

A Review on Intrusion Detection of Moving Objects at Sea

Dongliang Gu, Xin Jin

Dept. of Navigation, Dalian Naval Academy, Dalian Liaoning
Email: 670884240@qq.com

Received: May 13th, 2020; accepted: Jun. 4th, 2020; published: Jun. 11th, 2020

Abstract

Intrusion detection of moving targets at sea based on visible video is a very important research topic. In this paper, the research status of moving target detection and recognition is introduced from two aspects: the first is traditional moving target detection and recognition; the second is convolutional neural network based on deep learning. Because it is more difficult to extract the moving objects at sea, this paper introduces the research status of detecting and identifying moving objects at sea.

Keywords

Motion Detection, Target Identification, Maritime Intrusion Detection

海上运动目标入侵检测研究综述

谷东亮, 金鑫

海军大连舰艇学院航海系, 辽宁 大连
Email: 670884240@qq.com

收稿日期: 2020年5月13日; 录用日期: 2020年6月4日; 发布日期: 2020年6月11日

摘要

基于可见光视频的海上运动目标入侵检测是一个非常重要的研究课题。本文从传统运动目标检测与识别和基于深度学习中卷积神经网络的目标检测与识别这两个方面对其研究现状进行了介绍。由于海上运动目标的准确提取更加困难, 因此本文又单独介绍了海上运动目标检测与识别的研究现状。

关键词

运动检测, 目标识别, 海上入侵检测

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

我国现有的海上安全防护设施存在观测、监控及警戒手段稍微落后的缺陷, 其他国家侵犯我国海域的情况屡有发生。目前国内外文献主要集中于运动目标检测和运动目标识别这两个专题的研究, 对入侵检测等行理解方面的研究较少。因此, 本文从传统运动目标检测与识别和基于深度学习中卷积神经网络的目标检测与识别这两个方面对其研究现状进行了介绍。

但是相比于陆地场景, 复杂海面背景下的运动目标检测存在以下挑战:

1) 海面背景中大量不断运动的海浪经常会被误判为运动目标。

2) 海上经常会出现雨天、雾天的情况, 这使得运动目标检测的背景变得更加复杂多变。晴好天气也经常会出现光照突然变化的现象, 阳光直射到海面上会产生强烈的镜面反射现象, 光照强度突然变化的背景和不断运动的鱼鳞光经常会被误判为运动目标。

3) 由于有些需要监视的海域离岸较远, 需要将摄像机安装在舰船或无人艇上, 这种摄像机运动的情况对检测方法的鲁棒性提出了更高的要求。

近年来, 计算机视觉、模式识别、人工智能等技术发展得异常迅猛, 智能视频监控领域也随之取得了巨大的突破, 这无疑为海上监测技术突破提供了良好的契机。因此, 针对现有智能视频监控领域海上监测存在的难题, 本课题对基于可见光视频的海上运动目标入侵检测研究现状进行了介绍。

2. 传统运动目标检测与识别

运动目标检测是后续目标行为分析等技术的前提, 传统运动目标检测方法主要通过判断目标是否运动从而实现目标的检测, 按照其基本原理主要可以分为背景差分法、帧间差分法、光流法三类。

1) 背景差分法

这一类方法的基本原理是通过当前帧图像和预先得到的背景图像相减的方式实现运动目标的提取。例如 Wren 等人[1]针对没有前景目标和光照变化影响情况下提出了使用单高斯模型对背景进行更新。但是大多数的监测场景都是存在一定变化的, 针对这一问题 Stauffer 等人[2]提出了一种混合高斯模型(GMM), 通过使用多个不同权值的单高斯模型联合对背景进行更新。为了更好地适应动态背景变化, Droogenbroeck 等人[3]通过把每个像素点存储为若干采样值的方法提出了一种基于提取虚拟背景(ViBe)的运动目标检测方法。

这一类方法运算快, 适用于对实时性要求较高的系统, 但是背景模型构建得是否准确对检测结果会产生较大影响, 因此其对背景的变化比较敏感。

2) 帧间差分法

帧间差分法的基本原理是将视频序列中相邻两帧图像进行相减, 然后通过平滑去噪等方式得到差分图像, 从而实现运动目标的提取。Lipton 等人[4]通过这一方法实现了视频中运动目标的检测, 然后对目标

进行识别与跟踪。在二帧差分的基础上, 文献[5]通过三帧差分与自适应背景差分相结合的方法取得了更好的检测结果, 但是检测出的运动目标内部仍然容易出现空洞。为了更好地克服“空洞”现象, 文献[6]首先通过帧差法将每一帧视频进行分割, 然后对不同区域分别使用不同的更新率更新混合高斯模型, 从而实现背景的更新。

帧间差分法运算同样非常简单, 但是相邻帧时间间隔较短、目标运动缓慢等因素容易导致目标提取得不完整及目标的漏检。

3) 光流法

1981年裴巧娜[7]首次提出了光流场的概念, 创造性地将二维速度和灰度结合在了一起。针对光照突变和遮挡问题, Haussecker 等人[8]通过建立灰度变化模型实现了运动和灰度的自动同步估计, 但仍然无法消除阴影的影响。为了解决光流法计算速度慢这一问题, 文献[9]利用 GPU 可以高效处理视频这一能力, 通过使用 CUDA 体系框架实现了基于光流法的实时检测。

光流法相比于前两种方法运算量较大, 不适用于对实时性要求较高的系统, 并且需要特殊硬件的支持。

在运动目标识别方面的研究需要将视频问题转化成图像问题, 通过对每一帧的视频图像进行处理从而实现视频序列中目标的识别。传统的运动目标识别方法是通过人工设计图像中目标的 SIFT、HOG、Haar 等特征, 然后使用 SVM、AdaBoost 等分类器实现图像中目标的识别。

其中最具有代表性的是 Viola 等人[10]提出的基于 AdaBoost 算法框架的人脸检测方法, 这一方法通过提取人脸的 Haar 特征进行识别, 然后通过滑动窗口搜索策略实现在识别出人脸的同时准确地定位出人脸在图像中的位置。Dalal 等人[11]使用图像的局部梯度直方图(HoG)作为目标特征, 提出了一种基于支持向量机(SVM)的行人检测方法。基于这两种方法的各种改进算法被应用于不同种类目标的检测与识别中, 针对图像中目标非刚性形变问题 Felzenszwalb 等人[12]在继承 HoG 特征和 SVM 分类器的情况下提出了基于多尺度形变部件模型(DPM)的目标识别方法, 这一方法可以较好地适应目标的旋转变形。Ren 等人[13]提出使用稀疏编码直方图特征(HSC)代替基于 DPM 方法中的 HoG 特征, 进一步提高了目标识别的准确率。

在 2012 年之前, 目标识别任务中最具有影响力的方法就是基于 DPM 方法的各种改进方法。但是由于这些方法主要依靠人工设计的特征, 特征的好坏直接会影响识别的结果。自从 2012 年卷积神经网络应用于目标识别方面的研究中, 通过使用卷积神经网络提取图像中目标的深层特征, 图像识别的准确率获得了大幅度的提升, 同时目标检测方面的研究也取得了巨大的突破。

3. 基于卷积神经网络的目标检测与识别

卷积神经网络是受到生物视觉系统神经机制的启发而被特别设计出的一种生物物理模型, 可以被看成一种前馈神经网络或多层感知器。1962年, Hubel 等人[14]通过对猫的视觉皮层细胞进行研究提出了“感受野”这一概念。1979年, 日本研究人员 Fukushima [15]基于感受野提出了一种神经认知机模型, 这一模型被称为第一个卷积神经网络。1998年, LeCun 等人[16]通过将卷积层与下采样层相结合的方法提出了被称为现代卷积神经网络雏形的 LeNet。直到 2012年 Krizhevsky 等人[17]使用修正线性单元(ReLU)作为网络中的激活函数提出了 AlexNet, 卷积神经网络的发展从此取得了历史性的突破。自 AlexNet 之后演化生成了许多不同的网络结构, 例如以 VGGNet16 [18]、GoogLeNet [19]为代表的加深模型, 以 Faster R-CNN [20]为代表的区域模型, 以 PSPNet [21]为代表的分割模型。自从卷积神经网络取得突破性的进展, 受到了国内外学者的广泛关注, 在目标的检测与识别方面获得了广泛的成功应用。

在目标识别方面的研究中, 为了使用于训练和测试算法的图片数据可以更真实地反应我们这个世界, 进一步地丰富数据的多样性, Deng J 等人[22]建立了一套包含不同种类物体图片的数据库并将其命名为“ImageNet”。这是一件史无前例的事情, 这一数据库也被国内外的研究人员视为当今人工智能浪潮的

催化剂。“ImageNet”中共包含车、船、人、猫、狗等 1000 种物体, 100 多万张图片, 几乎包含了周围环境中的所有物体。研究人员发现, 不同的算法在经过“ImageNet”训练过后得到的模型都可以用作其他识别任务的启动模型, 然后再对模型进一步地进行微调就可以完成识别任务。Krizhevsky 等人[17]提出的基于 AlexNet 的图像分类方法在基于“ImageNet”数据库举办的 2012 年 ILSVR 图像分类挑战赛中取得了最好的成绩。在 2013 年至 2017 年的 ILSVR 目标分类挑战赛中基于 Claeifai [23]、GoogleNet [19]、残差网络[24]、六模型集成、双通道网络[25]的图像分类方法分别取得了最好的检测结果。其中在 2017 年 ILSVR 图像分类挑战赛中基于双通道网络的图像分类方法把错误率降低至了 3.41%。

在目标检测方面的研究中, 基于卷积神经网络的目标检测方法在识别出目标的同时可以通过边框回归等方法检测出目标的具体位置。Girshick 等人[26]在 2014 年通过将传统目标检测方法中的区域建议与卷积神经网络结合在一起提出了区域卷积神经网络模型(R-CNN), 取得了目标检测领域中里程碑式的突破。文献[27]在 R-CNN 的特征提取网络与全连接层之间加入了一个空间金字塔池化(SPP)层, 有效地提高了检测的速度。Girshick R 等人[28]借鉴 SPP-NET 的思想并引入了多任务损失函数, 提出了快速区域卷积神经网络(Fast R-CNN), 进一步地提高了目标检测的速度和精度。Ren S 等人[20]在此基础上提出了更快区域卷积神经网络(Faster R-CNN), 这一网络在 Fast R-CNN 的基础上加入了一个区域生成网络(RPN), 使用 RPN 来实现候选区域的提取, 实现了端到端的训练过程。基于这一网络的目标检测方法在 2015 年的 ILSVR 目标检测挑战赛中取得了 62.1%的平均准确率, 刷新了目标检测速度与精度的新高。Faster R-CNN 的杰出贡献在于完全使用卷积神经网络同时实现了目标的检测与识别, 但是其检测速度仍然没有完全达到实时。因此, 为了更进一步地提高目标检测的速度, 2016 年 Redmon J 等人[29]提出了 YOLO, 这种方法将目标检测的任务转化为一个回归问题, 很好地加快了检测的速度。在 YOLO 之后, Liu W 等人[30]结合 YOLO 和 Faster R-CNN 中的回归思想提出了基于 SSD 的目标检测方法, 检测速度与检测精度相比于 YOLO 都有所提高。YOLO 和 SSD 的检测速度虽然达到了实时性的要求, 但是由于这两种方法都没有对候选区域进行提取的过程, 因此其检测精度低于 Faster R-CNN。在 2016 年的 ILSVR 目标检测挑战赛中, Zeng 等人[31]使用基于门控双向卷积神经网络(gated bi-directional CNN)的检测方法获得了 66.28%的准确率。在 2017 年的挑战赛中, Shuai 等人通过将门控双向卷积神经网络和特征金字塔网络相结合得到了 73.14%的准确率。2018 年 Wang H 等人[32]提出通过多尺度位置感知核表示(MLKP)提取目标深层特征的高阶统计量, 然后将 MLKP 集成于 Faster R-CNN 中, 进一步地提升了目标检测性能。Singh B 等人[33]为了更好的检测多尺度目标提出了一种新的训练方法, 这一方法只回传和预训练模型所基于的训练数据尺寸相对应的 ROI 梯度, 并且通过图像金字塔对不同尺寸数据进行处理。Zhang Z 等人[34]针对 SSD 检测小目标容易误检的问题, 通过引入分段模块和全局激活模块在用于检测的特征图中加入了语义信息, 提升了小目标检测的准确性。

基于卷积神经网络的目标检测与识别方法, 与传统运动目标检测方法相比的优势在于它可以较好地克服复杂背景的影响, 并且无论感兴趣的目標是否运动, 只要目标在视频画面中出现, 即使目标处于静止状态也可以准确地将其检测出来。与传统人工提取特征的目标识别方法相比优势在于其对大量图片自动地提取特征并不断地进行学习, 以大量的图片为先验知识从而实现目标的检测与识别, 具有更强的鲁棒性。目前基于以上卷积神经网络模型的目标检测方法已被成功应用于行人检测、车辆检测、病理图像分析等领域中。

4. 海上运动目标检测与识别

相比于陆地场景中的运动目标检测与识别, 复杂海面背景下的运动目标检测与识别具有更多挑战, 现有的检测方法主要可以概括为基于建立背景模型、基于显著图视觉注意模型和基于边缘、纹理等特征三类:

1) 基于建立背景模型

基于建立背景模型的海上运动目标检测方法中, Borghraef 等人[35]提出一种基于像素的背景差分法, 通过引入最少量的先验目标信息实现对背景模型的评估与更新。文献[36]中提出了一种基于自适应背景更新和全局搜索的检测方法, 使用快速四连通分量标记方法降低了算法的计算量, 有效地抑制了海浪波动的影响, 实现了对海水养殖海域运动目标的检测。文献[37]提出了一种基于多背景模型的目标检测方法, 使用 Weibull 分布模型来表示海浪波动较大的海域, 使用 Gauss 分布模型来表示比较平稳的海域, 然后通过背景差分法实现运动目标的提取。陈慧敏[38]首先通过三帧差分法提取出运动目标的像素点, 然后对非运动像素点进行背景更新, 在运动像素点区域再进行背景差分获取更精确的运动目标区域, 最后通过膨胀、腐蚀等形态学处理方法实现运动目标的提取。李祎承等人[39]提出了一种基于背景差分和最大类间方差相结合的运动区域提取方法, 并且为了使算法可以适应不同角度的舰船目标引入了基于圆环划分的特征值算法, 使用该算法计算得到的舰船特征具有一定的平移、旋转及缩放的不变性。郭石磊等人[40]将背景差分法和改进的四帧差分法相结合, 通过区域的形状描述子和对运动目标的区域轮廓进行提取实现了海上运动目标的检测。

这类方法具有计算相对简单、实时性强的优点, 但是对背景的变化十分敏感, 例如在海浪波动较大时很容易将海浪误判为目标, 并且这类方法不适用于摄像机运动的情况, 只能在摄像机固定不动的情况下使用。

2) 基于显著图视觉注意模型

在基于显著图视觉注意模型这类海上运动目标检测方法中, 吴琦颖等人[41]首先通过视觉注意模型在静态图像中对感兴趣区域进行提取, 然后通过建立金字塔结构和迭代的倒三角小尺寸低通模板去除海面波浪噪声, 最后在感兴趣区域中使用变形的时间差分法对海上运动目标进行检测。谢小竹等人[42]在这一基础上提出了使用 Sobel 算子计算图像梯度从而实现感兴趣区域的提取, 并对变形的时间差分法做了进一步的改进, 改进后的方法通过相邻帧感兴趣区域的中心点是否移动来判断是否存在运动目标, 有效地降低了漏检率。王焜[43]首先使用视觉注意模型快速获得舰船目标显著性区域, 然后在显著性区域内使用 HoG 特征和学习机制结合实现了舰船目标的准确检测。臧风妮[44]首先在小波域建立了双尺度视觉选择注意模型, 并分别利用梯度法和相位谱法在粗分辨率低频子带上建立了视觉显著图, 然后将视觉选择注意模型和视觉显著图融合形成综合视觉显著图, 最后利用小波反向变换获得原始的高分辨率视觉显著图, 由此实现了海上运动目标的检测。文献[45]首先将彩色图像转变为灰度图像, 然后通过相位谱方法得到图像的临时显著图, 最后使用频率调谐法进一步得到最终的显著图。

这类方法引入了视觉注意模型, 通过对显著图信息中融合的不同尺度和特征进行分析, 可以实现海面目标的准确提取。但是计算较为复杂, 不适用于实时性要求较高的场景。并且当海浪波动较大时, 海浪引起的视觉显著性变化和海上目标引起的变化可能会处于相同量级, 因此容易出现将海浪误判为目标的情况。

3) 基于边缘、纹理等特征

在基于边缘、纹理等特征的海上运动目标检测方法中, 文献[46]首先使用多尺度边缘检测对海面背景中的海天线进行提取, 然后通过海面与目标的亮度值差异实现舰船目标的检测。何四华等人[47]提出在运动目标存在时对背景信号混沌特征的变化进行分析, 然后通过混沌特征变化的差别检测混沌背景信号中的舰船目标信号。李正周等人[48]设计了一个基于海杂波混沌动力系统相空间重构的海杂波动力模型, 使用 RBF 神经网络提取模型中的参数, 从而实现海杂波的预测和抑制, 为海上小目标的检测提供了新的思路。王立地等人[49]首先通过传统的边缘检测方法对舰船目标尾迹进行检测, 然后沿着航迹方向计算有向分维参数, 最后对图像进行二值化处理, 通过对二值图像进行分析确定舰船目标所在位置。李明兵[50]

通过彩色梯度和 Lab 彩色分量分别检测出了海天线和岸线, 实现了对海面区域的准确提取, 然后使用霍夫变换和斜线检测相结合确定出可能存在目标的平行四边形区域, 最后通过形态学处理和中值滤波对舰船目标进行标识。黄龙辉[51]设计了一个用于检测舰船目标的多特征学习框架, 为了弥补局部特征提取的缺点, 使用基于 Gabor 滤波的 MS-CLBP 特征对全局特征进行提取, 取得了很好的检测效果。

这类方法能否检测出目标在很大程度上取决于对视频场景中边缘、纹理等特征提取得是否准确, 因此只能适用于某些特定的场景中, 鲁棒性不强。并且这类方法计算复杂, 实时性需要进一步提升。

相比于海上运动目标检测的研究, 海上运动目标识别研究起步较晚, 研究成果较少, 主要可以分为基于 BP 神经网络和基于卷积神经网络两类:

1) 基于 BP 神经网络

在基于 BP 神经网络的海上舰船识别方法中, 金雪丹等人[52]首先人工提取舰船的边缘、颜色、高度、长度等特征, 然后把这些特征输入 BP 神经网络进行训练, 从而实现将海上舰船目标识别为大船、中船、小船三类。魏娜[53]使用模板匹配技术中的序贯相似性检测法对特征进行提取, 然后设计了一个三层的 BP 神经网络, 最后通过对网络的多次训练实现了舰船目标的识别。马忠丽等人[54]建立了一个岛屿、舰船、礁石的数据库, 然后构建了一个主分量分析降维和组合特征相结合的分级 BP 神经网络, 最后提取不同目标的几何、纹理等特征并将其输入神经网络进行训练, 实现了对岛屿、舰船、礁石三类目标的识别。

由于人工标记特征具有不稳定性和神经网络结构过于简单等原因, 基于神经网络的识别方法无法适应目标的平移、旋转、缩放以及复杂海面的变化, 鲁棒性较差。

2) 基于卷积神经网络

针对基于 BP 神经网络方法识别率较低这一问题, 赵亮等人[55]构造了一个 7 层的卷积神经网络自动地对舰船目标深层特征进行提取, 然后在提取出的特征中加入融合的边缘特征与颜色特征这一约束, 最后通过支持向量机实现了军舰、集装箱船、帆船、渔船、客船的识别。戚超等人[56]使用一个 8 层的卷积神经网络对运输船的特征进行提取, 然后将网络中第一个全连接层的特征用于支持向量机的训练, 实现了散装船、油气船、集装箱船的识别。

这类方法使用卷积神经网络自动提取图像中目标的特征, 既消除了人工标记特征的不稳定性又可以提取出目标的深层特征, 极大地提高了识别的准确率。但是这类方法模型复杂, 因此实时性较差。

5. 结束语

目前基于视频的海上运动目标检测与识别技术的研究仍然存在很多挑战, 未来的研究重点应该是在尽量提升检测速度的同时提出可以更好适应复杂海面背景变化、鲁棒性更强的方法, 并在准确地检测并识别出海上运动目标后进一步对目标的行为进行分析。

基金项目

辽宁省自然科学基金(2015020086); 辽宁省博士启动基金(201501029)。

参考文献

- [1] Wren, C.R., Azarbayejani, A., Darrell, T., *et al.* (1997) Pfunder: Real-Time Tracking of the Human Body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**, 780-785. <https://doi.org/10.1109/34.598236>
- [2] Stauffer, C. and Grimson, W.E.L. (1999) Adaptive Background Mixture Models for Real-Time Tracking. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, 246-252.
- [3] Van Droogenbroeck, M. and Barnich, O. (2014) ViBe: A Disruptive Method for Background Subtraction. In: Bouwmans, T., Porikli, F., Hoferlin, B. and Vacavant, A., Eds., *Background Modeling and Foreground Detection for Video Surveillance*, CRC Press, Boca Raton, 7.1-7.23. <https://doi.org/10.1201/b17223-10>

- [4] Lipton, A.J., Fujiyoshi, H. and Patil, R.S. (1998) Moving Target Classification and Tracking from Real-Time Video. *Fourth IEEE Workshop on Applications of Computer Vision*, Princeton, 19-21 October 1988, 8-14.
- [5] Collins, R.T., Lipton, A.J., Kanade, T., *et al.* (2000) A System for Video Surveillance and Monitoring. VSAM Final Report, 1-68.
- [6] 陈佳, 朱长仁, 钱智明. 一种基于帧差分法与快速图分割相结合的运动目标检测方法[J]. 现代电子技术, 2016, 39(3): 13-17+22.
- [7] 裴巧娜. 基于光流法的运动目标检测与跟踪技术[D]: [硕士学位论文]. 北京: 北方工业大学, 2009.
- [8] Haussecker, H.W. and Fleet, D.J. (2001) Computing Optical Flow with Physical Models of Brightness Variation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**, 661-673. <https://doi.org/10.1109/34.927465>
- [9] Chambolle, A. and Pock, T. (2011) A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *Journal of Mathematical Imaging and Vision*, **40**, 120-145. <https://doi.org/10.1007/s10851-010-0251-1>
- [10] Viola, P. and Jones, M. (2001) Rapid Object Detection Using a Boosted Cascade of Simple Features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, I-I.
- [11] Dalal, N. and Triggs, B. (2005) Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, 886-893.
- [12] Felzenszwalb, P.F., Girshick, R.B., McAllester, D., *et al.* (2010) Object Detection with Discriminatively Trained Part-Based Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**, 1627-1645. <https://doi.org/10.1109/TPAMI.2009.167>
- [13] Ren, X. and Ramanan, D. (2013) Histograms of Sparse Codes for Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Portland, 23-28 June 2013, 3246-3253. <https://doi.org/10.1109/CVPR.2013.417>
- [14] Hubel, D.H. and Wiesel, T.N. (1962) Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. *The Journal of Physiology*, **160**, 106-154. <https://doi.org/10.1113/jphysiol.1962.sp006837>
- [15] Fukushima, K. (1979) Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position-Neocognitron. *IEICE Technical Report A*, **62**, 658-665.
- [16] LeCun, Y., Bottou, L., Bengio, Y., *et al.* (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <https://doi.org/10.1109/5.726791>
- [17] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, Vol. 1, 1097-1105.
- [18] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition.
- [19] Szegedy, C., Liu, W., Jia, Y., *et al.* (2015) Going Deeper with Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [20] Ren, S., He, K., Girshick, R., *et al.* (2015) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in Neural Information Processing Systems*, Vol. 1, 91-99.
- [21] Zhao, H., Shi, J., Qi, X., *et al.* (2017) Pyramid Scene Parsing Network. *IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 2881-2890. <https://doi.org/10.1109/CVPR.2017.660>
- [22] Deng, J., Dong, W., Socher, R., *et al.* (2009) Imagenet: A Large-Scale Hierarchical Image Database. *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, 20-25 June 2009, 248-255. <https://doi.org/10.1109/CVPR.2009.5206848>
- [23] Zeiler, M.D. and Fergus, R. (2014) Visualizing and Understanding Convolutional Networks. In: *European Conference on Computer Vision*, Springer, Cham, 818-833. https://doi.org/10.1007/978-3-319-10590-1_53
- [24] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 26 June-1 July 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [25] Chen, Y., Li, J., Xiao, H., *et al.* (2017) Dual Path Networks. *Advances in Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 4467-4475.
- [26] Girshick, R., Donahue, J., Darrell, T., *et al.* (2016) Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **38**, 142-158. <https://doi.org/10.1109/TPAMI.2015.2437384>
- [27] He, K., Zhang, X., Ren, S., *et al.* (2014) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. In: *European Conference on Computer Vision*, Springer, Cham, 346-361. https://doi.org/10.1007/978-3-319-10578-9_23
- [28] Girshick, R. (2015) Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, 7-13 December 2015, 1440-1448. <https://doi.org/10.1109/ICCV.2015.169>

- [29] Redmon, J., Divvala, S., Girshick, R., *et al.* (2016) You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [30] Liu, W., Anguelov, D., Erhan, D., *et al.* (2016) SSD: Single Shot Multibox Detector. In: *European Conference on Computer Vision*, Springer, Cham, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [31] Zeng, X., Ouyang, W., Yang, B., *et al.* (2016) Gated Bi-Directional CNN for Object Detection. In: *European Conference on Computer Vision*, Springer, Cham, 354-369. https://doi.org/10.1007/978-3-319-46478-7_22
- [32] Wang, H., Wang, Q., Gao, M., *et al.* (2018) Multi-Scale Location-Aware Kernel Representation for Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 1248-1257. <https://doi.org/10.1109/CVPR.2018.00136>
- [33] Singh, B. and Davis, L.S. (2018) An Analysis of Scale Invariance in Object Detection-SNIP. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 3578-3587. <https://doi.org/10.1109/CVPR.2018.00377>
- [34] Zhang, Z., Qiao, S., Xie, C., *et al.* (2018) Single-Shot Object Detection with Enriched Semantics. Center for Brains, Minds and Machines (CBMM). <https://doi.org/10.1109/CVPR.2018.00609>
- [35] Borghgraef, A., Barnich, O., Lapierre, F., *et al.* (2010) An Evaluation of Pixel-Based Methods for the Detection of Floating Objects on the Sea Surface. *EURASIP Journal on Advances in Signal Processing*, **2010**, Article ID: 978451. <https://doi.org/10.1155/2010/978451>
- [36] Hu, W.C., Yang, C.Y. and Huang, D.Y. (2011) Robust Real-Time Ship Detection and Tracking for Visual Surveillance of Cage Aquaculture. *Journal of Visual Communication and Image Representation*, **22**, 543-556. <https://doi.org/10.1016/j.jvcir.2011.03.009>
- [37] 张明杰, 李翠华, 刘明业, 张龙飞. 基于多背景模型的海面运动目标检测[J]. 北京理工大学学报, 2005(6): 495-498+508.
- [38] 陈慧敏. 近岸海域运动目标检测与跟踪算法的设计与实现[D]: [硕士学位论文]. 成都: 电子科技大学, 2013.
- [39] 李祎承, 胡钊政, 初秀民. 基于图像处理的内河船舶目标提取与特征值计算[J]. 交通信息与安全, 2015, 33(3): 1-8.
- [40] 郭石磊, 饶进军. 基于动态背景下的海上运动目标检测[J]. 计量与测试技术, 2016, 43(12): 3-5+8.
- [41] 吴琦颖, 李翠华. 一种新颖的海上运动目标实时检测方法[J]. 计算机工程与应用, 2007(14): 213-216.
- [42] 谢小竹, 洪景新, 肖思兴. 有效的海上运动目标检测方法[J]. 计算机工程与应用, 2009, 45(4): 225-226+231.
- [43] 王焜. 基于视频的运动船只识别与跟踪技术研究[D]: [硕士学位论文]. 厦门: 厦门大学, 2009.
- [44] 臧凤妮. 智能视频监控中海面舰船目标检测算法研究[D]: [博士学位论文]. 青岛: 中国海洋大学, 2014.
- [45] 王金武, 姚志均, 于乃昭. 基于相位谱和频率调谐的海上场景显著性检测[J]. 计算机应用, 2013, 33(S1): 211-213.
- [46] Bouma, H., de Lange, D.J.J., van den Broek, S.P., *et al.* (2008) Automatic Detection of Small Surface Targets with Electro-Optical Sensors in a Harbor Environment. *Electro-Optical Remote Sensing, Photonic Technologies, and Applications II. International Society for Optics and Photonics*, Vol. 7114, Article No. 711402. <https://doi.org/10.1117/12.799813>
- [47] 何四华, 杨绍清, 石爱国, 李天伟. 基于图像序列区域混沌特征的海面舰船目标检测算法[J]. 交通运输工程学报, 2009, 9(1): 73-76.
- [48] 李正周, 陈静, 沈美容, 侯倩, 丁浩, 金钢. 基于混沌神经网络的海上目标图像的海杂波抑制方法[J]. 光电子激光, 2014, 25(3): 588-594.
- [49] 王立地, 黄莎白, 史泽林. 基于有向分维的海面运动目标自动检测方法[J]. 模式识别与人工智能, 2004, 17(4): 486-490.
- [50] 李明兵. 基于视频的海上运动目标检测技术研究[D]: [硕士学位论文]. 天津: 国家海洋技术中心, 2012.
- [51] 黄龙辉. 基于特征学习的场景图像分类和舰船识别研究[D]: [硕士学位论文]. 北京: 北京化工大学, 2017.
- [52] 金雪丹. 计算机视觉在实时船舶识别方法的研究与应用[D]: [硕士学位论文]. 上海: 上海海事大学, 2007.
- [53] 魏娜. 模板匹配和 BP 神经网络在船舶识别中的应用[J]. 舰船科学技术, 2016, 38(20): 133-135.
- [54] 马忠丽, 文杰, 梁秀梅, 陈虹丽, 赵新华. 无人艇视觉系统多类水面目标特征提取与识别[J]. 西安交通大学学报, 2014, 48(8): 60-66.
- [55] 赵亮, 王晓峰, 袁逸涛. 基于深度卷积神经网络的船舶识别方法研究[J]. 舰船科学技术, 2016, 38(15): 119-123.
- [56] 戚超, 王晓峰. 基于卷积神经网络的运输船舶分类识别方法[J]. 微型机与应用, 2017, 36(17): 52-55.