

密级:_____



中国科学院大学
University of Chinese Academy of Sciences

硕士学位论文

基于光学遥感图像的鲁棒性舰船检测

作者姓名: 刘嫣然

指导教师: 叶齐祥 教授 中国科学院大学

学位类别: 工程硕士

学科专业: 计算机技术

研究所 : 中国科学院大学电子电气与通信工程学院

2017 年 05 月

**Robust Warship Detection Based on Optical Remote-
Sensing Images**

By

Liu Yan Ran

A Thesis Submitted to

University of Chinese Academy of Sciences

In partial fulfillment of the requirement

For the degree of

Master of Computer Technology

School of Electronics, Electrical and Communication Engineering

University of Chinese Academy of Sciences

May, 2017

中国科学院大学直属院系 研究生学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

中国科学院大学直属院系 学位论文授权使用声明

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密的学位论文在解密后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘 要

现代化战争中，舰船作为一项重要的海上作战武器，成为世界各国重点观测甚至打击的目标，其自动检测技术成为亟待解决的问题。目前的舰船检测研究大多是围绕合成孔径雷达图像开展的，基于光学遥感图像的舰船目标检测研究起步较晚，还有待进一步探索和研究。在光学遥感图像舰船检测的研究中，主要存在以下三个问题：其一，岸上区域对检测器形成干扰，造成大量的虚警；其二，光学图像中的舰船目标俯拍视角下呈现多角度多姿态分布，传统的检测器不能适应目标的大角度旋转；其三，基于直立候选框的目标检测方法对并排舰船提取的候选区域高度重合，从而出现大量漏检。本文针对上述问题，展开了以下三个方面的研究：

（1）针对岸上疑似目标的误检问题，提出了一种基于图割的海陆分割方法，对岸上区域进行遮挡。首先利用基于图的分割方法将图像划分为许多个子区域，提取每个子区域的多类混合特征，训练随机森林分类器进行区域判别。

（2）针对目标多角度变化带来的性能下降问题，主动旋转卷积滤波器得到带有方向信息的特征图，通过对各个方向的特征求最大值来实现角度鲁棒性检测，提升深度卷积网络处理多角度目标的能力。

（3）针对并排停放舰船的漏检问题，采用分类激活特征图的方法找到图像中的关键区域，得到热成像图。通过一些形态学处理找到连通域，结合方向估计可得到紧致的检测结果框，从而减少目标漏检。

利用本文提出的检测框架对我们构建的真实场景下的舰船数据集进行实验，验证了本文检测算法的准确性和高效性。

关键词：深度学习；舰船检测；海陆分割；多角度目标检测

Abstract

In modern warfare, as important maritime warfare weapons, warships are observed and even targeted all over the world, and its detection technology has become a serious problem. The current researches of warship detection mostly rely on synthetic aperture radar image, and optical remote sensing images based technique needs to be further explored for its late development. In the research of warship detection, there were three problems: First, the bank region interfere with the detector, and cause many false alarms . Second, in optical remote sensing images, the warships are multi-angle distributed. Third, since vertical warships detection based methods easily lead to a large number of missing when warships are parked side by side. In view of the above questions, this paper has done the following aspects of the work:

(1) False positive problems against suspected targets ashore, and proposed a land and sea segmentation method based on graph cut, to occluded the bank region. Using methods based on graph divides the image into many subareas, extract multi-class mixed characteristics of each child area, random forest classifier trained regional discrimination.

(2) Performance degradation of object angle changes, initiative to revolving convolution filter to get feature maps with the direction information, the final features is the maximum value for each direction features, to achieve robustness of detection, enhance depth Convolutional network's ability in handling multi-angle targets.

(3) For the problem of missing detection of parked warships, a method called class activation mapping is used to find the critical regions in the image, and the thermal image is obtained. Through Morphological processing to find connected components, combined with the orientation estimated results can be obtained compact box, thereby reducing target detection.

We evaluate the proposed framework on the warship datasets which built from the normal environment, and then demonstrate the competitive accuracy and efficiency of the proposed framework.

Key Words: deep learning; warship detection; land-ocean segmentation;

目录

摘 要.....	I
Abstract.....	III
图目录.....	VII
表目录.....	IX
第 1 章 绪论.....	1
1.1 课题背景和研究意义.....	1
1.2 研究现状及问题.....	3
1.2.1 光学遥感图像舰船目标检测.....	3
1.2.2 舰船检测中的难点和问题.....	4
1.3 本文研究内容.....	6
1.4 本文的组织结构.....	7
第 2 章 相关工作与技术.....	9
2.1 舰船目标检测技术.....	9
2.1.1 离港舰船和靠港舰船的处理策略.....	9
2.1.2 预处理.....	10
2.1.3 感兴趣区域提取.....	11
2.1.4 并排舰船的处理.....	12
2.1.5 特征提取.....	13
2.1.6 目标判别.....	15
2.2 海陆分割算法.....	17
2.2.1 基于港口先验信息的方法.....	17
2.2.2 基于图像分析的方法.....	17
2.3 多角度目标检测.....	19
2.3.1 提取带角度的候选框.....	19
2.3.2 设计旋转不变性特征.....	20
2.3.3 深度学习中的角度自适应.....	20
2.4 本章小结.....	22
第 3 章 基于图割的海陆分割算法研究.....	25
3.1 问题描述及难点分析.....	25
3.1.1 问题描述.....	25
3.1.2 难点分析.....	25
3.2 海陆分割算法整体流程.....	26
3.3 图割算法分割区域.....	27
3.4 区域混合特征提取.....	30
3.4.1 灰度特征.....	30

3.4.2	几何特征	31
3.4.3	颜色特征	31
3.4.4	纹理特征	32
3.5	随机森林判别水域	33
3.6	形态学水域扩展	37
3.7	实验结果与分析	38
3.8	本章小结	39
第 4 章	基于深度学习的鲁棒性舰船检测	41
4.1	问题描述及解决方案	41
4.2	舰船检测研究框架	41
4.2.1	整体研究框架	41
4.2.2	可变角度滤波器实现旋转不变性	43
4.2.3	全卷积网络提取特征	45
4.2.4	CAM 定位重要区域	47
4.2.5	求连通域并估计参数得到检测框	49
4.2.6	难反例挖掘及正例扩充	50
4.3	实验结果与分析	51
4.3.1	数据集介绍	51
4.3.2	检测结果与分析	52
4.4	本章小结	54
第 5 章	结论与展望	55
	参 考 文 献	57
	致 谢	61
	个人简历、在学期间发表的论文与研究成果	63

图目录

图 1-1 舰船目标检测的应用	2
图 1-2 舰船检测的主要步骤	3
图 1-3 舰船角度多变	4
图 1-4 角度变化对检测结果的影响	5
图 1-5 舰船并排漏检	5
图 1-6 相似物体干扰	6
图 1-7 本文整体检测框架	6
图 2-1 舰船检测流程	9
图 2-2 图像去噪	10
图 2-3 云雾消除	10
图 2-5 RPN 提取候选区域	12
图 2-6 卷积神经网络结构	14
图 2-7 深度特征示例	15
图 2-8 支持向量机原理图	15
图 2-9 传统方法海陆分割示例	19
图 2-10 空间变换网络结构	21
图 2-11 空间变换网络方向矫正示例	22
图 3-1 区域内强度差异大于区域间强度差异	26
图 3-2 海洋与浅滩、草地等的相似性	26
图 3-3 海陆分割算法流程	27
图 3-4 图割算法基本原理	28
图 3-5 区域分割结果示例	30
图 3-6 区域标注工具	34
图 3-7 水域判别结果示例	36
图 3-8 水域扩展结果示例	38
图 3-9 本文海陆分割算法结果示例	38
图 4-1 整体研究框架	42
图 4-2 原始单方向滤波器角度冗余	43

图 4-3 滤波器旋转	43
图 4-4 ARF 提取特征图示例	45
图 4-5 残差网络结构	46
图 4-6 特征提取网络结构示意图	47
图 4-7 特征图加权求和得到激活响应图	48
图 4-8 CAM 定位重要区域	48
图 4-9 等高线法求连通域	49
图 4-10 估计参数画出检测框	50
图 4-11 难反例示例	50
图 4-12 初始模型未检测到的正例	50
图 4-13 数据标注工具	51
图 4-14 检测结果对比	53
图 4-15 本文算法检测结果示例	54

表目录

表 2-1 带方向的候选框提取算法	19
表 3-1 基于图的分割算法	29
表 3-2 区域混合特征	33
表 3-3 决策树算法	35
表 3-4 基学习器数量与测试准确率	36
表 3-5 分类算法准确率比较	36
表 3-6 海陆分割带来的检测性能变化	39
表 4-1 检测框计算流程	49
表 4-2 数据集信息统计	51
表 4-3 检测指标对比	52

第1章 绪论

1.1 课题背景和研究意义

近年来,随着对地观测技术的迅速发展,光学遥感成为获取高分辨率影像数据的一种重要手段。光学遥感图像符合人眼对物体的观察习惯,具有直观易理解、分辨率高等优点。在天气晴朗,光照充足的情况下,得到的图像内容较为丰富,地物结构特征明显,便于目标的检测识别。

为了得到画面清晰、细节信息丰富、更新时长更短的遥感图像,各国投入使用的遥感卫星不断向空间分辨率更高、重访周期更短的趋势发展。目前世界上分辨率最高的光学卫星 WorldView-3 分辨率已达 0.3 米,平均回访时间不到 1 天。我国的光学遥感卫星也已进入亚米级时代,2014 年发射的“高分二号”卫星分辨率达到 0.8 米,重访周期 5 天。此外,以飞机为载体得到的飞机航拍图像,随着航拍传感器、GPS 导航航拍等技术的发展,其多姿态、多视角、高分辨率的优势逐渐使其逐渐成为重要的数据来源。综上所述,通过遥感手段得到的数据细节越来越清晰,数据量也呈现指数爆炸式增长。如何快速有效地从海量的遥感图像中提取和检测出感兴趣目标,成为当前亟待解决的问题。

在此背景下,高效率、自动化的遥感图像目标检测成为遥感应用中重要的研究课题之一。遥感图像目标检测有效解决了人工判读效率低、周期长、主观性强的问题,可以在无需人工干预的情况下准确地定性定位目标,极大地满足了现代社会对高效信息的需求。遥感图像目标检测不仅可以在高科技军事对抗中实时准确地为作战部队提供敌方战场信息,捕获军舰、飞机等打击目标,还可以应用于城市规划、资源勘探、环境监测等民用领域。因此,遥感图像目标检测具有广泛的应用前景和研究价值。

舰船作为海上重要的军事作战武器,其自动检测与识别具有非常重要的现实意义。在军事领域,国际局势瞬息万变,不断升级的高科技武器对国家安全造成一定的威胁。舰船目标检测系统可以实时监视敌国舰船部署和动向,当别国舰

基于光学遥感图像的鲁棒性舰船检测

船进入我国领海时及时进行反应，分析战争形势，帮助军事专家和指挥官进行战略部署，维护我国领海安全。另外，舰船检测还可以迁移到民用船只的检测，帮助相关部门监控和疏导海上交通，对走私、非法倾倒油污等违法行为进行监视，突发自然灾害时对遇难船只进行定位，方便救援。

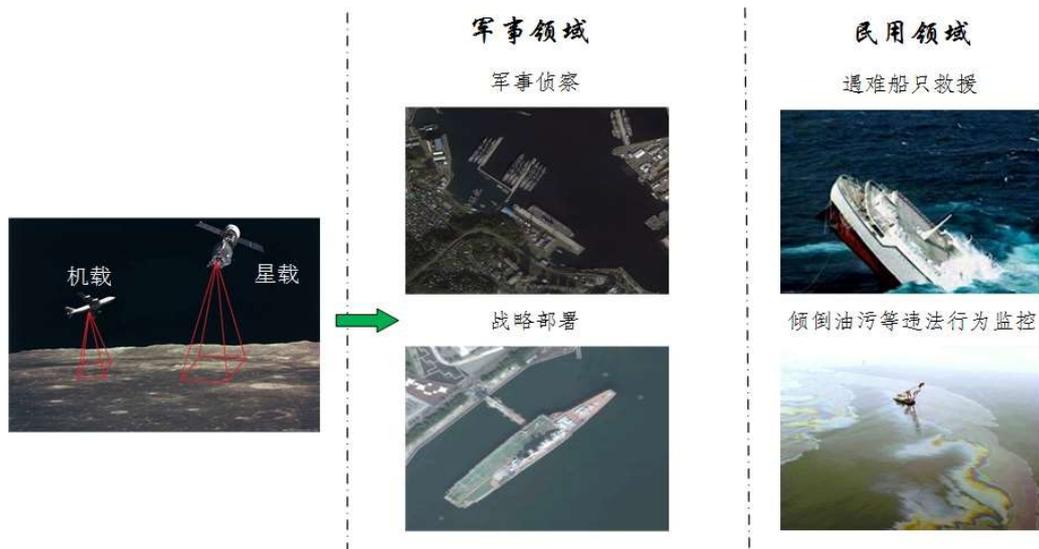


图 1-1 舰船目标检测的应用

目前，自动舰船目标检测的研究主要集中在合成孔径雷达图像，基于可见光舰船目标检测的研究起步较晚、相对滞后。随着光学遥感技术的发展，光学图像在舰船检测尤其是舰船目标分类与识别方面的突出优势引起了国内外学者及军事国防部门的广泛关注。近些年相关科研人员对舰船目标自动检测进行了一些探索，并取得了一定成果，多种基于光学遥感图像的舰船目标检测方法被提出，然而由于舰船这一目标具有类型多样、结构各异、大长宽比等特点及码头毗连、舰船并排停靠、背景复杂、拍照自然环境多样化带来的干扰，给舰船的检测与分类带来恶劣的影响，因此可见光舰船目标检测的研究仍然远不成熟，距离实际应用还有很大的差距。

本课题针对舰船目标检测的岸上虚警问题、多角度问题、并排停放问题分别给出了解决方案。致力于实际应用，提升光学遥感图像舰船目标检测的精度，探索适用于舰船的目标检测策略，搭建较为完善、性能较好的舰船目标检测框架。

本文所述框架对多角度、并排目标的检测具有一定的普适性，可推广到其他目标检测问题中，是一个颇具吸引力和挑战性的研究课题。

1.2 研究现状及问题

现有的舰船目标检测工作中，合成孔径雷达(Synthetic Aperture Radar, SAR)图像在探测和跟踪目标方面起着至关重要的作用^[1]，目前比较成熟的合成孔径雷达舰船目标检测系统有：欧盟的 JRC 系统和 VDS 系统、加拿大的 OMW 舰船探测系统、英国的 DERA 系统、中科院微波成像实验室的舰船目标监视系统等。合成孔径雷达成像受天气、光照等环境因素影响较小^[2]，然而合成孔径雷达遥感图像数据量少、成像代价大、图像细节不清晰。与之相比，光学卫星图像具有更高的分辨率，包含更详细的细节信息，因此更适用于舰船目标的检测或识别^[3]。近年来光学遥感场景下舰船检测的相关研究不断增多。本节将针对光学遥感图像舰船检测的相关技术及存在的难点和问题进行讨论。

1.2.1 光学遥感图像舰船目标检测

与其他目标检测问题类似，光学遥感图像舰船检测的相关工作大多采用“从粗到精”的策略，研究侧重于背景与目标的分离以及目标特征的提取与选择上。主要包括三个阶段：候选区域提取、特征提取、目标鉴别。

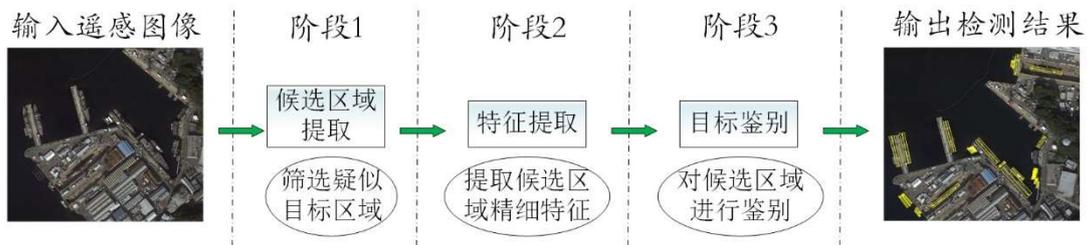


图 1-2 舰船检测的主要步骤

候选区域提取是检测过程中关键的一步，利用舰船与背景之间潜在的差异性从整个图像中得到疑似目标区域。最常用的算法是根据区域的强度差异^[4-5]找到候选区域；Guo^[6]提出采用自底向上的视觉注意机制对整个场景中具有显著性的区域进行选择；Corbane^[7]提出利用自适应阈值分割与灰度对比度增强的方法获得疑似目标区域。上述方法在纯水域场景下能够取得较好的性能，对于港口场景，会得到大量无关区域，极大地增加了后续模块的计算量。另外，实际场景下多数舰船通过缆绳与码头相连，轮廓粘连难以划分，导致得到的候选框极不精确。

得到待选区域之后，还需要对区域提取较为精细的特征，训练分类器来去除虚假感兴趣区域。针对舰船的“米粒”状长条结构，最直观的方法是提取区域的几何特征。**Buck**^[8]提出利用舰船长宽比以及舰船上层结构与整体舰船长度的比例来排除不满足条件的区域；**胡俊华**^[9]等提取区域的尺寸、面积、质心在区域中的相对位置及占空比等特征；**Zhu**^[4]等人对候选区域提取了形状特征、局部多模式特征、尺度高斯差分特征等多种鉴别特征。由于舰船与码头、岛礁等的相似性，利用这些“简单”特征可能会导致分类器不足以将舰船与其他疑似区域分离开。

在目标鉴别方面，当特征维数较少时通常采用距离度量或者多门限判别的方式进行分类鉴别，对于高维特征，林肯实验室提出了二项式鉴别算法，法国的IRD检测小组^[10]采用动态贝叶斯网络实现舰船的分类，**Zhang**^[11]等人提出使用误差后向传递多层感知器神经网络方法来区分舰船目标。针对过拟合问题，科研人员将支持向量机^[12](Support Vector Machine, SVM)引入舰船目标检测中，SVM迅速成为舰船检测中最常用的鉴别分类方法。

1.2.2 舰船检测中的难点和问题

由于舰船目标具有类型多样、结构各异、大长宽比等特点及码头毗连、舰船并排停靠、背景复杂、拍照自然环境多样化带来的干扰，可见光舰船目标检测的研究仍然远不成熟，距离实际应用还有很大的差距，在检测过程中遇到了一些难题，主要包括以下几点：

(1) 舰船角度多变

舰船在遥感俯拍视角下方向多变，给舰船检测造成了不利影响。我们对舰船数据集中的样本进行角度统计，如下图所示。

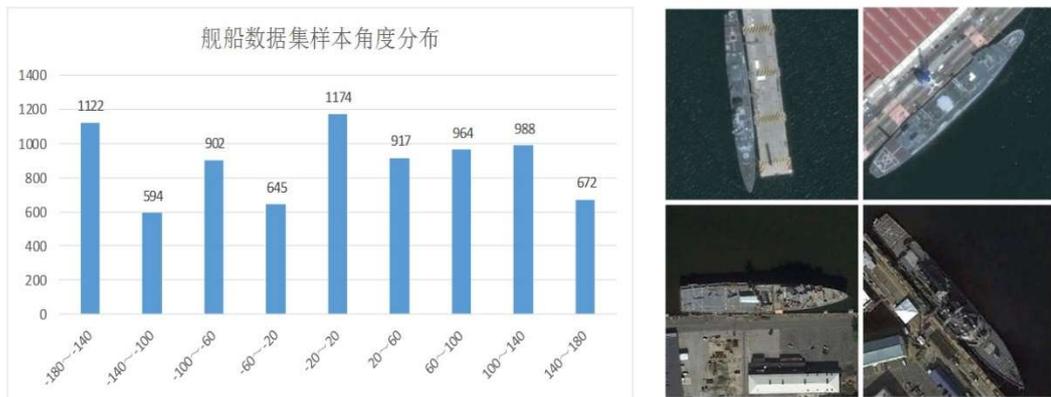


图 1-3 舰船角度多变

现有的目标检测算法对角度变化敏感，我们利用 DPM^[13]检测框架进行初步实验来说明舰船角度变化对检测结果的影响。



图 1-4 角度变化对检测结果的影响

图 1-4 中左图为直接利用 DPM 进行检测的实验结果，在阈值很低（出现大量虚警）的情况下仍有两艘舰船没有被检测到。为了验证这种漏检是由目标角度问题引起的，我们在单一角度下训练模型，测试时采用 18 个角度旋转图像，将各个角度的检测结果叠加，并进行多角度下非极大值抑制，得到图 1-4 中右图结果。检测网络对目标角度变化适应能力较差，目标方向变化引发检测性能退化。

(2) 舰船并排停放

港口中舰船会出现并排停放的情况，针对不同目标提取的候选框高度重叠，而非极大值抑制操作会删除重合度较高的候选框，最终多个并排目标只得到一个检测框，导致大量漏检。

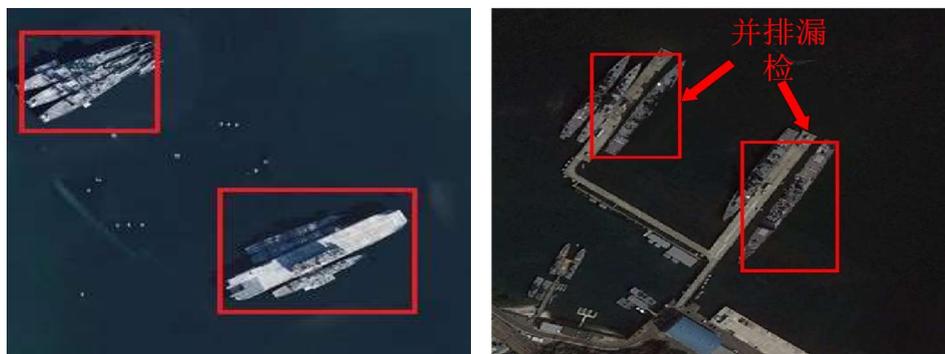


图 1-5 舰船并排漏检

(3) 相似物体干扰

舰船与码头等陆上目标非常相似，容易产生虚警。对于表面较平滑、不具有尖头形状的航母，更易与码头混淆。



图 1-6 相似物体干扰

1.3 本文研究内容

本文致力于在实际应用中能快速有效地在海量遥感数据中找到感兴趣的舰船目标。本文提出的目标检测方法整体框架如图 1-7 所示。

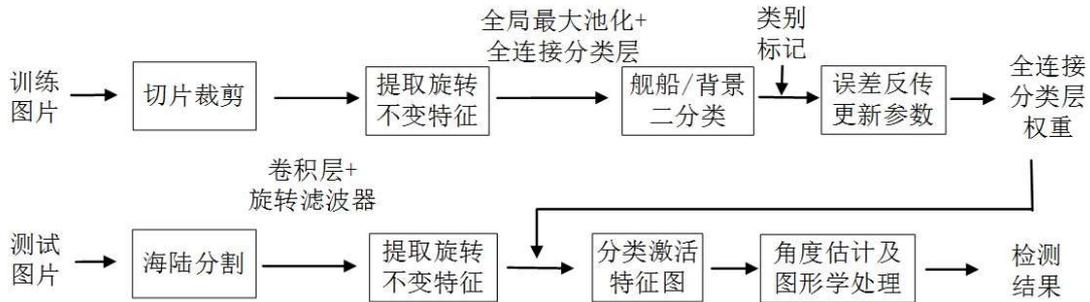


图 1-7 本文整体检测框架

检测过程分为训练阶段和测试阶段。训练阶段的处理流程如下：首先进行目标切片和反例切片裁剪；对深度神经网络中卷积层滤波器进行改进获得旋转不变性特征；在卷积层后加入全局最大池化层和全连接分类层，利用分类标记训练网络。测试阶段的处理流程如下：利用海陆分割算法进行水域划分，接着利用训练好的卷积网提取旋转不变深度特征；将得到的特征图与训练网络中分类层对应权重（权重表示该特征图对于特定类别的重要性）相乘求和得到响应图，通过角度估计及图形学后处理得到最终检测结果。

本文的主要研究内容总结如下：

- (1) 在舰船检测过程中，由于港口场景下的遥感图像背景较为复杂，容易

产生虚警，本文提出一种基于区域分割及多类区域特征提取的海陆分割方法，找到图像中的陆上区域，得到水陆划分的二值化掩膜图像，排除岸上虚警。

(2) 针对多角度舰船目标的检测问题，为了抵抗角度变化带来的性能下降，本文采用主动旋转滤波器的方式获得多方向特征，通过对各方向通道特征取最大值得到旋转不变深度特征，提升深度卷积网对多角度目标的适应能力，提高多角度目标的检测精度。

(3) 对于传统检测算法的并排舰船漏检问题，本文提出一种不依赖于候选框提取的目标检测框架，利用卷积层的定位能力找到有利于舰船判别的关键区域。结合角度估计及图形学处理得到紧致包络框，大量减少由并排摆放引发的舰船漏检。

1.4 本文的组织结构

本文共分五章，具体安排如下：

第一章，绪论。总结分析光学遥感图像舰船目标检测技术的课题背景、研究意义及研究现状。对现有可见光舰船目标检测算法存在的难点和问题进行总结，对本文针对舰船目标设计的目标检测框架进行了简单的阐述和说明。

第二章，相关技术综述。详细介绍传统舰船检测中各个环节，包括离港和靠港舰船的处理策略、云雾消除预处理、感兴趣区域提取、特征提取、目标鉴别，对每个处理过程中的已有方法进行综述。同时对现有的海陆分割方法、多角度目标检测方法进行列举总结。

第三章，基于图的海陆分割算法研究。本章首先对分割问题中的难点进行分析和说明，论述本课题所使用的海陆分割算法的具体实现步骤：基于图的区域划分、区域多类特征提取、随机森林分类器判别水域、基于图形学的水域扩展。最后对海陆分割结果进行评测和分析。

第四章，基于深度学习的鲁棒舰船检测。对本文提出的舰船目标检测框架进行详细阐述，具体介绍旋转不变性深度特征的提取过程，详细说明本课题提出的检测框架在处理多角度、并排目标方面的优势性。通过对比实验和结果分析说明本文算法的高效性和准确性。

第五章，结论和展望。对本文的研究成果进行总结，并针对研究过程中存在的问题及未来研究做分析和展望。

第2章 相关工作与技术

本章将系统介绍与本课题密切相关的前沿算法和研究工作。首先是舰船目标检测技术的论述，接着是对舰船检测衍生出的海陆分割问题和多角度目标检测问题的相关工作的综述。

2.1 舰船目标检测技术

本节将按照舰船检测流程中各个模块的处理顺序对每个步骤中用到的相关算法进行介绍。根据图像背景的差异将检测问题划分为离港舰船的目标检测和靠港舰船的目标检测，处理流程如下：

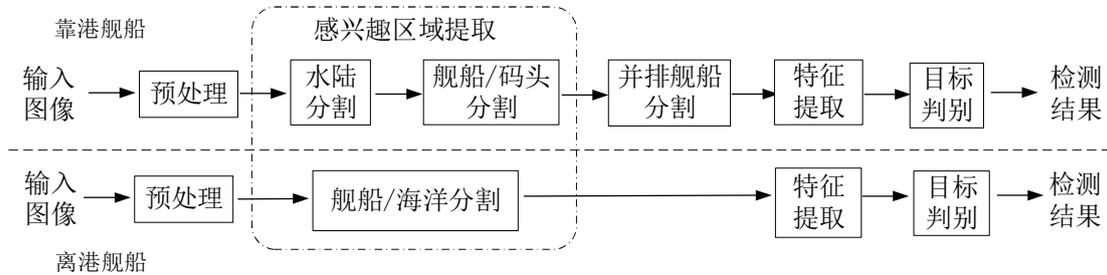


图 2-1 舰船检测流程

2.1.1 离港舰船和靠港舰船的处理策略

按照舰船所处位置可将包含舰船的遥感图像分为纯海洋场景及港口场景。对于出海作业状态下的舰船目标，图像多为纯海洋场景，背景较为简单，检测问题的关键是利用舰船和海洋的差异性找到舰船。而在实际情况下，舰船检测的目标多处于靠港状态，图像背景极为复杂。科研人员针对离港舰船及靠港舰船背景的差异，对两种场景下的舰船目标检测问题给出不同的处理策略。

由图 2-1 可以看出靠港舰船与离港舰船的处理都包括预处理、感兴趣区域提取、特征提取、目标判别四个步骤。靠港和离港舰船目标检测算法之间的不同主要体现在感兴趣区域提取阶段。对于海域中的船只，一般直接进行区域提取即可。而对于停靠在港口中的船只，一般采用二次分割的方式得到候选区域：。第一次分割划分海洋和陆地，此时舰船目标被归为陆地区域；。第二次分割划分舰船和码头，得到疑似目标区域。另外针对港口中舰船的并排码放，靠港船只的检测还需要对并排舰船进行分割操作。

2.1.2 预处理

遥感图像的视觉质量会受到噪声、环境等因素影响，因此在进行图像解译之前需要对遥感图像进行适当的预处理操作，一般包括图像去噪以及云雾消除。

图像去噪：在遥感图像中椒盐噪声和斑点噪声较为常见且干扰最大。文献^[14]利用中值滤波去除图像中的椒盐和斑点噪声，再通过高斯滤波滤除整个画面中服从正态分布的噪声。本论文试用了这种方法，得到图 2-2 所示结果。

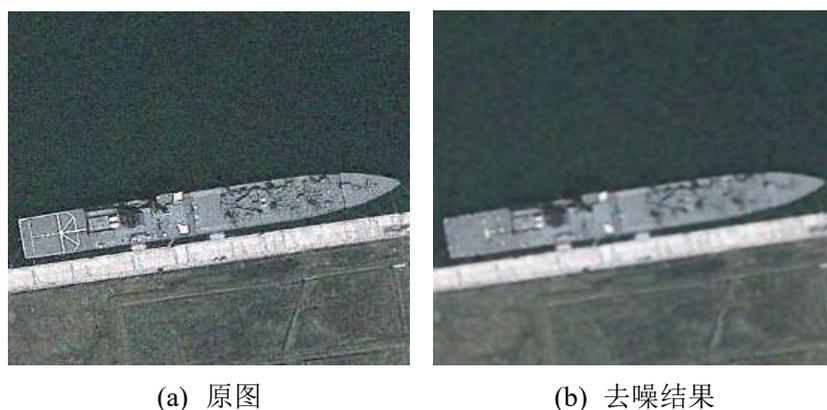


图 2-2 图像去噪

云雾消除：云雾去除算法主要分为替换法、滤波法以及缨帽变换法^[15]。本论文对缨帽变换法进行了初步尝试，利用顶帽开运算得到图 2-3 所示结果。由于本文所构建的数据集中只有极个别图像存在云雾遮挡问题，为避免信息损失，最终的实验框架中未加入去除云雾干扰这一环节。

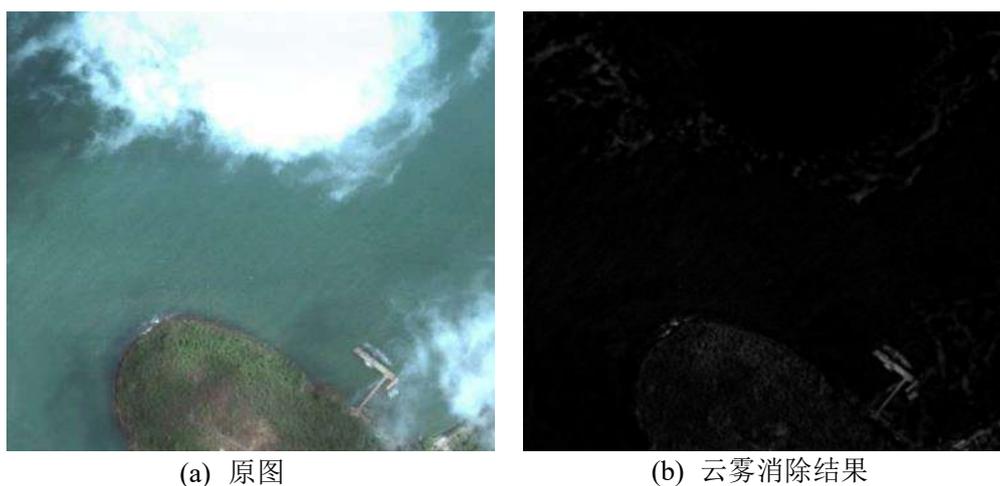


图 2-3 云雾消除

2.1.3 感兴趣区域提取

现有的检测方法一般采用由粗到精的策略得到检测结果，首先提取目标可能存在的区域再进行精细划分。如 2.1.1 中所述，舰船检测中的感兴趣区域提取一般分为离港舰船的感兴趣区域提取和靠港舰船的感兴趣区域提取。

2.1.3.1 离港舰船的感兴趣区域提取

针对离港舰船的目标与背景分离方法有四类：1) 基于灰度统计特征的方法^[6]，依据水体区域与目标区域的灰度统计特性差异进行分割，可应用的特征有：灰度、形态学对比度、类高阶梯度、局部统计方差、图像信息熵等；2) 基于边缘信息的方法^[4]，通过提取边缘结构信息，获得疑似目标区域；3) 基于分形特征的方法，相关研究表明海域等自然场景具有分形特征，而舰船等人工目标并不具备分形特征，文献^[17]利用分形方法和模糊理论进行离港舰船的区域提取；4) 基于视觉感知的方法，人的视觉系统可以快速聚焦到感兴趣区域，将此机理迁移到舰船检测领域，文献^[18]提出基于人眼多级视觉感知和视觉非对称机制的方法得到目标区域。

2.1.3.2 靠港舰船的感兴趣区域提取

大部分的靠港舰船候选区域提取都是在海陆分割的基础上进行的。海陆分割一般会将靠港舰船划分为陆地区域，因此需要对海陆分割得到的陆地区域进行二次分割，将码头等其他人工目标与舰船分离开来。由于码头和舰船在纹理和灰度特征上不存在明显的差异性，如何选择特征准确的分离两者是图像分割最大的难点。现有的文献中，一般采用基于全局几何形状的特征和基于局部结构显著性的特征来解决这一问题。文献^[19]通过陆地区域向水体区域凸出的部位大致圈定疑似靠港舰船区域，再根据该区域面积、长宽比等几何特征与各型号舰船比较，进一步确定该区域是否含有舰船。文献^[9]提出利用局部自相似来提取舰船显著性特征点，得到显著性区域，分析该区域显著性特征，确定此区域是否存在舰船。

2.1.3.3 普遍适用的候选框提取方法

除了上述针对舰船提出的候选区域提取方法外，还可利用一些普遍适用的区域提取算法来获取目标区域，如滑窗法、Selective Search^[20]、edgeboxes^[21]等。

随着深度卷积神经网络的兴起，区域提取网络（Region Proposal Networks,

RPN) 被提出。它是一种利用深度网络来提取候选框的方法, 利用卷积网络提取图像各个级别的特征, 得到特征图。在原有卷积层后再加入两个卷积层, 第一个卷积层对每个滑窗位置上相应的特征编码成一个低维向量, 第二个卷积层得到区域的分类得分和区域位置。训练过程中利用目标框真值分别求得分类和回归误差, 采用随机梯度下降法更新网络各层权重。测试过程将遥感图像输入训的网络即可得到感兴趣区域。RPN 网络提取目标区域的效果如下:



图 2-4 RPN 提取候选区域

无论是穷举式的扫窗方法还是先验式的 **Selective Search** 算法, 或者深度学习中的 RPN 方法, 它们都有一个共同的缺点: 只能提取直立的候选框。目标检测领域中大多数待检测目标都是接近直立的, 可应用上述方法获取候选区域。而对于舰船这种多角度、大长宽比目标, 直立的候选框并不适用。

2.1.4 并排舰船的处理

在实际拍摄的遥感图像中, 舰船常常会出现并排停靠、太阳斜射或天气原因导致舰船出现阴影, 遮挡其他舰船导致轮廓不全等问题, 给目标区域切分造成了极大的困难。为了实现粘连目标的划分, 文献^[22]采用基于轮廓区域统计的方法在并排舰船中提取单个目标。文献^[23]针对舰船并排和阴影遮挡问题, 利用舰船形状相似且有明显轴向、主轴与码头方向基本一致的特点, 通过投影找到相邻目标区域间的连通点, 实现并联舰船的轮廓切分。目前针对并排舰船划分的研究还比较少, 现有算法比较粗糙, 多利用舰船现有知识 (如同种类型舰船具有相近的长宽比等) 处理并排摆放问题, 对于舰船粘连很难取得理想的分离效果。

2.1.5 特征提取

2.1.5.1 手工设计特征

可用于表达和描述舰船目标特性的特征大致分为四类：

(1) 统计特征：舰船与其他目标相比在灰度、颜色、梯度等方面具有一定的差异性，可利用区域像素的均值、标准差、偏离度等统计特征进行目标鉴别。

(2) 纹理特征：通常情况下，舰船由于其特殊的构造呈现与其他区域不同的纹理结构，因此纹理特征是一种常用的鉴别舰船所使用的特征，包括通过灰度共生矩阵变换提取的纹理熵、纹理对比度等。

(3) 形状特征：舰船俯视图呈现轴对称、舰首较尖的狭长“米粒”状，是相对于虚警最为突出的差异特征，也是舰船检测与识别中最常用的一种特征。形状特征一般包括长宽比、周长、面积、相关矩、转动惯量、密实度、对称度等。

(4) 其他特征：主要包括尾迹特征、位置相关性。对于航行中的舰船，船身与水面相互作用产生尾迹，现有的研究表明尾迹特征对于行驶中的舰船具有重要的区分能力。

上述特征比较直观，易于理解。但手工提取特征十分费力，高度依赖先验知识。除了这些针对舰船特点设计的特征，还可采用一些普适性的特征提取算法获取特征。近年来 HOG、SIFT 这两种特征及其变形形式在计算机视觉领域得到了广泛使用，是公认的效果较好的两种手工设计特征。

尺度不变特征变换^[24] (Scale-invariant feature transform, SIFT) 算法通过高斯卷积构建高斯金字塔，对相邻层图像求差得到差分高斯金字塔，对每一个采样点和它在图像域和尺度域相邻点进行比较得到极值点作为特征点。对特征点邻域像素的梯度分布特性进行统计得到特征。SIFT 特征具有一定的旋转、尺度不变性，但 SIFT 算子只关注梯度信息而舍弃了其他信息，无法准确捕捉边缘光滑的特征点。

方向梯度直方图^[25] (Histogram of Oriented Gradient, HOG) 特征提取方法是对每个小区域内的像素梯度进行直方图统计，得到每个区域的特征描述子，将这些特征描述子串联得到图像的特征描述子。HOG 特征受噪声、光照变化影响较小，但 HOG 算子和 SIFT 算子一样只关注梯度，特征表示能力有一定的局限性。

2.1.5.2 深度特征

选择更优质更利于分类的特征对目标检测性能的提升至关重要。近年来通过卷积神经网络^[26]来提取图像特征逐渐成为计算机视觉领域中主流的提取特征的方法。卷积神经网络模拟生物视觉神经对输入图像进行分析、抽象、判定，在输入和输出层之间加入一些隐藏层，将特征提取和分类过程融合在一起，在有监督的情况下，根据标签分类和预测值之间的误差自动训练网络参数，其结构如下：

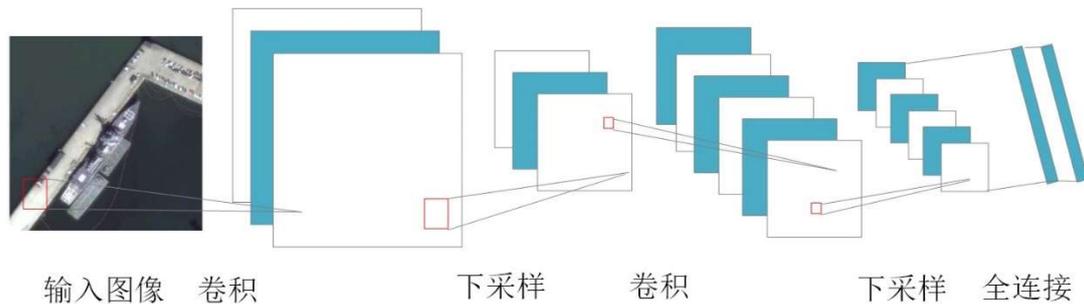


图 2-5 卷积神经网络结构

卷积层通过有偏的卷积运算得到图像的特征表示。下采样层通过求局部最大值或平均值操作对特征进行聚合统计，降低特征维度。全连接层将前边提取到的分布式特征投影到样本标记空间，得到便于分类的特征。卷积神经网络具有局部连接、权值共享、多层卷积等特点，能以较小的代价提取较为优质的特征。

(1) 局部连接：对一幅图像的某个像素来说，一般距离越近的像素与该像素的影响越大，距离越远的像素相关性越弱，因此只需要对局部区域进行感知，在更高层将局部的信息进行综合即可。

(2) 权值共享：权值共享即卷积核共享。图像局部统计特性与其他部分相同，对于整个可视域，都能使用相同卷积核抽取特征。每个核表示一种特征，将图像中符合条件的区域筛选出来，激活值越大表明对应局部感受野越符合条件。

(3) 多层卷积：特征提取过程中往往需要多个卷积层，一层卷积学习到的特征是局部的，较低级的，如第一层是颜色和边缘特征。通过多层卷积可以组合底层特征形成更加抽象的高层语义结合特征，形成对物体特性的认知。

我们利用深度卷积网提取特征，得到图 2-6 所示结果。图中第一列为原图，后三列为对应的浅层特征（第一个卷积层输出）、中层特征（第七个卷积层输出）、深层特征（第十三个卷积层输出）。可以看出，比较靠前的卷积层特征图涵盖的

细节信息较多，层数越深具有类别信息的区域会越来越突出。

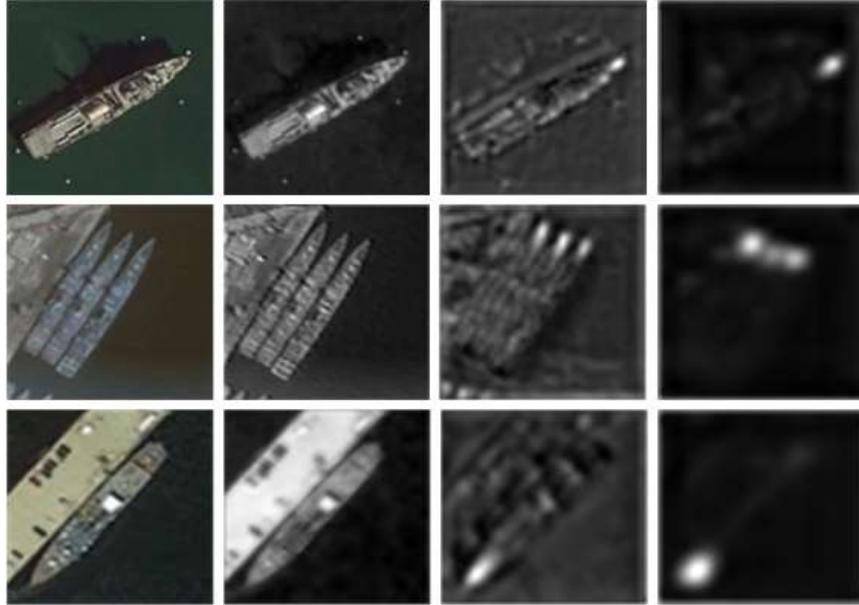


图 2-6 深度特征示例

2.1.6 目标判别

得到感兴趣区域之后，需要对每个候选区域进行目标判别，去除虚假目标。目前比较流行的判别方法有支持向量机和神经网络中常用的 softmax 分类。

2.1.6.1 支持向量机

支持向量机^[12]是一种建立在结构风险最小化基础上的分类器，在高维特征空间中找到一个分类超平面，将空间上的点分为两类。

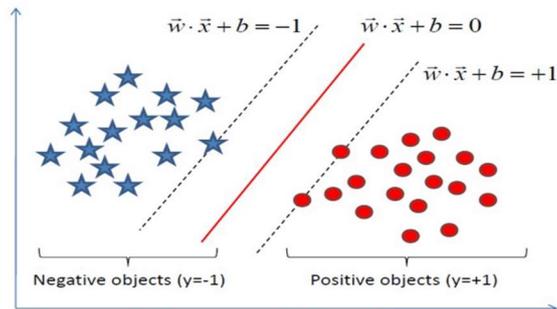


图 2-7 支持向量机原理图

支持向量机分类超平面为 $y = wx + b$ 时，正反例之间的距离达到最大。样本

空间中任意一点 x 到超平面的距离可写为: $r = \frac{|w^T x + b|}{\|w\|}$ 。假设超平面可以正确分类所有训练样本, 则对于样本空间中的点 (x_i, y_i) , 如果 $y_i = +1$, 则 $w^T x_i + b \geq +1$; 如果 $y_i = -1$, 则 $w^T x_i + b \leq -1$ 。离超平面最近的几个训练样本使等号成立, 被称为支持向量。两个异类支持向量到超平面的距离之和可表示为 $\gamma = \frac{2}{\|w\|}$ 。

通过最大化间隔 γ , 即最小化 $\|w\|^2$, 可得到最优超平面。其数学表达式为:

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} \|w\|^2 \\ \text{s.t.} \quad & y_i(w^T x_i + b) \geq 1, i = 1, 2, \dots, m \end{aligned} \quad (2-1)$$

上式是一个凸二次规划问题, 为了更高效地求解, 采用拉格朗日求解法得到其对偶问题, 再利用对偶原理进行参数求解。式 2-1 的对偶问题为:

$$\max_{\alpha} \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j) \quad (2-2)$$

如果原始特征空间是有限维的, 那么一定存在一个高维特征空间使样本可分, 但特征空间维数可能很高, 甚至是无穷维。这 SVM 使用核技巧来避免无穷维特征空间下 $\phi(x_i)^T \phi(x_j)$ 的计算问题。定义如下形式的核函数:

$$k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle = \phi(x_i)^T \phi(x_j) \quad (2-3)$$

即 x_i 和 x_j 在特征空间中的内积等于它们在原始样本空间中通过函数 $k(x_i, x_j)$ 计算的结果。常用的核函数有线性核、多项式核、高斯核、拉普拉斯核等。

2.1.6.2 softmax 分类

softmax 分类^[27]是 logistic 回归模型在多分类问题上的推广。对于给定的测试输入, 针对每一类别估算出概率值, 得到每一种分类结果出现的概率。设共有 k 个类别, 通过假设函数输出一个 k 维向量, 向量的元素之和为一。这个假设函数的定义如下:

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \dots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \dots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \quad (2-4)$$

其中 $\theta_1, \theta_2, \dots, \theta_k \in R^{n+1}$ 是模型参数, $1/\sum_{j=1}^k e^{\theta_j^T x^{(i)}}$ 这一项对概率分布进行归一化, 使所有概率之和为 1。为了求解参数 $\theta_1, \theta_2, \dots, \theta_k$, 定义如下交叉熵代价函数, 通过迭代的优化算法最小化代价函数。

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right] \quad (2-5)$$

上式中的 $1\{y^{(i)} = j\}$ 是示性函数，当 $y^{(i)} = j$ 成立时值为零，否则值为 1。通常采用梯度下降法使损失函数最小化，对式 2-6 求导得到如下结果：

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m [x^{(i)} (1\{y^{(i)} = j\} - p(y^{(i)} = j | x^{(i)}; \theta))] \quad (2-6)$$

求得偏导数后，每次迭代对参数 θ 进行更新。

$$\theta_j = \theta_j - \alpha \nabla_{\theta_j} J(\theta) (j = 1, \dots, k) \quad (2-7)$$

训练过程更新模型参数，测试时即可得到分类结果。

神经网络中多用 softmax 分类层实现目标鉴别。文献^[28]中对 SVM 和 softmax 分类器进行了比较，对于多分类问题，softmax 性能略优于 SVM，且在神经网络中 softmax 更有利于网络端到端的训练。

2.2 海陆分割算法

现有的舰船目标检测算法一般以海陆分割为基础，一方面借助海陆分割结果排除岸上相似目标，降低虚警率。另一方面提高检测效率，避免全图盲检带来的时间损耗和计算量的增加。

2.2.1 基于港口先验信息的方法

基于港口先验信息的方法是指通过港口及码头的先验地理位置信息来分离陆地（内陆和码头区域）和舰船（海洋和舰船）。由于港口建成后位置不变，利用港口先验信息可以比较准确地得到陆地区域。经典的方法有利用 eCognition 软件和 GIS 数据进行水陆分离^[29]，利用多光谱图像结合先验地理位置信息实现水域和港口分离的方法^[30]，基于地理信息的模板技术，将港口轮廓以模板的形式预先存储起来，通过匹配实现分离的方法^[31]。

然而这类方法有明显的弊端。基于港口先验信息的方法要求给出待检测图像对应的地理位置，在图像的经纬度信息缺失的情况下，此类方法完全失效。而实际应用过程中，并不能保证数据来源包含地理信息，另外即使是配准后的遥感图像，其位置信息仍然存在偏差，可能会导致分割结果不准确。

2.2.2 基于图像分析的方法

由于难以准确地获取遥感图像对应的地理位置信息，水陆分割多采用基于图像分析的方法。现有的海陆分割方法主要包括灰度阈值法和活动轮廓分割法。

2.2.2.1 灰度阈值法

目前海陆分割方法大多是基于灰度阈值法实现的。阈值的选取方法有很多，如最大熵法、直方图谷底法、自适应阈值法、最大类间方差法等。其中最大类间方差法^[31]（OTSU）由于具有较好的自适应性而被广泛应用于海陆分割中。

OTSU 算法寻找最佳门限将像素按灰度级分为背景和目标两类，使类间方差最大。设像素灰度级为 $1 \sim M$ ，总像素个数 N ，第 i 个灰度级像素个数 n_i 。则第 i 级灰度出现概率 $P_i = n_i/N$ ，图像平均灰度 $u = \sum_{i=1}^M i \cdot P_i$ 。

所有像素被分为 C_0 和 C_1 两类，设所求灰度门限为 k ，则 C_0 平均灰度为 $u(k) = \sum_{i=1}^k i \cdot P_i$ ，像素占全图比例为 $\omega_0 = \sum_{i=1}^k P_i$ 。 C_1 平均灰度为 $u - u(k)$ ，像素占全图比例为 $\omega_1 = 1 - \omega_0$ 。

对两类的均值做处理，得到 $u_0 = u(k)/\omega_0$ ， $u_1 = [u - u(k)]/[\omega_1]$ 。图像总均值 $u = \omega_0 u_0 + \omega_1 u_1$ ，可以得到类间方差为：

$$\sigma(k) = \omega_0(u - u_0)^2 + \omega_1(u - u_1)^2 = \omega_0 \omega_1 (u_0 - u_1)^2 \quad (2-8)$$

k 即为所求，将小于等于 k 的像素值置零，大于 k 的像素值为 255，得到最终的分割二值化图像。

即使将阈值调整到最优的理想状态，传统的自适应阈值法仍然不能得到令人满意的结果，根据灰度值一分为二策略的前提是水域区域灰度变化较小，岸上区域灰度普遍大于水域区域，但实际情况下这个前提条件不一定满足。

2.2.2.2 活动轮廓法

Liu^[32]等人提出了一种基于活动轮廓模型的方法。活动轮廓模型定义了可变形曲线及其对应的能量函数，以最小化能量目标函数为目标，控制曲线变形，具有最小能量的曲线即为轮廓。此方法将活动轮廓模型与 OTSU 算法结合，实现了 OTSU 分割结果的重新调整。其定义的能量函数为：

$$E(u, c) = \int_D |\nabla u(x)| dx + \lambda \int_D (c - I(x))^2 \cdot (1 - u(x)) dx + \lambda \int_D \alpha T(T - I(x))u(x) dx \quad (2-9)$$

上式中 λ ， α 为常数， T 为 OTSU 得到的阈值， I 表示图像，积分域 D 代表整个图像域。 $u(x)$ 表示分割结果，海洋区域值为零，陆地区域值为一。 $u(x)$ 初始化为 OTSU 二值分割结果。第一个积分项计算分割轮廓的长度，起到消除细小噪声区域的作用。第二项计算海洋区域类内方差，将误分到海洋区域的陆地区域进

行修正。第三项是约束项，防止过多的海洋区域被划分为陆地区域。

活动轮廓法依赖于 OTSU 阈值分割，因此存在相同的缺陷，当海面区域存在高亮的波浪、油污区域或光照造成海面区域灰度高于岸上区域时，海陆分割结果不准确，图 2-8 给出了灰度阈值法及活动轮廓法实现海陆分割的结果。

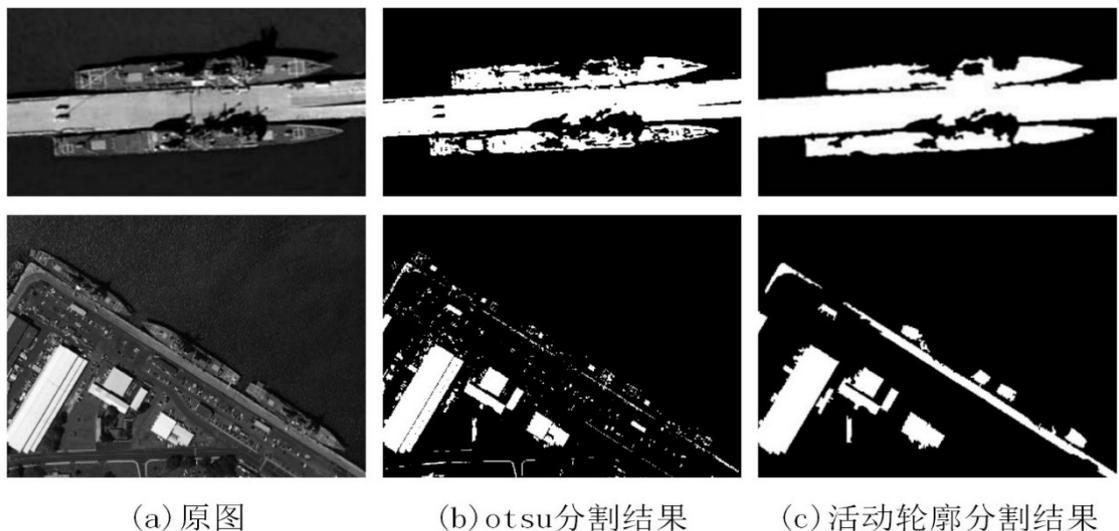


图 2-8 传统方法海陆分割示例

2.3 多角度目标检测

在检测任务中，为了适应目标的角度变化，近年来相关领域的研究人员提出了一些方法，大致分为三种思路：一是提取带角度的候选框，二是设计旋转不变性特征，三是对神经网络进行修改使其对旋转变换具备鲁棒性。

2.3.1 提取带角度的候选框

文献^[33]提出了一种最大化似然函数来获取倾斜候选框的方法。处理流程如下：

表 2-1 带方向的候选框提取算法

输入：待检测图像
初始化：随机选定一些种子点，设定初始候选框。
重复 1-2 步骤直至候选框调整幅度满足设定的限制条件。
1. 对每一个候选框，利用区域内的梯度信息、结构边缘信息及背景先验信息计算目标似然函数。
2. 更新提取框位置和形状使局部似然最大化。
输出：带候选框的图像

这种方法提取的候选框可以帮助检测器处理不同方位的目标，提取框形状与目标更匹配。与直立框相比，其与目标的覆盖更紧密。该算法可以避免出现大量的采样框，减少提取框的同时保持高召回率。然而有方向的候选框提取方式会得到倾斜的图像块，需要进行旋转处理才能进行特征提取和分类，增加运算量的同时给模块之间的连接带来问题，使整个检测架构变得不统一。

2.3.2 设计旋转不变性特征

一些经典的手工设计特征能够提取角度信息以实现对方角度变换的鲁棒性。Gabor^[34]特征可以在频域不同方向上提取相关特征，SIFT 描述子利用特征点邻域像素的梯度方向分布特性为每个关键点指定方向参数，将坐标轴旋转为关键点方向，确保旋转不变性。改进的 LBP 算子以窗口中心像素为阈值，将相邻圆形像素的灰度值与其比较，不断旋转圆形邻域得到一系列初始定义的 LBP 值，取最小值作为该邻域的 LBP 值，实现旋转不变性。改进的环形 HOG 采用环形方式划分细胞单元，并通过权重矩阵与特征向量相乘得到旋转特征。RI-HOG 利用径向梯度转换使提取的特征适应旋转变换。然而手工设计特征只能适应小尺度下的旋转变换，鲁棒性、非线性映射能力、自学习能力较差。

2.3.3 深度学习中的角度自适应

深度神经网络具有简洁有效、可以端到端处理等优势，如何提升卷积神经网络处理多方向目标的能力成为一个热门的研究课题。

2.3.3.1 TI-Pooling

针对神经网络角度变化敏感问题，TI-Pooling^[35]方法对每张输入图像进行一系列空间变换，利用一组平行结构分别对变换后图像进行特征提取，得到特征向量取最大值输入到全连接层。TI-Pooling 通过多示例学习和权值共享来选择输入图像最优的变换示例进行训练，与传统的深度卷积神经网络相比计算量更大，耗费时间更长，且要求有足够的训练样本。而世界范围内的军舰与其他目标相比数量较少，基于保密原则，大量的舰船样本更是难以获取。

2.3.3.2 小波散射卷积网络

小波散射卷积网络^[36] (Scatting Wavelet Network, SWN) 利用小波散射变换分层次地提取图像的高频信息，即图像细节信息，继而利用均值池化操作将高频

转化为低频，保持高频信息的稳定性。均值操作中丢失的高频信息，可在网络的下一层利用散射系数得到恢复。小波散射卷积网络能够提取在仿射变换、弹性形变下的不变性的图像特征表示。

2.3.3.3 空间变换网络

空间变换网络^[37] (Spatial Transform Network, STN) 在卷积神经网络中插入一个附加的网络模块，通过学习仿射矩阵使网络自主地实现平移、缩放、旋转矫正。该网络模块可以插入到卷积神经网络中的任意位置，将扭曲变形的目标调整成为接近直立、尺寸合适的形状，便于下一步的分类和识别。空间变换层由三部分组成，定位网络、网格生成器及采样器，其结构如图 2-9 所示。

定位网络输入为特征映射 U ，输出转换参数 $\theta = f_{loc}(U)$ ，定位网络函数 $f(U)$ 可以为任意形式，可以是全连接网络也可以是卷积网络。定位网络包含一个最终的回归层来产生参数 θ 。不同变换形式下参数 θ 大小不同，对于仿射变换 θ 为六维参数。

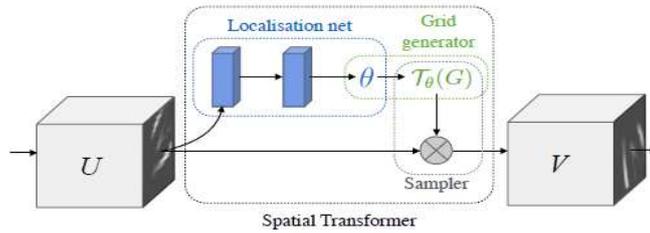


图 2-9 空间变换网络结构^[37]

利用定位网络可得到映射关系，找到转换后的整齐网格中每个像素对应的输入特征图上的位置和源像素。逐行扫描结果特征图，通过目标像素坐标位置找到其对应的源特征图像素位置。以仿射变换为例，坐标变换过程如下：

$$\tau(G_i) = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{bmatrix}' \cdot \begin{bmatrix} x_i^{Target} \\ y_i^{Target} \\ 1 \end{bmatrix} = \begin{bmatrix} x_i^{Source} \\ y_i^{Source} \end{bmatrix}$$

其中 $i = 1, 2, 3, 4, \dots, H * W$ (2-10)

其中 θ 表示转换参数， $(x_i^{Target}, y_i^{Target})$ 表示目标像素坐标， $(x_i^{Source}, y_i^{Source})$ 表示源像素位置。H 为图像高，W 为图像宽。

直接利用变换后的 (x_i^s, y_i^s) 从源特征图中复制像素值是不可行的，这是因为

经过变换后的 (x_i^s, y_i^s) 可能为实数，但是像素位置坐标必须为整数。为了解决像素值缺失的问题，还需要进行双线性插值。利用最大值和绝对值函数，将插值函数改写为式(2-11)，然后通过采样和插值得到最终空间转换后的特征图。

$$V_i^c = \sum_n^H \sum_m^W U_{nm}^c \max(0, 1 - |x_i^s - m|) \max(0, 1 - |y_i^s - n|) \quad (2-11)$$

其中 V_i^c 表示输出特征图第 i 个像素 c 通道的像素值， U_{nm}^c 表示输入特征图 (m, n) 位置 c 通道像素值， (x_i^s, y_i^s) 表示对应的源特征图中的坐标。

本论文使用 STN 对汽车、飞机、舰船切片做角度矫正，得到图 2-10 结果。图中可以看出 STN 矫正效果并不十分理想，分析可能有如下几个原因：1) 利用分类标记来训练网络，由于分类任务较为简单，不足以驱动网络学习到旋转变换；2) 定位网络中的结构对转换参数的学习尤为重要，需要大量的实验来寻找合适的结构。可见，STN 对于复杂背景下角度变化较大的物体适应性不强。

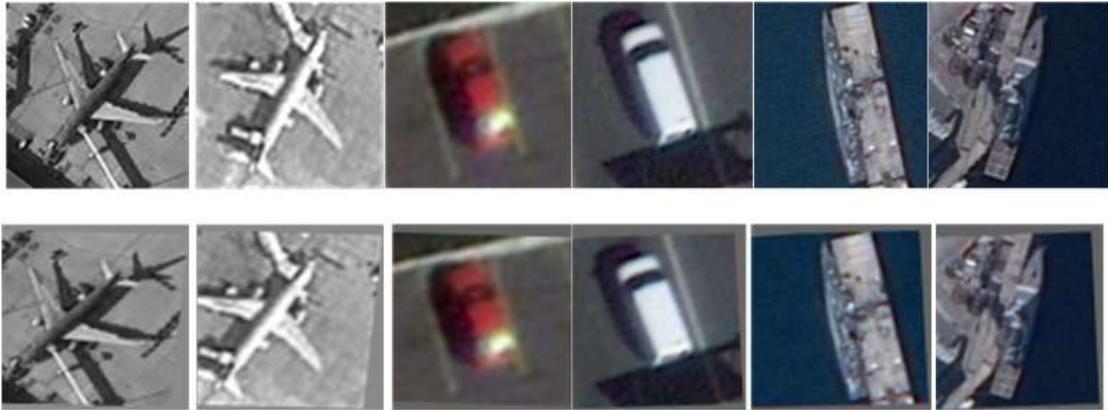


图 2-10 空间变换网络方向矫正示例

2.4 本章小结

本章主要介绍了光学遥感图像舰船目标检测问题中的各种相关技术，并对舰船检测引申出的海陆分割问题和多方向目标的处理进行了讨论。

首先是舰船检测相关技术的总结，说明了靠岸舰船和离港舰船的不同处理方式。舰船检测流程主要包括预处理、感兴趣区域提取、并排舰船的处理、特征提取、目标判别。本章分别对每个步骤中的一些经典算法进行了介绍和分析，简单阐述了现有方法的优缺点。

海陆分割问题可分为基于港口先验信息的方法和基于图像分析的方法。由于港口先验信息获取困难以及图像配准问题，目前水陆划分多采用图像分析法。

比较经典的图像分析法有灰度阈值分割法和活动轮廓法，本章给出了这两种经典方法得到的效果图，剖析了算法中存在的缺陷和问题。

对于多方向目标检测问题，本章总结了目前处理目标角度变化的三种思路：提取带角度的候选框、设计旋转不变性特征、深度学习中的角度自适应。由于本论文的重点是探索基于深度学习的方法，因此重点介绍了空间变换网络实现目标方向矫正的方法并分析了该方法的优点和不足之处。

第3章 基于图割的海陆分割算法研究

海岸线附近复杂的背景,尤其是舰船停泊位置附近的人工设施,给舰船图像解译造成极为不利的影晌。对于包含大量海岸信息的港口图像,海陆分割是遥感图像分析中一个重要的前处理问题,本章将从海陆分割难点入手,提出一种区域粗划分、多类特征精细判别的海陆分割方法。

3.1 问题描述及难点分析

3.1.1 问题描述

俯视视角下的舰船在颜色、灰度、纹理等方面与一些岸上目标极为相似,借助海陆分割算法可以排除大量岸上虚警,提高检测准确度。现有的海陆分割方法大致分为基于先验信息的方法和基于图像分析的方法。基于先验信息的方法是利用港口建成后位置的相对稳定性,利用海岸线地理信息进行水陆划分,但港口位置信息难以获取,且遥感图像配准后与真实地理位置之间仍然存在一定偏差,所以一般采用基于图像分析的方法。基于图像分析的方法根据海洋和陆地之间的差异性来划分区域,现有的图像分析的方法大多只考虑灰度,当海水中存在海浪、油污等高亮区域或光照影响造成亮度不均时,海洋区域灰度可能高于陆地区域灰度,海陆分割问题变得困难。另外,传统图像分析法多是基于像素级别的划分,缺乏全局信息,得到的海陆边界往往是不平滑不完整的(见图 2-9)。

针对以上问题,本论文提出一种将图像划分为区域再对区域进行判别的海陆分割方法。先不考虑海洋和陆地的差异性,利用像素点在特征空间中聚类的方式将图像分割成一些小区域,在再对每个小区域提取海陆差异性特征进行水域非水域的判别。这种不是基于像素而是基于区域的判别方式可以得到全局较好的分割效果,避免海洋区域变得支离破碎。

3.1.2 难点分析

本研究使用了粗划分精判别的方式进行海陆分割,这种海陆分割思路面临以下几个难点问题:

(1) 区域分割阶段——区域内强度差异大于区域间强度差异

对于图 3-1 左侧图像,有三个不同的区域:左半部分的渐变长方形区域,右半边的带孔洞的恒定强度区域和恒定区域中间的强度变化较大的区域。其中区

域 3 内的强度差异大于区域 1 与区域 2 边界上的强度差异，如果用像素值接近于常数或变化缓慢作为区域划分标准，则强度变化大的区域会被判定为多个不同区域，显然无法得到理想的分割结果。推广到真实遥感图像，如 3-1 右侧图像所示，受噪声影响的海洋区域和具有复杂背景的岸上部分都会被分割成很多细小区域。如何制定分割标准、捕获感知重要的属性成为亟待解决的问题。

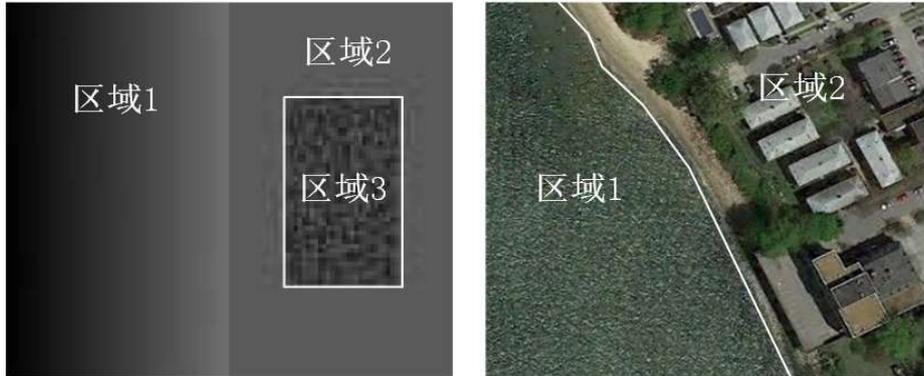


图 3-1 区域内强度差异大于区域间强度差异

(2) 区域判别阶段——海洋与沙滩、草地等的相似性

不同海域的海水在颜色、纹理等方面差异很大，同时某些海洋区域与沙滩、草地等极度相似。如图 3-2 所示，左图中海水呈现灰色而不是常见的深蓝色，与浅滩在颜色上相似。右图中草地区域与带噪声的海洋区域比较相像。即使对于极为智能的人脑也需要根据海面上有码头、船只这些已有知识来判定该区域属于海洋。对于缺乏先验知识的计算机来说，由于提取到的特征极为相似，很容易造成误判。

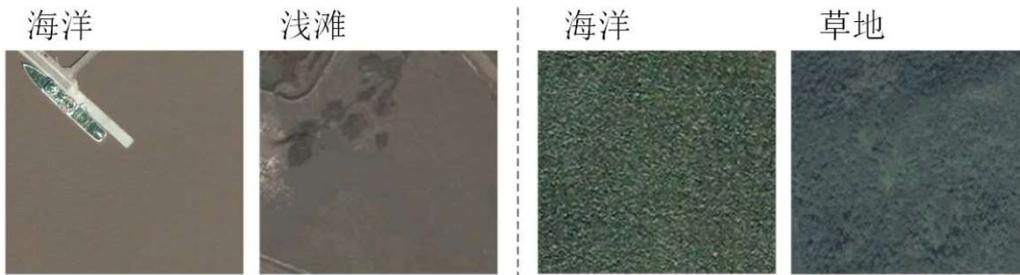


图 3-2 海洋与浅滩、草地等的相似性

3.2 海陆分割算法整体流程

本论文的海陆分割算法处理流程如下，主要包括区域分割、特征提取、区域

判别、水域扩展四个步骤。

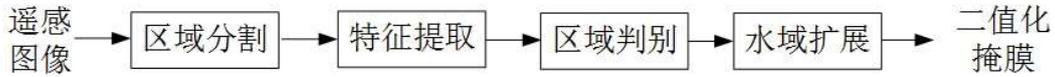


图 3-3 海陆分割算法流程

区域分割：利用图割算法在特征空间对图像中像素点进行聚类，通过迭代将符合设定标准的区域逐步合并，最终将图像分割成几个小区域。

特征提取：提取每个区域的特征。提取的特征有灰度均值、标准差、平滑度、区域面积、矩形度、偏心率、颜色一阶矩二阶矩、共生矩阵的熵和相关性。

区域判别：训练随机森林分类器对提取的多类区域特征进行分类，实现水域非水域的判别。

水域扩展：此时舰船被划分到陆地区域，根据舰船一定临近海洋的前提条件，对区域判别得到的结果进行一定的形态学处理，扩展海洋区域并消除孔洞。

经过上述四个步骤，可以得到最终的水陆分割二值化掩膜图像。

3.3 图割算法分割区域

本课题的海陆分割模块首先采用图割算法^[38]进行区域分割，得到同质性区域。设 $G = (V, E)$ 是一个无向图，其中 $v_i \in V$ 表示待分割的顶点集合， $(v_i, v_j) \in E$ 表示特征空间中连接一对顶点的边。每条边 $(v_i, v_j) \in E$ 有权重 $w(v_i, v_j)$ ，该权重是非负的，用来度量不同相邻元素 v_i 和 v_j 的差异性。树是一种简单意义上的图，树中任意两个结点都有路径相连，但不构成回路。最小生成树是图对应所有生成树中边权值之和最小的一棵树。图 3-4(a)就是一个无向图，棕色加粗的边连接成的树就是无向图对应的最小生成树。

在本实验中，用加权图抽象化表示图像，图像中每个像素代表一个顶点，初始化时每一个像素点都是一个区域，每个区域构成一个最小生成树。将每个像素映射到特征空间 (x, y, r, g, b) 中，其中 (x, y) 表示像素在图像中的位置， (r, g, b) 表示像素各颜色通道的值，使用 ANN^[39]算法找到特征空间 (x, y, r, g, b) 中最邻近的十个点，特征空间中相邻的点可能在图像中相距很远（如果他们的颜色高度相似，而它们之间的像素点与它们颜色不同）。特征空间中相邻的点 v_i, v_j 由边 (v_i, v_j) 连接生成图的边缘，利用顶点对应特征向量的欧几里得距离作为边的权重。当两个

区域之间的差异小于一定阈值时，将区域对应的最小生成树合二为一，利用区域生长的思想迭代合并满足约束条件的区域。直观的说，基于图的分割方式就是在到特征空间中，然后找到相似点的集群，如图 3-4(b)。

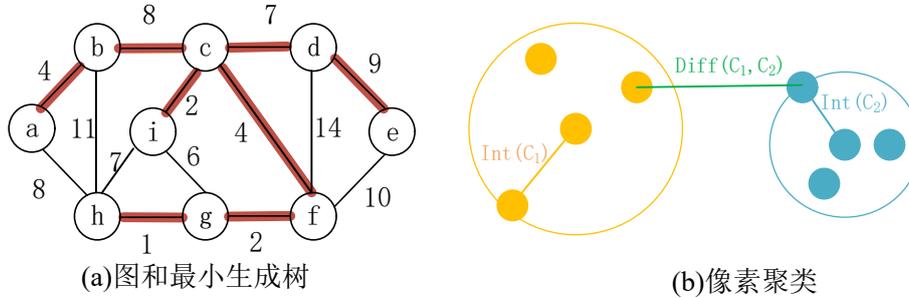


图 3-4 图割算法基本原理

此时需要定义一种划分区域边界的标准。分别计算边界割裂的分属于不同两个不同区域的像素之间的差异性以及区域内部像素之间的差异性。将区域间差异性与区域内差异性进行比较得到分割结果，从而自适应的保护数据的局部特性，判别区域是否需要合并。

对于一个区域 $C \subseteq V$ ，该区域对应的最小生成树表示为 $MST(C,E)$ 。选取区域内部差异性最大的边，即权重最大的一条边，该权重定义为区域的类内差异。以类内最长距离为半径，在特征空间构成一个超球体，该球体与其他类的点不相交，如图 3-4(b)所示。

$$Int(C) = \max_{e \in MST(C,E)} w(e) \quad (3-1)$$

两个区域 $C_1, C_2 \subseteq V$ 的类间差异定义为连接两个区域所有边中，权重最小的边，也就是两个区域最相似的地方的差异性。

$$Dif(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2, (v_i, v_j) \in E} w(v_i, v_j) \quad (3-2)$$

如果没有连接 C_1 和 C_2 的边缘，则将类间差设置为： $Dif(C_1, C_2) = \infty$ 。通过计算类间差 $Dif(C_1, C_2)$ 与类内差 $Int(C_1)$ 、 $Int(C_2)$ 的数值关系，来判断两个区域是否可以合并，区域类间差必须大于最小类内差，即

$$D(C_1, C_2) = \begin{cases} true & \text{if } Dif(C_1, C_2) \leq MInt(C_1, C_2) \\ false & \text{otherwise} \end{cases} \quad (3-3)$$

其中最小类内差 $MInt$ 定义为：

$$MInt(C_1, C_2) = \min(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2)) \quad (3-4)$$

上式中加入自适应阈值函数 τ 是为了解决极端情况下的过分割问题，如果直接定义最小类内差为 $\min(Int(C_1), Int(C_2))$ ，则可能出现特殊情况：当两个区域都是单个像素时， $Int(C) = 0$ ，此时出现“零容忍”情况。只有两个像素的值一样时才合并，最终导致过分割。所以刚开始的时候，应该给每个像素点设定一个可以容忍的范围，当区域增长到一定程度，再弱化该初始容忍值的作用。

阈值函数的形式为 $\tau(C) = k/|C|$ 。其中 $|C|$ 表示区域内包含像素点个数， k 为常数。 k 的值越大，最小类内差越大，越容易满足合并条件。实际应用中 k 设置了形成区域的大小，一般来说 k 越大得到的区域越大，得到的分割结果越粗，如图 3-5 所示。但 k 并不是图像中最小区域的面积，当一个小区域与其他临近区域的差异足够大时，可以得到面积小于 k 的区域。

算法的具体步骤如下：

表 3-1 基于图的分割算法

输入：待分割图像
<ol style="list-style-type: none"> 1. 使用高斯滤波器平滑图像，去除噪声，降低图像锐度。 2. 计算每一个像素点与其邻域在特征空间中的欧几里得距离得到权重。将边按照差异度非降序排列，得到序列 $\pi = (o_1, \dots, o_m)$。（权重相等的边处理顺序不影响最终分割结果） 3. 定义初始分割 S^0，每个像素点作为一个独立的区域。 <p>重复第 4 步，令 $q=1, \dots, m$。</p> <ol style="list-style-type: none"> 4. 通过 S^{q-1} 构造 S^q。选择边 o_q 进行合并判断，设其连接的顶点为 (v_i, v_j)，如果满足合并条件：点 v_i, v_j 在 S^{q-1} 中不属于同一区域且边的权重 $w(o_q)$ 不大于两个区域内部的差异度。则进行第 5 步合并两个区域，否则不进行处理。 5. 更新类别标号，将 v_i, v_j 所在区域类标号统一为 v_i 所在区域标号。更新该类差异性阈值为：$w_{ij} + \frac{k}{ C_i + C_j }$。 6. 返回最终分割结果 $S = S^m$
输出：分割结果图

图 3-5 中给出了图割算法进行区域分割的结果示例图像。基于图的分割方法产生的分割结果既不过粗也不过精。图 3-5 中图像的海洋区域中有许多海浪起的高亮部分，而图割算法将其划分为一个整体而不是将其切成一个个小的图像块。图割算法具有接近线性和低常数因子的运行速度。分割算法得到的海洋和陆地的边界比较精确，可以产生全局意义上的较好分割结果。

图像分割领域中，常用区域一致性、区域对比度等标准来评价分割算法的好坏。由于本文对分割算法的要求具有一定的特殊性，对于强度变化剧烈的背景区

域，我们不希望分割算法进行过于精细的划分，只需要从全局意义上将水和陆地分离开。因此本文只给出直观的分割结果而不对算法进行定量评价。

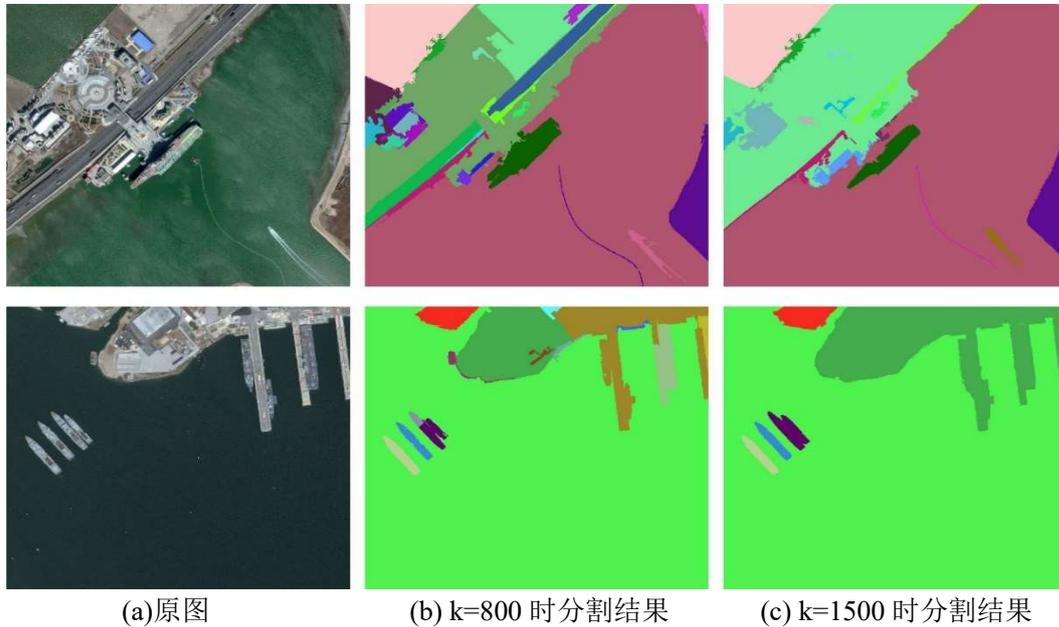


图 3-5 区域分割结果示例

3.4 区域混合特征提取

上一步分割得到很多小的区域，我们对每个小区域提取特征进行分类。海洋与陆地相比，具有灰度较暗、变换缓慢、形状多变等特点，陆地上多为刚性结构，形状较规则。本论文提取区域的灰度特征、几何特征、颜色特征、纹理特征等混合特征用于区域的分类。

3.4.1 灰度特征

首先利用公式 $N = 0.299 \times r + 0.587 \times g + 0.114 \times b$ 将 RGB 三通道彩色图像转换为灰度图像，再计算区域的灰度均值、标准差、平滑度。

灰度均值表征区域内像素点的平均亮度，一般水域灰度均值较小。计算公式如下：

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} \quad (3-5)$$

其中 N 表示区域内像素个数， x_i 表示第 i 个像素点的灰度值。

灰度标准差反应图中高频部分大小，度量区域内像素的对比度。总体来说陆

地区域的标准差大于海洋区域的标准差。灰度标准差计算公式为：

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (3-6)$$

灰度平滑度通过计算每个点与其 8 邻域像素灰度值之差来表示区域内灰度变化程度，一般情况下海洋区域灰度相对平滑。对应公式如下：

$$smooth = \frac{\sum_{i=1}^N |x_i - \sum_{k=1}^8 x_i^k / 8|}{N} \quad (3-7)$$

其中 x_i^k 表示第 i 个像素的第 k 个相邻点的灰度值。

3.4.2 几何特征

几何形状特征可以简单直观地表达图像的位置、方向、周长、面积等信息。本章中我们针对海洋区域面积在图像中具有所占比例较大、形状不规则、区域分布相对不紧凑等特点提取了区域归一化面积、矩形度、偏心率等鉴别特征，下面分别叙述。

归一化面积对区域内像素总数进行统计，并计算该区域占整幅图像的比例。其表达式如下：

$$S = \frac{S_0}{w \times h} \quad (3-8)$$

其中 w, h 分别表示图像的宽和高， S_0 表示区域内像素个数。

矩形度的大小表示了一个区域相对于其外接矩形的充满程度。区域越接近矩形则矩形度越大。用 S_{MER} 代表外界矩形面积，则矩形度的定义为：

$$R = \frac{S_0}{S_{MER}} \quad (3-9)$$

偏心率也称为伸长度，在一定程度上反映了区域的紧凑性。计算步骤如下：

$$(1) \text{ 计算水平和竖直坐标平均值: } x_0 = \frac{1}{M} \sum_{i=1}^M x_i, \quad y_0 = \frac{1}{N} \sum_{j=1}^N y_j \quad (3-10)$$

$$(2) \text{ 计算 } j+k \text{ 阶中心矩: } M_{jk} = \sum_{i=1}^M \sum_{j=1}^N (x_i - x_0)^j (y_i - y_0)^k \quad (3-11)$$

$$(3) \text{ 计算偏心率: } E = \frac{(M_{20} - M_{02})^2 + 4M_{11}}{S_0} \quad (3-12)$$

3.4.3 颜色特征

海洋区域一般呈现深蓝色，与陆地区域的颜色差异较大，因此可以利用颜色特征作为区域鉴别的依据。相对于 RGB 空间，HSV 颜色空间能够更加直观地表达色彩的鲜艳程度、明暗等，更符合人眼的视觉特性。所以，我们将图像转换到 HSV 颜色空间，再通过非等间隔量化将三个颜色分量合成为一维特征向量，最后计算区域的颜色矩作为鉴别特征。

首先将 RGB 图像转换为到 HSV 颜色空间，转换公式如下^[40]：

$$V = \max(R, G, B) \quad (3-13)$$

$$S = \begin{cases} 0 & \text{if } \max(R, G, B) = 0 \\ \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} & \text{otherwise} \end{cases} \quad (3-14)$$

$$H = \begin{cases} \text{undefined} & \text{if } \max(R, G, B) = \min(R, G, B) \\ \frac{60^\circ \times (G - B)}{\max(R, G, B) - \min(R, G, B)} + 0^\circ & \text{if } \max(R, G, B) = R \text{ and } G \geq B \\ \frac{60^\circ \times (G - B)}{\max(R, G, B) - \min(R, G, B)} + 360^\circ & \text{if } \max(R, G, B) = R \text{ and } G < B \\ \frac{60^\circ \times (B - R)}{\max(R, G, B) - \min(R, G, B)} + 120^\circ & \text{if } \max(R, G, B) = G \\ \frac{60^\circ \times (R - G)}{\max(R, G, B) - \min(R, G, B)} + 240^\circ & \text{if } \max(R, G, B) = B \end{cases} \quad (3-15)$$

HSV 图像包含的颜色信息十分丰富，在海洋和陆地鉴别过程中我们希望对颜色变化有一定的容忍能力，因此对 HSV 颜色进行量化^[41]降维，便于后续处理。将色调分成 8 份，饱和度和亮度分成 3 份进行非等间隔量化。用 H' 、 S' 、 V' 表示量化后的三个颜色通道的值，通过下式得到合并颜色通道后的颜色值^[40]：

$$L = 9H' + 3S' + V' \quad (3-16)$$

最后计算区域颜色的一阶矩(均值)： $\bar{L} = \frac{\sum_{i=1}^N L_i}{N}$ 、以及颜色二阶矩(标准差)：

$$\sigma_L = \sqrt{\frac{1}{N} \sum_{i=1}^N (L_i - \bar{L})^2}$$
 构成颜色特征向量。

3.4.4 纹理特征

水和岸上区域在纹理特征方面有显著的差别。灰度共生矩阵能反映图像中具有某种特殊灰度值的点对出现的频度，通过对这些点对的统计可描述区域的纹理。本文提取灰度共生矩阵的熵和相关性作为纹理鉴别特征。

为了减少计算量，将图像灰度量化为 16 级。给定距离和角度计算区域对应的灰度共生矩阵。一般取距离为 1，角度为 0° 、 45° 、 90° 、 135° 。灰度共生矩阵的计算公式如下^[42]：

$$p(i, j | 1, 0^\circ) = \#\{|k - m| = 0, |l - n| = 1, f(k, l) = i, f(m, n) = j\} \quad (3-17)$$

$$p(i, j | 1, 45^\circ) = \#\{(k - m) = 1, (l - n) = -1 \\ \text{or } (k - m) = -1, (l - n) = 1, f(k, l) = i, f(m, n) = j\} \quad (3-18)$$

$$p(i, j|1, 90^\circ) = \#\{|k - m| = 1, |l - n| = 0, f(k, l) = i, f(m, n) = j\} \quad (3-19)$$

$$p(i, j|1, 135^\circ) = \#\{(k - m) = 1, (l - n) = 1 \\ \text{or } (k - m) = -1, (l - n) = -1, f(k, l) = i, f(m, n) = j\} \quad (3-20)$$

上式中 (k, l) 、 (m, n) 表示区域中点的坐标， $f(\cdot)$ 表示像素点对应的灰度量值， $\#$ 表示符合条件的像素对数， $p(i, j)$ 表示在给定距离和角度的情况下，共生矩阵在 (i, j) 位置点上的值。

得到灰度共生矩阵后，一般不直接应用，而是在共生矩阵的基础上计算纹理特征量。常用的纹理特征量有反差、能量、熵、相关性等，本文为了更好地判别海洋区域，只应用了熵和相关性。

共生矩阵的熵可用来考量区域内灰度级的混乱程度。较为复杂的区域灰度变化较大，熵值也比较大，反之则熵值较小。其计算公式为：

$$En = -\sum_i \sum_j p(i, j) \log p(i, j) \quad (3-21)$$

共生矩阵的相关性也被称为同质性，可用于表示区域灰度级在行和列方向上的相似程度。区域内灰度值越接近，相关性越大。其表达式如下：

$$Corr = [\sum_i \sum_j ((ij)p(i, j)) - \mu_x \mu_y] / \sigma_x \sigma_y \quad (3-22)$$

其中 μ_x 、 σ_x 表示共生矩阵在行方向上的均值和标准差， μ_y 、 σ_y 表示共生矩阵在列方向上的均值和标准差。

将本节所述的灰度统计特征、几何形状特征、颜色特征、纹理特征组合起来，作为鉴别特征向量，特征向量共16维，见下表：

表 3-2 区域混合特征

灰度特征	几何特征	颜色特征	纹理特征
灰度均值	面积	一阶矩	共生矩阵的熵（4个方向）
灰度标准差	矩形度	二阶矩	共生矩阵的相关性（4个方向）
灰度平滑度	偏心率		

3.5 随机森林判别水域

利用基于图的分割算法进行区域划分后，还需要根据上一节提取的区域鉴别特征的差异性，训练分类器进行水域判别，得到水陆分割的二值图像。由于随机森林^[43]具有实现简单、训练速度快、可以自动选择某些重要特征、模型泛化能

力强等优势，本文选择随机森林分类器来进行区域分类。随机森林是基于决策树的预测分类方法，每次有放回的随机抽取训练样本的一个子集，再从该子集中随机选择一部分特征来构造决策树，最后对每棵决策树的预测分类结果进行“投票”，得到最终预测结果。

数据准备

为了利用区域的类别标号来监督分类器的训练，首先需要对图像进行区域标注。本论文编写了分割标注工具，其界面如图 3-6 所示。左侧图像为原图，右侧图像为 3.3 节得到的分割结果对应的轮廓图，在窗体上单击轮廓图对应区域输入数字进行标注，0 表示陆地，1 表示海洋。图 3-6 中副窗体中显示的二值图像即为海陆分割标注结果。通过标注我们得到了区域类别真值，跟着对分割出的每一个区域提取如 3.4 节中所述的 16 维混合特征。

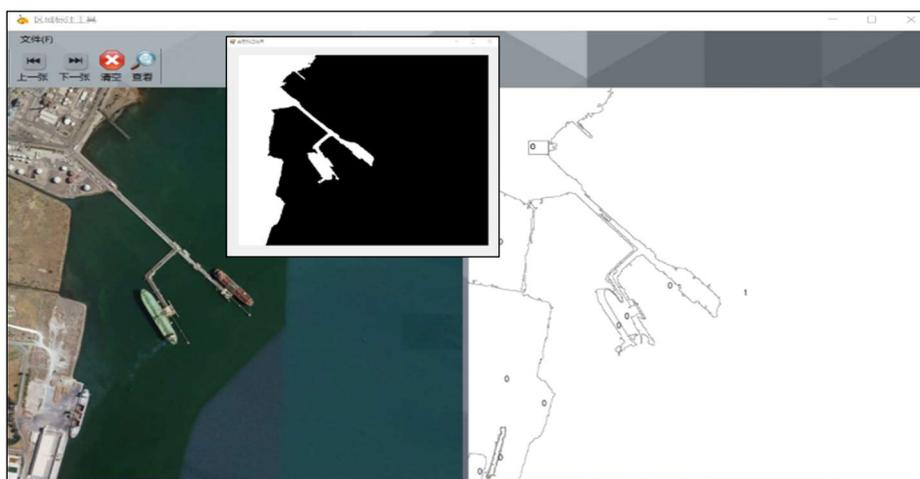


图 3-6 区域标注工具

用于训练的样本来源于民用港口和部分军事港口的 424 幅图像中 11,192 个区域的特征，测试样本为剩余 243 幅军事港口图像中 6,615 个区域对应的特征。

决策树构建

决策树利用树形结构模拟了人通过事物的某些属性来逐步判定结果的过程。树的叶子结点对应分类结果，非叶子结点对应了一个属性的测试，每个节点包含的样本集合根据测试结果被划分到子结点中；根结点到叶子结点的路径表示一个完整的测试序列。决策树利用分而治之的策略逐步形成对一个事件的判定结果。决策树的实现过程如下：

表 3-3 决策树算法

输入：训练集 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ 属性集 $P = \{p_1, p_2, \dots, p_d\}$
过程：函数 $\text{TreeGenerate}(D, P)$ <ol style="list-style-type: none"> 1. 创建根节点 node。 2. 如果 D 中的所有样本都属于同一类别 C，则将 node 标记为 C 类叶子结点。返回结果。 3. 如果 P 为空集或 D 中所有样本在属性集 P 上的取值都相同，则将 node 标记为叶子结点，使 $n_i W_i$ 达到最大的类别作为 node 的类别标记，其中 n_i 表示 D 中属于第 i 类的样本数量，W_i 表示第 i 类权重。返回结果。 4. 计算每个属性的基尼指数，从属性集中选择最优的属性 p_*。 5. 给属性 p_* 设定划分阈值 t_p，将属性值小于等于 t_p 的划分为左子树，否则划分为右子树。D_v 表示 D 中属性 p_* 取值在左子集或右子集的样本。 如果 D_v 为空，则将分支结点标记为叶子结点，使 $n_i W_i$ 达到最大的类别作为该结点的类别标记，返回结果。否则执行函数 $\text{TreeGenerate}(D_v, P \setminus \{p_*\})$。
输出：决策树

本文通过计算加权基尼不纯度来寻找属性的划分阈值。公式如下：

$$i(N) = \frac{\sum_{i=1}^C (n_i W_i)^2}{\sum_{i=1}^C n_i W_i} \quad (3-23)$$

$$\Delta i = i(N) - i(N_L) - i(N_R) \quad (3-24)$$

其中， N 表示划分前结点， n_i 表示结点内属于第 i 类的样本数量， W_i 表示第 i 类权重， N_L 、 N_R 分别表示划分后左结点和右结点。 Δi 表征不纯度的减少量，值越大代表分离效果越好。

基学习器集成

随机森林算法每次从训练集中随机地选择样本子集，再从 16 维中随机地选择 4 维特征构建基决策树，训练时树与树之间是相互独立的，可以并行化训练。每棵基决策树都会预测一个分类结果，为了得到最终的判决结果，需要将多个分类器进行结合。本论文采用相对多数投票法得到判定结果，其表达式如下：

$$H(x) = c_{\arg \max \sum_{i=1}^T h_i^j(x)} \quad (3-25)$$

其中 $h_i^j(x)$ 表示第 i 个基学习器在类别 j 上的输出。

单一学习器很可能因为误选导致泛化性能不佳，而将多个学习器进行组合则能提高分类准确度。本论文对不同数量基学习器得到的判定结果的准确度进行了比较，见表 3-4。基学习器数量为一时，随机森林退化为部分特征训练的决策树，对测试样本的分类准确率为 82.4%。随着基决策树数量的增加，准确率逐渐提高。当数量达到一定程度时，准确率趋于平稳不再提升。

表 3-4 基学习器数量与测试准确率

基学习器个数	1	5	30	50
测试准确率	82.4%	91.6%	98.6%	98.5%

判别结果

根据分类器预测的类别标号就可以得到海陆分割结果图了，如图 3-7 所示。利用多类混合特征训练的随机森林分类器能较好地鉴别区域。

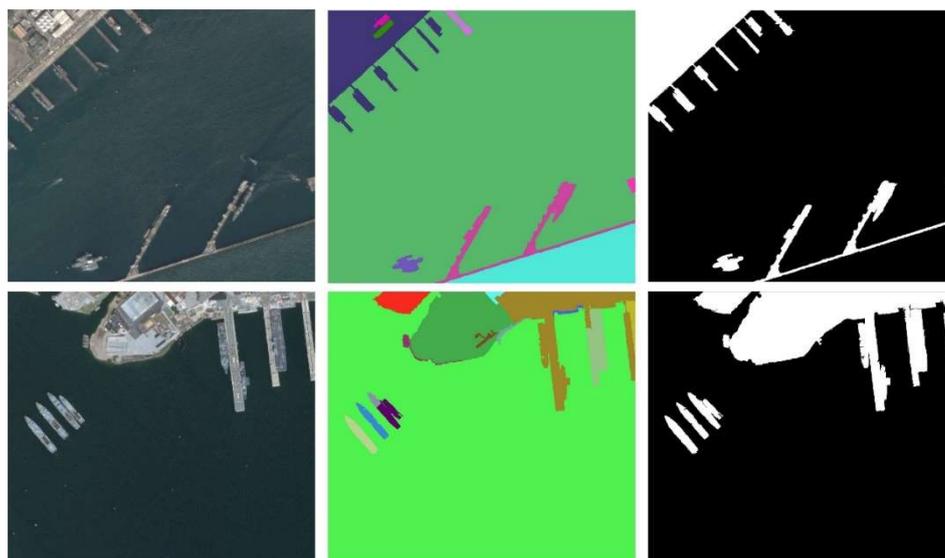


图 3-7 水域判别结果示例

为了说明随机森林分类器在水域判别方面的优势性，我们将该分类器与同样属于集成学习的 Boosting 分类器及流行的分类方法支持向量机进行比较，得到下表结果：

表 3-5 分类算法准确率比较

分类方法	支持向量机	Boosting	随机森林
测试准确率	93.6%	97.7%	98.6%

分析随机森林得到的分类结果优于 SVM、Boosting 算法的原因如下：

(1) 随机森林有放回地随机选择部分特征分别训练决策树，而不是所有特征参与训练。在海洋区域与陆地区域分类中，在某些场景下利用颜色特征可以较好地实现分类，而当海洋颜色较亮或者呈现灰色等特异性颜色时，利用区域的几何特征则会更容易鉴别区域。随机森林每次随机选择 4 维特征训练决策树，最后

采取多数表决的形式来提升分类准确率。

(2) 随机森林对于不平衡的数据集的处理方面有一定的优势。在分割算法得到的区域划分结果中,通常水域属于一个完整区域,陆地则被划分成几个区域,这会导致得到的区域数据不平衡,如图 3-5 所示。由于海洋区域数量远远少于陆地区域数量,这种不平衡会严重影响分类结果。这是因为由于海洋区域数量很小,因此,即使将海洋区域误划分为陆地,也能得到很高的准确率。在随机森林算法中,我们赋予海洋较高的权重,可以在一定程度上避免不平衡问题带来的性能下降。

3.6 形态学水域扩展

上一节得到二值化海洋-陆地分割图像,如图 3-7 所示。此时舰船被划分为陆地区域。舰船检测相关算法一般采取在海陆分割基础上进行二次分割的形式,由于舰船与浅海地形的相似性以及舰船与码头的粘连问题,得到的分割结果往往不是很理想。本文不直接提取舰船候选区域,而是利用舰船一定靠海(除了船坞中的舰船,这种情况极少)这一先验信息,对海洋区域进行扩展得到舰船可能存在的区域,辅助检测过程。

首先对海陆分割二值化图像进行形态学腐蚀操作来完成水域扩展。腐蚀实际上是一种求局部最小值的操作,通过卷积过程使图像中的黑色区域逐渐增长。对于海洋区域向岸上区域扩展多少个像素问题,我们对舰船宽度进行统计,设定了一个值 k ,使 99%以上的舰船在图像中宽度所占像素小于这个值。也就是说腐蚀操作迭代进行 k 次,扩展海洋部分区域,将舰船划分到水域。

对于个别尺寸很大的航母,在腐蚀操作后可能会出现孔洞,如图 3-8(c)所示。针对此情况我们进行填充处理。提取腐蚀后得到的二值化图像的轮廓,计算轮廓面积。当该轮廓面积与整幅图像面积之比小于预先设定的值时,则轮廓内像素点赋值为零。

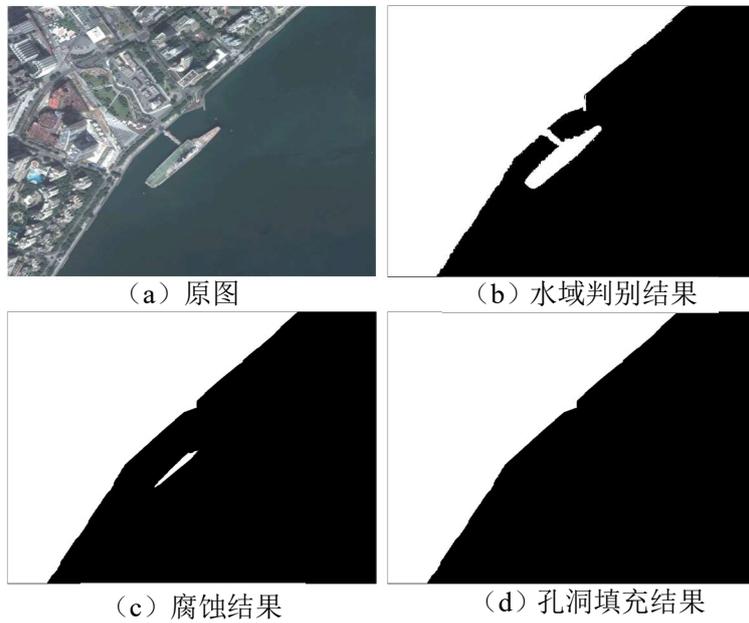


图 3-8 水域扩展结果示例

3.7 实验结果与分析

为了验证本文海陆分割算法的有效性，我们在舰船数据集上进行了实验。图 3-9 给出了本章海陆分割算法中区域分割、水域判别、水域扩展步骤得到的结果示例图。通过一系列处理获得二值化掩膜图像，与原图叠加得到图 3-9 第四列所示的分割结果。

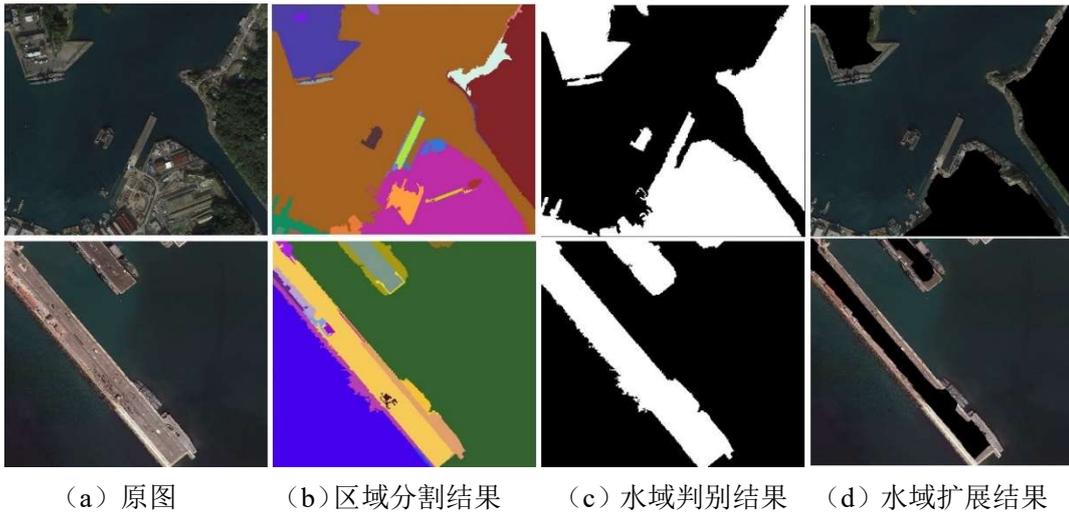


图 3-9 本文海陆分割算法结果示例

由图 3-9 可以看出，我们的分割方案可以正确分类海洋和陆地。值得一提的是，对于内部变化剧烈的岸上区域，我们可以将其完全标记为陆地，而对于传统的像素级别划分的海陆分割方法，如 2.2 节所述，得到的区域是不完整的碎片状。

利用海陆分割算法得到的二值化掩膜图像可对岸上区域进行遮挡，帮助目标检测算法排除大量虚警。我们对加入海陆分割模块前后得到的检测结果进行了对比，见表 3-6。其中虚警率表示检出虚假目标个数占全部检出目标的概率，查全率表示检测正确的目标与所有真实目标的比例。表中结果可以反映出，在目标检测过程中加入海陆分割模块，可在保证查全率基本不变的前提下使虚警率大幅下降。

表 3-6 海陆分割带来的检测性能变化

有无海陆分割	虚警率	查全率
✘	58.83%	86.97%
✓	33.21%	86.74%

3.8 本章小结

海陆分割是船舶检测中非常重要的辅助步骤，现有的逐像素判别的方法得到的分割结果容易陷入局部最优，得到碎片化的分割区域，很难应用于实际。针对当前算法存在的缺陷，本章提出了一种新的海陆分割方法。

首先用基于图的分割算法将图像分割成若干同质区域，得到全局效果较好的粗略分割结果。再对每个区域提取能够区分海洋和陆地的差异性特征，包括灰度特征、几何特征、颜色特征、纹理特征等，组合成 16 维混合特征向量。接着利用随机森林分类器进行区域鉴别，通过随机选择样本、随机选择特征的方式来训练基决策树，再对基决策树进行集成提高分类准确度。由于此时舰船可能被划分到陆地区域，通过形态学运算来扩展水域，由海洋区域扩展到近海区域（可能存在舰船的区域），得到最终的海陆分割掩膜图像。通过实验分析，我们发现本文提出的海陆分割算法得到了较为准确的分割结果，并且通过对岸上区域遮盖的方式可以极大地降低目标检测虚警率。

第4章 基于深度学习的鲁棒性舰船检测

目前相关领域的研究人员已提出一些适用于舰船的目标检测方法，然而由于遥感图像中舰船目标姿态多变、多目标并排停放等原因，现有的检测算法难以处理这些问题，检测效率不高，很难应用于实际。本课题提出了一种新的舰船检测思路，基于深度框架实现鲁棒性舰船检测。

4.1 问题描述及解决方案

如 1.2.2 中所述,传统光学遥感图像舰船检测中面临的难点问题主要有三个:多角度舰船漏检问题、并排舰船漏检问题、虚警问题。

针对以上问题,本论文提出以下解决思路:

针对问题一,由于在角度变化较大的情况下,深度卷积网络对多角度舰船目标的处理能力较差,因此,我们在卷积过程中,主动旋转滤波器以得到多方向通道特征,对各个方向特征求最大值获得旋转不变深度特征,从而实现角度鲁棒的舰船目标检测。

针对问题二,依赖于深度卷积网络中卷积层的定位能力直接找到图像中有利于舰船判别的关键区域,而不是像传统的目标检测算法那样提取候选框,这样可有效避免并排舰船候选框重叠,从而减少并排漏检。

针对问题三,由于绝大部分虚警都位于陆地区域,因此可通过海陆分割算法遮盖岸上区域,排除虚假目标。另外,还可以加入难反例挖掘步骤,把一些得分较高的非舰船图像块加入到训练集反例中,使检测网络的损失更小,从而达到提高检测性能的目的。

4.2 舰船检测研究框架

本小节将给出本章的整体研究框架,按照框架流程介绍本章舰船检测的关键环节方法。

4.2.1 整体研究框架

本课题的整体研究框架如图 4-1 所示。

训练阶段:首先需要准备训练样本。由于舰首的 V 型结构是航拍视角下舰船的最显著特征,因此在训练图像集中截取舰首切片作为正例样本。为了保证数

据集中角度的多样性,对训练图像进行随机的角度旋转,计算出旋转后的真值框,在真值框附近进行随机微小偏移得到正例切片。反例切片在背景中随机选取。将样本切片送入已在 CIFAR 数据集上预训练好的深层残差网络,根据类别标号对网络进行参数微调。其中角度不变性是通过将卷积滤波器替换为主动旋转滤波器^[44](Active Rotating Filter, ARF),对各个方向特征求最大值来实现的。

测试阶段:将训练过程使用的深度卷积网络去掉最后的全连接分类层和全局最大池化层得到特征提取网络,将测试图像送入网络得到特征图。采用分类激活特征图^[45](Class Activation Mapping, CAM)的方法得到舰船这一类别的响应图。利用地理等高线思想对得到的响应图求连通区域,得到初步检测框,最后对该区域的方向和长宽进行估计得到最终检测框。

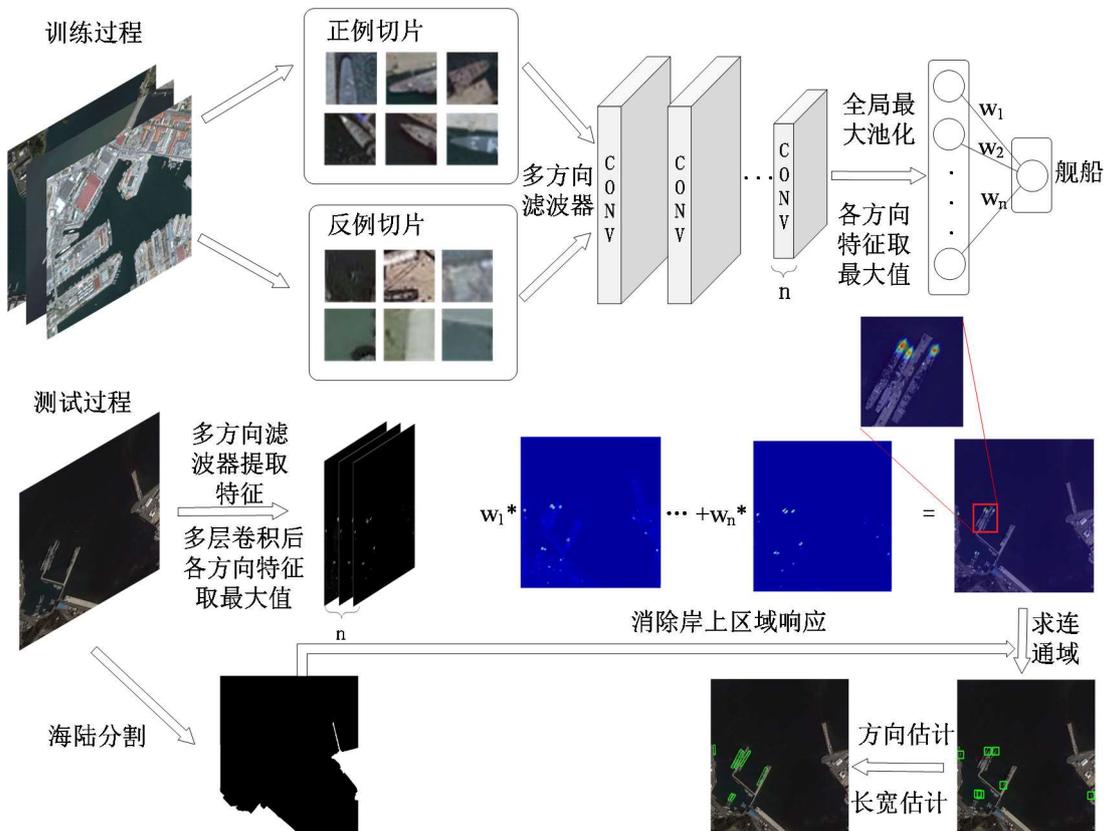


图 4-1 整体研究框架

4.2.2 可变角度滤波器实现旋转不变性

由于卷积神经网络缺乏对旋转的理解能力，为了适应目标复杂的角度变化，网络需要更多的滤波器来学习不同方向下的目标结构，这在增加训练时间的同时也降低了网络泛化能力。例如，现有网络模型中存在一些结构相似但角度不同的滤波器，图 4-2 所示为 AlexNet 网络在 ImageNet 数据集上训练后卷积滤波器的可视化结果。

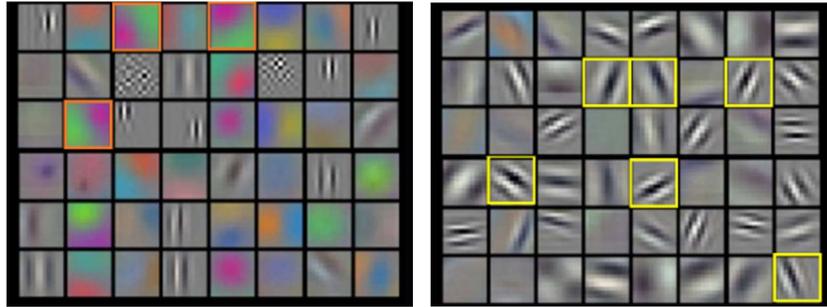


图 4-2 原始单方向滤波器角度冗余

为了解决卷积网络单方向滤波器角度冗余问题，增强网络的方向鲁棒性，本论文采用主动旋转滤波器^[44]的方式获取多个方向下图像特征。

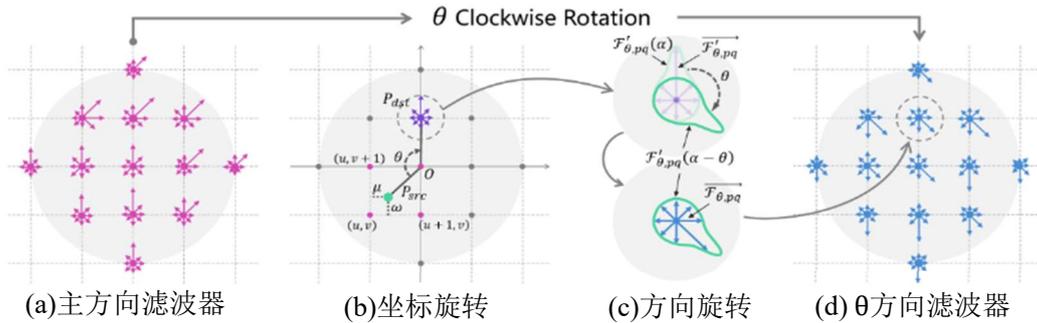


图 4-3 滤波器旋转^[44]

一组旋转滤波器 ARF 是由一个尺寸为 $W \times W$ ， N 个方向通道的滤波器旋转 $N-1$ 次得到的 N 个方向版本的滤波器，只有主方向滤波器 F 被实现和学习，剩余的 $N-1$ 个滤波器是未实现的复制版本。滤波器组中的第 $n(n \in [1, N - 1])$ 个滤波器是由主方向滤波器顺时针旋转 $\frac{2\pi n}{N}$ 度得到的。主方向滤波器上的每个元素可用 $\overline{F}_{ij}^{(n)}$ 表示，其中 $0 \leq |i|, |j| \leq \frac{W-1}{2}$, $0 \leq n \leq N - 1$ ，由标准方向滤波器旋转得到 θ 方向滤波器 F_{θ} 需要两步，坐标旋转和方向旋转，如图 4-3 所示：

(1) 坐标旋转

如图 4-3 所示, ARF 在一个圆形区域内旋转, F_θ 上的点 (p, q) , 是由 F 上对应的四个临近的点 (p', q') 插值得到的, $(p', q') = (p, q) \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$ 。插值方式为双线性插值, 表达式如下:

$$\begin{aligned} \overrightarrow{F'_{\theta,pq}} = & (1 - \mu)(1 - \omega)\overrightarrow{F_{uv}} + (1 - \mu)\omega\overrightarrow{F_{u,v+1}} + \mu(1 - \omega)\overrightarrow{F_{u+1,v}} \\ & + \mu\omega\overrightarrow{F_{u+1,v+1}} \end{aligned} \quad (4-1)$$

其中 $u = \lfloor p' \rfloor, v = \lfloor q' \rfloor, \mu = p' - u, \omega = q' - v$, 超出圆形范围的点补零。

(2) 方向旋转

F_θ 上每个元素的值 $\overrightarrow{F'_{\theta,pq}}$ 是滤波器关于方向的函数 $F'_{\theta,pq}(\alpha)$ 在 N 个期望方向上的采样, 其中 $F'_{\theta,pq}(\alpha)$ 是以 2π 为周期的连续周期函数。在坐标旋转阶段, 我们对滤波器中每一个元素相对于滤波器的位置进行了旋转, 但对于每个元素的 N 个方向并没有进行旋转, 因此还需要对每个点顺时针旋转 θ 得到 $\overrightarrow{F_{\theta,pq}}$, 量化为 $F'_{\theta,pq}(\alpha - \theta)$ 。这种旋转变换可以通过频域离散傅里叶变换的循环移位性质来实现:

$$X(k) \equiv DFT\{\overrightarrow{F'_{\theta,pq}}^{(n)}\} = \sum_{n=0}^{N-1} \overrightarrow{F'_{\theta,pq}}^{(n)} e^{-jk\frac{2\pi n}{N}}, k = 0, 1, \dots, N-1 \quad (4-2)$$

通过逆傅里叶变换求 N 个采样点的值:

$$\overrightarrow{F'_{\theta,pq}}^{(n)} \equiv IDFT\{X(k)e^{-jk\theta}\} = \frac{1}{N} \sum_{k=0}^{N-1} X(k)e^{jk(\frac{2\pi n}{N} - \theta)}, n = 0, 1, \dots, N-1 \quad (4-3)$$

为了平滑地处理所有旋转角度, ARF 需要足够多的方向通道。实际应用中, 由于网络中存在多个池化层, 池化操作能实现近似方向插值的效果, 因此可以使用有限的几个方向通道来保证准确度。为了在取得较高准确度的同时避免计算量过大, 本文使用 8 个方向的滤波器。

处理了滤波器旋转问题后, 接下来还需要解决误差反传的问题。在反向传播过程中, 误差信号 $\delta^{(k)}$ 由主方向滤波器及其所有旋转版本的误差 $\delta_{-\theta_k}^{(k)}$ 汇总得到, 利用总误差更新主方向版本的滤波器。

$$\delta^{(k)} = \frac{\partial L}{\partial F_{\theta_k}}, \theta_k = k \frac{2\pi}{N}, k = 0, \dots, N-1, F \leftarrow F - \eta \sum_0^{N-1} \delta_{-\theta_k}^{(k)} \quad (4-4)$$

其中 L 表示训练误差, η 表示学习率。后向传播过程只更新主滤波器, 当训练图像中大量存在特征相似方向不同的结构时, 利用多角度下误差更新主滤波器这种方式与传统卷积网络相比更具优势, 可以学习到更为“纯净”的滤波器。

实验中我们首先对原始图像进行复制, 扩充为 8 通道图像, 每个通道相同。

定义多角度卷积为 $\tilde{M} = ORConv(F, M)$ ，输出的特征图 \tilde{M} 由 8 个方向通道组成，第 k 个通道计算如下：

$$\tilde{M}^{(k)} = \sum_{n=0}^{N-1} F_{\theta_k}^{(n)} * M^{(n)}, \theta_k = k \frac{2\pi}{N}, k = 0, \dots, N-1 \quad (4-5)$$

其中 F_{θ_k} 为 F 顺时针旋转 θ_k 得到的旋转版本， $F_{\theta_k}^{(n)}$ 和 $M^{(n)}$ 分别表示 F_{θ_k} 和输入特征图的第 n 个方向通道。

在多方向卷积过程中，ARF 主动获取图像的多角度响应，多层卷积得到 8 个方向通道的特征图，对 8 个方向通道的特征图求最大值。图 4-4 给出了 ARF 对四个角度下舰船切片提取特征并对各方向通道特征取最大值后的特征图。可以看出尽管舰船方向变化很大，但我们得到的特征图是方向不同、响应相似的。

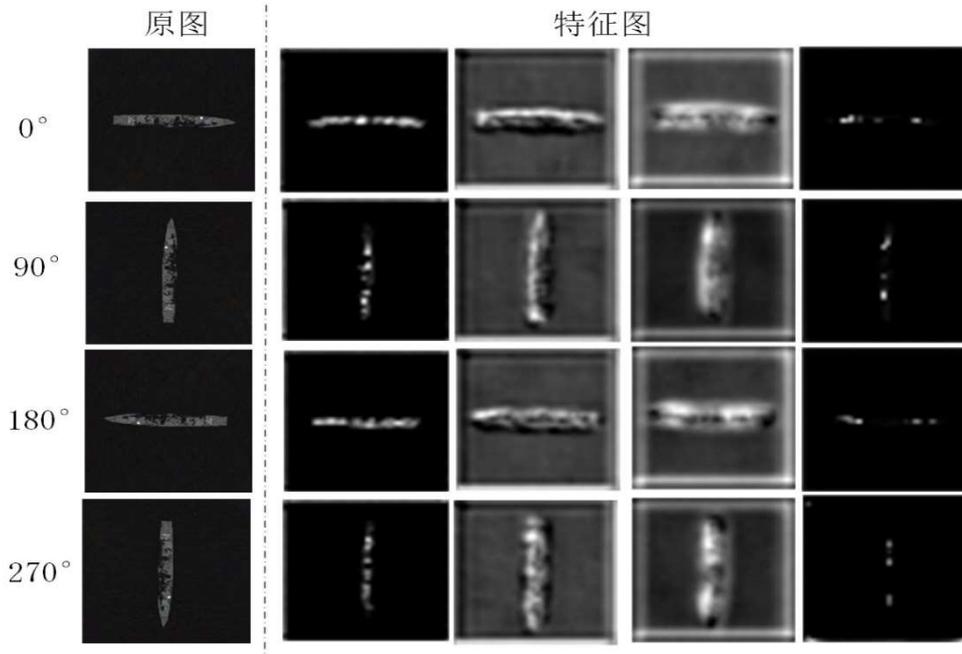


图 4-4 ARF 提取特征图示例

4.2.3 全卷积网络提取特征

传统的卷积神经网络一般来说是由卷积层、池化层、全连接层堆叠起来的网络。全连接层存在参数冗余问题，该层参数占整个网络参数 80% 左右。最近一些性能优异的全卷积神经网络如 NIN、GoogLeNet 等，提出使用全局平均池化替代全连接层来减少参数数量，同时保持高性能。全连接层逐渐被取代的另一个原因是全连接层会导致网络定位能力丧失。最近的文献^[46]表明，尽管没有提供目标位置的监督信息，卷积神经网络中的卷积层部分可作为目标检测器。卷积层所具有

的这种出色的定位能力，会在加入用于分类的全连接层之后丧失。基于上述两个原因，本章采用全卷积网络来提取特征。

4.2.3.1 残差网络

文献^[47]表明网络深度是视觉任务中一个至关重要的因素，一般情况下网络越深效果越好。本文中我们希望利用较深的网络模型来获取更高的检测性能。然而随着网络层数的增加，分类精度达到饱和，当层数超过一定数值时，分类精度显著降低。为了避免精度退化，本章利用残差网络^[48](ResNet)进行特征提取。

残差网络结构是在标准的前馈卷积网络上，增加捷径连接跳过神经网络中的一层或多个层而直接连接不相邻的层，每个捷径连接产生一个残差块，如图 4-5 所示。每层的输出不再是传统神经网络当中输入的映射，而是映射和输入的叠加。如果映射和输入的维度不等，则通过一个线性投影矩阵进行维度匹配再相加。在这种结构下，神经网络加层后至少能获得与之前相同的性能，而不会更差。

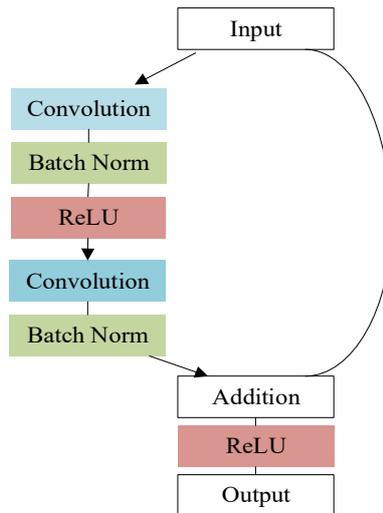


图 4-5 残差网络结构^[48]

残差网络结构中卷积层之后一般需要加入批归一化层(Batch Norm)和 ReLU 层。批归一化层进行归一化处理，去除数据相关性，ReLU 层对输入信号进行单侧抑制，修正数据分布。

4.2.3.2 特征提取网络结构

本章的特征提取网络结构如图 4-6 所示。网络主要由一些多角度卷积层（即图中的 ORConv）组成，批归一化层和 ReLU 激活层未在图中画出。卷积层滤波器替换为 4.2.2 中所述的可变角度滤波器即为多角度卷积层。网络中一些重复结构在图中用虚线框标出，虚线框右上数据为该结构重复次数。

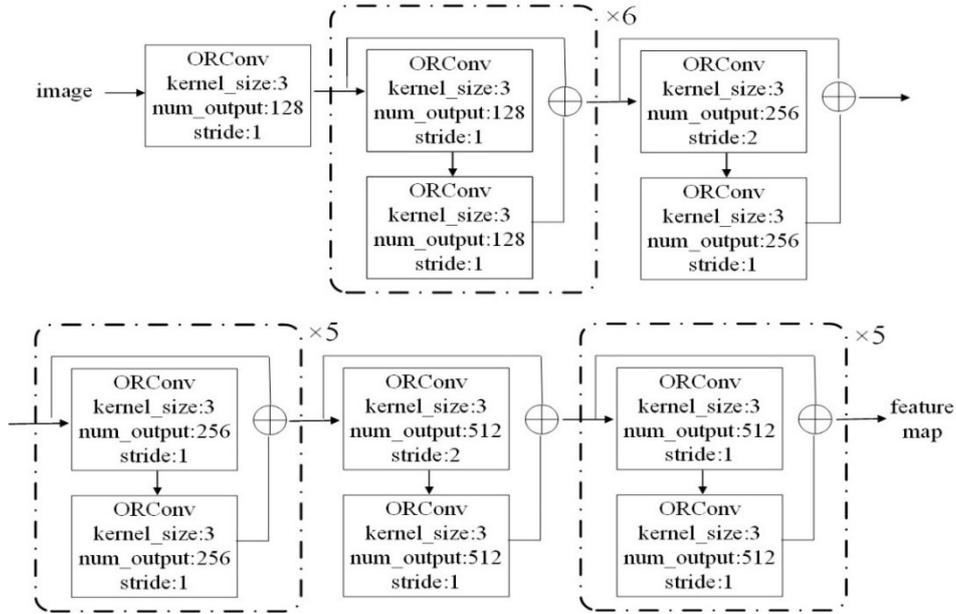


图 4-6 特征提取网络结构示意图

实验中发现使用数量较多的小滤波器在参数效率上优于使用大滤波器。因此本文使用的滤波器核都是 3×3 的，图中显示的卷积核个数中包括了 8 个方向的卷积核。另外，由于当卷积层有较高的空间分辨率时，网络的定位能力较高，因此我们使用较小的步长，减少特征图的分辨率损失。

4.2.4 CAM 定位重要区域

CAM 是一种弱监督定位算法。其利用类似于 NIN、GoogLeNet 的网络结构，对卷积特征图进行全局均值池化，将得到的特征输入给全连接分类层，得到期望的分类结果。在没有目标框标记的情况下，使用图像级别标记来训练卷积神经网络。并在测试过程中将分类权重映射到卷积特征图中，鉴别图像区域的重要性。

算法在卷积层后加入全局均值池化层，通过误差反传使网络找到所有具有判别力的目标区域。由于本课题中训练集图片为舰首切片，每个训练样本图像中

只有一艘舰船，因此我们将 CAM 算法中的全局均值池化层替换为全局最大池化层，训练时使网络只寻找舰首切片中最具判别力的区域，从而最大程度上抑制测试过程中图像非目标区域的响应。

本文由卷积特征图生成能表征区域重要性的激活响应图的具体方式如下：

对于卷积后得到的特征图，设 $f_k(x, y)$ 为卷积特征图第 k 个通道空间位置为 (x, y) 的特征值。训练过程中，对于第 k 个通道，全局最大池化层处理结果为 $F_k = \max f_k(x, y)$ 。因此对于类别 c ，分类层的输入为 $S_c = \sum_k \omega_k^c F_k$ ，其中 ω_k^c 为第 k 个特征图对于分类 c 的权重。 ω_k^c 说明了 F_k 对分类 c 的重要性，其值通过图像类别标号训练网络更新。

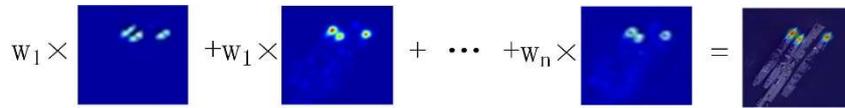


图 4-7 特征图加权求和得到激活响应图

测试过程中去掉全局最大池化层和分类层，将分类权重投影到特征层。每一个卷积特征图乘以对应的分类权重，相加得到舰船的激活响应图，如图 4-7 所示。定义 M_c 为分类 c 的类别激活响应图，每个空间位置的值由下式得到：

$$M_c(x, y) = \sum_k \omega_k^c f_k(x, y) \quad (4-6)$$

由此可得 $S_c = \sum_{x,y} M_c(x, y)$ ，因此 $M_c(x, y)$ 直接反映了图像上的点 (x, y) 对于分类 c 的重要性。

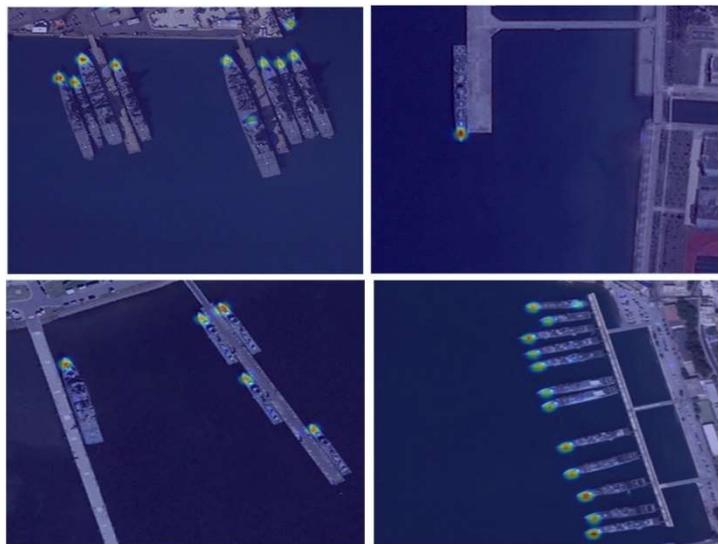


图 4-8 CAM 定位重要区域

直观地来说，我们希望感受野中的舰船能够激活每个特征图。 f_k 是卷积特征图，激活响应图简单地来说就是将不同通道特征图的线性加权求和。将激活响应图上采样到原图大小，就能得到对舰船有意义的图像区域。将响应图与原图加权叠加得到如图4-8所示的热图结果，图中高亮部分即为重要区域。

4.2.5 求连通域并估计参数得到检测框

上一步中找到了舰船所在区域，但响应集中在舰首区域，且没有像传统检测算法那样给出检测结果框。下面计算响应图中的连通域，并进行舰船参数估计求得最终得到紧致的检测框，处理流程如下：

表 4-1 检测框计算流程

输入：原图及 4.2.4 中得到的响应图
<ol style="list-style-type: none"> 1. 设定阈值，借助地理等高线思想在图中画出等响应线，得到连通区域，排除面积很小的区域。 2. 腐蚀操作断开靠的过近的响应区域，计算连通区域质心，得到以该质心为中心点的 32×32 大小（根据感受野制定）的图像块。 3. 在 4.2.3 节提取的多方向特征图中找到图像块对应区域，送入预先训练好的舰船参数估计网络模块，该模块由一个全局均值池化层和一个全连接层组成，全连接层输出为 $\sin \theta$、$\cos \theta$、长、宽四个参数。 4. 由 $\sin \theta$、$\cos \theta$ 推算 θ，根据质心位置和预测的角度、长、宽信息给出最终检测框。 5. 对检测结果框进行非极大值抑制。
输出：带检测框的检测结果图

本文利用地理等高线思想求连通域，即通过求响应灰度图中灰度相等的点连接而成的闭合曲线。通过实验选取一个阈值，对灰度图中像素值高于阈值的点进行八邻域比较，画出图中等灰度曲线，进而得到连通区域，结果如图4-9(c)所示。

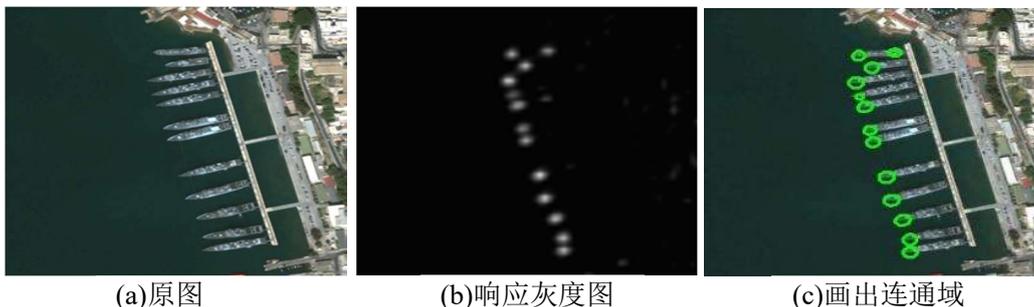


图 4-9 等高线法求连通域

利用训练样本切片的角度和舰船尺寸真值信息来训练表 4-1 中第四个步骤

所述的参数预测网络模块，测试过程中用得到的连通区域图像块来估计完整舰船的角度和尺寸信息，再根据连通区域质心位置画出舰船的紧致包围框，如下图所示：

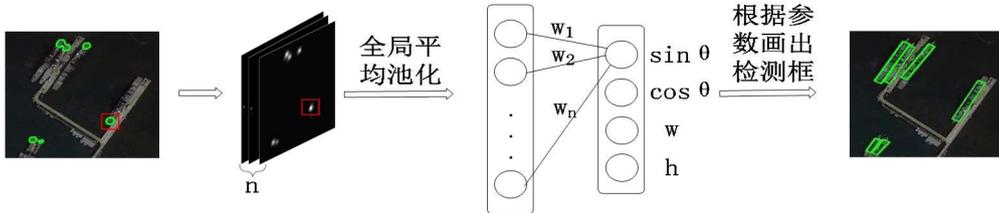


图 4-10 估计参数画出检测框

4.2.6 难反例挖掘及正例扩充

训练网络时，一般采用在训练集图像中随机选取与真值框覆盖面积小于一定值的图像块构成初始负样本集，这会导致训练集样本缺乏难例。舰船图像中某些易与舰船混淆的物体并未参与训练，如果加入这些样本可以提升检测准确性。

本文加入难反例挖掘步骤。先利用初始反例切片训练网络，然后将部分训练集图像送入网络得到响应图，收集虚警区域图像块形成负样本难例集。用此难例集继续训练网络，此过程可重复多次。实验得到的难反例切片如图 4-11 所示。



图 4-11 难反例示例

正例样本的数量和多样性在很大程度上决定了网络性能。我们希望利用类似于难反例挖掘的过程对正例样本集进行扩充。利用初始模型得到响应图与真值进行对比，找到响应值过小的舰船切片，加入到正例样本中，再次训练网络。

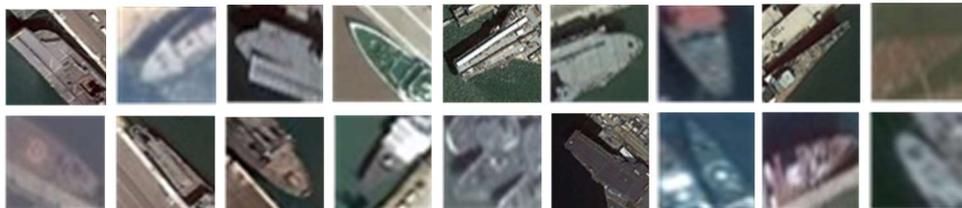


图 4-12 初始模型未检测到的正例

4.3 实验结果与分析

本节将对本文所用数据集进行介绍，并给出本文算法在该数据集上得到的实验结果并对结果进行分析。

4.3.1 数据集介绍

由于目前没有可用的公开舰船数据集，因此本文制作了一个光学遥感舰船数据集。数据来源于 Google Earth，利用 GEtScreen 截图工具得到遥感图像。舰船数据集中包括诺福克、圣迭戈、横须贺等 52 个军港以及著名民用港口在视角高度为 1km、1.2km、1.5km 下共 667 幅图像。数据集共包含舰船样本 7,978 个，舰船类型涵盖了航母、潜艇、驱逐舰、巡洋舰、护卫舰、战列舰、民用船只等。

表 4-2 数据集信息统计

数据集	舰船图像	舰船样本	反例图像
军用舰船数据集	565	6749	513
民用船舶数据集	102	1229	-
总计	667	7978	513

图像获取完成后，编写了标注工具对航拍图像中的舰船目标进行手工标定，标定结果为 14 字段格式: $x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, \theta, \text{class}$ 。

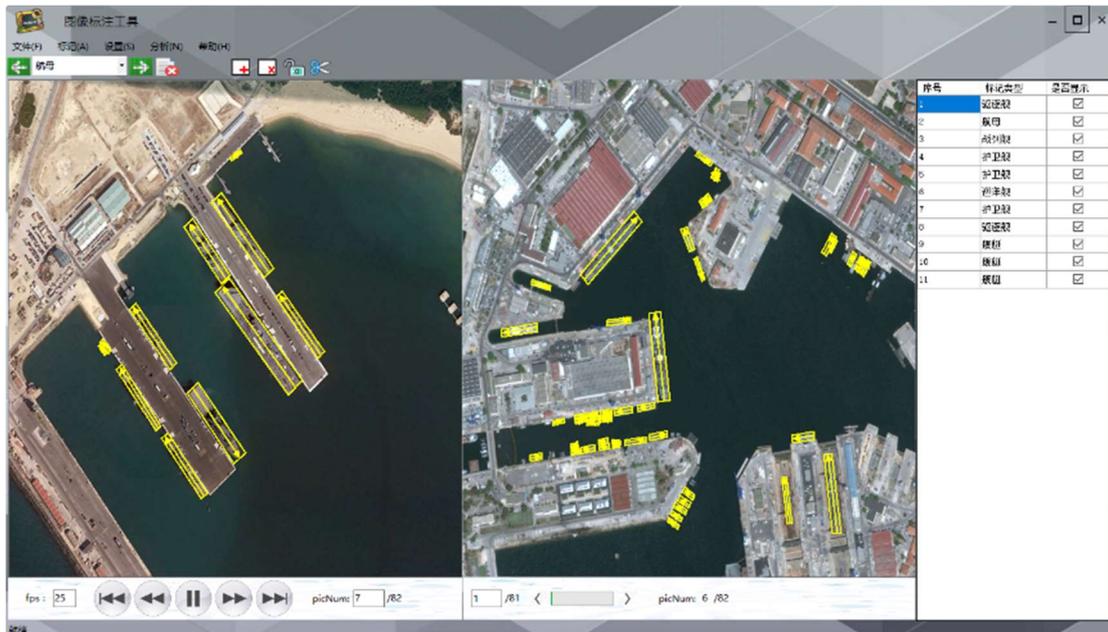


图 4-13 数据标注工具

4.3.2 检测结果与分析

本文在 4.3.1 所述的舰船数据集上进行实验，得到检测结果。选用 DPM、Faster R-CNN 检测算法作为对比方法，对比实验时采用规模一样的训练集和测试集，包括 475 幅训练图像和 90 幅测试图像。

采用检测准确率(Precision)和召回率(Recall)来综合评判检测算法的性能。首先介绍准确率和召回率的计算方式。得到检测结果后，对样本的正反例评判分为四种情况：

- (1) True Positive(TP):正确的正例，实际为正例，分类判定正确的样本。
- (2) False Positive(FP):错误的正例，实际为反例，分类判定错误的样本。
- (3) True Negative(TN):正确的反例，实际为反例，分类判定正确的样本。
- (4) False Negative(FN):错误的反例，实际为正例，分类判定错误的样本。

通过对上述四种情况的检测样本进行统计，可计算准确率和召回率，其表达式如下：

$$Recall = TP / (TP + FN) \quad (4-7)$$

$$Precision = TP / (TP + FP) \quad (4-8)$$

本文算法训练时每个样本图像只有一个目标，在卷积后对特征图进行求全局最大值操作，经过分类层可得到分类得分。然而测试图像中包含多个同类目标，对整幅图像对应特征图求最大值后，只能给出响应最大区域的分类得分，不能得到每个检测框的得分。在缺乏得分信息的前提下，显然不能像其他检测算法一样通过设置得分阈值来计算平均精度(mean Average Precision, mAP)。为了与其他算法进行比较，我们采用本文算法得到结果后，再对 DPM、Faster R-CNN 算法分别设置得分阈值，使得算法在尽可能查全的情况下虚警尽量少，得到检测结果，对三种算法的召回率和准确率进行比较。

表 4-3 检测指标对比

检测方法	DPM	Faster R-CNN	本文算法
召回率	0.673	0.805	0.867
准确率	0.296	0.312	0.668

由表 4-3 可以看出，本文算法得到的检测结果优于 DPM 算法和 Faster R-CNN 算法。另外在利用 DPM、Faster R-CNN 这两种方法进行检测的过程中，不

断降低得分阈值，召回率(也称查全率)仍然无法达到本文算法得到的召回率 0.867。

图 4-14 给出了本文算法和 Faster R-CNN 算法对同一幅测试图像的检测结果。可以看出 Faster R-CNN 算法对于相隔较近的并排舰船只得到一个检测框，出现漏检。另外对于小型舰艇的检测效果，本文算法也优于 Faster R-CNN 算法。

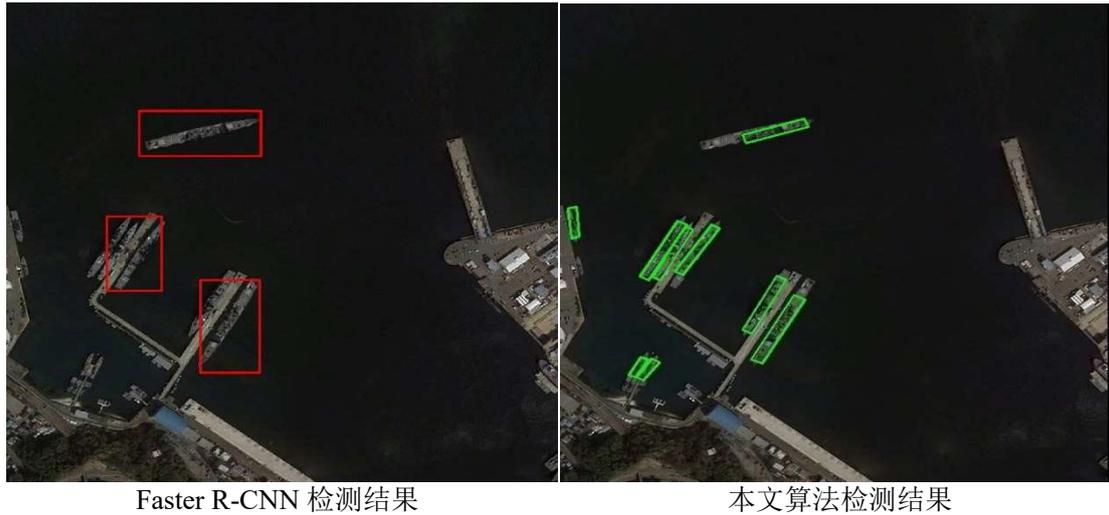
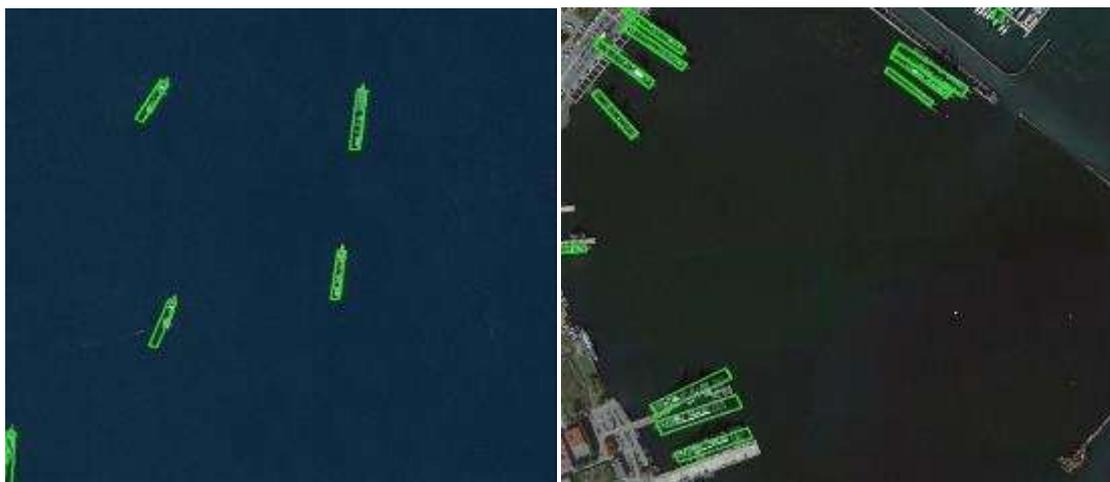


图 4-14 检测结果对比

图 4-15 给出了本文算法检测结果示例，本文算法在多角度、并排舰船的检测任务中，取得了优异的成绩。尽管存在一定的漏检，但这种漏检大多是由该类型舰船在训练集中不存在或数量极少导致的，可通过扩充数据集来改善。



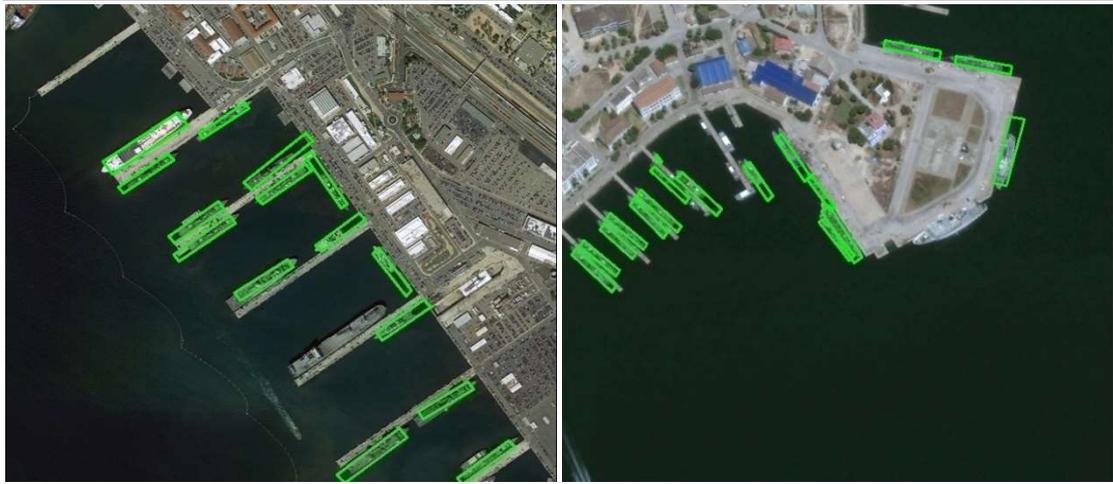


图 4-15 本文算法检测结果示例

4.4 本章小结

本章对基于深度学习的鲁棒性舰船检测方法的设计和实现进行了详细的论述。主要对旋转滤波器实现旋转不变性、全卷积网络提取特征、利用弱监督定位算法 CAM 找到图像中关键区域、求连通域并估计舰船参数得到检测框、难反例挖掘及正例扩充等环节进行具体介绍。通过对以上步骤的设计与分析，实现了基于深度学习的鲁棒性舰船检测。同时利用本文构建的舰船数据集上对本章算法进行了实验验证，证实了本文框架的有效性和准确性。

第5章 结论与展望

舰船自动目标检测在监视军舰、保护领海安全及疏导海上交通、救援遇险船只等方面有着广阔的应用前景。由于港口背景较为复杂，以及舰船本身多姿态多角度、并排停放等特点，使得光学遥感图像舰船检测成为一个极具挑战性的任务。已有的检测算法准确率不高，难以应用于实际，因此，如何制定适用于舰船的检测策略、提高实际应用中检测效率，是一个重要的研究方向。

本文针对舰船目标的特异性，提出了一种基于深度学习的鲁棒性舰船检测方法，同时，改善了神经网络对角度变化敏感的缺陷问题，并提出了相应方法极大程度上抵消复杂背景、舰船并排停放引发的性能下降。

第一章论述了研究背景和研究意义，介绍了国内外研究现状以及现有检测方法中存在的难点和问题。第二章主要介绍了舰船检测每个处理环节的相关技术，对舰船检测衍生出的海陆分割问题和多角度目标检测问题的已有研究成果进行了总结。第三章详细阐述了本文提出的基于图割的海陆分割算法。不同于以往基于像素判别的海陆分割方法，本文首先利用基于图的分割算法进行区域划分，提取区域特征，利用随机森林分类器进行区域鉴别，最后通过一些形态学方法进行水域扩展，得到水陆分割二值结果图像。加入水陆分割排除岸上区域后，检测过程中的虚警率大大降低。第四章是本文的核心章节，具体讲述了舰船检测中的每一个步骤所用的方法。为了使网络适应目标角度变化，通过可变角度滤波器实现特征的旋转不变性。由于传统的基于候选框的检测算法容易出现并排舰船漏检，我们提出了一种新的检测思路：全卷积网提取特征，利用 CAM 算法将分类权重映射到卷积特征图中，依赖卷积层的定位能力直接鉴别有识别力的图像区域，最后通过一些后处理将热图结果转换为检测任务所需的检测结果框。

综上所述，本文的贡献主要在于：

(1) 针对复杂港口场景下的岸上人工目标虚警问题，提出了一种基于图割的海陆分割算法，对陆地部分进行遮挡，减少疑似目标漏检。

(2) 对于舰船多角度变化问题，采用主动旋转滤波器的方式得到旋转不变深度特征，使学习过程中得到的滤波器更“纯净”，从而提高检测精度。

(3) 本文检测算法可解决传统目标检测方法无法解决的并排舰船漏检问题。本文检测算法不提取候选框，而是直接找到有识别力的重要区域，有效避免了并

排舰船候选框重叠问题造成的漏检。

本文在光学遥感图像舰船目标检测方面取得了一些阶段性成果。但由于时间有限，仍然存在一些不足之处，有待于进一步探索和研究。

(1) 本文的海陆分割算法不是基于深度学习的，采用一些深度框架下的分割算法可能会获得更准确的分割效果。文献^[49]提出利用全卷积网进行语义分割，此方法可应用于海陆分割中，如果实验性能较好，还可直接利用分割的方法找到舰船。由于此算法需要大量区域标注，本课题还未对此方法进行探索。

(2) V型舰首是舰船最容易与其他物体区分开的部分，为了更好地鉴别舰船，本文检测算法使用的训练样本为舰首切片，这导致我们并没有最大程度地利用舰船样本。尽管实验中发现舰中、舰尾特征区分度不高，但仍可考虑利用舰中、舰尾或舰船整体信息来辅助进行舰船检测。

参 考 文 献

- [1] Liu Z, Wang H, Weng L, et al. Ship Rotated Bounding Box Space for Ship Extraction From High-Resolution Optical Satellite Images With Complex Backgrounds[J]. IEEE Geoscience & Remote Sensing Letters, 2016, 13(8):1074-1078.
- [2] Zou Z, Shi Z. Ship Detection in Spaceborne Optical Image With SVD Networks[J]. IEEE Transactions on Geoscience & Remote Sensing, 2016, 54(10):5832-5845.
- [3] Yang G, Li B, Ji S, et al. Ship Detection From Optical Satellite Images Based on Sea Surface Analysis[J]. Geoscience & Remote Sensing Letters IEEE, 2014, 11(3):641-645.
- [4] Zhu C, Zhou H, Wang R, et al. A Novel Hierarchical Method of Ship Detection from Spaceborne Optical Image Based on Shape and Texture Features[J]. IEEE Transactions on Geoscience & Remote Sensing, 2010, 48(9):3446-3456.
- [5] Tang J, Deng C, Huang G B, et al. Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine[J]. IEEE Transactions on Geoscience & Remote Sensing, 2015, 53(3):1174-1185.
- [6] Guo C, Ma Q, Zhang L. Spatio-temporal Saliency detection using phase spectrum of quaternion fourier transform[C]// Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008:1-8.
- [7] Corbane C, Najman L, Pecoul E, et al. A complete processing chain for ship detection using optical satellite imagery[J]. International Journal of Remote Sensing, 2010, 31(22):5837-5854.
- [8] Buch H, Sharghi E, Guilas C, et al. Enhanced ship detection from overhead imagery[C]// Proceedings of SPIE, the International Society for Optical Engineering. Orlando, FL, USA: SPIE, 2008:1-12.
- [9] 胡俊华, 徐守时, 陈海林, 等. 基于局部自相似性的遥感图像港口舰船检测[J]. 中国图象图形学报, 2009, 14(4):591-597.
- [10] Wang Q J, Gao X, Chen D Q. Pattern Recognition for Ship Based on Bayesian Networks[C]// International Conference on Fuzzy Systems and Knowledge Discovery. IEEE, 2007:684-688.
- [11] Zhang Feng-Li, Zhang Lei, Wu Bing-Fang. Progress of ship detection technology and system based on remote sensing technology in European union. Journal of Remote Sensing, 2007,

11(4):552-562

[12] Chen P H, Lin C J, Scholkopf B. A tutorial on v-support vector machines[J]. Applied Stochastic Models in Business & Industry, 2005, 21(2):111-136.

[13] Forsyth D. Object Detection with Discriminatively Trained Part-Based Models[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2010, 32(9):1627-1645.

[14] 曾蕾, 王晓蕊, 郭冰涛等. 一种基于航天图像的舰船目标检测和船速估计方法 [J]. 光电工程, 2013, 40(12): 65-72.

[15] Yang F, Xu Q, Gao F, et al. Ship detection from optical satellite images based on visual search mechanism[C]// Geoscience and Remote Sensing Symposium. IEEE, 2015:3679-3682.

[16] Proia N, Page V. Characterization of a Bayesian Ship Detection Method in Optical Satellite Images[J]. Geoscience & Remote Sensing Letters IEEE, 2010, 7(2):226-230.

[17] He Si-Hua, Yang Shao-Qing, Shi Ai-Guo, Li Tian-Wei. Detection of ship target under sea background based on texture high-order fractal feature. Optics and Optoelectronic Technology, 2008, 6(4): 79-82

[18] Yang G, Li B, Ji S, et al. Ship Detection From Optical Satellite Images Based on Sea Surface Analysis[J]. Geoscience & Remote Sensing Letters IEEE, 2014, 11(3):641-645.

[19] WANG Min LUO Jiancheng MING Dongping. Extract Ship Targets from High Spatial Resolution Remote Sensed Imagery with Shape Feature[J]. Wuhan Daxue Xuebao, 2005, 30(8):685-688.

[20] Uijlings J R R, Sande K E A V D, Gevers T, et al. Selective Search for Object Recognition[J]. International Journal of Computer Vision, 2013, 104(2):154-171.

[21] Zitnick C L, Dollár P. Edge Boxes: Locating Object Proposals from Edges[M]// Computer Vision – ECCV 2014. Springer International Publishing, 2014:391-405.

[22] Jin B, Cong Y, Zhou W, et al. A new method for detection of ship docked in harbor in high resolution remote sensing image[C]// International Conference on Progress in Informatics and Computing. IEEE, 2014:341-344.

[23] Long Gang, Chen Xue-Quan. A method for automatic detection of ships in harbor area in high resolution remote sensing image. Computer Simulation, 2007, 24(5):198-201

[24] Lowe D G. Object Recognition from Local Scale-Invariant Features[C]// The Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, 1999:1150.

- [25] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]// Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE, 2005:886-893.
- [26] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
- [27] Gimpel K, Smith N A. Softmax-margin CRFs: training log-linear models with cost functions[C]// Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 2-4, 2010, Los Angeles, California, USA. DBLP, 2010:733-736.
- [28] Girshick R. Fast R-CNN[C]// IEEE International Conference on Computer Vision. IEEE, 2015:1440-1448.
- [29] Zhang Feng-Li, Zhang Lei, Wu Bing-Fang. Progress of ship detection technology and system based on remote sensing technology in European union. Journal of Remote Sensing, 2007, 11(4):552-562
- [30] You X. A New Method for Ship Detection Based on Feature Fusion in Optical Image[J]. Computer Engineering & Applications, 2005, 41(19):199-202.
- [31] 陈健, 郑绍华, 余轮,等. 基于方向的多阈值自适应中值滤波改进算法[J]. 电子测量与仪器学报, 2013, 27(2):156-161.
- [32] Liu G, Zhang Y, Zheng X, et al. A New Method on Inshore Ship Detection in High-Resolution Satellite Images Using Shape and Context Information[J]. IEEE Geoscience & Remote Sensing Letters, 2014, 11(3):617-621.
- [33] He S, Lau R W H. Oriented Object Proposals[C]// IEEE International Conference on Computer Vision. IEEE, 2015:280-288.
- [34] Haley G M, Manjunath B S. Rotation-invariant texture classification using modified Gabor filters[C]// International Conference on Image Processing, 1995. Proceedings. IEEE, 1995:262-265 vol.1.
- [35] Laptev D, Savinov N, Buhmann J M, et al. TI-POOLING: Transformation-Invariant Pooling for Feature Learning in Convolutional Neural Networks[J]. 2016:289-297.
- [36] Bruna J, Mallat S. Invariant Scattering Convolution Networks[J]. IEEE Transactions on Pattern

Analysis & Machine Intelligence, 2013, 35(8):1872-86.

[37] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu. Spatial transformer networks. In NIPS, 2015.

[38] Felzenszwalb P F, Huttenlocher D P. Efficient Graph-Based Image Segmentation[J]. International Journal of Computer Vision, 2004, 59(2):167-181.

[39] Mount D M, Arya S. ANN: A Library for Approximate Nearest Neighbor Searching[J]. in The 14th annual ACM-SIAM Symposium on Discrete Algorithms, 1998:783-792.

[40] D. S. Y. Kartika and D. Herumurti, "Koi fish classification based on HSV color space," 2016 International Conference on Information & Communication Technology and Systems (ICTS), Surabaya, Indonesia, 2016, pp. 96-100.

[41] 曹建芳, 陈俊杰, 赵青杉. 一种改进的 HSV 颜色空间量化方法及其应用[J]. 南京师范大学学报(工程技术版), 2014(2):68-73.

[42] Zulpe N S, Pawar V P. GLCM textural features for Brain Tumor Classification[J]. International Journal of Computer Science Issues, 2012, 9(3): 354–359.

[43] M. Pal. Random forest classifier for remote sensing classification[J]. International Journal of Remote Sensing, 2005, 26(1):217-222.

[44] Zhou Y, Ye Q, Qiu Q, et al. Oriented Response Networks[C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017.

[45] Zhou B, Khosla A, Lapedriza A, et al. Learning Deep Features for Discriminative Localization [C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 2921-2929.

[46] Liu Z, Luo P, Wang X, et al. Deep Learning Face Attributes in the Wild[C] // IEEE International Conference on Computer Vision. IEEE, 2014:3730-3738.

[47] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015:1-9.

[48] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015:770-778.

[49] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015:3431-3440.

致 谢

雁栖之滨，又见繁花满枝。今硕士即就，屡屡回望，求学三载，不过忽然而已。韶华已逝，学业已竟，获益良多，昔日乘兴而来，今日也可尽兴而归矣。

值此毕业到来之际，对所有帮助过我支持过我的老师、同学和家人，献上我最真挚的谢意。

首先，感谢焦建彬教授带我走入实验室，给予我珍贵的学习机会。感谢导师叶齐祥教授的悉心指导，帮助我顺利完成研究课题。感谢实验室老师们对我的指导和关怀，感谢他们提供的优越的科研条件，让我汲取新知，充实自己，更快地学习和成长。

同样要感谢与我并肩作战的师兄姐妹们。特别感谢魏鹏旭师姐对我的帮助和鼓励，对于我的研究课题，魏鹏旭师姐提出了很多指点和建议。当我在科研道路上迷茫不知所措时，指导引领我走出误区，树立信心；另外要感谢邓维建师弟，在和他讨论的过程中学习到很多专业知识；感谢周彦钊师弟对我的研究课题提供的帮助和技术支持。感谢实验室每一位成员，是你们朝夕相处地陪伴，让我的硕士生涯平添一抹色彩。愿你们每一个人都有一个灿烂前程，愿你们永远开心幸福。

感谢我的父母，你们的坚定支持、鼓励和无言奉献，是引领我一直走下去的最大动力，愿健康、平安、快乐一直与你们相伴。

最后感谢参与我开题、中期、结题评审和答辩的每一位专家组老师，你们以丰富的经验帮助我把握论文方向和研究进度，最终顺利完成论文。

刘嫣然

2017年4月

个人简历、在学期间发表的论文与研究成果

姓名：刘嫣然 性别：女 出生日期：1992年4月30日

教育经历：

- 2010年9月至2014年7月 北京工业大学 通信工程 学士
- 2014年9月至2017年7月 中国科学院大学 计算机技术 硕士

研究成果：

- 《基于深度学习的鲁棒性舰船目标检测》专利（已申请）。