

基于 Mask R-CNN 模型的变电设备 红外图像实例分割算法优化研究

陈彬彬

(国网江苏省电力有限公司无锡供电分公司)

摘要：本文针对变电设备红外图像的特点，提出基于 Mask R-CNN 模型的实例分割算法优化方法。主要通过引入可变形卷积和改进的注意力机制，优化模型的特征提取能力。可变形卷积增加了感受野的灵活性，更好地适应了目标的几何变形；改进的注意力机制通过并行连接通道和空间注意力模块，提升了对显著特征的提取能力。实验结果表明，在红外图像数据集上，该方法在不同主干网络下的分割精度均有所提升，尤其在 AP50、AP75 等评价指标上优于传统的 Mask R-CNN 模型。研究结果为变电设备红外图像分割算法的进一步优化提供了新的思路和方法。

关键词：Mask R-CNN；红外图像；可变形卷积；注意力机制；实例分割

2025.07.DQGY
51

0 引言

Mask R-CNN 自 2017 年由 He 等人提出后，迅速在实例分割领域中获得广泛应用。其通过引入 ROI Align 和全连接掩膜分支，使得在目标检测过程中能够同时生成高质量的实例掩膜。针对红外图像中变电设备几何形状复杂、目标与背景对比度低的挑战，传统卷积操作和掩膜预测存在一定的局限性。为提高模型对变电设备红外图像的分割效果，本文对 Mask R-CNN 模型进行改进，具体包括引入可变形卷积和基于注意力机制的残差模块优化。这些改进旨在提升对几何多样性和边缘模糊特征的适应性，从而更好地解决变电设备红外图像中的实例分割问题。

1 基于 Mask R-CNN 模型的实例分割算法

该模型包括卷积神经网络，RPN 网络，感兴趣区域对齐和结果输出四个模块。卷积神经网络：利用残差网络（Res-Net）、特征金字塔（FPN）等方法，实现多尺度特征的提取与融合，构建相应的特征图谱。RPN 网络：通过预先设定的各种宽度高度比的锚框架，在特征图上滑动，产生多个候选点，然后对候选点（前景与背景）进行双重分类，并对候选框的位置与尺寸进行校正，并采用 NMS 算法来选择更准确的候选盒子。感兴趣区域对齐：利用双线性内插法求出各样本点的像素值，并保持浮动点式坐标，将所得到的候选盒与原影像中的候选点进行主像素匹配。输出部分：采用全卷积神经网络构成的掩膜支路，对物体进行像素级划分，采用两个完全连通的层次对物体进行分类，并对物体进行边缘盒回归处理。Mask

R-CNN 模型结构如图 1 所示。

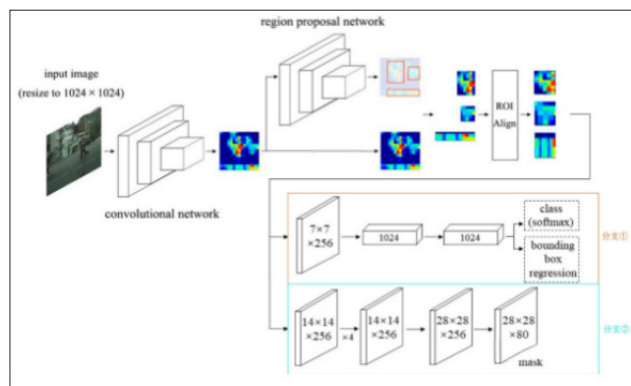


图 1 Mask R-CNN 模型结构图

2 Mask R-CNN 模型改进与优化

2.1 引入可变形卷积

传统卷积公式在传统的二维卷积运算中，卷积核通过在特征图的一个固定规则区域上滑动来进行采样。该区域由栅格 R 定义，栅格内的每个采样点 p_n 都对应着一个输入特征图上的点，计算时通过卷积核的权重 $w(p_n)$ 对这些点进行加权求和，从而得到输出值 $y(p_0)$ 。具体来说，对于输入特征图中的每个位置 p_0 ，输出值可以表示为：

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n) \quad (1)$$

式中， x 为输入特征图； w 为卷积核的权重； p_n 为卷积核在特定区域内的采样点。通过加权求和操作，卷积核能够在不同位置提取到局部的空间信息。然而，这种操作只是在固定的规则区域内进行，无法对复杂几何形状的目标自适应调整，尤其在红外图像中，当目标边界模糊且背景与目标差异不明显时，效果可能有限。可变形卷积公式为了增强卷积操作在复杂几何结构上的适应性，可变形卷积引入了偏移量， Δp_n 使得每个采样点 p_0 不再固定，而是根据物体的几何特征动态调整。这一调整可以有效扩大卷积核的感

知范围，从而适应目标形状的变化。可变形卷积的输出公式为：

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (2)$$

通过在采样点上加入偏移量 Δp_n ，卷积核能够对物体的边界特征进行自适应的调整，从而更好地提取出与背景有明显区别的目标特征。这种灵活的采样方式提高了特征表达的能力，使模型更好地适应红外图像中复杂的目标形态。加入偏移量后的感知野公式引入偏移量后，卷积核的感知野会变得更加灵活和广泛，但也会因此带来背景信息的干扰。为了进一步优化这种感知机制，可以在卷积过程中加入一个额外的权重系数 Δm_k ，以控制感知野的范围。这一权重系数会根据感知野的变化自适应调整，公式表示为：

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_k \quad (3)$$

可变形卷积的实现过程如图 2 所示。

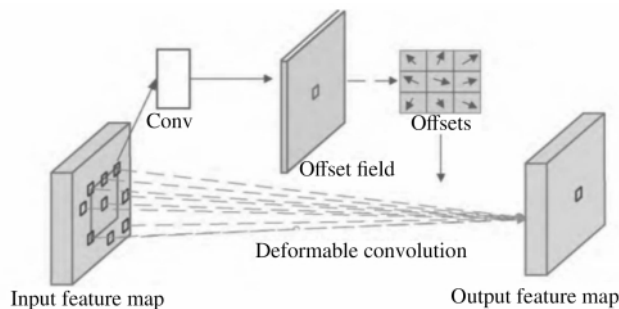


图 2 可变形卷积实现过程

通过对传统的 2D 卷积和形变卷积的感知野对比分析，可以看到，变形卷积可以更好地适应电力设备在红外图像中的不同几何特性。

2.2 基于注意力机制的残差模块构建

深度学习的注意力机制类似于人类的选择性注意机制，它使得模型能够从海量信息中选取对当前任务更有意义的信息。在图 3 中，把“通道注意”

与“空间注意”（CBAM）两种机制相结合，构建了 CBAM 网络。本课题针对红外图像中背景和目标对比强烈，样本之间仅依靠边缘特征进行识别的难题，提出将注意力机制引入到模型中，提高模型的高效特征提取能力。

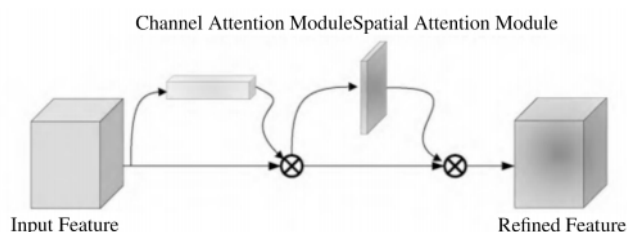


图 3 CBAM 实现原理

通过加入 Δm_k ，模型可以在特征图中选择更加重要的区域进行感知，从而减少背景信息对目标特征提取的干扰。这种改进使得卷积操作不仅能够自适应物体形状变化，还能够有效过滤掉不相关的背景信息，提高了红外图像中目标的边缘检测能力。ICBAM 改进卷积块注意力模型公式为了解决传统卷积注意力机制在信息提取过程中的不足，ICBAM 模型引入了并行的通道注意力和空间注意力模块。这些模块通过全局最大池化、全局平均池化以及可变形卷积对图像特征进行融合，最终输出加权的特征图。其计算公式为：

$$f_{out}^j = f^j \otimes M_c^j(f^j) + f^j \otimes M_s^j(f^j) \quad (4)$$

式中， f_{out}^j 为经过注意力机制加权后的输出特征图； M_c^j 和 M_s^j 分别为通道和空间注意力模块的输出； $\otimes$ 为逐元素相乘操作。这种融合方式能够充分利用图像中的局部和全局信息，有效提升对重要特征的提取能力，尤其是在复共的红外图像背景中，ICBAM 能够更好地定位与分割目标。

$$M_c^j(f^j) = \sigma_c^j \left(\begin{matrix} F_c^j(F_{cmax}^j(f^j)) \\ + F_c^j(F_{cavg}^j(f^j)) \end{matrix} \right) \quad (5)$$

其计算过程可以表示为：空间注意力模块的输

出 $M_c^j(f^j)$ 通过并行处理特征图的全局最大池化、全局平均池化以及可变形卷积得到三个相同大小的特征图。这三个特征图在通道维度上进行拼接，得到新的特征图。然后，拼接后的特征图通过一个卷积核大小为 7 的卷积层，最后经过 Sigmoid 激活函数生成最终的输出。① $F_{cmax}^j(f^j)$ 表示对输入特征图进行全局最大池化操作，得到特定的特征图。② $F_{cavg}^j(f^j)$ 表示对输入特征图进行全局平均池化操作，得到另一特征图。③ $F_{cdeform}^j(f^j)$ 表示通过可变形卷积处理输入特征图，得到第三个特征图。将这些特征图分别用函数处理后，将结果在通道维度拼接，生成新的特征图，经过卷积核大小为 7 的卷积操作和 Sigmoid 激活函数，输出 $M_c^j(f^j)$ 。

3 实验结果及分析

在本文中，使用 NVIDIA 2080 Ti 的专用加速卡来进行测试，所使用的操作系统为 Ubuntu18.04.5 LTS，使用 CUDA11.2 进行加速；Pytorch 和 Pytorch 都采用 Python3.8 语言和 Web 开发框架。本项目以变压器绝缘体、电流互感器、电压互感器、套管、避雷器这 5 种典型变压器关联设备为研究对象开展案例分割试验。其中训练集与验证集样本图像分别为 1548 张和 387 张，数量比例为 4:1。利用现有的典型案例分割方法（COCO）评价标准，对所建模型进行检验。本项目拟以 MaskR-CNN 为基本架构，训练过程 batchsize 为 2，优化器为 SGD，设定学习速率 2.5×10^{-3} ，动量值为 0.9，循环迭代 60 次。为了检验在算法中加入不同的改进方法的效果，设计了以下试验来比较和分析。

3.1 不同主干网络消融实验

在此基础上，利用 ResNet50-FPN、ResNet101-FPN 以及 ResNet152-FPN 等构建多目标 R-CNN 的骨干网，对红外影像中的变电设备进行分块，并开展模

块消融性试验。以骨干网为基础，用形变卷积代替局部卷积层；基于骨干网增加一种改进的注意力机制；基于骨干网，用形变卷积代替局部卷积，并在各残差模型中加入一种新的注意机制。

表 1 不同主干网络的分割性能对比

(单位: %)

主干网络	DCN	ICBAM	AP50: 95	AP50	AP75
ResNet50-FPN	√	√	63.6	90.1	77.4
			64.9	91.0	79.1
			64.7	90.3	80.7
ResNet101-FPN	√	√	65.5	91.6	80.8
			64.4	90.8	77.9
			65.3	91.6	79.3
ResNet101-FPN-DCN-ICBAM	√	√	65.1	91.2	80.7
			66.4	92.4	80.8
			64.9	91.5	77.6
ResNet152-FPN	√	√	65.6	91.2	79.1
			65.3	91.7	79.5
			66.2	91.4	80.4

在表 1 中，展示几种不同的方法在分割准确度上的比较。表格中的 AP50：95 代表 AP 在 0.05 的步进范围内，将 AP 从 0.50 至 0.95 进行计算，并对其进行平均值。在表格中，AP50 与 AP75 代表根据 IOU 的 0.50 与 0.75 所算得的 AP。从表 1 可以看出，在 MaskR-CNN 模型中，AP50：95，AP50，AP75 都比标准骨干网提高了很多，在采用两种方法后，AP50：95，AP50，AP75 都有较大提高。在采用 ResNet101-FPN-DCN-ICBAM 作为骨干的 MaskR-CNN 模型中，AP50：95，AP50，AP75 分别比 ResNet101-FPN 标准骨干网提高 2.0%，1.6%，AP75 提高 2.9%，并优于其他骨干网。鉴于变形卷积可以提升网络对物体几何特性的自适应性，降低背景信息的干扰，以及注意机制可以提升物体的表达能力，所以，本项目拟采用基于形变卷积的方法，结合改进的注意力机制，对模型进行优化。

3.2 损失函数消融实验

在此基础上，对多任务输出支路的损耗函数进行了优化。本项目拟以具有较高精度的 ResNet101-FPN-DCN-ICBAM 为研究对象，以 $\alpha=1$ ， β 值为 0.2， β 值为 0~1.0，以 AP50：95 为研究对象，对其进行评估。AP50：95 的值随着色温的升高而变小，AP50：95 的最大值为 67.9%，最小值为 64.1%。在 $\beta=0$ 的情况下，不计算所得到的输出单位的 Lbox 的丢失；在 $\eta=0$ 时，任务输出支路的损耗函数为刘-拉普拉斯；在 $\eta=1$ 时，mask 输出支路的损耗函数是 Lbce。在该模型下，评估指数 AP50：95 达到 67.9%，并分别达到 91.8% 和 82.1%，与 ResNet101-FPN 标准骨干网相比，分别提升 3.5%，1.0%，4.2%，表明该模型更注重对前景的挖掘，能更好地考虑到预测掩模与实际掩模的重叠度，从而提升网络分割的准确性。

4 结束语

本文通过对 Mask R-CNN 模型的优化，提升了其在变电设备红外图像实例分割中的性能。引入可变形卷积后，模型在处理目标几何特征的多样性方面表现出更强的适应性，降低了背景信息的干扰。改进的注意力机制有效提高了对关键特征的提取能力，增强了分割的准确性。实验结果表明，结合这两种策略后，模型在多个指标上均取得了显著提升，尤其是引入 ResNet101-FPN-DCN-ICBAM 主干网络时，分割效果最为显著。这一研究不仅验证了模型改进的有效性，也为进一步提升红外图像实例分割算法提供了理论支持和实践参考。

参考文献

[1] 刘传洋, 吴一全. 基于红外图像的电力设备识别及发热故障诊断方法研究进展 [J]. 中国电机工程学报, 2025, 45 (6): 2171-2196.

[2] 李雪寒, 刘沁怡, 杨晓彤, 等. 基于 Mask R-CNN 与改进 BP 神经网络联合算法的变压器套管红外热故障诊断 [J]. 上海电力大学学报, 2023, 39 (6): 591-598.

[3] 王坚俊, 孙林涛, 刘昌标, 等. 基于改进 Mask R-CNN 的变压器绝缘套管故障智能诊断 [J]. 浙

江电力, 2022, 41 (8): 87-94.

[4] 李冰, 王天, 杨珂, 等. 基于改进 Mask R-CNN 的变电设备红外图像实例分割算法 [J]. 华北电力大学学报 (自然科学版), 2023, 50 (2): 91-99.

[5] 叶剑涛, 杨为, 柯艳国, 等. 基于红外图像实例分割的敞开式开关设备发热缺陷智能诊断方法 [J]. 高压电器, 2021, 57 (12): 201-208.

(收稿日期: 2024-10-22)

(上接第 44 页)

[6] 董光德, 徐泽众, 孙鹏菊, 等. 弱电网下 LCL 型并网逆变器高鲁棒性并网电流反馈有源阻尼策略 [J]. 电网技术, 2024, 48 (4): 1672-1681.

[7] 丁欣, 施睿清, 薛睿南, 等. 弱电网下 LCL 型单相并网逆变器的鲁棒延时补偿策略 [J]. 太阳能学报, 2023, 44 (5): 162-170.

[8] 杨明, 杨杰, 赵铁英, 等. 基于数字控制延时的 LCL 型并网逆变器强鲁棒性加权平均电流控制策略 [J]. 电机与控制学报, 2023, 27 (2): 143-152.

[9] 杨明, 赵月圆, 杨杰, 等. 高渗透率下一种提高 LCL 滤波并网逆变器鲁棒性的新型锁相环设计方案 [J]. 中国电机工程学报, 2023, 43 (10): 3938-3949.

[10] 沈姝衡, 方天治, 章益凡. 高带宽数字控制 LCL 型并网逆变器及其提高并网系统鲁棒性的谐振抑制技术研究 [J]. 电工技术学报, 2022, 37 (21): 5548-5561.

(收稿日期: 2024-11-17)