

基于深度学习的复杂沙漠背景 SAR 目标检测

夏 勇^{1,2}, 田西兰^{1,2}, 常 沛^{1,2}, 蔡红军^{1,2}

(1. 中国电子科技集团公司第三十八研究所, 安徽合肥 230088;
2. 孔径阵列与空间探测安徽省重点实验室, 安徽合肥 230088)

摘 要: SAR 目标检测, 因成像场景大、背景复杂多变而极具挑战。传统基于恒虚警率的 SAR 目标检测方法极易受背景干扰。针对上述问题, 提出一种基于深度学习的复杂沙漠背景 SAR 目标端对端检测识别系统。即采用小规模沙漠背景下的 SAR 图像数据对 Faster-RCNN 网络进行迁移训练, 一体化完成典型目标的检测与识别。基于合成数据集 Desert-SAR 的试验结果表明, 与传统方法相比, 该方法检测速度更快、准确率更高、鲁棒性更强。

关键词: 深度学习; 沙漠背景; 合成孔径雷达; 目标检测

中图分类号: TN958; TP75 文献标志码: A 文章编号: 1672-2337(2019)03-0305-05

SAR Target Detection in Complex Desert Background Images Based on Deep Learning

XIA Yong^{1,2}, TIAN Xilan^{1,2}, CHANG Pei^{1,2}, CAI Hongjun^{1,2}

(1. The 38th Research Institute of China Electronics Technology Group Corporation, Hefei 230088, China;
2. Key Laboratory of Aperture Array and Space Application, Hefei 230088, China)

Abstract: Target detection in synthetic aperture radar (SAR) image is a challenge due to the large-scale and complex imaging scene. The classical methods based on CFAR are sensible to imaging scene. Aiming at this problem, we propose an end-to-end target detection method for SAR image in desert scene based on deep learning. That is, the transfer learning is employed to adjust the Faster-RCNN network for optical image to the SAR image. Experimental results of the Dessert-SAR data set show that the proposed method can achieve faster detection speed, higher accuracy and robustness compared with the classical ones.

Key words: deep learning; desert background; synthetic aperture radar; target detection

0 引 言

合成孔径雷达(SAR)是一种微波成像传感器,能够获得地物的二维图像,是一种重要的对地观测手段。其具有全天时全天候等多种特点,可以在恶劣的战场条件下完成侦察任务。它在国民经济和国防安全等领域有着极其广泛的应用,比如弹道导弹防御系统、灾情监测和地质矿物资源勘探等任务。随着雷达技术的不断发展,目标识别在雷达应用中的需求愈发强烈。SAR 目标自动检测识别已成为研究热点。随着 SAR 系统分辨

率的提高,有助于获得更加精细的目标信息。但是,图像分辨率的不断提高也给 SAR 目标检测识别带来了新的困难:场景变得更加复杂,并且数据量巨大,尤其是复杂沙漠背景图像。

1 SAR 目标检测识别国内外研究

在过去的数十年间,涌现出很多 SAR 目标自动检测识别的算法。传统方法主要分为两个步骤:目标检测和分类识别。

检测的目的是提取图像中可能存在的目标,最常用的是恒虚警率(Constant False Alarm

Rate, CFAR) 检测器^[1]。CFAR 算法一般含有 3 个输入因素:虚警率、CFAR 检测器和杂波背景统计分布模型。该算法在目标与背景具有较高的对比度且场景简单的情况下,能较好地从中分离出目标。但是,当面对种类繁多、散射特性迥异的杂波时,其检测性能会有所下降,得到的往往不仅仅包括真实目标,比如坦克、武装车辆;还有许多误检,如树木、建筑物、桥梁和普通车辆等。虚警过多会给后续操作带来更大的运算量,并会影响最终的识别结果;在接下来的分类任务中,需要提取许多不同的特征^[2],训练特定的分类器,用于去除误检以实现目标的精确识别。

当前 SAR 目标分类主要有两种方法:基于模板匹配的方法和基于机器学习的方法。其中,基于模板匹配的方法应用最为广泛。该方法对每一类生成一个模板分类器,将待识别目标输入到这些分类器中,以输出结果最优者为识别结果。基于模板匹配算法的计算效率低下,此外需要大容量存储设备,该方法在实时应用上存在较大挑战。

随着人工神经网络(Artificial Neural Network, ANN)和支持向量机^[3](Support Vector Machine, SVM)的出现,机器学习的方法已广泛应用于 SAR 目标自动检测识别任务中。不同于模板匹配,分类问题的首要任务被转化为提取一系列不同的特征去表征目标以及利用这些特征向量去训练分类器。通过最小化损失函数,可以找到训练集的最佳决策边界。这些手工提取的特征需要对不同种类具有较好的可区分性,且对不同场景有较强的鲁棒性,如何提取合适的特征成为其主要难点。

当前,随着深度学习的火热发展,卷积神经网络(Convolutional Neural Network, CNN)在雷达目标识别领域已取得广泛应用。深度学习与传统模式识别方法的最大不同点是可以自动从大量数据中学习特征,而不再需要设计手工特征。王思雨等^[4]将 CNN 引入到 SAR 飞机目标自动检测识别系统中,并使用 Terra SAR-X 数据集对算法进行评估。Ren 等^[5]提出 Faster-RCNN 用于自然场景图像目标检测与识别,并在 VOC 数据集上取得良好的效果。

在上述研究基础上,提出端对端沙漠背景 SAR 图像目标检测与识别,并且在合成沙漠背景

图像数据集上取得了很好的效果。该模型的框架如图 1 所示。

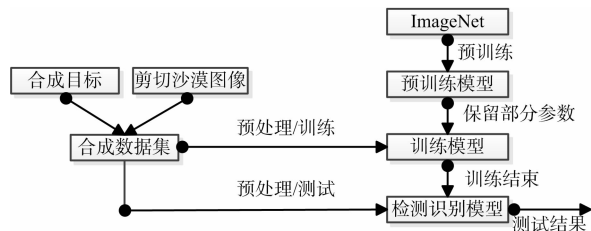


图 1 沙漠背景 SAR 目标自动检测识别框架流程

主要工作如下:

1) 为了提升目标检测识别的准确性,并提升检测的检测速度,基于深度学习提出了端对端的目标检测与识别算法。

2) 为了克服数据集的有限性,通过人工合成,得到沙漠背景数据集 Desert-SAR,用于训练和评估目标检测算法。通过将合成目标与复杂沙漠背景 SAR 图像融合得到 Desert-SAR,实现数据集的扩充。

此外,Desert-SAR 数据集的合成方法在第 2 节进行了介绍;第 3 节阐述了本方法的模型和训练过程;第 4 节展示了本方法在 Desert-SAR 数据集上的效果和分析;结论和未来工作在第 5 节进行了介绍。

2 Desert-SAR 数据集的生成

SAR 目标检测识别在很多场景中已取得广泛应用,但针对沙漠背景目标检测存在的问题仍亟待解决。当前沙漠背景的 SAR 图像规模较小,为此,通过人工合成的方法生成 Desert-SAR 数据集,用于检测识别网络模型的训练和测试。

传统训练样本的合成主要将 MSTAR 数据集的目标切片与背景图片融合,由于受到数据集样本数量、质量的限制,会导致目标数量不充分。此外,直接贴片会引入周围信息,导致融合的图像失真,如图 2 所示。通过对比发现,本方法更接近真实情况。

车辆目标通过 SAR 仿真成像软件合成,目标包含方位角、俯仰角、方位向分辨率、距离向分辨率等参数,可以根据需要生成不同种类和方向的目标。图像为复图像,对应每个像素包含实部和虚部两部分。示例模拟图像如图 3 所示。

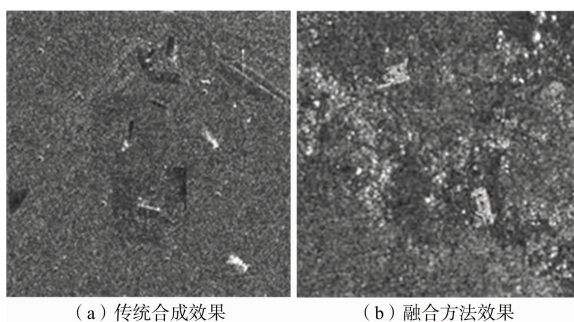


图2 传统与融合方法效果对比

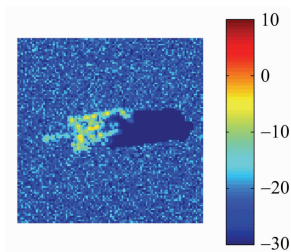


图3 示例模拟图像

通过仿真软件共生成不同方向的坦克目标144个,部分目标样本示意图如图4所示。图像块大小为 128×128 ,目标周围的像素值均为零,不会引入周边的信息,从而保证合成图片有着更好的效果。



图4 不同角度目标样本示意图

图5展示了从大型沙漠SAR图像中裁剪的复杂沙漠图像作为背景,共剪切10种大小为 $1\,100 \times 1\,300$ 的复杂沙漠背景图片,其中包含石块、道路等

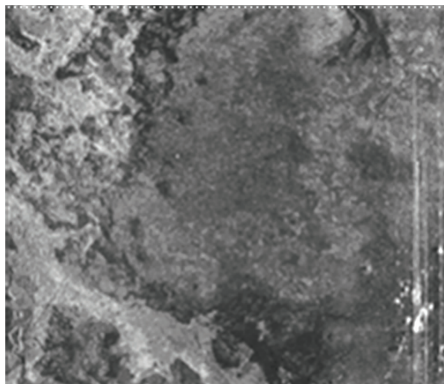


图5 沙漠背景样本图

各种复杂干扰因素。

随机在沙漠背景图片中放入5~10个合成目标,每个沙漠背景图生成100个样本。如此,共形成1 000张沙漠背景SAR图像的Desert-SAR数据集。如图6所示,展示了合成图像及其标注效果。

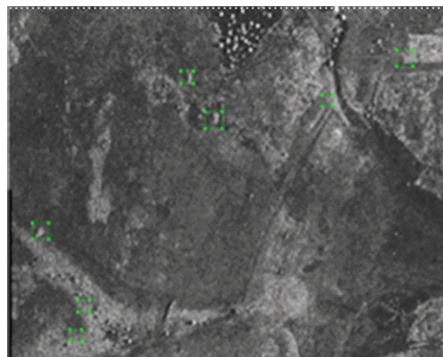


图6 合成沙漠背景 SAR 图像及标注

3 沙漠背景 SAR 图像目标检测

模型训练采用全监督训练模式,整个训练过程分为两部分:第一部分利用迁移学习对网络模型的部分参数进行预训练;第二部分利用合成的Desert-SAR数据集对模型进行再训练,完成对沙漠背景SAR图像中对目标的检测。

3.1 网络模型基本结构

该网络模型基于Faster-RCNN,其基本网络结构如图7所示。特征提取层通过多次卷积和池化操作,提取输入图像的特征图,特征提取层包含VGG网络和ZF网络两种,ZF网络是VGG网络的简化版。候选提取层在特征图上滑窗,提取候选框,为每个像素点设置9个大小长宽比不同的锚(anchor),作为基础候选框,并结合边框回归初步得到目标候选框。其中,锚的基础尺寸分别为 $[8, 16, 32]$,对应长宽比分别为 $[1:1, 1:2, 2:1]$ 。特征图经过多次池化,相对于原图缩小了一定比例,Base_size控制缩放比例,将候选框映射到原图对应的目标中。由于目标相对于图像而言较小,故需要对Base_size的大小进行适当修改,提升其对小目标的检测能力。后续的分类识别层对候选目标进行精细分类识别,并进一步对边框位置回归,最终实现对沙漠背景SAR图像的目标检测。

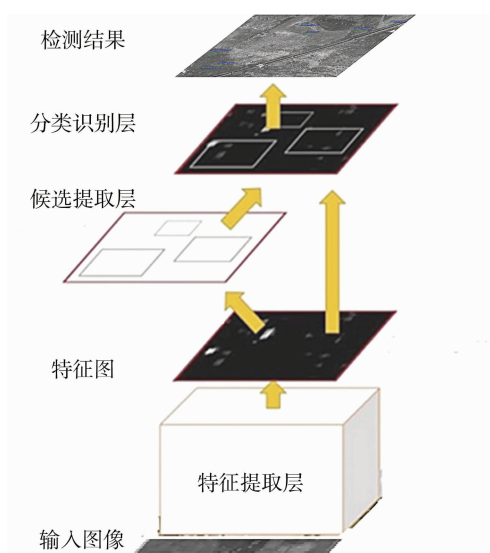


图 7 网络模型结构图

3.2 迁移学习进行预训练

迁移学习的定义是:运用已存在的知识对不同但相关领域问题进行求解的一种机器学习方法^[6]。对于卷积神经网络,迁移学习就是将在特定数据集上学到的“知识”应用到新的领域中。由于 SAR 图像的数据相对较少,无法充分从零开始训练网络模型,故引入迁移学习。训练过程中,迁移学习使用的具体步骤如下:

- 1) 利用大型数据集 ILSVRC-2012 对网络模型的随机初始化参数进行预训练。
- 2) 保留训练好的特征提取层参数,利用 Desert-SAR 数据集对候选提取层和分类识别层进行再训练。
- 3) 利用再训练好模型完成目标检测和识别。

预训练是光学图像,与 SAR 图像有差异,但通过大量训练使得模型参数拟合到可以检测识别目标,其效果远远优于初始化参数,并能大幅减少网络的训练开销。如图 8 所示,采用迁移学习的方法(虚线)与直接训练的方法(实线)相比明显加快了网络的收敛速度,而且收敛后的损失值要小于直接训练值。其次,该方法有效避免了网络对大量训练样本的需要,大大提升了模型的鲁棒性和泛化能力。

4 实验结果和分析

本方法在 Desert-SAR 数据集上进行训练和

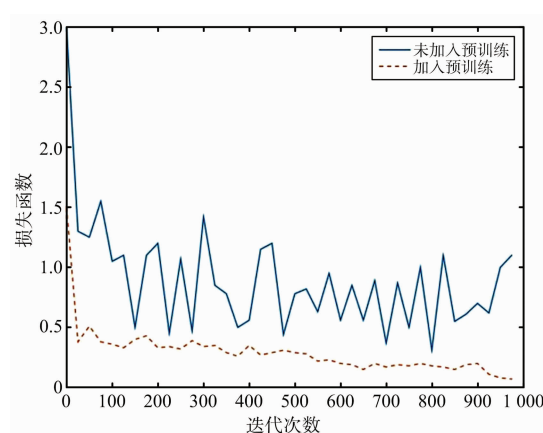


图 8 损失函数对比图

测试。每种背景图取 80 个,共 800 张图像作为训练集,剩余的 200 张图像作为测试集。

4.1 实验细节与评价标准

实验运行环境为 Windows7 系统, Intel(R) CPU E5, 16G RAM, NVIDIA K40 GPU 的工作站。采用平均准度(Average Precision, AP)作为检测效果评价标准,类别 AP 表示 P-R 曲线所围成下面积。AP 解决了传统评估标准查全率(Recall, R)、查准率(Precision, P)和 F -measure 的单点值局限性。因此,AP 可以更有效地综合评估算法的有效性和准确性。

对于目标 C , 查全率 $P = \text{图片中目标 } C \text{ 识别正确个数} / C \text{ 类别的总个数}$, 公式为

$$P_c = \frac{N(\text{TruePositives})_c}{N(\text{TotalObjects})_c}$$

对于目标 C , 查全率 $AP = \text{每张图的 } P \text{ 求和} / \text{含有类别 } C \text{ 的图片个数}$, 其公式为

$$AP = \frac{\sum P_c}{N(\text{TotalImages})_c}$$

4.2 实验结果与分析

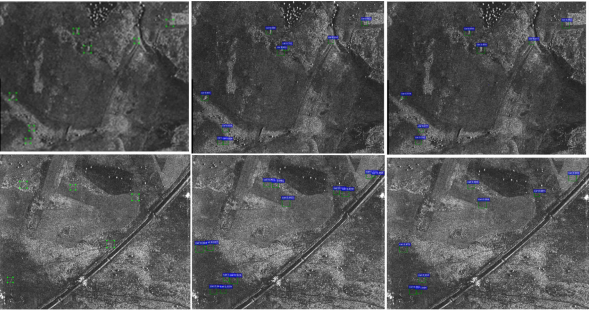
如表 1 所示,展示了本方法在 Desert-SAR 数据集上,不同参数对检测效果的影响。

表 1 各种参数对检测效果的影响

实验	预训练	模型	Base_size	AP	速度/s
T_1	无	ZF	8	0.05	0.1
T_2	有	ZF	8	0.57	0.1
T_3	有	ZF	16	0.52	0.1
T_4	有	VGG	8	0.60	0.2

通过对比 T_1 和 T_2 可以发现,采用预训练模型可极大地提升目标检测的性能。通过对比 T_2 和 T_3 ,可以发现锚的 Base_size 越大反而不利于小目标检测。通过对比 T_2 和 T_4 ,可以发现特征提取层采用 VGG 网络的检测效果优于 ZF 网络,但是其检测速度大约是 ZF 网络的一半,采取实验 T_4 的参数作为最终的检测效果。

图 9 展示了基于不同特征提取层训练得到的模型对沙漠背景 SAR 图像车辆目标的检测效果,图 9(a)、图 9(b)、图 9(c)分别表示标注图像、基于 ZF 网络的检测效果和基于 VGG 网络的检测效果。表 2 具体统计了这部分图片的检测效果。通过对比可以发现,基于 VGG 网络的检测效果优于基于 ZF 网络的效果,其可以有效降低误检的数量,这是因为 VGG 网络具有更好的特征提取能力,进而提升了其检测识别效果。



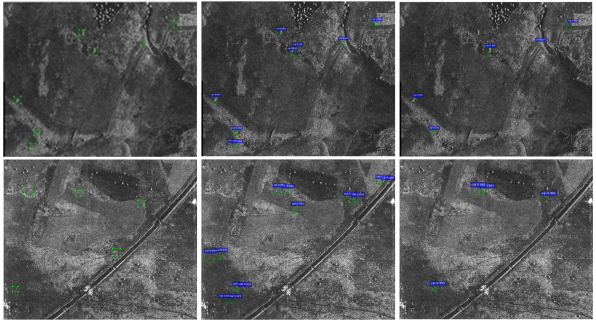
(a) 标注图像 (b) ZF 的网络效果 (c) VGG 的网络效果
图 9 部分图片不同模型的检测效果对比

表 2 部分图片不同模型的效果统计

网络	场景	目标数	检测数	虚警数	漏检数
ZF	场景 1	7	7	2	0
	场景 2	5	1	12	4
VGG	场景 1	7	7	0	0
	场景 2	5	1	8	4

图 10 展示了 VGG 网络 Base_size 不同时训练得到的模型对沙漠背景 SAR 图像车辆目标的检测效果。图 10(a)、图 10(b)、图 10(c)分别表示标注图像、锚 Base_size 为 8 的检测效果和锚 Base_size 为 16 的检测效果。表 3 具体统计了这部分图片的检测效果。通过对比可以发现,锚 Base_size 为 8 的时候可以检测到更多的真实目标,具有更好的检测效果。这是因为 SAR 目标在图像中相对较小,预设的锚主要针对自然场景中的一般目标,相对于 SAR 目标尺寸过大。因此,针对不同检测任务,为锚设置合适尺寸可有效提升检测

效果。



(a) 标注图像 (b) Base_size 为 8 效果 (c) Base_size 为 16 效果
图 10 不同 Base_size 的检测效果对比

表 3 不同 Base_size 的效果统计

Base_size	场景	目标数	检测数	虚警	漏检
8	场景 1	7	7	2	0
	场景 2	5	1	12	4
16	场景 1	7	5	0	2
	场景 2	5	0	4	5

5 结束语

提出了一种基于深度学习的端对端的复杂沙漠背景 SAR 目标自动检测识别系统。此外,提出了一种新的数据合成方法,生成沙漠背景 SAR 图像数据集 Desert-SAR 用于训练和测试。大量实验证明本文的方法不仅能有效提升复杂沙漠背景 SAR 目标检测识别的准确率,还能提升检测速度。然而,本文方法对一些特别小的目标、特别低的对比度和类似目标物体的检测效果较差。下一步研究将优化网络结构,进一步提升该系统的鲁棒性。

参考文献:

[1] 梁幸源. SAR 图像 CFAR 检测的快速算法研究[J]. 数字技术与应用,2016(10):134.
[2] 甄勇,刘伟,陈建宏,等. 高分辨率 SAR 图像舰船目标几何结构特征提取[J]. 信号处理,2016,32(4): 424-429.
[3] ZHAO Q, PRINCIPE J C. Support Vector Machines for SAR Automatic Target Recognition[J]. IEEE Trans on Aerospace and Electronic Systems, 2001,37(2):643-654.
[4] 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.