

改进YOLOX-S的红外舰船目标检测算法

娄树理 王岩 郭建勤 公维峰

Infrared ship target detection algorithm based on improved YOLOX-S

LOU Shuli, WANG Yan, GUO Jianqin, GONG Weifeng

引用本文:

娄树理, 王岩, 郭建勤, 等. 改进YOLOX-S的红外舰船目标检测算法[J]. 应用光学, 2023, 44(5): 1054–1060. DOI: 10.5768/JAO202344.0502006

LOU Shuli, WANG Yan, GUO Jianqin, et al. Infrared ship target detection algorithm based on improved YOLOX-S[J]. *Journal of Applied Optics*, 2023, 44(5): 1054–1060. DOI: 10.5768/JAO202344.0502006

在线阅读 View online: <https://doi.org/10.5768/JAO202344.0502006>

您可能感兴趣的其他文章

Articles you may be interested in

[基于注意力机制与图卷积神经网络的单目红外图像深度估计](#)

Depth estimation of monocular infrared images based on attention mechanism and graph convolutional neural network

应用光学. 2021, 42(1): 49–56 <https://doi.org/10.5768/JAO202142.0102001>

[基于双路多尺度金字塔池化模型的显著目标检测算法](#)

Salient target detection algorithm based on dual-channel multi-scale pyramid pooling model

应用光学. 2021, 42(6): 1056–1061 <https://doi.org/10.5768/JAO202142.0602007>

[基于音视频信息融合的目标检测与跟踪算法](#)

Object detection and tracking algorithm based on audio-visual information fusion

应用光学. 2021, 42(5): 867–876 <https://doi.org/10.5768/JAO202142.0502007>

[基于视觉SLAM和目标检测的语义地图构建](#)

Semantic SLAM based on visual SLAM and object detection

应用光学. 2021, 42(1): 57–64 <https://doi.org/10.5768/JAO202142.0102002>

[基于嵌入式GPU的红外弱小目标检测算法](#)

Infrared weak small target detection algorithm based on embedded GPU

应用光学. 2020, 41(5): 1089–1095 <https://doi.org/10.5768/JAO202041.0506004>

[结合视觉注意力机制和图像锐度的无参图像质量评价方法](#)

No-reference image quality assessment method based on visual attention mechanism and sharpness metric approach

应用光学. 2018, 39(1): 51–56 <https://doi.org/10.5768/JAO201839.0101009>



关注微信公众号，获得更多资讯信息

文章编号:1002-2082(2023)05-1054-07

改进 YOLOX-S 的红外舰船目标检测算法

娄树理¹, 王 岩¹, 郭建勤², 公维峰³

(1. 烟台大学 物理与电子信息学院, 山东 烟台 264005; 2. 山东电子职业技术学院 电子与通信工程系, 山东 济南 250200;
3. 高效能服务器和存储技术国家重点实验室, 山东 济南 250101)

摘要: 红外舰船目标检测与识别技术是反舰导弹红外成像制导的关键, 对于武器装备制导性能具有重大意义。针对在复杂环境下红外舰船目标检测的精度和速度问题, 提出了改进 YOLOX-S 的红外舰船目标检测算法。首先引入深度可分离卷积(depthwise separable convolution)代替 FPN (feature pyramid network) 及 YOLOHead 残差结构中的传统卷积, 降低模型的参数量; 其次引入 ECANet 通道注意力机制, 提高网络的注意力, 降低舰船目标的虚检率和漏检率; 最后使用 CIoU 损失函数, 进一步提高网络的检测准确率。实验表明, 改进后算法的检测平均精度(AP)达到 98%, 检测速度为 56 帧/s, 对比改进前 YOLOX-S 算法, 检测速度与平均精度分别提升 6 帧/s 和 3%, 且模型更加轻量化。实验结果充分证明本文提出的算法能够有效完成红外舰船目标检测任务。

关键词: 红外舰船; 目标检测; 注意力机制; 损失函数

中图分类号: TN219; TP391

文献标志码: A

DOI: 10.5768/JAO202344.0502006

Infrared ship target detection algorithm based on improved YOLOX-S

LOU Shuli¹, WANG Yan¹, GUO Jianqin², GONG Weifeng³

(1. School of Physics and Electronic Information, Yantai University, Yantai 264005, China; 2. Department of Electronics and Communication Engineering, Shandong College of Electronic Technology, Jinan 250200, China;
3. State Key Laboratory of High-end Server and Storage Technology, Jinan 250101, China)

Abstract: Infrared ship target detection and recognition technology is the key technology of infrared imaging guidance for anti-ship missile, which is of great significance to the guidance performance of weapon equipment. For the accuracy and speed in complex environment, an infrared ship target detection algorithm based on YOLOX-S was proposed. Firstly, the depthwise separable convolution (DSC) was introduced to replace the traditional convolution in the feature pyramid network (FPN) and YOLOHead residual structures, which could reduce the amount of parameters of the model. Secondly, the ECANet channel attention mechanism was introduced to improve the attention of network, which could reduce the false detection rate and missed detection rate of ship targets. Finally, the CIoU loss function was used to further improve the detection accuracy of the network. The experimental results show that the average precision (AP) of the optimized algorithm reaches 98% and the detection speed is 56 frame/s. Compared with the previous YOLOX-S algorithm, the detection speed and the average precision are improved by 6 frame/s and 3%, respectively, and the model is more lightweight. The experimental results fully prove that the proposed algorithm can effectively complete the infrared ship target detection task.

Key words: infrared ship; target detection; attention mechanism; loss function

引言

在现代海战中, 能否快速准确地识别出舰船目标并引导精确制导武器对其摧毁, 是削弱敌方战

斗力量的关键。红外成像制导具有制导精度高、抗干扰能力强、隐蔽性好等优点, 在各种精确制导体系中占据着重要的地位^[1]。由于红外舰船图像

收稿日期: 2022-09-21; 修回日期: 2022-12-12

基金项目: 山东省自然科学基金 (ZR2019LZH016)

作者简介: 娄树理 (1976—), 男, 博士, 副教授, 主要从事光电信息处理、机器视觉研究。E-mail: shulilou@sina.com

通信作者: 王岩 (1998—), 男, 硕士研究生, 主要从事图像处理研究。E-mail: kkudbj@163.com

受海杂波、雨雪等干扰影响严重,且目标和背景复杂多样,如何快速准确地检测识别出舰船目标,是红外成像制导武器的技术难题和关键问题,一直是国内外研究的热点。

传统的目标检测算法性能大多依赖于人工设计的特征和分类器^[2],且人工设置的特征适应性较差,难以很好地检测复杂场景下的舰船目标。随着深度学习的发展,许多基于深度学习的目标检测算法相继涌现,这类算法通过神经网络提取目标特征,无需人工设计特征,提高了目标检测的精度与效率。基于深度学习的目标检测算法分为两类:一类是双阶段目标检测算法,如 R-CNN(region-based convolutional neural networks)^[3] 和 Faster R-CNN(faster region-based convolutional neural networks)^[4] 等算法。顾佼佼^[5]等人在 Faster R-CNN 算法中将特征图进行拼接形成多尺度特征,提高特征图语义信息;AVOLA D^[6]等人在 Faster R-CNN 算法中引入多流(MS)架构,在每个流上施加不同的核大小来模拟多尺度图像分析。这类算法精度高,但检测速度较慢,难以达到实时检测的效果。另一类是单阶段目标检测算法,如 SSD(single shot multibox detection)^[7] 和 YOLO(you only look once)^[8-11] 系列等算法。CHOI H T^[12]等人在 SSD 算法中引入了由注意流和特征映射级联流组成的增强特征映射块(EMB);陈耀祖^[13]等人在 YOLOv4 的算法中使用模糊 ISODATA 动态聚类算法对先验框数目进行优化;MUKHIDDINOV M^[14]等人采用 h-swish

激活函数来减少 YOLOv4 的运行时间。这类目标检测算法为端到端的目标检测,相比于双阶段目标检测算法,其检测速度更快,但检测精度略低。

针对如何在复杂环境中快速准确检测出红外舰船目标的技术难题,本文提出了一种改进后的 YOLOX-S 红外舰船目标检测算法,将深度可分离卷积代替 FPN 及 YOLOHead 残差结构中的传统卷积,提高网络的运行与检测速度;引入 ECANet(efficient channel attention network)通道注意力机制,提高对舰船目标的识别能力,降低舰船目标的虚检率与漏检率;引用 CIoU(complete-IoU)损失函数,使目标框回归更加稳定,提高网络对舰船目标的检测精度。

1 改进的 YOLOX-S 算法模型

1.1 YOLOX-S 算法改进

YOLOX^[15] 算法是旷视科技研究院在 2021 年提出的新型单阶段目标检测算法,相较之前的 YOLO 系列算法, YOLOX 算法的改进主要包括数据增强、无锚点和预测分支解耦,不仅在平均精度(mAP)上超越了 YOLOv3、YOLOv4 和 YOLOv5,在检测速度上也极具竞争力。YOLOX 系列主要包括 YOLOX-S、YOLOX-M、YOLOX-L 和 YOLOX-X,其中 YOLOX-S 参数量最少,方便部署且具有代表性。YOLOX-S 的网络结构模型如图 1 所示,首先对输入图像进行 Mosaic 数据增强,其次经过 Backbone 部分进行浅层特征提取,输出的 3 个特

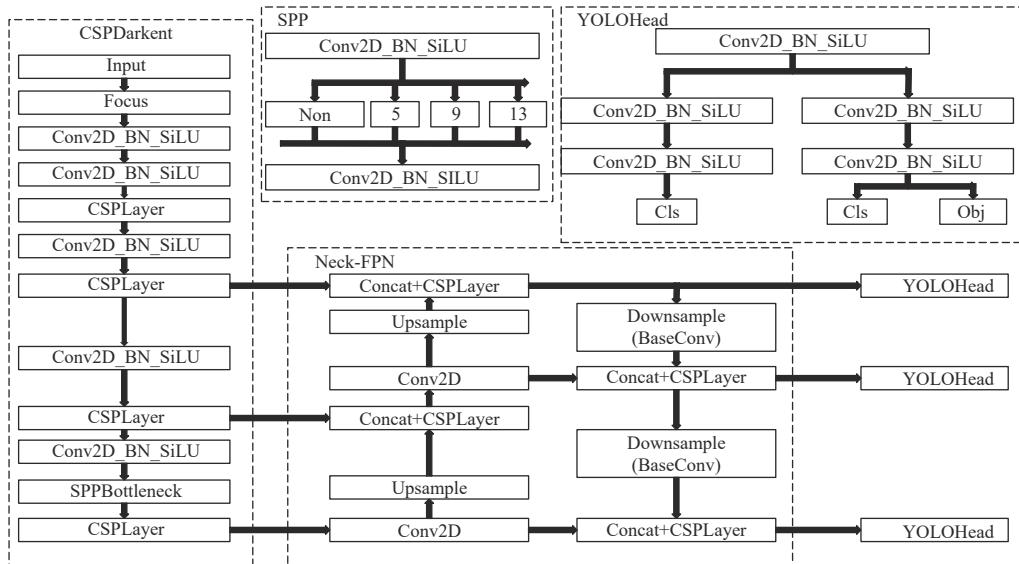


图 1 YOLOX-S 网络结构模型
Fig. 1 YOLOX-S network structure model

征层传入加强特征提取网络进行深度特征提取,最后将得到的 3 个有效特征层分别传入 Decoupled Head 进行目标边界框预测。

为提高网络检测速度与精度,本文对 YOLOX-S 算法进行了改进,改进后的网络结构图如图 2 所示,首先通过主干网络 CSPDarknet(cross stage par-

tial darknet) 对图像进行特征提取;将提取出的特征图作为输入,利用 ECANet 通道注意力机制,过滤冗余信息,加强模型重要信息提出能力;将增强后的特征图传入改进后的 Neck-FPN 中,进一步增强特征图的语义信息;最后将特征图分别传入改进后的 Decoupled Head 中进行目标框预测。

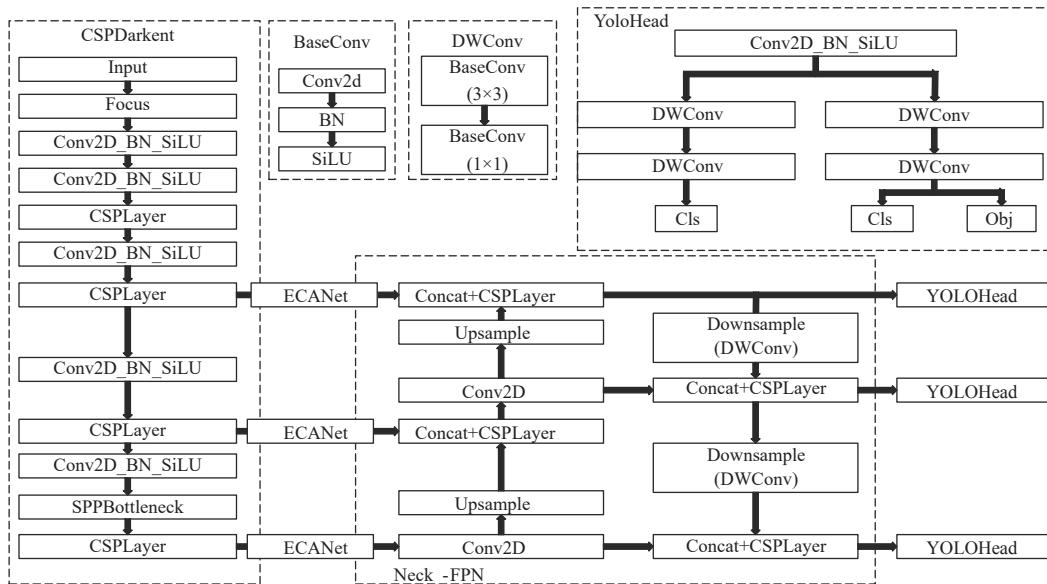


图 2 改进的 YOLOX-S 网络结构模型

Fig. 2 Improved YOLOX-S network structure model

1.2 深度可分离卷积

由于海战场环境瞬息万变,算法的检测速度尤为重要。YOLOX-S 网络中的残差结构均使用传统卷积^[16],传统卷积的参数量及计算量过大,严重影响算法的检测速度。为了提高算法的检测速度,使其更容易快速实现,本文引入深度可分离卷积,降低模型的参数量,提高算法的检测速度。为保证模型检测的准确率,将深度可分离卷积用于 FPN(feature pyramid networks)中下采样部分及 YOLOHead 残差结构中,主干部分残差结构及 CSPLayer 部分保持不变。

传统卷积先将各个通道输入的特征图与相应卷积核进行卷积相乘,最后累加得到输出特征,主要结构如图 3 所示。

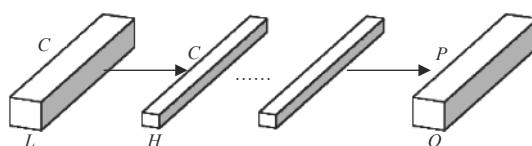


图 3 传统卷积结构

Fig. 3 Traditional convolution structure

图 3 中 L 和 Q 分别为输入和输出图像的大小, H 为卷积核大小, C 和 P 分别为输入和输出的通道数。传统卷积(SC)的参数量 W 为

$$W_{SC} = H^2 \times C \times P \quad (1)$$

计算量 O 为

$$O_{SC} = H^2 \times C \times P \times Q^2 \quad (2)$$

深度可分离卷积^[17]则是将提取特征和结合特征分为 3×3 的深度卷积和 1×1 的逐点卷积两部分,首先对每一个输入通道进行一个深度卷积核操作,然后利用 1×1 的卷积将深度卷积输出结果结合到特征中组合为新的特征,其结构图如图 4 所示。

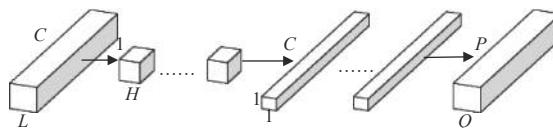


图 4 深度可分离卷积结构

Fig. 4 Depthwise separable convolution structure

深度可分离卷积(DSC)的参数量 W 为

$$W_{DSC} = H^2 \times C + C \times P \quad (3)$$

计算量 O 为

$$O_{DSC} = H^2 \times C \times Q^2 + C \times P \times Q^2 \quad (4)$$

两种卷积参数量与计算量比为

$$F_w = \frac{W_{DSC}}{W_{SC}} = \frac{1}{P} + \frac{1}{H^2} \quad (5)$$

$$F_o = \frac{O_{DSC}}{O_{SC}} = \frac{1}{P} + \frac{1}{H^2} \quad (6)$$

1.3 ECANet 通道注意力机制

红外舰船图像主要反映舰船目标与海天背景的辐射能量差异, 由于岛屿等物体辐射能量和海天背景也有较大差异, 实际检测时岛屿等物体会对目标检测形成干扰, 并且实际红外图像中噪声影响严重, 这些因素严重影响红外舰船目标的检测精度。为了增强网络的抗干扰能力, 降低网络对舰船目标的虚检率与漏检率, 本文在 YOLOX-S 主干网络后引入 ECANet 通道注意力机制^[18], 抑制岛屿等物体干扰信息的影响, 提高网络对舰船目标的检测精度。

在 ECANet 通道注意力机制过程中, 首先对输入的特征图进行全局平均池化, 其次利用 1D 卷积进行特征提取, 经过 Sigmoid 激活函数获得每一个特征点的权值, 最后与输入的特征图相乘, 得到新的特征图。ECANet 通道注意力机制结构如图 5 所示。

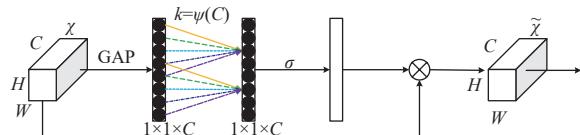


图 5 ECANet 通道注意力机制

Fig. 5 ECANet channel attention mechanism

图 5 中, H 、 W 、 C 分别为输入特征图的高度、宽度、通道数, GAP(global average pooling) 为全局平均池化, k 表示卷积核大小。 σ 为 Sigmoid 非线性激活函数, 具体表达式为^[19]

$$\sigma = f(z) = \frac{1}{1 + e^{-z}} \quad (7)$$

式中 z 表示神经网络上一层节点的输出。 k 的表达式为^[18]

$$k = \psi(C) = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right|_{\text{odd}} \quad (8)$$

式中: γ 取 2, b 取 1; $|t|_{\text{odd}}$ 表示距离 t 最近的奇数。

1.4 损失函数

在 YOLOX-S 算法中, 损失函数包括类别损失 (L_{cls})、置信度损失 (L_{obj}) 和边界框损失 (L_{IoU})。IoU(intersection over union) 为交并比, 反映目标检测中预测框与真实框的重叠程度, 其公式为^[20]

$$I_{IoU} = \frac{A \cap B}{A \cup B} \quad (9)$$

式中: A 表示预测边界框; B 表示真实边界框。由于 IoU 并没有考虑两框之间的距离, 当预测框与真实框无相交时, 无法反映两框的距离大小, 导致网络准确率下降。因此本文引入 CIoU(complete-IoU) 损失函数, CIoU 考虑预测框与真实框之间的距离、尺度以及重叠率等因素, 增加了目标框回归的稳定性, 提高了网络对舰船目标的检测精度。CIoU 的公式为^[11]

$$CIoU = I_{IoU} - \frac{\rho^2(b, b_{gt})}{c^2} - \alpha v \quad (10)$$

式中: b 、 b_{gt} 分别为预测框和真实框的中心位置; $\rho^2(b, b_{gt})$ 为两框中心位置的欧式距离; c 为两框外接矩形的对角线距离; α 为权重系数; v 表示衡量长宽比的相似性。 α 的计算公式为

$$\alpha = \frac{v}{1 - I_{IoU} + v} \quad (11)$$

v 的计算公式为

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w}{h} \right)^2 \quad (12)$$

式中: w 、 h 分别为预测框的宽度、高度; w_{gt} 、 h_{gt} 分别为真实框的宽度和高度。

CIoU 的计算公式为

$$CIoU = 1 - I_{IoU} + \frac{\rho^2(b, b_{gt})}{c^2} + \alpha v \quad (13)$$

类别损失包含检测目标的类别信息, 置信度损失包含图像的前景与背景信息, 两者损失使用 BCEWithLogitsLoss 函数进行计算, 该函数计算公式见式(14)、式(15), 类别与置信度损失使用该函数计算时网络稳定性更强。

$$L_B = \{l_1, \dots, l_N\} \quad (14)$$

$$l_n = -[y_n \cdot \log(\sigma(x_n)) + (1 - y_n) \cdot \log(1 - \sigma(x_n))] \quad (15)$$

式中: x 、 y 为输入的张量; N 为 Batch Size; n 为每批次预测的标签数; c 为目标类别; σ 为 Sigmoid 非线性激活函数。针对本文任务中舰船目标尺寸不规则等特点, 加大 CIoU 损失权重 w , 以此来增加模型对舰船目标尺寸变化的感知能力, 提高准确度。损失函数最终表达式为

$$L = w \times CIoU + L_{cls} + L_{obj} \quad (16)$$

2 实验验证与结果分析

2.1 实验环境及参数设置

本文所有仿真实验均在 Windows11 操作系

统; 搭载的 CPU 版本为 AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz; GPU 为 NVIDIA GeForce GTX 3060(6 GB); 深度学习框架为 Pytorch 1.10.1。训练过程分为冻结阶段和解冻阶段, 冻结阶段过程中的参数为: epoch=50, batchsize=6, momentum=0.937, Init_lr=1e-3, Min_lr = Init_lr×0.01, 优化器使用 adam。解冻后训练参数为: epoch=100, batchsize=4, momentum=0.937, Init_lr=1e-3, Min_lr = Init_lr×0.01, 优化器使用 adam。

2.2 数据集介绍

由于开源的红外舰船图像很少, 因此本文实验所用数据集来源于实际采集的红外舰船图像, 共 1 263 张舰船图像, 并且按照 PASCAL VOC 数据集的格式使用 LabelImg 图片标注工具对图像进行标注。该数据集的图像波段为 8 μm~12 μm 和 3 μm~5 μm, 图像大小为 320 像素×240 像素、621 像素×471 像素, 包含多种不同形状的舰船目标。该数据集按照 9:1 的比例分为训练集和验证集, 且图像中包含多种背景及干扰目标, 符合本文实验的需要。

2.3 评价标准

为了从不同方面评价所提算法对舰船目标检测的准确、实时性, 本文选取了平均精度(average precision, AP)和检测速度(frame per second, FPS)作为评价指标。AP 兼顾了精准率(precision)和召回率(recall), 常用来评估模型的精度有效性, FPS 为每秒检测图像的帧数, 用来评估模型的实时性。Params 为模型参数量的大小, 用来评估模型的复杂度。

精准率是指模型预测的所有舰船目标中, 预测正确的比例。召回率是指所有真实舰船目标中, 模型预测正确的目标比例。两者表达式为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (17)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (18)$$

式中: P 为精确率; R 为召回率; N_{TP} 为 $I_{oU}>0.5$ 的检测框数量, N_{FP} 为 $I_{oU}\leq 0.5$ 的检测框数量; N_{FN} 为没有检测到的真实检测框的数量。

平均精度(AP)可以用来评估舰船目标检测的检测效果, 它是 $P-R$ 曲线在 [0,1] 区间内精准率对召回率的积分, 即:

$$P_{AP} = \int_0^1 P(R) dR \quad (19)$$

式中 P_{AP} 表示舰船目标检测平均精度。所有类别

的 P_{AP} 的均值为 M_{AP} , 由于本文算法为单目标检测, 即 $P_{AP}=M_{AP}$ 。

2.4 算法性能分析与比较

为验证改进的 YOLOX-S 算法在舰船目标检测任务中的性能表现, 将其与现阶段主流的目标检测算法进行对比实验, 考虑到实验所用数据集存在制作误差等因素, 算法对比实验使用 $IoU=0.75$ 条件下的精度对比。对比结果如表 1 所示。

表 1 不同算法对比

Table 1 Comparison of different algorithms

Model	Backbone	Precision%	AP%	FPS	Params
Faster R-CNN	ResNet-50	0.98	0.97	4	108.6 M
SSD	VGG	0.95	0.92	47	99.76 M
YOLOv4	MobileNetv2	0.94	0.92	52	41.2 M
YOLOv5-s	CSPDarknet	0.94	0.94	54	27.76 M
YOLOX-S	CSPDarknet	0.96	0.95	50	34.21 M
Improved YOLOX-S	CSPDarknet	0.98	0.98	56	25.79 M

通过对比可知, 改进后的 YOLOX-S 算法 AP 达到 98%, FPS 达到 56 帧/s, 相较于改进前 YOLOX-S 算法, AP、FPS 分别提高了 3%、6 帧/s, 精度与速度都得到了一定的提升, 并且 Params 减少了 24.6%, 能够更容易快速实现。与其他算法相比, 改进后的 YOLOX-S 算法在精度、速度和参数量大小等方面, 具有明显的优势。

为验证每个模块对算法的影响, 本文设置消融实验, 使用相同的实验设备以及数据集对网络不同模块组合进行性能测试, 并对实验数据分析每个模块的性能, 其中“√”表示算法在实验中使用了改进模块, “-”表示未使用改进模块, 实验结果如表 2 所示。

表 2 消融实验结果对比表

Table 2 Comparison of ablation experiment results

DSC	ECANet	Loss_CIoU	AP%	FPS	Params
-	-	-	0.95	50	34.21 M
√	-	-	0.94	58	25.79 M
-	√	-	0.97	49	34.21 M
-	-	√	0.96	50	34.21 M
√	√	-	0.97	56	25.79 M
√	√	√	0.98	56	25.79 M

由表 2 可知, 对比原始算法, 引入深度可分离卷积模块 DSC 后, 虽然精度略低于原始算法, 但检测速度提升较大, 参数量也较大幅度的减少; 单独引入 ECANet 模块后, 虽然检测速度略有下降, 但

算法精度得到提升;单独引入 CIoU 损失函数模块,算法精度略有提升。多个模块组合实验中,对比引入单个模块,引入深度可分离卷积和 ECANet 模块后,检测精度和检测速度均有较大提升。实验数据表明,本文提出的改进算法,具有检测速度快和参数量小的优势,且算法平均精度得到进一步提升,达到 98%。

为验证本文算法的有效性,选取不同系列的红外舰船图像,与原始算法进行效果检测对比。选取 A、B、C 3 类典型海面目标红外图像,图像 A 类

为 $8 \mu\text{m} \sim 12 \mu\text{m}$ 的海面船只图像,图像中主要干扰是天气和海杂波的干扰,也存在探测器盲元带来的噪声干扰;图像 B 类为 $3 \mu\text{m} \sim 5 \mu\text{m}$ 的海面船只图像,此波段下海杂波干扰相对较为严重;图像 C 类为 $8 \mu\text{m} \sim 12 \mu\text{m}$ 的海面船只图像,相比前两类干扰源之外,此类还有岛岸背景干扰。两种算法的检测效果对比图如图 6 所示,其中图 A(g)、图 B(g)、图 C(g) 为本文改进的 YOLOX-S 算法检测的图像,图 A(y)、图 B(y)、图 C(y) 为原始算法检测的图像。

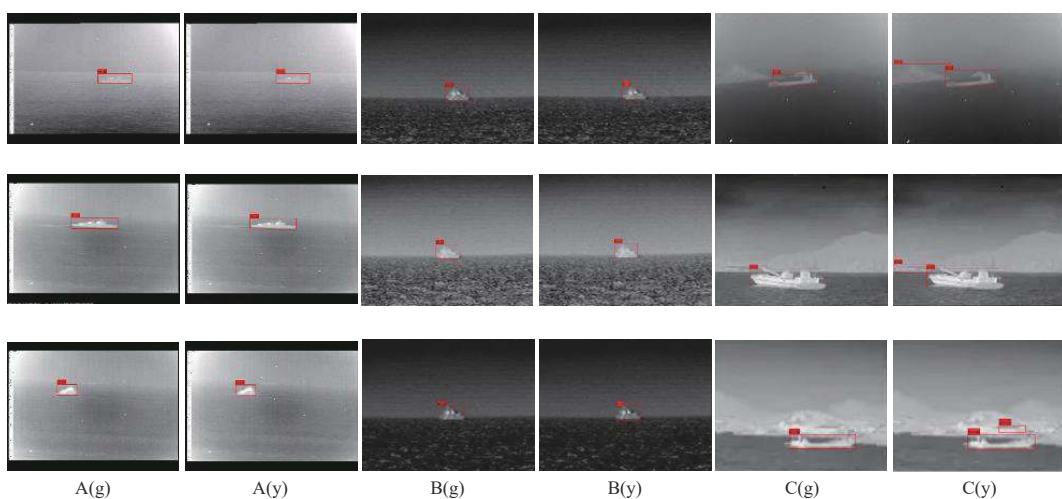


图 6 检测效果对比

Fig. 6 Comparison of detection effect

实验结果表明,原始算法对 A、B 类图像的检测效果较好,能够较完整地检测出舰船目标,由于 C 类图像中存在大面积岛屿干扰等因素,原始算法对 C 类图像检测出现将岛岸错检为目标的情况,检测效果较差;本文改进的 YOLOX-S 算法对 A、B、C 3 类图像的检测效果均良好,验证了本文算法检测精度更高、抗干扰性更强,尤其在大面积岛屿干扰的背景下,算法检测精度提升较大,能够充分胜任当前的检测任务。

3 结论

本文将 YOLOX-S 算法应用到红外舰船目标检测领域中,针对舰船目标检测任务提出一种改进的 YOLOX-S 算法,使用深度可分离卷积,在保证网络检测精度的前提下,提高检测速度,同时结合注意力机制与 CIoU 损失函数,进一步提升算法的检测性能。相较于改进前 YOLOX-S 算法,舰船目标检测平均精度提高了 3%,检测速度提高了

6 帧/s,算法参数量减少了 24.6%。同时,对比实验结果表明,本文提出的改进算法与现阶段主流的检测算法相比,具有较好的检测性能,能够胜任舰船目标检测任务。

参考文献:

- [1] 赵超, 杨号. 红外制导的发展趋势及其关键技术[J]. 电光与控制, 2008(5): 48-53.
- [2] 方路平, 何杭江, 周国民, 等. 目标检测算法研究综述[J]. 计算机工程与应用, 2018, 54(13): 11-18.
FANG Luping, HE Hangjiang, ZHOU Guomin, et al. Research overview of object detection methods[J]. Computer Engineering and Applications, 2018, 54(13): 11-18.
- [3] DUNDAR A, JIN J H, MARTINI B, et al. Embedded streaming deep neural networks accelerator with applications[J]. Neural Networks & Learning System, 2017, 28(7): 1572-1583.
- [4] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015

- IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 1440-1448.
- [5] 顾佼佼, 李炳臻, 刘克, 等. 基于改进Faster R-CNN的红外舰船目标检测算法[J]. 红外技术, 2021, 43(2): 170-178.
GU Jiaojiao, LI Bingzhen, LIU Ke, et al. Infrared ship target detection algorithm based on improved faster R-CNN[J]. Infrared Technology, 2021, 43(2): 170-178.
- [6] AVOLA D, CINQUE L, DIKO A, et al. MS-Faster R-CNN: multi-stream backbone for improved Faster R-CNN object detection and aerial tracking from UAV images[J]. *Remote Sensing*, 2021, 13(9): 1670.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// Proceedings of European Conference on Computer Vision. [S. l.]: Springer International Publishing, 2016: 21-37.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 779-788.
- [9] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2017: 6517-6525.
- [10] REDMON J, FARHADI A. Yolov3: an incremental improvement[EB/OL]. (2018-04-08)[2022-09-19]. <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]. (2020-03-23)[2022-09-19]. <https://arxiv.org/abs/2004.10934>.
- [12] CHOI H T, LEE H J, KANG H, et al. SSD-EMB: an improved SSD using enhanced feature map block for object detection[J]. *Sensors*, 2021, 21(8): 2842.
- [13] 陈耀祖, 谷玉海, 成霞, 等. 基于优化YOLOv4算法的行驶车辆要素检测方法[J]. *应用光学*, 2022, 43(2): 248-256.
CHEN Yaozu, GU Yuhai, CHENG Xia, et al. Driving vehicle elements detection method based on optimized YOLOv4 algorithm[J]. *Journal of Applied Optics*, 2022, 43(2): 248-256.
- [14] MUKHIDDINOV M, ABDUSALOMOV A B, CHO J. Automatic fire detection and notification system based on improved YOLOv4 for the blind and visually impaired[J]. *Sensors*, 2022, 22(9): 3307.
- [15] GE Z, LIU S, WANG F, et al. YOLOX: exceeding yolo series in 2021[EB/OL]. (2021-08-06)[2022-09-19]. <https://arxiv.org/pdf/2107.08430.pdf>.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 1-6.
- [17] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. New York: IEEE, 2018: 4510-4520.
- [18] MISHRA S K, RAI G, SAHA S, et al. Efficient channel attention based encoder-decoder approach for image captioning in hindi[J]. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 2022, 21(3): 1-17.
- [19] YIN X Y, GOUDRIAAN J, LANTINGA E A, et al. A flexible sigmoid function of determinate growth[J]. *Annals of botany*, 2003, 91(3): 361-371.
- [20] YU J, JIANG Y, WANG Z, et al. Unitbox: an advanced object detection network[C]//Proceedings of the 24th ACM international conference on Multimedia. [S. l.]: Association for Computing Machinery, 2016: 516-520.