

基于改进 RetinaNet 模型的接触网鸟巢检测

刘国文^{1,2,3}, 张彩霞^{1,2}, 李 斌^{1,2,3}, 杨 阳³, 张文生³

(1. 佛山科学技术学院, 佛山, 528000; 2. 广东省智慧城市基础设施健康监测与评估工程技术研究中心, 佛山, 528000; 3. 中国科学院自动化研究所, 北京, 100080)

摘 要: 鸟类活动故障已经成为高速铁路的主要隐患之一, 找出和清理接触网的鸟巢是一种应对手段。传统的鸟巢目标检测方法需要人工提取特征, 而手工设计的特征难以在复杂的接触网场景中保证泛化能力。针对该问题, 本文提出使用基于深度学习的目标检测算法识别接触网鸟巢, 并提出一种基于单阶段目标检测模型 RetinaNet 的改进模型, 增加 P2 特征层, 扩充网络的感受野范围, 以更好地检测出目标较小的鸟巢。最后使用高铁车载设备的数据集对基于深度学习的目标检测算法进行了训练和测试。实验结果表明: 基于深度学习的目标检测算法在接触网鸟巢检测任务上表现优秀, 且改进 RetinaNet 模型的 mAP 值达到了 90.4%, 优于原模型, 对于高速铁路的避障任务具有参考和应用价值。

关键词: 目标检测; 深度学习; 异常检测; 接触网

中图分类号: TP183

文献标志码: A

Detection of Birds' Nest in Catenary Based on Improved RetinaNet Model

LIU Guowen^{1,2,3}, ZHANG Caixia^{1,2}, LI Bin^{1,2,3}, YANG Yang³, ZHANG Wensheng³

(1. Foshan University, Foshan, 528000, China; 2. Guangdong Province Smart City Infrastructure Health Monitoring and Evaluation Engineering Technology Research Center, Foshan, 528000, China; 3. Institute of Automation, Chinese Academy of Sciences, Beijing, 100080, China)

Abstract: At present, bird activity failure has become one of the main hidden dangers of high-speed railway. Finding and cleaning the birds' nest of the catenary is a countermeasure. Traditional birds' nest object detection methods require manual extraction of features, but hand-designed features are difficult to ensure generalization in complex contact network scenarios. To solve this problem, this paper proposes to use the deep learning based object detection algorithm to identify the birds' nest on catenary. At the same time, an improved model based on the one-stage object detection model RetinaNet is proposed. The P2 feature layer is added to expand the receptive field range of the network, so that the smaller nest can be better detected. Finally, these deep learning based object detection algorithms are trained and tested using data sets collected by on-board equipment of high-speed railways. Experimental results show that the object detection algorithm based on deep learning is excellent in the catenary birds' nest detection task, and the improved RetinaNet model has a mAP value of 90.4%, which is better than the original model. This algorithm has certain both reference and application value for the obstacle avoidance task of high-speed

基金项目: 国家自然科学基金(61602484, 61702518)资助项目; 广东省智慧城市基础设施健康监测与评估工程技术研究中心和国家自然科学基金青年科学基金(61803087)资助项目。

收稿日期: 2019-10-22; **修订日期:** 2019-11-10

railway.

Key words: object detection; deep learning; anomaly detection; catenary

引言

随着国内高速铁路近年来的快速发展,以及因生态改善导致的鸟类数量的大幅增加,鸟类活动带来的接触网故障和安全隐患成为了不能忽视的铁路安全问题,由此产生了铁路上的鸟巢清理工作^[1-2]。最传统的识别接触网鸟巢方法通过人工分析拍摄的铁路车载设备视频,耗费大量的人力和时间。通过检测技术自动识别铁路接触网的鸟巢,具有重要的现实意义,对此国内进行了一些研究。段旺旺等^[3]对可能存在鸟巢的位置进行初步提取,并进一步通过方向梯度直方图(Histogram of oriented gradient, HOG)特征和支持向量机(Support vector machine, SVM)算法进行多次提取,算法精度不错,但自适应能力差,在高铁的复杂路途环境下泛化能力略有欠缺。金炜东等^[4]提出了使用双判别器生成对抗网络来解决鸟巢的数据不平衡问题,在判别上也取得了不错的效果,但对抗网络在训练时不容易把控两个网络的平衡和同步,训练过程可能不稳定,而且还存在崩溃模式现象等问题。祝振敏等^[5]提出一种基于鸟窝和硬横梁相对位置不变性的检测方法,将图像用Sobel算子水平边缘检测和概率霍夫变换以及Otsu二值化等方法处理后进行识别,该方法不需要很多训练样本就能实现自动识别,但需要先验信息,其阈值的设定直接影响最终识别效果。相比上述各算法所存在的不足,卷积神经网络在特征提取和归纳能力上表现优越,使用深度神经网络的目标检测算法也已被广泛地应用到图像视觉领域,相比传统的目标检测算法,其更加简洁高效。针对该现状,本文从深度学习的角度进一步地寻求解决高铁接触网鸟巢识别问题的更优方案,其主要贡献有:(1)引入基于深度神经网络的目标检测算法用于接触网的鸟巢检测,相比传统方法,无须人工提取特征,泛化性能更强;(2)将RetinaNet模型针对接触网的鸟巢检测问题进行了改进优化,检测效果有所提升。

1 相关工作

1.1 深度学习

深度学习^[6]是目前计算机科学的一个热门方向,作为机器学习的一个子类研究方向,受益于计算性能的提升和样本数据量的大增,在人工智能领域的很多方面都有所突破,如计算机视觉、自然语言处理等领域,表现较传统方法突出不少,目前于各行各业都有着广泛的应用。

深度学习是人工神经网络的进一步发展。传统的人工神经网络使用反向传播算法进行迭代,但随着神经网络层数的增加,这个迭代过程会存在梯度消失和梯度爆炸等问题,导致效果不佳。直到Hinton等^[7]首次提出深度网络和深度学习的概念,这个问题才慢慢得到解决,深度学习开始高速发展。He等^[8]提出的深度残差网络解决了传统人工神经网络的层数问题,达到了152层的深度。

1.2 基于候选区域的目标检测算法

基于深度神经网络的目标检测算法按实现过程可分为两类:基于候选区域进行分类的二阶段方法和基于回归进行分类的一阶段方法^[9]。Girshick等^[10]提出了R-CNN(Region convolution neural network)算法,也是深度学习第一次涉猎目标检测领域,此算法在使用传统SVM进行判别前,在图像上产生候选区域并使用卷积神经网络进行特征提取,效果提升明显。在其之后又发展出一系列的算法,如SPPNET(Spatial pyramid pooling convolution networks)等^[11]使用了能接受任意尺寸输入的SPPNet代替R-CNN的最后一层池化层,解决了R-CNN对于每个区域都要将图像缩放所带来的性能瓶颈,在提高准确率的同时大幅提升了R-CNN的速度。Grishick等^[12]提出的Fast R-CNN引用了SPPNET的工

作,并将R-CNN原来的串行结构改成并行结构,在对图像分类的同时进行回归,提升了速度和精度。Ren等^[13]提出的Faster R-CNN建立候选区域网络的概念,与Fast R-CNN相比,算法的生成方法大大提高了所生成候选区域的质量。He等^[14]将图像的mask信息加入到Faster R-CNN,算法精度得到了进一步的提升。

1.3 基于回归的目标检测算法

另一类基于回归的一阶段目标检测方法,由于直接对图像进行回归,相比前一类算法在速度上有较大的优势。YOLO(You only look once)算法^[15]对图像进行预处理后放入网络并通过全连接层输出,其结构简单且速度较快,但是最后的结果框有些粗糙。针对该问题,SSD(Single shot multibox detector)算法^[16]在YOLO的基础上增加了来源于Faster R-CNN的Anchor(锚)机制,同时融合不同卷积层的特征进行预测,相比YOLO算法来说提高了识别精度,尤其是在小目标的情况下,另外SSD算法没有全连接层,相比YOLO算法还有速度上的提升。YOLO9000^[17]在YOLO的基础上改进了网络,采用多尺度训练方法,并综合了ImageNet数据集和COCO数据集,采用联合训练的方式训练,可以识别9 000类物体。基于回归的目标检测方法基本上都存在一个类别不平衡问题,原因是算法中产生的bounding box检测框太多,而其中只有少数是前景,大多数都是背景,这会导致精度的下降,而RetinaNet^[18]通过Focal Loss损失函数解决了这一问题。

2 RetinaNet

RetinaNet主要由3部分构成:用于特征提取的深度残差网络(Residual network, ResNet)、特征金字塔网络(Feature pyramid networks, FPN)^[19]、分类和回归子网络。图像经过深度ResNet后特征被初步提取,特征在FPN中重组,最后被送入分类和回归子网络得出最后的检测结果。

2.1 ResNet(残差网络)

神经网络的层数越多,对特征的提取抽象程度就越高,可供调节的参数就越多,也就能产生更好的拟合效果。但随着网络层数的增加,会出现梯度爆炸和梯度消失等问题。针对这些问题,ResNet引入如图1所示的残差单元结构,由此构成残差网络。残差单元结构的表达式为

$$y_l = h(x_l) + F(x_l, W_l)$$
$$x_{l+1} = f(y_l)$$

(1)

式中: x_l 为第 l 个单元的输入, x_{l+1} 为第 $l+1$ 个单元的输入, y_l 为第 l 个单元的输出, F 为残差函数, $h(x_l)$ 为恒等映射, f 为ReLU激活函数。

从浅层 l 到深层 L 的学习特征为

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i)$$

(2)

根据反向传播原理,假设网络损失函数为loss,链式求导式(2)可得

$$\frac{\partial \text{loss}}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \cdot (1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i))$$

(3)

若

$$h(x_l) = \lambda_l x_l$$

(4)

则

$$x_{l+1} = \lambda_l x_l + F(x_l, W_l)$$

(5)

将式(5)代入式(3)可得

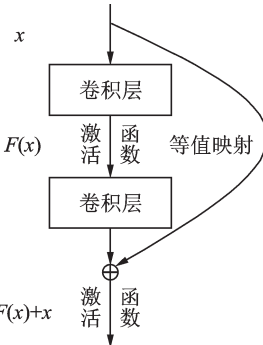


图1 ResNet残差单元结构

Fig.1 Residual unit structure of ResNet

$$x_L = \left(\prod_{i=l}^{L-1} \lambda_i \right) x_l + \sum_{i=l}^{L-1} F(x_i, W_i) \quad (6)$$

$$\frac{\partial \text{loss}}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \left(\left(\prod_{i=l}^{L-1} \lambda_i \right) + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i) \right) \quad (7)$$

由式(7)可知,当系数 $\lambda=1$ 时,网络不会产生梯度爆炸或消失的问题。

残差单元结构在卷积前馈网络中增加一个连接,由上述分析可知这种结构使得梯度不会消失,因此使深层次的神经网络训练成为可能。

2.2 FPN(特征金字塔网络)

在FPN被提出之前,基于深度神经网络的目标检测算法一般使用特征提取网络的顶层特征作为预测输入。特征提取网络顶层特征语义较为抽象,底层特征语义较为具体,FPN使用金字塔结构进行特征融合,且每层都是独立进行预测。

在RetinaNet模型中,FPN通过自下而上、自上而下的连接以及横向连接将ResNet不同层的特征图进行融合。自下而上和自上而下的路线分别生成C3,C4,C5和P3,P4,P5,P6,P7等特征图,P3到P5由C3到C5计算而得,P6和P7分别由P5和P6使用步长为2的 3×3 卷积核卷积而得。增加P6和P7是为了让模型更好地检测大物体,得益于更大的感受野。这样的操作可以保证每一层都有合适的分辨率以及强语义特征,配合目标检测算法,从而提升物体的检测性能。

2.3 分类和回归子网络

RetinaNet引入了Faster R-CNN里的Anchor思想:当卷积核滑动到特征图的某一个位置时,以当前窗口的中心点为中心,映射到原图的一个区域,以原图上这个区域的中心对应一个尺度和长宽比,成为一个Anchor。

分类和回归子网络对每个Anchor进行分类和回归,结构如图2所示,将FPN中某一层特征图输入到全卷积网络,最后通过Sigmoid函数得出物体的概率。对于框的回归也是类似的操作,区别仅在于最后不经过Sigmoid激活函数。

2.4 损失函数

神经网络中常见的二分类损失函数为交叉熵损失函数,即

$$CE(p, y) = \begin{cases} -\log p & y = 1 \\ -\log(1-p) & \text{其他} \end{cases} \quad (8)$$

RetinaNet针对样本不平衡的问题,提出了一种新的损失函数。为了平衡正负样本,RetinaNet最初提出的方案是在少数类别的样本上增大权重,即

$$CE(p_t) = -\alpha_t \log p_t \quad (9)$$

其中

$$p_t = \begin{cases} p & y = 1 \\ 1-p & \text{其他} \end{cases} \quad (10)$$

但是这并不能区分易分样本和难分样本,因此

$$FL(p_t) = -(1-p_t)^{\gamma} \log p_t \quad (11)$$

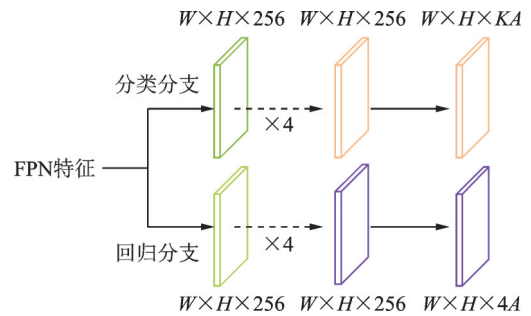


图2 分类和回归子网络结构

Fig.2 Classification and regression subnetwork structure

样本越容易被区分,则 p_i 越大,对于损失函数值的影响就越小,相应地提高了难样本的权重。在算法的实际应用中,由式(9, 11)可得 RetinaNet 的损失函数为

$$FL(p_i)=-\alpha_i(1-p_i)^{\gamma}\log p_i \quad (12)$$

2.5 RetinaNet模型的改进

为使 RetinaNet 模型更好地应用于接触网的鸟巢检测任务,本文对其进行了改进。原始的 RetinaNet 模型并没有使用到由 FPN 所得到的 P2 特征层,而根据对数据集的观察,与一般的目标检测任务略有不同,接触网的鸟巢大小占图像的比例较小,因此考虑在 RetinaNet 模型基础上增加 P2 特征层的使用,使得网络整体的感受野范围增加。增加原理与 RetinaNet 的 P6 和 P7 特征层的原理类似,最终使其对目标较小的鸟巢的识别检测能力增强,具体结构如图 3 和图 4 所示。

3 实验与结果

3.1 数据集

实验数据集取自高铁车载设备拍摄视频,选取其中有鸟巢的图像共 1 629 张,尺寸均为 1 620 像素×1 220 像素。根据实验需求将其标注为 PASCAL VOC 数据格式和 CSV 格式,其中 60% 作为训练集,40% 作为测试集。图 5 显示了数据集的部分图像。

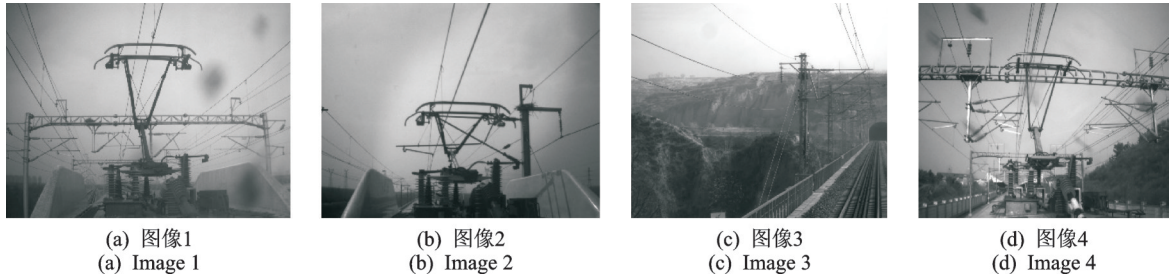


图 5 数据集的部分图像

Fig.5 Partial images of the dataset

3.2 运行环境设置及参数设置

运行环境为 Ubuntu16.04,CPU 为 Intel(R) Xeon(R) CPU E5-2640 v4,GPU 为 Nvidia GeForce GTX 1080 Ti,编程语言为 Python。利用 SSD,Faster R-CNN,原始 RetinaNet 模型及本文的改进模型对数据集进行训练,其中 SSD 模型有两种,分别为 SSD300 和 SSD512。SSD300 的输入图像尺寸为 300 像素×300 像素,SSD512 的输入图像尺寸为 512 像素×512 像素,它们都使用 VGG16 网络结构作为特征提取

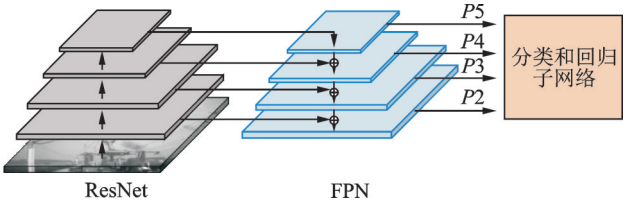


图 3 改进的 RetinaNet 模型结构

Fig.3 Improved RetinaNet model structure

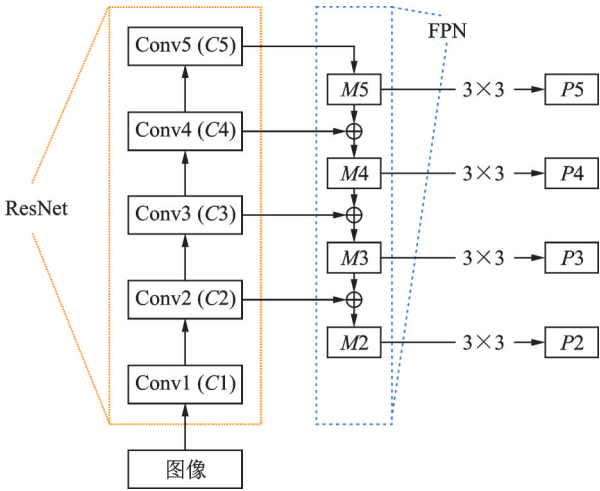


图 4 改进的 RetinaNet 模型结构细节

Fig.4 Improved RetinaNet model structure details

网络。RetinaNet模型采用ResNet50和ResNet101等网络作为特征提取网络的不同模型种类,其中有augment后缀的代表使用了数据增强方法,带P2字样的模型表示本文提出的改进模型。

训练参数设置如下:SSD300和SSD512的参数设置相同,冲量常数为0.9,初始学习率为0.002,权值衰减系数为0.000 5;Faster R-CNN的冲量常数为0.9,初始学习率为0.02,权值衰减系数为0.000 1;RetinaNet的梯度下降设置为随机梯度下降,冲量常数为0.9,初始学习率为0.02,权值衰减系数为0.000 1;改进RetinaNet模型、数据增强RetinaNet模型的参数设置与原RetinaNet模型的相同。

3.3 实验结果分析

使用训练好的SSD,Faster R-CNN,RetinaNet及改进的RetinaNet在数据集上进行测试,损失函数值见图6,结果如表1所示。

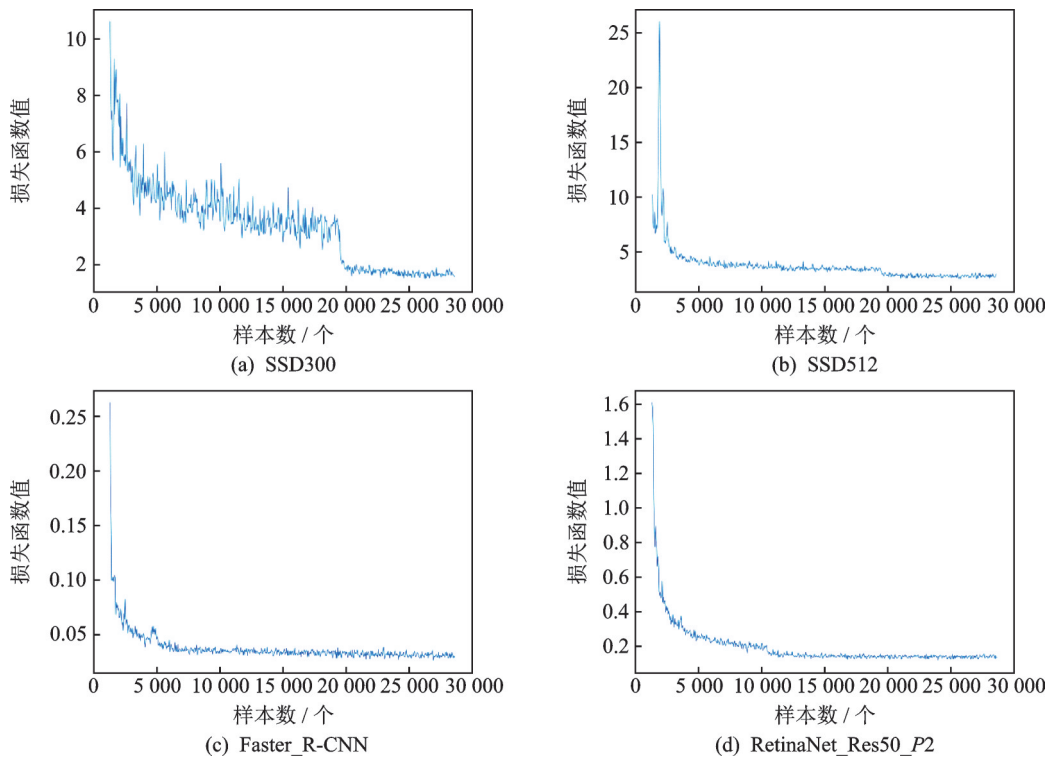


图6 训练的损失函数值

Fig.6 Loss function value during training

针对鸟巢样本,TP(True positive)为存在鸟巢且被检测到的鸟巢样本数量,FP(False positive)为不存在鸟巢且被检测到的鸟巢样本数量,FN(False negative)为存在鸟巢但没有检测出的鸟巢样本数量。召回率Recall和准确率Precision可分别表示为

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (14)$$

均值平均精度 mAP(Mean average precision)的计算公式为

$$\text{mAP} = \frac{\sum AP}{N(\text{Class})} \quad (15)$$

式中:AP(Average precision)定义为根据准确率和召回率绘制的曲线下方的面积大小。由于此数据集只有一个样本类别,因此样本类别 $N(\text{Class}) = 1$,因此由式(15)得

$$\text{mAP} = \text{AP} = \frac{\sum \text{Precision}}{N(\text{TotalImages})} \tag{16}$$

在统计学中,mAP是不同种类样本AP的平均值,比较能反应全局性能,因此本文以mAP作为主要评价指标。

由表1可知,RetinaNet模型在性能上总体优于其他算法模型;使用数据增强(图像翻转、图像剪切等)的RetinaNet模型在mAP上不但没有明显的提升,反而略有下降;使用层数更深的ResNet101作为特征提取网络同样在Map上没有提升,但召回率有所增加;而本文的改进RetinaNet模型将mAP值提高了近1%。

考虑到数据集的样本数量有限,为充分利用数据集的信息以及验证改进模型的有效性,设置了三折交叉验证实验,结果如表2所示。表2的结果为平均值,由结果可得,改进的RetinaNet模型在接触网鸟巢检测任务中各方面的效果都比原RetinaNet模型更优,其中mAP提升了1.2%。综合来看,本文方法是一种有效的接触网鸟巢目标检测方法。

通过对数据集结果的观察,绝大多数的鸟巢都能被检测出,如图7中的(a)—(d)的情况,在部分被水污遮挡的情况下也成功地检测出鸟巢,如图7(d)所示。同时也存在些许表现稍欠的情况,如图7(e)和(f)所示,鸟巢虽被检测出来但是有两个检测框,存

表1 基于深度学习的目标检测方法在数据集上的实验结果

Table 1 Experimental results of deep learning-based object detection methods on datasets

算法	召回率	准确率	mAP
SSD300	0.903	0.512	0.848
SSD512	0.901	0.807	0.872
Faster R-CNN	0.957	0.855	0.889
RetinaNet_Res50_augment	0.972	0.679	0.894
RetianNet_Res101	0.989	0.332	0.895
RetinaNet_Res50	0.985	0.303	0.896
RetinaNet_Res50_P2(Ours)	0.981	0.377	0.902

表2 改进 RetinaNet 模型和原始 RetinaNet 模型的三折交叉验证结果

Table 2 Tri-fold cross-validation results of the improved and the original RetinaNet models

算法	召回率	准确率	mAP
RetinaNet_Res50	0.989	0.545	0.892
RetinaNet_Res50_P2(Ours)	0.991	0.572	0.904

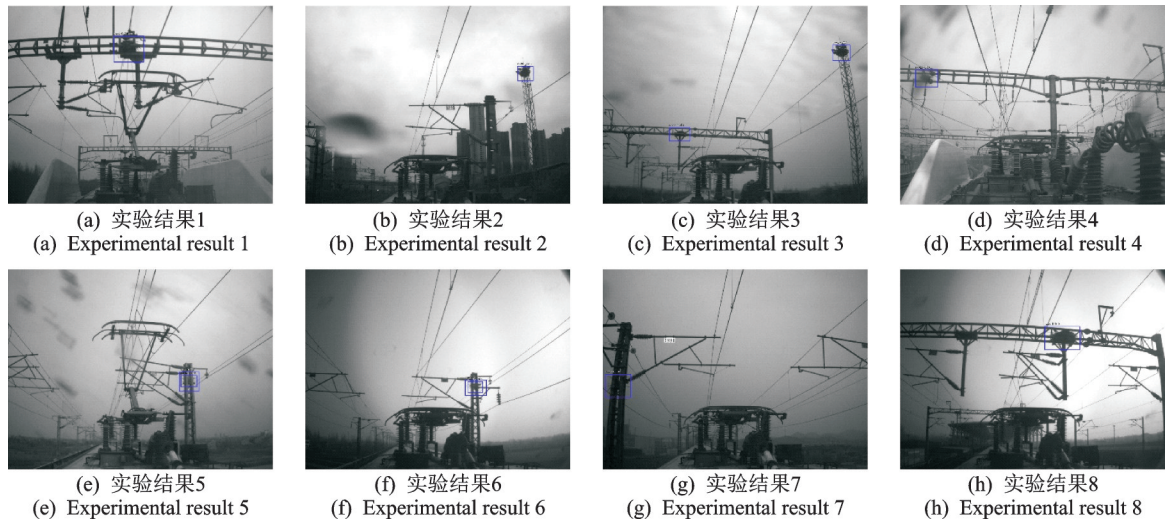


图7 改进的 RetinaNet 模型的部分实验结果

Fig.7 Partial experimental results of the improved RetinaNet model

在误检现象。另外还有少数漏检的情况,如图 7(g)和(h)所示,漏检的情况一般是因为鸟巢的筑巢进度较早,以至于鸟巢尚未完全成型或鸟巢过于小,难以被算法所检测。

相比其他算法,改进 RetinaNet 模型在检测未完全成型或体积过小的鸟巢的结果上仍存在优势。这点既可以从更高的召回率中看出,也可以从具体例子中得到验证,如图 8 所示。橙色框为真实数据(Ground truth),鸟巢体积偏小,Faster R-CNN 算法没有检测出鸟巢,SSD300(黄色框)和 SSD512(绿色框)算法虽然检测出了鸟巢,但其检测框的精确度和平均置信度不如 RetinaNet 模型(紫色框)。

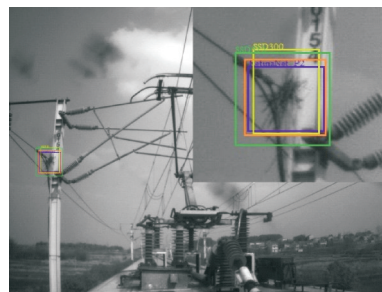


图 8 鸟巢偏小情况下的算法结果对比示例

Fig.8 Comparison example of algorithm results in the case of small nests

4 结束语

本文将基于深度神经网络的目标检测算法引入高速铁路接触网的鸟巢目标检测任务中,同时将 RetinaNet 模型针对鸟巢检测任务的特性进行了优化,并使用主流的基于深度神经网络的目标检测算法在高铁车载设备视频数据集上进行了训练和测试。实验结果表明,深度学习可以胜任高铁接触网鸟巢检测的任务,且本文提出的改进 RetinaNet 模型取得了最佳效果。

参考文献:

- [1] 张旭升. 输电线路运行安全影响因素分析及防治措施[J]. 中国战略新兴产业, 2017(48): 159.
ZHANG Xusheng. Analysis of factors affecting transmission line operation safety and prevention measures[J]. Tendering & Purchasing Management, 2017 (48): 159.
- [2] 黄美斐. 接触网鸟害防治措施研究[J]. 科技创新与应用, 2017(18): 234.
HUANG Meifei. Research on the prevention and control measures of bird damage in contact nets[J]. Technology Innovation and Application, 2017 (18): 234.
- [3] 段旺旺, 唐鹏, 金伟东, 等. 基于关键区域 HOG 特征的铁路接触网鸟巢检测[J]. 中国铁路, 2015(8): 73-77.
DUAN Wangwang, TANG Peng, JIN Weidong, et al. Bird nest detection of railway contact net based on HOG feature in key areas[J]. China Railway, 2015 (8): 73-77.
- [4] 金伟东, 杨沛, 唐鹏. 双判别器生成对抗网络及其在接触网鸟巢检测与半监督学习中的应用[J]. 中国科学: 信息科学, 2018, 48(7): 888-902.
JIN Weidong, YANG Pei, TANG Peng. Double discriminator generative adversarial networks and their application in detecting nests built in catenary and semisupervised learning[J]. Scientia Sinica(Informationis), 2018, 48(7): 888-902.
- [5] 祝振敏, 谢亮凯. 基于相对位置不变性的接触网鸟巢识别检测[J]. 铁道科学与工程学报, 2018, 15(4): 1043-1049.
ZHU Zhenmin, XIE Liangkai. Detection of birds' nest in catenary based on relative position invariance[J]. Journal of Railway Science and Engineering, 2018, 15(4): 1043-1049.
- [6] GUO Y, LIU Y, OERLEMANS A, et al. Deep learning for visual understanding: A review[J]. Neurocomputing, 2016, 187: 27-48.
- [7] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [8] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [9] ZHAO Zhongqiu, ZHENG Peng, XU Shoutao, et al. Object detection with deep learning: A review[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(11): 3212-3232.
- [10] 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. Columbus, OH, USA:

- IEEE, 2014: 580-587.
- [11] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, 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.
- [12] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1440-1448.
- [13] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [14] HE Kaiming, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2961-2969.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision And Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 779-788.
- [16] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single shotmultibox detector[C]//Proceedings of European Conference on Computer Vision. Las Vegas, NV, USA: Springer, 2016: 21-37.
- [17] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 7263-7271.
- [18] LIN Tsungyi, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2980-2988.
- [19] LIN Tsungyi, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision And Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 2117-2125.

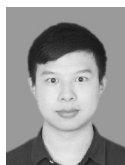
作者简介:



刘国文(1993-),男,硕士研究生,研究方向:机器视觉, E-mail: Keanu_L@outlook.com。



张彩霞(1976-),女,教授,研究方向:多元信息融合及智能控制系统。



李斌(1995-),男,硕士研究生,研究方向:机器视觉。



杨阳(1986-),男,副教授,研究方向:机器视觉及模式识别。



张文生(1965-),通信作者,男,教授,博士生导师,研究方向:人工智能、机器学习、大数据知识挖掘、概率图模型表示与推理、精密感知与智能控制、三维数字物理仿真、嵌入式视频图像处理, E-mail: zhangw-enshengia@ hotmail.com。

(编辑:王静)