

[引用格式] 郝紫霄, 王琦. 基于声呐图像的水下目标检测研究综述 [J]. 水下无人系统学报, 2023, 31(2): 339-348.

基于声呐图像的水下目标检测研究综述

郝紫霄, 王 琦

(江苏科技大学 计算机学院, 江苏 镇江, 212003)

摘 要: 通过处理声呐图像实现水下目标检测具有重大的军事与民用意义。文章对基于声呐图像的水下目标检测原理、方法、算法和发展趋势进行了全面阐述。将基于声呐图像的水下目标检测任务分为传统水下目标检测、基于深度学习的目标检测, 以及深度学习与迁移学习相结合的目标检测 3 个层面。又分别将传统目标检测分为基于数理统计、基于数学形态学以及基于像素的水下目标检测; 将基于深度学习的目标检测分为基于一阶段方法、基于二阶段方法以及基于 DETR 的目标检测; 将深度学习与迁移学习相结合的目标检测分为基于简单深度神经网络模型的迁移与基于复杂深度学习模型的迁移所实现的目标检测进行了具体讨论。最后总结归纳现有技术的优缺点, 并对该领域的未来发展方向作出进一步的展望。

关键词: 水下目标检测; 声呐图像; 深度学习; 迁移学习

中图分类号: U666.74; TJ630.34

文献标识码: R

文章编号: 2096-3920(2023)02-0339-10

DOI: 10.11993/j.issn.2096-3920.202205004

Underwater Target Detection Based on Sonar Image

HAO Zixiao, WANG Qi

(School of Computer, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: Underwater target detection by processing sonar images is of great military and civil significance. This paper comprehensively describes the principles, methods, algorithms, and development trends in underwater target detection based on sonar images. Initially, we divide the underwater target detection task based on sonar images into traditional, deep learning-based, and combined deep learning- and transfer learning-based underwater target detection. Traditional target detection is divided into underwater target detection based on mathematical statistics, mathematical morphology, and pixels. Deep learning-based target detection methods are primarily divided into one-stage, two-stage, and detection transformer(DETR) methods. Combined deep learning- and transfer learning-based target detection is primarily divided into target detection based on simple deep neural network model transfer and complex deep learning model transfer. Finally, the advantages and disadvantages of the existing technology are summarized, and the future development direction of this field is discussed.

Keywords: underwater target detection; sonar image; deep learning; transfer learning

0 引言

海洋的战略价值独特而重大: 海洋是资源富饶的宝藏, 是生物物种的摇篮, 是新时代经济的增长点, 是战略的新疆域, 是联系世界的通道^[1]。水下

目标检测是水下图像处理的基础, 也是利用与开发海洋资源等水域资源的基础任务之一。水下目标检测已经广泛应用于水中武器探测^[2]、水中兵器追踪与自导^[3]、海洋生物追踪^[4]、渔业养殖与打捞^[5]以及水下环境勘探^[6]等民用与军事领域。相比于

收稿日期: 2022-05-13; 修回日期: 2022-07-28.

基金项目: 江苏省教育厅 2019 年度江苏省高等学校自然科学研究面上项目(19KJB520031).

作者简介: 郝紫霄(1998-), 女, 在读硕士, 主要研究方向为图像处理.

光学成像等水下成像探测技术, 声呐成像具有探测距离远、穿透能力强、适用于浑浊水域等优点。成像声呐分为侧扫声呐^[7]、合成孔径声呐^[8]和多波束声呐^[9]等类型, 各类型的优缺点如表 1 所示, 一般以装载于潜艇等水下航行器的形式进行探测任务。

表 1 多种成像声呐类型及优缺点

Table 1 Advantages and disadvantages of various imaging sonars

成像声呐类型	优点	缺点
侧扫声呐	扫描分辨率高; 经济高效	存在声脉冲散射问题
合成孔径声呐	分辨率高; 抗干扰能力强	硬件要求高
多波束声呐	探测距离远; 精度高	操作难度高
前视声呐	低功耗; 实用价值高	存在高旁瓣现象

基于声呐图像的水下目标检测主要涵盖对水下目标进行定位与分类两方面。有效的特征提取与特征处理是目标检测任务实施的关键。声呐图像作为不同于自然图像的一种特殊图像, 具有更特殊的纹理特征、边缘特征、矩阵特征和统计特征等。基于声呐图像的水下目标检测的传统方法可分为基于数理统计的目标检测、基于数学形态学的目标检测与基于像素的目标检测。其中, 提取特征常应用尺度不变特征变换(scale-invariant feature transform, SIFT)算法^[10]及其基础上改进的加速稳健特征(speeded up robust features, SURF)算法^[11]、方向梯度直方图(histograms of oriented gradients, HOG)算法^[12]等, 特征处理常用的算法有支持向量机(support vector machine, SVM)算法^[13]和贝叶斯算法^[14]等。近年来, 随着深度学习与迁移学习的不断发展, 相关技术也被广泛应用于基于声呐图像的水下目标检测任务, 比如用于特征提取的卷积神经网络^[15]与用于特征处理的特征金字塔网络等^[16]。

文中结合声呐图像水下目标检测的原理、方法、算法和模型等, 从多方面多层次对声呐图像的水下目标检测进行了详细论述。首先给出基于声呐图像的传统目标检测以及基于深度学习、迁移学习与深度学习结合的目标检测的最新研究进展分析, 然后总结了相关方向亟待解决的问题, 最后对未来的发展趋势作出展望。

1 基于声呐图像的传统目标检测方法

1.1 基于数理统计方法的水下目标检测

声呐图像处理得到的直方图具有的均值、标

准差、峰值等数理统计特性与声呐图像处理得到的共生矩阵所具有的相关性、对比度、熵值等数理统计特性都可以反映声呐图像中的物体分布与物体种类。因此, 利用声呐图像的数理统计特性来实现水下目标检测具有可行性。

声呐图像具有显著的数理统计特性主要由于其特殊的纹理特征, 声呐图像中不同的底质呈现在声呐图像中的纹理特征不同。作为物体所具有的内在特征之一, 纹理特征越来越受到重视。由于纹理可以直接反映海底表面结构的粗糙度, 纹理特征常被用于声呐图像的分析, 各种纹理特征的对比如表 2 所示。陈强等^[17]将合成孔径声呐图像的统计特性和纹理特性相结合, 将提取到的数理统计特性输入到 SVM 分类器中进行训练, 并经实验验证取得良好效果。王涛等^[18]利用灰度共生矩阵提取侧扫声呐图像的纹理特征, 并利用 K-means 算法进行海洋底质分类, 从而实现对泥、砂和岩石 3 种海底底质类型的目标检测。然而常用于声呐图像目标检测的纹理特征也具有诸多局限性, 比如灰度共生矩阵的数理统计特性的计算量宏大且在处理过程中寻找特征无关的特征量组合比较困难。针对这一问题, 董凌宇等^[19]根据盒维数、毯维数与多重分形谱的侧扫声呐图像纹理提取算法, 设计了基于分形纹理特征的 Adaboost 分类器, 通过更便捷有效地处理纹理特征的数理统计特性获得更好的沉船目标检测效果。

表 2 不同纹理特征的对比

Table 2 Comparison of texture features

纹理特征	优点	缺点
灰度共生矩阵	描述具有方向性及灰度差巨大的纹理图像时效果好	计算量大; 特征量之间的统计相关性普遍存在
Tamura 纹理	具有良好的旋转不变性与尺度不变性	描述声呐图像的局部纹理特征时效果差
分形纹理特征	描述整体纹理与局部纹理的效果均良好; 识别效率高	精确率提升小

由于海洋环境复杂多变, 探测系统航行过程中可能产生气泡从而影响能量的发射和接收, 同时声呐载体平台的摇摆变化也可能造成声呐图像的失真, 这些问题的存在使声呐图像的处理存在诸多难点。为应对声呐图像中目标检测的挑战, 图像数理统计特性的有效利用尤为关键。王其林

等^[20]提出了优化后的 Hough 变换和线段检测算法用于侧扫声呐图像直线特征的提取, 建立图像边缘掩码矩阵使相应的特征点都能参与到直线特征的形成中, 并建立了一种新的断裂线特征拟合方法。但是, 利用较为单一的形状描述子进行特定特征的提取仍无法满足水下目标检测的精度要求。田晓东等^[21]定义了几种不同的形状特征描述子, 并统计计算描述子直方图, 根据直方图所反映的声呐图像的数理特性计算直方图间的相似性, 用于完成声呐图像中的目标检测与识别等任务。针对声呐图像的强干扰性致使二值化后的声呐图像在背景区中掺杂孤立区等问题, 卢逢春等^[22]将直方图的概念推广到属性直方图, 并将其应用到 Otsu 法中形成一种基于声呐图像数理统计特性的图像处理方法。

国内外在基于数理统计方法的声呐图像水下目标检测上也不断发展。Yang 等^[23]在所提技术的水下目标检测中, 利用物体几何特征的直方图计算物体属性的相关系数并进行聚类, 达到了较强鲁棒性的效果。但是仅通过一类数理统计特征无法充分反映声呐图像复杂的背景特征与纹理特征, 因此多种数理统计特征的综合使用成为趋势。王晓等^[24]分别采用 HSV (hue, saturation, value) 空间模型、方向梯度直方图和局部二值模式的方法处理声呐图像的色度、形状和纹理特征, 为水下自动目标检测任务的开展提供良好基础。将基于数理统计的方法与高效的目标检测模型结合也被证实效果良好。Sun 等^[25]将声呐图像纹理识别算法与 YOLO(you only look once)算法融合, 在海洋沉积物检测上取得了优于原始算法的结果。

1.2 基于数学形态学的水下目标检测

声呐图像在目标检测的特征提取阶段存在难点主要是由于海洋环境的复杂性与检测方式的特殊性所导致的特征地貌边缘检测困难、噪声干扰大等因素。为应对这些问题, 在水下目标检测任务中采用数学形态方法对声呐图像进行相应处理至关重要。数学形态学是分析几何形状和结构的科学, 是用集合论的方法定量描述目标几何结构的方法。罗进华等^[26]利用数学形态学方法对侧扫声呐图像进行处理, 用具有一定形态的结构元素度量和提取声呐图像中的对应形状, 并通过实验

验证了其方法可以有效得到连续化、粗化、圆滑的特征边缘。针对声呐图像对比度低、成像质量差、前景与背景的对比度低等问题所造成的自然目标与人造目标物的正确区分率低的问题, 邹岗等^[27]基于数学形态学算子, 通过定义多尺度形状的数学形态结构元素, 对各像素点进行形态开闭运算, 并且用阈值评判法实现对水中人造目标的快速有效检测。

声呐图像相比自然图像的水下目标检测更困难的原因还包括声呐图像存在散斑噪声等多种噪声综合作用的问题。目前的声呐图像目标检测任务广泛进行了数学形态学滤波的处理, 如表 3 所示。Mallet 等^[28]提出了一种新的基于小波的形态学滤波方法, 通过设计新的小波阈值方法, 解决合成孔径声呐图像被斑点噪声(颗粒乘法噪声)破坏的问题, 在合成孔径声呐图像的海底目标检测中取得了优于相关滤波方法的效果。Isar 等^[29]结合小波域散斑抑制方法与贝叶斯算法提出了一种新的声呐图像中水下目标的检测方法, 创新点为基于双边滤波器的变体获得一种对数据进行非线性运算的斑点抑制方法, 可以有效抑制散斑噪声, 提升目标检测精度。金凤来等^[30]结合改进的 Lee 滤波与中值滤波, 提出了一种新的声呐图像处理方法。然而, 基于小波理论的方法计算量较大, 实时性差, 缺乏实用性; 基于 Lee 滤波等常用的散斑抑制方法也存在无法平衡滤波效果与边缘检测的局限性。为解决之前方法存在的弊端, 郭海涛等^[31]设计了一种抑制声呐图像散斑噪声的多尺度多方向结构元素的两级级联形态学滤波器, 在一定程度上解决了散斑噪声严重影响基于声呐图像的海底小目标检测的问题。

表 3 应用于声呐图像水下目标检测的数学形态学方法对比

Table 3 Comparison of mathematical morphology methods applied to sonar image in underwater target detection

数学形态学方法	优点	缺点
中值滤波	部署简便; 速度快	抑制噪声效果较差
Lee滤波	无需精确建模; 能较好保持边缘	计算量大; 速度较慢
Kuan滤波	滤除散斑噪声、平滑图像时对边缘不产生影响	边缘保持能力差
Frost滤波	抑制噪声效果较好	边缘细节等结构信息模糊; 存在盲目平滑现象

1.3 基于像素的水下目标检测

在处理声呐图像进行水下目标检测特别是运动目标检测时,常用的基于像素的方法主要有背景帧差法和消减法等。帧差法是用图像流,即视频序列中的相邻帧作差分运算,根据阈值将得到的像素差值区域分为背景和前景,差值大于阈值的区域为前景,差值小于阈值的区域为背景。帧差法具有对动态背景不敏感、无需背景建模、计算简单及实时性好等优点。背景消减法又称为背景减除法,适用于背景相对静止情况下的水下运动目标检测。背景消减法将视频序列的当前帧与背景帧作差分运算,利用设置的阈值将像素差区域划分为背景与前景,差值大于阈值的区域划分为前景,小于阈值的区域划分为背景。因此,背景消减法也可以理解作为一种特殊的帧差法。崔杰等^[32]提出了一种基于帧差法的多波束前视声呐运动目标检测方法,采用三帧差分法,从三帧声呐数据中得到运动目标的二值图像,从而实现运动目标的检测。但是,帧差法也存在空洞现象,不易获取完整准确的水下运动目标的轮廓。为解决这一问题,崔杰等^[33]又提出了一种基于声呐图像序列的背景消减法的改进 Mean Shift 水下运动目标检测算法,经实验验证,该算法可以取得更高的检测精度。然而,背景减除法也具有对光照变化和背景变化十分敏感的缺陷。结合多种基于像素方法进行声呐图像的水下目标检测任务已成为趋势,多种算法的结合可以发挥不同算法的优势并在一定程度上弥补其中单一算法的局限性,Liu 等^[34]结合背景减除法与三帧差分法,提出了一种更适应水下复杂环境和光照条件的水下目标检测方法。

相比于帧差法等传统的基于像素的目标检测算法,恒虚警率(constant false alarm rate, CFAR)算法也常被应用于高分辨水下图像的目标检测。CFAR 算法在声呐图像处理领域已经非常成熟,并具有较高处理效率。Kalyan 等^[35]将 CFAR 技术应用到声呐目标检测任务中,使用平均决策阈值(average decision threshold, ADT)作为度量,该方法可以较好地克服目标干扰,实现水下目标检测。Li 等^[36]将 CFAR 应用于声呐图像的处理,提出了一种基于双参数自适应窗口的声呐图像目标检测方法,可以实现对复杂多目标的高精度检测。但由于

CFAR 算法是基于局部滑动窗口对声呐图像进行逐像素检测,会产生计算速度低的问题。为解决这一问题,CFAR 算法的诸多改进算法被提出,比如单元平均恒虚警率(cell averaging-constant false alarm rate, CA-CFAR)、有序统计恒虚警率(order statistics-constant false alarm rate, OS-CFAR)等。相比于原始 CFAR 算法,CA-CFAR 算法通过对图像中单个像素的能量与检测阈值比较来判定是否为检测目标,其自适应检测阈值根据背景杂波的分布函数和设定的虚警率计算得到,因此可以取得比 CFAR 算法更低的复杂度与更高的运算速度。Villar 等^[37]提出了一种基于 OS-CFAR 的侧扫声呐图像目标检测方法,在总任务的基础工作即目标检测阶段,应用 OS-CFAR 以克服环境波动,该算法相比传统 CFAR 算法速度更快,且在高分辨率下能有更优的性能。

2 基于深度学习的声呐图像水下目标检测

近年,基于深度学习的目标检测方法逐渐成为主流。基于深度学习的目标检测方法主要分为一阶段和二阶段,又分别被称为基于候选区域的目标检测方法与基于回归的目标检测方法。2020 年后,DETR(detection transformer)方法的提出也为声呐图像的水下目标检测领域带来了新活力。目前,基于深度学习的方法在声呐图像的水下目标检测任务中已取得了广泛应用与良好效果。

2.1 基于二阶段方法的声呐图像水下目标检测

以区域卷积神经网络(region-convolutional neural network, R-CNN)^[38]及其改进算法快速区域卷积神经网络(Fast R-CNN)^[39]、更快速区域卷积神经网络(Faster R-CNN)^[40]等为代表的二阶段目标检测算法先由算法生成一系列作为样本的候选框,再通过 CNN 进行样本分类,其算法流程图如图 1 所示。这种将目标检测任务分解为 2 个阶段的方式精度很高,但也存在实时性差的缺点。

曾文冠等^[41]将 Faster R-CNN 模型(见图 2)应用到声呐图像中的海星、贝类等水下目标的检测任务中,全类平均精度(mean average precision, mAP)可达 63.03%,单张图像的处理时长可达 1.172 s,均优于传统的目标检测方法。

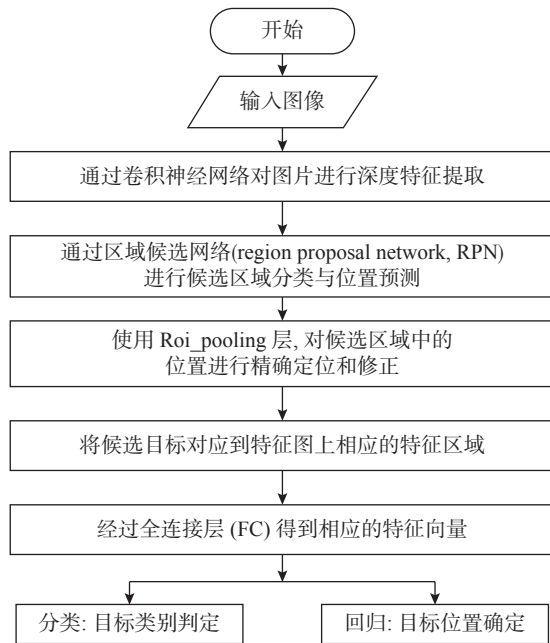


图 1 二阶段目标检测算法流程图

Fig. 1 Flow chart of two-stage target detection algorithm

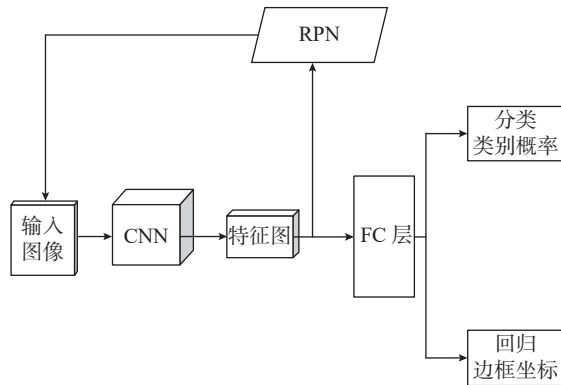


图 2 Faster R-CNN 示意图

Fig. 2 Diagram of Faster R-CNN

针对之前应用 Faster R-CNN 模型的声呐图像水下目标检测系统所存在的参数量过大、训练复杂度、检测速度慢等问题, Fang 等^[42]使用一系列 CNN 提取图像特征, 并构造 RPN 提取可能包含目标的区域, 提出了更为优化的基于 Faster R-CNN 的声呐图像目标检测方法。Ma 等^[43]也提出了基于优化的 Faster R-CNN 的目标检测方法, 引入了噪声对抗网络(noise adversarial networks, NAN)解决声呐图像质量受环境噪声干扰严重的问题。然而, 二阶段目标检测模型的噪声鲁棒性存在缺陷, 不利于声呐图像的目标检测, 马麒麟^[44]在二阶段模型 Faster R-CNN 的基准上进行改进, 引入对抗学

习方法提高了目标检测模型的噪声鲁棒性, 并利用生成对抗网络生成样本, 改善了声呐图像样本稀少导致的模型过拟合、泛化性差等问题, 相比原始二阶段目标检测模型, 精度提升了 4.8%^[45]。

2.2 基于一阶段方法的声呐图像水下目标检测

以 YOLO 系列、轻量级目标检测 SSD(single shot multi-box detector)系列为代表的一阶段目标检测方法直接回归物体的类别概率和位置坐标值^[46-47], 因此速度远快于二阶段目标检测方法, 其算法流程如图 3 所示, 最终会得到反映是否包含目标的置信度(confidence)参数与描述包围框位置的坐标参数。但一阶段目标检测方法直接通过主干网络给出物体类别和位置信息, 没有使用 RPN 网络, 所以其精度略低于二阶段目标检测方法。一阶段目标检测更适应于对实时性与检测效率要求高的任务, 因而在水下目标检测中效果显著。

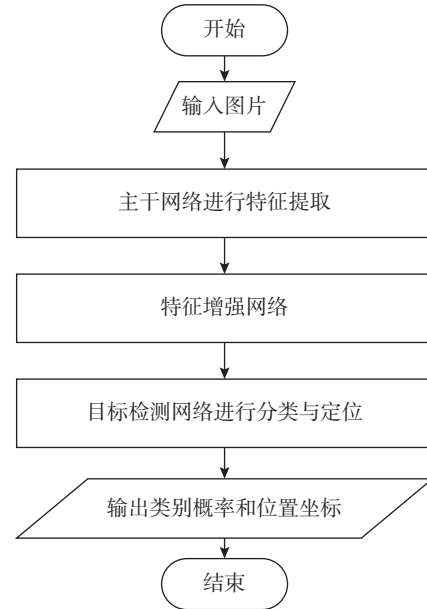


图 3 一阶段目标检测算法流程图

Fig. 3 Diagram of one-stage target detection algorithm

目前基于深度学习的声呐图像目标检测所面临的挑战主要有两方面: 一方面是相关数据集较少且图像样本不足; 另一方面是检测精度与速度之间的平衡。Wu^[48]将 YOLO-v3 模型应用于声呐图像水下目标检测任务中, 以数据增强策略解决样本不足问题, 以重新构建骨干网络的方式解决检测效率低的问题。然而, YOLO-v3 模型也存在局限, 因此研究者更倾向于将改进后的 YOLO 模

型应用到水下目标检测中。王霞等^[49]利用改进后的 Tiny-YOLO-v3 模型设计了一种更适应于水下小目标检测的方法,通过图像去噪的方式应对样本不足的问题,并通过采用轻量化 YOLO 模型的方式应对实时性不足的问题。Fan 等^[50]将 YOLO-v4 模型应用到声呐图像水下目标检测中,改进 YOLO-v4 的主干网络以解决原始网络模型参数量较大、模型深度过深引起的速度较慢的问题;改进特征增强模块以解决样本不足的问题。Yu 等^[51]提出了一种基于 YOLO-v5 的侧扫声呐图像水下目标检测算法 TR-YOLO-v5,通过引入注意力机制、添加 Transformer 模块,以适应声呐图像目标稀疏、特征贫乏的特点,从而提升检测精度;引入数据增强技术生成更多具有水下目标的侧扫声呐图像,以解决样本缺乏的问题;在训练中使用锚聚类方法,以提高收敛速度^[52]。

SSD 网络也是基于声呐图像的水下目标检测领域常用的一阶段目标检测网络。SSD 网络与 YOLO 网络的对比如表 4 所示。与 YOLO 类似,SSD 算法将目标检测任务的目标判定与目标识别合而为一,省去区域提名环节,对输入的声呐图像直接传入 CNN,在不同尺度的卷积层提取特征,从而完成水下目标检测任务。凡志邈等^[53]采用 SSD 网络实现基于声呐图像的水下单目标检测,并通过实验验证其精度与速度优于传统的水下目标检测方法。然而原始的 SSD 模型存在小目标检测精度低、模型参数过大等问题。因此对二阶段目标检测方法进行改进,再应用到水声图像的目标检测中更能满足特定应用场景的需求。李宝奇等^[54]利用可扩张、可选择的卷积核模块(ESK)重新搭建了 SSD 的基础网络与附加特征提取网络,并采用

SSD 的轻量化模型 SSD-MV2,最终构建了 SSD-MV2ESK 模型,在水声图像目标检测的精度上比原始算法提高了 4.71%,速度也得到了提升。

基于深度学习的声呐图像目标检测目前的研究难点包括数据集不足、小目标检测精度低、实时性差及准确率低等,这与水下环境的复杂性与声呐成像的特殊性有关。一阶段模型更为轻便快速,二阶段模型更为精确,在实际使用中往往根据应用需求进行选择与改进。

2.3 基于 DETR 的声呐图像水下目标检测

DETR^[55]是 Facebook 团队提出的基于 Transformer^[56]的端到端的目标检测算法,将目标检测视为集合预测的问题,从全局上将检测出的目标整体视为目标,而不是单独预测感兴趣物体。相比于主流检测算法 Faster-RCNN、YOLO 等,其舍弃了定义与编码锚框、非极大值抑制等操作,实现了端到端。汤寓麟等^[57]将 DETR 与 YOLO 模型融合,并引入通道注意力机制,加入多尺度特征复融合模块,提出一种适用于侧扫声呐图像的沉船目标检测算法 DETR-YOLO,并验证了 DETR 算法具有只用少量锚框即可完成目标预测的特征,可提升水下小目标检测精度,并使模型轻量化,实时性更强。

3 深度学习与迁移学习相结合的声呐图像水下目标检测

相比于传统算法,深度学习在目标检测的特征提取方面具有显著优越性,但深度学习是由数据驱动的,当数据集规模不够大时,容易造成过拟合的问题。而迁移学习的方法可以将大规模数据集上训练好的模型迁移到体量较小的数据集上,再进行微调^[58]。近些年,由于声呐图像相关数据集所存在的样本少、无法大量获取有标记的数据等问题,迁移学习被广泛应用到水下目标检测任务中。

朱兆彤等^[59]提出了一种利用迁移学习训练深度 CNN 实现声呐图像目标检测的方法,利用迁移学习的思想,通过对已有的训练网络进行微调与再训练,达到了更优的检测率与速度。武铄等^[60]提出了一种融合迁移学习与深度学习的侧扫声呐图像检测方法,先用强相关数据集与弱相关数据集训练网络,再用 Shipwreck 数据集进行网络训练,

表 4 YOLO 与 SSD 的对比
Table 4 Comparison between YOLO and SSD

	SSD	YOLO
损失函数	Softmax loss	Logistic loss
特征提取器	VGG19	Darknet-53
包围框预测	默认框直接偏移	网格单元通过 Sigmoid 激活函数偏移
对小目标的检测效果	底层的语义值不高,对小目标的检测效果差	分辨率较高的层有较高的语义值,对小目标的检测较好
对大目标的检测效果	较好	较差

经实验验证,采用迁移学习比未采用时正确率提高了9.81%,且小目标的检测正确率与泛化性提升明显。

与之前研究中利用结构较简单的深度卷积神经网络构建基于迁移学习和深度学习的声呐图像目标检测系统不同,于森^[61]利用迁移模型,在利用组合池化方式丰富特征层的改进后的YOLO-v3模型上训练数据,并通过生成对抗网络扩充数据集,最终实现准确快速的目标检测。盛子旗等^[62]将在ImageNet数据集中预训练好的VGG16网络、Darknet53网络的初始参数迁移到深度学习模型Faster R-CNN、SSD和YOLO-v3中,并通过真实样本与仿真样本结合的方法进行实验,提升了水雷目标检测精度。Tang等^[63]将在COCO(common objects in context)数据集上训练的部分卷积层进行YOLO-v3模型的迁移学习,并改进了特征金字塔网络与多尺度特征融合操作,完成针对侧扫声呐图像的水下目标检测。

为进一步兼顾精度与速度,许多结合深度学习、迁移学习以及传统方法的水下图像目标检测方法也被提出。付同强等^[64]将数理统计特性、形态学滤波的思想应用到深度学习框架中,并通过迁移学习进行微调,最终提出一种采用优化二维变分模态分解(two dimensions variational mode decomposition, 2D-VMD)与小波变换的,利用迁移学习再训练深度学习模型Inception V3的水下目标检测方法。

声呐图像的获取与标注十分困难且成本较高,而小样本的数据集用于深度学习训练时的局限性较大。针对此问题,联合迁移学习和深度学习的方法已取得较好的解决效果,在一定程度上解决了基于声呐图像的水下目标检测所面对的数据集不足、小目标检测精度低、效率较低等难题。

4 总结与展望

在目前国际形势严峻、经济发展高速、挑战与机遇并存的情况下,海洋资源的勘探与开发、海域军事部署、水下武器反爆破等民用与军事领域都需要水下目标检测技术提供支持。基于声呐图像的水下目标检测受到广泛关注,取得了诸多研究成果,同时也存在多项挑战。

在基于声呐图像的水下目标检测领域,传统方法可以利用声呐图像的数理统计特性、数学形态学处理、图像帧之间的像素差异等实现较有效的目标检测,但其精度与速度都存在很大的局限性;基于深度学习的目标检测由于其精度与速度上的显著优越性,已成为主流方法;但声呐图像相关数据集的不足又会限制深度学习的训练,为应对样本不足,迁移学习与深度学习被结合应用于声呐图像目标检测任务。现有技术可以有效地实现声呐图像的水下目标检测,但精度与速度都仍有较大提升空间,且存在受环境因素影响大、对声呐图像质量要求高、小目标漏检及误检率较高等问题。当前,我国在声呐图像的水下目标检测应用领域已走在世界前列,但在目标检测的模型构建和算法更新上原创性不足,理论支撑较薄弱,与国外仍存在客观差距。

水下环境复杂多变,声呐图像目标检测还存在亟待解决的难题,如小尺度目标的检测、实时检测等,未来可以对水下声呐图像的目标检测的发展趋势作出以下预测:1)由于水下小目标具有特征提取困难的特点,如何实现对其的精准检测是未来研究的重要课题;2)由于声呐图像在成像方式、图像特征、应用场景等方面具有区别于普通图像的特殊性,因此如何改进声呐图像的预处理对未来水下目标检测的发展具有重要意义;3)多种算法与模型的融合成为未来研究趋势,比如传统方法、深度学习方法、迁移学习方法的融合等;4)更轻量化的水下目标检测模型成为未来研究热点。

参考文献:

- [1] 贾宇. 关于海洋强国战略的思考[J]. 太平洋学报, 2018, 26(1): 1-8.
Jia Yu. On China's maritime power strategy[J]. Pacific Journal, 2018, 26(1): 1-8.
- [2] Bryner D, Huffer F, Srivastava A, et al. Underwater minefield detection in clutter data using spatial point-process models[J]. IEEE Journal of Oceanic Engineering, 2016, 41(3): 670-681.
- [3] Ceccarini C. The multi-role nato-interoperable light weight torpedo for the 21st century[J]. Information & Security an International Journal, 2004, 13(1): 89-97.
- [4] Chuang M C, Hwang J N, Ye J H, et al. Underwater fish tracking for moving cameras based on deformable multiple kernels[J]. IEEE Transactions on Systems, Man,

- and Cybernetics: Systems, 2016, 47(9): 1-11.
- [5] 张韞峰, 李娟, 黎明. 基于图像处理的水下海参识别和定位方法[J]. 水下无人系统学报, 2021, 29(1): 111-123.
Zhang Yunfeng, Li Juan, Li Ming. Underwater sea cucumber identification and localization method based on image processing[J]. Journal of Unmanned Undersea Systems, 2021, 29(1): 111-123.
 - [6] Byun S W, Kim J Y. Development and experiment of a hovering AUV tested-bed for underwater exploration[C]// Symposium on Underwater Technology & Workshop on Scientific Use of Submarine Cables & Related Technologies. Tokyo, Japan: IEEE, 2007.
 - [7] 李勇航, 牟泽霖, 万芃. 海洋侧扫声呐探测技术的现状及发展[J]. 通讯世界, 2015(3): 213-214.
 - [8] Hayes M P, Gough P T. Synthetic aperture sonar: A review of current status[J]. IEEE Journal of Oceanic Engineering, 2009, 34(3): 207-224.
 - [9] Komatsu T, Igarashi C, Tatsukawa K, et al. Use of multi-beam sonar to map seagrass beds in otsuchi bay on the Sanriku Coast of Japan[J]. Aquatic Living Resources, 2003, 16(3): 223-230.
 - [10] Lowe D G. Distinctive image features from scale-invariant key points[J]. International Journal of Computer Vision, 2003, 20: 91-110.
 - [11] Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features[C]//Proceedings of the 9th European Conference on Computer Vision-Volume Part I. Berlin, Heidelberg: Springer-Verlag, 2006.
 - [12] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//IEEE Computer Society Conference on Computer Vision & Pattern Recognition. San Diego, CA, USA: IEEE, 2005.
 - [13] Burges C. A Tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121-167.
 - [14] Rish I. An empirical study of the naive Bayes classifier[J]. Journal of Universal Computer Science, 2001, 1(2): 127.
 - [15] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60: 84-90.
 - [16] Lin T Y, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 2117-2125.
 - [17] 陈强, 田杰, 黄海宁, 等. 基于统计和纹理特征的SAS图像SVM分割研究[J]. 仪器仪表学报, 2013, 34(6): 214-221.
Chen Qiang, Tian Jie, Huang Haining, et al. Study on SAS image segmentation using SVM based on statistical and texture features[J]. Chinese Journal of Scientific Instrument, 2013, 34(6): 214-221.
 - [18] 王涛, 潘国富, 张济博. 基于侧扫声呐图像纹理特征的海底底质分类研究[C]//2020中国西部声学学术交流会论文集. 酒泉: 2020中国西部声学学术交流会, 2020: 403-406.
 - [19] 董凌宇, 单瑞, 刘慧敏, 等. 基于分形纹理特征的侧扫声呐图像沉船识别方法研究[J]. 海洋地质与第四纪地质, 2021, 41(4): 232-239.
Dong Lingyu, Shan Rui, Liu Huimin, et al. Shipwreck identification with side scan sonar image based on fractal texture[J]. Marine Geology & Quaternary Geology, 2021, 41(4): 232-239.
 - [20] 王其林, 王宏健, 李庆, 等. 侧扫声呐图像特征提取改进方法[J]. 水下无人系统学报, 2019, 27(3): 297-304.
Wang Qilin, Wang Hongjian, Li Qing, et al. Improved feature extraction method for side scan sonar images[J]. Journal of Unmanned Undersea Systems, 2019, 27(3): 297-304.
 - [21] 田晓东, 刘忠, 周德超. 基于形状描述直方图的声呐图像目标识别算法[J]. 系统工程与电子技术, 2007(7): 1049-1052, 1212.
Tian Xiaodong, Liu Zhong, Zhou Dechao. Mine target recognition algorithm of sonar image[J]. Systems Engineering and Electronics, 2007(7): 1049-1052, 1212.
 - [22] 卢逢春, 张殿伦, 郭海涛. 基于属性直方图的图像分割方法及其在声呐图像分割中的应用[J]. 哈尔滨工程大学学报, 2002(3): 1-3, 19.
Lu Fengchun, Zhang Dianlun, Guo Haitao. Image segmentation based upon bounded histogram and its application to sonar image segmentation[J]. Journal of Harbin Engineering University, 2002(3): 1-3, 19.
 - [23] Yang F, Du Z, Wu Z. Object recognizing on sonar image based on histogram and geometric feature[J]. Marine Science Bulletin, 2006, 25(5): 64.
 - [24] 王晓, 邹泽伟, 李勃勃, 等. 基于多特征融合的彩色图像声呐目标检测[J]. 计算机科学, 2019, 46(S1): 177-181.
 - [25] Sun C, Wang L, Wang N, et al. Image recognition technology in texture identification of marine sediment sonar image[J]. Complexity, 2021, 2021(2): 1-8.
 - [26] 罗进华, 蒋锦朋, 朱培民. 基于数学形态学的侧扫声呐图像轮廓自动提取[J]. 海洋学报, 2016, 38(5): 150-157.
Luo Jinhua, Jiang Jinpeng, Zhu Peimin. Automatic extraction of the side-scan sonar imagery outlines based on mathematical morphology[J]. Haiyang Xuebao, 2016, 38(5): 150-157.
 - [27] 邹岗, 田晓东, 刘忠. 基于几何特征的声呐图像人造目标检测算法[J]. 舰船科学技术, 2007, 29(6): 177-179, 183.
Zou Gang, Tian Xiaodong, Liu Zhong. Man-made target detection algorithm of sonar image based on geometric feature[J]. Ship Science and Technology, 2007, 29(6): 177-179, 183.
 - [28] Mallet J, Courmontagne P. A new wavelet thresholding approach for SAS images denoising[C]//Oceans, 2006. Boston, USA: IEEE, 2006: 1-6.

- [29] Isar A, Isar D, Moga S, et al. Multi-scale MAP despeckling of sonar images[C]//Oceans 2005-Europe. Brest, France: IEEE, 2005.
- [30] 金凤来, 钟何平. 一种改进的合成孔径声呐图像Lee滤波算法[J]. 舰船电子工程, 2017, 37(3): 35-37.
Jin Fenglai, Zhong Heping. An improved Lee filter algorithm for synthetic aperture sonar[J]. Ship Electronic Engineering, 2017, 37(3): 35-37.
- [31] 郭海涛, 徐雷, 赵红叶, 等. 一种抑制声呐图像散斑噪声的形态学滤波器[J]. 仪器仪表学报, 2015, 36(3): 654-660.
Guo Haitao, Xu Lei, Zhao Hongye, et al. A morphological filter for despeckling of a sonar image[J]. Chinese Journal of Scientific Instrument, 2015, 36(3): 654-660.
- [32] 崔杰, 胡长青, 徐海东. 基于帧差法的多波束前视声呐运动目标检测[J]. 仪器仪表学报, 2018, 39(2): 169-176.
Cui Jie, Hu Changqing, Xu Haidong. Moving target detection for multi-beam forward-looking sonar based on frame-difference method[J]. Chinese Journal of Scientific Instrument, 2018, 39(2): 169-176.
- [33] 崔杰, 胡长青, 徐海东, 等. 改进的Mean Shift算法在前视声呐运动目标跟踪中的应用[J]. 声学技术, 2020, 39(3): 279-283.
Cui Jie, Hu Changqing, Xu Haidong, et al. Application of improved mean shift algorithm in moving target tracking of forward-looking sonar[J]. Technical Acoustics, 2020, 39(3): 279-283.
- [34] Liu H, Dai J, Wang R, et al. Combining background subtraction and three-frame difference to detect moving object from underwater video[C]//Oceans-shanghai. Shanghai, China: IEEE, 2016.
- [35] Kalyan B, Balasuriya A. Sonar based automatic target detection scheme for underwater environments using CFAR techniques: A comparative study[C]//Proceedings of the 2004 International Symposium on Underwater Technology (IEEE Cat. No. 04EX869). Taipei: IEEE, 2005.
- [36] Li K, Liu Z, Hu S L, et al. Target detection of sonar image based on bis-parameter with adaptive windows[J]. Applied Mechanics & Materials, 2013, 347-350: 3407-3410.
- [37] Villar S A, Acosta G G, Solari F J. OS-CFAR process in 2-D for object segmentation from side-scan sonar data [C]//Information Processing & Control. Cordoba, Argentina: IEEE, 2016.
- [38] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014: 580-587.
- [39] Girshick R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1440-1448.
- [40] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [41] 曾文冠, 鲁建华. 基于卷积神经网络的声呐图像目标检测识别[C]//第十七届船舶水下噪声学术讨论会论文集. 衢州, 中国:《船舶力学》编辑部, 2019.
- [42] Fang J, Wang P. Target detection in sonar image based on Faster RCNN[C]//2020 International Conference on Information Science, Parallel and Distributed Systems (ISPDs). Xi'an, China: IEEE, 2020.
- [43] Ma Q, Jiang L, Yu W, et al. Training with noise adversarial network: A generalization method for object detection on sonar image[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Snowmass, USA: IEEE, 2020: 729-738.
- [44] 马麒麟. 基于深度学习的声呐图像目标检测算法研究[D]. 南京: 东南大学, 2020.
- [45] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems. [S.l.]: MIT Press, 2014.
- [46] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 779-788.
- [47] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]//Computer Vision ECCV 2016. European Conference on Computer Vision-Amsterdam. Cham: Springer, 2016, 9905: 21-37.
- [48] Wu Y. Sonar image target detection and recognition based on convolution neural network[J]. Mobile Information Systems, 2021, 2021(6): 1-8.
- [49] 王霞. 面向声呐图像的水下小目标识别方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2020.
- [50] Fan X, Lu L, Shi P, et al. A novel sonar target detection and classification algorithm[J]. Multimedia Tools and Applications, 2022, 81(7): 10091-10106.
- [51] Han K, Wang Y, Chen H, et al. A survey on vision transformer[C]//IEEE Transactions on Pattern Analysis and Machine Intelligence. Piscataway, USA: IEEE, 2022.
- [52] Yu Y, Zhao J, Gong Q, et al. Real-time underwater maritime object detection in side-scan sonar images based on transformer-YOLOV5[J]. Remote Sensing, 2021, 13(18): 3555.
- [53] 凡志邈, 李海林, 夏伟杰, 等. 基于深度学习的成像声呐水下目标的检测与分类[C]//中国声学学会水声学分会2019年学术会议论文集. 南京, 中国: 中国声学学会水声学分会, 2019.
- [54] 李宝奇, 黄海宁, 刘纪元, 等. 基于改进SSD的合成孔径声呐图像水下多尺度目标轻量化检测模型[J]. 电子与信息学报, 2021, 43(10): 2854-62.
Li Baoqi, Huang Haining, Liu Jiyuan, et al. Synthetic aperture sonar underwater multi-scale target efficient detection model based on improved single shot detector[J].

- Journal of Electronics & Information Technology, 2021, 43(10): 2854-62.
- [55] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//European Conference on Computer Vision. Cham: Springer, 2020: 213-229.
- [56] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
- [57] 汤寓麟, 李厚朴, 张卫东, 等. 侧扫声纳检测沉船目标的轻量化DETR-YOLO法[J]. 系统工程与电子技术, 2022, 44(8): 2427-2436.
- Tang Yulin, Li Houpu, Zhang Weidong, et al. Light-weight DETR-YOLO method for detecting shipwreck target in side-scan sonar[J]. Systems Engineering and Electronics, 2022, 44(8): 2427-2436.
- [58] Patel V M, Gopalan R, Li R, et al. Visual domain adaptation: A survey of recent advances[J]. Signal Processing Magazine, IEEE, 2015, 32(3): 53-69.
- [59] 朱兆彤, 付学志, 胡友峰. 一种利用迁移学习训练卷积神经网络的声呐图像识别方法[J]. 水下无人系统学报, 2020, 28(1): 89-96.
- Zhu Zhaotong, Fu Xuezhi, Hu Youfeng. A sonar image recognition method based on convolutional neural network trained through transfer learning[J]. Journal of Unmanned Undersea Systems, 2020, 28(1): 89-96.
- [60] 武铄, 王晓, 张丹阳, 等. 联合迁移学习和深度学习的侧扫声呐沉船识别方法[J]. 河南科技, 2021, 40(36): 36-40.
- [61] 于森. 基于深度学习的侧扫声呐图像目标检测方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2020.
- [62] 盛子旗, 霍冠英. 样本仿真结合迁移学习的声呐图像水雷检测[J]. 智能系统学报, 2021, 16(2): 385-392.
- Sheng Ziqi, Huo Guanying. Detection of underwater mine target in side scan sonar image based on sample simulation and transfer learning[J]. CAAI Transactions on Intelligent Systems, 2021, 16(2): 385-392.
- [63] Tang Y, Jin S, Bian G, et al. Shipwreck target recognition in side-scan sonar images by improved YOLOv3 model based on transfer learning[J]. IEEE Access, 2020, 8: 173450-173460.
- [64] 付同强, 胡桥, 刘钰, 等. 基于优化二维变分模态分解与迁移学习的水下目标识别方法[J]. 水下无人系统学报, 2021, 29(2): 153-163.
- Fu Tongqiang, Hu Qiao, Liu Yu, et al. Underwater target identification method based on optimized 2D variational mode decomposition and transfer learning[J]. Journal of Unmanned Undersea Systems, 2021, 29(2): 153-163.

(责任编辑: 许妍)

(上接第 332 页)

分析的基础上, 建立线导鱼雷出管纵向初始运动模型, 并进行计算机仿真, 分别针对不同横舵管制舵角和不同出管速度对鱼雷初始弹道以及航行姿态的影响进行了定量分析, 结果表明: 通过选取适当的 δ_e ($\delta_e=1^\circ$) 以及增大 v_0 , 可使线导鱼雷获得安全的管外激活初始弹道以及良好的姿态控制, 线导鱼雷管外激活发射方式是可行的。文中结论是在对鱼雷动力学方程进行简化的基础上研究得到的, 在实际中还存在若干干扰因素, 如由发射艇体引起的鱼雷流体动力干扰等, 对此还需开展进一步研究工作。

参考文献:

- [1] 丁浩, 王德石. 电力鱼雷动力电池管外激活发射弹道研究[J]. 鱼雷技术, 2009, 17(3): 52-55.
- [2] 文钢, 孙炯, 王国祥. 大型电力鱼雷动力电池管外激活发射对弹道的影响[J]. 海军工程大学学报, 2002, 14(2): 66-70.
- [3] 刘凯, 孙炯, 胡平. 管制舵角对某型鱼雷初始弹道的影响分析[J]. 海军工程大学学报, 2021, 33(2): 103-106.
- [4] 刘凯, 孙炯, 文钢. 鱼雷衡重特性对初始弹道的影响分析[J]. 海军工程大学学报, 2009, 21(6): 76-79.
- [5] 李宗吉, 程善政, 丁浩. 不同发射参数下潜射鱼雷初始弹道建模与仿真[J]. 舰船科学技术, 2011, 33(9): 78-80.
- Li Zongji, Cheng Shanzheng, Ding Hao. Research on initial trajectory of torpedo launched by submarine[J]. Ship Science and Technology, 2011, 33(9): 78-80.
- [6] 周杰, 王树宗. 无人水下航行器发射鱼雷初始弹道分析[J]. 弹道学报, 2009, 21(3): 86-89.
- [7] 文钢, 丁永忠, 玄兆林. Matlab/Simulink 在空投水雷水下弹道仿真中的应用[J]. 鱼雷技术, 2006, 14(3): 48-51.
- [8] 张岳青, 蔡卫军, 李建辰. 鱼雷斜入水过程非定常运动仿真研究[J]. 船舶力学, 2019, 23(1): 20-28.
- [9] 周德善. 鱼雷自导技术[M]. 北京: 国防工业出版社, 2009.
- [10] 潘光. 鱼雷力学[M]. 西安: 陕西师范大学出版总社有限公司, 2013.
- [11] 严卫生. 鱼雷航行力学[M]. 西安: 西北工业大学出版社, 2005.

(责任编辑: 许妍)