

基于计算机视觉的数据机房工业仪表识别研究

蔡登江

(中国海洋石油集团有限公司 信息技术中心, 北京 100010)

摘要: 数据机房作为保障数据安全的重要设施,其内部的各种仪器仪表对于监控数据机房的安全至关重要。然而,传统的仪表识别依赖人工操作,易出错且效率低下。为了解决数据机房中工业仪表在复杂场景下准确率低的问题,研究提出了一种结合渐进式注意力机制与目标检测网络的仪表识别算法。此外,研究还引入了幽灵模块以增强特征表达能力,提高模型的识别准确率。实验结果表明,在不同复杂场景下的仪表识别任务中该模型均表现出较高的准确率。例如,在光照不均或有遮挡的场景 A 中,准确率达到 92.15%,召回率为 89.47%, F_1 分数为 90.78%,处理时间为 34.21 ms;在背景复杂或多仪表密集的场景 B 中,准确率提升至 93.24%,召回率为 90.85%, F_1 分数达到 92.02%,处理时间为 35.17 ms。因此研究所提出的基于计算机视觉的数据机房工业仪表识别模型在复杂场景下的工业仪表识别任务中取得了显著成果,显示出较高的实用价值。

关键词: 计算机视觉;YOLOv5;仪表识别;工业机房

中图分类号: TP 998

文章编号: 1000-0682(2025)02-0099-06

文献标识码: A

DOI: 10.19950/j.cnki.CN61-1121/TH.2025.02.018

Research on industrial instrument recognition in data room based on computer vision

CAI Dengjiang

(Information Technology Center, China National Offshore Oil Corporation, Beijing 100010, China)

Abstract: As an important facility for ensuring data security, the various instruments and meters inside the data center are crucial for monitoring the security of the data center. However, traditional instrument recognition relies on manual operation, which is prone to errors and inefficient. In order to solve the problem of low accuracy of industrial instruments in complex scenarios in data centers, a instrument recognition algorithm combining progressive attention mechanism and object detection network is proposed. In addition, the study also introduced ghost modules to enhance feature expression ability and improve the recognition accuracy of the model. The experimental results show that the model exhibits high accuracy in instrument recognition tasks in different complex scenarios. For example, in scene A with uneven lighting and occlusion, the accuracy reached 92.15%, the recall rate was 89.47%, the F_1 score was 90.78%, and the processing time was 34.21 ms. In scene B with complex background and dense instruments, the accuracy is improved to 93.24%, the recall rate is 90.85%, the F_1 score reaches 92.02%, and the processing time is 35.17 ms. Therefore, the computer vision based industrial instrument recognition model proposed by the research institute has achieved significant results in industrial instrument recognition tasks in complex scenarios, demonstrating high practical value.

Keywords: computer vision; YOLOv5; instrument recognition; industrial computer room

0 引言

随着经济与科技的快速发展,中国的工业企业的智能化水平越来越高,数据的安全性也变得愈发重要。因此,许许多多的工厂建立起了自己的数据机房,以保证数据的安全性,但数据机房中往往存在

收稿日期:2024-08-06

第一作者:蔡登江(1980—),男,湖北孝感人,本科,信息通信工程高级工程师,研究方向为计算机网络传输及交换,数据中心基础设施建设等。E-mail:15533915062@163.com

着大小各式各样的仪器仪表,这些仪器仪表对于数据机房的安全起着重要的监控作用^[1-2]。但传统的仪表识别仅简单的靠人力进行,难免出现许多错误,因此越来越多的研究者开始关注仪表的识别问题^[3-4]。杨典针对工业仪表识别自动化智能化的问题,提出了一种基于 OpenCV 图像处理与深度学习结合的方法。对于数字式仪表,通过图像处理分割数字后,运用卷积神经网络进行识别。结果显示,该方法对实际拍摄仪表照片的识别准确度高、速度快,在工业仪表识别方面具有一定的应用价值^[5]。WENTING ZHAO 等人针对智能识别技术在仪表故障评估中的应用问题,提出结合细粒度图像特征与评估指标的方法,并使用注意力机制的细粒度图像识别技术。结果显示,基本仪器类型可有效识别,提供合理的评估建议^[6]。但相关方法在面对复杂场景的时候准确率显著下降。鉴于此研究创新性地将渐进式注意力机制与你只看一次版本 5 (You Only Look Once version 5, YOLOv5) 网络相结合,构建出 IYOLOv5 模型,用于数据机房工业仪表的识别。通过逐步计算和更新场景特征,该机制能够捕获场景中各个组成部分的关系,有效提升图像识别的准确性和鲁棒性,特别是在光照条件较差的情况下。研究旨在解决数据机房工业仪表识别在复杂场景下准确率降低的问题,提供一种高效、准确的仪表识别方法,以满足工业应用的需求。

1 结合渐进式注意力机制与 YOLOv5 网络的仪表识别

1.1 渐进式注意力机制的构建

数据机房中的工业仪表识别任务通常需要准确地检测和读取各种仪表上的数值,这对视觉识别系统提出了高要求^[7-8]。YOLOv5 作为一种先进的深度学习目标检测模型,能够快速且准确地识别图像中的目标物体,包括数据机房中的各类工业仪表。此外,YOLOv5 易于训练与部署,支持多种平台和设备,具有良好的可扩展性,非常适合用于工业场景下的实时仪表识别任务^[9-10]。但由于计算机机房通常情况下照明情况不是很理想,在暗光条件下,YOLOv5 出现漏检误检的问题,针对以上问题,研究在基础 YOLOv5 网络的基础上结合了一种新的渐进式注意力机制^[11-13]。首先,将场景特征和对象特征映射到同一个特征空间,假定 $F_s(\cdot)$ 是场景特征的映射函数, $F_o(\cdot)$ 是对象特征的映射函数,那么其映射函数表达式如式(1)所示:

$$\begin{cases} F_s(\cdot) = \ell_2(W_s(\cdot)) \\ F_o(\cdot) = \ell_2(W_o(\cdot)) \end{cases} \quad (1)$$

式中: W_s 是场景特征的可学习权重; W_o 是对象特征的可学习权重。

通过这个函数将 2 种特征统一的处理和分析;这种机制的核心思想是在每次计算得到一个对象的注意力权重后,根据这个权重来更新嵌入的场景特征,从而逐步构建整个场景的表示,渐进式注意力机制的结构示意图如图 1 所示。

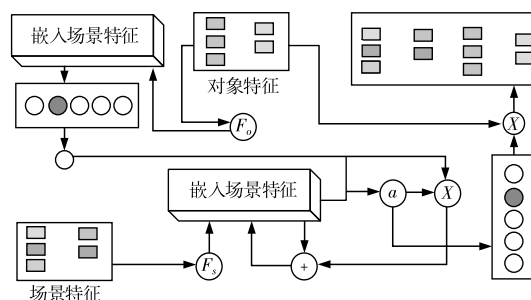


图 1 渐进式注意力机制的结构示意图

图 1 展示了渐进式注意力机制的结构示意图,这是一种用于处理复杂场景的机制,用于光照条件较差的情况下的图像识别问题。该机制的核心是通过逐步计算和更新场景特征来捕获场景中各个组成部分的关系,确保场景特征的时效性和准确性。首先,场景特征被嵌入到一个具有 512 维的向量空间中,每个维度代表了场景中的一个特定属性或特征。然后,每个对象的特征被提取出来,并计算相应的注意力权重,使用一个线性变换层,该层将对象特征映射到一个 128 维的向量,然后通过简洁的修正线性单元(Rectified Linear Unit, ReLU)激活函数来计算注意力权重。这些权重用于调整嵌入场景特征,形成新的场景特征表示。此过程反复进行,每次引入一个新的对象,直至所有对象都被考虑进去。如式(2)所示:

$$F_s = \sum_{i=1}^N \sum_{j=1}^N \alpha_{ij} (W_v x_j + b_v) \quad (2)$$

式中: W_v 和 b_v 是值(value)向量的线性变换参数。新的场景特征表示 F_s 是通过将所有对象的值向量(经过注意力权重加权)求和来得到的。 α_{ij} 表示对象 i 对对象 j 的注意力权重。

最终生成的场景特征包含了所有对象的重要信息和相互影响。利用线性变换与激活函数计算注意力权重,实现对象特征的加权处理,假定 o_i 表示第 i 个对象,那么对整体对象进行基于注意力权重的特征计算公式如式(3)所示:

$$f_{obj} = \sum_{i=1}^n \alpha_i \times o_i \quad (3)$$

式中: f_{obj} 表示整体对象权重特征; α_i 表示第 i 个对象的注意力权重。

加权后的对象特征再次与嵌入场景特征融合, 生成包含丰富场景与对象交互信息的输出特征。此外, 注意权重也参与到场景特征的更新过程中, 进一步优化了场景表示的质量, 如式(4)所示:

$$\begin{aligned} q_i &= W_q x_i + b_q \\ k_j &= W_k x_j + b_k \\ \alpha_{ij} &= \frac{\exp(q_i^T k_j)}{\sum_{m=1}^N \exp(q_i^T k_m)} \end{aligned} \quad (4)$$

式中: q_i 和 k_j 分别是对象 i 和对象 j 的查询(query)和键(key)向量, 它们通过线性变换(由权重矩阵 W_q 和 W_k 以及偏置向量 b_q 和 b_k 定义)从对象特征中提取出来。

通过这种方式, 渐进式注意力机制能够在光照条件较差的情况下, 有效提升图像识别的准确性和鲁棒性, 克服环境因素带来的挑战。

1.2 IYOLOv5 总体设计与 Ghost 模块

为了提升传统 YOLOv5 网络在复杂场景下的识别准确率, 研究设计出了一种渐进注意力机制, 并将其与 YOLOv5 结合构建出一种新的 IYOLOv5 结构, 模型核心由 Backbone、Neck 及 Prediction 以及注意力机制模块构成^[13]。Backbone 部分集成了 CSP1、CBL 及 Focus 等高效模块, 通过精细设计的网络结构有效提取图像特征。其中, CSP1 模块通过数据分割策略提升计算效率, CBL 模块则集成卷积、批量归一化与激活函数, 增强特征表达能力。Focus 模块则通过切片操作减少计算负担, 同时扩大网络视野。Neck 部分通过 SPP(空间金字塔池化)与 CBL 模块深度融合多尺度特征, 确保模型既能捕捉细节信息, 又能感知全局结构。渐进注意力机制模块的引入是 IYOLOv5 的关键模块。渐进注意力机制模块通过逐步计算和更新场景特征, 捕获场景中各个组成部分的关系。最终, Prediction 模块利用精细调整的特征图进行预测, 通过多个 Conv 与 CBL 模块的组合, 输出高精度的识别结果, IYOLOv5 的结构如图 2 所示。

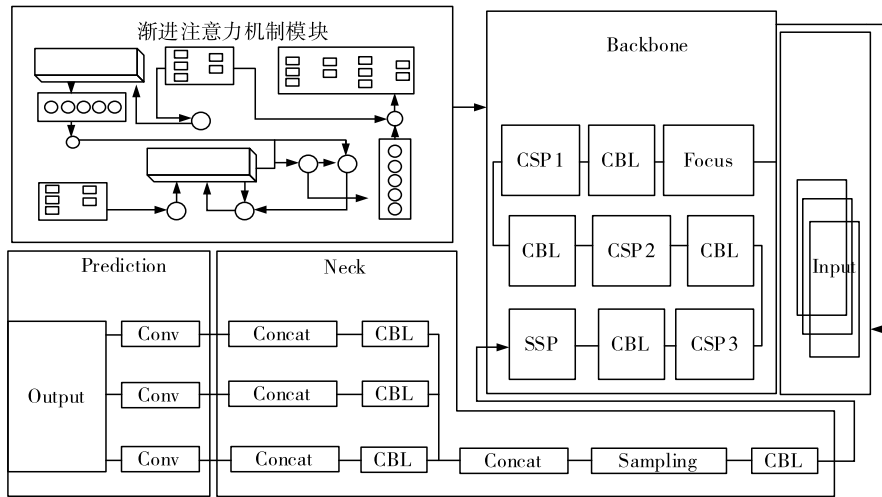


图 2 IYOLOv5 的结构示意图

IYOLOv5 模型虽然在目标检测任务中表现出色, 但其核心特征提取模块采用的卷积神经网络结构在某些方面存在局限性。特别是在处理高维度特征映射时, 卷积神经网络可能会遇到梯度消失或梯度爆炸的问题, 这限制了网络的深层学习能力和对复杂特征的表达能力^[15]。为了解决这一问题, 研究引入了 Ghost 模块来替换 IYOLOv5 网络中的 CNN 结构。Ghost 模块通过特征重用机制, 在保持计算效率的同时, 增强网络对特征的表达能力。它通过将一个卷积层分解为 2 个部分来实现。首先, 输入特

征图被分为 2 个子集, 1 个主干路径和 1 个附属路径。主干路径执行正常的卷积操作, 而附属路径则是一个轻量级的卷积操作, 通常具有较少的通道数, 假定 f 表示的是卷积核的大小, 那么常规的卷积计算公式如式(5)所示:

$$Y = X * f + b \quad (5)$$

式中: b 表示的是偏移量。

为了获取手势特征图, 还需要将其卷积计算公式进行线性变换, 假定 \bar{y}_i 是原始图中的第 i 个特征, 那么其计算公式如式(6)所示:

$$y_{ij} = \varphi_{ij}(\bar{y}_i) \quad (6)$$

式中: φ_{ij} 为线性变换。

这些相似性被视为一种有价值的信息,可以被进一步利用以增强模型的性能,Ghost 模块结构如图 3 所示。

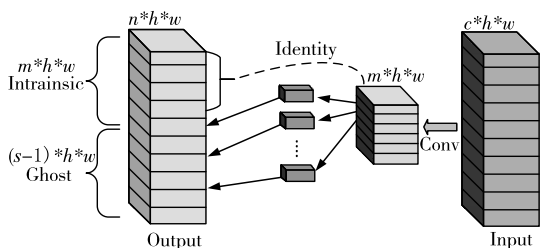


图 3 幽灵模块结构示意图

与传统的卷积神经网络相比,GhostNet 模型的核心创新之处在于其引入了独特的 Ghost 模块。该模块通过一种廉价的操作机制,能够有效地生成大量的幻影特征图(ghost maps)。具体而言,在 Ghost 模块中,首先使用较少数量的卷积核(例如,使用 m 组 $k \times k$ 的卷积核,相较于传统方法的 n 组卷积核, $m < n$)与输入进行卷积,生成本征图(intrinsic

maps)。随后,这些本征图经过一个线性变换 Φ ,进一步产生幻影特征图 G ,如式(7)所示:

$$G = \Phi(Y) = \text{Conv}(Y, W_{\text{ghost}}) \quad (7)$$

式中: W_{ghost} 表示生成幻影特征图的卷积核权重; Y 表示生成本征图。

将本征图与幻影特征图合并后作为该模块的输出。此外,GhostNet 利用深度分离卷积从生成的特征图中提取出冗余特征,进一步丰富了特征表示。最后,将这 2 步获取的特征图进行拼接(concat),并送入后续的网络环节进行处理。

2 IYOLOv5 数据机房工业仪表识别应用性分析

为了解决数据机房工业仪表识别在复杂场景下准确率低的问题,研究设计出一种新的 IYOLOv5 识别算法,为了验证该算法的在实际场景中的应用效果,研究在 Windows10 平台上利用 Python 语言实现了该算法,并在实际的数据机房中进行了实验,该数据机房的基本布局如图 4 所示。

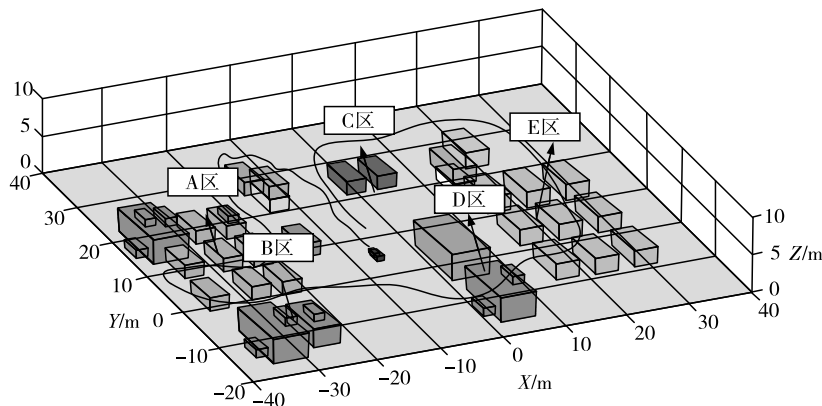


图 4 数据机房的基本布局图

从图 1 中可以看出,该数据机房的基本布局主要包含了 A,B,C,D 和 E 五个区域,其中 A 区光照不均,有遮挡,B 区背景复杂,多仪表密集,C 区仪表尺寸不一,角度多变。D 区域光照变化大,有阴影,

E 区域仪表部分损坏,污渍干扰,研究在这些区域进行了基于 IYOLOv5 模型的数据机房工业仪表识别系统应用性实验,以准确率、召回率、 F_1 分数和处理时间为评价指标。实验结果如表 1 所示。

表 1 基于 IYOLOv5 模型的数据机房工业仪表识别系统应用性实验结果

实验区域	场景描述	准确率/%	召回率/%	F_1 分数	处理时间/ms
A	光照不均,有遮挡	92.15	89.47	90.78	34.21
B	背景复杂,多仪表密集	93.24	90.85	92.02	35.17
C	仪表尺寸不一,角度多变	91.89	89.12	90.49	33.94
D	光照变化大,有阴影	92.73	90.36	91.53	34.68
E	仪表部分损坏,污渍干扰	90.98	88.65	89.79	33.72

从表 1 中可以看出,该系统在不同场景下的表

现均较为出色。在光照不均、有遮挡的场景 A 中,

系统准确率达到 92.15%, 召回率为 89.47%, F_1 分数为 90.78%, 处理时间为 34.21 ms。在背景复杂、多仪表密集的场景 B 中, 系统表现更佳, 准确率提升至 93.24%, 召回率为 90.85%, F_1 分数达到 92.02%, 处理时间略有增加, 为 35.17 ms。在仪表尺寸不一、角度多变的场景 C 中, 系统准确率为 91.89%, 召回率为 89.12%, F_1 分数为 90.49%, 处理时间为 33.94 ms。对于光照变化大、有阴影的场景 D, 系统准确率为 92.73%, 召回率为 90.36%, F_1 分数为 91.53%, 处理时间为 34.68 ms。即使在仪表部分损坏、有污渍干扰的场景 E 中, 系统仍保持

了较高的准确率(90.98%), 召回率为 88.65%, F_1 分数为 89.79%, 处理时间为 33.72 ms。总体来看, IYOLOv5 模型在不同复杂场景下的仪表识别任务中均表现出较高的准确率和效率。研究在这几个区域分布识别出指针式仪表、数字显示仪表、流量计、状态指示器四种仪表, 为了更好地表示识别结果, 研究引入了文献^[15]提出的智能机器人的变电站仪表识别系统(Intelligent Robot based Substation Instrument Recognition System, IR-SIRS)目标识别模型进行对比, 采用混淆矩阵来对比 2 个模型的分类性能, 结果如图 5 所示。

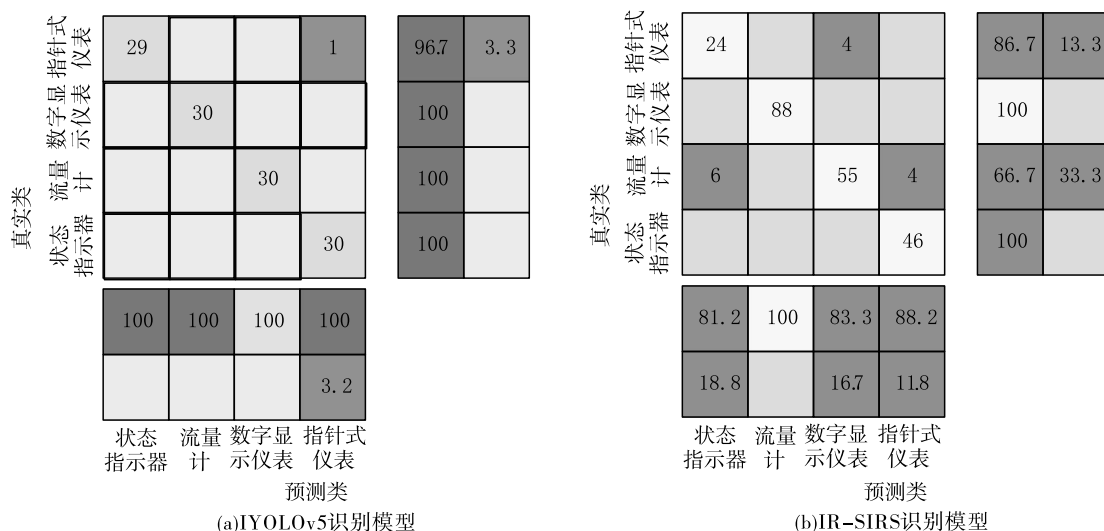


图 5 2 种不同模型的混淆矩阵情况

从图 5(a)中可以看出,在 IYOLOv5 模型中,大部分的指针式仪表和数字显示仪表都得到了正确的识别,分别有 96.7% 和 100% 的识别率。从图 5(b)中可以看出,在 IR-SIRS 模型中,除了数字显示仪表的识别率仍然维持在 100%,其他三类仪表的识别率都有所提高,特别是流量计达到了 88.2%。这主要是因为渐进式注意力机制通过逐步计算和更新场景特征来捕获场景中各个组成部分的关系。这种机制能有效提升图像识别的准确性和鲁棒性。通过将场景特征和对象特征映射到同一特征空间,并通过学习权重分配注意力,使得模型能够更准确地识别流量计和其他类型的仪表。

3 结论

研究提出了一种结合渐进式注意力机制与 YOLOv5 网络的仪表识别算法(IYOLOv5),旨在解决数据机房工业仪表识别在复杂场景下的准确率低的问题。通过在不同复杂场景下的实验验证,IY-

OLOv5 模型展现了优秀的识别性能。在光照不均、有遮挡的场景 A 中,系统准确率达到 92.15%,召回率为 89.47%, F_1 分数为 90.78%, 处理时间为 34.21 ms;在背景复杂、多仪表密集的场景 B 中,准确率提升至 93.24%, 召回率为 90.85%, F_1 分数达到 92.02%, 处理时间为 35.17 ms;在仪表尺寸不一、角度多变的场景 C 中,准确率为 91.89%, 召回率为 89.12%, F_1 分数为 90.49%, 处理时间为 33.94 ms;对于光照变化大、有阴影的场景 D,准确率为 92.73%, 召回率为 90.36%, F_1 分数为 91.53%, 处理时间为 34.68 ms;即使在仪表部分损坏、有污渍干扰的场景 E 中,系统仍保持较高准确率(90.98%), 召回率为 88.65%, F_1 分数为 89.79%, 处理时间为 33.72 ms。与文献[15]提出的 IR-SIRS 目标识别模型进行对比,IYOLOv5 模型在指针式仪表的识别率上达到了 96.7%, 数字显示仪表的识别率为 100%, 流量计的识别率为 88.2%。这些结果显示,IYOLOv5 模型在复杂环境下仍能保持较

高的识别准确率和处理速度,尤其在处理光照条件不佳和仪表类型多样化的场景时表现出色。综上所述, YOLOv5 在复杂场景下的工业仪表识别任务中取得了显著成果,展现出较高的实用价值。尽管 YOLOv5 模型在多个场景下表现出色,但在处理仪表损坏严重或存在重度遮挡的情况下,其识别准确率仍有待提高。

参考文献:

- [1] BOKAI ZHANG, DARRICK STURGEON, ARJUN RAVI SHANKAR, et al. Surgical instrument recognition for instrument usage documentation and surgical video library indexing[J]. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2023, 11(4): 1064 – 1072.
- [2] 陈开峰,俞伟聪,唐雁文,等. 基于轻量化深度学习网络的数字仪表识别[J]. 科学技术与工程, 2023, 23(2): 674 – 680.
- [3] KOSTIUCHIK GEORGII, SHARAN LALITH, MAYER BENEDIKT, et al. Surgical phase and instrument recognition: how to identify appropriate dataset splits[J]. International journal of computer assisted radiology and surgery, 2024, 19(4): 699 – 711.
- [4] 孙辉,李晓峰,程远方,等. 基于改进联合滤波与 CNN 模型的仪表识别算法[J]. 制造业自动化, 2023, 45(10): 166 – 172.
- [5] 杨典,李小燕,刘培焱,等. 基于 OpenCV 的变电站仪表识别方法研究[J]. 自动化与仪表, 2022, 37(4): 75 – 80.
- [6] ZHAO WENTING, WANG SHIGANG, ZHAO YAN, et al. A study on attention – based fine – grained image recognition: Towards musical instrument performing hand shape assessment[J]. Engineering Applications of Artificial Intelligence: The International Journal of Intelligent Real – Time Automation, 2024, 129(Mar.): 107512. 1 – 107512. 9.
- [7] 郝琨,韩冰,李志圣,等. 基于投影阈值分割和数字序列校正的高噪声数字仪表图像识别方法[J]. 电子科技大学学报, 2023, 52(5): 728 – 738.
- [8] 张森,万吉林,王慧芳,等. 基于注意力机制的卷积神经网络指针式仪表图像读数识别方法[J]. 电力自动化设备, 2022, 42(4): 218 – 224.
- [9] 钱玉宝,王紫涵,邱腾煌. 指针式仪表读数识别的研究现状与发展[J]. 电子测量技术, 2024, 47(8): 110 – 119.
- [10] 张琳. 基于传感器的工业仪表信号采集与处理技术研究[J]. 工业仪表与自动化装置, 2024(3): 111 – 116.
- [11] 赵伟达,陈海文,郭陆阳,等. 基于 YOLO – E 与改进 OCRNet 图像分割的变电站仪表读数自适应识别方法[J]. 电力建设, 2023, 44(11): 75 – 85.
- [12] 高志国,袁牧,童旸,等. 基于 PCBAM – YOLOv5 网络的变电站表计缺陷检测[J]. 电子设计工程, 2023, 31(4): 77 – 81.
- [13] PENG G, DU B, LI Z, et al. Machine vision – based, digital display instrument positioning and recognition[J]. 2022, 29(2): 230 – 243.
- [14] 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.
- [15] 孙顺远,魏志涛. 基于刻度轮廓拟合的指针式仪表自动识别方法[J]. 仪表技术与传感器, 2022, 2(8): 51 – 57.
- [16] 刘培焱,叶尔扎提·努尔德别克,孙敏,等. 基于智能机器人的变电站仪表识别系统[J]. 自动化与仪表, 2022, 37(12): 28 – 33.

(上接第 63 页)

5 结论

通过对基于 STM32 和 PyQt5 的分子泵监控系统的设计与实现,为高端分析仪器的小型化分子泵控制提供了一种有效的解决方案。该系统在 STM32 驱动控制基础上,结合串口 Modbus 通信、PyQt5、SQLite3 的便捷化、轻量化开发优势,实现了分子泵的实时监控及数据管理功能,为高端分析仪器用小型化分子泵的国产化应用奠定基础。

此外,该文所设计的分子泵监控系统还具有较

好的可扩展性和灵活性,可根据实际需求进行功能扩展和升级,未来,随着物联网、大数据等技术的不断发展,该系统可进一步与这些先进技术融合,实现更加智能化、自动化的分子泵监控与管理,为高端分析仪器的应用和发展提供更加全面的技术支持。

参考文献:

- [1] 李天柱,高皓天,王亚东. 高端科学仪器的创新特性与产业发展思路[J]. 科技和产业, 2021, 21(2): 109 – 114.
- [2] 卢文俊,冷杉,杨建军. 基于 Modbus 协议的控制器远程监控系统[J]. 电力自动化设备, 2003(06): 54 – 56.