A Feature Decomposition-based Method for Automatic Ship Detection Crossing Different Satellite SAR Images

Siyuan Zhao, Zenghui Zhang, Weiwei Guo and Tao Zhang

Shanghai Jiao Tong University and Tongji University, Shanghai, China
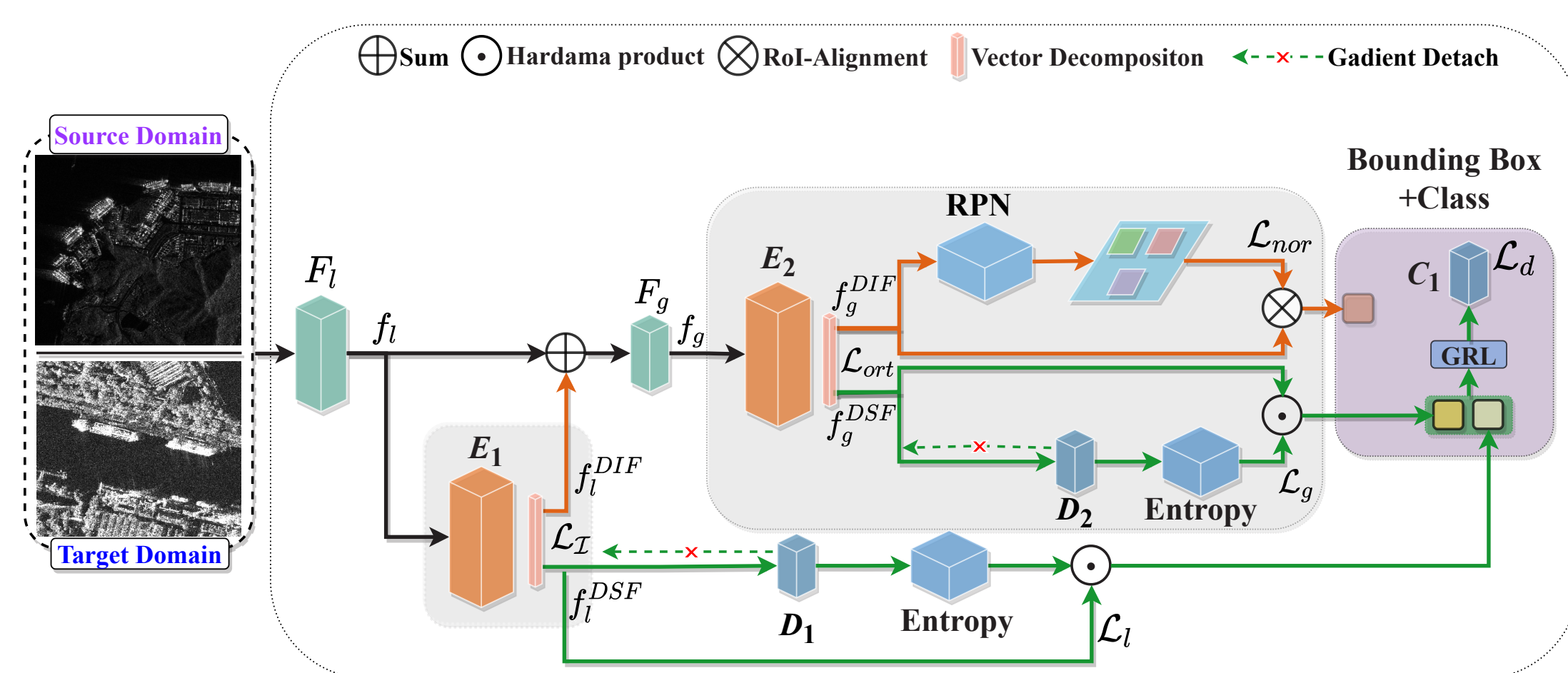
drickzsy@163.com

1. Abstract

In the face of Synthetic Aperture Radar (SAR) image object detection with different distributions of training and test data, traditional supervised learning methods cannot achieve good detection performance. Domain adaptation (DA) method has been shown to have the ability to solve this problem, but existing DA object detection algorithms all use adversarial DA theory for the detection task, which is ineffective in solving object regression localization in the detection task. In this article, to better solve the above problem, an automatic SAR image ship detection method based on feature decomposition crossing different satellites is proposed. The feature extraction layer of backbone network is divided into low level and high level, where domain-invariant feature extractors are designed for the local features extracted from the low level and the global features extracted from the high level, respectively. We argue that the local and global features extracted from source domain and target domain contain domain-specific features (DSF) for adversarial DA and domain-invariant features (DIF) that contribute to object regression localization. Then, we decompose the local features and global features into DSF and DIF via vector decomposition method. For DSF counterpart, we introduce adversarial DA attention for feature alignment. DIF from the local features are fused into the backbone network for high-level global feature extraction. Finally, by using region proposal network and adversarial domain classifier, we can get the accurate bounding box and object class of SAR image objects. Extensive experiments prove that the proposed method outperforms state-of-the-art methods in terms of detection performance.

2. Introduction

In general, with the increase of satellite-borne SAR imaging platforms, we can obtain more SAR image datasets, but it is very time-consuming and labor-intensive to re-label and train them. The DA method can be initially applied to the object detection problem of different satellite-borne SAR images, but because most of them copy the methods in natural images, they do not propose the DA network for the characteristics of the SAR image itself. Considering the SAR imaging system and SAR image features, it is necessary to propose a DA method that can meet the object class adversarial learning and object regression localization learning in large SAR image scenes.

3. Related Work

1) Traditional SAR Image Object Detection: Traditional SAR image object detection aims at obtaining the object position via utilizing detection systems combined independent functional components in a sequential cascade. A method for extracting ship features from SAR images is proposed, using a standard algorithm to generate a closed contour image of the ship, a two-step method to accurately estimate the orientation angle, and a modified centerline and minimum outer rectangle extraction criterion.
2) Deep Learning SAR Image Object Detection: SAR image object detection based on deep learning has become a popular topic in recent years, thanks to the potential to achieve better accuracy and manage bigger amounts of data through the extraction of object features.

6. Conclusions

We propose a feature decomposition-based method for automatic ship detection of SAR images crossing different satellites. We employ Faster RCNN as the baseline network, ResNet-101 as the backbone network, and divide the feature extraction layer of the backbone network into low level and high level. Experimental results on Gaofen-3 and SSDD show that the detection performance of the method is significantly better than that of the baseline network and other DA SOTA methods, which reflects the effectiveness and superiority of the method. More detailed technical details about this report can be found in our paper with the same titlle as this report in IEEE Transactions on Geoscience and Remote Sensing.

4. Method

Faster RCNN is utilized as the baseline network whose training input is a combination of labeled source domain SAR images and unlabeled target domain SAR images. The overall objective is to achieve good detection results on target domain SAR images. The feature extraction portion of the baseline network is divided into two levels: local feature extraction on the low level and global feature extraction on the high level. Specifically, the following figure presents the overall DA object detection framework.

5. Results

In this part, we select two SAR image object detection datasets with domain-shift, which are Gaofen-3 and SSDD. The former is imaged by Chinese C-band Gaofen-3, and the latter is the imaging result of German X-band TerraSAR-X, Canadian C-band Radarsat-2 and European Space Agency C-band Sentinel-1. Without loss of generality, we separately treat Gaofen-3 dataset and SSDD as the source domain and target domain in the course of subsequent experiments, and thereby define two DA SAR image object detection tasks: Gaofen-3 to SSDD (G $\rightarrow$ S) and SSDD to Gaofen-3 (S $\rightarrow$ G). The proposed DA SAR image object detection method is compared with conventional Faster RCNN and other seven SOTA DA object detection methods. For two SAR image object detection tasks, G $\rightarrow$ S and S $\rightarrow$ G, the experiments are conducted individually, and the quantitative results of the comparison experiments are shown in the following Table. As can be seen, the proposed method has the best detection accuracy for both detection tasks.

Method	G $\rightarrow$ S			S $\rightarrow$ G		
	PR	RE	mAP	PR	RE	mAP
Faster RCNN	62.5	77.8	67.0	57.7	71.0	57.9
FACL	74.6	82.9	78.1	69.8	79.9	68.4
UDA	73.5	81.9	76.2	66.9	78.0	67.3
SDA	68.6	77.2	70.1	63.7	75.3	61.9
SCL	69.4	78.6	71.4	65.9	76.8	63.2
MeGA	71.3	80.2	73.1	66.6	77.2	65.6
DTA	75.8	83.1	79.2	70.4	80.3	69.2
I³NET	73.8	82.2	77.6	67.2	78.4	67.6
Ours	**78.4**	**86.3**	**81.5**	**73.7**	**81.9**	**74.4**

Detection results and feature heat maps of a scene in G $\rightarrow$ S task are shown in the following figure. In (a)-(i), green indicates a correctly detected object, red indicates a wrong detected object, and yellow indicates a correctly detected object but the object's rectangular box is the wrong size.