

基于加强特征融合的轻量化船舶目标检测^①

王 林, 汪钰婷

(西安理工大学 自动化与信息工程学院, 西安 710048)

通信作者: 汪钰婷, E-mail: 939249401@qq.com



摘 要: 针对基于深度学习的海上船舶目标检测任务中存在检测网络复杂且参数量大、检测实时性差的问题, 提出一种加强特征融合的轻量化 YOLOv4 算法——MA-YOLOv4. 首先使用 MobileNetv3 替换主干网络, 引入新的激活函数 *SiLU* 并使用深度可分离卷积代替普通 3×3 卷积降低网络参数量; 其次加入自适应空间特征融合模块加强特征融合; 最后使用 MDK-means 聚类算法得到适用于船舶目标的锚框, 用 Ship7000 数据集进行训练和评估. 实验结果表明, 改进算法与 YOLOv4 相比, 模型参数量降低 82%, mAP 提高 2.57%, FPS 提高 30 帧/s, 能实现对海上船舶的高精度实时检测.

关键词: 船舶检测; YOLOv4; MobileNetv3; 自适应空间特征融合; 聚类算法; 目标检测

引用格式: 王林, 汪钰婷. 基于加强特征融合的轻量化船舶目标检测. 计算机系统应用, 2023, 32(2): 288–294. <http://www.c-s-a.org.cn/1003-3254/8948.html>

Lightweight Ship Target Detection Based on Enhanced Feature Fusion

WANG Lin, WANG Yu-Ting

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: A lightweight YOLOv4 algorithm, MA-YOLOv4, is proposed to enhance feature fusion for addressing problems of complex detection networks, a large number of parameters, and poor real-time detection in deep learning-based maritime ship target detection tasks. Firstly, MobileNetv3 is employed to replace the backbone network, and a new activation function *SiLU* is introduced. The depthwise separable convolution is applied to replace the ordinary 3×3 convolution to reduce the number of network parameters. Secondly, an adaptive spatial feature fusion module is added to enhance feature fusion. Finally, the MDK-means clustering algorithm is adopted to get the anchor frame suitable for the ship target, and the Ship7000 dataset is utilized for training and evaluation. Experimental results show that compared with YOLOv4, the improved algorithm can reduce the number of model parameters by 82% and increase mAP by 2.57% and FPS by 30 f/s, which can achieve high-precision real-time detection of ships at sea.

Key words: ship detection; YOLOv4; MobileNetv3; adaptive spatial feature fusion; clustering algorithm; object detection

近年来, 随着世界经济的发展和各国对外贸易的不断增长, 海上交通运输产业得以飞速发展, 保证海上交通运输安全, 实现对海上目标的实时监测刻不容缓. 船舶作为海上主要检测目标, 准确快速检测船舶所处位置、识别船舶类型, 对海上异常行为预警、保证船舶安

全航行有非常重要的意义. 船舶检测易受到海上恶劣天气、复杂基岸背景、海面反光及海浪尾迹等因素影响, 因此在可见光图像中开展船舰检测任务具有挑战性.

传统目标检测方法通过人工设计的尺度不变^[1]、方向梯度直方图^[2]等特征来判别滑动窗口, 人工设计的特

① 基金项目: 陕西省科技计划重点项目 (2017ZDCXL-GY-05-03)

收稿时间: 2022-06-24; 修改时间: 2022-07-25; 采用时间: 2022-08-19; csa 在线出版时间: 2022-11-04

CNKI 网络首发时间: 2022-11-16

征对于输入的多样性变化没有较强鲁棒性,难以应用于复杂的海洋环境感知任务.近年来,随着中央处理器(CPU)、图形处理单元(GPU)等计算机硬件的快速发展,计算机的计算能力得到了极大提高.此外,PyTorch等深度学习软件框架使得基于神经网络的目标检测算法逐渐普及,并广泛应用于各种实际工程项目中.与传统的目标检测算法相比,基于深度学习的目标检测算法能够自动提取输入图像中具有鉴别性和鲁棒性的特征,检测速度更快,检测结果更准确.基于深度学习的目标检测算法主要分为两阶段目标检测算法和单阶段目标检测算法.两阶段目标检测算法主要以R-CNN系列^[3]为代表,其中心思想是在第1阶段通过区域建议的方式产生大量候选框,在第2阶段将候选区域送入神经网络,对目标进行分类和定位.虽然两阶段目标检测算法精度较高,但由于其具有两个相互独立的步骤,模型结构复杂,需要更长时间生成候选框,因此检测速度慢.单阶段目标检测算法主要以SSD^[4]、YOLO^[5]为代表,将目标检测任务作为回归问题来处理,使用预定义的不同尺度的锚框来定位目标,实现端到端的训练,检测速度较快,但由于结构设计简单,对小目标的检测能力较弱.

自R-CNN问世以来,国内外学者尝试着将基于深度学习的目标检测算法应用于海上船舶目标检测任务中.Zwemer等^[6]开发了一个使用港口监控摄像机的实时船舶检测和跟踪系统,改进的SSD模型能支持对船舶尺寸和长宽比极端变化的检测.Kim等^[7]提出了一种概率船舶探测与分类系统,该算法利用Faster R-CNN^[8]网络对每一帧进行处理,将检测器的置信度作为概率,根据连续图像之间的相关性对连续帧的概率进行贝叶斯融合.Shao等^[9]构建了一个包含31455张图像的大规模船舶数据集Seaships(公开7000张,Ship7000),对Faster R-CNN、YOLOv2^[10]和SSD模型进行了训练和测试,其中YOLOv2模型具有较好的检测精度和计算速度.刘博等^[11]将残差网络的思想引入Darknet网络中提取船舶特征,在此基础上用加入了惩罚机制的YOLOv3^[12]进行目标区域和类别预测.Li等^[13]在YOLOv3-tiny的骨干网络中加入注意力模块,并使用卷积层代替最大池化层,提高了对小目标的检测精度.Liu等^[14]模拟了不同天气条件下的海上船舶图像,在此基础上训练神经网络以获得更加鲁棒的船舶检测模型.孔刘玲等^[15]首先用数据增强方式解决数据集不平衡的问题,并对YOLOv4^[16]算法提出改进,改

进后算法采用Softer-NMS^[17]对预测框进行后处理,有效解决因船舶遮挡导致的漏检问题.

上述算法虽能有效检测出船舶,但仍存在网络参数量较大、实时性较差、目标定位不准确的问题.针对现有船舶检测算法的问题,本文提出轻量化MA-YOLOv4算法,通过最大距离K-means算法获得初始锚框,用MobileNetv3进行初步特征提取,并添加自适应空间特征融合模块(adaptive spatial feature fusion, ASFF)加强不同尺度的特征融合,在保证检测精度的前提下,大幅提高检测速度,有更强的应用价值.

1 YOLOv4 算法概述

YOLOv4算法在YOLOv3的基础上进一步改进,它结合了近年来优秀的算法思想,在MS COCO目标检测数据集上,不降低每秒帧数(frames per second, FPS)的前提下,平均精度均值(mean average precision, mAP)提高到44%.YOLOv4网络主要组成部分有主干特征提取模块Backbone、特征融合模块Neck及检测模块Prediction.

特征提取模块将跨阶段局部网络(CSPNet)^[18]集成到Darknet53中,通过截断梯度流策略,使梯度流通过不同的网络路径传播,不但防止不同网络层学习重复的梯度信息,而且消除了计算瓶颈,减少了内存消耗,促进了低层次信息的融合,获得了更强的特征提取能力.在YOLOv4的Neck部分,原有的特征金字塔网络(feature pyramid network, FPN)被空间金字塔池化(space pyramid pooling, SPP)和路径聚合网络(path aggregation network, PANet)所取代,该模块首先通过卷积核大小为(1×1)、(5×5)、(9×9)、(13×13)的SPP模块增大感受野,再通过PANet进一步提取和融合不同层的特征,在一定程度上缓解了目标位置信息和轮廓信息的特征丢失问题.检测模块使用YOLOv3算法中原始的检测头YOLO head.

2 改进的YOLOv4网络

本文所提出的YOLOv4算法主要有3部分改进,一是在主干部分使用MobileNetv3替换CSPDarknet53,引入新的激活函数SiLU,并使用深度可分离卷积代替普通3×3卷积,二是加入ASFF模块,三是使用最大距离K-means聚类算法得到适用于船舶目标的先验框.改进后的MA-YOLOv4网络结构如图1所示.

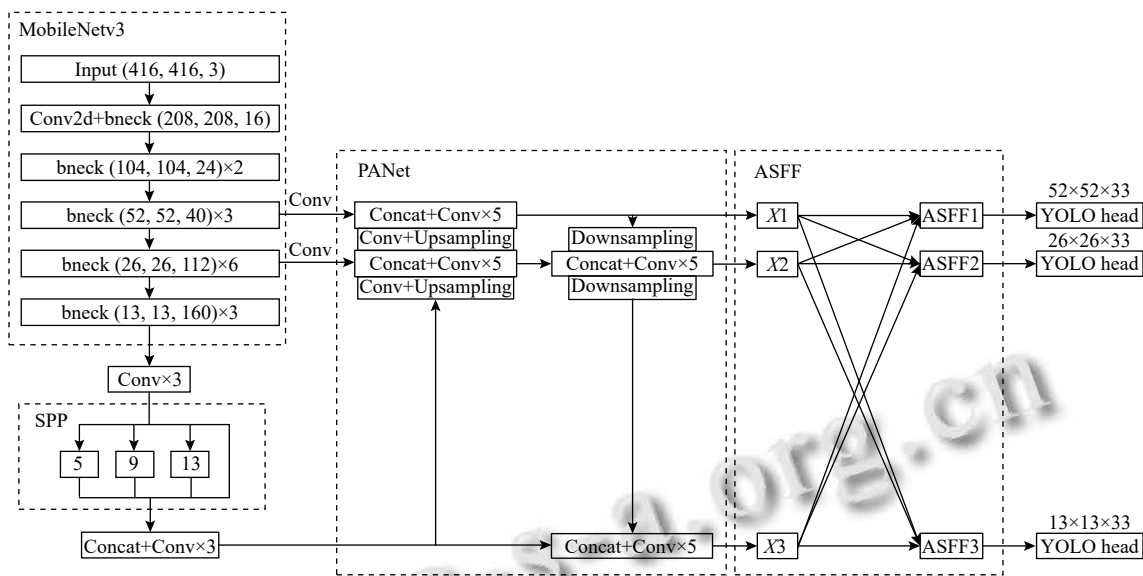


图1 MA-YOLOv4 网络结构

2.1 骨干网络轻量化

随着人们对网络的不断深化,模型计算带来的巨大存储压力和计算负担已经开始限制了深度学习模型的应用领域.传统卷积神经网络(CNN)内存需求大,计算量大,无法在移动设备和嵌入式设备上运行.因此引入轻量级深度神经网络 MobileNetv3 作为 YOLOv4 主干特征提取网络,它是一个模型尺寸较小,可训练参数和运算量较少的 CNN,适用于移动设备.

MobileNetv3 使用深度可分离卷积代替标准卷积块,深度可分离卷积由深度卷积和逐点卷积组成,这种操作极大程度减少了模型的理论参数量(params)和运算量(floating point operations, FLOPs).标准卷积和深度可分离卷积操作如图2所示.

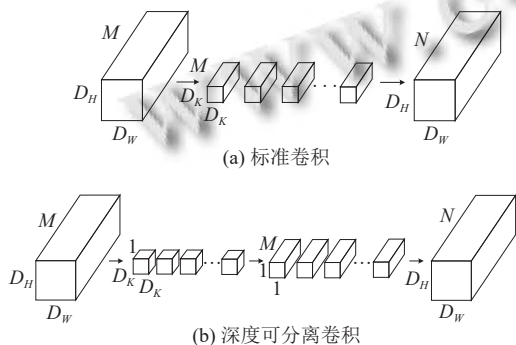


图2 标准卷积和深度可分离卷积

设输入特征图宽高为 D_W 、 D_H , 输入通道数为 M , 卷积核大小为 D_K , 输出通道数为 N , 标准卷积的参数数量和运算量如下:

$$P_s = D_K \cdot D_K \cdot M \cdot N \tag{1}$$

$$F_s = D_K \cdot D_K \cdot M \cdot N \cdot D_W \cdot D_H \tag{2}$$

深度可分离卷积的参数数量和运算量如下:

$$P_d = D_K \cdot D_K \cdot M + M \cdot N \tag{3}$$

$$F_d = D_K \cdot D_K \cdot M \cdot D_W \cdot D_H + M \cdot N \cdot D_W \cdot D_H \tag{4}$$

深度可分离卷积与标准卷积的参数数量和运算量之比均为 $\frac{1}{N} + \frac{1}{D_K^2}$, 若采用 3×3 的卷积核, 深度可分离卷积的参数数量和运算量将降为标准卷积的 $\frac{1}{8} \sim \frac{1}{9}$.

MobileNetv3 使用特殊的瓶颈结构——逆残差结构, 如图3所示. 每个逆残差网络由两个 1×1 普通卷积和一个 3×3 深度可分离卷积组成. 该网络首先将特征图的维数通过 1×1 卷积扩张, 然后通过 3×3 深度可分离卷积提取特征, 最后通过 1×1 卷积压缩通道. 并引入挤压和激励注意力机制, 放置于深度可分离卷积之后, 通过学习自动获取到每个特征通道的重要程度, 为每个特征通道赋予一定的权重, 让网络重点关注有用特征通道并抑制无用特征通道.

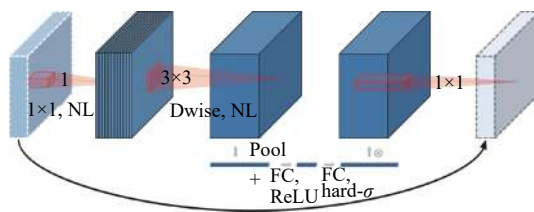


图3 逆残差结构

MobileNetv3 网络的浅层部分利用ReLU作为激活函数, 计算公式如下:

$$ReLU = \max(0, x) \quad (5)$$

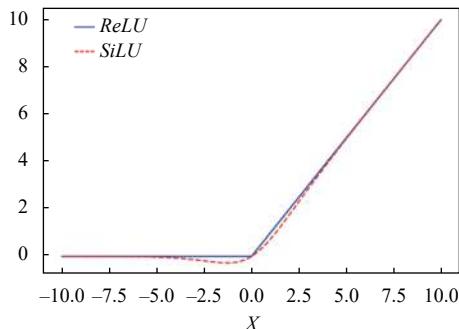
当激活函数的输入大于 0, 即 $x > 0$ 时, 激活函数的输入作为输出; 当 $x < 0$ 时, 输出为 0. 该函数收敛速度快, 能有效解决Sigmoid函数中的梯度消失问题, 但当输入都小于 0 时, 输出全部为 0, 梯度为 0, 神经元将全部“死亡”, 参数无法更新.

本文将ReLU函数替换为SiLU函数, SiLU函数计算公式如下:

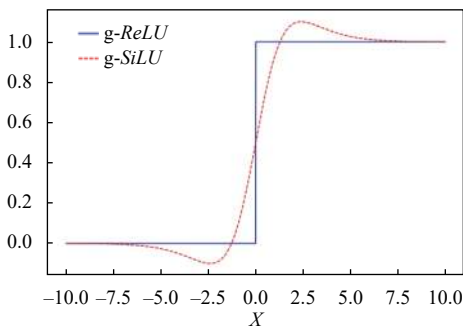
$$SiLU = x \cdot Sigmoid(x) \quad (6)$$

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

ReLU函数及SiLU函数如图4所示, 可以看出SiLU函数无上界有下界, 平滑且非单调, 在 $x > 0$ 时, 不存在梯度消失的情况; 在 $x < 0$ 时, 神经元也不会像ReLU一样出现“死亡”的情况, 且全局最小值能在梯度为零的情况下抑制大权值更新, 避免梯度爆炸.



(a) ReLU函数和SiLU函数



(b) ReLU函数导数和SiLU函数导数

图4 激活函数及其导数图像

本文利用改进后的MobileNetv3网络获得的 $52 \times 52 \times 40$ 、 $26 \times 26 \times 112$ 、 $13 \times 13 \times 160$ 这3个有效特征层来进行特征融合模块的构建. 并用深度可分离卷积替换PANet和YOLO Head中的标准卷积, 减少网

络参数量.

2.2 自适应空间特征融合

随着网络层次的深化, 特征的语义信息由低维向高维转变. 高级特征为对象分类提供了丰富的语义信息, 而低级特征为对象定位提供了丰富的细粒度信息. 虽然PANet进行了初始特征融合, 但对于小目标需要更多的信息融合. 因此, 本文构建了一种高效的特征融合结构, 引入ASFF模块, 在PANet进行初始的跨层特征融合后, 将特征导入ASFF模块, 为每个融合的特征图设置自学习权值来进行加权融合, 使船舶信息得到更深入的整合. 图5为ASFF模块结构.

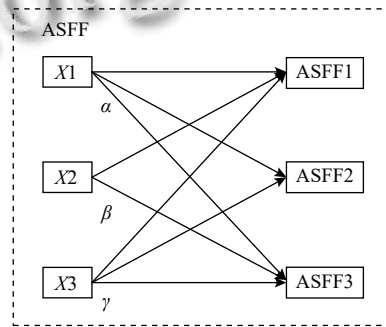


图5 ASFF模块结构

ASFF1、ASFF2、ASFF3分别对应PANet的输出 $X1$ 、 $X2$ 、 $X3$. 以ASFF3为例, 要得到ASFF3的输出, 首先通过 1×1 的卷积调整 $X1$ 、 $X2$ 的通道数, 然后通过上采样调整特征图的宽高, 使其与 $X3$ 一致, 之后将调整后特征图乘以相应的权重系数, 最后将具有位置信息的低分辨率特征图与具有语义信息的高分辨率特征图相加, 得到融合后的特征图. ASFF输出层计算公式如下:

$$y_{ij}^l = \alpha_{ij}^l \cdot x_{ij}^{1 \rightarrow l} + \beta_{ij}^l \cdot x_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \cdot x_{ij}^{3 \rightarrow l} \quad (8)$$

其中, y_{ij}^l 表示经过自适应特征融合后得到的新的第 l 层特征图; $x_{ij}^{n \rightarrow l}$ 表示第 n 层特征图经过大小、维度调整后与第 l 层特征图大小与维度相同的特征向量; α_{ij}^l 、 β_{ij}^l 、 γ_{ij}^l 表示网络通过自适应学习得到的第 l 层的权重系数, $\alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l = 1$, α_{ij}^l 、 β_{ij}^l 、 $\gamma_{ij}^l \in [0, 1]$.

权重系数计算公式如下:

$$\alpha_{ij}^l = \frac{e^{\lambda_{aij}^l}}{e^{\lambda_{aij}^l} + e^{\lambda_{bij}^l} + e^{\lambda_{cij}^l}} \quad (9)$$

其中, λ_{aij}^l 为第 l 层的自学习权重控制参数, 由调整后的特征向量经过 1×1 卷积得到.

2.3 锚框聚类算法优化

在目标检测任务中,选择合适的锚框可以显著提高检测速度和精度. YOLOv4 算法的锚框是基于 COCO 数据集,通过 K-means 算法聚类得到的通用目标锚框,对长宽比较大的船舶目标并不适用. 且 K-means 算法对初始聚类中心敏感,选择不同的初始值可能导致不同的簇划分规则. 为避免上述问题,本文用最大距离改进 K-means 算法,针对船舶数据集进行重新聚类. 最大距离 K-means (max distance K-means, MDK-means) 算法^[19]的核心思想是在初始化聚类中心时使初始聚类中心之间的相互距离尽可能远,确定初始聚类中心步骤如算法 1.

算法1. MDK-means算法确定初始聚类中心

初始化: 数据样本集合 $X=\{x_1, x_2, \dots, x_n\}$, 初始聚类中心数 k .

- (1) 计算各样本点之间的距离 $d(x_i, x_j), i, j=1, \dots, n$;
- (2) 取满足 $\max\{d(x_i, x_j)\}$ 的两个样本点作为前两个初始聚类中心, 记为 x^1, x^2 , 初始聚类中心集合 $C=\{x^1, x^2\}$, $X=\{x_1, x_2, \dots, x_{n-2}\}, x_1, x_2, \dots, x_{n-2} \notin C$, 已得到的初始聚类中心数 $c=2$.
- (3) 计算各样本与初始聚类中心 x^1, \dots, x^c 距离的乘积 $d(x_i, x^1) \times \dots \times d(x_i, x^c), i=1, \dots, n-c$.
- (4) 取满足 $\max\{d(x_i, x^1) \times \dots \times d(x_i, x^c)\}$ 的样本点作为第 $c+1$ 个初始聚类中心, 记为 $x^{c+1}, C=\{x^1, \dots, x^{c+1}\}, X=\{x_1, x_2, \dots, x_{n-c-1}\}, x_1, x_2, \dots, x_{n-c-1} \notin C, c=c+1$.
- (5) 当 $c < k$ 时, 迭代步骤(3)–(4), 否则迭代结束, 得到 k 个初始聚类中心 x^1, \dots, x^k .

本文首先通过 MDK-means 算法确定 k 个聚类中心, 之后通过传统 K-means 算法进行聚类, 选取 $k=9$, 即获得满足 YOLOv4 算法的 3 个不同尺度的 9 个先验框. 通过 MDK-means 算法得到的船舶锚框尺寸与原始锚框尺寸对比如表 1, 可以看出, 聚类后的锚框尺寸更符合船舶矮胖型特点.

表 1 聚类前后锚框尺寸对比

锚框	大尺度	中尺度	小尺度
原始锚框	(142, 110)	(36, 75)	(12, 16)
	(192, 243)	(76, 55)	(19, 36)
	(459, 401)	(72, 146)	(40, 28)
聚类后船舶锚框	(171, 87)	(84, 43)	(22, 11)
	(238, 57)	(116, 28)	(40, 16)
	(316, 105)	(140, 49)	(63, 26)

3 实验验证与分析

3.1 实验环境与超参数初始化

本文的硬件配置和实验环境如表 2 所示.

本实验利用 PyTorch 深度学习框架, 并部署表 2 实验环境. 在训练阶段, 需要初始化一些超参数. 凭借

迁移学习^[20], 网络采用冻结训练和解冻训练的策略, 可以加快训练效率, 防止权重被破坏, 并使用 Mosaic 数据增强技术丰富数据集, 模型训练时使用学习率余弦退火衰减算法和标签平滑. MA-YOLOv4 的超参数初始化如表 3 所示.

表 2 硬件配置和实验环境

名称	配置
操作系统	Ubuntu
处理器 (CPU)	Intel(R) Xeon(R) CPU E5-2640 v4 @ 2.20 GHz
显卡 (GPU)	GeForce GTX 1080Ti (x4), 11 GB
深度学习框架	PyTorch 1.4.0
集成开发环境	Anaconda, CUDA 11.2
编程语言	Python 3.7

表 3 MA-YOLOv4 超参数初始化

超参数	值	超参数	值
输入图片大小	416×416	冻结训练epoch	50
Batch size	32	冻结训练学习率	0.001
动量因子	0.9	解冻训练epoch	150
优化器	Adam	解冻训练学习率	0.0001

3.2 实验数据集

本文选择邵振峰等人提出的 Ship7000 船舶数据集^[9]进行实验. 该数据集基于河道船舶可见光监测视频获取, 均来自真实河道场景. 数据集内共有 7000 张分辨率是 1920×1080 的图片, 分为 6 大类: 矿砂船、集装箱船、散货船、普通货船、渔船和客船. 由于 Ship7000 涵盖了各种尺度、亮度和角度的目标, 因此其非常适用于多尺度舰船目标检测任务. 训练集、验证集和测试集按 8:1:1 的比例分配.

3.3 实验结果及分析

3.3.1 消融实验

为验证本文改进算法的有效性和可靠性, 用 mAP@0.5、Params、FPS 作为评估指标, 将 YOLOv4 算法作为基线算法, 并将第 2 节改进分为 4 组, 进行消融实验. 其中, “√”代表在网络中使用该改进策略, “×”代表在网络中不使用该改进策略.

由表 4 可知, 第 1 组为使用 MDK-means 聚类获得适合船舶目标先验框的 YOLOv4 算法, 模型 mAP@0.5 提升了 5.35%, 表明采用适合船舶数据集的先验框的必要性; 第 2 组在第 1 组的基础上将 YOLOv4 骨干网络替换为 MobileNetv3 网络, 检测精度下降 4.53%, FPS 提高了 89%, 这是由于 MobileNetv3 网络层数少, 导致特征提取能力变差, 但网络参数量减少使得 FPS

提高;第3组在第2组的基础上将 PANet 模块和检测头中的 3×3 普通卷积替换为深度可分离卷积,卷积层特征提取能力基本不变,网络参数减少 71%,在检测精度几乎持平的条件下使得 FPS 提升 44%;第4组在第

3组的基础上引入 ASFF 模块并添加至 PANet 模块之后,模型检测精度提高 1.83%,FPS 达到 49 帧/s,满足实时性要求,由此可以说明 MA-YOLOv4 算法的有效性和可靠性。

表4 消融实验结果对比

模型	MDK-means	MobileNetv3	深度可分离卷积	ASFF	mAP@0.5 (%)	Params	FPS (帧/s)
YOLOv4	×	×	×	×	87.45	69040001	19
1	√	×	×	×	92.80	69040001	19
2	√	√	×	×	88.27	39989933	36
3	√	√	√	×	88.19	11729069	52
4	√	√	√	√	90.02	12438716	49

3.3.2 MA-YOLOv4 算法与其他算法对比

为进一步验证模型性能,将改进后的算法与经典算法进行对比实验,并引入评估指标 mAP@0.75,实验结果如表5所示。可以看出,MA-YOLOv4 算法相较于 YOLOv3-tiny 算法,在 FPS 几乎持平的前提下 mAP@0.5 提高 11.63%,mAP@0.75 提高 11.47%;相较于 YOLOv4-tiny 和 YOLOv5s, mAP@0.5 分别提高 9% 和 4.46%,mAP@0.75 分别提高 8.83% 和 5.01%,但检测速度远不如二者;相较于 YOLOv4 算法,mAP@0.5 提高 2.57%,mAP@0.75 提高 2.49%,FPS 提高 30 帧/s。实验说明本文算法在检测精度方面达到最佳,定位准确度也较高,检测速度相较 YOLOv4-tiny 和 YOLOv5s 虽然还有很大差距,但仍满足实时性要求,模型性能良好。

3.3.3 检测结果可视化

为了更直观地展示 MA-YOLOv4 算法的检测效

果,在测试集中选取 3 张图片进行可视化展示。如图 6 所示,图 6(a) 为原图,图 6(b) 为 YOLOv4 算法检测结果,图 6(c) 为 MA-YOLOv4 算法检测结果。由第 1 行结果可以看出,YOLOv4 算法存在目标漏检问题,本文算法对小目标更敏感,能检测到更远处较小的船舶;由第 2 行结果可以看出,本文算法检测框定位更准确、更符合船舶真实形状;由第 3 行结果可以看出,YOLOv4 算法对被遮挡的船舶检测效果较差,本文算法能准确检测出被遮挡船舶,并有较高的置信度。

表5 不同算法在 Ship7000 数据集上的检测结果

算法	mAP@0.5 (%)	mAP@0.75 (%)	FPS (帧/s)
YOLOv3-tiny	78.39	64.57	48
YOLOv4-tiny	81.02	67.21	73
YOLOv5s	85.56	71.03	97
YOLOv4	87.45	73.55	19
MA-YOLOv4	90.02	76.04	49



图6 不同算法检测效果对比图

4 结论

本文针对YOLOv4网络参数量大、对船舶检测实时性不理想的问题,首先通过最大距离改进K-means从而得到更符合船舶目标的锚框,其次将原算法中的主干网络替换成浅层激活函数为SiLU的MobileNetv3网络,并用深度可分离卷积代替普通卷积降低网络参数,且引入ASFF模块加强特征融合,设计出MA-YOLOv4算法.本文算法的优势在于网络参数量较小,能更充分学习不同尺度的特征.实验结果表明MA-YOLOv4算法能较好平衡检测速度与精度,更适合船舶检测任务.

参考文献

- 1 Lowe DG. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, 60(2): 91–110. [doi: [10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94)]
- 2 Dalal N, Triggs B. Histograms of oriented gradients for human detection. *Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego: IEEE, 2005. 886–893. [doi: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177)]
- 3 Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014. 580–587. [doi: [10.1109/CVPR.2014.81](https://doi.org/10.1109/CVPR.2014.81)]
- 4 Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector. *Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer, 2016. 21–37.
- 5 Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 779–788.
- 6 Zwemer MH, Wijnhoven RGJ, de With PHN. Ship detection in harbour surveillance based on large-scale data and CNNs. *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. Madeira: CCIS, 2018. 153–160.
- 7 Kim K, Hong S, Choi B, et al. Probabilistic ship detection and classification using deep learning. *Applied Sciences*, 2018, 8(6): 936. [doi: [10.3390/app8060936](https://doi.org/10.3390/app8060936)]
- 8 Ren SQ, He KM, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
- 9 Shao ZF, Wu WJ, Wang ZY, et al. SeaShips: A large-scale precisely annotated dataset for ship detection. *IEEE Transactions on Multimedia*, 2018, 20(10): 2593–2604. [doi: [10.1109/TMM.2018.2865686](https://doi.org/10.1109/TMM.2018.2865686)]
- 10 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE, 2017. 6517–6525.
- 11 刘博, 王胜正, 赵建森, 等. 基于Darknet网络和YOLOv3算法的船舶跟踪识别. *计算机应用*, 2019, 39(6): 1663–1668. [doi: [10.11772/j.issn.1001-9081.2018102190](https://doi.org/10.11772/j.issn.1001-9081.2018102190)]
- 12 Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv: 1804.02767, 2018.
- 13 Li H, Deng LB, Yang C, et al. Enhanced YOLOv3 tiny network for real-time ship detection from visual image. *IEEE Access*, 2021, 9: 16692–16706. [doi: [10.1109/ACCESS.2021.3053956](https://doi.org/10.1109/ACCESS.2021.3053956)]
- 14 Liu RW, Yuan WQ, Chen XQ, et al. An enhanced CNN-enabled learning method for promoting ship detection in maritime surveillance system. *Ocean Engineering*, 2021, 235: 109435. [doi: [10.1016/j.oceaneng.2021.109435](https://doi.org/10.1016/j.oceaneng.2021.109435)]
- 15 孔刘玲, 刘秀文. 基于改进YOLOv4算法的船舶目标检测方法. *船舶工程*, 2022, 44(1): 96–103, 147. [doi: [10.13788/j.cnki.cbge.2022.01.17](https://doi.org/10.13788/j.cnki.cbge.2022.01.17)]
- 16 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. arXiv: 2004.10934v1, 2020.
- 17 He YH, Zhu CC, Wang JR, et al. Bounding box regression with uncertainty for accurate object detection. *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach: IEEE, 2019. 2883–2892. [doi: [10.1109/CVPR.2019.00300](https://doi.org/10.1109/CVPR.2019.00300)]
- 18 Wang CY, Liao HYM, Wu YH, et al. CSPNet: A new backbone that can enhance learning capability of CNN. *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Seattle: IEEE, 2020. 1571–1580.
- 19 翟东海, 鱼江, 高飞, 等. 最大距离法选取初始簇中心的K-means文本聚类算法的研究. *计算机应用研究*, 2014, 31(3): 713–715, 719. [doi: [10.3969/j.issn.1001-3695.2014.03.017](https://doi.org/10.3969/j.issn.1001-3695.2014.03.017)]
- 20 Pan SJ, Yang Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10): 1345–1359. [doi: [10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191)]

(校对责编: 牛欣悦)