

文章编号:1005-0523(2021)04-0072-08



# 基于改进的YOLOv3接触网鸟巢检测与识别

屈志坚,高天姿,池瑞,杨行

(华东交通大学电气与自动化工程学院,江西 南昌 330013)

**摘要:**接触网上鸟巢对铁路安全稳定运行存在严重威胁。为检测和识别接触网上鸟巢以解决鸟巢对铁路运行造成的不良影响,提出一种改进的YOLOv3算法。首先对接触网鸟巢图像进行前期预处理,去噪等操作能够加强对鸟巢本质特征的提取,数据增强一定程度上避免神经网络的过拟合现象产生。然后在网络结构中加入空间金字塔池化模块,对特征图进行不同尺度的池化操作再拼接,得到固定尺寸的输出,提取鸟巢多尺度特征。最后将衡量预测框与真实框距离的GIoU作为边界框损失函数,模型优化损失同时优化真实框与预测框的重叠度。实验结果表明,该方法在接触网鸟巢检测的平均准确率达到95.1%,在接触网鸟巢检测领域有较高的检测精度,能在复杂的接触网背景下较好的识别检测鸟巢。

**关键词:**接触网;深度学习;鸟巢检测;图像处理

**中图分类号:**TP317.4

**文献标志码:**A

**本文引用格式:**屈志坚,高天姿,池瑞,等.基于改进的YOLOv3算法的接触网鸟巢检测与识别[J].华东交通大学学报,2021,38(4):72-79.

DOI:10.16749/j.cnki.jecjtu.2021.04.009

## Detection and Recognition of Bird Nests in Overhead Catenary Systems Based on Improved YOLOv3

Qu Zhijian, Gao Tianzi, Chi Rui, Yang Hang

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

**Abstract:** Access to the bird nest is a serious threat to the safe and stable operation of overhead catenary system. An improved YOLOv3 algorithm was proposed to detect and identify the adverse effects of bird nests on railway operation. Firstly, the bird nest image of overhead catenary system is preprocessed in the early stage. The denoising and other operations can enhance the extraction of the essential features of the bird nest, and the data enhancement can avoid the over-fitting phenomenon of the neural network to a certain extent. The spatial pyramid pooling module is added into the network structure, and the feature map is pooled at different scales and then stitched together to get the output of fixed size, and the multi-scale features of the nest are extracted. Finally, the GIoU, which measures the distance between the prediction box and the real box, is used as the bounding box loss function, and the model optimizes the loss and the overlap degree between the real box and the pre-

收稿日期:2021-05-11

基金项目:江西省自然科学基金(20202BAB204023);国家自然科学基金项目(51867009);江西省重点研发计划项目(20192BBEL50006)

作者简介:屈志坚(1978—),教授,博士,研究方向为智能监控理论与信息处理技术。E-mail:quzhijian781231@163.com。

通信作者:高天姿(1997—),女,硕士研究生,研究方向为高铁供电图像检测。E-mail:2019028085222003@ecjtu.edu.cn

diction box. The experimental results show that the average accuracy of this method in the overhead contact net nest detection reaches 95.1%, which has a high detection accuracy in the overhead contact net nest detection field, and can better identify and detect the nest under the complex overhead contact net background.

**Key words:** overhead catenary systems; deep learning; the bird nest detection; image processing

**Citation format:** QU Z J, GAO T Z, CHI R, et al. Detection and recognition of bird nests in overhead catenary systems based on improved YOLOv3[J]. Journal of East China Jiaotong University, 2021, 38(4): 72-79.

接触网是电气化铁路的动力输电网络,接触网稳定的工作状况与铁路安全稳定运行密切相关<sup>[1-2]</sup>。鸟类在接触网上筑巢已不罕见,若及时发现并处理鸟巢,容易引起输电线路跳闸或绝缘子击穿等,对铁路运输安全稳定运行造成严重威胁<sup>[3]</sup>。传统检测识别鸟巢的方式是人工检测法<sup>[4]</sup>,即采用人工对车载视频进行标记,该方法费时费力且效率较低。

近几年,随着计算机图像技术的发展,利用图像处理技术对拍摄的车载视频进行自动检测,能提高接触网故障检测的效率。祝振敏等<sup>[5]</sup>针对鸟巢的结构特性,对接触网上疑似鸟巢区域进行粗提取,得到鸟巢可能存在的关键区域,再利用HOG特征和分类器训练对鸟巢进行精确提取。金炜东等<sup>[6]</sup>分析鸟类筑巢特点,针对杆塔的中上部、图片中最前景的硬横梁结构,提出一种基于相对位置不变性的接触网鸟巢识别检测。罗婷婷等<sup>[7]</sup>通过分析发现接触网图像中正常数据与异常数据样本数量存在严重的不平衡,针对鸟巢的检测问题,提出利用双判别器生成对抗网络来解决鸟巢样本数据不平衡问题。上述方法均能够检测到接触网上的鸟巢,但存在如对于遮挡问题性能不太稳定;或在鸟巢检测前需要设定阈值;阈值的选择影响鸟巢的检测结果等问题。近年来,以卷积神经网络为基础的深度学习目标检测算法主要分为一阶段检测和二阶段检测<sup>[8-9]</sup>。FasterR-CNN<sup>[10]</sup>等是基于目标候选框思想的二阶段检测,SSD<sup>[11]</sup>,YOLO<sup>[12]</sup>等是基于回归思想的一阶段检测,具有更高的检测速度。

在接触网鸟巢的检测上,改进的YOLOv3算法与目前比较热门的SSD、RetinaNet<sup>[13]</sup>等算法相比,有更高的检测精度

## 1 鸟巢图像预处理

鸟巢图像预处理主要包含灰度化,小波去噪,维纳滤波,数据增强4个阶段,其中灰度化是鸟巢图像预处理的首要步骤。小波去噪,维纳滤波,数据增强各阶段顺序不影响,可独立完成对鸟巢图像检测。由于拍摄到的鸟巢图像是彩色图像,每个像素点有 $256^3$ 个颜色变化区间,若直接对彩色图像处理,计算量大且训练速度慢。将鸟巢图像灰度化后,计算量大大减小<sup>[14-15]</sup>,且仍能反映图像亮度等级分布以及整体和局部色度。本文采用平均值法对鸟巢图像进行灰度化处理,将彩色的鸟巢图像RGB三分量取均值作为灰度图的灰度值。平均值灰度化结果图如图1中平均值灰度化结果图所示。

列车车载摄像机在拍摄过程中由于受到环境影响,导致采集到的鸟巢图像具有各种类型噪声。为了将无效的噪声信号过滤掉,本文采用小波阈值去噪,其中阈值的选择主要有硬阈值和软阈值函数<sup>[16]</sup>,如图1软硬阈值去噪结果图所示,分别为软硬阈值的去噪效果。

列车的高速行驶导致车载摄像机拍摄到的鸟巢图片存在一定程度的运动模糊,需要将带有运动模糊的鸟巢图片复原成原始图像。采用维纳滤波方法,使得恢复后的鸟巢图像与最初不带运动模糊的鸟巢图像间均方误差最小。维纳滤波的传递公式为

$$Q = \frac{H^*}{|H|^2 + Z/S} \quad (1)$$

式中: $H^*$ 为分母的复共轭; $Z$ 为噪声功率谱; $S$ 为原图像功率谱。结合式(1)可以得出

$$\hat{Q} = \frac{1}{H} \cdot \frac{|H|^2}{|H|^2 + Z/S} \cdot N \quad (2)$$

式中： $\hat{Q}$ 为去掉运动模糊后的图像做傅里叶变换。

利用维纳滤波去除运动模糊的结果图如图 1(f)所示。

一般来说,拍摄到的高铁接触网鸟巢图像仅占很少一部分,若直接对样本数量如此匮乏的鸟巢图

像进行训练,神经网络易出现过拟合现象,即出现在训练集上效果好,而对于未参与训练的数据集上效果不好。数据增强在一定程度上能避免过拟合现象的产生,基于此,采用数据增强<sup>[16]</sup>方法增加鸟巢数据集。部分扩增结果如图 1(f)~图 1(h)所示。

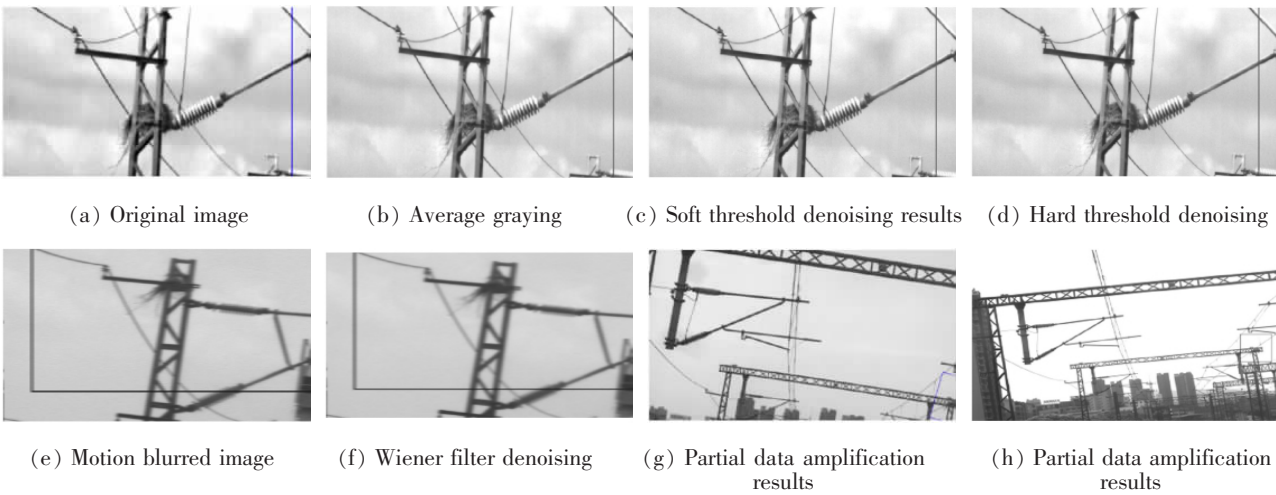


图 1 鸟巢图像预处理结果图  
Fig.1 Bird nest image preprocessing result map

如图 1 所示,经灰度化后的鸟巢图像与原始图像相比仍能反映鸟巢的形状特征、边缘特征;经过软硬阈值小波去噪处理后的鸟巢图像显示,软阈值去噪后的图像更平滑,去噪效果更好;维纳滤波结果图与原图相比,能够将拍摄高铁接触网鸟巢图像的过程中产生的运动模糊过滤;数据增强采用翻转、旋转、改变亮度对比度等方式增加鸟巢数据集数量。

2 深度学习结构

2.1 深度学习算法原理

典型的深度学习模型有卷积神经网络、深度信任网络、堆栈自编码网络三类模型。卷积神经网络在计算机视觉等领域取得了很好的效果。一般的卷积神经网络结构主要包含卷积层、池化层、全连接层,基本结构如图 2 所示。

高铁接触网鸟巢图像经卷积神经网络输入层并传递给卷积层,卷积层利用卷积核从上到下,从左到右对输入鸟巢图像进行卷积操作,能够提取数据的抽象特征和边缘特征;池化层能够对数据进行降维,对特征进行压缩,减少参数量;全连接层综合提取到的特征并映射为固定长度向量,一般包含类

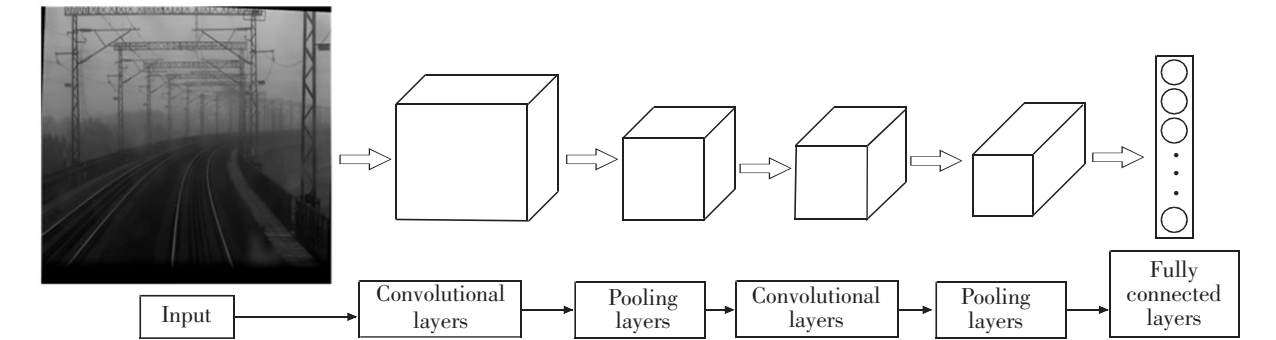


图 2 卷积神经网络的基本结构

Fig.2 Basic structure of convolutional neural network

别概率及坐标位置等。本文使用的深度学习算法首先直接将图片作为输入,将图片分割为  $G \times G$  个单元格,在每个单元格上通过边界框预测,最终输出每个边界框预测的目标概率和坐标,包含边界框的横坐标,纵坐标,宽度,高度以及类别的共 5 个置信度信息,分别用  $a_x, a_y, a_w, a_h, a_{\text{conf}}$  表示,其中置信度  $a_{\text{conf}}$  如下所示

$$a_{\text{conf}} = \text{Pre} \times S_{\text{IOU}} \quad (3)$$

其中  $\text{Pre}$  表示单元格是否含有目标,令

$$\text{Pre} = \begin{cases} 0, & \text{该单元格不包含目标中心} \\ 1, & \text{该单元格包含目标中心} \end{cases} \quad (4)$$

$S_{\text{IOU}}$  表示预测框  $T_p$  与真实边界框  $T_r$  的交并比,即

$$S_{\text{IOU}} = \frac{T_p \cap T_r}{T_p \cup T_r} \quad (5)$$

若目标中心落在某一块单元格内部,则由该单元格预测该目标,以这个单元格为中心生成多尺度先验框,比对先验框与真实框的交并比选择最合适尺度的先验框进行尺度缩放和平移边界框来判断物体的位置。记边框预测值为  $(p_x, p_y, p_w, p_h)$ ,目标中心在单元格中相对于图像左上角偏移量为  $(s_x, s_y)$ ,

同时先验框具有宽度和高度  $(A_w, A_h)$ ,则修正后(amend)的边框为

$$\begin{cases} b_{\text{amend}_x} = \sigma(p_x) + s_x \\ b_{\text{amend}_y} = \sigma(p_y) + s_y \\ b_{\text{amend}_w} = A_w e^{p_w} \\ b_{\text{amend}_h} = A_h e^{p_h} \end{cases} \quad (6)$$

## 2.2 深度学习网络结构

本文使用的深度学习算法 YOLOv3 网络模型以卷积神经网络作为基础结构,主要由特征提取模块 DarkNet-53<sup>[17]</sup>、FPN(feature pyramid network)组成<sup>[18]</sup>。其中,特征提取模块 DarkNet-53 共含有 53 层卷积层,除去最后一层由  $1 \times 1$  卷积组成的全连接层。网络整体结构如图 3 所示。

增强后的鸟巢图片尺寸为  $640 \times 640$ ,首先经过 Reshape 将尺寸统一到  $416 \times 416$  后再输入到 Darknet-53 网络模型中。经过 Darknet-53 结构将后 3 个残差模块输出的特征图输入到 FPN 中。其中,  $13 \times 13$  的特征图经过 5 次卷积输入到检测网络中,再将该特征图经过一次卷积和上采样与  $26 \times 26$  的特征图拼接融合,将融合后的特征图经过 5 次卷积输入

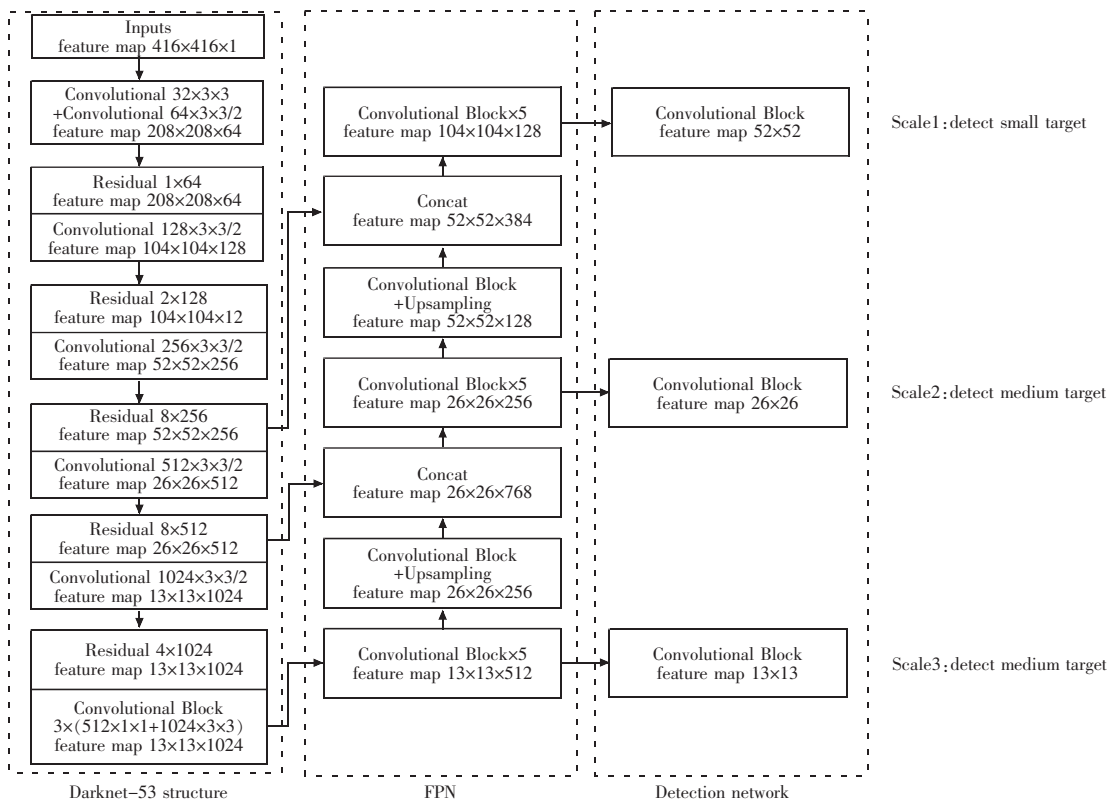


图3 YOLOv3 网络整体结构

Fig.3 The overall structure of the YOLOv3 network



到检测网络中,再将该特征图经过一次卷积和上采样与  $52 \times 52$  的特征图拼接融合,经过 5 次卷积输入到检测网络中。最终检测网络输出  $52 \times 52, 26 \times 26, 13 \times 13$  特征图分别适合检测小目标,中等目标,大目标鸟巢。

### 3 改进的深度学习算法

#### 3.1 引入空间金字塔池化处理

由于高铁接触网鸟巢图像的鸟巢目标轮廓不清晰,降维过程易发生信息丢失等问题,鸟巢图像在预处理和多尺度预测过程中,存在拉伸成固定尺寸等操作导致信息失真影响后续检测结果。为了解决这个问题,引入空间金字塔池化。空间金字塔池化的网络结构图如图 4 所示。

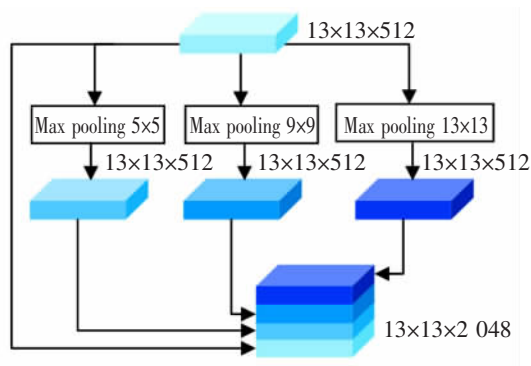


图 4 空间金字塔池化处理模块

Fig.4 Spatial pyramid pooling module

空间金字塔池化处理模块如图 4 所示。经 5 次残差模块后,鸟巢的输入图像尺寸从  $416 \times 416 \times 1$  变成  $13 \times 13 \times 1\,024$ ,为了提取到鸟巢的多尺度信息,在经过 3 次卷积后,对特征图进行分块池化,本文选择的池化核为 5,9,13,其中为了不使输入特征图尺寸发生变化,采用步长为 1 且填充的池化方式。经过 3 个不同尺度的池化操作后的特征图和原始全局特征图融合后能够获得更多特征,消除由于鸟巢尺寸大小不一样导致有效信息不一致的影响,有助于提高对鸟巢的检测精度。

#### 3.2 广义交并比及其损失函数

目标检测常把均方误差作为损失,但又使用交并比(IoU)<sup>[19]</sup>即真实框与预测框的交并比来衡量预测结果的好坏,模型对两者的优化并不等价,当真实框与预测框的均方误差损失相等时,两者交并比并不一定相等。同时交并比能被反向传播,若把交

并比作为损失函数直接被优化即直接优化评价指标比选择替代的损失函数好。但交并比有两个缺点:① 当真实框与预测框不相交时,交并比为 0,此时梯度也为 0,无法进行优化;② 若真实框与预测框不相交时,如图 5 所示,两者不相交,交并比为 0,但无法反映哪个预测结果更好。

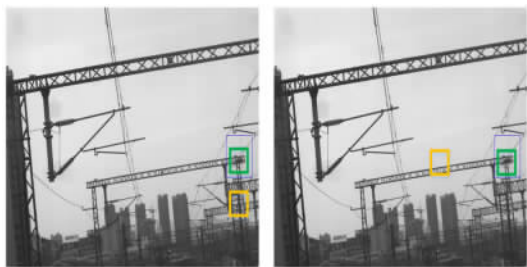


图 5 交并比为 0 时两种可能的情况

Fig.5 Two possible situations when IoU is 0

如图 5 所示,虽然绿色的真实框与黄色预测框都不相交,但明显左边的预测框与真实框距离更近,预测效果更好。

为了更好的评价预测框的质量,本文利用 GIoU (Generalized IoU, GIoU, 令为  $G_{IoU}$ ) 即广义交并比作为新的预测框与真实框的距离的评价方法。GIoU 不仅关注真实框预测框交集,也关注两者并集,更能评价两者重叠关系。

GIoU 评价方法的计算步骤为:

- 1) 计算真实框与预测框的面积分别用  $T_T, T_P$  表示;
- 2) 计算两者交集  $T_T \cap T_P$  和并集  $T_T \cup T_P$ ;
- 3) 找到包含真实框与预测框的最小框,并计算其面积大小,用  $T_{\min}$  表示;
- 4) 计算  $S_{IoU} = \frac{T_T \cap T_P}{T_T \cup T_P}$ ;
- 5) 计算  $G_{IoU} = S_{IoU} = \frac{|T_{\min} / T_T \cup T_P|}{|T_{\min}|}$ 。

GIoU 也能被反向传播,当作为目标函数被优化时,损失函数的表达式为

$$L_{GIoU} = 1 - G_{IoU}$$

这种评价方式当真实框与预测框不相交时,也能进行优化,改善了当 IoU 为 0 时难以优化的问题。

### 4 实验分析

#### 4.1 实验环境的配置

本文使用 Pytorch 平台进行鸟巢检测识别,使用 GPU 加速,实验环境配置如表 1 所示。

表 1 实验环境配置  
Tab.1 Experimental environment configuration

Name	Type
System	Win10
Language	Python3.7.6
CPU	Intel(R)Xeon(R)W-2123CPU
GPU	Nvidia Quadro P4000
Memory	32 GB
Hand disk	256 GB
Frame	Pytorch

4.2 网络的训练及测试

针对某高铁接触网鸟巢图像,经 6 种数据增强方法后得到的鸟巢图像共 716 张,其中训练集共 585 张,验证集 70 张,测试集 61 张。将鸟巢图片尺寸归一化至 640×640 像素。根据实验需求将鸟巢数据集标注为数据格式并保存。模型训练开始阶段将学习率设置为 0.001,在 step 模式下更新学习率,训练迭代次数达到 7 000 次时学习率下降为初始学习率的 10%,训练迭代次数达 8 000 次时,学习率降为初始学习率的 1%。训练过程对损失函数值进行记录。

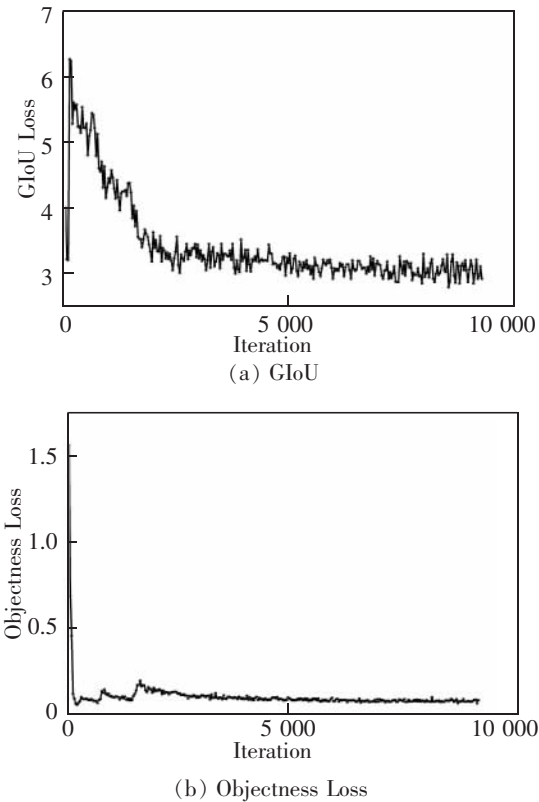


图 6 损失变化曲线图  
Fig.6 Loss change graph

如图 6 所示,损失随迭代次数增加逐渐趋于稳定,GIOU 损失达到 3.5 左右基本不超过 4,Objectness 损失达到 0.1 左右基本不超过 0.25。根据曲线收敛情况所示,9 000 次训练后可停止迭代。

在高铁接触网鸟巢图像检测中,网络模型精度与召回率可由下列两个公式计算得出。

$$P=\frac{n_{TP}}{sum_1}$$
 (7)

$$R=\frac{n_{TP}}{sum_2}$$
 (8)

式中: $P$  表示网络模型精度; $R$  表示召回率; $n_{TP}$  表示鸟巢被正确框出的数量; $sum_1$  表示所有被框出数量的总和; $sum_2$  表示图片中鸟巢的数量总和。平均精度均值  $mAP$  是衡量模型在不同类别中检测效果的重要指标,计算方法为

$$mAP=\frac{\sum AP}{num(class)}$$
 (9)

式中: $num(class)$  表示图像检测类别数目总和; $AP$  (average precision) 表示精度与召回率关系曲线的有限积分值,与模型整体表现呈正相关。改进后的 YOLOv3 网络在接触网鸟巢数据集上鸟巢检测的  $AP$  曲线如图 7 所示。

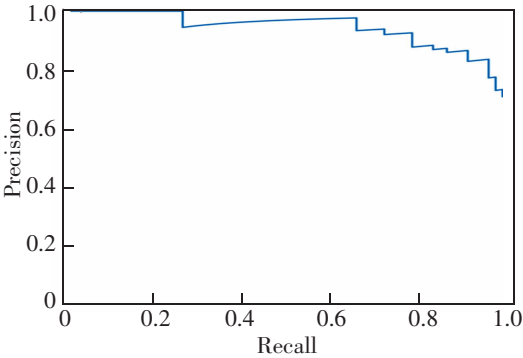


图 7 改进后的 YOLOv3 网络在鸟巢数据集上的 AP 曲线  
Fig.7 AP curves of the improved YOLOv3 network on the bird nest dataset

如图 7 所示,改进后的 YOLOv3 网络在鸟巢数据集上的  $AP$  达到 95.1%。 $mAP$  是  $AP$  在不同类别下的均值,在本文单一类别数据集中,可作为鸟巢检测的评价指标。测试结果如表 2 所示。

由表 2 可知,基于层数 50 层的 RetinaNet 算法的鸟巢检测, $mAP$  达到 65.7%,基于 SSD300 算法的鸟巢检测, $mAP$  达到 90.2%,原始 YOLOv3 算法在鸟巢上的检测, $mAP$  达到 90.4%,改进的 YOLOv3

算法在鸟巢上的检测,  $mAP$  达到 95.1%。结果表明,改进的 YOLOv3 算法与层数 50 层的 RetinaNet 算法、SSD300 算法相比,在接触网鸟巢上的检测有更高的平均准确度。

表 2 不同检测算法在本文数据集上的检测结果  
Tab.2 Detection results of different detection algorithms on the dataset in this paper

Detection algorithms	$mAP/\%$
RetinaNet-depth=50	65.7
SSD300	90.2
YOLOv3	90.4
Improred YOLOv3	95.1

用改进后的 YOLOv3 网络对高铁接触网鸟巢图片进行测试,部分检测结果如图 8 所示。检测结果显示,在高铁接触网复杂背景下,改进的 YOLOv3 网络能检测到部分不清晰、有遮挡的鸟巢,对于远距离、部分形状较小的鸟巢也能检测到,具有较好

的泛化能力。

4.3 接触网鸟巢特征图的提取与转换

卷积神经网络由多个卷积层组成,每个卷积层利用卷积核对图像进行卷积运算,得到特征图,每层提取的特征不同。浅层卷积层提取较为局部、细节的信息,高层卷积层感知图像更复杂、抽象的信息。通过对部分卷积层输出的鸟巢图像特征图的分析,得到卷积神经网络对鸟巢信息的提取过程。如图 9 所示。

可以看出,神经网络对鸟巢图像的信息提取过程与人认知图像相似。人在认知图像时先感知图片亮度信息与颜色信息,再是边缘、轮廓等局部细节,然后是纹理特征等更加复杂的信息,最后将所有信息综合形成整体的概念。如图 9 所示,神经网络先对杆塔边缘仔细辨认,突出夹角位置,随着网络层数加深,将杆塔与背景分开并寻找目标可能的重点部位。



(a) Test result



(b) Test result



(c) Test result



(d) Test result

图 8 改进后的 YOLOv3 网络在接触网鸟巢数据集上的测试结果

Fig.8 Test results of the improved YOLOV3 network on the Catenary Nest dataset

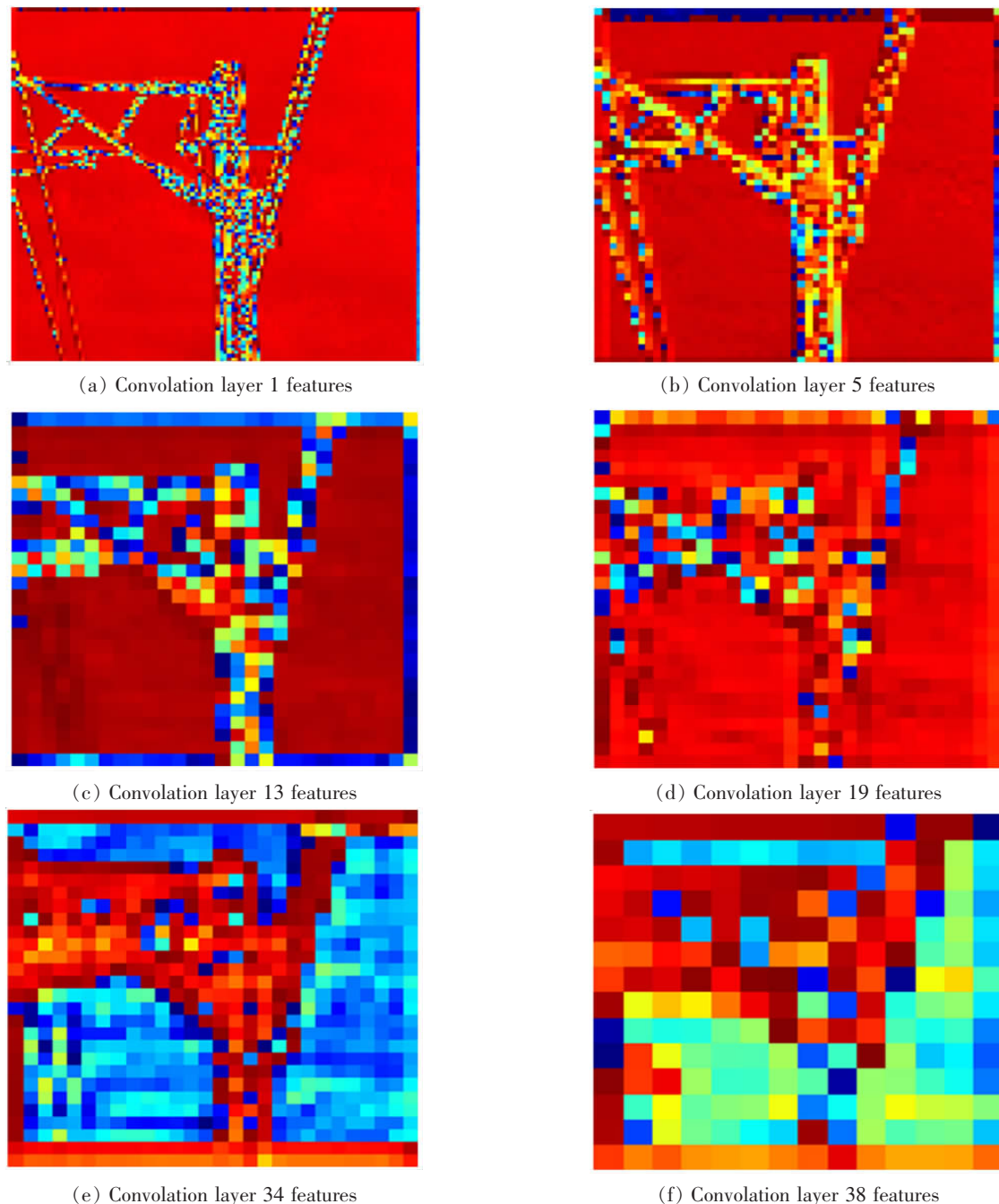


图 9 卷积层输出特征图

Fig.9 The convolutional layer outputs the characteristic graph

## 5 结论

该文基于 Pytorch 平台使用以卷积神经网络为基础的深度学习模型,对高铁接触网鸟巢图像进行检测与识别。

1) 前期对鸟巢图像的去噪处理,能有效降低各种噪声对鸟巢目标的干扰,突出鸟巢目标;数据增强能提高鸟巢样本数目,一定程度上解决了神经网络

出现过拟合问题。

2) 将空间金字塔池化与改进的深度学习模型结合,有效提高对鸟巢的多尺度特征信息,同时引入广泛交并比作为损失函数,与均方误差相比,直接将评价真实框与预测框距离的方法作为损失函数,改善了 IoU 作为评价方法时,真实框与预测框不相交无法优化的问题,更适合作为损失函数进行优化。



3) 与传统检测鸟巢的方法相比,该方法检测鸟巢精度更高,有更高的检测速度,能在复杂的接触网背景下较好的识别检测鸟巢,对接触网鸟害的处理有重要意义。

### 参考文献:

- [1] 王少华,叶自强. 架空输电线路鸟害故障及其防治技术措施[J]. 高压电器,2011,47(2):61-67.  
WANG S H, YE Z Q. Analysis of bird damage accidents on overhead transmission lines and prevention techniques[J]. High Voltage Apparatus, 2011, 47(2): 61-67.
- [2] 张栋梁,韩志伟,刘志刚,等. 基于集对分析和证据理论的高铁接触网健康状态评估[J]. 铁道学报,2020,42(5):58-65.  
ZHANG D L, HAN Z W, LIU Z G, et al. Assessment of health status of high-speed railway catenary based on set-pair analysis and evidence theory[J]. Journal of the China Railway Society, 2020, 42(5): 58-65.
- [3] 周俊,陈剑云. 基于 DSSD 的接触网鸟窝识别检测研究[J]. 华东交通大学学报,2019,36(6):70-78.  
ZHOU J, CHEN J Y. Research on recognition and detection of catenary bird's nest based on DSSD[J]. Journal of East China Jiaotong University, 2019, 36(6): 70-78.
- [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(4):1043-1049.  
ZHU Z M, XIE L K. Detection of birds' nest in catenary based on relative position invariance[J]. Journal of Railway Science and Engineering, 2018(4): 1043-1049.
- [5] 段旺旺,唐鹏,金炜东,等. 基于关键区域 HOG 特征的路接触网鸟巢检测[J]. 中国铁路,2015(8):73-77.  
DUAN W W, TANG P, JIN W D, et al. Nest detection of railway catenary based on HOG feature in key areas[J]. Chinese Railways, 2015(8): 73-77.
- [6] 金炜东,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 W D, YANG P, TANG P. Double discriminator generative adversarial networks and their application in detecting nets built in catenary and semisupervised learning[J]. Scientia Sinica(Informationis), 2018, 48(7): 888-902.
- [7] 罗婷婷. 两种改进的彩色图像灰度化算法研究[D]. 杭州:浙江工商大学,2014.  
LUO T T. Research on two improved decolorization algorithms[D]. Hangzhou: Zhejiang Gongshang University, 2014.
- [8] 马旗,朱斌,程正东,等. 基于双通道的快速低空无人机检测识别方法[J]. 光学学报,2019,39(12):97-107.  
MA Q, ZHU B, CHENG Z D, et al. Detection and recognition method of fast low-altitude unmanned aerial vehicle based on dual channel[J]. Acta Optica Sinica, 2019, 39(12): 97-107.
- [9] 鞠默然,罗海波,王仲博,等. 改进的 YOLOv3 算法及其在小目标检测中的应用[J]. 光学学报,2019,39(7):245-252.  
JU M R, LUO H B, WANG Z B, et al. Improved YOLOv3 algorithm and Its application in small target detection[J]. Acta Optica Sinica, 2019, 39(7): 245-252.
- [10] REN S Q, HE K M, GIRSHICK R, et al. FasterR-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[C]//European Conference on Computer Vision. Springer, Cham, 2016.
- [12] 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, 2016.
- [13] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017(99): 2999-3007.
- [14] JOSE M P, ERNESTO M, GERMAN O, et al. Freezing front velocity estimation using image processing techniques[J]. Measurement, 2020, 151: 107085.
- [15] 晋健,刘育,王琴慧,等. 基于小波去噪和 FA-SVM 的中长期径流预报[J]. 人民长江,2020,51(9):67-72.  
JIN J, LIU Y, WANG Q H, et al. Mid-long term runoff forecasting based on wavelet de-noising and FA-SVM[J]. Yangtze River, 2020, 51(9): 67-72.
- [16] 雷飞,朱林,王雪丽. 改进小波软硬折衷法在水下图像去噪中的应用[J]. 计算机技术与发展,2017,27(11):150-153.  
LEI F, ZHU L, WANG X L. Application of Improved compromise soft-hard threshold algorithm in underwater image denoising[J]. Computer Technology and Development, 2017, 27(11): 150-153.
- [17] 孙晓凯,倪卿元,陈文强. 图像增强方法在深度学习图像识别场景应用中的可行性研究[J]. 电信科学,2020,36(S1):172-179.  
SUN X K, NI Q Y, CHEN W Q. Feasibility study of image enhancement method in deep learning image recognition scene application[J]. Telecommunications Science, 2020, 36(S1): 172-179.
- [18] REDMON J, FARHADI A. YOLO9000: Better faster stronger[J]. IEEE Conference on Computer Vision & Pattern Recognition, 2017: 6517-6525.

(责任编辑:姜红贵)