

# 基于 YOLOv5 的电力线和杆塔实时检测算法研究

叶树芬<sup>1</sup>, 施振华<sup>2</sup>, 苏成悦<sup>1</sup>, 梁立翀<sup>1</sup>, 黄海润<sup>1</sup>, 吴家华<sup>3</sup>

(1. 广东工业大学 物理与光电工程学院, 广州 510006; 2. 广州杰超科技有限公司, 广州 510006;  
3. 广东电网有限责任公司佛山供电局, 广东 佛山 528000)

**摘要:** 针对当前电力线路检测中存在深度学习网络参数量大、计算复杂度高问题; 在 YOLOv5 的基础上提出一种电力线和杆塔的实时检测算法; 通过减少 Bottleneck 数量来简化特征提取层网络结构, 使用深度可分离卷积技术实现模型计算量的降低; 分析电力线目标框筛选机制, 改进 (non-maximum suppression) NMS 算法, 提升模型目标检测精度; 实验结果表明, 对 Bottleneck 的改进在识别精度有所提高的情况下能有效降低模型的参数量, 模型检测准确率和召回率分别达到 94% 与 95%, 体积压缩了 20.7%, 在 Jetson Nano 嵌入式平台上检测速度达到 17.2 fps, 对两类电力线路目标检测达到较高的识别率和实时性, 对无人机电力巡检导航有较好的参考价值。

**关键词:** 电力线路检测; Bottleneck; 深度可分离卷积; NMS; 嵌入式平台

## Research on Real-time Detection Algorithm of Power Line and Pole Tower Based on YOLOv5

YE Shufen<sup>1</sup>, SHI Zhenhua<sup>2</sup>, SU Chengyue<sup>1</sup>, LIANG Lichong<sup>1</sup>, HUANG Hairun<sup>1</sup>, GUAN Jiahua<sup>3</sup>

(1. School of Physics and Optoelectronic Engineering, Guangdong University of Technology, Guangzhou 510006, China;  
2. Guangzhou Jiechao Technology Co., Ltd., Guangzhou 510006, China;  
3. Foshan Power Supply Bureau of Guangdong Power Grid Co., Ltd., Foshan 528000, China)

**Abstract:** Aiming at the problems of large number of parameters and high computational complexity of deep learning network in current power line detection; Based on YOLOv5, a real-time detection algorithm for power lines and towers is proposed. The network structure of feature extraction layer is simplified by reducing the number of Bottleneck, and the depth-separable convolution technique is used to realize the reduction of model computational amount. The mechanism of power line target box screening is analyzed, and the non-maximum suppression (NMS) algorithm is used to improve the model target detection accuracy. The experimental results show that the improvement of Bottleneck can effectively reduce the number of parameters of model when the recognition accuracy increases. The model detection accuracy and recall rate reach 94% and 95%, respectively, and the volume is compressed by 20.7%. The detection speed on Jetson Nano embedded platform reaches 17.2 FPS. The detection of two kinds of power line targets achieves high the recognition rate and real-time performance, which has a good reference value for UAV power inspection and navigation.

**Keywords:** transmission line detection; Bottleneck; depth separable convolution; NMS; embedded platform

## 0 引言

电力线路巡检是电网稳定安全运行的重要环节, 人工操作的无人机巡检替代传统人工巡检可节约大量的时间和人力成本<sup>[1]</sup>。智能无人机自动巡检成为当前电力巡检的研究热点, 导航是智能无人机巡检的关键<sup>[2]</sup>。由于导航精度不够, 巡检无人机飞偏的概率较大<sup>[3]</sup>, 对电力线和杆塔准确且高效的识别是无人机沿电力线路实时巡检的基础<sup>[4]</sup>。

电力线识别的传统数字图像处理算法需要预先对目标特征进行数值化定义, 通过某种固定特征进行匹配。Zhu 等<sup>[5]</sup>通过 Radon 变换以及线段平行的约束条件识别电力线。骆顺平等<sup>[6]</sup>改进 Ratio 算子并结合轮廓特征和利用 Hough 变换的直线编组拟合算法识别出电力线。数字图像处理算

法对与电力线形状相似的道路和房屋等有边缘特征的图像, 以及多样性场景存在识别干扰, 容易产生漏检和误检<sup>[7]</sup>。

检测电力线路的深度学习算法主要有电力杆塔识别的目标检测算法和电力线识别的语义分割算法。目标检测双阶段算法以 R-CNN 系列<sup>[8-10]</sup>为代表, 该算法产生候选区域再分类和回归。单阶段算法直接预测目标的类别与位置, 算法速度更快, 常见有 YOLO<sup>[11-14]</sup>, SSD<sup>[15]</sup> 等。孙乐杨等<sup>[16]</sup>增加 YOLOv5 模型特征层尺度以及通过切割高分辨图像后进行识别, 提升遥感图像里小目标电力塔检测效果。杨知等<sup>[17]</sup>提出结合 YOLOv2 和 VGG 模型级联的目标检测算法来识别高分辨率遥感影像中的电力塔。有研究人员针对电力线的特征提出语义分割算法, 但这类算法复杂度

收稿日期: 2022-08-08; 修回日期: 2022-08-29。

作者简介: 叶树芬(1998-), 男, 广东东莞人, 硕士研究生, 工程师, 主要从事电子信息技术、深度学习方向的研究。

通讯作者: 苏成悦(1961-), 男, 湖南长沙人, 博士研究生, 教授, 主要从事应用物理、应用电子方向的研究。

引用格式: 叶树芬, 施振华, 苏成悦, 等. 基于 YOLOv5 的电力线和杆塔实时检测算法研究[J]. 计算机测量与控制, 2022, 30(11): 77-84.

高<sup>[18-19]</sup>。深度学习算法泛化能力更强,但算力和内存空间是模型在嵌入式平台应用的瓶颈,复杂模型实时性较差,已知的算法还不能同时实现电力杆塔和电力线两类目标的检测任务。

本文在 YOLOv5 网络模型基础上,采用减少特征提取层 Bottleneck 数量和深度可分离卷积,以减少模型整体参数量和计算量;通过改进 NMS 算法,提升电力线和电力杆塔的检测精度;部署于嵌入式平台验证表明本算法能兼顾检测精度和速度,具有较好的实用性。

## 1 YOLOv5 算法

YOLOv5 包含 5 个版本的模型,它们的区别在于模型的深度和宽度不同,考虑在嵌入式平台上使用,本文选用模型体积较小,精确度适中的 YOLOv5s 作为基础模型,其由输入端、Backbone、Neck 和 Prediction 四个部分组成。

### 1.1 输入端

输入端部分包含对图像的预处理和数据增强的方式,自适应锚框计算和自适应图片缩放。

自适应锚框计算是 YOLOv5s 网络在训练之前,会对锚框的长宽进行初始设定,训练过程中,网络会根据初始锚框大小输出预测框,和真实框比较计算出两者的差距,再反向更新迭代网络参数。YOLOv5s 会在训练和参数迭代的过程中,根据不同的数据集,自适应计算出最佳的锚框值。

自适应图片缩放是 YOLOv5s 网络会自动采用相同比例将图片分辨率缩放到长宽均为 32 的倍数,不满足比例的一边进行填充,在保证不失真的情况下,减少计算量,加快目标检测速度。

### 1.2 Backbone

Backbone 包含 C3 层和 SPPF (spatial pyramid pooling, 快速空间金字塔池化模块) 结构,用于图像的特征提取,其网络结构如图 1 所示。



图 1 Backbone 结构图

YOLOv5s 新版中 C3 层是采用 CSP (cross stage partial connections, 跨阶段局部网络) 架构,包含三个标准卷积层以及多个 Bottleneck 模块,替换了原先设计应用于 Backbone 中的 CSP1\_X 和 Neck 中的 CSP2\_X 两种 CSP 结构,可以增强网络的学习能力,降低模型的尺寸,同时保持目标检测的精度。

SPPF 模块采用多种尺寸的最大池化方式进行多尺度特征的融合,池化后的输出将作为下一池化的输入,速度比 SPP (spatial pyramid pooling, 空间金字塔池化模块) 更快。

### 1.3 Neck

Neck 部分包含 FPN (feature pyramid networks, 特征金字塔网络) 和 PAN (path aggregation network, 路径聚合网络) 的组合结构,其中 FPN 是通过对高低层特征进行融合,可以提高细长型电力线等小目标检测效果。PAN 是

自底向上方向的增强,使得深层的特征图可以获得浅层丰富的细节信息,提高大目标的检测效果。

### 1.4 Prediction

YOLOv5s 用到的损失函数为定位损失 (box\_loss)、置信度损失 (object\_loss) 和分类损失 (class\_loss) 三部分的加权和。YOLOv5s 的定位损失使用的是 CIoU\_loss<sup>[20]</sup>,其计算过程公式如下:

$$\text{CIoU} = \text{IoU} - \frac{d_c^2}{d_s^2} - \alpha\nu \quad (1)$$

其中:  $d_c$  为预测框和真实框中心点间的距离,  $d_s$  为预测框和真实框最小外接矩形的对角线距离, IoU 为预测框和真实框的交并比,其计算公式如式 (2) 所示,  $\alpha$  为平衡参数,  $\nu$  用于衡量宽高比的倾斜角度,计算公式如式 (3) 和式 (4) 所示:

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

$$\alpha = \frac{\nu}{1 - \text{IoU} + \nu} \quad (3)$$

$$\nu = \frac{4}{\pi^2} \left( \arctan \frac{w^{gr}}{h^{gr}} - \arctan \frac{w}{h} \right)^2 \quad (4)$$

CIoU\_Loss 计算公式如下:

$$\text{CIoU\_Loss} = 1 - \text{CIoU} \quad (5)$$

## 2 数据集制备与分析

为了验证本文提出的电力线和电力杆塔算法识别的效果,且现有公开的电网数据集较少,实验数据集采用自建的方式。

### 2.1 数据采集

通过无人机实地拍摄获取,拍摄地点在广东省佛山市南海区部分备用电力配网线路,再将这些视频数据进行取帧得到实验所需的图片。经过筛选整理后得到 2 000 张图片,并统一分辨率为  $1\,280 \times 720$ ,方便后续模型的训练。数据集示例如图 2 所示,背景包含草地、公路、建筑和水塘等,每张图片都包含多个复杂背景下电力线目标以及多角度的电力线和电力杆塔目标,根据深度学习领域常用的 8:1:1 比例随机将图片分为训练集、验证集和测试集。

### 2.2 标注

数据集使用在线标注工具对图片进行标注,标注的结果利用格式转换脚本转换为 yolo\_txt 格式进行保存,每个标注 txt 文件存放一个图片的目标信息,文件的每一行存放一个目标的信息,包括标注标签的种类 class,目标框的中点  $x$  坐标  $x\_center$ ,中点  $y$  坐标  $y\_center$ ,宽度 width 以及高度 height。标注标签的种类包含电力杆塔 tower 和电力线 line。

为了通过目标检测的方式对电力线进行识别,本文对电力线采用分段标注的方式,如图 3 所示,将图片中电力线逐一分段,每一小段打上一个电力线的标签,这种方式相比较于一个框标注一整条电力线,可以减少大量不必要的背景信息引入,同时分段小框还能用于在电力线走向的



图 2 数据集示例

判别上, 将连续的分段小框中心点连接起来, 可以形成一条带方向的电力线。



图 3 标注示意图

### 2.3 样本分析

自制电力线数据集过程中, 电力线的长度和不同倾斜角度都会影响分段真实框的大小, 采用人工分段不可避免存在真实框长宽不完全统一的情况, 这种偏差会在一定程度上使得模型训练出来的目标框和真实框大小也会有存在偏差, 并且一个分段电力线真实框尺寸本身也比较小。当且仅当目标框同样大小且位置高度重合, 目标框和真实框的 IoU 才可能较高, 而在其他大部分情况下, 由于上述标注误差和电力线目标较小, 目标框和真实框是较难高度重合, IoU 偏小。对于常见长度和角度的电力线, 在标注之前, 提前固定真实框长度和宽度在一个区间内, 尽可能减少引入较大的标注误差。

## 3 检测算法改进

本文主要从两个方面改进检测算法: (1) 优化特征提取层网络结构, 加快算法移植在嵌入式平台上的部署推理; (2) 结合实际电力线识别的场景, 改进 NMS 算法, 提高识别效果。

### 3.1 轻量化特征提取网络

对特征提取网络轻量化的优化工作主要分为修改各 C3 层中 Bottleneck 数量比例和简化特征提取网络中参数量大的卷积结构两个部分, 在不损失过多模型精度前提下, 减少模型的参数量和计算复杂度。

#### 3.1.1 修改 Bottleneck 数量比例

YOLOv5s 特征提取网络的 C3 层主要是对残差特征进行学习, 比 CSP 结构更简单和更快, 但是 C3 层采用多路分离卷积并且通道数较高, 容易占用较多的缓存空间, 降低算法运行速度, C3 层的工作逻辑如下:

- 1) 原始输入  $i$  进入一分支执行标准卷积模块, 输出结果赋值给  $a$ ;
- 2) 将  $a$  传入由 Bottleneck 组成的模块, 使用 Bottleneck 的数量在定义网络的时候确定, 输出结果赋值给  $b$ ;
- 3) 原始输入  $i$  进入另一分支只执行标准卷积模块, 输出结果赋值给  $c$ ;
- 4) 特征拼接  $b$  和  $c$ , 输出结果为  $d$ ;
- 5) 对结果  $d$  执行标准卷积操作, 最后返回输出。

其中 Bottleneck 结构如图 4 所示, 通过  $1 \times 1$  的标准卷积将输入特征图的通道数减小一半, 再通过  $3 \times 3$  的标准卷积将通道数扩大一倍, 获得需要的特征并且通道数不发生改变, 还有一个 shortcut 参数控制是否进行残差连接, 在特征提取层默认为 True, 最后使用 add 操作进行特征融合。

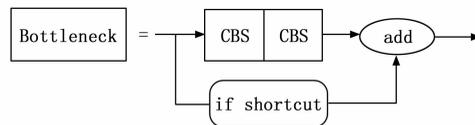


图 4 Bottleneck 结构图

神经网络中模块的宽度、深度和数量是可以根据不同数据集需要进行定制化的设计, 减少卷积的操作可以有效的减少网络的计算量, 提升网络的目标检测速度。

ResNet 网络中跨阶段计算分布的原始设计在很大程度上是经验值, 借鉴 ConvNeXt 网络修改方法, 将 layer0 到 layer3 层中 block 数量从 ResNet-50 的 (3, 4, 6, 3) 改为 (3, 3, 9, 3) 后降低计算复杂度, 提高识别精度<sup>[21]</sup>。YOLOv5s 的特征提取网络中包含 4 个 C3 层, 分别位于第二到第五层下采样卷积层的下一层, 本文将各 C3 层中 Bottleneck 数量从 (3, 6, 9, 3) 改为 (1, 1, 3, 1), 数量比例也是 1: 1: 3: 1, 修改后特征提取网络中各个 C3 层的结构如图 5 所示。

从图 4 和图 5 可以看出, 减少各个 C3 层中 Bottleneck 模块的数量, 可以减少 Bottleneck 中卷积和特征融合的参数量和计算量, 精简网络结构, 减少缓存的使用, 提高模型的目标检测速度。

#### 3.1.2 引入深度可分离卷积

在特征提取的过程中, 随着网络深度的增加, 特征图的通道数也在不断增加, 使用传统卷积对深层特征图提取特征时, 必然会有较大的参数量产生, 导致算法计算运行速度较慢。本文采用深度可分离卷积<sup>[22]</sup>来替换 YOLOv5s 特征提取网络中最后一个卷积层结构, 这个卷积的参数量是最大的, 可以有效减少参数量和计算量, 并且这样不会影响上层卷积的特征提取效果。

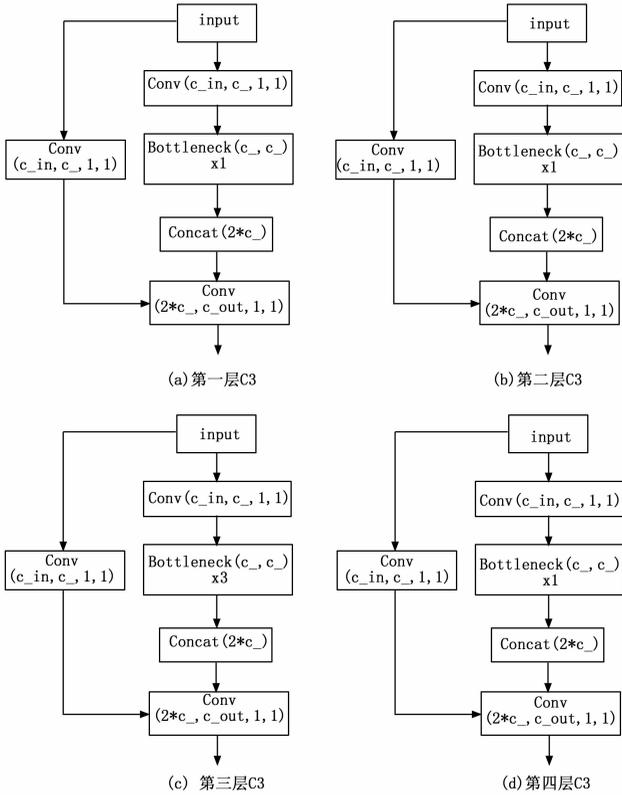


图 5 修改后各 C3 层结构

深度可分离卷积是将一个标准卷积分解成一个深度卷积 (depthwise convolution) 和一个点卷积 (pointwise convolution), 如图 6 所示, 拆分了卷积中通道和空间相关性的联合映射, 可以实现与传统卷积同样的效果且不会损失过多的识别精度。

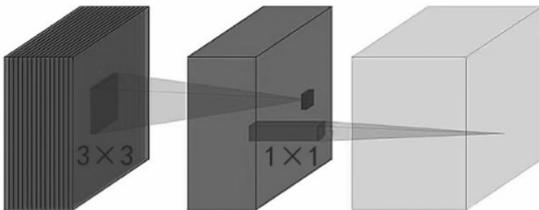


图 6 深度可分离卷积

假设输入特征图大小为  $W \times H \times C$ , 标准卷积核大小为  $K \times K \times C$ , 个数为  $N$  个, 输出特征图大小为  $W \times H \times N$ , 那么该标准卷积的参数量为卷积核大小与输入输出特征图通道数的乘积, 计算过程如式 (6) 所示, 计算量的计算过程如式 (7) 所示:

$$Para_{sc} = K \times K \times C \times N \quad (6)$$

$$O_{sc} = W \times H \times C \times K \times K \times N \quad (7)$$

标准卷积核拆分之后得到深度卷积核大小为  $K \times K \times 1$ , 个数与上一层通道数相同为  $C$  个, 是在二维平面上进行卷积, 点卷积核大小为  $1 \times 1 \times C$ , 个数为  $N$  个, 得到该深度可分离卷积的参数量为深度和点卷积核参数量的相加, 计算过程如式 (8) 所示, 同理计算量也为两者相加, 如式 (9) 所示:

$$Para_{dsc} = K \times K \times C + C \times N \quad (8)$$

$$O_{dsc} = W \times H \times C \times K \times K + W \times H \times C \times N \quad (9)$$

因此, 该深度可分离卷积与标准卷积的参数量和计算量压缩比分别为:

$$F_{Para} = \frac{K \times K \times C + C \times N}{K \times K \times C \times N} = \frac{1}{N} + \frac{1}{K^2} \quad (10)$$

$$F_O = \frac{W \times H \times C \times K \times K + W \times H \times C \times N}{W \times H \times C \times K \times K \times N} = \frac{1}{N} + \frac{1}{K^2} \quad (11)$$

可以看出, 对于一个  $3 \times 3$  大小的标准卷积, 在输出通道数为 4 的情况下, 经过深度可分离卷积的替换, 参数量和计算量均可以减少到原来的 36.1%, 当输出通道数远远大于卷积核大小乘积, 一个标准卷积的计算量压缩比可以大约为  $1/9$ 。本文针对的卷积结构为输入特征图通道数为 256 个, 输出特征图通道数为 512 个的卷积层, 是特征提取网络中下采样的最后一层卷积, 替换为深度可分离卷积后, 参数量有较大下降, 模型计算复杂度降低。

### 3.1.3 特征提取网络搭建

本文经过改进后 YOLOv5s 特征提取网络结构如表 1 所示, 其中网络的深度超参数设置为 0.33, 宽度超参数设置为 0.50。

表 1 特征提取网络结构

层名称	输出特征图通道数	参数配置
卷积层 Conv1	32	size=6, stride=2, padding=2
卷积层 Conv2	64	size=3, stride=2
C3 层 1	64	Bottleneck_number=1
卷积层 Conv3	128	size=3, stride=2
C3 层 2	128	Bottleneck_number=1
卷积层 Conv4	256	size=3, stride=2
C3 层 3	256	Bottleneck_number=3
卷积层 DWConv5	512	size=3, stride=2
C3 层 4	512	Bottleneck_number=1
SPPF 层	512	size=5, stride=1

从表 1 可见, 本文算法特征提取网络包含 5 个卷积层, 4 个 C3 层和 1 个 SPPF 模块, 第 5 层卷积层为深度可分离卷积, 卷积核大小  $3 \times 3$ , 卷积的步长为 2, 最后经过特征提取后得到通道数为 512 的特征图, 将输入到 Neck 和 Prediction 用于后续目标预测。

### 3.2 改进 NMS 算法

当一幅待测试的图像输入到训练好的目标检测网络, 每个目标网格区域都会产生多个候选框, 这些候选框之间会有重叠部分, NMS 是一种非极大值抑制算法, 用于将冗余的候选框去除, 尽可能找到最佳的预测框, 即一个目标对应一个预测框, 原算法步骤如下: (1) 将候选框集合根据置信度得分进行排序; (2) 选择置信度最高的候选框添加到最终输出列表中, 将其从候选框列表中删除; (3) 计算所有候选框的面积, 并计算置信度最高的候选框与其它



它们的特征向量和标签，对应地加权相加，得到新样本的特征向量和标签，新样本融合了两张原始图片的信息，同样可以增强模型的抗干扰能力。

### 4.2 训练参数及策略

在整个训练过程中，训练参数设置如下：最大训练迭代次数 (epoch) 为 100, batch size 为 16, 初始学习率为 0.01, 动量 (momentum) 为 0.937, 权重衰减 (decay) 为 0.000 5, 在一定程度上降低模型出现过拟合的可能。

训练策略用 warmup 来优化训练的效果<sup>[26]</sup>, 在模型刚开始训练时, 模型的权重 (weights) 是随机初始化的, 若选择预先设定好的学习率, 可能会比较大, 导致模型出现振荡的不稳定情况, 而选择 warmup 预热学习率的方式, 将前三个 epoch 设置为预热学习轮数, 预热学习初始动量设置为 0.8, 学习率较小, 模型可以降低训练的误差, 慢慢趋于稳定, 等模型相对稳定后再使用预先设置的学习率 0.01 进行训练, 使得模型收敛速度变得更快, 模型效果更佳。

本文在多次调整训练参数后得到如图 10 所示的损失下降曲线, 其中横坐标表示训练 epoch, 纵坐标为损失值, 从图中可以看到, YOLOv5s 损失函数中三个方面的损失整体均趋于收敛, 通过数据表明, 模型达到较好的拟合效果, 呈现稳定状态, 可以用于接下来进行验证和测试。

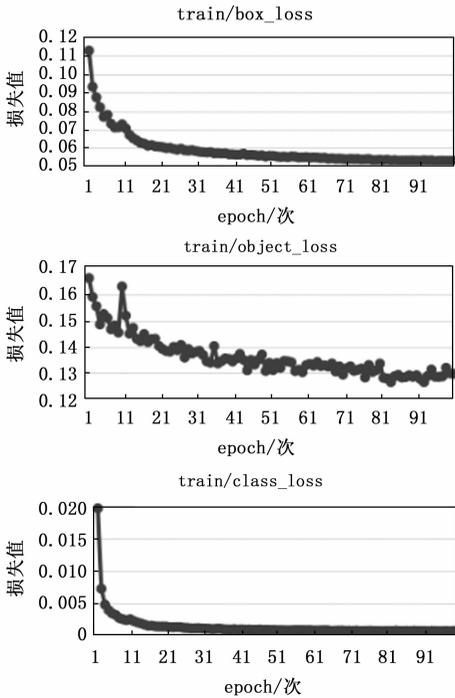


图 10 训练损失下降曲线

### 4.3 评价指标

在评估目标检测算法识别效果的时候, 通常根据数据集特性, 应用场景侧重等不同, 合理选择使用一些特定的量化指标进行评判。

本文在电力线和电力杆塔检测任务当中, 对算法的识别精确度, 采用准确率 (P, precision) 和召回率 (R, re-

call) 来衡量不同算法的分类和检测性能, 除此之外, 针对无人机携带的嵌入式平台算力有限的算法效率优化, 采用模型参数量, 模型体积和每秒帧数 (FPS, frames per second) 作为改进算法的衡量指标。P 和 R 的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (13)$$

$$R = \frac{TP}{TP + FN} \quad (14)$$

其中: TP 表示算法检测为电力线和电力杆塔目标当中判断正确的数目, FP 为判断错误的数目, FN 表示算法没有检测出真实电力线和电力杆塔目标的数目。在判断目标是 TP 还是 FP 时, 首先是网络产生大量的预测框, 通过 NMS 算法筛选出符合要求阈值的预测框, 再分别计算真实框和这些预测框的 IoU, 此时如果存在符合 IoU 条件的预测框, 那么 IoU 最大的预测框就确定为这个真实框的唯一正确判断结果, 最后没有判断正确的预测框, 均确定为错误判断结果。

从实际工程应用的角度来看, 分段标注的目的在于能将分段的目标框用于无人机沿着线飞行的指引, 而不是将分段的目标框和真实框一一对应起来, 因此本文将上述的 IoU 条件设置为 0.3, 充分考虑到手工标注的误差存在和工程应用的需求, 允许目标框位置上存在一定的偏差, 更能反映不同目标检测算法的检测效果。

### 4.4 实验设计与结果分析

本节将从特征提取网络 Bottleneck 数量比例选取, 不同目标检测算法对比和嵌入式平台部署三个方面设计实验并分析改进对实验结果的影响, 其中对于算法测试过程中 NMS 算法用到的参数设置如下: IoU 阈值为 0.01, 置信度阈值为 0.25。

#### 4.4.1 Bottleneck 数量比例改进对比

在特征提取网络轻量化的改进中, 本文针对各 C3 层中 Bottleneck 数量进行了修改。为了对比使用不同 Bottleneck 数量对算法的影响, 在 YOLOv5s 原始特征提取网络上采用 4 种不同的数量比进行实验, 得到对应的准确率、召回率和参数量指标, 最终算法在测试集上结果如表 3 所示。

表 3 不同 Bottleneck 比例效果对比

比例	P/%	R/%	参数量/MB
(3,6,9,3)	96.6	88.9	6.69
(3,3,9,3)	96.6	88.3	6.65
(2,2,6,2)	96.6	89.1	6.59
(1,1,3,1)	97.0	89.4	6.34

原始 YOLOv5s 特征提取网络中 Bottleneck 数量比为 (3, 6, 9, 3), 从表 3 可以看出, 其准确率为 96.6%, 召回率为 88.9%, 参数量为 6.69 MB, 不断减少各层 Bottleneck 数量, 参数量不断下降, 但是对准确率和召回率的影响并不明显, 最后当数量比为 (1, 1, 3, 1) 时, 参数量为 6.34 MB, 减少了 5.2%, 准确率和召回率均有少量提升。

从实验结果可以看到, Bottleneck 数量比例改进是可行的, 在识别精度有所提高的情况下能有效降低模型的参数量。

#### 4.4.2 目标检测算法对比

为了验证本文算法的有效性, 将本文算法与双阶段算法 Faster R-CNN, 单阶段算法 SSD, 轻量化的 YOLOv4-tiny<sup>[27]</sup> 和 YOLOv5s 进行对比, 计算各个算法的准确率、召回率、模型体积和 FPS 指标, 结果如表 4 所示。

表 4 不同算法性能和参数对比

算法	P/%	R/%	模型体积/MB	FPS
Faster R-CNN	16.9	56.7	1042.0	6.9
SSD	58.8	97.7	91.1	33.2
YOLOv4-tiny	64.6	78.7	22.4	66.7
YOLOv5s	96.6	88.9	14.0	136.1
本算法	94.0	95.0	11.1	153.8

表 4 的数据均在自制电力线路数据集的训练集上进行训练, 通过测试集得出的结果。从表中可以看出, YOLOv5s 获得了 96.6% 的准确率, 88.9% 的召回率, 两者均优于 Faster R-CNN 和 YOLOv4-tiny, 相比于 SSD 的召回率虽然低了 8.8%, 但是高了 SSD 的准确率 37.8%, 再看模型体积和 FPS, YOLOv5s 相比于其他算法具有绝对的优势, 模型体积为 14.0 MB, 仅为 SSD 模型体积的 15.4%, FPS 为 YOLOv4-tiny 的两倍以上。

本文算法将 YOLOv5s 作为基础算法, 在特征提取网络中引入深度可分离卷积并且减少 Bottleneck 数量, 同时也改进了 NMS 算法, 优化目标框的筛选机制, 从表中可以看出本算法模型体积进一步压缩 20.7%, 有效降低了算法的参数量和计算量, 在 GTX1080Ti 上 FPS 提高了 17.7, 虽然准确率有 2.6% 的下降, 但是召回率有 6.1% 的提升, 弥补了 YOLOv5s 在召回率上的不足, 在实际电力线路检测当中, 出于防止无人机巡检事故发生的角度, 更加看重召回率。从对比实验可以看到, 本文提出的算法在平衡识别精度和识别速度上更具有优势。

#### 4.4.3 嵌入式平台实验

为了进一步验证本算法在实际场景中的实用性, 将目标检测算法移植部署到嵌入式平台 Jetson Nano 上。

Jetson Nano 是一款嵌入式系统主板, 体积较小, 适合搭载在无人机上。硬件系统其 CPU 为 4 核的 Cortex-A57, GPU 为基于 Maxwell™ 架构的 128 核集成 CPU, 有 64 位的 4 GB LPDDR4 内存, 在算力方面具有 472 GFLOPs, 可以满足小型目标检测算法的实时推理过程; 软件系统本文安装的是 JetPack 4.4.1 SDK。

目标检测模型在训练阶段, 更加注重识别精度的提升, 但在模型部署阶段, 部署端平台的性能会有较大的下降, 因此在算法实际部署到嵌入式平台这个过程中, 一般都需要加速优化算法的推理过程, 包括降低计算精度、运算合并以及内存使用的优化等。本实验利用 TensorRT 加速推理, 通过自制的测试集进行测试, 同时还设置了与

YOLOv5s 的对比实验, 对比部署后的算法模型体积和每一秒检测图片的数量, 实验结果如表 5 所示, 其中精度设置为半精度浮点型 (FP16), 输入图片大小为  $640 \times 640$ 。

表 5 嵌入式平台算法性能对比

算法	模型体积/MB	FPS
YOLOv5s	18.8	15.6
本算法	14.6	17.2

从表 5 可以看到, 本文改进算法, 模型体积为 14.6 MB, 下降了 22.3%, 这意味着嵌入式平台可以减少这部分模型的计算, 能有更多算力完成其他功能。在识别速度上, FPS 达到 17.2, 在原算法基础上增加了 1.6, 在运算资源有限的情况下基本满足实时检测的需求, 测试效果如图 11 所示, 电力线和电力杆塔能同时准确的识别出来。



图 11 Jetson Nano 测试效果图

为了进一步验证使用目标检测识别电力线来使无人机沿线路巡检的效果, 本文将电力线目标框的中心点连接起来, 如图 12 所示, 给无人机的飞行提供电力线的方向, 从而可以更好完成巡检的工作。



图 12 测试验证图

实验结果表明, 本文在模型结构上压缩计算量和参数量的方式, 能有效优化算法在嵌入式平台上推理的速度, 提高实时性。

## 5 结束语

针对电力线和电力杆塔检测任务, 为了解决深度学习算法计算量大的问题, 本文构建并标注了一个新的数据集, 使用 YOLOv5s 作为基础目标检测算法, 通过使用数量更少

的 Bottleneck 完成对输入目标的特征提取,采用深度可分离卷积技术降低卷积计算的参数量,有效降低算法的计算复杂度,优化 NMS 算法,进一步提升了目标检测的召回率。通过多组对比实验表明,改进的 YOLOv5s 算法比原始算法在检测召回率、模型体积和 FPS 指标均有一定的提升,能在嵌入式平台上同时完成电力线和电力杆塔的检测,具有较好的实时性。现有数据集所能覆盖的场景有限,接下来需要获取更多原始电力线路样本数据,从而进一步验证算法的鲁棒性。

#### 参考文献:

- [1] MATIKAINEN L, LEHTOMAKI M, AHOKAS E, et al. Remote sensing methods for power line corridor surveys [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 119: 10-31.
- [2] GAO F, WANG S, WU Y, et al. A novel inverse method for automatic UAV line patrolling with magnetic sensors [C] // IEEE International Symposium on Electromagnetic Compatibility and 2018 IEEE Asia-Pacific Symposium on Electromagnetic Compatibility (EMC/APEMC). Piscataway: IEEE, 2018: 481-485.
- [3] AVIZONIS P, BARRON B. Low cost wire detection system [C] // Proceedings of the 18th Digital Avionics Systems Conference. Piscataway: IEEE, 1999: 3. C. 3.
- [4] MEJIAS L, CORREA J F, MONDRAGON I, et al. COLIBRI: A vision-guided UAV for surveillance and visual inspection [J]. IEEE International Conference on Robotics and Automation, 2007.
- [5] ZHU L L, CAO W R, HAN J D, et al. A double-side filter based power line recognition method for UAV vision system [C] // Proceedings of the 2013 IEEE International Conference on Robotics and Biomimetics. Piscataway: IEEE, 2013: 2655-2660.
- [6] 骆顾平, 陈静, 苏成悦, 等. 复杂地物背景下的电力线识别算法 [J]. 应用科技, 2021, 48 (5): 23-28, 34.
- [7] 李晓峰, 胡川, 尹洪, 等. 基于机载 LiDAR 点云的电力线自动提取 [J]. 计算机测量与控制, 2018, 26 (7): 220-224.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation [J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 38 (1): 142-158.
- [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 r-cnn: Towards real-time object detection with region proposal networks [J]. Advances in neural information processing systems, 2015, 28: 91-99.
- [11] 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, 2016: 779-788.
- [12] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.
- [13] REDMON J, FARHADI A. Yolov3: an incremental improvement [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018, 87 (8): 101-104.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: optimal speed and accuracy of object detection [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020, 57 (5): 9-12.
- [15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C] // European conference on computer vision. Springer, Cham, 2016: 21-37.
- [16] 孙乐杨, 凌振宝, 王永志. 基于改进的 YOLOv5 识别遥感影像中输电塔的方法 [J]. 实验技术与管理, 2022, 39 (4): 19-24.
- [17] 杨知, 欧文浩, 费香泽, 等. 基于高分辨率 SAR 影像和深度学习的输电杆塔智能识别研究 [J]. 电测与仪表, 2020, 57 (4): 71-77.
- [18] ZHANG H, YANG W, YU H, et al. Detecting power lines in UAV images with convolutional features and structured constraints [J]. Remote Sensing, 2019, 11 (11): 1342.
- [19] 王树文. 基于图像的输电线路语义分割技术研究 [D]. 杭州: 浙江大学, 2019.
- [20] ZHENG Z, WANG P, REN D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [J]. IEEE Transactions on Cybernetics, 2021, 26 (8): 1-13.
- [21] LIU Z, MAO H Z, WU C Y, et al. A ConvNet for the 2020s [J/OL]. IEEE Conference on Computer Vision and Pattern Recognition, 2022. [2022-05-02]. <https://arxiv.org/abs/2201.03545>.
- [22] SANDLER M B, HOWARD A G, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks [C] // Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [23] 巢渊, 刘文汇, 唐寒冰, 等. 基于改进 YOLO-v4 的室内人脸快速检测方法 [J]. 计算机工程与应用, 2022, 58 (14): 105-113.
- [24] 张秀再, 邱野, 张晨. 改进 YOLOv5s 算法的地铁场景行人目标检测 [J/OL]. 激光与光电子学进展: 1-20 [2022-08-08]. <http://kns.cnki.net/kcms/detail/31.1690.TN.20220713.1944.609.html>
- [25] 侯媛媛, 李江红, 薛军印. 基于深度学习的航空发动机滑油磨粒检测研究 [J]. 计算机测量与控制, 2022, 30 (4): 14-22, 127.
- [26] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas: IEEE Press, 2016: 770-778.
- [27] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Scaled-YOLOv4: Scaling cross stage partial network [C] // Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 13024-13033.