

基于 YOLOv11 的电表读数识别系统的设计与实现

23122721 冯俊佳

摘要：如今，传统的人工抄表方式逐渐暴露出效率低、易出错、成本高等局限性，难以满足当下对电力管理高效性与智能化的要求。因此，借助计算机视觉和深度学习等新兴技术，本课题利用一个包含电表读数与标签的数据集，基于 PyTorch 深度学习框架，结合 YOLOv11 网络模型和基于 Paddle 平台的 PPOCR 算法，开展模型训练、参数评估、可视化检测输出与数字识别等过程，最终构建起了一套完整的电表读数识别模型。

关键词：深度学习，目标检测，数字识别，参数评估

1. 项目背景和意义

随着社会经济与智能电网技术迅猛发展，使得电表数据采集从传统人工抄表逐步向自动化识别转型。电表读数采集经历了从人工抄录，到目标检测算法辅助识别的转变。人工方式存在抄录效率低、出错率高等问题，无法满足实时监测和动态预警的实际需求。因此，随着深度学习在目标检测领域的快速发展，基于 YOLO 系列等目标检测模型的应用，凭借高效的端到端架构，以及在复杂环境下展现出的多尺度检测能力，为电表图像的指针式、滚轮式数字识别提供了更为高效解决方案^[1]。

基于此类系统，电力企业能够实时掌握用户的用电变化趋势，从而优化发电与供电策略，降低能耗浪费；同时，用户方也可通过系统查看用电详情，以促进节能降耗。因此，电表读数识别系统作为能源数字化转型的重要一环，正在逐步成为建设智能、高效、可持续电力系统的支撑技术^[2]。

2. 相关原理与技术

2.1 YOLOv11 目标检测模型原理

YOLOv11 以端到端的结构实现高效目标检测，在模型结构上兼顾精度与速度的双重保障。其模型通过多尺度特征融合、增强型注意力机制与精细化损失函数设计，有效提升了对小目标与密集目标的检测能力。该模型在本系统中为电表图像中 EM_Data 与 EM_Label 区域的精准识别提供了坚实的技术基础，其具体原理如图 1 所示。

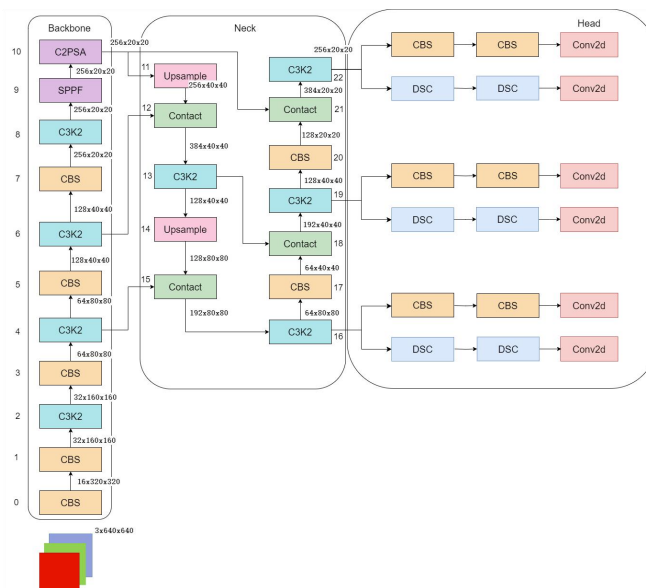


图 1 YOLOv11 网络结构图^[3]

结合图 1 可以知道, YOLOv11 网络结构主要由六大模块组成: 输入层 (Input)、主干网络 (Backbone)、特征融合网络 (Neck)、检测头 (Head)、损失函数 (Loss)、后处理 (Post-processing)。

由于篇幅限制, 本文仅对每个模块的基本操作进行介绍, 具体的操作内容不多赘述。

(1)输入模块 (Input Layer)

输入层负责对原始图像进行预处理, 使其满足模型输入要求。具体操作包括:

1)尺寸统一; 2)像素归一化: 将像素值标准化到 $[0, 1]$; 3)数据增强; 4)标签编码。

(2)主干网络 (Backbone)

YOLOv11 采用改进的轻量化主干网络作为特征提取模块, 其核心思想在于平衡参数规模与特征表达能力, 主要技术特性包括:

1)CSP (Cross Stage Partial) 结构: 提升网络收敛效率; 2)RepVGG 重参数卷积: 折叠为单一 3×3 卷积; 3)ELAN (Efficient Layer Aggregation Network) 结构: 捕获多级语义信息; 4)DropBlock 正则化: 随机遮挡特征区域, 提升网络模型对目标遮挡与局部缺失的有效性自动识别。

主干网络最终输出三层不同尺度的语义特征图, 供后续模块使用。

(3)特征融合模块 (Neck)

Neck 模块用于多尺度特征融合, 增强模型对不同大小目标的检测能力。YOLOv11 中采用改进的 FPN + PAN 架构, 具体包括: 1)FPN (Feature Pyramid Network): 增强低层特征的语义表达; 2)PAN (Path Aggregation Network): 保留细粒度的定位特征; 3)SPPF (Spatial Pyramid Pooling - Fast): 进行最大池化并拼接其中特征。

此外, 为提升轻量性与注意力机制融合效果, YOLOv11 支持使用 SimAM、ECA 等算法模块对特征图进行增强, 并有效提升对目标区域的聚焦能力。

(4)检测头 (Head)

YOLOv11 的检测头采用 anchor-based 多分支输出结构, 对每个候选框的回归目标位置、预测类别概率和置信度评分, 核心输出包括:

1)边界框坐标 (x, y, w, h) ; 2)置信度 (P_{obj}) ; 3)类别概率 (P_{class}) : 目标属于每一类的概率。

每个检测头输出的通道数为: 输出通道数 = anchor 数量 \times $(5 + \text{类别数})$ 。在本系统中, 类别数为 2, 则每个 anchor 输出 7 个值 (4 个边框参数, 1 个置信度, 2 个分类概率)。

(5)损失函数 (Loss Function)

YOLOv11 使用多分量联合损失函数进行端到端训练, 定义为:

总损失 = 边界框损失 \times 权重 1 + 置信度损失 \times 权重 2 + 分类损失 \times 权重 3

典型权重配置为: 边界框损失权重 = 0.05, 置信度损失权重 = 1.0, 分类损失权重 = 0.5。

(6)推理与后处理 (Inference & Post-processing)

YOLOv11 在推理阶段首先输出大量候选框, 为了得到最终检测结果, 需要通过以下步骤处理:

1)置信度过滤: 本课题设为 0.25; 2)非极大值抑制 (NMS); 3)改进的 NMS 算法。

最终, YOLOv11 输出一组边框、类别及对应置信度, 作为检测结果用于后续分析或展示。

2.2 PaddleOCR 数字识别模块原理

PaddleOCR 构建了一个高效的文本检测与识别框架, 具备识别精度高、速度快、模型轻量化等优势。在本课题中, PaddleOCR 被用于对电表图像中的数字区域进行精准识别与结果输出, 并与

YOLOv11 检测模块协同构成完整的电表读数识别系统。此外，本系统部署的字符增强机制与结构性处理策略，有效应对了真实场景中的模糊、遮挡、翻滚字符等问题，显著提升了识别准确率。

(1)文本检测模块 (Text Detection)

该模块用于定位图像中所有可能包含文字的区域，输出一系列矩形或多边形的文本框坐标。PaddleOCR 默认使用 DB 算法进行文本检测，其结构和原理如下：

1)主干网络 (Backbone)：提取图像特征；2)特征融合 (FPN)：提高小字体检测效果；3)分支输出 (Head)：生成二值图；4)后处理：从二值图中提取文字边界框。

以上方法相较其他传统的检测算法具有更高的速度与端到端可训练性，尤其适合复杂背景下的长文本和弯曲文本检测任务。

(2)文本识别模块 (Text Recognition)

文本识别模块的核心任务是将检测到的文字区域图像转化为可读的字符序列，也就是本课题所使用 OCR 算法所主要实现的功能。PaddleOCR 默认采用 CRNN 结构，具有如下流程：

1)特征提取 (CNN)：使用卷积神经网络提取字符图像的二维特征；2)序列建模 (RNN 或 Transformer)：提取字符间的上下文信息；3)转录解码 (CTC 或 Attention)：使用 CTC 损失进行无对齐字符识别。

(3)后处理与纠错模块 (Post-processing)

经过识别后的结果可能存在误识别或置信度较低的情况，因此 PaddleOCR 通常会增加一系列处理步骤以优化识别结果，主要包括：

1)置信度过滤：剔除低置信度输出；2)规则约束：提升合理性；3)结构性校验：对电表读数位数进行长度修补与补全；4)增强模块嵌入：采用反色增强等自定义增强方式，提高识别质量。

(4)PaddleOCR 模型部署与调用机制

PaddleOCR 支持多种部署方式，包括 Python API、C++ 预测库、Paddle Lite 移动端部署等。本系统通过 Python 方式集成 PaddleOCR 模型，调用方式如下：

1)加载检测模型与识别模型；2)输入图像，依次进行文本检测与裁剪；3)对每个文本框执行预处理后送入识别模型；4)输出字符序列与置信度得分。

系统将检测到的目标区域与识别得到的结果进行联合处理，提升整体读数准确率。

3. 模型训练与模型评估

3.1 数据集扩充与标注

本课题所使用的数据集主要为黑白电表图像，包含了亮度、模糊程度、黑白对比度以及电表磨损状态等多方面的变化。为进一步提升数据多样性，本课题主要在图像亮度、对比度、模糊程度、尺寸与分辨率等方面进行调整，对原始数据集进行了扩充处理。本课题采用开源图像标注工具 LabelImg 对图像进行边界框标注，包含五项信息：类别标签（电表读数或标签）、边界框中心点的纵横坐标、宽度和高度。所有标注信息以与图像文件同名的.txt 文件进行保存。标注示例如图 2 所示。

3.2 数据划分与训练流程设计

为提升模型对电表图像中多样化目标的检测能力，本文在已完成数据扩充与标注的基础上，采用 8:2 的比例将数据集进行随机划分训练集与测试集，其中训练集包含 2650 张图像，验证集为 663 张图像。训练过程中启用 Mosaic 拼接、随机旋转、颜色扰动等数据增强策略，以增强模型对多种复杂条件下的适应能力。具体训练流程如图 3 所示。



图 2 电表数据集边界框标注示例

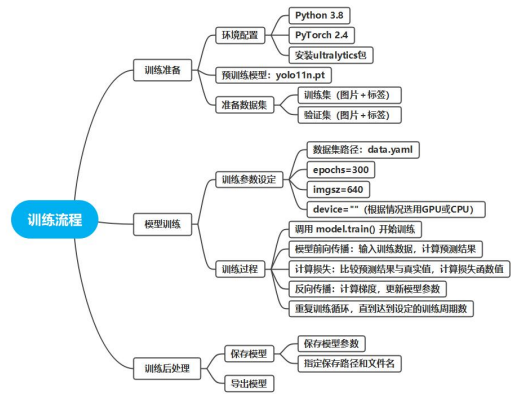


图 3 训练流程思维导图

3.3 模型性能评估

如图 4 所示, 展示了 YOLOv11 模型在整个训练过程中的各项损失函数与评估指标的变化趋势。

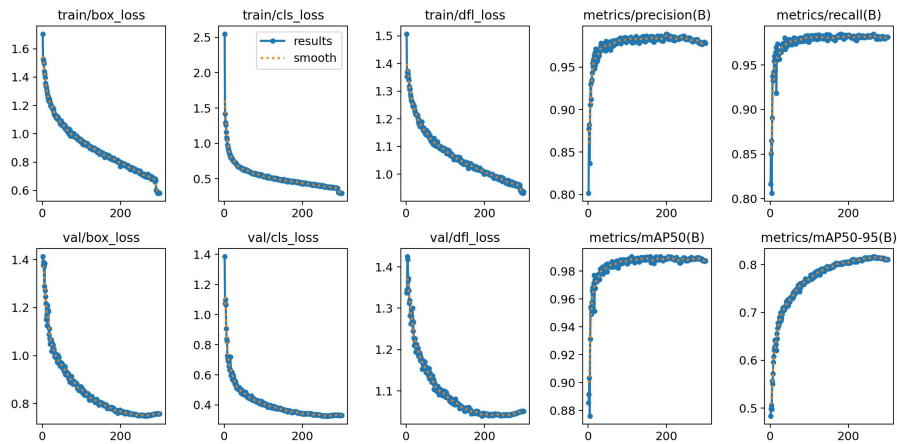


图 4 损失曲线图

其中, 横轴为训练轮数, 纵轴为对应指标的数值。该图中包含以下内容:

(1)左上三图 (train/box_loss, train/cls_loss, train/dfl_loss) 为训练集中三种损失函数的下降曲线, 这三种曲线均整体呈指数型下降趋势, 表明模型在目标框回归、类别分类与分布式回归上的拟合能力在训练阶段持续增强。

(2)左下三图 (val/box_loss, val/cls_loss, val/dfl_loss) 对应验证集损失, 此类曲线也呈现稳定下降趋势, 表明了训练过程中, 模型未发生明显过拟合, 具备良好的泛化能力。

(3)右侧四图为 Precision、Recall、mAP@0.5、mAP@0.5:0.95 等四个指标随训练轮数而提升的曲线图: Precision 与 Recall 在前 50 轮迅速提升, 随后波动趋于稳定, 收敛于 0.98 附近; mAP@0.5 及 mAP@0.5:0.95 分别收敛到 0.99 与 0.81 附近, 表明模型具备极高的整体检测精度与细粒度定位能力。

该损失曲线图清晰展示了模型训练过程的有效性与指标收敛性, 验证了 YOLOv11 架构在训练过程中展示出来的适用性与高性能表现。

此外, 为进一步量化模型在不同类别下的检测性能, 本课题基于验证集分别计算了每类目标的 Precision、Recall、F1 分数与 AP 指标, 结果如表 1 所示, 具体结果图片可在压缩包中查看。

表 1 各类标签的评估值

类别 ID	类别名称	Precision	Recall	F1 Score	AP
0	EM_Data	0.9895	0.9836	0.9865	0.9929
1	EM_Label	0.9757	0.9847	0.9802	0.9841

注: COCO 风格下的 $mAP@[0.5:0.95] = 0.8334$ 。

对于 EM_Data 类别, 模型表现出极高的检测准确性与召回能力, 其 Precision 为 0.9895, Recall 为 0.9836, F1 分数达到 0.9865, AP 指标高达 0.9929。对于 EM_Label 类别的检测能力亦表现良好, 各项指标均维持在 0.97 以上水平。

从总体角度看, 模型在验证集上取得了平均 AP 值为 0.9885 的优异成绩。此外, 采用 COCO 风格的 $mAP@[0.5:0.95]$ 综合评价方式获得的分数为 0.8334, 说明模型在不同 IoU 阈值下的泛化能力与定位精度同样表现稳定, 具备较强的实用性与迁移性。

4 数字识别模块实验结果分析

在完成目标检测模块的训练与验证后, 在目标检测的基础上, 本课题进一步设计并实现了数字识别模块, 用于对检测到的 EM_Data 与 EM_Label 区域执行数字文本提取与呈现。该模块基于 PaddleOCR 框架构建, 并通过图像增强与多版本回退策略提升文本识别准确率。

如图 5 示例所示, 最终识别结果叠加在原图上进行可视化标注, 标注内容包括类别、识别值与置信度等级, 其中, 框体颜色为绿色表示高置信度, 橙色表示中等, 红色表示较低; 同时保存所有识别失败样本的图像与结果, 如图 6 所示, 便于后期人工分析与模型改进。



图 5 数字识别模块可视化识别图 (例)



图 6 数字识别模块失败样本图 (例)

该模块在验证集上实现了较高的识别精度, 尤其是在复杂环境下表现稳定。相比基础 OCR 直接识别方案, 系统在读数与标签两个类别中的识别准确率稳定提升, 平均置信度小幅度提高。该模块的集成显著增强了整个电表读数系统在实际部署场景中的鲁棒性与实用性。

参考文献

- [1] 张建华, 李春梅. 基于深度学习的电表自动读数系统研究[J]. 电测与仪表, 2021, 58(10): 114-119.
- [2] Howard J, Gugger S. Deep learning for coders with fastai and PyTorch[M]. O' Reilly Media, 2020.
- [3] CooVally_AI. YOLOv11 网络模型原理讲解 [EB/OL]. (2024-04-19)[2025-04-28]. https://blog.csdn.net/CooVally_AI/article/details/144721514.
- [4] 王世玮, 王昊, 王强. 基于 YOLO 的智能电表读数识别系统设计[J]. 电力系统保护与控制, 2020, 48(10): 150-155.
- [5] 李凌峰, 杨振东, 刘川. 基于图像识别技术的智能电表自动抄表系统研究[J]. 电测与仪表, 2019, 56(14): 107-111.