

基于深度卷积神经网络的红外船只目标检测方法

王文秀^{1,2,3}, 傅雨田^{1,2}, 董峰^{1,2*}, 李锋^{1,2,3}

¹中国科学院红外探测与成像技术重点实验室, 上海 200083;

²中国科学院上海技术物理研究所, 上海 200083;

³中国科学院大学, 北京 100049

摘要 针对红外船只图像较模糊导致的识别率低、识别速度慢等问题, 提出了一种基于深度卷积神经网络(CNN)的检测算法。首先采用标记分水岭分割算法提取红外船只图像中的连通区域, 并对原图相应的目标位置进行标记和归一化处理, 提取候选区域。采用改进的 AlexNet(一种深度 CNN 模型)进行船只目标识别, 将提取的候选区域送入改进的 AlexNet 进行特征提取和预测, 得到最终检测结果。分水岭方法可大大减少候选区域检测时间, 以及减少深度 CNN 识别时间。利用实验室自制的红外成像系统获取近千张红外船只图像数据, 并对其平移缩放形成的数据集进行仿真实验。结果表明, 标记分水岭与深度 CNN 的结合, 可有效识别船只目标, 所提方法具有良好的性能, 能够更加快速准确地识别红外船只目标。

关键词 测量; 红外船只目标检测; 标记分水岭; 卷积神经网络; 图像分割

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201838.0712006

Infrared Ship Target Detection Method Based on Deep Convolution Neural Network

Wang Wenxiu^{1,2,3}, Fu Yutian^{1,2}, Dong Feng^{1,2}, Li Feng^{1,2,3}

¹Key Laboratory of Infrared System Detection and Imaging Technology, Chinese Academy of Sciences, Shanghai 200083, China;

²Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

³University of Chinese Academy of Sciences, Beijing 100049, China

Abstract Aiming at the problems of low recognition accuracy and slow recognition speed due to the fuzzy image of infrared ship targets, a classification algorithm based on deep convolution neural network (CNN) is proposed. By using the marker-controlled watershed segmentation algorithm, the connected regions in infrared ship image are extracted and the corresponding target positions of the original image are marked and normalized to extract the candidate regions. The improved AlexNet (a deep CNN model) is used for ship targets identification. The extracted candidate regions are sent to the improved AlexNet for feature extraction and prediction to obtain the final detection result. The marker-controlled watershed segmentation method can greatly reduce the number of candidate regions and reduce the classification time of deep CNN. The data of nearly one thousand infrared ship images are obtained by the laboratory-made infrared imaging system, and the simulation experiment on the dataset formed by its translation and scaling is performed. The simulation results show that the combination of the marker-controlled watershed segmentation algorithm and the deep CNN can effectively identify the ship targets. The proposed method has good performance and can identify infrared ship targets more quickly and accurately.

Key words measurement; infrared ship target detection; marker-controlled watershed; convolution neural

收稿日期: 2018-02-06; 收到修改稿日期: 2018-03-09

基金项目: 天基视频探测技术(2015AAxxx5097)

作者简介: 王文秀(1991-), 女, 博士研究生, 主要从事红外成像、图像处理分析等方面的研究。

E-mail: wenxiu@mail.ustc.edu.cn

导师简介: 傅雨田(1967-), 男, 博士, 研究员, 博士生导师, 主要从事红外成像遥感技术等方面的研究。

E-mail: yutianfu@mail.sitp.ac.cn

* 通信联系人。E-mail: dongfeng@mail.sitp.ac.cn

network; image segmentation

OCIS codes 120.0280; 150.0155; 110.2970; 100.3008

1 引 言

船只检测在军事和民用领域都有着重要的应用。为了满足船舶各种碰撞灾难的观测、入侵者检测以及海上事故搜救等需求,快速、精准地检测不同海面环境中的船只目标变得越来越重要^[1]。目前,最为常见的船只目标识别方法主要集中于可见光和合成孔径雷达(SAR)方面,本文尝试在长波红外船只目标检测方面进行探索。红外船只目标的检测与识别为重点海域和重要港口进行监控提供重要信息,可全天候提高海防预警以及海运监测管理、调度的能力,同时,可以监督渔业及管理海洋运输,以确保海岸和海洋的安全。计算机视觉算法为船只的监管提供了强有力的技术手段。

船只目标的检测方法常分为两个步骤:候选区域提取和识别。候选区域的选取常采用图像分割的方式实现。图像分割是指将图像分成若干个小区域,优质的图像分割方法可有效分离图像中目标区域与背景。较少的合理候选区域的提取,可大大增加算法的速率。目前图像分割技术主要包括基于区域、基于边缘和基于边界的分割。本研究中在桥上拍摄的船只目标,船只结构较为清晰,传统的分割方法很容易出现漏分割现象。Sun L 等^[2]对靠岸舰船目标进行了识别,采用船头的角点特征和船身边界检测进行识别,得到较好的识别效果,本研究中船只目标较大,船只再分割时分块问题仍较为严重,算法适用性不佳。随着机器学习、深度学习等技术的快速发展,卷积神经网络(CNN)在图像分类、目标识别等领域都取得出色的成绩。Ren S Q 等^[3]提出的基于区域的 CNN(R-CNN),采用 Selective Search 提取候选区域,并统一大小送入 CNN。Selective Search 图像分割算法提取 2000 个候选区域,在本研究的应用场景中,视场内的红外船只目标数量偏少,运动的船只目标有一定的间距,因此,红外船只目标的检测不需要过多的候选区域。Zhang R Q 等^[4]采用 S-CNN 将船头“V”形结构和船身的“||”形结构结合,在遥感船只目标识别中取得较好的效果。但在本研究的场景中,船头“V”形及船身“||”

形状并不明显。

分水岭分割属于区域分割领域,其保留了传统分割方法的多种优点。分水岭能够准确地定位边缘,算法是自动的,不需要设定参数用于终止^[5],其分割思想也一直是广大学者研究的热点。相比于传统的算法,分水岭算法能够准确获取前景物体边缘信息,为后期目标识别提供较好的分割结果。分水岭算法可得到简单、闭合的完整分割线,易于并行化处理^[6]。

参考以上思路,提出一种基于标记分水岭和深度 CNN 结合的红外船只目标识别方法,基于标记分水岭的候选区域提取和深度 CNN 的网络实现船只目标识别。采用图像处理单元(GPU)对训练和测试过程加速,实现较好的效果。

2 标记分水岭候选区域提取

分水岭方法把图像看成一幅“地形图”,通过寻找集水盆地和分水岭对图像进行分割:首先,在各个局部极小值表面刺穿一个小孔,紧接着,将模型浸入水中,每个局部极小值的影响域会随着浸入的加深不断向外扩散,在两个集水盆地汇合处构筑大坝,进而形成分水岭^[5]。在图像处理过程中,首先将原始图像转为灰度图,其中的每个局部极小值及其影响的区域被称作集水盆地,分水岭则由集水盆地的边缘形成。分水岭可以较好地提取图像的微弱边缘,得到封闭连续边缘,其中封闭的集水盆地可用于更好地分析图像的区域特征。

直接应用分水岭分割算法实现效果不佳,采用滤波的方法也不能完全去除纹理细节和干扰噪声,梯度图像中也会存在不稳定的区域最小值,直接采用标准分水岭变换会出现过分割问题^[6]。为了更好地实现图像区域的分割,本研究拟采用标记提取方法,对梯度图像中感兴趣区域的最小值进行标记,屏蔽去除其他最小值,进而分割出感兴趣区域。实现过程分为闭重建、求差、开重建和二值化 4 个步骤,具体如图 1 所示。

由图 1 可知,为了得到图像边缘信息,通常把梯度图像作为输入图像^[7],梯度图像为:

$$g(x, y) = \text{grad}[f(x, y)] = \sqrt{[f(x, y) - f(x - 1, y)]^2 + [f(x, y) - f(x, y - 1)]^2}, \quad (1)$$

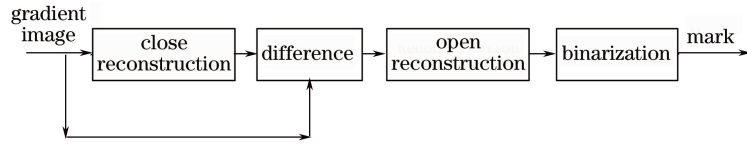


图 1 标记分水岭分割图像获取步骤

Fig. 1 Acquisition step marker-controlled watershed segmentation image

式中 $f(x, y)$ 表示原始图像, $\text{grad}[f(x, y)]$ 表示梯度运算。具体实现如图 2 所示。

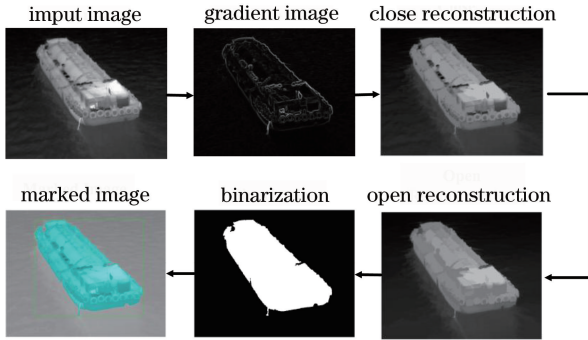


图 2 具体实现

Fig. 2 Realization in detail

图 2 的具体说明如下:

闭重建: 填补低灰度区域中完全被结构元素包含的部分。相比于原始梯度图像, 闭重建后的区域最小值点数目减少, 两者之差是被消除的区域。把比结构元素小的缺口或孔填充上, 搭接短的间隔, 起到连接作用^[8]。

开重建: 去除比结构元素小的特定图像细节, 同时保证不产生全局几何失真; 滤除比结构元素小的突刺, 切断细长搭接, 实现分离效果。

开重建及其相应的闭重建定义如下:

$$\gamma = \rho_f[\varphi(f)], \quad (2)$$

$$\varphi = \rho_f[\psi(f)], \quad (3)$$

式中 $\varphi(f)$ 、 $\psi(f)$ 分别表示形态学中的开运算和闭运算, f 表示掩模图像, ρ 为获取梯度图像的标记。

二值化过程采用大津 (OTSU) 算法 (最大类间差法) 将标记图像转化为二值图像。OTSU 算法是自动阈值选取中最优的方法之一^[9]。图像中, 方差是灰度分布均匀性的一种度量, 方差值越大, 构成图像的两部分差别越大。鉴于此, OTSU 算法对灰度图像的直方图进行分析, 利用目标与背景间的方差最大值确定分割的阈值, 将错分概率降到最小。

3 深度 CNN 船只检测

3.1 CNN 简介

CNN 是神经学影响深度学习的典例, 是当今图

像理解领域的研究热点。近年来, 随着计算机性能的提升以及神经网络的不断发展, 诸多基于 CNN 的方法都取得了较好的效果。常见的有 AlexNet、VGG-Net 和 ResNet 等。CNN 是专用于处理具有类似结构数据的神经网络, 可直接将多维图像用于网络输入, 更高效地提取特征, 在图像处理及机器学习领域得到广泛的应用。神经网络输出层公式为:

$$X_j^l = f\left(\sum_{i=1}^d X_i^{l-1} \omega_{ij}^l + b_j^l\right), \quad (4)$$

式中 i 表示第 i 个输入单元, j 表示第 j 个隐含层, ω_{ij}^l 表示第 i 个输入层与第 j 个隐层权重, b_j^l 表示第 j 个隐层的偏置, d 为对应一层滤波器个数, $f(\cdot)$ 为激活函数, l 表示网络层。

CNN 中采用池化操作, 减少了模型中的神经元数量, 在平移不变性上保持较好的稳健性。CNN 的权值共享网络结构降低了模型复杂度, 减少了参数数量, 提升了泛化能力。在 CNN 中, 输入多维图像时, 避免了传统方法特征提取和数据重建过程^[10-11]。CNN 的基本结构包括: 输入层、卷积层、下采样层 (池化层)、全连接 (FC) 层和 softmax 分类器层, 如图 3 所示。

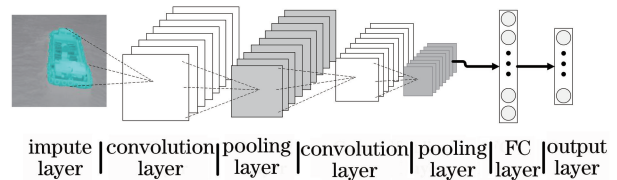


图 3 CNN 示意图

Fig. 3 CNN diagram

子采样又称为池化。池化函数使用相邻区域总体统计特征代替网络中对应位置的输出。池化对输入有一定的局部平移不变性, 对于存在少量平移的输入, 池化可实现近似不变, 可在一定程度上提高网络的统计效率^[12]。常见的池化方法有: 最大池化法、均值池化法、L2 范数池化法及加权平均池化法。CNN 输出层公式为:

$$X_j^l = f[B_j^l D(X_j^{l-1}) + b_j^l], \quad (5)$$

式中 $D(X_j^{l-1})$ 表示池化函数, B_j^l 为增益, b_j^l 为卷积层偏置。

3.2 改进的 AlexNet 模型

2012年, Girshick R 等^[13]提出深度 CNN 模型 AlexNet, 其为 LeNet 的更深更宽版本。AlexNet 将网络分成上下两个部分, 原文中两部分完全相同。AlexNet 全部使用最大池化(除了全连接层), 每层处理后得到的特征图都独立^[14]。

AlexNet 主要由 5 个 CNN 层(convolution1~convolution5)、两个全连接层(Full)和 softmax 组成。AlexNet 中的局部响应归一化(LRN)层, 创建了局部神经元间的创新机制, 响应大的值相对变得更大, 但效果不明显, 同时, LRN 会造成前馈、反馈的速度大大下降, 因此本研究不使用 LRN。同时, 随着卷积层数的加深, CNN 提取的特征越复杂, 对物体的识别性能就越好, 但细节部分体现少。而层数越浅, 细节表现越好, 但同时也会引入背景杂波。

鉴于此, 新增 Fine-Grained Features 技巧, 参考特征金字塔和 ResNet, 把高分辨率与低分辨率特征结合在一起, 从而增加对小物体的识别精度。参考

YOLO 网络结构^[15], 采用 reorg 和 route 对 AlexNet 网络进行改进, 对浅层和深层特征进行抽取, 对复杂的特征与细节特征同时进行检测。

改进的 AlexNet(图 4)结构具体实现如表 1 所示。卷积核尺寸大小根据 AlexNet 网络经验选取, Pool 层均采用 2×2 的 MaxPooling, 在降低维度、提高训练速度的同时, 确保不出现过拟合现象。船只目标占据视场较大, 在输入图像中所占区域偏大, 在不断地池化过程中容易丢失细节, 因而浅层的边缘等细节特征的加入, 能提高检测网络对船只性能的检测, 加强特征表征能力^[16]。

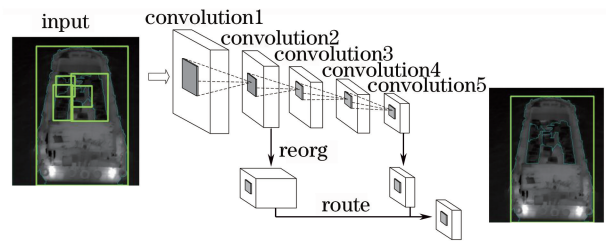


图 4 改进的 AlexNet

Fig. 4 Improved AlexNet

表 1 改进的 AlexNet 结构

Table 1 Structure of improved AlexNet

Layer	Net	Input	Convolution size	Stride	Padding	Output
1	Convolution	$224 \times 224 \times 2$	$11 \times 11 \times 96$	2	Same	$112 \times 112 \times 96$
2	Pool	$112 \times 112 \times 96$	—	2	—	$56 \times 56 \times 96$
3	Convolution	$56 \times 56 \times 96$	$5 \times 5 \times 256$	1	Same	$56 \times 56 \times 256$
4	Pool	$56 \times 56 \times 256$	—	2	—	$28 \times 28 \times 256$
5	Convolution	$28 \times 28 \times 256$	$3 \times 3 \times 384$	1	Same	$28 \times 28 \times 384$
6	Convolution	$28 \times 28 \times 384$	$3 \times 3 \times 384$	1	Same	$28 \times 28 \times 384$
7	Convolution	$28 \times 28 \times 384$	$3 \times 3 \times 256$	1	Same	$28 \times 28 \times 256$
8	Pool	$28 \times 28 \times 256$	—	2	—	$14 \times 14 \times 256$
9	Reorg	4	—	2	—	$14 \times 14 \times 1024$
10	Route	9, 8	—	—	—	$14 \times 14 \times 1280$
11	Convolution	$14 \times 14 \times 1280$	3×3	1	—	$14 \times 14 \times 1280$
12	FC	$14 \times 14 \times 1280$	—	—	—	4096
13	FC	4096	—	—	—	4096
14	FC	—	—	—	—	2

图 4 及表 1 中的 reorg 选取 stride 为 2。reorg 的实现上, 采用加上一个 Passthrough Layer 来取得之前的某个 28×28 分辨率的层的特征。这个 Passthrough layer 把 28×28 的特征图与 14×14 的特征图联系在一起, 把相邻的特征堆积在不同的 Channel 之中, 类似于 ResNet 的 Identity Mapping, 从而把 $28 \times 28 \times 256$ 变成 $14 \times 14 \times 1024$ 。实现过程如图 5 所示, 左侧表示 Reorg 的输入, 数字 1, 2, 3, 4 分别代表第一行前两个像素和第二行前两个像

素, 右侧为 reorg 结果, 输入像素分别位于第一个像素。reorg 层用于匹配特征图尺寸, route 层起连接作用, 为第 9 层和第 8 层输出的叠加。

4 实验结果与分析

4.1 红外船只数据采集

实验采用实验室自制长波红外探测器对红外船只进行数据采集, 面阵规格为 384×288 , 工作波长为 $7.7 \sim 9.5 \mu\text{m}$, 探测器像元尺寸为 $25 \mu\text{m} \times$

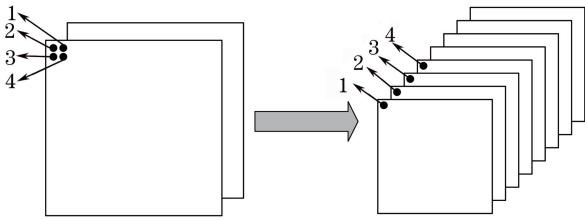


图 5 Reorg 实现过程

Fig. 5 Reorg implementation

25 μm , 光学系统的入瞳直径为 20 mm; 有效像元率 $>99\%$ 。在某重要航运枢纽附近的桥上采集船只数据近千张, 并采用平移缩放等操作实现数据集的扩充。

4.2 算法实现过程

利用 TensorFlow 神经网络学习框架, 对近千张长波红外相机拍摄的船只图像进行训练。算法实现平台为 64 位 Ubuntu16.04 LTS, 基于 i7、NVIDIA GTX1060、8G 内存, Python 3.5 OpenCV 3.0.0, 搭建 AlexNet 网络进行红外窗子目标识别的模型训练和测试。具体流程如图 6 所示。

主要过程描述如下: 拍摄图像, 准备训练样本和测试样本; 标记船只样本, 制作训练机和测试数据集; 利用训练样本训练船只模型; 对测试集进行标记分水岭分割; 用训练好的模型对分割后的区域进行识别; 选择船只目标识别的评价指标, 对识别结果进行评价。

4.3 图像分割效果评估

实验 1: 采用标记分水岭方法对红外图像数据

集进行分割效果分析, 从扩展后的数据集中随机抽取近千张红外图像数据, 其中近 50% 数据可以得到较好的分割效果, 具体分割效果如图 7 所示。

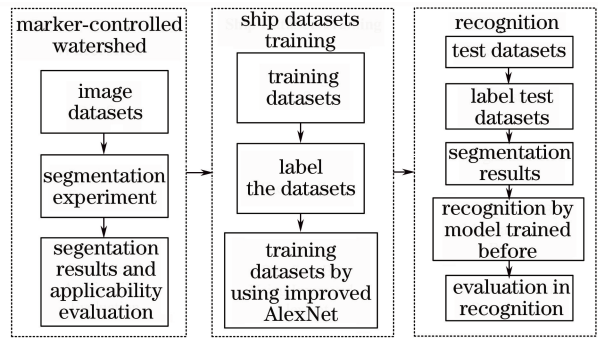


图 6 标记分水岭与改进的 AlexNet 红外船只目标实现流程图

Fig. 6 Flow chart of infrared ship target implementation with marker-controlled watershed and improved AlexNet

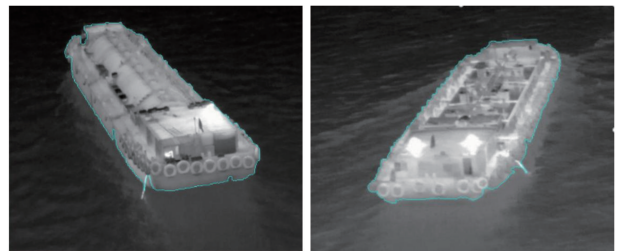


图 7 两幅分水岭分割效果图

Fig. 7 Two effect pictures by marker-controlled watershed segmentation

由图 7 可知, 标记分水岭中较好的分割效果可达到近似人工标记边界的情况。分水岭分割也存在多区域分割的情况, 具体如图 8 所示。

input image	segmentation result(one)	input image	segmentation result(one)	input image	segmentation result(more)	input image	segmentation result(more)

图 8 标记分水岭红外船只目标分割效果图

Fig. 8 Effect pictures of infrared ship targets by marker-controlled watershed segmentation

图 8 包括较好分割的单块区域和多区域的分割效果, 对比单块区域检测结果, 近 98% 的船只目标在分割标记框内。拟采用对单块区域的检测结果直

接识别, 对多块区域的分割结果送入深度 CNN 网络的方式进行研究。

实验 2: 基于标记分水岭方法与 Selective

Search 方法的耗时比较。对标记分水岭分割方法进行耗时性能分析,如图 9 所示,其中蓝色为 Selective Search 每帧图像分割用时,绿色为标记分水岭分割用时。标记分水岭分割方法每帧图像平均耗时 0.0137 s, Selective Search 每帧图像平均耗时 0.421 s。基于标记分水岭的分割方法比 Selective Search 方法耗时节省约 1/3,且稳定性能优良。Selective Search 方法对每帧图像平均选取 10 个红外船只候选区域,标记分水岭方法则平均选取 3 个,且船只目标被选出的概率相当。基于标记分水岭的方法在红外船只目标覆盖率及耗时上均有一定的优势。

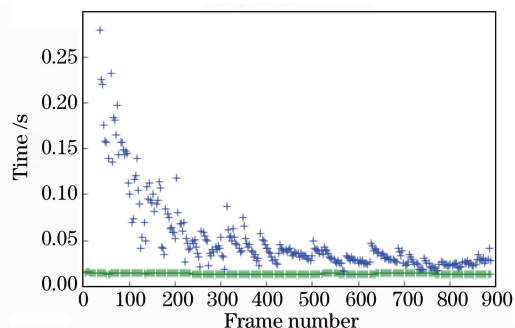


图 9 分水岭方法与 Selective Search 方法分割用时对比
Fig. 9 Segmentation time comparison between marker-controlled watershed method and Selective Search method

4.4 红外船只最终识别效果评估

实验 3:对检测结果进行性能评估。所提出的标记分水岭和深度 CNN 方法基于开源神经网络框架 TensorFlow 实现。

本实验中,首先将标记分水岭分割后的多块候选区域图送入改进的 AlexNet 网络。实现效果如图 10 所示。

由图 10 可以看出,所提方法能较好地对单块目标区域进行判别,对于多块目标候选区域图像则能较好地去除非目标区域,实现较好的检测效果。

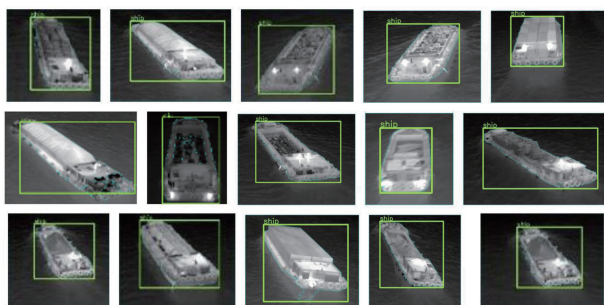


图 10 最终识别结果
Fig. 10 Final recognition results

为了进一步验证本方法的有效性,利用准确率(P)、召回率(R)和 F1 调和平均(F)等精度评价指标对船只最终的识别结果进行定量评价。准确率是测量被正确提取的信息的比例,而召回率用来测量提取的信息中心有多少是正确的,准确率和召回率均为红外船只目标识别中重要的参数,船只目标的识别既要有较高的准确率,也要有较高的召回率,因而也引用 F1 调和平均对其进行了度量。三者的公式分别为:

$$P = \frac{X_{TP}}{X_{TP} + X_{FP}}, \quad (6)$$

$$R = \frac{X_{TP}}{X_{TP} + X_{FN}}, \quad (7)$$

$$F = \frac{X_{TP} \times 2}{X_{TP} \times 2 + X_{FP} + X_{FN}}, \quad (8)$$

式中 X_{TP} 为正确提取的建成区目标, X_{FN} 为漏的船只目标, X_{FP} 为错误提取的船只目标。最终得到本方法的准确率为 93.27%, 召回率为 96.8%, 调和平均为 94.70%。

5 结 论

提出了一种基于深度 CNN 的红外船只目标识别方法。首先采用标记分水岭对图像进行分割,提取红外船只目标候选区域。再将候选区域的图像送入深度 CNN 进行识别。最后,输出识别结果,并标注目标区域。为了验证本算法的有效性,利用实验室自制的长波红外相机拍摄近千张照片及其扩展变换,并分为训练集和测试集。实验结果表明,本方法的准确率为 93.27%, 召回率为 96.8%, 调和平均为 94.70%。多种实际场景实验结果表明,相比于传统检测方法,所提出的算法不仅检测性能好,而且耗时少,工程适用性强。

参 考 文 献

- [1] Wang Y Q, Ma L, Tian Y. State-of-the-art of ship detection and recognition in optical remotely sensed imagery[J]. Acta Automatica Sinica, 2011, 37(9): 1029-1039.
王彦情, 马雷, 田原. 光学遥感图像舰船目标检测与识别综述[J]. 自动化学报, 2011, 37(9):1029-1039.
- [2] Sun L, Zhou Z Q, Wang B, *et al.* A novel inshore ship detection via ship head classification and body boundary determination [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(12): 1920-1924.
- [3] Ren S Q, He K M, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region

- proposal networks [C]. International Conference on Neural Information Processing Systems, 2015: 91-99.
- [4] Zhang R Q, Yao J, Zhang K, *et al.* S-CNN-based ship detection from high-resolution remote sensing images[J]. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2016, XLI-B7: 423-431.
- [5] Zhu S P, Yang L. Stereo matching algorithm with graph cuts based on adaptive watershed [J]. Acta Optica Sinica, 2013, 33(3): 0315004.
祝世平, 杨柳. 基于自适应分水岭的图割的立体匹配算法[J]. 光学学报, 2013, 33(3): 0315004.
- [6] Cai Q, Liu Y Q, Cao J, *et al.* A watershed image segmentation algorithm based on self-adaptive marking and interregional affinity propagation clustering[J]. Acta Electronica Sinica, 2017, 45(8): 1911-1918.
蔡强, 刘亚奇, 曹健, 等. 一种基于自适应标记与区域间近邻传播聚类的水分水岭图像分割算法[J]. 电子学报, 2017, 45(8): 1911-1918.
- [7] Qu L, Wang K R, Cheng L L, *et al.* Fast road detection based on RGBD images and convolution neural network [J]. Acta Optica Sinica, 2017, 37(10): 1010003.
曲磊, 王康如, 陈利利, 等. 基于 RGBD 图像和卷积神经网络的快速道路检测 [J]. 光学学报, 2017, 37(10): 1010003.
- [8] Yu W S, Hou Z Q, Wang C Y, *et al.* Watershed algorithm based on modified filter and marker-extraction[J]. Acta Electronica Sinica, 2011, 39(4): 825-831.
余旺盛, 侯志强, 王朝英, 等. 基于改进滤波和标记提取的分水岭算法 [J]. 电子学报, 2011, 39(4): 825-831.
- [9] He Z Y, Sun L N, Chen L G. Fast computation of threshold based on Otsu criterion [J]. Acta Electronica Sinica, 2013, 41(2): 267-272.
何志勇, 孙立宁, 陈立国. Otsu 准则下分割阈值的快速计算 [J]. 电子学报, 2013, 41(2): 267-272.
- [10] Chang L, Deng X M, Zhou M Q, *et al.* Convolutional neural networks in image understanding[J]. Acta Automatica Sinica, 2016, 42(9): 1300-1312.
常亮, 邓小明, 周明全, 等. 图像理解中的卷积神经网络 [J]. 自动化学报, 2016, 42(9): 1300-1312.
- [11] Ye G L, Sun S Y, Gao K J, *et al.* Nighttime pedestrian detection based on faster region convolution neural network [J]. Laser & Optoelectronics Progress, 2017, 54(8): 081003.
叶国林, 孙韶媛, 高凯珺, 等. 基于加速区域卷积神经网络的夜间行人检测研究 [J]. 激光与光电子学进展, 2017, 54(8): 081003.
- [12] Xin P, Xu Y L, Tang H, *et al.* Fast airplane detection based on multi-layer feature fusion of fully convolutional networks [J]. Acta Optica Sinica, 2018, 38(3): 0315003.
辛鹏, 许悦雷, 唐红, 等. 全卷积网络多层特征融合的飞机快速检测 [J]. 光学学报, 2018, 38(3): 0315003.
- [13] Girshick R, 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, 2014, 580-587. [14] Krizhevsky A. One weird trick for parallelizing convolutional neural networks [J]. Neural and Evolutionary Computing, 2014, arXiv: 1404. 5997.
- [15] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[J]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517-6525.
- [16] Liu F, Shen T S, Ma X X. Convolutional neural network based multi-band ship target recognition with feature fusion[J]. Acta Optica Sinica, 2017, 37(10): 1015002.
刘峰, 沈同圣, 马新星. 特征融合的卷积神经网络多波段舰船目标识别 [J]. 光学学报, 2017, 37(10): 1015002.