

基于改进 CenterNet 的航拍绝缘子缺陷实时检测模型

李发光, 伊力哈木·亚尔买买提

引用本文

李发光, 伊力哈木·亚尔买买提. [基于改进 CenterNet 的航拍绝缘子缺陷实时检测模型](#)[J]. 计算机科学, 2022, 49(5): 84-91.

LI Fa-guang, YILIHAMU·Yaermainaiti. [Real-time Detection Model of Insulator Defect Based on Improved CenterNet](#)[J]. Computer Science, 2022, 49(5): 84-91.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[结合注意力机制的多尺度特征融合图像去雾算法](#)

Multi-scale Feature Fusion Image Dehazing Algorithm Combined with Attention Mechanism

计算机科学, 2022, 49(5): 50-57. <https://doi.org/10.11896/jsjx.210400093>

[结合注意力机制与几何信息的特征融合框架](#)

Feature Fusion Framework Combining Attention Mechanism and Geometric Information

计算机科学, 2022, 49(5): 129-134. <https://doi.org/10.11896/jsjx.210300180>

[面向对话的融入知识的实体关系抽取](#)

Dialogue-based Entity Relation Extraction with Knowledge

计算机科学, 2022, 49(5): 200-205. <https://doi.org/10.11896/jsjx.210300198>

[基于多特征融合的重叠组套索脑功能超网络构建及分类](#)

Construction and Classification of Brain Function Hypernetwork Based on Overlapping Group Lasso with Multi-feature Fusion

计算机科学, 2022, 49(5): 206-211. <https://doi.org/10.11896/jsjx.210300049>

[基于用户关联的立场检测](#)

Stance Detection Based on User Connection

计算机科学, 2022, 49(5): 221-226. <https://doi.org/10.11896/jsjx.210400135>

基于改进 CenterNet 的航拍绝缘子缺陷实时检测模型

李发光 伊力哈木·亚尔买买提

新疆大学电气工程学院 乌鲁木齐 830047

(2434740150@qq.com)

摘要 针对无人机在电力巡检过程中对绝缘子及其缺陷检测的准确率较低、实时性较差的问题,提出一种改进 CenterNet 的绝缘子缺陷检测模型。首先,使用轻量级网络 EfficientNet-B0 代替原始模型的特征提取网络 ResNet18,在保证模型提取能力的同时加快了其推理速度;其次,搭建特征加强模块 (Feature Enhancement Module, FEM),并对经过上采样后的特征通道权重进行合理分配,抑制无效特征,并借鉴 FPN (Feature Pyramid Networks) 融合浅层与深层特征,使特征层信息更加丰富;然后在 CenterNet-Head 中引入空间和通道混合的注意力机制 CA (Coordinate Attention),使类别和位置信息的预测更加准确;最后,使用 Soft-NMS 解决在模型检测过程中由中心点预测不准导致的“单目标多框”问题。实验结果表明,改进的 CenterNet 比原始模型的精度提高了 11.92%,速度提高了 8.95 FPS,模型大小减小了 54 MB。与其他检测模型相比,改进模型的精度与速度均有提高,证明了其实时性和鲁棒性。

关键词: 绝缘子;缺陷检测;特征融合;注意力机制;CenterNet

中图分类号 TM933;TN919.8;TM769

Real-time Detection Model of Insulator Defect Based on Improved CenterNet

LI Fa-guang and YILIHAMU · Yaermaimaiti

College of Electrical Engineering, Xinjiang University, Urumqi 830047, China

Abstract Aiming at the problem that it is difficult to detect insulators and their defects in real time and efficiently in the course of electric patrol inspection of UAV, an improved insulator defect detection model based on CenterNet is proposed. Firstly, light-weight network EfficientNet-B0 is used to replace the original model's feature extraction network ResNet18, which ensures the model extraction ability and speeds up its reasoning speed. Then, a feature enhancement module FEM is built, which distributes the weight of the feature channels after upsampling reasonably and suppresses invalid features. Using FPN (feature pyramid networks) for reference, the shallow and deep features are integrated to enrich the information of feature layer. Secondly, the coordination attention (CA) mechanism, which is a mixture of space and channel, is introduced into CenterNet-Head, which makes the prediction of category and location information more accurate. Finally, Soft-NMS is used to solve the problem of "single target and multiple frames" caused by inaccurate prediction of center points in the process of model detection. Experimental results show that the precision of the improved CenterNet is improved by 11.92%, the speed is increased by 8.95FPS, and the model size is reduced by 54 MB. Compared with other detection models, the accuracy and speed are improved, which proves the real-time performance and robustness of the improved model.

Keywords Insulator, Defect detection, Feature fusion, Attention mechanism, CenterNet

1 引言

随着人们生活水平的不断提高,社会对电力的依赖程度大大提升,保证输电线路网安全运行显得尤为重要。绝缘子作为输电线路上的关键部件,其运行状况直接影响到输电线路网的安全性及可靠性。由绝缘子故障引发电力系统瘫痪的现象在电力系统故障中的占比最高,若不对其进行及时排查并采取补救措施,则会酿成严重后果^[1]。

由于电网规模不断扩大,利用传统人工巡检的方式难以按时完成电力巡检任务,因此电力巡检任务逐步向“机巡为主,人巡为辅”的方向发展^[2]。使用无人机对绝缘子故障进行自动检测与分析成为了当今热门的研究方向^[3]。对于绝缘子识别及其缺陷的检测,主要采取基于图像处理的方法和基于深度学习的目标检测方法。随着计算机硬件性能的提升,后者取得了较好的发展,前者因其效率低下逐步被后者替代。基于深度学习的目标检测方法主要有双阶段 R-CNN^[4-7] 系列

到稿日期:2021-04-15 返修日期:2021-09-05

基金项目:国家自然科学基金(61866037,61462082)

This work was supported by the National Natural Science Foundation of China(61866037,61462082).

通信作者:伊力哈木·亚尔买买提(65891080@qq.com)

检测算法和单阶段 SSD^[8-10], YOLO^[11-13] 和 CornerNet^[14-16] 系列算法。使用基于深度学习的目标检测方法对绝缘子定位及其缺陷进行检测,国内外学者在这方面进行了大量的研究。Zhao 等利用 CenterNet 对输电线路中绝缘子自爆、防震锤脱落和鸟巢 3 类常见的故障进行检测,其虽然可以进行实时高效的检测,但是在特征提取阶段使用了 DLANet,使模型所占内存较大且不易部署到无人机设备中^[17]。Yi 等利用 Faster RCNN 对输电线路中的航拍绝缘子进行检测,但并未对绝缘子缺陷进行检测^[18]。基于锚框(Anchor-Boxes)的 RCNN, SSD 和 YOLO 系列的检测模型会在一张图片上产生大量的锚框,而一张图片上存在的目标却较少,这会造成正负样本数量不平衡。此外,引入锚框的同时也会产生更多的参数,从而降低模型的训练效率。因此,本文将无锚框(Anchor-Free)检测模型 CenterNet^[19] 作为基准模型,但其会舍弃目前双阶段及单阶段检测模型所用到的 Anchor-Boxes。

综上所述,本文对 CenterNet 进行针对性的改进,使其对绝缘子缺陷检测更有鲁棒性。首先将特征提取网络替换为 EfficientNet-B0^[20], 并构建 FEM 特征加强模块;其次借鉴 FPN^[21] 融合深层特征和浅层特征,使最后得到的特征图信息更加丰富;然后在检测模块(CenterNet-Head)中引入空间和通道混合的 CA^[22] 注意力机制;最后,为解决“单目标多框”问题,在模型后处理阶段引入 Soft-NMS^[23] 对多余的预测框进行筛选。实验结果证明,本文模型对绝缘子及其缺陷检测有较好的鲁棒性。

2 CenterNet 目标检测模型

2.1 CenterNet 网络结构

CenterNet 是一种无锚框(Anchor-Free)的单阶段检测模型,其网络结构简单且检测性能优异。与基于锚框的目标检测模型相比,其利用目标的中心点来代表锚框,将特征提取网络得到的热力图的前 100 个峰值作为待检测目标中心点,通过设置阈值来筛选得到最终的目标中心点,然后根据该中心点的图像特征回归得到目标所属的类别及位置信息。整个过程不基于 Anchor 机制,因此不需要提前设置 Anchor 的超参数,同时舍弃 NMS(Non Maximum Suppression)^[24] 后处理操作,显著地缩减了网络的计算量和训练时间。原模型分别使用了 ResNet18^[25], DLA-34^[26] 和 Hourglass-104^[27] 这 3 种卷积神经网络来对特征进行提取,将其提取到的特征图传递给检测模块,由 3 个卷积块分别对目标进行中心点与类别预测、目标宽高预测和中心点偏置预测。CenterNet 模型简图如图 1 所示。

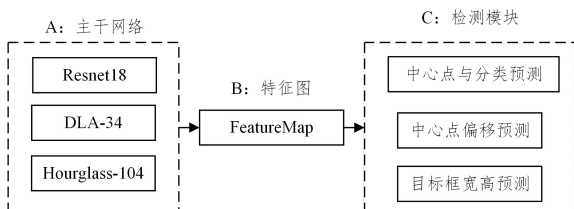


图 1 CenterNet 网络模型简图

Fig. 1 CenterNet network model diagram

2.2 损失函数

CenterNet 检测模型通过 3 个独立的卷积块分别对中心点与分类、中心点偏执(Offset)和目标框尺寸(Size)进行预测,因此其损失函数由 3 部分构成:中心点与分类的损失函数、目标中心偏置的损失函数和目标框尺寸的损失函数。其中心点与分类的损失函数 L_k 由根据 CenterNet 网络结构修改的 FocalLoss^[28] 函数构成,其计算式如式(1)所示:

$$L_k = -\frac{1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^a \lg(\hat{Y}_{xyc}), & Y_{xyc} = 1 \\ (1 - \hat{Y}_{xyc})^b (\hat{Y}_{xyc})^a \times \lg(1 - \hat{Y}_{xyc}), & \\ \text{others} & \end{cases} \quad (1)$$

其中,中心点与分类损失函数 L_k 的下标 k 表示第 k 张输入图像, N 为图片中关键点数量,下标 xyc 为图像的正负样本, Y_{xyc} 是真实值的标签, \hat{Y}_{xyc} 为预测值的标签, a 和 b 为超参数,分别设置为 2 和 4。

中心点偏置损失函数 L_{offset} 采用 L1 loss 函数,如式(2)所示:

$$L_{\text{offset}} = \frac{1}{N} \sum_p \left| \hat{O}_p - \left(\frac{p}{R} - \tilde{P} \right) \right| \quad (2)$$

其中, p 为原图中目标真实值的坐标; $\frac{p}{R} - \tilde{P}$ 为实际中心点偏移量; R 为下采样倍数,其值为 4; \hat{O}_p 为预测的偏移量; \tilde{P} 为模型预测的中心点坐标。

目标框尺寸损失函数 L_{size} 也采用 L1 loss 函数,如式(3)所示:

$$L_{\text{size}} = \frac{1}{N} \sum_{k=1}^N |S_{pk} - \hat{S}_k| \quad (3)$$

其中, S_k 为原目标框尺寸, \hat{S}_k 为回归后的目标框尺寸。总损失函数是将 3 个损失函数分别乘以相应的系数再相加,如式(4)所示:

$$L_{\text{sum}} = L_k + \lambda_{\text{size}} L_{\text{size}} + \lambda_{\text{off}} L_{\text{off}} \quad (4)$$

其中, λ_{size} 为 0.1, λ_{off} 为 1。

3 改进的 CenterNet 模型设计

3.1 特征提取网络 EfficientNet-B0

EfficientNet 是一种利用神经网络架构搜索(Neural Architecture Search)而设计的网络模型,其通过缩放因子来控制模型的宽度和深度,以获取 EfficientNet B0~B7 模型。由于 EfficientNet-B0 在该系列中模型最小、推理速度最快,因此本文将其作为改进的 CenterNet 目标检测模型的特征提取网络。EfficientNet-B0 网络模型的结构如表 1 所列。由表可知, EfficientNet-B0 网络结构由 1 个普通卷积层及 16 个移动倒置瓶颈卷积模块(Mobile Inverted Bottleneck Convolution, MBConv)构成。其中 MBConv(1/6) 模块为其核心模块,其主要借鉴 MobileNet v2^[29] 倒残差结构。首先使用 1×1 的逐点卷积对输入特征图通道数进行一定比例的扩大,然后使用 3×3 或 5×5 大小的深度可分离卷积进行特征提取,最后使用 1×1 大小的逐点卷积对得到的特征图的通道数进行缩小。同时,在 3×3 或 5×5 大小的深度可分离卷积操作之后引入 SENet 注意力机制,通过自动学习赋予各通道不同的权重,使网络模型更关注信息量较大的通道。MBConv

(1/6)的网络结构如图2所示。

表1 EfficientNet-B0网络模型

Table 1 EfficientNet-B0 network model

stage	module	module numbers	kernel/stride	outputs
1	Conv0	1	(3,3)/2	(256,256,32)
2	MBCConv1	1	(3,3)/1	(256,256,16)
3	MBCConv6	2	(3,3)/2	(128,128,24)
4	MBCConv6	2	(5,5)/2	(64,64,40)
5	MBCConv6	3	(3,3)/1	(64,64,80)
6	MBCConv6	3	(5,5)/2	(32,32,112)
7	MBCConv6	4	(5,5)/2	(16,16,192)
8	MBCConv6	1	(3,3)/1	(16,16,320)

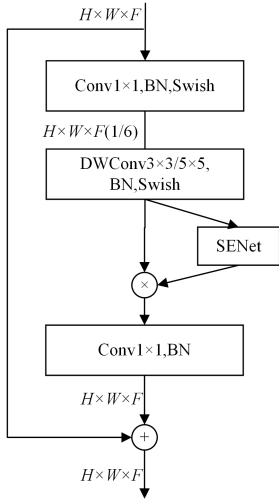


图2 MBCConv(1/6)模块
Fig. 2 MBCConv(1/6) model

图2中, SENet (Squeeze-and-Excitation)^[30] 首先进行Squeeze操作,将尺寸大小为 $H \times W \times C$ 的特征图通过全局平均池化(Global Average Pooling)对每个通道的 FeatureMap 进行压缩得到 $1 \times 1 \times C$ 的实数列;然后引入两个全连接层进行Excitation操作,通过 Sigmoid 激活函数将其归一化到 $0 \sim 1$ 之间,其代表每一通道的权重;最后将权重与原始输入对应相乘,实现对特征信息重新分配。SENet的网络结构如图3所示。

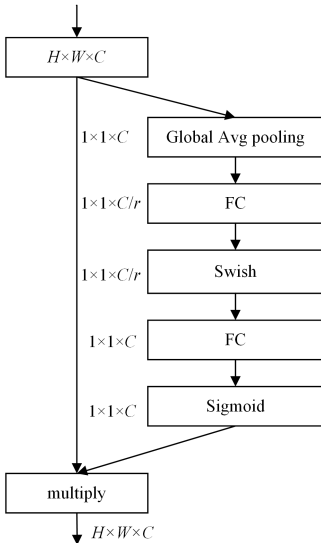


图3 SENet的网络结构
Fig. 3 SENet network structure

3.2 FEM 特征加强模块

原始 CenterNet 检测模型直接使用反卷积对尺寸为 16×16 的特征图进行3次上采样,得到尺寸为 108×108 的特征图,将其传递给检测模块,对目标类别及位置信息进行成功预测。但是直接进行上采样得到的特征图会存在一些对当前任务用处不大或无用的特征,因此设计 FEM 特征加强模块来加强有效特征,抑制无效或者效果微乎其微的特征,使网络的检测效果更好。在特征加强模块中,首先利用 1×1 卷积对输入特征的通道数进行扩增,利用 3×3 普通卷积对特征图尺寸进行调整,利用 1×1 卷积对通道数进行压缩;然后使用 SENet 注意力机制,通过自动学习获取当前任务的重要特征;接着使用 3×3 卷积对提取到的特征进行整合以减少噪声;最后使用 1×1 卷积对得到的特征图的通道数进行调整,以便于在通道维度上拼接后续特征图。FEM 模块的网络结构如图4所示。

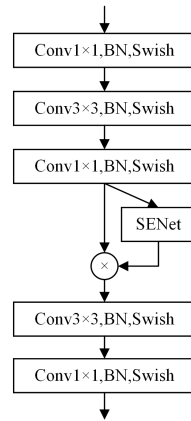


图4 FEM 模块的网络结构
Fig. 4 FEM module network structure

3.3 CA-CenterNet-Head 检测模块

虽然通道注意力机制 SENet 对提升模型性能具有显著的效果,但是其通常会忽略位置信息,而位置信息对目标的检测也非常重要。因此,本文引入轻量级 CA 注意力机制,该注意力机制是在 SENet 通道注意力机制的基础上嵌入空间位置信息,其网络结构如图5所示。

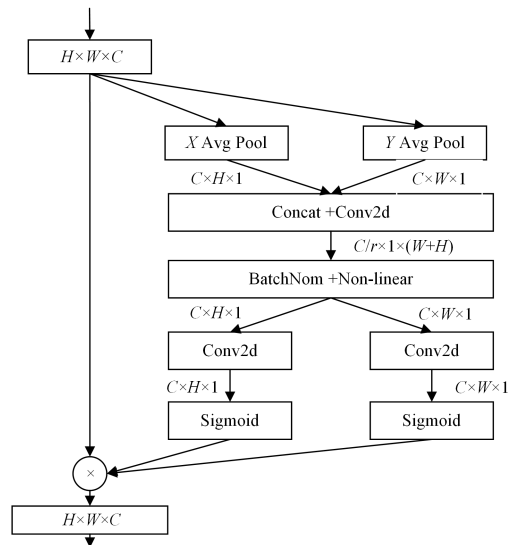


图5 CA 注意力机制
Fig. 5 CA attention mechanism

该机制将通道注意力分解为两个 1 维特征编码过程。首先,分别沿着 w 和 h 方向,通过 $(1 \times w)$ 和 $(h \times 1)$ 全局平均池化来聚合特征。这样既可以沿着一个空间方向捕获远程依赖关系,又可以沿着另一个空间方向保留精确的位置信息,将生成的特征图分别编码为一对方向感知和位置敏感的特征图。然后使用 Sigmoid 激活函数,对 w 和 h 两个维度的特征进行权重重分配。最后让输入与这两个维度的权值相乘得到输出,并将其互补地应用到输入特征图中,以增强关注对象的表示。

为了让模型在预测目标类别信息的同时也关注位置信息,本文将加入 CenterNet 的检测模块 (CenterNet-Head) 1×1 卷积与 3×3 卷积中间,构建 CA-CenterNet-Head 检测模块,其网络结构如图 6 所示。

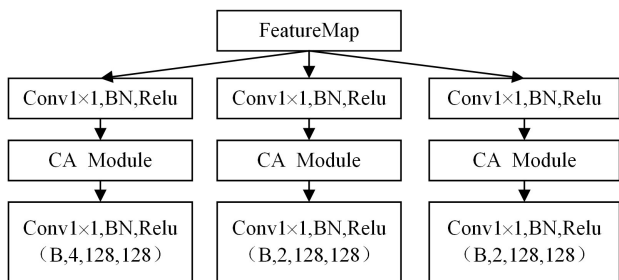


图 6 CA-CenterNet-Head 模块的网络结构

Fig. 6 Network structure of CA-CenterNet-Head module

3.4 FPN 特征融合

因浅层网络更加关注位置信息,而深层网络更加关注语义信息,若只用深层的特征来进行检测,会造成绝缘子缺陷这种小目标检测精度不高,目标的预测框不能完全贴合真实框,因此本文借鉴特征金字塔 (Feature Pyramid Networks, FPN) 来对 EfficientNet-B0 提取到的浅层与深层特征进行融合,以丰富特征层信息和提高检测精度。首先对尺寸大小为 16×16 的特征图进行上采样及 FEM 特征加强,与 32×32 的特征图在通道维度上进行拼接;然后将得到的 32×32 大小的特征图重复上述操作,得到尺寸大小为 64×64 的特征;最后将尺寸大小为 64×64 的特征图经过上采样及 FEM 特征加强与大小为 128×128 的特征图进行通道维度拼接,得到富含目标信息的特征图。

本文模型首先将轻量级卷积神经 EfficientNet-B0 作为特征提取网络,其主要由 16 个移动倒置瓶颈卷积模块构成,在保证特征提取能力的前提下,极大地减少了模型的参数量,加快了模型的推理速度。首先进行引入了 FEM 特征加强模块的 FPN 特征融合操作,通过自动学习的方式来获取对当前绝缘子及其缺陷检测任务的重要特征,削弱复杂背景信息的干扰,并把经过 FEM 特征加强的深层特征与浅层特征融合,在提高绝缘子缺陷检测精度的同时,使负责预测的特征层拥有深层的语义信息及浅层的位置信息;然后在检测模块中加入空间与通道混合的注意力机制 CA 来构建 CA-CenterNet-Head,使该模型在预测绝缘子及其缺陷类别的同时也关注其位置信息。本文模型的网络结构如图 7 所示,其中虚线框 A 表示特征提取网络 EfficientNet-B0, B 表示 FPN 特征融合

网络, C 表示检测模块 CA-CenterNet-Head。

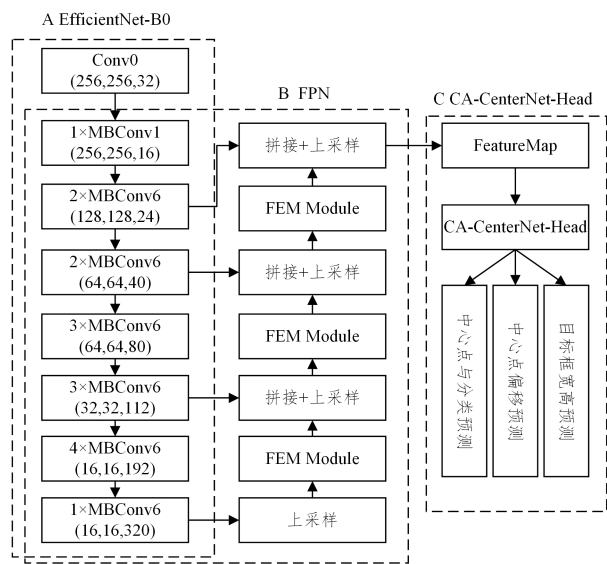


图 7 本文模型的网络结构

Fig. 7 Network structure of the proposed model

3.5 Soft-NMS 协同后处理

虽然 CenterNet 检测算法是通过热力图找到中心点,然后根据该中心点的图像特征来预测目标的类别及位置信息,但是对于绝缘子这种呈长条状的目标,难以通过热力图找到正确的中心点,可能会找到多个中心点,进而出现“单目标多框”问题,即不止一个预测框对应同一个真实物体的情况。因此,引入非极大值抑制 (Non-Maximum Suppression, NMS) 对检测结果进行后处理操作, NMS 的计算式如式 (5) 所示:

$$s_i = \begin{cases} s_i, & iou(M, b_i) < N_t \\ 0, & iou(M, b_i) \geq N_t \end{cases} \quad (5)$$

其中, s_i 代表每个边框的得分, M 为当前得分最高的框, b_i 为剩余框中的某一个, N_t 为设定的阈值, $iou(M, b_i)$ 表示两个目标框的重叠面积。可以看出,当 $iou \geq N_t$ 时,该边框的得分直接置零。

由于本文数据集中存在部分互相遮挡且连接紧密的绝缘子,因此使用传统的 NMS 进行后处理操作会将本属于两个物体但得分低的框删除,从而降低模型的召回率。为了避免出现错误抑制的情况,使用 Soft-NMS 对检测结果进行协同后处理,其未将 iou 大于阈值的边框的得分直接置 0,而是降低该边框的得分。Soft-NMS 的计算式如式 (6) 所示:

$$s_i = \begin{cases} s_i, & iou(M, b_i) < N_t \\ s_i(1 - iou(M, b_i)), & iou(M, b_i) \geq N_t \end{cases} \quad (6)$$

从式 (6) 中可以看出,利用边框的得分与 iou 来确定新的边框得分,当前边框与边框 M 的 iou 超过阈值 N_t 时,不会直接置 0,而是将它按一定比例进行衰减。

4 实验

4.1 实验环境

本实验环境为: Windows10 操作系统, 计算机 CPU 为 i7/4 核, 内存为 12 GB, GPU 为 GTX1660Ti/6 GB。运行程序软件为 Pycharm2019, Python 3.7, 安装了 Cuda10.1 和 Cudnn7.5.1

来进行深度学习加速运算,深度学习框架是 Pytorch1.4,安装 Opencv3.4 和 numpy1.18.4 等一系列第三方库来协助代码运行。

4.2 数据集

实验使用的绝缘子数据集是由无人机巡检拍摄的 3000 张包含玻璃绝缘子和复合型绝缘子的图片构成。从中随机选取 600 张作为测试集,将剩余的 2400 张作为训练集。为了提高模型的泛化能力,避免过拟合,对训练集图片进行数据增强处理,主要包括旋转、随机裁剪、色彩变换、加入高斯噪声及椒盐噪声,将其数量扩充至 6000 张。然后使用 Labelimg 目标检测图像标注软件对数据集中的玻璃绝缘子(Insulator)、玻璃绝缘子缺陷(Flaw)、复合型绝缘子(Insulator1)和复合型绝缘子缺陷(Flaw1)这 4 类目标进行标注。

其中,一张图片中大多包含多个目标,它们背景信息复杂,涵盖了山地、森林、农田、农庄、城市等多个场景。由于晴天、阴天、雾天等天气状况的影响,图片中光线强度变化较大。无人机在拍摄过程中偶尔会出现抖动现象,导致图片模糊,且存在拍摄角度随机性导致的绝缘子及其缺陷被杆塔等电力部件遮挡的情况。在这些因素的影响下,检测任务变得更有挑战性。数据集中部分玻璃绝缘子、复合型绝缘子如图 8 所示。

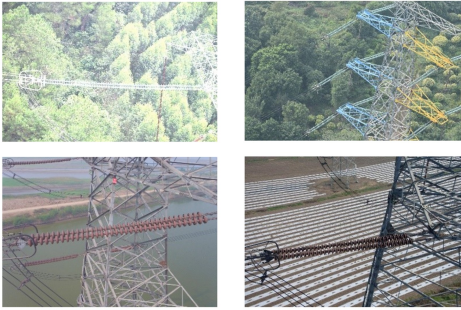


图 8 绝缘子样本

Fig. 8 Insulator samples

4.3 评价指标

本文采用模型大小(Model Size)、推理速度(FPS)、平均精度(Average Precision, AP)、均值平均精度(Mean Average Precision, mAP)作为模型的评价指标。AP 和 mAP 的计算式如式(7)~式(10)所示:

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

$$AP = \int_0^1 P(R) dR \quad (9)$$

$$mAP = \frac{\sum_i^k AP_i}{k} \quad (10)$$

其中,TP(True Positive)代表原本为正样本且被检测为正样本的目标,FP(False Positive)代表原本为负样本但被检测为正样本的目标,FN(False Negative)代表原本为正样本但被检测为负样本的目标,P 为精确率,R 为召回率。AP 代表某一类 $P \sim R$ 曲线下阴影部分的面积,k 代表待检测目标的类别数,mAP 代表所有类 AP 值总和的平均值。

4.4 模型训练

按照改进部分完成本文网络模型的构建,将批处理大小设置为 8,选用 Adam 作为优化器,分两步完成训练,共对模型训练 327 个 epoch。其中,前 10 个 epoch 对模型进行预热,将初始学习率设置为 10^{-3} ,学习率调整方式为阶梯式下降:每经过 1 个 epoch,学习率下降为原来的 90%,将剩余 317 个 epoch 的学习率设置为 10^{-4} 。为避免模型陷入局部最优,学习率调整方式设置为余弦退火衰减。最终在第 327 个 epoch 附近,训练集、验证集损失趋于稳定,损失函数值收敛于 0.9 左右。训练中的损失变化过程如图 9 所示,其中 train_loss 为训练集损失, val_loss 为验证集损失。

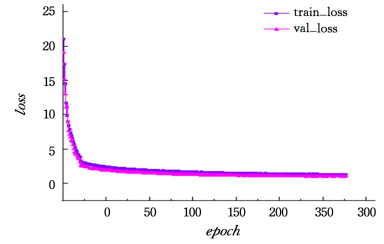


图 9 损失变化过程

Fig. 9 Loss of change process

5 实验结果及对比分析

5.1 验证各改进策略

为了验证本文模型中特征提取网络 EfficientNet-B0、FEM 特征加强模块、CA-CenterNet-Head 检测模块,FPN 特征融合及 Soft-NMS 协同后处理的有效性,采用消融实验的方式对它们进行实验对比分析。消融实验结果如表 2 所列。

表 2 消融实验结果

Table 2 Ablation test results

Experiment	Efficient-B0	FEM	CA	FPN	Soft-NMS	mAP/%	FPS/(f/s)	Model size/MB
1	—	—	—	—	—	80.33	19.92	88.00
2	✓	—	—	—	—	84.86	30.30	32.40
3	✓	✓	—	—	—	86.13	30.00	32.80
4	✓	✓	✓	—	—	88.94	29.40	33.20
5	✓	✓	✓	✓	—	91.90	29.00	34.00
6	✓	✓	✓	✓	✓	92.25	28.87	34.00

从表 2 可以看出引入不同的改进策略对检测效果的影响。与实验 1 相比,实验 2 将原始检测模型 CenterNet 的特征提取网络 ResNet18 替换为轻量级卷积神经网络 Efficient-

Net-B0,由于其网络层数较深,特征提取能力较强,并使用了移动倒置瓶颈卷积模块,因此参数量减小得较多,mAP 提高了 4.53%,FPS 提高了 10.38 f/s。与实验 2 相比,实验 3 在

实验 2 的基础上引入了 FEM 特征加强模块,其将注意力机制 SENet 与卷积操作融合,通过自动学习获取到当前任务的重要特征,抑制无效或者效果微乎其微的特征,以提高背景较为复杂的绝缘子及其缺陷的检测效果;在略微降低检测速度的情况下, mAP 提高了 1.27%。与实验 3 相比,实验 4 在实验 3 的基础上构建了 CA-CenterNet-Head 检测模块,在检测模块中引入嵌入空间位置信息的轻量级通道注意力机制 CA,使该模型在预测绝缘子及其缺陷类别信息的同时也关注位置信息。CA 的引入使模型 FPS 降低了 0.6 f/s, mAP 提高了 2.81%。与实验 4 相比,实验 5 在实验 4 的基础上对 EfficientNet-B0 提取到的特征进行 FPN 特征融合操作,将浅层与深层的特征进行融合,以丰富特征层,提高绝缘子位置的准确性及绝缘子缺陷的召回率;特征融合使 mAP 提高了 2.96%, FPS 降低了 0.4 f/s。与实验 5 相比,实验 6 在实验 5 的基础上加入 Soft-NMS 来协同后处理操作,其主要解决检测阶段对中心点预测不准导致的“单目标多框问题”,并对多余的框进行剔除,使模型 FPS 降低了 0.13 f/s, mAP 提高了 0.35%。通过上述实验对比分析得出,5 种改进策略的检测精度与速度均优于原始算法。

5.2 本文模型与原始 CenterNet 检测模型的对比结果

将本文改进的 CenterNet 模型与原始模型 ResNet18-CenterNet 进行比较,各类评价指标 AP , mAP , FPS 及 Model size 如表 3 所列。

表 3 本文模型与原始检测模型的对比

Table 3 Comparison between the proposed model and the original detection model

Model	insulator AP/%	flaw AP/%	insulator1 AP/%	flaw1 AP/%	mAP /%	FPS f/s	Model size/ MB
ResNet18-CenterNet	69	79	83	90	80.33	19.92	88
ours	94	92	92	92	92.25	28.87	34

由表 3 可知,本文模型的 mAP 比原始模型提高了 11.92%, FPS 提高了 8.95 f/s,模型大小减小了 54 MB。由于玻璃绝缘子拍摄场景大多处于山地、森林等,与背景颜色较为相似,因此 ResNet18-CenterNet 在进行检测时难以区分,导致对其检测的精度较低。但对于复合型绝缘子,其颜色与拍摄场景的背景相差较大,其检测精度相比玻璃绝缘子略有提高。对此,本文构建特征加强模块 FEM,通过合理分配通

道权重,使模型更好地区分出前景与背景,对玻璃绝缘子、复合型绝缘子这两类大目标的检测精度分别提高了 25%,9%。原始模型直接使用 3 次上采样之后的特征图来对目标进行检测,这会造成对绝缘子缺陷这样的小目标的检测效果不好,因此在本文改进的模型中借鉴 FPN 网络,将深层特征经过上采样之后与浅层特征融合,提高了小目标的检测效果。相比原始模型,本文改进的模型对玻璃绝缘子缺陷、复合型绝缘子缺陷这两类小目标的检测精度分别提高了 13%,2%。为了使模型在对绝缘子及缺陷精确分类的同时关注它们的位置信息,加入通道与空间混合的 CA 注意力机制,使模型预测目标类别信息的同时关注其位置信息。相比 ResNet18,本文使用的特征提取网络 EfficientNet-B0 使用深度可分离卷积来进行特征提取,在降低模型参数数量的同时保证了特征的提取能力,进而使模型的推理速度 FPS 提高 8.95 f/s,模型所占内存大小减少 54 MB。

从数据集中选取相同的图片并分别用两种模型进行测试,图 10 为检测效果对比图。其中,左侧为原始模型 ResNet18-CenterNet 的检测结果,右侧为本文模型的检测结果,不同颜色的矩形框代表检测到的不同目标。受恶劣天气、强光照、复杂背景、杆塔遮挡等多种客观因素和无人机拍摄距离较远、拍摄抖动等主观因素的影响,同一张航拍图像上存在不同大小及形态各异的绝缘子及其缺陷。由图 10 可知,在检测条件较好的情况下,原始算法虽可以准确检测出绝缘子及其缺陷,但本文算法的置信度更高,且预测框也更加贴合真实位置。在绝缘子较为密集和拍摄距离较远的复杂背景下,绝缘子与自然背景差异较小,不易区分,且原始算法存在对绝缘子漏检的情况。在河流背景下,由于复合型绝缘子的颜色与河流较为相像,因此较难检测出图片中的绝缘子,导致原始算法对绝缘子的定位不准确。受强光照及无人机拍摄抖动的影响,原始算法可以检测到尺寸较大的绝缘子,但是对于尺寸较小的绝缘子缺陷仍存在漏检的情况。在绝缘子被杆塔遮挡的情况下,原始模型不仅存在绝缘子漏检的情况,还存在将电力杆塔误检为绝缘子的情况。由于阴雨天气使得玻璃绝缘子的缺陷颜色与天空的颜色较为相近,不易区分,因此原始算法对尺寸较小且背景较复杂的绝缘子缺陷存在漏检。从图 10 右侧可以看出,本文模型可以很好地对此 8 种情况下的绝缘子及其缺陷进行检测,证明了本文模型的鲁棒性。

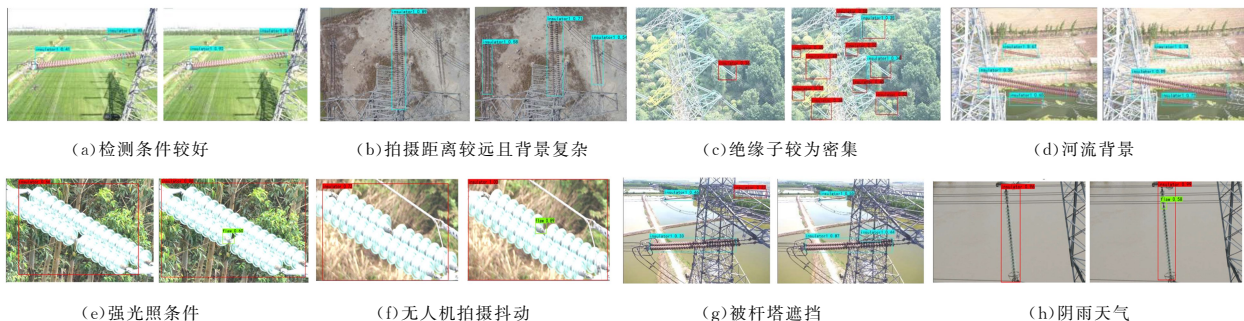


图 10 本文模型与原始模型的检测结果(电子版为彩图)

Fig. 10 Detection results of the proposed model and original model

我们使用 Soft-NMS 来解决“单目标多框”问题, iou 阈值

设置为 0.5。该模型对于同一个目标上产生的多余的框可以

进行准确的过滤,进而提高模型的检测精度,其处理效果如图 11 和图 12 所示。其中,图 11 为出现“单目标多框”问题的绝缘子图片,图 12 为经过 Soft-NMS 协同操作的绝缘子图片。

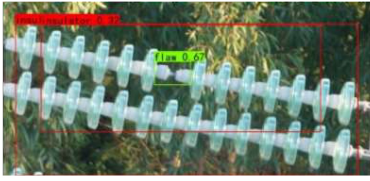


图 11 单目标多框

Fig. 11 Single target multiple boxes

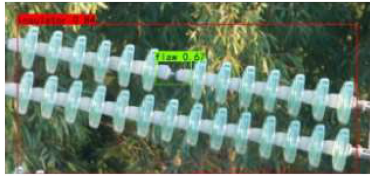


图 12 经过 Soft-NMS 操作

Fig. 12 After Soft-NMS operation

5.3 本文模型与其他检测模型的对比结果

为进一步验证本文模型的鲁棒性,在同一实验环境下分别与 ResNet50-CenterNet, YOLOv3, YOLOv4, M2det, RetinaNet, Efficientdet 和 YOLOv4-Tiny 这 7 种检测模型进行对比分析,实验结果如表 4 所列。

表 4 本文模型与其他检测模型的对比

Table 4 Comparison of the proposed model with other detection models

Model	insulator AP/%	flaw AP/%	insulator1 AP/%	flaw1 AP/%	mAP/%	FPS f/s	Model size/MB
ResNet50-CenterNet	84	87	85	87	85.75	16.07	124
YOLOv3	72	73	72	81	74.47	16.03	235
YOLOv4	89	90	90	89	89.23	17.86	244
M2det	87	86	88	91	87.99	13.47	228
RetinaNet	85	84	88	84	84.99	18.42	145
Efficientdet	82	67	73	73	73.62	34.23	16
YOLOv4-Tiny	73	72	79	74	74.50	31.28	23
Ours	94	92	92	92	92.25	28.87	34

从表 4 可以看出,本文模型与同属 Anchor-Free 系列的 ResNet50-CenterNet 检测模型相比,在模型尺寸减小 90 MB 的情况下,速度与精度均有所提升。虽然 ResNet50 可以提取到很好的特征,但是并未将提取到的特征作进一步的融合,导致负责预测的特征图包含的绝缘子信息并不丰富,获得的检测精度略低。ResNet50 的参数数量也比本文使用的 EfficientNet-B0 高,这会导致模型的推理速度降低,模型尺寸增大。

与 Anchor-Boxes 系列检测模型 YOLOv3、YOLOv4 相比,本文模型的检测速度与精度均有所提高。YOLOv3 和 YOLOv4 分别使用了 FPN 网络、PANet 网络来对浅层特征和深层特征进行融合,丰富负责预测的特征图信息,但是由于其是对特征直接进行融合,可能会融合一些对当前任务无用的特征,进而影响检测效果。M2det^[31]是一种将 MLFPN 与 SSD 相结合的检测模型,其使用 VGG16 和 TUM 模块进行特

征提取,虽然可以有效提取特征,但是相比本文模型其参数量较大,导致模型的推理速度降低,模型尺寸增大。RetinaNet^[32]直接将 ResNet50 提取到的特征经过 FPN 特征融合,没有对融合的特征进行合理删选。其在检测模块分两部分来分别预测绝缘子的类别信息及位置信息,导致目标的预测位置不能很好的贴合其真实位置。相比上述检测模型,本文构建 FEM 特征加强模块,对 FPN 融合的特征进行合理删选,在加强有效特征的同时过滤无用特征,使模型更好地区分出前景中的绝缘子及其缺陷,削弱背景信息的干扰。在检测模块中加入空间与通道混合的 CA 注意力机制,以构建 CA-CenterNet-Head 检测模块,使该模型在准确预测绝缘子及其缺陷类别信息的同时也可以关注其位置信息。

由表 4 可知,本文模型与同属轻量级检测模型的 Efficientdet^[33]和 YOLOv4-Tiny^[34]相比,mAP 分别提高了 18.63%,17.75%;FPS 分别降低了 5.36 f/s,2.41f/s;Model size 分别增加了 18 MB,11 MB。虽然 Efficientdet 检测网络使用特征提取网络 EfficientNet-B0,模型尺寸及检测速度均优于本文模型,但是对提取到的绝缘子特征并未进行特征加强处理,使其对两类绝缘子及缺陷的识别精度远低于本文模型。YOLOv4-Tiny 作为 YOLOv4 检测模型的精简版,相比 YOLOv4,其检测速度有较大的提高,但检测精度远低于本文模型。

结束语 为了使无人机在电力巡检过程中对绝缘子缺陷进行实时高效的检测,本文提出了一种改进 CenterNet 的绝缘子缺陷检测模型。在特征提取方面,本文使用参数量更小、特征提取能力更强的 EfficientNet-B0 代替原始模型中的 ResNet18,并搭建 FEM 特征加强模块,增强对当前任务重要的特征。借鉴 FPN 网络将深层与浅层的特征相融合,使负责预测的特征图信息更加丰富,以提高小目标的检测精度。在检测部分构建 CA-CenterNet-Head 模块,使该模型在预测目标类别信息的同时可兼顾其位置信息。针对中心点预测不准确导致的“单目标多框”问题,本文使用 Soft-NMS 对冗余的框进行过滤。实验结果表明,本文改进的 CenterNet 模型的 mAP 达到 92.25%,FPS 达到 28.87 f/s,模型大小为 34 MB,可以实时高效地完成对航拍绝缘子缺陷的检测。本文实验均在实验室设备上完成,下一步工作将继续优化网络模型,提高其检测精度及速度,并将其部署到无人机嵌入式设备中来完成对绝缘子缺陷的检测。

参考文献

- [1] LING S, NGUYEN K, ROUX L A, et al. A lattice E based group signature scheme with verifier-local revocation [J]. Theoretical Computer Science, 2018, 730(19): 1-20.
- [2] ZHANG M J, LI H W, ZHAO W H, et al. Application of deep learning in unmanned aerial vehicle patrol inspection of military optical cable line [J]. Optical Communication Research, 2018 (6): 57-61.
- [3] YANG G, SUN C W, WANG D W, et al. A Comparative Study of Transmission Line Component Detection Models Based on UAV Front-end SSD Algorithm [J]. Journal of Taiyuan University of Technology, 2020, 51(2): 212-219.
- [4] REN S Q, HE K M, GIRSHICK R, et al. Faster rcnn: Towards

- real-time object detection with region proposal networks[C]// Advances in Neural Information Processing Systems. Montreal: MIT press, 2015: 91-99.
- [5] DAI J, LI Y, HE K M, et al. R-FCN: Object detection via region-based fully convolutional networks[C]// Advances in Neural Information Processing System. 2016: 379-387.
- [6] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 580-587.
- [7] GIRSHICK R. Fast R-CNN[C]// Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1440-1448.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// Proceedings of the 14th European Conference on Computer Vision (ECCV). Amsterdam, Netherlands: Springer, 2016: 21-37.
- [9] FU C Y, LIU W, RANGA A, et al. DSSD: Deconvolutional single shot detector[J]. arXiv:1701.06659, 2017.
- [10] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// Computer Vision and Pattern Recognition. IEEE, 2016: 779-788.
- [12] REDMON J, FARHADI A, GIRSHICK R. YOLO900: faster, stronger[C]// Computer Vision and Pattern Recognition. IEEE, 2017: 6517-6525.
- [13] REDMON J, FARHADI A. Yolov3: An incremental improvement[OL]. [2019-06-03]. <https://arxiv.org/abs/1804.02767>.
- [14] LAW H, DENG J. Cornernet: Detecting object as paired keypoints[C]// Proceedings of the 15th European Conference on Computer Vision. Munich, Germany, 2018: 734-750.
- [15] LAW H, TENG Y, RUSSAKOVSKY O, et al. CornerNet-Lite: Efficient keypoint based object[J]. arXiv:1904.08900, 2019.
- [16] TIAN Z, SHEN C H, CHEN H, et al. FCOS: Fully Convolutional one-stage object detection[J]. arXiv:1904.01355, 2019.
- [17] ZHAO R, ZHAO G W, ZHANG J, et al. Real-time fault detection method for high voltage transmission line based on improved algorithm[J]. Computer Engineering and Applications, 2021, 57(17): 7.
- [18] YI J Y, CHEN C F, GONG G Q. Aerial Insulator Detection of Transmission Line Based on Improved Faster RCNN [J]. Computer Engineering, 2021, 47(6): 292-298, 304.
- [19] DUAN K W, BAI S, XIE L X, et al. CenterNet: Keypoint triplets for object detection[J]. arXiv:1904.08189, 2019.
- [20] AN M, LE Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks [C]// International Conference on Machine Learning. Long Beach, USA: ICML, 2019: 6105-6114.
- [21] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// IEEE Conference on Computer Vision and Pattern Recognition. 2017.
- [22] HOU Q B, ZHOU D Q, FENG J S. Corinate Attention for Efficient Mobile Network Design [C] // Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition. 2021.
- [23] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-nms improving object detection with one line of code[C]// ICCV. Venice, USA: IEEE Press, 2017: 5562-5570.
- [24] NEUBECK A, VAN G L. Efficient non-maximum suppression [C]// Proceedings of 18th International Conference on Pattern Recognition. 2006: 850-855.
- [25] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // CVPR. Las Vegas, USA: IEEE Press, 2016: 770-778.
- [26] ZHANG T, TAO D, YANG J, et al. Discriminative Locality Alignment [C] // European Conference on Computer Vision. 2008: 725-738.
- [27] NEWELL A, YANG K Y, DENG J. Stacked hourglass networks for human pose estimation[C]// Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands, 2016: 483-499.
- [28] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C] // Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2980-2988.
- [29] MARK S, ANDREW H, MENGLONG Z, et al. MobileNet v2: Inverted residuals and linear bottlenecks[C]// CVPR. Salt Lake City, USA: IEEE Press, 2018: 4510-4520.
- [30] JIE H, LI S, GANG S. Squeeze-and-excitation networks[C]// CVPR. Salt Lake City, USA: IEEE Press, 2018: 7132-7141.
- [31] ZHAO Q, SHENG T, WANG Y, et al. M2det: A single-shot object detector based on multi-level feature pyramid network[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2019: 9259-9266.
- [32] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]// Proceeding of the IEEE International Conference on Computer Vision. 2017: 2980-2988.
- [33] TAN M, PANG R, LE Q, et al. EfficientDet: Scalable and efficient object detection [J]. arXiv:1911.09070, 2020.
- [34] WANG C Y, LIAO H Y. YOLOv4: Optional speed and accuracy of objection detection [OL]. [2020-04-23]. <https://arxiv.org/abs/2004.10934>.



LI Fa-guang, born in 1996, postgraduate. His main research interests include deep learning and electric power inspection.



YILIHAMU Yaermaimaiti, born in 1978, master, associate professor, master supervisor. His main research interests include pattern recognition and artificial intelligence.