

# 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】实验验证了本文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

## 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, a small target detection algorithm LM-YOLO based on YOLOV5 is proposed. 【Method】Firstly, increase the small target detection head and K-DBSCAN clustering algorithm is 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 is designed as a feature extraction network to reduce the size of the network model. Finally, the large kernel selective attention mechanism LSK is 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.

**Keywords:** Uav image; small target detection; clustering algorithm; YOLOv5s; attention mechanism

【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。Yang<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。Xie 等人<sup>[13]</sup>在 YOLOv5 的基础上增加专门检测小目标的预测层,

对特征融合层进行改进，重新分配权重并优化了损失函数，但新的特征融合层不能自适应分配权重，提取到的小目标特征信息较少。谢椿辉等人<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等人<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。**【关键问题】**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(Backbone)、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。

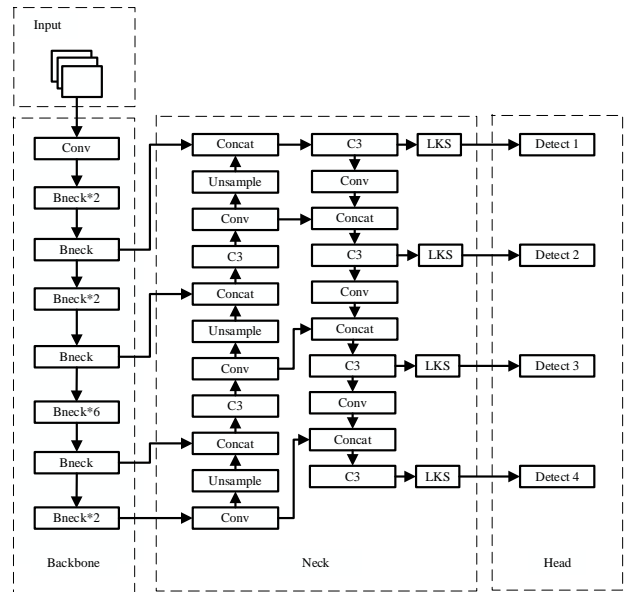


图1 LM-YOLO网络结构图

Fig.1 LM-YOLO network structure diagram

### 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<sup>[19]</sup>。因此本文使用了基于密度的K-DBSCAN算法(K-means and DBSCAN)对目标真实框进行聚类，其核心思想是使用DBSCAN<sup>[20]</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。

表1 预设锚框尺度分配表

Tab.1 anchor boxes scale assignment

目标大小	锚框尺度		
极小目标	(2, 5)	(4, 8)	(5, 6)
小目标	(7, 12)	(10, 10)	(9, 18)
中目标	(20, 18)	(15, 31)	(34, 42)
大目标	(28, 40)	(61, 46)	(59, 101)

### 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,为模型未来植入无人机嵌入式设备做准备,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<sup>[21]</sup>包含了 MobileNetV3-large 和 MobileNetV3-small 两个版本,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,这四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 层、第 6 层、第 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>[22]</sup>,而忽视了小目标检测所需的空间特征,在文献<sup>[23]</sup>中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。

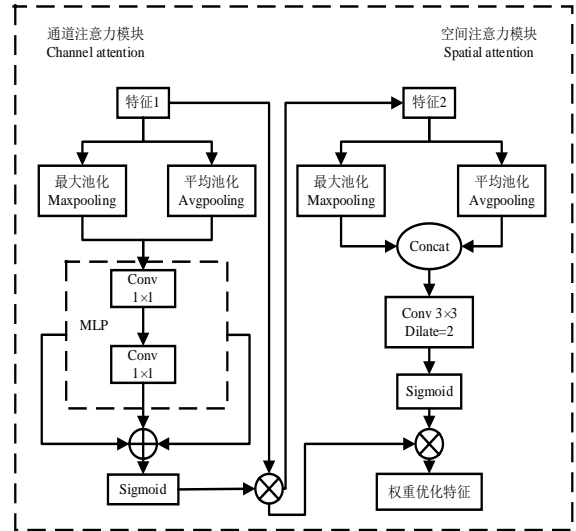


图 2 CBAM 注意力机制

Fig.2 CBAM attention mechanism

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>[24]</sup>,经过 Sigmoid 权重归一化后与原始图进行特征进行相乘,最终生成权重优化的特征图。本文在 MobileNetV3-large 的 Bneck 层嵌入了 CBAM 模块,并删去了原有的通道注意力模块,设计了 MobileNetV3-CBAM 网络作为 YOLOv5s 的 Backbone 网络,该网络结构如表 2 所示。

表 2 MobileNetV3-CBAM 网络结构

Tab.2 MobileNetV3-CBAM network structure

序号	输入尺寸	操作类型	扩展尺寸	输出通道数	CBAM 模块	激活函数	步长
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

#### 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, 才能正确的进行识别。LSKNet<sup>[25]</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的图像特征的能力, 从而提高航拍小目标的检测精度。

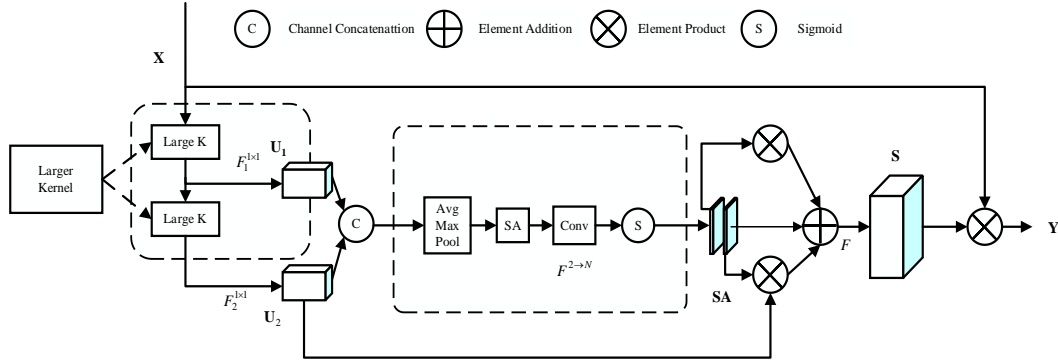


图3 LSK 模块结构示意图

Fig.3 Schematic diagram of LSK module structure

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  $\mathbf{X}$  依次通过这些小卷积核, 并在每个小卷积核后添加一个  $1 \times 1$  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:

$$\mathbf{U}_0 = \mathbf{X}, \mathbf{U}_{i+1} = F_i^{dw}(\mathbf{U}_i) \quad (1)$$

$$\mathbf{U}_i = F_i^{1 \times 1} \mathbf{U}_i, \text{for } i \text{ in } 1, N \quad (2)$$

其中  $F_i^{dw}()$  指第  $i$  个卷积核的深度可分离卷积操作 (Depth-wise),  $\mathbf{U}_i$  指提取后的特征图。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。首先要对来自不同感受野的特征图进行拼接:

$$\mathbf{U} = [\mathbf{U}_1, \dots, \mathbf{U}_i] \quad (3)$$

然后对拼接后的特征图  $\mathbf{U}$  进行最大池化和平均池化操作:

$$\mathbf{SA}_{avg} = P_{avg}(\mathbf{U}), \mathbf{SA}_{max} = P_{max}(\mathbf{U}) \quad (4)$$

其中  $P_{avg}()$  和  $P_{max}()$  分别代表平均池化和最大池化操作,  $\mathbf{SA}_{avg}$  和  $\mathbf{SA}_{max}$  分别代表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  $F^{2 \rightarrow N}(\bullet)$  将两通道的池化特征转换为  $N$  通道空间注意力:

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

对于每个空间注意力图  $\mathbf{SA}$ , 采用 sigmoid 激活函数得到每个分解大核的独特空间选择掩膜, 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  $F(\bullet)$  进行融合, 得到注意力特征  $\mathbf{S}$ 。

$$\mathbf{SA}_i = \sigma(\mathbf{SA}_i) \quad (6)$$

$$\mathbf{S} = F\left(\sum_{i=1}^N (\mathbf{SA}_i \cdot \mathbf{U}_i)\right) \quad (7)$$

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

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{S} \quad (8)$$

## 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 Windows10 系统下的 Pytorch 深度习框架, 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 VisDrone2019<sup>[26]</sup>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 540000 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,需要考虑每秒传输帧数 FPS 和模型的大小。

在本试验中,传输帧数 FPS 表达式如公式(9)所示,平均检测精度  $mAP$  的表达式如公式(10)所示。

$$FPS = \frac{1000}{P_{reTime} + I_{nferTime} + N_{msTime}} \quad (9)$$

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

式中,  $P_{reTime}$  是图像的预处理时间,  $I_{nferTime}$  是网络推理时间,  $N_{msTime}$  是非极大值抑制时间,单位都是毫秒;  $n$  为检测的类别种类,  $P$  (Precision) 表示检测精确度,  $R$  (Recall) 表示检测召回率。

## 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。

表 3 改进 YOLOv5s 消融实验结果

Tab.3 Results of improved YOLOv5s ablation test

方	K-DBSCAN	MobileNetv3 -CBAM	小	LSK	mAP@0.5	模/MB	FPS
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

从表 3 中可看出,本文提出的模块均可提高算法的检测精度。在单独使用 K-DBSCAN 聚类算法时,模型  $mAP@0.5$  提升了 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,  $mAP@0.5$  提升了 0.4%,模型 FPS 获得了极大的提升,提升了 31 帧/s,同时模型的大小减少了 6.1MB;在单独增加检测头时,由于融合了浅层信息,模型的  $mAP@0.5$  有了 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 0.3MB, FPS 减少了 9 帧/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  $mAP@0.5$  提升了 4.3%,但模型大小增加了 0.7MB, FPS 下降了 17 帧/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。

表 4 不同算法性能对比

Tab.4 Comparison of the selected detection algorithm

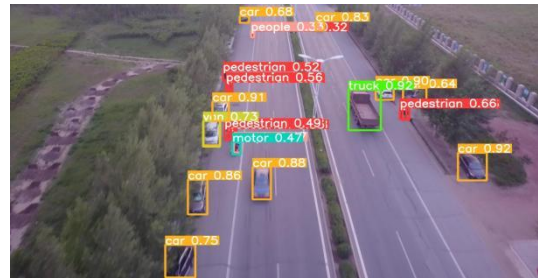
算法	mAP@0.5	FPS
FasterRCNN	22.8	6
SSD	29.1	35
Swin-transformer	31.4	53
YOLOX	24.6	49
YOLOV5m	32.9	46
CornerNet	23.7	15
Drone-YOLO	36.5	31
TPH-YOLOv5	37.2	22
LM-YOLO	36.3	54

从表 4 可以看出, 本文提出的 LM-YOLO 与 FasterRCNN、SSD、Swin-transformer、YOLOX、和 YOLOV5m、CornerNet 相比, mAP@0.5 分别提升了 13.5 个百分点、7.2 个百分点、4.9 个百分点、11.7 个百分点、3.4 个百分点、12.6 个百分点。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 13 帧/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, 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 算法可效性分析

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) 是在模糊失真场景下, 可以发现改进的模型, 能够精准的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) 是光线变化的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, 可以看出算法可以检测出很多小目标, 即使摩托车被遮

挡也能被算法精确的检测到, 能够有效地应对实际场景下的检测需求。



(a) Fuzzy distortion scene



(b) Light change scene

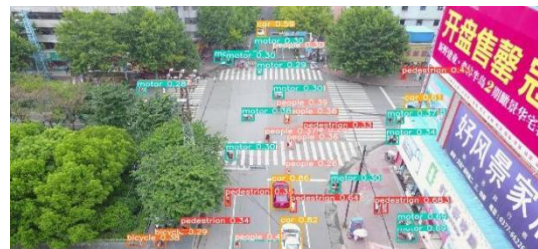


(c) Small target dense scene

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

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 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。



(a) LM-YOLO daytime scene detection



(b) YOLOv5s daytime scene detection



(c) LM-YOLO night scene detection shooting



(d) YOLOv5s night scene detection shooting

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

### 3 结论

1)本研究改进 YOLOv5s,提出了 LM-YOLO 算法。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,检测精度提升了 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 提升了 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,检测速度达到了 54 帧/s,满足了无人机检测的时效性要求。

4)本文所提出的改进方法虽然大幅提高了 YOLOv5s 对小目标的检测精度,但也导致了检测速度下降了 42 帧/s。未来计划对模型进行轻量化研究,进一步降低模型的计算量和参数量,提升算法的检测速度。

#### 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(07):251-265.FENG Z Q XIE Z J, BAO Z W, et al.Real-time dense small object detection algorithm for UAV based on improved YOLOv5[J].Acta Aeronautica et Astronautica Sinica,2023,44(07):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(09):2563-2586.LENG J X, MO M J C, ZHOU Y H, et al.Recent advances in drone-view object detection. Journal of Image and Graphics, 2023,28 (09) : 2563-2586
- [3] GIRSHICK B R ,DONAHUE J ,DARRELL T,et al.Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[J].IEEE Computer Society,2014.
- [4] GIRSHICK R .Fast R-CNN[J].Computer Science, 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(06):1137-1149.
- [6] REDMON J ,DIVVALA K S ,GIRSHICK B R,et al.You Only Look Once: Unified, Real-Time Object Detection[C]//Computer Vision & Pattern Recognition.IEEE, 2016.
- [7] WANG Q , WANG Z , LI B ,et al.An Improved YOLOv3 Object Detection Network for Mobile Augmented Reality[C]//2021 IEEE 7th International Conference on Virtual Reality (ICVR).IEEE, 2021.
- [8] GE Z ,LIU S ,WANG F ,et al.YOLOX: Exceeding YOLO Series in 2021[J].2021.
- [9] WANG C Y ,BOCHKOVSKIY A ,LIAO H Y M .YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J].arXiv e-prints, 2022.
- [10] BERG A C , FU C Y , SZEGEDY C ,et al.SSD: Single Shot MultiBox Detector[J].2015.

- [11] ZHANG S , WEN L , BIAN X ,et al.Single-Shot Refinement Neural Network for Object Detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition.IEEE, 2018.
- [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 Syst. Appl.2023, 225:120065.
- [13] XIE L C,XUE Y L,JINZE Y E.UAV Aerial Photography Target Detection Algorithm Based On Improved YOLOv5[J].Journal of Physics: Conference Series,2022, 2284(01):
- [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(09):198-206.XIE C H,WU J M,XU H Y.Small Object Detection Algorithm Based on Improved YOLOv5 in UAV Image[J/OL].Computer Engineering and Applications. 2023,59(09):198-206.
- [15] QIU M L,HUANG L,TANG B H.ASFF-YOLOv5: Multielement Detection Method for Road Traffic in UAV Images Based on Multiscale Feature Fusion[J].Remote Sensing,2022,14(14):3498-3498.
- [16] HOWARD A G,ZHU M,CHEN B,et al.MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[J].2017.
- [17] SANDLER M,HOWAD A,ZHU M,et al. MobileNetV2:Inverted Residuals and Linear Bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).IEEE, 2018.
- [18] MODHA D S,SPANGLER W S.Feature Weighting in k-Means Clustering[J].Machine Learning, 2003,52(03): 217-237.
- [19] 王森,刘琛,邢帅杰.K-means 聚类算法研究综述[J].华东交通大学学报, 2022, 39(05): 119-126.WANG S,LIU C,XING S J.Review on K-means Clustering Algorithm [J]. Journal of East China Jiaotong University,2022, 39(05):119-126
- [20] GAN J ,TAO Y.DBSCAN Revisited: Mis-Claim, Un-Fixability, and Approximation[C]//ACM.ACM, 2015.
- [21] 刘永胜,熊吉光,游志杰,等.基于改进 SSD 算法的结构表面裂缝病害检测[J].华东交通大学学报, 2023, 40(06): 1-7.LIU Y S,XIONG J G,YOU Z J,et al.Structural Apparent Crack Disease Detection Based on Improved SSD Algorithm[J].Journal of East China Jiaotong University,2023,40(06):1-7.
- [22] WOO S, PARK J, LEE J Y,et al.CBAM: Convolutional Block Attention Module[J].Springer, Cham, 2018.
- [23] HOWAR A,SANDLER M,CHU G,et al.Searching for MobileNetV3[J].2019.
- [24] 张长乐,金钧.基于深度学习的绝缘子故障检测仿真研究[J].华东交通大学学报, 2023, 40(05): 41-48.ZHANG C L,JIN J.Simulation Study on Insulator Fault Detection Based on Deep Learning[J].Journal of East China Jiaotong University, 2023, 40(05): 41-48.
- [25] 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,(01):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,(01):49-54.
- [26] DU D, ZHU P,WEN L,et al.VisDrone-DET2019: The Vision Meets Drone Object Detection in Image Challenge Results[C]//ICCV visdrone workshop.2019.



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研究。E-mail:huajundong4025@163.com。



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