

基于改进的 U-Net 和 YOLOv5 的绝缘子掩模获取与缺陷检测

唐小煜^{1,2,3,4}, 熊浩良^{1,2,3,4}, 黄锐珊^{1,2,3,4}, 林威霖^{1,2,3,4}

(1. 华南师范大学物理与电信工程学院, 广州 510006; 2. 广东省量子调控工程与材料重点实验室, 广州 510006; 3. 广东省光电检测仪器工程技术研究中心, 广州 510006; 4. 华南师范大学物理国家级实验教学示范中心, 广州 510006)

摘要: 输电线路的绝缘子定期巡检是必不可少的一项任务, 而传统的人工巡检存在着效率低、工作强度大等问题。因此, 本文设计了一种改进的 U-Net 模型实现对绝缘子的分割, 并使用改进的 YOLOv5 实现在复杂背景下对爆破绝缘子的定位。本文基于 U-Net 图像语义分割模型, 提出一种改进的网络结构 SERes-Unet。模型引入残差结构减少卷积过程中存在的梯度消失、结构信息损耗的影响, 引入注意力机制对特征权重进行校正, 从而提升网络性能。为实现对高分辨率图像的爆破绝缘子检测, 提出将图片进行切割再进行检测, 再通过非极大值抑制 (Non-maximum suppression, NMS) 进行筛选, 获取图像全部爆破绝缘子的位置。本文设计的多组实验验证了模型的有效性和高效性。本文方法绝缘子分割精度达到 0.96, 爆破绝缘子检测精确率达到 0.97, 召回率达到 0.99。

关键词: 爆破绝缘子; 图像语义分割; 目标检测; U-Net 模型

中图分类号: TP391

文献标志码: A

Insulator Mask Acquisition and Defect Detection Based on Improved U-Net and YOLOv5

TANG Xiaoyu^{1,2,3,4}, XIONG Haoliang^{1,2,3,4}, HUANG Ruishan^{1,2,3,4}, LIN Weilin^{1,2,3,4}

(1. School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou 510006, China; 2. Guangdong Provincial Key Laboratory of Quantum Engineering and Quantum Materials, Guangzhou 510006, China; 3. Guangdong Provincial Engineering Technology Research Center for Optoelectronic Instrument, Guangzhou 510006, China; 4. National Demonstration Center for Experimental Physics Education, South China Normal University, Guangzhou 510006, China)

Abstract: Regular inspection of insulators of transmission lines is an indispensable task, while traditional manual inspections have problems such as low efficiency and high work intensity. Therefore, this paper designs an improved U-Net model to realize the segmentation of insulators, and uses an improved YOLOv5 to realize the positioning of blasting insulators in complex backgrounds. Based on the U-Net image semantic segmentation model, this paper proposes an improved network structure SERes-Unet. The model introduces residual structure to reduce the influence of gradient disappearance and structural information loss in the convolution process, and introduces an attention mechanism to correct feature weights, thereby improving network performance. In order to realize the detection of blasting insulators on

high-resolution images, it is proposed to cut the pictures and then detect them, and then filter through Non-Maximum suppression (NMS) to obtain the positions of all blasting insulators in the image. The article designs multiple sets of experimental controls to verify the effectiveness and efficiency of the model. In the end, the method achieves an insulator segmentation accuracy of 0.96, a blasting insulator detection accuracy of 0.97, and a recall rate of 0.99.

Key words: blasting insulator; image semantic segmentation; object detection; U-Net model

引 言

为了保证输电线路的可靠运行,输电线变电系统需要进行定期的巡检、维修以及维护来确保消除故障或者隐患。绝缘子在输电线路中起着很大的作用,由于长期遭受恶劣天气的影响,以及器件的使用寿命及劳损,绝缘子很容易产生缺陷,严重威胁输电线路的安全可靠运行。传统的电网巡查方式通过人工进行巡检,该方法存在劳动强度大,工作效率低,在危险地段会危及到巡查工人生命安危等问题。

随着近些年来计算机技术的飞速发展及计算机处理性能的增强,利用计算机图像处理的方法对不同领域的缺陷定位与检测已经成为主要趋势^[1-4]。

而针对绝缘子的分割和爆破绝缘子的检测,目前已有不少研究通过图像处理技术与机器学习算法进行实验并取得了一定的成果。朱文天等^[5]提出红绿色差灰度化方法,并结合最大类间方差法(Otsu)、红蓝色差灰度化、K均值聚类算法(K-means)和投影特征实现绝缘子分割;王森等^[6]提出通过阈值分割、图像边缘检测等方法检测爆破绝缘子,最终实现在复杂背景下的识别正确率为85%;商俊平等^[7]采用分割提取特征的方法,利用最大类间方差法对绝缘子进行分割并提取绝缘子的不变特征值,使用Ada-boost分类器^[8]定位绝缘子位置并计算相邻绝缘子的欧氏距离来对爆破绝缘子进行定位,最终爆破绝缘子的识别准确率达到87%。

基于卷积神经网络的算法模型在目标检测及图像分割领域的蓬勃发展,为绝缘子的分割提取和缺陷检测提供了更为高效的方法与思路。U-Net模型常被应用于医学图像的语义分割,本文通过引入注意力机制和残差结构对网络结构进一步优化,使网络能够捕获更多高低层的特征信息,进一步提高绝缘子的分割精度;同时结合引入注意力机制的改进YOLOv5,实现对爆破绝缘子的快速和精确定位。

1 基于U-Net改进的绝缘子分割算法

1.1 U-Net算法原理

U-Net是一个能够使用少量训练图像进行训练并且做到精准分割的网络。获取绝缘子的掩模需要耗费大量的人工以及时间,使用U-Net则可以克服这一点。U-Net主要由收缩路径和扩张路径组成,收缩路径和扩张路径基本上是对称的,因此是一个U形的网络,网络结构如图1所示。收缩路径是经典的卷积网络架构,用于特征提取,包含4层,每层由2个卷积层和1个最大池化层以及ReLU激活函数组成。每层下采样后特征通道数加倍。扩张路径的每一层都先使用反卷积,每次反卷积后特征通道数减半,特征图大小加倍。反卷积后的结果和收缩路径对应步骤的特征图拼接起来后做2次卷积。通过将收缩过程中产生的高分辨率的特征拼接在上采样的结果,可以更好地利用浅层的局部信息,获取更精确的结果。

1.2 算法的改进

对于绝缘子掩模的获取,可以将高分辨率的图像切割后作为原始输入图片,将绝缘子掩模提取视

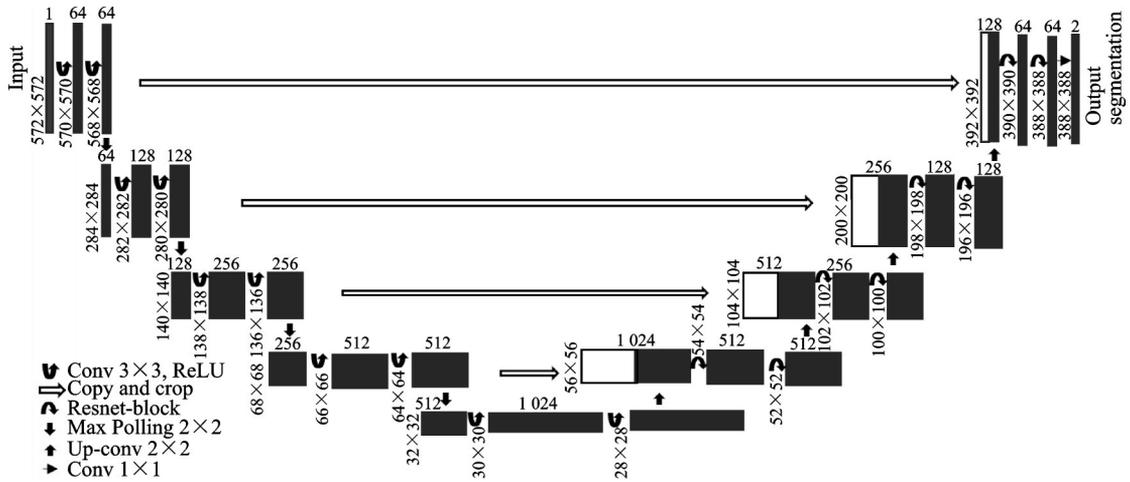


图1 U-Net的网络结构

Fig.1 Network structure of U-Net

为一个与绝缘子与背景的图片二分类分割问题。本文利用Ronneberger等^[9]提出的U-Net、He等^[10]提出的Resnet block以及Hu等^[11]提出的SE block来实现这种分割。网络的整体结构如图2所示。

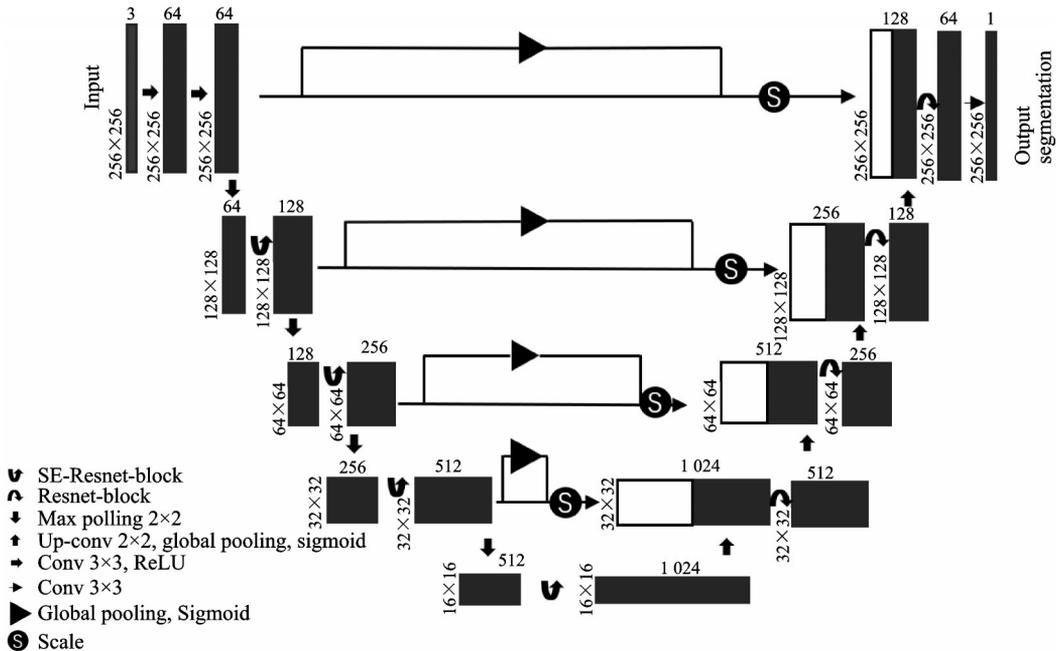


图2 SERes-Unet网络结构

Fig.2 Network structure of SERes-Unet

模型在U-Net的基础上,将收缩路径和扩张路径的两个3x3的卷积层替换成残差结构(Resnet-block),模块的组成如图3(a)所示,由两个重复的结构,3x3的卷积层以及批量归一化层(Batchnormalization, BN)和ReLU激活函数组成,使用Resnet-block可以减少卷积过程中存在梯度消失、结构信息损耗的影响,加入BN层对输入进行标准化处理,可以使网络的训练变得更加容易。在收

缩路径的 Resnet-block, 添加了压缩-奖惩模块 (Squeeze-and-Excitation block, SE block), SE block 是一种注意力机制, 通过支路结构学习如何评估通道间的关联, 作用到原特征图上, 实现对输入的校正。收缩路径模块的组成如图 3(b) 所示。同时, 对收缩路径每一层产生的特征图以及扩展路径每一层产生的特征图也使用了由全局池化层 (Global pooling) 全连接层 (Fully connected layer, FC) 以及 Sigmoid 激活函数组成的注意力机制, 通过学习选择性地强调有用的特征和抑制不太有用的特征, 再将特征图进行拼接。

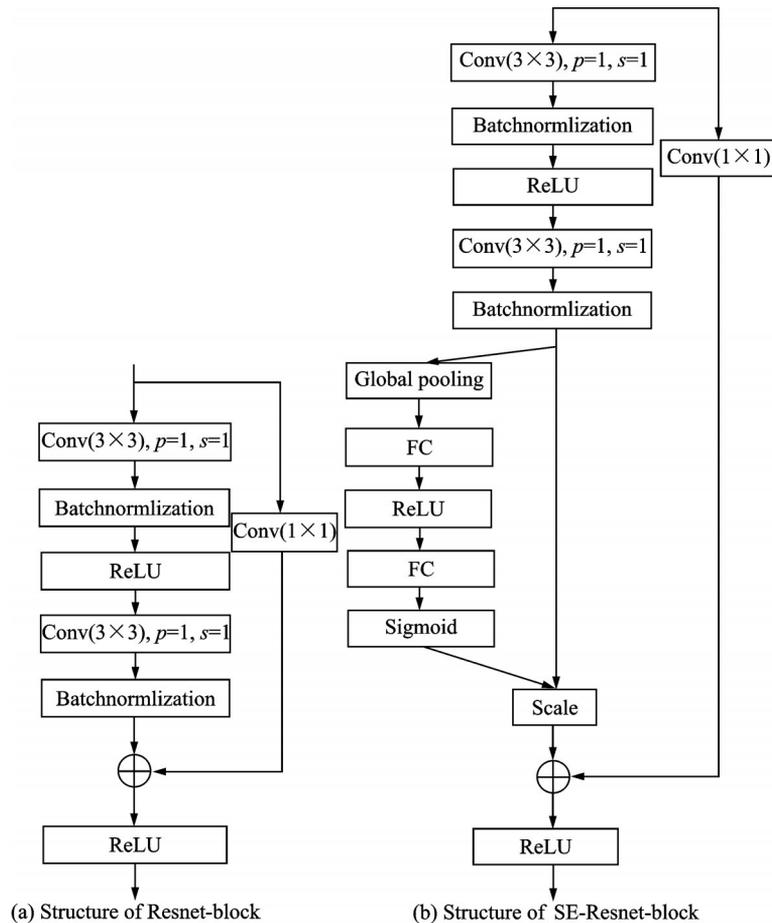


图 3 Resnet-block 的结构

Fig.3 Structure of Resnet-block

2 基于改进 YOLOv5 的绝缘子缺陷检测

2.1 YOLOv5 算法原理

目前在目标检测领域主要分为单阶段目标检测器和两阶段目标检测器。而经过优化和改进, 以通用目标检测器 (Single shot multiBox detector, SSD)^[12]、目标检测算法 (You only look once, YOLO)^[13] 为代表的单阶段目标检测在检测速度和精度上都有了较大的提升, 特别是如 EfficientDet^[14]、YOLOv5 等目标模型在检测精度和速度均能超越以 Faster-RCNN 为代表的两阶段目标检测器, 具有更广阔的应用场景。

YOLOv5算法在YOLOv4的基础上作进一步的优化,在输入端采取Mosaic增强、自适应图片缩放等方法,不仅能够完成对数据的增强,更能通过减少由图片缩放所带来的冗余信息,从而提高了推理速度。此外,YOLOv5在骨干网络中结合了Focus结构和局部跨阶段结构(Cross stage partial, CSP)结构以及特征金字塔(Feature pyramid networks, FPN),路径聚合网络(Path aggregation network, PAN)等特征融合技巧,利用跨阶段层次结构将特征信息融合,在减少计算量的同时保证了正确率。通过修改网络的深度和宽度,最终YOLOv5迭代出YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x共4种不同深度的模型,并随着模型的复杂度增加,模型的识别准确率不断提高。

2.2 算法的改进

改进的YOLOv5模型在YOLOv5s的基础上,在模型的预测层前加入SE block,将经过前面卷积层后的融合特征分别作为SE block的输入,通过SE block分支结构上的全连接层与ReLU、Sigmoid激活函数,实现利用ReLU激活函数增加网络的非线性复杂度的同时,也能利用Sigmoid激活函数将输出缩放到0-1的区间,从而能表现出各特征维度上(Channel-wise)的信息重要性。最终通过引入注意力机制实现对输入的校正、增强模型对有效特征信息的选择,进一步提高检测的精度,提高网络的泛化能力。改进后模型结构如图4所示。

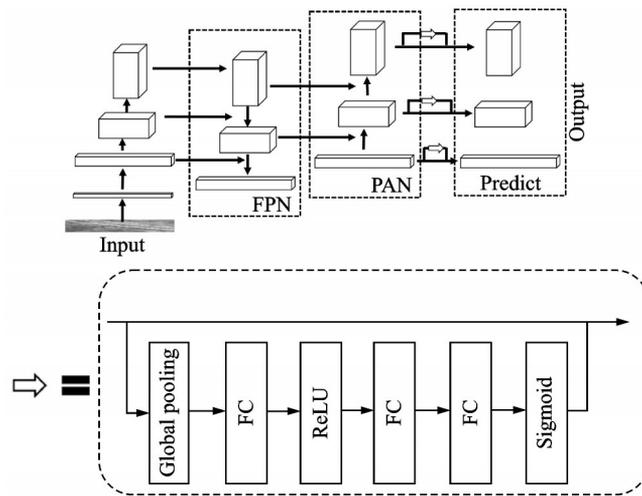


图4 改进的YOLOv5网络结构

Fig.4 Network structure of improved YOLOv5

3 实验和实验结果

本实验软件运行的操作系统为Ubuntu16.04,硬件服务器的配置为:CPUi7-7800X、两块NVIDIA 1080Ti GPU显卡、内存为16 GB。

3.1 图像分割模型

3.1.1 模型训练

本文实验选取由无人机拍摄获得的,分辨率为 $4\ 912 \times 7\ 360$ 、 $2\ 624 \times 3\ 936$ 、 $4\ 016 \times 6\ 016$ 的输电线路图像进行训练。在分割绝缘子阶段,数据被视为一个二分类问题,其中绝缘子为前景,除绝缘子外的其他物体皆为背景。

为了降低模型过拟合的风险,对原始数据进行了数据增强,包括水平翻转增强、垂直翻转增强、高

斯噪声增强、像素乘法增强、裁剪填充增强。同时,为了标准化输入数据,对输入的图像以及其对应掩模图像进行了归一化处理。

模型选取二进制交叉熵函数和 Sigmoid 函数的结合作为损失函数,Adam 算法作为学习率的优化算法。其中初始学习率 L_r 设置为 $4e-4$,并使用自适应调整对学习率进行调整,超过 4 个 Epochs 准确率没有提升则将学习率 L_r 调整为 $0.8 \times L_r$ 。训练的批量大小(Batch_size)设置为 16。

3.1.2 实验结果及分析

本文实验对原始数据进行多种方式的数据增强,并将数据增强前后的数据在 U-Net 模型上进行训练和效果验证,最终效果对比如表 1 实验结果表明经数据增强后,训练数据量从 5 000 张图像增加到 12 000 张图像,并且使用经过增强后的数据进行训练的模型能在任务上取得更佳的效果。

本文将 U-Net、Res-Unet、SERes-Unet 进行了对比,其中,Res-Unet 是指在 U-Net 的基础上将卷积层替换为 Resnet-Block 以及添加了 SE Block 的 Resnet-Block 的网络,SERes-Unet 是指在 Res-Unet 的基础上,在拼接的特征图上添加由全局池化层以及 Sigmoid 激活函数组成的注意力机制的网络。随着迭代步数的增加,模型 Dice 系数的变化如图 5 所示。最终实验结果如表 2 所示,实验结果表示 SERes-Unet 在 Dice 系数上达到了最佳。

预测结果如图 6 所示,图 6(a)为输入的原始图像,将原始图像切割成 256×256 的图像加载到模型中进行推理,再将得到的结果拼接合成原始图像的预测图像。其中图 6(b)表示 U-Net 网络的预测结果,图 6(c)表示 Res-Unet 网络的预测结果,图 6(d)表示 SERes-Unet 的预测结果。从预测结果图像上看,相较于 Res-Unet 以及 SERes-Unet,U-Net 网络的识别精度较低,存在更多背景的错误,对于较小的绝缘子的识别,边缘的完整性较差。SERes-Unet 相较于 Res-Unet,识别精度进一步提升了。从指标上也可以得出,SERes-Unet 的精度相较于 Res-Unet 的精度提升了 0.013 7,因此本文使用 SERes-Unet 进行绝缘子掩模的提取。

对于较小的、颜色与背景接近的绝缘子,相比 U-Net 网络的预测结果,改进后的模型对于绝缘子的识别更加准确,掩模的边缘更加完整,图 7 显示了原始模型与改进模型预测结果的细节对比。

3.2 目标检测模型

3.2.1 模型训练

在目标检测部分,实验中使用无人机拍摄的图像作为原始数据,由于原始数据的分辨率较大,直接

表 1 数据增强前后的在 U-Net 模型上的效果对比

Table 1 Performance comparison on the U-Net model before and after data augmentation

条件	训练数据量	Dice 系数
数据增强前	5 000	0.860 1
数据增强后	12 000	0.927 5

表 2 原始模型与改进优化模型的效果对比

Table 2 Performance comparison between the original and improved models

模型	Dice 系数
U-Net	0.927 5
Res-Unet	0.954 8
SERes-Unet	0.968 5

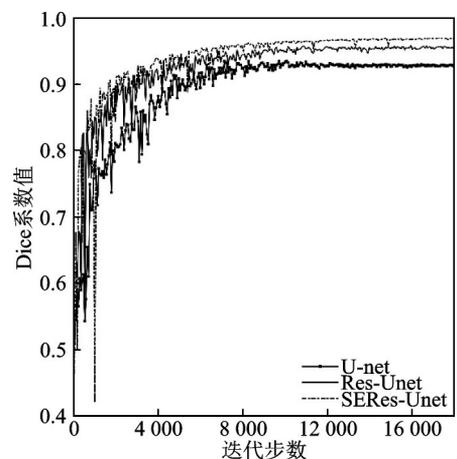


图 5 原始模型与改进模型的 Dice 系数
Fig.5 Dice coefficient of the original and improved models

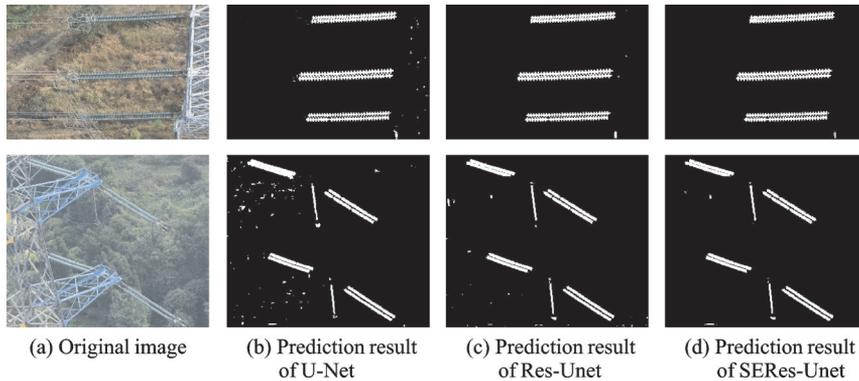


图 6 原始模型与改进模型预测结果对比

Fig.6 Comparison of prediction results between original and improved models

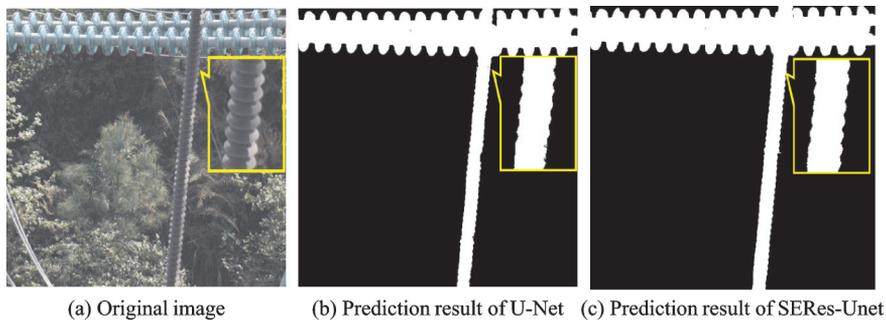


图 7 原始模型与改进模型预测结果的细节对比

Fig.7 Detail comparison of the prediction results between original and improved models

输入模型中存在大量细节丢失的问题,因此在实验中首先对原始数据集中的爆破绝缘子的边界框(Bounding box)的坐标信息进行提取,通过设置爆破绝缘子区域和切割边界之间的间隔,只保存边界框与切割边界的距离大于该间隔的图像,否则丢弃。使用该方法对数据进行清洗。

通过对原始数据的切割,确保了数据集中的图像均包含爆破绝缘子,且有效降低了图像的分辨率,更有利于模型的训练。

模型使用随机梯度下降算法(Stochastic gradient descent,SGD)算法作为优化算法,初始学习率设置为 $1e-2$,训练的批量大小(Batch_size)设置为 32。使用自适应锚定框自动学习锚定框的尺寸,共训练 400 个 Epochs。

3.2.2 实验结果与分析

本文将 YOLOv5 和改进的 YOLOv5 以及目前最先进的一阶段算法之一 EfficientDet 进行比较。实验结果如表 3 所示,结果表明,针对爆破绝缘子的定位问题,YOLOv5 相对于 EfficientDet 召回率以及精确率更高。引入注意力机制优化后的 YOLOv5 算法提升了爆破绝缘子的定位精确率。通过实验验证,优化后的模型降低了爆破绝缘子的误识率,优化 YOLOv5 训练曲线如图 8 所示。

表 3 目标检测模型与改进模型的效果对比

Table 3 Performance comparison between target detection model and improved model

目标检测算法	精确率	召回率	平均处理时间/ms
EfficientDet	0.811	0.892	67
YOLOv5	0.953	0.990	14
改进 YOLOv5	0.973	0.990	15

4 结束语

实验通过采集高分辨率的高空绝缘子图像,采用裁剪、增强的方式对原始数据集进行扩充,同时提出一种改进的U-Net模型——SERes-UNet,并通过实验与原始U-Net等模型进行对比,结果表明引入注意力机制和残差结构能有效提高对绝缘子的分割精度;同时针对爆破绝缘子定位问题,本文也利用目前先进的一阶段检测算法YOLOv5与EfficientDet进行对比,并在YOLOv5中引入注意力机制,提升了模型的精确率。经实验验证,YOLOv5算法能够以较快检测速度和较高检测精度完成对绝缘子缺陷的定位,可大大减少高强度的人工作业,对于输电线路的绝缘子定期巡检任务有一定参考价值。

无论是图像切割网络还是目标检测网络,大量的训练数据不可或缺。实验中虽已经采用了多种数据预处理的方法来对数据进行扩充,但由于原始数据场景较少,因此即便经过数据增强,仍具有一定的局限性。为进一步提升此方法的鲁棒性和实用性,未来的研究计划将会拓宽途径去收集更多的数据样本,或者通过生成对抗网络等智能算法去生成数据,使技术可以达到真正用于现实场景的标准。

参考文献:

- [1] 陈泽斌,罗文婷,李林.基于改进U-Net模型的路面裂缝智能识别[J].数据采集与处理,2020,35(2): 260-269.
CHEN Zebin, LUO Wenting, LI Lin. Automatic identification of pavement crack using improved U-net model[J]. Journal of Data Acquisition and Processing, 2020, 35(2): 260-269.
- [2] 李琛,骆汉宾,魏威,等.基于图像的混凝土表面裂缝和孔洞检测[J].土木工程与管理学报,2020,37(6): 118-123.
LI Chen, LUO Hanbin, WEI Wei, et al. Detection of concrete surface cracks and bugholes based on image[J]. Journal of Civil Engineering and Management, 2020, 37(6): 118-123.
- [3] 张磊,郎贤礼,王乐.基于图像融合与YOLOv3的铝型材表面缺陷检测[J].计算机与现代化,2020(11): 8-15.
ZHANG Lei, LANG Xianli, WANG Le. Surface defect detection of aluminum profile based on image fusion and YOLOv3[J]. Journal of Computer and Modernization, 2020(11): 8-15.
- [4] 董家顺,王兴东,李殿杰,等.基于改进K-means算法的钢管表面缺陷视觉检测方法[J].武汉科技大学学报,2020,43(6): 439-446.
DONG Jiashun, WANG Xingdong, LI Dianjie, et al. Visual inspection method for surface defects of steel pipes based on improved K-means algorithm[J]. Journal of Wuhan University of Science and Technology. 2020, 43(6): 439-446.
- [5] 朱文天,胡立坤,王帅军.改进色差法的绝缘子图像分割方法[J].广西大学学报(自然科学版),2019,44(5): 1276-1283.
ZHU Wentian, HU Likun, WANG Shuaijun. Insulator images segmentation method based on improved color difference[J]. Journal of Guangxi University(Natural Science Edition), 2019, 44(5): 1276-1283.
- [6] 王淼,杜毅,张忠瑞.无人机辅助巡视及绝缘子缺陷图像识别研究[J].电子测量与仪器学报,2015,29(12): 1862-1869.
WANG Miao, DU Yi, ZHANG Zhongrui. Study on power transmission lines inspection using unmanned aerial vehicle and image recognition of insulator defect[J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(12): 1862-1869.

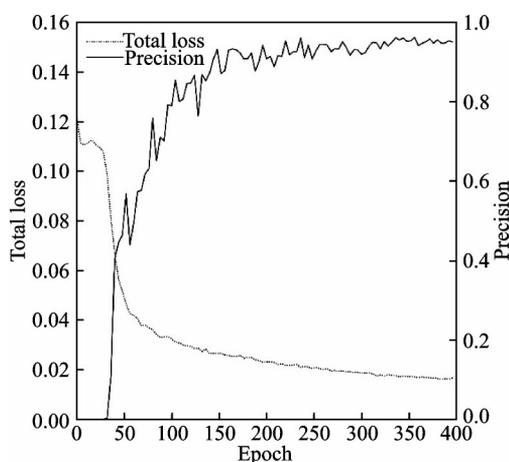


图8 优化YOLOv5训练曲线图

Fig.8 Improved YOLOv5 training curves

- [7] 商俊平,李储欣,陈亮.基于视觉的绝缘子定位与自爆缺陷检测[J].电子测量与仪器学报,2017,31(6):844-849.
SHANG Junping,LI Chuxin,CHEN Liang. Location and detection for self-explode insulator based on vision[J]. Journal of Electronic Measurement and Instrumentation, 2017,31(6):844-849.
- [8] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997,55(1):119-139.
- [9] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention.Cham: Springer, 2015:234-241.
- [10] 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. [S.l.]: IEEE, 2016:770-778.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [S.l.]: IEEE, 2016:7132-7141.
- [12] LIU Wei. SSD: Single shot multibox detector[C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2016:21-37.
- [13] SEMICONDUCTOR Z. GP2000 GPS receiver hardware design [EB/OL]. (2003-12-07)[2020-07-20]. <http://www.zarlink.com/appnotes>.
- [14] TAN M, PANG R, LE Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE,2020:10781-10790.

作者简介:



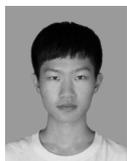
唐小煜(1980-),通信作者,男,博士,讲师,研究方向:人工智能与智能系统,E-mail: tangxy@snu.edu.cn。



熊浩良(1998-),男,本科,研究方向:计算机视觉、图像处理,E-mail: 229145399@qq.com。



黄锐珊(1998-),女,本科,研究方向:计算机视觉,E-mail:jeoisaaan@163.com。



林威霖(1998-),男,本科,研究方向:计算机视觉,E-mail:619382349@qq.com。

(编辑:陈珺)