

• 通信与信息工程 •

改进 YOLOX 的 SAR 近岸区域船只检测方法



刘霖^{1,2,3}, 肖嘉荣^{2,3*}, 王晓蓓², 张德生², 喻忠军^{2,3}

(1. 电子科技大学信息与通信工程学院 成都 611731; 2. 中国科学院空天信息创新研究院 北京 海淀区 100094;

3. 中国科学院大学电子电气与通信工程学院 北京 怀柔区 100049)

【摘要】针对 SAR 近岸区域船只检测准确率低与虚警率高的问题, 提出一种基于改进注意力机制与旋转框的 SAR 近岸区域船只检测方法。该方法首先通过改进坐标注意力机制并引入至特征提取网络中, 提升网络的特征提取能力; 其次, 增加角度分类头, 并引入二维高斯分布, 计算预测分布与目标分布的 KL 散度评估旋转框损失值, 完成目标的角度信息提取; 再基于 YOLOX 算法中的无锚框(AF)机制, 减少候选框冗余, 使模型轻量化并进一步提高定位精度。最后在公开数据集 official-ssdd 上进行测试, 在嵌入式平台 (NVIDIA Jetson AGX Xavier) 上对模型进行推理验证。该算法模型计算参数仅 1.14 M, 在近岸情况下平均检测精度较 YOLOX 模型提高了 18.77%, 总体检测精度达到 94.2%。验证结果表明, 该算法适用于复杂场景下任意方向的密集船只目标检测, 满足实时处理需求。

关键词 改进坐标注意力机制; 近岸区域; 旋转目标框; SAR; 船只检测

中图分类号 TN957.52; U675.79 文献标志码 A doi:10.12178/1001-0548.2022039

Improved YOLOX SAR Near-Shore Area Ship Detection Method

LIU Lin^{1,2,3}, XIAO Jiarong^{2,3*}, WANG Xiaobei², ZHANG Desheng², and YU Zhongjun^{2,3}

(1. School of Information and Communication Engineering, University of Electronic Science and Technology of China Chengdu 611731;

2. Aerospace Information Research Institute, Chinese Academy of Sciences Haidian Beijing 100094;

3. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences Huairou Beijing 100049)

Abstract To solve the problem of low accuracy and high false alarm rate of synthetic aperture radar (SAR) nearshore area vessel detection, a new SAR nearshore area vessel detection method based on improved attention mechanism and rotating frame is proposed. Firstly, the feature extraction capability of the network was enhanced by improving the coordinate attention mechanism and introducing it into the feature extraction network. Secondly, the angle classification head was added and the two-dimensional Gaussian distribution was introduced to calculate the KL divergence between the prediction distribution and the target distribution, so as to evaluate the loss value of the rotating frame and complete the angle information extraction of the target. Then, based on the anchor frameless (AF) mechanism of YOLOX algorithm, the model can be made lightweight and the positioning accuracy can be further improved by reducing the redundancy of candidate frames. Finally, the model was tested on the open dataset Official - SSDD, and the inference verification was performed on the embedded platform (NVIDIA Jetson AGX Xavier). The calculation parameter of the algorithm model is only 1.14M, and the average detection accuracy of the algorithm model is 18.77%, higher than that of the YOLOX model in the nearshore condition, and the overall detection accuracy reaches 94.2%. The verification results show that the algorithm is suitable for dense ship target detection in any direction in complex scenes and can meet the requirements of real-time processing.

Key words improved coordinate-attention; inshore region; rotation anchor; SAR; ship detection

合成孔径雷达 (synthetic aperture radar, SAR) 具有全天时、全天候对地观测能力, 针对广阔的海洋区域可以大范围地获取海面 SAR 图像, 有效探测海面船只, 对我国海洋事业起着至关重要的作用。

但是 SAR 图像在视觉上是灰度图像, 目标纹理特征不明显, 导致基于 SAR 图像的目标信息提取困难。同时, 由于 SAR 相干成像的特点, 近岸区域中存在的各类金属建筑会产生与船只特征相近的强

收稿日期: 2022-01-27; 修回日期: 2022-06-11

作者简介: 刘霖 (1982-), 男, 副研究员, 主要从事合成孔径雷达数字信号处理方面的研究。

*通信作者: 肖嘉荣, E-mail: 524737474@qq.com

散射点, 仅靠人工进行目标检测难度很高, 效率较低, 容易产生误检现象, 且船只目标具有不定朝向、密集排列特点, 易产生漏检现象, 因此近岸区域下的 SAR 图像船只检测是一个具有挑战性的问题。

传统的 SAR 图像船只目标检测方法多采用恒虚警率法, 其依赖于人工建模, 通过设计海杂波模型提取船只特征, 无法有效区分船只与岛屿、近岸建筑等虚警目标, 因此在背景复杂的近岸情形下效果不佳^[1]。随着深度学习技术的不断发展, 其在 SAR 图像目标检测方面的应用具有良好的发展前景, 目前基于深度学习方法进行 SAR 船只检测大体分为两个技术路线, 前者对以 YOLO^[2] 和 SSD^[3] 为代表的单阶段目标检测算法进行适应性改进, 以实现实时高效船只检测。如文献 [4] 针对船只在近岸场景下方向多变问题, 对 YOLOv3 算法进行改进, 结合水平框和旋转框的预测结果实现船只目标方位角估计, 提升了 YOLOv3 算法在复杂场景下的船只检测效果。文献 [5] 针对单阶段目标检测算法在复杂场景下船只检测性能不佳问题, 在 YOLOv3 上引入了通道注意力机制, 通过加强各网络层通道间的特征融合, 提高模型在近岸场景下的检测性能。后者以 Faster R-CNN^[6] 为代表的双阶段目标检测算法进行适应性改进, 以实现高精度船只目标检测, 如文献 [7] 通过多尺度特征融合与特征通道校准方法, 提高 Faster-RCNN 算法在复杂场景下的船只目标检测性能。

单阶段目标检测算法由于缺少了双阶段算法的精细处理, 在面对 SAR 图像中尺度大小不一、视觉特征不明显的船只目标时表现不佳。另外, 单阶段目标检测算法 YOLOv3 作为 Anchor Based 算法虽然在一定程度上缓解了双阶段目标检测算法, 如在 Faster R-CNN 中选择性搜索带来的候选框计算量爆炸的问题, 但其每个网格中不同尺寸锚框的生成依赖于人工参数设置, 目标的定位精度严重依赖于手动调参。同时为兼顾目标检测实时处理平台资源功耗受限等情形, 需要避免主流算法模型参数量大、网络推理耗时长的问题。

针对以上问题, 本文提出了一种基于改进 YOLOX 的 SAR 近岸区域密集船只高精度检测方法, 有效地解决 SAR 近岸情形下算法模型准确率低、虚警率高且参数冗余问题, 主要创新点如下:

1) 改进坐标注意力机制, 进一步提升对空间域信息的学习, 并引入至特征提取网络, 提升网络对船只目标信息的聚焦, 实现目标信息的有效提取;

2) 采用旋转目标框进行目标检测, 增加角度分类头学习目标的角度信息, 并通过将旋转目标框转化为二维高斯分布的方法, 计算预测分布与目标分布的 KL 散度作为旋转框损失值训练网络, 提升网络在近岸区域下的检测精度;

3) 基于 YOLOX 的 Anchor Free 机制, 结合上述两点设计, 有效减少由于角度分类头导致的冗余候选框激增问题, 避免冗余框带来的无效计算。

1 算法结构

当前最新提出的 YOLO 系列算法中效果最优的单阶段目标检测算法 YOLOX^[8], 其大致分为 3 个组成部分: 主干网络、特征聚合网络与检测头。主干网络采用 CSPDarknet 来对图像进行特征提取, 其主要组成部分 CBS 由卷积 (convolution)、批量归一化 (batch-normalization, BN) 和激活函数 (silu) 组成, YOLOX 在主干网络前通过采用 Focus 模块将输入的高分辨率图像拆分成多个低分辨率的特征图, 以减少传统下采样操作产生的信息损失, 而在主干网络提取特征后通过空间金字塔池化层 (spatial pyramid pooling, SPP) 来提取不同尺寸的空间特征信息, 提升模型对不同尺寸目标的鲁棒性。此外, 主干网络中的 CSPLayer 通过其残差结构可以将特征图的梯度变化集成到了最后的输出中, 减少网络参数与运算量, 并同样用于特征聚合网络 PAFPN 中以加强网络特征聚合能力, 是目标特征提取的关键网络。

而 YOLOX 检测头采用 Decoupled Head 对提取的特征进行解耦, 使回归与分类任务分成两部分, cls 特征图负责学习目标的类别信息, reg 与 obj 特征图则负责学习目标的位置及存在信息, 以有效提高网络的检测精度, 实现高精度的船只目标检测。最后在预测端将不同尺度的特征向量进行融合, 通过 SimOTA 方法对融合的特征向量进行正样本筛选匹配, 得到与真实样本近似的预测候选框, 以消除先前 YOLO 系列检测算法中对先验 Anchor 设置的依赖, 减少冗余候选框造成的无意义计算资源消耗, 其网络结构如图 1 所示。

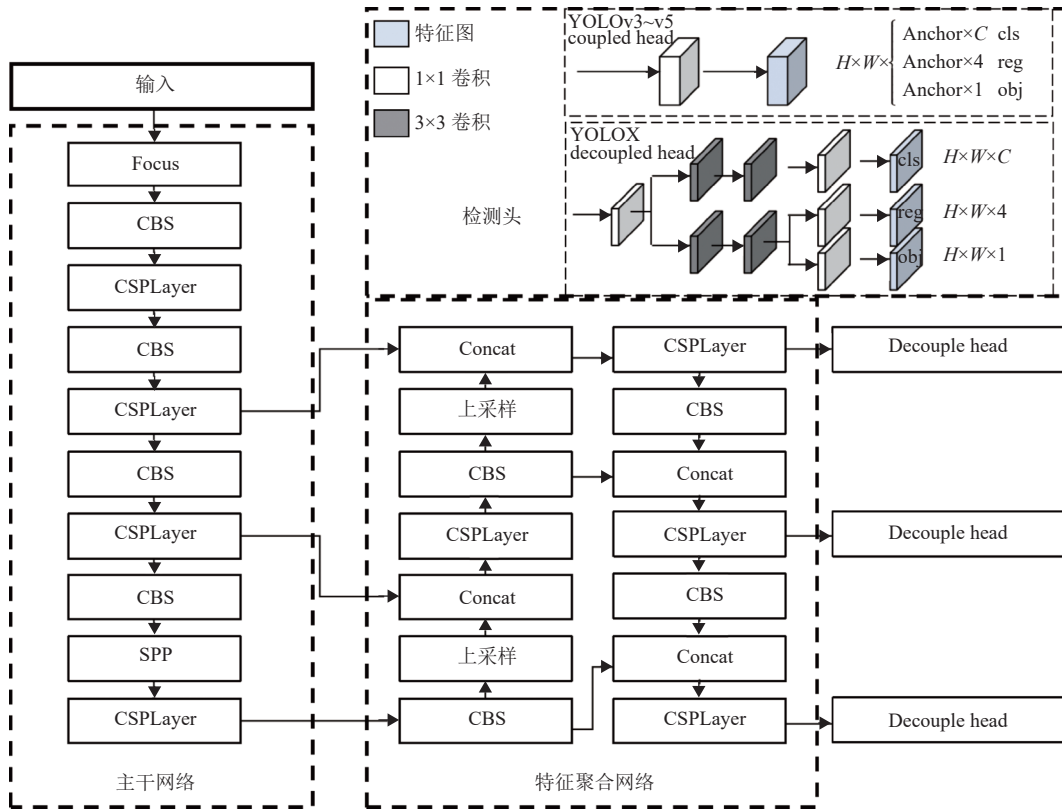


图 1 网络结构示意图

1.1 改进坐标注意力机制引入

CSPDarknet 主干网络的设计能够有效提取船只目标的细节信息，但设有对细节信息做出有效聚焦，忽略了能有效提升网络检测性能的目标空间结构信息与内部通道信息。因此引入注意力模块对原有网络结构进行改进，目前主流注意力机制模块有 SE^[9]、CBAM^[10]、CA^[11] 等，其中模块 CA 作为通道域的最新进展，对输入的特征图沿水平坐标与垂直坐标方向进行平均池化以对每个通道进行编码，此步骤相较于 SE 通道注意力模块而言对不同方向的空间信息进行了筛选，以嵌入到通道信息中，实现了空间信息的引入。进一步地，坐标注意力机制采用 1*1 卷积进行通道信息学习，该步骤结合了 ECANet 注意力机制^[12] 的思想，采用 1*1 卷积替代了原 SE 通道注意力模块中的全连接层，以提升注意力模块的运行效率与精度，但该模块对空间域信息特征的利用仍不够全面。因此，本文结合 BAM 注意力模块^[13] 的构造思想，将坐标注意力模块加以改进并引入到特征提取网络中，学习目标空间结构信息与通道信息，引导神经网络更有效地提取目标特征，提高网络对 SAR 图像船只目标的特征表达能力，实现近岸船只目标信息的有效提取，

改进后的坐标注意力模块如图 2 所示。

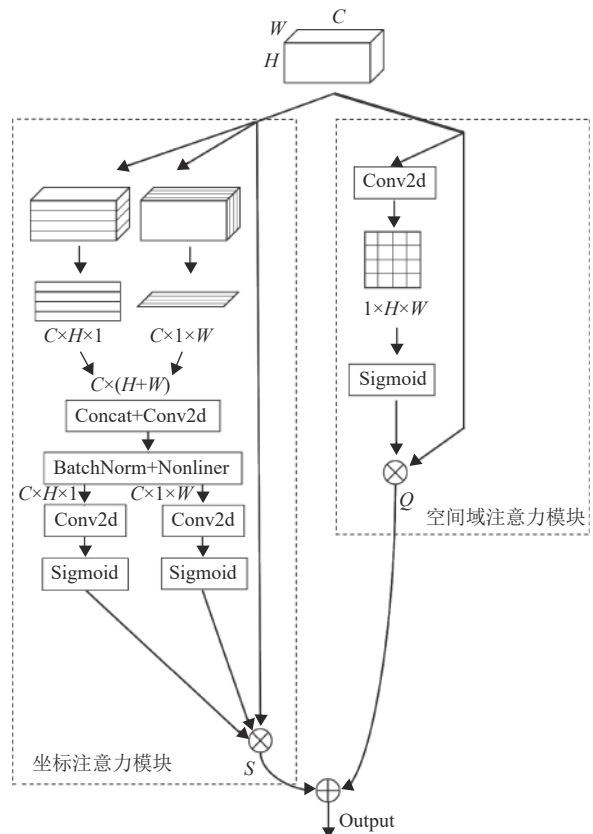


图 2 改进坐标注意力机制模块示意图

改进后的坐标注意力模块 BCA 通过结合空间域信息学习模块分支对输入维度为 $C \times H \times W$ 的特征图 m 进行空间域与通道域的特征聚焦, 原坐标注意力机制作为通道分支沿水平坐标与垂直坐标方向进行尺寸为 $(H, 1)$ 和 $(1, W)$ 的平均池化, 对每个通道进行编码, 得到两个一维向量, 当高度为 h 时, 向量第 c 个通道的输出如式 (1) 所示; 当宽度为 w 时, 第 c 个通道的输出如式 (2) 所示。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} m_c(h, i) \quad (1)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq i < H} m_c(i, w) \quad (2)$$

两个一维向量沿着空间维度进行特征聚合压缩通道, 而后级联生成两个特征图 n^h 和 n^w , 用大小为 1 的卷积核进行卷积, 如式 (3) 所示, 生成水平方向与垂直方向的空间信息特征图 t ; 再沿空间维度将特征图 t 分成两个单张量 t^h 和 t^w , 利用两个大小为 1 的卷积对张量进行卷积, 如式 (4) 和式 (5) 所示, 得到与输入特征图同样的通道数; 最后将经过激活函数运算后的值作为通道域注意力权重, 与输入特征图相乘, 得到权重调整后的特征图 S 如式 (6) 所示, 实现对特征图通道域信息的加强。

$$t = \delta(F_1([n^h, n^w])) \quad (3)$$

$$g^h = \sigma(F_h(t^h)) \quad (4)$$

$$g^w = \sigma(F_w(t^w)) \quad (5)$$

$$S_c(i, j) = m_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

式中, δ 为卷积运算; σ 为 sigmoid 激活函数运算。

空间分支则通过引入 1×1 卷积将输入维度为 $C \times H \times W$ 的特征图 m 进行通道挤压, 变化为 $1 \times H \times W$ 的空间特征图, 再将经过激活函数运算后的值作为空间域注意力权重, 与输入特征图进行相乘得到权重调整后的特征图 Q , 实现对特征图空间域信息的加强。最终将通道分支输出的加强特征图 S 与空间分支输出的加强特征图 Q 进行相加, 得到空间信息与通道信息融合的最终特征图。

改进前后的特征提取网络 AttCSPLayer 如图 3 所示, 在原始 CSPLayer 网络初步卷积提取产生的特征图后, 将改进坐标注意力机制作用于输出的特征图中, 根据通道域信息与空间域信息调整特征图的权重, 再进行各层级特征融合提取。

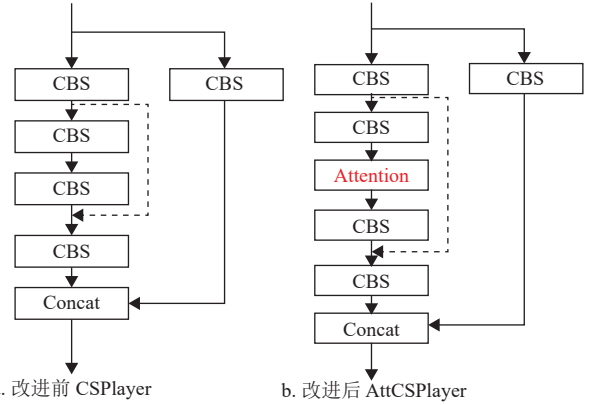


图 3 特征提取网络示意图

1.2 旋转目标框转化

基于水平框的船只检测在近岸情形下表现不佳, 且无法提供船只的方位角信息, 不利于近岸船只调度任务进行。为此本文采用旋转目标框进行目标检测, 增加角度分类头学习目标的角度信息, 使检测头能同时输出船只目标位置与方位角估计信息, 改进后的旋转检测头如图 4 所示。

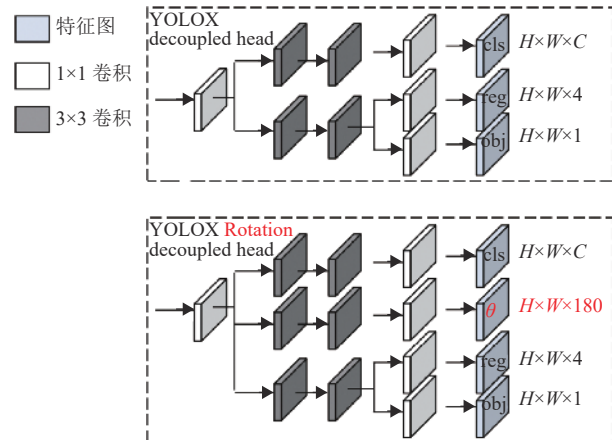


图 4 增加角度分类头的旋转检测头示意图

图 4 中旋转检测头的 reg 特征图为目标的位置信息, 其输出为 (x, y, w, h) , θ 特征图为目标的角度信息, 其输出为目标在 $0^\circ \sim 180^\circ$ 之间的概率, 取最大概率的角度作为预测角度值, 构成预测旋转目标框 (x, y, w, h, θ) , 通过式 (7) 将旋转目标框 (x, y, w, h, θ) 转化为二维高斯分布 (μ, Σ) , 以计算目标真实分布与模型预测分布之间的 KL 散度^[14] 作为损失值进行反向传播优化, 其中 KL 散度计算公式如式 (8) 所示, 特别地, 设目标角度值为 0, 计算位置信息的损失梯度, 可得式 (9)。

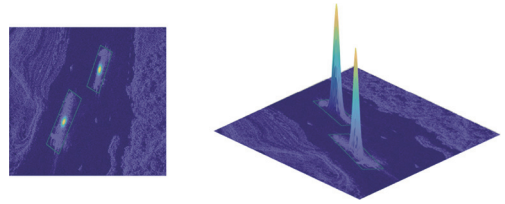
$$\begin{cases} \boldsymbol{\mu} = (x, y)^T \\ \boldsymbol{\Sigma}^{1/2} = \mathbf{R}\boldsymbol{\Lambda}\mathbf{R}^T = \\ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \frac{w}{2} & 0 \\ 0 & \frac{h}{2} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \end{cases} \quad (7)$$

$$D_{kl}(N_p \parallel N_t) = \frac{1}{2}(\boldsymbol{\mu}_p - \boldsymbol{\mu}_t)^T \boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\mu}_p - \boldsymbol{\mu}_t) + \frac{1}{2}\text{Tr}(\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\Sigma}_p) + \frac{1}{2}\ln\frac{|\boldsymbol{\Sigma}_t|}{|\boldsymbol{\Sigma}_p|} - 1 \quad (8)$$

$$\frac{\partial f_{kl}(\boldsymbol{\mu}_p)}{\partial \boldsymbol{\mu}_p} = \left(\frac{4}{w_t^2} \Delta x, \frac{4}{h_t^2} \Delta y \right) \quad (9)$$

式中, D_{kl} 为两分布间的 KL 散度距离; N_p 代表模型预测旋转框转化的二维高斯分布; N_t 代表真实数据旋转框转化的二维高斯分布; Tr 为求矩阵的迹。目标旋转框转化为二维高斯分布的示意图如图 5 所示, 通过式 (7) 将旋转目标框转化为二维高斯分布, 每个目标框转化的高斯分布受目标的长宽与角度影响, 大小形状不尽相同。由式 (9) 可知其梯度权重受 $1/w^2$ 和 $1/h^2$ 影响, 使得模型可根据目标的尺度来动态调整权重, 当目标尺度较小或纵横比较大时, 模型通过反向传播学习损失, 会增大相应方向的梯度, 便于有效地进行位置偏移的优化,

提高船只目标中心位置的定位精度。



a. 二维高斯分布俯视示意图 b. 二维高斯分布斜视示意图
图 5 旋转框转化二维高斯分布示意图

1.3 改进后的算法结构

引入改进坐标注意力机制与旋转目标框后的算法结构如图 6 所示, 在原有 YOLOX 网络结构基础上, 采用改进坐标注意力机制改进特征提取网络结构, 提高网络对船只目标信息的聚焦; 增加角度分类头学习目标的角度信息。在训练阶段读取数据集的目标旋转框信息 (x, y, w, h, θ) 并将其转化为二维高斯分布 $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 而后与旋转检测头输出的特征向量进行特征匹配, 得到与目标分布相近的预测分布信息, 通过计算预测分布与目标分布的 KL 散度作为旋转框损失值训练网络, 使网络模型根据目标的尺度与角度来动态调整位置偏移权重, 提升网络在近岸区域下对船只目标位置的定位精度, 实现高精度的船只目标检测。

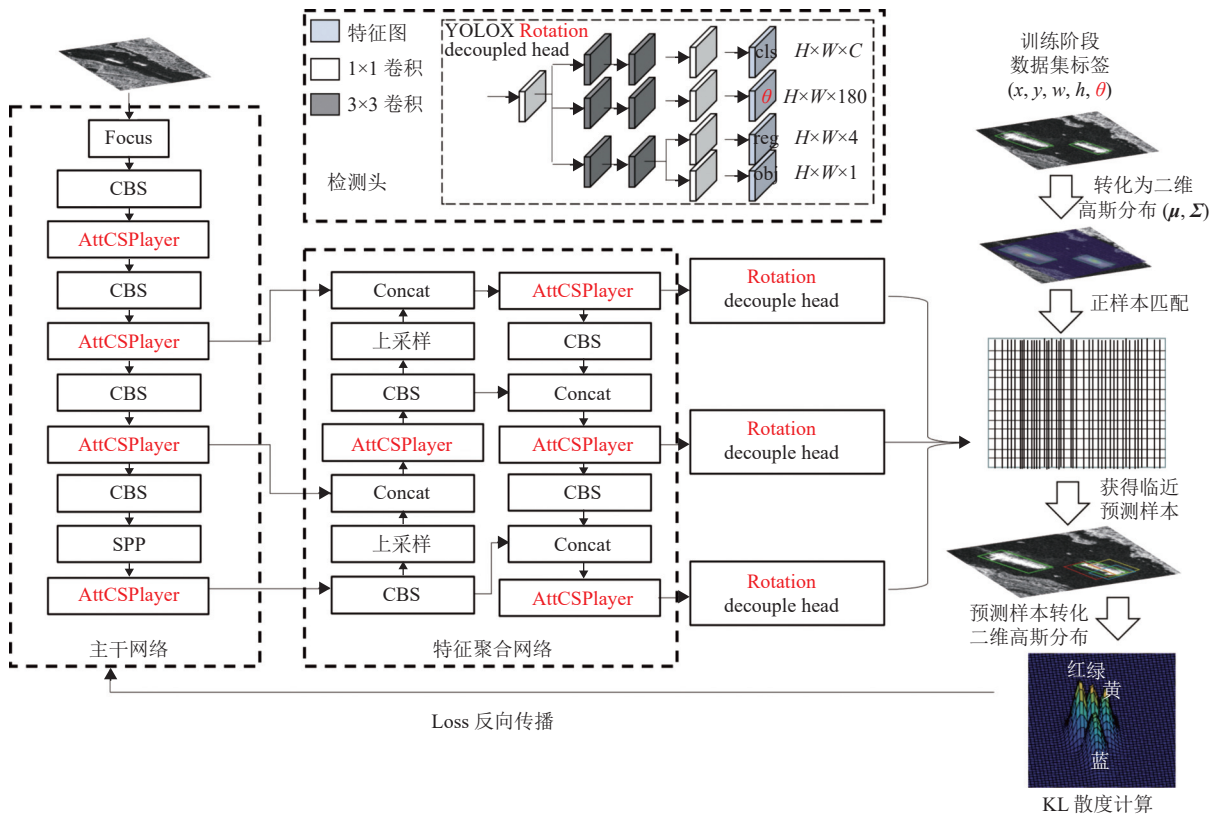


图 6 改进后的 YOLOX 算法结构示意图

2 实验结果与分析

2.1 数据集介绍与实验设置

本文使用公开的 official-ssdd 数据集^[15], 该数据集对近岸船只目标与远海船只目标进行了划分, 并提供了经过二次校验后的船只目标精确标注, 数据集样本采用 C 波段的 Radarsat-2、Sentinel-1 卫星以及 X 波段的 TerraSAR-X 卫星的 SAR 图像的裁剪得到数据, 共包含 1 160 张图像、2 587 只船舶目标, 其中 928 张图像、2 041 只船舶作为训练集, 232 张图像、546 只船舶作为测试集, 平均每张图像有 2.23 个船只, 包含 5×4 像素的小目标船只到 384×251 像素的大目标船只, 涉及了大片远海区域和近岸区域目标, 测试集共 172 只近岸区域船舶目标, 374 只远海区域船舶目标, 背景多样。

实验采用 Pytorch 深度学习框架, 基于 CUDA11.4 和 cuDNN8.2.4 加速训练, 基于 TensorRT 部署推理, YOLOX 根据不同的网络深度与宽度, 由浅到深可以分为 YOLOXnano、YOLOXtiny、YOLOXs 和 YOLOXm 等多个模型, 本文采用最轻量的模型 YOLOXnano 进行实验, 在此基础上进行改进并与其原始网络对比, 最后再跟其他一些改进方法进行测试对比。输入到模型的图像大小设为 416×416 , batch size 大小为 128, 训练优化器采用 Adam 优化器更新网络参数, 其中学习率 (learning rate) 和权值衰减 (weight decay) 参数设为 0.001 和 0.000 5, 旋转目标框损失函数采用 KLD_Loss, 损失权重为 5, 目标角度损失采用 CSL_Loss^[16], 损失权重为 0.1, 置信度损失采用 BCE_Loss, 损失权重为 1; 计算机具体配置为 Intel® Xeon® E52640, 2.40 GHz, 显卡为 2*P5000 [Pascal Quadro], 显存 2×16 GB; 部署平台选用嵌入式平台 (NVIDIA Jetson AGX Xavier), 显存 32 GB; 算法的训练和测试均在 GPU 加速下完成。

2.2 实时数据增强

深度学习需要足够的训练样本作为支撑, 相较于光学图像, SAR 图像数据难以大量获取, 为了充分利用有限的训练数据, 加强模型泛化能力, 抑制过拟合现象^[17], 本文通过随机缩放、随机旋转、mosaic 拼接、mixup 变化等方法对训练数据进行实时数据增强, 效果如图 7 所示。

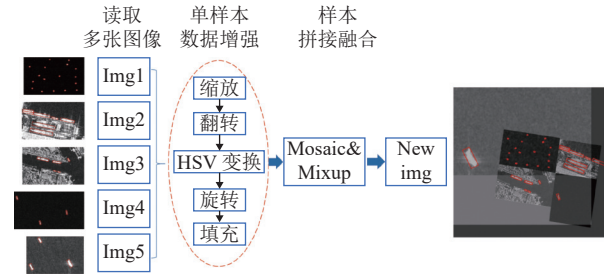


图 7 实时数据增强示意图

2.3 实验结果与分析

本文采用平均精度 (average precision, AP) 作为 SAR 图像船只检测模型性能的评价指标, 计算式为:

$$\begin{cases} P = \frac{TP}{TP + FP} = \frac{TP}{\text{所有检测结果}} \\ R = \frac{TP}{TP + FN} = \frac{TP}{\text{所有真实结果}} \end{cases} \quad (10)$$

$$AP = \int_0^1 PRdR$$

式中, TP 为模型正确检测的船只目标数量; FP 为误检船只目标数量; FN 为漏检船只目标数量; 精确率 P 表示正确检测的船只目标数量在全部检测结果中的比率; 召回率 R 表示正确检测的船只目标数量占所有船只目标真实数量的比率。以 P 为纵轴, R 为横轴, 绘制出 PR 曲线, 曲线与横坐标围成的面积为 AP 值, Ap_{50} 为交并比阈值设为 0.5 时的 AP 值。

利用本文算法在 official-ssdd 数据集上进行检测, 对采用不同改进方案的 YOLOX 检测算法结果进行对比, 以检测精度 Ap_{50} 作为算法检测精度性能指标, 实验结果如表 1 所示。表中 YOLOX 为原始网络不加任何改进的检测结果; CA_YOLOX 为引入了坐标注意力机制到特征提取网络中进行改进的 YOLOX 算法; BCA_YOLOX 为引入了改进后坐标注意力机制到特征提取网络中进行改进的 YOLOX 算法; R_YOLOX 为引入了旋转框方法到检测头中进行改进的 YOLOX 算法; 本文方法则是结合了改进坐标注意力机制与旋转框的 YOLOX 算法。从表 1 可以看出, 引入坐标注意力机制后算法模型的精确率有所提升, 召回率差异减少, 误检严重问题得到改善, 引入改进坐标注意力机制后算法模型的精确率进一步提升。引入旋转框后算法模型在近岸区域的精确率有所提升, 召回率基本保持一致, 能更有效地定位出目标的正确位置。通过结合改进坐标注意力机制与旋转框方法, 本文方法在近

岸区域的检测精度高于其他算法, 拥有更好的检测性能, 在 official-ssdd 数据集近岸区域检测的 Ap50 值相较于 YOLOX 算法提升了 18.77%, 引入注意

力机制与旋转框方法都提高了近岸区域下船只目标的检测平均精度, 同时引入改进坐标注意力机制与旋转框这两种改进方法后, 网络的检测性能最好。

表 1 SAR 船只检测结果评价指标

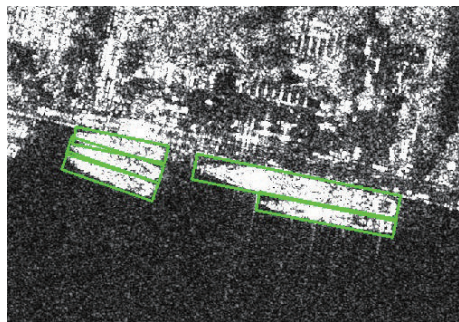
| 算法 | 近岸区域 | | | 远海区域 | | | 总计 |
|-----------|----------------|---------|----------------|---------|---------|---------|----------------|
| | P/% | R/% | Ap50/% | P/% | R/% | Ap50/% | Ap50/% |
| YOLOX | 0.668 6 | 0.912 7 | 0.659 3 | 0.958 2 | 0.981 3 | 0.979 2 | 0.878 8 |
| CA_YOLOX | 0.802 3 | 0.926 2 | 0.798 4 | 0.992 0 | 0.989 3 | 0.990 7 | 0.930 6 |
| BCA_YOLOX | 0.854 7 | 0.803 3 | 0.826 2 | 0.981 3 | 0.960 1 | 0.979 0 | 0.935 1 |
| R_YOLOX | 0.744 2 | 0.748 5 | 0.733 3 | 0.941 3 | 0.943 9 | 0.937 4 | 0.872 8 |
| 本文算法 | 0.848 8 | 0.979 9 | 0.847 0 | 0.989 3 | 0.986 7 | 0.987 0 | 0.942 6 |

为验证本文所提检测网络的性能, 在 official-ssdd 数据集上完成本文方法与 YOLOv3^[18]、YOLOv5^[19] 和 Faster R-CNN 算法的对比实验, 结果如表 2 所示。从表中可以看出, 本文方法相较除 YOLOX 以外的其他算法, 模型计算参数最少, 总体测试检测精度最优, 网络推理延时最低, 且能获取船只目标方位角信息, 满足实时处理需求。

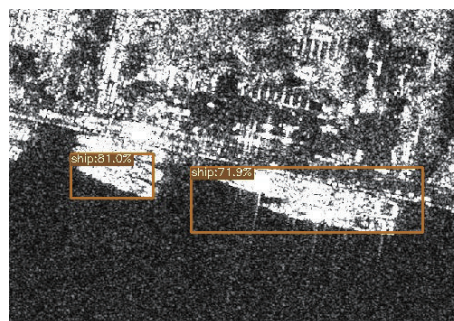
表 2 不同算法检测性能对比

| 算法 | 参数量×10 ⁶ | Ap50/% | 网络推理耗时/ms |
|-------------|---------------------|--------------|-----------|
| YOLOX | 0.91 | 0.879 | 8 |
| YOLOV3 | 36.3 | 0.871 | 28 |
| YOLOV5 | 7.2 | 0.901 | 30 |
| Faster RCNN | 272.7 | 0.890 | 119 |
| 本文算法 | 1.14 | 0.942 | 10 |

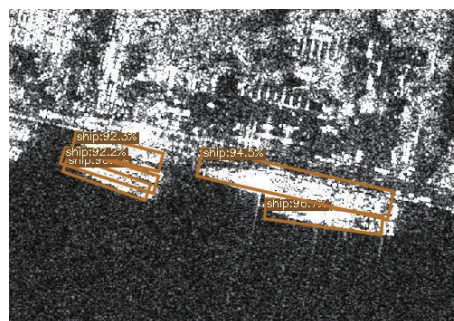
本文实验检测结果如图 8 所示。对于紧密排列的邻岸大目标, 传统 YOLOX 算法会出现漏检现象, 或如图 8b 所示只输出一个大的目标框, 而本文算法则能有效地检测并区分邻近排列的船只, 如图 8c 所示。



a. 真值图



b. YOLOX 检测结果



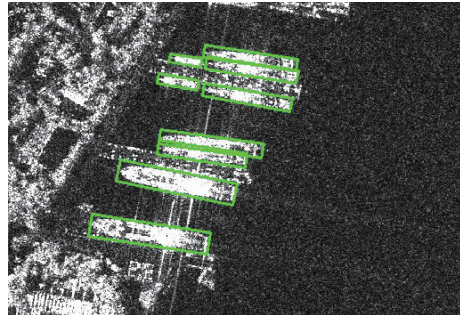
c. 本文算法检测结果

图 8 邻岸大目标检测对比示意图

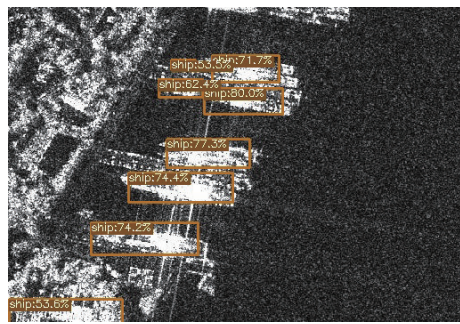
对于在港口中停靠的船只 (图 9a), 采用传统 YOLOX 算法进行检测时, 出现了误检与漏检现象, 将陆地建筑误检为船只目标, 对港口停靠紧密的船只则只输出一个较大的目标框, 无法有效区分排列紧密的船只目标, 如图 9b 所示。而本文方法则有效检测并区分各船只目标, 没有错检现象, 如图 9c 所示。

从图 10 可以看出, 传统 YOLOX 算法采用水平框检测, 船只越大水平框引入的背景信息越多, 模型对大船只目标的关注度不够, 导致输出预测框

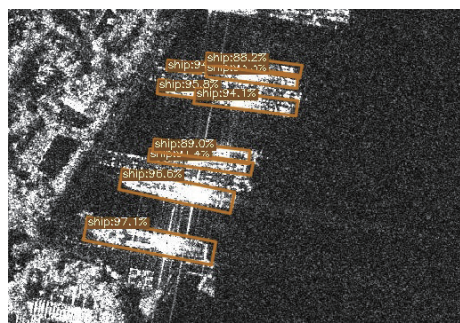
置信度较低。本文算法预测框置信度明显提升, 且预测框可以更精确地涵盖目标区域, 输出更精确的目标位置信息, 检测更为准确。



a. 真值图

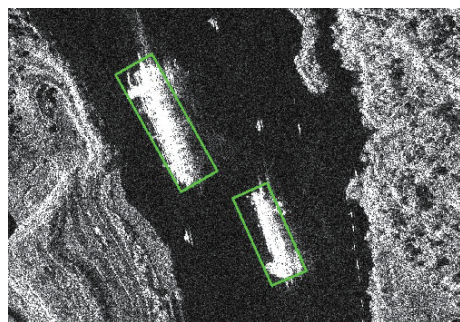


b. YOLOX 检测结果

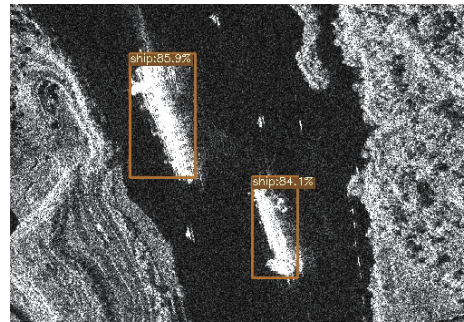


c. 本文算法检测结果

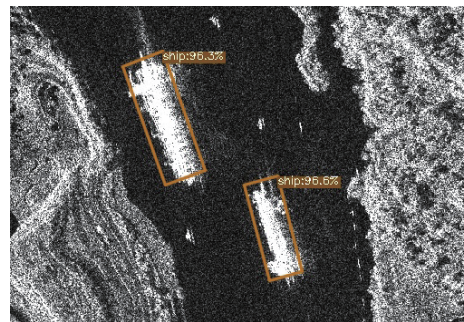
图 9 港口大目标检测对比示意图



a. 真值图



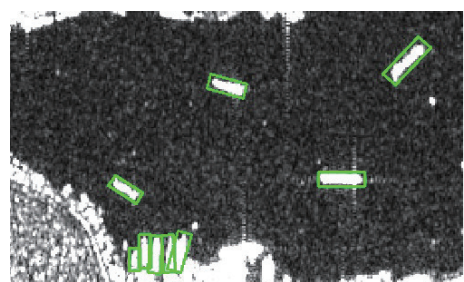
b. YOLOX 检测结果



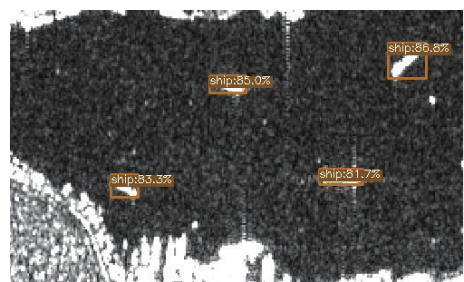
c. 本文算法检测结果

图 10 河道大目标检测对比示意图

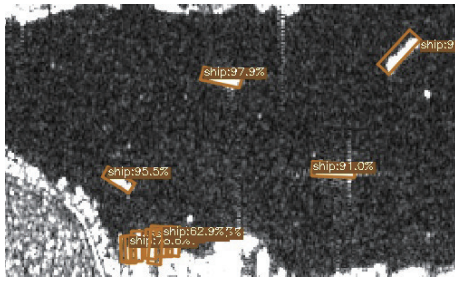
对于近岸小目标, 传统 YOLOX 算法会出现漏检现象, 如图 11 所示, 难以检测出停靠在岸边的小目标船只。而本文方法则能对近岸小目标进行有效检测, 但由于近岸场景下, 近岸建筑产生的强散射点与临近的船只特征相近, 在特征提取的过程中引入背景干扰信息过多, 导致近岸区域小目标的检测置信度较低, 仅在 50%~60% 之间。



a. 真值图



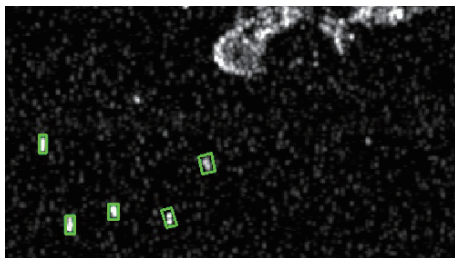
b. YOLOX 检测结果



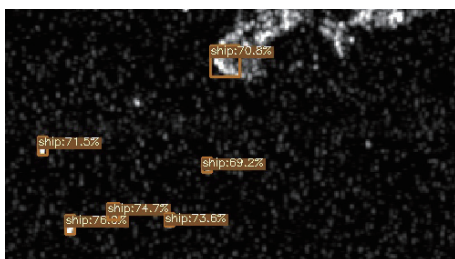
c. 本文算法检测结果

图 11 近岸小目标检测对比示意图

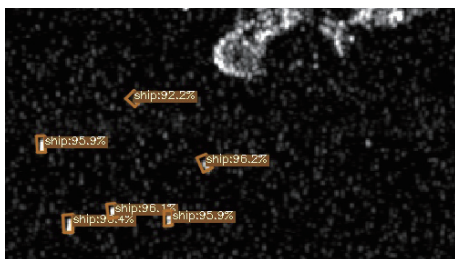
在海岛周围区域，从图 12 可以看出，传统 YOLOX 算法将岛屿误检为船只目标，而本文方法则是将海面杂波散射点误检为船只目标。这是由于本文方法采用旋转框检测，对异于周围海面背景下的杂波散射点更为敏感，导致误检现象。而传统 YOLOX 采用水平框检测，在周围背景干扰下对海面杂波散射点的关注度降低，其对船只目标输出的预测框置信度也较低，在检测阈值设为 0.5 的情况下不会将杂波散射点视为船只目标，但会将区域亮度信息较为明显的岛屿错检为船只目标。



a. 真值图



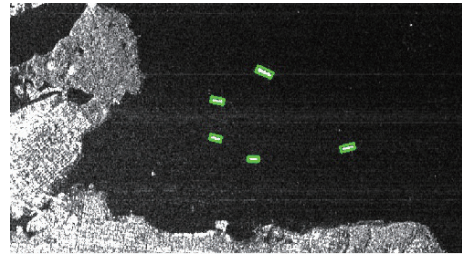
b. YOLOX 检测结果



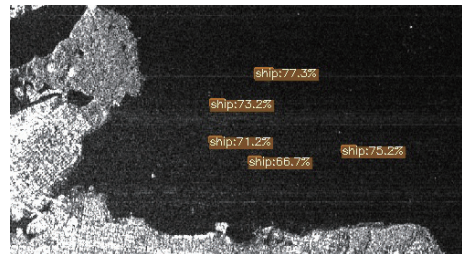
c. 本文算法检测结果

图 12 邻岛小目标检测对比示意图

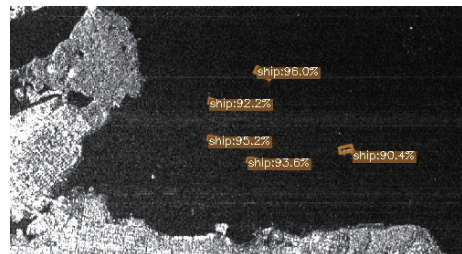
从图 13 也可以看出，本文方法对异于周围海面背景下的散射信息更为敏感，将海面杂波散射点误检为船只目标，而传统 YOLOX 算法对海面散射信息的关注度较低，因此未产生误检现象。在近海区域，本文方法预测框置信度较传统 YOLOX 算法有明显提升。



a. 真值图



b. YOLOX 检测结果



c. 本文算法检测结果

图 13 近海小目标检测对比示意图

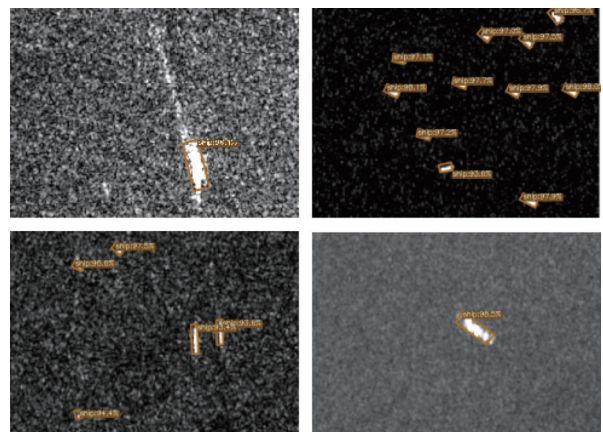


图 14 远海船只目标检测结果示意图

本文方法在远海场景不同海况情形下检测效果

如图 14 所示, 可以看到, 对于不同海况情形下的远海场景目标, 本文方法依旧可有效进行检测。综上所述, 对于 SAR 船只目标检测任务, 本文方法在近岸和远海场景下都能取得良好的效果, 能够在复杂背景下识别船只目标, 并且能更好地处理多尺度问题, 得到更加精准的检测框。

3 结束语

本文提出了基于改进坐标注意力机制与旋转目标框的单阶段检测网络方法。在构建特征提取网络时加入改进坐标注意力机制准确提取感兴趣区域, 同时选用旋转目标框获得船只方位角信息, 提高检测精度。经 official-ssdd 数据集实验验证, 改进的 YOLOX 检测算法的检测精度在近岸情形下较 YOLOX 模型提升了 18.77%, 与主流 YOLOv5 检测算法相比, 本文模型计算参数量减少了 84.2%, 推理耗时缩减至 1/3, 总体检测精度提升了 4.1%, 满足实时处理需求, 验证了本方法的有效性。本文方法虽较基于水平框检测方法能得到更精准的目标框, 但对 SAR 图像近岸场景中的大船只目标中心定位仍有待进一步提升。后续仍需进一步扩充数据集, 优化网络结构, 减小网络对近岸船只的定位偏移量。

参 考 文 献

- [1] 孙显, 王智睿, 孙元睿, 等. AIR-SARShip-1.0: 高分辨率 SAR 舰船检测数据集[J]. 雷达学报, 2019, 8(6): 852-862.
SUN X, WANG Z R, SUN Y R, et al. AIR-SARShip-1.0: High-Resolution SAR ship detection dataset[J]. *Journal of Radars*, 2019, 8(6): 852-862.
- [2] 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. [S.l.]: IEEE, 2016: 779-788.
- [3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]//European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [4] 徐英, 谷雨, 彭冬亮, 等. 面向合成孔径雷达图像任意方向舰船检测的改进 YOLOv3 模型[J]. 兵工学报, 2021, 42(8): 1698-1707.
XU Y, GU Y, PENG D L, et al. Improved YOLOv3 model for ship detection of synthetic aperture radar images in any direction[J]. *Acta Armamentarii*, 2021, 42(8): 1698-1707.
- [5] 杨杰. 基于深度学习的 SAR 舰船目标检测方法研究[D]. 秦皇岛: 燕山大学, 2021.
YANG J. Research on SAR ship target detection method based on deep learning[D]. Qinhuangdao: Yanshan University, 2021.
- [6] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [7] 周雪珂, 刘畅, 周滨. 多尺度特征融合与特征通道关系校准的 SAR 图像船舶检测[J]. 雷达学报, 2021, 10(4): 531-543.
ZHOU X K, LIU C, ZHOU B. Multi-Scale feature fusion and feature channel relationship calibration for SAR image ship detection[J]. *Journal of Radars*, 2021, 10(4): 531-543.
- [8] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. (2021-08-06). <https://arxiv.org/abs/2107.08430>.
- [9] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 7132-7141.
- [10] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: [s.n.], 2018: 3-19.
- [11] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 13713-13722.
- [12] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 11531-11539.
- [13] PARK J, WOO S, LEE J Y, et al. BAM: Bottleneck attention module[EB/OL]. [2022-01-23]. <https://arxiv.org/abs/1807.06514>.
- [14] YANG X, YANG X J, YANG J R, et al. Learning high-precision bounding box for rotated object detection via kullback-leibler divergence[J]. *Advances in Neural Information Processing Systems*, 2021, 34: 18381-18394.
- [15] ZHANG T W, ZHANG X L, LI J, et al. Sar ship detection dataset (ssdd): Official release and comprehensive data analysis[J]. *Remote Sensing*, 2021, 13(18): 3690.
- [16] YANG X, YAN J C. Arbitrary-Oriented object detection with circular smooth label[C]//European Conference on Computer Vision. Cham: Springer, 2020: 677-694.
- [17] 阮晨, 郭浩, 安居白. 复杂背景下 SAR 近岸舰船检测[J]. 中国图象图形学报, 2021, 26(5): 1058-1066.
RUAN C, GUO H, AN J B. SAR nearshore ship detection under complex background[J]. *Journal of Image and Graphics*, 2021, 26(5): 1058-1066.
- [18] REDMON J, FARHADI A. YOLOv3: An incremental improvement[EB/OL]. [2022-01-20]. <https://arxiv.org/abs/1804.02767>.
- [19] ZHU X K, LYU S C, XU W, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2021: 2778-2788.