

DOI:10.16136/j.joel.2023.03.0147

基于细节增强的级联多分类光电船舶检测

徐志京*, 谢安东

(上海海事大学 信息工程学院, 上海 201306)

摘要:为提升无人船航行环境中船舶目标的感知精度,提出一种基于细节增强的级联多分类船舶检测模型。首先,提出混合平移数据增强(pan-mixed data augmentation, PMDA)算法,减少模型对船舶整体轮廓的依赖;其次,设计可变卷积平衡特征金字塔(deformable convolution-balanced feature pyramid, DC-BFP),提高模型对船体细节特征的提取能力;再次,将全连接层和卷积层联合构成级联交叉检测器(cascaded cross detector, CCD),提高模型对船体细节特征的解析能力;最后,采用标签平滑正则化(label smoothing regularization, LSR)方法,改善多分类检测的过拟合问题。在自建 11 分类光电船舶检测数据集 MCSD11 上进行消融和对比实验,特征提取结果和实验结果数据可视化表明,模型的各个改进部分能够提升船舶检测效果,平均精度达到了 91.53%,相比主流的检测模型,算法得到大幅提升。

关键词:光电船舶检测;可变卷积平衡特征金字塔(DC-BFP);MCSD11 数据集;细节增强;混合平移数据增强算法(PMDA)

中图分类号:TP391.4 文献标识码:A 文章编号:1005-0086(2023)03-0241-09

Cascaded multi-classification photoelectric ship detection based on detail enhancement

XU Zhijing*, XIE Andong

(College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China)

Abstract:In order to improve the perception accuracy of unmanned ships while sailing, a cascade multi-class ship detection model based on detail enhancement is proposed. First, a pan-mixed data augmentation (PMDA) algorithm is proposed to reduce the network's dependence on the overall outline of the ship. Secondly, a deformable convolution-balanced feature pyramid (DC-BFP) is designed to improve the model's ability to extract detail features of the hull. Thirdly, combine the fully connected layer and the convolutional layer to form a cascaded cross detector (CCD) to improve the model's ability to analyze the detail features of the hull. Finally, the label smoothing regularization (LSR) is used to improve the overfitting problem of multi-class detection. Ablation and comparison experiments are carried out on the self-built 11-category photoelectric ship detection dataset MCSD11. The feature extraction results and the data visualization of the experimental results show that the improved parts of the model can improve the ship detection effect, with an average accuracy of 91.53%, which is higher than that of the mainstream ones. The detection model algorithm has been greatly improved.

Key words:photoelectric ship detection; deformable convolution-balanced feature pyramid (DC-BFP); MCSD11 dataset; detail enhancement; pan-mixed data augmentation (PMDA)

0 引 言

随着航运业和人工智能技术的蓬勃发展,无

人船态势感知的研究已成为航运领域的研究热点,而船舶感知精度的提升是该领域的一大难题。船舶自动识别系统(automatic identification sys-

* E-mail: zjxu@shmtu.edu.cn

收稿日期:2022-03-10 修订日期:2022-05-12

基金项目:国家重点研发计划(2019YFB1600605)和上海市扬帆计划项目(20YF1416700)资助项目

tem, AIS)和自动雷达标绘仪(automatic radar plotting aid, ARPA)是目前使用较为广泛的船舶驾驶辅助设备,但这些设备的使用在无人船感知中存在一定的局限。例如:AIS只能获取已安装AIS应答器的船舶信息;导航雷达探测优势是感知距离远,但船舶周围会存在一定距离的感知盲区。近年来,随着机器视觉的飞速发展,越来越多的研究者采用光电船舶检测来替代人工识别,获取更为直接明了的船舶辨识与分类结果,解决雷达探测盲区的难题。精确辨别周边船舶的类型对于实施有效避碰策略至关重要,然而很多船舶往往存在同种类船舶的外形差异细微,而不同类船舶外形差异较大的问题,这给光电船舶精准检测带来了巨大的挑战。

近年来,KRIZHEVSKY^[1]使用深度卷积神经网络赢得ImageNet竞赛,使深度学习得到高速发展,船舶检测技术也借助深度学习迎来了大变革。基于深度学习的船舶检测技术^[2,3]不断在准确率和速度上提高的同时,应用场景逐渐增多。HUO等^[4]针对光电图像中舰船分类检测困难的问题,提出了基于改进循环注意卷积神经网络的舰船目标识别方法。ZHOU等^[5]提出了一种深度注意力机制,用于显著性船舶检测。ZHAO等^[6]将混合注意力机制引入YOLOv4中,满足了海事领域对实时监管的要求。

在多分类船舶检测领域,SHAO等^[7,8]等提出了6分类船舶数据集SeaShips,将基于显著性感知的卷积神经网络应用在船舶检测中。GUPTA等^[9]提出了一种基于支持向量机、特征包和卷积神经网络的船舶分类方法。LI等^[10]对YOLOv3进行改进,在SeaShips数据集上获得了96.6%的均值平均精度。SeaShips数据集只提供了6种常见河运类型船舶的检测样本,这没有考虑到无人船航行环境中远多于6类船舶的实际场景,随着船舶种类的增多,类内差异的精准检测矛盾难以化解,因此迫切需要解决针对多分类船舶目标的精准检测算法。

同时,HAN等^[11]重新设计了感受野放大模块,提出了ShipYOLO,相比于YOLOv4在速度和准确度上都有提升。ZHOU等^[12]对YOLOv5s算法进行了改进,将其应用于无人船视觉。HU等^[13]针对小型舰船目标,提出了PAG-YOLO。虽然,这些方法在检测速度上有一定的提升,但在处理目标时存在比例变化和视点变化时效果欠佳的

问题,船舶种类的多样性和复杂性更会大大降低网络的普适性。

船舶检测不同于通用目标检测,船舶一般都拥有相似的整体外形轮廓,仅仅从整体轮廓对船舶进行多分类检测,是极其困难的。因此,船舶检测的难点也主要聚集在船体细节特征模糊和整体轮廓差异大的中小型船舶检测上,只有聚焦于船舶的船体细节特征^[14,15],才能准确地区分不同类别的船舶。因此本文建立了一个包含11分类的常见船舶检测数据集,为多分类船舶检测提供数据样本支撑。同时根据数据集的特点,设计了一种基于细节增强的级联多分类光电船舶检测模型。

相较于AIS和ARPA,光电船舶检测对于提升无人船感知精度也有着重要意义。但在实际中,少量分类检测的实用性差,覆盖常见类型的多分类船舶检测更具有实际意义。本文模型在经典的二阶段目标检测框架上,提出了混合平移数据增强算法(pan-mixed data augmentation, PMDA),通过对船舶样本的平移和融合,减少网络对船舶整体轮廓的特征依赖。设计了可变卷积平衡特征金字塔(deformable convolution-balanced feature pyramid, DC-BFP),强化了模型的船体细节提取能力。结合全连接层和卷积层,从分类和回归两个方面同时改进,构成了级联交叉检测器(cascaded cross detector, CCD)。使用标签平滑正则化方法(label smoothing regularization, LSR)^[16],将硬标签替换为软标签,改善检测中的多分类过拟合问题。

1 模型设计

本文设计的基于细节增强的级联多分类光电船舶检测模型如图1所示。该模型由4个部分组成:PMDA、DC-BFP、CCD、LSR。首先,针对船舶外形特点,设计了PMDA完成数据增强,对于输入网络的训练样本进行平移混合操作,减少网络对船舶整体轮廓的依赖。其次,将混合增强后的样本数据送入DC-BPN中,利用可变形卷积的特点提取不规则船体特征,再通过特征交叉精炼方法,对交叉聚合的特征,使用非局部操作增强船体细节特征。然后,通过CCD,对提取的船舶特征使用全连接层和卷积层交叉的方式,分别完成分类和回归任务。最后,利用LSR将硬标签替换为软标签,改善多分类任务的过拟合现象。

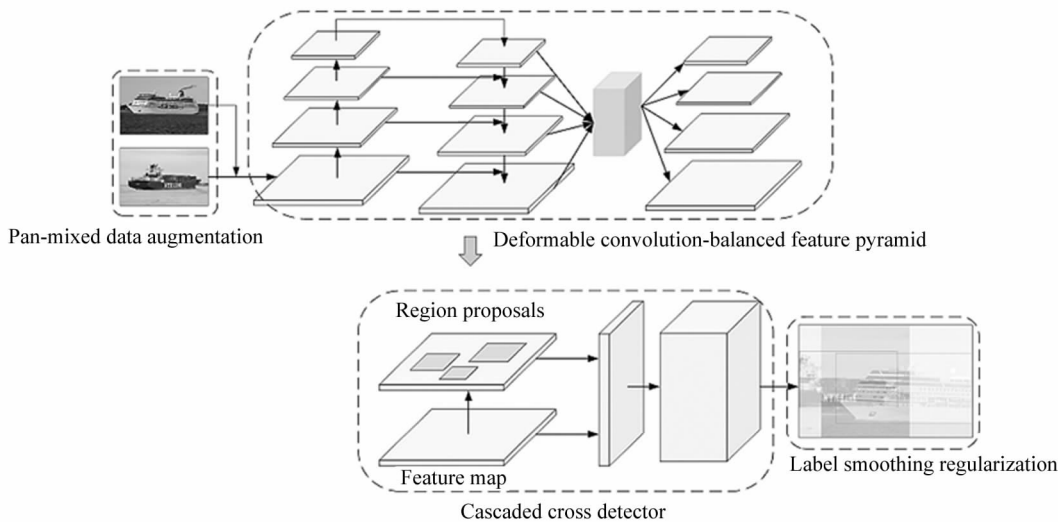


图 1 基于船体细节增强的级联多分类光电船舶检测模型

Fig. 1 Multi-class photoelectric ship detection model based on hull detail enhancement

1.1 PMDA

船舶检测不同于其他检测任务,船舶整体轮廓特征难以为中小型船舶检测提供有力的依据,而船体细节才是检测不同种类船舶的根本依据。深度卷积神经网络的训练过程往往需要大量的数据集才能取得理想的检测效果,但针对船体细节获取数据样本是困难的。这导致能够提供用于训练网络提取船体细节能力的数据量是有限的,数据增强算法能够通过提高样本的多样性解决这个问题。

船舶根据其类别不同在外形上有不同的细节差异,如:1) 航空母舰在右舷处有一个小而集中的岛形上层建筑,有一个供飞机起降的宽阔甲板;2) 集装箱船的最大特点是它所装载的标准规格集装箱,以及具有水密舷边舱的双舷侧;3) 驱逐舰普遍装备舰空导弹、舰舰导弹、反潜导弹、鱼雷等武器系统;4) 干货船的驾驶室和机舱一般布置在艉部,拥有宽大的货舱口,以向上倾斜的边板连接内底板与舷侧;5) 多数捕鱼船只的船体较小,配备作业时使用的捕捞、保鲜和加工设备;6) 拖船拥有较短的船体,配备拖钩等拖拽设备;7) 客船全船具有多层甲板的上层建筑,用于布置旅客舱室;8) 油船上层建筑和机舱一般布置在艉部,甲板纵中部位,全船纵通输油管及步桥;9) 帆船拥有不同种类的帆(主帆、前帆、球帆),以及具有明显的视觉特征的桅杆、横杆等。

本文根据船舶的船体细节特征,综合船舶检测任务中的难点,提出 PMDA。像素点的混合,能够创造出大量实际中很难获取的船舶重叠的场景。PMDA 在图像像素混合前,根据数据样本标注的信息,

针对不同目标的尺度对图像定量平移,同时生成新的样本数据标注文件。这样可以减少网络对船舶整体轮廓的依赖,更多聚焦于不同类别的船体细节特征。PMDA 原理如式(1)–(4)所示:

$$\lambda = \text{Beta}(\alpha, \beta), \tag{1}$$

$$\begin{aligned} \text{Scale} = & \\ & \begin{cases} [O_{xmin} + \lambda O_{width} : batch_{xmax}] , \lambda \leq 0.5 \\ [batch_{xmin} : O_{xmax} - \lambda O_{width}] , \lambda > 0.5 \end{cases} \end{aligned} \tag{2}$$

通过样本目标的尺度和位置信息,根据贝塔分布决定整体平移方向和距离,得到平移结果 *Scale*。其中 *Beta* 指的是贝塔分布, *O_{xmin}* 和 *O_{xmax}* 表示船舶目标的位置, *O_{width}* 则表示船舶目标的尺度大小, *batch_{xmin}* 和 *batch_{xmax}* 表示对应数据样本的尺寸大小。

$$\begin{aligned} mixed_batch_x = & \lambda \cdot batch_{x_1} [Scale] + \\ & (1 - \lambda) \cdot batch_{x_2} [Scale], \end{aligned} \tag{3}$$

$$mixed_batch_y = batch_{y_1} + batch_{y_2}, \tag{4}$$

式中, *mixed_batch_x* 是混合后的 *batch* 样本, *mixed_batch_y* 是混合后的 *batch* 样本对应的标签。PMDA 效果如图 2 所示,船舶样本经过增强后能创



图 2 混合平移数据增强算法效果展示

Fig. 2 Display of pan-mixed data augmentation effect

造出船体细节重叠的场景,为训练模型的船体细节特征提取能力提供数据支撑。

1.2 DC-BFP

船体细节特征提取是多分类船舶检测的关键,而船体细节的形状不规则和尺度差异大是特征提取的难点。DC-BFP 是将可变形卷积网络(deformable convolution network, DCN)结合平衡特征金字塔

(balanced feature pyramid, BFP)的增强型特征金字塔网络(feature pyramid network, FPN)。其中 DCN 通过灵活改变卷积核的形状和尺度,能够应对船体细节形状不规则的问题。BFP 区别于传统 FPN,使用融合强化的方法,融合特征尺度差异较大的非相邻层,加强模型提取不同尺度船体细节特征的能力。

DC-BFP 输出特征图的过程如图 3 所示,需要通

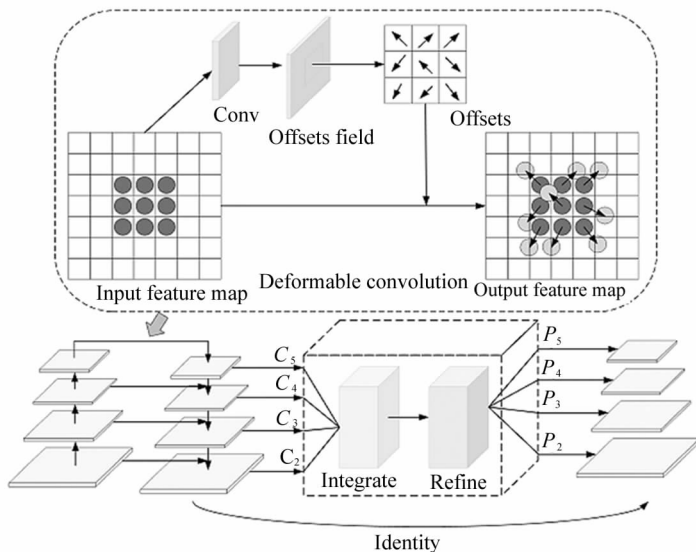


图 3 可变卷积平衡特征金字塔

Fig. 3 Deformable convolution-balanced feature pyramid

过提取、融合和强化 3 个步骤完成。

1.2.1 提取

标准卷积计算得到的同级特征层具有同样的感受野,对于将位置信息进行编码的浅层卷积并不可取。由于船舶结构的复杂性,具有尺度或形态差异的船体特征会存在于图像的不同位置,能够自动调整感受野大小或者形状的方法能更好应对这种情况。DCN 在感受野中通过学习偏移量,使感受野与船舶实际形状贴近^[17]。引入 DCN,卷积核形状大小会根据船舶空间环境自适应调整。

具体地,计算输入图像的每个像素点的二维偏移量,创建可变形采样点位置,将具有相似特征的生成采样点覆盖其相邻像素的位置。其次利用生成采样点,将具有相似结构信息的相邻像素压缩成固定的网格,生成可变形特征矩阵。假设标准卷积是使用一个 3×3 大小,膨胀系数为 1 的标准卷积核 $\mathbf{R}$ 进行计算,其表达式如式(5)–(6)所示:

$$y(p_0) = \sum_{p_n \in \mathbf{R}} w(p_n) \cdot x(p_0 + p_n), \quad (5)$$

$$\mathbf{R} = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}, \quad (6)$$

式中, p_0 表示特征图上的每个位置, p_n 是对 $\mathbf{R}$ 中所

列位置的枚举。对 $\mathbf{R}$ 进行可变形卷积运算,就是将每个点都增加一个偏移参数 Δp_n ,可变形卷积表达式如式(7)所示:

$$y(p_0) = \sum_{p_n \in \mathbf{R}} w(p_n) \cdot x(p_0 + p_n + \Delta p_n). \quad (7)$$

1.2.2 融合

为了将各层级统一到特征层 C_4 的大小,均衡不同特征层的语义特征,分别对特征层 C_2 、 C_3 进行渐进式上采样操作,对特征层 C_5 进行差值操作,完成特征图尺度的统一。再通过累计平均的方式,融合同尺度的特征图,不同层级特征融合如式(8)所示:

$$C = \frac{1}{L} \sum_{l_{\min}}^{l_{\max}} C_l, \quad (8)$$

式中, C_l 表示不同层级的特征图, L 表示融合的特征图层总数, $l_{\max}$ 和 $l_{\min}$ 分别表示最高和最低层级的层级数。

1.2.3 强化

局部的卷积层堆叠可以增大感受野,但是某一特征层上的卷积核映射到原图上的感受野是有限的。为了获取原图上的更多信息,得到更具有鉴别

性的语义特征,需要引入全局的信息。这里使用非局部操作,该操作会计算每一像素点与特征层内其他像素点的相似性,然后对相似性加权求和来表示对应像素点的特征信息,这样能通过全局特征达到增强局部特征的效果。其表达式如式(9)所示:

$$y_i = \frac{1}{C(x)} \sum_{\forall j} f(x_i, x_j) g(x_j), \quad (9)$$

式中, x , y 分别表示具有相同维度的输入特征和输出特征, i, j 分别表示输入和输出的所有位置索引, $f(x_i, x_j)$ 函数用于计算 x_i 和 x_j 之间的位置依赖关系, $g(x_j)$ 函数计算输入特征在位置 j 的特征表示, $C(x)$ 为归一化参数。最终,将精炼之后的特征图与对应层级的原始特征图相加,得到 DC-BFP 结构输出 $\{P_2, P_3, P_4, P_5\}$ 的特征图。

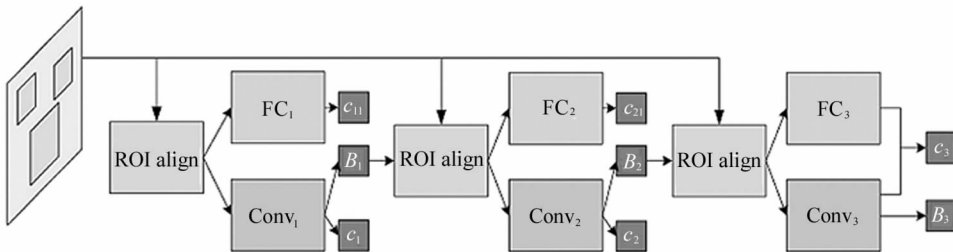


图4 级联交叉检测器

Fig. 4 Cascade cross detector

1.4 LSR

实际船舶检测数据集中,有些样本标注信息会因为 AIS 信息缺失而出现标注错误的情况,LSR 作为深度神经网络的正则化方法,可有效地减轻此类问题对模型的影响。同时 LSR 也能改善多分类检测任务的过拟合问题,进一步提高网络分类性能。

对于含 K 类目标的数据集 $D_{train} = \{(x_i, y_i)\}$ 中的某个样本 (x_i, y_i) ,检测网络使用 Softmax 函数输出每一类别的概率 $p(k | x_i)$ 。硬标签 y_i 的分布 q 如式(10)所示:

$$\begin{cases} q(k = y_i | x_i) = 1 \\ q(k \neq y_i | x_i) = 0 \end{cases}, \quad (10)$$

式中, x_i 表示网络中的输入图像, y_i 表示输入图像对应的真实标签。那么使用硬标签进行图像分类的标准交叉熵损失如式(11)所示:

$$L_{hard} = - \sum_{k=1}^K q(k | x_i) \log p(k | x_i) = - \log p(k = y_i | x_i). \quad (11)$$

LSR 通过使用均匀分布方法生成的软标签,代

1.3 CCD

CCD 使用全连接层和卷积层级联交叉的结构,完成船舶检测的最终任务。相较于二阶段检测模型只使用全连接层完成分类回归任务,而卷积层在输入特征图的所有位置上使用共享转换,对位置信息比全连接层更敏感^[18]。CCD 通过将两种结构获取的信息互补,提升检测器的整体性能。CCD 的整体流程如图4所示。

图4中, FC_1, FC_2 和 FC_3 代表全连接层, $Conv_1, Conv_2$ 和 $Conv_3$ 代表卷积层, B_i 和 c_i 代表不同阶段的回归和预测结果, ROI Align 代表池化操作。同时, CCD 器分为3个阶段,前一阶段输出的偏移量和 ROI 池化结果作为下一阶段的输入,且每个阶段的 IoU 阈值不断上升,能为下个阶段提供更好的正样本,使检测器效果逐渐变好。

替硬标签进行模型训练。具体而言,软标签中出现 K 类的概率如式(12)所示:

$$q'(k | x_i) = (1 - \epsilon) q(k | x_i) + \frac{\epsilon}{K}, \quad (12)$$

式中, ϵ 表示标签平滑因子, $q'(k | x_i)$ 为标签平滑操作后的样本标签。标签平滑的过程就相当于将噪声加入到真实分布中,减小正负样本预测的输出差值,从而避免模型过于依赖正确标签,提高模型的泛化能力。

2 实验结果与分析

2.1 光电船舶检测数据集 MCS11

目前,流行的目标检测数据集是 VOC 和 COCO,这些数据集仅将船舶分类为一类,但在特定应用中,通常需要更精确地对船舶进行分类。因此,为了展现多分类船舶检测的有效性,本文在自行采集和标注的基础上,从互联网上获取了大量船舶图像,构建了一个多分类的光电船舶检测数据集 MCS11,数据集部分样本展示如图5所示。

数据集共包含 11 795 张高清船舶图片,涵盖 11

种常见船舶类型,各类型的占比如图 6 所示,由于军用舰船数据公开较少,航空母舰和驱逐舰占比分别

为 1.78% 和 2.60% 外,其他类型船舶的占比皆在 10% 左右。

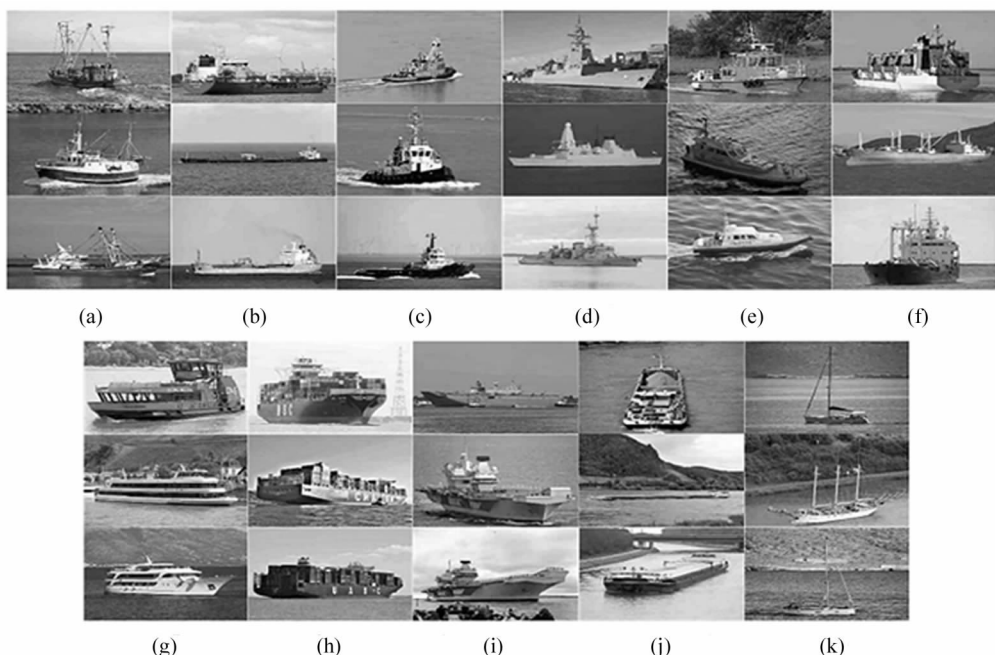


图 5 MCSD11 数据集展示:(a) 渔船;(b) 油船;(c) 推轮;(d) 驱逐舰;(e) 领航船;
(f) 冷藏船;(g) 客轮;(h) 集装箱船;(i) 航空母舰;(j) 干货船;(k) 帆船
Fig. 5 MCSD11 data set display:(a) FV;(b) OT;(c) PT;(d) DE;(e) PB;(f) RE;
(g) PS;(h) CS;(i) AC;(j) DC;(k) SS

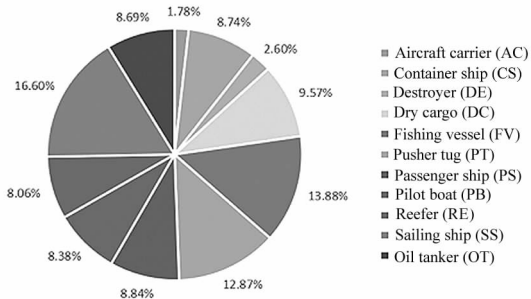


图 6 MCSD11 各类别船舶占比
Fig. 6 Proportion of ships of various categories in MCSD11

由于图像是从互联网上找到的,因此像素分辨率不同,图像大小也不同,例如 $2560 \times 1500 \times 3$ 和 $1920 \times 1080 \times 3$ 。在每幅图像中,船舶比例与整个图像的比例也不同,甚至非常不同,这可以从图 5 中的数据图像看到。这能提升船舶检测网络的鲁棒性,所适用的检测场景更广,实用性更强。相比于特定场景获取的图片,MCSD11 数据集的数据来源地包含港口、湖泊、运河等地点,拥有早晨、傍晚、夜晚等不同时间段,也考虑到船只拍摄远近、重叠等特殊情况,自然生成的不同规格和质量图像更有利于训练效果和泛化能力,所训练的网络更具有实际应

用价值。特别的,MCSD11 同时拥有图像中对应目标的 AIS、IMO、MMSI 等船舶身份信息,船舶长宽比例、最大航速、所属国、自重等船舶外形特征,部分样本附有拍摄地点、日期以及拍摄工具等图片信息。相较于依靠经验的人工判断,利用这些信息能够极大提升数据集的准确性。

2.2 实验环境配置

本实验的实验环境配置如下:操作系统是 Ubuntu18.04,CPU 是 I7-7700x,GPU 是 NVIDIA GTX1080Ti,内存为 32 G,所有实验均是在 Pytorch 框架下进行的。3 层 CCD 的 IoU 分别设置为 0.5、0.6 和 0.7,使用 SGD 优化器,学习率为 0.02,momentum 为 0.9,权重衰减为 0.0001。

本文采用的评价指标是平均精度均值(mean average precision, mAP),即各类别 AP 的平均值, mAP 越高代表网络检测性能更好。 mAP 式如(13)所示:

$$mAP = \frac{\sum_{c=1}^C \frac{TP + TN}{TP + TN + FP}}{C}, \quad (13)$$

式中, TP 是检出正样本的概率, FP 是检出错误样本

的概率, TN 是检出负样本的概率。

2.3 特征提取结果可视化

本节直观展示了使用 ResNet50 网络作为主干网络的 FPN 和 DC-BFP 的特征可视化。其中,使用 ResNet50 输出的特征图分辨率依次为原始图片的 $1/4$ 、 $1/8$ 、 $1/16$ 、 $1/32$, 通道数依次为 256、512、1024、2048,使用 FPN 和 DC-BFP 得到的特征图 C_i 。

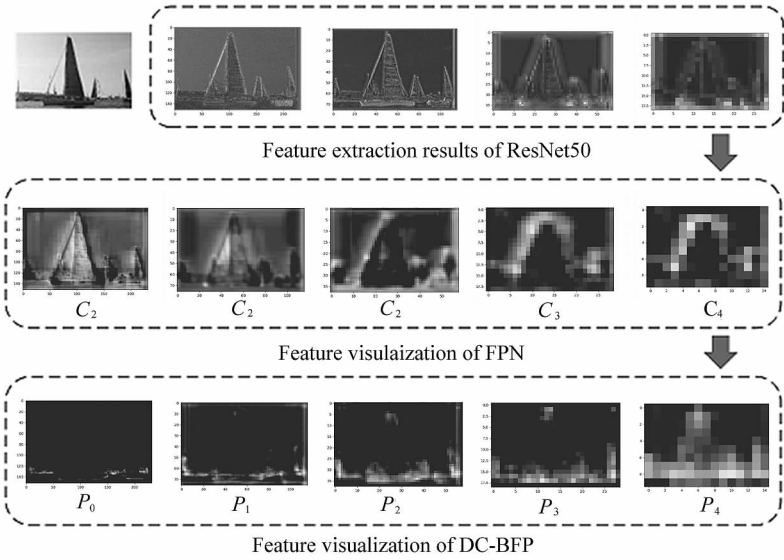


图 7 特征提取结果
Fig. 7 Feature extraction results

图 8 展示了同一张图片分别使用 FPN 和 DC-BFP 的检测结果,两种方法的回归效果基本一致,但是对船舶具体类别判定的性能存在差距(左 FPN: 0.33、0.31,右 DC-BFP:0.93、0.81)。帆船作为典型的中小型船舶,拥有该类型船舶普遍缺乏船体细节信息的特点,DC-BFP 相较于 FPN 具有更强的船体细节提取能力,能更好地完成中小型船舶的检测任务。

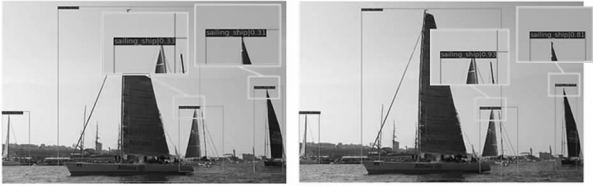


图 8 FPN 和 DC-BFP 检测结果对比
Fig. 8 Comparison of FPN and DC-BFP test results

2.4 消融实验

为了证明 PMDA、LSR、DC-BFP 和 CCD 对于多分类船舶检测任务的有效性,以及各个部分对整体提升的贡献度,本文以使用 ResNet50 特征提取网络

和特征金字塔的 Faster RCNN 为作为基准线,在 MCSD11 船舶数据集上进行对比实验。实验结果如表 1 所示。

通过图 7 的特征提取结果可以直观地看出,卷积神经网络提取边缘等低级特征,高层捕获语义特征。相较于 FPN,每层更注重相邻层的特征融合,DC-BFP 对 FPN 操作输出的所有特征层进行了分局部操作,融合各个距离的特征层,强化了更多有助于分类的船体细节特征信息。

和特征金字塔的 Faster RCNN 为作为基准线,在 MCSD11 船舶数据集上进行对比实验。实验结果如表 1 所示。

表 1 消融实验对比表					
Tab. 1 Comparison table of ablation experiments					
Baseline	PMDA	LSR	DC-BFP	CCD	$mAP/\%$
✓					88.10
✓	✓				89.15
✓		✓			88.86
✓			✓		90.02
✓				✓	89.54
✓	✓	✓	✓	✓	91.53

由表 1 可以看出,改进部分在多分类船舶检测中分别拥有 1.05、0.76、1.92、1.34 个百分点的提升,相比于 Faster RCNN,模型综合提升3.43个百分点,可以很好地完成船舶的多分类检测任务。其中,作为特征提取网络的 DC-BFP 能更好地提取船体细节特征,对网络的提升效果最明显,这也验证了船体细节对于船舶检测的重要性。如图 9 所示,在同时

包含大小型船舶的场景中,模型能够完成重叠度较高、目标咬合、目标体积差别大的不同类型船舶检测任务,这对提升无人船感知能力具有实际意义。



图 9 模型效果展示
Fig. 9 Mode effect display

2.5 对比实验

为了展现本文模型的效果,分别与 Faster RCNN、SSD、YOLOv3 等目标检测模型进行对比。由

表 2 可以看出,模型对比其他目标检测模型,拥有最高的 mAP ,相比于其他模型分别提高 1.4、2.2、2.3、3.1、3.4、5.7 个百分点,这表明本文提出的模型在整体性能上优于其他模型,同时每个分类的 mAP 也保持较高的水平。相较于同属二阶段网络的 *Faster RCNN* 和 *Cascade RCNN*,本文模型通过优化船体细节解析能力,提高了每一个船舶类别的 mAP 。而相较于一阶段网络,本文模型仅在船体轮廓特征丰富的大型船舶中有细微差距,对于缺乏船体细节的中小型船舶检测效果依然是最好的。

各类别船舶在本文模型上的检测结果如图 10 所示,结合对比实验结果分析可知,本文所提出的模型能够提升多分类船舶检测的性能,拥有比其他目标检测模型更强的分类能力。由于网络更加注重船体细节的特点,帆船(SS)、领航艇(PB)、渔船(FV)和拖船(PT)这类小型船舶的检测能力较其他网络具有较大的提升。对于干货船这类纵横比差异较大的船舶种类的检测,模型也能够通过聚焦船体细节得到提升。同时,其他类型的船舶检测能力也能保持

表 2 对比实验结果

Tab. 2 Compaing experimental results

Model	AC	CS	DE	DC	FV	PT	PS	PB	RE	SS	OT	$mAP/\%$
Faster	86.9	91.2	89.2	80.7	92.4	82.3	86.9	91.0	95.6	85.3	88.0	88.1
SSD	93.0	89.9	89.2	71.4	91.3	78.3	81.6	92.0	91.1	80.9	85.1	85.8
YOLOv3	93.0	90.6	93.8	81.9	89.2	84.5	85.7	93.8	95.4	82.6	91.3	89.3
Cascade	89.4	91.1	89.5	81.9	92.8	81.5	85.5	92.9	95.1	83.7	88.5	88.4
RetinaNet	92.1	92.4	90.3	84.5	92.4	84.6	90.1	91.8	95.4	86.3	89.2	89.9
CenterNet	94.3	91.6	87.4	80.1	91.5	86.3	89.6	91.6	91.3	86.9	90.7	89.2
Ours	94.7	91.2	93.7	87.2	93.6	88.5	88.3	94.9	95.4	89.6	89.8	91.5

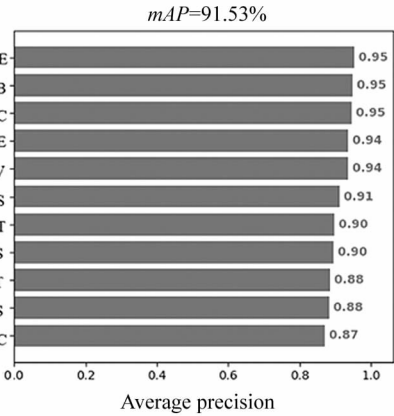
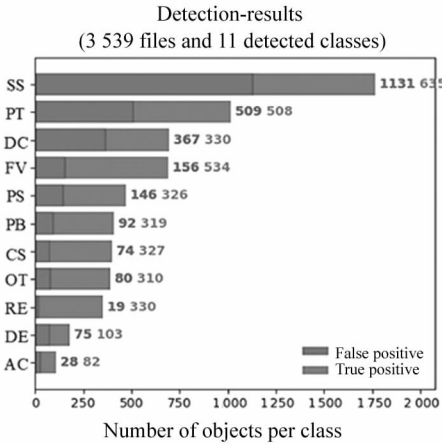


图 10 模型测试结果
Fig. 10 Mode test results

较高的精度度。

3 结 论

为提升无人船航行环境中目标的感知精度,有效保障无人船航行安全,本文提出了基于细节增强的级联多分类光电船舶检测模型。针对船舶整体轮廓相似的特点,利用 PMDA,降低模型对船舶整体轮廓的依赖;将可变形卷积引入 BFP 中,强化模型的船体细节特征提取能力,解决船体细节特征形状和尺度差异大的问题;CCD 融合全连接层和卷积层的分类信息进一步加强模型的船舶分类能力,使用 LSR,提升了模型的多分类泛化性。在自建的 MCSD11 数据集上进行对比实验,相比其他现存主流目标检测模型,模型取得了较高的平均召回率和均值平均精度,尤其是在船体细节不丰富和整体轮廓差异较大的中小型船舶样本中,效果十分突出,验证了模型的有效性,为无人船航行环境精准感知提供了较好的解决方案。

参考文献:

- [1] KRIZHEVSKY A, STSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]//25th International Conference on Neural Information Processing Systems, December 3-6, 2012, Lake Tahoe, Nevada, USA. New York: Curran Associates Inc., 2012:1097-1105.
- [2] CHEN Z J, CHEN D P, ZHANG Y S, et al. Deep learning for autonomous ship-oriented small ship detection[J]. Safety Science, 2020, 130:104812.
- [3] WANG L Q, FAN S Z, LIU Y X, et al. A review of methods for ship detection with electro-optical images in marine environments[J]. Journal of Marine Science and Engineering, 2021, 9(12): 1408.
- [4] HUO Y H, XU Z J. Photoelectric ship target identification method based on improved RA-CNN [J]. Journal of Shanghai Maritime University, 2019, 40(3): 38-43.
霍煜豪, 徐志京. 基于改进 RA-CNN 的舰船光电目标识别方法[J]. 上海海事大学学报, 2019, 40(3): 38-43.
- [5] ZHOU W N, CHEN P Q. A deep attention mechanism method for maritime salient ship detection in complex sea background[J]. Optoelectronics Letters, 2021, 17(7): 438-443.
- [6] ZHAO Y R, GUO H M, JIAO H, et al. Application of YOLOv4 with mixed-domain attention in ship detection [J]. Computer and Modernization, 2021(9): 75-82.
赵玉蓉, 郭会明, 焦函, 等. 融合混合域注意力的 YOLOv4 在船舶检测中的应用[J]. 计算机与现代化, 2021(9): 75-82.
- [7] SHAO Z F, WU W J, WANG Z F, et al. SeaShips: A large-scale precisely annotated dataset for ship detection[J]. IEEE Transactions on Multimedia, 2018, 20(10): 2593-2604.

- [8] SHAO Z F, WANG L G, WANG Z Y, et al. Saliency-aware convolution neural network for ship detection in surveillance video[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30(3): 781-794.
- [9] GUPTA V, GUPTA M, SINGLA P. Ship detection from highly cluttered images using convolutional neural network [J]. Wireless Personal Communications, 2021, 121(1): 287-305.
- [10] LI H, DENG L, YANG C, et al. Enhanced YOLO v3 tiny network for real-time ship detection from visual image [J]. IEEE Access, 2021, 9:16692-16706.
- [11] HAN X, ZHAO L N, NING Y, et al. ShipYOLO: an enhanced model for ship detection[J]. Journal of Advanced Transportation, 2021, 2021(10): 1-11.
- [12] ZHOU J C, JIANG P, ZOU A R, et al. Ship target detection algorithm based on improved YOLOv5[J]. Journal of Marine Science and Engineering, 2021, 9(8): 908.
- [13] HU J M, ZHI X Y, SHI T J, et al. PAG-YOLO: a portable attention-guided YOLO network for small ship detection [J]. Remote Sensing, 2021, 13(16): 3059.
- [14] XIONG W, LE L, ZHOU L, et al. Pedestrian re-identification algorithm based on multi-level feature fusion [J]. Journal of Optoelectronics • Laser, 2021, 32(8): 872-878.
熊炜, 乐玲, 周蕾, 等. 基于多层级特征融合的行人重识别算法[J]. 光电子 • 激光, 2021, 32(8): 872-878.
- [15] XU G Y, YIN M Y. Improved SSD object detection algorithm based on space-channel attention [J]. Journal of Optoelectronics • Laser, 2021, 32(9): 970-978.
许光宇, 尹孟园. 基于空间-通道注意力的改进 SSD 目标检测算法[J]. 光电子 • 激光, 2021, 32(9): 970-978.
- [16] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016:2818-2826.
- [17] LIU H Q, KANG X D, ZHANG F Q, et al. Improved image detection algorithm of cerebral arterial stenosis by Libra region-convolutional neural networks [J/OL]. Journal of Computer Applications. <http://kns.cnki.net/kcms/detail/51.1307.TP.20210926.1548.022.html>.
刘汉卿, 康晓东, 张福青, 等. 改进的 Libra 区域卷积神经网络脑动脉狭窄影像学检测算法[J/OL]. 计算机应用. <http://kns.cnki.net/kcms/detail/51.1307.TP.20210926.1548.022.html>.
- [18] LIU Y P, LIU T. Improved cascade RCNN pedestrian detection algorithm research[J]. Computer Engineering and Applications, 2022, 58(4): 229-236.
刘艳萍, 刘甜. 改进的 Cascade RCNN 行人检测算法研究[J]. 计算机工程与应用, 2022, 58(4): 229-236.

作者简介:

徐志京 (1972—), 男, 博士, 副教授, 硕士生导师, 主要从事无线通信和导航技术、人工智能及其应用方面的研究。