

DOI:10.19651/j.cnki.emt.2211593

高分辨率的航拍输电线路绝缘子检测

柳方圆^{1,2} 任东^{1,2} 王露^{1,2} 杨军³ 郑朋³

(1. 三峡大学湖北省农田环境监测工程技术研究中心 宜昌 443000; 2. 三峡大学电气与新能源学院 宜昌 443000;
3. 湖北省宜昌市水产技术推广站 宜昌 443000)

摘要: 绝缘子是输电线路上的重要部件之一,利用无人机巡检准确的检测出绝缘子及其缺陷是保障电力安全输送的重要手段。针对目前主流目标检测网络处理高分辨率图像时直接缩放原图带来目标的细节信息丢失或者将原图切块再检测导致目标丢失整体信息的问题,在残差网络(ResNet 50)的基础上设计了一个双分支结构的主干网络(RC Net)同时兼顾绝缘子的位置信息和细节信息,能减少目标上下文信息和局部信息的丢失。同时引入可变形卷积替换部分常规卷积来改变采样点,使采样点能更贴合目标本身的几何形状,提高网络的特征表达能力,并根据绝缘子本身的大小和形状重新设计锚框的参数,使锚框更适合目标本身的尺度,边框回归更精确。在扩增的中国输电线路绝缘子数据集(CPLID)进行实验,结果表明,本文提出的算法的平均精度达到88.3%,相比于目前主流的检测算法在高分辨率图像背景下具有更好的检测效果。

关键词: 无人机巡检;目标检测;高分辨率图像;可变形卷积;锚框

中图分类号: TM75;TP391 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Transmission line insulator detection based on high resolution uav image

Liu Fangyuan^{1,2} Ren Dong^{1,2} Wang Lu^{1,2} Yang Jun³ Zheng Peng³

(1. Hubei Engineering Technology Research Center for Farmland Environmental Monitoring, China Three Gorges University, Yichang 443000, China; 2. Collage of Electrical & New Energy, China Three Gorges University, Yichang 443000, China;
3. Aquatic Technology Promotion Station of Yichang City, Hubei Province, Yichang 443000, China)

Abstract: Insulator is one of the important components on the transmission line. It is an important means to ensure the safe transmission of power to accurately detect the insulator and its defects using UAV patrol inspection. In order to solve the problem that the main target detection network directly scales the original image when processing high-resolution images, which leads to the loss of target details or re detects the original image by cutting it into blocks, which leads to the loss of the overall information of the target, a dual branch structure backbone network (RC Net) is designed based on the residual network (ResNet 50), which can reduce the loss of target context information and local information. At the same time, the deformable convolution is introduced to replace part of the conventional convolution to change the sampling points, so that the sampling points can more closely fit the geometric shape of the target itself, improve the feature expression ability of the network, and redesign the parameters of the anchor frame according to the size and shape of the insulator itself, so that the anchor frame is more suitable for the scale of the target itself, and the frame regression is more accurate. The experimental results on the expanded Chinese transmission line insulator dataset (CPLID) show that the average accuracy of the algorithm proposed in this paper reaches 88.3%, which is better than the current mainstream detection algorithm in the high-resolution image background.

Keywords: UAC inspection; object detection; high resolution image; deformable convolution; anchor

0 引言

电力是现代社会必不可缺的能源之一,保证电力的安全生产、输送是不可忽视的。绝缘子是输电线路上的重要组成

部分,检测并修复缺陷绝缘子对于保证输电线路的正常运行起着重要作用。由于长时间暴露在各种自然环境下,绝缘子不可避免的会有损坏,破裂、缺损、锈蚀是绝缘子常见的缺陷类型^[1]。因此,相关工作人员会定期对输电线路绝缘子进行

收稿日期:2022-10-04

巡检。随着无人机技术的日渐成熟,巡检过程从传统的人工巡检发展到无人机拍摄绝缘子影像、人工筛查缺陷绝缘子^[2]。近些年由于深度学习的发展,电力巡检的方式逐渐向无人机巡检拍摄图片,计算机识别缺陷绝缘子的方式转型,这样可以极大的节省人力并且提高巡检效率^[3-4]。

目前,一个至少具有 $2\,048\text{ pixel}\times 1\,080\text{ pixel}$ 的图像被视为2K高分辨率图像,至少具有 $3\,840\text{ pixel}\times 1\,080\text{ pixel}$ 的图像就达到了4K分辨率的标准,但是4K超高清清晰度通常指最小分辨率为 $3\,840\times 2\,160$ 的图像^[5]。而本文的输电线路绝缘子图像除了从中国输电线路绝缘子数据集(CPLID)获取的848张 $1\,152\text{ pixel}\times 864\text{ pixel}$ 的图像,还有自己用无人机拍摄的1478张 $5\,472\text{ pixel}\times 3\,648\text{ pixel}$ 的输电线路绝缘子图像,本文的原始图像构成存在大量的高分辨率图像。如果对原始图像不做处理直接放入网络训练,会直接占用计算机大量显存,训练精度也不理想。一般对于高分辨率图像检测的做法有两种:一是将图片调整至固定大小(resize)后,再输入到主干网络^[6],但是由于原图分辨率太高,直接resize会导致检测目标的部分细节信息丢失。另一种做法是先将原始图像切割成固定大小的块,再将切块后的图像输入网络进行训练^[7],这种做法能有效地避免丢失部分上下文信息,但是也不可避免会有其他的弊端:一张原图切图后会产生本身25倍数量的切割后的图(以 $5\,000\times 3\,000$ 切割成 $1\,000\times 600$ 举例),要对这些图片进行重新的筛选和人工标注,放在整个数据集中就是一个相当大的工作量。并且当检测的目标刚好在切图的边缘时,切割后的图片中的目标还会丢失上下文信息^[8]。

目前目标检测的算法可以分成一阶段检测网络和二阶段检测网络。一阶段网络直接将候选框进行分类、回归,完成对目标的检测,二阶段网络则是将候选区域和特征提取这两步分开进行。以SSD^[9]和YOLO系列^[10-11]为代表的一阶段检测网络比二阶段网络具有更快的检测速度,自2014年R-CNN^[12]的提出,二阶段检测算法相继出现了SPP-Net^[13]、Fast R-CNN^[14]、Faster R-CNN^[15]等经典算法,二阶段算法在检测速度上不如一阶段算法快,检测精度上一般比一阶段算法略高。R-CNN提出通过选择性搜索(selective search)获得候选区域,SPP-Net算法主要提出了一个空间金字塔池化层,解决了输入图像尺寸受限问题。继SPP-Net后,Fast R-CNN提出在分类和回归中将坐标引入,进一步提高了检测的速度。2016年,Faster R-CNN被提出,其相比于Fast R-CNN的核心创新在于提出了RPN网络,RPN网络替代selective search算法来产生候选区域,将特征提取、区域提议、分类、回归等过程整合到一起,进一步提高了检测的速度和精度。近些年大量学者在Faster R-CNN结构上做改进,目前主流的Faster R-CNN的特征提取网络已经由VGG16替换为ResNet^[16]和FPN^[17]的结构。Faster R-CNN具有较高的检测精度,作为目前二阶段的主流检测算法,已被广泛应用于电力绝缘

子的检测中。

本文提出一种同时兼顾目标细节信息和位置信息的网络—RC Net,能有效的解决网络直接将图片缩放带来的细节信息丢失和直接切图带来的位置信息丢失的问题。RC Net基于Faster R-CNN的框架,设计了一个以ResNet为主干,FPN为特征融合层的双分支结构,一个分支负责提取目标的全局语义信息,一个分支负责提取目标的局部位置和上下文信息,再将两个分支提取的信息进行融合,实现对高分辨率图像的特征提取。此外,绝缘子背景复杂问题也存在于航拍影像中,为解决此问题本文在RC Net的R分支的ResNet 50的 $C_3\sim C_5$ 层采用可变形卷积(deformable convolution network,DCN)^[18],通过改变卷积采样点的方式,使网络的深层在提取绝缘子特征时能让采样点更精确的落在绝缘子上,抑制背景信息的干扰。另外根据绝缘子本身的大小重新设计锚框,使先验框更适合绝缘子的大小,在网络最后的分类和边框回归时能更精确,最终实现对无人机航拍图像下的绝缘子检测精度的提升。

1 本文方法

1.1 主干网络的改进—RC Net

RC Net是基于Faster R-CNN的网络架构的改进,在骨干网络层采用双分支ResNet 50结构代替原来的单分支ResNet 50结构。一个分支负责提取目标的整体信息,一个分支负责提取目标的上下文位置信息,再将两个分支提取到的语义信息深度融合,将融合的特征逐层送入FPN层进行特征融合,后续经过RPN层用anchor对特征图上的目标进行初步的预测,经过ROI Pooling层后将特征图输出为固定的 7×7 大小,最后分别对目标做分类和边框回归。图1是RC Net的整体流程图,图2、3分别是RC Net两个分支具体的流程图。

图1中,输入一幅分辨率为 $5\,472\times 3\,648$ 大小的图像,R分支对图像缩放至 $1\,200\times 800$ 大小,保留目标的上下文信息,但同时会丢失目标部分像素。C分支将数据集中每一个分辨率为 $5\,472\times 3\,648$ 的图像都切成 $1\,200\times 800$ 大小的块,这样做可以更好保留目标的细节信息,但是不可避免会丢失上下文信息。最后R分支的FPN和C分支的FPN通过一个 3×3 的卷积来实现聚合,在不改变特征图大小和通道数的情况下,对R分支和C分支的FPN的特征图逐层相加(add),分层输出特征图。

图2和3描述了R分支与C分支的相互融合过程。图像经过R分支resize到 $1\,200\times 800$,经过ResNet 50后, C_1 至 C_5 层分别得到 600×400 、 300×200 、 150×100 、 75×50 、 37×25 大小的特征图。同时,图像经过C分支crop到 $1\,200\times 800$ 大小的块,每一块图像也经过ResNet 50,输出的 $C_1\sim C_5$ 层分别得到2、4、8、16、32倍下采样的图。由于只采用ResNet 50的 $C_2\sim C_5$ 层作为后续FPN层的输入,所以只在 $C_2\sim C_5$ 层上做两个分支的特征融合,每一层特征

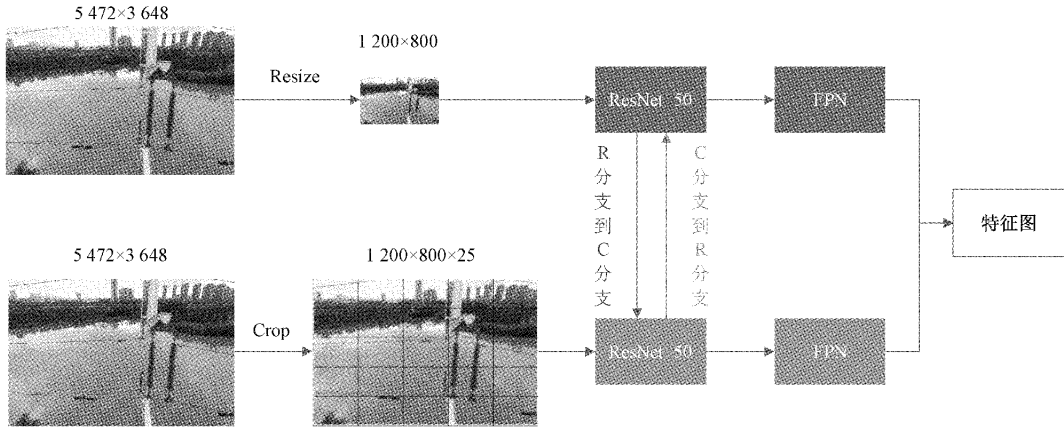


图 1 RC Net 整体流程图

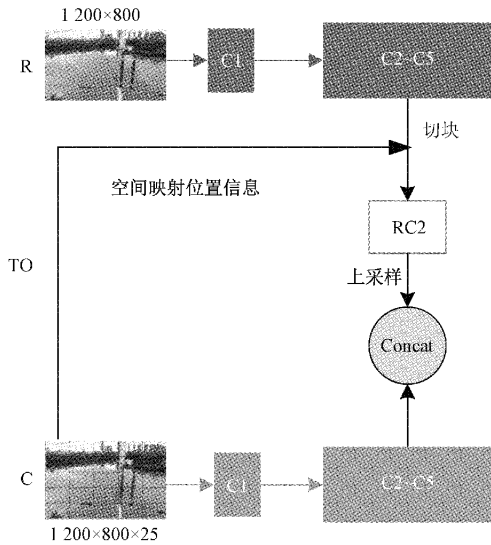


图 2 R 分支到 C 分支的过程图

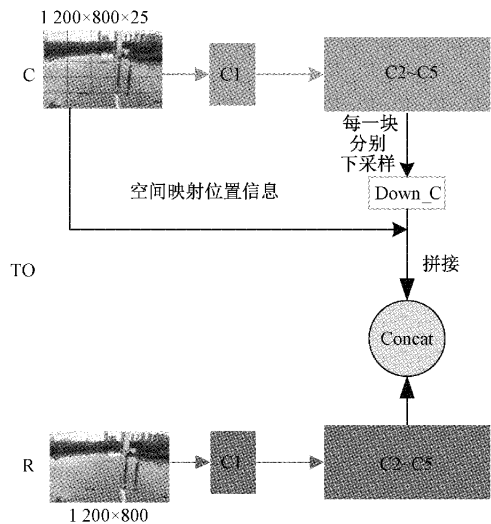


图 3 C 分支到 R 分支的过程图

融合的过程都一样。以 C_2 层为例,在 R 分支将图像下采

样到 300×200 ,C 分支将图像切割成 300×200 大小的块。在 R 分支向 C 分支融合的过程中,先通过空间映射的关系,找到 C 分支上每一个切块后的图像对应 R 分支上图像的具体位置,然后将 R 分支上与之对应的图像一一裁剪出来,分别上采样到和 C 分支特征图相同的大小,也就是 300×200 ,最后将上采样后的图和 C 分支上的对应的切块图做通道相加(concat)操作。在 C 分支向 R 分支融合的过程中,先对 C 分支每一块的切块图像下采样,通过空间映射关系,每一块分别匹配到 R 分支图像上对应的位置,然后将这些块按裁剪的顺序重新拼接(combine)回去,得到和 R 分支大小相同的特征图,也就是 300×200 ,再与 R 分支上的特征图做 concat 操作。图 4 是 RC Net 的具体结构图。

1.2 基于 DCN 的特征提取网络改进

DCN 是 2017 年由 Dai 等人提出的,DCN 在传统卷积核的基础上,添加了一个偏移量 offset,由此改变传统规则卷积核的特征提取点,使得可变形卷积能够根据目标本身的几何形状进行特征提取。

传统的卷积过程可由式(1)表示:

$$y(a_0) = \sum_{a_n \in R} w(a_n) \cdot x(a_0 + a_n) \quad (1)$$

式中: y 表示卷积输出, R 表示卷积核的大小,并且定义了卷积核上每个点坐标,例如一个 3×3 的卷积核 R 的定义就是 $R = \{(-1, -1), (-1, 0), (-1, 1), (0, -1), (0, 0), (0, 1), (1, -1), (1, 0), (1, 1)\}$, a_n 表示以 a_0 为中心的附近的像素点, $x(a_0 + a_n)$ 表示使用规则网格 R 采样的特征, w 是权重系数,对 x 进行加权求和。

DCN 在传统卷积上引入了一个偏移量 offset,使得网络在进行卷积操作时可以根据物体本身的形状去定义输出的特征图,能让特征图更贴合物体本身的几何形状。式(2)表示可变形卷积的过程,式中 Δa_n 表示的就是偏移量 offset。

$$y(a_0) = \sum_{a_n \in R} w(a_n) \cdot x(a_0 + a_n + \Delta a_n) \quad (2)$$

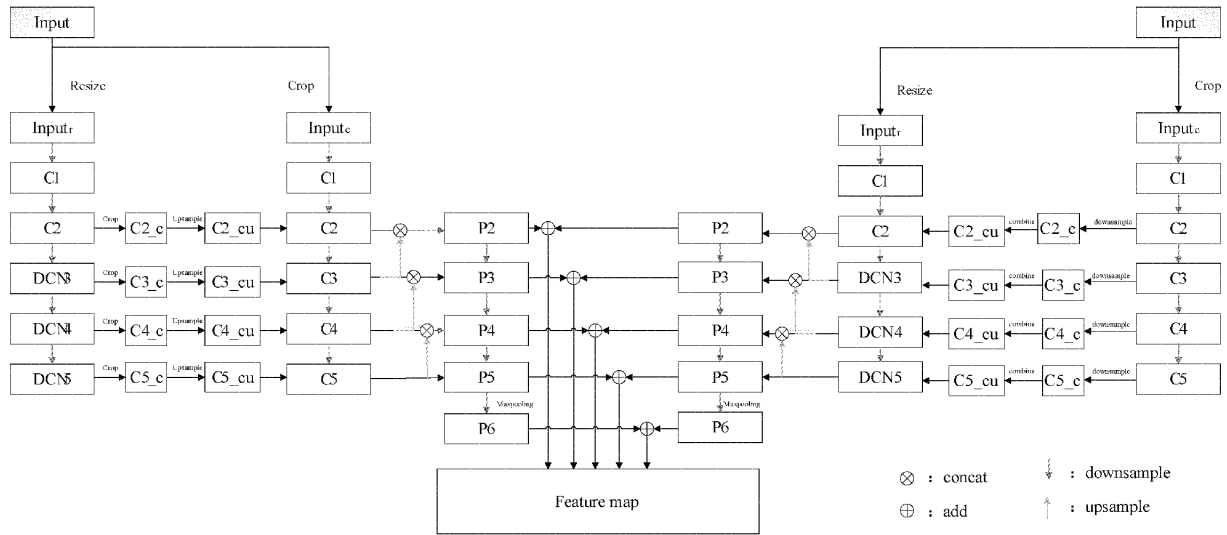


图 4 RC Net 整体结构

一开始,和规则的卷积核一样提取目标的特征图,然后在这个特征图上再施加一个卷积层,以此来得到偏移量 offset,这个 offset 就用来指导规则卷积核卷积出来的特征上的 x 和 y 位置偏移。DCN 和规则的卷积在网络的同一层具有相同的超参数,DCN 在输出规则特征图的同时也输出偏移量,对每个采样点同时输出 x 、 y 两个方向的偏移量。

在前人研究的基础上,在 ResNet 50 的 $C_3 \sim C_5$ 层都用上了 DCN,图 5 是 ResNet 中任意一个 3×3 卷积替换成 DCN 的结构图,图中左边分支加上右边的残差结构就是原本的 ResNet,DCN 是在其基础上加上右边一条分支,引入 DCN 用来学习目标相对于原本 3×3 卷积核的偏移量,然后通过 concat 和 add 的操作将常规卷积学习到的特征和可变形卷积学习到的偏移量融合,得到最后的特征输出。

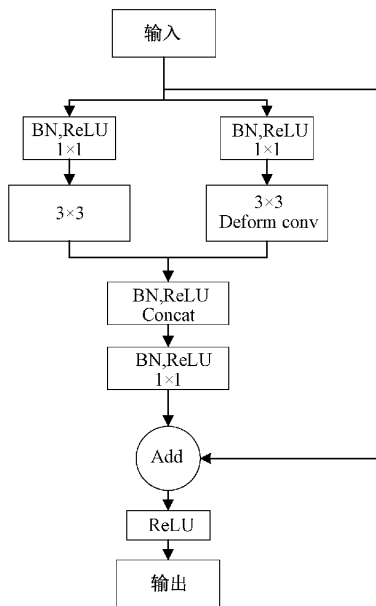


图 5 ResNet+DCN 结构

1.3 锚框的优化改进

Faster R-CNN 原本设置的锚框(anchor)的基础边长值是 128、256、512,有 1:1,1:2,2:1 这三种宽高比,因此一共会产生 9 种不同尺度大小的 anchor。式(3)~(5)分别给出了 anchor 的宽高和面积的定义。式中, w 代表 anchor 的宽, h 代表 anchor 的高, S 代表 anchor 的面积, R_a 代表 anchor 的宽高比, B_s 代表基础边长值。

$$w = \sqrt{S \times R_a} \tag{3}$$

$$h = \sqrt{S \div R_a} \tag{4}$$

$$S = B_s^2 \tag{5}$$

在原来的 3 种基础边长下增加了 2 个尺度的基础 anchor 值,并且改变了 anchor 的宽高比。重新设置锚框的参数如下:设定 32、64、128、256、512 为 anchor 的基础边长值,设置三种宽高比例(1:1,1:2,4:1)。因为每种基础尺度的 anchor 都会产生 3 种变形尺度的 anchor,所以一共会产生 15 个尺度的 anchor 去匹配原图中的标注框。本文统计了原图标注框的面积大小和长宽比,图 6 是标注框在原图上的面积大小分布图,横坐标表示像素点个数,纵坐标表示标注框个数。图 7 是标注框在原图上的长宽比分布图,横坐标表示长宽比,纵坐标表示标注框的个数。图 6 和 7 中 max、min、mean、var 分别代表面积和长宽比的最大值、最小值、平均值、方差。

由于 anchor 是在原始图像上与标注框进行匹配,而原始图像在经过网络前会 resize,根据原图分辨率是 5472×3648 大小的特点,设置原图 resize 后的尺寸为 1200×800 ,这样在宽高两个方向可以刚好保证缩小相同的倍数(4.56 倍),可以避免更多的像素丢失。所以在匹配 anchor 和原图的标注框时,需要对 anchor 框的面积再乘以大约 20.8 倍(4.56^2),才是 anchor 在原图上的面积大小。改进前的最小的 anchor 匹配的标注框面积是 $128 \times 128 \times$

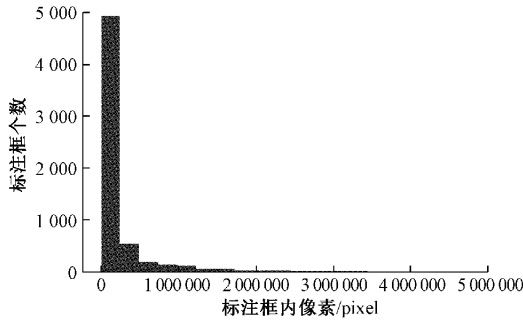


图 6 标注框的面积分布图

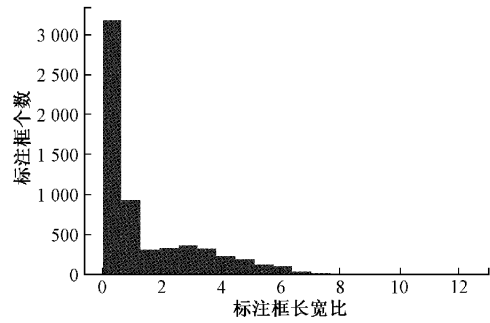


图 7 标注框的长宽比分布图

20.8=340 787.2,最大的 anchor 匹配的标注框面积是 $512 \times 512 \times 20.8 = 5\,452\,595.2$,根据图 6 可以看到大量的标注框面积分布在 0~1 000 000 pixel 之间,所以增加了 32 和 64 这两个 anchor_base 值来关注原图中较小的标注框。改进后的 anchor 对应到原图标注框最小的面积是 $32 \times 32 \times 20.8 = 21\,299.2$,更适合本文的标注框。

原图标注框的宽高比最大是 12.79,最小是 0.07,大多数分布在 0.1~4 之间,根据图 7 的统计重新设计 anchor 的长宽比为(1:1,1:2,4:1),这样的长宽比的 anchor 更适合框出绝缘子的特征。表 1 展示了改进后的 anchor 宽高和在不同大小特征图下的原图标注框的个数分布情况。

表 1 改进后特征图各尺寸下 anchor 的长宽参数以及对应的原图标注框的统计

FPN 层数	P2	P3	P4	P5	P6
特征图大小	300×200	150×100	75×50	37×25	18×12
	32×32	64×64	128×128	256×256	512×512
anchor 长宽	23×46	45×91	90×181	181×362	362×724
	64×16	128×32	256×64	512×128	1 024×256
标注框个数	1 919	1 611	1 695	703	202
总计	6 130				

2 实验结果及分析

本实验采用一块 NVIDIA GeForce RTX 3090 的显卡进行实验,显存为 24 GB,基于 linux 操作系统和 Pytorch 框架搭建的深度学习平台,Pytorch 版本为 1.8.1,Python 版本为 3.6.13,CUDA 版本为 11.1。

2.1 数据集及评价指标

本文数据集来源于中国输电线路绝缘子数据集(CPLID)的 848 张分辨率为 $1\,152 \times 864$ 图片和国家电网某公司的分辨率为 $5\,472 \times 3648$ 的输电线路巡检图片 1 082 张,一共 1 930 张图片。使用 labelimg 工具对图像进行标注,标注出绝缘子(insulator)、缺陷绝缘子(defect)、鸟窝(nest)三类,最后共有 6 130 个标注框,其中绝缘子有 5 115 个标注框,缺陷绝缘子有 615 个标注框,鸟窝有 400 个标注框。将图片按 8:1:1 的比例随机分为训练集、验证集、测试集。本文主要采用每一类别的准确率(AP)、召回率(Recall)、所有类别的平均准确率(mAP)这三项作为评价指标。

2.2 消融实验

1)RC Net 特征融合

设计双分支结构时,为了尽量使图片保留原本的信

息,在 R 分支让图片的长宽等比缩放,在 C 分支保证相邻的两个切块之间至少保留 50 个像素点的重叠。在选定切块尺寸大小时,考虑应该尽量使切块个数少,这样能减少边缘信息的丢失次数。同时在下采样时,下采样倍数越大,丢失的细节信息就越多,但是切块尺寸越大,对每一块特征提取的计算量也会增加,综合考虑上面几点因素的影响,最终选择切块尺寸和下采样的尺寸为 $1\,200 \times 800$ 。此尺寸下 ResNet 50 的 $C_1 \sim C_5$ 层输出特征图大小依次为 $600 \times 400, 300 \times 200, 150 \times 100, 75 \times 50, 37 \times 25$ 。表 2 给出了各切块尺寸下切块的个数和下采样倍数的统计。

表 2 切块尺寸、个数和下采样倍数的统计

C 分支 切块尺寸	C 分支 切块个数	R 分支下 采样倍数
$1\,368 \times 912$	25	4
$1\,200 \times 800$	25	4.56
$1\,000 \times 666$	36	5.472
800×533	49	6.84

为了验证 RC Net 特征融合的有效性,设计了 3 组实验。第 1 组是只对图像进行下采样处理,第 2 组是只对图像进行切块处理,第 3 组是双分支结构,同时对图像下采

样和切块处理,并将两个分支的特征图做特征融合,实验结果如表 3 所示。表 3 中下采样就是 Faster R-CNN+ResNet 的结构做的实验,而切块方法的检测精度全面降

低,主要因素还是因为切块导致目标上下文信息丢失,而 RC Net 的结果精度相较下采样和切块都有提升,证明本研究对主干网络的改进是有效的。

表 3 双分支融合的检测结果

检测方法	准确率			召回率			平均准确率
	绝缘子	缺陷绝缘子	鸟窝	绝缘子	缺陷绝缘子	鸟窝	
下采样	75.1	86.0	80.1	84.8	94.3	84.5	80.4
切块	72.5	78.4	77.4	81.2	89.4	79.3	76.1
RC Net	78.8	88.8	86.1	86.5	94.7	91.5	84.6

2)DCN 模块的添加

为了验证 DCN 模块的添加对网络的改进是有效的,尝试在 ResNet 的不同层将原来的 3×3 卷积替换成 DCN,此消融实验是基于原本的 Faster R-CNN +

ResNet50+FPN 上进行的,实验结果如表 4 所示。结果表明,在不同层添加 DCN 后,模型的检测精度都有不同程度的提高,在 C3~C5 层同时将常规卷积替换为 DCN 时效果最好。

表 4 不同层添加 DCN 的检测结果

主干网络模型	准确率			召回率			平均准确率
	绝缘子	缺陷绝缘子	鸟窝	绝缘子	缺陷绝缘子	鸟窝	
ResNet 50	75.1	86	80.1	84.8	94.3	84.5	80.4
C3 层添加 DCN	74.7	84.7	82.9	83.5	94.7	90.1	80.8
C4 层添加 DCN	75.6	86.6	82.0	85.1	94.7	90.1	81.4
C5 层添加 DCN	75.6	88.4	82.3	81.6	93.8	93.0	82.1
C3 层~C5 层添加 DCN	78.8	88.4	82.6	86.1	95.6	93.0	83.3

3)anchor 的优化改进

Faster R-CNN 原本的锚框参数设置如下:anchor_base 是 128,256,512,有 $1:1, 1:2, 2:1$ 三种宽高比,共 9 种大小的 anchor。改进后的 anchor 参数如下:anchor_

base 增加了两个尺度 32 和 64,宽高比根据绝缘子的特征改成了 $1:1, 1:2, 4:1$ 。此消融实验还是基于 Faster R-CNN+ResNet 50+FPN 上进行的,实验结果如表 5 所示。

表 5 anchor 改进结果

anchor 参数	准确率			召回率			平均准确率
	绝缘子	缺陷绝缘子	鸟窝	绝缘子	缺陷绝缘子	鸟窝	
原始的	75.1	86.0	80.1	84.8	94.3	84.5	80.4
改进后的	79.2	87.9	81.3	85.7	97.3	91.5	82.8

2.3 对比实验

最终将主干网络替换为 RC Net,并将 R 分支上的 ResNet 50 主干网络的 C3~C5 层将常规卷积替换为 DCN,将 anchor 改用改进后的参数。为了验证本文算法在高分辨率航拍输电线路绝缘子检测上的优越性,将本文算法与目前一些先进的目标检测算法做对比,包括 Faster R-CNN、Cascade R-CNN^[19]、Atss^[20]、HRNet^[21]。表 6 统计了各模型在测试集上的精度、召回率和平均精度。

由表 6 可知,本文的算法相比于 Faster R-CNN 在绝缘子、缺陷绝缘子、鸟窝这三类目标检测的准确率上分别

提升了 7.6%,9.5%,9.5%,分别达到了 82.7%,95.5%,89.6%;召回率分别提升了 1.8%,1.3%,5.7%,分别达到了 86.6%,95.6%,90.2%;mAP 提升了 8.9%,达到了 89.3%,证明改进是有效的。Cascade R-CNN 也是基于 Faster R-CNN 的一种改进算法,它利用不同的 IOU 阈值选取,将多级目标检测器进行级联,进一步的优化检测结果;ATSS 提出能够自动根据目标物体的统计特征选择正负训练样本,在正负样本的策略上做了改进;HRNet 相比于 ResNet 的改进是将不同分辨率之间的串联结构改成了并联结构,通过连接高到低分辨率的表征,并且在并行卷

表 6 本文模型与其他模型对比

检测方法	准确率			召回率			平均准确率
	绝缘子	缺陷绝缘子	鸟窝	绝缘子	缺陷绝缘子	鸟窝	
Faster R-CNN	75.1	86	80.1	84.8	94.3	84.5	80.4
Cascade R-CNN	81.6	88.7	83.5	85.9	91.6	87.5	84.6
ATSS	79.4	86.2	87.1	86.1	94.5	88.5	84.2
HRNet	79.1	89.5	89.2	85.7	92.7	87.3	85.9
本文	82.7	92.5	89.6	86.6	95.6	90.2	88.3

积之间反复的进行多尺度融合,来维持高分辨率的特征。本文算法取得最好的结果得益于设计的 RC 双分支融合结构能同时注意到目标的细节信息和上下文信息,在主干网络的 $C_2 \sim C_5$ 使用 DCN 替换常规卷积则可以更好的根据目标本身的几何形状去提取特征,然后通过目标本身的大小和宽高比去设计适合目标的 anchor 参数以达到更好的分类和回归效果。

2.4 结果分析

为了更直观的展示本文检测方法的结果,在测试集上将结果可视化,因为原图太大,所以只截取含绝缘子部分的区域,如图 8 所示,左列是 Faster R-CNN 的检测结果,右列是本文算法的检测结果,展示了复杂背景、遮挡等环境因素下的测试结果对比图,图 8(a)~(f)是复杂背景下的测试实验结果,图 8(g)、(h)是遮挡环境下的测试实验结果。从测试结果图可以看出,Faster R-CNN 算法对高分辨率图像在复杂背景以及目标存在遮挡的情况时检测效果受限,存在漏检、错检的情况,而本文改进的算法能有效提高对于高分辨率图像在复杂背景下以及遮挡的情况下的准确率,验证了本文算法在检测高分辨率航拍绝缘子图像的有效性。

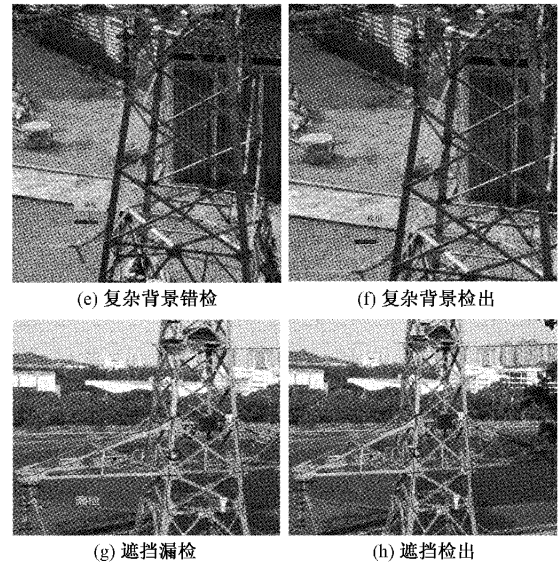


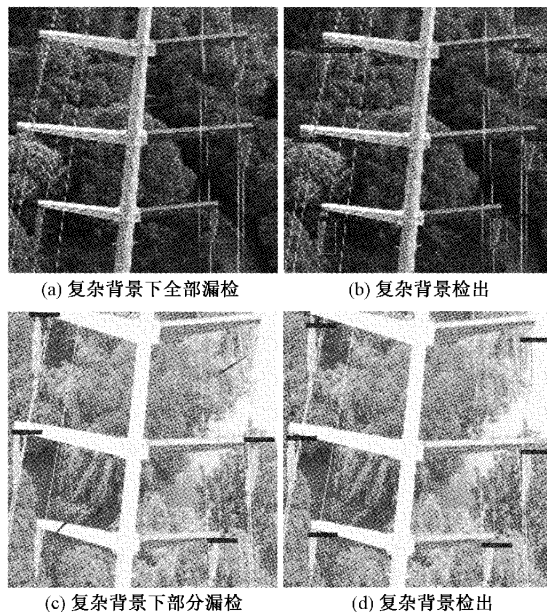
图 8 测试集上的对比结果

3 结 论

目前对高分辨率航拍图像的预处理无论是下采样还是切图都会带来细节或者上下文信息的丢失,一些先进的检测算法一般在输入图像经过下采样的处理后也难以取得较好的效果。本文提出的 RC Net 可以同时提取目标的细节信息和上下文信息并有效地融合,同时根据目标本身设计合适的锚框参数并且引入可变形卷积,通过实验结果验证取得了比较好的结果,相比于原本的 Faster R-CNN,本文算法 mAP 提升了 7.9%,达到了 88.3%,能够满足对高分辨率航拍绝缘子及其缺陷检测任务的要求。

参考文献

- [1] 李鑫,刘帅男,杨楨,等. 基于改进 Cascade R-CNN 的输电线路多目标检测[J]. 电子测量与仪器学报, 2021,35(10):24-32.
- [2] 张培培,王昭,王菲. 基于深度学习的图像目标检测算法研究[J]. 国外电子测量技术, 2020, 39(8):21-39.
- [3] 高伟,周宸,郭谋发. 基于改进 YOLOv4 及 SR-GAN 的绝缘子缺陷辨识研究[J]. 电机与控制学报, 2021, 25(11):93-104.
- [4] 侯涛,蒋瑜. 改进 YOLOv4 在遥感飞机目标检测中的



- 应用研究[J]. 计算机工程与应用, 2021, 57(12): 224-230.
- [5] TANG Y, GAO R, LEE H H, et al. High-resolution 3D abdominal segmentation with random patch network fusion[J]. Medical Image Analysis, 2020, 69(3): 101894, DOI:10.1016/j.media.2020.101894.
- [6] 陈庆, 闫斌, 叶润, 等. 航拍绝缘子卷积神经网络检测及自爆识别研究[J]. 电子测量与仪器学报, 2017, 31(6): 942-953.
- [7] VAN ET TEN A. You only look twice: Rapid multi-scale object detection in satellite imagery [J]. Arxiv Preprint, 2018, ArXiv:1805.09512.
- [8] 汤踊, 韩军, 魏文力, 等. 深度学习在输电线路中部件识别与缺陷检测的研究[J]. 电子测量技术, 2018, 41(6): 60-65.
- [9] WEI L, DRAGOMIR A, DUMITRU E, et al. SSD: Single shot MultiBox detector [C]. European Conference on Computer Vision, Springer, Cham, 2016.
- [10] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. Arxiv Preprint, 2018, ArXiv:1804.02767.
- [11] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[J]. Arxiv Preprint, 2020, ArXiv:2004.10934.
- [12] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [13] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [14] GIRSHICK R. Fast r-cnn [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [15] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]. NIPS, 2016.
- [16] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [17] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [18] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 764-773.
- [19] CAI Z, VASCONCELOS N. Cascade r-cnn: Delving into high quality object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6154-6162.
- [20] ZHANG S, CHI C, YAO Y, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 9759-9768.
- [21] SUN K, XIAO B, LIU D, et al. Deep high-resolution representation learning for human pose estimation [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 5693-5703.

作者简介

柳方圆, 硕士研究生, 主要研究方向为深度学习、图像处理。

E-mail: 2429777956@qq.com

任东, 博士, 教授, 博士生导师, 主要研究方向为计算机视觉、图像处理、模式识别、3S技术、物联网。

E-mail: rendong5227@163.com

王露, 博士研究生, 主要研究方向为深度学习、图像处理。

E-mail: 276379755@qq.com

杨军, 高级工程师, 主要研究方向为深度学习、图像处理。

E-mail: 1010893056@qq.com

郑朋, 工程师, 主要研究方向为深度学习、图像处理。

E-mail: 1020677019@qq.com