

基于改进 YOLOv7 的电力调度控制系统图元识别方法研究

张文广¹ 曾祥玖^{1,2} 刘重阳^{1,2}

(1. 新能源电力系统全国重点实验室(华北电力大学), 北京 102206;
2. 华北电力大学控制与计算机工程学院, 北京 102206)

摘要 针对电力调度控制系统图中图元分布密集、图元相似和小尺寸图元较多导致识别效果差的问题, 本文提出一种基于改进 YOLOv7 的电力调度控制系统图元识别方法。首先, 使用轻量化的 DRBNCSPPELAN 模块替换主干网络中的 ELAN 模块, 以同时捕获不同尺度的空间模式; 其次, 提出一种信息引导融合模块, 替代颈部网络中的 Concat, 并融合 SE 注意力机制, 以增强信息全局交互能力; 接着, 引入 MPDIoU 损失函数, 以改善图元边界框的识别效果; 最后, 利用电力调度控制系统图数据集进行验证。结果表明, 与基准模型相比, 所提方法的精确率、召回率和平均精确率均值分别提高了 5.1 个百分点、3.1 个百分点和 3.5 个百分点, 有助于实现对电力调度控制系统图元的精准识别。

关键词: 深度学习; 电力调度控制系统图; 图元识别; YOLOv7; SE 注意力机制

Research on element recognition method of power dispatching control system based on improved YOLOv7

ZHANG Wenguang¹ ZENG Xiangjiu^{1,2} LIU Chongyang^{1,2}

(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources,
North China Electric Power University, Beijing 102206;

2. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206)

Abstract Aiming at the problems of dense distribution of elements, similar elements and large number of small-sized elements in power dispatching control system diagrams, which lead to poor recognition effect, an improved you only look once v7 (YOLOv7) element recognition method for the power dispatching control system diagrams is proposed. Firstly, the lightweight dilated reparam block net with cross stage partial and efficient layer aggregation network (DRBNCSPPELAN) module is used to replace the efficient layer aggregation network (ELAN) module in the backbone network to capture spatial patterns of different scales simultaneously. Secondly, an information-guided fusion module is proposed to replace the Concat in the neck network, and the squeeze-and-excitation (SE) attention mechanism is integrated to enhance the global information interaction ability. Then, the minimum point distance intersection over union (MPDIoU) loss function is introduced to improve the recognition effect of the element bounding box. Finally, experimental validation is performed via the power dispatch control system diagram dataset. The results show that compared with the baseline model, the precision, recall and mean average precision of the proposed method are improved by 5.1 percentage points, 3.1 percentage points and 3.5 percentage points respectively, which is helpful to achieve accurate recognition of elements in the power dispatching control system diagrams.

Keywords: deep learning; power dispatching control system diagrams; element recognition; you only look once v7 (YOLOv7); squeeze-and-excitation (SE) attention mechanism

0 引言

电力调度控制系统作为电力调度的重要支撑,其稳定性、安全性和运行效率直接关系到电力行业的正常运作,在新型电力系统构建中具有不可替代的作用^[1]。电力调度控制系统图作为电力调度控制系统的基础元素,广泛用于电网拓扑结构展示、实时监控运行状态、辅助调度决策和安全与稳定性管理^[2]。电力调度控制系统图元识别是实现电力调度控制系统图高效转换为结构化数据的关键任务之一。通过精准识别图元,可以为电网智能运维管理工作提供更加高效的数据支持。因此,提高图元识别能力具有重要意义。

针对图元识别任务,研究学者使用传统的规则和机器学习方法开展了相应的研究。文献[3]提出一种基于图元分割和模板匹配的电气图元识别方法,取得了较好的识别效果。文献[4]提出一种基于支持向量机和决策树的组合符号识别模型,能够有效识别电气符号。文献[5]提出一种基于塔式梯度方向直方图(pyramid histogram of oriented gradients, PHOG)的电气图元识别方法,提高了对图元的识别率。然而,上述方法对同类电气图元符号的不同表示形式的泛化能力不强,影响了其在工程场景中应用的性能。

近年来,研究人员将深度学习应用于电力系统,例如线路绝缘子缺陷检测^[6]、输电线路缺陷识别^[7]、无人机巡检^[8]和电力调度控制系统图的图元识别等,并在图元识别方面取得了一定的进展。文献[9]提出一种基于深度学习和图匹配的接线图检测与校核方法,实现了厂站一次接线图的自动识别和自动校核。文献[10]提出一种基于更快区域卷积神经网络(faster region convolutional neural network, Faster-RCNN)的电网图元识别方法,该方法对电力调度控制系统图中的图元有较好的识别效果。文献[11]提出一套基于人工智能的完整电网厂站接线图识别方法,提升了接线图识别的效率和准确性。以上方法未考虑不同图元的相似性,对小尺寸图元的识别准确率和图片识别速率有待提高。

针对上述图元识别方法泛化能力不强、对相似图元及小尺寸图元的识别效果差等问题,本文提出一种基于改进 YOLOv7 (you only look once v7) 的电力调度控制系统图元识别方法。首先,使用轻量化的 DRBNCSPELAN(dilated reparam block net with

cross stage partial and efficient layer aggregation network) 模块替换主干网络中的 ELAN (efficient layer aggregation network) 模块;其次,提出一种信息引导融合模块 (information-guided fusion module, IGFM) 替代颈部网络中的 Concat;接着,将 YOLOv7 的 CIoU (complete intersection over union) 损失函数替换为 MPDIoU (minimum point distance intersection over union) 损失函数;最后,通过电力调度控制系统图数据集验证所提方法的有效性。

1 YOLOv7 模型

YOLOv7 作为单阶段检测算法的代表之一^[12],在目标检测任务中具有较高的识别精度及较快的识别速度。它包含 4 部分:输入 (Input)、主干网络 (Backbone)、颈部网络 (Neck) 和预测头 (Head Prediction)。

输入部分对输入图片进行预处理,包括马赛克数据增强、锚框的自适应计算及图像的自适应缩放^[13],将图片转换成张量形式,并进行归一化等操作,以满足 YOLOv7 模型的输入要求。

主干网络又称为特征提取网络,包含 3 部分: CBS 模块、MP 模块和 ELAN 模块。CBS 模块由卷积 (convolution) 层、BN (batch normalization) 层和 SiLU (sigmoid linear unit) 激活函数构成。MP 模块由最大池化 (max-pooling) 层和 CBS 模块拼接而成。ELAN 模块由多个 CBS 模块拼接而成,主要用于提取图像特征和调整网络的通道数。

颈部网络又称为特征融合网络,主要包括 SPPCSPC (spatial pyramid pooling with cross stage partial connection) 模块^[12]、CBS 模块、ELAN-N 模块、上采样模块和 MP 模块。颈部网络采用自底向上和自顶向下的方式将上述模块组合在一起,通过对输入的特征图片使用卷积操作来计算目标的类别、位置和尺寸信息,以实现特征提取与特征融合。

预测头部分通过处理来自主干网络和颈部网络的特征图,使用 CIoU 损失函数计算预测框与真实框之间的几何位置^[14],并应用非极大值抑制去除冗余框,保留置信度最高的检测框。

2 改进的 YOLOv7 模型

尽管 YOLOv7 是一种精度和速度均较高的目标检测算法,但它在电力调度控制系统图中的图元识别方面存在以下不足:主干网络未能充分提取相似

图元的特征，对相似图元的识别效果有待提高；特征融合网络中的 Concat 结构未能充分利用特征图的多尺度信息，对小尺寸图元的识别效果较差；CIoU 损失函数在真实框和预测框的宽高比值相同的情况下，对边界框尺寸差异不敏感。针对上述问题，本文进行如下改进：使用 DRBNCSPELAN 模块替换

主干网络中的 ELAN 模块，以提高相似图元的识别效果；提出一种 IGFM 替代颈部网络中的 Concat，以提高小尺寸图元的识别效果；将 CIoU 损失函数替换为 MPDIoU，以改善图元边界框的识别效果。改进后的 YOLOv7 网络结构如图 1 所示，其中虚线框内为改进部分。

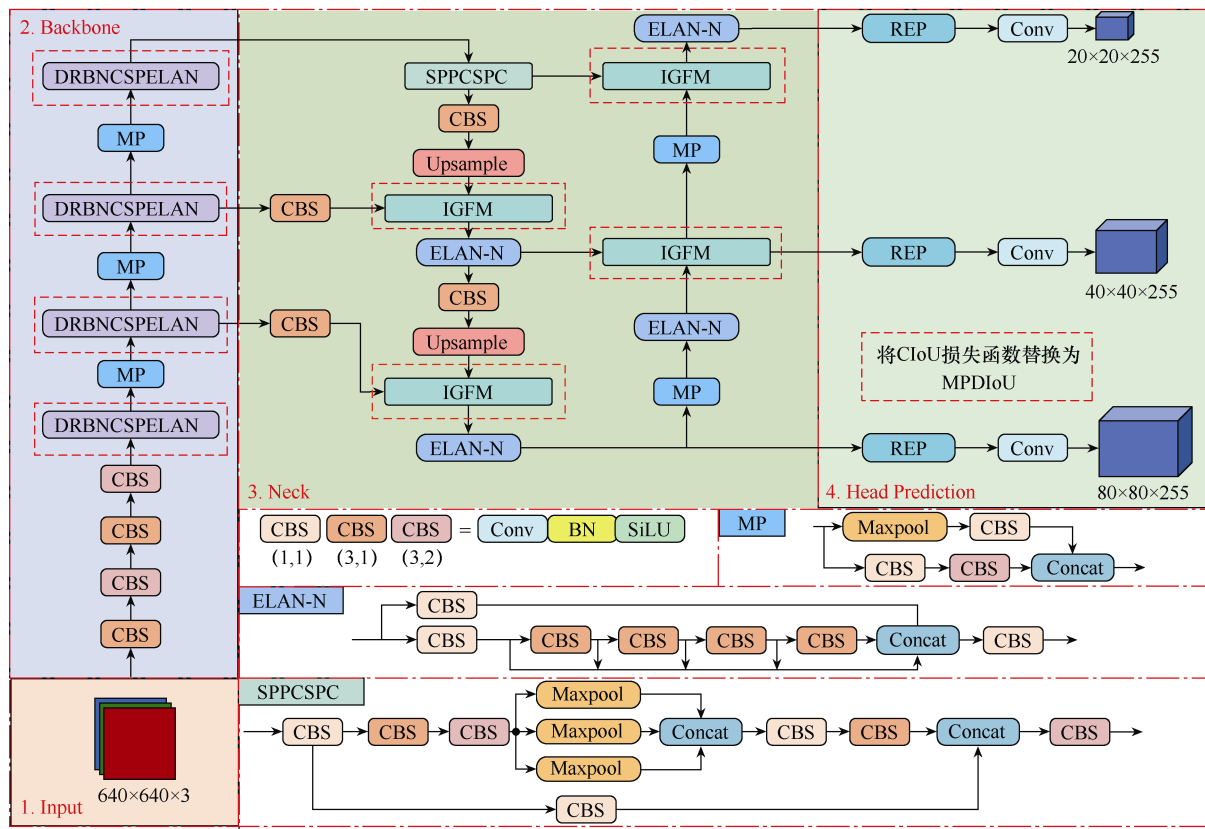


图 1 改进后的 YOLOv7 网络结构

2.1 特征提取网络的改进

CSPNet (cross stage partial network) 通过梯度分流策略将特征图一分为二，一部分直接跳跃连接，另一部分经深度特征提取后融合。该设计降低了计算成本并优化了梯度流，但导致一些浅层特征信息丢失。YOLOv7 的 ELAN 模块在 CSPNet 的基础上增加了多分支结构，通过将每个特征提取模块的中间结果进行融合，增强了浅层和深层特征的交互。GELAN (generalized efficient layer aggregation network) 结合 CSPNet 和 ELAN，通过改进特征融合方式，解决了浅层信息丢失的问题，同时保持了对多种任务和模块的适应能力^[15]，其结构如图 2 (a) 所示。本文提出的 DRBNCSPELAN 模块则将 GELAN 中的计算块 Any Block 替换为 DRBNCSP (dilated reparam block net with cross stage partial)

模块和 CBS 模块，结构如图 2 (b) 所示。

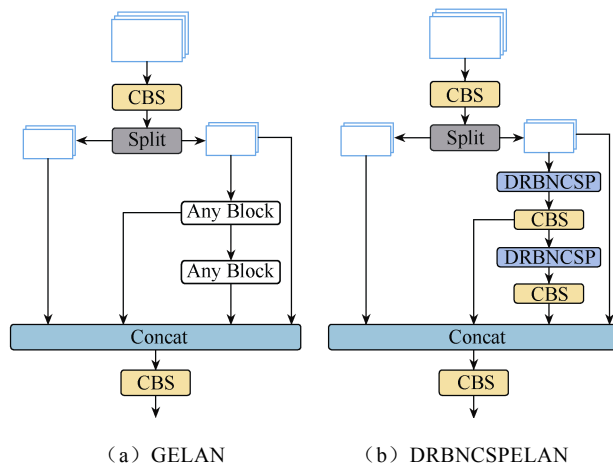


图 2 GELAN 和 DRBNCSPELAN 结构

在电力调度控制系统图的图元识别任务中,考虑到提升相似图元识别效果及模型的轻量化,本文用 DRBNCSPELAN 模块替换 ELAN 模块,作为新的特征提取模块。DRBNCSP 模块将输入特征图经过 CBS 和 DRBNBottleneck(dilated reparam block net bottleneck) 两个分支后进行特征融合,如图 3(a)所示。DRBNBottleneck 由核心的 DRBlock(dilated reparam block) 模块和 CBS 模块的残差连接构成,如图 3(b)所示。DRBlock 的结构如图 3(c)所示,它使用一个小核非膨胀卷积层(卷积核大小 $k=5$, 膨胀率 $r=1$)和多个小核(卷积核大小 $k=5, 3, 3$, 膨胀率 $r=1, 2, 3$)膨胀卷积层来增强大核非膨胀卷积层(卷积核大小 $k=7$, 膨胀率 $r=1$),即使用膨胀小核来重参数化非膨胀大核^[16]。大核卷积拥有更大的有效感受野,小核卷积更容易捕捉纹理特征,DRBlock 则是将二者结合,在增强大核捕捉稀疏模式的能力的同时,使网络可以提取高质量的特征,进一步提高对相似图元的识别能力。

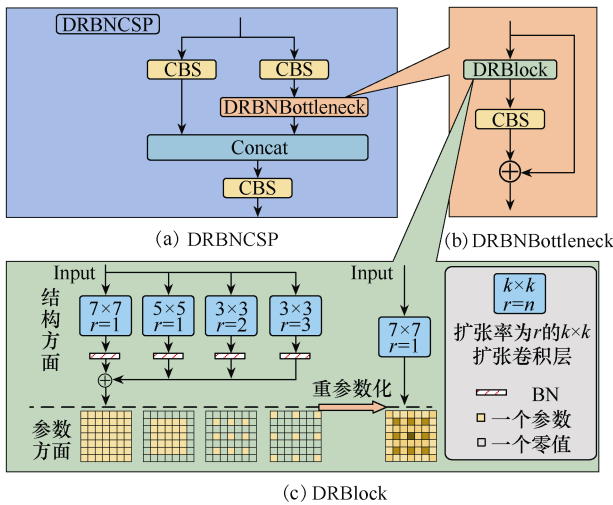


图 3 DRBNCSP、DRBNBottleneck 和 DRBlock 结构

DRBlock 模块通过重参数化技术,在训练阶段使用多个由卷积层构成的分支促进训练过程中的信息流动和梯度传播。在推理阶段将每个小核分支的参数重新参数化到大核主分支。这不仅减少了计算量和内存消耗,还提高了模型的推理效率。

2.2 特征融合网络的改进

由于电力调度控制系统图中小尺寸图元在整个图中占据的像素区域较小,导致网络无法充分捕捉小尺寸图元的语义信息,图元识别效果差。针对该问题,本文提出一种 IGFM,以一种更精细的特征融合方法来优化 Concat 结构。IGFM 包含 Concat 结

构、SE(squeeze-and-excitation)模块、Split 操作和双向信息交换操作,如图 4 所示。首先,将包含不同尺度信息的特征图作为输入,引入 SE 注意力模块,该模块能够在特征融合过程中捕捉并利用重要的上下文信息,从而增强特征表示的有效性,并引导模型学习小尺寸图元的信息^[17];接着,通过 Split 操作根据不同尺度信息的重要性分配权重,这种加权方式使融合后的特征图更能反映原始特征图中对于小尺寸图元重要的信息;最后,通过双向信息交换操作,促进不同尺度特征之间的信息交流,使网络能够提取到更加丰富的小尺寸图元的特征表示。

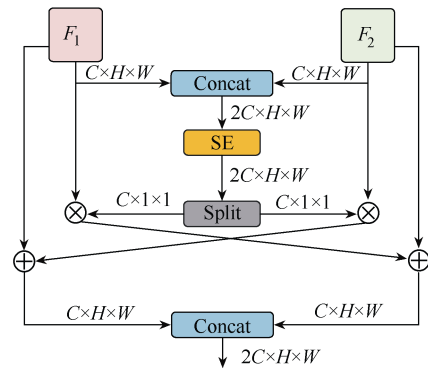


图 4 IGFM

SE 是一种专注于通道信息的轻量型注意力机制,通过捕捉通道间的相互依赖关系,动态地重新校准每个通道对特征的响应权重。本文利用 SE 注意力机制来提高模型对小尺寸图元重要特征的关注意度,同时抑制不相关的背景噪声,进一步提升小尺寸图元的识别能力。SE 注意力机制的原理如图 5 所示。

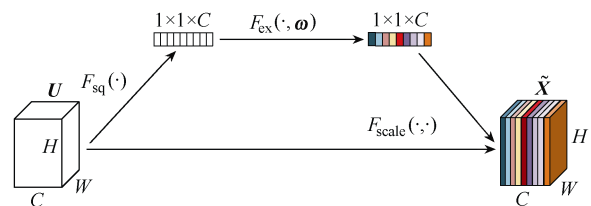


图 5 SE 注意力机制原理

Squeeze 部分将特征图 $U = [u_1 \ u_2 \ \dots \ u_c]$ 进行全局平均池化,生成一个向量 $Z = (z_1, z_2, \dots, z_c)$,其中 $U \in \mathbf{R}^{H \times W \times C}$, $Z \in \mathbf{R}^{1 \times 1 \times C}$, 计算式为

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

式中: $u_c(i, j)$ 为特征图 U 的第 c 个元素, $c=1, 2, \dots, C$;

$F_{\text{sq}}(\cdot)$ 为 Squeeze 操作; H 、 W 分别为特征图的高度、宽度; z_c 为 \mathbf{Z} 的第 c 个元素。

Excitation 部分由两层全连接层构成, 用于学习通道间的重要性。计算式为

$$\mathbf{s} = F_{\text{ex}}(\mathbf{Z}, \boldsymbol{\omega}) = f_1[\boldsymbol{\omega}_2 f_2(\boldsymbol{\omega}_1 \mathbf{Z})] \quad (2)$$

式中: $F_{\text{ex}}(\cdot)$ 为 Excitation 操作; \mathbf{Z} 为输入向量; $\boldsymbol{\omega}$ 为权重矩阵; $\boldsymbol{\omega}_1 \in \mathbf{R}^{r \times C}$ 和 $\boldsymbol{\omega}_2 \in \mathbf{R}^{C \times r}$ 为两个全连接层权重矩阵; f_1 为 ReLU (rectified linear unit) 函数; f_2 为 sigmoid 函数; \mathbf{s} 为每个通道重要程度的权重。

Scale 部分计算 \mathbf{u}_c 和 \mathbf{s}_c 的乘积, 即特征图 \mathbf{U} 和权重 \mathbf{s} 通过 Scale 操作生成 SE 的输出 $\tilde{\mathbf{X}} = [\tilde{x}_1 \ \tilde{x}_2 \ \cdots \ \tilde{x}_C]$, 其中 $\tilde{\mathbf{X}} \in \mathbf{R}^{H \times W \times C}$ 。计算式为

$$\tilde{x}_c = F_{\text{scale}}(\mathbf{u}_c, \mathbf{s}_c) = \mathbf{s}_c \mathbf{u}_c \quad (3)$$

式中: $F_{\text{scale}}(\cdot)$ 为 Scale 操作; \mathbf{u}_c 和 \mathbf{s}_c 分别为通道 c 中的特征映射和权重; \tilde{x}_c 为 $\tilde{\mathbf{X}}$ 的第 c 个元素。

2.3 MPDIoU 损失函数

YOLOv7 模型采用 CIoU 损失函数来计算边界框损失。CIoU 损失函数为

$$f_{\text{IoU}} = \frac{B_{\text{gt}} \cap B_{\text{pred}}}{B_{\text{gt}} \cup B_{\text{pred}}} \quad (4)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w^{\text{pred}}}{h^{\text{pred}}} \right)^2 \quad (5)$$

$$\alpha = \begin{cases} 0 & f_{\text{IoU}} < 0.5 \\ \frac{v}{(1-f_{\text{IoU}})+v} & f_{\text{IoU}} \geq 0.5 \end{cases} \quad (6)$$

$$L_{\text{CIoU}} = 1 - f_{\text{IoU}} + \frac{\rho^2(B_{\text{gt}}^{\text{ctr}}, B_{\text{pred}}^{\text{ctr}})}{l^2} + \alpha v \quad (7)$$

式中: B_{gt} 和 B_{pred} 分别为真实边界框和预测边界框; f_{IoU} 为真实边界框和预测边界框的重叠率; w^{gt} 和 h^{gt} 分别为真实边界框的宽和高; w^{pred} 和 h^{pred} 分别为预测边界框的宽和高; v 为用于评估宽高比是否一致的函数; α 为权重函数; $\rho^2(B_{\text{gt}}^{\text{ctr}}, B_{\text{pred}}^{\text{ctr}})$ 为真实框和预测框的中心点之间的欧氏距离; l 为真实框和预测框的最小外接矩形的对角线长度; L_{CIoU} 为 CIoU 损失函数值。

根据式(4)~式(7), 如果真实框和预测框的宽高比值相同, 则 $v=0$ 。这种情况下, CIoU 损失中的宽高比损失 αv 恒为 0, 导致这部分损失不会贡献任何梯度用于更新模型参数, 使模型无法学习宽高比方面的特征, 这将限制收敛速度和识别精度。为了解决这个问题, 文献[18]提出了 MPDIoU 损失函数, 通过最小化预测边界框和真实边界框之间的左上和右下点的欧式距离, 更好地训练深度学习模型。MPDIoU 示意图如图 6 所示。

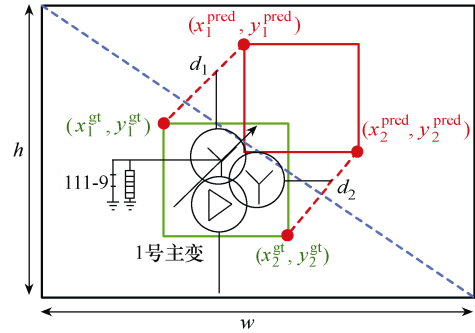


图 6 MPDIoU 示意图

MPDIoU 的计算式为

$$d_1^2 = (x_1^{\text{pred}} - x_1^{\text{gt}})^2 + (y_1^{\text{pred}} - y_1^{\text{gt}})^2 \quad (8)$$

$$d_2^2 = (x_2^{\text{pred}} - x_2^{\text{gt}})^2 + (y_2^{\text{pred}} - y_2^{\text{gt}})^2 \quad (9)$$

$$L_{\text{MPDIoU}} = 1 - \left(f_{\text{IoU}} - \frac{d_1^2}{h^2 + w^2} - \frac{d_2^2}{h^2 + w^2} \right) \quad (10)$$

式中: $(x_1^{\text{pred}}, y_1^{\text{pred}})$ 和 $(x_1^{\text{gt}}, y_1^{\text{gt}})$ 分别为预测框和真实框的左上点坐标; d_1 为预测框和真实框左上点之间的欧式距离; $(x_2^{\text{pred}}, y_2^{\text{pred}})$ 和 $(x_2^{\text{gt}}, y_2^{\text{gt}})$ 分别为预测框和真实框的右下点坐标; d_2 为预测框和真实框右下点之间的欧式距离; w 和 h 为输入图片的宽和高; L_{MPDIoU} 为 MPDIoU 损失函数值。

MPDIoU 损失函数不仅考虑了 CIoU 损失函数的所有因素, 而且解决了预测边界框与真实边界框具有相同宽高比的问题, 能够改善图元边界框的识别效果。

3 实验结果与分析

基于上述改进的 YOLOv7 模型, 利用电力调度控制系统图数据集验证改进的有效性。

3.1 数据集说明及预处理

数据集是一组 JPG 或 PNG 格式的电力调度控制系统图, 共 91 张, 图像分辨率跨度较大, 大多数集中在 $1\,000 \times 1\,000 \sim 8\,700 \times 6\,100$ 像素范围内。对数据集进行标注, 产生了 19 444 个标注框, 包含 13 种不同的类别。图元类别及标签见表 1。

表 1 图元类别及标签

图元类别 (英文名称)	标签
断路器 (cbreaker)	0
隔离开关 (disconnecter)	1
接地开关 (ground disconnecter)	2
两卷变压器 (transformer2)	3
三卷变压器 (transformer3)	4
电容器 (capacitor)	5
电抗器 (reactor)	6
发电机 (generator)	7
线路末端 (line end)	8
避雷器 (lightning arrester)	9
电压互感器 (potential transformer)	10
文本 (text)	11
功率补偿器 (static var generator)	12

由于原始图片的分辨率大多在 $1\,000 \times 1\,000$ 以上, 而 YOLOv7 网络需要输入图片的尺寸为 640×640 , 直接将原始图片输入网络中会导致图元识别效果较差。为了解决这一问题, 文献[19]在图像处理阶段采用重叠滑窗分割。在此基础上, 本文采用一种双向重叠滑窗分割方法, 具体做法是: 设定窗口重叠率为 50%, 滑窗的大小为 640×640 ; 一方面从左上角开始, 依次从左到右、从上到下进行滑窗分割; 另一方面从右下角开始, 依次从右到左、从下到上进行滑窗分割。设置窗口重叠率可避免一些图元从中间分割, 采用双向重叠滑窗分割可以避免边界的图元被分割掉, 从而保证数据的完整性。为方便理解该过程, 双向滑窗分割过程示意图如图 7 所示。经数据预处理后, 得到 5 748 张图片, 67 089 个标注框。

本文将经过预处理的自建数据集按照 8:1:1 划分训练集、验证集和测试集, 先后用于模型的训练、验证和测试, 进行模型的训练、参数的调整及预测结果的输出。

3.2 实验参数设置及评价指标

输入图片尺寸为 640×640 , 训练轮数设置为 300, 批次大小设置为 8, 选择 SGD (stochastic gradient descent) 优化器, 初始化学学习率为 0.01, 使用余弦函数来调整学习率, 使用马赛克进行数据增强, 开启自动锚框。

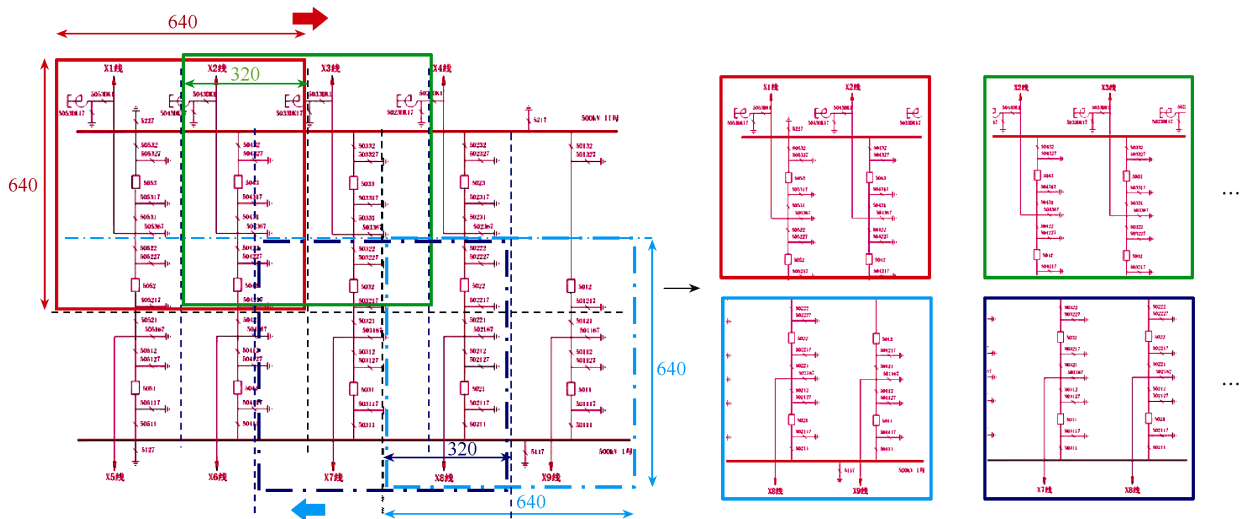


图 7 双向滑窗分割过程示意图

本文选择的评价指标包括精确率 (precision) P 、召回率 (recall) R 、精确率和召回率的调和平均数 F_1 、参数量 (parameters, Paras)、每秒浮点运算次数

(floating point operations per second, FLOPs)、画面每秒帧数 (frames per second, FPS) 和 $f_{IoU}=0.5$ 时的平均精确率均值 (mean average precision@0.5,

mAP@0.5)。其中, P 、 R 和 F_1 的计算公式为

$$P = \frac{T_p}{T_p + F_p} \quad (11)$$

$$R = \frac{T_p}{T_p + F_N} \quad (12)$$

$$F_1 = \frac{2PR}{P+R} \quad (13)$$

式中: T_p 为预测和真实样本均为正样本的样本数; F_p 为预测样本为正样本但真实样本为负样本的样本数; F_N 为预测样本为负样本但真实样本为正样本的样本数。

平均精确率均值是衡量目标检测模型性能的重要指标, 其值越大, 识别效果越好。通常, 性能较好的算法的上述指标均高, 综合分析上述指标可较好地反映其性能。计算公式为

$$A_p = \int_0^1 P(R)dR \quad (14)$$

$$m_{AP} = \frac{1}{N} \sum_{j=1}^N A_p(j) \quad (15)$$

式中: A_p 为 P - R 曲线围成的面积; m_{AP} 为所有类别 A_p 的平均值; N 为样本的类别总数。

3.3 实验结果

为了评估对 YOLOv7 网络所做各项改进的有效性, 分别对改进部分进行消融实验, 结果见表 2, 其中“√”表示使用了该改进方法。

表 2 消融实验结果

A	B	C	$P/\%$	$R/\%$	$F_1/\%$	Paras	FLOPs	mAP@0.5/%
			86.4	92.4	89.3	36.5×10^6	103.4×10^9	94.1
√			87.4	95.7	91.4	35.2×10^6	97.5×10^9	97.0
√	√		88.9	95.9	92.3	35.4×10^6	97.5×10^9	97.4
√	√	√	91.5	95.5	93.5	35.4×10^6	97.5×10^9	97.6

注: A 表示替换为 DRBNCSPPLAN 模块; B 表示替换为 IGFM; C 表示替换为 MPDIoU 损失函数。

从表 2 数据可知, 本文改进后模型与基准模型相比, P 、 R 、 F_1 和 mAP@0.5 分别提升了 5.1 个百分点、3.1 个百分点、4.2 个百分点和 3.5 个百分点, 参数量减少了 3.0%, 计算量减少了 5.7%, 显著提升了模型对于电力调度控制系统图的图元识别精度。将主干网络中的 ELAN 模块替换为 DRBNCSPPLAN 模块, P 、 R 、 F_1 和 mAP@0.5 均有提升, 参数量和计算量也均降低, 这得益于 DRBNCSPPLAN 模块内部使用膨胀小核来重参数化非膨胀大核, 增强了

大核捕捉稀疏模式的能力, 使网络可以提取高质量的特征, 进而提高相似图元的识别能力。在此基础上引入 IGFM, 结果表明, 在少量增加参数的情况下, P 、 R 、 F_1 和 mAP@0.5 分别提升了 1.5 个百分点、0.2 个百分点、0.9 个百分点和 0.4 个百分点, 说明引入 IGFM 能够使网络在特征融合过程中捕捉并利用重要的多尺度信息, 并有效引导模型学习小尺寸图元的信息, 从而提高了模型对小尺寸图元的识别效果。在此基础上, 将基准模型的 CIoU 替换为 MPDIoU, 除 R 降低 0.4 个百分点外, 其他各项精度指标均有提升, 参数量和计算量不变, 表明这一改进能够进一步提升模型对电力调度控制系统图元的识别精度。

为了验证 MPDIoU 损失函数的综合性能, 与 GIoU (generalized intersection over union)、SIoU (scylla intersection over union)、EIoU (efficient intersection over union) 和 DIoU (distance intersection over union) 4 种损失函数进行对比实验, 实验结果见表 3。

表 3 不同损失函数的对比实验结果

损失函数	$P/\%$	$R/\%$	mAP@0.5/%
GIoU	90.4	95.2	97.0
SIoU	90.3	95.6	97.0
EIoU	89.1	95.3	97.3
DIoU	89.8	95.1	96.5
MPDIoU	91.5	95.5	97.6

由表 3 可知, 使用 MPDIoU 损失函数改进后的模型的 P 和 mAP@0.5 均达到了最高。与 SIoU 相比, 本文方法的 R 降低了 0.1 个百分点, P 和 mAP@0.5 分别提高了 1.2 个百分点和 0.6 个百分点。由于 P 和 mAP@0.5 这两项指标通常更能反映模型的综合性能, 所以本文方法的综合性能优于 SIoU。综上所述, 本文方法的综合性能最优。引入 MPDIoU 损失函数改善了图元边界框的回归损失, 提高了模型对图元的识别精度。改进后的 YOLOv7 模型, 对各类图元的识别结果见表 4。

由表 4 可知, 每一类图元的 mAP@0.5 都达到了 92% 以上, 改进后的模型能够精准识别电力调度控制系统图的图元。

最后, 为进一步验证改进的 YOLOv7 模型对电力调度控制系统图中图元的识别性能, 本文选择 Faster-RCNN、SSD (single shot multibox detector)、

表 4 各类图元的识别结果

图元类别	mAP@0.5/%
断路器 (cbreaker)	97.8
隔离开关 (disconnecter)	94.9
接地开关 (ground disconnecter)	96.7
两卷变压器 (transformer2)	95.9
三卷变压器 (transformer3)	98.1
电容器 (capacitor)	98.3
电抗器 (reactor)	99.6
发电机 (generator)	99.5
线路末端 (line end)	92.5
避雷器 (lightning arrester)	98.0
电压互感器 (potential transformer)	99.1
文本 (text)	98.5
功率补偿器 (static var generator)	99.5

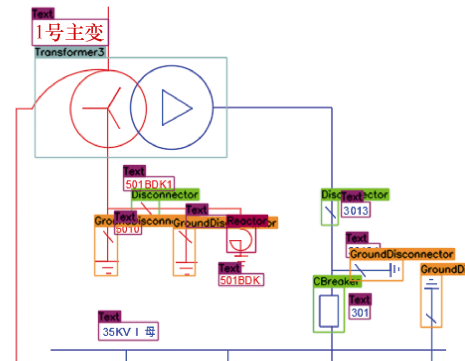
YOLOv5n (you only look once v5 nano) 作为对比模型。不同模型的识别结果见表 5。

表 5 不同模型的识别结果

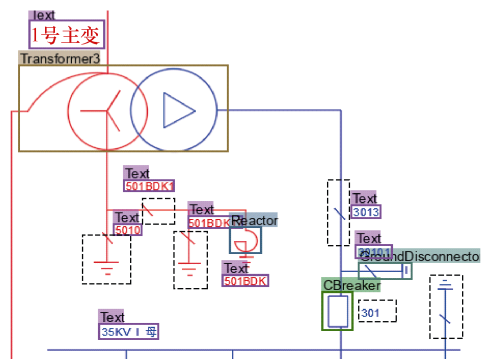
模型	P/%	R/%	Paras	FLOPs	FPS	mAP@0.5/%
Faster-RCNN	56.2	80.7	28.4×10^6	948.3×10^9	14.4	69.8
SSD	60.4	82.4	26.2×10^6	89.5×10^9	46.9	88.8
YOLOv5n	82.0	93.7	1.8×10^6	4.2×10^9	138.9	93.7
改进 YOLOv7	91.5	95.5	35.4×10^6	97.5×10^9	137.0	97.6

由表 5 可知, Faster-RCNN 和 SSD 算法对图元的识别精度较低, 识别速度较慢。YOLOv5n 虽然具有最快的识别速度, 并且参数量和计算量都很小, 但在识别精度方面还有待提升。本文方法的 P 、 R 和 $mAP@0.5$ 都达到了最高, 识别速度较快, 能够在保证图元识别精度的同时, 满足实时性的要求。为了直观展示所提方法的优越性, 选取一张具有多种图元的电力调度控制系统图, 将所改进的模型与 SSD、Faster-RCNN 和 YOLOv5n 进行识别效果对比, 如图 8 所示, 图中方形虚线框表示未被识别的图元, 椭圆形虚线框表示识别错误的图元。

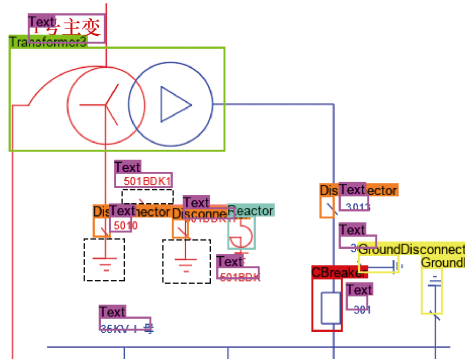
由图 8 可知, 改进的 YOLOv7 模型实现了全部图元的准确识别, 预测框定位准确, 如图 8 (a) 所示。SSD 识别到的图元最少, 有 6 个图元未被识别, 预测框定位准确, 如图 8 (b) 所示。Faster-RCNN 有 1 个隔离开关和 2 个接地开关未被识别, 且预测框定位不够准确, 如图 8 (c) 所示。YOLOv5n 有 1 个三卷变压器、1 个文本和 1 个接地开关未被识别,



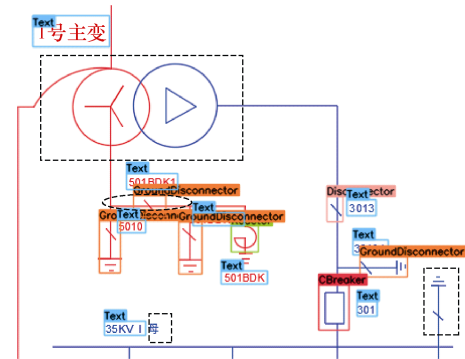
(a) 改进的 YOLOv7



(b) SSD



(c) Faster-RCNN



(d) YOLOv5n

图 8 不同识别模型效果对比

中间的隔离开关出现错检, 预测框定位准确, 如图 8 (d) 所示。综上所述, 本文所改进的 YOLOv7 具

有较好的电力调度控制系统图元识别能力。

4 结论

本文提出了一种基于改进YOLOv7的电力调度控制系统图元识别方法。与基准模型相比,所提方法的参数量和计算量分别减少了3.0%和5.7%, P 、 R 和 $mAP@0.5$ 分别提高了5.1个百分点、3.1个百分点和3.5个百分点,每种图元的 $mAP@0.5$ 都在92%以上,提高了模型对相似图元及小尺寸图元的识别效果,有效改善了模型对电力调度控制系统图元的识别能力。主要结论如下:

1) 利用轻量化的DRBNCSPELAN模块,在减少了模型参数量和计算量的同时,使网络可以提取高质量的特征,提高了相似图元的识别能力。

2) 通过引入一种融合SE注意力机制的IGFM,增强了小尺寸图元的信息在特征融合网络中的流动,改善了小尺寸图元的识别效果。

3) 通过MPDIoU损失函数,改善了图元边界框的回归损失,能够实现更精准的电力调度控制系统图中的图元识别。

参考文献

- [1] 张柯,李金旗.数字站建设方案及技术探究[J].电气技术,2023,24(12):68-74.
- [2] 李都红,郭树林,解凯,等.基于“四型机场”理念的某机场10kV变电站综合自动化管理系统设计与应用[J].电气技术,2021,22(2):84-90.
- [3] 陈晓杰,方贵盛.一种基于图元结构关系的电气草图符号识别方法[J].机电工程,2017,34(8):823-828.
- [4] 王玉豪,方贵盛.基于DAGSVM和决策树的电气草图符号识别[J].轻工机械,2017,35(4):56-59.
- [5] 肖豆,侯晓荣.基于PHOG特征的电路图中电气符号识别[J].舰船电子工程,2017,37(1):90-93.
- [6] 苟军年,杜慷慷,刘力.基于改进掩膜区域卷积神经网络的输电线路绝缘子自爆检测[J].电工技术学报,2023,38(1):47-59.
- [7] 郑含博,胡思佳,梁炎桑,等.基于YOLO-2MCS的输电线路走廊隐患目标检测方法[J].电工技术学报,2024,39(13):4164-4175.
- [8] 仲林林,胡霞,刘柯好.基于改进生成对抗网络的无人机电力杆塔巡检图像异常检测[J].电工技术学报,2022,37(9):2230-2240,2262.
- [9] 李昊,管萸,王杉,等.电力系统厂站接线图拓扑关系检测技术[J].北京航空航天大学学报,2021,47(3):531-538.
- [10] 徐剑,张皓,徐航,等.一种基于Faster RCNN的电网图元识别方法[J].计算机与现代化,2021(12):53-57.
- [11] 王梓耀,罗庆全,萧文聪,等.电网厂站接线图人工智能识别关键方法[J].电力系统自动化,2023,47(2):115-124.
- [12] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC Canada, 2023: 7464-7475.
- [13] 林丽霞,吴悦园.基于改进YOLOv7的电力设备红外过热缺陷检测方法[J].电气技术,2024,25(1):42-47.
- [14] ZHENG Zhaohui, WANG Ping, REN Dongwei, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2022, 52(8): 8574-8586.
- [15] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information[C]//European Conference on Computer Vision, Milan, Italy, 2024: 1-18.
- [16] ZHANG Yiyuan, DING Xiaohan, YUE Xiangyu. Scaling up your kernels: large kernel design in ConvNets towards universal representations[EB/OL]. [2025-03-10]. <https://doi.org/10.48550/arXiv.2410.08049>.
- [17] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018: 7132-7141.
- [18] MA Siliang, XU Yong. MPDIoU: a loss for efficient and accurate bounding box regression[EB/OL]. [2025-03-10]. <https://doi.org/10.48550/arXiv.2307.07662>.
- [19] LAVANYA K, KARNICK S, GHALIB M R, et al. A novel method for vehicle detection in high-resolution aerial remote sensing images using YOLT approach[J]. Multimedia Tools and Applications, 2022, 81: 23551-23566.

收稿日期:2024-12-27

修回日期:2024-12-30

作者简介

张文广(1975—),男,山东省烟台市人,教授,博士,研究方向为人工智能技术在电力系统的应用。