

文章编号: 1005-0523(2024)04-0118-09



## 基于改进YOLOv5s的无人机小目标检测算法研究

董华军<sup>1,2</sup>, 王宇栖<sup>1</sup>

(1. 大连交通大学机械工程学院, 辽宁 大连 116028; 2. 大连交通大学自动化与电气工程学院, 辽宁 大连 116028)

**摘要:**【目的】针对无人机航拍图像中目标尺度多样、背景复杂、小目标密集的特点,提出了基于YOLOv5s的小目标检测算法LM-YOLO。【方法】首先,增加小目标检测头并采用K-DBSCAN聚类算法优化锚框,生成更适合小目标检测的锚框,提高算法对小目标的检测精度;然后,设计更高效的MobileNetV3-CBAM作为特征提取网络,减小网络模型大小;最后,在特征融合网络引入大核选择性注意力机制LSK,增加模型对相似目标的分辨率。【结果】在公开数据集VisDrone2019上的实验结果表明,与基准模型YOLOv5s相比,LM-YOLO对所有目标的平均检测精度提升了7.6%,模型大小压缩了45%。【结论】文章算法可以在降低模型大小的同时保持良好的检测精度。

**关键词:** 无人机图像; 小目标检测; 聚类算法; YOLOv5s; 注意力机制

中图分类号: TP391.4

文献标志码: A

本文引用格式: 董华军, 王宇栖. 基于改进YOLOv5s的无人机小目标检测算法研究[J]. 华东交通大学学报, 2024, 41(4): 118-126.

## Research on UAV Small Target Detection Algorithm Based on Improved YOLOv5s

Dong Huajun<sup>1,2</sup>, Wang Yuqi<sup>1</sup>

(1. School of Mechanical Engineering, Dalian Jiaotong University, Dalian 116028, China; 2. School of Automation and Electrical Engineering, Dalian Jiaotong University, Dalian 116028, China)

**Abstract:** 【Objective】Aiming at the characteristics of various target scales, complex background and dense small targets in aerial images of unmanned aerial vehicles (UAV), a small target detection algorithm LM-YOLO based on YOLOv5 is proposed. 【Method】Firstly, the number of small target detection head was increased and K-DBSCAN clustering algorithm was used to optimize the anchor frame, so as to generate an anchor frame more suitable for small target detection and improve the detection accuracy of the algorithm. Then, a more efficient MobileNetV3-CBAM was designed as a feature extraction network to reduce the size of the network model. Finally, the large kernel selective attention mechanism LSK was introduced into the feature fusion network to increase the resolution of the model to similar targets. 【Result】The experimental results on the public data set VisDrone2019 show that the average detection accuracy of LM-YOLO for all targets is improved by 7.6% and the model size is reduced by 45% compared with the benchmark model YOLOv5. 【Conclusion】Experiments show that the proposed algorithm can reduce the model size while maintaining good detection accuracy.

**Key words:** UAV images; small target detection; clustering algorithm; YOLOv5s; attention mechanism

**Citation format:** DONG H J, WANG Y Q. Research on UAV small target detection algorithm based on

收稿日期: 2024-03-15

基金项目: 辽宁“百千万人才工程”培养经费资助项目; 辽宁省教育厅科学研究计划资助项目(LJKMZ20220835)

improved YOLOv5s[J]. Journal of East China Jiaotong University, 2024, 41(4): 118–126.

**【研究意义】**随着深度学习技术的蓬勃发展,无人机与目标检测技术的结合已在智能巡检、环境勘测、交通监管等领域有着广泛应用<sup>[1]</sup>。由于无人机拍摄距离远、视野宽广,导致航拍图像具有背景复杂、邻近目标相互遮挡、拍摄光线不足、小目标众多等问题,现有的算法在无人机视角下的检测效果并不理想<sup>[2]</sup>。因此,研究适用于无人机航拍图像的目标检测算法有着重要意义。

**【研究现状】**目前目标检测算法大致可以分为两类,第一类是双阶段检测算法,主要以区域卷积神经网络 R-CNN<sup>[3]</sup>, Fast-RCNN<sup>[4]</sup>, Faster-RCNN<sup>[5]</sup>为代表。该算法在第一阶段生成候选区域,提取目标的候选框,在第二阶段对目标物体进行分类和回归,该类算法平均目标检测精度高,但是检测速度较差,难以达到实时性要求。第二类是单阶段检测算法,主要以 YOLO 系列算法<sup>[6-9]</sup>、SDD 算法<sup>[10-11]</sup>为代表。其中 YOLOv5 具有检测速度快、模型相对较小的优点,在单次检测中直接获得结果,被广泛应用于无人机图像目标检测中。Jiang 等<sup>[12]</sup>在 YOLOv5 的特征融合网络增加了上采样方法,在通道维度对特征图进行拼接,增强了算法对小目标的检测精度,但降低了颈部网络提取图像特征的能力,检测速度也大幅下降,不足以满足实际需求。吴萌萌等<sup>[13]</sup>在 YOLOv5s 的基础上,设计了融合特征融合因子的自适应双向特征融合模块(M-BiFPN),使网络的特征表达能力提升,由于未考虑数据集长尾分布的特点,该算法对目标的检测性能提升不够明显。谢椿辉等<sup>[14]</sup>将 YOLOv5 的头部网络进行改进,将检测头解耦并新增小目标预测头,增强网络对小目标的检测率,但对 4 个检测头解耦降低了算法的检测速度。Qiu 等<sup>[15]</sup>对 YOLOv5 的空间金字塔进行改进,使用 SPPF 代替原有的 SPP,增强了骨干网络提取多尺度特征的能力,但没有提升小目标周围的感受野,忽略了小目标背景的先验知识。

**【创新特色】**针对现有改进方法存在的问题,本文从轻量化和高效化角度对 YOLOv5s 算法进行改进研究,提出了 LM-YOLO 算法。首先增加小目标检测层,其次采用 K-DBSCAN 算法重新聚类锚框,

提高模型对小目标的检测精度。从特征融合角度,设计 MobileNetV3-CBAM 网络代替 YOLOv5s 骨干网络,增强网络对空间特征的提取能力,减少模型大小。从特征融合角度,在颈部网络插入大核选择性的注意力机制 LSK,获取小目标更多的背景上下文信息,增加网络对相似小目标的辨识度。**【关键词】**研究基于 YOLOv5s 的航拍图像小目标检测算法,能够充分利用 YOLOv5s 算法的速度优势,融合优秀的改进策略后,可以有效避免检测小目标时出现的误检或漏检问题。

## 1 LM-YOLO

### 1.1 LM-YOLO 网络结构

针对无人机航拍图像具有背景复杂、小目标多且密集的特点,LM-YOLO 在满足实时检测和模型轻量化的条件下,能够有效提升模型对复杂背景下密集小目标的检测精度,LM-YOLO 模型结构如图 1 所示。图中, Bneck 是 MobileNetV3 的瓶颈结构,融合了深度可分离卷积<sup>[16]</sup>和倒置残差结构<sup>[17]</sup>, LSK 是大核选择注意力模块。LM-YOLO 模型由融合 MobileNetV3-CBAM 的特征提取网络(Backone),添加 LSK 注意力机制的特征融合网络(Neck),以及增加小目标检测头的目标检测层(Head)组成。

### 1.2 聚类算法改进

YOLOv5s 对默认先验框的聚类算法是 K-means<sup>[18]</sup>,此算法适合聚类大目标的数据集,但对于本文研究的无人机航拍小目标来说,由于数据集孤立点的干扰和初始聚类中心的随机性,可能会导致聚类算法陷入局部最优的情况。因此本文使用了基于密度的 K-DBSCAN 算法对目标真实框进行聚类,其核心思想是使用 DBSCAN<sup>[19]</sup>算法消除孤立点的干扰并确定初始聚类中心个数  $K$ ,再使用 K-means 算法对初始聚类中心进行聚类分析,生成适合无人机小目标检测的锚框。

同时,为了能够提取更多的小目标位置信息,本文对 YOLOv5 中的检测层进行改进,增加了一个微小目标检测头,选取 12 个聚类中心,等分到 4 个大小不同的预测头上,聚类后的锚框大小,如表 1 所

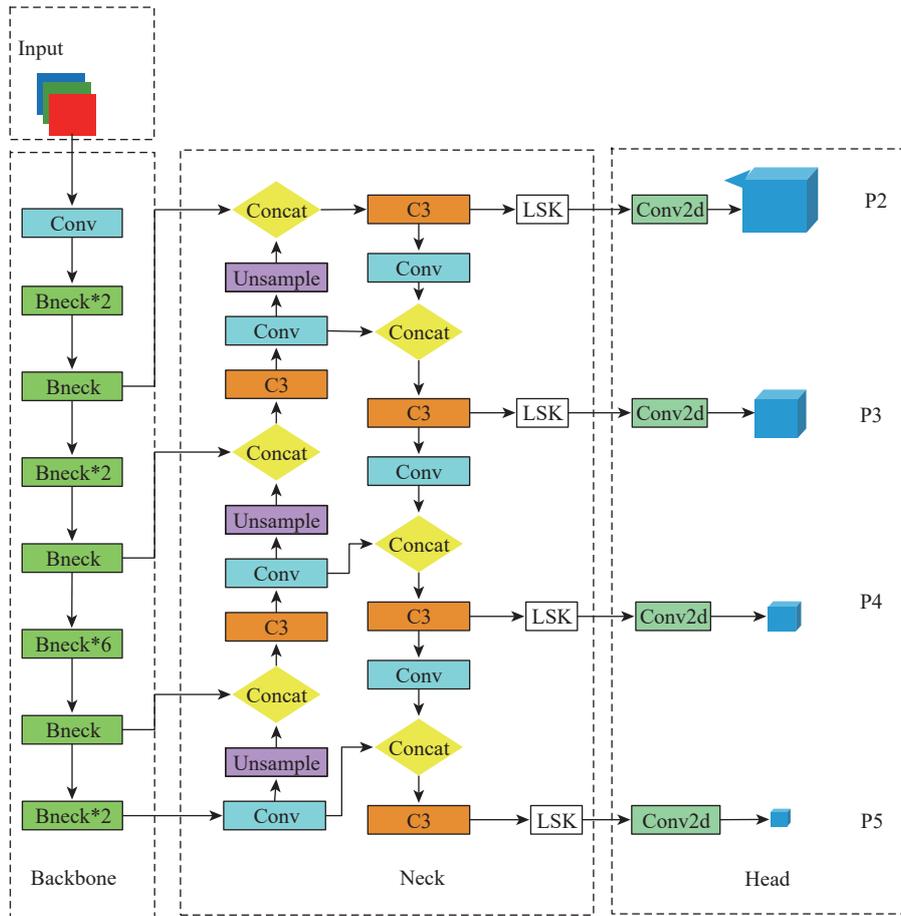


图1 LM-YOLO网络结构图  
 Fig. 1 LM-YOLO network structure diagram

表1 预设锚框尺度分配表

Tab.1 Anchor boxes scale assignment

Target size	Anchor frame size		
Minimal target	(2, 5)	(4, 8)	(5, 6)
Small target	(7, 12)	(10, 10)	(9, 18)
Medium target	(20, 18)	(15, 31)	(34, 42)
Large target	(28, 40)	(61, 46)	(59, 101)

示。这样新增的小目标检测头能够检测更小的目标,也避免了因锚框过大所带来的小目标漏检问题。

### 1.3 特征提取网络改进

原YOLOv5s的骨干网络参数量较大,对硬件配置要求较高。本文基于轻量化思想,设计了融合MobileNetV3网络和CBAM注意力的MobileNetV3-CBAM网络,替换原有的骨干网络,在保证模型检测能力的同时降低模型大小。

MobileNetV3包含了MobileNetV3-large和Mo-

bileNetV3-small两个版本<sup>[20]</sup>,本研究选取检测精度更高的MobileNetV3-large作为特征提取层的基准网络,移除了MobileNetV3-large原网络的最后4层,因为这4层是分类层对目标提取无用。此外,增加微小目标检测头后,YOLOv5s的特征融合网络所需输入由3个变为4个,选取第3、6、13层和第15层特征作为Backbone的输出,输入到特征融合网络之中,输入尺寸设置为640×640×3。

MobileNetV3网络中的注意力模块是SEnet,网络会重点关注信息量较大的通道特征<sup>[21]</sup>,而忽视了小目标检测所需的空间特征,文献[22]提出了一种同时关注通道和空间特征的注意力模块(convolutional block attention module,CBAM),如图2所示。

由图2可知,CBAM模块先后使用两个注意力模块,推断通道和空间的权重系数<sup>[23]</sup>,经过Sigmoid权重归一化后与原始图特征进行相乘,最终生成权重优化的特征图。本文在MobileNetV3-large的

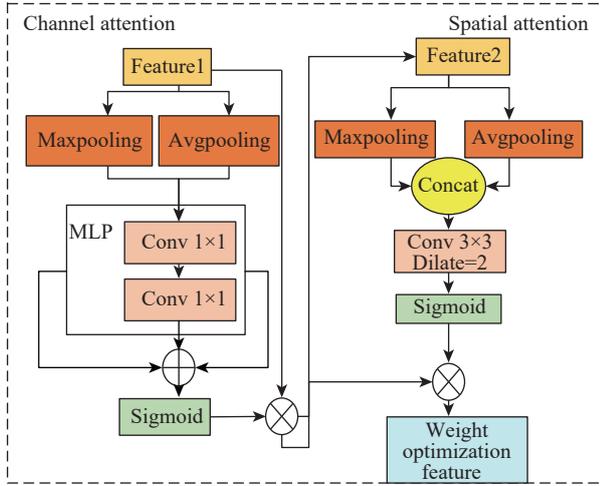


图2 CBAM注意力机制  
Fig. 2 CBAM attention mechanism

Bneck层嵌入了CBAM模块,设计了MobileNetV3-CBAM网络作为YOLOv5s的Backone网络,该网络结构如表2所示。

1.4 特征融合网络改进

无人机航拍图像中的目标尺寸很小,仅靠表面特征容易导致误检,需要结合大范围的背景上下文信息,才能正确地进行识别。LSKNet<sup>[24]</sup>考虑到了目标周围背景的先验知识,可以通过空间注意力机制,动态改变其感受野,利用大的感受野来提取航拍图像中的背景上下文信息。

LSKNet的核心是大核选择性注意力机制LSK,本文将LSK注意力模块添加到YOLOv5s的特征提取网络,增强网络动态捕捉高级图像特征的能力,从而提高航拍小目标的检测精度。

表2 MobileNetV3-CBAM网络结构

Tab.2 MobileNetV3-CBAM network structure

Number	Input size	Operation type	Extended size	Output channels	CBAM module	Activation function	Step length
0	640×640×3	Conv2d	-	16	-	HS	2
1	320×320×16	Bneck, 3×3	16	8	-	RE	1
2	320×320×8	Bneck, 3×3	64	16	-	RE	2
3	160×160×16	Bneck, 3×3	72	16	-	RE	1
4	160×160×16	Bneck, 5×5	72	16	✓	RE	2
5	80×80×16	Bneck, 5×5	120	16	✓	RE	1
6	80×80×16	Bneck, 5×5	120	16	✓	RE	1
7	80×80×16	Bneck, 3×3	240	32	-	HS	2
8	40×40×32	Bneck, 3×3	400	32	-	HS	1
9	40×40×32	Bneck, 3×3	184	32	-	HS	1
10	40×40×32	Bneck, 3×3	184	32	-	HS	1
11	40×40×32	Bneck, 3×3	480	10	✓	HS	1
12	40×40×40	Bneck, 3×3	672	40	✓	HS	1
13	40×40×40	Bneck, 5×5	672	56	✓	HS	2
14	40×40×56	Bneck, 5×5	960	56	✓	HS	1
15	20×20×56	Bneck, 5×5	960	56	✓	HS	1

LSK模块结构如图3所示,LSK模块将一个大的卷积核顺序拆解成多个小卷积核,将输入数据X依次通过这些小卷积核,并在每个小卷积核后添加一个1×1的卷积核,进行特征向量的通道融合。

假设一个大卷积核分解为N个小卷积核,通道融合过程为

$$U_0 = X, U_{i+1} = F_i^{dw}(U_i) \quad (1)$$

$$U_i = F_i^{1 \times 1}(U_i), \text{ for } i \text{ in } [1, N] \quad (2)$$

式中:  $F_i^{dw}()$  为第*i*个卷积核的深度可分离卷积操作(depth-wise);  $U_i$  为提取后的特征图。为了让网络更加关注目标背景上下文信息,利用SA空间注意力机制从不同尺度的卷积核中对特征图进行空间维度选择。首先对来自不同感受野的特征图进行拼接为

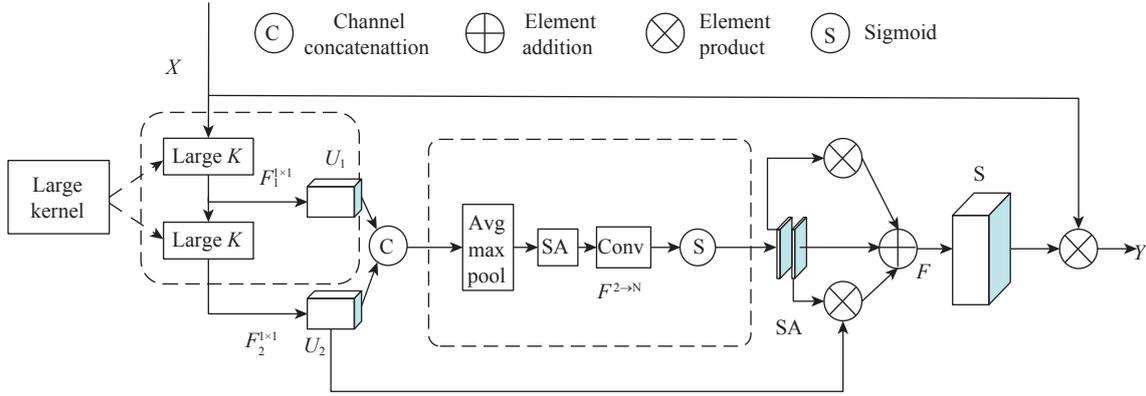


图3 LSK模块结构示意图

Fig. 3 Schematic diagram of LSK module structure

$$U = [U_1, \dots, U_i] \quad (3)$$

然后对拼接后的特征图  $U$  进行最大池化和平均池化操作为

$$SA_{\text{avg}} = P_{\text{avg}}(U), SA_{\text{max}} = P_{\text{max}}(U) \quad (4)$$

式中:  $P_{\text{avg}}()$  和  $P_{\text{max}}()$  分别为平均池化和最大池化操作;  $SA_{\text{avg}}$  和  $SA_{\text{max}}$  分别为平均池化和最大池化的空间特征描述符。接着利用SA空间注意力机制,将两个特征描述符拼接起来,并利用卷积层  $F^{2 \rightarrow N}(\cdot)$  将两通道的池化特征转换为  $N$  通道空间注意力为

$$SA = F^{2 \rightarrow N}(SA_{\text{avg}}; SA_{\text{max}}) \quad (5)$$

对于每个空间注意力  $SA$ , 采用Sigmoid激活函数得到每个分解大核的独特空间选择掩码,然后将解耦后的大卷积核序列的特征与对应的空间选择掩膜进行加权处理,通过卷积层  $F$  进行融合,得到注意力特征  $S$  为

$$SA_i = \sigma(SA_i) \quad (6)$$

$$S = F\left(\sum_{i=1}^N (A_i \cdot U_i)\right) \quad (7)$$

最后将输入特征  $X$  和注意力特征  $S$  进行逐元素点乘,得到最终输出  $Y$  为

$$Y = X \cdot S \quad (8)$$

## 2 实验验证与结果分析

### 2.1 实验环境与数据集

实验使用的硬件配置为 NVIDIA GeForce RTX3060(6GB) GPU 和 Intel(R) Core(TM) i7-10750H CPU。软件环境为 Windows 10 系统下的 Pytorch 深度学习框架, CUDA 版本为 11.3。设置训练初始学习率为 0.01, 循环学习率为 0.01, 训练轮数

为 200 轮, 训练批次为 8。为验证 LM-YOLO 算法模型的有效性, 本文所有的实验均用 VisDrone2019 数据集进行实验。该数据集由天津大学机器学习与数据挖掘实验室团队负责, 有 10 类标签, 大约有 540 000 个标注信息, 是无人机视角下具有代表性的数据集。

### 2.2 性能评价指标

为评估 LM-YOLO 算法对小目标检测的有效性, 本文采取  $mAP@0.5$  作为模型对多个目标检测精度的评价指标,  $mAP@0.5$  表示当候选框与真实框的交并比为 0.5 时, 模型对所有目标类别的平均检测精度。为了衡量网络是否适合部署到无人机上, 还需要考虑每秒传输帧数  $FPS$  和模型的大小。

试验中, 传输帧数  $FPS$  和平均检测精度  $mAP$  的表达式如下

$$FPS = \frac{1\ 000}{P_{\text{reTime}} + I_{\text{inferTime}} + N_{\text{msTime}}} \quad (9)$$

$$mAP = \frac{\sum_{i=0}^n \int_0^1 P(R) dR}{n} \quad (10)$$

式中:  $P_{\text{reTime}}$  为图像的预处理时间, ms;  $I_{\text{inferTime}}$  为网络推理时间, ms;  $N_{\text{msTime}}$  为非极大值抑制时间, ms;  $n$  为检测的类别种类;  $P$  为检测精确度;  $R$  为检测召回率。

### 2.3 消融实验

为了验证锚框重新聚类、改进的 MobileNetV3-CBAM 特征提取网络、增加小目标检测头以及引入 LSK 注意力机制的有效性, 本文在 YOLOv5s 原模型的基础上逐一引入了改进方案, 对于改进部分进行消融实验, 实验结果如表 3 所示。

从表 3 中可看出, 本文提出的模块均可提高算

表3 改进YOLOv5s消融实验结果

Tab.3 Results of improved YOLOv5s ablation test

Method	K-DBSCAN	MobileNetV3-CBAM	Small target detection head	LSK	$mAP@0.5/\%$	Model size/MB	FPS/(frame/s)
A	×	×	×	×	28.7	13.4	96
B	√	×	×	×	29.0	13.4	96
C	×	√	×	×	29.1	7.3	127
D	×	×	√	×	31.6	13.7	87
E	×	×	×	√	33.0	14.1	79
F	√	√	×	×	29.2	7.3	127
G	√	√	√	×	31.5	10.6	93
H	√	√	√	√	36.3	12.7	54

法的检测精度。在单独使用K-DBSCAN聚类算法时,模型 $mAP@0.5$ 提升0.3%,说明K-DBSCAN可以优化锚框,生成更适合小目标任务的候选框;在单独将骨干网络换为MobileNetV3-CBAM轻量级网络时, $mAP@0.5$ 提升0.4%,模型FPS获得了极大的提升,提升31 frame/s,同时模型的大小减少6.1 MB;在单独增加检测头时,由于融合了浅层信息,模型的 $mAP@0.5$ 有2.9%的提升,但模型计算量因检测头的增加而有所增大,导致模型大小增加了0.3 MB,FPS减少了9 frame/s;在单独引入LSK注意力机制后,由于充分利用了背景信息的先验知识,极大增强了模型的检测精度,模型的 $mAP@0.5$ 提升4.3%,但模型大小增加0.7 MB,FPS下降17 frame/s。通过叠加效果的观察可知,经过改进的模型在检测效果上达到最优。

#### 2.4 对比实验

为了验证LM-YOLO算法的优越性,本文将此算法与其他经典目标检测算法进行对比,实验结果如表4所示。

从表4可以看出,本文提出的LM-YOLO与Faster R-CNN, SSD, Swin-transformer, YOLOX和YOLOV5m, CornerNet相比, $mAP@0.5$ 分别提升13.5%, 7.2%, 4.9%, 11.7%, 3.4%, 12.6%。与针对航拍小目标的检测算法Drone-YOLO相比,LM-YOLO的 $mAP@0.5$ 差距仅为0.2%,但是FPS高出23 frame/s。虽然TPH-YOLOv5的检测精度略高于本文算法,但是TPH-YOLOv5以YOLOv5x作为基准模型,计算量大且推理速度较慢,FPS低于以YOLOv5s为基准模型的LM-YOLO。对比结果表明:LM-YOLO在检测精度和速度之间的均衡性更强,

表4 不同算法性能对比

Tab.4 Comparison of the selected detection algorithm

Target detection algorithm	$mAP@0.5/\%$	FPS/(frame/s)
Faster R-CNN	22.8	6
SSD	29.1	35
Swin-transformer	31.4	53
YOLOX	24.6	49
YOLOV5m	32.9	46
CornerNet	23.7	15
Drone-YOLO	36.5	31
TPH-YOLOv5	37.2	22
LM-YOLO	36.3	54

再次证明了本文所设计算法在无人机航拍检测中的有效性。

#### 2.5 算法有效性分析

为了验证本文改进的算法模型在实际应用场景下的检测效果,在VisDrone测试集中选取不同场景下拍摄的图片进行测试。检测效果图4(a)是模糊失真场景,可以发现改进的模型能够精准地检测出远处的模糊车辆,展现出较好的鲁棒性和检测率。图4(b)是光线变化场景下的检测结果图,可以发现LM-YOLO模型受光线变化影响较小,对处在阴影下的行人和车依然有很高的检测精度。图4(c)是在小目标密集场景下的检测结果,可以看出LM-YOLO可以检测到被遮挡的摩托车,对小目标的检测精度较高,能够有效地应对实际场景下的检测需求。

为了评述本文所提算法LM-YOLO与原YOLOv5s算法在无人机小目标检测任务上的性能差异,选取VisDrone2019测试集中白天和夜间场景进行对比实验,可视化结果如图5所示。可以看出,与

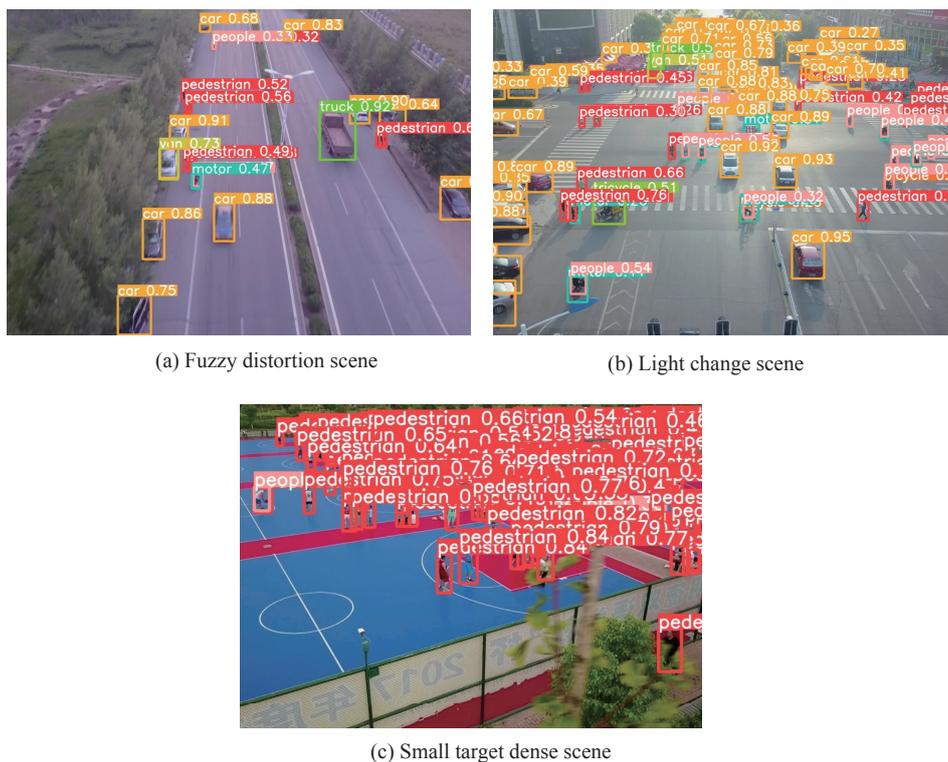


图4 LM-YOLO算法在不同场景检查结果  
Fig. 4 Detection effect of LM-YOLO algorithm in complex scenes

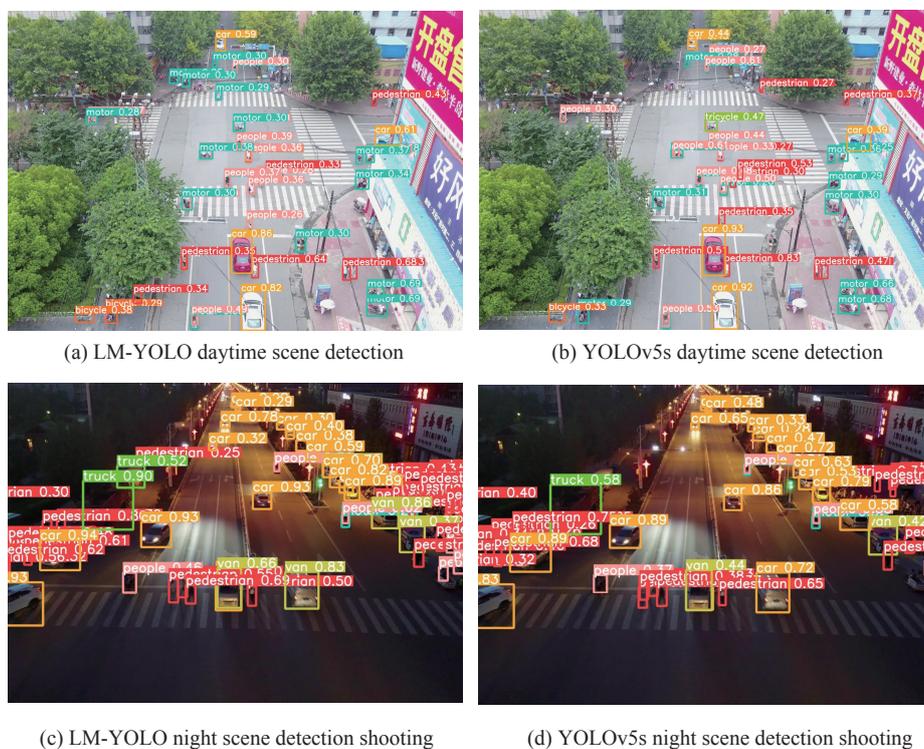


图5 LM-YOLO与YOLOv5s检测效果对比图  
Fig. 5 Comparison of detection effects between LM-YOLO algorithm and YOLOv5s

YOLOv5s相比,LM-YOLO对密集小目标检测性能更好,能精确检出被遮挡的小目标,漏检和误检情况较少。

### 3 结论

1) 提出了基于YOLOv5s的小目标检测算法LM-YOLO。针对无人机目标检测的轻量化要求,使用MobileNetV3-CBAM重构骨干网络,在获取更多小目标空间特征的同时,将模型大小压缩了45%,为目标检测模型向无人机部署提供了一定的理论依据。

2) 针对无人机图像背景复杂的特点,在特征融合网络引入使用LSK注意力模块,通过动态调整感受野,增强了模型对背景信息的提取能力,检测精度提升4.3%。针对无人机图像小目标数量多的问题,添加了小目标检测分支,并使用K-DBSCAN算法重新聚类锚框,改进后模型的检测精度达到了36.3%。

3) 在VisDrone数据集上的实验结果表明,LM-YOLO算法检测精度较YOLOv5s提升7.6%。该算法与多种主流的目标检测算法相比,在精度-检测速度方面拥有更好的均衡性,检测速度达到54 frame/s,满足了无人机检测的时效性要求。

#### 参考文献:

- [1] 奉志强,谢志军,包正伟,等.基于改进YOLOv5的无人机实时密集小目标检测算法[J].航空学报,2023,44(7):251-265.  
FENG Z Q XIE Z J, BAO Z W, et al. UAV real-time dense small target detection algorithm based on improved YOLOv5[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(7): 251-265.
- [2] 冷佳旭,莫梦竟成,周应华,等.无人机视角下的目标检测研究进展[J].中国图象图形学报,2023,28(9):2563-2586.  
LENG J X, MO M J C, ZHOU Y H, et al. Research progress of target detection from the perspective of UAV[J]. Journal of Image and Graphics, 2023, 28(9): 2563-2586.
- [3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Columbus: 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014.
- [4] GIRSHICK R. Fast R-CNN[C]//Santiago: 2015 International Conference on Computer Vision, 2015.
- [5] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] REDMON J, DIVVALA K S, GIRSHICK B R, et al. You only look once: Unified, real-time object detection [C]//Las Vegas: 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [7] WANG Q, WANG Z, LI B, et al. An improved YOLOv3 object detection network for mobile augmented reality [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [8] WANG C Y, BOCHKOVSKIY A, LIAO H Y M, et al. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//Vancouver: 2023 IEEE Conference on Computer Vision and Pattern Recognition, 2023.
- [9] GAI R, CHEN N, YUAN H. A detection algorithm for cherry fruits based on the improved YOLOv4 model[J]. Neural Computing and Applications, 2023, 35(19): 13895-13906.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Berlin: 2016 European Conference on Computer Vision, 2016.
- [11] SUN J, YANG Y, HE X F, et al. Northern maize leaf blight detection under complex field environment based on deep learning[J]. IEEE Access, 2020, 8: 33679-33688.
- [12] JIANG J, YANG Z, WU C, et al. A compatible detector based on improved YOLOv5 for hydropower device detection in AR inspection system[J]. Expert Systems With Applications, 2023, 225: 120065.
- [13] 吴萌萌,张泽斌,宋尧哲,等.基于自适应特征增强的小目标检测网络[J].激光与光电子学进展,2023,60(6):65-72.  
WU M M, ZHANG Z B, SONG Y Z, et al. Small target detection network based on adaptive feature enhancement [J]. Laser & Optoelectronics Progress, 2023, 60(6): 65-72.
- [14] 谢椿辉,吴金明,徐怀宇.改进YOLOv5的无人机影像小目标检测算法[J].计算机工程与应用,2023,59(9):198-206.  
XIE C H, WU J M, XU H Y. Small object detection algorithm based on improved YOLOv5 in UAV image[J].

- Computer Engineering and Applications. 2023, 59(9): 198-206.
- [15] QIU M L, HUANG L, TANG B H, et al. ASFF-YOLOv5: Multielement detection method for road traffic in UAV images based on multiscale feature fusion[J]. Remote Sensing, 2022, 14(12): 3498.
- [16] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3[C]//Seoul: 2019 IEEE/CVF International Conference on Computer Vision, 2019.
- [17] SANDLER M, HOWAD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks [C]//Salt Lake: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.
- [18] MODHA D S, SPANGLER W S. Feature weighting in  $k$ -means clustering[J]. Machine Learning, 2003, 52(3): 217-237.
- [19] 耿超, 彭余华. 基于动态分段和DBSCAN算法的交通事故黑点路段鉴别方法[J]. 长安大学学报(自然科学版), 2018, 38(5): 131-138.
- GENG C, PENG Y H. Traffic accident black spot identification method based on dynamic segmentation and DBSCAN algorithm[J]. Journal of Chang'an University(Natural Science Edition), 2018, 38(5): 131-138.
- [20] 宋非洋, 吴黎明, 郑耿哲, 等. 基于MobileNetV3的结构性剪枝优化[J]. 自动化与信息工程, 2019, 40(6): 20-25.
- SONG F Y, WU L M, ZHENG G Z, et al. Optimization of structural pruning based on MobileNetV3[J]. Automation & Information Engineering, 2019, 40(6): 20-25.
- [21] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Nashville: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Munich: 2018 European Conference on Computer Vision, 2018.
- [23] 王欣, 王美丽, 边党伟, 等. 融合MobileNetv2和注意力机制的轻量级人像分割算法[J]. 计算机工程与应用, 2022, 58(7): 220-228.
- WANG X, WANG M L, BIAN D W, et al. Algorithm for portrait segmentation combined with MobileNetv2 and attention mechanism[J]. Computer Engineering and Applications, 2022, 58(7): 220-228.
- [24] 范美楷, 方志, 晏宇, 等. 一种基于LSKNet的绝缘子缺陷检测方法研究[J]. 通信与信息技术, 2024(1):49-54.
- FAN M K, FANG Z, YAN Y, et al. Research on an insulator defect detection method based on LSKNet[J]. Communications and Information Technology, 2024(1): 49-54.



第一作者:董华军(1978—),男,教授,博士,博士生导师,研究方向为机器视觉技术应用。E-mail: huajundong4025@163.com。



通信作者:王宇栖(1998—),男,硕士研究生,研究方向为机器视觉与智能制造。E-mail: 1289511017@qq.com。

(责任编辑:吴海燕)