# 联合结构重参数和 YOLOv5 的航拍红外目标检测

邵延华[1*]，张兴平[1]，张晓强[1]，楚红雨[1]，吴亚东[2]

（1. 西南科技大学 信息工程学院，绵阳 621010；2. 四川轻化工大学 计算机科学与工程学院，宜宾 644000）

**摘要**　无人机进行红外航拍目标检测在交通、农业和军事等方面有着广泛应用。该领域的主要挑战有目标较小、相互遮挡、非刚体形变大以及红外成像纹理信息少、边缘特征弱等。针对以上问题，基于 YOLOv5 和结构重参数化优化思想，提出了一种针对航拍场景的目标检测模型 Rep-YOLO。首先，在主干网络中引入 RepVGG 模块，提升模型特征提取能力；在模型推理时对 RepVGG 模块的多分支进行结构重参数化，减少网络分支和结构复杂度。其次，结合数据特征，改进检测网络颈部的路径聚合网络，提升检测算法在机载平台的精度–速度均衡能力。最后，在两个公开红外数据集进行对比实验，表明该算法的有效性。以南航 ComNet 航拍数据集为例，统计结果显示主要检测指标各类平均精度（mean Average Precision，mAP）提升 5.9%，同时参数量和模型大小分别减少约 29.7% 和 23.2%。另外，对 Rep-YOLO 在典型机载平台 Jetson Nano 上进行了模型部署验证，为航拍场景的检测算法改进和实际应用提供了可靠的技术支撑。

**关键词**　深度学习；红外图像；航拍目标检测；YOLOv5；结构重参数化

**中图分类号**　TP39　　**文献标志码**　A　　DOI　10.12178/1001-0548.2022070

# Infrared Aerial Object Detection Based on Structural Re-Parameterization and YOLOv5

SHAO Yanhua[1*], ZHANG Xingping[1], ZHANG Xiaoqiang[1], CHU Hongyu[1], and WU Yadong[2]

(1. School of Information Engineering, Southwest University of Science and Engineering, Mianyang 621010, China;

2. School of Computer Science & Engineering, Sichuan University of Science & Engineering, Yibin 644000, China)

**Abstract**　Infrared aerial object detection has been widely used in transportation, agriculture, military security, and other areas. The main challenges are small objects, mutual occlusion, little texture information, weak edge features, and large deformation of non-rigid bodies. To address these problems, based on YOLOv5 and structural Re-Parameterization (Rep), an improved object detection network Rep-YOLO is proposed for infrared aerial object detection. Firstly, the RepVGG module is introduced in the backbone network to improve the model feature extraction capability. During the model inference, the branches of the RepVGG module are structurally re-parameterized to reduce the branch and the complexity of the network structure. Secondly, the path aggregation network (PANet) in the neck of the detection network is improved by combining the priori feature, to increase the accuracy and speed balance capability. Finally, experiments are conducted on two publicly available infrared datasets, showing that the algorithm can effectively detect aerial infrared objects. Compared with the baseline (YOLOv5s), the statistical results on ComNet dataset show the mean Average Precision (mAP) is increased by 5.9%, while the parameters and model size are reduced by about 29.7% and 23.2%, respectively. In addition, the model deployment verification of our Rep-YOLO is carried out on the airborne platform Jetson Nano. It provides reliable technical support for the improvement of the detection algorithm and its practical application with UAV platforms.

**Key words**　deep learning; infrared imaging; aerial object detection; YOLOv5; structural re-parameterization

收稿日期：2022−03−10；修回日期：2022−08−31

基金项目：国家自然科学基金（6160382）；国防科工项目（20zg6108）；四川省科技厅项目（2019YJ0325, 2020YFG0148）

作者简介：邵延华，博士，主要从事计算机视觉方面的研究。

*通信作者 E-mail: syh@cqu.edu.cn

目标检测是计算机视觉的一项重要研究课题。无人机（Unmanned Aerial Vehicles, UAV）航拍目标检测在交通[1]、农业[2]和军事[3]等领域都有广泛的应用。航拍场景中，目标具有姿态变化快、遮挡严重和不规则运动等特点，同时考虑到监控视频场景的多样性以及景深、分辨率、天气、光照等条件变化，准确鲁棒的航拍目标检测仍然是一项具有挑战性的任务[3-4]。

目前，目标检测任务主要面向可见光图像，易受环境照明、复杂背景等外部因素影响。红外成像可在夜间或弱光环境下工作，避免了光照因素的影响[4]。依赖手工设计特征的传统检测方法鲁棒性差，无法应对复杂场景的航拍目标检测任务[5-6]。

当前，基于深度学习的算法是红外航拍目标检测的主流[7]。算法主要分为两类，一类是以 R-CNN[8]系列为代表的双阶段模型，另一类是以 YOLO（You Look Only Once）系列[9-12]为代表的单阶段模型。综合考虑检测速度、模型大小和机载平台的计算能力与功耗，后者是航拍目标检测的主流[4]。文献 [4] 提出 ComNet 航拍红外数据集，使用 BASNet 生成显著图对航拍红外图像进行图像增强，作为 YOLO 的注意力机制，提高了检测性能。文献 [13] 提出一种改进的 SSD 算法，将密集连接网络作为主干网络，并构建特征金字塔网络实现遥感目标检测。

卷积神经网络（Convolutional Neural Network, CNN）的结构设计是深度模型的关键内容之一[14-18]。为提高网络的特征提取能力，常见设计是构建更深更宽，或更多分支的拓扑结构，如 Inception 模块中的旁支、残差网络（Residual Network, ResNet）的残差分支等，尽管复杂结构的准确率较高，但会降低推理速度。对标复杂、多分支网络的精度，直接训练一个无分支的简单网络非常困难[14]。

因此，业界提出了结构重参数化（Re-parameterization, Rep），其基本思想是通过参数的等效转换，将复杂的拓扑结构线性变换为简单的无分枝结构[14]。文献 [15] 提出 ResRep[15]，基于模型重参数化技术，通过解耦重要参数与不重要参数实现模型剪枝。文献 [14] 提出 RepVGG，训练时采用具有多分枝的拓扑结构，推理结构类似于 VGG 的简约拓扑结构，便于优化、部署和应用。

针对红外航拍图像中目标尺寸小、纹理信息少和边缘特征弱等问题[19]，本文以典型 YOLO 的最新版 YOLOv5s[12,17]为基准模型，基于重参数化构建 Rep-YOLO。

# 1　YOLOv5

YOLO[9-12]系列是最具代表性的单阶段目标检测算法之一，它直接对图像进行特征提取，将目标检测作为回归任务。YOLOv2[9]引入锚框（Anchor）机制，解决了 YOLOv1 检测目标少，小目标表现欠佳的问题。YOLOv3[10]进一步提升了速度−精度的均衡能力，是目标检测的里程碑算法之一。YOLOv4[11]采用 CSP[18]（Cross Stage Partial）结构提高推理速度，颈部网络引入空间金字塔池化（Spatial Pyramid Pooling, SPP）与 PANet[20]，进一步提升检测效果。

YOLOv5 是 YOLO 系列算法的最新发展[12,17]，与 YOLOv4[11]相比，二者精度基本相当，但 YOLOv5 模型尺寸更小、部署成本更低、运算速度更快[17]。YOLOv5 在 YOLO 的预处理、输入端、主干网络、颈部网络和输出端等基本模块分别进行了改进。

预处理阶段采用 Mosaic 数据增强提升数据多样性，并采用自适应锚框获取更优锚框。YOLOv5 的主干和颈部采用基于 CSP 的基本组件 C3，替换常用的 CSP-Bottleneck[18]模块。C3 由 3 个卷积和多个 Bottleneck 模块级联组成，可精简网络结构，减少计算量。颈部网络与 YOLOv4 类似，采用特征金字塔（FPN）和路径聚合网络（PANet）相结合的方式进一步提升特征提取的鲁棒性。后处理阶段采用加权非极大值抑制（Non-Maximum Suppression, NMS）。

# 2　Rep-YOLO 方法

训练时本文 Rep-YOLO 的整体网络结构如图 1 所示。

## 2.1　使用结构重参数化改进主干网络

为提升特征提取能力，在本文 Rep-YOLO 训练时，采用 RepVGG[14]模块代替 YOLOv5 主干网络中的所有单独的卷积（Conv）模块，代替后的主干网络示意图如图 2a 所示。推理时，将训练模型中 RepVGG 模块的对应参数通过参数等效变换转化成 3×3 卷积，推理时的主干网络示意图如图 2b 所示。

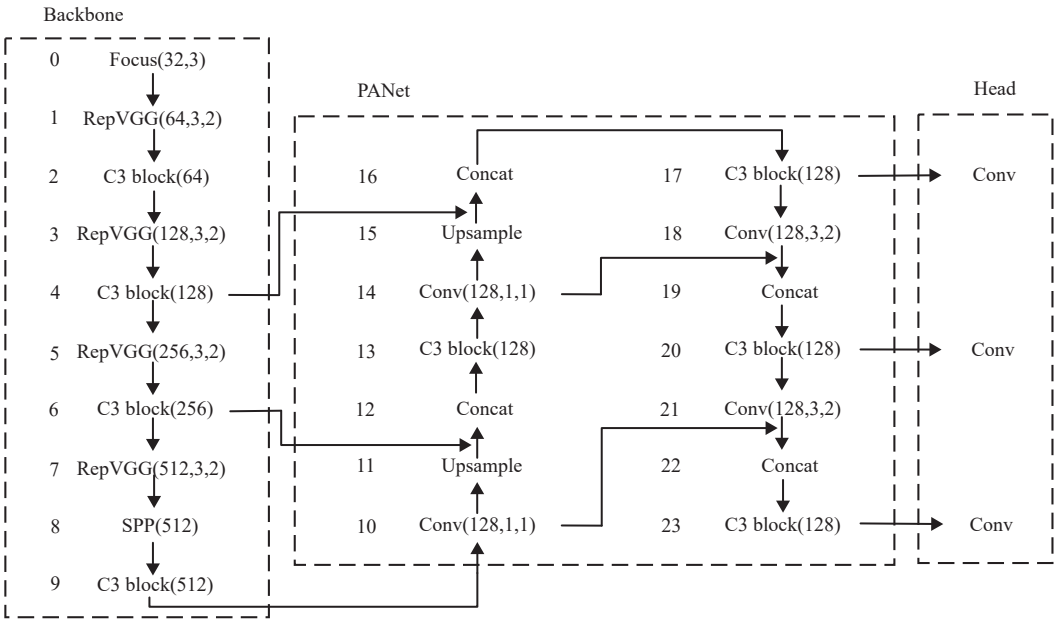
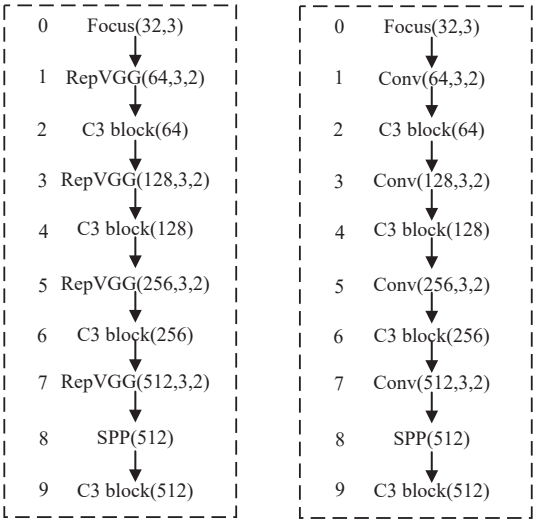
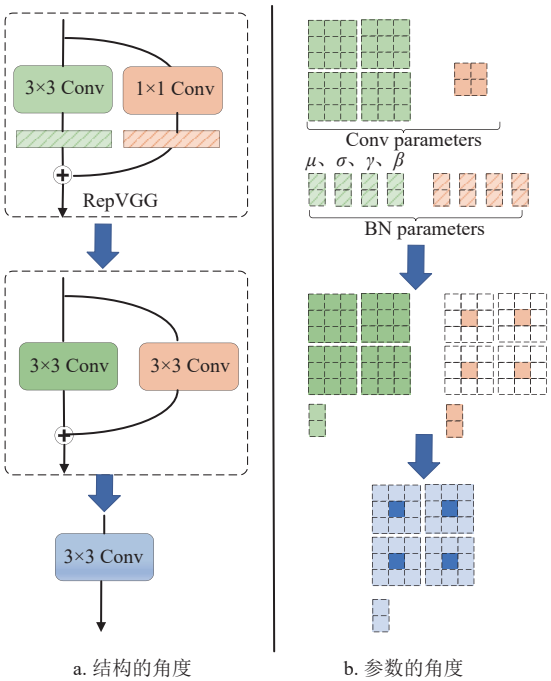



图 1 Rep-YOLO 的网络结构




a. 训练时的主干网络 b. 测试时的网络结构

图 2 Rep-YOLOv5s 的主干网络结构




a. 结构的角度 b. 参数的角度

图 3 RepVGG 模块结构重参数化过程

本文采用的 RepVGG 模块由 1 个3×3卷积、1 个1×1卷积、2 个批归一化层（BN）和 2 个 ReLU 激活函数并行连接组成，具体结构以及相应的参数等效转换过程如图 3 所示。

先将 RepVGG 模块的3×3卷积分支上的 Conv-BN 结构等效变换为一个带偏置的3×3卷积。针对1×1卷积分支，先进行 Conv-BN 结构的等效变换，再通过零填充操作转换为3×3卷积，转换完成后得到两个并行的3×3卷积分支。然后对上述两个分支的参数进行点对点相加，从而将 RepVGG 模块的参数等效转换为3×3卷积。

从数学角度分析，将卷积–批归一化–激活损失函数 CBL 结构等效转换为卷积层的过程如下：假设$K \in \mathbb{R}^{C_2 \times C_1 \times k \times k}$代表 RepVGG 模块的卷积核参数，其中，$C_1$为输入通道数，$C_2$为输出通道，$k$为卷积核大小，一般取 3 或 1。模型输入输出分别为$\boldsymbol{I} \in \mathbb{R}^{N \times C_1 \times H \times W}$和$\boldsymbol{O} \in \mathbb{R}^{N \times C_2 \times H' \times W'}$，$N$代表样本数，$W$和$H$表示输入特征图的宽和高，$W'$和$H'$表示输出特征图的宽和高。卷积运算可表示为：

$$\boldsymbol{O} = \boldsymbol{I} * K + B(b)$$

式中，$b \in \mathbb{R}_2^c$表示偏置；$B$是广播函数，可将偏置$b$映射到$N \times C_2 \times H' \times W'$；$*$代表卷积操作。针对卷积层连接一个 BN 层，假设 BN 层的参数包括均值$\mu$、方差$\sigma$、尺度因子$\gamma$和偏置$\beta$，那么 Conv-BN 结构的第$i$个通道的输出$O'_{:,i,:,:}$表示为：

$$O'_{:,i,:,:} = ((\boldsymbol{I} * K)_{:,i,:,:} - \mu_i)\frac{\gamma_i}{\sigma_i} + \beta_i \quad i \in [1, C_2]$$

假设$K'$和$b'_i$表示上述公式等效的卷积核参数核和偏置参数，那么：

$$K' = \frac{\gamma_i}{\sigma_i} K_{:,i,:,:} b'_i = -\frac{\mu_i \gamma_i}{\sigma_i} + \beta_i$$

最终转换后的卷积层输出为：

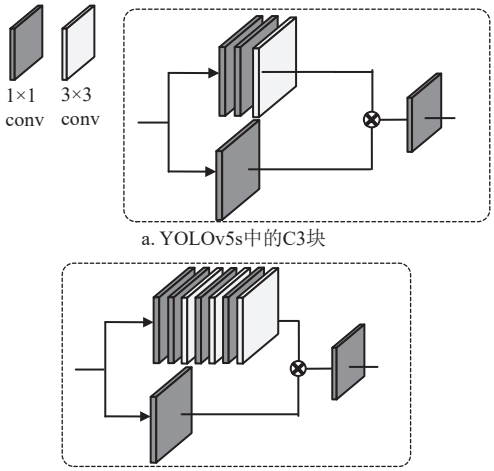
$$O'_{:,i,:,:} = (\boldsymbol{I} * K')_{:,i,:,:} + b'_i \quad i \in [1, C_2]$$

经上述变换，可将两个分支分别等效转换为一个 3×3 卷积和一个 1×1 卷积。针对 1×1 卷积分支，通过补零的方式转换成 3×3 卷积参数，最终两个分支的 3×3 卷积参数进行点对点相加，实现将 RepVGG 结构转换为 3×3 卷积层，详细流程参见图 3。

### 2.2　改进颈部网络

YOLOv5s 的颈部网络中的主要结构为 C3 块，是 CSP-Bottleneck[18] 的改进版本，C3 块更简单、更轻量化，二者在相似损耗的情况下，C3 块可取得更好的效果。采用多分支的 C3 结构如图 4a 所示，主要包括两个分支，次分支只有 1×1 卷积，主分支包括 2 个 1×1 卷积和 1 个 3×3 卷积，两分支经合并（Concat）之后再连接一个 1×1 卷积。

红外图像纹理信息较少，为增强底层信息传播能力和更好地进行特征聚合，本文在原始 C3 块主分支上增加了两个 1×1 卷积和 3×3 卷积结构，如图 4b 所示。同时，C3 块中主分支上卷积层和通道的增加会占用更多缓存空间以及增加运行时间。应避免使用通道数过多的 C3 块，保持检测速度优势。对场景数据分析可知，针对红外航拍场景的目标检测，相较于 COCO 等大型数据集，航拍场景中需检测的类别较少，因此可适当减小颈部的网络宽度。本文减少了整个颈部 PANet 的输入通道数，将各个 C3 模块的输入通道数量修改为 128，在一定程度上减少了参数量，提高检测网络的速度。




a. YOLOv5s中的C3块

b. 本文的C3块

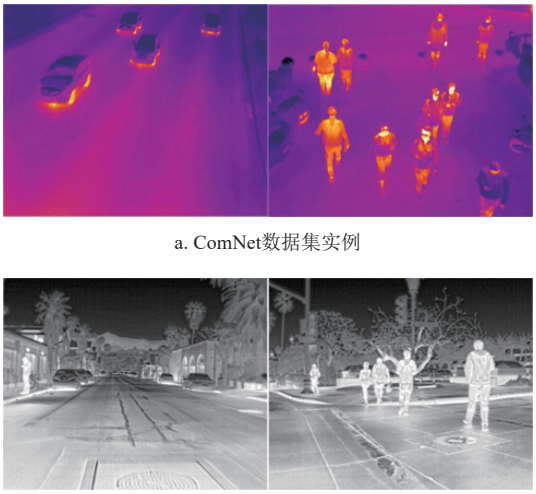
图 4　PANet 中的 C3 块

## 3　实验结果与分析

### 3.1　数据集

为验证本文算法的性能，在两个公开红外航拍数据集上进行对比实验，分别为 ComNet[4, 16] 数据集和 FLIR 数据集[19]。

ComNet[16] 红外航拍数据集针对无人机载热红外场景下行人及车辆的目标检测，包括 2 975 张红外图像，分辨率为 640×512。使用 DJIM600 PRO 无人机搭载 FLIR 热红外相机 Vue Pro 采集，无人机飞行高度为 20～40 m。将热红外图像经过温度映射后转化成 RGB 三通道的伪彩色图像，目标尺度变化大，多数目标的大小为 $10^3$～$10^4$ 像素，平均每张图片 3 个样本，数据集部分实例如图 5a 所示。



a. ComNet数据集实例

b. FLIR数据集实例

图 5　数据集实例

FLIR[19] 红外数据集主要面向高级驾驶辅助场景，其中白天场景的图像占比为 60%，夜间场景的图像占比为 40% [19]，实例如图 5b 所示。图像大小为 512×640，平均每张图像中包括 8 个目标。其中，“person”“bicycle”和“car”类目标大小均值像素数分别为 50×20、43×30、43×49。场景中目标尺度变化大，以“car”类为例，其目标像素大小范围为 $10^2$～$10^6$。两个数据集的训练集与测试集划分如表 1 所示。

**表 1　ComNet 与 FLIR 红外数据集划分**

| Dataset | ComNet[16] | | FLIR[19] | |
|---|---|---|---|---|
| | 样本数 | 实例数 | 样本数 | 实例数 |
| 训练集 | 2 434 | 6 744 | 8 862 | 68 621 |
| 测试集 | 541 | 1 880 | 1 366 | 11 688 |

### 3.2　实验环境及评估方法

实验操作系统为 Windows 10 （×64），采用的深度学习框架为 Pytorch，GPU 为 NVIDIA GeForce RTX3060Ti （8 GB），采用的软件包括 Python 3.7、CUDA 11.0。

目标检测场景中，COCO 数据集[21] 的平均精度（Average Precision, AP）是最主流的评价指标[22]。AP 指标中的各类平均精度（mean AP, mAP）最为常用。其中 mAP@0.5 来自典型的 VOC 数据集，表示将检测器交并比 （Intersection Over Union, IOU） 阈值设为 0.5 时，计算每一类目标的 AP 得到的均值，能够充分衡量算法的特征提取能力。mAP@[.5:.95] 来自 COCO 数据集，表示在 IOU 阈值为 0.5～0.95，步长为 0.05 时每一类目标 AP 值的均值[21]。

### 3.3　ComNet 数据集上的对比实验

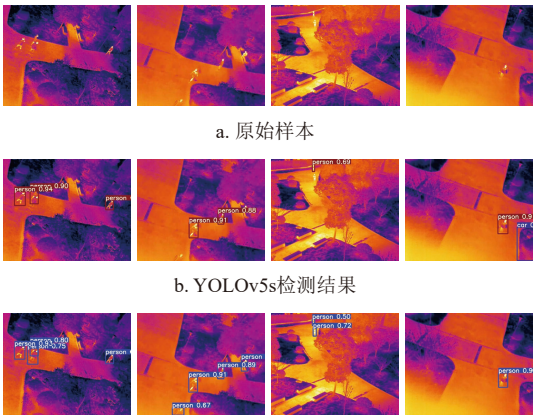
为了验证算法的有效性，将本文算法 Rep-YOLO 与主流目标检测算法在 ComNet 数据集[16] 分别进行测试。使用基于大型数据集 COCO 的预训练权重，使用随机梯度下降法优化学习率，batch 设置为 16，学习率设置为 0.01，epoch 设置为 100，输入大小为 640×640。

表 2 展示了本文算法与主流目标检测算法在 ComNet 数据集上的表现。本文算法在各个精度指标上表现更优，其中在当前更主流的 mAP@[.5:.95] 指标上对比尤其明显，与原始 YOLOv5s 比较，由 75.5% 增加到 81.4%，说明本文 Rep-YOLO 检测目标的置信度更高，提取特征信息能力更强。

**表 2　在 ComNet 数据集上的结果比较**

| 模型 | AP/% | | mAP@0.5/% | mAP@[.5:.95]/% |
|---|---|---|---|---|
| | Person | Car | | |
| YOLOv3-tiny | 79.9 | 99.4 | 89.7 | 52.3 |
| YOLOv3[10] | 88.8 | 99.8 | 94.3 | 62.5 |
| ComNet[16] | 90.3 | 93.0 | 91.7 | — |
| Faster-RCNN | 93.5 | 98.8 | 96.1 | 66.7 |
| Mask-RCNN[8] | 94.3 | 99.9 | 97.1 | 72.7 |
| YOLOv5s | 98.3 | 99.6 | 99.0 | 75.5 |
| Rep-YOLOv5s | **99.4** | **99.6** | **99.5** | **81.4** |

为更直观地观察两个算法的检测差异，图 6 给出了不同典型场景和典型对象的检测结果。图 6a 为原始图像，图 6b 为 YOLOv5s 的检测结果，图 6c 为本文 Rep-YOLO 的检测结果，图中数字表示检测结果的置信度，且 IOU 阈值都取 0.5。图 6 样本从左到右分别代表红外航拍场景中的典型挑战：1）相似目标互相遮挡；2）非刚性目标形变大；3）纹理缺失仅依靠亮度信息；4）遮挡与出视野。



a. 原始样本

b. YOLOv5s检测结果

c. 本文算法检测结果

图 6　ComNet 数据集检测结果对比

由图 6a 左边第一张图可知，该场景存在相似目标互相遮挡，如图 6c 左边第一张所示，本文算法可检测存在遮挡的目标。航拍人体目标尺度变化大，这对目标检测提出更高的挑战。由 6a 第二张图可知，本文 Rep-YOLO 可检测出路边长椅上形变较大的“person”目标，而 YOLOv5s 漏检。由图 6a 第三张可知，详细分析道路上的“person”目标区域发现，该区域背景亮度和目标亮度比较接近，再加上纹理信息缺失，造成 YOLOv5s 漏检，如图 6b 第三张所示。航拍目标检测中的另一类主

要挑战“出视野”，该问题和遮挡类似。造成目标有效区域面积过小，易引起误检。如图 6b 第四张图中，YOLOv5s 将路边的花坛误检为“car”。

因此，由图 6 中示例可知，面向红外航拍场景，本文算法表现更优，更多的数据和分析见 https://github.com/sunriselqq/Rep-YOLOv5。

### 3.4　在 FLIR 数据集上的对比实验

为进一步验证本文 Rep-YOLO 的性能，同时在 FLIR 红外数据集上进行了消融实验。为了保证模型测试的公平性，算法的训练参数一致，模型输入大小为 640×640，初始学习率为 0.01。为分析本文提出的两个改进对 YOLOv5 性能的影响，设计消融实验对不同的改进结果进行分析，不同改进方法对算法的影响对比如表 3 所示。其中“√”代表在网络中采用的对应的策略，“×”则代表未采用该策略。

**表 3　在 FLIR 数据集上的消融实验**

| 方法 | | | 结果 | |
|---|---|---|---|---|
| YOLOv5s（baseline） | RepVGG | Improved Neck | mAP@0.5/% | Model volume/MB |
| √ | × | × | 79.7 | 14.2 |
| √ | √ | × | 79.9 | 14.8 |
| √ | × | √ | 79.9 | 10.9 |
| √ | √ | √ | 80.4 | 10.9 |

由表 3 可知，训练时主干网络引入 RepVGG 结构，可增加主干网络提取信息的能力，提升 mAP 值，但模型体积不会增加。由于加深了颈部网络 C3 块，颈部网络利用网络的浅层信息将更加充分。再通过减少通道数目（宽度），使得模型占用内存减少。该改进的 C3 块是本文算法计算成本减少的主要来源之一。最终改进后算法的 mAP@0.5 为 80.4%，模型大小减少了 3.3 MB，且在训练过程中需要的显存更小。

同时在 FLIR 数据集上与目前典型的目标检测算法在多个指标上进行了比较，结果如表 4 所示。

文献 [23] 中， MMTOD-UNIT 通过图像转换框架将红外图像生成配对的伪 RBG 图，构造“伪多模态”目标检测器提升红外图像检测性能，在基线模型 Faster-RCNN 的基础上，精度增加了 7% 以上。ThermalDet[24] 是在 RefineDet 的基础上，通过融合不同层次的特征提高检测效果。YOLO-FIR[25] 在 YOLOv5 的基础上进行改进，通过压缩通道、优化参数等方式获得较好的表现。实验结果表明，本文算法在各个指标上都实现了更先进的检测结果，在不增加模型体积的情况下，进一步提高了检测精度，与之前的工作比较，Rep-YOLO mAP@0.5 检测结果达到 80.4%，从图像中学习了更多的特征。

**表 4　不同检测模型在 FLIR 数据集的比较**

| 模型 | AP /% | | | mAP@0.5/% |
|---|---|---|---|---|
| | Person | Bicycle | Car | |
| Faster-RCNN[23] | 54.7 | 39.7 | 67.6 | 54.0 |
| SSD[26] | 62.0 | 45.0 | 75.6 | 60.9 |
| MMTOD-UNIT[23] | 64.5 | 49.4 | 70.7 | 61.5 |
| YOLOv3 | 73.3 | 49.2 | 84.3 | 68.9 |
| RefineDet[25] | 77.2 | 57.2 | 84.5 | 72.9 |
| ThermalDet[25] | 84.5 | 60.0 | 85.5 | 74.6 |
| YOLO-FIR[26] | 85.2 | 70.7 | 84.3 | 80.1 |
| YOLOv5s | 84.9 | 63.5 | 90.7 | 79.7 |
| Rep-YOLOv5s | **85.8** | **64.6** | **90.6** | **80.4** |

### 3.5　机载平台模型定量对比分析

针对无人机内存、算力和功耗有限的特点，在航拍检测任务中，机载算法的检测速度、计算量和参数量和模型大小等需综合考量[3, 12, 27]。典型的边缘计算平台有英伟达 Jetson TX2、Jetson Nano、大疆妙算和树莓派等。本节测试平台选择典型的边缘计算平台 Jetson Nano[28]，测试时不采用任何其他软件加速手段，随机选取 ComNet 数据集中多张图片分别对两个模型进行测试，取平均值作为时间开销，输入大小为640×640。表5 从参数量（Parameters）、十亿浮点运算数（Giga Floating Point Operations, GFLOPs）、模型大小 （Model size） 和图像检测耗时（Time overhead） 这 4 个方面对算法进行对比。

**表 5　算法在 Jetson Nano 上的定量对比**

| Method | Parameters/MB | GFLOPs | Model size/MB | Time overhead/ms |
|---|---|---|---|---|
| YOLOv5s | 7.54 | 16.4 | 14.2 | 201 |
| Rep-YOLO | 5.30 | 15.2 | 10.9 | 192 |

由表 5 可知，在参数量上，总体上本文 Rep-YOLO 比原始 YOLOv5s 减小了 2.24 MB；在浮点运算数方面，比原始 YOLOv5s 少了 0.8 GFLOPs；在模型大小方面，本文算法也表现更优，减小了 23.2%；在 Jetson Nano 上，本文算法比 YOLOv5s 检测耗时减少 10 ms。提出的改进 C3 块有效地实现了模型参数的压缩。本文算法在推理中进行结构重参数化后，实现了复杂网络拓扑结构向简单结构

的等价转换，参数量、浮点运算数、计算耗时都明显减少，该结构重参数化方法可以减少算法的计算花销。

## 4 结束语

针对航拍红外图像目标检测，本文提出了以结构重参数化思想为核心的 Rep-YOLO 网络。以 YOLOv5s 为基础网络，在训练时的主干网络中采用具有多分支的 RepVGG 模块替换典型的 CBL 模块，以提升检测精度，在推理阶段将 RepVGG 重参数化为典型的 3×3 卷积。同时，结合数据特点，对颈部网络 PANet 进行通道缩减的同时增加网络深度以保持良好的性能。实验表明，和原始 YOLOv5s 相比，本文方法在两个公开数据集上的检测性能都有了明显提升，并且计算量更少，模型更小。机载平台的部署实验也验证了本文 Rep-YOLO 的比较优势，表明本文算法可以更好地应用于无人机航拍目标检测场景。当然，面向更具体的场景和任务，工程上可以结合其他剪枝或量化等加速技术进一步提升模型性能。

### 参 考 文 献


[1] 柳长源, 王琪, 毕晓君. 多目标小尺度车辆目标检测方法[J]. 控制与决策, 2021, 36(11): 2707-2712.
LIU C Y, WANG Q, BI X J. Multi-target and small-scale vehicle target detection method[J]. Control and Decision, 2021, 36(11): 2707-2712.

[2] 杨蜀秦, 刘江川, 徐可可, 等. 基于改进 CenterNet 的玉米雄蕊无人机遥感图像识别[J]. 农业机械学报, 2021, 52(9): 206-212.
YANG S Q, LIU J C, XU K K, et al. Improved CenterNet based maize tassel recognition for UAV remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 206-212.

[3] 吕洋, 康童娜, 潘泉, 等. 无人机感知与规避: 概念、技术与系统[J]. 中国科学: 信息科学, 2019, 49: 520-537.
LYU Y, KANG T N, PAN Q, et al. UAV sense and avoidance: Concepts, technologies, and systems[J]. Sci Sin Inform, 2019, 49: 520-537

[4] 赵兴科, 李明磊, 张弓, 等. 基于显著图融合的无人机载热红外图像目标检测方法[J]. 自动化学报, 2021, 47(9): 2120-2131.
ZHAO X K, LI M L, ZHANG G, et al. Object detection method based on saliency map fusion for UAV-borne thermal images[J]. Acta Automatica Sinice, 2021, 47(9): 2120-2131.

[5] ZHANG L, WU B, NEVATIA R. Pedestrian detection in infrared images based on local shape features[C]//2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Minneapolis: IEEE, 2007: 1-8.

[6] WU X, LI W, HONG D, et al. Deep learning for unmanned aerial vehicle-based object detection and tracking: A survey[J]. IEEE Geoscience and Remote Sensing Magazine, 2021, DOI: 10.1109/MGRS.2021.3115137.

[7] 许延雷, 梁继然, 董国军, 等. 基于改进 CenterNet 的航拍图像目标检测算法[J]. 激光与光电子学进展, 2021, 58(20): 192-201.
XU Y L, LIANG J R, DONG G J, et al. Aerial image target detection algorithm based on lmproved centerNet[J]. Laser and Optoelectronics Progress, 2021, 58(20): 192-201.

[8] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Trans Pattern Anal Mach Intell, 2020, 42(2): 386-397.

[9] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 7263-7271.

[10] REDMON J, FARHADI A. Yolov3: An incremental improvement[EB/OL]. [2022-01-21]. https://arxiv.org/abs/1804.02767.

[11] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [EB/OL]. [2022-01-21]. https://arxiv.org/abs/2004.10934.

[12] ZHU X, LYU S, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops. Montreal: IEEE, 2021: 2778-2788.

[13] 王俊强, 李建胜, 周学文, 等. 改进的 SSD 算法及其对遥感影像小目标检测性能的分析[J]. 光学学报, 2019, 39(6): 373-382.
WANG J Q, LI J S, ZHOU X W, et al. Improved SSD algorithm and its performance analysis of small target detection in remote sensing images[J]. Acta Optica Sinica, 2019, 39(6): 373-382

[14] DING X H, ZHANG X Y, MA N, et al. Repvgg: Making vgg-style convnets great again[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 13733-13742.

[15] DING X H, HAO T X, TAN J C, et al. ResRep: Lossless CNN pruning via decoupling remembering and forgetting [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 4510-4520.

[16] LI M L, ZHAO X K, LI J S, et al. ComNet: Combinational neural network for object detection in UAV-borne thermal images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(8): 6662-6673.

[17] 郭磊, 王邱龙, 薛伟, 等. 基于注意力机制的光线昏暗条件下口罩佩戴检测[J]. 电子科技大学学报, 2022, 51(1): 123-129.
GUO L, WANG Q L, XUE W, et al. Detection of mask

wearing in dim light based on attention mechanism[J]. Journal of University of Electronic Science and Technology of China, 2022, 51(1): 123-129.

[18] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. Seattle: IEEE, 2020: 390-391.

[19] LI C, XIA W, YAN Y, et al. Segmenting objects in day and night: Edge-conditioned cnn for thermal image semantic segmentation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(7): 3069-3082.

[20] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City: IEEE, 2018: 8759-8768.

[21] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context[C]//Proceedings of the European Conference on Computer Vision (ECCV). Zurich: Springer, 2014: 740-755.

[22] 何泽文, 张文生. 保持高分辨率信息的无锚点框检测算法[J]. 计算机辅助设计与图形学学报, 2021, 33(4): 580-589.
HE Z W, ZHANG W S. High resolution information reserved anchor-free detection algorithm[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(4): 580-589.

[23] DEVAGUPTAPU C, AKOLEKAR N, M SHARMA M, et al. Borrow from anywhere: Pseudo multi-modal object detection in thermal imagery[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Los Angeles: IEEE, 2019: 1029-1038.

[24] CAO Y, ZHOU T, ZHU X, et al. Every feature counts: An improved one-stage detector in thermal imagery [C]//2019 IEEE 5th International Conference on Computer and Communications (ICCC). Chengdu: IEEE, 2019: 1965-1969.

[25] LI S, LI Y, LI Y, et al. YOLO-FIRI: Improved YOLOv5 for infrared image object detection[J]. IEEE Access, 2021, 9: 141861-141875.

[26] MUNIR F, AZAM S, RAFIQUE M A, et al. Exploring thermal images for object detection in underexposure regions for autonomous driving[EB/OL]. [2022-02-03]. https://arxiv.org/ abs/:2006.00821.

[27] MURSHED M G S, MURPHY C, HOU D, et al. Machine learning at the network edge: A survey[J]. ACM Computing Surveys (CSUR), 2022, 54(8): 1-37.

[28] 李玉华, 刘全程, 李天华, 等. 基于 Jetson Nano 处理器的大蒜鳞芽朝向调整装置设计与试验[J]. 农业工程学报, 2021, 37(7): 35-42.
LI Y H, LIU Q C, LI T H, et al. Design and experiments of garlic bulbil orientation adjustment device using Jetson Nano processor[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(7): 35-42.

编　辑　税　红