

A-FPN算法及其在遥感图像船舶检测中的应用

于野^{1,2}, 艾华¹, 贺小军^{1,3}, 于树海³, 钟兴^{1,4}, 朱瑞飞^{1,4}

1. 中国科学院 长春光学精密机械与物理研究所, 长春 130033;

2. 中国科学院大学, 北京 100049;

3. 长光卫星技术有限公司, 长春 130102;

4. 长光卫星技术有限公司 吉林省卫星遥感应用技术重点实验室, 长春 130039

摘要: 光学遥感图像船舶检测主要面临两个挑战: 光学遥感图像背景复杂, 船舶检测易受海浪、云雾及陆地建筑等多方面干扰; 遥感图像分辨率低, 船舶目标小, 对于其分类与定位带来很大困难; 针对上述问题, 在FPN的基础上, 提出一种融入显著性特征的卷积神经网络模型A-FPN (Attention-Based Feature Pyramid Networks)。首先, 利用卷积提取图像特征金字塔; 然后, 利用顶层金字塔逐级构建显著特征层, 抑制背景信息, 通过金字塔顶层的细粒度特征提高浅层特征的表达能力, 构建自上而下的多级显著特征映射结构; 最后利用Softmax分类器进行多层次船舶检测。A-FPN模型利用显著性机制引导不同感受下的特征进行融合, 提高了模型的分辨能力, 对遥感图像处理领域具有重要应用价值。实验阶段, 利用公开的遥感目标检测数据集NWPU VHR-10中的船舶样本进行测试, 准确率为92.8%, 表明A-FPN模型适用于遥感图像船舶检测。

关键词: 光学遥感, 船舶检测, 吉林一号卫星, 神经网络, 显著性特征

引用格式: 于野, 艾华, 贺小军, 于树海, 钟兴, 朱瑞飞. 2020. A-FPN算法及其在遥感图像船舶检测中的应用. 遥感学报, 24(2): 107-115

Yu Y, Ai H, He X J, Yu S H, Zhong X and Zhu R F. 2020. Attention-based feature pyramid networks for ship detection of optical remote sensing image. Journal of Remote Sensing (Chinese), 24(2): 107-115 [DOI: 10.11834/jrs.20208264]

1 引言

随着交通运输的日益发展, 船舶已经成为人们开发海洋, 利用海洋的重要工具之一。对海面上行驶船舶进行监视与检测, 可以有效的对特定海域、海湾和港口内的水运交通进行管制, 提高我国海运监控, 管理和调度能力, 以便快速处理或者尽量避免海运交通事故。

早期的船舶检测都是使用孔径雷达 (SAR) 图像 (田已睿等, 2007; 陈鹏等, 2010), 虽然SAR图像具有抗干扰能力强, 穿透力强等优点, 但同时其缺点也较为突出, 例如, 图像处理和检测难度偏大, 图像分辨率较低等。随着遥感技术的发展, 高分辨率光学遥感图像的出现, 遥感图像细节更加丰富, 利用卫星可见光图像进行海洋

目标检测引起了越来越多的关注。但由于天气、光照等自然条件的影响, 加上成像设备及卫星姿态变化等多种因素的作用, 卫星可见光图像容易产生模糊、阴影、噪声及亮斑等问题, 同时受海洋特殊地理条件 (如海水波动、浅滩等), 以及云块和陆地复杂边缘的干扰, 船舶目标往往难以和背景区分开来, 甚至会淹没在复杂的海面背景当中, 这给船舶目标的检测带来了一定的困难。

传统的检测方法, 是利用目标自身的物理参数, 通过手工设定参数的约束范围, 进而筛选出待测目标, 例如Lure和Rau (1994) 通过识别船体尾迹, 进行船舶检测, 但此类方法抗干扰能力低, 难以满足实际使用需求。Corbane等 (2010) 使用统计和形态学滤波的方法确定候选目标, 再通过小波变换与Radon变换的方法去除伪目标; Yang

收稿日期: 2018-07-06; 预印本: 2018-09-29

基金项目: 国家重点研发计划 (编号: 2016YFB0502600); 吉林省优秀青年人才基金 (编号: 20170520166JH)

第一作者简介: 于野, 1992年生, 男, 博士研究生, 研究方向为光学遥感图像处理、目标检测与识别。E-mail: 957458545@qq.com

通信作者简介: 贺小军, 1983年生, 男, 副研究员, 研究方向为光学卫星总体技术及遥感图像处理技术。E-mail: hexiaojun6@163.com

等(2014)综合图像的灰度信息与纹理信息,构建能突出水体中异常信号的特征图,分割特征图获得候选目标,通过紧致度与最小外接矩形长宽比这两个几何特征鉴别真实目标。Qi等(2015)首先利用基于频域的显著性增强提取出船舶候选区域,再对候选区域提取梯度直方图特征(HOG),分析其对称性,进而筛选出待检船只。然而,遥感图像背景复杂,大多数传统算法在遥感领域的应用方面都存在一定程度的局限性。

近年来,深度学习(Hinton和Salakhutdinov, 2006; LeCun等, 2015)在目标检测、识别等领域取得了巨大进展,以Alexnet、VGG为代表的卷积神经网络结构,在特征提取方面具有传统人工特征工程无法比拟的优势,例如:Suleymanova等(2018)将基于深度学习的目标检测方法用于星形胶质细胞检测,推动了医疗领域的智能化发展。因此,研究深度学习算法在遥感领域中的应用具有重要意义。基于卷积神经网络的目标检测算法主要分为One-Stage和Two-Stage两类。最早出现的Two-Stage目标检测算法是R-CNN(Girshick等, 2014),通过Selective Search算法选择候选区域,然后再利用CNN进行分类。2015年, Fast-RCNN(Girshick, 2015)在R-CNN基础上提出一种Roi-Pooling层,减少R-CNN在提取特征时的冗余计算。同年, Ren等(2015)提出Faster-RCNN模型,利用RPN结构代替R-CNN中的搜索方法,实现神经网络的端到端训练,进一步提高网络检测速度。随后,又出现了很多Faster-RCNN的改进算法R-FCN(Dai等, 2016)、RON(Kong等, 2017)、MSCNN(Cai等, 2016)、MLKP(Wang等, 2018)等,其中R-FCN中通过分析分类问题于检测问题的区别,提出一种新的卷积层Score-Map,并且将全部的卷积层共享,进一步降低网络复杂度;RON利用特征融合与多层预测的方式提高了小目标检测准确率;MLKP算法利用目标检测中的高阶统计量来提高检测器的性能。与Two-Stage思路不同,One-Stage直接利用回归的方法检测目标,具有更快的检测速度,但同时其精度普遍低于Two-Stage方法。One-Stage主要包括Yolo(Redmon等, 2016)、Yolov2(Redmon等, 2016)、Yolov3(Redmon和Farhadi, 2018)、SSD(Liu等, 2016)、FPN(Lin等, 2017a)、RetinaNet(Lin等, 2017b)等,其中Yolo和Yolov2只在最后一层特征

上进行预测,对于小目标检测效果一般,而SSD、FPN、RetinaNet、Yolov3都采用多级预测,提高模型对小目标的检测率,其中FPN与Yolov3采用金子塔特征融合方法,提高特征表达能力;RetinaNet在FPN的基础上,提出一种新的损失函数focal loss,用于解决正负样本不均衡问题。此外,近几年研究表明,将显著性机制融入神经网络模型,可以有效提高目标分类或检测的准确率。Fu(2017)定义了一种新的损失函数rank loss,并强迫精细尺度得到的分类结果优于上一尺度,从而网络提取出更有利于精细分类的局部区域。Wang(2017)利用显著性机制对残差模块(He等, 2016)进行调整,突出对分类更有利的特征。Lin等(2017c)提出一种针对遥感船舶的分割网络,利用深层的网络能得到船头/船尾的大概位置即显著图,然后浅层网络根据显著图,预测出船舶的精细的位置。相比于传统算法,卷积神经网络可以有效提高目标检测准确率,但是其自身结构也存在一种矛盾:浅层网络的特征表达能力弱,随着网络层数的加深,网络的特征表达能力随之提高,但由于池化处理,其空间定位能力会逐渐下降。目前较为有效的解决方法就是进行特征融合,例如上述的FPN与RetinaNet,但其融合方法仅仅是将深层特征上采样处理后与浅层特征进行相加处理,因此本文在此基础上提出一种新的特征融合方法,将显著性理论融入神经网络算法,提高网络的特征描述能力。此外,本文算法与Wang(2017)不同的是:Wang(2017)算法是在单一特征层级内,利用显著性机制,突出对分类更有利的特征,而本文采用的显著图是在整个多级特征金字塔内,利用深层细粒度特征,学习用于浅层特征融合的掩码图,使得融合后的网络层同时具备鲁棒的特征表达能力与准定位能力,最后将A-FPN应用于船舶检测,验证A-FPN模型在遥感船舶检测领域的实用性。

2 模型方法

本文的模型方法是利用融入显著性机制的深度卷积神经网络检测光学遥感图像中的船舶目标,主要流程图如图1所示:首先利用残差模块提取图像特征金字塔,获取目标多级映射特征,然后利用深层特征的高表达能力提取浅层显著特征,提高映射特征的鲁棒性,然后,为降低小目标检测

的漏检率, 分别在处理后的多级特征金字塔上, 利用 Ren 等 (2015) 提出的 Anchor 结构, 检索遥感图像中不同尺度、不同位置的船舶目标, 然后利用 Softmax 分类器进行船舶目标确认, 最后利用非极大值抑制的方法去除冗余信息。

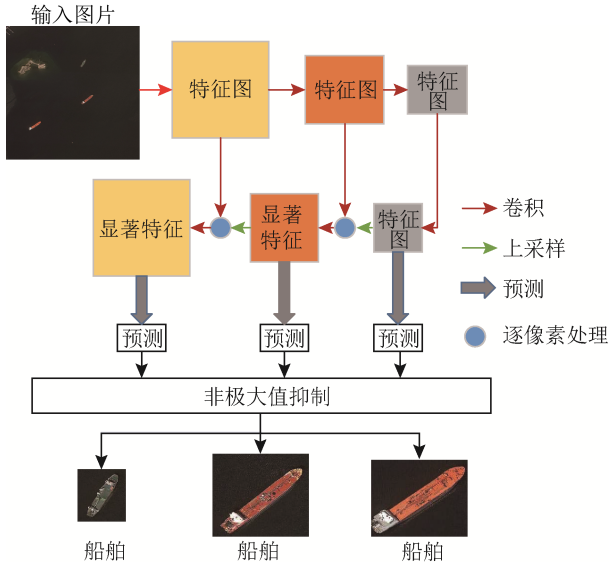


图1 算法流程图

Fig. 1 Workflow of our proposed method

2.1 构建卷积特征金字塔

本文在构建卷积特征金字塔时参考 FPN, 采用 ResNet 结构, 其残差模块有模块 1 和模块 2 两种形式, 如图 2 所示, 模块 1 将输入 X_1 直接与卷积后的输出 $F(X_1)$ 相加, 模块 2 采用 $\{1 \times 1, 3 \times 3, 1 \times 1\}$ 形式的卷积, 并且由于模块 2 内部存在升维操作, 因此对输入 X_2 做 1×1 卷积的升维处理, 然后与输出 $F(X_2)$ 相加。本文在构建特征金字塔时选择模块 2。

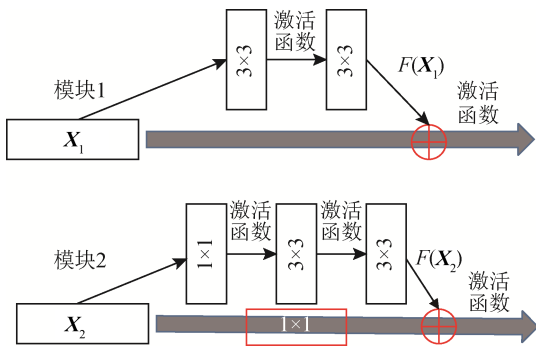


图2 残差结构

Fig. 2 Two structures of residual block

2.2 显著性特征网络原理

显著区域是图像中最能引起人们兴趣、最能

表现图像内容的区域。认知心理学的研究表明, 图像中有些区域能显著的吸引人的注意, 这些区域含有较大的信息量。事实上, 显著区域的选择是非常主观的, 由于任务和知识背景的不同, 对于同一幅图像, 可能会选择不同的区域作为显著区域。因此, 出现越来越多的显著性分析算法, 但对于复杂度极高的遥感图像, 传统显著算法难以提取有效目标, 本文将深度神经网络与显著性检测思想结合, 提出一种 A-FPN 模型, 利用图像的高级卷积特征通过前馈的方式, 学习特征层的显著图, 抑制背景信息, 提高有效目标的响应参数。

A-FPN 主要分为两个部分: (1) 计算图像卷积特征金字塔; (2) 逐级构建显著特征。显著特征的核心思想是利用深层特征, 通过反向传播学习出浅层特征的掩码图 $N_{p,c}(x)$, 然后将掩码图作用于卷积特征图上, 得到显著特征, 具体式 (1)

$$H_{p,c}(x) = N_{p,c}(x) \otimes F_{p,c}(x) \quad (1)$$

式中, $F_{p,c}(x)$ 是卷积金字塔的浅层特征, $N_{p,c}(x)$ 表示掩码图, $H_{p,c}(x)$ 是显著特征, p 表示二位空间中所有点的位置, c 表示特征维度, 符号 ' $\otimes$ ' 表示 $N_{p,c}(x)$ 作用于 $F_{p,c}(x)$ 。

显著特征的处理方法 ' $\otimes$ ' 可以选用对应位置的元素相乘, 但掩码图中可能产生大量 $[0, 1)$ 区间的掩码点, 有可能导致整个网络模型陷入局部极值, 产生过拟合现象, 并且这种暴力相乘的方式很有可能破坏原有特征层之间的相关性, 因此提出利用类似于残差学习的思想, 式 (2)

$$H_{p,c}(x) = (1 + N_{p,c}(x)) * F_{p,c}(x) \quad (2)$$

式中, ' $*$ ' 表示对应位置元素相乘, $N_{p,c}(x)$ 的取值范围是 $[0, 1)$, 这样做的好处在于没有直接暴力拆解原特征层之间的联系, 用一种平缓的处理方式, 保证了模型的精度。

本文设计了两种不同的掩码图 (分别对应图 3 中的方法 1 与方法 2) $N_{p,c}(x)$

$$N_{p,c}(x)_1 = \text{Sigmoid}_{p,c}(x) \quad (3)$$

$$N_{p,c}(x)_2 = \text{Soft max}_p(x) \quad (4)$$

图 3 绿色虚线和紫色虚线分别表示两种不同掩码图的计算原理。其中 $G_{p,c}(x)$ 表示利用深层特征上采样得到与原浅层特征尺寸相同的上采样特征, 图中红色虚线的部分表示 $G_{p,c}(x)$ 中 2 维空间一点在其所有通道上的数据。

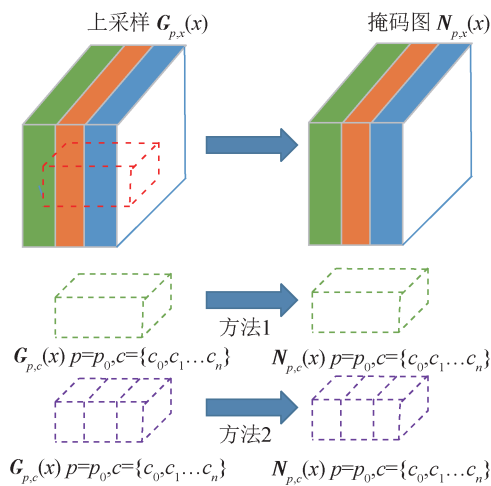


图3 两种掩码图计算原理

Fig.3 Principle of two mask-maps

图3中的Sigmoid函数属于归一化‘S’形曲线，将多维空间（2维几何空间 p 与‘多维光谱空间’ c ）每一点 x 归一化，因此 $G_{p,c}(x)_1$ 属于混合全空间掩码， $N_{p,c}(x)_2$ 在每一个2维空间位置，在其‘多维光谱空间’ c 中，利用Softmax函数将其归一化，属于‘多维光谱掩码’，最后经过实际测试，如表1所示， $N_{p,c}(x)_1$ 在NWPU VHR-10测试集的调和均值 F 为92.8%，而 $N_{p,c}(x)_2$ 只有88.0%，因此本文采用第一种掩码图。

表1 两种掩码图测试结果

Table 1 The results of two mask-maps

方法	N	N_{tp}	N_{fp}	$F/\%$
$N_{p,c}(x)_1$	308	283	19	92.8
$N_{p,c}(x)_2$	308	262	26	88.0

注： N_{tp} 为正确检测到的船舶目标， N_{fp} 为错误检测到的目标。

2.3 A-FPN网络结构

本设计的A-FPN具体结构如图4所示，首先利用卷积、池化提取图像的浅层特征，然后利用4个残差模块继续提取目标深层特征，接着利用掩码图提取目标显著特征，通过抑制背景响应，提高有效目标表达能力，最后分别利用3个特征层的输出检测船舶目标，图4中残差一模块[2, 3, 3, 2]代表残差模块数量的设置。

此外，由于本算法利用多层特征进行预测，将导致同一船舶目标被多次检测，因此引入NMS（非极大值抑制）算法，利用置信度与IOU信息，排除冗余目标。

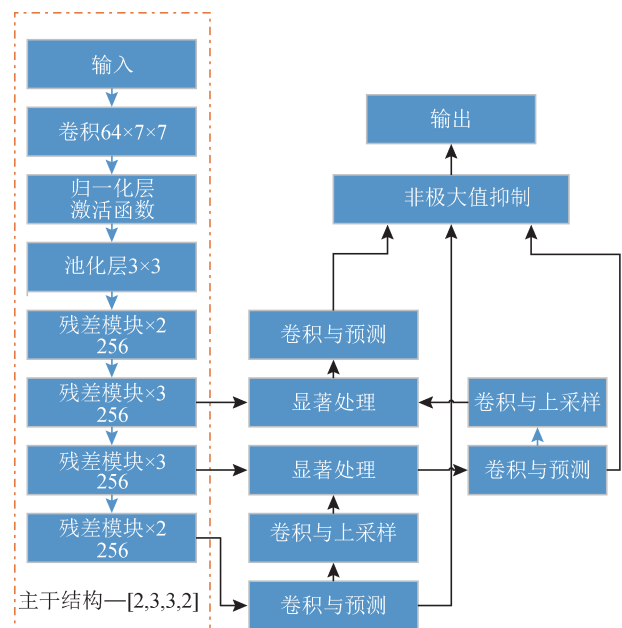


图4 本文算法的网络结构

Fig.4 The network structure of this algorithm

3 实验与分析

3.1 实验数据收集及增强处理

实验数据来自国内首颗商业光学遥感卫星吉林一号实拍影像与谷歌地图影像。吉林一号分辨率为0.72 m，谷歌卫星数据来源于谷歌地图，其分辨率为1 m，经有效裁剪后得到尺寸为2000×2000的1300张带有船舶目标的图像，其中，谷歌卫星数据约为830张。此外，为保证模型的通用性，在实验阶段，利用西北工业大学Cheng等（2014）公开的数据集NWPU VHR-10中所有包含船舶的61张数据即进行测试。

3.1.1 数据增强

深度学习算法不仅依赖于网络结构和参数的设计，数据集的质量也会影响最终的船舶检测准确率，而遥感目标图像与常规图像不同，难以大量获取，因此，设计如下的人工数据增强方法：

（1）平移增强。平移即移动样本在原图像中的位置，虽然卷积神经网络在特征提取方面具有平移不变性，但平移增强可避免网络学习定位参数时发生过拟合。

（2）旋转增强。遥感图像中，即便同一船舶目标在不同时刻可能出现不同的姿态，因为采集到的数据有限，不可能囊括所有可能的目标角度，而神经网络算法只具有部分旋转不变性，因此对

样本目标进行旋转增强处理, 提高模型的泛化能力。

(3) 颜色变换增强。卫星成像受天气影响, 船舶目标在不同时刻可能呈现出不同的颜色, 因

此颜色变换也可以提高模型泛化能力。

在训练过程中, 对于每一批次的训练样本首先进行数据增强处理, 然后将增强处理之后的数据传入网络进行训练。图5是3种样本增强的效果图。

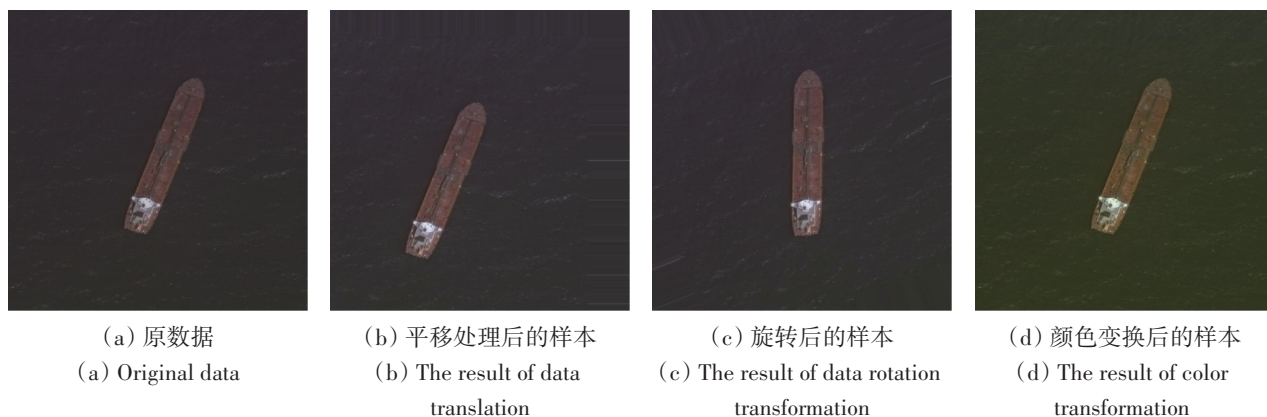


图5 样本增强效果

Fig. 5 The results of data argument

3.1.2 数据集拆分

将数据集拆分成训练集、验证集和测试集, 其中训练集用来拟合模型, 寻找算法的全局最优解。为防止过拟合现象, 通过设立验证集进行验证, 当模型在验证集上的拟合结果远低于训练集时, 即为过拟合。验证集间接参与的模型的训练及参数调优过程, 因此为保证算法的鲁棒性, 设立测试集用于最后的性能评估测试, 本文在收集的1300张吉林一号与谷歌卫星拍摄的数据中, 随机选择800张作为数据集1, 随机选择250张作为数据集2, 剩余250张作为数据集3, NWPU VHR-10中的船舶样本作为数据集4。实验分为两部分:

实验一: 为提高算法效率, 在FPN基本架构下, 进行算法裁剪, 在保证准确率的同时, 寻找最简单的结构。利用数据集1作为训练集, 数据集2作为验证集, 然后为保证模型的鲁棒性, 利用数据集4作为测试集。

实验二: 由于NWPU VHR-10中的船舶数据远小于训练集(训练率高于90%), 为保证模型不受训练率影响, 并且验证本文算法在遥感船舶检测领域的应用价值, 利用数据集1作为训练集, 数据集2作为验证集, 数据集3作为测试集, 对3种不同算法进行对比实验。

3.2 实验平台与算法评价准则

实验采用的计算机配置是 Nvidia Titan Xp 显

卡, 16 G内存, 利用Keras框架进行训练。实验环境为: ubuntu14.04, Keras-2.1.1在cuda8.0的背景下调用GPU训练与测试。

为客观评估目标检测算法的性能, 本实验采用 Precision-Recall 曲线(后文用 P-R 代替)、Ap Average precision 值对本文算法进行全面测试。

P-R 曲线中 Precision 直观的反映出算法的误检率为

$$\text{Precision} = \frac{N_{tp}}{N_{tp} + N_{fp}} \quad (5)$$

式中, N_{tp} 是正确检测到的船舶目标, N_{fp} 是错误检测到的目标。

P-R 曲线中的 Recall 可以代表算法的漏检率为

$$\text{Recall} = \frac{N_{tp}}{N_{tp} + N_{fn}} \quad (6)$$

式中, N_{fn} 是漏检的船舶目标。P-R 曲线以 Precision 为纵轴, Recall 为横轴, 通过改变阈值条件拟合而成, 另外对于目标检测任务, 为体现目标位置预测的准确程度, 计算 P-R 曲线时, 需要考虑预测位置与目标真实位置的 IOU, 在本实验中计算 P-R 曲线时 IOU 设置为 0.5。

3.3 实验结果与分析

3.3.1 超参数选择

在训练卷积神经网络时, 需要设定一些超参数, 例如学习率、训练迭代次数, 损失函数的选

择等, 具体参数设置如表2所示。

表2 优化算法Adam模型超参数
Table 2 The hyperparameters of Adam method

参数	值
初始学习率	0.0001
训练批次	16
训练轮数	200
目标函数	Cross-entropy、MSE

优化算法Adam是带有动量项的SGD(梯度下降)算法, 利用梯度的一阶矩估计和二阶矩估计动态调整每个参数的学习率, 每次迭代更新时, 每一次迭代学习率都有个确定范围, 使得参数比较平稳, 并且引入临近迭代中的梯度方向的惩罚项, 提高模型收敛速度。

目标函数分为分类与回归两部分, 其中MSE(均方误差)算法通过计算预测位置与真实位置的最小二乘值, 实现目标定位; Cross-entropy(交叉熵函数)计算目标分类预测置信度与目标真实类别的概率差异。另外, 交叉熵作为损失函数还有一个好处是使用sigmoid函数在梯度下降时能避免均方误差损失函数学习速率降低的问题。

3.3.2 实验一

神经网络的结构直接影响算法的复杂度, 对于海量遥感图像, 目标检测算法不仅需要较高的准确率, 还应具备较快的检测速度, 避免资源浪费, 因此本节通过实验与分析, 通过调整残差结构对A-FPN主干结构进行网络裁剪与参数寻优。图6中给出3种主干架构对应的P-R曲线, 综合准确率和检测速度, 最终选择残差一模块[2, 3, 3, 2]作为A-FPN的主干结构。

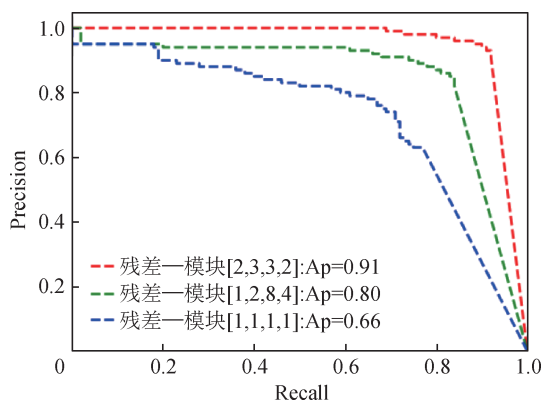


图6 3种结构对应的P-R曲线

Fig. 6 P-R curves of three structures

3.3.3 实验二

本文对比9种主流的目标检测算法: Qi等(2015)、Yolov2 (Redmon等, 2017)、Faster-RCNN (Ren等, 2015)、FPN (Lin等, 2017a)、Mask-RCNN (He, 2017)、R-FCN (Dai等, 2016)、RON (Kong等, 2017)、MSCNN (Cai等, 2016)、SSD (Liu等, 2016), 分别按顺序对应表3中的方法1-9, 具体实验结果如表3与图7所示。

表3 算法对比结果

Table 3 Results of three methods

方法	Precision	Recall	<i>F</i>
A-FPN	93.7	90.8	92.2
方法1	84.4	47.6	60.8
方法2	87.2	78.0	82.3
方法3	94.7	87.8	91.1
方法4	93.0	83.1	87.8
方法5	94.6	92.8	93.7
方法6	92.7	90.6	91.6
方法7	93.0	90.8	91.9
方法8	91.9	89.7	90.8
方法9	93.7	85.1	89.2

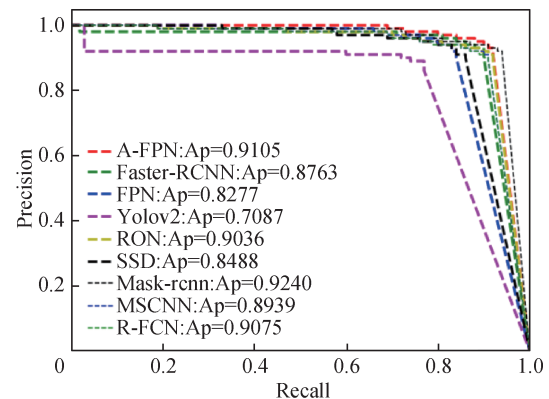


图7 P-R曲线

Fig. 7 P-R curves

图7是实验二中对比较算法的P-R曲线, 其中方法1是非监督学习算法, 在本实验的测试集表现一般, 因此没有统计其P-R曲线。表3为具体实验结果, 可以看出A-FPN模型的综合评价指标, 即调和均值F高于除方法5的其他算法, 而方法5算法在Faster-RCNN模型上添加一个用于目标分割的分支结构, 检测速度相对较慢, 不利于海量遥感图像快速处理。

表3中的调和均值 F 给出各算法的综合性能, 其定义为

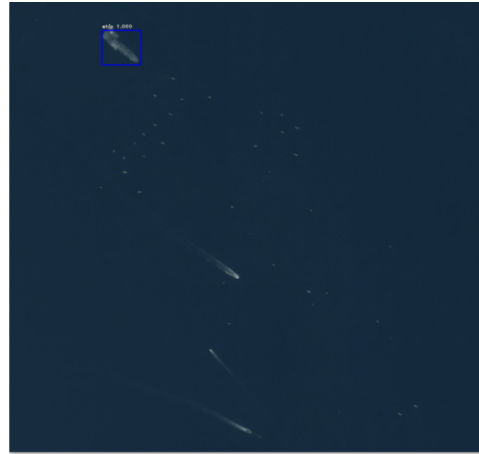
$$F = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

图8是A-FPN检测的部分结果, 图8(a)来自吉林一号拍摄数据, 图8(b)来自NWPU VHR-10, 图8(c)为谷歌卫星数据。



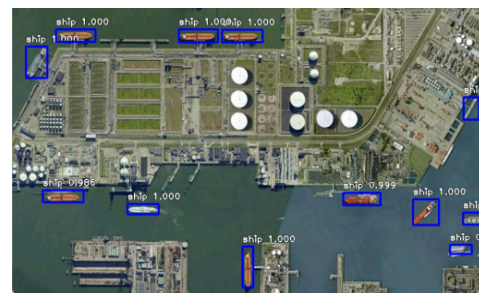
(a) 吉林一号卫星影像的检测结果

(a) Test result from Jilin-1 satellite imagery



(b) NWPU VHR-10数据的检测结果

(b) Test result from NWPU VHR-10



(c) 谷歌卫星影像的检测结果

(c) Test result from Google satellite imagery



图8 本文算法部分检测结果

Fig. 8 Part of the test results of our algorithm

4 结 论

本文在FPN的基础上,提出了一种基于显著性特征的神经网络模型A-FPN,用于遥感图像船舶检测。首先利用残差模型提取图像特征金字塔,然后设计一种显著图模式,利用深层卷积特征的抽象表达能力,通过掩码层,对有效目标特征进行增益,同时抑制背景信息,提高浅层特征层的表达能力,因此,经过显著性处理后的浅层特征同时具备强定位能力与稳定的特征描述能力,最后利用多尺度预测的方法检测船舶。A-FPN的优点在于利用显著性机制将深层卷积提取的全局特征与浅层卷积提取的局部特征相融合,提高了模型的鲁棒性,适用于背景极度复杂的遥感图像,并且利用国产商业卫星吉林一号收集大量真实数据,通过充分实验与分析,在FPN主干结构的基础上对模型进行寻优与裁剪,在保证检测准确率的同时,降低了算法的复杂度。与传统船舶检测算法相比,本文算法具有更好的鲁棒性,能够适用于各类遥感图像背景,对于云雾、海浪及人工建筑等干扰物具有较高的抗干扰能力,对遥感图像处理领域具有重要应用价值。

志 谢 此次实验的数据获取得到了长光卫星技术有限公司的支持,在此表示衷心的感谢!

参考文献(References)

- Cai Z W, Fan Q F, Feris R S and Vasconcelos N. 2016. A unified multi-scale deep convolutional neural network for fast object detection//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer [DOI: 10.1007/978-3-319-46493-0_22]
- Chen P, Liu R Y and Huang W G. 2010. A ship detection model based on multi-distribution on SAR imagery. *Journey of Remote Sensing*. 14(3): 546-557 (陈鹏, 刘仁义, 黄韦良. 2010. SAR图像复合分布船只检测模型. *遥感学报*, 14(3): 546-557) [DOI: 10.11834/jrs.20100312]
- Cheng G, Han J W, Zhou P C and Guo L. 2014. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS Journal of Photogrammetry and Remote Sensing*, 98: 119-132 [DOI: 10.1016/j.isprsjprs.2014.10.002]
- Corbane C, Najman L, Pecoul E, Demagistri L and Petit M. 2010. A complete processing chain for ship detection using optical satellite imagery. *International Journal of Remote Sensing*, 31(22): 5837-5854 [DOI: 10.1080/01431161.2010.512310]
- Dai J F, Li Y, He K M and Sun J. 2016. R-FCN: object detection via region-based fully convolutional networks//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain: Curran Associates Inc
- Fu J, Zheng H, Mei T. Look closer to see better: Recurrent attention convolutional neural network for fine-grained image recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4438-4446.
- Girshick R, Donahue J, Darrell T and Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE: 580-587 [DOI: 10.1109/CVPR.2014.81]
- Girshick R. 2015. Fast R-CNN//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE [DOI: 10.1109/ICCV.2015.169]
- He K M, Gkioxari G, Dollár P and Girshick R. 2017. Mask R-CNN//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE [DOI: 10.1109/ICCV.2017.322]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hinton G E and Salakhutdinov R R. 2006. Reducing the dimensionality of data with neural networks. *Science*, 313(5786): 504-507 [DOI: 10.1126/science.1127647]
- Kong T, Sun F C, Yao A B, Liu H P, Lu M and Chen Y R. 2017. RON: reverse connection with objectness prior networks for object detection//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE [DOI: 10.1109/CVPR.2017.557]
- LeCun Y, Bengio Y and Hinton G. 2015. Deep learning. *Nature*, 521(7553): 436-444 [DOI: 10.1038/nature14539]
- Lin H N, Shi Z W and Zou Z X. 2017c. Fully convolutional network with task partitioning for inshore ship detection in optical remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 14(10): 1665-1669 [DOI: 10.1109/LGRS.2017.2727515]
- Lin T Y, Dollár P, Girshick R, He K M, Hariharan B and Belongie S. 2017a. Feature pyramid networks for object detection//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE [DOI: 10.1109/CVPR.2017.106]
- Lin T Y, Goyal P, Girshick R, He K M and Dollár P. 2017b. Focal loss for dense object detection//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE [DOI: 10.1109/ICCV.2017.324]
- Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y and Berg A C. 2016. SSD: single shot multibox detector//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer [DOI: 10.1007/978-3-319-46448-0_2]
- Lure F Y M and Rau Y C. 1994. Detection of ship tracks in AVHRR cloud imagery with neural networks//Proceedings of 1994 IEEE

- International Geoscience and Remote Sensing Symposium. Pasadena: IEEE: 1401-1403 [DOI: 10.1109/IGARSS.1994.399451]
- Qi S X, Ma J, Lin J, Li Y S and Tian J W. 2015. Unsupervised ship detection based on saliency and S-HOG descriptor from optical satellite images. *IEEE Geoscience and Remote Sensing Letters*, 12 (7): 1451-1455 [DOI: 10.1109/LGRS.2015.2408355]
- Redmon J and Farhadi A. 2018. Yolov3: an incremental improvement. arXiv preprint arXiv: 1804.02767
- Redmon J, and Farhadi A. 2017. YOLO9000: better, faster, stronger// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE [DOI: 10.1109/CVPR.2017.690]
- Ren S Q, He K M, Girshick R and Sun J. 2015. Faster R-CNN: towards real-time object detection with region proposal networks// *Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal, Canada: MIT Press: 91-99
- Suleymanova I, Balassa T, Tripathi S, Molnar C, Saarna M, Sidorova Y and Horvath P. 2018. A deep convolutional neural network approach for astrocyte detection. *Scientific Reports*, 8: 12878 [DOI: 10.1038/s41598-018-31284-x]
- Tian S R, Sun G Y, Wang C and Zhang H. 2007. A Ship detection method in SAR image based on gravity enhancement. *Journey of Remote Sensing*. 11(4): 452-459 (田巳睿, 孙根云, 王超, 张红. 2007. 基于引力场增强的SAR图像船舶检测方法研究. *遥感学报*, 11(4): 452-459) [DOI: 10.11834/jrs.20070463]
- Wang H, Wang Q L, Gao M Q, Li P H and Zuo W M. 2018. Multi-scale location-aware kernel representation for object detection// *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA: IEEE [DOI: 10.1109/CVPR.2018.00136]
- Wang F, Jiang M, Qian C, et al. Residual attention network for image classification[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017: 3156-3164.
- Yang G, Li B, Ji S F, Gao F and Xu Q Z. 2014. Ship detection from optical satellite images based on sea surface analysis. *IEEE Geoscience and Remote Sensing Letters*, 11(3): 641-645 [DOI: 10.1109/LGRS.2013.2273552].

Attention-based feature pyramid networks for ship detection of optical remote sensing image

YU Ye^{1,2}, AI Hua¹, HE Xiaojun^{1,3}, YU Shuhai³, ZHONG Xing^{1,4}, ZHU Ruifei^{1,4}

1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Chang Guang Satellite Technology Co., Ltd., Changchun 130102, China;

4. Chang Guang Satellite Technology Co., Ltd, Key Laboratory of Satellite Remote Sensing Application Technology of Jilin Province, Changchun 130039, China

Abstract: Ship detection on spaceborne optical images is a challenging task that has attracted increasing attention because of its potential applications in many fields. Although some ship detection methods have been proposed in recent years, many obstacles still exist because of the large-scale and high complexity of optical remote sensing images. Identifying ships from interferences, such as the features of clouds, waves, and some land architectures that are similar to ships, is difficult. Therefore, an accurate and stable deep-learning based method is proposed in this work.

The method involves three steps: First, the image feature pyramid is extracted using convolution to detect multiscale ship targets. Second, a multilevel attention feature mapping structure is constructed from top to bottom using the fine-grained features of the top layer from the pyramid to improve the expressive ability of shallow features. Finally, Softmax classifier is used for multilevel ship detection.

The experimental results based on real remote sensing images are shot by "JL-1" satellite, Google satellite, and NWPU VHR-10. The result proves that the performance of our algorithm is better than the three other state-of-the-art methods. In addition, the network was cut while ensuring accuracy. The complexity of our algorithm is reduced, and its practicality is improved by experiments and analysis.

This work proposes an attention-based method called A-FPN. However, unlike traditional algorithms, A-FPN has higher robustness and wider range of use. Furthermore, we effectively cut the network to reduce the complexity of the algorithm, thereby exhibiting the significance of our algorithm in practical applications.

Key words: optical remote sensing, ship detection, JL-1 satellite, neural network, attention features

Supported by National Key Research and Development Program of China (No. 2016YFB0502600); Outstanding Young Talent Fund Project of Jilin Province (No. 20170520166JH)