

DOI:10.19651/j.cnki.emt.2209514

# 基于改进 YOLOv4 算法的高压塔鸟巢检测<sup>\*</sup>

谢国波<sup>1</sup> 郑晓锋<sup>1</sup> 林志毅<sup>1</sup> 林立<sup>1</sup> 文刚<sup>2</sup>

(1. 广东工业大学计算机学院 广州 510006; 2. 云南电网有限责任公司电力科学研究院 昆明 650000)

**摘要:** 针对现有算法对高压塔上鸟巢检测存在参数量过大,实时性不足及对小目标检测能力较弱的问题,提出了一种改进的 YOLOv4 算法。首先使用 Mobilenetv2 网络代替 CSPDarknet53 网络作为主干网络,减少算法的参数量且提升检测速度;同时在 Mobilenetv2 网络的逆残差网络中嵌入注意力 Coordinate Attention 模块,增强网络对目标特征提取能力。然后,对 PANet 网络进行改进,获取更多的细节特征信息,提高对小目标鸟巢的检测能力。最后,使用 Focal Loss 函数优化损失函数,降低大量简单背景样本训练的权重,提升对小目标鸟巢困难样本训练的侧重,进一步提高对小目标鸟巢的检测能力。实验结果表明,较原始的 YOLOv4 算法,改进后的 YOLOv4 算法的参数量减少了 48.1%,检测速度和精度分别提高了 12.9 fps 和 2.33%。即改进后的 YOLOv4 算法大幅度减少了算法参数量,且对鸟巢的检测拥有更好的检测性能。

**关键词:** 参数量;逆残差网络;细节特征;检测能力

中图分类号: TP391.4 文献标识码: A 国家标准学科分类代码: 520.6040

## Bird's nest detection of high voltage tower based on improved YOLOv4 algorithm

Xie Guobo<sup>1</sup> Zheng Xiaofeng<sup>1</sup> Lin Zhiyi<sup>1</sup> Lin Li<sup>1</sup> Wen Gang<sup>2</sup>

(1. School of Computer Science, Guangdong University of Technology, Guangzhou 510006, China;

2. Electric Power Research Institute of Yunnan Power Grid Co., Ltd., Kunming 650000, China)

**Abstract:** Aiming at the problems of excessive parameters, insufficient real-time performance and weak detection ability of small targets in the existing algorithms for bird's nest detection on high-voltage tower, an improved YOLOv4 algorithm is proposed. Firstly, Mobilenetv2 network is used to replace CSPDarknet53 network as the backbone network, which reduces the amount of parameters of the algorithm and improves the detection speed. At the same time, the Coordinate Attention module is embedded in the inverse residual network of Mobilenetv2 network, which enhance the ability of the network to extract target features. Then, the PANet network is improved to obtain more detailed feature information and improve the detection ability of small target bird's nest. Finally, the Focal Loss function is used to optimize the loss function, reduce the weight of a large number of simple background samples, and improve the focus on the difficult sample training of small target bird's nest, which further improves the detection ability of small target bird's nest. The experimental results show that compared with the original YOLOv4 algorithm, the parameters of the improved YOLOv4 algorithm are reduced by 48.1%, and the detection speed and accuracy are improved by 12.9 fps and 2.33% respectively. That is, the improved YOLOv4 algorithm greatly reduces the amount of algorithm parameters, and has better detection performance for bird's nest detection.

**Keywords:** amount of parameters; inverse residual; detailed feature; detection ability

## 0 引言

随着高压输电线路数量增加,鸟类对电力系统造成的破坏也越来越多。有些鸟在高压塔上进行筑巢。在下雨天

或天气湿度大的时候,鸟巢容易充当导体,从而触发跳闸。同样,在比较干燥的天气,鸟巢的树枝容易着火,这不仅给电力传输造成了影响,还容易造成巨大的安全风险<sup>[1]</sup>。因此,为了保证电网系统的正常运行,需要对高压塔上是否存在

收稿日期:2022-04-06

<sup>\*</sup> 基金项目:国家自然科学基金(61802072)项目资助

在鸟巢进行定期巡检清除。然而使用传统人工巡检的方法往往效率低且成本高,为此有必要研究自动检测鸟巢的方法。

针对鸟巢的检测,有学者尝试使用传统图像检测方法对鸟巢进行检测。徐静等<sup>[2]</sup>通过在图像中搜索符合鸟巢样本的连通区域并对连通区域分析,最后通过描述鸟巢的特征量来实现鸟巢检测。Lu 等<sup>[3]</sup>通过组合特征获得更多的鸟巢特征,使用级联分类器的方法实现对鸟巢的检测。祝振敏等<sup>[4]</sup>使用位置不变性对鸟巢进行检测,先对图像进行处理,再通过统计白色区域来判断图像中是否存在鸟巢。以上方法在特定条件下对鸟巢都有较好的检测效果,然而传统图像检测方法往往对复杂场景中目标特征的提取与描述能力不足,当鸟巢的背景较为复杂时,传统图像检测方法的检测效果往往不佳。

近年来,基于深度学习的目标检测方法凭借其具备更强的特征提取与描述能力,适用于越来越复杂的自然环境中,逐渐成为鸟巢检测研究的必然趋势。Li 等<sup>[5]</sup>使用各种图像处理方法对鸟巢数据集进行处理,最后得到的 Faster RCNN 算法对鸟巢检测的具有较强的适用性。王纪武等<sup>[6]</sup>通过在特征提取,建议区域以及检测三个方面改进 Faster RCNN 算法,改进后的算法更适用于鸟巢的检测。上述检测算法均拥有较高的检测精度,但由于这些算法存在参数量较大、实时性差的缺点,导致这些算法很难部署在嵌入式

设备中实现鸟巢检测。丁建等<sup>[7]</sup>为了兼顾算法的精度与速度,提出了双尺度 YOLOv3 网络对鸟巢进行检测。虽然在改进后的 YOLOv3 能够有较好的检测精度与速度,但其双尺度的检测方法并不适合部署在无人机等嵌入式设备中。然而与上述算法相比, YOLOv4 算法同样拥有较高的检测精度,且模型参数量较少,检测速度更高,因此本文选取 YOLOv4 算法作为研究对象。

虽然 YOLOv4 算法拥有优秀的检测精度与速度,但是应用于无人机等嵌入式设备进行鸟巢检测还存在以下问题:1)为了使得算法更好的部署在无人机等嵌入式设备中,算法的权重文件不能过大同时需要有良好的实时性;2)大部分图像中高压塔上鸟巢目标为小型目标,故需要对算法进行适当改进,提高算法对小目标鸟巢检测的性能。针对上述问题,本文对 YOLOv4 算法进行了主干网络的替换、PANet 的改进和损失函数的优化。实验结果表明,与原有的 YOLOv4 算法比较,改进后的 YOLOv4 算法大幅度减少了算法的参数量,且提升了对鸟巢的检测性能。

## 1 YOLOv4 算法

### 1.1 YOLOv4 算法原理

YOLOv4 算法是 Bochkovskiy 等<sup>[8]</sup>提出的目标检测方法,其网络结构主要由三部分组成,分别是主干网络,颈部网络和头部网络<sup>[9]</sup>,算法的网络结构如图 1 所示。

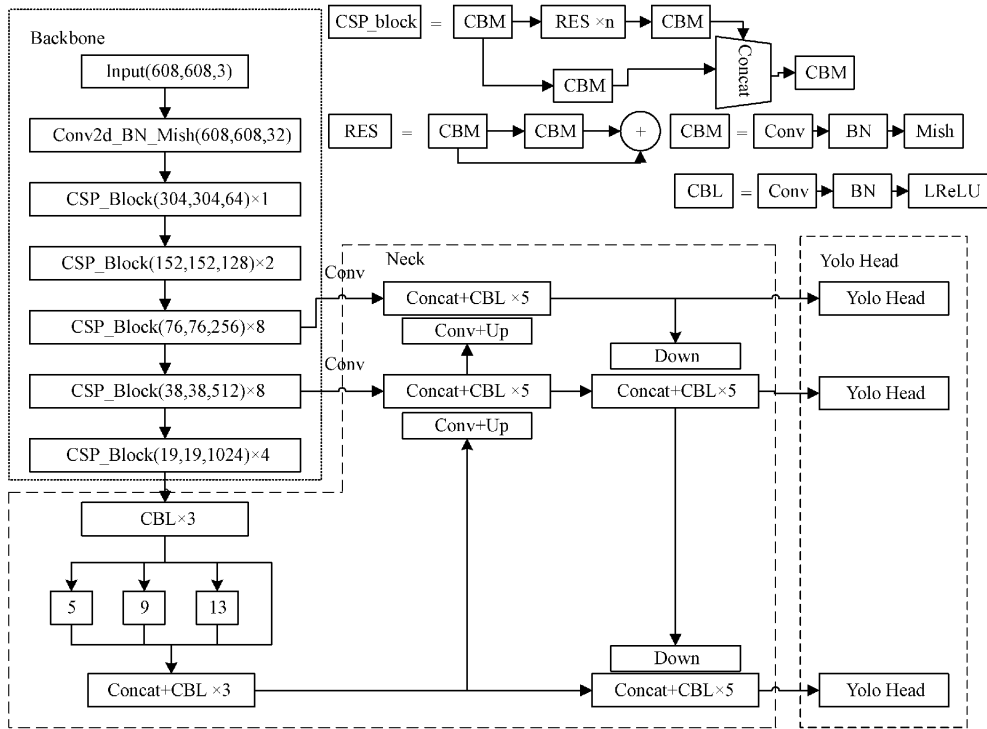


图 1 YOLOv4 结构

主干网络上, YOLOv4 算法的主干网络为 CSPDarknet53 网络, CSPDarknet53 网络主要由五个

CSPNet 残差块组成,引入 CSPNet 残差不仅降低了网络计算量,而且增强了特征的提取<sup>[10]</sup>。在激活函数上,

CSPDarknet53网络则使用Mish激活函数,进一步增强网络深层信息的传播<sup>[11]</sup>。

颈部网络上,YOLOv4算法则使用了SPP(空间金字塔)模块和PANet结合的结构。SPP模块对获得的特征图进行 $13 \times 13, 9 \times 9, 5 \times 5$ 和 $1 \times 1$ 四种不同尺度的最大池化操作,有效的增加了网络的感受野,提取出显著的上下文特征<sup>[12]</sup>。PANet主要是基于FPN(特征金字塔网络)进行改进,在FPN基础上增加了一条自下而上的路径,增强了不同特征图之间的融合<sup>[13]</sup>。

YOLOv4使用多尺度预测的方式,头部网络最终输出3个大小为 $13 \times 13, 26 \times 26, 52 \times 52$ 的检测头分别检测大、中、小三种目标<sup>[14]</sup>。最后通过预先设置好的先验框大小,在此基础上利用相对偏移量求得预测框的位置与大小。

## 1.2 损失函数

YOLOv4算法的损失函数主要由三部分组成,分别是置信度损失函数( $loss_{conf}$ ),分类损失函数( $loss_{class}$ )和回归框损失函数( $loss_{ciou}$ ),相关公式如式(1)~(6)所示。

$$Loss = loss_{conf} + loss_{class} + loss_{ciou} \quad (1)$$

$$loss_{conf} = - \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} [\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j)] - \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{noobj} [\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j)] \quad (2)$$

$$loss_{class} = - \sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in class} [\hat{P}_i^j \log(P_i^j) + (1 - \hat{P}_i^j) \log(1 - P_i^j)] \quad (3)$$

$$loss_{ciou} = \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} (2 - w^{gt} \times h^{gt}) \left[ 1 - \text{IOU} + \frac{\rho^2(b, b^{gt})}{m^2} + \alpha v \right] \quad (4)$$

$$\alpha = \frac{v}{[1 - \text{IOU}] + v} \quad (5)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (6)$$

式中: $S^2$ 表示网格数量, $B$ 为网格中先验框的个数, $I_{ij}^{obj}$ , $I_{ij}^{noobj}$ 表示是否有目标在网络中, $\lambda_{noobj}$ 为置信度权重参数。

$\hat{C}, C$ 为真实框与预测框的置信度, $\hat{P}, P$ 为真实框与预测框的类别概率。 $w^{gt}, h^{gt}$ 分别表示真实框宽高, $\rho^2(b, b^{gt})$ 为真实框与预测框的中心点之间的欧式距离, $a$ 表示权重函数, $v$ 用于度量长宽比的相似性,相关公式分别为式(5)、(6)。

## 2 改进YOLOv4算法

为了解决YOLOv4算法在鸟巢检测中存在的问题,本文对YOLOv4算法进行了相应的改进,主要改进如下:1)为了减少算法的参数数量且满足实时性的要求,本文使用轻量级Mobilenetv2网络代替CSPDarknet53网络,同时为了避免使用轻量化网络造成网络特征提取能力不足,本文将Coordinate Attention注意力模块嵌入主干网络的逆残差

网络中。2)为了增强对小目标鸟巢的检测能力,在PANet网络中,增加了尺寸为 $(152, 152)$ 特征图的融合,五次连续卷积模块替换为三次连续卷积模块。使用Focal Loss函数对损失函数进行优化。改进后的YOLOv4算法其网络结构如图2所示。

### 2.1 主干网络的优化

#### 1) Mobilenetv2网络

本文选取轻量级Mobilenetv2网络作为算法的特征提取网络,Mobilenetv2网络是Howard等<sup>[15]</sup>提出的轻量级网络,相较于CSPDarknet53网络,它拥有更少的参数,同时拥有较强的特征提取能力,它的特点主要分为两方面:使用深度可分离卷积和逆残差网络。

(1)深度可分离卷积由深度卷积(Depthwise Conv)和逐点卷积(Pointwise Conv)组成<sup>[16]</sup>。如图3所示,深度卷积在不改变输入特征图通道数的情况下对每个通道进行卷积操作。逐点卷积用于对特征图进行升维与降维操作。当输入通道数为 $c_1$ ,大小为 $w \times h$ 的特征图转换到通道数为 $c_2$ 的同尺寸特征图时,在卷积核大小为 $k \times k$ 的标准卷积中,参数数量为: $c_1 c_2 k^2$ 。而在深度可分离卷积中,其参数数量为: $c_1 k^2 + c_1 c_2$ ,其参数数量降低为标准卷积的 $1/c_2 + 1/k^2$ 。当使用 $3 \times 3$ 卷积核大小时,使用深度可分离卷积进行操作后可使参数数量下降到原来的 $1/8 \sim 1/9$ 。

(2)传统的残差网络使用卷积对特征图进行降维,经过卷积提取特征后再进行升维。由于输入更多的特征图通道数有利于深度可分离卷积提取更多特征信息,逆残差网络先使用卷积对特征图进行升维,经过卷积提取更多有效的特征,再压缩通道数。同时,使用Linear函数代替ReLU函数,减少特征信息的损失。

#### 2) Coordinate Attention模块嵌入逆残差网络

为了避免使用轻量化网络造成网络特征提取能力不足,本文在Mobilenetv2网络中的逆残差结构中嵌入Coordinate Attention注意力模块<sup>[17]</sup>,相较于通道注意力模块,CA(Coordinate Attention)注意力模块在通道注意力中嵌入了位置信息,从而使目标检测定位更加精确<sup>[18]</sup>。如图4所示,CA注意力模块的运算过程主要分为两步:坐标信息嵌入和坐标注意力生成。

(1)坐标信息嵌入:不同于通道注意力模块使用全局池化将输入转化为单个特征向量的方式,为了促进注意力模块能够捕获具有精确位置信息的空间长程依赖,CA注意力模块将全局池化分解为每个通道上两个不同方向上的平均池化。即对于输入特征图,使用不同尺寸的池化核沿着水平与垂直坐标方向对每个通道进行编码处理,因此高度为 $h$ 的第 $c$ 个通道的输出可以用式(7)表示,宽度为 $w$ 的第 $c$ 个通道的输出可以用式(8)表示。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (7)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < W} x_c(j, w) \quad (8)$$

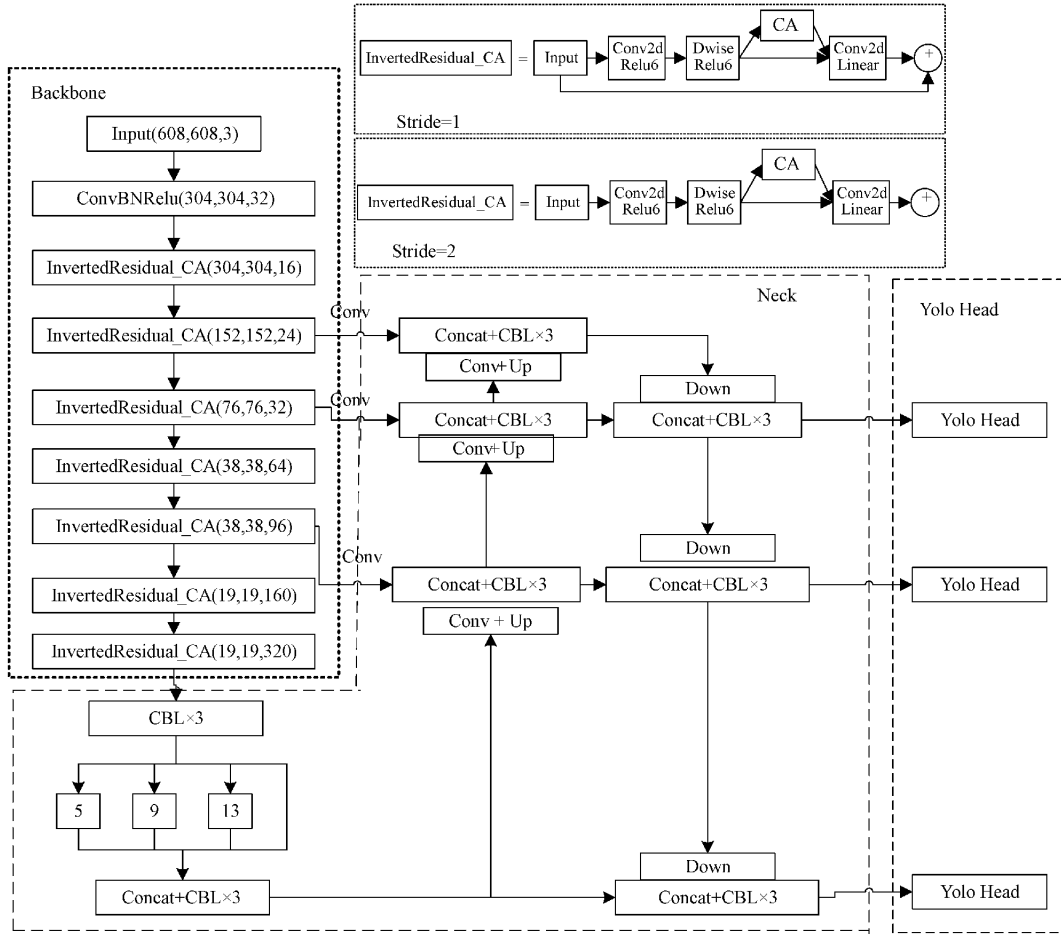


图 2 改进后的 YOLOv4 结构

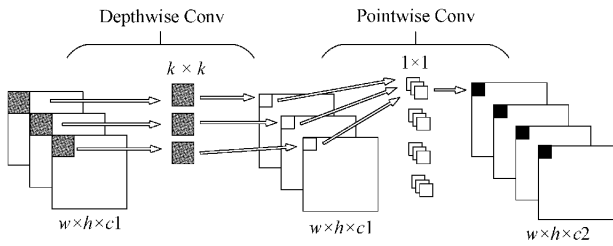


图 3 深度可分离卷积

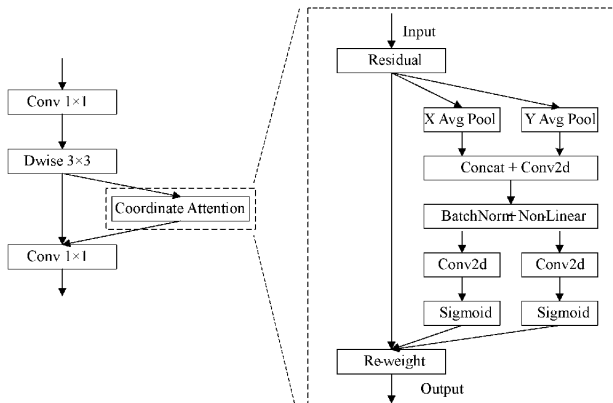


图 4 逆残差网络中嵌入 CA 注意力

式中:  $W$  与  $H$  分别表示特征图的宽度与高度,  $x_c$  表示第  $c$  个通道的特征图。上述两种变换沿着水平与垂直坐标方向编码每个通道, 最终能获得一对具有方向感知特征的特征图。

(2) 坐标注意力生成: 将坐标信息嵌入过程得到的特征图进行拼接, 卷积和激活函数的操作, 得到同时具有水平与垂直方向空间信息的特征图, 如式(9)所示。

$$f = \delta(F([\mathbf{z}^h, \mathbf{z}^w])) \quad (9)$$

式中:  $\delta$  表示 ReLU 激活函数功能,  $F$  表示进行  $1 \times 1$  卷积操作,  $[\cdot]$  表示拼接操作。再将特征图  $f$  按空间维度切分成两个独立张量, 对张量分别进行卷积与激活函数的操作, 将其通道数变换同原始输入特征图一致, 将结果进行扩展作为水平与垂直坐标方向的注意力权重, 如式(10)~(11)所示。

$$g^h = \sigma(F_h(f^h)) \quad (10)$$

$$g^w = \sigma(F_w(f^w)) \quad (11)$$

式中:  $\delta$  表示 sigmoid 函数,  $F_h, F_w$  分别表示对张量  $f^h$  与  $f^w$  进行  $1 \times 1$  卷积操作。将得到的权重与输入的特征图相乘, 得到具有权重的特征图, 增强聚焦特征能力。

### 2.2 PANet 网络的改进

当输入图像尺寸为  $608 \times 608$  时, YOLOv4 算法通过

主干网络提取 3 个特征图尺寸分别为  $(76 \times 76)$ ,  $(38 \times 38)$  和  $(19 \times 19)$ , 这 3 个特征图分别被命名为中间层, 中底层和底层<sup>[19]</sup>。将提取到的 3 个不同尺寸特征图输入 PANet 网络中, PANet 网络通过自下而上与自上而下融合深层特征图与浅层特征图, 使得最终输出的 3 个特征图拥有更多的有效特征信息, 从而提高算法的检测性能。

在深度卷积神经网络中, 深层的特征图往往拥有更多的语义信息, 但是细节特征信息会出现丢失, 此时更有利于检测大目标; 相反, 浅层的特征图所包含的特征更趋向于细节信息, 此时更有利于检测小目标。为了增强算法对小目标的检测性能, 本文将尺寸为  $(152, 152)$  的浅层特征图融入 PANet 网络中, 使得最终的检测特征图获得更丰富的细节信息, 更有利于小目标的检测。同时, 在保证检测性能的前提下, 为避免过多的计算开销, 本文将 PANet 网络的五次连续卷积模块代替为三次连续卷积模块。改进后的 PANet 网络如图 5 所示, PANet 网络的上采样与下采样次数都增加到三次, 五次连续卷积模块都被代替为三次连续卷积模块。

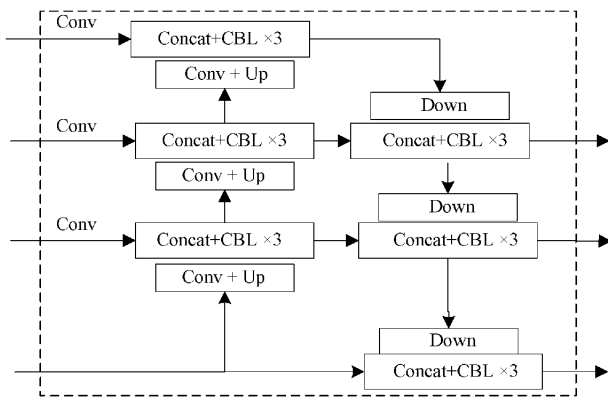


图 5 改进的 PANet 网络结构

### 2.3 损失函数的优化

YOLOv4 算法在训练时需要处理大量的先验框, 当先验框中绝大多数不包含待检测目标时, 会造成训练正负样本数量不平衡的问题。通过分析鸟巢数据集, 目标鸟巢大多占据图像的比例较小, 该情况容易导致算法在训练过程中出现正负样本数量不平衡。基于此, 本文引入了 Focal Loss 函数代替交叉熵损失函数, 使用该函数减少简单样本背景所占的权重, 让模型专注于目标物体的检测, 避免算法预测时倾向于数量较大的背景样本, Focal Loss 函数如式(12)所示。

$$FL(p_t) = -\alpha_t(1 - p_t)^{\gamma} \log(p_t) \quad (12)$$

$$p_t = \begin{cases} p & y = 1 \\ 1 - p & y = -1 \end{cases} \quad (13)$$

式中:  $\alpha_t$  ( $\alpha_t \in [0, 1]$ ) 用于调整正负样本权重, 通过设置  $\alpha_t$  参数可以降低负样本(背景)在损失函数中的权重, 相对增加正样本(目标鸟巢)的权重, 本实验  $\alpha_t$  设置为 0.5。 $r$  ( $r \geq 0$ ) 用于调整难易样本的权重, 通过设置  $r$  参数提升

算法对小目标鸟巢困难样本训练的侧重, 从而提高对小目标鸟巢的检测能力, 本实验  $r$  设置为 2。 $p_t$  则表示预测框为目标鸟巢的概率, 定义如式(13)所示。

## 3 实验

### 3.1 实验环境及训练方法

本文实验采用的是台式计算机, 系统为 64 位的 windows10 系统, 处理器为 Intel(R) Core(TM) i9-10900K, 显卡为 Nvidia GeForce GTX 3080/10 G, 运行内存为 64 GB。

改进后的 YOLOv4 算法沿用 YOLOv4 算法中的余弦退火学习率方法来调整学习率。训练过程中, 初始学习率为  $1 \times 10^{-4}$ , 最大学习率与最小学习率分别设置为  $1 \times 10^{-3}$  和  $1 \times 10^{-5}$ , 训练轮次为 100, 批次大小为 8。图 6 为训练轮次与损失值下降关系, 从图 6 中可看出, 训练轮次 30 次之后损失值下降开始变缓, 且当训练轮次在 75 次之后, 算法的损失值趋于稳定。由此可知, 改进后的 YOLOv4 算法学习效果理想。

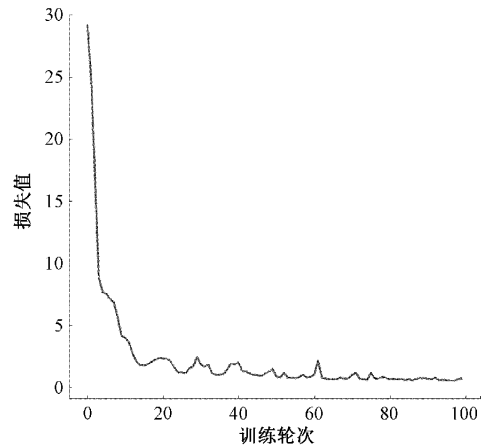


图 6 改进后 YOLOv4 算法的损失值变化

### 3.2 数据预处理

原始的鸟巢数据集共有 760 张, 分别来自于公共鸟巢数据集<sup>[20]</sup>与供电局无人机巡检航拍得到的图像。但该鸟巢图像数据量难以训练一个可靠的算法, 为了使得训练得到的算法具有鲁棒性, 本文尝试对鸟巢图像进行数据增强处理。本文随机选取一定数量的图像对图像进行仿射变换, 高斯噪声等数据增强处理, 最终获得了 2 016 张鸟巢图像, 按照训练集与测试集比例为 8 : 2 对数据集进行划分, 得到 1 612 张图像用于算法的训练, 404 张图像则用于测试算法的效果。

### 3.3 评价指标

在目标检测中, 准确率与召回率是一对矛盾的指标, 单一的指标往往无法完整衡量算法的性能, 通常使用 AP (average precision) 平均精度值作为算法检测精度指标, 相关公式如式(14)~(16)所示。

$$AP = \int_0^1 P(R) dR \quad (14)$$

$$P = \frac{TP}{TP + FP} \quad (15)$$

$$R = \frac{TP}{TP + FN} \quad (16)$$

式中： $P, R$  分别表示准确率与召回率。 $TP$  代表准确检测出鸟巢的数量， $FP$  代表非鸟巢的样本检测为鸟巢的数量， $FN$  表示未检测出的鸟巢的数量。

在实际应用中，网络往往需要部署在嵌入式设备中，故本文也关注了模型的大小与检测速度。模型的大小由模型参数量或权重大小表示。而检测速度由 FPS 决定，FPS 表示每秒能检测的图像数量。

### 3.4 实验结果分析

为了验证本文使用轻量级网络 Mobilnet2 网络作为原始 YOLOv4 算法主干网络的优势，本文选取了 Mobilenet 网络系列的 Mobilenet1, Mobilenet3 作为原始 YOLOv4 算法的主干网络进行对比实验。实验结果如表 1 所示。由表 1 可知，使用 Mobilnet2 网络作为原始 YOLOv4 算法的主干网络，算法的检测精度略低于使用 Mobilnet3 网络作为主干网络，而算法的检测速度优于其他算法，且算法的权重文件最小。综合检测精度，检测速度与权重文件大小来看，Mobilenet2-YOLOv4 算法的综合性能最好。

表 1 YOLOv4 轻量化主干网络实验对比

模型	AP/%	FPS	权重/MB
YOLOv4	85.02	36.4	244
Mobilenet1-YOLOv4	80.88	50.4	151
Mobilenet3-YOLOv4	84.19	47.2	147
Mobilenet2-YOLOv4	83.88	51.3	144

为了进一步验证改进后 YOLOv4 算法中各个改进的有效性，本文以 Mobilenet2-YOLOv4 算法为基础算法进行消融实验。表 2 给出了使用不同改进后算法的实验结果。其中 CA 表示 Coordinate Attention 注意力模块，P 表示对 PANet 网络结构进行改进，F 表示使用 Focal Loss 函数。由表 2 可知，将 Coordinate Attention 注意力嵌入逆残差网络后，在略微降低检测速度的情形下，算法的检测精度提升了 1.62%，说明加入 Coordinate Attention 注意力模块增强了主干网络的特征提取能力，同时增加了少量的计算开销，因此检测速度略微降低。在此改进上，对 PANet 网络进行改进，此时，算法的检测精度与速度提升了 1.03% 与 1.8 fps，说明改进后的 PANet 网络不增加额外的计算开销，且检测层能够获取更多的细节信息。最后，对算法的损失函数进行优化，算法的检测精度提升了 0.82%，说明改进后的损失函数能够有效的降低大量简单背景环境样本训练的权重，提升对小目标鸟巢困难样本训练的侧重，从而提高对小目标鸟巢的检测能力。

表 2 消融实验结果

基础模型	CA	P	F	AP/%	FPS
√	×	×	×	83.88	51.3
√	√	×	×	85.50	47.5
√	√	√	×	86.53	49.3
√	√	√	√	87.35	49.3

表 3 为原始 YOLOv4 算法与改进后 YOLOv4 算法的检测性能与算法参数量，权重文件大小对比。由表 3 可得，较原始的 YOLOv4 算法，改进后的 YOLOv4 算法在检测精度与速度分别提升了 2.33% 和 12.9 fps，模型的参数量缩减了 48.1%，且模型的权重文件大小缩减为 121 MB。即改进后的 YOLOv4 算法拥有更好的检测性能且更适合部署在嵌入式设备中。

表 3 YOLOv4 算法与改进后的 YOLOv4 算法对比

模型	AP/%	FPS	参数量/( $\times 10^7$ )	权重/MB
YOLOv4	85.02	36.4	6.40	244
本文模型	87.35	49.3	3.32	121

### 3.5 不同算法结果对比

为了验证改进后的 YOLOv4 算法相比于其他算法的优劣，本文将改进后的 YOLOv4 算法与 Faster RCNN, YOLOv3, EfficientDet, YOLOv4 以及文献 6 中改进的 Faster RCNN 算法，文献 7 中改进的 YOLOv3 算法进行了比较。表 4 给出了同一实验环境与测试数据集下，不同算法的实验结果。由表 4 可知，相比于传统的 Faster RCNN, YOLOv3, EfficientDet 和 YOLOv4 算法，改进后的 YOLOv4 算法的检测精度与速度拥有一定的优势，且模型的权重文件较小。而相比于文献[6]中改进后的 Faster RCNN 算法，虽然改进后的 Faster RCNN 算法的检测精度略高于本文的算法，但是其检测速度远不及本文的算法，需进一步进行优化，且模型的权重大于本文的算法。而本文中改进后的 YOLOv4 算法在检测精度与速度，权重大小都要优于文献 7 中改进后的 YOLOv3 算法。综合检测精度，检测速度与权重文件大小而言，改进后的 YOLOv4 算法拥有更优秀的综合性能，更适用于部署在嵌入式设备中实现鸟巢检测任务。

表 4 不同算法对比

模型	AP/%	FPS	权重/MB
Faster RCNN	84.20	15.3	521
YOLOv3	82.87	30.9	234
EfficientDet	82.47	35.5	128
YOLOv4	85.02	36.4	244
文献[6]	88.30	18.9	207
文献[7]	86.37	41.3	234
本文模型	87.35	49.3	121

### 3.6 实验效果测试

图7(a)为原始YOLOv4算法的检测效果,图7(b)为改进后YOLOv4算法的检测效果。可见改进后的YOLOv4算法改善了算法对鸟巢小目标漏检的情况,并且检测置信度值都得到明显提高。

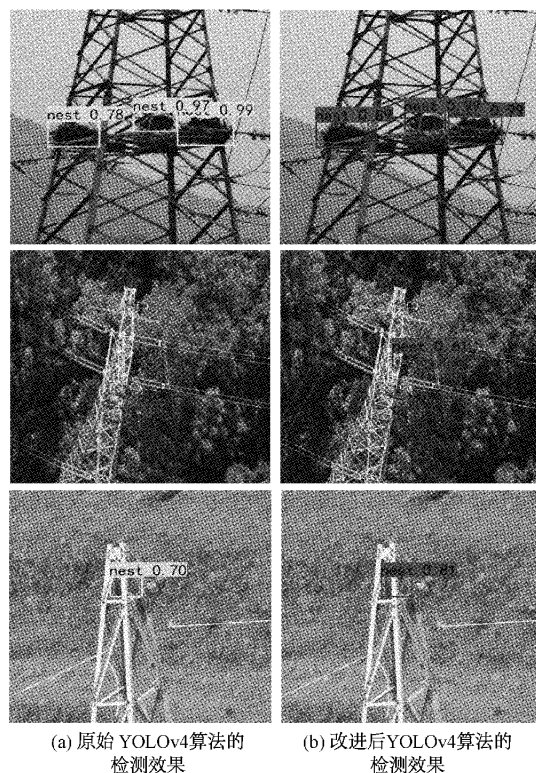


图7 算法改进前后检测效果对比

## 4 结 论

针对现有算法对高压塔上鸟巢检测存在的问题,本文提出了一种改进的YOLOv4鸟巢目标检测算法。为了减少算法的参数且提高检测速度,使用轻量级Mobilenetv2网络作为算法的主干网络;同时在保证轻量化的前提下,在Mobilenetv2网络的逆残差网络嵌入Coordinate Attention注意力,提高网络特征提取能力。对PANet网络进行改进,提高对小目标鸟巢的检测能力。最后,使用Focal Loss函数优化损失函数,进一步提高对小目标鸟巢的检测能力。实验结果表明,与YOLOv4算法相比,改进后的YOLOv4算法的检测精度与速度都得到了提升,且模型权重文件更小,更适合用于部署在无人机等嵌入式设备中实现鸟巢检测。但本文研究内容仍有改进的空间,例如,本文所提出的算法都运用了先验框的思想,目前存在着无先验框的目标检测算法,该算法舍弃了非极大值抑制的思想,拥有较高的检测速度,因此可以尝试借鉴该思想,进一步对算法进行优化。

## 参考文献

- [1] QIU Z, ZHU X, LIAO C, et al. Detection of bird species related to transmission line faults based on lightweight convolutional neural network [J]. IET Generation, Transmission & Distribution, 2022, 16(5): 869-881.
- [2] 徐晶,韩军,童志刚,等.一种无人机图像的铁塔上鸟巢检测方法[J].计算机工程与应用,2017,53(6):231-235.
- [3] LU J, XU X, LI X, et al. Detection of bird's nest in high power lines in the vicinity of remote campus based on combination features and cascade classifier [J]. IEEE Access, 2018, 6: 39063-39071.
- [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] LI F, XIN J, CHEN T, et al. An automatic detection method of bird's nest on transmission line tower based on Faster\_RCNN[J]. IEEE Access, 2020, 8: 164214-164221.
- [6] 王纪武,罗海保,鱼鹏飞,等.基于Faster R-CNN的多尺度高压塔鸟巢检测[J].北京交通大学学报,2019,43(5):37-43.
- [7] 丁建,黄陆明,朱迪锋,等.面向高空塔架鸟巢检测的双尺度YOLOv3网络学习[J].西安理工大学学报,2021,37(2):253-260.
- [8] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection [J]. ArXiv Preprint, 2020, ArXiv: 2004.10934.
- [9] TAN L, LV X, LIAN X, et al. YOLOv4\_Drone: UAV image target detection based on an improved YOLOv4 algorithm [J]. Computers & Electrical Engineering, 2021, 93: 107261.
- [10] LIU H, FAN K, OUYANG Q, et al. Real-time small drones detection based on pruned YOLOv4 [J]. Sensors, 2021, 21(10): 3374.
- [11] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390-391.
- [12] 张明路,郭策,吕晓玲,等.改进的轻量化YOLOv4用于电子元器件检测[J].电子测量与仪器学报,2021,35(10):17-23.
- [13] ZHANG M, ZHANG Y, ZHOU M, et al. Application of lightweight convolutional neural network for damage detection of conveyor belt [J].

- Applied Sciences, 2021, 11(16): 7282.
- [14] 解尧婷,张丕状. 基于改进的 YOLOv4 输电线路小目标检测[J]. 国外电子测量技术,2021,40(2):47-51.
- [15] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [16] 王亚朝,赵伟,徐海洋,等. 基于轻量型卷积神经网络的非固定场景天气识别算法[J]. 电子测量技术,2019, 42(17):152-156.
- [17] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [18] XIE C, ZHU H, FEI Y. Deep coordinate attention network for single image super-resolution[J]. IET Image Processing, 2022, 16(1): 273-284.
- [19] ZHU Q, ZHENG H, WANG Y, et al. Study on the evaluation method of sound phase cloud maps based on an improved YOLOv4 algorithm[J]. Sensors, 2020, 20(15): 4314.
- [20] LI J, YAN D, LUAN K, et al. Deep Learning-Based Bird's Nest Detection on Transmission Lines Using UAV Imagery[J]. Applied Sciences, 2020, 10(18): 6147.

### 作者简介

谢国波,博士,教授,主要研究方向为深度学习,遥感大数据等。

郑晓锋,硕士研究生,主要研究方向为深度学习,目标检测等。

林志毅(通信作者),博士,讲师,主要研究方向为深度学习,生物信息等。

E-mail: lzy291@gdut.edu.cn

林立,硕士研究生,主要研究方向为深度学习,目标检测等。

文刚,硕士,研究员,主要研究方向为深度学习,数据分析等。