



DOI: 10.12382/bgxb.2022.0736

# 复杂战场环境下改进 YOLOv5 军事目标识别算法研究

宋晓茹<sup>1\*</sup>, 刘康<sup>1</sup>, 高嵩<sup>1</sup>, 陈超波<sup>1,2</sup>, 阎坤<sup>1</sup>

(1. 西安工业大学 电子信息工程学院, 陕西 西安 710021; 2. 机电动态控制重点实验室, 陕西 西安 710065)

**摘要:** 复杂战场环境下军事目标识别技术是提升战场情报获取能力的基础和关键。针对当前军事目标识别技术在复杂战场环境下漏检误检率高、实时性差等问题,提出一种基于改进 YOLOv5 模型的 PB-YOLO 军事目标识别算法。将改进的目标识别算法对于陆战场军事单元的识别锚框进行重新聚类,以提升模型对于目标大小适应度,加速模型收敛;采用通道-空间并行注意力机制,增加模型对复杂战场环境下目标特征信息与位置信息关注度;在特征融合网络部分使用 BiFPN 以提升模型对于特征的融合能力与速度;采用 Alpha\_IoU 损失函数加速模型收敛,解决当真实框与预测框重合时 IoU 计算退化问题。实验结果表明,在自建军事目标数据集下,改进算法与主流目标识别算法相比,在保证模型空间复杂度的同时,mAP 值达到了 90.17%。消融实验对比结果表明,改进后网络较原模型精度提升 11.57%,具有较好的识别性能,能够为战场情报获取提供有效的技术支撑。

**关键词:** 军事目标识别; 通道-空间并行注意力机制; 特征融合; 损失函数

**中图分类号:** TP391.4 **文献标志码:** A **文章编号:** 1000-1093(2024)03-0934-14

## Research on Improved YOLOv5-based Military Target Recognition Algorithm Used in Complex Battlefield Environment

SONG Xiaoru<sup>1\*</sup>, LIU Kang<sup>1</sup>, GAO Song<sup>1</sup>, CHEN Chaobo<sup>1,2</sup>, YAN Kun<sup>1</sup>

(1. School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, Shaanxi, China;

2. Science and Technology on Electromechanical Dynamic Control Laboratory, Xi'an 710065, Shaanxi, China)

**Abstract:** Military target recognition technology used in complex battlefield environment is the basis and key to improve the battlefield intelligence acquisition capability. A PB-YOLO military target recognition algorithm based on the improved YOLOv5 model is proposed to solve the problems of high missed and false detection rates and poor real-time performance of current military target recognition technology in complex battlefield environments. The improved target recognition algorithm is re-clustered for the identification anchor boxes of military units in the land battlefield to improve the model's fitness for target size and accelerate the convergence of model, and the channel-spatial parallel attention mechanism is used to increase the model's attention to the feature information and location information of the targets in complex battlefield environments. BiFPN is used in the feature fusion network part to improve the fusion ability and speed of the model for features, and the Alpha\_IoU loss function is used to accelerate the convergence of model, and solve the problem of IoU calculation degradation when the real frame and the

收稿日期: 2022-08-21

基金项目: 国家自然科学基金项目(62103315); 陕西省重点研发计划项目(2021GY-287)

\* 通信作者邮箱: masha0422@163.com

predicted frame overlap. The experimental results show that, compared with the mainstream target recognition algorithm, the mAP value obtained by the improved algorithm reaches 90.17% while ensuring the model space complexity under the self-built military target data set. Through the comparison of ablation experiments, the results show that, compared with the original model, the accuracy of the improved network is improved by 11.57%, and it has better recognition performance, which can provide effective technical support for battlefield intelligence acquisition.

**Keywords:** military target recognition; channel-spatial parallel attention mechanism; feature fusion; loss function

## 0 引言

战场情报获取是通过光学照相、红外激光、雷达等设备对作战部队的战略部署、装备分布以及作战环境信息进行实时捕捉的过程<sup>[1]</sup>。在现阶段的信息战联合作战中,通过获取战场情报对敌方作战意图进行分析和预测,采取针对性措施来优先掌握战场主动权、提升己方作战优势。在未来无人战争中,提升战场情报获取能力可以有效提升战争的全局掌控能力,不同军事强国都在提升相关技术的研究与开发<sup>[2]</sup>。目前,复杂战场环境下军事目标识别是战场情报获取的关键技术,因此开展复杂战场环境下军事目标识别算法的研究,对于新形势多元化作战下的战场情报获取具有重要意义。

近年来,研究者们将深度学习引入战场情报获取中,使军事目标识别技术不断发展。而军事目标识别技术大致可以分为6个阶段:经典的统计模式识别方法、基于知识的自动目标识别方法、基于模型的自动目标识别方法、基于多传感器信息融合的自动目标识别方法、基于人工神经网络和专家系统混合应用的自动目标识别方法、基于人工智能和深度学习的自动目标识别方法<sup>[3]</sup>,前五阶段的军事目标识别技术无法与人类视觉识别效果相提并论。近年来,随着国防科技水平不断提高,利用深度卷积神经网络在军事应用领域对复杂战场环境下的军事目标识别研究逐渐增多,其主要分为两大类,一类是基于候选框的目标识别算法<sup>[4-9]</sup>,另一类是基于回归的目标识别算法<sup>[10-15]</sup>。

潘浩<sup>[16]</sup>提出用于军事目标识别的密集连通区域快速分类网络(Dense connected Region Fast Classification Network, DRFCN)算法,通过共享卷积层特征提高军事目标识别的准确率。Zeng等<sup>[17]</sup>提出在回归思想的基础上增加先验框,既能够保证精度要求,又满足了军事目标识别的实时性。Chen等<sup>[1]</sup>利用基于掩码区域的卷积神经网络(Mask Region-

based Convolutional Neural Network, Mask R-CNN),通过对Anchor尺度的优化以及引入迁移学习来解决原网络漏检率高和由于军事目标数据集少而导致的精度低问题。陈龙等<sup>[18]</sup>提出了小样本条件下基于深度森林学习模型的典型军事目标识别方法,该方法利用对抗生成网络与深度森林相结合,采用主动学习中基于后验概率的启发式主动学习算法对样本进行优化,使其能够准确识别军事目标。林洋等<sup>[19]</sup>提出了一种基于条件生成对抗网络(Conditional Generative Adversarial Nets, CGAN)和GcForest的军事目标识别方法。杨朝红等<sup>[20]</sup>提出基于优化单次多边框检测(Single MultiBox Detector, SSD)的小尺度典型军事目标识别方法,通过增加default box种类数量来提高密集小目标的识别能力。李鑫等<sup>[21]</sup>将YOLOv3与深度卷积生成对抗网络(Deep Convolutional Generative Adversarial Networks, DCGAN)相结合,提出了基于非监督网络的军事目标识别算法,有效解决了由于数据集不足而导致军事目标识别精度低的问题。于博文等<sup>[2]</sup>在YOLOv3的结构中引入ResNet50-D残差网络、双注意力机制与特征重构模块,提出一种基于改进YOLOv3的军事目标检测算法,增加了模型对于复杂环境下军事目标的表征能力,提升了其检测精度与速度。

由于战场情报具有高准确性与实时性,对军事目标识别的识别精度与实时性也有较高的要求。现有的军事目标识别方法归结起来存在以下问题:1)基于候选框的目标识别算法通常精度高,但算法冗余度高、识别速度慢;基于回归的目标识别算法虽然识别精度相对较低,但实时性高;2)对于复杂背景下的军事目标图像,识别精度不高;3)对军事目标的识别需要具有高实时性与高精度,但目前许多算法要么精度高、实时性差,要么实时性强、精度低,难以做到二者的平衡。

基于此,本文从战场情报获取中复杂战场环境下的军事目标识别技术出发,通过深度学习的方式,

采用目前主流基于回归的目标识别算法 YOLOv5 为基础网络并进行改进,提出一种复杂战场环境下的军事目标识别(Parallel attention BiFPN-YOLO, PB-YOLO)算法。针对在复杂战场环境下军事目标识别中所出现的漏检误检率高、精度低、速度慢等问题,首先在模型的每个 BottleNeck 处引入通道-空间并行注意力机制模块<sup>[22]</sup>,通过融合多感受野,使网络学习到的信息更加丰富,提升网络在复杂战场环境下对于目标特征的提取能力;其次加入 BiFPN 加权特征融合网络<sup>[23]</sup>,加速模型特征融合,提升模型的推理速度;然后使用 Alpha\_交并比(Intersection over Union, IoU)<sup>[24]</sup>对目标识别分类器的损失函数进行改进,解决当目标框与预测框重叠时的 IoU 退化问题,并加快模型收敛速度,提升识别精度;最后通过自建包含多目标的陆战场军事目标数据集,使改进后的模型在自建军事目标数据集上进行训练与测试,与主流的 6 种目标识别算法进行对比,验证改

进算法对复杂环境下陆战场军事目标单元识别的性能,再通过模型拆分进行消融实验,对比体现改进算法的有效性;最终利用无人机与终端完成通信,进行实物测试,验证其在实际场景下的识别效果,为战场情报获取提供有效的技术支撑。

### 1 YOLOv5 算法模型设计与分析

YOLOv5 是目标识别算法 YOLO 系列的较新一代算法,其兼顾识别精度与识别速度,是单阶段识别算法中实施高效的目标识别算法。针对本文的特殊性,考虑精度的同时需要满足实时性,因此选取 YOLOv5s 作为军事目标识别算法的基础模型。YOLOv5s 的网络结构主要由输入端、主干网络、Neck 网络和预测输出端四部分组成,网络结构如图 1 所示。输入端输入图像并对其进行随机缩放、批处理归一化、图像增强等预处理操作;主干网络用来提取图像特征;Neck 网络用来提升网络多样性与鲁

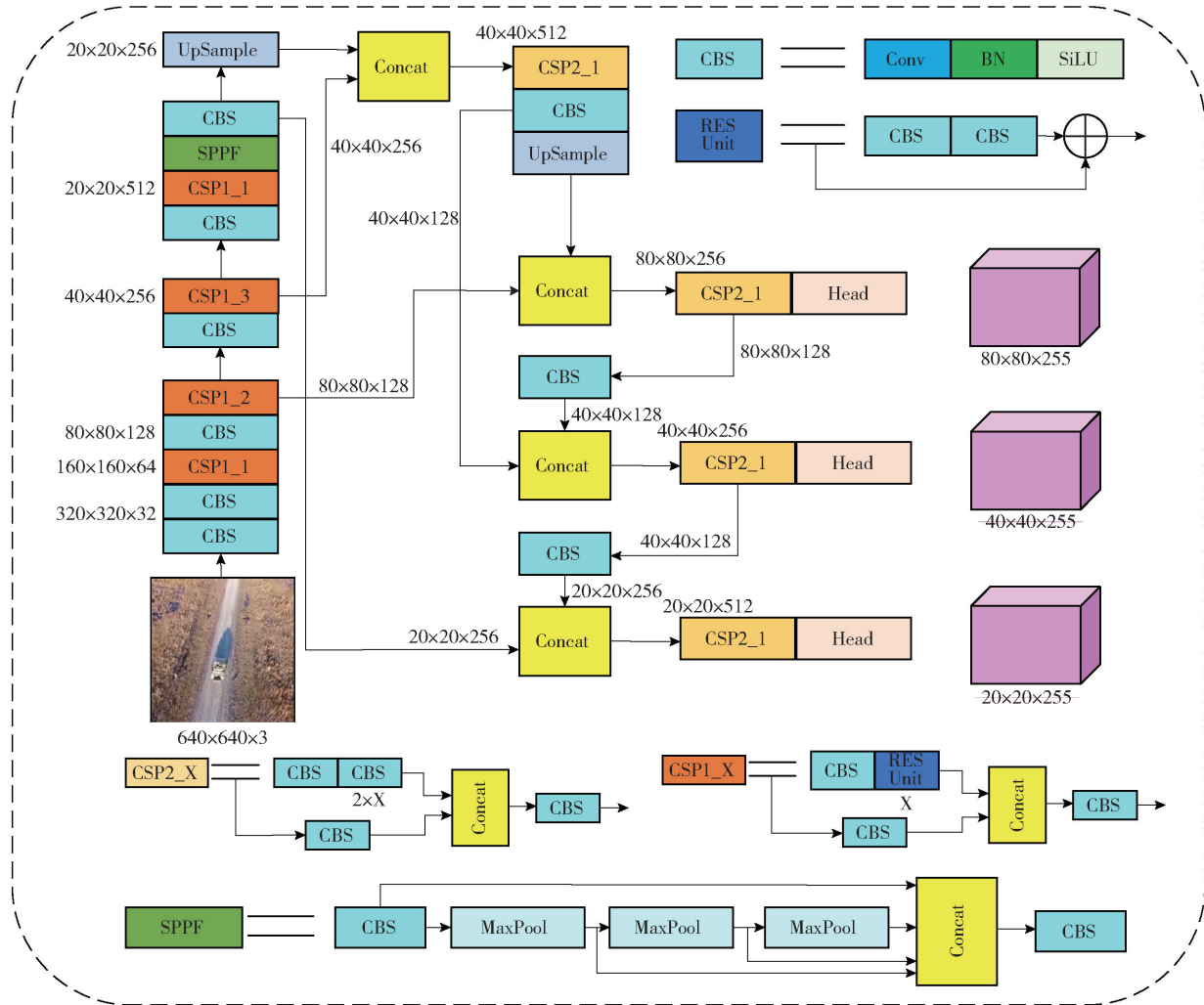


图 1 YOLOv5s 网络结构图

Fig. 1 YOLOv5s network structure diagram

棒性;预测输出端完成对多目标的分类预测。

YOLOv5s 在输入端随机对输入图像按照缩放、裁剪、排布的方式进行 Mosaic 数据增强,将多张图片拼接组合成一张,以丰富样本、降低模型内存需求;主干网络采取 Focus 结构与跨阶段部分(Cross Stage Partial, CSP)结构,并且用到了切片操作,采用 CSP 结构的主要目的在于提升模型的学习能力,既可以使模型轻量化,又可以确保网络的识别准确率,减少模型开销与内存占用;Neck 结构部分采用特征金字塔网络(Feature Pyramid Networks, FPN)与路径聚合网络(Path Aggregation Network, PAN)的结构,FPN 层自顶向下传达语义特征,而 PAN 层自底向上传达强定位特征,二者结合,从不同的主干层对不同检测层的参数进行聚合,提升了网络的特征融

合能力;输出端检测网络部分采用 K-means 聚类提供目标位置的先验信息。

### 2 改进 YOLOv5 算法模型的构建

面向战场的烟雾避障等恶劣场景,针对 YOLOv5s 模型在复杂环境下小目标识别性能较低的问题,本文提出通道-空间并行注意力机制,解决模型对小目标特征提取能力差的问题,采用 BiFPN 多尺度特征融合来提高模型推理速度,采用 Alpha\_IoU 损失函数解决模型在识别过程中面对目标框与预测框相互重合而导致精度降低等问题,在实现上述策略方法对 YOLOv5 目标识别算法改进的基础上,提出 PB-YOLO 复杂战场环境下的军事目标识别算法,其算法具体网络结构如图 2 所示。

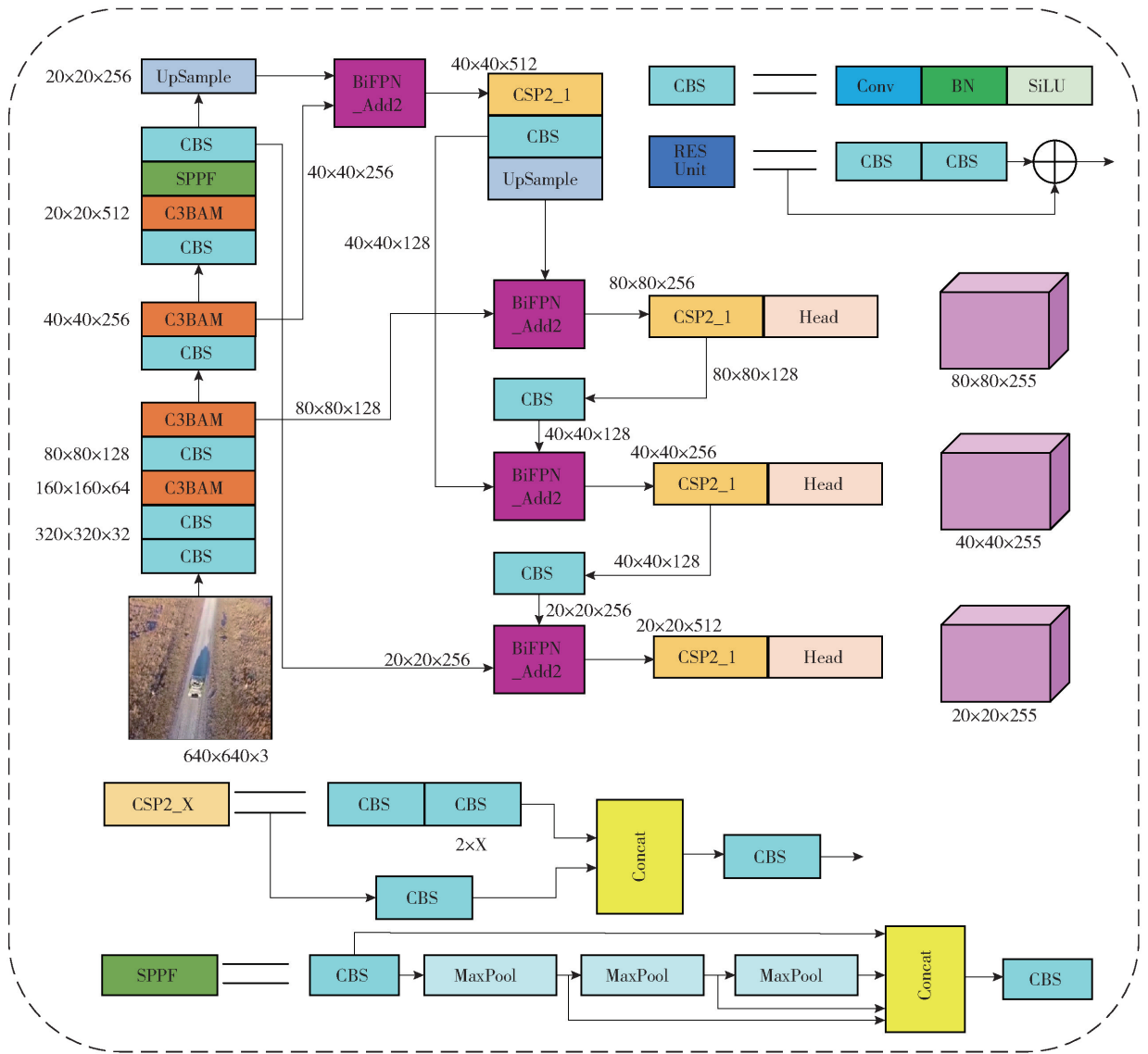


图 2 PB-YOLO 网络结构图  
Fig. 2 PB-YOLO network structure diagram

2.1 通道-空间并行注意力机制

由于小目标具有分辨率低、像素有限等特点,往往只占图像很小部分,尤其遇到恶劣环境或烟雾遮挡时识别效率大大降低,在空对地复杂场景下的军事目标识别中这种情况尤为普遍。注意力机制是机器学习中的一种特殊结构,是基于输入图像特征的动态权重调整过程,使模型更加关注识别对象。利用注意力机制,调整不同维度的分值,依据分值大小

给出不同的权重信息,权重占比大的特征即对后续模块影响较大,决定分配信息的资源即越多。目前,常见的注意力机制模块为 SE<sup>[25]</sup> (Squeeze-and-Excitation) 模块,其过程主要分为压缩 (Squeeze) 与激发 (Excitation) 两个步骤,是采用通道轴之间关系的注意力机制,但未考虑空间信息也能够推断目标特征准确位置。本文基于此,引入通道-空间并行注意力机制,其引入后的具体结构如图 3 所示。

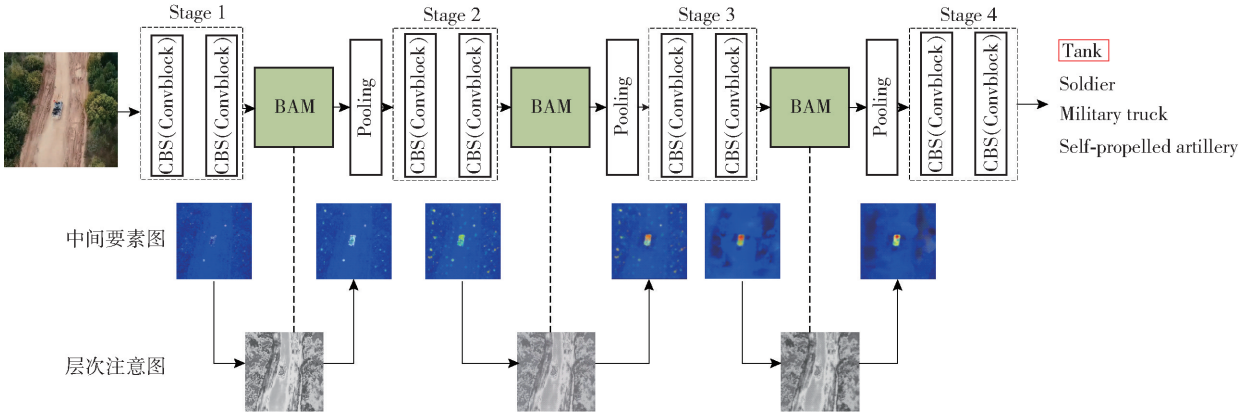


图 3 通道-空间并行注意力机制引入位置图

Fig. 3 Channel-Spatial parallel attention mechanism introduces location map

该模块主要由通道注意力分支与空间注意力分支组成,通过高效梯度流<sup>[26]</sup>元素求和,实现通道与空间注意力机制的结合,其原理结构图如图 4 所示。图 4 中,  $F$  为输入特征图维度,  $F \in \mathbf{R}^{C \times H \times W}$

( $C, H, W$  分别为通道数、高度、宽度),  $M(F)$  为三维注意力图,  $M(F) \in \mathbf{R}^{C \times H \times W}$ ,  $M_c(F)$  为通道注意力图,  $M_c(F) \in \mathbf{R}^C$ ,  $M_s(F)$  为空间注意力图,  $M_s(F) \in \mathbf{R}^{H \times W}$ 。

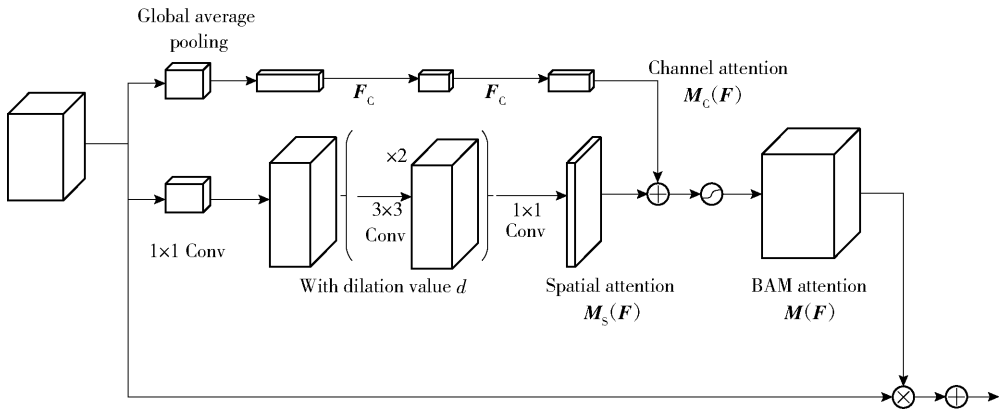


图 4 通道-空间注意力机制结构图

Fig. 4 Structure diagram of channel-spatial attention mechanism

对于输入特征图维度  $F \in \mathbf{R}^{C \times H \times W}$ , 该通道-空间注意力机制能够得出总的三维注意力图  $M(F) \in \mathbf{R}^{C \times H \times W}$ , 故重新定义的特征图维度  $F'$  为

$$F' = F + F \otimes M(F) \quad (1)$$

式中:  $\otimes$  表示元素级乘法, 通过采用残差学习和注意力机制促进梯度流。为了使模型的效率和检测精度得到进一步提升, 将通道注意力机制与空间注意力

机制变为两个独立的分支: 通道注意力图  $M_c(F) \in \mathbf{R}^C$  和空间注意力图  $M_s(F) \in \mathbf{R}^{H \times W}$ , 两个独立分支在输出求和之前将其维度都调整为  $\mathbf{R}^{C \times H \times W}$ , 因此总的三维注意力图  $M(F)$  为

$$M(F) = \text{sigmoid}(M_c(F) + M_s(F)) \quad (2)$$

1) 通道注意力分支

由于每个通道包含一个特定的特征响应, 首先

利用通道分支间的关系来聚合每个通道的特征图。通过对特征图  $F$  进行全局平均池化,产生一个通道向量  $F_c \in \mathbf{R}^{C \times 1 \times 1}$ ,该向量对每个信道中的全局信息进行软编码;然后将一个带有隐藏层的多层感知机 (Multilayer Perceptron, MLP) 根据通道向量  $F_c$  估计跨通道的注意力;其次,为了减少参数运算量,激活函数大小被设置为  $\mathbf{R}^{C/r \times 1 \times 1}$ ,其中  $r$  为衰减率。在 MLP 中结束后,为了通过空间分支输出调整尺度,增加了一个批处理归一化 (Batch Normalization, BN<sup>[27]</sup>) 层。故通道注意力的计算方式为

$$M_c(F) = BN(MLP(AvgPool(F))) = BN(W_1(W_0 AvgPool(F) + b_0) + b_1) \quad (3)$$

式中:  $W_0 \in \mathbf{R}^{C/r \times C}$ ;  $b_0 \in \mathbf{R}^{C/r}$ ;  $W_1 \in \mathbf{R}^{C \times C/r}$ ;  $b_1 \in \mathbf{R}^C$ 。

### 2) 空间注意力分支

空间分支产生空间注意力图  $M_s(F) \in \mathbf{R}^{H \times W}$ ,用来强调或抑制不同空间位置的特征,通过得到大的接受域以有效地利用上下文信息,获取到空间中应该被重点关注的位置;其次采用扩展卷积高效扩大接受野,通过与标准卷积相比,扩张卷积在构建空间映射方面更有优势;最后在空间分支中采用 ResNet<sup>[28]</sup> 中所采用的“瓶颈结构”,使参数量与计算量得到大幅减少。

特征  $F \in \mathbf{R}^{C \times H \times W}$  通过  $1 \times 1$  卷积被投影到一个降维  $\mathbf{R}^{C/r \times H \times W}$  中,对跨通道维的特征图进行继承和

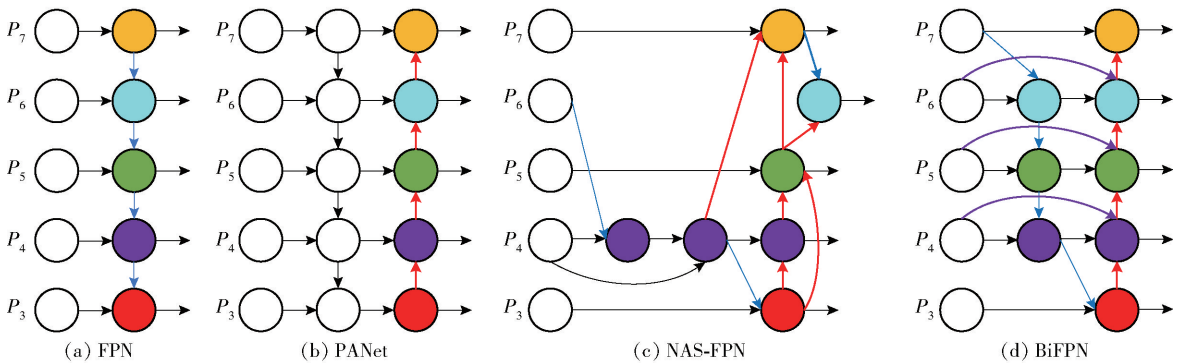


图 5 第 3 层 ~ 第 7 层采用不同特征融合网络示意图

Fig. 5 Schematic diagram of different feature fusion networks from Layer 3 to Layer 7

图 5 (a) 显示了自顶向下的 FPN,其需要 3 ~ 7 层的输入特征  $P^{in} = (P_3^{in}, \dots, P_7^{in})$ ,其中  $P_i^{in}$  表示特征层,分辨率为输入图像的  $1/2^i$ 。例如,输入图像的分辨率是  $640 \times 640$ ,则  $P_3^{in}$  表示第 3 层特征的分辨率为  $80 \times 80$  ( $640/2^3 = 80$ ),  $P_7^{in}$  表示第 7 层的分辨率为  $5 \times 5$ 。传统的 FPN 以自顶向下的方式聚集多尺度特征:

$$P_7^{out} = Conv(P_7^{in})$$

压缩。首先对每个通道分支使用相同的缩减比  $r$ ;然后利用两个  $3 \times 3$  扩张卷积有效利用上下文信息;最后利用  $1 \times 1$  卷积再次将特征简化为  $\mathbf{R}^{1 \times H \times W}$  的空间注意力图,并在空间分支末尾使用批处理归一化层进行尺度调整。空间注意力的具体计算公式为

$$M_s(F) = BN(f_3^{1 \times 1}(f_2^{3 \times 3}(f_1^{3 \times 3}(f_0^{1 \times 1}(F)))))) \quad (4)$$

式中:  $BN$  表示批处理归一化;  $f$  表示卷积运算;  $f$  符号的上标表示卷积滤波器的大小,有两个  $1 \times 1$  卷积用于信道缩减,中间的  $3 \times 3$  扩张卷积用于聚合具有更大接受域的上下文信息。

### 2.2 BiFPN 加权特征融合

针对军事战场场景下的识别算法,对模型的推理速度都有较高要求, YOLOv5s 模型中采用 FPN 和 PAN 完成特征融合,二者结合虽然提升了网络的特征融合能力,但由于 PAN 是在 FPN 上额外增加的,增大了网络训练时间与计算量。为此本文采用 BiFPN 结构加速特征融合与双向跨尺度连接的时效性。

#### 2.2.1 问题公式化

多尺度特征融合在形式上按照分辨率大小的差异对其特征进行聚合,给出一列多尺度特征  $P^{in} = (P_{l_1}^{in}, P_{l_2}^{in}, \dots)$ ,其中  $P_{l_i}^{in}$  表示  $l_i$  层的特征,使其得到一个变换  $f$ ,能够输出一列新特征并对其特征进行融合:  $P^{out} = f(P^{in})$ ,不同特征网络结构见图 5。

$$\begin{aligned} P_6^{out} &= Conv(P_6^{in} + Resize(P_7^{out})) \\ &\vdots \\ P_3^{out} &= Conv(P_3^{in} + Resize(P_4^{out})) \end{aligned} \quad (5)$$

式中:  $Resize$  普遍使用在分辨率匹配上采样或下采样操作;  $Conv$  为卷积操作。

#### 2.2.2 Cross-Scale 连接

经典 FPN 从根本上受单向信息流的制约,为此 PANet 在其基础上扩增一个自底向上的路径聚合网

络,如图5(b)所示。近年来,为探寻更优的跨尺度特征网络拓扑,神经网络架构搜索-特征金字塔网络(Neural Architecture Search-Feature Pyramid Network, NAS-FPN)通过神经体系结构进行搜索,但由于其网络结构复杂且不易阐明,导致花费数千个GPU小时进行搜索,如图5(c)所示。

针对上述问题,本文采用3种跨尺度连接改善形式:首先,由于单输入边节点对特征融合网络贡献低,对其进行剔除,构建出一个简化双向网络;然后,在同一等级输入输出节点上增加一条额外的边,能够在不扩增计算量的同时使更多的特征进行融合;最后,与PANet不同,一个特征网络层由一个双向(自上而下和自下而上)路径构成,进行多次利用,完成高级特征融合,称为双向特征金字塔网络(Bidirectional Feature Pyramid Network, BiFPN)。

### 2.2.3 加权特征融合

当融合不同分辨率的特征时,通常先将它们调整到相同的分辨率,然后相加。由于不同的输入特征具有不同的分辨率,它们对输出特征的贡献通常是不平等的,针对这个问题,为每个输入添加额外的权重,并让网络学习每个输入特征的重要性。常见的融合方式有无界融合、基于Softmax融合与快速归一化融合3种。基于识别对象的特殊性,本文拟采用快速归一化融合方法:

#### 1) 无界融合

$$\mathbf{O} = \sum_i \mathbf{w}_i \mathbf{g} \mathbf{I}_i \quad (6)$$

式中: $\mathbf{O}$ 为融合后的特征值。 $\mathbf{w}_i$ 为可学习的权重,可以是标量(每个特征)、向量(每个通道)或多维张量(每个像素); $\mathbf{g}$ 为一种采样变换方式,用来对特征进行变换或处理,以调整到统一的尺寸; $\mathbf{I}_i$ 为第*i*个特征的取值。由于标量权重是无界的,可能会导致训练不稳定,因此采用权值归一化的方法限定各个权值的取值范围。

#### 2) 基于Softmax的融合

$$\mathbf{O} = \sum_i \frac{e^{w_i}}{\sum_j e^{w_j}} \cdot \mathbf{I}_i \quad (7)$$

直观的想法是对每个权重应用softmax,这样所有权重都被标准化为一个概率,其取值范围为0~1,代表每个输入的重要性。然而,额外的softmax会导致GPU硬件的显著减速,为最小化额外的延迟成本,进一步提出了一种快速归一化融合方法。

#### 3) 快速归一化融合

$$\mathbf{O} = \sum_i \frac{w_i}{\varepsilon + \sum_j w_j} \cdot \mathbf{I}_i \quad (8)$$

式中: $w_i \geq 0$ 为确保每个 $w_i$ 之后应用Relu激活函数; $\varepsilon = 0.0001$ 是一个较小的值,用于避免数值的不稳定。类似地,每个BN完成后的权值同样位于0~1范围内,但由于没有softmax操作,效率更高。

最终,BiFPN将双向跨尺度连接与快速归一化融合进行聚合。式(9)为描述图5(d)中所示BiFPN在第6级的两个融合特性:

$$P_6^{\text{ld}} = \text{Conv} \left( \frac{w_1 \cdot P_6^{\text{in}} + w_2 \cdot \text{Re size}(P_7^{\text{in}})}{w_1 + w_2 + \varepsilon} \right)$$

$$P_6^{\text{out}} = \text{Conv} \left( \frac{w'_1 \cdot P_6^{\text{in}} + w'_2 \cdot P_6^{\text{ld}} + w'_3 \cdot \text{Re size}(P_5^{\text{out}})}{w'_1 + w'_2 + w'_3 + \varepsilon} \right) \quad (9)$$

式中: $P_6^{\text{ld}}$ 为自顶向下路径第6级的中间特征; $P_6^{\text{out}}$ 为自底向上路径第6级的输出特征;其余特征均以类似方式构建; $w'_1, w'_2, w'_3$ 为前一级特征权重、当前级特征权重、下一级特征权重。本文通过深度可分离卷积操作聚合特征,同时引入BN与激活操作,以提升模型计算速度。

### 2.3 Alpha-IoU 损失函数

在识别军事目标时,常会遇到目标框与预测框相互包裹而导致网络收敛速度降低,广义交并比(Generalized Intersection over Union, GIoU)损失函数会退化成IoU,并且当两个框相交时模型在水平和垂直方向上收敛较慢,因此本文采用Alpha-IoU,该损失函数通过将幂变换应用于基于IoU的损失函数中,以获取更准确的框回归和对象检测。

普通的IoU损失定义为 $L_{\alpha\text{-IoU}} = 1 - \text{IoU}$ ,首先应用Box-Cox变换,由IoU损失推出 $\alpha\text{-IoU}$ 损失函数的表达式:

$$L_{\alpha\text{-IoU}} = \frac{1 - \text{IoU}^\alpha}{\alpha}, \alpha > 0 \quad (10)$$

通过调整 $\alpha\text{-IoU}$ 中的 $\alpha$ 参数,可以推出现有损失函数如 $\log_2(\text{IoU})$ 、 $\text{IoU}$ 、 $\text{IoU}^2$ 等。

当 $\alpha \rightarrow 0$ 时,得出

$$\lim_{\alpha \rightarrow 0} L_{\alpha\text{-IoU}} = -\log_2(\text{IoU}) = L_{\log_2(\text{IoU})} \quad (11)$$

因此,当 $\alpha = 1$ 时 $L_{1\text{-IoU}} = 1 - \text{IoU} = L_{\text{IoU}}$ ,弥补了IoU的损失;当 $\alpha = 2$ 时 $L_{2\text{-IoU}} = \frac{1}{2}(1 - \text{IoU}^2) = \frac{1}{2}L_{\text{IoU}^2}$ 。故当 $\alpha$ 值不同时,所对应的IoU损失函数亦不同。利用 $\alpha > 0$ 和 $\alpha \rightarrow 0$ 简化 $\alpha\text{-IoU}$ 损失函数表达式:

$$L_{\alpha\text{-IoU}} = \begin{cases} -\log_2(\text{IoU}), & \alpha \rightarrow 0 \\ 1 - \text{IoU}^\alpha, & \alpha > 0 \end{cases} \quad (12)$$

然后将 $\alpha \rightarrow 0$ 的 $\alpha\text{-IoU}$ 损失扩展到更一般的形式,在公式中引入幂指数惩罚项/正则化项:

$$L_{\alpha-IoU} = 1 - IoU^{\alpha_1} + P^{\alpha_2}(B, B^{st}) \quad (13)$$

式中:  $\alpha_1 > 0; \alpha_2 > 0; P^{\alpha_2}(B, B^{st})$  表示基于  $B$  和  $B^{st}$ , 该扩展项直接概括出现有的 IoU 基础损失到其  $\alpha$ -IoU 版本。根据式(13), 可以将常用的 IoU 损失包括  $L_{IoU}$ 、 $L_{GIoU}$ 、 $L_{DIoU}$  和  $L_{CIoU}$ , 对 IoU 和惩罚项使用相同的参数  $\alpha$ :

$$L_{IoU} = 1 - IoU \Rightarrow L_{\alpha-IoU} = 1 - IoU^{\alpha}$$

$$L_{GIoU} = 1 - IoU + \frac{|C \setminus (B \cup B^{st})|}{|C|} \Rightarrow$$

$$L_{\alpha-GIoU} = 1 - IoU^{\alpha} + \frac{|C \setminus (B \cup B^{st})|}{|C|}$$

$$L_{DIoU} = 1 - IoU + \frac{\rho^2(b, b^{st})}{c^2} \Rightarrow$$

$$L_{\alpha-DIoU} = 1 - IoU^{\alpha} + \frac{\rho^{2\alpha}(b, b^{st})}{c^{2\alpha}}$$

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{st})}{c^2} + \beta v \Rightarrow$$

$$L_{\alpha-CIoU} = 1 - IoU^{\alpha} + \frac{\rho^{2\alpha}(b, b^{st})}{c^{2\alpha}} + (\beta v)^{\alpha} \quad (14)$$

式中:  $C$  为包围  $B$  和  $B^{st}$  的最小凸形;  $b$  和  $b^{st}$  为  $B$  和  $B^{st}$  的中心点;  $\rho(\cdot)$  为欧氏距离;  $c$  为最小框的对角线长度;  $v = \frac{4}{\pi^2} \left( \arctan \frac{w^{st}}{h^{st}} - \arctan \frac{w}{h} \right)^2$ ,  $w$  为预测框的宽,  $h$  为预测框的高,  $\beta = \frac{v}{(1 - IoU) + v}$ ; 当  $\alpha = 1$  时,  $L_{IoU}$ 、 $L_{GIoU}$ 、 $L_{DIoU}$  和  $L_{CIoU}$  的变换形式变为原型。

综上所述可知,  $\alpha > 1$  时, 容易的例子先学习;  $IoU = 1$  时, 难的例子会逐渐开始学习;  $0 < \alpha < 1$  时, 趋向于降低最终性能, 减少高  $IoU$  对象的损失和梯度会导致对象的局部化更差。  $L_{\alpha-IoU}$  的不同变化形式都与目标的 IoU 值相适应, 这种方式对模型的训练提供了更大的灵活性, 能够实现不同水平的目标框回归精度。

综上所述可知,  $\alpha > 1$  时, 容易的例子先学习;  $IoU = 1$  时, 难的例子会逐渐开始学习;  $0 < \alpha < 1$  时, 趋向于降低最终性能, 减少高  $IoU$  对象的损失和梯度会导致对象的局部化更差。  $L_{\alpha-IoU}$  的不同变化形式都与目标的 IoU 值相适应, 这种方式对模型的训练提供了更大的灵活性, 能够实现不同水平的目标框回归精度。

### 3 实验结果与分析

#### 3.1 实验环境与参数设置

本文实验测试环境为 Windows10 操作系统, CPU 参数为 Intel(R) Core(TM) i9-10920X CPU @ 3.50 GHz, 32 GB 内存, Nvidia RTX3090 显卡  $\times 2$ 。网络框架基于 Pytorch1.8.1, CUDA11.1, 使用 Python 进行编程实现, 初始参数设置如表 1 所示。实物测试采用 P450-NX 科研无人机, 该无人机配备 T265 相机、二维激光雷达、深度相机 D435i 与 200 万像素单目摄像头, 无人机图片如图 6 所示, 其

飞行高度为 50 ~ 100 m, 属于低空任务范畴。

表 1 初始参数设置

Table 1 Initial parameter setting table

参数	数值
lr	0.01
momentum	0.937
weight_decay	0.0005
batch_size	16
image_size	640 $\times$ 640
epoch	300



图 6 P450-NX 科研无人机

Fig. 6 P450-NX UAV

#### 3.2 构建军事目标数据集

根据作战环境中的常见军事行动单元, 本文重点对坦克、士兵、军事卡车和自行火炮四类陆战场军事目标进行识别。由于真实军事目标目前并没有大量公开的数据集, 本文针对复杂战场环境下的军事目标识别任务, 结合项目合作单位某研究所提供内部资料的陆战场军事目标与互联网下载构建陆战场军事目标数据集共 6 322 张, 数据集标注使用 LabelImg 工具, 将其所标注的格式与 PASCAL VOC 数据集标签格式保持一致, 图片样本如图 7 所示, 并按照 7:2:1 的比例将数据集分为训练集、验证集和测试集。其中 4 种目标的具体数量及尺寸参数如表 2 所示。



图 7 部分数据集样图

Fig. 7 Sample drawings of some datasets

表 2 目标数量及尺寸统计

Table 2 Statistical table of target quantity and size

数量及尺寸	坦克	士兵	装甲车	自行火炮
数量/个	1 652	2 497	1 086	1 087
尺寸/m	11 × 3.4 × 2		5.5 × 2.5 × 2.3	8 × 3 × 2.1

3.3 评价指标

将算法的检测精度与检测速度作为主要衡量本文所提算法性能的标准,其次检测精度包括精度(Average Precision, AP)与平均精度(mean Average Precision, mAP),而 AP 由精确率(Precision)与召回率(Recall)二值指标进行评估。AP 值由集合中某一类别所有准确率之和除以含有该类别目标的图像数量  $N_C$  得到,AP 的计算公式如式(15)所示:

$$AP = \frac{\sum Precision}{N_C} \quad (15)$$

由于军事目标识别任务中不仅只有一个类别,而且每一种类别都应统计其 AP 均值,mAP 的表达式如式(16)所示:

$$mAP = \frac{\sum_{k=0}^{ClassNum} AP_k}{ClassNum} \quad (16)$$

式中:  $\sum_{k=0}^{ClassNum} AP_k$  表示每一个类别的平均精度之和;  $ClassNum$  为总类别数。

检测速度方面,使用 FPS(每秒可以处理的图片数量)来衡量,通过计算一定数量测试图片总的测试时间,再求平均值,即可求出 FPS 值。

结合以上评价标准,本文采用消融对比实验,通过拆分网络、加减改进模块,对比改进后的模型;还可以在同一实验环境下采用其他的 one stage 目标识别算法做对比。

3.4 模型测试结果分析

3.4.1 先验框重聚类

在目标识别中,由于先验框的聚类,网络模型可以不对目标位置直接进行预估,根据先验框与真实框之间的距离偏差来预估样本。表 3 为本文利用识别算法对先验框采用 K-means 重聚类,该方法能够在一定程度上提升召回率与识别精度。

3.4.2 模型对比实验

主流算法的 mAP 值对比如图 8 所示,将主流识别算法与 PB-YOLO 算法采用同一段军事目标视频进行预测识别对比,IoU 阈值、置信度阈值分别设置为 0.5、0.3,其中某一帧的具体识别效果对比如图 9 所示。

表 3 重聚类后先验框宽高值

Table 3 Prior box width and height values after reclustering

序号	先验框宽高值
1	[12.734 28.496]
2	[18.768 39.405]
3	[52.218 42.258]
4	[35.064 83.506]
5	[102.03 61.18]
6	[153.82 115.14]
7	[232.96 306.19]
8	[409.64 177.11]
9	[526.23 396.53]

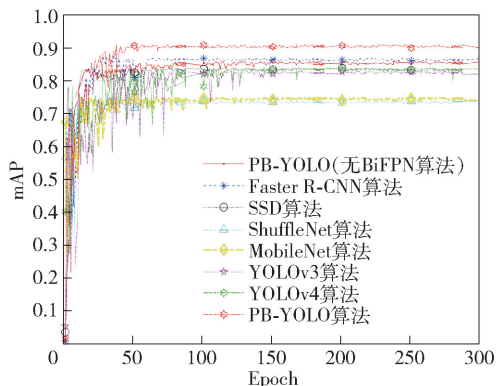


图 8 主流算法精度对比图

Fig. 8 Precision comparison diagram of mainstream algorithms

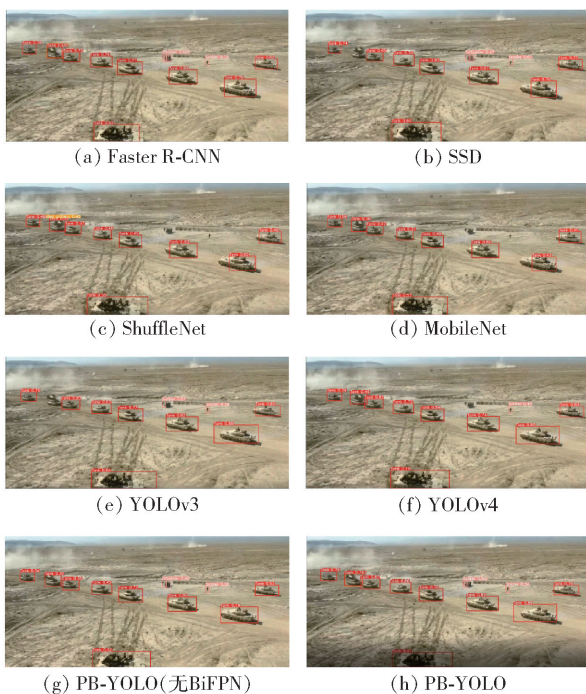


图 9 主流算法识别效果对比图

Fig. 9 Comparison of recognition effects of mainstream algorithms

从图8的主算法精度对比图中可以看出,本文提出的PB-YOLO算法识别精度最高,Faster R-CNN算法紧随其后,ShuffleNet算法的精度最低;从图9具体识别效果图中能够发现SSD、ShuffleNet、MobileNet共3种算法的识别精度低,而且对于中间士兵的识别出现了不同程度的漏检和误检;YOLOv3、YOLOv4、Faster R-CNN与PB-YOLO共4种算法对于小目标的识别均表现出较好的效果,其中本文提出的PB-YOLO算法对于坦克以及小目标士兵的识别效果最好。

表4为基于自建军事目标数据集的不同主流目标识别算法的实验结果,主要对比的值分别为模型体积、FPS与识别精度。从表4中可以看出,在PB-YOLO算法未加入BiFPN模块之前,Faster R-CNN算法的mAP值最高,但是其FPS值只有6.5;在加入BiFPN模块后,PB-YOLO算法相比于Faster R-CNN算法在模型体积方面仅占其40.07%左右,FPS值是其8.7倍,mAP值提升了3.47%。相比于其他算法,虽然在mAP值上接近本文所提算法,但是在模型体积上却远大于PB-YOLO算法,例如YOLOv3算法、YOLOv4算法,其模型体积是PB-YOLO算法的3.7倍之多;YOLOv5m算法、YOLOv5l算法、YOLOv5x算法由于其网络的残差组件个数与卷积核个数逐渐增多,导致网络的计算量、参数量不断增加,虽然在识别精度上相比于YOLOv5s算法有所提升,但是其模型大小却是YOLOv5s算法的2~7倍,而且由于计算量的增加使模型的推理速度变长,导致FPS值下降。由此可见,本文提出的PB-YOLO算法在FPS值与mAP值上相比于其他算法更突出。

表5 PB-YOLO消融实验对比

Table 5 PB-YOLO ablation experiment comparison

模型	识别算法及改进	Parameters/M	GFLOPs	Training time/h	mAP/%
A	YOLOv5	7.021	15.8	4.12	78.6
B	A + 通道注意力机制	6.932	15.5	4.5	80.4
C	A + 空间注意力机制	6.879	15.7	4.37	81.2
D	A + 并行注意力机制	6.713	15.3	5.12	84.3
E	A + Alpha-IoU	7.021	15.8	3.495	79.3
F	B + Alpha-IoU	6.932	15.5	4.39	80.8
G	C + Alpha-IoU	6.879	15.7	4.22	81.5
H	D + Alpha-IoU	6.713	15.3	4.75	85.0
I	A + BiFPN	7.169	16.5	2.126	83.6
J	D + BiFPN	6.862	15.9	3.304	89.47
K	E + BiFPN	7.169	16.5	2.061	85.5
L	H + BiFPN	6.862	15.9	3.261	90.17

表4 主流算法识别结果对比

Table 4 Comparison of recognition results of mainstream algorithms

主流算法	模型体积/MB	FPS	mAP/%
Faster R-CNN	159.7	6.5	86.7
SSD	23.1	38	83.4
ShuffleNet	3.0	28	73.3
MobileNet	8.0	34	74.8
YOLOv3	235.6	20	82.1
YOLOv4	244.8	48	83.1
YOLOv5m	118.3	44.3	81.5
YOLOv5l	226.7	43.2	83.4
YOLOv5x	401.0	40.9	84.0
YOLOv5s + SE	60.0	42	81.1
YOLOv5s + CBAM	63.0	37	84.0
PB-YOLO(无 BiFPN)	63.3	51	85.0
PB-YOLO	64.0	57	90.17

### 3.4.3 消融实验对比

为了更好地理解PB-YOLO算法中各个改进模块对识别效果的影响,本节将对模型进行拆分,并通过逐步加入改进模块进行消融实验对比。

如表5所示,模型A为原始的YOLOv5模型,模型B为引入通道注意力机制的YOLOv5模型,模型C为引入空间注意力机制的YOLOv5模型,模型D为引入通道-空间并行注意力机制的YOLOv5模型,模型E、F、G、H分别为在模型A、B、C、D上加入Alpha-IoU改进损失函数,模型I、J、K、L分别为在A、D、E、H上加入BiFPN模块。下面对各种改进策略详细说明。

1) 模型 A→模型 B、C、D: 在 YOLOv5 网络结构的特征融合阶段分别加入通道注意力机制、空间注意力机制与通道-空间并行注意力机制。从表 5 中可以看出, 融合了多感受野后的网络其 Training time 与 mAP 值均有所增加, 在加入本文所用到的通道-空间并行注意力机制后, 训练时间增加 1 h, 但是 mAP 值从 78.6% 上升到 84.3%, 结果表明引入并行注意力机制, 模型识别准确率得以明显提升。

2) 模型 A、B、C、D→模型 E、F、G、H: 针对加入不同注意力机制的网络引入新的损失函数, 由表 5 的 Training time 一栏可以看出, 在加入 Alpha-IoU 之后, 优化了模型的 Training, 加快模型收敛, 并且使模型的 mAP 值上升 0.3% ~ 0.7%。

3) 模型 A、D、E、H→模型 I、J、K、L: 针对改进的主要模型分别加入 BiFPN 模块, 由表 5 可以看出, 加入该模块后模型的特征融合速度加快, 最明显的反映就是训练时间缩短, 并且精度也有了进

一步提升。

具体消融实验 mAP 值对比如图 10 所示。图 11 所示为主要改进模型 A、D、J 采用 Alpha-IoU 损失函数后精度提升对比图。从图 11 中可以看出, 在加入 Alpha-IoU 之后网络精度均有一定幅度提升; 其单目标与多目标识别对比图如图 12、图 13 所示。

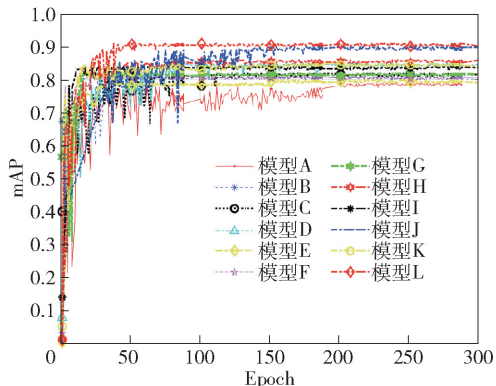
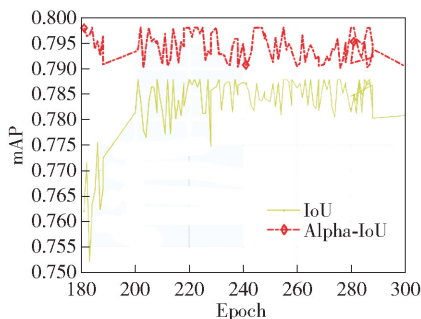
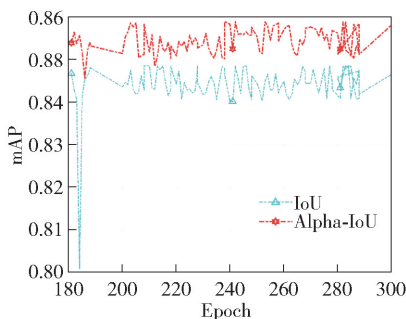


图 10 消融实验对比图

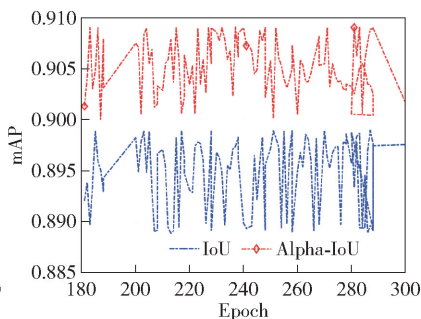
Fig. 10 Comparison of ablation experiments



(a) 模型 A 与 E 局部 IoU 提升对比图  
(a) Model A and Model E local IoU improvement comparison diagram



(b) 模型 D 与 H 局部 IoU 提升对比图  
(b) Model D and Model H local IoU improvement comparison diagram



(c) 模型 J 与 L 局部 IoU 提升对比图  
(c) Model J and Model L local IoU improvement comparison diagram

图 11 Alpha-IoU 精度提升对比图

Fig. 11 Alpha-IoU accuracy improvement comparison chart

从图 12 中能够看出, YOLOv5 算法及相关改进算法对于单个目标的识别效果都能够满足要求, 能够准确识别并进行分类。图 13 验证了相关算法对于多目标的识别效果, 从中能够看出本文提出的 PB-YOLO 算法对于多目标的识别精度能够满足实际战场的需求。

图 14 所示为将改进算法应用在无人机上进行实物验证, 通过与无人机建立通信, 实现实时的陆战场军事目标识别。从图 14 中可以看出, 本文提出的算法能够在有烟雾的复杂战场环境下对陆战场军事目标进行有效识别。

#### 4 结论

本文从影响战场情报获取的关键问题军事目标

识别技术出发, 针对目前复杂战场环境下军事目标识别中所遇到的漏检误检率高、实时性差的问题, 提出了一种 PB-YOLO 目标识别算法并应用于军事目标识别领域中。通过自建军事目标数据集, 得出以下主要结论:

1) 在主干网络中引入通道-空间并行注意力机制, 使模型更好地捕捉目标在图像中的空间以及位置信息, 提升了模型对于战场中小目标特征的提取能力和识别准确率, 但使模型的训练时间更冗长。

2) 采用 BiFPN 加权特征融合网络替换原有的 FPN 与 PANet, 加快了模型对于目标语义信息与位置信息的融合速度, 减少了模型的训练时间, 增加了目标的识别精度, 同时只增加了少量的网络参数。

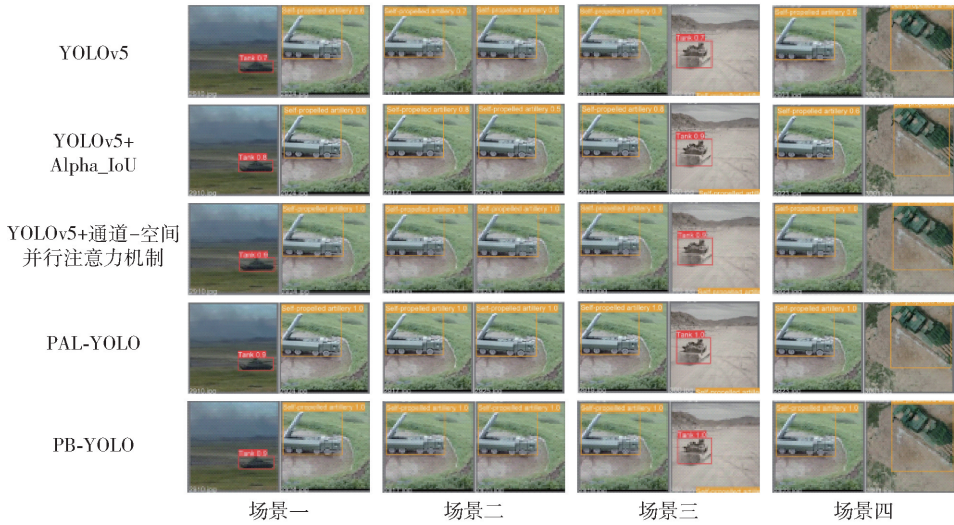


图 12 单个目标不同算法识别对比图

Fig. 12 Comparison of identification of different algorithms for a single target

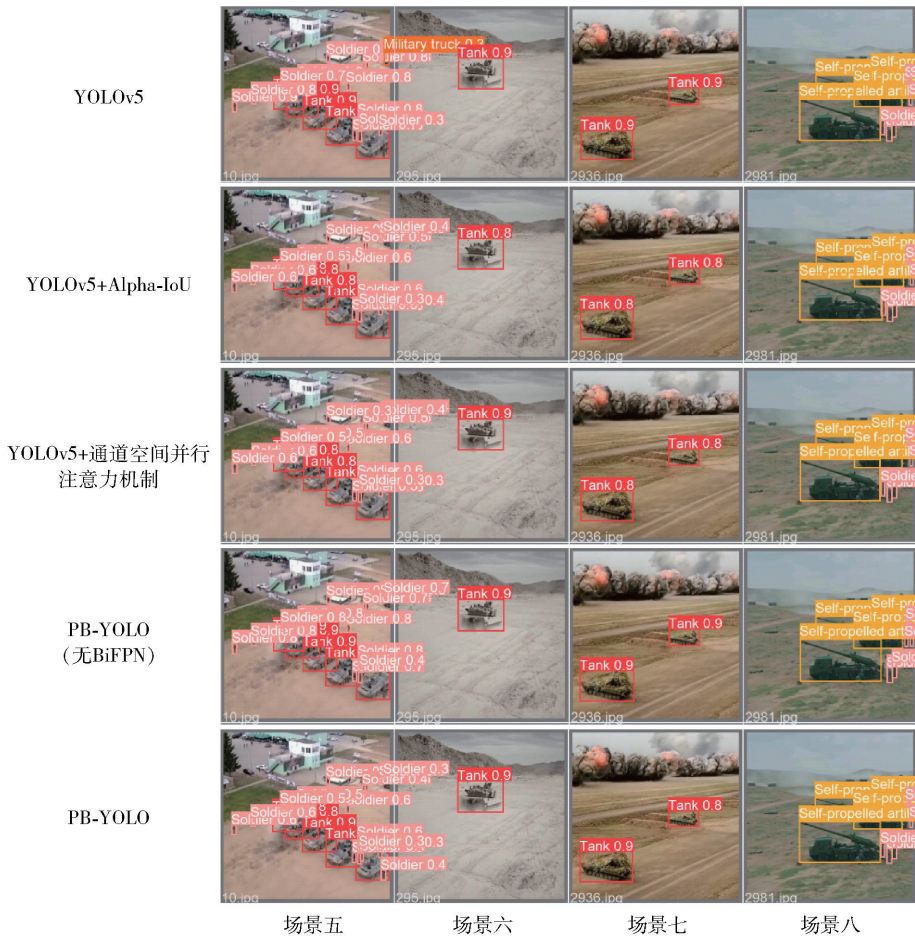


图 13 多目标识别对比图

Fig. 13 Multi-target recognition comparison chart

3) 使用 Alpha\_IoU 损失函数代替 GIoU 损失函数,使模型在解决了在对军事目标进行识别时目标框与预测框相互包裹而导致网络收敛速度降低、

GIoU 退化成 IoU 的问题,而且能够在不同程度上提升 mAP 值 0.3% ~0.7%。

4) 本文提出的目标识别算法需要在自建军事

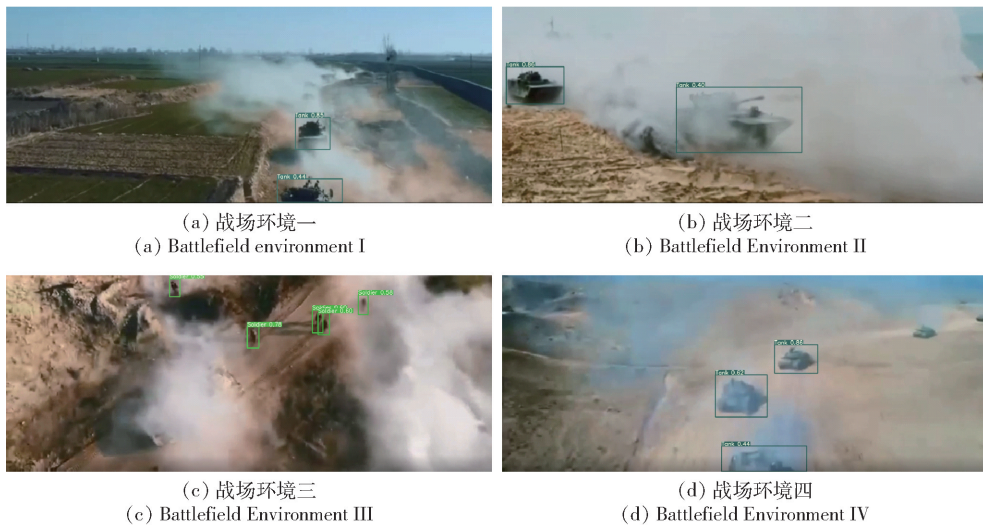


图 14 陆战场军事目标实时识别图

Fig. 14 Real-time identification map of military targets in land battlefield

目标数据集标注构建好后进行训练与测试。识别结果表明,相较于原始算法,本文提出的改进算法在识别精度与识别速度上均有提升,为复杂战场环境下的军事目标识别提供了参考。目前,针对复杂战场环境下的军事目标数据集还不够充足,后续可以考虑对其进行扩充以增加样本丰富性;后续研究还可以针对本文提出的模型,在保证精度的前提下缩小模型体积,进行轻量化处理,使其应用在低算力设备中提升模型普适性,以满足实际战场需求。

### 参考文献 (References)

- [1] CHEN C, HUANG J, PAN C Y, et al. Military image scene recognition based on CNN and semantic information [ C ] // Proceedings of the 2018 3rd International Conference on Mechanical, Control and Computer Engineering. Huhhot, China; IEEE, 2018: 573 - 577.
- [2] 于博文,吕明.改进的YOLOv3算法及其在军事目标检测中的应用[J].兵工学报,2022,43(2):345-354.  
YU B W, LÜ M. Improved YOLOv3 algorithm and its application in military target detection[J]. Acta Armamentarii, 2022, 43(2): 345 - 354. (in Chinese)
- [3] 刘俊,孟伟秀,余杰,等.面向军事目标识别的DRFCN深度网络设计及实现[J].光电工程,2019,46(4):21-30.  
LIU J, MENG W X, YU J, et al. Design and implementation of DRFCN deep network for military target recognition [ J ]. Optoelectronic Engineering, 2019, 46 ( 4 ): 21 - 30. (in Chinese)
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [ C ] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, US; IEEE, 2014:580-587.
- [5] WANG Z Z, XIE K, ZHANG X Y, et al. Small-object detection based on YOLO and dense block via image super-resolution [ J ]. IEEE Access, 2021, 9: 56416 - 56429.
- [6] GIRSHICK R. Fast R-CNN [ C ] // Proceedings of International Conference on Computer Vision. Piscataway, NJ, US: IEEE, 2015:1440 - 1448.
- [7] OTHMANI M. A vehicle detection and tracking method for traffic video based on faster R-CNN [ J ]. Multimedia Tools and Applications, 2022, 81(20):28347 - 28365.
- [8] SI J X, LIN J J, JIANG F, et al. Hand-raising gesture detection in real classrooms using improved R-FCN [ J ]. Neurocomputing, 2019, 359:69 - 76.
- [9] ZHANG C, ZHOU J W, WANG H W, et al. Multi-species individual tree segmentation and identification based on improved mask R-CNN and UAV imagery in mixed forests [ J ]. Remote Sensing, 2022, 14(4):874 - 874.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [ C ] // Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NE, US; IEEE, 2016:779 - 788.
- [11] 孙剑明,韩生权,沈子成,等.基于双卷积链的双目人体姿态距离定位识别[J/OL].兵工学报;1-8[2022-08-23]. DOI: 10.12382/bgxb.2021.0629.  
SUN J M, HAN S Q, SHEN Z C, et al. Binocular human body pose distance localization and recognition based on double convolution chain [ J/OL ]. Acta Armamentarii; 1 - 8 [ 2022-08-23 ]. DOI:10.12382/bgxb.2021.0629. (in Chinese)
- [12] XU M J, WANG X H, ZHANG S, et al. Detection algorithm of aerial vehicle target based on improved YOLOv3 [ J ]. Journal of Physics: Conference Series, 2022, 2284(1): 012 - 022.
- [13] KONG H, CHEN Z, YUE W J, et al. Improved YOLOv4 for pedestrian detection and counting in UAV images [ J ]. Computational Intelligence and Neuroscience, 2022, 2022: 6106853 - 6106853.

- [14] LIU G, CAO Z X, LIU S, et al. An improved SSD method for infrared target detection based on convolutional neural network [J]. *Journal of Computational Methods in Sciences and Engineering*, 2022, 22(4):1393-1408.
- [15] LIN T Y, GOYAL P, GRISHICK R, et al. Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2):318-327.
- [16] 潘浩. 基于深度学习的军事目标识别[D]. 杭州:杭州电子科技大学, 2018.  
PAN H. Military target recognition based on deep learning[D]. Hangzhou:Hangzhou Dianzi University, 2018. (in Chinese)
- [17] ZENG G Z, SONG R, HU X, et al. Applying convolutional neural network for military object detection on embedded platform[J]. *Computer Engineering and Technology*, 2018, 22(1):131-141.
- [18] 陈龙, 张峰, 蒋升. 小样本条件下基于深度森林学习模型的典型军事目标识别方法[J]. *中国电子科学研究院学报*, 2019, 14(3):232-237.  
CHEN L, ZHANG F, JIANG S. Typical military target recognition method based on deep forest learning model under small sample conditions [J]. *Journal of Chinese Academy of Electronic Sciences*, 2019, 14(3):232-237. (in Chinese)
- [19] 林洋, 董宝良, 刘泽平. 一种基于 CGAN 和 GcForest 的军事目标识别方法[J]. *信息技术*, 2020, 44(3):134-138.  
LIN Y, DONG B L, LIU Z P. A military target recognition method based on CGAN and GcForest [J]. *Information Technology*, 2020, 44(3):134-138. (in Chinese)
- [20] 杨朝红, 王伟男. 基于优化 SSD300 的小尺度典型军事目标识别方法研究[J]. *电脑与信息技术*, 2020, 28(4):19-22.  
YANG C H, WANG W N. Research on small-scale typical military target recognition method based on optimized SSD300 [J]. *Computer and Information Technology*, 2020, 28(4):19-22. (in Chinese)
- [21] 李鑫, 王晟全, 李昂. 基于非监督网络的军事目标识别算法的研究[J]. *电光与控制*, 2021, 28(10):36-39.  
LI X, WANG S Q, LI A. Research on military target recognition algorithm based on unsupervised network[J]. *Electro Optic and Control*, 2021, 28(10):36-39. (in Chinese)
- [22] PARK J, WOO S, LEE J Y, et al. Bam: Bottleneck attention module; arXiv:1807.06514 [R]. Ithaca, NY, US: Cornell University, 2018:1807.06514.
- [23] TAN M X, PANG R M, LE Q V. Efficientdet: Scalable and efficient object detection [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, WA, US:IEEE, 2020:10781-10790.
- [24] HE J B, ERFANI S, MA X J, et al. alpha-IoU: a family of power intersection over union losses for bounding box regression [J]. *Advances in Neural Information Processing Systems*, 2021, 34:20230-20242.
- [25] LÜ H H, YAN H B, LIU K Y, et al. YOLOv5-AC: attention mechanism-based lightweight YOLOv5 for track pedestrian detection[J]. *Sensors*, 2022, 22(15):5903.
- [26] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, US:IEEE, 2016:770-778.
- [27] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C] // *Proceedings of International Conference on Machine Learning*. Miami, FL, US:ICMLA, 2015:448-456.
- [28] 任江涛, 李定主, 屠惠琳. 基于训练网络的目标检测方法及应用[J]. *火力与指挥控制*, 2020, 45(4):173-177.  
REN J T, LI D Z, TU H L. Target detection method and application based on training network [J]. *Firepower and Command and Control*, 2020, 45(4):173-177. (in Chinese)