

基于改进 YOLOv8n 的数字式仪表读数识别

李笑笑¹, 焦良葆^{1,2}, 顾嘉炜¹, 陈治锐¹, 孟琳¹

(1. 南京工程学院 计算机工程学院, 南京 211167;

2. 江苏省智能感知技术与装备工程研究中心, 南京 211167)

摘要: 针对数字式仪表读数中人工抄表误差大和识别准确率低的问题, 提出一种基于改进 YOLOv8n 的数字式仪表自动识别算法; 在网络结构中, 将 SCConv 模块嵌入主干网络的 C2f 结构, 以增强模型在空间与通道维度的特征提取能力, 提高对细粒度目标和边缘细节的感知效果; 使用 Focal Modulation 结构替换原有的 SPPF 模块, 通过分层上下文化与门控聚合机制实现局部与全局特征的高效融合, 从而提升检测精度与特征表达能力; 在损失函数设计方面, 将 CIoU 损失替换为 EIou 损失, 以优化边界框回归的收敛速度与定位精度; 实验结果表明, 该方法在自建仪表数据集上的 $mAP@0.5$ 较原始 YOLOv8n 模型提升 1.5%, 其中小数点类别精度提升 14.3%; 结果验证了所提算法在数字式仪表读数任务中的有效性与优越性。

关键词: 目标检测; 仪表读数; 深度学习; 损失函数

Digital Meter Reading Recognition Based on Improved YOLOv8n

LI Xiaoxiao¹, JIAO Liangbao^{1,2}, GU Jiawei¹, CHEN Zhirui¹, MENG Lin¹

(1. School of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, China;

2. Jiangsu Intelligent Perception Technology and Equipment Engineering Research Center, Nanjing 211167, China)

Abstract: To address the problems of large manual reading errors and low recognition accuracy in digital meter reading, this paper proposes an improved YOLOv8n-based automatic recognition algorithm for digital instruments. In the network structure, the SCConv module is embedded into the C2f blocks of the backbone to enhance feature extraction capability in both spatial and channel dimensions, thereby improving the perception of fine-grained targets and edge details. Meanwhile, the original SPPF module is replaced with the Focal Modulation structure, which efficiently fuses local and global features through hierarchical contextualization and gated aggregation mechanisms, thus enhancing its detection accuracy and feature representation capability. In terms of the loss function, the CIoU loss is replaced by the EIou loss to accelerate the convergence of bounding box regression and improve its localization precision. Experimental results show that compared with the original YOLOv8n model, the proposed method improves the $mAP@0.5$ by 1.5% on a self-built meter dataset, with the accuracy of decimal point class increasing by 14.3%, verifying the effectiveness and superiority of the proposed algorithm in digital meter reading tasks.

Keywords: object detection; meter reading; deep learning; loss function

0 引言

数字式仪表凭借其读数精度高、数据稳定性强、抗干扰能力优的特性, 已广泛应用于工业控制、电力系统

运维、交通设备监控、医疗设备参数显示等关键领域^[1]。在这些场景中, 数字式仪表的读数结果直接关联设备运行状态评估、故障预警时效性及生产操作安全性, 因此, 数字式仪表自动化读数是推动工业智能化监

收稿日期:2025-10-14; 修回日期:2025-11-24。

基金项目:江苏省产学研合作项目(BY20230656)。

作者简介:李笑笑(2001-),女,硕士研究生。

焦良葆(1972-),男,教授,硕士研究生导师。

孟琳(1989-),女,副教授,硕士研究生导师。

引用格式:李笑笑,焦良葆,顾嘉炜,等. 基于改进 YOLOv8n 的数字式仪表读数识别[J]. 计算机测量与控制, 2026, 34(5):284-291.

测与无人化运维的关键技术。传统数字式仪表读数方式主要依赖人工巡检与基于传统图像处理方法^[2]。人工巡检需工作人员现场观察记录, 不仅消耗大量人力成本, 且存在读数效率低、主观误差大等问题, 尤其在高压、高辐射等危险环境中, 难以满足实时监测需求^[3]; 传统图像处理常采用图像二值化、边缘提取和字符模板匹配等方法^[4], 文献 [5] 表明在夜间或暴雨环境下, 图像处理方法的误差达到了日间的近两倍, 说明其在低光或恶劣环境下的可靠性不足。文献 [6] 指出在仪表读数真实场景中, 拍摄角度倾斜、反光严重、低照度等问题仍难以通过传统方法有效解决, 导致识别误差大、漏检率高。

针对这些问题, 开始引入机器学习方法以提升仪表读数的智能化水平。文献 [7-8] 将 HOG 特征与线性 SVM 耦合, 在实验室光照下取得较好精度, 但在自然环境中光照强弱变化、反光或背景杂乱时, 识别性能急剧下降, 显示其对复杂工业场景的适应能力有限。随着深度学习技术的发展, 卷积神经网络 (CNN, convolutional neural network) 被引入以增强特征表示能力。文献 [9] 采用卷积神经网络对机械水表进行读数识别, 但在自然场景中仍难以应对小尺寸字符、倾斜视角及背景杂乱等干扰, 识别精度仍有提升空间。文献 [10] 基于 LeNet-5 改造轻量卷积神经网络, 能够处理部分噪声干扰, 但网络深度不足, 难以捕获小尺寸字符的细粒度特征, 尤其在雾天、低照度场景中识别效果不理想。文献 [11] 继续加深主干, 引入可变形卷积自适应扩展感受野, 改善了非均匀光照下的特征捕获能力, 但模型结构复杂, 推理速度降低, 无法满足实时检测需求。

随着目标检测算法的迅速发展, 其强大的特征表示与自适应学习能力, 为解决复杂场景下的仪表读数难题提供了关键技术支撑。目标检测算法主要分为双阶段 (two-stage) 和单阶段 (one-stage) 两类^[12]。双阶段检测算法通过生成候选框再进行分类, 如 R-CNN、Fast R-CNN、Faster R-CNN^[13]; 单阶段检测算法直接基于回归完成目标检测, 以 YOLO、SSD^[14] 为代表。当前, YOLO 系列模型因检测速度快、实时性强的优势, 在工业仪表读数场景中展现出良好的适用性^[15]。

为进一步提高数字仪表读数的精度和鲁棒性, 本文以 YOLOv8n 模型为基础进行改进。在原骨干网络中嵌入 SCConv 模块, 增强空间与通道特征提取能力, 提高细节的感知; 然后引入 Focal Modulation 模块替换原有的快速空间金字塔池化结构模块, 增强对关键区域的特征调制能力, 提高在复杂环境下的检测稳定性; 同时将原始回归损失函数替换为 EIou Loss, 优化回归精度, 从而实现工业数字仪表的高精度、自动化读数。

1 YOLOv8 网络结构

YOLOv8 继承了 YOLO 系列的核心思想, 将目标检测任务视为回归问题, 通过直接预测目标边界框的位置和类别来完成检测。相较于前代 YOLO 模型, YOLOv8 在检测精度和推理速度方面均有所提升。根据网络的宽度与深度不同, Ultralytics 官方提供了五个规格: YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 与 YOLOv8x。其中, YOLOv8n 具有最小的模型尺寸和最快的推理速度, 既能满足数字式仪表读数场景中的识别需求, 又能降低对部署硬件的资源依赖, 因此成为本研究中数字式仪表读数任务的优选模型规格。

YOLOv8 整体网络分为 4 部分, 分别是 Input、Backbone、Neck 和 Head。其结构如图 1 所示。在 Input 部分延续了 YOLOv5 中 Mosaic 与 MixUp 增强策略, 并在训练后期关闭 Mosaic 以稳定模型收敛^[16]。YOLOv8 采用 anchor-free 检测机制, 利用任务对齐样本分配器 Task-Aligned Assigner 在特征图的格点上动态分配正负样本, 从而摆脱对预定义 anchor 的依赖。该机制依据分类得分与回归质量综合判断正样本, 实现了更高效的训练样本选择, 降低了正负样本失衡问题。

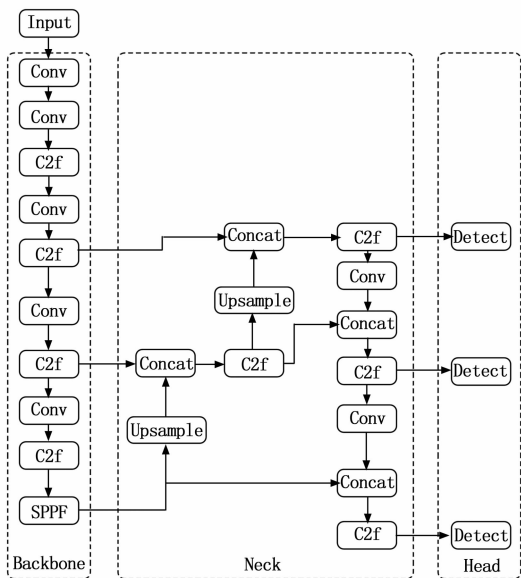


图 1 YOLOv8 网络结构图

在 Backbone 部分, YOLOv8 以 C2f (Cross Stage Partial with 2 Fast connections) 模块为核心构成单元^[17]。C2f 模块对 YOLOv5 中的 C3 结构进行了轻量化设计, 通过双支路结构实现梯度分流, 一支经过多个 Bottleneck 堆叠, 另一支直接跳跃连接, 最后在通道维度拼接融合。该结构在保持丰富特征表达的同时, 显著减少了参数量与计算量, 提高了网络的训练稳定性。空间金字塔聚合层仍采用 SPPF (Spatial Pyramid Pooling-

Fast) 模块, 通过堆叠小核最大池化操作有效扩大感受野, 增强模型的多尺度特征提取能力。Backbone 阶段采用五次 stride=2 的卷积进行下采样, 输出 P3-P5 多层特征图, 以更好地检测小目标。

Neck 部分基于 FPN+PAN 结构进行了轻量化优化^[18]。YOLOv8 采用深度可分离卷积 DWConv 与逐点卷积 PWConv 组合, 既保证了特征融合的充分性, 又显著降低了计算量与延迟^[19]。训练阶段采用多分支并行结构, 推理阶段通过结构重参数化融合为单路卷积。此外, Neck 中引入了轻量化注意力机制, 使网络能够自适应地分配多尺度特征的融合权重, 从而进一步增强小目标检测性能与整体特征表达能力。

Head 部分实现了从耦合结构到解耦结构的转变。分类与回归分支被独立设计, 分别预测类别概率与目标框偏移量, 从而降低了任务间的相互干扰。YOLOv8 在回归任务中引入 DFL (Distribution Focal Loss) 以分布建模的方式优化边界框预测, 使定位更为精确。与以往的 YOLO 系列不同, YOLOv8 取消了 objectness 分支, 分类得分与回归质量通过任务对齐机制联合确定最终置信度, 输出通道由 (5+C) 减少为 (4+C), 进一步降低了显存占用并提高了计算效率。

2 算法改进

2.1 C2f-SCConv 模块

在仪表数字识别任务中, 由于拍摄角度、光照条件以及仪表表面反光等外部因素的影响, 数字边缘模糊、局部反光区域亮度过高等问题容易导致网络模型对关键信息提取不足, 从而影响检测精度。为解决该问题, 本文将原 YOLOv8 网络中的 C2f 模块替换为 C2f-SCConv 模块, 以增强网络的特征表达能力与抗干扰能力。

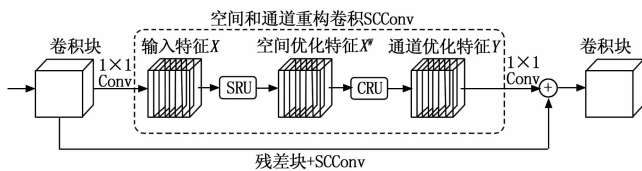


图 2 SCConv 模块示意图

SCConv 模块通过在卷积神经网络 CNN 中引入空间重建单元 SRU (Spatial Reconstruction Unit) 与通道重建单元 CRU (Channel Reconstruction Unit), 实现了空间与通道层面的动态特征建模^[20-23]。模块示意图如图 2 所示。其核心思想是在不同尺度和方向的特征之间建立选择性关系, 从而使网络能够自适应地关注目标区域的重要特征, 抑制背景噪声和无效信息, 提高检测的鲁棒性。

C2f-SCConv 模块的结构如图 3 所示。首先, 输入特征图经过 1×1 卷积进行通道压缩, 以减少计算量并保留

主要语义信息。随后, 特征图被送入 SCConv 模块内部的 SRU 和 CRU 结构中。SRU 通过引入门控权重 W 对输入特征 X 进行空间筛选。首先, 对输入特征进行归一化处理, 获得标准化结果 $GN(X)$: 表达式见式 (1):

$$X_{out} = GN(X) = \gamma \frac{X - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (1)$$

其中: μ 与 σ 分别为输入特征 X 的均值与标准差, γ, β 为可训练参数。随后, 根据归一化结果的通道缩放因子计算特征权重分布, 计算式如式 (2) 所示:

$$W_Y = \{w_i\} = \frac{y_i}{\sum_{j=1}^C y_j}, i, j = 1, 2, \dots, C \quad (2)$$

式 (2) 中, i 表示当前要计算权重的通道, j 表示用于遍历所有通道以计算总和, C 表示当前处理的特征图通道数。通过空间门控机制学习得到初始空间权重分布, 其表征输入特征图中各空间位置的重要性系数。采用 Sigmoid 激活函数对 W_Y 进行非线性映射, 得到重新加权后的连续值权重图 W , 表达式如式 (3) 所示:

$$W = \text{Gate}\{\text{Sigmoid}\{W_Y[GN(X)]\}\} \quad (3)$$

最后, 通过门控操作, SRU 将输入特征分离为重要特征和次要特征两个互补子集, 并进行交叉重构, 表达式:

$$\begin{cases} X_1^w = W_1 \odot X \\ X_2^w = W_2 \odot X \\ X_{11}^w \oplus X_{22}^w = X^{w1} \\ X_{21}^w \oplus X_{12}^w = X^{w2} \\ X^{w1} \cup X^{w2} = X^w \end{cases} \quad (4)$$

其中: \odot 表示逐元素乘, \oplus 为逐元素相加, \cup 表示通道拼接操作。通过这种交叉重构方式, SRU 在保持主要特征完整性的同时, 最大化了空间维度信息的重用与融合。

在完成空间重构后, CRU 进一步在通道维度进行特征融合。其核心思想是结合组卷积 GWC 与逐点卷积 PWC, 在减少计算复杂度的同时增强通道间特征交互。设输入特征为 X_{up} , 则 CRU 的通道变换可表示为:

$$Y_1 = M^G X_{up} + M^P X_{up} \quad (5)$$

其中: M^G 和 M^P 分别代表 GWC 与 PWC 的变换矩阵, 该式表明, 通道特征通过组卷积捕获局部相关性, 并通过逐点卷积实现全通道信息融合, 两种卷积结构的叠加使得 CRU 能兼顾细粒度特征和全局依赖, 有效强化通道间的交互关系。最后, SRU 与 CRU 协同工作, 实现空间与通道双重重构, 从而获取更具判断力的特征映射, 表达式:

$$Y = \text{CRU}[\text{SRU}(X)] \quad (6)$$

2.2 Focal Modulation 模块

在 YOLOv8 的颈部网络中, 原有的空间金字塔池

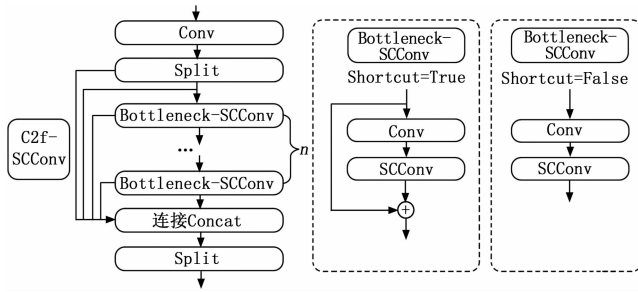


图 3 C2f-SCConv 结构图

化快速版 SPPF 模块虽能够通过多尺度空间池化获得固定长度的特征向量,但在仪表读数场景中,其对小目标特征的保持能力较弱。由于表盘中存在污渍、反光等复杂干扰因素, SPPF 在上下文信息建模方面的局限性会导致小目标特征被淹没,从而影响检测性能。

为提升网络在复杂背景下的特征表达能力,本文将 SPPF 模块替换为 Focal Modulation 结构,以显式建模长程依赖关系并增强局部与全局特征之间的交互^[24]。其网络结构如图 4 所示。Focal Modulation 模块首先对输入特征进行线性投影,生成查询向量以及用于构建上下文的内容特征。在“分层上下文化”阶段,内容特征依次经过多个不同感受野的深度可分离卷积,从而获取由局部到全局逐步扩展的多尺度上下文表示,同时结合全局平均池化进一步补充全局语义。

随后,模块引入门控聚合机制对各层级的上下文特征进行选择融合。该机制通过一个轻量线性映射直接从输入特征生成空间相关的门控权重,使网络能够针对每个空间位置自适应地选择更重要的上下文信息。经门控融合后的上下文特征在空间上具有更高的语义聚焦能力,并有效抑制复杂背景噪声的干扰。融合结果再经过线性映射转化为调制器,并通过逐通道乘法注入查询向量,实现查询—上下文之间的显式交互,最终输出表达式如式(7)所示:

$$y_i = q(x_i) \odot m(i, X) \quad (7)$$

其中: $q(x_i)$ 为输入 x_i 特征的查询映射, $m(i, X)$ 为基于多尺度上下文加权融合得到的调制器, \odot 为逐通道乘法。

与 SPPF 模块相比, Focal Modulation 避免了固定池化核带来的分辨率损失,并能够自适应整合多尺度上下文,对仪表场景中尺度较小、对背景敏感的数字区域具备更强的特征增强能力。同时该结构保持较低的计算复杂度,适用于实时仪表读数任务。实验结果表明,引入该模块后网络在整体 mAP 和小目标类别上的检测精度均获得明显提升,验证了其在复杂场景下的有效性与鲁棒性。

2.3 引入 EIoU 损失函数

YOLOv8n 默认采用 CIoU Loss 进行边界框回

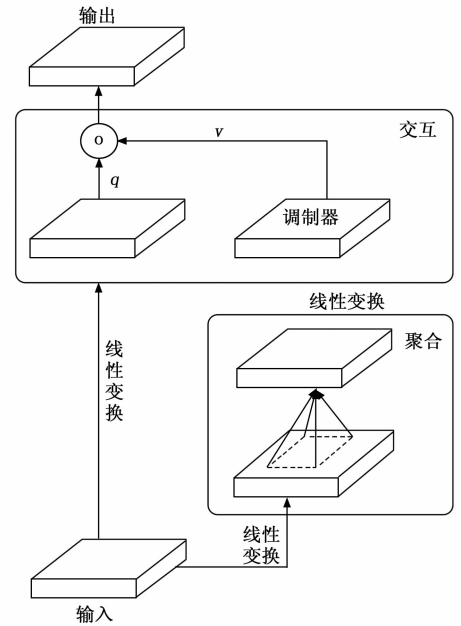


图 4 Focal Modulation 网络结构

归^[25],其表达式如式(8)~(10)所示。CIoU Loss 在 IoU 的基础上引入了预测框与真实框中心点距离与纵横比一致性约束,通过权重函数对宽高比差异进行惩罚,并以同时覆盖真实框与预测框的最小外接矩形对角线距离进行归一化。然而,CIoU 中的 v 仅反映纵横比的相对差异,未能直接建模宽高的真实偏差,导致部分场景下优化受阻:

$$L_{CIoU} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + av \quad (8)$$

$$a = \frac{v}{(1 - IOU) + v} \quad (9)$$

$$v = \frac{4}{\prod^2} \left(\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h} \right) \quad (10)$$

其中: a 是权重函数, v 用来度量宽高比的一致性,代表中心点的欧式距离, c 代表能同时覆盖真实框与预测框重合区域的对角线距离。

为缓解上述缺陷,文献[26]提出 EIoU Loss,该方法在 CIoU 的基础上将纵横比一致性因子拆分为宽度和高度两个独立优化项,同时引入宽高直接差异约束,从而更精确地对预测框的几何形状进行约束,加速收敛过程。EIoU 由重叠损失、中心距离损失及宽高损失 3 部分构成,其中前两部分沿用 CIoU 策略,宽高损失则直接最小化预测框与真实框的宽、高之差,从而加速收敛。其定义如式(11)所示:

$$L_{EIoU} = L_{IOU} + L_{dis} + L_{asp} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(\omega, \omega^{gt})}{c_\omega^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (11)$$

其中: c_ω^2 和 c_h^2 分别代表能同时覆盖预测框与真实框的最小包围区域的宽和高, ω 和 h 分别为实际框的宽

和高, w^{gt} 和 h^{gt} 分别为预测框的宽和高, b 和 b^{gt} 分别表示实际框与预测框的中心点。 L_{EIoU} 表示预测框与真实框的重叠损失, L_{dis} 表示预测框与真实框的中心点距离损失, L_{asp} 表示预测框与真实框的宽和高损失。

鉴于 EIoU 直接优化宽、高偏差, 收敛更快且定位更精确^[27], 本文将原网络中的 CIoU Loss 替换为 EIoU Loss, 以提升模型预测框与真实框的一致性, 进而提高检测框定位精度。

3 实验结果与分析

3.1 数据集的构建

由于目前缺乏权威数字式仪表公开数据集, 本实验选择自建数据集, 该数据集包含自然光照场景下拍摄的数字式仪表图像, 涉及不同光照条件、远近距离、单目标与多目标等多种复杂场景。该数据集共有 3 140 张图片, 数据集包含 11 类标签, 分别是数字 0~9 和小数点。本实验使用的部分数据集图片如图 5 所示。



图 5 部分数据集

3.2 数据集的处理

所有图像都使用 Labelimg^[28] 工具进行标注, 标注格式采用标准 YOLO 格式。标注完成后按 8 : 1 : 1 的比例, 将图片和 TXT 文件随机拆分为训练集、验证集和测试集, 数量分别为 2 512 张、314 张和 314 张。为保证数据集的可靠性与标注一致性, 本文在标注完成后进行了严格的人工复核与一致性检验。随机抽样复查 10% 的样本图像, 核对标注框位置、类别标签及边界覆盖情况。未发现明显的漏标或误标样本, 说明标注结果具有较高的准确性与一致性, 可有效支撑模型训练与验证实验。

3.3 实验环境

本实验在 Ubuntu 18.04.6 操作系统上基于 PyTorch 框架、Python 语言完成, 硬件配置为 Intel Core i7-7800X CPU 与 2 × NVIDIA GeForce RTX 1080 Ti GPU 总显存 22 GB, 内存 31 GB, 具体参数见表 1。训练时, 训练 epoch 数设为 300, 批次大小 batch-size 为 16, 初始学习率 lr0 设为 0.01, 交并比阈值 (IoU) 设置为 0.7, 输入图像尺寸为 640 × 640。为保证结果的

可比性, 所有消融与对比实验均未加载预训练权重, 并在同一设备、相同超参设置下重复进行。

表 1 实验环境

名称	配置
操作平台	Ubuntu18.04.6
CPU	IntelCorei7-7800XCPU
GPU	NVIDIA GeForce RTX1080Ti 显卡 × 2
显存	22G
内存	31G
编译环境	Python

3.4 模型参数指标及参数设置

本实验使用查准率 (P)、召回率 (R)、平均精度均值 (mAP, mean average precision)、参数量 (Params)、计算量 (GFLOPs) 和模型大小 (Model size) 作为评价模型的指标。

查准率 (P) 计算公式:

$$P = \frac{TP}{TP + FP} \quad (12)$$

召回率 (R) 计算公式:

$$R = \frac{TP}{TP + FN} \quad (13)$$

式 (12) ~ (13) 中 TP 代表模型正确预测为正类别的样本数量, FP 代表模型错误地将负类别样本预测为正类别的数量, FN 代表模型未能检测到的正类别样本数量。

在式 (12) 与式 (13) 中, TP (True Positives) 表示被模型正确判别为正类的样本数; FP (False Positives) 表示原本属于负类却被误判为正类的样本数; FN (False Negatives) 则表示本应被识别为正类却未被检出的样本数。

平均精度 AP (Average Precision,) 定义为 PR (Precision-Recall) 曲线与坐标轴所围区域的面积, 其数值可通过对该曲线下的区域进行积分获得, 计算公式如式 (14) 所示:

$$AP = \frac{1}{r} \sum_{i=1}^r P_i \quad (14)$$

平均精度均值计算公式为:

$$mAP = \frac{\sum_1^c AP}{C} \quad (15)$$

式中, C 为数据集中的检测类别。

3.5 实验结果分析

1) YOLOv8n 预测结果:

表 2 展示了 YOLOv8n 在各类别目标检测上的性能指标, 包括查准率、查全率、mAP@0.5 及 mAP@0.5 : 0.95。从整体结果来看, YOLOv8n 对大部分类别的检测表现较为优秀, 查准率和查全率也保持在较高水

平。然而, 仍存在明显不足, 尤其是小目标的检测表现欠佳, 如类别小数点 “.” $mAP@0.5$ 仅为 66.2%, 查全率低至 64.3%, 说明模型在小尺寸目标上的特征提取能力不足。

表 2 YOLOv8n 检测结果

Class	查准率/%	查全率/%	$mAP@0.5$ /%	$mAP@0.5:0.95$ /%
all	97.5	90.4	93.5	76.6
0	98.8	92.8	98.2	84.1
1	97.9	89.5	94.5	65.1
2	97.4	95.8	98.3	83.1
3	97.8	91.4	97.7	83.5
4	97.1	92.5	96.7	82.7
5	98.7	86.9	92.5	82.9
6	97.2	99.9	97.7	84
7	98.8	96.4	97.8	80.4
8	96	89.9	94.7	80.3
9	97	95.4	94.2	84.4
.	95.4	64.3	66.2	42

2) 多层特征嵌入 C2f-SCConv 的性能对比:

为进一步验证 C2f-SCConv 模块在不同层级中的应用效果, 本文分别在 YOLOv8n 网络的 P2、P3、P4、P5 层以及全部 4 个层级 (记为 “All”) 同时引入该模块, 并在相同训练参数与数据集条件下进行对比实验。实验结果如表 3 所示。

表 3 各特征层引入 C2f-SCConv 模块后的检测结果比较

实验方案	查准率 /%	查全率 /%	$mAP@0.5$ /%	$mAP@0.5:0.95$ /%
YOLOv8n	97.5	90.4	93.5	76.6
C2f-SCConv@P2	97.2	90.7	93.5	76.5
C2f-SCConv@P3	98.1	90.6	93.7	77.0
C2f-SCConv@P4	95.6	91.8	93.6	76.9
C2f-SCConv@P5	96.7	90.4	93.2	76.3
C2f-SCConv@ALL	97.7	90.2	94.2	77.3

从表中可以看出, C2f-SCConv 模块在不同特征层的嵌入效果存在一定差异。当模块分别作用于中层特征 (P3、P4) 时, 检测性能较原始 YOLOv8n 模型略有提升, 说明该模块能在细节特征与语义信息之间实现更优平衡。然而, 当模块仅作用于深层特征 (P5) 时, 性能略有下降, 表明在高语义层特征中引入该模块可能削弱小目标的细节表征能力。将 C2f-SCConv 模块同时嵌入 P2、P3、P4、P5 四个特征层时, 模型性能最优, $mAP@0.5$ 提升至 94.2%。这一结果充分验证了, 多层特征联合嵌入能够通过跨层信息融合与空间通道协同建模, 有效增强特征表达的判别性, 从而显著提升模型的整体检测精度与鲁棒性。

3) Focal Modulation 模块可视化分析:

为直观验证本文提出的 Focal Modulation 模块在上

下文信息建模与特征筛选方面的优势, 本节采用梯度加权类激活映射 (Grad-CAM) 技术, 对原始 YOLOv8n 网络与改进后网络在浅层与中层输出的特征激活图进行可视化, 并将结果以热力图形式展示, 如图 6 所示。

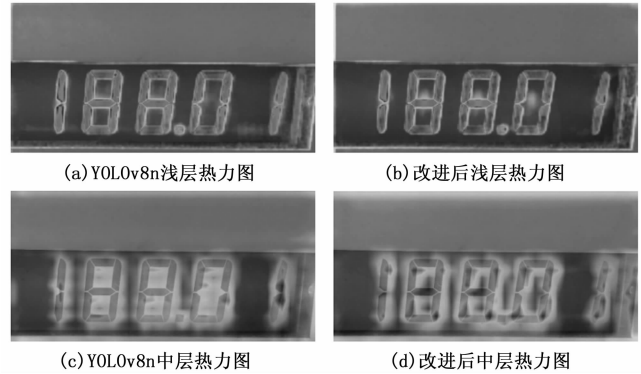


图 6 原 YOLOv8 与改进后热力图

从热力图结果可以看出, 原始 YOLOv8n 的浅层激活主要集中在数字及小数点区域, 能够通过激活强度分布大致勾勒出字符轮廓, 但对小数点等小目标的边界表征仍较模糊, 目标区域与表盘背景之间的激活强度差异不明显。改进后网络的浅层热力图在保持数字轮廓完整性的同时, 小数点区域的激活强度梯度更明显, 边界过渡更清晰。原始 YOLOv8n 的中层热力图中, 激活区域多围绕字符边缘分布, 存在明显的响应扩散现象, 部分激活超出字符本体边界, 容易引入背景纹理的无效干扰。改进后网络的中层热力图虽然因感受野扩大仍存在激活区域外延, 但数字主体区域的激活强度显著增强, 响应更集中于目标内部, 整体目标聚焦性优于原网络。上述特征响应差异验证了改进方案的有效性, Focal Modulation 模块通过分层上下文聚合与门控筛选机制, 能够在浅层提升小目标细节的特征表征精度, 在中层强化数字主体的响应强度, 同时有效抑制背景噪声, 使网络特征表达更集中于目标区域。

4) 消融实验:

对比 YOLOv8n 原网络, 本实验进行了三点改进: 首先, 将 SCConv 模块集成到 Backbone 的 C2f 模块中, 以增强网络在空间和通道维度的特征表达能力, 从而提升对小目标和细节信息的感知能力; 其次, 用 Focal Modulation 结构替换原有的 SPPF 模块, 用于调制重要特征区域, 增强全局上下文信息的感知能力, 进一步改善密集目标和小目标的检测效果; 最后, 将原始的回归损失函数替换为 EIoU Loss, 以优化边界框回归精度, 尤其是针对宽高比差异较大的目标。为了系统评估各改进模块对模型性能的贡献, 本实验分别对 C2f+SCConv、FocalModulation 以及 EIoU 进行组合消融实验, 表 4 为所有类别消融实验结果, 表 5 为标签小数点 “.” 消融实验结果。

表 4 总类别消融实验

算法	查准率	查全率	mAP	$mAP@0.5:$
	/%	/%	@0.5/%	0.95/%
YOLOv8n	97.5	90.4	93.5	76.6
YOLOv8n+C2f-SCConv	97.7	90.2	94.2	77.3
YOLOv8n+FocalModulation	95.4	90.7	94.3	77.4
YOLOv8n+EIoU	95.6	92.1	94.9	76.7
YOLOv8n+C2f-SCConv+Focal Modulation+EIoU	97.2	91.7	95.0	77.4

表 5 小数点消融实验

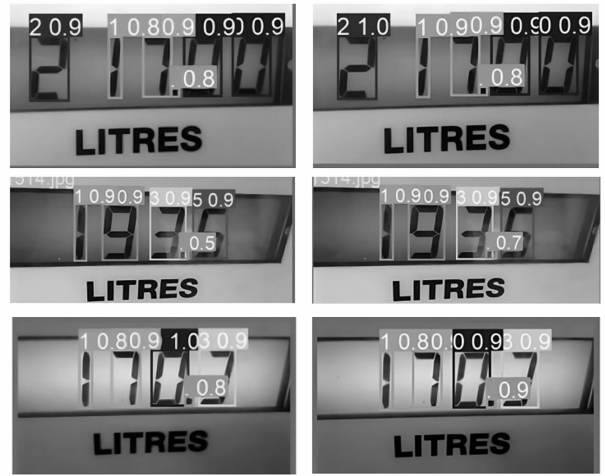
算法	查准率	查全率	mAP	$mAP@0.5:$
	/%	/%	@0.5/%	0.95/%
YOLOv8n	95.4	64.3	66.2	42.0
YOLOv8n+C2f-SCConv	95.4	71.3	76.7	46.7
YOLOv8n+FocalModulation	95.9	66.1	67.9	43.8
YOLOv8n+EIoU	95.6	69.0	66.6	44.4
YOLOv8n+C2f-SCConv+FocalModulation+EIoU	96.8	71.2	80.5	47.0

由表 4 和表 5 的结果可知, 所提出的三项改进均在不同程度上提升了模型的整体检测性能。将 C2f-SCConv 模块嵌入 Backbone 后, 网络的特征提取能力得到增强, $mAP@0.5$ 由 93.5% 提升至 94.2%; 引入 Focal Modulation 模块替换原有的快速空间金字塔池化结构模块, 模型的全局上下文建模能力进一步提高, $mAP@0.5$ 达到 94.3%; 替换为 EIoU 损失函数后, 边界框回归精度显著改善, $mAP@0.5$ 提升至 94.9%。当 3 种改进模块联合使用时, 模型的 $mAP@0.5$ 进一步提升至 95.0%, 说明各模块在特征提取与回归优化方面具有良好的互补性。

在小目标检测实验中, C2f-SCConv 改进效果较好, $mAP@0.5$ 由 66.2% 提升至 76.7%, 查全率提升 7 个百分点, 表明其在优化目标定位方面具有显著优势。当三者联合应用时, 模型在小目标检测任务中的 $mAP@0.5$ 提升至 80.5%, 验证了所提改进方案的有效性与鲁棒性。

为了清楚展示改进后网络相对于原网络的性能提升, 本文分别对两者在测试集上的表现进行了对比。如图 7 所示, 图 7 (a) 和图 7 (b) 分别展示了两种网络的测试结果。

从对比分析结果可以明显观察到, 原始网络在读数识别任务中对小数点目标检测精度较低。这主要是由于小数点尺寸极小、亮度对比度低, 且在复杂背景或光照不均的情况下易与周围字符产生混淆, 从而影响检测的准确性。经过网络结构改进后, 模型在特征提取与全局感知方面的能力显著增强, 对小尺度目标的表征更加充分, 能够有效区分小数点与背景区域。实验结果表明,



(a) 原YOLOv8n检测结果

(b) 改进后YOLOv8n检测结果

图 7 YOLOv8n 网络与改进后网络效果对比图

改进后的模型在小数点检测中漏检率显著降低, 识别稳定性和精度均得到明显提升。

4 结束语

本文针对复杂仪表读数场景中目标识别精度低、特征提取能力不足及边界框回归不准确等问题, 提出了一种基于改进 YOLOv8n 的数字式仪表读数方法。从特征提取、回归优化等方面对网络结构进行了系统性改进。首先, 将 SCConv 模块嵌入 Backbone 中的 C2f 结构中, 以增强网络在空间与通道维度的特征表达能力, 从而有效提升了对小目标和细节特征的感知性能。其次, 用 Focal Modulation 结构替换原有的 SPPF 模块, 用以强化全局上下文信息的建模能力, 使网络能够自适应聚焦于关键特征区域, 进一步提升了在复杂背景下的识别准确率。最后, 将原有的 CIoU 损失函数替换为 EIoU 损失函数, 有效优化了边界框的回归精度。实验结果表明, 3 种改进模块均在不同程度上提升了模型整体检测性能, 在保证总体检测精度的情况下, 使得小目标检测精度由 66.2% 提升至 80.5%, 优于原始 YOLOv8n 模型, 表明该方法在复杂背景及小目标场景下具有更强的鲁棒性与泛化能力。由于仪表表盘反光、表盘污渍堆积和环境光照不足等因素, 仍会出现漏检、误检等问题, 本文提出的算法在应对这些挑战仍有提升空间。后续将深入探讨在表面反光、环境光照不足、表面污渍堆积等不同背景下识别效果, 进一步优化检测系统精度。

参考文献:

- [1] 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 (08): 110-119.
- [2] ZHANG Z, HUA Z, TANG Y, et al. Recognition method of digital meter readings in substation based on connected

- domain analysis algorithm [C] //Actuators. MDPI, 2021, 10 (8): 170.
- [3] 沈云青,熊卫华,黄为民,等. 基于模板匹配与Hough圆检测的仪表识别[J]. 计算机技术与发展, 2021, 31 (4): 69-73.
- [4] 高菊,叶桦. 一种有效的水表数字图像二次识别算法[J]. 东南大学学报:自然科学版, 2013, 43 (A01): 153-157.
- [5] CHEN G, BAI K, LIN Z, et al. Method on water level ruler reading recognition based on image processing [J]. Signal, Image and Video Processing, 2021, 15 (1): 33-41.
- [6] PENG J, ZHOU W, HAN Y, et al. Deep learning-based autonomous real-time digital meter reading recognition method for natural scenes [J]. Measurement, 2023, 222: 113615.
- [7] XIAO X C, HUA F, GUO Q Y, et al. A new method of digital number recognition for substation inspection robot [C] //2016 4th International Conference on Applied Robotics for the Power Industry (CARPI), IEEE, 2016: 1-4.
- [8] QUINTANILHA D B P, COSTA R W S, DINIZ J O B, et al. Automatic consumption reading on electromechanical meters using HoG and SVM [C] //7th Latin American Conference on Networked and Electronic Media (LACNEM 2017), IET, 2017: 57-61.
- [9] LI J, SHEN J, NIE K, et al. Reading recognition method of mechanical water meter based on convolutional neural network in natural scenes [J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2024, 28 (1): 206-215.
- [10] 卜令正,王洪栋,朱美强,等. 基于改进卷积神经网络的多源数字识别算法[J]. 计算机应用, 2018, 38 (12): 3403-3408.
- [11] 郭兰英,韩睿之,程鑫. 基于可变形卷积神经网络的数字仪表识别方法[J]. 计算机科学, 2020, 47 (10): 187-193.
- [12] YU T, REN J, CHANG X, et al. A survey of target detection algorithms in deep learning [C] //International Conference on Algorithms, High Performance Computing, and Artificial Intelligence (AHPCAI 2023), SPIE, 2023, 12941: 187-199.
- [13] LI W, FENG X S, ZHA K, et al. Summary of target detection algorithms [C] //Journal of Physics: Conference Series. IOP Publishing, 2021, 1757 (1): 012003.
- [14] ABDUSALOMOV A, UMIRZAKOVA S, KUTLIMURATOV A, et al. Lightweight UAV-based system for early fire-risk identification in wild forests [J]. Fire, 2025, 8 (8): 288.
- [15] LI J, CHEN M. DEW-YOLO: an efficient algorithm for steel surface defect detection [J]. Applied Sciences, 2024, 14 (12): 5171.
- [16] 张绪康,朱硕. 自适应复杂环境噪声的多重关注联合优化检测算法[J]. 电子测量技术, 2025, 48 (1): 175-185.
- [17] LI Z, QIN Z, LUO W, et al. Gasoline engine misfire fault diagnosis method based on improved YOLOv8 [J]. Electronics, 2024, 13 (14): 2688.
- [18] ZENG B, ZHOU Y, HE D, et al. Research on lightweight method of insulator target detection based on improved SSD [J]. Sensors, 2024, 24 (18): 5910.
- [19] LI Y, ZHENG Q, XU S, et al. YOLO-EDGE: an object detection algorithm for traffic scenarios [J]. The Journal of Supercomputing, 2025, 81 (7): 840.
- [20] LI J, WEN Y, HE L. Sconv: Spatial and channel reconstruction convolution for feature redundancy [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 6153-6162.
- [21] YANG Z, SHAO Y, WEI Y, et al. Precision-boosted forest fire target detection via enhanced YOLOv8 model [J]. Applied Sciences, 2024, 14 (6): 2413.
- [22] CHANG Z, XU M, WEI Y, et al. UNeXt: an efficient network for the semantic segmentation of high-resolution remote sensing images [J]. Sensors, 2024, 24 (20): 6655.
- [23] YU M, LI Y, LI Z, et al. Dense detection algorithm for ceramic tile defects based on improved YOLOv8 [J]. Journal of Intelligent Manufacturing, 2024, (prepublish): 1-16.
- [24] ZHANG J, HUANG W, ZHUANG J, et al. Detection technique tailored for small targets on water surfaces in unmanned vessel scenarios [J]. Journal of Marine Science and Engineering, 2024, 12 (3): 379.
- [25] XUE J, CHENG F, LI Y, et al. Detection of farmland obstacles based on an improved YOLOv5s algorithm by using CIoU and anchor box scale clustering [J]. Sensors, 2022, 22 (5): 1790.
- [26] CAI D, ZHANG Z, ZHANG Z. Corner-point and foreground-area IoU loss: better localization of small objects in bounding box regression [J]. Sensors, 2023, 23 (10): 4961.
- [27] YANG Z, WANG X, LI J. EIou: an improved vehicle detection algorithm based on vehiclenet neural network [C] //Journal of physics: Conference series, IOP Publishing, 2021, 1924 (1): 012001.
- [28] SANG H, LI Z, SHEN X, et al. Rapid identification of mangrove leaves based on improved YOLOv10 model [J]. Forests, 2025, 16 (7): 1068.