

doi: 10.7690/bgzdh.2023.04.015

# 基于改进 YOLOv5 与嵌入式平台的多旋翼无人机检测算法

程江川<sup>1</sup>, 王伟<sup>2</sup>, 康林<sup>3</sup>, 牛小明<sup>1,4</sup>, 李亚南<sup>1</sup>

(1. 中国兵器装备集团自动化研究所有限公司信控中心, 四川 绵阳 621000;

2. 中国兵器装备集团自动化研究所有限公司无人机事业部, 四川 绵阳 621000;

3. 陆装驻广元地区军代室, 四川 广元 628000; 4. 重庆大学数学与统计学院, 重庆 400044)

**摘要:** 针对在嵌入式平台上检测无人机时面临的资源占用率高、实时性差的问题, 提出一种改进 YOLOv5 网络的目标检测算法。以 YOLOv5s 网络为基础模型, 使用 MobileNetV3 网络代替 CSP-Darknet53 作为骨干网络进行特征提取, 并优化改进特征加强网络以及算法的回归框损失函数。基于自建无人机数据集分别在 PC 机和嵌入式平台 RK3399 上进行测试, 实验结果表明: 改进后的 YOLOv5 算法与原算法相比, 在保持较高检测精度的同时, 检测速度提升了 38%, 模型大小降低了 45%, 有效提升了算法的检测性能, 满足应用于嵌入式设备的实际需求。

**关键词:** 目标检测; 轻量化网络; 无人机检测; YOLOv5

**中图分类号:** V279 **文献标志码:** A

## Multi-rotor UAV Detection Algorithm

### Based on Improved YOLOv5 and Embedded Platform

Cheng Jiangchuan<sup>1</sup>, Wang Wei<sup>2</sup>, Kang Lin<sup>3</sup>, Niu Xiaoming<sup>1,4</sup>, Li Ya'nan<sup>1</sup>(1. *Weapon Equipment Information and Control Technology Innovation Center, Automation Research Institute Co., Ltd. of China South Industries Group Corporation, Mianyang 621000, China;* 2. *Department of UAV, Automation Research Institute Co., Ltd. of China South Industries Group Corporation, Mianyang 621000, China;*3. *Military Representative Office in Guangyuan District, Army Equipment Department, Guangyuan 628000, China;*4. *College of Mathematics and Statistics, Chongqing University, Chongqing 400044, China)*

**Abstract:** Aiming at the problems of high resource occupancy rate and poor real-time performance when detecting UAV on embedded platform, a target detection algorithm based on improved YOLOv5 network is proposed. Based on YOLOv5s network, MobileNetV3 network is used to replace CSP-Darknet53 as the backbone network for feature extraction, and the feature enhancement network and the regression box loss function of the algorithm are optimized and improved. The improved YOLOv5 algorithm is tested on PC and embedded platform RK3399 based on the self-built UAV data set, and the experimental results show that compared with the original algorithm, the improved YOLOv5 algorithm improves the detection speed by 38% and reduces the model size by 45% while maintaining a high detection accuracy, which effectively improves the detection performance of the algorithm. Meet the actual needs of embedded devices.

**Keywords:** target detection; lightweight network; UAV detection; YOLOv5

## 0 引言

多旋翼无人机具有体积小、易于操作、生产成本低等优点, 其在公共事务安全、军事侦察打击等多个行业都有了较大规模的应用<sup>[1]</sup>。无人机的普遍应用不可避免地带来了一些安全隐患和公共威胁, 对其进行有效的反制与管控则十分重要, 如何快速且准确地检测出无人机是当前主要的任务<sup>[2]</sup>。

目前, 常用的无人机检测手段主要有无线电检测、声纹检测、雷达检测以及光电检测 4 种<sup>[3]</sup>。相比前 3 种方法, 光电检测方式具有成本低、检测速度快和检测结果直观等优点, 更适用于小型化的无

人机检测系统<sup>[4-5]</sup>。

YOLOv5<sup>[6-7]</sup>算法具有良好的检测性能, 但实际的无人机光电检测系统需要设备具有可移动性和便携携带的特点, 而 YOLOv5 模型参数量较多, 网络模型较大, 算法在实际工作中部署到嵌入式设备会遇到运行速率慢和加载时间长的问题。笔者提出一种改进的 YOLOv5 模型, 通过更换轻量级网络 MobileNetV3 作为特征提取网络, 大幅度降低网络的参数量, 并优化特征加强网络以及算法的回归框损失函数, 获得更加精确的锚框定位并且加快回归框的收敛速度, 有效提升算法的检测精度和速度, 可在实际应用的嵌入式平台上更好地部署目标检测

收稿日期: 2022-12-06; 修回日期: 2023-01-05

作者简介: 程江川(1998—), 男, 重庆人, 硕士, 高级工程师, 从事计算机视觉、图像处理、目标检测研究。E-mail: 1154017495@qq.com。

算法，以此来解决检测无人机中面临的问题。

## 1 YOLOv5 算法及改进

### 1.1 YOLOv5 网络模型

YOLOv5 网络模型主要由 Input(输入端)、Backbone(骨干网络)、Neck(颈网络)和 Output(输出端)4部分网络组成。YOLOv5 算法实现了网络版本的系列化，开发者提供的该目标检测网络共有4个版本，分别是 Yolov5l、Yolov5m、Yolov5s、Yolov5x4 个模型。通过对比这几个版本的模型大小、检测速率后发现，YOLOv5s 十分适用于对无人机的实时目标检测，故本次实验采用的 YOLOv5 版本为 v5s。笔者分别从 YOLOv5 的骨干网络、特征融合网络以及损失函数对算法特点进行介绍。

#### 1.1.1 骨干网络

YOLOv5 的骨干网络采用 CSPDarkNet53 表征学习，在检测性能上优于 ResNet 基准算法。其深度、宽度均衡化的特性兼容了不同设备、数据集。残差网络避免了深度网络学习中梯度消失的问题，以及 CSPNet<sup>[8]</sup>在不丢失模型精度的条件下，加速推理 44%。

#### 1.1.2 特征融合

YOLOv5 算法沿用了特征金字塔网络(feature pyramid networks, FPN)<sup>[9]</sup>作为特征融合网络，增添了 PANet<sup>[10]</sup>网络进行多尺度特征融合，实现整合不同特征层的检测结果，增加多层目标特征信息，提高了多尺度目标的定位和分类精度。并且采用 CSPNet 网络对各个输出层的特征进行融合，有效地提升了多尺度特征融合的速度。

#### 1.1.3 损失函数

YOLOv5 算法总的损失函数由回归框损失 box\_loss、置信度损失 obj\_loss 和分类损失 cls\_loss 3部分组成，共同作用使得网络参数调整到最小化误差。回归框损失采用 GIOU 进行计算，其作用是计算预测框与标定框之间的误差；置信度损失用于计算网络的置信度，分类损失用于计算锚框与对应的标定分类是否正确，这2种损失函数均采用交叉熵损失进行计算。

### 1.2 改进 YOLOv5 算法

尽管 YOLOv5 算法在公开数据集上表现优异，在自然场景下有巨大优势，但在实际检测无人机的系统中，由于嵌入式设备的算力受到硬件资源的限

制，算法的检测效果则没有在 PC 机上表现好。笔者从轻量化网络模型对 YOLOv5 进行改进处理，并对算法的部分网络和回归框损失函数进行改进，在保持算法精度的同时尽可能地提升网络检测速度。

#### 1.2.1 轻量化骨干网络

MobileNetV3 网络采用了“轻量级注意力模型”“深度可分离卷积”“逆残差结构”这3种思想，删除了特征提取网络中最后一个卷积层并且前移最后一个平均池化层。原 swish 激活函数存在计算复杂、不利于量化过程的问题，故将其更新为 h-swish 激活函数，减少了运算量，提高网络运行性能。

为降低网络模型的参数量、降低模型大小，使算法更高效地部署在嵌入式开发平台上，笔者采用 MobileNetV3<sup>[11]</sup>轻量化网络代替原始骨干网络 CSP-DarkNet53 进行特征提取工作，大幅降低网络的参数量和运算内存。构建了 MobileNetV3-YOLOv5 网络，改进后的网络模型如图1所示。

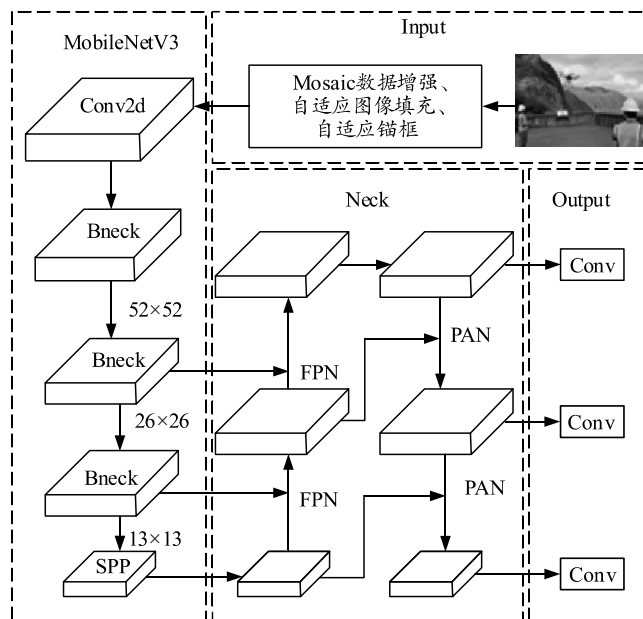


图1 改进后YOLOv5网络结构

#### 1.2.2 特征融合网络优化

PANet 作为 YOLOv5 检测模型的特征加强网络，其具有2个创新点：引入了自底向上的路径，使得底层信息更容易传递到高层顶部；PANet 的引入，可以增强整个特征提取网络的定位能力。

深度可分离卷积可以在对精度影响较小的同时大大减小网络参数量和计算量<sup>[12]</sup>。图2为深度可分离卷积的结构示意图，深度可分离卷积由逐通道卷积和逐点卷积2个模块组合而成。逐通道卷积层使用3x3逐深度卷积核对每一个输入通道进行卷积；

逐点卷积层使用  $1 \times 1$  卷积对来自逐通道卷积层的卷积输出进行组合。正则化项 BN 层用于对深度卷积层后输出的数据进行分布调整，其后的 ReLU 激活函数可以提升网络的计算速率。

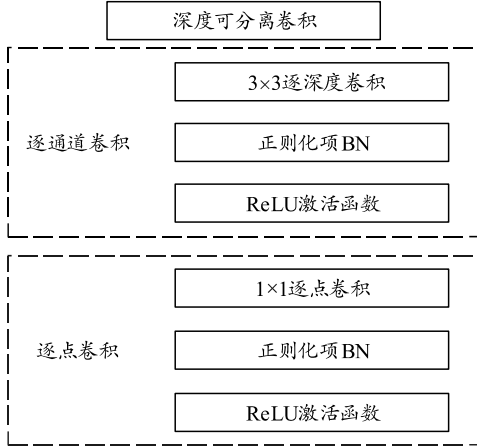


图 2 深度可分离卷积结构

在对比深度可分卷积和常用卷积计算方式不同点的基础上，为进一步降低网络的参数量、提升计算速度，本次实验将 PANet 网络中的卷积方式改进为深度可分离卷积，力求提升 YOLOv5 算法的运行效率。

1.2.3 损失函数优化

YOLOv5 算法采用 GIoU 作为回归框损失函数。GIoU Loss 在原函数的基础上引进预测框和真实框的最小外接矩形。假设 A 为预测框，B 为真实框，令 C 为包含 A 与 B 的最小凸闭合框。其计算公式如下：

$$GIoU = IoU - |C / (A \cup B)| / |C|; \quad (1)$$

$$L_{GIoU} = 1 - GIoU. \quad (2)$$

GIoULoss 损失函数值无法很好地衡量预测框和真实框相对的位置关系，同时存在预测框在水平或垂直方向优化困难，导致收敛速度慢。而 CIoULoss 针对 GIoULoss 面临的问题，综合考虑了中心点距离、覆盖面积以及长宽比这 3 点。其计算公式如下：

$$CIoU = IoU - p^2(b, b^{gt}) / c^2 - \alpha v; \quad (3)$$

$$L_{CIoU} = 1 - CIoU. \quad (4)$$

式中： $b$  和  $b^{gt}$  分别为真实框和预测框的中心点； $p^2(\bullet)$  为欧式距离的计算公式； $c$  为真实框和预测框的最小包围框所对应的的最短对角线距离； $\alpha$  为 2 个框长宽比的正平衡参数。 $\alpha$  和  $v$  的定义如下：

$$v = 4(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h})^2 / \pi^2; \quad (5)$$

$$\alpha = v / ((1 - IoU) + v). \quad (6)$$

CIoU 损失函数较原始算法的损失函数相比，能提升网络训练中的收敛速度，并得到更高的回归框定位精度，故笔者将 CIoULoss 作为检测无人机算法的损失函数。

2 实验准备与设计

2.1 数据集准备

本文中无人机数据集所需图片来源于网络搜集，选取不同背景下的多种型号以及尺寸的无人机图片。在进行数据集训练之前需对该数据集进行处理，使其样本标准化。选用 LabelImg 人工标注软件对图片中的无人机目标进行注释，完整并精确地框选出无人机的目标区域并匹配上对应的目标名称，软件会自动生成一个包含图片各种信息的 xml 文件。无人机数据集格式为 Pascal VOC，共整理得到 3 000 张图片，将其划分为训练集、验证集和测试集，比例为 7:2:1。

2.2 实验环境

笔者选用 PyTorch 深度学习框架，PyTorch 具有易于使用的 API，可实现强大的 GPU 加速同时包含自动求导系统的深度神经网络，这是许多其他主流框架不支持的。本次实验具体的软硬件开发环境如表 1 所示。

表 1 本实验软硬件开发环境

名称	信息
处理器	Intel Corei7-7800X@3.5 GHz
显卡	NVICIA GeForce GTX 2080 Ti
内存	16GB
硬盘	Intel 760P M.2 SSD(256G)
操作系统	Ubuntu18.04
深度学习框架	PyTorch
CUDA	Cuda10.0
Anaconda	Anaconda3(Python3.7)
OpenCV	OpenCV3.0

2.3 实验指标

为更好地对算法的性能进行评估，本文中选用平均精度均值 (mean average precision, mAP) 和检测速度 FPS 这 2 个常用的目标检测评价指标。mAP 是测试集中每一目标类别检测准确率的平均值。准确率 (Precision) 是指所有预测中预测正确的比例。

$$Precision = TP / (TP + FP). \quad (7)$$

式中： $TP$  为算法成功预测出正确样本的数量； $FP$  为算法失败预测出正确样本的数量。单类别的平均精准度 AP 的计算公式为：

$$AP = \sum \text{Precision} / N. \quad (8)$$

式中： $\sum \text{Precision}$  为算法在所有样本中检测出某个目标类别准确率的总和； $N$  为检测数据集中的总样本数。所有类别目标的平均检测精度  $mAP$  的计算公式为：

$$mAP = \left( \sum_{i=1}^k AP \right) / K. \quad (9)$$

### 3 实验结果与分析

为对改进后的算法模型进行客观和全面的效果展示以及性能分析，分别将表现良好的双阶段检测算法 Faster R-CNN、原始 YOLOv5 以及笔者改进后的 YOLOv5 算法在 PC 机上使用相同的无人机训练集和训练方法进行训练。在实验训练过程中通过查看训练日志可以发现，在训练达到 200 次左右时，损失函数基本不再下降逐渐趋于平稳；因此，在训练达到 200 次时终止训练。将训练日志中的各个算法的损失函数进行统计，具体结果如图 3 所示。

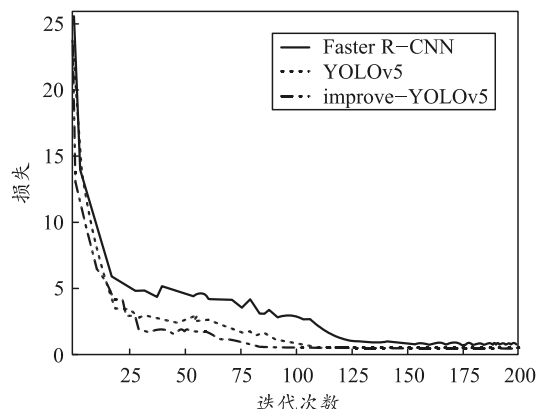


图 3 模型损失曲线对比

上图为 Faster R-CNN、原始 YOLOv5 以及改进后的 YOLOv5 这 3 个算法模型在训练中的损失函数变化曲线。由图可知，在相同的训练参数下，Faster R-CNN 的收敛速度最慢，在 125 个 Epoch 后开始收敛，损失值波动较大；原始 YOLOv5 在 100 个 Epoch 后开始收敛，损失值波动大；而改进后的 YOLOv5 的收敛速度最快，在 75 个 Epoch 后开始收敛，损失值的波动也最小。改进后 YOLOv5 的损失函数下降趋势更明显，表明了改进后算法的有效性，可见笔者提出 CIoU Loss 对算法速度提升效果较优。

为体现改进后算法的检测泛化能力，从测试集中选出不同的测试图片在改进后的 YOLOv5 算法上进行测试，图片分别包含单个无人机目标和多个无人机目标，以及分别进行复杂场景和夜晚背景下的

测试，算法检测无人机的测试结果如图 4 所示。从测试结果可以看出，当图片中检测目标仅为 1 个时，算法的检测具有较高的精度，而当图片中存在多个无人机目标时，无人机的尺寸存在大小的变化，算法检测精度随着无人机距离的增大而降低；当无人机处于复杂场景下时，如周围存在烟雾、火焰以及部分遮挡时，检测精度较高，说明本文中算法具有较强的特征提取能力，受其他背景的影响较小；但当夜晚背景时对无人机进行检测则效果较差，说明光照对于检测结果的影响较大。



(a) 单个无人机检测



(b) 多个无人机检测



(c) 复杂场景下无人机检测



(d) 夜晚背景无人机检测

图 4 多种场景下无人机检测效果

再分别从测试集中选取 150 张图片在 PC 机以及嵌入式设备 RK3399 平台上进行测试，对上述 3 种算法分别进行 15 次测试后取其平均值，再计算其平均检测精度，最后通过不同算法的模型大小、 $mAP$  以及检测速率 FPS 这 3 个指标对模型进行性能对比，结果如表 2 所示。

表 2 模型性能对比

算法	主干网络	模型大小/MB	设备	Map/%	检测速率/FPS
Faster R-CNN	ResNet50	109	PC 机	90.6	2.8
			RK3399	90.2	1.1
YOLOv5s	Darknet-53	14.1	PC 机	89.6	61.5
			RK3399	89.1	32.3
改进 YOLOv5	MobileNetV3	7.8	PC 机	88.4	85.0
			RK3399	88.1	44.6

由上表可以看出,改进后的 YOLOv5 算法相较于改进前虽然 mAP 下降了 1.3%,但仍保持着较高的平均检测精度,且检测速率较改进前提升了 38%,轻量化策略使得网络模型大幅度地缩小,模型大小从 14.1M 下降到 7.8M,降低幅度为 45%;RK3399 平台上的结果显示检测速率从 32.3FPS 提升到了 44.6FPS,检测速度明显提升,同时保持着较高的检测精度,达到了 88.1%,和常用的两阶段算法 Faster R-CNN 相比,Faster R-CNN 的模型大小过大,导致检测速率在 RK3399 上仅为 1.1FPS,无法满足实际的实时无人机检测需求,故改进后的 YOLOv5 算法在检测精度和检测速度上综合比较具有明显优势。综上所述,经过不同算法的性能对比分析,可得到改进后 YOLOv5 模型适合部署至嵌入式开发平台,其性能最能满足实际无人机检测系统的需求。

#### 4 结束语

笔者提出的改进 YOLOv5 模型选用 MobilenetV3 替换原 YOLOv5 的主干特征提取部分 CSP-Darknet53,降低了网络的参数量和计算量,有效缓解了因嵌入式设备资源受限造成的检测实时性差的问题;优化改进特征加强网络以及回归框损失函数,改进了特征融合网络部分卷积方式,加快了模型的收敛速度且提高了回归框定位的精度,提高了算法的运行速度。实验结果表明,改进后的算法在 PC 机以及嵌入式设备上均能满足实时检测无人机的需求。

虽然笔者所提改进方法在性能上已具有较好的提升,但从检测结果图中可以看出,当同时有多个无人机目标并且无人机尺度变化较大时,算法对于图中较小的无人机目标检测效果不佳;在夜晚背景下算法检测精度较低,甚至偶尔出现了如图 5(d)的误检现象。查阅文献后得知:这是因为该算法对于小目标的特征提取能力还有待加强,对于夜晚场景下,图像具有噪声高、对比度低、特征信息不明显的特点,加大了检测难度;且本实验训练数据集中的小目标样本数以及夜晚场景样本数量都较少,算法未对其进行充分训练学习也会导致检测效果欠

佳。后续将在现有模型基础上改进多尺度特征网络,增加注意力机制以及进行小目标和夜晚环境下的图像数据增强,力图在保证现有检测性能的同时增强算法对于小目标的检测能力。

#### 参考文献:

- [1] 王守银. 多旋翼无人机在公共安全领域应用构想[J]. 机器人产业, 2022(3): 36-41.
- [2] 罗俊海, 王芝燕. 无人机探测与对抗技术发展及应用综述[J]. 控制与决策, 2022, 37(3): 530-544.
- [3] 蒋榕圻, 白若楷, 彭月平. 低慢小无人机目标探测技术综述[J]. 飞航导弹, 2020(9): 100-105.
- [4] 齐咏生, 孟学斌, 高学金, 等. 基于视觉标志检测的旋翼无人机姿态估计方法[J]. 农业机械学报, 2019, 50(6): 28-40, 139.
- [5] BEHERA D K, RAJA B. Drone Detection and Classification using Deep Learning[C]//2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2020.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.
- [7] 刘思诚, 李嘉琛, 邓皓, 等. 基于 YOLOv5 改进的小目标检测算法[J]. 兵工自动化, 2022, 41(12): 78-82.
- [8] WANG C Y, LIAO H, WU Y H, et al. CSPNet: A New Backbone that can Enhance Learning Capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2020.
- [9] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.
- [10] LIU S, QI L, QIN H, et al. Path Aggregation Network for Instance Segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018.
- [11] HOWARD A, SANDLER M, CHUG, et al. Searching for Mobile NetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2019.
- [12] CHOLLET F. Xception: Deep Learning with Depthwise Separable Convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017.