

基于生成对抗网络和线上难例挖掘的SAR图像舰船目标检测

李健伟^{*①} 曲长文^① 彭书娟^① 江源^②

^①(海军航空大学 烟台 264001)

^②(92228部队 北京 100044)

摘要: 基于深度学习的SAR图像舰船目标检测算法对图像的数量和质量有很高的要求, 而收集大体量的舰船SAR图像并制作相应的标签需要消耗大量的人力物力和财力。该文在现有SAR图像舰船目标检测数据集(SSDD)的基础上, 针对目前检测算法对数据集利用不充分的问题, 提出基于生成对抗网络(GAN)和线上难例挖掘(OHEM)的SAR图像舰船目标检测方法。利用空间变换网络在特征图上进行变换, 生成不同尺寸和旋转角度的舰船样本的特征图, 从而提高检测器对不同尺寸、旋转角度的舰船目标的适应性。利用OHEM在后向传播过程中发掘并充分利用难例样本, 去掉检测算法中对样本正负比例的限制, 提高对样本的利用率。通过在SSDD数据集上的实验证明以上两点改进对检测算法性能分别提升了1.3%和1.0%, 二者结合提高了2.1%。以上两种方法不依赖于具体的检测算法, 且只在训练时增加步骤, 在测试时候不增加计算量, 具有很强的通用性和实用性。

关键词: SAR图像; 舰船检测; 生成对抗网络; 线上难例挖掘

中图分类号: TN957.51

文献标识码: A

文章编号: 1009-5896(2019)01-0143-07

DOI: 10.11999/JEIT180050

Ship Detection in SAR images Based on Generative Adversarial Network and Online Hard Examples Mining

LI Jianwei^① QU Changwen^① PENG Shujuan^① JIANG Yuan^②

^①(Naval Aeronautical University, Yantai 264001, China)

^②(The Unit 92228 of PLA, Beijing 100044, China)

Abstract: Deep learning based ship detection method has a strict demand for the quantity and quality of the SAR image. It takes a lot of manpower and financial resources to collect the large volume of the image and make the corresponding label. In this paper, based on the existing SAR Ship Detection Dataset (SSDD), the problem of insufficient utilization of the dataset is solved. The algorithm is based on Generative Adversarial Network (GAN) and Online Hard Examples Mining (OHEM). The spatial transformation network is used to transform the feature map to generate the feature map of the ship samples with different sizes and rotation angles. This can improve the adaptability of the detector. OHEM is used to discover and make full use of the difficult sample in the process of backward propagation. The limit of positive and negative proportion of sample in the detection algorithm is removed, and the utilization ratio of the sample is improved. Experiments on the SSDD dataset prove that the above two improvements improve the performance of the detection algorithm by 1.3% and 1.0% respectively, and the combination of the two increases by 2.1%. The above two methods do not rely on the specific detection algorithm, only increase the time in training, and do not increase the amount of calculation in the test. It has very strong generality and practicability.

Key words: SAR images; Ship detection; Generative Adversarial Network (GAN); Online Hard Examples Mining (OHEM)

收稿日期: 2018-01-15; 改回日期: 2018-09-26; 网络出版: 2018-10-22

*通信作者: 李健伟 lgm_jw@163.com

基金项目: 国家自然科学基金(61571454)

Foundation Item: The National Natural Science Foundation of China (61571454)

1 引言

近年来计算机视觉领域内的目标检测技术得到了飞速的发展,这主要得益于深度学习在特征提取上的巨大优势^[1]。基于深度学习的目标检测方法可以分成两大类:第1类是以Fast R-CNN(Region-based Convolutional Neural Network)^[2,3]为代表的两阶段检测器;第2类是以YOLO^[4]和SSD^[5]为代表的单阶段检测器。两阶段检测器包括候选区域提取和对候选区域进行分类与定位两个阶段。单阶段检测器将原图划分成多个区域,直接进行分类和回归定位,具有较高的计算效率。但是整体而言以Fast R-CNN系列为代表的两阶段的检测器精度更高一些,因此本文将其作为基本检测框架。

由于深度学习在目标检测领域的巨大优势,将其应用于SAR图像舰船目标检测是大势所趋^[6-8]。针对本领域公开数据集缺失、无法统一评判的问题,本文建立了SAR舰船目标检测的数据集(SAR Ship Detection Dataset, SSDD),2.1节进行了详细介绍。该数据集目前得到了国内许多研究所和高校的使用。从数据集统计特点可以看到SAR图像中的舰船目标具有多种尺寸和观测角度,但数据集仍然难以包含各种条件下的舰船目标,这对舰船目标的检测和识别造成了很大的困难。目前较好的解决方法是尽量对数据集进行扩充,增加不同尺寸和角度的舰船目标的SAR样本,增强对其适应性,但这种方法需要消耗大量的人力、物力和财力。

通过对深度学习目标检测算法和数据集SSDD的分析,本文认为对SAR图像中的舰船目标检测尚存在以下两个问题:第一,难以检测不同角度、不同尺寸的舰船目标;第二,训练样本利用不充分,简单样本和困难样本没有区别对待,导致模型训练不充分,整体检测效果不好。

目前在深度学习领域^[9],解决问题1的方法基本上就是数据驱动的方法,即通过收集大量的各种尺度和观察角度的SAR图像,增加到数据集,通过多次训练模型即可具备对此类目标的检测能力。但是,实际收集此类目标并制作相应的标签,人力物力消耗较大。针对这种情况,本文训练一个对抗器,其通过在特征图上产生仿射变换来生成难例样本,这个对抗器会预测什么样的样本是对Fast R-CNN检测较难的。反过来,Fast R-CNN会调整自身去对这些对抗样本进行分类。随着训练的进行,目标检测器会慢慢适应对抗网络所产生的各种困难目标,本文通过这种对抗学习策略来提高检测器对各种尺寸和旋转角度目标的适应性。

解决问题2的常用方法是采用bootstrapping方法,由于bootstrapping方法需要对模型和数据集进

行交替优化更新,而由于深度学习目标检测模型的限制,对训练样本(候选窗口)进行bootstrapping需要很大的训练代价。因此,这里采用在训练时进行难样本挖掘^[10,11]的技术,在前向传播时根据不同样本的损失值将样本分成容易样本与困难样本,在进行后向传播时,对于困难样本进行重点关注,高效地实现了对样本的高利用效率。它同时也去除了Fast R-CNN中对正负样本按比例(正负样本比例1:3)采样的环节,避免少量的难例样本在采样时被忽略。

同时,由于二者在训练过程中是选择不同类型的特征,因此本文将以上两点改进进行组合,可对Fast R-CNN的检测性能有进一步的提升。

2 算法介绍

2.1 数据集

基于深度学习的目标检测技术的关键是构建合适量级的数据集及其标签。本文建立了SAR图像数据集,称作SSDD。该数据集借鉴了PASCAL VOC数据集的构建流程,包括了不同分辨率、极化、海况、大片海域和靠岸等条件的SAR图像。SSDD一共有1160个图像,2456条船,平均每幅图像有2.12条船,我们将这个数据按照7:2:1的比例划分成了训练集、测试集和验证集。

后续根据任务需要会增加样本个数,扩充数据集。相比于有9000多幅图像、20类目标的PASCAL VOC数据集,SSDD足够大到可以对舰船这一类目标进行训练。由于一些小船在低分辨率下只有极少量的像素个数,有时很难通过人眼确定是不是船。因此,我们将像素个数大于3个的视作船的像素,并对其做标注。标注所用的是开源软件“labelimg”,该软件很容易就能对船做标签 (x, y, h, w) ,其中 (x, y) 是舰船所在矩形框左上角的坐标, h 是矩形框的高度, w 是矩形框的宽度。数据一部分来源于3景RadarSat-2烟台港海域的SAR图像、小部分来自于印度维沙卡帕特南港口TerraSAR-X、剩下的大部分来自于欧空局网站下载的哨兵SAR图像。对每一景的SAR图像进行人工裁剪成大约500像素×500像素的尺寸,并进行标注。分辨率有以下几种:1 m, 3 m, 5 m, 7 m和10 m,分辨率的多样性是为了保证训练出来的模型具有更好的适应性。极化同样也具有多样性。

2.2 Fast R-CNN目标检测器

Fast R-CNN目标检测器包括两个步骤:第一,用Selective Search^[12]实现的候选区域提取;第二,用于分类与定位的卷积神经网络(Convolutional Neural Network, CNN)。网络结构有两部分组成:提取特征的卷积神经网络和RoI(Region of Interest)池化层连接着的全连接层。Fast R-CNN借

鉴了SPP-net的池化层并对其进行简化,使输入的图片大小可变,因为RoI池化层可以将不同大小的特征图重新采样成固定大小的向量,从而作为全连接层的输入。全连接的输出是每个目标类别的概率(包括背景)和目标边框的坐标。Fast R-CNN原理框图如图1所示,图中FC代表全连接层。

每个RoI被分配了真实的类别标签 u 和真实的边框回归坐标 v ,对每个RoI用多任务损失函数 L 去联合训练分类和边框回归。

$$L(p, u, t^u, v) = L_{\text{cls}}(p, u) + \lambda[u \geq 1]L_{\text{loc}}(t^u, v) \quad (1)$$

这里, $L_{\text{cls}}(p, u) = -\lg p_u$ 是真类别 u 的对数损失函数。

第2项损失函数 L_{loc} 表示类别 u 的边框回归, $v = (v_x, v_y, v_w, v_h)$ 预测的坐标为 $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$,前者表示当 $u \geq 1$ 时其值为1,反之为0,对于背景RoI,没有类别坐标的概念,所以 L_{loc} 会被忽略。对于边框回归:

$$L_{\text{loc}}(t^u, v) = \sum_{i \in \{x, y, w, h\}} s_{L_1}(t_i^u - v_i) \quad (2)$$

其中,

$$s_{L_1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & |x| \geq 1 \end{cases} \quad (3)$$

对于与真实边框交并比大于0.5的RoI视为正样本,与任一真实边框的最大交并比介于0.1和0.5之间的视为负样本,余下的RoI被抛弃不参与训练模型。为了处理候选区域提取得到的数据类别不平衡的问题,作者在每个训练的小批量控制正样本与负样本数量为1:3,这是有效训练Fast R-CNN的关键,增大或减小这个比例都会或多或少地降低检测准确率。但简单粗暴的丢弃掉部分RoI不是最好的方法,因为它会忽略掉一些不常见的但是很重要的困难样本。

2.3 对抗空间变换网络

生成对抗网(Generative Adversarial Networks, GAN)^[13]是一种概率生成模型,就是找出给定观测数据内部的统计规律,并且能够基于所得到的概率分布模型,产生全新的,与观测数据类似的

数据。为了学习生成器关于数据 x 上的分布 p_g ,定义输入噪声的先验变量 $p_z(z)$,然后使用 $G(z; \theta_g)$ 来代表数据空间的映射。这里 G 是一个由含有参数 θ_g 的多层感知机表示的可微函数。定义一个多层感知机 $D(x; \theta_d)$ 用来输出一个单独的标量, θ_d 是多层感知机的参数。 $D(x)$ 代表 x 来自于真实数据分布而不是生成数据的概率,本文训练 D 来最大化分配正确标签给不管是来自于训练样例还是 G 生成的样例的概率,同时训练 G 来最小化 $\lg(1 - D(G(z)))$ 。换句话说, D 和 G 的训练是关于值函数 $V(D, G)$ 的极小极大化的二人博弈问题:

$$\min_G \max_D V(D, G) = E_{x \sim p_{\text{data}}(x)} [\lg D(x)] + E_{z \sim p_z(z)} [\lg [1 - D(G(z))]] \quad (4)$$

训练过程中, GAN采用了一种非常直接的交替优化方式,它可以分为两个阶段:第1个阶段是固定判别模型 D ,然后优化生成模型 G ,使得判别模型的准确率尽量降低,而另一个阶段是固定生成模型 G ,来提高判别模型的准确率。

生成图像(或像素)已经有很多方法,一种学习这些稀有样本的方法是通过从长拖尾分布(指部分尺寸和旋转角度的样本在数据集的数量比较少)中产生真实图像,但这不是一个可行的方案,因为这需要大量的稀有样本来训练GAN。借鉴文献[14]的思路在特征图上产生不同条件下的难例样本,不产生像素本身,这对于GAN模型来说会更加容易实现。

卷积神经网络功能强大,但是对输入图像的空间变换(尺度和旋转角度等)不变性性能一般,为了提高对空间变换的适应性,目前一些算法会对数据做一些数据扩充的预处理比如切分、放大和旋转等。空间变换网络(Spatial Transformer Network, STN)^[15]模型允许在网络中对图像进行操作,网络能够产生具有空间变换特性的特征图,使训练处的模型不用收集大量的样本就具有非常强的空间变换不变性。空间变换网络由定位网络、网格产生器和采样器3个组成部分,如图2所示。以特征图 U 作为

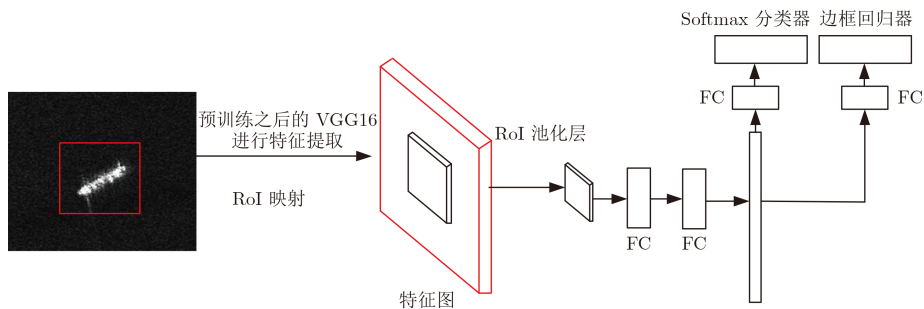


图1 Fast R-CNN原理示意图

输入,定位网络会估计变形参数(例如旋转角度、变换距离和尺度因子)。这些变量会被作为网格产生器和采样器的输入,输出是变形之后的特征图 V 。本文仅仅需要学习定位网络的参数,STN的一个重要贡献是使整个过程可微,因此定位网络可以直接通过后向传播,优化分类损失函数。在输入数据的空间差异较大的情况下,这个网络可以加入到现有的卷积神经网络中,使得神经网络能够主动对特征进行变换,学习数据在各种空间变换下的不变性,提高分类的准确性。

本文引入对抗空间变换网络,中心思想是在目标特征上产生变形,使检测器Fast R-CNN较难检测此类样本。本文的网络是基于使分类更加容易,能适应各种角度变化和尺度变化等仿射变换的空间变换网络,通过与对抗空间变换网络进行斗争,可以训练一个更好的检测器,使得其检测变形物体更加稳健,对抗空间变换网络结构如图2所示。

本文将对抗空间变换网络与Fast R-CNN进行联合训练,为了训练检测器,RoI池化之后的特征 U 首先被STN转换成 V ,然后前向传播到更高层计算Softmax损失。为了训练对抗空间变换网络,优化检测器使它会对前景和背景进行分类,由于空间变换网络是可微的,可以直接用分类损失进行后向传播,然后微调预训练之后的定位网络的参数。

用 $A(X)$ 表示特征 X 下的对抗网络,用于产生新的对抗样本。检测器的损失函数不变,只不过是每个小批量会包含更少的原始样本,且多一些产生的对抗样本。但是,对抗网络必须学习到使检测器失效的特征。通过式(5)的损失函数训练对抗网络:

$$L_A = -L_{st}(L_{cls}(A(X)), C) \quad (5)$$

式中, X 是其中的一个候选区域,检测器有两个输出: L_{cls} 代表分类器的输出, L_{loc} 代表预测边框的输出。如果对抗网络产生的特征对于检测器来说容易进行分类,给对抗网络较高的损失,如果对抗网络

生成的特征对于检测器来说不容易分类,给检测器较高的损失值,给对抗网络较低损失值。

由于遥感图像各个角度都有可能存在且上下翻转也是很正常的,因此将旋转角度不进行限制。同时,也不会把所有的特征图在同一个方向旋转,对特征图的每个通道分别进行旋转,对每个特征图估计不同的选择角度(这是因为每个通道对应其单独类型特征的激活值),如果对所有的特征图用同一个旋转角度,对抗空间变换网络会经常预测最大的角度,通过用多个不同角度而不是一个,增加了任务的复杂性,使产生的对抗样本更加丰富。

2.4 线上难例挖掘(OHEM)

Bootstrapping^[16]利用有限的样本经由多次重复抽样,重新建立起足以代表母体样本分布的新样本,可以提高对数据集样本的利用率,使模型得到更充分的训练,使得到的检测器鲁棒性更强,在多种机器学习方法中得到了大量的应用,R-CNN和SPP-net等算法在后面的SVM中也使用了该技术。

但是,目前较先进的目标检测器没有使用bootstrapping,由于bootstrapping方法需要交替训练,即固定数据集来训练模型,固定模型寻找难例样本更新数据集,而用SGD(Stochastic Gradient Descent)训练CNN需要几十万次SGD,在进行迭代训练时,固定模型参数去选择难例样本会使整个训练过程变的极其慢。因此,采用OHEM(Online Hard Examples Mining)方法用以训练基于CNN的目标检测模型。算法对SGD进行了微小的改动,训练样本根据每个例子的损失值进行重新采样,选择那些高损失值的难例样本用于后向传播,其它样本不进行后向传播。其后向传播计算梯度仅仅利用小部分候选框,因此这个方法是高效的。

传统难例挖掘算法包括两个步骤:模型和数据集交替优化,上述交替方法可以通过线上SGD的方式与Fast R-CNN结合。之所以称为线上,主要是

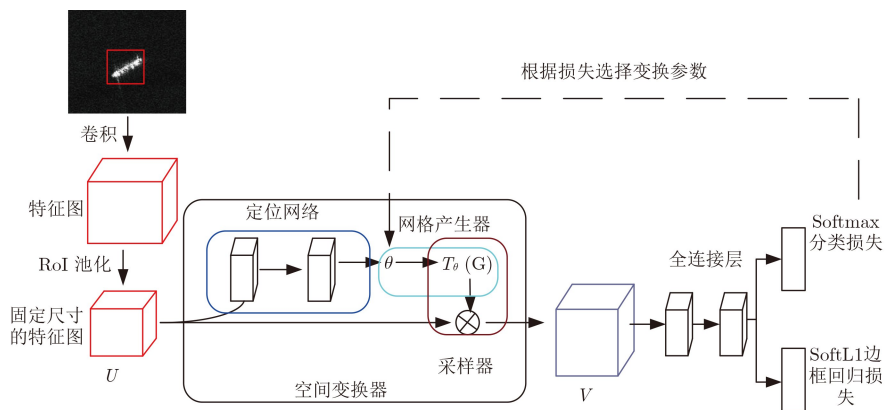


图2 对抗空间变换网络示意图

因为操作是在后向传播过程中进行的, 不像以前的方法交替优化网络模型和数据。

在原有的Fast-RCNN里的分类损失 L_{cls} 里面对所有的候选窗口计算其损失值(包括分类损失 L_{cls} 和边框回归损失 L_{loc}), 根据分类损失 L_{cls} 的大小, 从高到低排序, 并利用非极大值抑制(Non-Maximum Suppression, NMS), 选出前64个难例样本(候选框), 对它们进行反向传播, 而其他的候选框的梯度不进行回传, 流程图如图3所示。

进行NMS操作的原因是由于特征图分辨率的降低, 交叠的RoI会在特征图上映射成同一个区域, 因此导致损失值的重复计算。为了处理这类问题, 用标准的非极大值抑制去除重复的计算。输入一系列RoI和它们的损失, NMS通过迭代式选择损失值较大的RoI, 同时去除那些与所选择的区域重叠的, 损失值较低的区域, 本文设置交并比的门限为0.7, 抑制那些相交程度很高的RoI。将选出的64个候选窗口(这里看成难例样本), 进行回传, 计算64个候选窗口的损失值, 并更新卷积网络的参数。

大部分的前向计算是共享卷积层的, 所以计算量不大, 同时由于后向传播仅仅对那些少量的损失比较大的RoI, 因此也不会增加计算量。

上述过程不需要像Fast R-CNN那样对每个小批量内的数据做类别均衡(前景与背景的比例为1:3)。假设一个小批量内类别不均衡, 舰船目标被忽略掉, 那么它的损失值会增加, 直到含有舰船目标的候选区域被选择为止。即简单样本的分类损失值 L_{cls} 较低, 对梯度贡献较小, 训练会自动地聚焦于那些难例样本。

2.5 有限样本下检测方法

对抗空间变换网络可以产生数据集中不存在的样本的特征图, 而OHEM只能利用数据集存在的样

本, 而且由于OHEM从真实图像获得, 所以有更真实的特征。由于二者在训练过程中是选择不同类型的特征, 二者是可以互补的, 即通过训练之后, 检测性能高于二者任一。因此先通过联合训练方式训练对抗空间变换网络和Fast R-CNN, 在训练达到稳定之后, 再训练OHEM和Fast R-CNN。对于第1次训练, 本文在SSDD数据集上用SGD方法训练 10^5 次, 前 8×10^4 次学习率是0.0010, 后 2×10^4 次学习率是0.0001。对于第2个模型, 本文在SSDD数据集上用SGD方法训练 8×10^4 次, 前 6×10^4 次学习率为0.0010, 后 2×10^4 次学习率是0.0001。

3 实验结果与分析

实验环境: 本文利用Caffe框架在SSDD上去训练检测器, VGG16是基本架构, 已在ImageNet进行了预训练, 计算环境是64位Ubuntu 14.04操作系统, 计算机CPU Intel(R) Core (TM) i7-6770K @ 4.00 GHz×8和NVIDIA GTX1080 GPU (8G 内存) × 4, 用CUDA8.0和cuDNN5.0加速。

评价指标是mAP(mean Average Precision), 定义为

$$mAP = \int_0^1 P(R) dR \quad (6)$$

其中 R 为召回率, P 为精度。

为了对所提方法性能进行验证, 我们在数据集SSDD上进行试验, 需要验证的方法有标准的Fast R-CNN、标准的Fast R-CNN+GAN、标准的Fast R-CNN+OHEM和标准的Fast R-CNN+GAN+OHEM 4种方法, 评价指标包括mAP、每次迭代(batchsize为128)所需训练时间和每幅图像所需的测试时间, CNN模型采用的是VGG16, 实验结果如表1所示。

由表1可以看出: 对于标准的Fast R-CNN检测

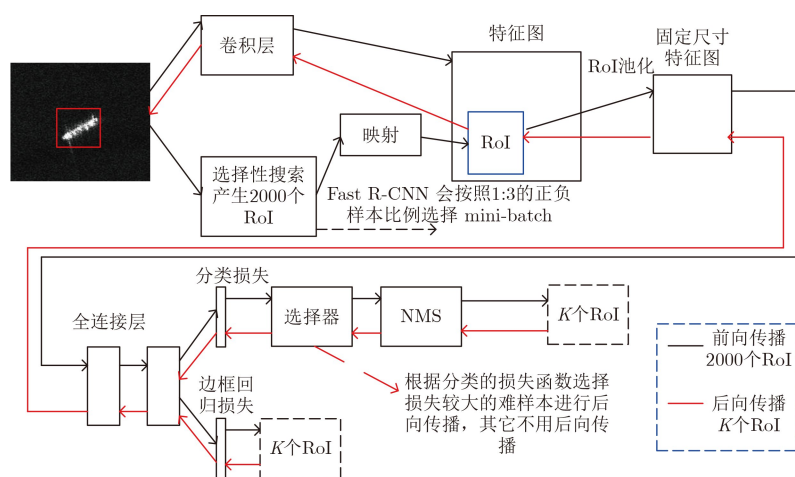


图3 线上难样本挖掘流程(K=64)

表 1 4种方法检测性能

方法	mAP (%)	训练时间 (s)	测试时间 (s)
标准的 Fast R-CNN	68.0	0.610	0.328
标准的 Fast R-CNN+ GAN	69.4	0.823	0.326
标准的 Fast R-CNN+OHEM	69.1	1.152	0.321
标准的 Fast R-CNN+GAN+OHEM	70.2	2.109	0.330

算法,其batchsize为128,即每次迭代有128个候选样本,检测的mAP为68.0%,每次迭代所需时间是0.61 s,对于每幅SAR图像(大约500×300像素),检测时间为0.328 s。而通过对抗空间变换网络进行训练之后,mAP增加了1.4%,到69.4%,每次迭代所需要的时间变为0.326,测试时间几乎没有发生变化。通过增加OHEM环节改进SGD过程,得到的检测结果的mAP为69.1%增加了1.1%,训练时间1.152 s,测试时间几乎没有发生变化。通过同时使用空间对抗网络和OHEM技术我们发现,检测的mAP增加到了70.2%,每次迭代所需时间也增加到了2.109 s,同时检测时间几乎不变。提出的两点改进,对检测算法的测试时间没有发生变化,这是因

为仅仅对训练过程进行改进,提高对数据集的利用效率,本文将训练好的Fast R-CNN模型存储下来,在进行测试时候,没有用到空间对抗网络和OHEM,因此测试时间没有发生变化。

图4是从数据集SSDD中选出的几个不同尺寸图像,从图中可以看出,检测算法可以适应这些样本。图4(g),图4(h)和图4(i)是从数据集SSDD中选出的几个难例样本的检测结果,从图中可以看出,检测算法可检测出大部分的难例样本(主要是一些靠岸舰船目标),尤其是对于港口区域不需要预处理、海陆分割杂波建模等中间环节,可直接检测出靠岸的舰船目标。图4(c)和图4(g)中是典型的靠岸目标,由于训练集包含了大量的靠岸目标,所以训练出来的模型可以检测到这类靠岸的舰船,相比于传统方法一步一步地处理(海陆分割等),本文方法可进行端到端的处理。而且传统方法基本上只对大片海域效果好,本文方法不仅对大片海域效果好,而且还可以端到端的检测到靠岸目标。需要指出的是,在图4(d)存在虚警(最上面那个框),图4(h)存在漏检的情况(左上那个框),本文通过平均精度来衡量检测性能的好坏,会不可避免的存在一些虚

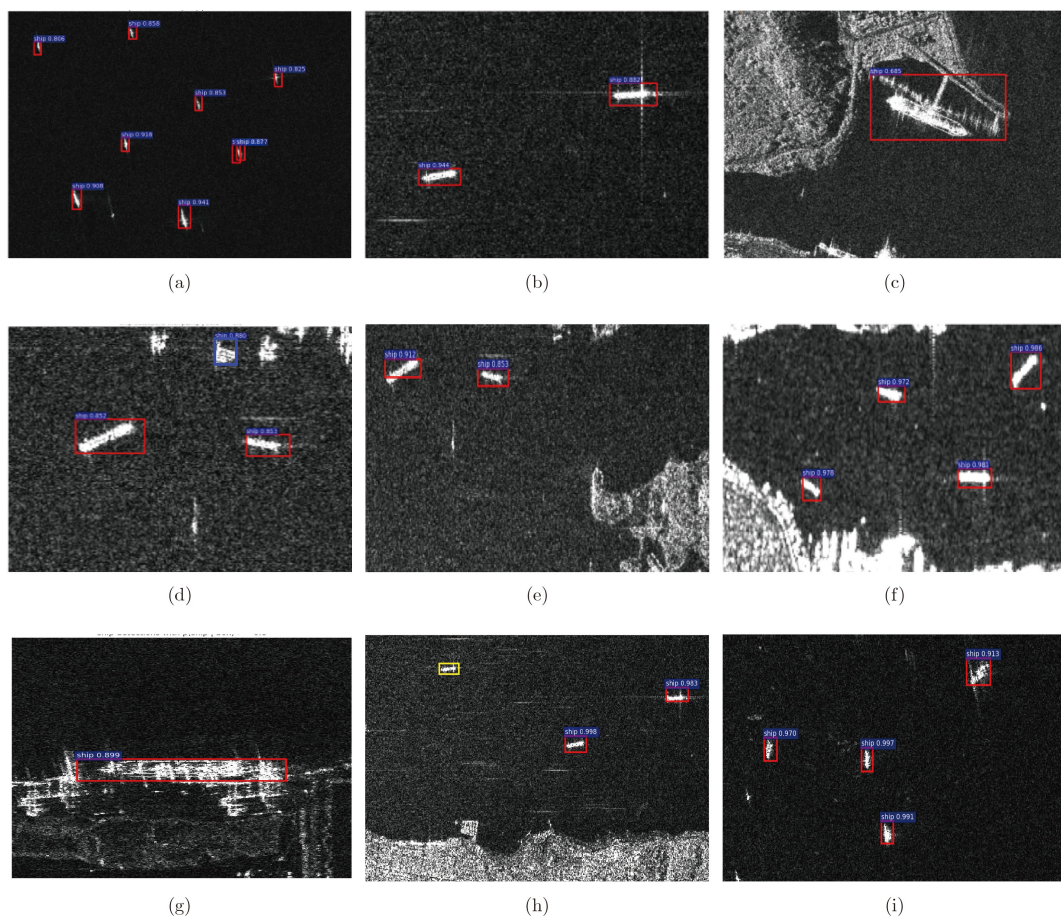


图 4 数据集SSDD中不同尺度和角度舰船检测结果

警和漏警, 不过这不能代表整体情况。图中框上的字表示目标类别, 这里全是ship(舰船)。

4 结论

本文提出了在不对数据集样本进行扩充的基础上, 通过提高样本利用率的方法来提升检测性能的方法。通过对抗空间变换网络在特征图上生成各种尺寸和旋转角度的检测目标的特征图, 与检测器Fast R-CNN形成对抗训练, 最终使检测器能够提高对各种尺寸和旋转角度样本的鲁棒性; 通过在SGD过程中使用OHEM技术, 自动选择难例样本进行回传计算梯度, 使检测器能够更充分地利用难例样本。通过在数据集SSDD上的实验, 验证了以上两项技术对检测结果的提升。由于两点改进只是对训练阶段进行改进, 对预测阶段没有影响, 因此不会增加检测算法的耗时。

参考文献

- [1] KRIZHEVSKY A, SUTSKEVER I and HINTON G E. ImageNet classification with deep convolutional neural networks[C]. International Conference on Neural Information Processing Systems. Nevada, USA, 2012: 1097–1105. doi: [10.1145/3065386](https://doi.org/10.1145/3065386).
- [2] GIRSHICK, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 580–587. doi: [10.1109/CVPR.2014.81](https://doi.org/10.1109/CVPR.2014.81).
- [3] REN Shaoqing, HE Kaiming, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 39(6): 1137–1149. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
- [4] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 779–788. doi: [10.1109/CVPR.2016.91](https://doi.org/10.1109/CVPR.2016.91).
- [5] LIU Wei, ANGUELOV D, ERHAN D, *et al.* SSD: Single shot multiBox detector[C]. IEEE European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 21–37.
- [6] 王思雨, 高鑫, 孙皓, 等. 基于卷积神经网络的高分辨率SAR图像飞机目标检测方法[J]. 雷达学报, 2017, 6(2): 195–203. doi: [10.12000/JR17009](https://doi.org/10.12000/JR17009).
Wang Siyu, Gao Xin, Sun Hao, *et al.* An aircraft detection method based on convolutional neural networks in high-resolution SAR images[J]. *Journal of Radars*, 2017, 6(2): 195–203. doi: [10.12000/JR17009](https://doi.org/10.12000/JR17009).
- [7] 徐丰, 王海鹏, 金亚秋. 深度学习在SAR目标识别与地物分类中的应用[J]. 雷达学报, 2017, 6(2): 136–148. doi: [10.12000/JR16130](https://doi.org/10.12000/JR16130).
- Xu Feng, Wang Haipeng, and Jin Yaqiu. Deep learning as applied in SAR target recognition and terrain classification[J]. *Journal of Radars*, 2017, 6(2): 136–148. doi: [10.12000/JR16130](https://doi.org/10.12000/JR16130).
- [8] 刘泽宇, 柳彬, 郭炜炜, 等. 高分三号NSC模式SAR图像舰船目标检测初探[J]. 雷达学报, 2017, 6(5): 473–482. doi: [10.12000/JR17059](https://doi.org/10.12000/JR17059).
- Liu Zeyu, Liu Bin, Guo Weiwei *et al.* Ship detection in GF-3 NSC mode SAR images[J]. *Journal of Radars*, 2017, 6(5): 473–482. doi: [10.12000/JR17059](https://doi.org/10.12000/JR17059).
- [9] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770–778. doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [10] HUANG Gao, LIU Zhuang, WEINBERGER K Q, *et al.* Densely connected convolutional networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, 2017: 4700–4708. doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243).
- [11] SUNG K K and POGGIO T. Example-based learning for view-based human face detection[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2002, 20(1): 39–51. doi: [10.1109/34.655648](https://doi.org/10.1109/34.655648).
- [12] SHRIVASTAVA A, GUPTA A, and GIRSHICK R. Training region-based object detectors with online hard example mining[C]. IEEE Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 761–769. doi: [10.1109/CVPR.2016.89](https://doi.org/10.1109/CVPR.2016.89).
- [13] UIJLINGS J R R, SANDE K, GEVES T, *et al.* Selective search for object recognition[J]. *International Journal of Computer Vision*, 2013, 104(2): 154–171. doi: [10.1007/s11263-013-0620-5](https://doi.org/10.1007/s11263-013-0620-5).
- [14] WANG Xiaolong, SHRIVASTAVA A, and GUPTA A. A-Fast-RCNN: Hard positive generation via adversary for object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, 2017. doi: [10.1109/CVPR.2017.324](https://doi.org/10.1109/CVPR.2017.324).
- [15] JADERBERG M, KAREN S, and ANDREW Z. Spatial transformer networks[C]. Advances in Neural Information Processing Systems, Montreal, Canada, 2015: 2017–2025.
- [16] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial nets[C]. International Conference on Neural Information Processing Systems, Montreal, Canada, 2014: 2672–2680.

李健伟: 男, 1989年生, 博士生, 研究方向为SAR图像处理、机器学习及深度学习。

曲长文: 男, 1964年生, 教授, 研究方向为雷达信号处理, 信息对抗, 信号与信息处理等。

彭书娟: 女, 1980年生, 博士生, 研究方向为SAR图像处理。