

DOI: 10.13382/j.jemi.B2407771

适应遥感船舶图像的轻量化旋转小目标检测网络*

孙伟^{1,2} 沈欣怡¹ 张小瑞³ 管菲¹

(1. 南京信息工程大学自动化学院 南京 210044; 2. 江苏省大气环境与装备技术协同创新中心 南京 210044;
3. 南京工业大学计算机与信息工程学院 南京 211816)

摘要: 遥感图像船舶目标小、背景复杂、姿态变化大,传统船舶检测算法为提升精度,往往忽视了模型规模和实时性,难以应用到资源受限的设备上。针对上述问题,提出一种适应遥感船舶图像的轻量化旋转小目标检测网络(RFDNet)。考虑到遥感船舶图像拍摄距离远而导致图片中目标较小且图像中包含丰富的背景信息,设计注意力卷积融合双分支网络(ACFNet),通过对局部特征信息与全局空间感知信息的充分提取,提高船舶小目标检测精度;为避免船舶目标姿态各异而导致检测时的精度下降,利用旋转目标方向信息引入旋转边界框损失函数,获得更准确的边界框回归损失,提升任意方向旋转船舶目标的检测性能;针对为提高模型精度而带来的参数量增加问题,在特征融合部分引入轻量级卷积,将卷积、深度可分离卷积以及通道混洗相结合,减少模型的参数量。通过对比实验和消融实验证明,RFDNet在HRSC2016数据集和DOTA数据集上的平均精度均值(mAP)分别达到了97.63%和81.63%,模型参数降到了 2.99×10^6 ,在有效提升检测精度的同时实现了模型的轻量化设计,为遥感船舶目标检测算法在资源受限设备上的应用提供了新思路。

关键词: 遥感图像;旋转目标检测;深度学习;特征融合;损失函数

中图分类号: TP391;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 510.4

Lightweight rotating small target detection network adapted to remote sensing ship images

Sun Wei^{1,2} Shen Xinyi¹ Zhang Xiaorui³ Guan Fei¹

(1. School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China;

2. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing 210044, China;

3. School of Computer and Information Engineering, Nanjing Tech University, Nanjing 211816, China)

Abstract: Remote sensing images of ships is characterized by small target sizes, complex backgrounds, and significant attitude changes. Traditional ship detection algorithms focus on improving detection accuracy while neglecting model size and real-time performance, thereby limiting their practical application on resource-constrained devices. A lightweight rotated fusion detection network RFDNet adapted to remote sensing ship images is proposed to address the above problems. Considering that the remote sensing ship images are taken at a long distance, resulting in small target sizes and rich background information in the images, ACFNet is designed to improve the detection accuracy of small ship targets by fully extracting local feature information and global spatial sensing information. To avoid accuracy degradation when detecting ship targets with different attitudes, a rotating bounding box loss function is introduced, which utilizes the orientation information of rotating targets for obtaining a more accurate bounding box regression loss, thereby improving the detection performance of ship targets rotating in any direction; for the problem of increasing parameter counts brought about by increasing the accuracy of the model, a lightweight convolution is introduced into the feature fusion part, which combines the convolution, the depth separable convolution, and the channel blending to reduce the number of parameters in the model. Through comparative and ablation experiments, it has been demonstrated that RFDNet achieved mAPs of 97.63% and 81.63% on the HRSC2016 and DOTA datasets, respectively, while reducing the model parameters to 2.99×10^6 . This not only effectively improved detection accuracy but also realized a

收稿日期: 2024-08-17 Received Date: 2024-08-17

* 基金项目: 国家自然科学基金(62376128, 62272236)项目资助

lightweight model design, providing a new insight for the application of remote sensing ship detection algorithms to resource-constrained devices.

Keywords: remote sensing image; rotating target detection; deep learning; feature fusion; loss function

0 引言

近年来,随着计算机视觉的快速发展和应用,船舶图像的目标检测成为海洋监测、安全管理和资源勘探等领域的重要技术手段。通过从卫星、飞机和无人机等平台获取的遥感图像,能够实现对广阔海域的实时监测和数据采集。遥感图像能够提供大量高分辨率的海洋表面信息,使得人们能够更精确地掌握海洋动态和船舶活动。

船舶目标检测技术的发展经历了从传统的人工特征提取到现代的自动特征提取的演变。早期的船舶目标检测主要依赖于人工设计的船舶特征,有以下几种方法^[1]:对船舶图像进行边缘检测^[2]、形状分析^[3]和纹理特征提取^[4]等对船舶进行检测。这些方法在简单背景下能够取得一定效果,但在复杂海洋背景中往往表现不佳。随着卷积神经网络(convolutional neural networks, CNN)等深度学习技术的崛起^[5],基于深度学习的目标检测算法,如YOLO(you only look once)算法、Faster R-CNN和RetinaNet等,在船舶目标检测中展现了卓越的性能。这些算法通过大规模标注数据进行训练,能够自动学习图像中的复杂特征,大幅提高了检测的准确性和鲁棒性。

在船舶检测任务中,由于船舶图片拍摄距离较远,船舶目标较小^[6],检测结果对位置误差的敏感度高、依赖性强^[7],很小的位置误差就会导致目标检测错误。现有的一些检测方法基于卷积神经网络提出了一系列改进模型^[8]。针对小目标的问题,Zhao等^[9]利用卷积操作进行改进,使用多分辨率卷积特征以提高小型和密集船舶的检测精度,这些卷积操作通过局部感受野提取特征,从而提取到更多的细节特征。然而卷积层逐层传递信息,远距离像素之间的信息传递效率低,无法有效捕捉图像中的长程依赖关系,在聚焦小目标特征的同时会忽略对背景信息的提取和利用。除了卷积操作的改进外,Zhang等^[10]在检测模型中引入注意力机制来捕捉全局信息,注意力机制能够学习输入数据中不同部分的重要性。然而,这种方法在检测小目标的任务中易丢失目标的细节特征。

虽然卷积操作和注意力机制可以在全局和局部上捕捉到更多的特征信息,但是由于遥感设备的移动性,拍摄出的遥感图像中船舶目标姿态变化大,再加上船舶目标较小,容易对检测精度造成影响^[11]。传统的水平框目标检测算法对旋转目标的检测能力有限^[12],并且在检测多角度旋转船舶目标时,水平标注框内会包含很多冗余的

背景信息,对检测结果造成干扰^[13]。因此,一些针对旋转目标检测的方法被提出^[14],Ding等^[15]针对遥感图像中船舶方向各异的问题,将任意方向的船舶表示为带有头部点的旋转框,通过关键点估计来确定舰船的中心,回归出目标船舶的中心点和尺寸。这种方法虽然通过回归中心点确定了目标的位置,但并没有完全利用目标的旋转角度信息,并且为检测模型带来了巨大的模型参数。

为了应对背景复杂、检测目标小以及船舶目标旋转变化等挑战,大部分方法通过调整网络结构和修改损失函数来提高精度的方法在船舶检测领域虽然取得了一定的效果,但也为模型引入了大量的参数,并且忽视了模型的规模和实时性,不利于将检测算法部署到卫星或无人机等嵌入式设备中。因此,在实际应用中,迫切需要研究一种新的更加精确、轻量且能适应目标旋转变化的遥感图像船舶检测方法。

综上所述,为解决遥感图像船舶目标小容易被复杂背景噪声干扰、船舶姿态变化大以及模型计算复杂度高的难题,研究提出了一种基于遥感图像的旋转船舶目标检测网络(rotated fusion detection network, RFDNet)。针对船舶遥感图像中存在的部分船舶目标较小且包含丰富背景信息的问题,提出具有全局分支和局部分支的注意力卷积融合的双分支网络(attention convolution fusion network, ACFNet)。通过有效捕捉图像中的长距离依赖关系,该网络利用卷积操作提取图像中的局部特征细节信息的同时使用注意力机制获取全局信息,实现卷积和注意力机制的互补,提取更加丰富的特征信息。由于遥感设备拍摄角度任意,为了充分考虑旋转船舶目标的角度信息,引入概率交并比(IoU)来衡量预测边界框与真实边界框的重叠程度,并提出旋转边界框损失函数 $Loss_{RB}$ 作为检测模型的边界框损失,实现对不同角度和方向目标的精确检测,增强检测模型对旋转目标的适应能力。为降低模型精度改进带来的参数量增加,提出一种创新的轻量化卷积方法LDCConv并用其替代网络中的原始标准卷积。LDCConv利用普通标准卷积和深度可分离卷积分别对特征进行提取和细化,再结合通道混洗重组特征以增强信息交换。通过将输入通道分组及重新排列减少计算量,有助于模型的轻量化。

1 方法

改进后的适应遥感船舶图像的轻量化旋转小目标检测网络框架如图1所示。该网络由主干网络、颈部网络

以及头部网络组成。将输入图片送入主干网络进行特征提取,主干网络中提取到的船舶目标特征在颈部网络进行更深层次的特征传递和融合。其中,在颈部网络和头部网络之间加入 ACFNet,将全局特征和局部信息进行融合;并在头部网络进行边界框损失和回归损失的计算,通过引入旋转边界框损失 $Loss_{RB}$,充分考虑船舶旋转角度信息对检测的影响,得到更加准确的预测目标框的坐标信息,并将颈部网络中的普通卷积方法替换为轻量级卷积 LDCConv 减少模型参数。

1.1 注意力卷积融合双分支网络

卷积和注意力机制是改进模型时最常用的两种方法。为了结合卷积使用局部感受野提取特征的特点以及注意力机制更关注全局信息的优点,通过将卷积操作和注意力机制进行融合提出了一个注意力卷积融合的双分支网络 ACFNet 以实现特征的多样化表示和提升网络的整体性能,如图 2 所示。

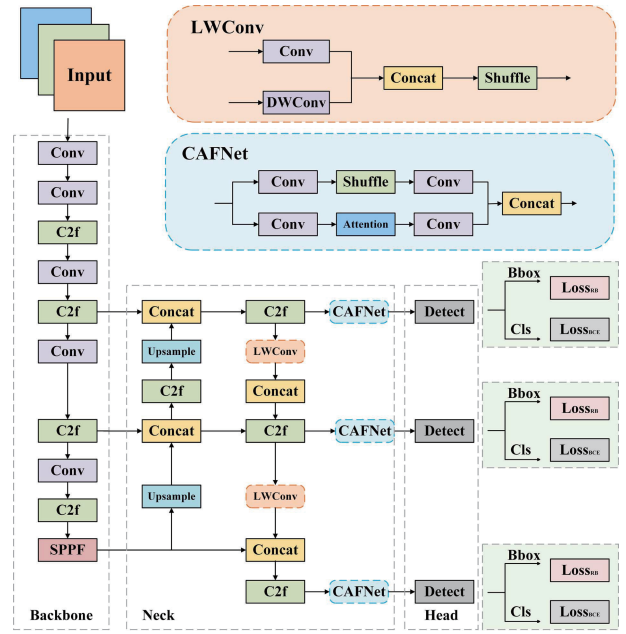


图 1 网络框架

Fig. 1 Network framework

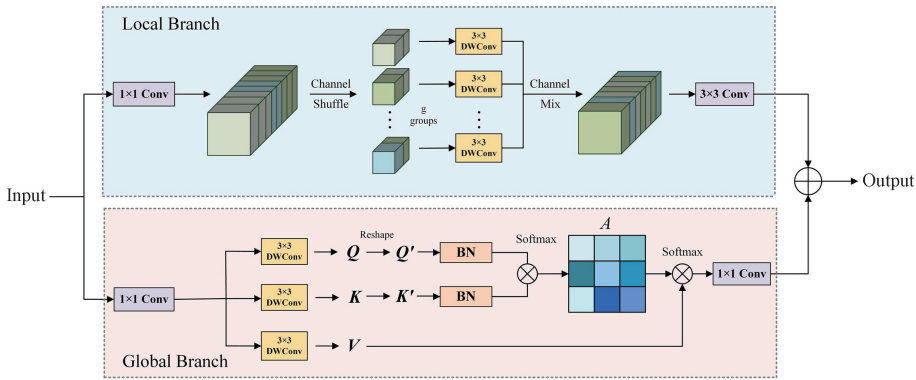


图 2 ACFNet 网络结构

Fig. 2 Network framework of ACFNet

ACFNet 由全局分支和局部分支组成。局部分支利用卷积操作的局部感受野特性,可以高效地捕捉和处理细粒度的特征,负责局部特征提取和增强。输入特征图 X 首先经过 1×1 卷积调整通道尺寸,再对调整后的特征图应用信道洗牌,将输入张量划分为 g 个组,并在每个组内进行深度可分离卷积操作,实现信道信息的混合和重排。最后通过 3×3 卷积提取特征得到局部分支输出 F_{local} ,增强特征表示的多样性和有效性,其计算公式如式(1)所示。

$$F_{local} = \text{Conv}_{3 \times 3}(\text{CS}(\text{DWConv}_{3 \times 3}(\text{Conv}_{1 \times 1}(X)))) \quad (1)$$

卷积操作天然具备局部感受野的特性,这意味着它可以专注于输入图像的局部区域,捕捉到细微的纹理和边缘等局部特征。这些局部特征在图像处理和分析中至

关重要,特别是在船舶检测中,能够帮助模型更好地识别物体的形状和细节。

全局分支中,为了调整空间尺寸,提取到丰富的空间特征,通过 1×1 卷积调整通道尺寸后的输入特征图 $X \in \mathbb{R}^{H \times W \times C}$ 在每个通道上独立进行 3×3 深度可分离卷积操作以得到 Q (query), K (key), V (value) 3 个张量。接下来,将 Q 重构为形状为 $Q' \in \mathbb{R}^{H' \times W' \times C}$ 的张量,并将 K 重构为形状为 $K' \in \mathbb{R}^{C' \times H' \times W'}$ 的张量。重构后的 Q' 和 K' 在 BN 层经过标准化处理后通过 Softmax 操作以得到注意力权重 $A \in \mathbb{R}^{C \times C}$,如式(2)所示。

$$A = \text{BN}(Q') \otimes \text{BN}(K') \quad (2)$$

注意力机制通过计算输入特征图各部分之间的关系,可以捕捉到全局范围内的信息。这与卷积操作的局部感受野形成互补,使得全局分支能够获取图像中的长

距离依赖关系和全局特征。注意力权重矩阵 $\mathbf{A}$ 表示了输入特征图不同位置之间的关系,从而实现全局信息的捕捉。注意力权重与张量 $\mathbf{V}$ 经过 Softmax 和 1×1 卷积进行通道整合得到全局分支输出 $\mathbf{F}_{global}$,其公式如式(3)所示。

$$\mathbf{F}_{global} = \text{Conv}_{1 \times 1}(\mathbf{A} \otimes \mathbf{V}) \quad (3)$$

最终,全局分支的注意力输出和局部分支的输出相加,得到最终的特征图输出 $\mathbf{F}_{output}$,如式(4)所示。

$$\mathbf{F}_{output} = \mathbf{F}_{local} + \mathbf{F}_{global} \quad (4)$$

算法1总结了注意力卷积融合双分支网络输出特征图的过程。

算法1 卷积注意力融合双分支网络 ACFNet 流程

Input: $\mathbf{X}$ to ACFNet
Output: $\mathbf{F}_{output}$
 $\mathbf{F}' = \text{Conv}(\mathbf{X})$
for $c = 1$ to g
 $\mathbf{F}'' = \text{dwConv}(\mathbf{F})$
 $\mathbf{F}_{local} = \text{Conv}(\text{CS}(\mathbf{F}''))$ (局部特征图生成)
end for
 $\mathbf{Q} = \text{reshape}(\text{Conv}(\mathbf{F}'))$
 $\mathbf{K} = \text{reshape}(\text{Conv}(\mathbf{F}'))$
 $\mathbf{V} = \text{Conv}(\mathbf{F})$
 $\mathbf{A} = \text{softmax}(\text{BN}(\mathbf{Q}) \otimes \text{BN}(\mathbf{K}))$ (注意力权重输出)
 $\mathbf{F}_{global} = \text{Conv}(\text{Conv}(\mathbf{A} \otimes \mathbf{V}))$ (全局特征图生成)
 $\mathbf{F}_{output} = \mathbf{F}_{local} + \mathbf{F}_{global}$

通过组合策略融合卷积和注意力机制可以分别利用卷积和深度可分卷积的优势,卷积层的局部特征提取能力与注意力机制的全局信息捕捉能力相结合,使得网络能够更全面地理解图像内容。ACFNet 的设计既能捕获全局信息,又能保留局部细节,确保特征的多样性和网络的计算效率,旨在提升目标检测的性能和鲁棒性。

1.2 旋转边界框损失 (RotatedBboxLoss)

在遥感船舶检测任务中,准确表示船舶的姿态和位置是提高检测性能的关键^[16]。传统的 YOLOv8 模型使用 CIoU 损失函数^[17]来进行边界框回归,但该方法在处理具有复杂几何特征和方向的目标时存在一定的局限性^[18]。RFDNet 网络结合旋转边界框和概率 IoU (probIoU) 损失函数来优化检测性能。该旋转框损失函数通过增加用于处理边界框旋转角度信息的参数,从而捕捉到更多船舶的几何特性和姿态信息。

概率 IoU 用于计算两个旋转边界框间的相似度。为了计算 probIoU,需要先将旋转边界框表示为高斯边界框 (GBBs),并利用高斯分布之间的重叠程度来计算相似度。五参数旋转边界框 (OBB) 参数包括中心坐标 $(rbox_x, rbox_y)$ 、宽度 $rbox_w$ 、高度 $rbox_h$ 和旋转角度 θ ,如图3所示。这种边界框定义法覆盖了更广的旋转范围,能够更好地适应各种旋转情况。

在 probIoU 的计算过程中,首先,通过边界框的宽度、高度和旋转角度计算协方差矩阵的组成元素 a 、 b 、 c 为:

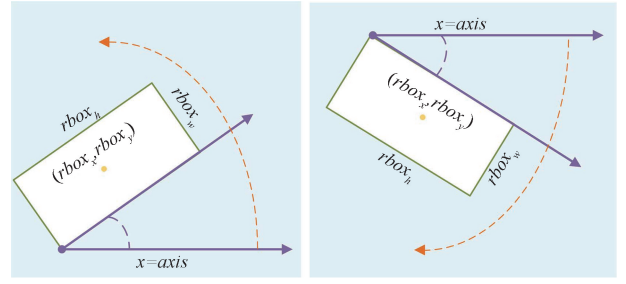


图3 长边定义法旋转边界框标注

Fig. 3 Long edge definition method rotated bounding box labeling

$$a = \frac{rbox_w^2}{12} \cos^2 \theta + \frac{rbox_h^2}{12} \sin^2 \theta \quad (5)$$

$$b = \frac{rbox_w^2}{12} \sin^2 \theta + \frac{rbox_h^2}{12} \cos^2 \theta \quad (6)$$

$$c = \left(\frac{rbox_w^2}{12} - \frac{rbox_h^2}{12} \right) \cos \theta \sin \theta \quad (7)$$

通过预测边界框和目标边界框的宽度、高度和旋转角度信息分别计算出各自的协方差矩阵 Σ_1 和 Σ_2 ,公式如式(8)、(9)所示。

$$\Sigma_1 = \begin{pmatrix} a_1 & c_1 \\ c_1 & b_1 \end{pmatrix} \quad (8)$$

$$\Sigma_2 = \begin{pmatrix} a_2 & c_2 \\ c_2 & b_2 \end{pmatrix} \quad (9)$$

Bhattacharyya 距离是一种度量两个概率分布之间相似性的度量。它用于评估两个分布的重叠程度,常用于模式识别和分类等领域。Bhattacharyya 距离被用来计算预测边界框和真实边界框之间的相似度,距离越小,两个边界框越相似。Bhattacharyya 距离 (bd) 计算公式如式(10)所示。

$$bd = \frac{1}{8} (\mu_1 - \mu_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\mu_1 - \mu_2) + \frac{1}{2} \ln \left(\frac{\det \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)}{\sqrt{\det \Sigma_1 \det \Sigma_2}} \right) \quad (10)$$

式中: Σ_1 和 Σ_2 分别是两个边界框的协方差矩阵; $\det \Sigma$ 表示矩阵 Σ 的行列式; μ_1 和 μ_2 表示预测边界框和目标边界框的均值向量;其计算公式如式(11)所示。

$$\mu_1 = \begin{pmatrix} rbox_{x1} \\ rbox_{y1} \end{pmatrix}, \mu_2 = \begin{pmatrix} rbox_{x2} \\ rbox_{y2} \end{pmatrix} \quad (11)$$

由于 Bhattacharyya 距离的计算涉及二维向量和矩阵,逆矩阵和行列式可以解析计算,计算表示为 t_1 和 t_2 的和。中心点差异 t_1 反映了预测边界框和目标边界框在中心点坐标上的差异,如式(12)所示。

$$t_1 = \frac{1}{4} \frac{(a_{1+} a_2)(rbox_{y1} - rbox_{y2})^2 + (b_{1+} b_2)(rbox_{x1} - rbox_{x2})^2 + 2(c_1 + c_2)(rbox_{x2} - rbox_{x1})(rbox_{y2} - rbox_{y1})}{(a_{1+} a_2)(b_{1+} b_2) - (c_1 + c_2)^2} \quad (12)$$

交叉项 t_2 考虑旋转角度差异的交叉项计算过程如式(13)所示。

$$t_2 = \frac{1}{2} \ln \frac{(a_{1+} a_2)(b_{1+} b_2) - (c_1 + c_2)^2}{4\sqrt{(a_1 b_1 - c_1^2)(a_2 b_2 - c_2^2)}} \quad (13)$$

Hellinger 距离 hd 是基于综合距离 bd 的一个指标距离,其计算公式如式(14)所示。

$$hd = \sqrt{1 - \exp(-bd)} \quad (14)$$

其中, $\exp$ 表示指数函数。

基于 Hellinger 距离 hd 的 probIoU 计算公式如式(15)所示。

$$probIoU = 1 - hd \quad (15)$$

式(15)保证了 probIoU 的值范围在 0~1 之间,其中 1 表示预测边界框和真实边界框完全重叠;0 表示它们完全不重叠。

最终的旋转边界框损失 $Loss_{RB}$ 用于衡量预测边界框与真实边界框之间的重叠程度。公式如式(16)所示。

$$Loss_{RB} = 1 - probIoU \quad (16)$$

在此旋转船舶目标检测的损失函数中, probIoU 在计算两个边界框之间的相似度时,通过 Bhattacharyya 距离和 Hellinger 距离,综合考虑了中心点差异、形状和方向的差异。

相比于大部分使用 CIoU 作为边界框损失函数的检测算法, RFDNet 引入的旋转边界框损失 $Loss_{RB}$ 通过衡量预测边界框与真实边界框的重叠程度,通过增加旋转角度参数,更精确地拟合船舶的实际形状和姿态,捕捉到更多的几何特性,进一步优化模型的回归性能。

1.3 轻量级卷积 LDCnv

在深度学习模型中,深度可分离卷积被广泛应用于各种视觉任务中,大多数方法选择使用深度可分离卷积替换所有的普通标准卷积来搭建网络模型。这种方法虽然可以减少模型参数,但是也会带来模型检测精度的降低。为了平衡模型参数和检测精度的问题,本文引入了一种改进的轻量级卷积方法 LDCnv,该卷积方法结合普通标准卷积、深度可分离卷积以及通道混洗的优点,逐通道对输入特征进行处理。标准卷积层具备强大的特征提取能力,能够有效捕捉复杂特征和局部模式,并且能够压缩通道数,保留重要特征,减少后续层的计算负担。深度可分离卷积在提取特征时更加高效,通过独立处理每个通道并进行逐点卷积,显著减少了参数数量和计算量。这种卷积方式在保留细节的同时降低了计算成本。LDCnv 卷积构成如图 4 所示。

通道数为 c_1 的输入张量通过标准卷积层将通道数减少到 $c_2/2$ 后逐通道输入 3×3 的深度可分离卷积层中。

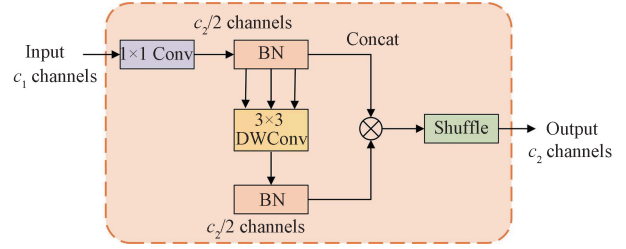


图 4 LDCnv 结构

Fig. 4 The structure diagram of LDCnv

深度可分离卷积包含两个步骤:深度卷积和逐点卷积。在深度卷积的每个通道独立进行 3×3 的卷积操作,参数量仅为 $3 \times 3 \times (c_2/2)$ 。接着使用 1×1 的逐点卷积将每个通道的信息重新组合,参数量为 $(c_2/2) \times (c_2/2)$ 。深度可分离卷积通过分解标准卷积的计算过程,显著减少了参数数量和计算量。具体来说,原始标准卷积的参数量为 $3 \times 3 \times c_1 \times c_2$,而深度可分离卷积的参数量为 $3 \times 3 \times c_1 + c_1 \times c_2$,尤其当 c_1 和 c_2 较大时,参数减少非常显著。

将标准卷积和深度可分离卷积的输出在通道维度上进行拼接,标准卷积和深度可分离卷积各自的特长得以充分发挥,使特征表示更加丰富。通道混洗操作通过重新排列通道,增强不同特征之间的交互。这种操作使得信息在不同通道间得以更好地融合,进一步提升了模型的表达能力和准确性。

使用轻量级卷积方法 LDCnv 来代替原始网络中的标准卷积,其计算成本约为标准卷积的 60%~70%,这种轻量化的卷积方式不仅在保持模型精度的同时显著减少计算量和参数数量,解决了单纯使用深度可分离卷积可能导致的特征提取能力下降的问题,还可以提升计算效率,降低硬件资源的消耗,有效平衡参数数量和检测精度。

2 实验

2.1 实验环境及参数设置

实验使用 Python 编程语言进行基于遥感图像的船舶目标检测任务。

实验过程中使用随机梯度下降(SGD)优化器来对所提出 ADFNet 网络的训练,并且,所有进行的实验中统一将动量设置为 0.937, weight decay 数值均设置为 5×10^{-4} ,批量大小设置为 16, epoch 数值总数设置为 300,具体实验环境如表 1 所示。

2.2 数据集及处理

实验选用 HRSC2016 数据集和 DOTA 数据集进行模型训练和实验验证。在进行模型训练前需将其转换成

YOLO 检测模型适用的标签形式,数据集处理过程如图 5 所示。

表 1 实验平台及环境配置
Table 1 Experimental platform and environment configuration

名称	配置
操作系统	Ubuntu 16. 04. 7
GPU	NVIDIA GeForce RTX 2080ti
内存	12 GB
框架	Pytorch1. 21. 1
CUDA	CUDA10. 2

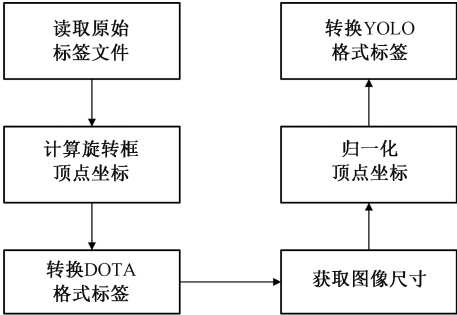


图 5 数据集标签处理流程
Fig. 5 The process of dataset labeling

HRSC2016 数据集是西北工业大学于 2016 年公开的光学遥感船舶目标检测数据集。该数据集图像包含来自 Google Earth 的 1 680 张图像,目标密集且尺度变化较大,船舶目标角度各异。在实验中使用该数据集的 L2 检测任务,即检测航母、军舰、商船、潜艇 4 类目标。

DOTA 数据集由华中科技大学和武汉大学遥感国家重点实验室联合提出,由包含 188 282 个目标实例的 2 806 张卫星图像组成。数据集中包含 15 个类别的目标,分别是:飞机、船舶、储罐、棒球场、网球场、游泳池、田径场、港口、桥梁、大型车辆、小型车辆、直升机、环岛、足球场和篮球场,采用旋转框标注。

由于 HRSC2016 为遥感船舶图像数据集,而同样是遥感图像的 DOTA 数据集中目标种类为包括船舶在内的 15 个类别,因此,本次实验以 HRSC2016 为主要实验数据集,分别进行对比试验和消融实验。DOTA 验证集仅用于对比试验验证。

2.3 评价指标

在船舶检测任务中,平均精度均值 (mean average precision, mAP) 是一项用于评估检测算法准确性的关键指标,它反映了模型在给定数据集上的检测精度。其计算公式如式 (17) 所示。

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

(17)

式中: P 表示准确率; R 表示召回率; n 表示该检测任务中的类别数量。令 IoU 阈值 >0.5 ,即 mAP@ 0. 5;令 IoU 阈值 >0.95 ,即 mAP@ 0. 95。

为全面评估模型性能,使用参数量 (Params) 以及计算量等多个指标来评估所提出方法的综合性能。

2.4 实验结果及分析

1) RFDNet 网络检测结果

改进后的 RFDNet 网络检测 HRSC2016 数据集中航母 (aircraft carrier)、军舰 (waicraft)、商船 (merchant ship) 以及潜艇 (submarine) 4 种类别船舶的各个性能指标值,检测结果如表 2 所示。

表 2 RFDNet 网络在 HRSC2016 数据集上检测结果

Table 2 RFDNet network detection results on HRSC2016 dataset (%)

种类	P	R	mAP@ 0. 5	mAP@ 0. 95
所有	93. 81	96. 54	97. 63	84. 23
航母	98. 22	100. 00	99. 57	93. 88
军舰	91. 13	97. 21	97. 71	88. 92
商船	86. 35	88. 95	93. 92	81. 74
潜艇	99. 64	100. 00	99. 50	72. 79

从表 2 可得,所提出网络 RFDNet 对于 4 种类别的船舶检测的整体平均精度均值为 97. 62%。其中,对航母和潜艇具有最高的检测准确率,mAP@ 0. 5 分别为 99. 57% 和 99. 50%;商船较其他 3 种类别船舶检测准确率较低,mAP@ 0. 5 为 93. 92%,原因是相比于航母、潜艇和军舰,商船外形较小,且细节特征不如其他类型船舶丰富,因此检测中易出现误检的情况。

由以上检测结果可以看出,改进后的检测网络 RFDNet 在 HRSC2016 遥感船舶数据集上具有较高的检测精度。

2) 对比实验及结果分析

为证明改进网络 RFDNet 的先进性,现将其与目前几种主流的目标检测算法 RoI-Transformer、Gliding Vertex^[19]、R³Det^[20]、YOLOv5、YOLOv7^[21]、YOLOv8 以及最新的 YOLOv10 在 HRSC2016 数据集上进行了比较,其中,RoI-Transformer、Gliding Vertex 和 R³Det 为旋转标注框检测方法^[22],YOLOv5、YOLOv7、YOLOv8、YOLOv10 为较先进的水平标注框检测方法,对比实验结果如表 3 和 4 所示。

由表 3 可知改进的基于遥感图像的旋转船舶目标检测网络 RFDNet 在保持较高检测精度的情况下降低了模型的复杂度,在大部分指标上均表现出最佳性能。mAP@ 0. 5 指标比次优方法 YOLOv8 高了 1. 72%,可以看出,相比于原始 YOLOv8 使用的基于水平标注框的算法,RFDNet 使用的基于旋转标注框的算法在船舶检测性能

上有较大幅度的提升。同时在 Parameters 指标上明显优于 RoI-Transformer、Gliding Vertex 和 R³Det 3 个基于旋转标注框的检测方法,仅需要 2.99×10^6 的模型参数,实现了模型的轻量化。相比于最新发布的检测算法 YOLOv10,RFDNet 具有更高的检测精度,但模型参数略多于 YOLOv10 算法,原因是 YOLOv10 算法更专注于端到端的实时目标检测任务,对模型进行轻量化的同时牺牲了一定的检测准确性,导致其平均精度结果值均低于改进后的网络 RFDNet。

表 3 不同目标检测算法在 HRSC2016 数据集下的对比实验

Table 3 Comparative experiments of different target detection algorithms under HRSC2016 dataset				
方法	mAP@ 0. 5/%	mAP@ 0. 95/%	P/%	计算量/GFLOPs
RoI-T	90. 21	77. 25	55. 13	122. 61
Gliding-V	88. 20	67. 28	41. 13	121. 50
R ³ Det	89. 26	76. 44	41. 58	200. 92
YOLOv5	87. 94	66. 35	86. 23	204. 7
YOLOv7	94. 12	80. 51	37. 21	105. 2
YOLOv8	95. 91	82. 23	3. 01	8. 2
YOLOv10	91. 93	75. 48	2. 71	8. 4
本文	97. 63	84. 23	2. 99	9. 2

为进一步验证所提出方法的有效性,在 DOTA 数据集上同样进行了一系列对比试验。分别和 RoI-Transformer、Gliding Vertex、R³Det、YOLOv5、YOLOv7、YOLOv8 以及 YOLOv10 等一系列相关以及最新的检测算法在 DOTA 数据集上进行了对比实验,结果如表 4 所示。

表 4 不同目标检测算法在 DOTA 数据集下的对比实验

Table 4 Comparative experiments of different target detection algorithms under DOTA dataset				
方法	mAP@ 0. 5/%	mAP@ 0. 95/%	P/%	计算量/GFLOPs
RoI-T	69. 56	58. 74	55. 13	122. 61
Gliding-V	75. 02	65. 31	41. 13	121. 50
R ³ Det	73. 79	65. 35	41. 58	200. 92
YOLOv5	73. 17	64. 22	86. 23	204. 7
YOLOv7	72. 54	63. 74	37. 21	105. 2
YOLOv8	79. 78	68. 96	3. 01	8. 2
YOLOv10	77. 41	64. 23	2. 71	8. 4
本文	81. 63	70. 14	2. 99	9. 2

从表 4 可以看出,相比于 RoI-Transformer、Gliding

Vertex、R³Det 3 个旋转检测网络,改进后的 RFDNet 在 mAP@ 0. 5 指标上达到了最高,相比于次高算法 YOLOv8 的检测精度提升了 1. 85%,原因是 RoI-Transformer 更注重基于位置敏感对齐的特征提取,对小目标的检测并不完全适用,而 Gliding Vertex 以及 R3Det 检测方法虽然都引入了方向回归的方法,但遇到目标角度近似水平的情况时,容易出现误差从而导致检测精度的下降,尤其 R³Det 检测方法更是带来了巨大的模型参数,不适用于嵌入式设备中。而 RFDNet 的参数数量达到了 2.99×10^6 ,仅落后于 YOLOv10,但其检测精度远高于 YOLOv10 检测模型。

通过以上对比实验结果可以看出,所提针对遥感图像的船舶检测网络 RFDNet 具有最好的综合检测性能。

3) 消融实验及分析

所提出的网络针对遥感图像中船舶角度各异、目标较小且模型参数过多的问题,分别提出了旋转边界框损失函数、注意力卷积融合双分支网络 ACFNet 以及轻量级卷积 LDConv。

为证实所提出的船舶检测方法的有效性,本节通过 5 组消融实验来检验各改进模块对检测性能的影响。以原始未改进的 YOLOv8 检测模型为原始模型,依次在此原始模型的基础上引入旋转边界框损失函数

$Loss_{RB}$ 、注意力卷积融合双分支网络 ACFNet 以及轻量级卷积 LDConv 来进行消融实验。

训练过程中仍然采用相同的配置参数与实验平台进行消融实验。其中,“√”表示“进行此项改进”,实验结果如表 5 所示。

由表 5 可知,各改进部分均有效果。旋转边界框损失函数的效果显著比原始水平边界框损失函数更好,对 mAP@ 0. 5 和 mAP@ 0. 95 分别有着 0. 52% 和 0. 23% 的提升;在使用旋转边界框损失的情况下,单独引入注意力卷积融合网络 ACFNet 以及轻量级卷积 LDConv 对 mAP@ 0. 5 和 mAP@ 0. 95 分别具有 0. 82%、2. 17% 和 0. 36%、0. 17% 的提升,但加入轻量级卷积 LDConv 后,模型参数量降低 0.1×10^6 ;而在引入 ACFNet 的基础上同时引入 LDConv 时,相较于原始基线网络的增益为 mAP@ 0. 5 提高 2. 14% 且 mAP@ 0. 95 提高 2. 51%,与未加入 ACFNet 的实验组相比,仅有 0.01×10^6 的参数数量增加。

表 5 各改进方法消融实验对比

Table 5 Comparison of ablation experiments by improved methods							
原始网络	改进模块			结果			
	$Loss_{RB}$	ACFNet	LDConv	mAP@ 0. 5/%	mAP@ 0. 95/%	P/%	计算量/GFLOPs
YOLOv8				95. 49	81. 72	3. 01	8. 2
YOLOv8	√			96. 01	81. 95	3. 08	8. 4
YOLOv8	√	√		96. 83	84. 12	3. 59	9. 6
YOLOv8	√		√	96. 37	82. 12	2. 98	8. 3
YOLOv8	√	√	√	97. 63	84. 23	2. 99	9. 2

2.5 检测结果可视化

为了能够更加直观地分析 RFDNet 的改进效果,随机选取数据集中部分图片检测结果进行对比,如图 6 所示,依次为原图、基于水平框标注的基线网络 YOLOv8 检测效以及改进网络 RFDNet 检测效果。

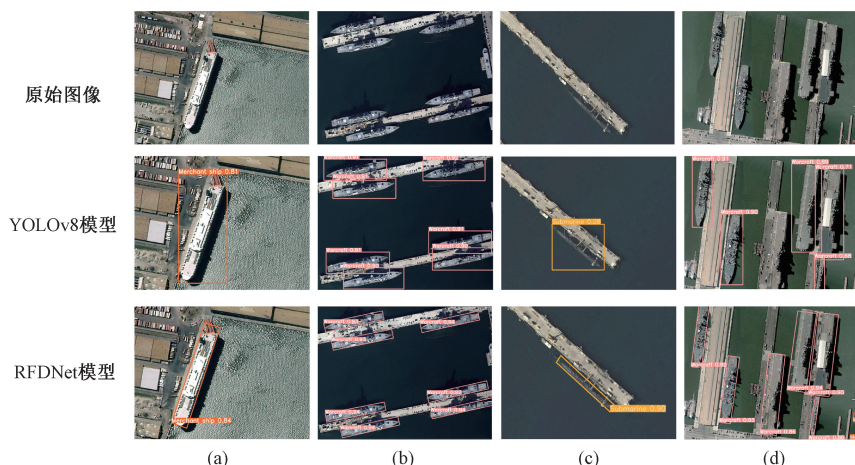


图6 检测结果可视化对比

Fig. 6 Comparative visualization of test results

相比于原始 YOLOv8 检测算法,RFDNet 除了提高了遥感船舶图像中的小目标检测精度外,还检测出了原始基准网络漏检的船舶目标,如图 6(d)所示。尤其在复杂背景下的小目标旋转角度的船舶检测中,RFDNet 的检测精度高于检测算法 YOLOv8 约 62%,如图 6(c)所示。RFDNet 检测网络对不同大小形状以及旋转角度的船舶目标具有更高的检测精度,并且基准网络未检测到的对象也在改进后的方法中检测到了,更适用于遥感图像中的船舶旋转小目标检测。

3 结 论

本文针对遥感图像中船舶目标检测任务面临的姿态变化、船舶目标较小以及参数量大的难题,提出了一种名为 RFDNet 的旋转船舶目标检测网络,旨在提升检测精度的同时实现模型轻量化。RFDNet 结合概率 IoU 提出旋转边界框损失函数,从而有效捕捉船舶的几何特性和姿态,提高模型定位和分类的准确性;通过注意力卷积融合双分支网络 ACFNet 实现卷积和注意力机制的互补,获取更丰富的特征信息;此外,引入轻量级卷积 LDConv 替换标准卷积方法,利用普通标准卷积、深度可分离卷积和通道混洗的优势,在增强特征交互的同时减少计算量。实验结果表明,改进后的 RFDNet 在 HRSC2016 数据集和 DOTA 数据集上的 mAP 分别达到了 97.63% 和 81.63%,具有较优的检测效果,模型参数也减少到 2.99×10^6 。

手工设计的神经网络结构往往需要在性能和效率之间进行权衡,耗时且成本高。未来可以进一步探索神经

网络结构搜索方法在遥感船舶检测任务中的应用,以自动搜索出在特定任务下性能和效率最优的网络结构,既能保证检测精度,又能提高推理速度,为实际应用提供更可靠的解决方案。

参考文献

- [1] WANG W, REN J, SU C, et al. Ship detection in multispectral remote sensing images via saliency analysis[J]. Applied Ocean Research, 2021, 106(5): 102448.
 - [2] CHEN X, XU X, YANG Y, et al. Visual ship tracking via a hybrid kernelized correlation filter and anomaly cleansing framework [J]. Applied Ocean Research, 2021, 106: 102455.
 - [3] HUANG X, YANG W, ZHANG H, et al. Automatic ship detection in SAR images using multi-scale heterogeneities and an a contrario decision[J]. Remote Sensing, 2015, 7(6): 7695-7711.
 - [4] LENG X, JI K, ZHOU S, et al. Ship detection based on complex signal kurtosis in single-channel SAR imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(9): 6447-6461.
 - [5] 郑自立,徐健,刘秀平,等.联合多注意力和 C-ASPP 的单目 3D 目标检测[J].电子测量与仪器学报,2023,37(8):241-248.
- ZHENG Z L, XU J, LIU X P, et al. Monocular 3D object detection combining multi-attention and C-ASPP [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8): 241-248.

- [6] 孙伟,徐凡,张小瑞,等. 多视图融合和全局特征增强的车辆重识别网络[J]. 电子测量与仪器学报, 2023, 37(1):78-86.
- SUN W, XU F, ZHANG X R, et al. Vehicle re-identification network with multi-view fusion and global feature enhancement [J]. Journal of Electronic Measurement and Instrumentation, 2023,37(1):78-86.
- [7] WANG Y, WANG C, ZHANG H. Combining a single shot multibox detector with transfer learning for ship detection using sentinel-1 SAR images [J]. Remote Sensing Letters, 2018, 9(8): 780-788.
- [8] ZHANG S, HE G, CHEN H B, et al. Scale adaptive proposal network for object detection in remote sensing images [J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(6): 864-868.
- [9] ZHAO J, GUO W, ZHANG Z, et al. A coupled convolutional neural network for small and densely clustered ship detection in SAR images [J]. Science China Information Sciences, 2018, 62(4): 351-362.
- [10] ZHANG S, WU R, XU K, et al. R-CNN-based ship detection from high resolution remote sensing imagery [J]. Remote Sensing, 2019, 11(6): 631.
- [11] 童小钟,魏俊宇,苏绍璟,等. 融合注意力和多尺度特征的典型水面小目标检测[J]. 仪器仪表学报, 2023, 44(1):212-222.
- TONG X ZH, WEI J Y, SU SH J, et al. Detection of typical small surface water targets with attention and multi-scale feature fusion [J]. Chinese Journal of Scientific Instrument, 2023, 44(1): 212-222.
- [12] 张珂,余鲲鹏,石怀涛. 改进 YOLOv5 的绝缘子旋转目标检测[J]. 电子测量技术, 2023, 46(15):52-58.
- ZHANG K, YU K F, SHI H T. Rotating insulator object detection based on improved YOLOv5 [J]. Electronic Measurement Technology, 2023, 46(15): 52-58.
- [13] FIRDIANTIKA I M, KIM S. Is-YOLO: A YOLOv7-based detection method for small ship detection in infrared images with heterogeneous backgrounds [J]. International Journal of Control, Automation and Systems, 2024, 22(11): 3295-3302.
- [14] HUANG Y, HAN D, HAN B, et al. ADV-YOLO: Improved SAR ship detection model based on YOLOv8[J]. The Journal of Supercomputing, 2024, 81(1): 34-34.
- [15] DING J, XUE N, LONG Y, et al. Learning RoI transformer for oriented object detection in aerial images [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 2849-2858.
- [16] 单慧琳,吕宗奎,付相为,等. 改进 YOLOv5s 的交通多目标检测方法[J]. 国外电子测量技术, 2023, 42(4): 8-15.
- SHAN H L, LYU ZH K, FU X W, et al. Traffic multi-object detection method based on improved YOLOv5s [J]. Foreign Electronic Measurement Technology, 2023, 42(4): 8-15.
- [17] 王明阳,王江涛,刘琛. 基于关键点的遥感图像旋转目标检测 [J]. 电子测量与仪器学报, 2021, 35(6): 102-108.
- WANG M Y, WANG J T, LIU CH. Keypoint-based rotated object detection in remote sensing images [J]. Journal of Electronic Measurement and Instrumentation, 2021,35(6):102-108.
- [18] TAN M, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 10778-10787.
- [19] XU Y, FU M, WANG Q, et al. Gliding vertex on the horizontal bounding box for multi-oriented object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(4): 1452-1459.
- [20] YANG X, YAN J, FENG Z, et al. R3Det: Refined single-stage detector with feature refinement for rotating object [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(4): 3163-3171.
- [21] 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 [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [22] FUPEI W U, JIAYE L, SHENGPING L I. Fast image segmentation algorithm based on salient features model and spatial-frequency domain adaptive kernel [J]. Instrumentation, 2022(2): 33-46.

作者简介



孙伟(通信作者),2004 和 2006 年分别于河南科技大学获得学士和硕士学位,2010 年于东南大学获得博士学位,2014 年至 2015 年,任美国普渡大学 NEXTRANS 中心博士后研究员,现为南京信息工程大学教授,主要研究方向为车辆再识别、计算机视觉、深度学习、智能车辆环境感知等。

E-mail:sunw0125@163.com

Wei Sun (Corresponding author) received his B. Sc. and M. Sc. degrees from Henan University of Science and Technology in 2004 and 2006, and Ph. D. degree from Southeast University in 2010, respectively. From 2014 to 2015, he was a post-doctoral

researcher with the NEXTRANS Center, Purdue University, USA. He is currently a professor at Nanjing University of Information Science and Technology. His main research interests include vehicle re-identification, computer vision, deep learning, and environment perception for intelligent vehicles.



沈欣怡, 2022 年于安庆师范大学获学士学位, 现为南京信息工程大学硕士研究生, 主要研究方向为船舶目标检测、计算机视觉和深度学习。

E-mail: shenxinyi20@163.com

Shen Xinyi received her B. Sc. degrees from Anqing Normal University in 2022. She is a M. Sc. candidate in Nanjing University of Information Science and Technology. Her main research interests include ship object detection, computer vision, and deep learning.



张小瑞, 2004 和 2006 年分别于河南科技大学获得学士和硕士学位, 2010 年于东南大学获得博士学位, 2013 年至 2014 年, 在美国宾夕法尼亚州费城宾夕法尼亚大学 ViDi 中心担任博士后研究员, 现任南京工业大学教授, 主要研究方向为虚拟现实与人机

交互、触觉感知、模式识别。

E-mail: zxr365@126.com

Zhang Xiaorui received her B. Sc. and M. Sc. degrees from Henan University of Science and Technology in 2004 and 2006, respectively, and the Ph. D. degree from Southeast University in 2010. From 2013 to 2014, she was a post-doc toral researcher with the ViDi Center, University of Pennsylvania, Philadelphia, PA, USA. She is currently a professor at Nanjing Tech University. Her main research interests include virtual reality and human-computer interaction, haptic perception, and pattern recognition.



管菲, 2022 年于常熟理工学院获学士学位, 现为南京信息工程大学硕士研究生, 主要研究方向为船舶重识别、计算机视觉和深度学习。

E-mail: 202212490112@nuist.edu.cn

Guan Fei received her B. Sc. degrees from Changshu Institute of Technology in 2022. She is a M. Sc. candidate in Nanjing University of Information Science and Technology. Her main research interests include ship re-identification, computer vision, and deep learning.