

文章编号: 1005-0523(2019)06-0070-09

基于 DSSD 的接触网鸟窝识别检测研究

周俊, 陈剑云

(华东交通大学电气与自动化工程学院, 江西 南昌 330013)

摘要:随着中国电气化铁路里程的迅速增长,接触网的安全稳定运行面对巨大的压力,对其进行监测拥有重要意义。针对影响电气化铁路接触网正常工作的鸟害问题进行研究,通过对不同深度学习模型进行分析比较,选择 DSSD 模型以自动识别高速铁路接触网沿线。同时使用迁移学习的方法,运用 Caffe 平台,在数据集不足的情况下,通过微调训练好的 DSSD 网络,提高了鸟窝识别训练网络的泛化性和稳定性。训练完的模型具有更快的识别速度和更好的鲁棒性,对于接触网安全稳定运行拥有重要的参考意义。

关键词:DSSD 网络;鸟窝识别;迁移学习;深度学习

中图分类号:U225 **文献标志码:**A

DOI:10.16749/j.cnki.jecjtu.2019.06.011

随着高速铁路电气化运营里程的快速增长,接触网的安全稳定运行日益受到重视,对接触网相关检测的研究层出不穷。鸟类活动引起的接触网线路短路,电气控制部件损坏、跳闸、机械补偿装置卡滞失灵等问题,严重影响了铁路运输安全。鸟害成为仅次于雷击、外力破坏^[1]以外,又一对接触网的安全运行造成的严重威胁的因素。典型的鸟害包括鸟类筑巢、鸟类飞行和鸟粪闪络,其中鸟窝材料(树枝、铁丝等)对接触网形成短路,导致跳闸断电,严重时还可能导致接触网断线和设备损毁^[2],对接触网的危害最大。

目前,对鸟窝的监测主要是通过人工分析视频图像并判断和标记,这种人工的监测方式耗时耗力,且无法及时准确的发现并排除鸟窝,因此对鸟窝进行自动识别检测技术是必要的。近年来,基于图像处理的接触网状态监测研究已经在国内开展,通过拍摄接触网沿线的车载视频,应用图片处理技术自动识别视频中到故障信息。文献[3]根据鸟巢通常出现在电线杆塔横担区域这一特点,利用鸟窝担颜色特征和接触网塔架结构特征建立分类模型,能够取得较好检测精度和抗干扰性。文献[4]分析了接触网设备上的鸟巢图片,发现几乎所有的鸟巢通常搭建在接触网设备的几个固定位置,对疑似鸟窝对关键区域初步提取 HOG 特征,在通过 SVM 分类器进行特征对精确提取,能取得较好的检测效果。文献[5]提出了引入了支持向量机对接触网故障识别方法,并从鸟窝对颜色,形状,相对空间位置等方面分析鸟巢等特征并建模。文献[6]结合鸟巢筑巢特点,提出了一种基于鸟窝和硬横梁相对位置不变性等检测方法。虽然上述方法能够识别图像上的鸟窝,但是都是根据一定的先验知识建立某种模型来完成的。其无法满足大样本等特征提取训练,识别种类与精度受到限制。深度学习技术目前越来越多地应用于图像识别领域,如:车牌检测,人脸识别,手写体识别等,可以有效解决图像识别在种类,准确度和速度上的难点,显示出自动特征提取和分装等优点^[7]。

本文重点分析几种常用的深度学习识别算法,得出 DSSD 在对鸟窝的识别精度更佳的结论,识别精度相对较高,且识别速度快,实时性强。同时结合迁移学习技术,在数据集不足的情况下,可以进一步提高模型的识别精度,缩短模型的训练时间。

收稿日期:2019-04-07

作者简介:周俊(1992—),男,硕士研究生,主要研究方向为铁路微服务技术。

通讯作者:陈剑云(1962—),男,教授,博士,博士生导师,主要研究方向为电力系统监控自动化运动技术。

1 鸟窝识别模型的分析与选择

1.1 鸟窝图像特点分析

鸟窝数据集来自于南昌、沈阳、太原铁路局车载设备拍摄的接触网支柱图片,图片的分辨率为 5 184*3 435 像素。如图 1 所示,鸟窝图像的背景十分复杂,其对于整个场景而言只有很小的一部分,在自然图像采集过程中很容易受到天气变化,局部障碍物遮挡等自然环境或人为因素的干扰,从而导致检测目标产生形状变化,使检测难度大幅增加。



图 1 鸟窝图像

Fig.1 Bird's nest images

1.2 鸟窝识别模型分析与选择

传统的图像识别算法受限于自身算法的局限性,无法找到通用的特征提取模型,只能针对个别特征进行识别模型的构建,泛化性不强。近年来,相关研究人员在 CNN(convolutional neural network)的基础上不断改进,刷新了图像分类识别的精度与识别速度。Krizhevsky 等^[7]提出 AlexNet 卷积神经网络架构的图像对识别精度具有里程碑式的贡献。CNN 具有非线性特征表达能力,泛化性能好等优势,能够很好地完成物体等识别检测任务,并且 CNN 网络可以自主提取特征图像特征进行学习,池化等特性使之可以有效地降低神经网络复杂度,减少训练次数,避免了复杂的人工特征选择和提取过程,提高检测效率。

基于 CNN 的目标识别算法主要分类两大类:基于候选区域的深度学习目标检测算法和端对端学习的深度学习目标检测算法。基于候选区域的深度学习目标检测算法主要有:R-CNN(region-convolutional neural network)^[8],Fast R-CNN(fast region-convolutional neural network)^[9],Faster R-CNN(faster region-convolutional neural network)^[10],R-FCN(region-based fully convolutional network)^[11]。其中,R-CNN 使用 Selective Search 候选区域的思路,通过卷积神经网络来分类目标候选框,进行目标的特征提取,再通过 SVM 的特征向量进行分类,算法的识别准确度较高,但检测速度较慢。Fast R-CNN 在 R-CNN 的基础上,采用了自适应

尺度池化对整个网络进行了优化,避免了 R-CNN 过于冗余对特征提取,并使用感兴趣区域池化层进行特征提取,同时,使用 SoftMax 非线性分类器进行分类回归,算法的精度和速度得到了进一步提升。Faster R-CNN 构建区域建议网络提取候选框,取代了 Selective Search 候选区域搜索的方式,进一步提升了速度。

端到端的目标检测算法主要有 YOLO(you only look once)^[12],SSD^[13](single shot multibox detection)。其中, YOLO 优势在于检测速度快,但对识别目标的定位误差较大。SSD(single shot detector)深度学习物体检测算法,识别速度快,且识别精度高,但是 SSD 对小目标的检测性能却不好。DSSD(deconvolutional single shot detector)模型^[14]是 SSD 模型的继续发展,改进了对小目标物体的检测精度。在拍摄到的鸟窝图片中,鸟窝只占很小的一部分,加大了检测识别的难度。在比较了各个深度学习的目标检测算法后,本文选用 DSSD 构建鸟窝的目标识别模型。

2 鸟窝识别模型构建

2.1 DSSD 网络结构

DSSD 的网络结构^[14]如图 2 所示,相比于 SSD,最大的改变是把 Vgg-16 网络结构换成了 Resnet-101,利用中间层的上下文信息,将卷积层的特征图信息与反卷积层的特征图信息融合一起做预测。由于越往后的反卷积层,分辨率越高,且包含的上下文信息越丰富,综合在一起使得对小目标的检测精度得以提升。

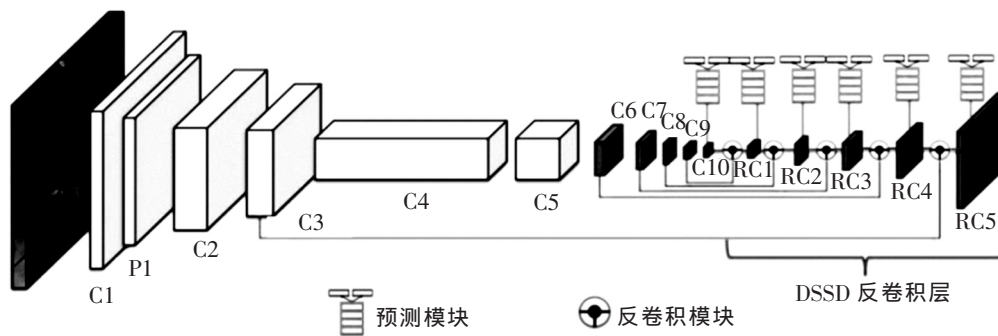


图 2 DSSD 网络模型
Fig.2 DSSD network model

2.2 模型构建过程

DSSD 采用了特征金字塔结构进行检测,如图 1 所示,从卷积层的 C3,C6,C7,C8,C9,C10 和反卷积层的 RC1,RC2,RC3,RC4,RC5 提取出作为检测所用的 feature maps 且 C3 与 RC1,C6 与 RC2,C7 与 RC3,C8 与 RC4,C9 与 RC5 模块融合一起作检测。对于每一个 feature maps,设置先验框,例如第 k 个 feature maps,先验框比例公式^[13]如下

$$S_k = S_{\min} + \frac{S_{\max} - S_{\min}}{m-1} (k-1), k \in \{1, 2, \dots, m\} \quad (1)$$

其中: m 为特征图的个数; S_{\min}, S_{\max} 分别为比例的最小值和最大值, S_{\min} 取 0.2, S_{\max} 取 0.95,分别代表特征图上先验框占输入图像的比例。

在每一层 feature maps 上,SSD 加入了 Faster R-CNN 的 anchors 机制,对每一特征图上的先验框采取不同宽高比,以增强先验框对物体形状的鲁棒性,这里采用的宽高比 $r = \{1, 2, 1/2, 3, 1/3\}$ 。每一个先验框的中心^[13]为

$$\left(\frac{i+0.5}{|f_k|}, \frac{j+0.5}{|f_k|} \right), i, j \in [0, |f_k|) \quad (2)$$

其中 $|f_k|$ 为特征图的大小。

得到特征图后,需要对特征图进行卷积得到检测结果,这些大小不同的 feature maps 同时进行 softmax 分类和位置回归。检测框架的总体目标损失函数由中心位置损失^[13](L_{loc})和置信损失^[13](L_{conf})的加权和表示。公式如下,

$$L(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \quad (3)$$

式中: N 为验框的正样本数量,这里 $x_{ij}^p \in \{1, 0\}$ 为一个指示参数,当 $x_{ij}^p=1$ 时表示第 i 个先验框与第 j 个 ground truth 匹配; c 为类别置信度预测值; l 为先验框的所对应边界框的位置预测值; g 为 ground truth 的位置参数; p 为 ground truth 的类别。IOU 计算的是“预测的边框”和“真实的边框”的交集和并集的比值。

$$L_{loc}(x, l, g) = \sum_{i \in Pos}^N \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k smooth_{L1}(|I_i^m - \hat{g}_j^m|) \quad (4)$$

$$\hat{g}_j^{cx} = (g_j^{cx} - d_i^{cx}) / d_i^w \quad (5)$$

$$\hat{g}_j^{cy} = (g_j^{cy} - d_i^{cy}) / d_i^h \quad (6)$$

$$\hat{g}_j^w = \log\left(\frac{g_j^w}{d_i^w}\right) \quad (7)$$

$$\hat{g}_j^h = \log\left(\frac{g_j^h}{d_i^h}\right) \quad (8)$$

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & otherwise \end{cases} \quad (9)$$

式中: cx, cy, h, w 分别为预测框的中心点和长宽的值。由于 x_{ij}^p 的存在,所以位置误差仅针对正样本进行计算。置信度误差采用 softmax loss^[13]

$$L_{conf}(x, c) = - \sum_{i \in Pos}^N x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in Neg} \log(c_i^0) \quad (10)$$

$$\hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \quad (11)$$

权重系数 α 通过交叉验证设置为 1。

对于每一个预测框,根据类别置信度确定图片中是否有鸟窝,并过滤掉属于背景的预测框。根据置信度阈值,例如取 0.5。过滤掉阈值较低的预测框。最后利用 NMS(non-maximum suppression)算法,过滤掉那些重叠较大的预测框,最后剩余的预测框就是检测结果。

3 基于迁移学习的鸟窝识别模型构建

3.1 迁移学习

迁移学习基本思想是利用学习目标和已有知识之间的相关性,把知识从已有的模型和数据中迁移到目标上去。以减少对标记数据的依赖,通过和已有数据和模型之间的迁移,更好地完成任务。

迁移学习最开始属于机器学习的一个分支,通过使用获取的数据、任务或模型之间的相似性,把之前学习实现的模型使用在当前需要的模型上。其中关于深度神经网络的迁移学习是近年来的研究热点^[15-16]。

训练一个深度学习网络需要有足够多的标签数据,实际的鸟窝识别模型训练中,如果在训练数据集的数量不足的情况下进行模型训练,会直接影响训练模型的识别精度。而采用迁移学习的方法可以解决样本数据不足的问题。本文将迁移学习应用在 DSSD 模型训练,实现对鸟窝自动检测识别。

3.2 DSSD 网络模型迁移

本文的识别模型是基于 Resnet-101 网络模型,主要由 Resnet-101 网络,卷积层网络,反卷积层网络组

成, 该网络模型被广泛的用作特征提取, 对图像的基础特征有很好的提取效果, 且该网络在 PASCAL-VOC2012 数据集中实现了对 20 类不同对象的分类, 本文在该网络的基础上用于鸟窝识别。因为原 DSSD 模型的训练分类任务的类别有 20 种, 而目标的任务只有两类, 即图像中是否有鸟窝, 因此需要对该网络的输出层做出相应的修改, 即将原 DSSD 模型的全连接层的输出结果改成本次模型识别的种类。

具体模型的迁移如图 3 所示, 将 PASCAL-VOC2012 数据集用作训练 DSSD 网络结构, 通常对于迁移的权重训练主要有微调和冻结两种方式, 本文选择冻结的方式进行模型的迁移, 其通过反向传播更新所有层的权重。加入鸟窝数据后, 继续对 DSSD 模型进行训练, 将之前的网络权重用作当前的模型训练, 不用重新对网络结构参数重新训练, 每次训练, 只对原网络结构的参数进行微调。通过后续的训练进行学习, 这样可以大大减少鸟窝识别模型训练等开销。

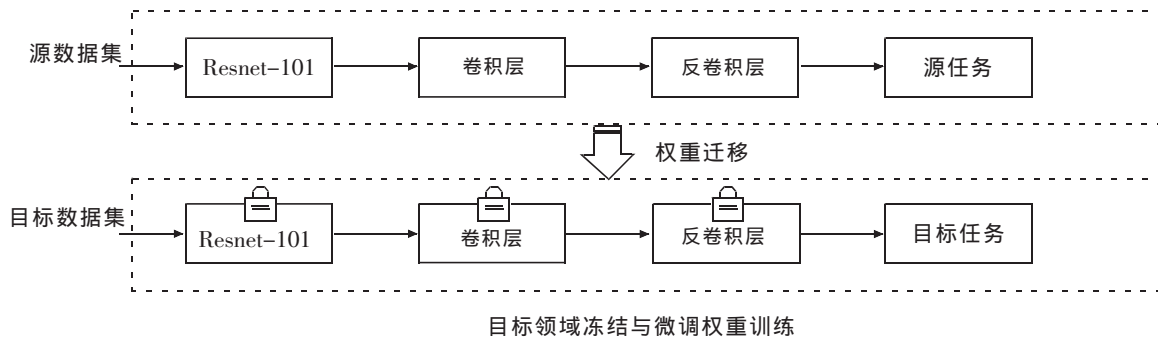


图 3 DSSD 网络模型迁移

Fig.3 DSSD network model migration

4 实验准备与模型训练

4.1 实验平台搭建

本文使用 Caffe 平台进行检测, 并采用 GPU 加速, 具体实验硬件配置如表 1 所示。

表 1 模型训练环境配置

Tab.1 Model training environment configuration

Name	Type
System	Linux64 Ubuntu14.04
Language	Python3.5.2
CPU	Intel i7-7700, 内存 16 GB
GPU	NVIDIA GeForce GTX1060, 内存 6 GB
Memory	8 GB
RAM	16 GB
Hard disk	256 GB
Frame	Caffe

4.2 鸟窝图像数据整理

在原始数据集的基础上经过预处理缩放为 512*512 分辨率的大小, 以减少后期特征提取时对硬件的处理的压力。配置好开发环境后, 需要将鸟窝图片数据整理为 Caffe 制定的数据格式——TFRecord。

本次采集到的图片总数为 1 000 张, 其处理步骤如下

第一步,对带有检测目标的图片进行数据扩增研究表明,在数据足够多前提下,模型训练结果通常更佳,同时,也能防止训练时出现过拟合的情况出现。本文主要采用的数据扩增技术有水平翻转、随机裁剪、随机采集块域。扩增后,图像的总数量为 2 000 张。

第二步,数据标记。对带有检测目标的图片进行手动标记。标记后图片被生成与之对应的后缀名为.xml 的文件,该文件主要用于记录图片中鸟窝的实际位置。

第三步,数据分类。将图片数据集和对应的.xml 文件分别保存到 train 和 test 文件夹中。train 作为训练集,test 作为测试集。训练集的数量为 1 600 张图片,测试集的数量为 400 张图片。

4.3 鸟窝识别模型训练

为了验证 DSSD 模型与其它深度学习模型的检测精度,检测速度的优势,分别将上一节的数据集,用于训练 Faster R-CNN,SSD,YOLO,DSSD 模型。而基于迁移学习的目标检测模型训练主要分为三个步骤:预训练模型;训练模型;模型测试。具体的训练流程如图 4 所示。为了验证迁移学习是否产生了作用,选用相同的数据集,学习率,训练迭代次数,训练出预训练模型,然后使用上一节整理出的鸟窝数据集进行鸟窝识别模型的训练。由于篇幅原因,本节重点介绍引入率迁移学习的 DSSD 模型的训练过程。

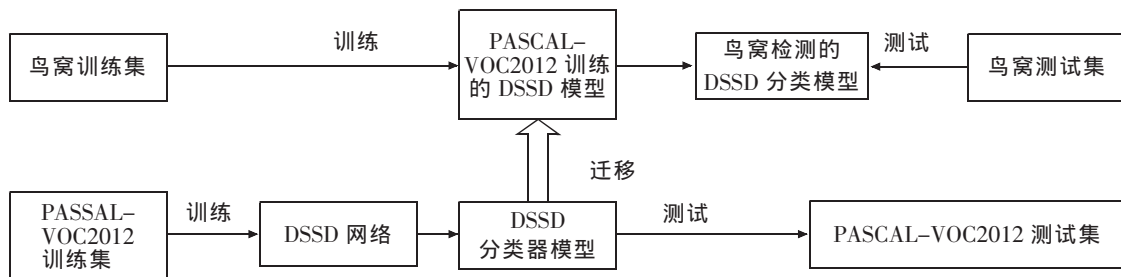


图 4 模型训练流程

Fig.4 Model training process

预模型训练:DSSD 的网络结构主要由两部分组成,基础网络层,卷积层和反卷积层。其中基础网络层是 Resnet-101。将 DSSD 网络用作为迁移学习的预训练模型训练模型,并选用 PASCAL-VOC2012 作为预训练模型的数据集。图片大小为 512*512,学习率为 10^{-3} ,迭代次数为 10 000 次,耗时 36 h。

模型训练:预模型训练结束后,加入鸟窝训练数据集进行 DSSD 模型的训练。预训练阶段产生的分类模型将根据迁移学习的思想,转变为预训练模型的初始化网络权值。本次实验模型训练的学习率为 10^{-3} ,迭代次数为 10 000 次,最终得到最优的 DSSD 模型。耗时 15 h。

测试阶段:将鸟窝数据测试集对新产生的分类模型进行测试。

5 实验结果对比与分析

5.1 无迁移学习的实验结果对比

在目标分类中,通常最关心的是存在的目标识别效果准确率与识别速度。 mAP 和 FPS 是目标检测算法中衡量算法精确度重要指标。 mAP 用于衡量模型的目标识别的精度,公式如下

$$mAP = \frac{\sum Average\ precision}{N_i} \quad (12)$$

式中: N 为检测次数; $Average\ precision$ 为检测目标的平均准确率。

mAP 值介于 0 到 1 之间,值越大说明算法的检测精度越好。 FPS 衡量模型对目标的识别速率,即一秒内识别图像数量。输入测试数据集,对比当前数据集,学习率,迭代次数相同的前提下,Faster R-CNN,SSD,YOLOv2,DSSD 的 mPA 和 FPS ,结果如表 2 所示。

表 2 模型训练结果对比
Tab.2 Comparison of training results of model

算法模型	$mAP/\%$	FPS frame/s
Faster R-CNN	86.15	9.91
SSD	88.17	15.09
YOLO	80.16	21.52
DSSD	90.12	12.90

可以看出,以FPS的大小评估算法的检测速度可知,Faster R-CNN 的识别检测速度较差,其它三个模型的识别速度相差不大,且都能满足实时性检测的要求。以 mPA 值评估算法的精确度,DSSD 对鸟窝的检测精度最高。本次实验使用测试集进行测试,如图 5 所示,模型的识别率可以到达近 80%,对于有遮挡物的情况下,该模型也能很好的识别带有鸟窝的图像;因此 DSSD 模型可以满足对接触网鸟窝的识别,也能满足实时性的检测要求。

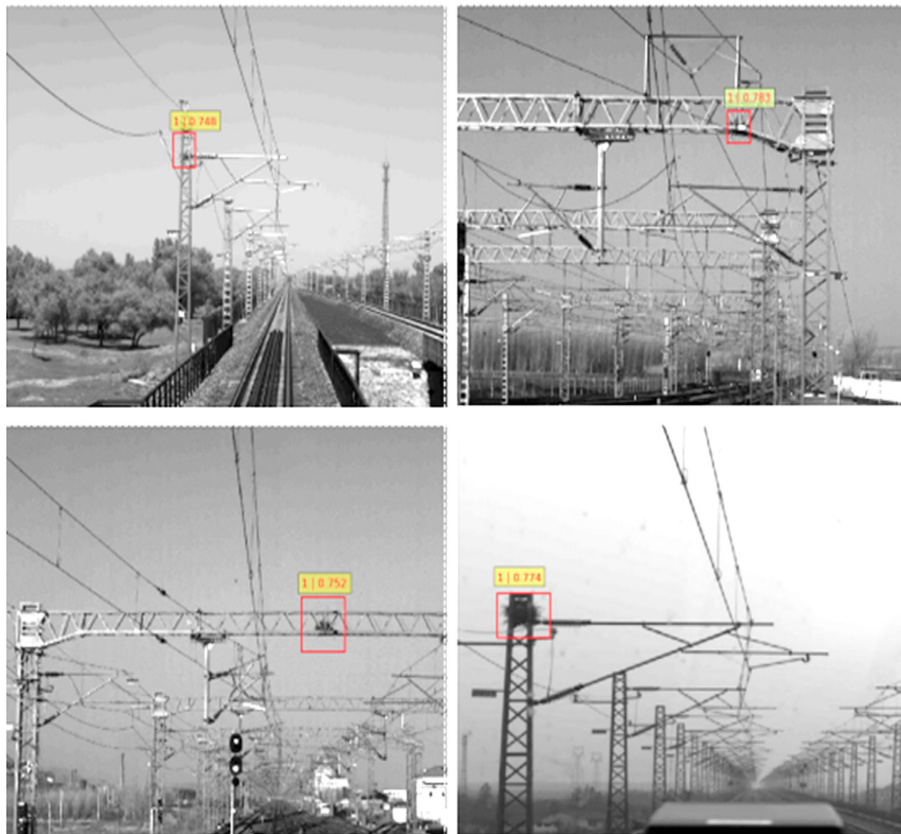


图 5 DSSD 模型识别结果
Fig.5 DSSD model recognition results

5.2 基于迁移学习的实验结果对比

对比其它深度学习模型,DSSD 对铁路接触沿线对鸟窝的识别准确率和速度效果都很好。将迁移学习应用在各模型的训练,在数据集不够充分的前提下,每个模型都识别精度都普遍提高了 1 到 2 个百分点,结果如表 3 所示。

表 3 迁移学习模型结果对比
Tab.3 Result comparison of transfer learning model

算法模型	<i>mAP</i> /%	<i>FPS</i> frame/s
Faster R-CNN	88.13	10.47
SSD	90.12	16.14
YOLO	83.75	22.56
DSSD	92.35	13.87

实验在测试集上进行测试,模型的识别效果如图 6 所示。配合迁移学习,*mPA* 得到提高,模型的识别准确性得到提高,可以达到近 85%,模型的训练时间缩短了近一半,大大地减少了鸟窝识别模型训练的开销。从 *FPS* 来看,模型的识别速度几乎没有受到模型迁移带来的影响,能够满足实时性检测要求。说明将迁移学习应用在 DSSD 鸟窝识别模型是可行的,能够在数据集不充分的情况下加快模型的训练,并提高模型对鸟窝识别的准确率。

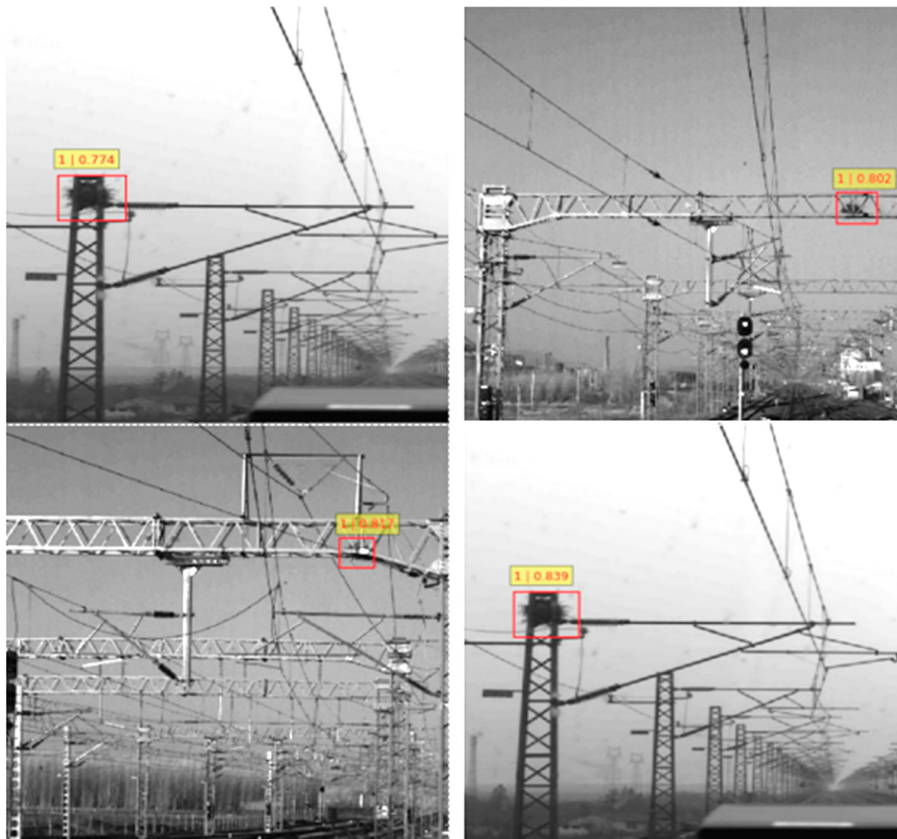


图 6 迁移学习识别效果

Fig.6 Recognition effect of migration learning

6 结论

该文基于 Caffe 平台使用 DSSD 深度学习模型,对高速铁路接触网沿线的鸟窝进行自动识别的研究。

1) 对比其它深度学习模型,DSSD 的识别速度较快,准确率高,且对于复杂背景下,有部分障碍物遮挡的鸟窝也能识别出来,模型的鲁棒性较好。

2) 使用迁移学习技术,在有限的鸟窝数据条件下训练模型,进一步提高了模型的识别精度,且模型的训练时间大大缩短。

3) 对于鸟害故障的处理,相对与传统的处理方法,该算法能够带来较高的识别精度和识别速度,且模型的鲁棒性较好,对于鸟害处理的应用有重要的参考意义。同时,也为牵引接触网可靠性分析提供了十分重要的参考意义。

参考文献:

- [1] 赵红玉,魏博. 接触网鸟害防治方案研究[J]. 铁道标准设计,2012(2):98-101.
- [2] 李文豪,刘航. 接触网鸟害防治措施研究[J]. 电气化铁道,2013(5):15-17.
- [3] 黄一鸣,袁天辰,杨俭. 基于支持向量机的接触网故障识别方法研究[J]. 计算机仿真,2018,35(11):145-152.
- [4] 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.
- [5] 葛雯,宫婷,王媛,等. 基于深度学习的目标检测算法[J]. 微处理机,2019,40(3):29-33.
- [6] 张义莲,陈咏,万迪明,等. 基于粗-精搜索策略的电线塔鸟巢检测[J]. 信息技术,2017(3):104-109.
- [7] KRIZHEVSKY A,SUTSKEVER L,HINTON G E. Image net classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems. Curran Associates Inc,2012:1097-1105.
- [8] TIANYU T,SHILIN Z,ZHIPENG D,et al. Vehicle detection in aerial images based on region convolutional neural networks and hard negative example mining[J]. Sensors,2017,17(2):336-339.
- [9] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision.2015:1440-1448.
- [10] REN S,HE K,GIRSHICK R,et al. Faster Rcnm: Towards Realtime Object Detection with Region Proposal Networks[J]. Ieee Trans Pattern Anal Mach Intell,2017,39(6):1137-1149.
- [11] LI Y,HE K,SUN J,et al. R-FCN: Object Detection via Region-based Fully Convolutional Networks[C]//Proc of Conference on Advances in Neural Information Processing Systems,2016:379-387.
- [12] REDMON J,DIVVALA S,GIRSHICK R,et al. You Only Look Once: Unified, Real-Time Object Detection[J/OL]. (2015-06-01)[2019-04-01]. https://www.researchgate.net/publication/278049038_You_Only_Look_Once_Unified_Real_Time_Object_Detection.
- [13] LIU W,ANGUELOV D,ERHAN D,et al. SSD: Single Shot MultiBox Detector[C]//European Conference on Computer Vision. Springer. Cham,2016:21-37.
- [14] FU C Y,LIU W,RANGA A,et al. DSSD: Deconvolutional Single Shot Detector[J/OL]. (2017-02-01)[2019-04-01]. https://www.researchgate.net/publication/312759848_DSSD_Deconvolutional_Single_Shot_Detector.
- [15] 刘鑫鹏,栾悉道,谢毓湘,等. 迁移学习研究和算法综述[J]. 长沙大学学报,2018,32(5):28-31.
- [16] 刘颖,张师,范九伦,等. 基于迁移学习及特征融合的轮胎花纹图像分类[J]. 计算机工程与设计,2019,4(5):1401-1406.

Research on Recognition and Detection of Catenary Bird's Nest Based on DSSD

Zhou Jun, Chen Jianyun

(School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: With the rapid growth of China's electrified railway mileage, the safe and stable operation of catenary is facing tremendous pressure, so it is of great significance to monitor. In this paper, the avian hazards affecting the normal operation of the catenary of electrified railway were studied. Through the analysis and comparison of different depth learning models, DSSD model was selected to automatically identify the catenary of high-speed railway. At the same time, this paper used transfer learning method and Caffe platform to improve the generalization and stability of bird's nest recognition training network by fine-tuning the trained DSSD network under the condition of insufficient data set. The trained model has faster recognition speed and better robustness, which has important reference significance for the safe and stable operation of OCS.

Key words: DSSD network; bird nest recognition; transfer learning; deep learning