

基于 Mask R-CNN 算法的遥感图像处理技术及其应用



凌晨¹ 张鑫彤^{2,3} 马雷²

1 军事科学院系统工程研究院后勤科学与技术研究所 北京 100166

2 中国科学院自动化研究所 北京 100190

3 河北工业大学人工智能与数据科学学院 天津 300401

(179897839@qq.com)

摘要 遥感技术的发展使得遥感影像被应用于农业、军事等诸多领域,而深度学习方法的融入使得该项技术在目标检测、场景分类、语义分割方面取得了重大突破。与自然场景下的舰船检测不同,遥感图像中的舰船为俯视图,舰船较为密集,且容易与港口混合。当前对舰船检测的输出结果主要是检测框,缺少对舰船掩码的输出,使得无法全面分析出模型存在的不足;同时,由于遥感图像中的舰船停靠密集,容易产生漏检问题。为解决上述问题,利用 Mask R-CNN 对舰船进行目标检测,较全面地分析模型的训练情况、掩码和检测框的输出结果;通过对目标边缘的学习及参数的调整,使模型与舰船目标相适应。通过实验分析得出了适用于舰船检测的网络模型参数,从而有效降低了舰船停靠密集所产生的误检和漏检问题。

关键词: Mask R-CNN 算法;遥感图像处理技术;深度学习;舰船目标检测;影像提取与识别

中图分类号 TP751

Remote Sensing Image Processing Technology and Its Application Based on Mask R-CNN Algorithms

LING Chen¹, ZHANG Xin-tong^{2,3} and MA Lei²

1 Institute of Logistics Science and Technology, Beijing 100166, China

2 Institute of Automation, Chinese Academy of Science, Beijing 100190, China

3 School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China

Abstract With the development of remote sensing, there are many fields using remote sensing image, such as agriculture, military and so on. At the same time, deep learning, now, is applying in computer vision and image processing widely. It is successful in object detection, classification and semantic segmentation. Unlike fighting ship detection in natural scenes, fighting ships in remote sensing images are overhead views, dense and easy to mix with ports. The main result on fighting ship is taking bounding box as output, which is lacking the mask of the fighting ship, so may not analyzing the weakness in model. Meanwhile, because of the tight fighting ships in remote sensing images, there are easy to have missed detection. For solving the problems, this paper uses Mask R-CNN to detect fighting ships, analyzing training situation and the results of mask and bounding box. By learning the edges of objects and modifying parameter, making model more suitable to fighting ship. After experiment, it can be concluded that the appropriate parameters can effectively reduce the false positive and false negatives caused by compact berthing of fighting ships.

Keywords Mask R-CNN algorithm, Remote sensing image processing technology, Deep learning, Warship detection, Image extraction and recognition

随着我国卫星通信技术的发展,遥感图像在气象、农业、环境、军事等领域的应用越来越广,通过分析遥感图像来解决实际问题的方式已经逐步被人们认可。与自然图像不同,遥感图像具有实时性、全局性和地理性,携带的信息更丰富,这些优势促使遥感图像在各个领域的应用越来越成熟,其作用越来越突出。在气象领域,通过分析遥感图像预测未来的气

候变化;在农业领域,利用遥感图像可检测农作物生长是否正常;在环境监测方面,遥感图像可用于检测植物茂盛的地区是否发生火灾等。在社会各个领域,尤其在重视全局信息的军事领域中,问题的复杂性使得利用自然图像无法得到完整的敌方军队情况,因此无法正确分析敌方情况,导致作战命令错误,给国家造成严重损失。遥感图像的实时性和全局性为军

收稿日期:2019-09-18 返修日期:2019-12-11 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61806199)

This work was supported by the Youth Program of National Natural Science Foundation of China (61806199).

通信作者:马雷(lei.ma@ia.ac.cn)

事作战胜利奠定了基础,为敌情分析提供了重要的资料。与自然图像不同,遥感图像更具有全局性和实时性,图像携带的信息也更丰富,因此,分析、处理遥感图像是获得目标信息的重要途径。舰船检测作为我国海洋事业的重要分支,可及时有效地监视敌方港口和设备,及早发现军事情报,为我方军事指挥提供作战依据,使其掌握战场主动权。因此,舰船目标检测的研究对维护我国海洋权益及保护我国海洋领域安全具有重要的经济和战略意义。舰船目标检测研究的样本来源主要有两种:可见光遥感影像和 SAR(合成孔径雷达)影像。由于 SAR 卫星遥感图像存在图像特征不明显、解译不方便等问题,可见光遥感图像的优势逐渐突出,舰船检测的研究重点渐渐从 SAR 图像转向光学图像。

1 已有研究成果

遥感图像的处理可分为边缘特征提取、图像校正和波段融合等方向。通过提取图像边缘特征细化遥感图像,得到更清晰的遥感图像,常用的方法有 Sobel 算子^[1]、Log 算子和 Canny 算子^[2]。由于遥感图像在拍摄时受大气影响,导致数据存在误差,需要对其进行校正。常用的图像校正方法有直方图最小值去除法和回归分析法。为了得到更准确的遥感数据,通常采用精几何校正的方式进行重采样。常用的重采样方法有最近邻法^[3]、双线性内插法^[4]和三次卷积内插法^[5]。在实际情况下,需要将遥感图像可视化。为了得到理想的可视化结果,通常采用波段融合的方式。常用的波段融合方法有 HIS、PCA、Brovey 变换法、简单加权法、小波变换法和高通滤波法等^[6]。实验前对遥感图像预处理,有利于得到更准确、更真实的图像数据。

深度学习主要分为两大类:无监督学习和监督学习。无监督学习的方式有自编码器、降噪自编码器和稀疏自编码器等^[7]。无监督学习最大的特点是训练时输入不带标签的样本集,将样本集输入到网络中,通过训练编码器,尽可能输出与图像相似的特征值。自编码器通过特征映射,使输出结果接近输入。在自编码器的训练过程中,可以通过主动修改或限制隐藏层的大小来提高编码器的鲁棒性。降噪编码器在自编码器的基础上增加了噪声限制条件,以提高自编码器的特征提取能力,避免自编码器过拟合。将带有噪声的特征输入自编码器中训练,最后输出无噪声的特征。降噪自编码器主要被应用在特征提取方面。稀疏自编码器主要用来将高维特征映射为低维特征,起到降维和减少存储空间的作用。稀疏自编码器通过计算各层输入与输出中间的差异,自动调节自编码器隐藏层中的相关参数,得到最优的自编码器模型,从而提取出低维度的有效特征。自编码器、降噪自编码器和稀疏自编码器主要依赖于对图像的编码学习。由于缺少标签限制,通过无监督学习无法获取到图像的高层特征,因此使用无监督学习可能会存在“语义鸿沟”问题。卷积神经网络是最具代表性的监督学习,主要由卷积层、池化层、全连接层及一些激活函数组成,目前被广泛应用于计算机视觉领域中,如图像分类、目标检测和语义分割。LeNet 模型^[8]的提出掀起了卷积神经网络研究的热潮;AlexNet 模型^[9]的实验结果证明了卷

积神经网络的有效性;Inception 网络模型^[10]增加了网络的深度和宽度,并且在将深度增加到 22 层的同时将参数减少到 AlexNet 模型的 1/12,在当时的图像分类任务中表现最优。卷积神经网络将图像和对应标签一起输入到网络中训练,提高了网络的泛化性,且有效解决了“语义鸿沟”的问题。深度学习能够有效利用当今大数据的优势,提取大数据中的有效信息,方便人们分析、解决问题。

传统方法^[10-19]需要人工设定有效的图像特征,才能使目标算法拟合。随着人工智能的不断推进,神经网络的优势不断凸显,其最突出的特点是图像输入到网络中自主提取图像特征并学习,有效减小了人为因素造成的误差。近年来,深度学习不断成熟,在处理图像方面取得了重大成就,引起了遥感领域的广泛关注。目前,深度学习在遥感图像领域不断深入发展。为避免复杂的特征选择和提取过程,文献[20]使用卷积神经网络提取特征,并使用支持向量机(Support Vector Machine, SVM)对特征进行分类而得到结果,最后利用非极大值抑制(Non-Maximum Suppression, NMS)去除多余候选框。文献[21]为提高网络模型的鲁棒性提出 SVD Network(SVD Net),即卷积神经网络与奇异值分解(Singular Value Decomposition, SVD)结合,降低了时间复杂度,提高了模型的计算速度。文献[22]使用多光谱图像得到舰船目标,再对全色遥感影像切片,得到舰船影像,以减小遥感影像中的云层、波浪、阳光等因素对舰船特征的影响。将舰船影像输入神经网络中得到多分类结果,通过旋转卡壳法优化检测框,得到与目标最接近的检测框。此方法基本可以排除多种因素的干扰。文献[23]针对停靠在港口的舰船目标特征与岸上目标特征相似的问题,提出使用尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)进行配准,增强特征表达,再利用随机抽样一致(Random Sample Consensus, RANSAC)算法实现海陆分割,通过 Sobel 算子实现候选目标的提取,最后利用 VGGNet-16 进行目标识别。文献[24]在 Faster R-CNN 模型区域建议网络中的检测框大小设置中引入了 K-Means 聚类算法,并根据目标大小进行聚类,再将聚类结果输入区域建议网络进行目标检测。此方法提高了模型检测目标的准确度,同时加快了检测速度。文献[25]提出了一种基于 CNN 的靠岸舰船检测方法,其使用分类网络来检测船头位置,有效减少了候选区域,能更加准确地定位到舰船的位置。

基于神经网络方法的舰船检测是当前的研究热点,使用神经网络检测舰船目标,可以有效降低遥感影像中存在的干扰,提高舰船检测模型的有效性。当前对舰船检测的研究主要分析输出的检测框结果,缺少对舰船掩码的输出,因此无法确定模型是否可靠;同时,针对停靠密集存在的漏检问题,主要采用旋转检测框的策略,忽略了网络模型参数的问题。针对以上问题,本文提出使用 Mask R-CNN 模型对舰船进行检测,将 ResNet101 作为图像特征提取器,利用 FPN 融合高层和底层特征,使用 RoIAlign 实现像素对齐,使模型实现像素级特征的学习,从而实现舰船掩码的输出,通过掩码的输出分析模型的整体性能;同时,对 NMS 进行多次实验,得到适用于舰船检测的 NMS 值,有效减少了漏检情况。

2 Mask R-CNN 算法及其损失函数

Mask R-CNN 在 Faster R-CNN 的基础上添加了掩码预测分支,将其改进成一个实例分割框架,实现了目标检测与分

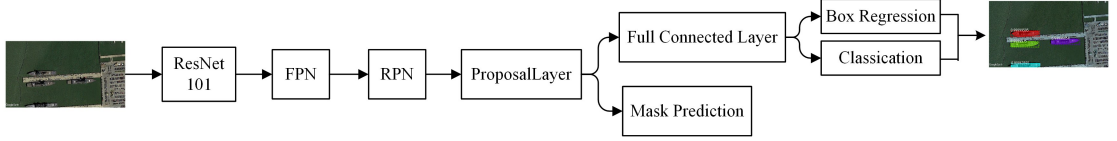


图 1 网络基本结构

Fig.1 Structure of model

在特征提取部分,本文将 ResNet101 作为 Mask R-CNN 的骨干网络结构,主要用来提取图像特征;使用 FPN^[27] 作为骨干网络结构的拓展,采用由高到低的连接方式提取图像的特征,融合从 ResNet101 提取到的高层和底层特征,避免了图像特征信息的丢失,使得网络能够更详细地描述不同尺寸的图像。受图像金字塔的影响,FPN 在此基础上添加了旁路连接,降低了网络的计算量,同时保留了检测场景中的小目标特征,并减少了模型的漏检情况。FPN 通过自底向上、自顶向下和横向连接的方式,融合了高层特征和底层特征,并保留了小目标特征,提高了网络检测的准确度。

RPN 用滑动窗口扫描 FPN 处理的图像特征,寻找图像中存在的目标区域,生成目标边框。通过使用 RPN 进行扫描预测,可以得到包含目标的最好区域;在此基础上对检测框的位置和尺寸进行调节;当重叠的检测框较多时,使用 NMS 保留得分最高的边框及其包含的区域。

为了适应像素级图像特征,Mask R-CNN 使用 RoIAlign 取代 Faster R-CNN 的 RoI Pooling,以解决区域不匹配的问题。与 Faster R-CNN 不同,Mask R-CNN 需要精确到每个像素点,将提取的特征和像素准确对齐,因此用于物体检测的 RoI Pooling 并不适用于 Mask R-CNN。RoIAlign 提取从 Proposal 层输出的不同尺寸的特征图中的感兴趣区域,再将提取到的目标区域输入到 softmax 分类器对目标进行分类并对检测框进行校准。同时,RoIAlign 提取的目标区域也被输入到掩码分支中,用于预测目标的掩码。

2.1 Mask R-CNN 算法的网络组成

(1)ResNet101 模块。本文使用 ResNet101 模型提取舰船图像特征,ResNet101 模型中的 101 指的是卷积层和全连接层共 101 层。选择 ResNet101 模型的目的是提取到高分遥感图像中舰船的复杂特征。网络模型的基本结构是 5 个卷积部分:Conv1 是 64 个 7×7 的卷积核;Conv2_x 由 3 组卷积核组成,每组包括 128 个 1×1 、128 个 3×3 和 256 个 1×1 的卷积核;Conv3_x 由 4 组卷积核组成,每组包括 128 个 1×1 、128 个 3×3 和 512 个 1×1 的卷积核;Conv4_x 由 23 组卷积核组成,每组包括 256 个 1×1 、256 个 3×3 和 1024 个 1×1 的卷积核;Conv5_x 由 3 组卷积核组成,每组包括 512 个 1×1 、512 个 3×3 和 2048 个 1×1 的卷积核。卷积核特征的映射公式如式(1)所示。

$$f_k^i = w_k^T x_{i,j}^l + b_k \quad (1)$$

割。经过实验,Mask R-CNN 可以与多种 R-CNN 框架结合^[26],网络的基本结构如图 1 所示。Mask R-CNN 基于 Faster R-CNN,将图像分解成像素级别,提取了目标更精细的特征空间分布,实现了多任务输出。

其中, f_k^i 表示第 l 层的特征经过第 k 层卷积核后的特征, w_k^T 为第 k 层的权重, $x_{i,j}^l$ 是大小为 (i,j) 的第 l 层特征, b_k 为第 k 层偏置。

ResNet101 的基本结构如图 2 所示,各卷积的分组如表 1 所列。

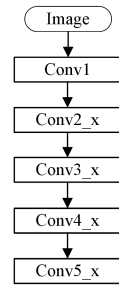


图 2 ResNet101 的基本结构

Fig.2 Basic structure of ResNet101

表 1 各卷积的分配情况

Table 1 Distribution of convolution

| 卷积层名字 | 卷积核分组 | 输出大小 |
|---------|--|------------------|
| Conv1 | $7 \times 7, 64$ | 112×112 |
| Conv2_x | $\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$ | 56×56 |
| Conv3_x | $\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$ | 28×28 |
| Conv4_x | $\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 23$ | 14×14 |
| Conv5_x | $\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$ | 7×7 |

(2)FPN 模块。早期主要使用图像金字塔来提取图像中不同尺寸的目标及其存在的特征,主要是在浅层获得详细的特征信息,在高层获得目标的语义信息。但由于需要大量的内存来计算图像特征,因此这些研究成果并不适用于神经网络这种计算量大的算法。FPN 的提出不仅缓解了金字塔算法计算图像特征时占用内存大的问题,而且解决了小目标特征被忽略的问题。FPN 的基本模块如图 3 所示。在 Mask R-CNN 中,FPN 将 ResNet101 中每组卷积层提取到的特征作为输入,通过上采样方法将特征尺寸转换成对应大小;采用自底向上、自顶向下和横向连接的方式,在保留小目标的情况下,将底层信息和高层信息进行融合。FPN 的基本架构如图 4 所示。

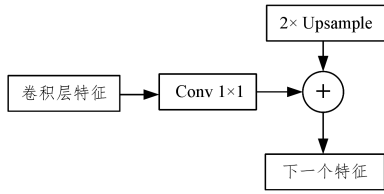


图3 FPN的基本模块

Fig. 3 Basic model of FPN

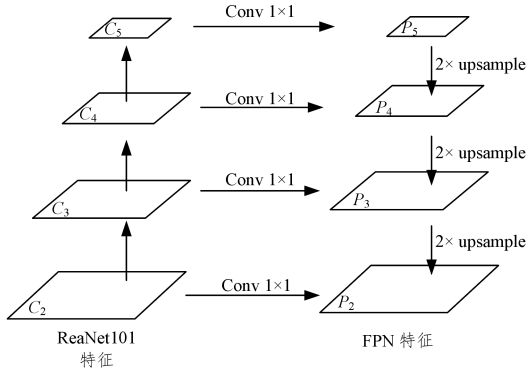


图4 FPN的基本架构

Fig. 4 Basic frame of FPN

自底向上指的是在 ResNet101 中特征由底层到高层传播,但在特征映射的过程中,特征尺寸会越来越小,如图4中的左半部分,其中底层特征反映图像的颜色、形状等;处于较高层的特征反映出图像的语义特征,即所属类别等。自顶向下指从高层特征向底层特征进行映射,如图4中右半部分所示。高层特征虽然尺寸比较小,但能表示更准确的语义信息。高层特征信息在物体分类及目标检测中占据着重要位置,因此使用上采样方法将小尺寸的高层特征放大,放大后再与当前 ResNet 中的特征相加。横向连接指使用 1×1 的卷积核提取 ResNet 模型输出的特征,在生成更有效特征的同时,减少中间层的通道数目,使不同维度的特征具有相同的通道数。

(3) RPN 模块。RPN 模型主要通过无分类滑动的方式寻找图像中的目标区域。将 FPN 提取到的特征 RPN 模型作为输入,使用滑动窗口,为每个特征点(像素)设置 3 个锚点。网络主要由卷积核构成,图像特征首先经过 64 个 3×3 的卷积核,然后生成两个分支,其中一个分支是 6 个 1×1 的卷积核,另一个分支是 12 个 1×1 的卷积核,最后得到预测检测框、真实检测框和交叉熵的值。RPN 中的 Softmax 主要用来区分前景和背景,RPN 的模型结构如图 5 所示。

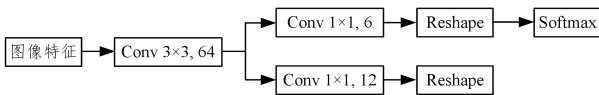


图5 RPN的模型结构

Fig. 5 Model structure of RPN

RPN 为每个像素点生成对应的锚点和以锚点为中心的滑动窗口,滑动窗口经过图像中的每一个像素点。如果得到的检测框与真实检测框之间的重叠度大于 0.7,则认为此部分像素是当前图像的前景;如果与所有真实检测框的重叠度小于 0.3,则为背景。滑动窗口如图 6 所示,图中

方框为锚点,圆点为锚点中心。

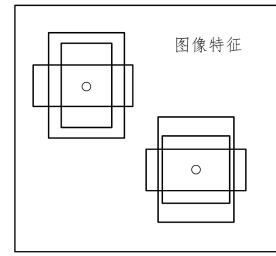


图6 滑动窗口

Fig. 6 Sliding window

ROI align 作为 RPN 的校正部分,使用双线性内插值法来减少像素点经过网络造成的偏差。假设最接近实际目标的 4 个像素点为 $(x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2)$,通过 4 个点预测的位置为 (x, y) ,则得到 x 方向上的线性插值如式(2)所示。

$$\begin{cases} f(x, y_1) = \frac{x_2 - x}{x_2 - x_1} f(x_1, y_1) + \frac{x - x_1}{x_2 - x_1} f(x_2, y_1) \\ f(x, y_2) = \frac{x_2 - x}{x_2 - x_1} f(x_1, y_2) + \frac{x - x_1}{x_2 - x_1} f(x_2, y_2) \end{cases} \quad (2)$$

与式(2)同理,得到 y 方向上的线性插值如式(3)所示。

$$f(x, y) = \frac{y_2 - y}{y_2 - y_1} f(x, y_1) + \frac{y - y_1}{y_2 - y_1} f(x, y_2) \quad (3)$$

将式(2)和式(3)合并,得到校正后的特征点位置,如式(4)所示。

$$f(x, y) = \frac{f(x_1, y_1)}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y_2 - y) + \frac{f(x_2, y_1)}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y_2 - y) \quad (4)$$

(4) Proposal Layer 模块。Proposal Layer 主要用于对 RPN 生成的候选框和目标前景/背景进行筛选。首先,排除长度过小的检测框,对所有前景检测框进行排序,得到前 k 个检测框;然后,使用 NMS 算法删除重叠的检测框,再从剩余检测框中选择 $m (m < k)$ 个检测框,得到与真实检测框最接近的候选框。候选框的表达式如式(5)所示。

$$\begin{cases} \hat{G}_x = P_w d_x(P) + P_x \\ \hat{G}_y = P_h d_y(P) + P_y \\ \hat{G}_w = P_w \exp(d_w(P)) \\ \hat{G}_h = P_h \exp(d_h(P)) \end{cases} \quad (5)$$

其中, $d_x(P)$ 和 $d_y(P)$ 为偏移值的中心点坐标, $d_w(P)$ 和 $d_h(P)$ 为偏移尺寸。 $d(p)$ 的表达式如式(6)所示。

$$d(P) = \mathbf{W}^T \phi(P) \quad (6)$$

其中, \mathbf{W} 为模型的权重,优化公式如式(7)所示。

$$\mathbf{W} = \underset{\mathbf{W}}{\operatorname{argmin}} \sum_i (t^i - \hat{\mathbf{W}}^T \phi(P^i))^2 + \lambda \|\mathbf{W}\|^2 \quad (7)$$

其中:

$$\begin{cases} t_x = (G_x - P_x) / P_w \\ t_y = (G_y - P_y) / P_h \\ t_w = \log(G_w / P_w) \\ t_h = \log(G_h / P_h) \end{cases} \quad (8)$$

得到推荐框后,设置 NMS 的阈值以筛选出与目标最接近的候选框。

(5) Fully Connected Layer 模块。Fully Connected Layer 为全连接层,特征由底层到高层所代表的含义不同,在底层主要获取图像的颜色、纹理等,在高层可得到图像的语义信息。使用全连接层对特征进行映射,可得到图像的语义信息;把特征向量映射成一维特征向量,为损失函数计算做准备。在 Mask R-CNN 模型中,一维特征向量主要作为分类损失函数和检测框回归函数的输入。经过全连接层后,图像特征被映射到标记空间中,减少了特征位置对分类造成的影响。

2.2 损失函数

为实现网络多任务功能,Mask R-CNN 在分类损失函数和检测框损失函数的基础上添加了掩码损失函数。Mask R-CNN 允许网络为不同类别的目标生成掩膜,而不同类别之间不存在干扰。使用分类器来预测类别标签并用于掩膜输出,从而解耦了分类和掩膜预测。基于 FCN 的语义分割使用像素级 Softmax 和多项式交叉熵损失进行分割实验,导致掩膜与分类之间相互影响;而 Mask R-CNN 使用了像素级 Sigmoid 和二进制损失有效降低了不同类别掩膜间的干扰。

L_{cls} 为分类损失函数,用来区分图像中存在的各个类别和背景。

$$L_{cls} = -\frac{1}{N_{cls}} \sum_i \log [p_i^* p_i + (1 - p_i^*) (1 - p_i)] \quad (9)$$

其中, p_i 为第 i 张图像中的目标, p_i^* 为第 i 张图像中的非目标。

L_{box} 是检测框回归函数,主要用来修正前景的锚点坐标,从而得到最佳的检测框。

$$L_{box} = \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (10)$$

其中, $p_i^* L_{reg}$ 仅在在有前景锚点即 $p_i^* = 1$ 的情况下才有回归损失;当没有前景锚点即 $p_i^* = 0$,没有回归损失。 L_{reg} 是回归损失函数,其公式为:

$$L_{reg} = R(t_i - t_i^*) = \begin{cases} \frac{0.5(t_i - t_i^*)^2}{\sigma}, & |t_i - t_i^*| < \frac{1}{\sigma} \\ |t_i - t_i^*| - 0.5, & \text{otherwise} \end{cases} \quad (11)$$

其中, t 为预测的目标检测框坐标, t_i^* 为真实的检测框坐标; $\sigma = 0.3$,用来控制损失函数的平滑区域。

L_{mask} 为平均二进制交叉熵损失,使用 Sigmoid 函数对每个像素进行特征映射,为不同类别的 RoI 提供独立的掩膜预测,将分类和语义分割任务进行解耦。

最终的损失函数为:

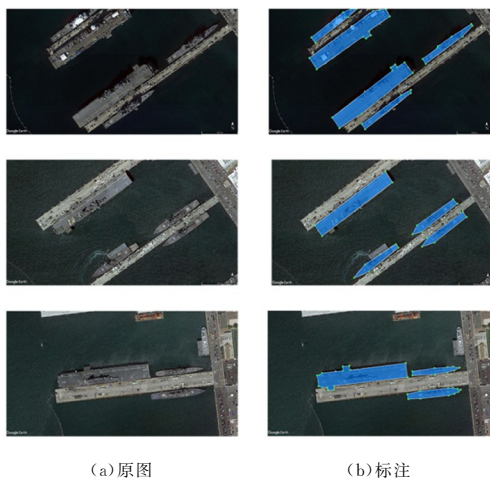
$$L = L_{cls} + L_{box} + L_{mask} \quad (12)$$

3 实验数据集的构建

本文使用的数据样本选自谷歌地球工具中拍摄的圣迭戈军港、珍珠港军港和布鲁克林造船厂。通过对国外各个军港的对比筛选,发现舰船存在一定优势。1)数量多:在 2018 年的统计中,舰船数量为 415 艘,其中航母数量为全球最多,共 20 艘,护卫舰、驱逐舰、潜艇、巡逻舰和水雷战舰艇共 165 艘,

因此在实验过程中可为模型训练提供充足的数据量。2)种类多:舰船不仅数量多,在种类上也占有优势,使用此舰船训练网络模型可有效提高网络模型检测出舰船的能力,减少模型存在的误检和漏检问题,同时也为后续舰船识别工作做准备。3)特征明显:这些舰船在颜色上与民用船有一定的区别,除医院船是白色外,其他舰船的颜色基本是灰色,虽然有些民用船只与舰船大小类似,但通过颜色判断仍然能区分出舰船。4)有具体的停放位置:通过舰船的停靠位置,可以较方便地区分出舰船与民用船,在制作目标数据时可有效减少舰船与民用船的对比时间,从而提高数据集的整理效率。基于以上 4 方面的优势,我们选择此舰船队作为实验研究的样本。

样本制作时,本文将图像的分辨率设定为 0.5m 和 1m。将分辨率设置为 0.5m,主要是为了让网络能够学习到样本的局部特征(主要为舰船的特征),如舰船上的标志线、舰船的具体轮廓及停靠港口的初步特征;选择分辨率为 1m,则是为了让网络可以学习到全局特征,如舰船港口的特征信息、舰船的主要形状,避免将陆地的目标检测成舰船。将两种不同分辨率的遥感图像输入到网络中交替训练,有利于提高网络的鲁棒性和泛化能力,降低网络的误检率。实验中的样本尺寸为 1366×632 ,共 300 张舰船图像,每张图像中至少有 1 艘舰船,共 1000 个舰船目标。本文使用 Labelme 开源工具制作样本标签,根据 Mask R-CNN 网络对标签的要求,在标注舰船时,通过密集的标注点连成线,将舰船的轮廓准确地显示出来,使网络模型可以根据舰船的轮廓生成准确的掩码和检测框。样本与标签如图 7 所示,其中左侧为样本原图,右侧为样本标签。在截取样本时,有些舰船可能在图像中显示不完整,如图 7(a)所示。针对此情况,设定当舰船有 50% 的船身在图像中进行标注,如图 7(b)所示;反之,则忽略。使用 Labelme 标注样本后,标签文件名称与样本名称一样,样本命名格式为“军港名称(首字母简称)-日期-第几张”,标签保存格式为 json,标签文件按目标分组,其中每组主要包括了坐标点、边缘线颜色、填充颜色、形状类型等,如图 8 所示。本文主要使用标签文件中每组的标签名称和坐标点信息。



(a) 原图

(b) 标注

图 7 样本与标签

Fig. 7 Samples and labels

```
[
  "line_color":null,
  "fill_color":null,
  "points":[
    [
      1040.0,
      337.75
    ],
    [
      1005.0,
      367.75
    ],
    [
      946.25,
      469.0
    ],
    [
      936.25,
      495.5
    ],
    [
      956.25,
      511.5
    ],
    [
      1005.75,
      367.5
    ],
    [
      946.25,
      469.0
    ],
    [
      956.25,
      511.5
    ],
    [
      1003.75,
      375.25
    ]
  ],
  "shape_type":"polygpr",
  "flags":{}
]
```

图8 标签保存形式

Fig. 8 Style of saving labels

使用谷歌地球工具截取 300 张圣迭戈图和珍珠港的军舰图分别作为训练集和验证集。为减少人为参与带来的实验误差,本文调用 python 中的随机函数,从 300 张舰船图像中随机选取 230 张作为训练集,70 张作为验证集,其中训练集有 728 个舰船目标,验证集有 272 个目标。截取 20 张布鲁克林造船厂的舰船作为测试集,按 1:1 的比例分别截取 10 张 0.5m 分辨率的测试图像,10 张 1m 分辨率的测试图像。为防止产生过拟合现象,本文对训练集进行 10 倍的数据增广,令每张图片从 0° 旋转到 360° ,间隔 36° 。

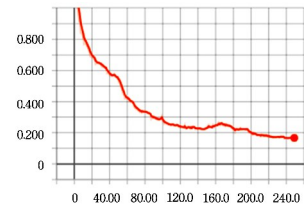
4 实验结果分析

本文使用 Mask R-CNN 模型对舰船目标进行检测。以 tensorflow 框架为基础;使用 python3.6 进行编译;显卡为英伟达 GP102。实验一共迭代 250 次,每 1000 步迭代一次。网络训练分 3 个阶段,第一阶段训练网络头部,迭代 60 次,学习率为 0.001;第二阶段微调 ResNet 第 4 部分及以上,迭代 90 次,学习率为 0.001;第三阶段微调全部层,迭代 100 次,学习

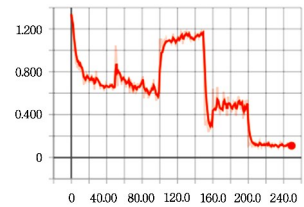
率为 0.0001。学习动量为 0.9,权重衰减为 0.0001。网络的初始权重为 COCO 数据集下的网络参数。

4.1 实验训练过程

在训练过程中,本文使用 TensorBoard 查看网络中损失函数的变化趋势,以便了解模型的训练效果。网络的总函数如图 9 所示。图 9(a)的曲线为网络训练时的损失函数曲线,总体呈下降趋势。图 9(b)是验证网络时的损失函数曲线,在网络开始训练至迭代到第 100 次的过程中,验证的损失函数总体呈下降趋势;从第 100 次到第 150 次这一过程中,网络在验证时的损失函数总体呈上升趋势,发生在训练的第二阶段微调 ResNet 部分,说明当前网络在训练过程中遇到瓶颈,需要减小学习率;从第 150 次迭代起,学习率缩小为原来的 1/10,损失函数直线下降;在第 160 次到第 200 次的过程中,网络验证时的损失函数稍有上升,然后趋于平稳;从第 200 次到第 250 次的过程中,验证时的损失函数首先直线下降,然后趋于平缓。在此过程中,验证的损失函数总体出现两次大的震荡,最后曲线趋于平稳,证明网络不再收敛,此时网络为最优状态。本文实验结果均是在迭代第 250 次下的模型参数。



(a) loss-训练

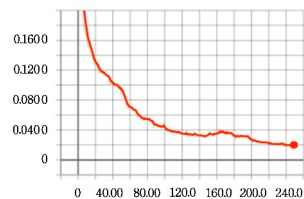


(b) loss-验证

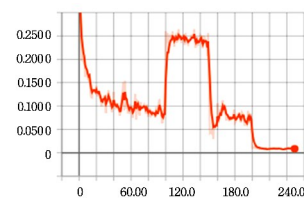
图9 Mask R-CNN的总损失函数曲线

Fig. 9 Loss curve of Mask R-CNN

Mask R-CNN 模型的检测框损失函数如图 10 所示。



(a) mrcnn_bbox_loss-训练



(b) mrcnn_bbox_loss-验证

图10 Mask R-CNN 的检测框损失函数曲线

Fig. 10 Loss curve of bounding box in Mask R-CNN

由图 10(a)可知,模型在训练时的损失函数总体呈平稳下降趋势。在图 10(b)中,模型迭代到第 100 次之前,检测框的验证损失函数总体呈下降趋势;从第 100 次开始,损失函数直线上升到 0.25 左右,然后平缓震荡;直到迭代至第 156 次时,网络趋于稳定,参数不再大幅度变化;但在网络迭代到第 157 次进行验证时,曲线呈直线下降趋势,这主要是由于网络的学习率发生了变化,同时训练层数加深,拓展了模型学习的空间;在第 157 次到 198 次间,检测框曲线稍稍上升;在第 199 次时,损失曲线直线下降,然后趋于平稳,网络不再优化。

Mask R-CNN 的分类损失函数如图 11 所示。由图 10(a)可知,网络的分类损失函数在训练过程中呈下降趋势。图 10(b)是在验证情况下的分类损失函数变化曲线,从开始迭代到第 148 次时,曲线在震荡中总体呈下降趋势,属于模型训练的正常现象;由于从第 150 次迭代开始微调网络的所有层,网络原有参数发生变化,影响了特征表达,从而造成分类结果存在偏差;但从第 200 次迭代开始,损失曲线直线下降,损失曲线小幅度震荡并趋于稳定,证明模型参数不再更新。

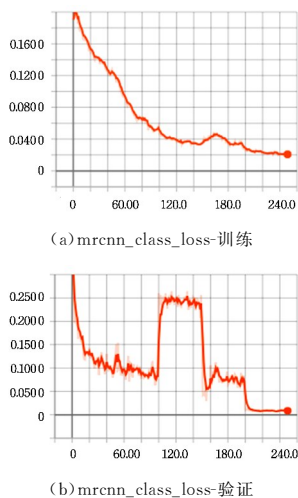


图 11 Mask R-CNN 的分类损失函数曲线

Fig. 11 Loss curve of classification in Mask R-CNN

并逐渐趋于平缓。网络在验证时的掩码损失函数如图 12(b)所示,从开始迭代到第 15 次的过程中,损失函数值从 0.4 左右下降到 0.2 左右,损失曲线趋于稳定,证明在当前学习率下网络模型已经趋于饱和,参数不再更新;从第 100 次迭代开始,损失曲线直线上升到 0.4 左右;从第 150 次起,由于学习率缩小为原来的 1/10,扩大了网络的训练空间,经过 10 次迭代后,学习率下降到 0.05;迭代到第 207 次时,损失值直接下降到 0.02,损失曲线趋于平缓,网络参数不再更新。

Mask R-CNN 中包含一个小型网络 RPN,此网络有自己的损失函数,主要用来区分图像的前景和背景。RPN 的损失函数如图 13 所示。由图 13(a)可知,RPN 在训练时总体呈下降趋势。在验证过程中,如图 13(b)所示,损失曲线大幅度震荡,震荡区间为第 1 次到第 150 次迭代;从第 151 次迭代起,损失函数直线下降,这主要是由于模型的学习率产生了变化;迭代到第 160 次时,损失值降低到 0.05;由于网络层数也发生了变化,从第 161 次到第 196 次迭代时,损失值上升到 0.1,这在模型训练中属于正常情况;迭代到第 204 次时,损失值下降到 0.03,损失曲线平稳下降;迭代结束时,损失值为 0.02。

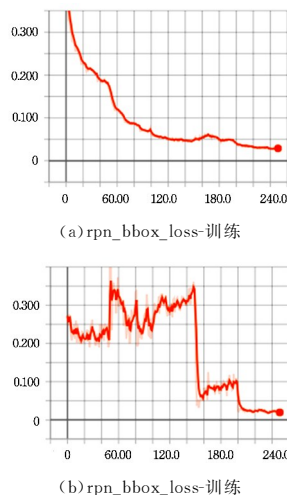


图 13 RPN 的检测框损失函数曲线

Fig. 13 Loss curve of bounding box in RPN

Mask R-CNN 的掩码损失函数如图 12 所示。

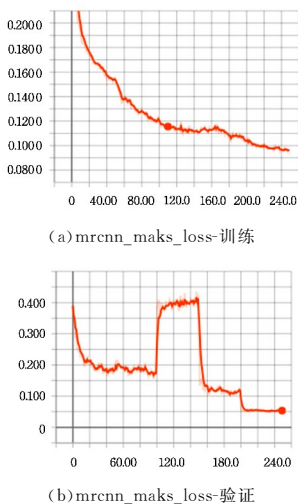


图 12 Mask R-CNN 的掩码损失函数曲线

Fig. 12 Loss curve of mask in Mask R-CNN

由图 12(a)可知,掩码损失函数在训练时,曲线总体下降

RPN 的分类函数损失曲线如图 14 所示。

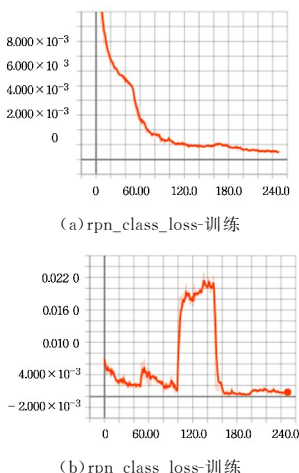


图 14 RPN 的分类损失函数曲线

Fig. 14 Loss curve of classification in RPN

由图 14(a)可知,RPN 在训练时的损失曲线总体呈下降

趋势,从第 120 次迭代开始,损失曲线趋于平缓。图 14(b)为 RPN 在验证时的损失函数曲线,从迭代开始到第 99 次,损失值总体在 0~0.004 范围内浮动;从第 100 次迭代开始,损失曲线呈直线上升状态;迭代到 140 次时,损失值增加到 0.02;从第 148 次到第 160 次的迭代过程中,由于学习率发生了变化,模型在当前基础上继续优化参数,损失曲线直线下降,损失值从 0.02 下降到 0.003;从第 161 次起,损失曲线平缓下降;在迭代结束时,损失值为 0.0008,达到最低。

Mask R-CNN 在训练时,各个损失函数的曲线总体呈下降趋势,而在验证时,损失曲线出现了大幅度震荡,主要集中在第 100 次迭代到第 180 次迭代之间,此时模型参数与训练的网络层均发生变化,出现损失曲线震荡现象;在第 180 次以后,损失曲线呈小幅度下降趋势并趋于平稳。

4.2 评价公式

本文主要使用精确率、召回率和 F 分数来评估网络模型。精确率和召回率越高,代表模型越好;但在实际情况中,精确率随召回率的增大而减小。因此,使用 F 分数对两者进行平衡,得到了较全面的模型评估。

精确率的计算公式如下所示:

$$precision = \frac{TP}{TP + FP} \quad (13)$$

其中,TP 表示舰船目标被正确识别为舰船的数量,FP 表示非舰船目标被识别为舰船的数量。

召回率的计算公式如下所示:

$$recall = \frac{TP}{TP + FN} \quad (14)$$

其中,TP 与式(13)中 TP 的含义相同;FN 表示未识别出的舰船数量。

F 分数的计算公式如下所示:

$$F\text{-score} = \frac{2 \cdot precision \cdot recall}{precision + recall} \quad (15)$$

其中,precision 为精确率,recall 为召回率。

4.3 实验结果

为得出适用于舰船检测的 NMS 参数值,本文对 NMS 参数进行多次实验。使用 DetEval^[28] 评估 Mask R-CNN 模型的召回率和精确率,实验结果如表 2 所列。

表 2 不同 NMS 参数下的召回率和精确度

Table 2 Recall and precision of different NMS parameters

| NMS | recall | precision |
|-----|--------|-----------|
| 0.1 | 0.66 | 0.82 |
| 0.2 | 0.68 | 0.83 |
| 0.3 | 0.7 | 0.83 |
| 0.4 | 0.7 | 0.82 |
| 0.5 | 0.7 | 0.83 |
| 0.6 | 0.7 | 0.82 |
| 0.7 | 0.69 | 0.81 |
| 0.8 | 0.69 | 0.81 |
| 0.9 | 0.69 | 0.8 |
| 1 | 0.69 | 0.8 |

由表 2 可以得出:在 NMS 增加的过程中,精确率和召回率大体呈先上升后下降的趋势。当 NMS 参数值为 0.1 时,模型的精确率和召回率最低;当 NMS 参数值从 0.1 增加到 0.3 时,模型的精确率呈上升趋势,从 0.82 上升到 0.83;当 NMS 为 0.4 时,网络模型的召回率保持稳定,但精确率降低

了 0.01;当 NMS 的值为 0.5 时,Mask R-CNN 的召回率和精确率与 NMS 为 0.3 时的一样;当 NMS 为 0.6 时,网络的召回率保持不变,但精确率降低到 0.82;当 NMS 参数值分别为 0.7 和 0.8 时,Mask R-CNN 的召回率和精确率均为 0.69 和 0.81;当 NMS 的值分别为 0.9 和 1.0 时,召回率与 NMS 为 0.8 时的召回率相等,但精确率降低了 0.01。

通过设置不同的 NMS 值,得到不同的精确率和召回率。当 NMS 参数值过小时,距离较近的舰船之间的检测框会因为重叠率大于阈值(NMS 参数值)而被误判为一个舰船目标下的检测框,导致其中一艘舰船的检测框被误删,造成召回率较低,产生漏检现象;当 NMS 参数值过大时,则会将两艘停靠较近的舰船误检为一个目标,即一个检测框将两艘舰船框在一起,影响了模型的准确度。当 NMS 参数为 0.3 或 0.5 时,模型的召回率和精确率均较高。

本文在实验时将 NMS 设置为 0.5,选取验证集作为模型的评估样本,将评估样本输入到训练好的 Mask R-CNN 中得到预测出的舰船坐标点,使用 DetEval 开源工具将预测出的舰船坐标点与真实的舰船坐标点进行比较,得到的结果如图 15 所示。图中的横坐标 tr 表示预测出的坐标点与真实坐标点之间的重叠度。由图 15 可知,3 条曲线在 $tr=0.8$ 之前均处于平稳状态,在 0.8 以后呈下降趋势。

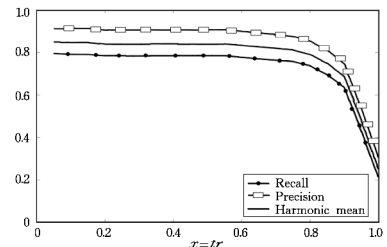


图 15 模型评估

Fig. 15 Model evaluation

为直观验证模型的检测效果,从测试集中选取 4 张有代表性的图像进行检测,分别为舰船停靠稀疏、舰船停靠稍密集和舰船停靠密集,并对检测结果进行分析,如图 16 所示。

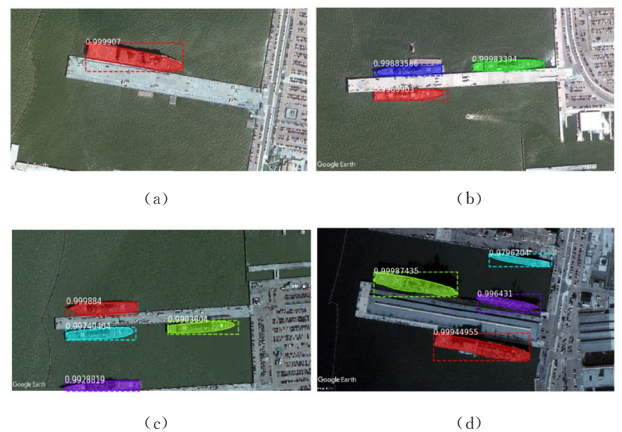


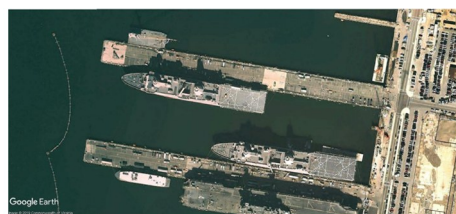
图 16 舰船检测

Fig. 16 Detection of fighting ships

由图 16(a) 的检测结果可以看出,当舰船停靠稀疏时,Mask R-CNN 的检测效果较好,检测框与掩码均为较优状态,即检测框刚好将舰船全部包围在框内,掩码将舰船完整覆盖。

图 16(b)为舰船停靠较密集的情况,根据图中的掩码显示,可以得到舰船的停靠方向和船头/船尾,并且当图像中存在民用船只或其他小型军用船只(非作战船只)时,模型将舍弃此类型目标。图 16(c)和图 16(d)为舰船停靠密集的情况,从中可以看出,当舰船停靠数量较多时,Mask R-CNN 输出的掩码效果相比停靠稀疏的图像效果稍差,但根据掩码对舰船的描述,仍然能区分出舰船的停靠方向和船头/船尾,同时没有出现漏检、误检的情况。

在解决舰船与港口特征易混合的问题时,我们发现模型对船影较明显的图像不敏感,导致漏检,如图 17 所示。其中,图 17(a)为待检测图像;图 17(b)为模型的检测结果。导致漏检的主要原因是此类型的图像样本较少,即包含运输舰和航母,并且含有船影的图像。当前制作的样本中,大多数为不存在倾斜角的驱逐舰和巡洋舰图像,此类型船只在形状上与运输舰有一定区别:驱逐舰和巡洋舰的形状较为细长,而运输舰则较宽。因此,网络不能通过学习大量驱逐舰检测出运输舰。除舰船类别不均匀以外,当舰船排列较为密集时,检测框重叠度较高,导致检测框被误删,从而产生漏检问题。



(a) 原图



(b) 漏检

图 17 漏检情况

Fig. 17 Missed detection

本文以召回率和精确率为模型评价标准进行实验。1)通过设置不同的 NMS 参数得到适合舰船检测的参数,由实验结果对比可知,当 NMS 参数为 5 时,模型的召回率和准确率最高,分别为 0.70 和 0.83。对于 NMS 的设置,过高和过低都会影响模型的检测结果。2)本文将模型的召回率和准确率以曲线图的形式显示,发现在 $tr=0.8$ 之前,模型的召回率与准确率保持平缓;在 0.8 之后则直线下降,最终停止在 0.2 和 0.3 左右。3)为了使结果更加清晰,本文将检测结果可视化输出,发现模型在拍摄时无倾斜角的图像及不存在船影的检测结果较好,没有出现误检和漏检现象;当图像存在船影时,会产生漏检,主要原因是此类型的训练图像较少,导致模型无法在此情况下将舰船检测出来。

结束语 本文使用谷歌地球软件下载圣迭戈、珍珠港和布鲁克林造船厂的舰船图片进行目标检测研究,针对港口停靠舰船较密集、舰船易与港口混淆的问题,提出使用 Mask R-CNN 模型,此模型在分类和检测分支的基础上加入掩码分支,使目标同时有检测框和掩码,为今后的研究工作提供了更

丰富的目标信息。Mask R-CNN 模型实现了多功能化,即同时输出目标的检测框和掩码。

(1)本文使用 Mask R-CNN 模型减少了因舰船与港口特征相似而导致的误检问题。网络模型在提取特征部分采用 FPN,融合了底层与高层特征,增强了特征的表达;在掩码生成部分,使用 RoIAlign 取代原有的 RoI Pooling,使得网络模型适合在像素级特征的条件学习并适用于分割;Mask R-CNN 避免了多分类下生成的掩码相互干扰的问题,通过结合分类器与掩码生成部分,提高了不同类别之间掩码的差异性。

(2)本文在实验过程中对多个损失函数进行分析,确保选择本实验中最优的网络参数进行实验。通过实验可以得出,在舰船停靠稀疏、稍密集和密集的情况下,Mask R-CNN 都表现出良好的检测效果,即掩码可以比较完整地显示出舰船的整体轮廓和船头/船尾,检测框能够将舰船目标完整地显示在框内。同时,本文在不同 NMS 参数下进行实验,综合分析精确率和召回率,得出适合舰船目标检测的 NMS 参数。对 NMS 进行多次实验的主要原因是舰船与其他船只相比体积较大,停靠港口时会有多艘舰船并排的现象,导致检测框重叠度较高。当 NMS 过大时,会将多艘停靠紧凑的舰船误检为一艘舰船;当 NMS 过小时,容易将重叠度较高的检测框误删,造成漏检,从而影响模型的检测效果。

通过实验发现,当一张图像中舰船种类较多时,会产生漏检问题,主要原因是数据不够多样化,图像中包含的舰船类别较少。同时,本文认为在增加数据多样化、扩大训练网络数据样本的同时,应提高模型对图像的特征表达;在图像特征提取部分,考虑使用其他网络模型提取目标特征,使特征表达更清晰,尤其当舰船目标种类较多时,如果特征表达不清晰,则会造成误检和漏检问题。针对军港舰船密集排列停靠的现象,在未来工作中,考虑使用旋转检测框的方法来减少网络当前存在的漏检问题,主要根据舰船停靠方向画出对应的检测框。但是,由于当前检测框是非旋转的,当舰船停靠紧凑时,检测框重叠率较高,导致网络模型在非旋转检测框的情况下无法将排列密集的舰船检测出来。

参考文献

- [1] WU H B, HE Z, LI W J. Image Edge Detection Based on Improved Sobel Operator and Embedded Realization[J]. Machine Tool & Hydraulics, 2009, 37(10): 133-136.
- [2] WANG N, LI X. An improved edge detection algorithm based on the Canny operator[J]. Journal of Shenzhen University (Science and Engineering), 2005, 22(2): 149-153.
- [3] LIN X F, DING X Q, WU Y S. Theoretical Analysis of Confidence Estimation of Nearest Neighbor Classifier[J]. Chinese Science Bulletin, 1998, 43(3): 322-325.
- [4] YANG G D, GUO C, ZHANG X Q. Research on Geometric Correction Method of GF-2 Image[J]. Geomatics & Spatial Information Technology, 2018, 41(10): 25-27.
- [5] CHEN J W, CHANG S L, ZHANG Y Y. Multiscale time domain full waveform inversion based on cubic convolution interpolation[J]. Journal of China Coal Society, 2018, 41(10): 25-27.
- [6] FU Z F, ZENG Z Y, ZHANG Z L. Wavelet and IHS transform combined to realize band fusion of ETM images[J]. GEO-Infor-

- mation Science, 2005, 7(4):29-32.
- [7] ZHANG X, ZHOU W, DUAN Z M. Convolutional sparse auto-encoder for image super-resolution reconstruction[J]. Infrared and Laser Engineering, 2019, 48(1):324-330.
- [8] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks[C]// NIPS. Curran Associates Inc., 2012:1097-1105.
- [10] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016:2818-2826.
- [11] YANG L, SU L H, WU B G, et al. Target Detection Algorithm of Hyperspectral Remote Sensing Imagery Combined with CEM [J]. Journal of Chongqing University of Technology (Natural Science), 2017(12):152-156, 178.
- [12] WANG W X, FU Y T, DONG F, et al. Remote sensing ship detection technology based on DoG preprocessing and shape features[C]// 2017 3rd IEEE International Conference on Computer and Communications (ICCC). 2017.
- [13] BI F K, CHEN J, ZHUANG Y, et al. A model based hierarchical method for inshore ship detection in high-resolution remote sensing images[C]// 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2017.
- [14] HUANG J, JIANG Z G, ZHANG H P. Region proposal for ship detection based on structured forests edge method[C]// 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2017.
- [15] XIE X Y, Xu Q, Lei H. Fastership detection from optical satellite images based on ship distribution probability analysis[C]// International Workshop on Earth Observation & Remote Sensing Applications. IEEE, 2016.
- [16] LIU S T, JIANG K H, LIU Z X. Ship Target Detection of Aerial Reconnaissance Image Base on Region Covariance and bhicetness[J]. Systems Engineering and Elecornics, 2019(5).
- [17] YE G Q, ZHA X, LI H B. High Resolution Remote Sensing Image Ship Detection Based on Visual Significance[J]. Hydrographic Surveying And Charting, 2018, 38(4):51-55.
- [18] YANG G, WEI P Z, DENG X. Detection Method of Ships in Port Based on SIFT Characteristics[J]. Electronic Technology & Software Engineering 2017(5).
- [19] ZHOU Y G, WANG P, GAO Y H. Remote Sensing Image Classification with Bag-of-Visual-Words Model [J]. Journal of Chongqing University of Technology (Natural Science), 2015, 29(5):71-77.
- [20] HUANG J, JIANG Z G, ZHANG H P, et al. Ship object detection in remote sensing images using convolutional neural networks[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(9):132-139.
- [21] ZOU Z, SHI Z. Ship Detection in Spaceborne Optical Image with SVD Networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(10):5832-5845.
- [22] CHEN L Q, SHI W X, FAN E C. Ship detection in optical remote sensing images based on multi-class learning[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2019(5).
- [23] ZHANG X. A method for ship detection and recognition in port area based on template matching and deep learning[J]. Information Technology & Informatization, 2019(4):59-63.
- [24] ZHAO C H, ZHOU Y. Ship Target Detection and Recognition Based on Improved Faster R-CNN Algorithm [J]. Journal of Shenyang University (Natural Science), 2018, 30(5).
- [25] WU F, ZHOU Z Q, WANG B. Inshore Ship Detection Based on Convolutional Neural Network in Optical Satellite Images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(11):4005-4015.
- [26] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39(6):1137.
- [27] LIN T Y, PIOTR D, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection[C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017:2117-2125.
- [28] WOLF C, JOLION J M. Object count Area Graphs for the Evaluation of Object Detection and Segmentation Algorithms[J]. International Journal on Document Analysis and Recognition, 2006, 8(4):280-296.



LING Chen, born in 1979, B. S., assistant researcher. His research interests include internet of things and big data intelligence analysis.



MA Lei, born in 1980, Ph. D, associate professor. His main research interests include intelligent interpretation of remote sensing images.