未经 SR 清晰度增强，直接从目标检测模型得到的表盘图像送入 OCR 模型后的字符检测结果，图中下方的为对应经过 SR 清晰度增强之后的表盘图像送入 OCR 模型后的字符检测结果。可以看出图像超分辨率重建能够改善仪表图像的清晰度，OCR 模型的识别效果得到了显著增强，很多 SR 清晰度增强前无法识别的字符经过处理之后都被准确检测了出来，这对于提高仪表读数识别算法的准确率起到了有益作用。

#### 2.4.3 大角度倾斜表盘检测

由于机器人定位精度和摄像头调整范围有限，拍摄存在仪表无法正对摄像头的情况，导致拍摄得到的表盘是倾斜的。当倾斜程度较小的时，对仪表读数影响较小，主要体现在读数计算的误差上，此时即便不做校正处理，也能得到基本可靠的读数结果。当表盘倾斜程度较大时，则可能产生严重不利影响。

由于本文的研究目标要求算法脱离表盘先验信息也能实现正确读数，因此在本阶段仅有表盘图像信息，缺乏理想表盘信息，因此本文未实现表盘校正环节。但在实际应用中多数巡检机器人拍摄的仪表表盘都是基本正对相机的，大角度倾斜是少数情况。为提前发现上述问题，本文依据表盘图像的宽高比提出一种大角度倾斜表盘检测策略。由于多数表盘宽高比都为 1，本文在分析了大量表盘图像后得出当宽高比大于 1.2 的图像有较大概率属于大角度倾斜表盘的结论，对于这些表盘算法在给出读数结果后应当标记检测结果供人工后续核验。本方法虽不完善，但在测试中验证的确能够提前发现大角度倾斜仪表，从而在检测阶段就将其发现，避免算法的错误读数，具有一定的参考价值。

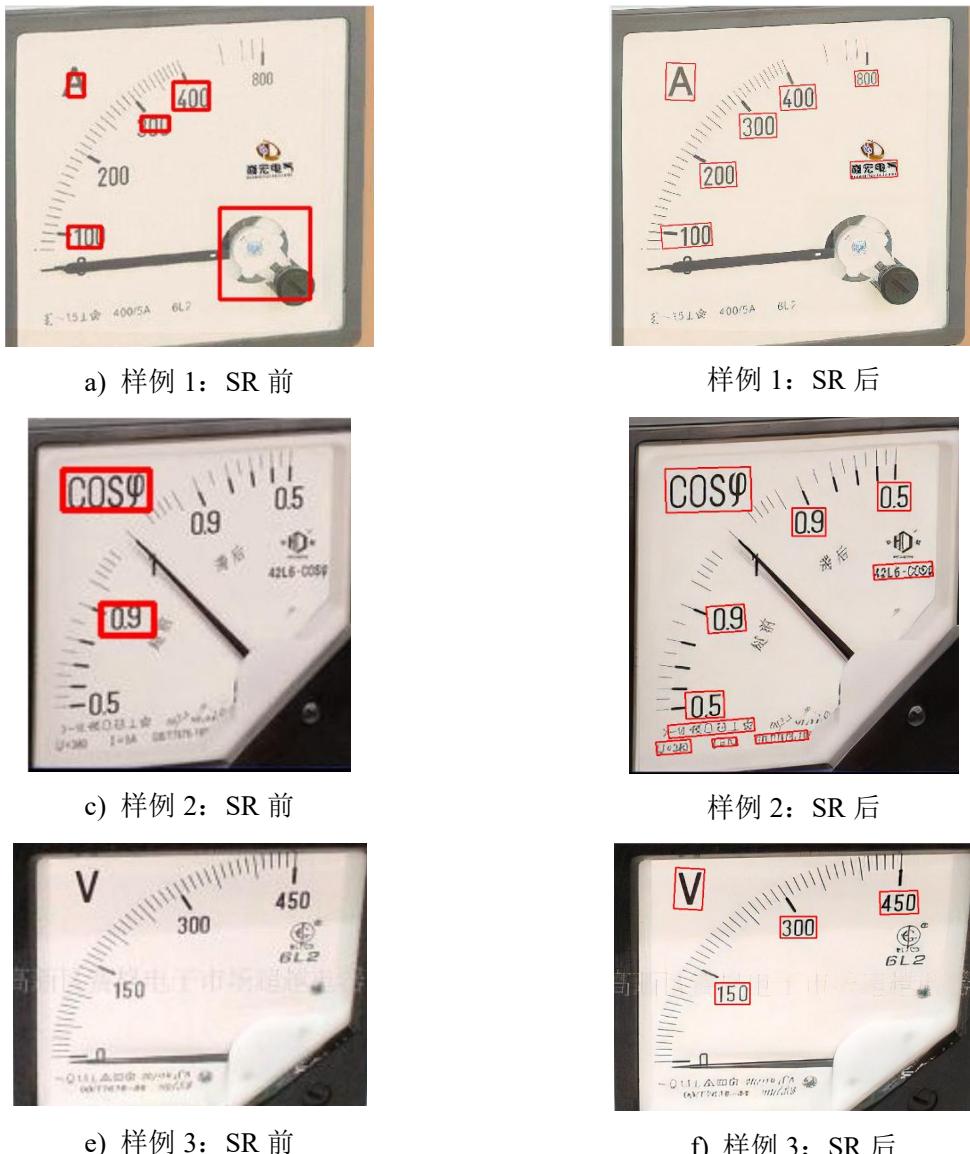


图 2-12 SR 清晰度增强前后对比图

## 2.5 本章小结

本章针对仪表自动读数识别过程中的表盘图像提取问题做了深入研究，提出了一种用于表盘检测的轻量化目标检测模型——Lite-FCOS 模型，该模型基于 FCOS 模型改进的，应用了本文提出的 FGCNet 轻量化主干网络，并对模型的特征融合和检测头部分也做了轻量化改进和优化。利用不同的主干网络和 SOTA 目标检测模型对 Lite-FCOS 模型进行了充分的实验，验证了其保持了较好检测精度，mAP 可达 94.4%。同时，在检测速度和参数量方面相比同类算法有了很大改进，检测速度可达 56.2FPS，参数量仅为 4.2M，效果显著好于同类算法。为增强表盘图像质量，本文提出了应用双边滤波和图像超分辨率重建模型对表盘进行图像增强的方法，并且展示了实验效果图。最后，在章节末尾提出了对大角度倾斜表盘进行检测的策略。

### 第三章 标度信息提取

为实现仪表的正确读数需要准确获取表盘信息，包括标度信息和指针信息。由于本文采用基于角度法的仪表读数计算方法，因此标度信息在本文中指表盘上所有标度数字的语义和位置信息，语义信息指标度数字对应的标度数值，位置信息指标度数字的中心点坐标，提取结果示意图如图 3-1 所示。针对当前大多数文献都将标度信息视为先验信息导致仪表读数识别算法泛化能力不足的问题，本文分别从 OCR 和目标检测两种不同技术视角切入，探索并实现了两种标度信息提取方法。基于 OCR 的方法先提取表盘全部字符信息，然后通过筛选的方式获取标度信息，基于目标检测的方法先提取表盘全部单个数字字符，然后通过数字融合的方式获取标度信息。

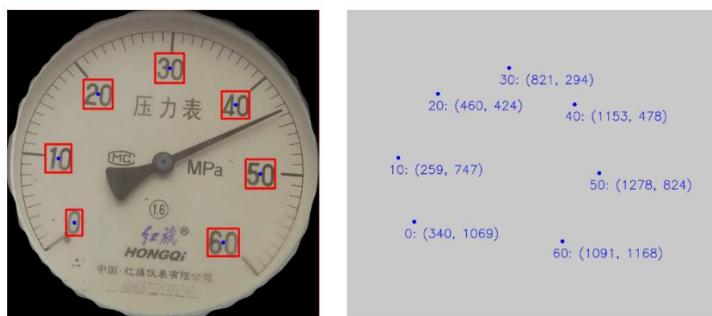


图 3-1 标度信息示意图

#### 3.1 OCR 提取标度信息

OCR 的全称为 Optical Character Recognition，意为光学字符识别，是指提取图像中的文字信息。早期 OCR 技术主要用于文档识别等机器视觉系统中，对拍摄清晰度，对正程度和光照条件都有较高要求。随着深度学习技术发展，深度神经网络在 OCR 领域成为研究热点，其强大的特征提取能力大幅推动了相关技术的快速发展。如今 OCR 技术已经能够实现自然场景下工作，也涌现了多个高质量开源代码库。OCR 模型能够在复杂的自然场景中准确对字符进行检测和识别，使其具备了自动获取表盘图像中标度信息的潜力。本节对基于 OCR 模型的表盘图像标度信息提取进行了深入研究。

目前的主流 OCR 模型一般由两个具体的深度神经网络组成，分别是文本检测模型和文本识别模型。本文检测模型主要解决从待测图像中获取文本所在区域并对这些区域进行提取和分割的问题；文本识别模型主要解决从文本检测模型获取得到的字符图像中预测该图像所对应的字符本文的问题。如图 3-2 所示，图片左侧展示了文本检测的效果，图片右侧展示了基于文本检测结果得到的文本识别效果。

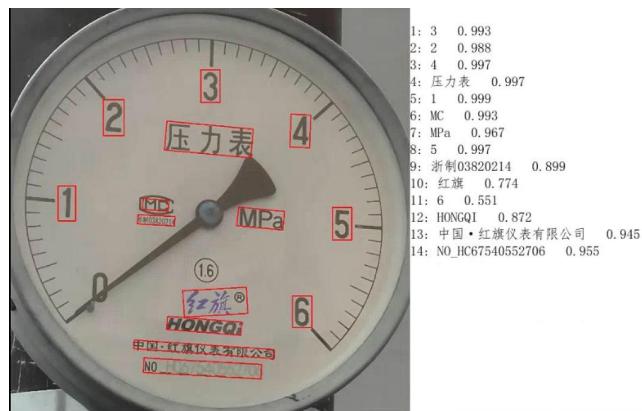


图 3-2 OCR 模型效果图

### 3.1.1 DBNet 文本检测模型

文本检测模型主要分为基于传统方法和基于深度学习方法两个发展阶段。传统文本检测方法大致分为预处理、版面处理、图像切分、特征提取与模型训练和识别后处理等几个步骤，通过提取图像特征然后基于经典机器学习算法实现文本检测。此类方法对复杂背景的干扰十分敏感，仅在 OCR 技术早期用于从文档图像这种背景干净的图像中提取字符区域，对背景复杂的自然场景图像中的字符检测能力有限。

基于深度学习的文本检测方法是自然场景 OCR 技术的基础，通过深度学习模型对图像中可能存在文字的区域进行检测和分割，其具体实现主要分为两类方法，一种是基于回归的文本检测，一种是基于分割的文本检测。

基于回归的文本检测受目标检测算法的启发，将文本区域作为一种目标，然后对目标检测模型进行改进以提高检测能力。由于文本区域一般在图像中呈现为长宽比不定的长矩形，且文本区域由多个相近字符组成，不具备闭合的边缘轮廓，这些特点和通用目标检测模型所处理的目标有一定区别，因此直接使用 Faster R-CNN 等目标检测模型直接对文本区域进行预测效果并不好。因此，一些研究者针对文本区域目标的特点对通用目标检测模型进行改进从而提出了相应的文本检测模型，如 Tian 等<sup>[67]</sup>提出的 CTPN 模型是对 Faster R-CNN 模型进行的改进，Liao 等<sup>[68]</sup>提出的 TextBoxes 模型是对 SSD 模型进行的改进。基于回归的文本检测主要的问题在于其对倾斜文本和弯曲文本的检测能力不足，且计算过程较为繁琐。

基于分割的文本检测模型受到实例分割模型的启发，用掩模分割而不是边界框回归的方式从图像中提取文本区域。旷视团队提出的 PSENet 模型<sup>[69]</sup>利用渐进式尺度扩张的方法进一步解决了分割过程中的字符粘连问题，进一步提高了此类算法的效果，但是该解决思路也导致了网络的计算速度较慢。

DBNet 是当下主流的文本检测模型，具有良好的性能，本文选用 DBNet 作为文本检测环节的实现模型。DBNet 的网络结构图如图 3-3 所示，首先通过一个主干网络提取图像特征，然后利用一个 FPN 网络对图像特征进行特征融合以实现更好的特征提取效果。FPN 输出的不同尺寸的公共特征层经过上采样统一到其中最大的尺寸上，然后经过特征堆叠形成公共特征层。接着，分别用两个分支基于公共特征层生成文本区域的概率图和阈值图，然后基于 DB 模块利用上述的概率图和阈值图生成文本区域的二值图。其中，概率图中的每个像素都是指输入图像中对应位置在属于文本区域的概率，阈值图中的每个像素指对应像素位置上进行二值化过程中概率的阈值，若概率图中该像素值高于阈值则输出二值图该位置为 1，否则为 0，二值图就是利用上述阈值图对概率图进行处理之后得到的结果，也就是最后用于分割文本区域的掩模图像。

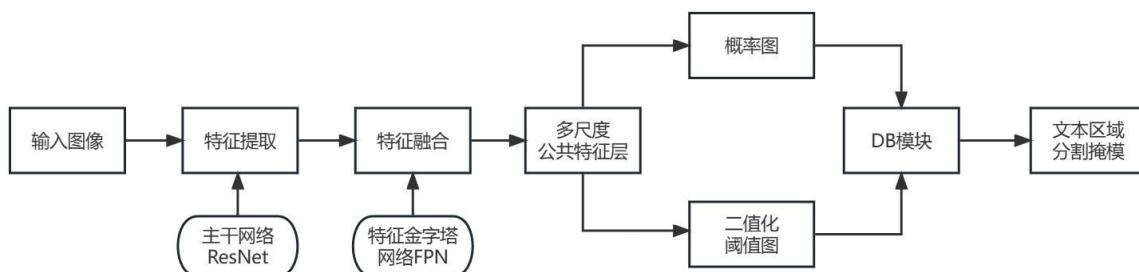


图 3-3 DBNet 网络结构图

相比其他文本检测模型，DBNet 最大的创新点在于其能够自适应的给文本检测得到的字符区域掩膜设定二值化阈值。一般情况下二值化的公式如公式(3-1)所示：

$$B_{ij} = \begin{cases} 1 & P_{ij} \geq t \\ 0 & \text{其他} \end{cases} \quad (3-1)$$

式中， $B_{ij}$  表示二值化之后  $(i, j)$  位置的像素值， $P_{ij}$  表示二值化之前  $(i, j)$  位置的像素值， $t$  表示二值化的分割阈值。

该公式并不是一个可以求导的公式因而无法应用在神经网络的训练中，因此无法使用深度学习模型训练得到二值化阈值。DBNet 中的 DB 模块通过引入了一个可微分二值化函数对标准二值化函数进行了替代，从而使二值化阈值的网络训练成为了可能。该可微分二值化函数的公式如公式(3-2)所示：

$$\hat{B}_{ij} = \frac{1}{1 + e^{-k(P_{ij} - T_{ij})}} \quad (3-2)$$

式中  $\hat{B}_{ij}$  表示二值化之后  $(i, j)$  位置的像素值， $T_{ij}$  表示  $(i, j)$  位置的分割阈值， $k$  表示放大

系数，在文献中取固定值 50。

公式(3-1)和(3-2)的函数图像高度相似，因此都具有二值化的功能，但是公式(3-2)是可微分的，因而也就可以利用神经网络的进行训练。

### 3.1.2 CRNN 文本识别模型

CRNN 模型是一种结合了卷积神经网络 CNN 和循环神经网络 RNN 的端到端不定长文本识别模型，通过将文本识别转化为时间序列的序列学习问题避免了对单个文字进行切割的步骤。作为一种相对早期提出深度学习模型，CRNN 的模型结构并不复杂，从上到下依次包括卷积层，循环层和 CTC 层，有些文献中也称其为转录层，其具体的网络结构图如图 3-4 所示：

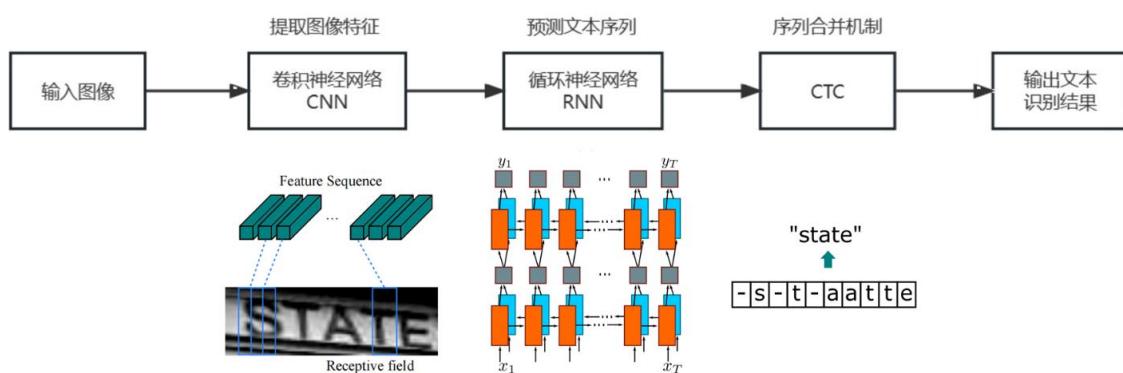


图 3-4 CRNN 网络结构图

CNN 模型被用于提取图像特征，在 VGG 主干网络<sup>[70]</sup>的基础上改进之后得到，主要是对第二层和第三层的最大池化层卷积核的尺寸进行调整，这样做的目的是为了是 CNN 输出的特征层能够和 RNN 的输入相匹配并且更加符合字符目标的特点，同时在第五层和第六层的卷积层之后增加了一个批归一化层（Batch Normalization），可以加快网络的收敛速度。

RNN 层用于基于 CNN 层得到的图像特征预测文本序列，完成从图像到文本的转换，图像特征在送入 RNN 之前应该首先从图像特征中生产特征向量序列，特征向量序列中的每个元素都是图像特征的一个小矩形区域。RNN 相比 CNN 能够更好的提取序列的上下文信息，并且能够处理不定长序列，这都是直接使用 CNN 所无法实现的效果。常规的 RNN 在训练过程中遇到过长的序列很容易出现梯度消失或者梯度爆炸的问题，因此作者使用了双向 LSTM 层作为传统 RNN 层的代替，在一定程度上解决了这个问题。双向 LSTM 层的另一个好处就是能够更加充分的利用上下文信息，提升识别的准确率。

CTC 层用于将 RNN 层输出的预测序列转化为最终的识别结果，主要是用于去除预测序列中的冗余信息，具体是通过 blank 机制实现的。CTC 层的关键是要在对重复的冗余字符进行合并的同时又避免将并不应该合并的字符误处理，比如“book”中的“oo”。解决办法是 RNN 在输出序列时会在文本标签中重复的字符之间插入一个 blank “-”，有 blank 出现的地方就不进行合并，而模型的训练使得 CTC 层能够正确的找到插入 blank 的序列位置，从而解决了字符重复冗余的问题。

### 3.1.3 OCR 模型的训练和优化

现有的 OCR 模型已经有很多优秀的开源框架实现包括 mmocr, PaddleOCR 等。完整的 OCR 模型训练需要足量图像和语料作为训练数据集，同时也需要较高的计算资源开销，受到条件限制本文难以实现 OCR 模型的完整训练，而是选择利用自己标注的表盘数据集对基于 PaddleOCR 框架的中的文本检测模型进行模型微调以改进其检测性能。

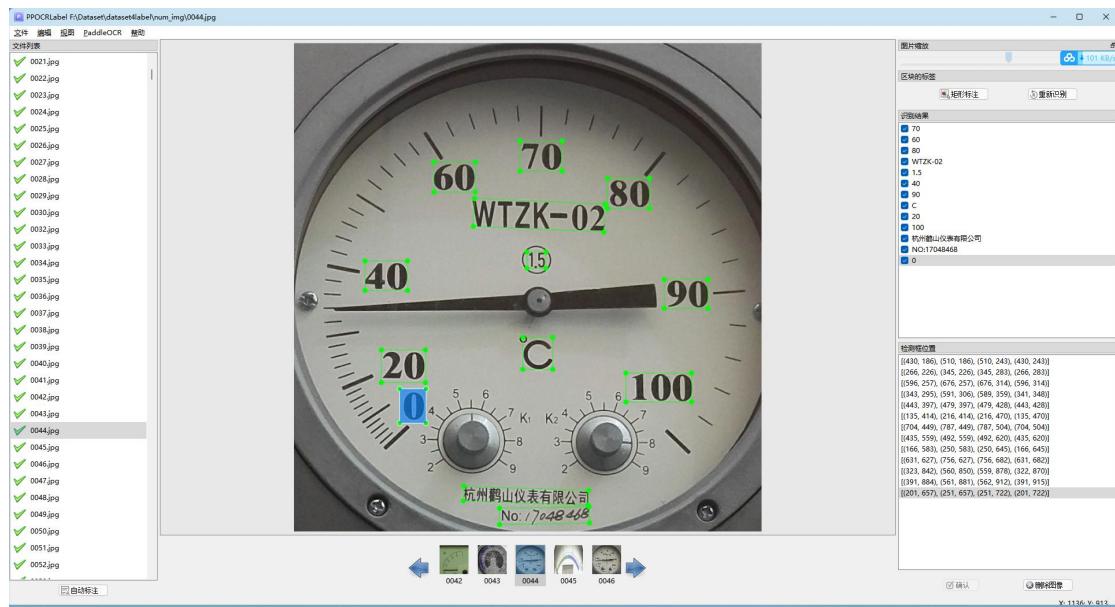


图 3-5 PPOCRLLabel 标注软件示意图

PaddleOCR 框架中的模型针对得是通用场景下的 OCR 识别问题，因此理论上也可以直接使用框架开源的通用 OCR 模型实现表盘字符的检测和识别任务，然而实践中发现检测效果并不好。因此，本文收集表盘图像并基于 PaddleOCR 框架中的 PPOCRLLabel 标注工具制作了表盘图像字符数据集——DRID-OCR，用于模型的微调训练，共收集表盘图像 1448 张，其中 1159 张作为训练数据集，289 张作为验证测试集，使用 OCR 标注工具对表盘中的字符目标都做了非常详尽的标注，标注软件的界面如图 3-5 所示。

经过微调训练之后，OCR 模型的检测效果有了十分显著的提升，如图 3-6 所示，图 a 所示为通用 OCR 模型的识别效果，图 b 为经过微调训练之后的 OCR 模型的识别效果，

可以从图中看出模型的检测能力有了很大改善，可以对表盘中那个所有字符得到较好的检测结果，更多的仪表读数实验也验证了这一点，详细实验信息见第四章的仪表读数相关小节。

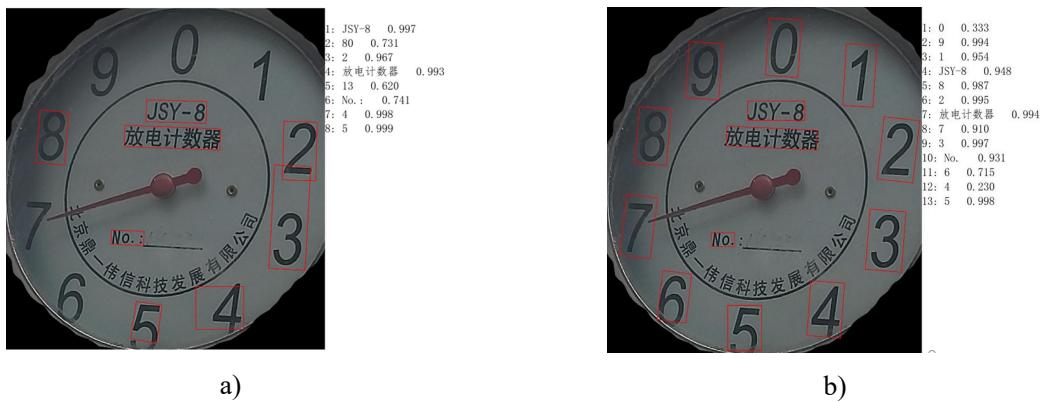


图 3-6 OCR 模型训练前后效果图

### 3.1.4 OCR 识别结果筛选

通过 OCR 模型获取得到的信息是表盘图像中所有字符的信息，大体可以包含三类信息，标度数字、仪表类型相关字符和其他字符。标度数字是表盘图像中用于展示指针偏转角度和仪表读数关系的数字。仪表类型相关字符用于标识仪表测量的物理量种类，一般以名称或者物理量单位的形式标注在表盘上。其他字符指上述标度数字和仪表类型相关字符之外的所有其他表盘图像中的字符，常常包括仪表的厂家，厂商的型号编号等。

如本小节引言所述，本阶段要解决的核心问题是仪表标度数字的语义及位置信息，因此在获取了 OCR 模型的识别结果之后需要将标度数字从识别结果中筛选出来，筛选的结果就是标度信息。为解决达到上述目的，本文依据指针式仪表表盘图像的特征提出了一种简单快速的标度数字筛选办法。

表盘图像特征 1：表盘图像中，除了标度数字之外其他的字符内容很少有其他的纯数字字符。依据上述特征，首先将 OCR 识别结果中的非数字字符全部剔除，保留 OCR 识别结果中的数字字符作为标度数字的预筛选结果。

表盘图像特征 2：表盘图像中，标度数字的数值按照一定的规律排列，且由于本文的研究对象限定在单指针单标度且刻度均匀的仪表，这些标度数字的数值经过排序之后应该是一个等差数列。依据上述表盘特征，可以对所有纯字符数字排列后计算和相邻字符数值之差，然后利用字符左右侧字符差值判断该字符是否需要被剔除。记升序排列之后的纯数字字符构成的数组为 N，

$$N = \{n_1, n_2, \dots, n_i, \dots, n_m\} \quad (3-3)$$

式中,  $n_i$  为数组中第  $i$  个字符数值,  $m$  为数组的总长度,

可以计算其和相邻字符数字之差的绝对值, 称为标度数字跨度, 记为  $\Delta n_i = |n_i - n_{i-1}|$ ,  
由标度数字跨度可以构建一个数组为  $\Delta N$ ,

$$\Delta N = \{\Delta n_2, \Delta n_3, \dots, \Delta n_i, \dots, \Delta n_m\} \quad (3-4)$$

在理想情况下, 所有的标度数字都被正确的检测到且没有多余的数字被误判为标度数字的情况下,  $\Delta N$  中的所有标度数字跨度应该都是一样的。在算法实验过程中允许存在个别标度数字被漏检的情况, 因此仅剔除标度数字跨度明显过大的标度数字。具体使用正态分布的  $2\sigma$  原则作为判断依据, 若该标度数字跨度  $\Delta n_i$  位于  $(\mu - 2\sigma, \mu + 2\sigma)$  区间之外, 则判定该标度数字属于离群点标度数字, 予以剔除, 其中  $\mu$  表示数组  $\Delta N$  的均值,  $\sigma$  表示数组  $\Delta N$  的标准差。

经过上述两个筛选步骤处理之后, 保留下的 OCR 识别结果就是最终得到的仪表标度数字信息, 可用于后续的仪表读数计算环节。

### 3.2 OCR 提取仪表类型

绝大多数仪表自动读数识别相关文献的重点关注的都是读数计算, 而不包括仪表类型识别, 通常直接忽略该步骤或通过精准的表盘先验信息获取方法, 如模板匹配或者深度学习预测直接得到具体仪表的先验信息, 从而确定仪表类型。OCR 模型的优势在于可以检测识别表盘上标度数字外字符信息, 这其中就包含了仪表类型相关信息, 因此能够面向未知仪表实现自动识别仪表类型和读数单位。

本文提出采用字符串匹配的方法实现上述目标。尽管仪表标度盘设计会有不同, 但是仪表的类型是有限的, 可以人工枚举场景中所有仪表, 然后根据仪表的名称关键字对 OCR 模型得到的字符串进行匹配。

本文实现了一个常用仪表关键字列表如下表 3-1 所示, 在实践中可以通过添加关键字的模式将新引入的仪表类型加入其中而不需额外的训练和构建模板。虽然这里的仪表关键字列表也可以视为先验信息的一种, 但是这种列表信息的构建成本极低能够快速进行增删改操作, 且考虑到常用仪表的种类非常有限, 理论上可以将所有仪表类型都加入下方的表格中, 因而这种先验信息并不会对算法的泛化能力产生较大负面影响。

设定名称和单位两类读数关键字作为判断依据。首先对 OCR 识别的各个字符串进行检验, 若一个字符串包含名称关键字, 则该字符串被判定为名称相关字符串, 若该字

表 3-1 仪表类型匹配关键字表

仪表类型	名称关键词	单位关键词
压力表	压力	MPa
温度表	温度	C
电流表	电流	A
电压表	电压	U
油位计	油位	L
放电计数器	放电	F

字符串和单位关键字一致，则该字符串被判定为单位相关字符串，以上两种统称为类型相关字符。若该仪表仅得到一个类型相关字符或者多个类型相关字符对应同一类型仪表，则直接根据上表确定仪表类型；若该仪表对应多个相关类型字符且对应不同仪表类型，则判定该仪表类型暂时无法确定，交由人工进行进一步核验。

如图 3-7 所示，OCR 识别仪表类型所示，仪表读数检测得到的字符串中“温度计”字符串包括了名称关键字“压力”，因此被检测得到用绿色边界框标识出来，字符串中的“MPa”和单位关键字一致，因此被检测用红色边界框标识出来。两个读数关键字对应的仪表类型是一致的，因此该仪表被判定为温度表。



图 3-7 仪表类型识别示意图

### 3.3 目标检测提取标度信息

在实现了基于 OCR 模型的仪表标度信息提取算法后，本文又尝试了应用目标检测模型解决这一问题，探索并实现了第二种具有强泛化能力的仪表标度盘信息提取算法。字符区域长度不定是制约目标检测模型直接应用于 OCR 文本检测领域的主要因素，然而表盘图像中的所有标度数字都是由 0 至 9 这 10 个基础数字构成的，且通常标度数字都是印刷体阿拉伯数字，长宽比较为稳定，因此利用目标检测模型对表盘图像中的单个

数字字符进行检测和提取是可行的。在此基础上本文进一步提出一种基于目标检测的标度信息提取算法，将标度信息提取过程分为单个数字字符检测和数字字符融合成为标度数字两个步骤解决。

### 3.3.1 数字字符检测和跨类别 NMS

对 3.1.3 节中的 DRID-OCR 中的表盘图像进行 COCO 目标检测格式的重新标注，标注目标为 0 至 9 一共 10 个数字字符，得到的新数据集命名为 DRID-DET，其训练集和验证集的划分比例和 DRID-OCR 中保持一致。基于 DRID-DET 数据集训练 2.2 节中 Lite-FCOS 模型，得到一个表盘数字目标检测模型，具体的模型设计细节这里不在重复赘述。利用该模型对表盘图像进行检测，得到检测结果如图 3-8 所示，蓝色边界框为目标检测模型输出的边界框结果，将单个字符目标的预测标签打印在了边界框下方：

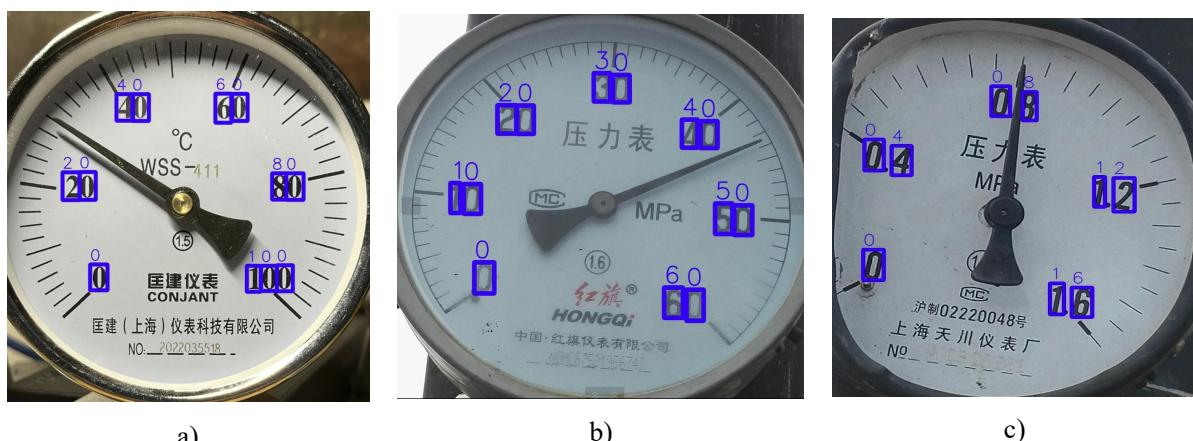


图 3-8 目标检测提取单个字符效果图

目标检测模型常常会在目标周围生成若干置信度同时满足设定阈值的边界框，这时就需要对这些边界框进行筛选保留一个作为最终识别结果否则就会出现重复检测的问题。NMS( Non-maximum suppression)，非极大值抑制，是一种目标检测模型的常用后处理环节，就是用于解决上述问题。对于多类别的目标检测，常规 NMS 方法是对不同类别的目标分别进行独立的 NMS 处理，然后呈现最终的结果。这种设计有效避免了相近的两个不同类别目标由于边界框重合度过高导致一个目标被漏检的问题。

在表盘图像中，各个标度数字都是相互独立的，正常情况不会出现两个标度数字重叠在一起的情况，而部分字符可能由于较为相似或者受到干扰而出现了字符变形，导致同时达到多个不同字符的置信度阈值被目标检测模型识别出来。根据常规 NMS 的计算规则，尽管这些字符的边界框已经高度重合了但还是会被全部保留。如图 3-9 所示，指针距离标度数字“40”非常近以至于遮盖了字符，导致其中的字符“4”因受到干扰而

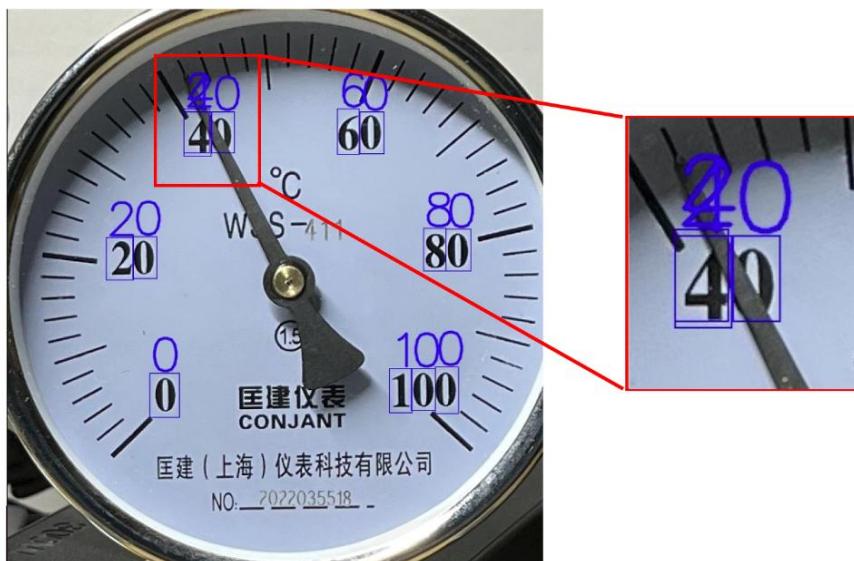


图 3-9 指针重复检测示意图

被同时识别为“2”，“4”这两种字符。如果此时进行数字融合，则会得到标度数值为“240”的标度数字，导致识别错误。

为避免上述检测误差的发生，需要对目标检测模型输出的所有字符再进行一次不区分类别的跨类别 NMS 处理。首先将所有字符分为三类进行讨论：经过 NMS 处理的保留字符，被剔除的删除字符和未经过 NMS 处理的待处理字符。另外定义 IoU 检验过程为对两个字符计算其边界框的交并比 IoU，然后依据 IoU 是否超过预先设定的阈值判定两个字符是否属于重叠字符的过程，当 IoU 过大则表示两个字符的边界框高度重复。IoU 的计算如公式(3-5)所示：

$$IoU = \frac{S_1 \cap S_2}{S_1 \cup S_2} \quad (3-5)$$

式中  $S_1$  和  $S_2$  分别表示了两个正在计算 IoU 的边界框的面积

跨类别 NMS 的具体流程如下，流程图如图 3-10 所示：

- 1) 任取一个待处理字符作为本循环待处理字符；
- 2) 从剩余未和当前字符进行过 IoU 检验的待处理字符中任取一个与当前字符进行 IoU 检验；
- 3) 如果该字符未被判定为重合字符，则继续重复步骤 2) 直至所有剩余待处理字符都已经和该字符完成 IoU 检验，该字符被分类为保留字符；若被判定为重合字符，则比较该字符和重合字符的置信度，并剔除置信度低的字符，即将其分类为删除字符，如果本循环被处理的字符被剔除则回到步骤 1)，否则继续重复步骤 2) 直至该字符完成 NMS。

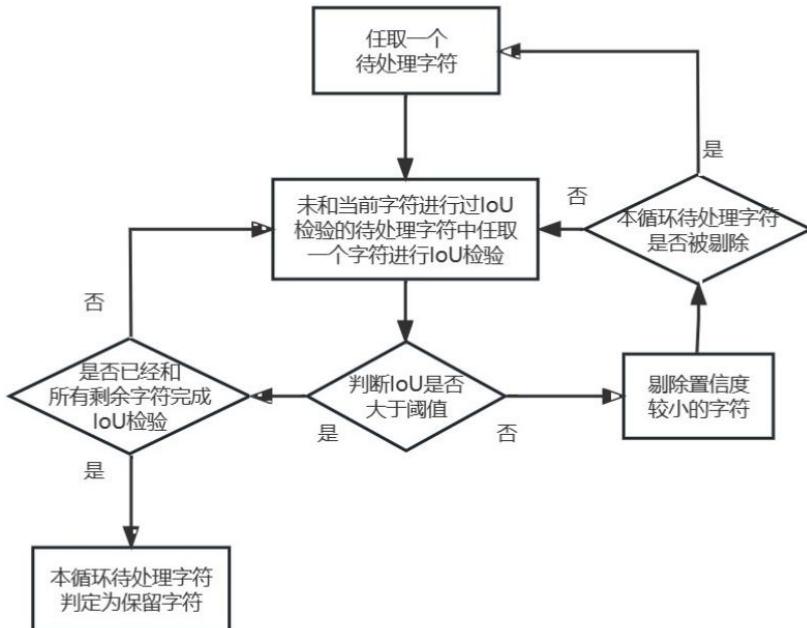


图 3-10 跨类别 NMS 流程图

4) 最后对经过 NMS 处理之后的字符进行基于边界框面积的筛选，剔除面积明显过小或者过大的字符。

经过上述处理，可以得到较为可靠的仪表单个字符检测结果，接下来就可使用这些仪表的单个字符检测结果提取仪表的标度数字信息。

### 3.3.2 待融合字符组合

标度数字可能包含若干位数字字符，也可能包含小数，从表盘图像中的单个数字字符计算标度数字信息的过程本质就是找到“组成同一个标度数字”的所有单个数字字符，然后生成新的标度数字。依据表盘图像的特征，“同一个标度数字”的组成数字字符最直接的特征就是他们之间都是相邻的，距离很近，而不构成同一个标度数字的数字字符通常距离相对较远。由此，本文提出将获取标度数字的过程分为两个阶段，首先根据单个数字字符的之间距离进行划分组合，得到“可能构成同一个标度数字”的单个数字字符的字符组，本文称为“待融合字符组”；然后基于待融合字符组生成各个标度数字，所有标度数字都确定之后也就得到的表盘图像的标度盘信息。

待融合字符组组合流程图如图 3-11 所示，具体计算过程如下：

1) 将所有数字字符划分为未融合字符和已融合字符两类；定义字符位置坐标为字符边界框的中心点坐标，两个字符之间的字符间距离为两个字符位置坐标之间的欧氏距

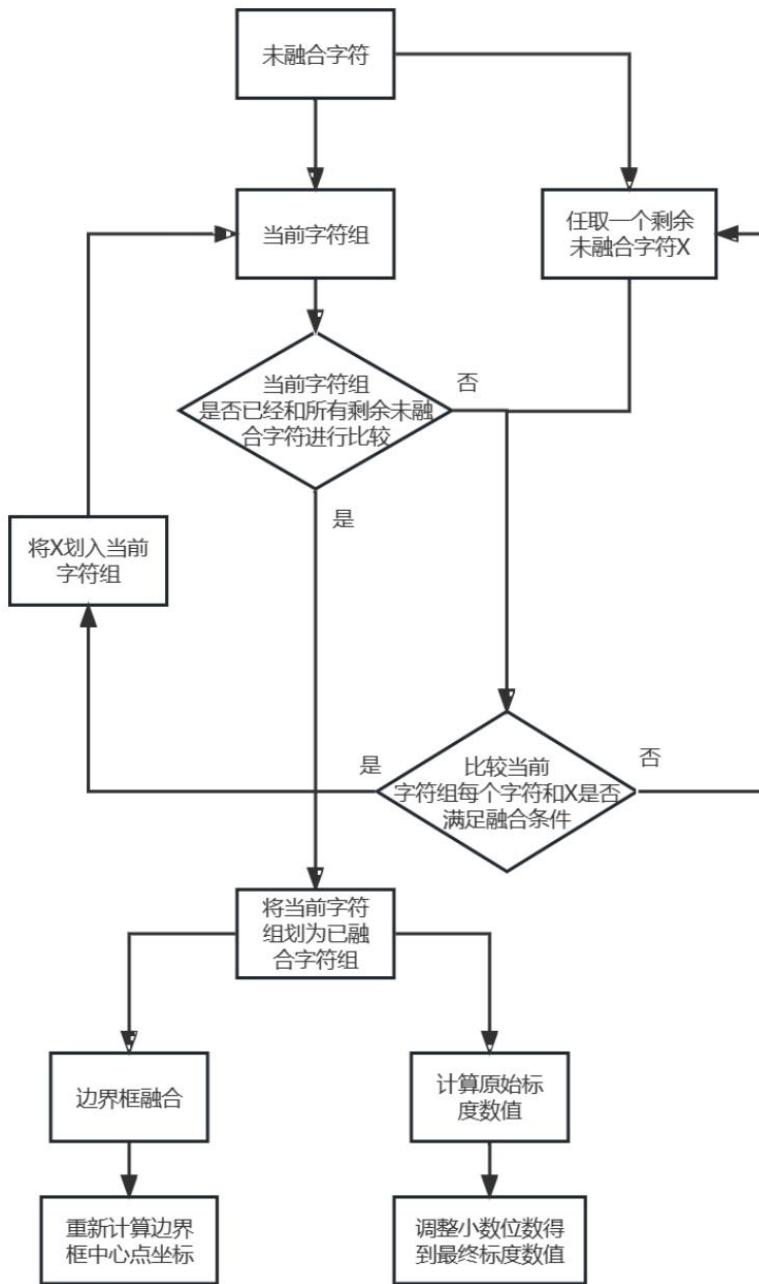


图 3-11 数字融合流程图

离；定义字符边界框距离为字符边界框水平方向距离之和的一半，如图 3-12 所示；

- 2) 任取一个未融合字符作为当前待融合字符组的初始字符，该待融合字符组在下文中简称为“当前字符组”；
- 3) 任取剩余未融合字符中一个字符和当前字符组中的字符进行字符边界框距离和字符间距离的比较。若字符间距离与字符边界框距离之差大于设定的阈值，则判定两个字符归属于同一个待融合字符组，判定标准如公式(3-6)所示；

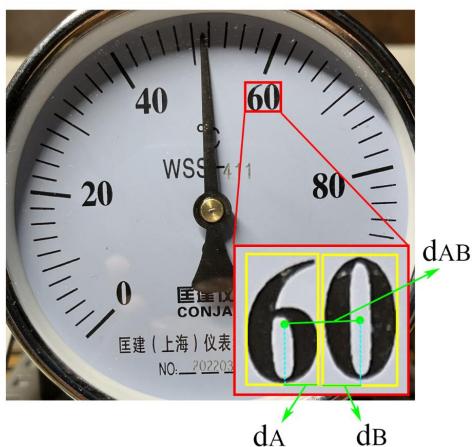


图 3-12 待融合字符组组合示意图

$$d_{AB} - 0.5(d_a + d_b) > \alpha \quad (3-6)$$

式中  $d_p$  表示字符间距离,  $d_A$  和  $d_B$  分别表示两个字符边界框的水平长度,  $\alpha$  表示设定的阈值, 本文经过实验比较后设定这个阈值取  $0.25(d_A + d_B)$ , 即字符边界框距离的一半;

如图 3-12 中右侧所示, 字符“6”和字符“0”之间的字符间距离为  $d_{AB}$ , 两个边界框水平长度分别为  $d_A$  和  $d_B$ , 字符边界框距离为。根据上述计算过程, “6”和“0”被划分为同一个待融合字符组, 而其他字符到“6”或“0”的字符间距离都远大于和这两字符之间的字符边界框距离, 因而不被划入该字符组中。

4) 对待融合字符组内部元素按照字符位置横坐标升序进行排列, 如图 3-12 所示, 因此字符“6”和字符“0”构成的待融合字符组内部的字符顺序必须是“6”在前, “0”在后, 而不能是相反的情况;

5) 重复步骤 3) 至 4) 所有的未被融合字符都已经和当前字符组进行过比较时, 该待融合字符组最终确定; 需注意, 当一个新的字符被划归为当前字符组, 则以后的再重复步骤 3) 至 4) 中后取的未融合字符都要和被新划入的字符进行上述判断过程, 以保证只要一个未融合字符和当前字符组中的任何一个字符距离足够近时, 都会被划入当前字符组;

6) 重复步骤 2) 至步骤 5), 直到所有的字符都已经完成待融合字符组的划分操作;

至此, 已经将表盘图像中的每个单个数字字符都被划分到了一个待融合字符组中, 用于下一步进行数字字符融合过程。

### 3.3.3 数字字符融合

数字字符融合过程具体指利用一个待融合字符组中若干字符生成一个新的标度数

字的过程，所谓生成新的标度数字具体指确定标度数字的语义和位置信息，语义信息指通过待融合字符组中的字符预测标签确定新的标度数字所代表的标度数值，位置信息指通过待融合字符组中的字符边界框信息确定新的标度数值的中心点坐标。该过程流程图如图 3-11 所示，具体包括以下步骤：

1) 边界框的融合。选取所有待融合字符组中各个字符最外部的坐标作为最终边界框的左上角和右下角的坐标，所谓最外侧坐标指左上角坐标中横坐标和纵坐标最小的，右下角坐标中横坐标和纵坐标最大的，对应使得融合后的边界框的面积尽可能大，本质上是求待融合字符组内部各个字符边界框的最小公共外接矩形。计算如公式(3-7)所示：

$$\begin{aligned}x_1 &= \min(x_{11}, x_{12}, x_{13} \dots, x_{1i} \dots, x_{1n}) \\y_1 &= \min(y_{11}, y_{12}, y_{13} \dots, y_{1i} \dots, y_{1n}) \\x_2 &= \max(x_{21}, x_{22}, x_{23} \dots, x_{2i} \dots, x_{2n}) \\y_2 &= \max(y_{21}, y_{22}, y_{23} \dots, y_{2i} \dots, y_{2n})\end{aligned}\quad (3-7)$$

式中  $(x_1, y_1)$  和  $(x_2, y_2)$  分别表示融合生成的新标度数字左上角和右下角的坐标， $(x_{1i}, y_{1i})$  和  $(x_{2i}, y_{2i})$  分别表示有  $n$  个字符的待融合字符组中第  $i$  个字符的左上角和右下角坐标；

如图 3-13 所示，子图 a) 中的各个单个字符经过融合后生成了子图 b) 中新的标度数字边界框和标度数值；

2) 根据融合后生成的边界框坐标计算标度数字的中心点坐标  $(x, y)$ ，也就是该标度数字的位置信息，计算公式如(3-8)所示：

$$\begin{aligned}x &= \frac{x_1 + x_2}{2} \\y &= \frac{y_1 + y_2}{2}\end{aligned}\quad (3-8)$$

3) 由于待融合字符组中的字符排列是按照空间上从左到右的顺序排列的，因此可以利用公式(3-9)计算标度数字的原始标度数值：

$$L = \sum_{i=1}^n (l_i \times 10^{n-i+1}) \quad (3-9)$$

式中  $L$  表示待融合字符组的原始融合后标签，也就是对应标度数字表示的原始标度数值， $l_i$  表示待融合字符组中第  $i$  个字符的标签即该字符的字符数值。

所谓原始标度数值指在忽略小数点的情况下只对数字字符读数得到的数值。如图 3-13 所示，待融合字符组的排列顺序为“1”，“0”，“0”，带入公式(3-9)得到读数结果如公式(3-10)所示为 100，该读数结果和人眼观察一致。

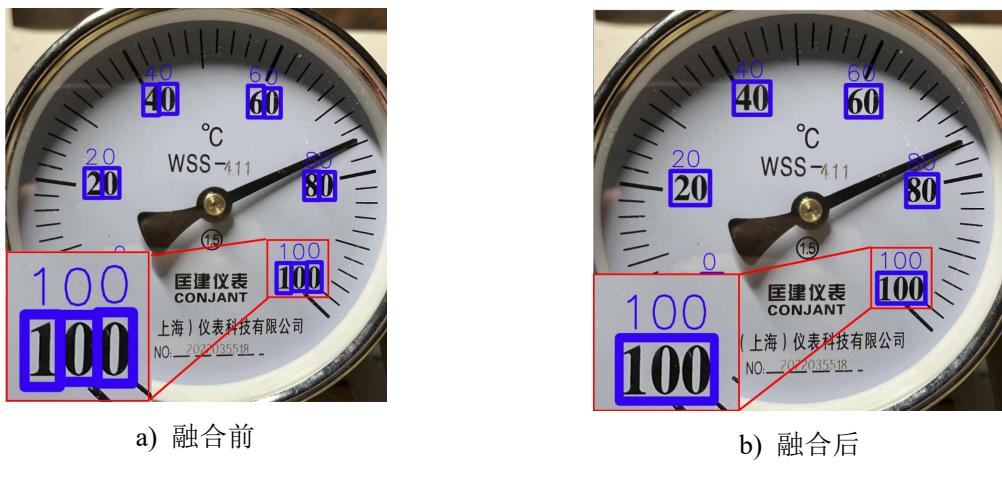


图 3-13 数字融合示意图

$$L = 1 \times 10^{3-1} + 0 \times 10^{3-2} + 0 \times 10^{3-3} = 100 \quad (3-10)$$

### 3.3.4 小数标度数字的处理机制

在实际的工程场景中存在一类仪表的标度数字为小数，如图 3-14 所示。如果直接使用上述待融合字符组的原始标度数值；由于目标检测难以直接识别小数点这样的小目标，同时小数点也非常容易和表盘上的污渍等元素混淆。因此本文放弃使用目标检测模型直接识别小数点的思路，而是依据待融合字符组内部排序特点设计了小数标度数字判断方法，并基于此方法设计了含小数标度数字仪表的标度信息提取算法。

依据表盘图像的特点，本文提出当一个待融合字符组的字符排列中第一个字符为“0”且该待融合字符组的字符个数大于 1，即 0 出现在了最大位数而该字符的原始标度数值不为 0 时，即可判断该待融合字符组对应标度数字的数值为小数。若一个表盘图像中出现小数标度数字，则该表盘图像就判定为含有小数标度数字的表盘。根据表盘图像是否为小数标度类型的判断结果，确定标度数字的标度数值识别结果，具体计算过程如下：

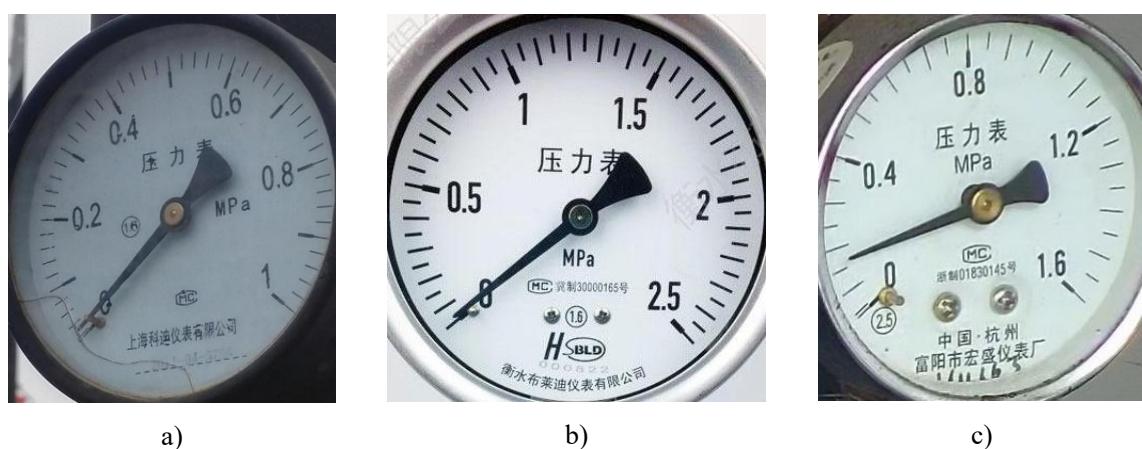


图 3-14 小数仪表示意图

1) 根据待融合字符组的中字符个数计算潜在小数点后位数, 潜在小数点后位数  $\beta$  计算公式如(3-11)所示:

$$\beta = \text{len}(nums) - 1 \quad (3-11)$$

式中  $\text{len}(N)$  表示统计数组 N 的长度,  $nums$  表示一个待融合字符数组;

2) 若待融合字符组的第一个元素为 0 且该待融合字符组的字符个数大于 1, 即  $\beta > 0$ , 则判定该待融合字符组表示一个小数标度数字, 若表盘中存在小数标度数字, 则判定该表盘为含有小数标度数字的表盘;

3) 若表盘未被判定为含有小数标度数字的表盘, 则原始标度数值就是该标度数字的最终识别标度数值。若表盘被判定为含有小数标度数字的表盘, 则需要对表盘上所有的标度数字的标度数值进行调整, 经过调整后的标度数值为最终的识别标度数值  $L_{new}$ , 具体调整公式如下:

$$L_{new} = L \times 10^{-\beta} \quad (3-12)$$

如图 3-15 所示, 子图 b) 所示待融合字符组 “0, 4” 满足步骤 2) 中 “第一个元素为 0 且该待融合字符组的字符个数为 2 大于 1” 的情况, 因此这是一个小数标度数字。根据步骤 1) 该字符的潜在小数点后位数  $\beta$  为待融合字符组中字符个数减一, 即。因此根据步骤 3) 所述公式, 将原始标度数值  $L = 4$ , 带入公式(3-12)对标度数值进行调整:

$$L_{new} = 4 \times 10^{-1} = 0.4 \quad (3-13)$$

因为存在小数标度数字 “0.4”, 所以整个仪表被判定为小数标度类型仪表, 因此所

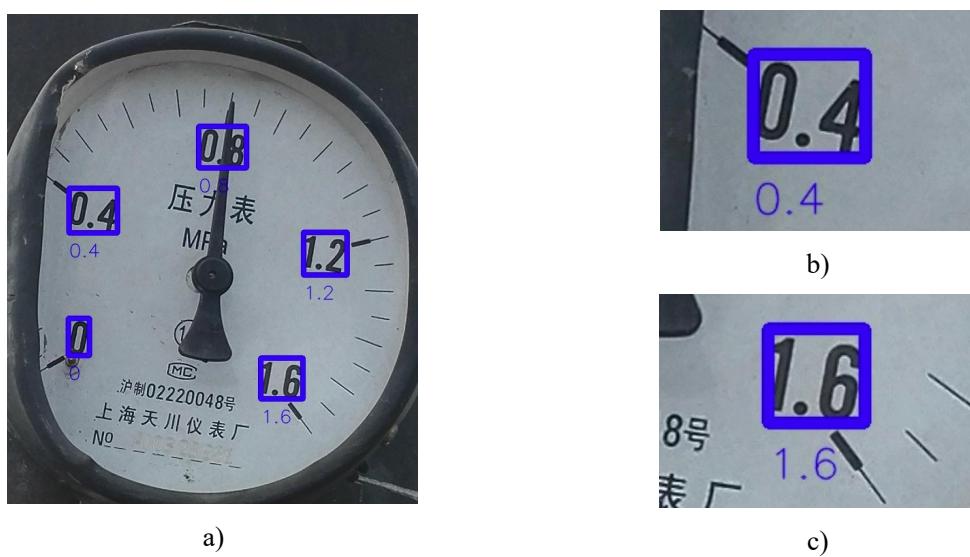


图 3-15 小数仪表数组融合效果图

有字符都要进行步骤 6) 的调整, 如果 8 中 c) 所示, 待融合字符组“1、6”原本并不满足小数标度数字的条件, 但是因为仪表被判定为小数标度类型仪表, 所以也进行了读数结果的调整, 最终的标度数值结果为 1.6, 和人眼的读数一致。

### 3.4 本章小结

本章针对表盘图像标度信息提取问题做了深入研究。基于表盘图像特征和读数识别计算任务的需求, 本文分别提出了基于 OCR 模型和基于目标检测模型的两种方法。基于 OCR 模型的方法主要研究了 OCR 文本检测模型的微调训练以及 OCR 模型识别结果分析, 提出了一种从所有 OCR 字符检测结果中获取标度信息的字符筛选方法。基于目标检测的方法重点研究了如何从目标检测模型的得到的单个数字字符检测结果中获取标度信息, 提出了跨类别 NMS、待融合字符聚类和数字字符融合等方法将表盘图像图像中的单个字符融合生成标度数字进而得到标度信息。此外, 还利用了 OCR 模型能够识别表盘中所有字符的优势, 提出了一种从 OCR 模型识别结果中获取仪表类型信息的方法。实验结果表明, 本章提出的两种方法都达到了直接从表盘图像提取标度信息的目的, 具体实验数据及分析将在第六章中集中进行讨论。

## 第四章 指针信息提取

本章研究的是指针信息提取问题，基于读数识别算法的需求，本文选择指针尖端点和指针中心点作为指针信息提取的求解目标。指针信息提取是同方向文献中被广泛研究和改进的环节，然而多数方法仍然依赖传统图像处理方法，导致算法鲁棒性和抗干扰能力较弱，制约了面向未知仪表的泛化能力。针对上述问题，本章从传统图像处理和关键点检测两种不同的技术视角切入，提出了针对 Hough 直线检测的指针信息提取方法改进，以及使用关键点检测直接预测指针信息关键点的方法。

### 4.1 指针区域提取

早期的仪表自动读数算法都是针对表盘图像直接进行二值化以提取指针所在直线的，然而表盘图像中包含的各种字符指针等干扰信息过于复杂，因此对读数计算不利。表盘直接二值化的示意图如图 4-1 所示，从图中可以看出大量的指针之外的区域存在干扰，这种干扰很难通过固定的参数设定进行去除，且表盘图像中也有可能出现其他的直线目标，这对指针检测会更加不利。

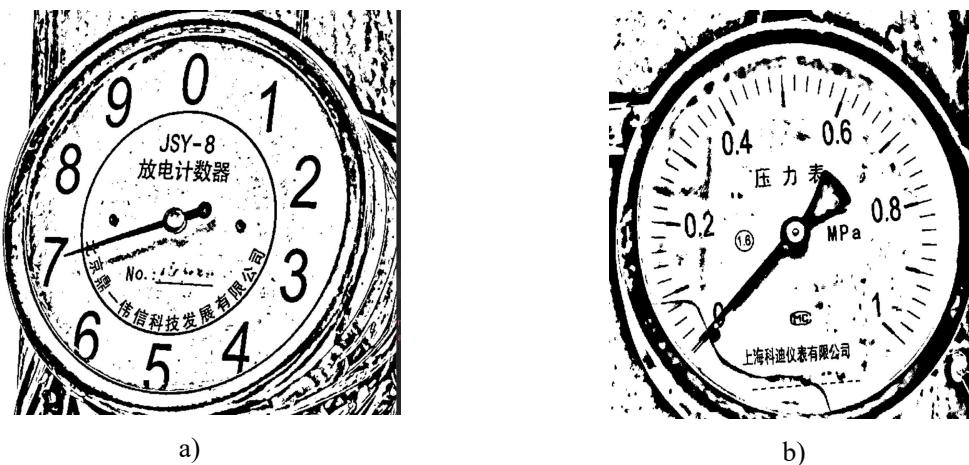


图 4-1 表盘二值化效果图

因此，本文受 Fang 等研究者<sup>[7]</sup>的工作启发，也选择先用目标检测模型从表盘中检测出相应的指针区域然后在指针区域的基础上继续进行指针信息的提取。本文直接应用在 2.2 节中提出的 Lite-FCOS 模型作为指针检测模型。在标度信息提取使用基于目标检测的方法的情况，只需要在 DRID-DET 数据集中新增一个指针的标注类别，将模型预测的种类扩充到 11 类即可用训练得到的模型同步提取指针区域和数字字符，两者不会发生冲突。对于标度信息提取使用基于 OCR 模型方法的情况，需对 DRID-OCR 数据集中的图像重新进行标注，标注一个单独的指针目标类，得到的数据集称为 DRID-PT，其训练集和验证集的划分比例和 DRID-OCR 数据集保持一致。

经过训练的 Lite-FCOS 检测表盘指针效果如图 4-2 所示，图中右侧得到的就是从表盘图像中分割下来的指针图像，指针图像即为后续指针信息提取的分析基础。

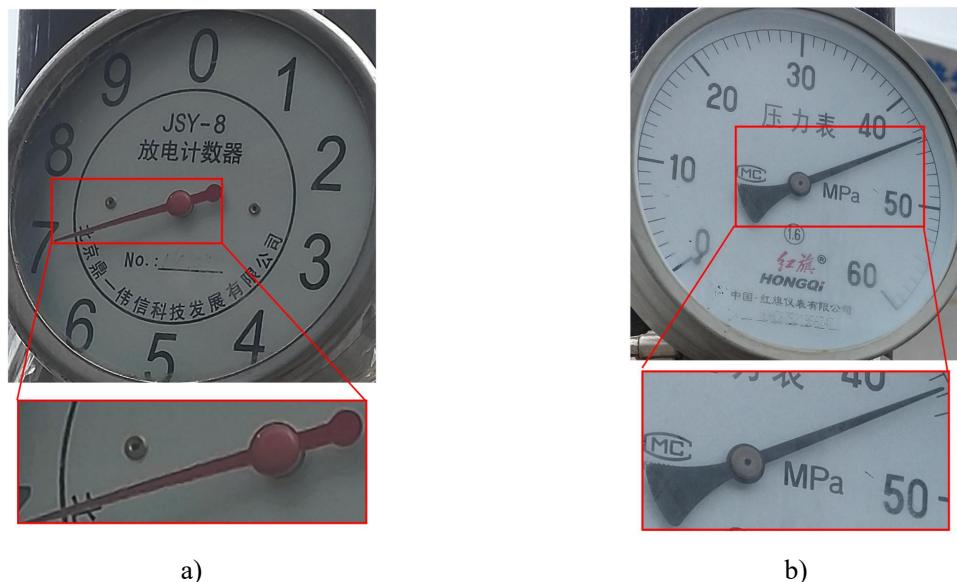


图 4-2 指针图像提取效果图

## 4.2 基于 Hough 直线检测提取指针信息

### 4.2.1 拟合指针中心点

本文中指针中心点被定义为指针旋转轴所在点，是指针旋转过程的转动中心，而不是指针的几何中心。根据一般仪表特征，各个标度数字和标度盘刻度都分布在一个以指针中心点为圆心的圆上。基于表盘图像的上述特征，本文提出了一种基于仪表标度信息的指针中心点拟合方法。

由几何知识可知，平面内不共线的三点可以确定唯一的圆，因此每不共线的三个标度数字都可以用其位置坐标得到一个唯一的拟合圆。在理想情况下，表盘图像中的任意三个标度数字得到的拟合圆都应该是相同的且拟合圆的圆心就是指针中心点。然而在实际情况下，由于标度数字位置的绘制和检测都会存在偏差以及表盘本身可能倾斜等原因，这些拟合圆不可能完全相同，但是对于一个正常的仪表，这些拟合圆的圆心总应该分布在指针中心点的附近。如图 4-3 所示黄色圆就是一个利用标度数字“1”、“3”和“5”得到的拟合圆，其中  $P_1P_3P_5$  分别表示对应标度数字的位置所在点， $P_c$  表示拟合圆的圆心。

依据上述分析，本文通过对标度数字分组进行圆拟合，并求这些拟合圆的圆心坐标，去除其中质量不佳的拟合圆心坐标之后求剩余拟合圆心坐标的中心位置，并将该位置作为指针中心点。该方法具体流程如下：



图 4-3 拟合圆示意图

1) 每三个标度数字计算一个拟合点, 得到一个拟合点集, 首先定义拟合圆的一般方程如公式(4-1)所示:

$$\begin{aligned} (x^2 + y^2) + Bx + Cy + D = 0 \\ B^2 + C^2 - 4 \times D^2 > 0 \end{aligned} \quad (4-1)$$

式中圆上任意一点坐标为  $(x, y)$ ,  $B, C, D$  分别为方程的参数矩阵,

在三个点不共线的情况下该方程有关于唯一的解

$$\begin{bmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (4-2)$$

根据圆的一般方程可以推导得到圆心坐标  $P_c(x_c, y_c)$  计算公式, 带入上述计算得到的系数  $B, C, D$ , 可以得到圆心坐标也就是拟合点坐标如公式(4-3)所示:

$$\begin{aligned} x_c &= -\frac{B}{2} \\ y_c &= -\frac{C}{2} \end{aligned} \quad (4-3)$$

2) 筛选并去除拟合点中不可靠的离群点。由于标度数字检测过程和表盘拍摄视角可能造成的误差, 可能存在距离指针中心点较远, 对于指针中心点估计容易带来误导作用的不可靠拟合点, 这类拟合点通常距离其他正常的拟合点较远, 因此本文称其为离群拟合点。本文参考 k-最近邻法的思想定义反应一个拟合点和拟合点集离散程度的距离指标, 称为离群指数, 记为  $\eta$ , 具体计算过程如下:

2.1) 计算拟合点到拟合点集中所有其他拟合点的距离集合;

$$D_i = \{d_{i1}, d_{i2}, \dots, d_{ij}, \dots, d_{in}\} \quad (4-4)$$

式中  $D_i$  表示拟合点集中第  $i$  个拟合点的距离集合,  $d_{ij}$  表示拟合点即中第  $i$  个拟合点和第  $j$  个拟合点之间的欧氏距离。

2.2) 定义拟合点的离群指数为所有距离中最近的前  $1/4$  个距离的平均数, 具体计算流程为首先对  $D_i$  进行升序排序并统计  $D_i$  中的元素个数将其记为  $n$ , 然后取排序后的集合  $D_i$  中前  $n/4$  个元素并求其平均值, 得到的结果就是离群指数  $\eta$ 。

对每个拟合点都进行独立判断, 计算其对应的离群指数并和预先设定的离群指数阈值进行比较判断, 若拟合点的离群指数大于设定的阈值, 则将其判定为离群拟合点, 则将其从拟合点集中剔除, 不参与指针中心点的估计过程。



图 4-4 拟合指针中心点效果图

3) 计算筛选后拟合点集的中心点坐标, 即对所有的拟合点横纵坐标分别求平均值, 得到的结果就是指针旋转中心点坐标。如图 4-4 所示, 图像中外围的绿色圆点表示被剔除的离群点, 在指针旋转中心处密集的青色圆点为保留的拟合点; 粉色圆点表示了最终得到的指针旋转中心点, 黄色圆点为不进行离群点剔除情况下得到的指针中心点, 可以看出剔除离群点后的拟合效果明显得到提高。本例中设定离群指数阈值  $\sigma$  由公式(4-5)得到, 若调高这一阈值则更多的点会被判定为保留的拟合点, 若调低这一阈值, 则更多的拟合点会被判定为离群拟合点。

$$\sigma=0.08 \times \frac{(w+h)}{2} \quad (4-5)$$

式中,  $w$  表示表盘图像的宽度,  $h$  表示表盘图像的高度。

#### 4.2.2 拟合指针中心线

利用指针图像进一步求取指针尖端点信息。顾名思义, 指针尖端点表示指针尖端所

在的位置，该点指针中心点共同确定了指针偏转角度。本文通过指针中心线和指针图像边沿的交点确定指针中心点，所以首先需要拟合得到指针中心线的相关信息。

本文使用 Hough 变换这一最常用给直线检测器作为指针中心线检测的基本算法。Hough 变换是一种空间变换，笛卡尔坐标系中的一个直线可以经 Hough 变换得到 Hough 空间中的一个点。例如， $A(x_1, y_1)$  和  $B(x_2, y_2)$  是空间中两个点，其连线的直线方程为  $y = kx + q$ ，若将 AB 两点带入直线方程中如公式(4-6)所示，则直线方程仅有唯一的解  $k$  和  $q$ ，可将其视为  $k$  和  $q$  所构成的另一个笛卡尔空间中的一个点  $(k_1, q_1)$ ，而这个  $k$  和  $q$  所构成的笛卡尔空间就是  $x$  和  $y$  对应的 Hough 空间。如图 4-5 所示，与之相反，Hough 空间中的一条直线也对应笛卡尔空间中的一个点。由此可知，笛卡尔空间中的两个点对应了 Hough 空间中的两条直线，在 Hough 空间中会形成一个交点，进一步可知，在笛卡尔空间中共线的点必然在 Hough 空间中相交于一个公共点。

$$\begin{aligned} q &= -kx_1 + y_1 \\ q &= -kx_2 + y_2 \end{aligned} \quad (4-6)$$

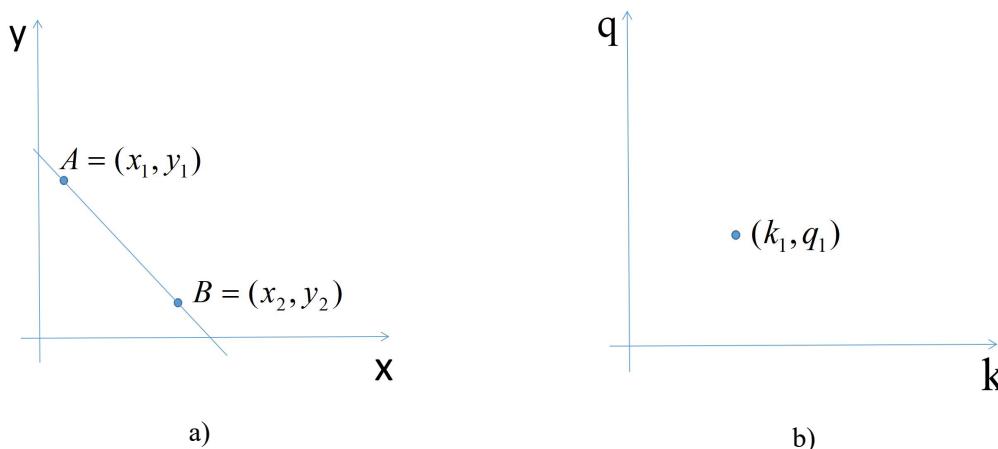


图 4-5 Hough 变换示意图

由此可以得到 Hough 直线检测的原理：首先将图像变为二值图像，然后将二值图像中的白色部分视为笛卡尔坐标系中的点，将这些点全部转化为 Hough 坐标系中的直线。此时 Hough 空间中会有大量的直线并形成大量的交点，统计各个交点经过的直线数目，假设一个交点有  $a$  条直线经过，则该交点就对应了笛卡尔坐标系中共线的  $a$  个点。根据设定的 Hough 变换直线检测阈值  $k$  进行判断，保留 Hough 空间中交点经过直线数目大于该阈值的图像作为检测结果，这其中每个交点都对应了笛卡尔坐标系中的一条直线，这些直线就是 Hough 变换直线检测的检测结果。Hough 直线检测的直观解释就是根据笛卡尔空间中的点找到了所有满足有  $k$  个点以上共线的直线。

由于传统的图像二值化和 Hough 直线检测算法均基于固定参数进行, 而算法处理效果又对参数的设定非常敏感, 因而很难找到一组参数能够适应于复杂多变的场景, 这是制约指针中心线检测泛化能力进一步提高的核心问题。如图 4-6 所示, 子图 a) 和 c) 分别是子图 b) 和 d) 的对应原图, 使用了相同的阈值参数进行二值化处理, a) 的二值化效果良好, 但是 c) 的二值化效果则非常差, 无法用于后续的指针中心线拟合。

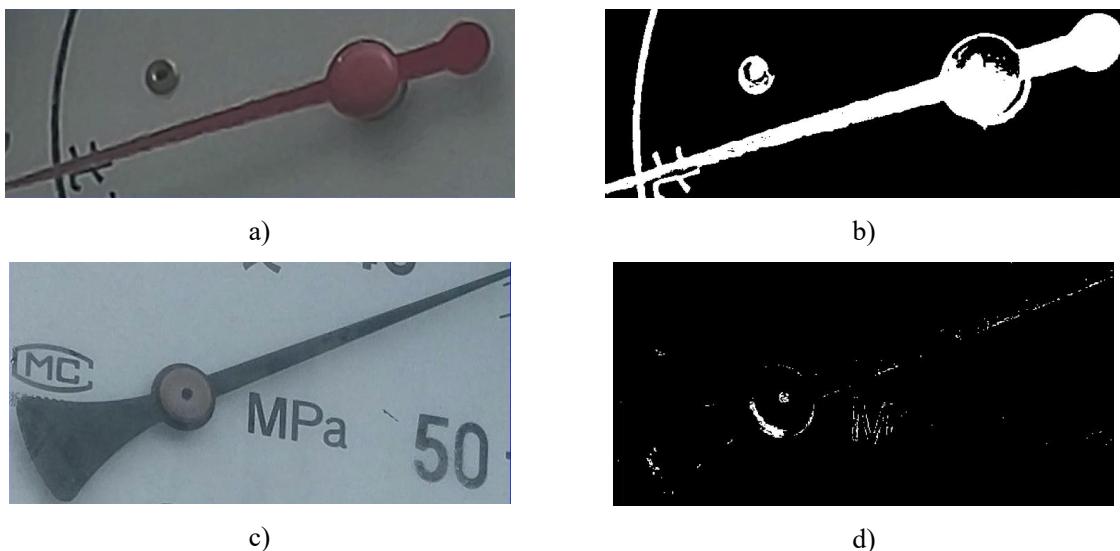


图 4-6 基于固定阈值的二值化效果图

为使指针中心线检测能够更加更好的泛化能力, 本文对基于 Hough 直线检测的指针中心线拟合算法进行改进以使其能够自动适应不同条件下的指针图像。改进后的指针中心线检测算法主要包括自适应二值化、自适应 Hough 变换直线检测两个部分, 结合图 4-7 所示, 该过程具体计算流程如下:

- 1) 根据指针图像的长宽比分段设定不同二值化终止阈值, 记为  $t_{end}$ 。由于部分长宽比较高的指针区域中指针所占比例更高, 如指针接近水平或者竖直状态的时候, 因此本文分段设定阈值可以增强算法的鲁棒性, 达到一定的自适应效果, 以指针图像的长宽比作为分类依据, 对长宽比大于 4 的指针图像设定  $t_{end} = 30$ , 对长宽比小于 4 的指针区域设定为  $t_{end} = 90$ ,

- 2 初始化用于二值化的分割阈值  $t$ , 本文经过实验比较不同阈值后设定  $t = 10$ , 初始化一个二值化的分割阈值更新步长  $\Delta t$ , 本文设定为  $\Delta t = 10$ ;

- 3) 对指针图像进行自适应二值化。对指针图像进行基于分割阈值  $t$  的二值分割, 并求指针区域二值图像的平均图像灰度  $g$ ; 比较平均灰度  $g$  和二值化终止阈值  $t_{end}$  的大小, 若  $t_{end} > g$  则用分割阈值的步长  $\Delta t$  更新二值化阈值  $t_{end}' = t_{end} + \Delta t$ , 并用新的二值化阈值

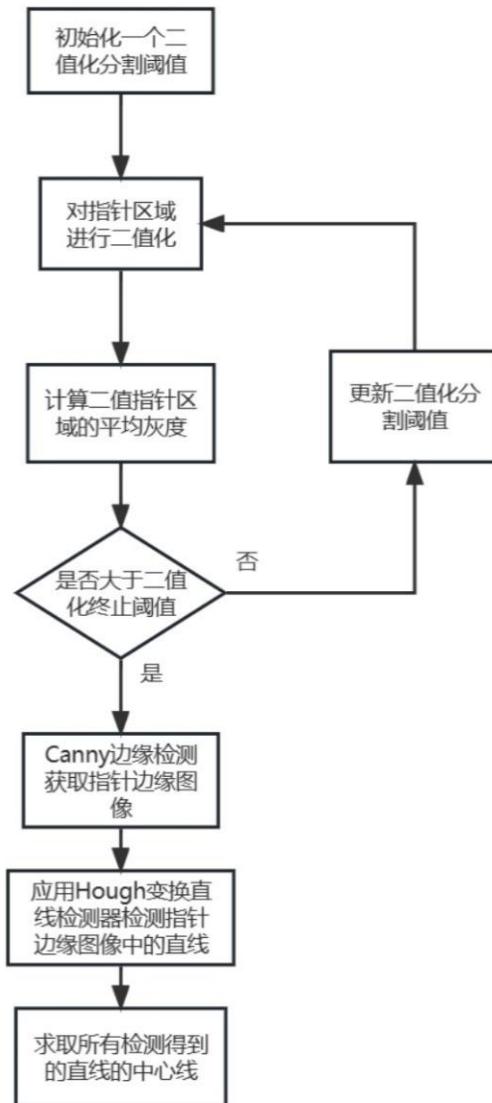


图 4-7 自适应 Hough 直线检测流程图

$t_{end}$ '再次计算，直到 $t_{end} < g$ 时结束二值化的过程；

如图 4-8 所示为基于目标检测模型从本实施例表盘图像中提取出来的指针图像，应用上述步骤 6.2) 至步骤 6.4) 中方法迭代了 6 轮后，最终分割阈值  $t$  被设定为 90，得到指针区域二值图像的平均灰度约为 42.86，得到的指针区域二值图像如图 4-8 中子图 (d) 所示，子图(b)和(c)展示了迭代过程中的指针二值化效果。

4) 对二值化之后的指针区域进行运用 Canny 算法进行边缘检测，得到二值化的指针边缘图像，如图 4-9 所示。由于 Hough 变换是统计所有可能共线的白色像素，并不要求这些像素点严格相邻，因此多个不连续的实心区域可能会被检测成一条直线，在 Hough 直线测之前预先进行边缘检测可以降低检测过程的计算量，同时也减少误检测的可能性。

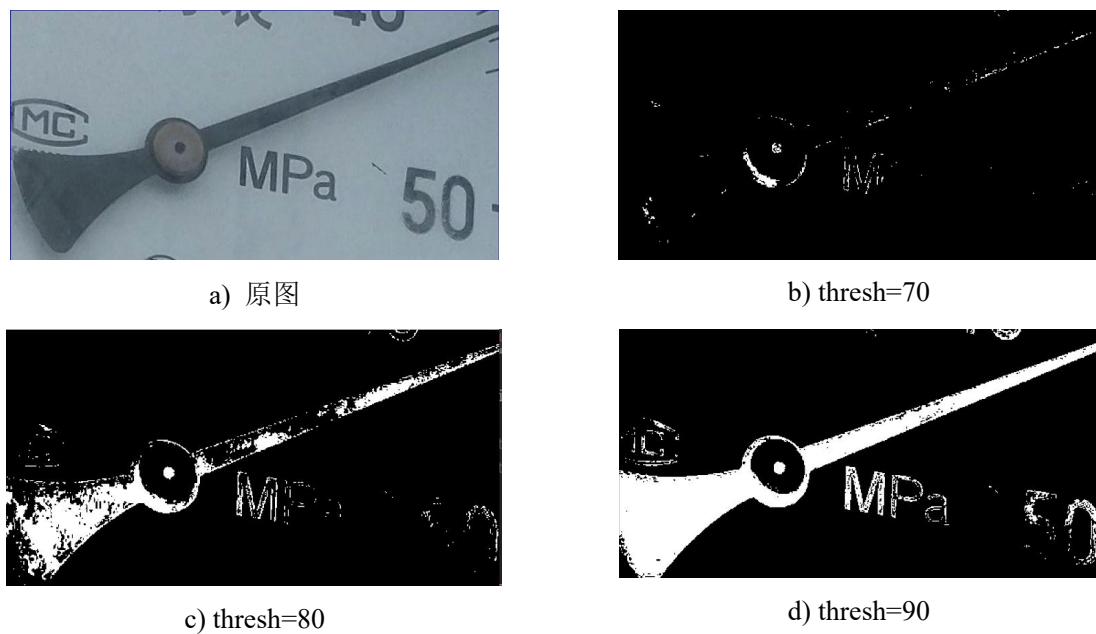


图 4-8 自适应阈值二值化效果图

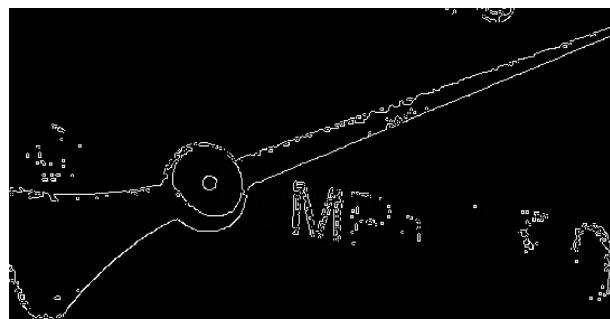


图 4-9 指针边缘图像

5) 运用 Hough 变换直线检测的方法检测得到的指针边缘图像, 为适应不同大小尺寸的输入表盘图像, 按尺寸设定不同 Hough 直线检测阈值, 对于尺寸较大的图像设定检测的阈值适当大一点, 反之则适当小一点, 经过实验测试, 本文选择分两段设定 Hough 直线检测的阈值。

6) 经直线检测后得到若干条直线方程的检出结果, 求出这些直线的中心线方程, 就是指针中心线, 效果如图 4-10 所示, 其中绿色直线表示检测得到的各个直线, 红色直线表示拟合得到的指针中心线。

#### 4.2.3 拟合指针尖端点

上一节提取指针中心线的目的在于进一步计算指针尖端点, 指针尖端点和指针中心点的连线表示了指针所指示的方向。在指针图像中, 指针尖端点通常位于指针图像的边缘位置附近, 且在表盘图像中, 指针尖端点是指针上距指针中心点最远的位置。基于上

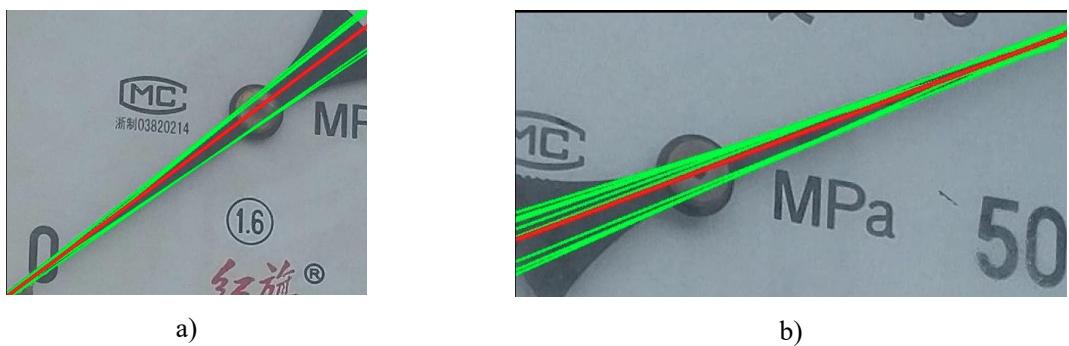


图 4-10 指针中心线拟合示意图

述表盘图像特点，本文提出一种指针尖端点计算方法，首先使用指针中心线和指针图像边线求指针中心线的两个端点位置，然后比较两端点和指针中心点之间的距离，保留其中距离较远的端点作为最终得到的指针尖端点，具体计算过程如下：

- 1) 分别计算指针中心线和指针区域 4 条遍所在直线的交点坐标；
- 2) 判断交点坐标是否在指针区域的内部，筛选出指针中心线和指针区域边沿的两个交点作为指针端点；如图 4-10 所示，两个交点坐标分别为；
- 3) 根据指针区域左上角的横纵坐标计算指针图像坐标系下的指针端点在表盘坐标系下对应的坐标如公式(4-7)所示，其中  $(x_{tl}, y_{tl})$  为指针区域左上角坐标， $(x_p, y_p)$  为指针区域坐标系下的一个指针端点坐标， $(x_m, y_m)$  为对应的表盘区域坐标系下的指针尖端点坐标。

$$\begin{aligned} x_m &= x_p + x_{tl} \\ y_m &= y_p + y_{tl} \end{aligned} \quad (4-7)$$

- 4) 分别计算表盘图像下指针中心线两端点和指针中心点之间的距离，保留距离较远的点作为最终得到的指针尖端点坐标。如图 4-11 所示， $P_{e1}$  点的到指针中心点的距离大于  $P_{e2}$  点到指针中心点的距离，所以点  $P_{e1}$  为最终求得的指针尖端点。

### 4.3 关键点检测提取指针信息

由于本文中指针信息由指针尖端点和指针旋转中心点构成，本质是表盘图像中的两个关键点的获取。由于指针的旋转中心点和指针尖端点都具有比较明显的图像特征，由此，本文尝试利用计算机视觉领域中的关键点检测技术，通过训练指针关键点检测模型直接从表盘图像中预测两关键点坐标，从而得到指针信息。经过实验测试，该方案具有可行性，能够满足任务的需求。

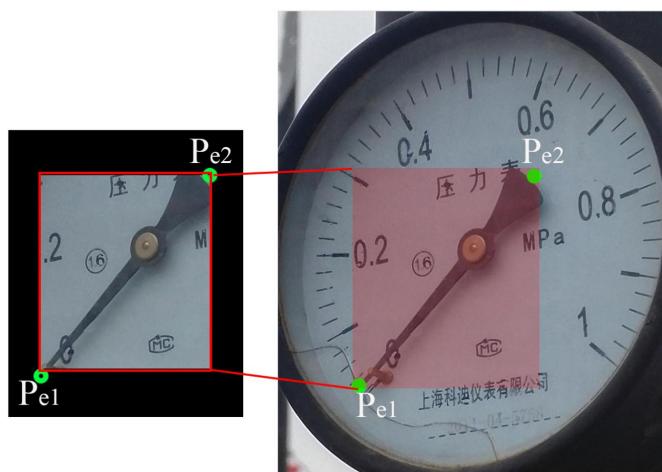


图 4-11 指针尖端点计算示意图

#### 4.3.1 关键点检测技术简介

关键点检测技术是计算机视觉领域重要的基础任务之一，是众多更高次计算视觉任务的基础，如行为识别，人机交互等。该技术源于人体姿态估计任务，主要目标是从输入图像中求解人体骨骼关键点信息，然后基于该相应的关键点信息可以进一步对人的姿态和行为进行分类和研究，随着该技术的不断发展，现在也广泛应用于人脸识别，手势识别和服饰识别等更多领域。

早期关键点检测主要是利用专家系统实现的，通过肢体之间相互作用的知识体系人工设计图片的特征表达，并基于上述特征结合经典机器学习方法实现检测功能。然而人体的姿态具有多变性，因此设计足够有效的专家系统和特征提取办法是十分困难的，随着检测器使用的规则越来越复杂效果也难以得到根本性的突破。

随着深度学习技术的兴起和发展，卷积神经网络在图像识别、目标检测等领域大放异彩，展现出了惊人的特征提取能力，经过训练的 CNN 的特征提取效果远好于人类专家手工设计出的特征。于是研究者们转向关注基于深度学习方法实现关键点检测。Toshev 等<sup>[72]</sup>在 2014 年提出了 DeepPose，首先将深度学习的方法引入了关键点检测的领域，开启了这一领域的新时代。

基于深度学习的关键点检测网络的人工标注 ground truth 大致可以分为两类，基于坐标回归、基于热力图 heatmap。基于坐标回归是指直接从输入的待处理图像中回归预测关键点的位置坐标，利用 CNN 提取图像特征然后利用全连接层输出每个关键点的坐标，每个关键点的计算是相对独立的。这种方法比较直观，但是效果相对不佳，因为长距离回归问题对于深度神经网络是较为困难的，因此这种方法缺乏足够的空间泛化能力，

且全连接层得到的权重分布高度依赖训练数据的分布，所以非常容易造成过拟合。基于 heatmap 的方法指通过深度神经网络输出模型各个关键点的预测概率图，也是 heatmap 图像，然后基于该图像确定最终的关键点。基于 heatmap 的方法首先由 Tompson 等<sup>[73]</sup>在 2014 提出，后来得到的这一领域的广泛应用。这一方法相对精度更高，不过确定是坐标点的输出不是一个全微分的模型，而是通过 argmax 的方式得到的，而这一过程不可求导也就无法使用深度学习的方法进行学习。不过实验证明，基于 heatmap 的方法效果上总体优于坐标回归的方法。在此基础上，另一些研究者如 George 等<sup>[74]</sup>提出了综合了 heatmap 和坐标回归的方法构建 ground truth，也取得了比较好的效果。

基于深度学习的关键点检测从流程上分类大致可以分为自下而上和自上而下两大类。自下而上的关键点检测流程为首先对图像中的所有关键点进行提取，然后再对关键点进行聚类以确定哪些关键点对应同一个具体的目标，典型方法包括 openpose<sup>[75]</sup>和 deepcut<sup>[76]</sup>等。自上而下的关键点检测流程为首先利用目标检测模型对图像中的目标进行检测，然后在各个目标的边界框内部分别进行关键点检测，以得到各个目标的关键点骨架信息，典型方法包括：HourglassNet<sup>[41]</sup>和 CPN<sup>[77]</sup>等，这种先提取图像中的感兴趣区域然后再对区域内部的图像上完成检测识别任务的做法和 OCR 模型中分为文本检测和文本识别两个环节也有着高度的类似。

由于本文已经拥有现成的指针目标检测模型可以用于指针区域，因此可以认为本文至此已经完成自上而下的关键点检测技术的前半部分，因此，本文选择了此类方法来实现具体的指针关键点检测任务。

### 4.3.2 HRNet 关键点检测网络

HRNet 是一种优秀的关键点检测模型，由 Sun 等<sup>[55]</sup>在 2019 年提出，已经成为了多种姿态估计框架中的主要关键点检测模型。该模型也属于自上而下类型，即首先通过目标检测模型提取指针部分，然后通过 HRNet 模型对得到的目标区域图像提取关键点信息。本文中目标检测模型使用的是 2.2 节中的 Lite-FCOS 模型，这里不再赘述，本节重点关注 HRNet 模型的原理。

HRNet 模型的整体网络结构沿用同类的自上而下方法的基本结构，都包括主干特征提取网络和关键点回归网络。主干特征网络用于提取图像的特征信息，然后关键点回归网络将特征层映射为各个关键点的 heatmap，最后从 heatmap 中计算出局部极大值点作为最终的关键点识别结果。

HRNet 模型的主要改进点在于主干网络，作者通过改进主干网络提高了模型的特征提取能力，保留了高分辨率下的图像特征，从而实现了更好的关键点检测效果。在传统的 CNN 中，随着网络深度的不断加深，特征层的通道数在增加的同时，其尺度也变得越来越小，这会导致特征层能够反映更好的语义信息，但是却会丢失一些空间信息。关键点检测问题需要从整幅图像中精确的预测一个具体的关键点，因此对高分辨的空间信息要求较高。之前的解决方法主要是通过反卷积等方法将小尺寸的特征层进行尺寸的放大，以调整到正常尺寸中，如 Hourglass 模型的做法。然而 HRNet 通过并行多个分辨率的分支的同时进行不同分支之间的信息交互，在保持了足够强的语义信息提取能力的前提下也全程保留了高分辨的特征层，从而保留了高分辨率的信息，达到的提取更好特征的目的。HRNet 模型的网路结构图如图 4-12 所示，其全程保持了不同分辨率下的特征层，并且每隔一层卷积层就进行一次特征层之间的相互融合。

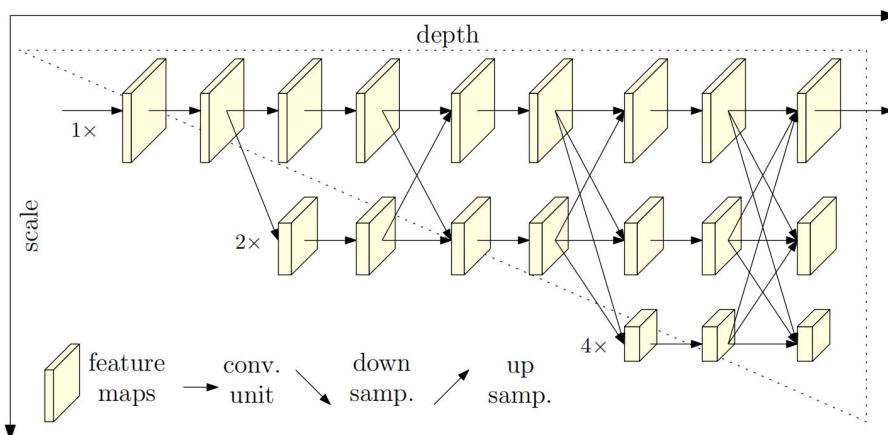


图 4-12 HRNet 主干网络模块示意图<sup>[55]</sup>

### 4.3.3 关键点检测提取指针信息

对 3.1.3 节中的 DRID-OCR 中的表盘图像进行 COCO-KeyPoint 关键点检测格式的重新标注，标注目标为指针尖端点和指针中心点，得到的新数据集命名为 DRID-KP，其训练集和验证集的划分比例和 DRID-OCR 中保持一致。利用 DRID-KP 数据集对 HRNet 关键点检测模型进行重新微调训练，得到 HRNet 指针关键点检测模型，运动该模型对表盘图像进行指针关键点检测效果如图 4-13 所示：

图中红色圆点表示检测得到的指针尖端点，绿色圆点表示检测得到的指针中心点，用蓝色线段将两个检测得到的关键点连接起来，该蓝色直线反映了指针的中心线，从效果图中可以看出关键点检测模型的检测结果比较可靠，能够准确找到反映指针信息的两个关键点。

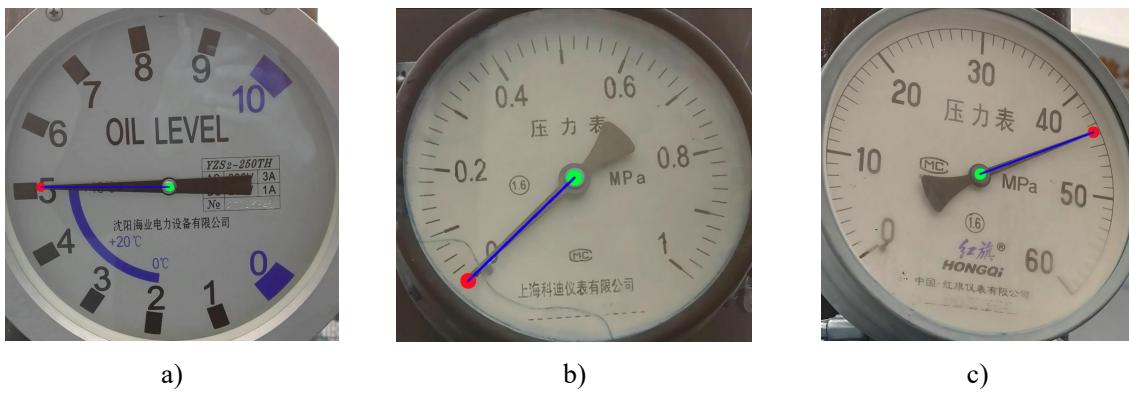


图 4-13 指针关键点检测效果图

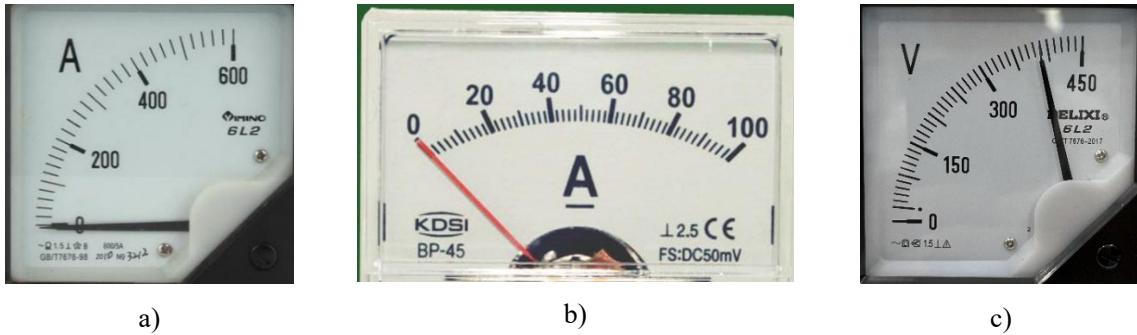


图 4-14 方形仪表示意图

由于方形仪表的指针中心点常常是无法直接观察到的，如图 4-14 所示。因而也就无法使用关键点检测网络直接预测，对于方形仪表，本文仍然采用 4.2.1 节中的基于标度信息拟合的方法来确定指针中心点，仅使用关键点检测网络提取其指针尖端点。

#### 4.4 基于边界框的指针尖端点近似

尽管本文提出的指针尖端点检测算法已经相比同类算法在鲁棒性和泛化能力方面有了很大进步，然而仍然难以避免存在算法无法得到指针尖端点的情况。为进一步提升仪表读数算法的泛化能力，本文依据指针目标在表盘图像中的特点提出一种在无法得到指针尖端点情况下的应对策略。

由于指针目标是一个长窄矩形形状的目标，因此，通常目标检测得到的指针图像中，指针都位于对角线的位置上，因此，指针中心线和指针边界框的交点通常都位于指针图像四个角附近，也就是指针边界框的四个矩形顶点位置附近。又因为指针尖端点是指针上距离指针中心点最远的点，所以指针尖端点通常位于距离指针中心点最远的指针区域边界框矩形顶点的附近。

如图 4-15 所示，蓝色边界框表示目标检测模型输出的指针目标边界框，红色直线表示穿过边界框距离指针中心点最远顶点的对角线，绿色圆点表示距离指针中心点最远

的



图 4-15 指针尖端点拟合效果图

顶点，从图中可以看出，对角线和指针中心线的方向非常相近，且该顶点和指针尖端点的位置也非常相近。

基于上述认知，本文提出一种十分简单的针对指针中心线检测缺失情况的备选方案：直接用指针边界框距离指针中心点欧氏距离最远的顶点坐标作为指针尖端点的近似拟合坐标，然后用该点作为指针尖端点继续后续计算步骤。此方法看似朴素且缺乏理论基础，但是经本文后续的实验验证，该策略给算法泛化能力带来提升效果非常显著，因此具有较强的工程应用价值，在指针信息提取算法无法准确计算出指针尖端点坐标的情况下是一个较好的后备方案。

#### 4.5 指针尖端点坐标标准化

由于不同仪表有不同的样式，其指针和标度数字的相对位置关系也并不一致，有些指针则相对较短，指针尖端点位于标度数字内侧，如图 4-16 中 a)所示；有些指针相对较长，指针尖端点位于标度数字外侧，如图 4-16 中 b)所示。

由于本文是面向未知仪表进行的读数识别，为方便后续的仪表读数计算，同时减小因仪表本身特点而带来的读数计算误差，对提取的指针尖端点进行标准化处理。为达到上述目的，本文提出一种基于仪表标度数字圆的指针尖端点标准化方法。首先定义所有

标度数字到指针旋转中心的平均距离为仪表标度数字圆半径，利用该仪表标度数字圆半径优化指针尖端位置坐标，将指针尖端点调整到距离指针中心点距离为标度数字圆半径的位置上。标度数字圆半径计算如公式(4-8)所示：

$$r = \frac{1}{n} \sum_{i=1}^n d_i \quad (4-8)$$

式中，为  $r$  为所述标度数字圆半径， $n$  为标度数字个数， $d_i$  表示标度数字中的第  $i$  个。



图 4-16 不同类型表盘示意图

进一步的仪表的指针尖端点优化公式如(4-9)所示：

$$\begin{aligned}x_{e21} &= x_c + \sqrt{\frac{r \times r}{1+k \times k}} \\y_{e21} &= y_c + k \times \sqrt{\frac{r \times r}{1+k \times k}} \\x_{e22} &= x_c - \sqrt{\frac{r \times r}{1+k \times k}} \\y_{e22} &= y_c - k \times \sqrt{\frac{r \times r}{1+k \times k}}\end{aligned}\quad (4-9)$$

式中， $P_c$  为标度数字圆心坐标，即指针中心点坐标， $k$  为  $P_c$  和  $P_{el}$  所连直线构成的斜率。

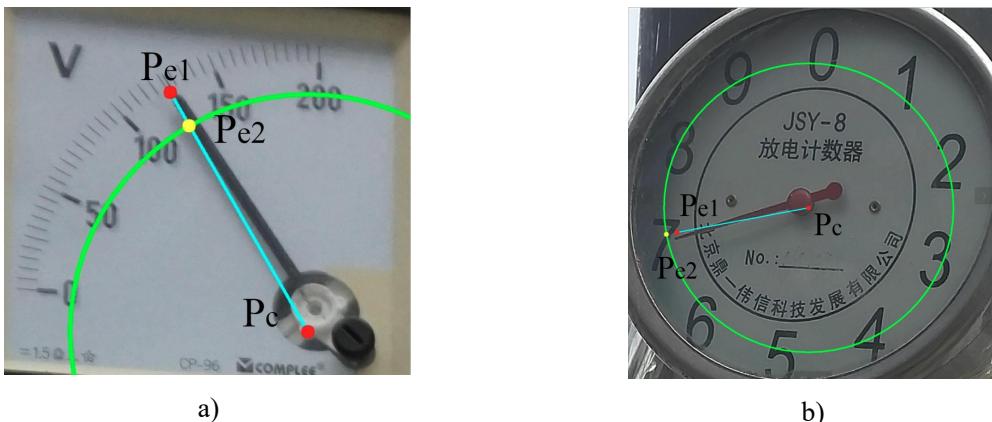


图 4-17 指针尖端点标准化效果图

最终分别计算得到的两点  $(x_{e21}, y_{e21})$  和  $(x_{e22}, y_{e22})$  分别与  $P_c$  之间的距离，保留距离较小的一个点作为优化后的指针尖端点。如图 4-17 所示，绿色的圆形表示了该表盘的标度数字圆，该圆以指针旋转中心点为圆心，以各个标度数字到的平均距离为半径得到。图中表示指针尖端的原始位置，表示经过优化后的指针尖端点位置，从图中可以看出和共线，且位于标度数字圆上。

## 4.6 本章小结

本章针对仪表自动读数识别过程中的表盘信息提取环节的表盘图像指针信息提取问题做了深入研究。首先基于 2.2 节提出的 Lite-FCOS 模型训练了指针目标检测模型用以从表盘图像中提取指针图像。然后分别提出了基于传统图像处理和基于深度学习模型直接预测的两种指针信息提取方法。其中，基于传统图像处理的方法中提出了一种改进的自适应二值化和 Hough 直线检测的方法求解指针中心线，并基于指针中心线进一步结合指针图像边界框求解得到指针尖端点坐标；基于深度学习的方法应用了关键点检测模型，通过微调训练之后的 HRNet 关键点检测模型直接从指针图像中预测指针中心点和指针尖端点的坐标。

为进一步提高算法的鲁棒性，本文提出了一种指针中心线检测算法失效的情况下，直接利用指针目标边界框估计指针尖端点位置的策略。最后，对指针尖端点坐标进行标准化处理，使不同样式仪表得到的指针尖端点统一在标度数字圆上，便于增强读数计算算法的泛化能力。实验结果表明，本章提出的两种方法都达到了提高指针信息提取过程鲁棒性和抗干扰能力的目的，具体实验数据及分析将在第六章集中进行讨论。

## 第五章 读数计算与总体框架

本文旨在实现仪表全过程自动读数识别并计算，最大限度降低对先验信息的依赖，同时保证读数正确率。为达到上述目的，本章整合全文各章节的算法，提出了仪表全过程自动读数识别的总体框架。首先针对角度法有效范围小等问题研究了针对性的读数计算公式优化与改进策略；接着，针对本文提出的读数计算公式，研究了读数基准标度区间的建立方法；最后，对第二章至本章提出的所有方法进行梳理，得到了仪表读数总体框架，并提出了基于目标检测模型和基于多深度学习模型的两种仪表自动读数识别算法作为本文研究的综合性成果。

### 5.1 仪表读数计算公式

在正式介绍仪表读数计算方法前，首先定义表盘图像中两点之间的夹角指表盘图像中任意两点和指针中心点所连的线段构成的，以指针中心点为顶点的劣角，即小于 180 度的角。可以基于上述三点构成表盘图像中的一个三角形，然后通过解三角形的相关几何知识基于这三点的坐标算得所述夹角的角度值。

基于角度法的仪表读数方法是文献中最常见的仪表读数计算方法，利用指针在表盘中的偏转角度结合仪表标度信息计算仪表读数结果。

原始的角度法读数依靠的是仪表的零标度和最大量程刻度，提取指针相对于零标度线偏角，然后将表盘零标度线到表盘最大量程刻度线之间的夹角视为已知量，然后利用角度和量程之间的比例关系计算仪表读数。如图 5-1 所示，两个红色圆点  $P_e$  和  $P_c$  分别代表指针尖端点和指针中心点，两个绿色圆点  $P_d$  和  $P_u$  分别代表零标度和最大量程标度位置，图中绿色的标度区域代表零标度和最大量程标度之间的角度范围，记为  $\theta_{\max}$ ，红色的标



图 5-1 仪表读数计算示意图

度区域代表零标度和指针中心线之间的角度范围，记为 $\theta$ ，则指针偏转角度和仪表示数比例关系如公式(5-1)所示：

$$\frac{\theta}{\theta_{\max}} = \frac{v}{v_{\max}} \quad (5-1)$$

式中， $v$ 表示指针指向位置的仪表读数数值， $v_{\max}$  表示指针式仪表的量程，即最大读数数值。

可以得出仪表读数计算公式如(5-2)所示：

$$v = \frac{\theta}{\theta_{\max}} \times v_{\max} \quad (5-2)$$

在原始角度法的基础上，一些研究者对其进行了改进，实现了一种基于读数基准标度区间的角度法读数计算方法。本文对此类改进的总结如下，该方法利用两个标度数字构成一个读数参照系，本文称之为读数基准标度区间，指针位于读数基准标度区间的内部，如图 5-2 所示。

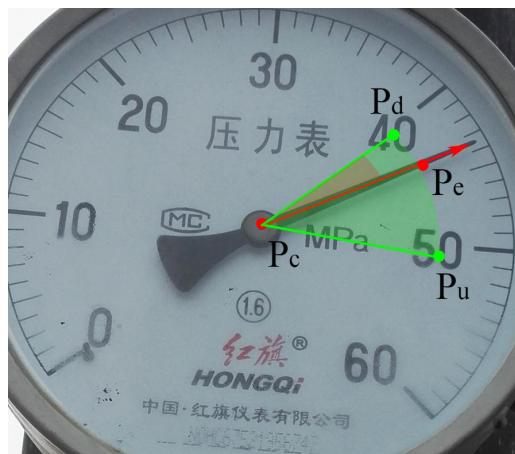


图 5-2 仪表读数计算示意图

记指针中心线和读数基准标度区间下界 $P_d$ 的夹角为 $\theta_d = \angle P_e P_c P_d$ ，读数基准标度区间上下界标度数字所构成的夹角为 $\theta_s = \angle P_u P_c P_d$ ，读数基准标度区间上下界标度数字的标度数值分别为 $v_u$  和  $v_d$ 。则指针在读数基准标度区间内部的偏转角度和仪表示数的比例关系如公式(5-3)所示：

$$\frac{\theta_d}{\theta_s} = \frac{\Delta v}{v_u - v_d} \quad (5-3)$$

进一步可以得出仪表的读数计算如公式(5-4)所示：

$$v = v_d + \Delta v = v_d + \frac{\theta_d}{\theta_s} \times (v_u - v_d) \quad (5-4)$$

使用公式(5-4)所示的仪表读数算法使得仪表读数识别过程初步具备了脱离对先验信息的依赖，能够从表盘图像中提取得到的标度信息中匹配的读数基准标度区间建立仪表读数计算的比例关系基础。不过上述方法要求必须获取指针位于两个标度区间的内部，由于标度信息的提取难以保证任何情况下都能完整获取每一个标度数字，因此当指针所指位置若有一侧的标度数字未能被识别出来，则读数计算将无法进行。例如图 5-2 中的指针若指向了标度数字 50 至 60 之间，而标度数字 60 未被检测出来则会出现这种情况。

为进一步增强仪表读数计算的泛化能力，提高算法的鲁棒性，本文提出对仪表读数计算公式做进一步的改进优化方法，使其能够对指针位于读数基准标度区间左右侧时，也能够得出正确的读数计算结果。将仪表读数计算划分为三个区间进行讨论，具体划分方法如图 5-3 所示，指针在读数标度区间内部记为绿色的 I 区域，指针在读数基准标度区间外部靠近标度区间下界记为红色的 S 区域，指针在读数基准标度区间外部靠近标度区间上界记为蓝色的 L 区域。

利用此方法，本文大幅扩展了一个读数基准标度区间所能够计算的指针偏转角度之外的区域，不再局限于指针标度区间的内部。但是需要说明，并不是表盘图像中的任何一个位置都可以用于指针读数结果的计算，受限于本文所用的平面内角度计算方法，本文要求任何一个夹角均为劣角，即小于 180 度的角，否则计算结果会出错，也就是说，图中 S、I、L 之外的一小部分黑色区域 N 是不能用于仪表读数计算的，若指针中心线位于该区域，则仪表读数计算公式失效，然而在实践过程中，这种情况是十分罕见的。

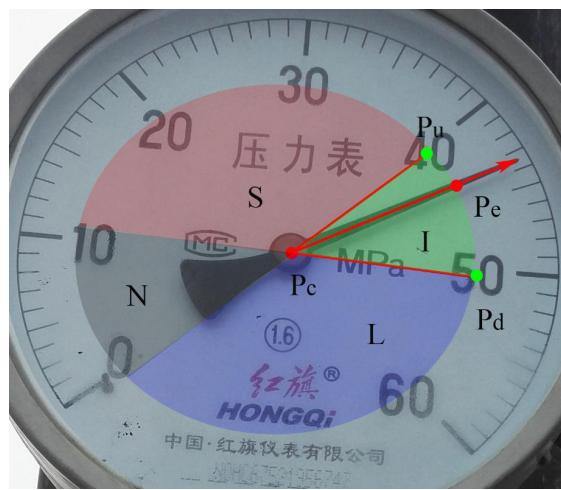


图 5-3 仪表读数计算示意图

通过指针中心线和读数基准标度区间的上下界之间的夹角判断指针所处的读数计算区间位置。当指针位于读数区域 I 时，指针中心线到读数基准标度区间上下界的夹角之和等于读数基准标度区间上下界之间的夹角  $\theta_u + \theta_d = \theta_s$ ；当指针位于区域 L 时，指针

中心线到读数基准标度区间上界的夹角和读数基准标度区间上下界之间的夹角之和等于指针中心线到读数基准标度区间下界的夹角  $\theta_u + \theta_s = \theta_d$ ；当指针位于区域 S 时，指针中心线到读数基准标度区间下界的夹角和读数基准标度区间上下界之间的夹角之和等于指针中心线到读数基准标度区间上界的夹角  $\theta_d + \theta_s = \theta_u$ ；

根据仪表的指针偏转角度和仪表示数之间的比例关系如公式(5-5)所示：

$$\begin{cases} \frac{\theta_d}{\theta_s} = \frac{\Delta v_d}{v_u - v_d} \\ \frac{\theta_u}{\theta_s} = \frac{\Delta v_u}{v_u - v_d} \end{cases} \quad (5-5)$$

式中， $\Delta v_d$  表示指针所指位置相对于读数基准标度区间下界的偏移量， $\Delta v_u$  表示指针所指位置相对于读数基准标度区间上界的偏移量。

由此可以进一步得到仪表读数的计算如公式(5-6)所示：

$$v = \begin{cases} v_d + \Delta v = v_d + \frac{\theta_d}{\theta_s} \times (v_u - v_d) & (1) \\ v_d + \Delta v = v_d + \frac{\theta_u}{\theta_s} \times (v_u - v_d) & (2) \end{cases}$$

由仪表表盘的读数特征，本文规定当指针位于 I 和 L 区域时，使用公式(5-6)中的式(1)计算仪表读数结果，当指针位于区域 S 时，使用公式(5-6)中的式(2)计算仪表读数结果。

## 5.2 读数基准标度区间匹配

上一节所述的仪表读数计算公式需要基于标度信息和指针信息确定一个读数基准标度区间才能得到完整的读数计算所需信息。理论上基于标度信息中任取两个标度数字所构成的区间都可以求得正确的仪表读数结果。然而在实际应用中，表盘轻微的倾斜，标度数字检测过程得到的标度数字位置精确程度有限、甚至是表盘本身的印刷效果都使得获取的表盘信息不可能达到理想状态。因此，为尽可能减小读数误差，读数基准标度区间跨度，即读数基准标度区间上下界之间的夹角应尽可能小。

本文提出最小距离优先原则用于读数基准标度区间匹配，即匹配的读数基准标度区间两端标度数字应该距离指针尖端点尽可能近。具体方法为计算指针尖端点到各个标度数字的欧式距离，然后选择距指针尖端点欧式距离最小的两个标度数字作为指针的读数基准标度区间，其中标度数值较大者为上界，较小者为下界。然而，仅使用上述最小距

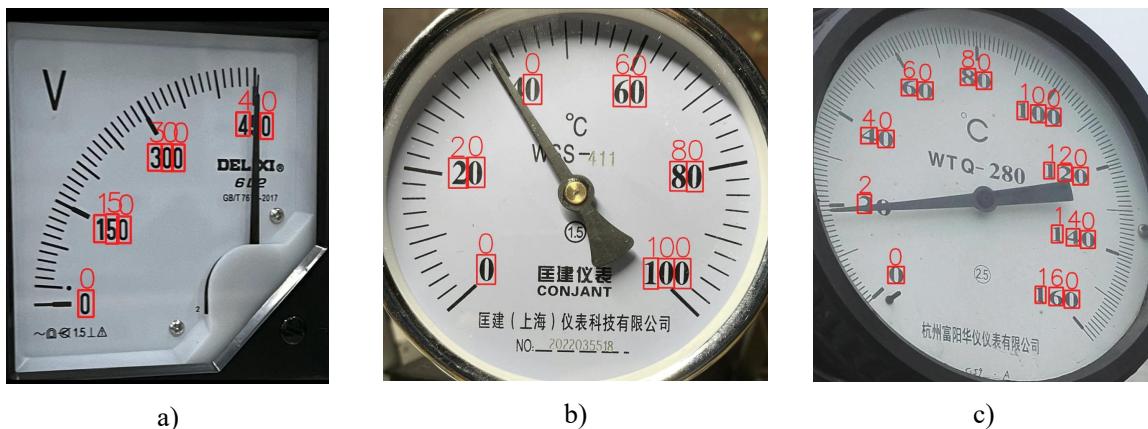


图 5-4 指针遮挡示意图

离优先原则可能出现几种错误匹配的情况，因此本文又针对性地提出了两条优化策略。

指针式仪表难免出现指针遮挡标度数字的情况，指针的遮挡会影响标度数字识别可靠性。如图 5-4 所示，子图 a)、b)、c) 中指针分别穿过标度数字“450”、“40”和“20”，导致这些标度数字被 OCR 模型错误地识别为“40”、“0”和“2”，如果继续使用这些标度数字匹配的读数基准标度区间，则必然会导致严重的读数识别错误，而且这种物理遮挡的效果会从根本上改变字符的图像特征，因此很难依靠改进文本识别模型来解决这一问题。

为解决此问题，本文提出将指针附近标度数字视为不稳定标度数字，在进行读数基准标度区间匹配之前将其从标度信息中剔除。由于本文改进的仪表读数计算公式对读数基准标度区间的鲁棒性很好，剔除该标度数字之后只需选择其他更大范围的标度数字构成读数基准标度区间即可，并不会给读数计算带来负面影响。具体计算过程如下，首先计算标度数字位置到指针所在直线的最近距离，若此最近距离小于该标度数字的指针干扰阈值，则判定其距指针过近，在匹配的过程中忽略这个标度数字。本文经过实验确定标度数字的边界框对角线长度一半作为该标度数字的指针干扰阈值，距离表读数数字距离小于指针干扰阈值的区域为指针干扰区。

如图 5-5 所示，用一个黄色圆形区域绘制了标度数字“7”和“40”的指针干扰区。在图 a) 中，由于指针中心线有一部分位于指针干扰区内，则说明标度数字“7”是一个不稳定标度数字，需要被剔除；在图 b) 中，指针不在指针干扰区内，则无需剔除任何标度数字，直接得到标度数字“40”和“50”构成的读数基准标度区间。子图 a) 利用不稳定标度数字剔除之后的标度信息进行读数基准标度区间匹配，最终确定标度数字 6 和标度数字 8 所构成的区间为最终的读数基准标度区间。



图 5-5 指针干扰区示意图

在 5.1 小结中，本文提到了尽管改进之后的仪表读数计算公式鲁棒性非常强，但是受限于两点之间角度的计算方法，要求本文所有的两点之间夹角必须为劣角，因此如果指针中心线落在图 5-3 中黑色的 N 区域内部，则不满足仪表读数计算公式的使用条件。由于 S、I、L 已经覆盖了表盘的大多数区域，所以这种情况一般并不常见也不会对仪表读数识别构成问题。但是在仪表标度数字分布较为稀疏的情况下，如果指针非常靠近零标度或者最大量程标度，由于上述指针干扰区机制，该标度被剔除后，则存在读数基准标度区间同时匹配了剩余标度数字中最大和最小标度的情况，会导致无法满足仪表读数计算公式条件的情况。如图 5-6 所示，标度数字 0 被提出后，标度数字 16 的距离相比标度数字 8 的距离略近一点，此时标度数字 4 和 16 构成了读数基准标度区间，由图中可以看出，此时指针落在黑色的 N 区域中，此时若强行使用仪表读数公式进行计算，则会得出错误的读数计算结果。

经过分析，上述情况的出现典型特征为：表盘读数基准标度区间是由剩余标度数字中标度数值最大和最小的标度数字共同构成的。由此，本文提出对匹配完成的读数基准标度区间进行检验，判断是否出现了上述由剩余标度数字中标度数值最大和最小的标度



图 5-6 仪表读数计算公式失效示意图

较近的端点，剔除另外一个，并从剩余标度数字中再取其中距离指针尖端点最近的标度数字再次构成标度区间。对新的标度区间也要重复上述过程，直到不再存在上述问题。本文针对这一现象提出的解决方案无法从理论上避免所有指针中心线落入 N 区域的情况，但是再仪表自动读数识别这样的具体应用场景下，从实验效果中可以得出结论，本方法基本能够解决测试数据集中所有出现此类问题的图像，因此仍然可以认为是有效的。

改进效果如图 5-7 所示，子图 a) 和 b) 中，指针分别穿越了标度数字“5”和标度数字“20”，遮挡了这两个字符的结构，因此这两字符被判定为不稳定标度数字，没有参与读数基准标度区间的匹配；子图 c) 中标度数字分布较为稀疏，标度数字“0”被判定为不稳定标度数字剔除后，标度数字“2.5”距离指针尖端点的欧式距离相比“1”更小，但是“0.5”至“2.5”分别是最大和最小标度数字此时指针处于 N 区域中，根据 5.2 节所述策略“2.5”会被剔除，然后在剩余标度数字中重新进行匹配，最后得到“0.5”到“1”的读数基准标度区间并在后续的计算中得到了正确的读数。

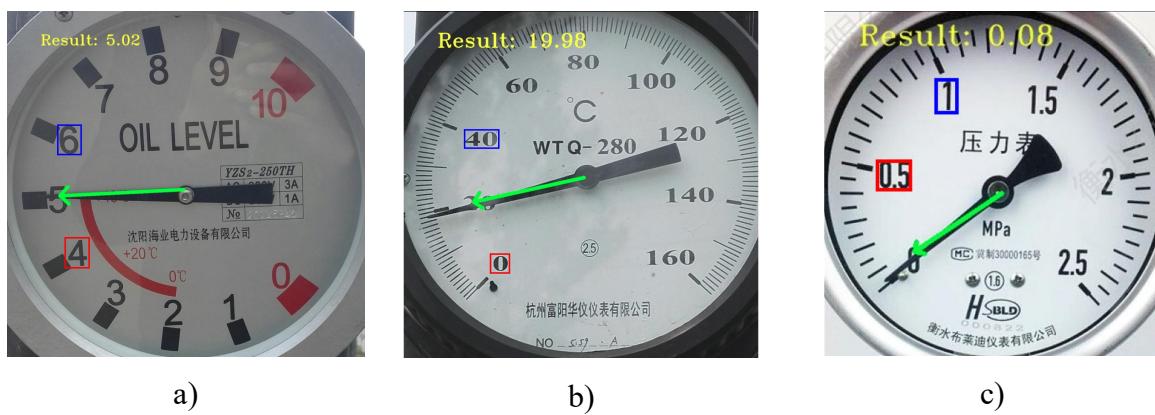


图 5-7 标度区间匹配优化示意图

### 5.3 仪表读数计算总体框架

本节综合上文第二、三、四、五章之前的所有内容，整合之后提出完整的仪表自动读数识别算法，并得出总体方案的框架。本文提出巡检机器人仪表自动读数识别框架如图 5-8 所示，包括场景图像获取、表盘图像提取、表盘信息提取和仪表读数计算等环节，表盘图像提取环节可进一步细分为表盘检测和表盘图像增强两个环节、表盘信息提取环节可进一步细分为指针信息提取和标度信息提取，仪表读数计算环节可进一步细分为读数基准标度区间匹配和仪表读数计算公式。每个上述环节都需要解决一个仪表读数识别的关键任务，并得到对应的仪表读数识别中间信息，分别是：高质量表盘图像，表盘信息和仪表读数识别结果。

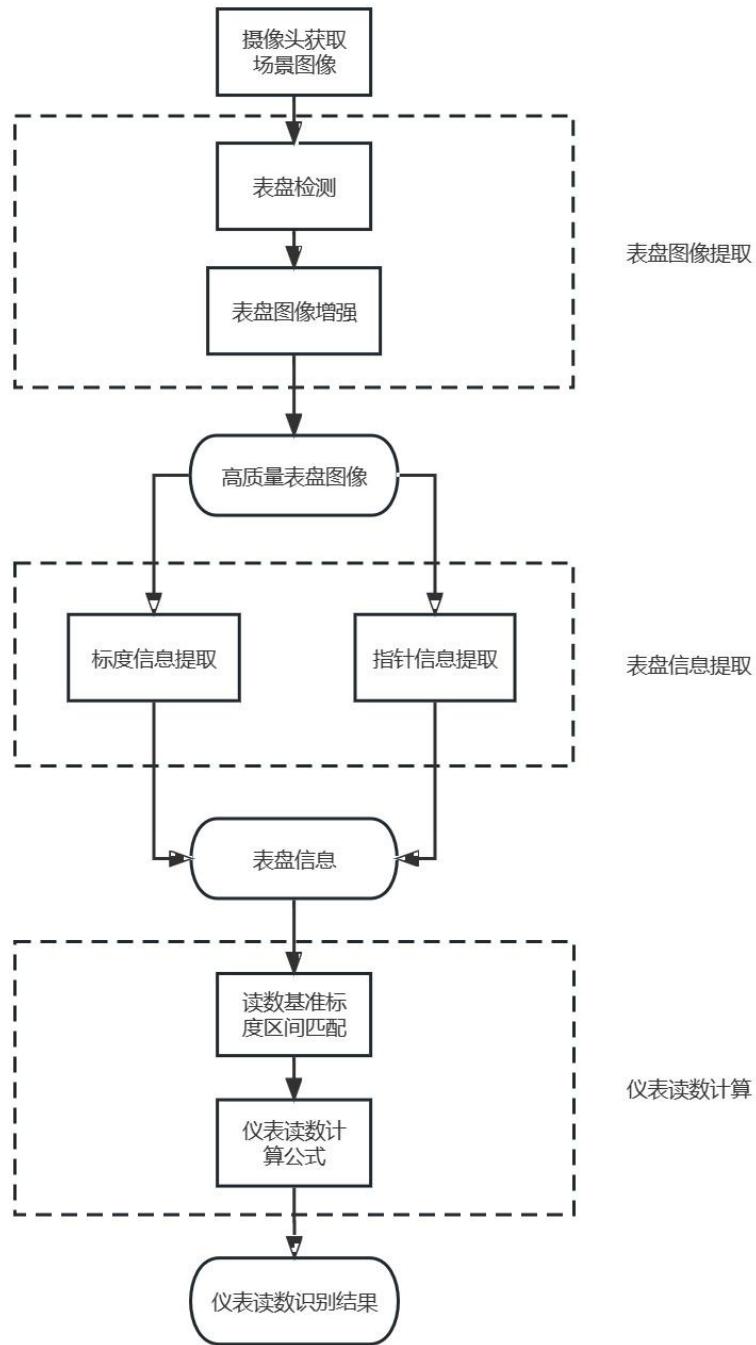


图 5-8 仪表自动读数识别总体方案流程图

基于上述仪表自动读数识别方法框架图,本文综合了上文所述的各个环节的具体实现方法提出了两种具体仪表全过程自动读数识别算法的流程。

第一种算法综合目标检测、OCR 和关键点检测等多种深度学习模型,利用 3.1、3.2 和 4.3 写章节所述方法获取表盘信息,本文将其命名为基于多深度学习模型的指针式仪表读数算法,英文名为 Meter Reading based on Multiple deep learning models Algorithm,为简化描述下文将其简写为 MRM 算法。如图 5-9 所示为该算法的流程图。

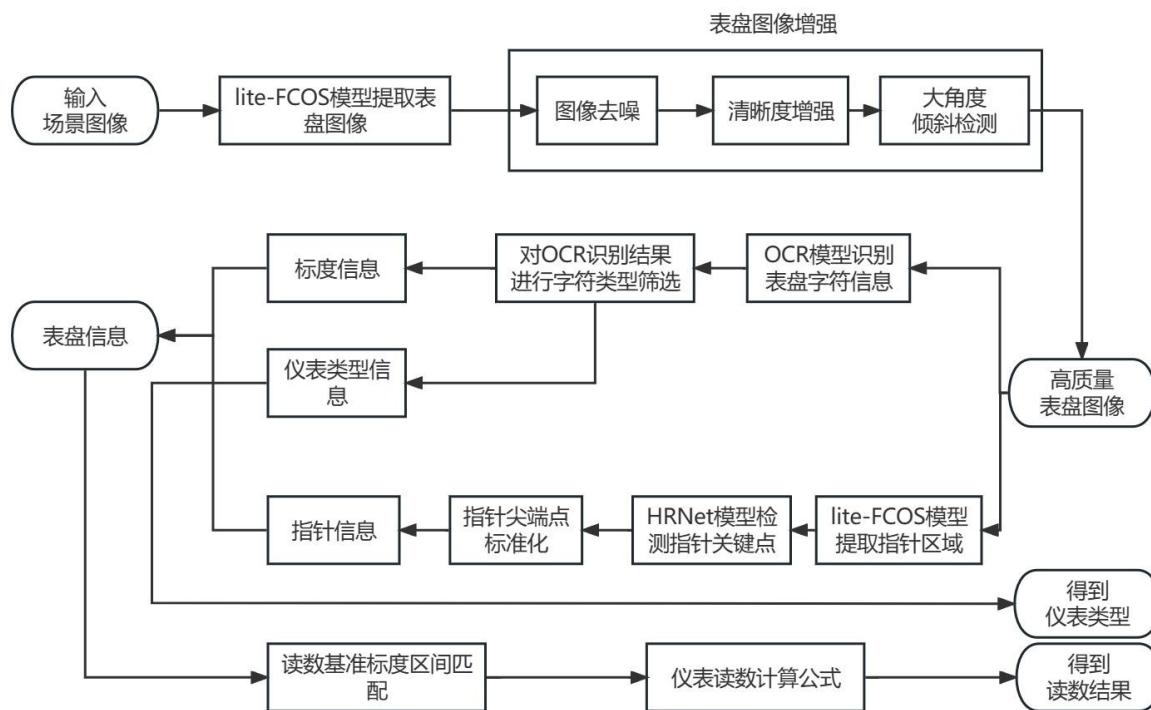


图 5-9 MRM 算法流程图

具体算法流程如下：

- 1) 巡检机器人获取巡检点位拍摄得到的场景图像并输入算法；
- 2) Lite-FCOS 模型对场景图像进行表盘检测，并用边界框提取表盘图像；依次进行双边滤波图像去噪、使用 RealSR 模型对低分辨率图像进行超分辨率重建清晰度增强，最后检验表盘图像是否存在严重倾斜；
- 3) 经过上述表盘图像处理之后得到的图像为高质量表盘图像；
- 4) 使用 OCR 模型获取表盘上字符信息，得到表盘字符 OCR 识别结果，然后对其进行筛选处理，分别得到标度信息和仪表类型信息；
- 5) 使用 Lite-FCOS 模型对表盘图像进行指针检测，并用边界框提取指针图像；
- 6) 使用 HRNet 关键点检测模型检测指针中心点和指针尖端点，并对指针尖端点进行标准化处理；
- 7) 基于标度信息和指针信息进行读数基准标度区间的匹配，确定计算仪表读数用的读数基准标度区间；
- 8) 基于指针中心点、指针尖端点和仪表读数基准标度区间利用仪表读数计算公式计算仪表的读数结果；
- 9) 基于仪表类型信息确定仪表类型；

上述流程中步骤 4) 至步骤 5 和步骤 6) 至步骤 8) 可以并行处理。

第二种算法仅使用目标检测模型，利用 3.3 节和 4.2 节所述方法获取表盘信息，本文将其命名为基于目标检测模型的指针式仪表读数算法，即 Meter Reading based on Object Detection algorithm，为简化描述下文将其简写为 MROD 算法。如图 5-10 所示为该算法的流程图。

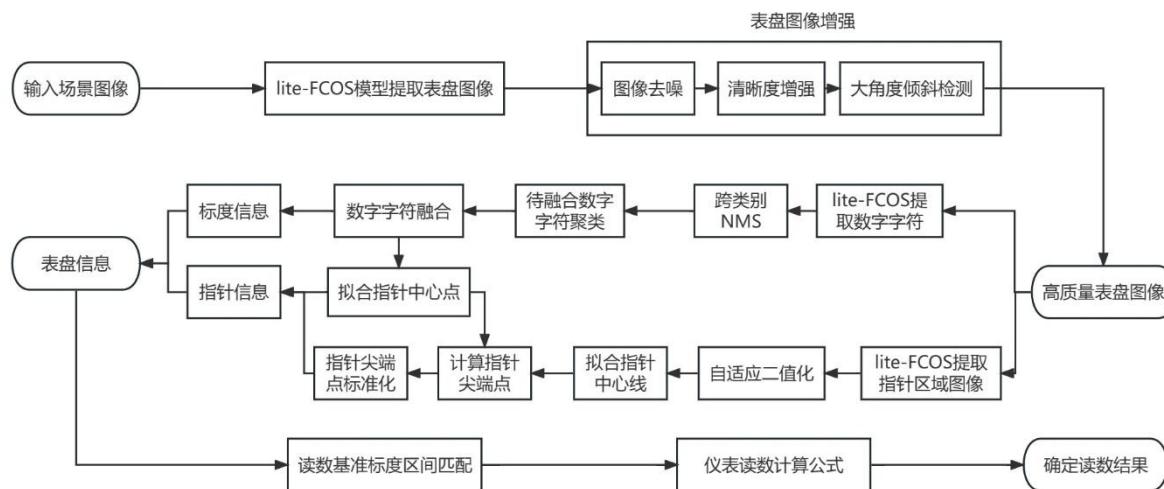


图 5-10 MROD 算法流程图

具体算法流程如下：

- 1) 巡检机器人获取巡检点位拍摄得到的场景图像并输入算法；
- 2) Lite-FCOS 模型对场景图像进行表盘检测，并用边界框提取表盘图像；
- 3) 依次进行双边滤波图像去噪、使用 RealSR 模型对低分辨率图像进行超分辨率重建清晰度增强，最后检验表盘图像是否存在严重倾斜；
- 4) 经过上述表盘图像处理之后得到的图像为高质量表盘图像；
- 5) 使用 LiteFCOS 模型获取表盘上单个字符信息和指针图像；
- 6) 基于表盘图像单个字符信息通过数字融合获取标度信息，具体包括单个字符目标的跨类别 NMS 处理，待融合数字字符聚类和数字融合等步骤。基于指针图像通过拟合指针中心线获取指针信息，其中包括：对指针图像进行自适应二值化处理、使用改进的 Hough 变换直线检测算法拟合指针二值图像中的指针中心线，然后进一步确定指针尖端点并对其进行标准化；
- 7) 基于标度信息和指针信息进行读数基准标度区间的匹配，确定计算仪表读数用的读数基准标度区间；
- 8) 基于指针中心点、指针尖端点和仪表读数基准标度区间利用仪表读数计算公式

## 计算仪表的读数结果

上述流程中步骤 5) 和步骤 6) 可以并行处理。

总体来说两种算法的主要区别在于表盘信息提取环节，MRM 算法尝试更激进的深度学习模型应用策略，通过广泛采用 OCR 和关键点检测等计算机视觉的其他领域的相关方法以追求更好的鲁棒性和更强的表盘信息提取泛化能力，同时还基于 OCR 模型能够提取仪表类型相关信息的优势，实现了自动识别位置仪表类型的功能。MROD 算法相对保守，仅仅使用目标检测模型，重点在于模型的轻量化改进，通过结合本文提出的数字字符融合与改进 Hough 变换指针检测等算法，也实现了具有较强泛化能力的仪表读数识别性能，同时该方法具有速度和参数量上的优势非常利于实际应用场景下的部署和嵌入式边缘端硬件设备的后续移植开发。

## 5.4 本章小结

本章针对仪表自动读数识别过程中的读数计算环节进行了深入研究。提出了一种针对角度法的改进方案，大大增强了角度法读数计算公式有效性的鲁棒性。针对改进后的读数计算公式，提出了所需的读数基准标度区间匹配方法，并分别对指针遮挡字符和指针中心线不满足读数计算公式条件两类情况提出了改进措施。综合本章节和上文各章节提出的算法与技术，整合出基于目标检测模型和基于多深度学习模型的两种仪表自动读数识别算法，分别命名为 MRM 算法和 MROD 算法，上述两种算法为本文针对仪表自动读数识别问题所提出的最终解决方案。

## 第六章 仪表读数实验及分析

前面章节已经对本文提出的两种仪表自动读数识别算法的技术细节进行了详细的介绍。在此基础上，本文收集并制作了测试实验数据集以验证算法的有效性和泛化能力为主要目标，从检测精度、检测速度等方面对本文提出的方法进行验证，分别设计了读数识别性能验证实验、消融实验、读数对比实验和抗扰能力分析实验，通过充分的实验验证本文所提出方法的有效性，并总结存在的不足。由于 MRM 和 MROD 算法的表盘检测部分完全相同，且该环节相关实验已经在 2.3 小节进行了详细分析和讨论，因此本章实验重点聚焦表盘图像增强到最终读数计算的过程。

### 6.1 实验说明

为便于读者了解实验的背景和细节，本小节从数据集，评价指标、实验硬件环境与模型训练配置等几个角度对实验总体情况进行说明。

#### 6.1.1 数据集

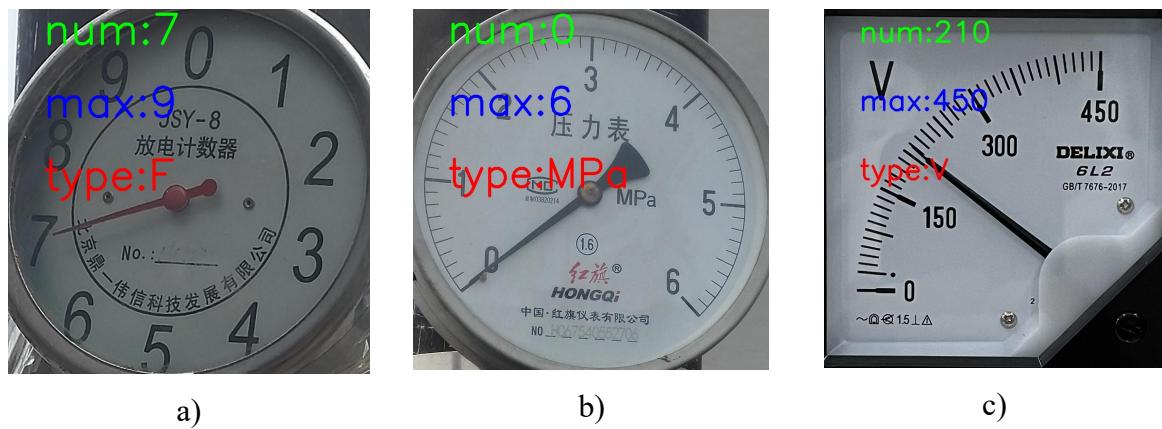


图 6-1 MRD 数据集标注示意图

由于深度学习模型训练使用的数据集已经在对应章节进行过介绍，因此这里不再赘述，本节主要介绍仪表读数识别测试用数据集。本文使用自行标注的测试数据集来检验仪表读数算法的读数性能，本文从 DRID 系列数据集中选用了 280 张表盘图像并分别对每张图片标注仪表读数、仪表最大量程标度和仪表类型三种信息，制作了仪表读数测试数据集，命名为 MRD，为英文名称 Meter Reading Dataset 的缩写，其中部分图像和标注如图 6-1 所示，图片中的“num”、“max” 和 “type” 分别对应该表盘图像的仪表读数、仪表最大量程标度和仪表类型的标注结果。

为有效验证算法的泛化能力，本数据集中 80% 以上图像均由自行实际拍摄的场景图像分割得到，这些图像包括了不同远近程度，左右小角度倾斜，明暗变化等多种实际中

可能出现的扰动因素，同时测试数据集中所有表盘都未曾在上文所述的用于深度学习模型训练和验证的数据集中出现过，属于同类文献中很少出现的无先验信息表盘图像。

### 6.1.2 评价指标

针对仪表读数识别算法，为突出检验仪表读数算法的泛化能力，本文提出以下仪表读数评价指标：

- 1) 检出率：指仪表读数算法能够输出读数结果的样本数；
- 2) 读数准确率：参考上海工业自动化仪表研究所指定的《工业过程测量和控制用检测仪表和显示仪表精确度等级》<sup>[78]</sup>，本文取读数引用误差低于 5%，即精度等级 5 及以上的图像为读数正确图像，统计读数算法能够输出正确读数结果的样本数目为读数正确数，其中仪表读数引用误差计算如公式(6-1)所示；

$$e = \frac{|v - v_{real}|}{v_{max}} \times 100\% \quad (6-1)$$

式中， $e$  为仪表读数的引用误差， $v$  为算法得到的读数数值， $v_{real}$  为仪表的标注真实值， $v_{max}$  代表仪表的最大量程标度数值

- 3) 类型准确率：指仪表读数算法能够输出正确仪表类型识别结果的样本数；
- 4) 帧率：指算法每秒能处理的图像张数，也是图像处理算法的常见速度指标，本文基于单张测试图像平均检测时间计算算法的处理帧率；
- 5) 平均引用误差：由于部分读数失败的图像会产生偏差极大的读数结果，因此在本文中使用读数成功图像的引用误差平均值作为检测精度的指标，当读数准确率较高时该指标同样能够较好反映算法的读数精度。

### 6.1.3 硬件环境与模型训练配置

本文实验使用的 CPU 为 i9-9900x 3.5Ghz，操作系统为 Ubuntu16.04LTS，使用显卡为 NVIDIA Geforce 2080ti。同时使用了 PyTorch 和 PaddlePaddle 两种深度学习框架。

对本文中应用的目标检测模型，统一设定实验优化器为 SGD，动量 momentum 设定为 0.9，学习率设定为 0.01；对本文中的 OCR 模型，设定训练的 epoch 为 1200 轮，使用 Adam 优化器，设定学习率为 0.001，使用 L2 正则化防止过拟合；对本文中的 HRNet 关键点检测模型，设定训练的 epoch 为 210 轮，使用 Adam 优化器，设定初始学习率为 0.0005，并在训练过程中使用了 warmup 的策略以提升学习的效果。

## 6.2 读数识别性能验证

本节实验用于验证本文所提出的 MRM 算法和 MROD 算法的具备预期的面向未知仪表进行读数识别的能力，使用 MRD 数据集进行读数实验，测试集总数为 280 张图像，实验结果如下表 6-1 所示：

表 6-1 算法性能验证实验

算法	检出率	读数准确率	类型准确率	帧率	平均引用误差
MRM	99.6	84.6	79.3	4.0	1.2
MROD	98.5	88.6	-	13.5	1.7

注：测试集总数 280 张

分析实验结果可知，MRM 和 MROD 两种算法均初步具备了面向未知仪表进行读数识别的能力，两个算法的读数准确率分别能够达到 84.6% 和 88.5%，且成功读数的仪表平均引用误差分别仅为 1.2% 和 1.7%，距离精度等级 5 所要求的 5% 的基准线较远，说明读数成功的图像中总体误差较小，读数结果具备较高精度水平。

对比两种算法可知，MROD 算法性能相对 MRM 算法性能更优，读数正确率更高，同时帧率相比 MRM 方法也要更快。本文认为，精度更高是由于 MROD 算法利用了目标检测模型训练 10 类数字字符目标和 1 类指针目标，目标的种类相对固定使得模型训练效果更好，同时模型直接预测字符的标签也省去了字符识别的环节，而 OCR 模型面向任意字符识别，且需要使用 CRNN 模型进行文字识别，因此基于 Lite-FCOS 模型识别字符结果进行数字融合的算法标度信息提取效果优于基于 OCR 识别结果进行筛选的方法，进而导致读数识别算法性能更优。MROD 算法速度更快是由于使用的网络较为简单，仅使用了两次本文提出的 Lite-FCOS 模型，且该模型本身就是进行轻量化改进的，其速度和参数量相比整合了 Lite-FCOS 模型、OCR 模型、HRNet 关键点检测模型的 MRM 算法有着显著优势。MRM 算法的在于 OCR 模型的全表盘字符识别能力，使其具有了面向未知仪表识别其类型的功能，该功能的识别准确率可达约 80%，具备了实用价值。

如图 6-2 所示是一些 MRM 和 MROD 算法的读数效果图，图中绿色箭头表示算法提取得到的指针信息，箭头方向从指针中心点指向指针尖端点（标准化后），图中使用红色边界框表示匹配得到的读数基准标度区间下界，使用蓝色边界框表示匹配得到的读数基准标度区间上界，在图片的左上角使用黄色字体显示了仪表自动读数识别算法得到的读数结果。图 6-2 中可以看出，子图 a)、b)、c) 均匹配了距离指针尖端点最近的两个标度数字构成了读数基准标度区间，并且得到了可靠的读数计算结果。

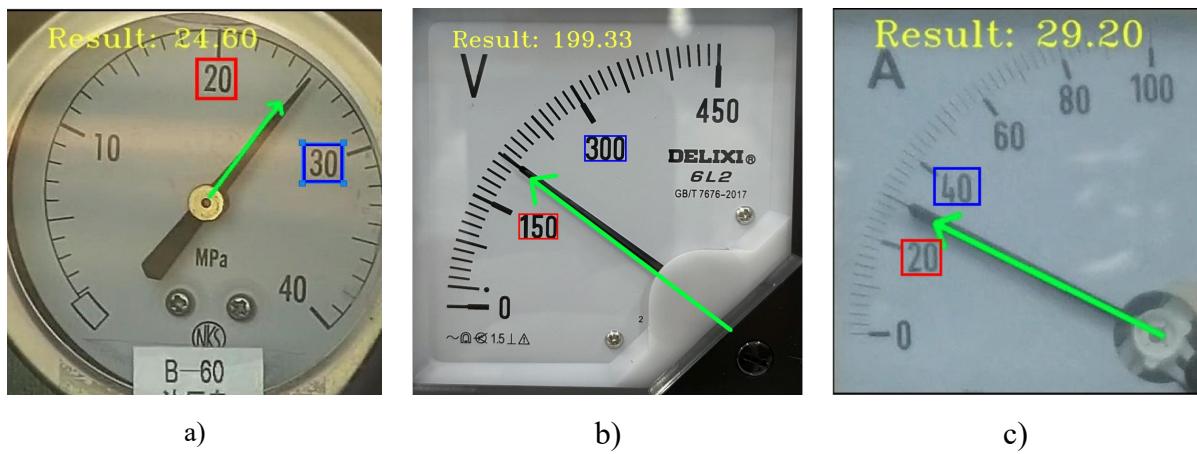


图 6-2 仪表读数结果示意图

### 6.3 消融实验

为进一步研究本文提出的各项改进效果，本文又对各项改进进行了消融实验。消融实验指通过对改进点设定对照实验，用去除改进的方法和包含改进的方法进行对比，以验证改进模块的有效性，常用于深度学习的相关研究中。如果改进去除后，总体性能下降显著，则说明该模块发挥了作用。由于 MRM 和 MROD 两种算法在标度信息提取和指针信息提取都有较大差异，因此本文对每种改进都分别在两种算法上进行了实验测试。

**1) 超分辨率重建清晰度增强：**本文 2.4.2 节中使用了 RealSR<sup>[65]</sup>超分辨率重建模型进行清晰度增强后，而对比方法没有该环节。

实验结果如表 6-2 所示，实验结果显示两种算法的读数准确率和检出成功率都有了显著提升，其中 MRM 算法的读数准确率提升了 6.4%，MROD 算法的读数准确率提升了 4.6%。改进对 MRM 算法的性能提升效果更加明显，可能是 OCR 模型对图像的细节更加敏感，因此受图像清晰度影响较大。这一点可以从 MRM 算法的类型识别准确率看出，图像超分使得类型识别准确率提升了 9.6%，因为字符的识别相比数字对图像清晰度的要求更高，同时下文的抗干扰性能验证实验中还将再次验证此观点。

表 6-2 SR 清晰度增强消融实验

算法	图像超分	检出率	读数准确率	类型成功数
MRM	✓	99.6	84.6	79.3
		91.4	78.2	69.6
MROD	✓	98.5	88.6	-
		96.8	83.9	-

注：测试集总数 280 张

**2) 近似指针尖端点:** 本文 4.4 节中提出一种使用边界框四个顶点中距指针中心点最近的顶点作为指针尖端点的近似点，并在指针信息提取算法无法找到准确指针尖端点的情况下作为替代完成后续计算，对比方法没有该环节。

实验结果如表 6-3 所示，实验结果显示，两种算法的检出率和读数准确率都有显著提升，其中 MRM 算法的读数准确率和检出率分别提高了 1.8%，MROD 算法的读数成功成功率和检出率分别提高了 7.1% 和 8.9%。本文在消融实验的基础上还针对指针尖端点进行了深入实验，研究了仅使用近似指针尖端点作为指针尖端点信息提取方法的仪表读数识别性能，MRM 算法读数准确率可达 78.4%，MROD 算法读数准确率可达 81.4%，使用如此简单的策略就能够达到较高的仪表读数性能令人印象深刻。由此可以得出结论，使用边界框中的顶点在指针尖端点检测失效的情况下是具有可行性的，能够有效增强算法的泛化能力，提高算法的鲁棒性。99.6

表 6-3 近似指针尖端点消融实验

算法	近似尖端点	检出率	读数准确率	类型成功数
MRM	✓	99.6	84.6	79.3
		99.6	82.8	79.3
MROD	✓	98.5	88.6	-
		89.6	81.4	-

注：测试集总数 280 张

**3) 标度区间匹配优化:** 本文 5.2 小结中提出了在最小距离匹配原则的基础上进一步对读数基准标度区间的匹配过程进行优化的策略，重点针对指针带来的字符干扰和标度数字分布稀疏导致匹配失效的情况，对比方法仅使用最小距离匹配原则进行读数基准标度区间的匹配。

具体实验结果如表 6-4 所示，实验结果显示，MRM 算法的读数准确率提升了 11.1%，MROD 算法的读数准确率提升了 3.9%，说明仅基于最小距离匹配原则，的确会带来一定的不利影响，本文的分析是正确的，对应提出的针对性改进策略也发挥了预期效果。这些改进策略针给 MRM 算法能够带来更大的性能提升，这可能是由于 OCR 模型提取表盘字符相比基于 Lite-FCOS 模型结合数字融合的方法漏检的几率更高，也就更容易造成提取得到的标度数字在表盘分布稀疏的现象。

**4) 读数计算公式优化:** 本文 5.1 小结中提出了针对仪表读数计算公式的改进，得到了优化后的仪表读数计算公式(5-6)，大幅拓展了基于读数基准标度区间的仪表读数有效

表 6-4 标度区间匹配优化消融实验

算法	匹配优化	检出率	读数准确率	类型成功数
MRM	✓	99.6	84.6	79.3
		99.6	73.5	79.3
MROD	✓	98.5	88.6	-
		98.5	84.6	-

注：测试集总数 280 张

范围，对比方法为 5.1 节中改进前的公式(5-4)。

实验结果如表(6-5)所示，实验结果显示，MRM 算法的读数准确率提升了 14.3%，MROD 算法 的读数准确率提升了 8.5%，均为所有公共改进点中最大的性能提升，可见对仪表读数识别计算公式的改进具有重要作用，大幅降低了算法对于表盘信息的要求，给表盘信息提取阶段带来了更多的容错空间，因而显著提升了算法的泛化能力。

表 6-5 读数计算公式优化消融实验

算法	改进公式	检出率	读数准确率	类型成功数
MRM	✓	99.6	84.6	79.3
		99.6	70.3	79.3
MROD	✓	98.5	88.6	-
		98.5	83.6	-

注：测试集总数 280 张

**5) OCR 微调训练：**本文 3.1.3 小结介绍了对 OCR 模型进行微调训练的过程，得到表盘图像专用的模型权重参数，对比方法为相同 OCR 模型在 PaddleOCR 计算机视觉框架中开源的通用 OCR 模型权重参数。

实验结果如表 6-6 所示，实验结果显示，经过训练的模型相比通用模型检出率提高了 8.9%，读数准确率提高了 16.8%，类型识别准确率提高了 32.9%。实验结果显示，通用 OCR 模型直接迁移到仪表读数识别任务效果比较一般，针对表盘字符重新标注数据集然后对模型进行专门的微调训练对提高 OCR 模型检测效果至关重要，对于 MRM 算法中的仪表类型识别效果，这一过程更加重要。

表 6-6 OCR 微调训练消融实验

算法	微调 OCR 模型	检出率	读数准确率	类型成功数
MRM	✓	99.6	84.6	79.3
通用模型		90.7	67.9	46.4

**6) 跨类别 NMS:** 本文 3.3.1 小节提出针对 Lite-FCOS 模型得到的单个字符检测结果还需要进行跨类别 NMS 以避免某个字符被识别成不同字符而导致的重复检出，对比方法没有该环节。

实验结果如表 6-7 所示，实验结果显示，使用跨类别 NMS 算法能够使 MROD 算法的读数准确率提高 8.2%，显著提升了仪表读数识别准确率。实验结果表明，由于干扰等因素，单一数字字符的确会被重复识别成多个不同字符带来对读数的不利影响，本文提出的跨类别 NMS 算法能够有效解决这一问题。

表 6-7 跨类别 NMS 消融实验

算法	跨类别 NMS	检出率	读数准确率
MROD	✓	98.5	88.6
		98.5	80.3

注：测试集总数 280 张

**7) 剔除离群点：**本文 4.2.1 小结提出了针对指针中心点拟合过程中的离群拟合点进行剔除的策略以避免不合适的拟合点带来指针中心点拟合位置偏移，对比方法没有这一环节。

实验结果如表 6-8 所示，实验结果显示，剔除离群拟合点能够给读数准确率带来 10% 的提升，对仪表读数识别效果提升显著。实验结果表明，表盘倾斜、标度信息提取过程中产生的误差等种种干扰因素的确会导致部分拟合点偏离指针中心点的情形，但是本文提出的剔除拟合点方法能够有效的改善这一问题，带来算法性能提升。

表 6-8 剔除离群点消融实验

算法	剔除离群点	检出率	读数准确率
MROD	✓	98.5	88.6
		99.6	78.6

注：测试集总数 280 张

**8) 自适应二值化指针图像：**本文在 4.2.1 节提出了针对传统 Hough 直线检测的改进，对直线检测中的指针图像二值化进行了改进以提高算法的适应能力，对比方设定了不同的固定二值化阈值；

实验结果如下表 6-9 所示，使用本文提出的自适应 Hough 直线检测方法效果相比其他使用固定阈值 Hough 直线检测方法都有一定的提升，相比二值化采用阈值 60, 80, 100 的情况分别提升了 3.3%, 2.2% 和 2.6%。实验结果表明，本文提出的自适应 Hough

表 6-9 自适应二值化指针图像消融实验

算法	二值化阈值设定	检出率	读数准确率
MROD	自适应	98.5	88.6
	60	98.5	85.3
	80	98.5	86.4
	100	98.5	86.1

注：测试集总数 280 张

直线检测是有效的改进，能够对仪表自动读数识别算法的泛化性能提升起到积极作用。

**9) 总体改进对比实验：**由于不同改进点之间也会存在相互影响，为验证本文提出的一系列改进的总体效果，本小节进行一次总体改进对比实验，即放弃使用上述所有改进点进行仪表读数实验测试，与本文提出的完整 MRM 算法和 MROD 算法进行对比。

实验结果如表 6-10 所示，实验结果显示，全面应用本文所述的一系列改进使得 MRM 算法的读数准确率相比不应用这些改进提升了 48.9%，MROD 算法的读数准确率提升了 28.4%。实验结果表明，本文提出的一系列改进大幅提高了仪表读数的泛化能力，具有显著的改善效果和较高的实用意义。

表 6-10 总体改进对比实验

算法	应用改进	检出率	读数准确率	类型成功数
MRM	✓	99.6	84.6	79.3
		90.7	35.7	46.4
MROD	✓	98.5	88.6	-
		90.7	60.0	-

注：测试集总数 280 张

## 6.4 读数对比实验

现有文献中多数算法都高度依赖表盘先验信息，导致无法实现完成面向未知仪表的读数识别，因此无法应用本文提出的 MRD 数据集对其进行测试，而这些文献本身的数据集也不是开源的，获取难度较大。上述因素导致本文难以直接和同类方法进行对比实验。指针信息提取在仪表读数识别相关研究中受到最多的重视，这其中指针中心线提取又是核心步骤，目前已经形成了多种典型方法。因此，本文聚焦该环节作为对比实验切入点，复现了主流的几种指针中心线拟合方法和本文对应的指针信息提取方法进行比较，包括：最小二乘法 LS<sup>[38, 79]</sup>、主成分分析法 PCA<sup>[80]</sup>，随机采样一致算法 RANSAC<sup>[31, 81]</sup>。

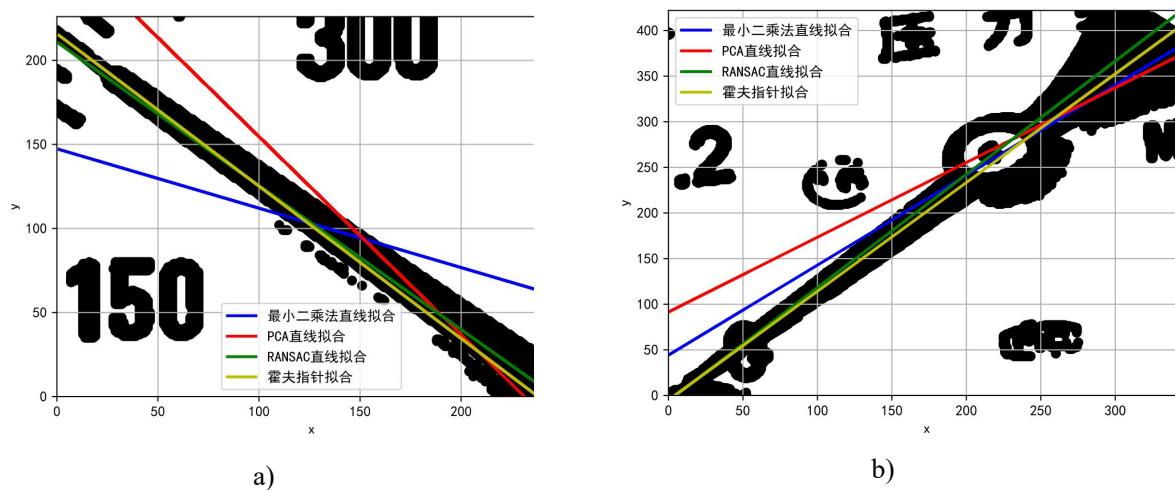


图 6-3 指针中心线提取效果图

上述算法的指针中心线拟合效果对比如图 6-3 所示, 图中红色直线表示 PCA 算法的拟合结果, 蓝色直线表示 LS 算法的拟合结果, 绿色直线为 RANSAC 算法的拟合结果, 黄色直线表示本文改进后的 Hough 直线检测算法提取结果, 关键点检测提取指针信息的算法由于本身并不对指针中心线进行拟合, 因此下图并未展示其拟合效果。

本文使用上述算法来代替 MRM 和 MROD 算法中对应环节进行仪表读数识别的对比实验。由于本文提出的两种算法实现差异较大, 本文分别进行了实验以达到控制实验变量和获取更可靠实验结果的目的。对于 MROD 算法, 分别使用上述指针中心线拟合算法替换本文使用的自适应 Hough 直线检测算法, 以得到对比算法; 对于 MRM 算法, 需要将关键点检测模型整体替换为 MROD 算法的对应部分, 然后在此基础上进一步替换指针中心线拟合算法, 以得到对比算法。为使实验结果的呈现简洁美观, 本文在表格

表 6-11 读数对比实验结果

算法框架	指针信息提取	检出率	读数准确率	帧率 FPS
MRM	LS <sup>[79]</sup>	90.0	58.9	5.0
	PCA <sup>[23]</sup>	90.0	81.1	5.1
	RANSAC <sup>[81]</sup>	90.0	72.8	1.2
	KP	99.6	84.6	3.9
MROD	LS	98.5	71.8	12.2
	PCA	98.5	85.7	14.7
	RANSAC	98.5	84.6	1.7
	HOUGH	98.5	88.6	13.7

注: 测试集总数 280 张

中使用上述三种指针中心线拟合算法的英文缩写代表相应的对比算法，使用 KP 表示本文提出的基于关键点检测的指针信息提取算法，使用 HOUGH 表示基于本文提出的改进后 Hough 变换指针信息提取算法。

基于指针中心线拟合的读数对比实验结果如上表 6-11 所示。由实验结果可知，使用基于关键点检测指针信息提取算法的 MRM 算法在读数准确率方面相比使用 LS 算法提高 25.7%，相比使用 PCA 算法提高 3.6%，相比 RANSAC 算法提高 11.8%；使用自适应 Hough 直线检测指针信息提取算法的 MROD 算法读数准确率方面相比 LS 算法提高 16.8%，相比 PCA 算法提高 3.9%，相比 RANSAC 算法提高 4%。实验结果表明本文提出的两种算法都在各自的对比实验组内都取得了最优性能。

进一步分析实验结果可知，在使用相同指针中心线拟合算法的情况下，表中基于 MROD 算法框架实现的对比算法的读数准确率和检出率相比基于 MRM 算法框架实现的对比算法总是更高，帧率要更高。观察三种不同的指针中心线提取算法都可得出上述结论，说明 MROD 算法在读数计算方面相比 MRM 算法性能总体上更加优越，具备更强的可靠读数能力。

## 6.5 抗扰能力分析

为进一步验证 MRM 和 MROD 算法在不同干扰场景下的工作性能，本文进一步实施了干扰条件下的读数对比实验。选用高斯噪声、椒盐噪声、泊松噪声和低光照条件设定了 4 种干扰场景，并在每一种干扰场景下调节不同的干扰强度进行 6.4 节中的对比实验，重点研究不同种类以及不同强度的噪声干扰对算法检测精度的影响。

首先介绍本文选用的四种干扰场景。高斯噪声指概率密度函数服从高斯分布的一类噪声，产生原因一般包括拍摄图像时视场亮度不够均匀和图像传感器温度过高问题；椒盐噪声指图像中随机位置出现的黑白脉冲像素，因其形态类似黑胡椒和白盐粒而得名，会在数据传输、图像编解码等过程中产生，一般是环境中的电磁干扰和传感器内部时序错误导致的；泊松噪声是一种概率密度函数服从泊松分布的噪声，其存在的根本原因是光是由离散的光子构成的，光源和 CMOS 之间可能存在某些因素导致单个光子没有被 CMOS 接收或者某段时间内发射的光子特别多导致灰度值波动；低光照条件则是由于缺乏光照导致图像整体偏暗，从而影响图像效果。

高斯噪声干扰可通过高斯函数生成一张噪声图像，然后用噪声图像和原图叠加得到，高斯噪声的概率密度函数如公式(6-2)所示，通过控制公式中的标准差  $\sigma$  可以控制噪声的

强弱；椒盐噪声由在图像中随机生成黑色和白色像素点得到，通过控制像素点的数量可以控制噪声的强弱；泊松噪声干扰由泊松概率密度函数得到，通过噪声函数生成噪声图像然后和原图叠加得到，通过控制公式(6-3)中的参数 $\lambda$ 可以控制噪声的强弱；低光照条件可以通过调整图像的平均灰度得到。

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (6-2)$$

$$P(x=k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad (6-3)$$

本文通过给 MRD 数据集中的图像定量施加不同强度的干扰生成干扰图像测试集，干扰图像测试集的标注内容和 MRD 数据集一致，区别仅在于图像部分。本文具体的生成干扰所用参数设定如下表 6-12 所示，每次使用一个参数生成干扰图像测试集，然后用更新步长更新参数继续重复生成环节：

表 6-12 干扰类型及其参数

干扰类型	干扰参数	参数最小值	参数最大值	参数更新值
高斯噪声	$\sigma$	0.01	0.21	0.02
椒盐噪声	$\alpha$	0.01	0.21	0.02
泊松噪声	$\lambda$	1.00	51.00	5.00
光照强度	$\beta$	0.10	1.00	0.10

注：测试集总数 280 张

为直观展示各种干扰给图像带来的实际影响，本文对每种噪声选用三种不同强度，生成了图像受到干扰后效果，生成的受干扰图像效果如图 6-4 所示。

使用不同干扰类型，不同噪声强度的干扰图像测试集分别对 6.4 中基于 MRM 和 MROD 算法实现的每种算法进测试并记录算法的读数准确率，以干扰强度为自变量，读数准确率为因变量可以分别绘制每种算法在不同干扰条件下的折线图。图像中黄色折线代表本文提出的算法在 MRM 中为基于关键点检测的方法，在 MROD 中为自适应 Hough 直线检测的方法，绿色折线代表 RANSAC 算法，红色折线代表 PCA 算法，蓝色折线代表 LS 算法，黄色折线代表本文提出的算法。

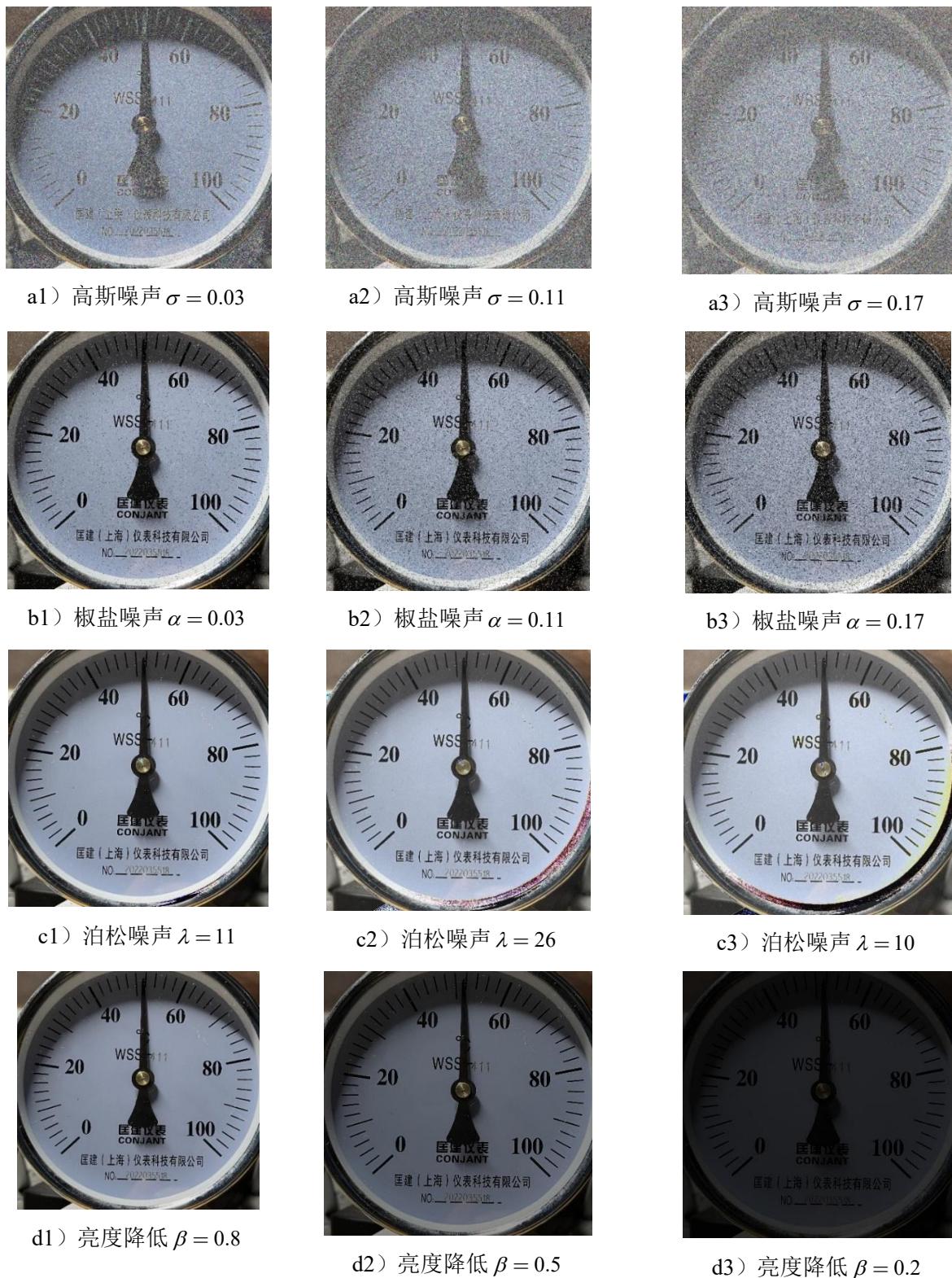


图 6-4 干扰图像效果图

在 MRM 算法框架的实验中指基于基于关键点检测的指针信息提取算法，在后文中为了叙述简洁，将其简称为 KP 算法；在 MROD 算法中指改进的 Hough 直线检测的指针信息提取算法，在后文中为了简洁叙述，将其简称为 HOUGH 算法。

如图 6-5 所示为 MRM 算法框架下的抗干扰性能对比实验结果。

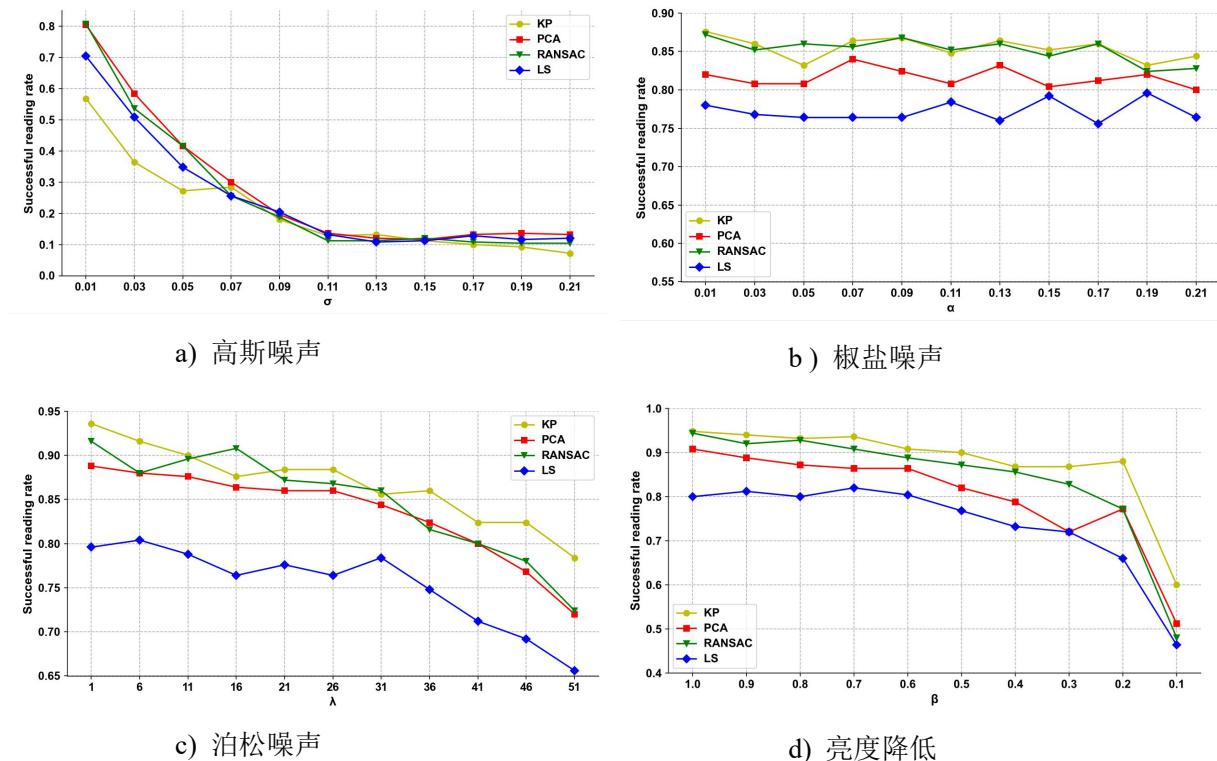


图 6-5 MRM 算法噪声实验结果

观察实验结果可知，KP 算法在高斯噪声实验中表现不佳，尤其是  $\sigma$  较小的部分，性能相对其他三种算法是最差的，随着干扰增强，这种差距在逐渐减小；在椒盐噪声实验中 KP 算法和 RANSAC 算法表现最佳，PCA 和 LS 算法表现较差；在泊松噪声实验中 KP 算法和 PCA、RANSAC 算法表现相差不大，都显著优于 LS 算法，KP 算法总体性能比 PCA 和 RANSAC 稍好；在低照度增强实验中，KP 算法的性能优于其他三种算法，随着照度逐渐降低，这种优势也在逐渐增大。总体来看，KP 算法对于高斯噪声干扰适应不佳，可能是由于关键点检测网络对于高斯噪声更加敏感的缘故，但是在其他干扰场景下相对对比算法仍然具有更好的性能表现，验证了该算法较好的抗干扰能力。

如图 6-6 所示为 MROD 算法框架下的抗干扰性能对比实验结果。

观察实验结果可知，HOUGH 算法在四种干扰场景下性能都要明显优于对比算法，尤其在面对椒盐噪声的干扰时和其他算法的性能拉开了较大差距；在另外三种干扰场景下，HOUGH 算法也在各个干扰参数强度上都表现出了优于其他对比算法的实验结果。

两组实验结果显示，在绝大多数情形下，随着干扰强度的增大算法的检测性能降低，这符合一般认知。在两组实验中性能和本文提出方法最接近相近的对比算法都是 RANSAC 算法，但是本文使用的 RANSAC 算法的迭代次数设定为 10，此时算法的计算

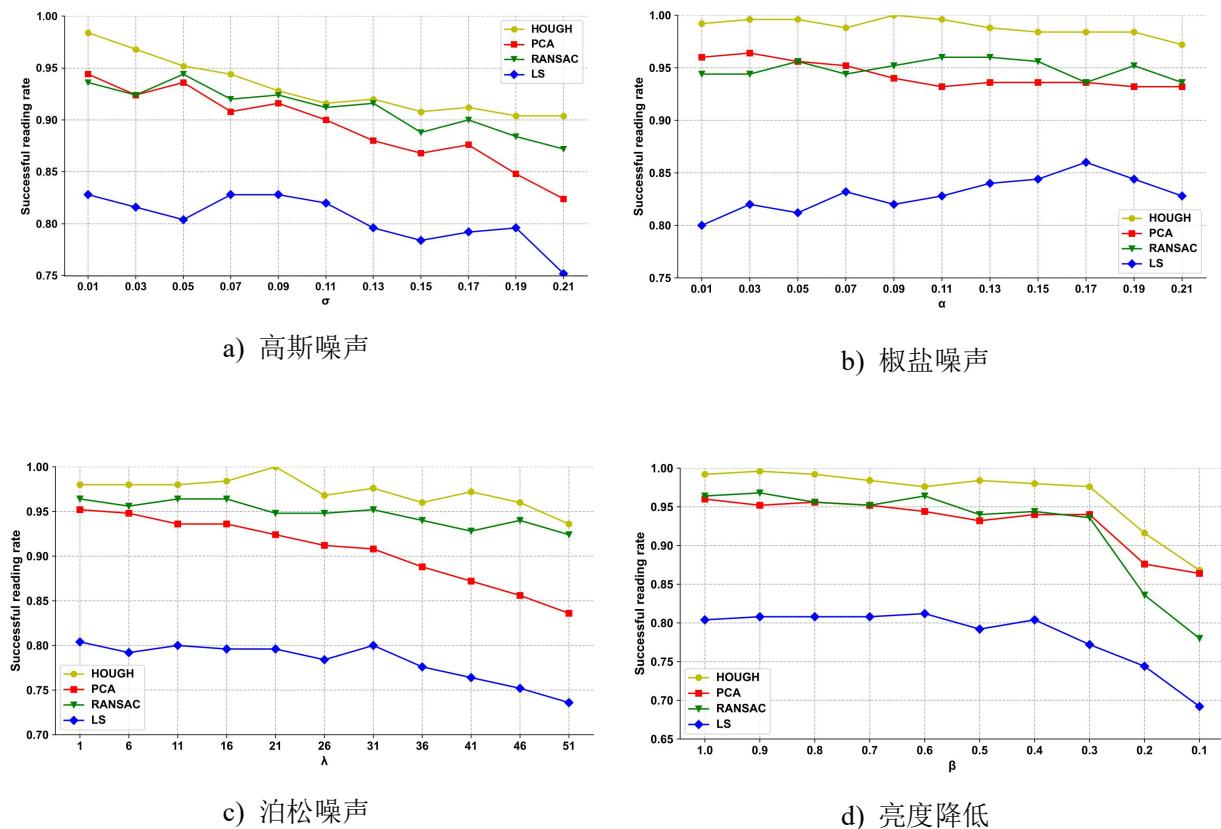


图 6-6 MROD 算法噪声实验结果

速度很慢，如果降低迭代次数则会导致识别成功率的下降，导致该算法并不具备很好的实用价值，而对比 PCA 和 LS 算法本文提出的算法均展现出了明显优势。众多论文中采用的最小二乘法在实验中效果最差，基本不具备面向未知仪表进行可靠读数的泛化能力。

实验结果表明，文本提出的 MRM 算法和 MROD 算法在对比实验中都表现出了优于同类文献中方法的性能，在检测精度和抗干扰能力方法都有着较好表现，在抗干扰实验中验证了两种算法的抗干扰能力，具备在复杂环境下正常工作的能力，其中 MROD 算法相比 MRM 算法表现除了更好的抗干扰能力，本文认为这是基于目标检测模型的方法在标度信息提取方面的能力由于基于 OCR 模型的方法所导致的。

## 6.6 智能巡检系统界面设计

综上所述，本文所述算法具备良好的泛化能力，因此也具备在生产生活中实际部署的潜力。实际的巡检机器人在巡检过程中，也会将实时的检测结果回传给生产单位的控制中心，供人工进行复核与应急处理。针对仪表自动读数识别算法的落地需求，为了进一步降低控制中心工作人员的使用门槛，本文提出一种智能仪表巡检系统，并基于 PyQt 框架完成了上述系统的用户管理界面设计，将本文所述算法的实现代码整合在一个 PC 端应用程序中。

### 6.6.1 界面设计

如图 6-7 所示为本文提出的智能仪表巡检系统的用户管理界面示意图，整个界面大致分为菜单栏、图像监视区域、待检测巡检图像列表、控制列表和信息显示区域这 5 个主要部分。该用户管理界面可以实现验证图像选择、对仪表读数识别的中间过程进行可视化显示和获取巡检图像信息的功能，巡检图像信息包括图像基本信息和读数相关信息等两部分，下面将展开对各个区域进行详细介绍。

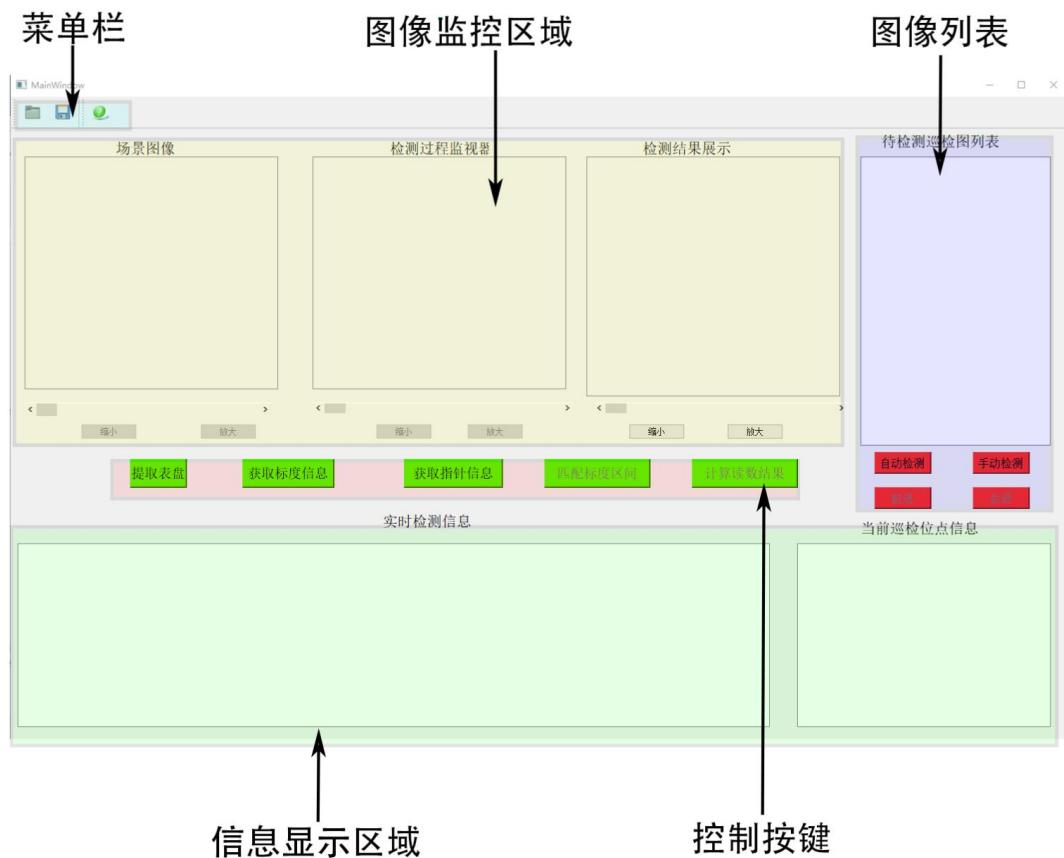


图 6-7 智能仪表巡检系统用户管理界面示意图

**菜单栏区域：**菜单栏主要包括三个菜单，分别是：“选择文件夹”、“保存结果”和“使用说明”。“选择文件夹”菜单可以让使用者自由选择图像回传链接的文件夹或历史数据文件夹，从而接收到实时回传或者先前保留的图像进行查看。“保存结果”按钮用于将检测得到的数据保存到 csv 格式的数据文件中。“使用说明”菜单可以让使用者获取软件使用的说明书以快速解决典型的问题，快速掌握使用技巧。

**图像监视区域：**用于控制中心的人员观察图像处理的可视化结果，分三个不同窗口分别展示不同的内容。一个场景图像窗口用于显示巡检直接得到的场景图像；一个检测结果展示窗口用于显示最终得到的仪表读数识别检测结果；一个检测过程监视器窗口用

于动态显示标度信息、指针信息等仪表读数过程中各个可视化图像。

**信息显示区域：**用于显示仪表巡检过程中的各种实时信息，具体细分为两个模块，当前巡检点位信息和实时检测信息两个部分。当前巡检点位信息部分获取并显示图像相关信息，包括图片录入名称、巡检时间记录、巡检地点记录、检测到仪表个数和识别异常仪表个数；实时检测信息用于显示仪表读数计算的相关信息，包括表盘编号、标度信息、指针信息和读数结果。

**待检测巡检图像列表：**用于控制中心的人员快速选择要进行识别的图像，当使用菜单栏中的“选择文件夹”菜单打开文件夹之后，其中的所有 jpg 格式的图像就都会出现在该列表中。该列表下预设了自动模式和手动模式两个红色模式选择按钮，在接入实际系统后能够让工作人员自行选择和切换模式，在手动模式中可以使用红色前进和后退按钮顺序进行检测识别。

**控制按键：**用于工作人员手动观察可视化的各个步骤，分为提取表盘，获取标度信息、获取指针信息、匹配标度区间和计算读数结果等，按下对应的按钮之后信息显示区域和图像显示区域就会出现对应的检测结果。

## 6.6.2 案例展示

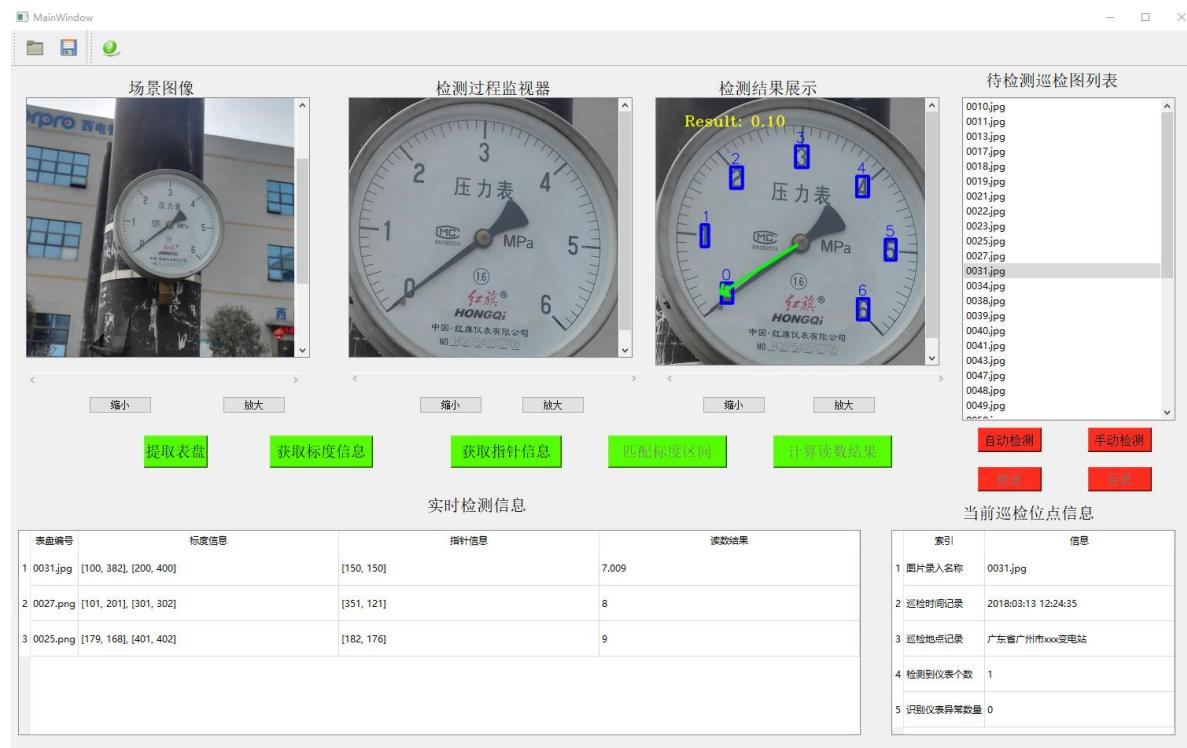


图 6-8 智能仪表巡检系统界面案例图

图 6-8 为一个实际运行中的智能仪表巡检系统用户管理界面截图。工作人员选择了图像 0031.jpg，然后一次按下绿色控制按键，场景图像窗口保持显示场景图像不变，检测过程监视器目前显示表盘图像，检测结果展示窗口展示了仪表的读数结果，图像中绘制了识别得到的标度和指针信息，同时标注了读数结果，下方的信息显示区域也成功得到的了当前巡检点位信息和实时检测信息。后续再进一步点击菜单栏中的“保存结果”按钮即可将检测得到的数据保存下来。

## 6.7 本章小结

本章基于上文提出的 MRM 算法和 MROD 算法利用自制数据集设计了充分的实验，对两种算法的读数识别性能进行了验证；针对本文提出的各项改进点进行了细致的消融实验；以指针信息提取环节为研究对象对比了本文提出的方法和多个主流方法之间的读数识别性能，同时进行了抗干扰性能的对比测试。实验结果显示，本文提出的方法具有良好的读数性能、具备面向未知仪表读数的强泛化能力，MRM 和 MROD 算法在测试集上的读数准确率分别可达 84.6% 和 88.6%，检测帧率可达 4FPS 和 12FPS。同时，对比同类方法本文算法具有显著的性能优势，面对各种典型干扰环境表现出了更好的抗干扰能力。最后基于 PyQt 开发出了巡检机器人指针式仪表智能读数识别系统的用户端管理界面，可以可视化展示仪表自动读数识别算法的各个中间步骤识别效果，同时可以获取图像的相关检测数据和算法回传读数计算数据，便于人工进行复核，降低技术的应用门槛。

## 结论与展望

本文围绕仪表全过程自动读数识别算法进行研究，旨在提出具有强泛化能力、高检测精度和较快检测速度的仪表自动读数识别算法，取得的研究成果和结论如下：

1) 综合当前仪表自动读数识别算法的研究背景，说明文本的研究价值和意义，并对仪表自动读数算法、表盘图像提取算法、指针信息提取算法和标度信息提取算法进行了文献综述，介绍了国内外研究现状。结合仪表自动读数识别的实际特点明确了本文的研究内容确定了技术路线和研究方案，绘制了本文的技术路线图。

2) 针对表盘图像提取问题，提出利用目标检测算法进行表盘检测和提取，选择 anchor-free 的 FCOS 模型作为基础模型并对其进行了轻量化的改进；提出了 FGCNet 轻量化主干网络在保持模型检测精度的同时大幅降低了参数量、提高了检测速度；针对 FCOS 模型的特征融合层和检测头进行了轻量化改进，最终得到 Lite-FCOS 轻量化目标检测模型，经过实验测试 mAP 达到 94.4%，帧率可达 56.2FPS，参数量仅为 4.2M；同时针对检测得到的表盘图像增强从图像去噪、清晰度增强和倾斜表盘检测三个角度提出了改进策略，重点介绍了基于 RealSR 模型的表盘图像清晰度增强方法。

3) 针对标度信息提取问题，提出了基于 OCR 模型和基于目标检测模型的两种方法；基于 OCR 模型的方法利用重新微调训练之后的 OCR 模型提取表盘图像中的所有字符信息，然后从中筛选分离出标度信息和仪表读数类型信息；基于目标检测模型的方法利用训练之后的 Lite-FCOS 模型提取表盘图像中的所有单个字符信息，然后利用本文提出的数字字符融合方法，包括跨类别 NMS、待融合字符组合和数字融合等算法，利用表盘中的单个数字字符正确还原出完整的标度数字，最终实现标度信息提取。

4) 针对指针信息提取问题，提出了基于 Hough 直线检测和基于关键点检测的两种方法；首先利用 Lite-FCOS 目标检测模型检测并提取指针图像。基于 Hough 直线检测的方法运用标度信息拟合指针中心点，运用自适应二值化和分段 Hough 直线检测提取指针中心线，进而计算指针中心线和指针图像的边界交点得到指针尖端点；基于关键点检测的方法直接进行微调训练之后的 HRNet 关键点检测模型从指针图像中提取指针中心点和指针尖端点坐标。

5) 针对仪表读数计算问题，提出了一种仪表读数基准标度区间的匹配方法，包括最近距离原则和防止错误匹配的策略，提出了一种基于扩展了标度区间有效范围的仪表读数计算公式；利用自制的 MRD 数据集对本文提出的两种仪表读数算法进行了包括读

数识别性能验证、消融实验、读数对比实验和抗扰能力分析实验等详细的实验，验证了算法具有良好的检测性能。最后基于本文所述算法设计并实现了一种智能巡检系统界面。

但是由于个人研究时间有限等因素，本文还存在一些不足之处，在现有工作的基础上，提出几点可供未来进一步开展研究的工作：

1) 本文关注到仪表倾斜校正的问题，但是并没有能够给出实质性的校正解决方案，目前停留在对大倾角表盘的检测阶段。这一问题对于进一步增强仪表读数识别的鲁棒性和泛化能力有着重要意义，值得进一步研究并得出更加通用的解决方案。

2) 标度盘信息的识别能力还有提升空间，当下不论是基于 OCR 模型的方法还是基于目标检测模型的方法在标度信息识别的能力上仍然还有提升空间，可以对相关模型和方法进一步进行优化以提高检测的效果。

3) 可以进一步拓展算法的适用范围，本文的研究对象限定在单指针单标度盘且刻度均匀的表盘，未来可以进一步探索和研究将仪表全过程自动读数识别算法推广到多指针、多标度表盘或者非均匀刻度仪表的应用中，从而进一步推动仪表自动读数识别技术的泛化能力和通用性能。

## 参考文献

- [1] Sablatnig R, Kropatsch W G. Application constraints in the design of an automatic reading device for analog display instruments; proceedings of the Proceedings of 1994 IEEE Workshop on Applications of Computer Vision, F 5-7 Dec. 1994, 1994 [C]. 205-212.
- [2] 黎明, 王厚枢. 飞机零位检测系统中的座舱仪表读数识别法[J].南京航空航天大学学报,1996(01):106-112.
- [3] 郭子海, 郭健, 谭志强. 用图象识别概念判读指针式仪表读数的研究[J].仪器仪表学报,1996(03):327-329.
- [4] 王涛, 王晓霞. 多尺度边缘检测在仪表读数识别中的应用 [J]. 工业控制计算机, 2006, (07): 58-59.
- [5] 陶冰洁, 韩佳乐, 李恩. 一种实用的指针式仪表读数识别方法[J].光电子工程,2011,38(04):145-150.
- [6] Ni T, Miao H, Wang L, et al. Multi-meter intelligent detection and recognition method under complex background; proceedings of the 2020 39th Chinese Control Conference (CCC), F, 2020 [C]. IEEE, 7135-7141.
- [7] 徐发兵, 吴怀宇, 陈志环, 等. 基于深度学习的指针式仪表检测与识别研究[J].高技术通讯,2019,29(12):1206-1215.
- [8] 郭月阳. 基于变电站巡检机器人的仪表读数识别研究 [D]; 华南理工大学, 2020.
- [9] Alegria E C, Serra A C. Automatic calibration of analog and digital measuring instruments using computer vision [J]. IEEE Transactions on Instrumentation and Measurement, 2000, 49(1): 94-99.
- [10] Hung M, Hsieh C. Automatic Pointer Meter Reading Based on Machine Vision; proceedings of the 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC), F 5-7 July 2019, 2019 [C]. 32-35.
- [11] Fan Z, Shi L, Xi C, et al. Real Time Power Equipment Meter Recognition Based on Deep Learning [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-15.
- [12] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation; proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, F, 2015 [C]. Springer, 234-241.

- [13] 刘葵. 基于深度学习的指针式仪表示数识别 [D]; 华中科技大学, 2017.
- [14] Zhang Z, Chen G, Li J, et al. The research on digit recognition algorithm for automatic meter reading system; proceedings of the 2010 8th World Congress on Intelligent Control and Automation, F, 2010 [C]. IEEE, 5399-5403.
- [15] Gao J, Guo L, Lv Y, et al. Research on Algorithm of Pointer Instrument Recognition and Reading Based on the Location of the Rotation Center of the Pointer; proceedings of the 2018 IEEE International Conference on Mechatronics and Automation (ICMA), F 5-8 Aug. 2018, 2018 [C]. 1405-1410.
- [16] Zhang L, Fang B, Zhao X, et al. Pointer-type meter automatic reading from complex environment based on visual saliency; proceedings of the 2016 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR), F 10-13 July 2016, 2016 [C]. 264-269.
- [17] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International journal of computer vision, 2004, 60: 91-110.
- [18] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features; proceedings of the European conference on computer vision, F, 2006 [C]. Springer, 404-417.
- [19] Mai X, Li W, Huang Y, et al. An Automatic Meter Reading Method Based on One-dimensional Measuring Curve Mapping; proceedings of the 2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE), F 24-27 Aug. 2018, 2018 [C]. 69-73.
- [20] 邢浩强, 杜志岐, 苏波. 变电站指针式仪表检测与识别方法[J]. 仪器仪表学报, 2017, 38(11):2813-2821.
- [21] Zhang X, Dang X, Lv Q, et al. A Pointer Meter Recognition Algorithm Based on Deep Learning; proceedings of the 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), F 24-26 April 2020, 2020 [C]. 283-287.
- [22] 王俊凯. 基于面部类化的仪表读数识别方法研究 [D]; 华中科技大学, 2019.
- [23] Zuo L, He P, Zhang C, et al. A robust approach to reading recognition of pointer meters based on improved mask-RCNN [J]. Neurocomputing, 2020, 388: 90-101.
- [24] 孙绍祖, 刘今越, 史宝军. 基于改进 PSF 参数估计的仪表模糊图像复原与识别[J]. 传感器与微系统, 2022, 41(09):109-112+120.
- [25] 胡东华, 闫坤, 辛文凯, 等. 基于轮廓的圆形仪表自动透视校正[J]. 电子测量与仪器学报:1-8.

- [26] 胡鑫, 欧阳华, 侯卓成, 等. 特征增强的 SSD 算法及在仪表读数识别上的应用[J]. 电子测量技术, 2021, 44(17):154-159.
- [27] 岳国义, 李宝树, 赵书涛. 智能型指针式仪表识别系统的研究[J]. 仪器仪表学报, 2003(S2):430-431.
- [28] Zhang H, Chen Q, Lei L. A YOLOv3-Based Industrial Instrument Classification and Reading Recognition Method [J]. Mobile Information Systems, 2022, 2022.
- [29] 金爱萍, 袁亮, 周德勤, 等. 基于 YOLOv5 和 U-net 的指针式仪表读数识别方法[J]. 仪表技术与传感器, 2022(11):29-33.
- [30] 李娜, 王军, 董兴法, 等. 基于改进 Hough 变换的指针式仪表识别方法[J]. 液晶与显示, 2021, 36(08):1196-1203.
- [31] 李宝, 程志全, 党岗, 等. 一种基于 RANSAC 的点云特征线提取算法[J]. 计算机工程与科学, 2013, 35(02):147-153.
- [32] He P, Zuo L, Zhang C, et al. A Value Recognition Algorithm for Pointer Meter Based on Improved Mask-RCNN; proceedings of the 2019 9th International Conference on Information Science and Technology (ICIST), F 2-5 Aug. 2019, 2019 [C]. 108-113.
- [33] Liu J, Wu H, Chen Z, et al. Automatic Identification Method of Pointer Meter under Complex Environment [Z]. 2020 12TH International Conference on Machine Learning and Computing. 2018: 276-282
- [34] 张雪飞, 黄山. 多类指针式仪表识别读数算法研究 [J]. 电测与仪表, 2020, 57(16): 147-152.
- [35] 莫文雄, 裴利强, 黄青丹, 等. 基于模版匹配和查表法的高精度指针式仪表自动检测系统研制[J]. 电测与仪表, 2017, 54(12):100-105.
- [36] 孙慧媛, 杨晓城, 蒋明峰, 等. 基于二维码匹配的指针式仪表读数识别方法[J]. 计算机系统应用, 2019, 28(03):208-214.
- [37] Gao H, Yi M, Yu J, et al. Character Segmentation-Based Coarse-Fine Approach for Automobile Dashboard Detection [J]. IEEE Transactions on Industrial Informatics, 2019, 15(10): 5413-5424.
- [38] 靳一丹. 基于机器视觉的模拟仪表自动识别与监控方法研究 [D]; 电子科技大学, 2021.
- [39] Meng X, Cai F, Wang J, et al. Research on Reading Recognition Method of Pointer Meters Based on Deep Learning Combined with Rotating Virtual Pointer; proceedings

- of the 2020 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT), F 13-15 Nov. 2020, 2020 [C]. 115-118.
- [40] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2016 [C]. 770-778.
- [41] Newell A, Yang K, Deng J. Stacked hourglass networks for human pose estimation; proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VIII 14, F, 2016 [C]. Springer, 483-499.
- [42] Tian Z, Shen C, Chen H, et al. Fcos: Fully convolutional one-stage object detection; proceedings of the Proceedings of the IEEE/CVF international conference on computer vision, F, 2019 [C]. 9627-9636.
- [43] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2015 [C]. 3431-3440.
- [44] Xie S, Girshick R, Dollár P, et al. Aggregated Residual Transformations for Deep Neural Networks; proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), F 21-26 July 2017, 2017 [C]. 5987-5995.
- [45] Lin T-Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2017 [C]. 2117-2125.
- [46] Lin T-Y, Goyal P, Girshick R, et al. Focal loss for dense object detection; proceedings of the Proceedings of the IEEE international conference on computer vision, F, 2017 [C]. 2980-2988.
- [47] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 39(6): 1137-1149.
- [48] Howard A G, Menglong Z, Bo C, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [arXiv] [J]. arXiv, 2017.
- [49] Cao Y, Xu J, Lin S, et al. Gcnet: Non-local networks meet squeeze-excitation networks and beyond; proceedings of the Proceedings of the IEEE/CVF international conference on computer vision workshops, F, 2019 [C]. 0-0.

- [50] Wang X, Girshick R, Gupta A, et al. Non-local neural networks; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2018 [C]. 7794-7803.
- [51] Hu J, Shen L, Sun G. Squeeze-and-excitation networks; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2018 [C]. 7132-7141.
- [52] Dai J, Qi H, Xiong Y, et al. Deformable convolutional networks; proceedings of the Proceedings of the IEEE international conference on computer vision, F, 2017 [C]. 764-773.
- [53] Zhu X, Hu H, Lin S, et al. Deformable convnets v2: More deformable, better results; proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, F, 2019 [C]. 9308-9316.
- [54] Wang W, Xie E, Li X, et al. Pyramid vision transformer: A versatile backbone for dense prediction without convolutions; proceedings of the Proceedings of the IEEE/CVF international conference on computer vision, F, 2021 [C]. 568-578.
- [55] Sun K, Xiao B, Liu D, et al. Deep high-resolution representation learning for human pose estimation; proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, F, 2019 [C]. 5693-5703.
- [56] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2018 [C]. 4510-4520.
- [57] Cai Z, Vasconcelos N. Cascade r-cnn: Delving into high quality object detection; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2018 [C]. 6154-6162.
- [58] Pang J, Chen K, Shi J, et al. Libra r-cnn: Towards balanced learning for object detection; proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, F, 2019 [C]. 821-830.
- [59] Kong T, Sun F, Liu H, et al. Foveabox: Beyond anchor-based object detection [J]. IEEE Transactions on Image Processing, 2020, 29: 7389-7398.
- [60] Zhang S, Chi C, Yao Y, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection; proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, F, 2020 [C]. 9759-9768.

- [61] Tomasi C, Manduchi R. Bilateral filtering for gray and color images; proceedings of the Sixth International Conference on Computer Vision (IEEE Cat No98CH36271), F 7-7 Jan. 1998, 1998 [C]. 839-846.
- [62] 唐艳秋, 潘泓, 朱亚平, 等. 图像超分辨率重建研究综述 [J]. 电子学报, 2020, 48(07): 1407-1420.
- [63] Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks [J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 38(2): 295-307.
- [64] Zhao H, Kong X, He J, et al. Efficient image super-resolution using pixel attention; proceedings of the Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16, F, 2020 [C]. Springer, 56-72.
- [65] Ji X, Cao Y, Tai Y, et al. Real-world super-resolution via kernel estimation and noise injection; proceedings of the proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, F, 2020 [C]. 466-467.
- [66] Wang X, Yu K, Wu S, et al. Esrgan: Enhanced super-resolution generative adversarial networks; proceedings of the Proceedings of the European conference on computer vision (ECCV) workshops, F, 2018 [C]. 0-0.
- [67] Tian Z, Huang W, He T, et al. Detecting text in natural image with connectionist text proposal network; proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VIII 14, F, 2016 [C]. Springer, 56-72.
- [68] Liao M, Shi B, Bai X, et al. Textboxes: A fast text detector with a single deep neural network; proceedings of the Proceedings of the AAAI conference on artificial intelligence, F, 2017 [C]. 4161-4167.
- [69] Wang W, Xie E, Li X, et al. Shape Robust Text Detection With Progressive Scale Expansion Network; proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), F 15-20 June 2019, 2019 [C]. 9328-9337.
- [70] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv:14091556, 2014.
- [71] Fang Y, Dai Y, He G, et al. A Mask RCNN based Automatic Reading Method for Pointer Meter; proceedings of the 2019 Chinese Control Conference (CCC), F 27-30 July 2019, 2019 [C]. 8466-8471.

- [72] Toshev A, Szegedy C. DeepPose: Human Pose Estimation via Deep Neural Networks; proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, F 23-28 June 2014, 2014 [C]. 1653-1660.
- [73] Tompson J J, Jain A, Lecun Y, et al. Joint training of a convolutional network and a graphical model for human pose estimation [J]. Advances in neural information processing systems, 2014, 27.
- [74] Papandreou G, Zhu T, Kanazawa N, et al. Towards accurate multi-person pose estimation in the wild; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2017 [C]. 4903-4911.
- [75] Cao Z, Hidalgo G, Simon T, et al. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields [J]. IEEE transactions on pattern analysis and machine intelligence, 2021, 43(1): 172-186.
- [76] Pishchulin L, Insafutdinov E, Tang S, et al. Deepcut: Joint subset partition and labeling for multi person pose estimation; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2016 [C]. 4929-4937.
- [77] Chen Y, Wang Z, Peng Y, et al. Cascaded pyramid network for multi-person pose estimation; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2018 [C]. 7103-7112.
- [78] 上海工业自动化仪表研究所. 工业过程测量和控制用检测仪表和显示仪表精确度等级: [S]. 中华人民共和国国家质量监督检验检疫总局;中国国家标准化管理委员会, 2008: 8.
- [79] York D. Least-squares fitting of a straight line [J]. Canadian Journal of Physics, 1966, 44(5): 1079-1086.
- [80] 何配林. 基于深度学习的工业仪表识别读数算法研究及应用 [D]; 电子科技大学, 2020.
- [81] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6): 381-395.