

特征图强化网络：利用特征图强化船舶检测模型训练的网络结构

FMRNet: A network structure for enhanced ship detection model training using feature maps

张泽琨;谭震彪;余坤;方斌;黄骁;马杰

引用本文:

张泽琨, 谭震彪, 余坤, 方斌, 黄骁, 马杰. 2023. 特征图强化网络：利用特征图强化船舶检测模型训练的网络结构. 遥感学报, 27 (12): 2697-2705 DOI: 10.11834/jrs.20221656.

Zhang Z K, Tan Z B, Yu K, Fang B, Huang X and Ma J. 2023. FMRNet: A network structure for enhanced ship detection model training using feature maps. National Remote Sensing Bulletin, 27 (12): 2697-2705 DOI: 10.11834/jrs.20221656.

在线阅读 View online: <http://www.ygxb.ac.cn/thesisDetails?columnId=23459820>

您可能感兴趣的其他文章

Articles you may be interested in

移动地理信息系统技术发展的3个时代

Three development stages of mobile geographic information system technology

遥感学报. 2022, 26(12):2399-2410.

遥感知识图谱创建及其典型场景应用技术

Remote sensing knowledge graph construction and its application in typical scenarios

遥感学报. 2023, 27(2):249-266.

新一代人工智能驱动的陆地观测卫星遥感应用技术综述

A review of land observation satellite remote sensing application technology with new generation artificial intelligence

遥感学报. 2022, 26(8):1530-1546.

人工智能时代的遥感变化检测技术：继承、发展与挑战

Remote sensing change detection technology in the Era of artificial intelligence: Inheritance, development and challenges

遥感学报. 2023, 27(9):1975-1987.

全球湿地遥感研究综述：1975年—2020年

Review of global studies on the remote sensing of wetlands from 1975 to 2020

遥感学报. 2023, 27(6):1270-1280.

# 特征图强化网络:利用特征图强化船舶检测模型训练的网络结构

张泽琨<sup>1</sup>, 谭震彪<sup>1</sup>, 余坤<sup>3</sup>, 方斌<sup>1</sup>, 黄骁<sup>2</sup>, 马杰<sup>1</sup>

1. 华中科技大学 人工智能及自动化学院 多谱信息技术处理国家重点实验室, 武汉 430072;

2. 中国舰船研究设计中心, 武汉 430064;

3. 中铁第四勘察设计院集团有限公司, 武汉 430063

**摘要:** 随着人工智能技术的进一步发展, 深度学习方法在船舶检测领域发挥着重要作用。然而, 深度学习方法出现的虚警和漏检, 对船舶检测领域技术的应用存在一定的阻碍。虽然经典的深度学习方法能够有效处理单一背景的海面, 但是当面对复杂背景之下的数据时, 经典模型很容易得出岸上的虚警。并且在常规训练中, 模型常常对一些显著特征过于关注, 出现特征过拟合现象, 当这些显著特征发生改变时极易出现漏检。在模型对输入进行前向传播的过程中, 模型中不同网络层会对输入生成对应的映射, 也就是特征图。充分利用特征图的语义信息和空间信息是一种有效减少虚警和漏检的方法。与传统模型相比, 我们提出的特征图强化网络可以充分利用特征图生成自适应特征图掩码与水陆分割掩码, 避免模型的特征过拟合与削弱复杂背景造成的影响, 最终达到减少虚警与漏检的目的。在现有公共数据集上与算法模型的对比实验结果表明, 本文所提出的方法的性能更为出色, 超过了其他SOTA算法。

**关键词:** 遥感成像, 人工智能, 船舶目标检测, 神经网络, 特征图, 网络强化, 旋转目标检测, 抑制过拟合

**中图分类号:** P2

**引用格式:** 张泽琨, 谭震彪, 余坤, 方斌, 黄骁, 马杰. 2023. 特征图强化网络: 利用特征图强化船舶检测模型训练的网络结构. 遥感学报, 27(12): 2697-2705

Zhang Z K, Tan Z B, Yu K, Fang B, Huang X and Ma J. 2023. FMRNet: A network structure for enhanced ship detection model training using feature maps. National Remote Sensing Bulletin, 27(12): 2697-2705 [DOI: 10.11834/jrs.20221656]

## 1 引言

近年来为了达到对海洋资源的进一步开发的目的, 领海内的船舶信息获取技术越来越受到相关机构的重视。卫星遥感信息是海洋中船舶信息的重要来源, 如何有效地在遥感图像上准确检测船舶的类型和位置是一项关键的任务 (Goetz 等, 1983; Lunetta 等, 1991; Gu, 2010)。早期的方法是传统图像处理方法, 例如许开宇等 (2006) 提出基于 Canny 算子的方法进行定位, 然而这些算法的效果并不令人满意。

随着人工智能技术的发展, 以卷积神经网络为代表的机器学习技术逐渐取代传统图像处理方法,

成为图像处理任务的主流, 例如 VGG (Simonyan 和 Zisserman, 2015)。随着卷积神经网络结构的优化与发展, 其性能得到了极大地提高, 在图像处理任务上得到了不俗的效果。例如 He 等 (2016) 设计的 Resnet 由于良好的特征提取能力, 在公开数据集取得了不错的效果。这时卷积神经网络为基础诞生了许多单阶段的算法, 如 Redmon 等 (2016) 提出的 YOLOv1, Redmon 等 (2018) 提出的 YOLOv2, Redmon 等 (2018) 提出的 YOLOv3。通过利用这些卷积神经网络, 一些双阶段的算法, 如 Girshick (2015), 设计的 Fast-rcnn, 也被发明并且广泛用于检测任务。

近年来, 通过充分利用模型中网络层生成的

收稿日期: 2021-10-20; 预印本: 2022-03-01

基金项目: 国家自然科学基金 (编号: U1913602)

第一作者简介: 张泽琨, 研究方向为遥感条件下船舶检测的理论和应用。E-mail: m202072950@hust.edu.cn

通信作者简介: 马杰, 研究方向为导航与制导、目标检测与识别。E-mail: majie@hust.edu.cn

特征图, 注意力机制为卷积神经网络的性能提升做出了重大的贡献。Hu 等 (2018) 提出了一个挤压激活模块 (SE), 通过学习通道之间的权重对应关系, 利用权重来重新增强模型的特征图。这些通过 SE 模块挤压与激活生成的特征图, 使模型对于存在潜在目标的区域更加敏感。杨星等 (2021) 依照这个原理, 设计光谱空间注意力模块从原始数据中增强有用特征, 抑制无效特征, 来提升模型的性能。Bi 等 (2019) 在神经网络中嵌入了一个视觉注意力增强网络来提取视觉特征, 提升模型准确率。在目标检测方面, Nie 等 (2020) 充分利用了注意力机制模块, 将其引入 Mask-rcnn, 使算法可以在像素水平上准确地检测和分割目标船只。在复杂的背景下, 一个 SE 模块使模型在处理特征图时, 能够更快、更准确地计算得出目标可能出现的区域。

尽管现阶段算法在船舶检测中大多情况下取得了不错的效果, 但是检测网络的漏检和误报仍然是目标检测领域的一个重要问题。本文认为船舶检测任务中漏检和误报主要与陆地环境和特征过拟合有关。首先, 当环境从简单的水面, 变为海岸线或港口等涉及复杂陆地环境时, 神经网络模型就容易受到这些复杂背景环境的影响。在这种情况下模型容易从陆地区域中获得与船舶相似的错误特征, 最终导致了虚警。其次, 在模型训练过程中, 模型对同一输入的注意力会逐渐集中在特征图中的一些独特特征上, 通过这些显著的特征进行船舶的检测与定位。这种处理方法导致模型过于依赖某些突出的特征, 对容易关注的区域进行特征过拟合, 导致模型停止梯度更新, 陷入局部最优问题。如果一些独特的特征有微小的变化, 模型很可能无法正确识别这些变化的特征, 造成漏报的错误。如果背景中出现了相似的特征, 模型大概率会引入错误的结果, 造成虚警。受 Ding 等 (2017)、Ding 等 (2018) 等在目标检测的特征工作启发, 本文从特征图的角度入手, 提出了特征图强化网络 FMRNet (Feature Map Reinforcement Network)。该网络利用自适应特征图掩码模块减少特征过拟合的同时, 并联特征图海陆分割模块减弱复杂背景对模型的影响, 达到提高准确和召回的目的。它包含一个自适应特征图掩码模块 SFMM (Self Feature-map Mask Module) 和一个特征图海陆分割模块 FSSM (Feature-map Sea-land Segmentation

Module)。在现有公共数据集上与其他算法模型的对比实验结果表明, FMRNet 的性能更为出色, 在 HRSC2016 数据集中得到了 90.7% 的 mAP, 比原算法提高了 4.5%, 并且超过了其他 SOTA 算法。

## 2 研究方法或原理

特征图 (图 1) 展现了输入在模型中的状态变化, 直观地反应了模型在输入中关注的特征。从特征图的角度入手, 本文设计了一种利用特征图进行训练增强的骨干网络解决特征过拟合与复杂背景的难点。

### 2.1 特征图增强网络

图 1 完整展示了特征图增强网络 FMRNet 的训练结构。本文提出的 FMRNet, 包含自适应特征图掩码模块 (SFMM) 和特征图海陆分割模块 (FSSM), 它利用特征图进行遥感图像的特征图强化训练。其主干网络是经典的 Resnet50, 我们对它的结构进行改进。它包含了四层网络层 ( $Layer_1$ ,  $Layer_2$ ,  $Layer_3$ ,  $Layer_4$ ), 并且每层网络层会生成特定的特征图 ( $Feature_1$ ,  $Feature_2$ ,  $Feature_3$ ,  $Feature_4$ )。在图 1 中展示了不同网络层的特征图可视化, 浅层次的特征主要关注位置信息, 深层次的特征则更注重于语义信息。我们把特征图增强网络生成的特征图称作 *Output*, 它的维度是 ( $H$ ,  $W$ ,  $C$ ),  $C$  代表特征图的通道数,  $W$  和  $H$  分别代表特征图的宽和高。

每次训练的迭代时, 把维度为 ( $H_0$ ,  $W_0$ ,  $C_0$ ) 的输入送进特征图增强网络进行特征提取最终生成输出特征图  $Output_0$ 。它的维度与输出一样。考虑到分辨率, FMRNet 选取了 ( $Layer_1$ ,  $Layer_2$ ,  $Layer_3$ ) 所生成的特征图。我们把这些特征图送入自适应特征图掩码模块获得一个自适应特征图掩膜 AFM, 它的维入度是 ( $H_0$ ,  $W_0$ )。AFM 由特征图决定, 将与输入相乘生成一个新的  $Input'$ , 维度是 ( $H_0$ ,  $W_0$ ,  $C_0$ )。  $Input'$  的每个像素值计算方法如下:

$$Input'_{ijk} = Input_{0ijk} \times AFM_{ij} \quad (1)$$

式中,  $i \in (0, W_0)$ ,  $j \in (0, H_0)$ ,  $k \in (0, C_0)$ 。由于 AFM 的存在, 该模型在  $Input'$  上不再仅仅关注输入特定特征, 从而寻找其他可关注的点, 得到  $Output_1$ 。  $Input'$  经过 ( $Layer_1$ ,  $Layer_2$ ,  $Layer_3$ ), 生成与  $Output_0$  拥有相同的维度大小的  $Output_1$ 。最后通过融合  $Output_1$  与  $Output_0$  得到特征  $Output'$ :

$$Output' = \alpha \times Output_0 + (1 - \alpha) \times Output_1 \quad (2)$$

式中,  $\alpha$  控制两个特征参与特征融合的比例。由于保持了特征的多样性, 这种融合特征可以防止模

型过于依赖  $Output_1$  或者  $Output_0$ 。SFMM 生成自适应掩码并且进行梯度更新。

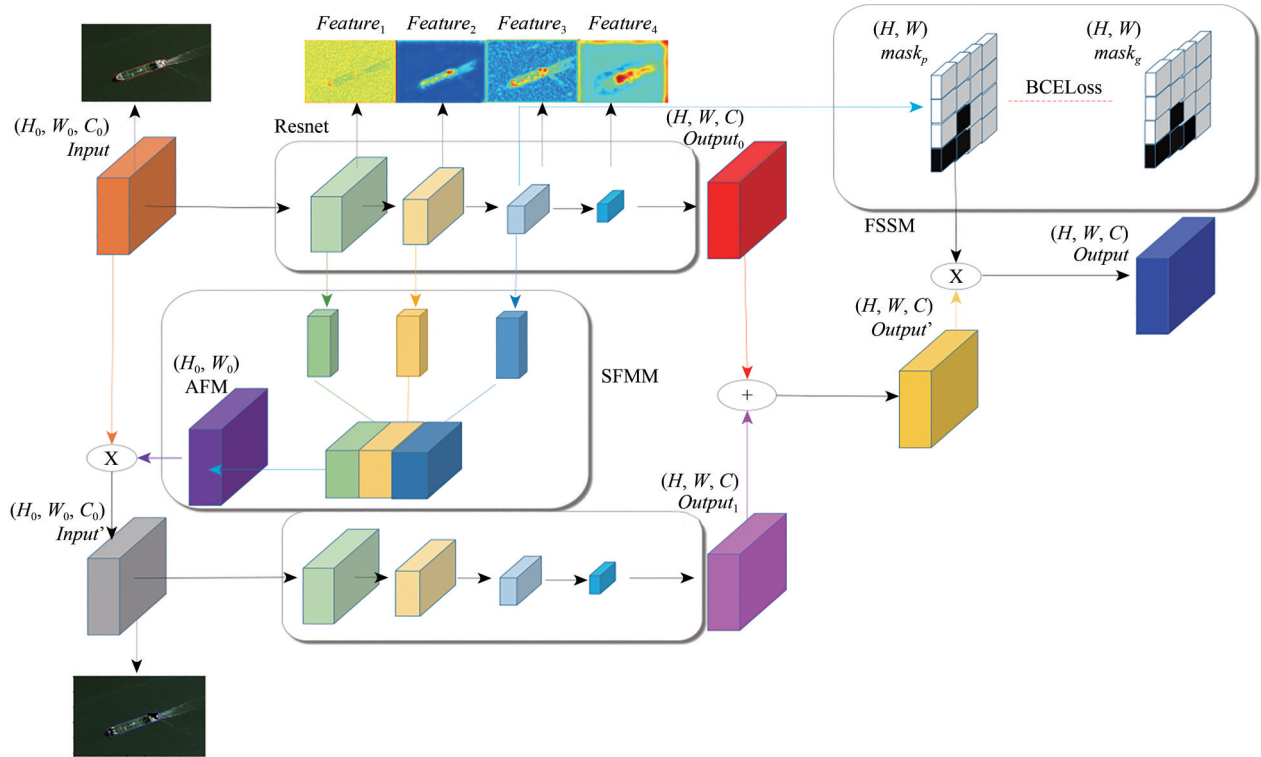


图1 特征图增强网络(FMRNet)的训练结构

Fig.1 The training structure of FMRNet

并联 FSSM 模块的任务是利用  $Feature_3$  生成一个海陆分割掩码。与  $Feature_1$ ,  $Feature_2$  相比,  $Feature_3$  含有更加丰富的语义信息。与  $Feature_4$  相比  $Feature_3$  拥有更多的像素, 位置信息与原图更加符合。因此  $Feature_3$  兼顾语义信息与位置信息, 能生成高质量的海陆分割掩码。

该分割掩码的取值范围为 0—1, 表示每个像素属于海面的可能性。之后将分割掩码与  $Output'$  相乘, 得到  $Output$ 。如果一个区域属于海面的概率很低, FSSM 将抑制对该区域的预测。 $Output$  是由 FMRNet 生成的特征图, 将用于获得检测结果。

当进行推理任务时, SFMM 不参与处理,  $Output'$  等于  $Output_0$ 。结构图如图 2 所示。

## 2.2 自适应特征图掩码模块

在神经网络训练过程中, 模型的“关注”将会逐渐集中在特征图中的一些独特的特征上, 导致模型过分关注此类特征从而陷入局部最优解。我们提出的 SFMM 可以在训练时将模型不同尺度的特征图

根据标签中大小目标的样本分布进行融合, 生成自适应掩码对模型在特征图中容易关注的特征进行遮掩, 防止模型过于关注这些特征造成过拟合。

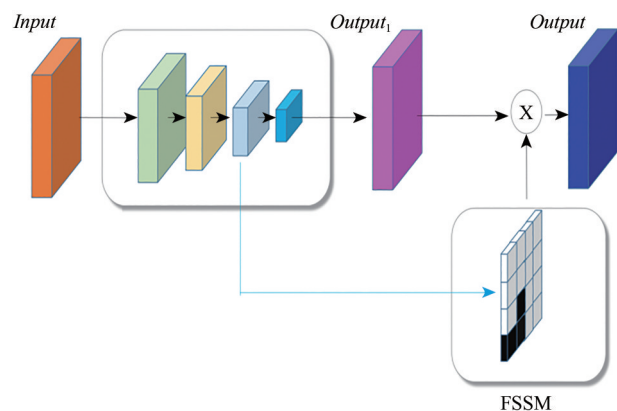


图2 特征图增强网络(FMRNet)的预测结构

Fig.2 The inference structure of FMRNet

SFMM 的结构如图 3 所示。所选中来自 3 个网络层的 3 个特征图都被送入 SFMM, 首先每个特征图都被上采样到  $(H_0, W_0, C_0)$  尺寸, 并由一个



编码器 Encoder 进行编码。每个特征图都对应于一个特定的编码器，该编码器包含卷积层 conv、池化层 pooling 和全连接层 FC。所有的编码器都有相同的结构，但有各自的参数，被用来从不同

的尺度上提取特征。每个编码器将特征图分别转换为 3 个向量，然后堆叠成一个向量，最后被送到全连接层，生成一个 (1, 3) 维度注意力权重 *Attention Weight*。

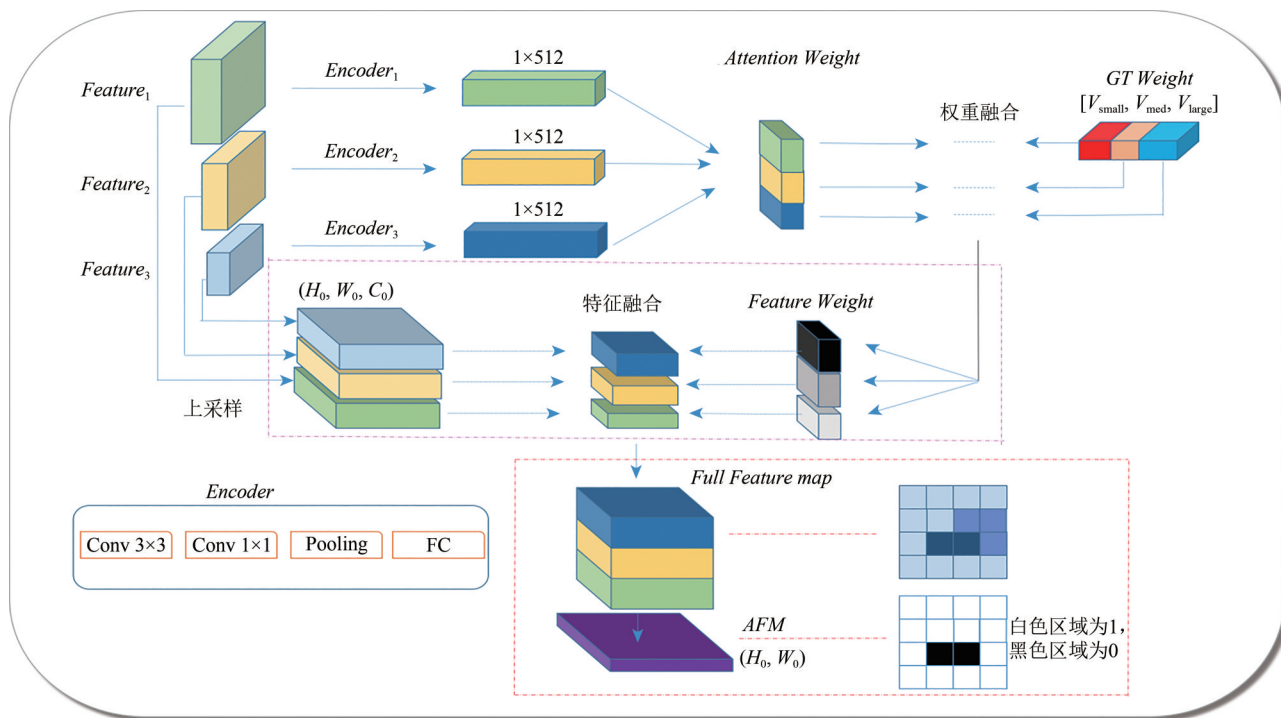


图3 自适应特征图掩码模块(SFMM)的整体结构

Fig. 3 The structure of SFMM

根据输入数据的标签，我们为每一个输入分配一个标签权重 *GT Weight*。 *GT Weight* 包含  $[V_{small}, V_{med}, V_{large}]$ ，按顺序依次对应 *Feature<sub>1</sub>*，*Feature<sub>2</sub>*，*Feature<sub>3</sub>*。  $V_{small}$  的计算方法如下所示：

$$V_{small} = \frac{n_{small}}{n_{targets}} \quad (3)$$

不同的特征图拥有不同的细粒度，低分辨率特征图语义信息丰富，对大目标的效果好；高分辨率特征图位置信息丰富细粒度高，则有益于小目标。如果目标的目标框在图像的面积中所占据的百分比低于4%，则为小目标。如果它大于9%，它就是一个大目标。中等目标的大小介于两者之间。  $n_{small}$  表示样本中的小目标数，  $n_{targets}$  表示样本中的目标数。  $V_{med}$ ，  $V_{large}$  与  $V_{small}$  的计算方法相同。 *GT Weight* 的作用是根据训练样本中3种目标在总目标中的比例，决定各个特征的比重。通过融合 *GT Weight* 和 *Attention Weight* 得到特征权重 *Feature Weight*，用于控制3个特征比例的权重。然后用特征权重融合 *Feature<sub>1</sub>*， *Feature<sub>2</sub>*， *Feature<sub>3</sub>*，得到一个全特征图 *Full feature map*，计算公式如下：

$$Feature\ Weight = GT\ Weight + Attention\ Weight \quad (4)$$

$$Full\ feature\ map = \sum_{i=1}^3 Features'_i \times Weight_i \quad (5)$$

我们在通道层面上对整个全特征图进行平均，得到一个原始掩码 *OM*，其维度为  $(H_0, W_0)$ 。该处理方式可以写成如下格式：

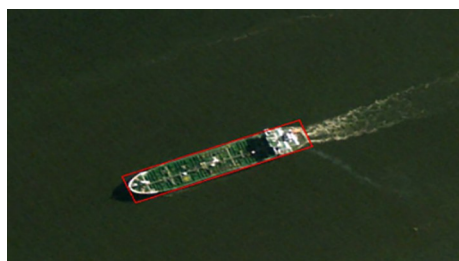
$$OM = \frac{\sum_{i=0}^{C'} Full\ feature\ map_i}{C'} \quad (6)$$

全特征图的平均值被当作一个阈值。我们使用阈值处理 *OM* 来获得自适应特征图掩码 *AFM*，使模型在推理中无法依赖这些“容易发现”的激活点进行判断。处理过程如下：

$$threshold = \frac{\sum_{i=0}^W \sum_{j=0}^H OM_{ij}}{H \times W} \times \mu \quad (7)$$

$$AFM(u, v) = \begin{cases} 0, OM(u, v) > threshold \\ 1, OM(u, v) \leq threshold \end{cases} \quad (8)$$

式中，  $(u, v)$  是一个坐标位置，  $\mu$  表示掩码的保守程度，实验中设定  $\mu$  为0.75。自适应特征图掩码可视化如图4所示。



(a) 原始掩码图

(a) Original mask map



(b) AFM掩码图

(b) AFM mask map

图4 自适应特征图掩码可视化

Fig. 4 The visualization of AFM

### 2.3 特征图海陆分割模块

与海面相比, 陆地环境更加复杂, 容易发生误报。主要原因是陆地区域的噪声误导了特征图, 导致了错误的激活。我们为特征图增强网络设计了一个水陆语义分割的辅助网络, 通过减少模型对陆地的激活来辅助检测任务。我们引入了一个特征图分割模块来实现对特征图的海陆分割预测。考虑到分割任务只存在单分类, 过于复杂的结构是不合适的。由于推理速度和模型大小的问题, 我们设计了一个更轻量级的模块结构, 命名为特征图海陆分割模块 (FSSM)。FSSM 由一个 conv 卷积模块、一个 Batchnorm 层和一个 Relu 激活函数组成。为了保证生成的水陆分割掩码不受 AFM 干扰, FSSM 利用不经过 FSSM 处理的  $Feature_3$  生成  $mask_p$ , 抑制陆地上的虚警。其结构如图 5 所示。

数据集的海陆分割标签  $mask_g$  是通过人工标注得到的, 船舶和水面区域标注为 1, 其他区域标注为 0。在训练时, 通过二分类交叉损失函数 BCELoss 计算  $mask_p$  和  $mask_g$  之间的损失  $Loss_{aux}$  来辅助更新梯度。 $mask_p$  和  $mask_g$  的尺寸均为  $(H, W)$ 。损失函数计算如下:

$$Loss_{aux} = BCELoss(mask_g, mask_p) \quad (9)$$

FSSM 的原理是经过引入陆地上存在船舶的可能性很小这一先验概念来增强网络预测的可靠性。FSSM 的目的是学习样本中的海陆分布, 通过海陆

分布预测与特征图进行融合来达到抑制特征图的误报的目的。FSSM 是一个独立的模块, 可以灵活加入其他的目标检测网络。

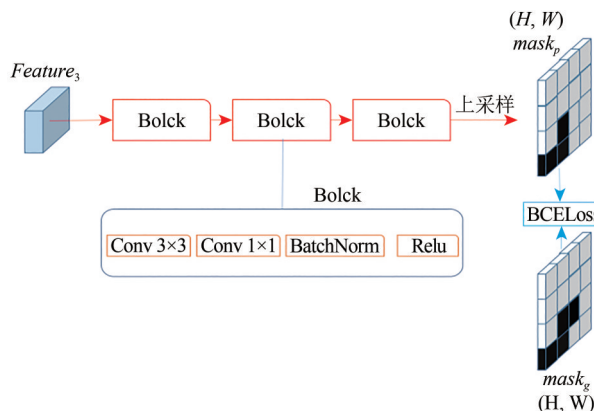


图5 水陆分割掩码模块(FSSM)的整体结构

Fig. 5 The structure of FSSM

## 3 数据结果处理与分析

### 3.1 实验条件

HRSC2016数据集由西北工业大学收集并且使用旋转框标注, 包含 1070 张图片 (Google Earth) 和 2976 个实例。HRSC2016 数据集包含尺寸从  $300 \times 300$  像素到  $1500 \times 900$  像素的图像, 拥有多种不同的尺度。(https://sites.google.com/site/hrsc2016/[2021-10-20]) 从类别上看, HRSC2016 数据集包含“船舶”、“驱逐舰”、“商船”和“航母”4 个类别。为了测试算法在不同细粒度条件下的效果, 我们选用两个任务 L1 与 L2。L1 任务中, 所有船舶目标的统一标注为“船舶”。而 L2 任务中目标的标签类别更加详细, 其中所有的船舶目标包含“船舶”、“驱逐舰”、“商船”和“航母”4 个类别标签。

本文所提出的方法是根据 Resnet 的代码实现, 并在一台有 GeForce 1080 和 8 G 内存的服务器上进行实验。该模型被训练了 40 个迭代, 批次大小为 4。

平均准确率 (mAP) 被采用来定量评估所提出方法的性能。采用 PASCAL VOC2007 指标来计算 mAP。我们选择 0.5 的 IOU 阈值, 如果样本标注和预测之间的 IOU 超过 0.5, 并且它们具有相同的类别, 则预测结果为真阳性 (TP)。否则, 预测结果为假阴性 (FN)。

通过十折交叉验证的策略将所提方法与其他 SOTA 算法进行了比较, 结果列于表 1。加入所提

方法后，检测算法 ROI trans 的十折平均 mAP 数值有着较大的提升，将基线 mAP 的平均值从 86.1% 提升到 90.8%，并超过了其他 SOTA。

表1 SOTA 算法与本文算法对比  
Table 1 Comparison between SOTA and ours

算法	mAP/%
RoI trans (Ding 等, 2018)	86.1±0.1
RSDet (Qian 等, 2019)	86.6±0.4
Gliding Vertex (Xu 等, 2021)	88.4±0.3
R3Det (Yang 等, 2020)	88.9±0.1
FMRNet + RoI trans (本文)	90.8±0.2

注:黑体表示最优结果。

3.2 自适应特征图掩码模块

HRSC2016 数据集上的对照实验将 SFMM 与 FMN 和基线（第一行）进行了比较，以证明我们方法的先进性。

同时在不同的骨干网络上进行了消融实验，从表 2 中可以看到，SFMM 在船舶检测任务中有最好效果。如图 6 所示，通过模型在 B、C 和 D 的样例中特征图的可视化，可以得出 SFMM 比其他模型从同一目标获得的关注点密度更大，并且不再过

拟合单一明显的特征得出检测结果。在样本 A 中，SFMM 更容易聚焦于目标，并明显抑制了非目标区域。受益于自适应掩码，我们的方法比基线在 mAP 方面取得了 4.2% 的改进。FMN 只利用了全连接层的特征，它有更深的语义信息，但失去了一些位置信息，缺乏可解释性。我们的方法同时考虑了位置和语义信息，比 FMN 高出 2.7%。从表 2 中 Resnet101 的结果可以得出，通过合理选择融合不同层次的特征图，更深层次的 SFMM 结构具有好的结果。

表2 在不同条件下,基线,FMN与本文算法实验对比  
Table 2 Comparison between ours, baseline and FMN in different situations

骨干网络	FMN	SFMM(本文算法)	mAP/%
Resnet50			86.1±1.2
Resnet50	√		87.6±1.4
Resnet50		√	90.3±0.9
Resnet101			87.0±1.0
Resnet101	√		88.5±0.8
Resnet101		√	90.4±1.1
Resnet101 细粒度		√	90.7±1.0

注:黑体表示最优结果。

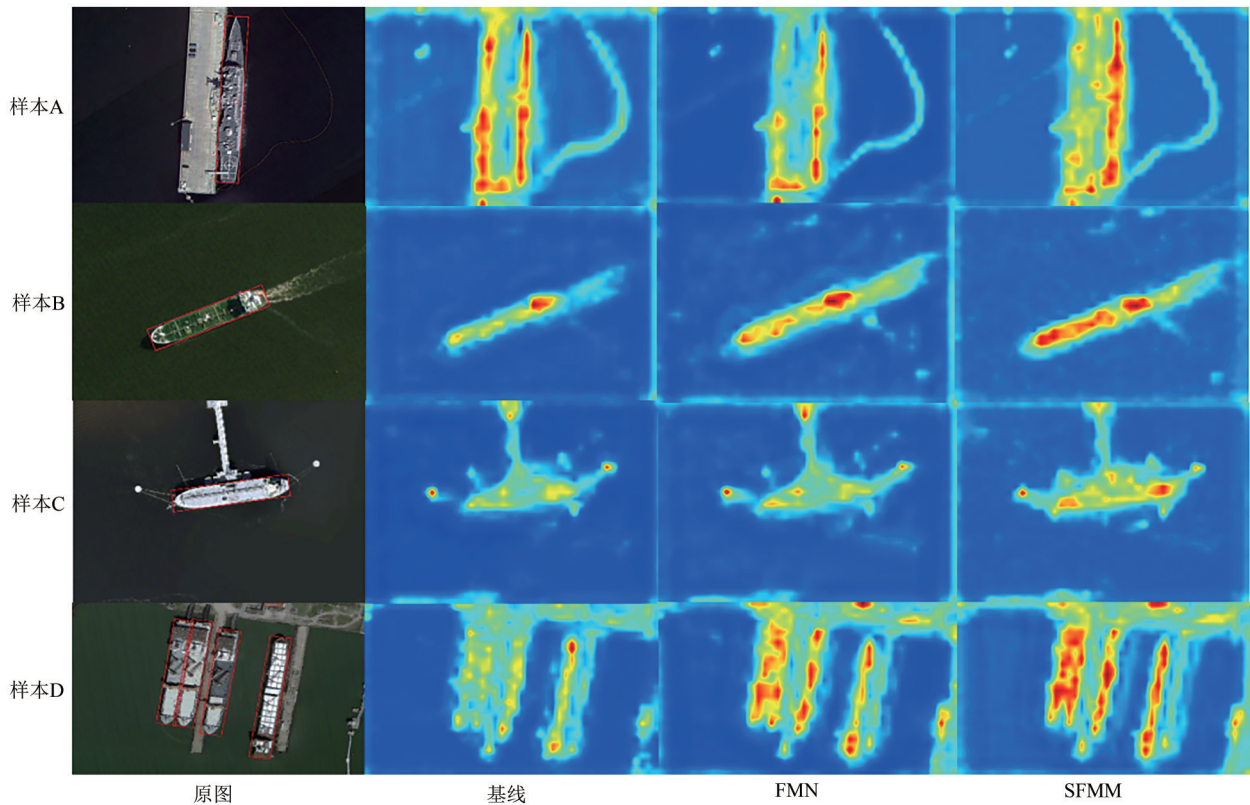


图6 基线,FMN,与SFMM的输出与特征图的对比

Fig. 6 Difference between baseline, FMN and SFMM in results and feature maps



考虑到船舶识别需要对具体型号的研判，我们在细粒度方面进行了相关实验。实验中用27类小类代替了“船舶”、“驱逐舰”、“商船”和“航母”4个大类别进行模型训练。为了保持结果的一致性，验证函数中采用了大类别进行计算mAP数值。由2表中可知，学习到更加细粒度的特征后，SFMM能生成更合适的自适应掩码，使模型的效果得到提升。

3.3 特征图海陆分割模块

我们的对照组实验是在HRSC2016的taskL1和taskL2上进行，同时在表3中比较了加入FSSM的基线和经过后处理的基线来证明FSSM方法的有效性。

详细的特征图和结果显示在图7。加入了FSSM之后，我们可以发现模型在特征图中抑制了图7中样本A和样本B的虚警。传统的消除虚警的方法是后处理，它根据距离消除了远离海陆边缘的预测。

后处理的特征图与原图相同，因为后处理是在模型输出之后，它并没有从根本上解决特征图中的误报来源。此外，由于决策边界不明确，海陆边缘的一些误报不容易清除，如样本C。此外，后处理方法在清洗结果的过程中极易去除处于岸上例如船坞中的正确标签，如样本D，FSSM能很好的避免这个错误。在减少误判和提高模型的特征提取能力方面，我们的结果比后处理中的暴力清除更好。

表3 在不同任务情况下,后处理与本文算法实验对比

Table 3 Comparison between ours and postprocessing on different tasks

任务	后处理	FSSM	P/%	R/%	mAP/%
Task1	√		81.0	93.2	88.5±0.6
Task1		√	82.1	93.6	90.0±0.4
Task1			79.9	93.4	86.2±0.3
Task2	√		76.7	90.0	83.2±0.4
Task2		√	77.3	90.5	84.6±0.1
Task2			76.3	90.3	82.5±0.7

注:黑体表示最优结果。

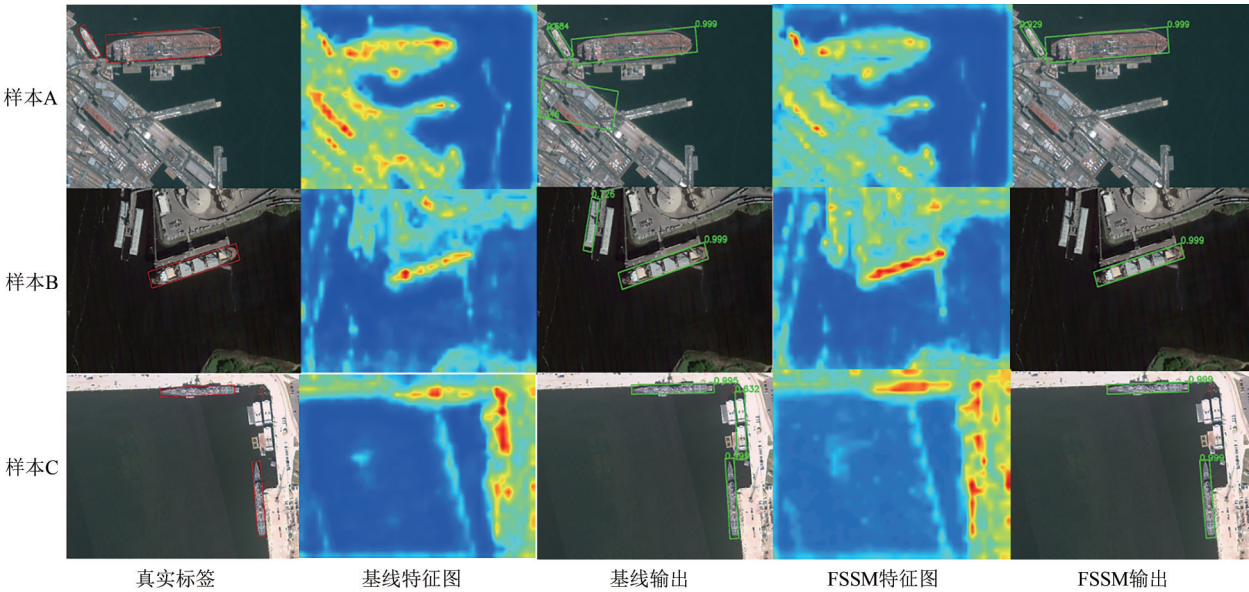


图7 基线与FSSM的特征图对照

Fig. 7 Comparison of the results and feature maps of baseline and FSSM

如表3所示，P表示精度，R表示召回率。可以看出后处理方法增加了1.1%和0.4%的精度，但没有改善甚至降低了召回率。我们的模块在准确率和召回率方面都有提高。从结果来看，由于从海陆分布中学到的先验知识，所提出的模块能有效抑制来自陆地的虚警。

4 结 论

在船舶检测的任务上，本文以Resnet为基础，提出了一种新型的骨干网络FMRNet。在解决模型对显著特征过拟合的问题上，我们提出的SFMM模块能让模型从多个特征对目标进行判断。为了减



少复杂背景导致的虚警,我们设计了FSSM模块,通过对非水面区域进行抑制,降低出现在非水面的目标置信度,在达到去除不合理的虚警的目的同时,也提高了模型的精度。在这两种模块的增强下,基线的mAP平均值分别有3.8%与1.4%的提升。加入FMRNet之后,基线的mAP平均值有4.7%的提升,比其他的SOTA算法有更好的表现。综上所述,本项工作的主要成果如下:

(1) 提出了FMRNet,这是一种新的骨干网络结构。

(2) 在FMRNet中设计了一个SFMM,该模块可以通过注意力机制选择性地利用特征图,生成一个自适应掩码,防止模型关注单一的特征点,从而防止特征过拟合。

(3) 同时也提出了一个与SFMM并行的特征FSSM,通过引入水陆分割掩码与特征图之间的融合,使模型聚焦在水面区域,减少船舶目标出现在陆地区域的误报。

**志 谢** 此次实验的数据获取得到了中国舰船设计研究中心的支持,在此表示衷心的感谢!

## 参考文献(References)

- Bi F K, Hou J Y, Chen L, Yang Z H and Wang Y P. 2019. Ship detection for optical remote sensing images based on visual attention enhanced network. *Sensors*, 19(10): 2271 [DOI: 10.3390/s19102271]
- Ding G D, Kha S, Tang Z M and Porikli F. 2017. Let features decide for themselves: feature mask network for person re-identification. *arXiv*: 1711.07155
- Ding J, Xue N, Long Y, Xia G S and Lu Q K. 2018. Learning RoI transformer for detecting oriented objects in aerial images. *arXiv*: 1812.00155
- Girshick R. 2015. Fast R-CNN//2015 IEEE International Conference on Computer Vision. Santiago: IEEE: 1440-1448 [DOI: 10.1109/ICCV.2015.169]
- Goetz A F H, Rock B N and Rowan L C. 1983. Remote sensing for exploration; an overview. *Economic Geology*, 78(4): 573-590 [DOI: 10.2113/gsecongeo.78.4.573]
- Gu Z Y. 2010. Research on high performance computing platform of ocean environmental information processing//2010 International Conference on Optics, Photonics and Energy Engineering. Wuhan: IEEE: 298-300 [DOI: 10.1109/OPEE.2010.5508128]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]
- Lunetta R S, Congalton R C, Fenstermaker L K, Jensen J R, McGwire K C and Tinney L R. 1991. Remote sensing and geographic information system data integration: error sources and research issues. *Photogrammetric Engineering and Remote Sensing*, 57(6): 677-687
- Nie X, Duan M Y, Ding H X, Hu B L and Wong E K. 2020. Attention mask R-CNN for ship detection and segmentation from remote sensing images. *IEEE Access*, 8: 9325-9334 [DOI: 10.1109/ACCESS.2020.2964540]
- Qian W, Yang X, Peng S L, Guo Y and Yan J C. 2019. Learning modulated loss for rotated object detection. *arXiv*: 1911.08299
- Redmon J, Divvala S, Girshick R and Farhadi A. 2016. You only look once: unified, real-time object detection//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE: 779-788 [DOI: 10.1109/CVPR.2016.91]
- Redmon J and Farhadi A. 2018. YOLOv3: an incremental improvement. *arXiv*: 1804.02767v1
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition. *arXiv*: 1409.1556
- Xu K Y, Hu W H and Chen R. 2006. A vessel targets detection method based on Canny algorithm. *Infrared Technology*, 28(6): 315-318 (许开宇, 胡文骅, 陈仁. 2006. 一种基于Canny算子的船舶目标检测算法. *红外技术*, 28(6): 315-318) [DOI: 10.3969/j.issn.1001-8891.2006.06.002]
- Xu Y C, Fu M T, Wang Q M, Wang Y K, Chen K, Xia G S and Bai X. 2021. Gliding vertex on the horizontal bounding box for multi-oriented object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(4): 1452-1459 [DOI: 10.1109/TPAMI.2020.2974745]

# FMRNet: A network structure for enhanced ship detection model training using feature maps

ZHANG Zekun<sup>1</sup>, TAN Zhenbiao<sup>1</sup>, YU Kun<sup>3</sup>, FANG Bing<sup>1</sup>, HUANG Xiao<sup>2</sup>, MA Jie<sup>1</sup>

*1. National Key Laboratory of Science and Technology on Multispectral Information Processing, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430072, China;*

*2. China Ship Development, and Design Center, Wuhan 430064, China;*

*3. China Railway Siyuan Survey and Design Group, Wuhan 430063, China*

**Abstract:** With the ongoing development of artificial intelligence technology, deep learning methods have become increasingly important role in the field of ship detection. However, the false alarms and missed detections that appear in deep learning algorithms hinder the application of technology in the field of ship detection. Although the classical deep learning methods can effectively deal with a single-background sea surface, the classical models can easily yield false alarms on shore when faced with data under complex backgrounds. In custom training, the model often tends to overly emphasize some salient features, which leads to feature overfitting. Detection can easily be missed when these salient features change. In the process of forward propagation of the model to the input, different network layers in the model generate corresponding mappings or feature maps from the input. Fully utilizing the semantic and spatial information of the feature maps is an effective way to reduce false alarms and missed detections. Compared with the traditional model, our proposed Feature Map Reinforcement Network (FMRNet) can fully utilize feature maps to generate adaptive feature map masks and water – land segmentation masks. This method ultimately reduces false alarms and missed detections by avoiding feature overfitting of the model and weakening the effects caused by complex backgrounds. In FMRNet, we design the Self Feature-map Mask Module (SFMM), which can selectively utilize the feature map through the attention mechanism for generating an adaptive mask. The mask prevents the model from focusing on a single feature point, which prevents feature overfitting. We also propose a Feature-map Sea-Land Segmentation Module (FSSM) that is parallel to SFMM. It reduces the false alarms of ship targets appearing in the land area by introducing the fusion between the water-land segmentation mask and the feature map. The experimental results, when compared with SOTA algorithms on publicly available datasets, show that the performance of the proposed method in this study is excellent and outperforms that of other SOTA algorithms. After FMRNet is added, the 10-fold average mAP value of the detection algorithm ROI trans largely improves. This enhancement increases the mean value of baseline mAP from 86.1% to 90.8%, which surpasses that of other SOTA algorithms. Benefiting from the adaptive mask, the mAP value of the model including the SFMM module is 90.4%, which achieves a 4.2% improvement over the baseline. Owing to the priori knowledge learned from the water-land distribution, FSSM improves the precision and recall of the model, which results in a MAP value of 86.4%. For the task of ship detection, we propose a novel backbone network, that is, FMRNet based on Resnet. Our proposed SFMM module enables the model to discern the target from multiple features for avoiding the overfitting of salient features. We design the FSSM module to reduce the false alarms caused by complex backgrounds. Suppressing the non-water surface area reduces the confidence level of targets appearing in non-water surface. FSSM achieves the purpose of removing unreasonable false alarms while improving the accuracy of the model.

**Key words:** remote sensing imaging, artificial intelligence, ship detection, neural networks, feature maps, network reinforcement, rotational object detection, overfitting suppression

**Supported by** National Natural Science Foundation of China (No.U1913602)