

【引用格式】陈悦, 罗逸豪, 李锦. 基于多模态感知的水下目标检测应用构想[J]. 数字海洋与水下攻防, 2024, 7(3): 334-341.

基于多模态感知的水下目标检测应用构想

陈悦^{1,2}, 罗逸豪^{1,2}, 李锦³

- (1. 清江创新中心, 湖北 武汉 430200;
2. 中国船舶集团有限公司第七一〇研究所, 湖北 宜昌 443003;
3. 海军工程大学 基础部, 湖北 武汉 430033)

摘要 水下目标检测在海洋生物研究、考古探索、军事防御等多领域广泛应用, 随着人工智能快速发展, 水下目标检测也朝着无人化、智能化发展。深度学习采用神经网络挖掘信息特征, 在速度和精度上均表现出优异的性能, 成为了计算机视觉技术的主流算法, 然而水下环境复杂, 将其应用于水下图像目标检测仍存在较大的挑战。水下目标各模态信息互补, 特征丰富, 有利于目标检测识别, 因此结合应用场景调研现有技术, 然后设计基于深度学习的多模态水下目标检测系统, 同时对比分析了现有关键技术的优缺点, 最后对多模态目标检测系统未来发展进行总结与展望, 具有重要意义。

关键词 水下目标检测; 深度学习; 多模态

中图分类号 TP391.4

文献标识码 A

文章编号 2096-5753(2024)03-0334-08

DOI 10.19838/j.issn.2096-5753.2024.03.012

Research on Application of Underwater Object Detection Based on Multimodal Perception

CHEN Yue^{1,2}, LUO Yihao^{1,2}, LI Jin³

- (1. Qingjiang Innovation Center, Wuhan 430200, China;
2. No. 710 R&D Institute, CSSC, Yichang 443003, China;
3. Department of Basic Courses, Naval University of Engineering, Wuhan 430033, China)

Abstract Underwater object detection has been widely applied in various fields, such as marine biology research, archaeological exploration, and military defense. With the rapid development of artificial intelligence, underwater object detection has also become unmanned and intelligent. Deep learning uses neural networks to mine information features, demonstrating excellent performance in both speed and accuracy, and has become the mainstream algorithm in computer vision technology. However, in complex underwater environments, there are still significant challenges in applying it to underwater image object detection. The complementary information and rich features of various modalities of underwater targets are beneficial for target detection and recognition. Therefore, this article combines application scenarios to investigate existing technologies, and then designs a multimodal underwater target detection system based on deep learning. At the same time, the advantages and disadvantages of existing core technologies are compared and analyzed. Finally, a summary and outlook on the future development of multimodal object detection systems are carried out, which is of great significance.

Key words underwater object detection; deep learning; multimodal

收稿日期: 2024-03-11

作者简介: 陈悦 (1997-), 女, 硕士, 助理工程师, 主要从事深度学习、计算机视觉方向研究。

0 引言

随着国际形势变化和海洋强国战略的提出,我国对海上安防尤为重视。漂雷、锚雷、沉底雷等水下武器凭借着威力大、隐蔽性好、易布难除的优势,常用于基地防御、航道封锁、抗登陆^[1-3]。反水雷装备及技术是降低水雷威胁的关键,也是当前各国海军的研发焦点。

传统反水雷作战需要作战人员涉入雷区操纵反水雷装备,危险性高,效率低。以人工智能(Artificial Intelligence, AI)为代表的新兴技术在快速发展,逐渐应用在各种反水雷装备^[4-5]。猎扫雷舰艇、无人水面艇(Unmanned Surface Vehicle, USV)、无人水下航行器(Underwater Unmanned Vehicle, UUV)等平台,研发无人化远程遥控反水雷系统,可提前侦察敌方水雷武器等,具有强机动性、无伤亡、易隐身等优点,极大地提高了反水雷作战效率^[6-9]。

目标检测技术是反水雷无人系统探测水下目标的关键技术之一,为舰船安全航行提供有力保障,因此其相关研究迫在眉睫。然而海底环境复杂,锚雷、沉底雷等固定式水下目标形态小,水下探测难度系数高^[10]。目前大多研究集中于前视声呐、侧扫声呐等单一设备开展水下小目标探测^[11]。BALTRU等^[12]学者研究认为,单模态描述事物具有相对独立性,而多模态数据表述事物能起到互补作用,使其呈现信息更立体、更全面。

近年来,多模态融合感知技术应用在计算机视觉、自然语言处理和生物医学等前沿方向取得了重要突破。为了进一步提高反水雷无人系统探测水雷等水下目标的性能,本文研究了多模态融合感知技术及水下目标探测技术现状,提出了基于深度学习的多模态水下目标检测系统,提升水下目标定位精度和分类识别能力,减少水下虚假目标数量。

1 多模态融合感知现状

模态是从某特定角度对事物进行描述的一种表现形式,获取目标特征有限。多模态从多个视角描述事物,各模态信息互补,获取特征更加全面。近年来,深度学习在文本识别、图像处理、音频识

别等感知领域表现优秀^[13],采用多模态深度学习能进一步提高系统感知能力。

多模态深度学习已应用于智能医学^[14-15]、智能驾驶^[16-17]、智慧办公^[18-19]、公共安全^[20-21]等多个领域。2000年,在机器翻译领域,人类融合语音和视觉信息进行多模态学习,理解语义^[22]。2017年,在航天领域,LIN等^[23]研究太阳射电频谱时,将不同频率信道捕获的太阳射电频谱作为不同的模态,学习其多样性。2019年,在智能驾驶领域,常昕^[24]提出了一种激光雷达与相机融合的目标检测及跟踪算法,相对单模态算法提高了系统鲁棒性。自动驾驶的先驱 Point Painting^[25]在2020年验证了融合点云数据和对应投影区域的图像分割网络输出信息的有效性,提高了目标识别性能。而后,程腾等^[26]提出从不同层次融合多模态特征策略,充分挖掘了目标不同模态、不同层次的特征。2021年,在医学领域,DUAN等^[27]研究的医学多模态分割方法 HyperdenseNet,并行使用2个全卷积网络处理不同模态的图像,有效提高模型对病变区域的感知能力。同年,杨舜翔^[28]基于多模态影像数据,采用卷积神经网络进行疾病早筛,提高了医学任务识别的准确率,对辅助医生诊断决策具有重要的意义。2021年,在公共安全领域的行人识别任务中,CHEN等^[29]提出了基于特征搜索策略的神经网络,融合可见光和红外成像2种模态,特征互补,提高识别行人的准确度。多模态深度学习其研究成果在农业、医学、军事、安防等诸多场景都具有重要的应用前景^[30]。

多模态深度学习相较于单模态深度学习能有效挖掘单模态缺失特征,应用于多个领域,然而水下目标数据模态差异大、易受到噪声干扰、信号采集难等问题,导致目前针对水下多模态目标探测的研究较少。

2 水下目标探测现状

目前水下目标探测方式多样,其中声探测应用最广泛,其他探测方式,如光探测、磁探测也备受关注。

2.1 声探测

声探测利用声波远距离传播和强穿透力的

特点在浑浊水域形成的声呐图像发现目标。基于声呐图像的水下目标检测可实现对水下目标定位分类功能,有效提取目标特征。声呐图像相较于光学图像,具有更特殊的纹理、边缘、矩阵和统计特征等。

基于声呐图像的水下目标检测的传统方法主要有数学形态学、数理统计的目标检测^[31]。近年来,深度学习不断发展,相比传统目标检测算法在速度和精度上表现更加优秀,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^[32]。20 世纪初期,RAO 等^[33]分析侧扫声呐的回波信号,转换获得图形符号,基于此检测水下目标。英国海军发明了水下探测声呐,用于探测水雷、潜水艇等水下目标^[34]。20 世纪 60 年代,美国 Raytheon 公司开始研究合成孔径技术更好地检测海水中的小目标。美国的轻型宽带变深声呐具有抗海流干扰能力,已应用至 DDG1000^[35]。2021 年,英国 Wavefront 公司开发的 Viglant 前视声呐可装载于 UUV 和 USV,具备水下自动探测能力。在国内的研究中,21 世纪初,中科院声学研究所研制出了高分辨率测深侧扫声呐实用样机,有效提高了侧扫声呐的测深精度,能适应更复杂的地形^[36]。

声波在水体中衰减幅度相对较小,能够传播距离较远,不受水体可视度影响,信号损失低,水下目标声探测手段是当前进行水下探测最主要、最有效的技术手段^[37]。

2.2 磁探测

磁探测技术目前是搜寻水下铁磁性目标的有效技术之一^[38]。地球具备天然磁性,现代水下目标如航行器、水雷、潜艇等目标壳体都含铁磁性物质,受到地磁场的磁化作用,会相应影响地磁场的分布从而引起地磁异常。基于磁异信号的目标磁探测技术(Magnetic Anomaly Detection, MAD),检测多个位置传感器的磁通变化,判定目标的方位,寻找水下目标。

近年来,国内外都非常重视磁探测技术,多国海军舰船相继装备了各类型号、灵敏度的磁力仪用于探测水下目标。自 20 世纪 70 年代以来,美国开始关注磁力测量,美国海岸系统局为探测埋藏在水下 6 m 内的目标,建立了一套针对磁性目标探

测磁异常信号感测方法^[39]。加拿大海军专门建立了水下电磁信号研究机构,负责水下常规目标的电磁分布^[40]。我国的水下目标磁探测技术在不断进步,复杂情况下的磁异常解释推理、反演技术等方面研究具有创新性^[41]。2003 年,蒋敏志等^[42]研究了一种可探测水面水下的磁性物质被动磁定位技术,已应用于一种机动舰船物理场测量系统。2012 年,黄玉等^[43]利用磁矩信息解算载体位置,实现水下地磁连续定位。

水下磁探测将捕获的磁场信号物理量经数据处理转换为所需目标信息,实现对目标的非接触测量,然而对铁磁性物质种类识别的准确度还较低。

2.3 光探测

可见光成像探测目标在水面、陆地是常用高效识别定位目标的技术之一。然而光在水介质传播过程中,水的吸收和散射使得光发生严重衰减,极大限制了光在水中的传播距离^[44]。且水下介质分布不均匀,易导致成像质量差、特征模糊。研究表明:光在水中的传播衰减特性主要与水质及光的波长相关,波长范围在 470~580 nm 的蓝绿光在海水中衰减最小^[45]。因此蓝绿光在海水中相较其他光波能传播更远,在水下可将蓝绿激光照射到目标表面,高分辨率相机接收该目标回波进行成像,从而检测识别水下目标。

自 20 世纪 60 年代以来,水下光探测技术在军事应用方面飞速发展^[46]。目前多个国家进行了激光水下探测系统的研制,并将其应用到潜艇和水下机器人^[47]。1968 年,在美国 Syracuse 大学第 1 台激光海洋探测系统诞生,从此奠定了海洋激光探测技术的理论基础^[48]。20 世纪 90 年代,美军 Kaman 公司在海湾战争期间研制的机载激光探雷系统“魔灯”可实现浅水区域探雷^[49]。美国 Westinghouse 公司研发的 SM2000 型水下激光成像系统最大探测深度为 45 m^[50]。1993 年,俄罗斯研制了机载蓝绿激光潜艇探测系统“紫石英”,来搜索沿海潜艇和水雷^[51]。国内也开展了光探测水下目标的研究。2008 年华中科技大学成功研制了机载激光雷达海洋探测系统^[51]。2010 年,孙剑锋^[52]自主研制了一套单狭缝条纹管激光成像雷达,在水质较浑浊的黄海海域进行了若干次水下目标探测实验,实现了远

距离目标成像。

基于可见光图像的目标检测效果直观可靠, 光学图像成像粒度细且图像信息丰富, 但复杂的水下环境导致的色彩失真、图像模糊、对比度低等问题亟待解决。

3 多模态水下目标检测系统

目前水下目标探测方法丰富多样, 但海洋环境复杂多变, 受水流、光线、水质等多方面影响, 单模态水下目标探测各有优缺点, 如表 1 所示。

表 1 水下单模态探测优缺点对比

Table 1 Comparison of advantages and disadvantages of underwater single-mode detection

探测类型	优点	缺点
磁探测	扫描范围大, 扫描速度快	无法探查磁性较弱或无磁性物体, 精度较低, 易受背景磁场干扰。只能判断物体大致位置, 无法得到物体大小和轮廓等信息, 探测距离短
光探测	获得水下目标的真实图像, 提供外观色彩等信息丰富	对能见度要求较高, 受水质影响大, 探测距离短
声探测	精度较高, 对水体可视度无要求, 可扫描远距离物体, 可实时观察扫描情况	受水流和目标距离影响, 无法获取目标色彩, 难以扫描过于微小的物体

声探测探测距离更远, 抗干扰能力相对较好, 但其检测精度仍较低。由于每种模态的异构性决定了其存在不同的学习模型, 光探测形成的可见光图像, 可与声呐图像从不同方面感知目标特征, 磁探测可感知物体铁磁性质, 因此采用多模态水下目标检测方法可丰富获取的目标信息。

主流成像声呐分为前视声呐和侧扫声呐, 前视声呐使用常规主动探测原理, 具有探测距离远、低功耗的优点, 侧扫声呐具有分辨率良好、经济性好、图像连续的优点但受发射孔径限制, 存在声脉冲散射问题。合成孔径声呐 (Synthetic Aperture Sonar, SAS) 与普通侧扫声呐相比, 方位向分辨率更高, 不受声波频段影响, 在实际水下目标探测装备中可与前视声呐共同使用, 从不同视角探测水下目标。因此可采用前视声呐和合成孔径声呐作为 2 种模态探测目标。

多模态深度学习需将各模态进行融合、信息对齐, 数据和特征层进行融合难度及硬件要求较高, 因此采用决策融合。分别训练各模态模型, 对每种模态得到输出结果 (如分类或回归) 采用不同的规则 (如平均值融合、集成学习等规则) 进行决策融合, 判定目标类别及定位, 具体流程如图 1 所示。

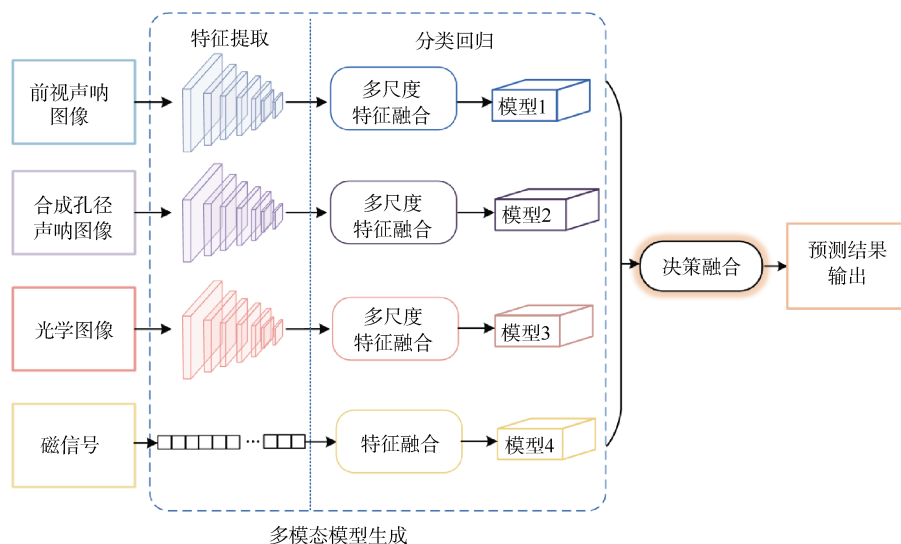


图 1 多模态水下目标检测流程

Fig. 1 Multimodal underwater object detection process

声模态提取目标纹理特征、几何特征、回波强度等, 光模态提取目标边缘特征、颜色特征、光照

特征等, 磁模态提取磁场强度、梯度、分布等信息。声光模态数据通过特征提取网络提取目标 3 级尺

度特征图，分别对应原始图像尺度的 1/8、1/16 和 1/32 的特征图，然后进行上采样、下采样获得对应输入尺度的 1/8、1/16 和 1/32、1/64 和 1/128 大小的特征图，低层特征图细节信息丰富，高层特征图语义信息丰富，各尺度特征层将顶层的语义特征融合至低层，不同尺度的目标需求匹配不同特征层进行检测，提高检测精度，最后通过损失函数反向更新权重获取模型。磁模态通过磁传感器，获取各方向的磁信号，进行信号处理、分解，提取磁异常数据，进行融合匹配，根据不同目标的几何特征判断磁场及梯度变化情况，预测水下目标类别并定位。综合各模型预测结果形成最终的预测，可较简单地处理数据的异步性。系统可随实际应用扩展其他模

态数据，各模态模型相互之间无干扰，若缺失某一模态依然能进行预测，维护系统稳定性。

多模态水下目标检测系统通过多传感器采集数据，对多源数据进行融合管理，建立样本数据库；然后采用深度学习目标检测算法，进行多模态特征融合，获取丰富特征，加强模型鲁棒性；通过训练网络，获取模型，将其部署，应用至客户端进行交互使用，系统结构图如图 2 所示，通过研究多源数据的信号级融合识别和任务级协同分析，解决二次识别目标匹配、同一目标多次重复标识问题。多模态水下目标检测系统联合利用多类传感器进行异构水下目标信息关联与融合处理，得到水下目标融合信息，更精确地探测、判别目标。

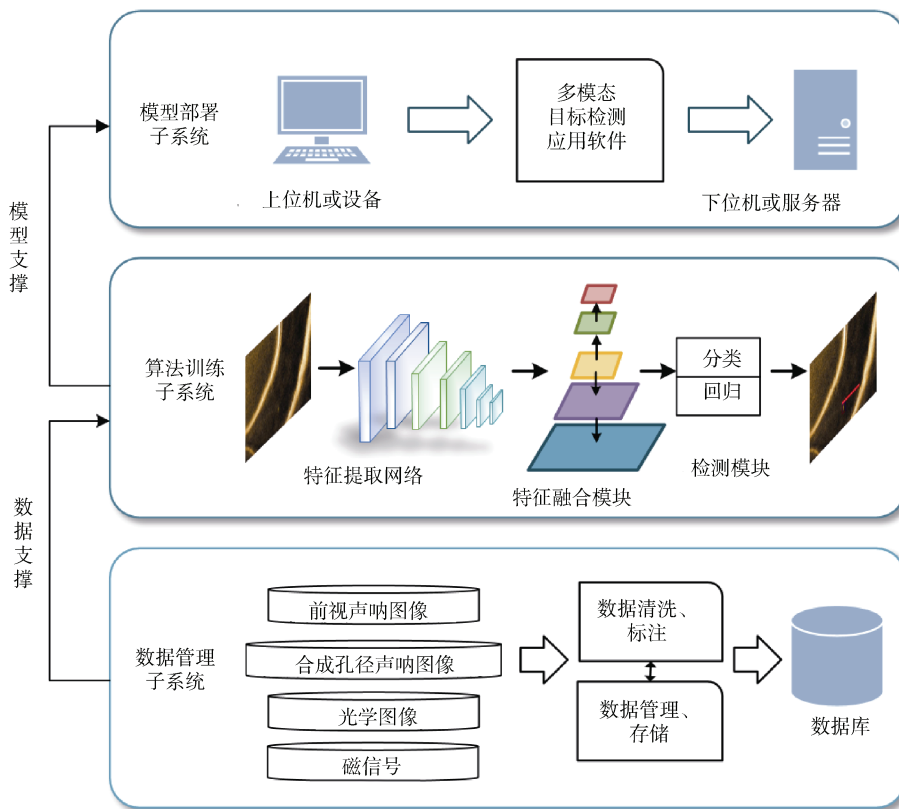


图 2 系统结构
Fig. 2 System structure

4 关键技术

4.1 数据处理

水下目标多模态数据采集和成像过程中易受外界环境影响，产生噪声，可能会影响数据处理的结果。对数据进行预处理能有效降低噪声干扰。目

前常见的去噪方法可归纳为滤波、偏微分、字典和稀疏表示、低秩矩阵和深度学习去噪方法等，优缺点对比如表 2 所示。用于单一图像、特定领域的去噪算法已经较为成熟，因此对多模态数据去噪可融合多种方法，结合不同算法的优势，获得更优的去噪效果。

表 2 数据去噪算法的优缺点对比

Table 2 Comparison of advantages and disadvantages of data denoising algorithms

去噪算法	优点	缺点
滤波去噪	空间域和变换域去噪, 方法简单, 易实现	去噪效果有限, 对图像的全局联系关注较少
偏微分去噪	能够有效平滑噪声并保留边缘信息	缺乏较严谨的理论论证, 计算成本较高
字典和稀疏表示去噪	特征提取、边缘细节保留效果好, 效率较高	自适应效果差
低秩矩阵去噪	挖掘全局信息, 有效恢复原始数据结构	灵活性不足; 训练样本不足时, 效果较差
深度学习去噪	自主挖掘深层特征信息, 具有良好去噪性能	依托大量数据, 计算代价较高

在样本数据较少的情况下, 为充分利用多模态数据可进行数据增强, 对原始数据集进行增广, 通过对现有的数据进行处理从而产出新数据, 扩张数据的可使用性。多模态领域内的数据增广方法总结如表 3 所示。

表 3 数据增广算法的优缺点对比

Table 3 Comparison of advantages and disadvantages of data augmentation algorithms

分类	常见算法	优点	缺点
几何变换	旋转、缩放、翻转、裁剪、填充、平移和仿射变换	简单易操作、信息真实	对数据重复记忆, 增加的信息量有限
光学变换	色彩变换、清晰度变换	简单易操作、色彩信息增加	增加的信息量有限, 可能会损失部分颜色信息
深度学习	对抗生成网络	学习源样本集的潜在分布	操作复杂, 硬件要求高

数据增广能够使得数据集涵盖更多的信息, 增强数据刻画模型的能力, 使得基于数据驱动的检测模型获取更强的检测性能。水下图像目标色彩特征较少, 多采用几何变换, 基于深度学习的数据增广网络软硬件要求较高, 但也是未来可观的使用方法。

4.2 深度学习目标检测

声光探测主要通过生成图像进行探测识别, 深度学习算法相较传统图像处理算法和机器学习算法更快更准, 因此水下目标探测可采用深度学习算法。

深度学习检测算法主要分为双阶段 (Two-stages) 和一阶段 (One-stage) 算法^[53]。双阶段目标检测算法基于区域提议网络 (Region Proposal Network, RPN) 分 2 步进行检测, 以 Faster-RCNN^[54]

为代表, 其检测精度较高, 但存在算法结构复杂, 推理时间长的缺点, 不适用于实时水下目标检测系统。单阶段目标检测算法不需要提取目标区域, 直接在特征图上生成预测框进行回归检测, 具有更好的实时性能, 如 YOLO 系列^[55-57]、SSD^[58]、RetinaNet^[59]算法等。但上述算法需根据先验知识设计调整预设锚定框的大小、纵横比和数量, 在多尺度目标检测中产生的锚定框数量大, 系统需计算大量超参数, 占用运行内存。而后 CornerNet^[60]、FCOS^[61]、ObjectBox^[62]等无锚定框的研究出现, 极大地优化了网络结构模型, 且具有较好的准确率和实时性。在水下目标探测应用中, 可根据采用无锚定框网络进行训练, 实现实时准确探测目标。

各模态获取传感器数据, 进行数据处理后, 通过深度学习神经网络提取主要特征, 并进行多尺度融合, 集成特征, 然后通过损失函数反向传播更新模型参数, 不断迭代训练, 最终得到检测模型, 将其部署在人机交互软件用于直接实时识别水下目标, 具体流程如图 3 所示。

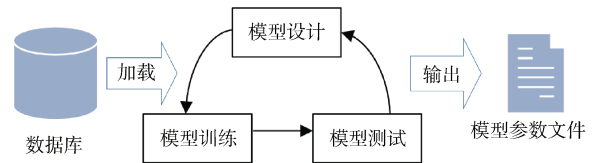


图 3 模型训练流程

Fig. 3 Model training process

5 结束语

水下目标检测技术在我国军事应用中有着巨大的发展前景, 本文结合现有技术, 提出了多模态感知水下目标检测系统, 通过多源传感器接收光探测数据和声呐探测数据, 获取目标多模态信息, 采用智能识别算法融合分析特征, 将定位识别结果发送给人机交互软件。本文所提出的多模态水下目标检测系统能有效提高准确率及抗干扰能力, 但存在一些可提升的改进空间, 总结如下。

1) 多模态目标检测的发展。

本文提出的多模态感知水下目标检测系统主要从多数据决策层面进行融合, 提供了不同的视角和信息, 有助于提高目标检测的准确性和鲁棒性, 避免了在中间特征上进行复杂的交互。然而要进一步丰富各模态语义信息需从特征层面进行融合, 可

采用早期融合、中期融合和混合融合等方法,但对于非同步采集的数据如何进行时间与空间配准,以及进行各模态特征数据对齐是当前的研究难点。

2) 提升模型推理速度。

当前主流目标检测算法精度相比较好,但检测速度未达最优,且多模态数据量大、计算复杂度高,如何适配更多硬件设备,提升检测速度是亟待解决的问题。后续研究需要通过模型结构轻量化、知识蒸馏和模型减枝等技术,减小模型参数量,提高算法实时性。

3) 国产化平台部署。

深度学习算法推理运算量大,训练数据多,对部署平台硬件要求较高,之前多采用英伟达、英特尔、高通等国外 AI 芯片。将模型转换至国产芯片通用模型,响应国家战略,实现自主可控,减少对外部供应链的依赖,对部署平台进行国产化升级具有积极意义。

参考文献

- [1] 张起. 新形势下的反水雷需求与对策[J]. 水雷战与舰船防护, 2015, 23 (2): 1-5.
- [2] 周穗华, 张晓兵, 蒋培, 等. 水雷总体技术[M]. 北京: 兵器工业出版社, 2015, 1-22.
- [3] 赵治平, 官红, 艾艳辉, 等. 无人化时代反水雷装备体系构想[J]. 数字海洋与水下攻防, 2018, 1 (1): 1-6.
- [4] 傅金祝. 美国海军反水雷能力和无人反水雷平台装备分析[J]. 现代舰船, 2012 (6): 44-47.
- [5] 张鹏, 何心怡, 林杨. 反水雷与 UxV[J]. 舰船科学技术, 2009, 31 (1): 29-33.
- [6] DJAPIC V, NAD D. Using collaborative autonomous vehicles in mine countermeasures[C]// OCEANS 2010-MTS/IEEE Seattle. Sydney: IEEE, 2010.
- [7] 刘平, 杨洋, 王青. 国外反水雷舰艇装备现状及发展趋势[J]. 船舶工程, 2004, 26 (6): 1-3.
- [8] 王进, 王慧. 反水雷无人艇的发展趋势分析[J]. 数字海洋与水下攻防, 2021, 4 (2): 113-116.
- [9] 王久法, 吴乔, 高频, 等. 国外无人反水雷特点及关键技术分析[J]. 数字海洋与水下攻防, 2020, 3 (5): 382-386.
- [10] 张晓鹏, 周利生. 掩埋小目标声探测技术研究[J]. 声学技术, 2012, 31 (1): 30-35.
- [11] COURMONTAGNE P. A new approach for mine detection in SAS imagery[C]// OCEANS 2008-MTS/IEEE Kobe Technology. Kobe: IEEE, 2008.
- [12] BALTRU T, AHUJA C, MORENCY L P. Multimodal machine learning: a survey and taxonomy[J]. Transactions of Pattern Analysis and Machine Intelligence, 2018, 41 (2): 423-443.
- [13] 刘建伟, 丁熙浩, 罗雄麟, 等. 多模态深度学习综述[J]. 计算机应用研究, 2020, 37 (6): 1601-1614.
- [14] ZHOU T X, RUAN S, CANU S. A review: deep learning for medical image segmentation using multi-modality fusion[J]. Array, 2019, 3-4: 100004.
- [15] ZHOU T X, RUAN S, VERA P. A tri-attention fusion guided multi-modal segmentation network[J]. Pattern Recognition, 2022, 124: 108417.
- [16] VORA S, LANG A H, HELOU B, et al. PointPainting: sequential fusion for 3D object detection[J]. Computer Vision and Pattern Recognition, 2020, 2019 (1): 10150.
- [17] PANG S, MORRIS D, RADHA H. Fast-CLOCs: fast camera-LiDAR object candidates fusion for 3D object detection[EB/OL]. [2022-03-08]. https://openaccess.thecvf.com/content/WACV2022/papers/Pang_Fast-CLOCs_Fast_Camera-LiDAR_Object_Candidates_Fusion_for_3D_Object_Detection_WACV_2022_paper.pdf.
- [18] KIM W, SON B, KIM I. ViLT: vision-and-language transformer without convolution or region supervision [EB/OL]. [2021-07-10]. <https://arxiv.org/abs/2102.03334>.
- [19] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[C]// 38th International Conference on Machine Learning. San Francisco: IEEE, 2021.
- [20] YE M, LAN X Y, LI J W, et al. Hierarchical discriminative learning for visible thermal person re-identification[C]// 32nd AAAI Conference on Artificial Intelligence. New Orleans: AAAI, 2018.
- [21] LU Y, WU Y, LIU B, et al. Cross-modality person re-identification with shared-specific feature transfer[EB/OL]. [2020-03-12]. <https://arxiv.org/abs/2002.12489>.
- [22] JANSEN B J, GOODRUM A, SPINK A. Searching for multimedia: analysis of audio, video and image Web queries[J]. World Wide Web: Internet and Web Information Systems, 2000, 3 (4): 249-254.
- [23] MA L, ZHUO C, XU L, et al. Multimodal deep learning for solar radio burst classification[J]. Pattern Recognition, 2017, 6 (1): 573-582.
- [24] 常昕, 陈晓冬, 张佳琛, 等. 基于激光雷达和相机信息融合的目标检测及跟踪[J]. 光电工程, 2019, 46 (7): 91-101.
- [25] VORA S, LANG A H, HELOU B, et al. PointPainting: sequential fusion for 3D object detection[EB/OL]. [2020-05-06]. <https://arxiv.org/abs/1911.10150>
- [26] 程腾, 孙磊, 侯登超, 等. 基于特征融合的多层次多

- 模态目标检测[J]. 汽车工程, 2021, 43(11): 1602-1610.
- [27] DUAN W T, ZHANG L, COLMAN J, et al. Multi-modal brain segmentation using hyper-fused convolutional neural network[C]// 4th International Workshop on Machine learning in Clinical Neuroimaging. Strasbourg: Springer, 2021.
- [28] 杨舜翔. 基于多模态医学影像的疾病早筛方法研究[D]. 济南: 山东大学, 2021.
- [29] CHEN Y, WAN L, LI Z H, et al. Neural feature search for RGB-infrared person re-identification[EB/OL]. [2021-06-06]. <https://arxiv.org/abs/2104.02366>.
- [30] 孙影影, 贾振堂, 朱昊宇, 等. 多模态深度学习综述[J]. 计算机工程与应用, 2020, 56(21): 1-10.
- [31] 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.
- [32] 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17届船舶水下噪声学术讨论会论文集. 衢州: 《船舶力学》编辑部, 2019.
- [33] RAO C, MUKHERJEE K, GUPTA S, et al. Underwater mine detection using symbolic pattern analysis of side scan sonar images[C]// 2009 American Control Conference. St. Louis: IEEE, 2009.
- [34] JEONG D M, SON K, LEE Y G, et al. Echo energy estimation in active sonar using fast independent component analysis [J]. Neural Information Processing, 2011, 5863: 381-388.
- [35] 杜召平, 陈刚, 王达, 等. 国外声呐技术发展综述[J]. 舰船科学技术, 2019, 41(1): 145-151.
- [36] 梁开龙. 水下地形测量[M]. 北京: 测绘出版社, 1995.
- [37] 康路遥, 夏开奇. 图像声呐技术现状及其在水下探测中的应用浅析[C]// 第10届中国国际救捞论坛论文集. 上海: 中国航海学会救捞专业委员会, 2018.
- [38] 张合, 江小华. 目标探测与识别技术[M]. 北京: 北京理工大学出版社, 2015.
- [39] 徐燕, 周昌剑. 磁技术在水下密封和探测中的应用[J]. 电子世界, 2014(10): 163-164.
- [40] 张明. 水下磁性目标探测系统研制[D]. 杭州: 杭州电子科技大学, 2017.
- [41] 卞光浪, 翟国君, 黄谟涛, 等. 利用剖面测线磁场数据反演三度磁性体参数[J]. 武汉大学学报: 信息科学版, 2012, 37(1): 91-95.
- [42] 蒋敏志, 林春生. 船舶水下被动磁性定位研究[J]. 武汉理工大学学报: 信息与管理工程版, 2003, 25(6): 168-170, 187.
- [43] 黄玉, 郝燕玲. 基于磁矩信息的水下地磁连续定位算法[J]. 上海交通大学学报, 2012, 46(3): 390-393.
- [44] 朱彩霞, 闫亚东, 沈满德, 等. 基于水下反恐的微光成像系统[J]. 应用光学, 2008, 29(2): 220-224.
- [45] 冯包根. 国外水下激光成像技术现状[J]. 红外与激光技术, 1995, 24(2): 11-16.
- [46] DUNTLEY S Q. Light in the sea[J]. Journal of the Optical Society of America, 1963, 52(1): 214-233.
- [47] 赵涛. 国外激光告警技术的设备与发展[J]. 舰船电子工程, 2009, 29(2): 27-29.
- [48] HULST H C. Light Scattering by Small Particles[M]. New York: Wiley, 1957.
- [49] 冯包根. 美国海军“魔灯”机载激光探雷系统[J]. 水雷战与战舰防护, 1998, 17(5): 30-32.
- [50] 刘雪明, 胡正荣. 同步扫描水下激光成像系统主要光学参数的理论分析[J]. 电子器件, 1997, 20(3): 47-51.
- [51] 李哲, 邓甲吴, 周卫. 水下激光探测技术及其进展[J]. 舰船电子工程, 2008, 28(12): 13-16, 53.
- [52] 孙剑峰, 郜键, 魏靖松, 等. 条纹管激光成像雷达水下探测成像研究进展[J]. 红外与激光工程, 2010, 39(5): 811-814.
- [53] 蒋弘毅, 王永娟, 康锦煜. 目标检测模型及其优化方法综述[J]. 自动化学报, 2021, 47(6): 1232-1255.
- [54] REN S Q, HE K M, 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.
- [55] 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. New York: IEEE, 2016.
- [56] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017.
- [57] REDMON J, FARHADI A. YOLOv3: an Incremental Improvement[EB/OL]. [2018-04-08]. <https://arxiv.org/pdf/1804.02767>.
- [58] LIU W, ANGELOV D, ERHAN D, et al. SSD: single shot multibox detector[EB/OL]. [2016-12-29]. <https://arxiv.org/abs/1512.02325>
- [59] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[EB/OL]. [2018-02-07]. <https://arxiv.org/abs/1708.02002>
- [60] LAW H, DENG J. CornerNet: detecting objects as paired keypoints[J]. International Journal of Computer Vision, 2020, 128(3): 642-656.
- [61] TIAN Z, SHEN C H, CHEN H, et al. FCOS: fully convolutional one-stage object detection[EB/OL]. [2019-08-20]. <https://arxiv.org/abs/1904.01355>.
- [62] ZAND M, ETEMAD A, GREENSPAN M. ObjectBox: from centers to boxes for anchor-free object detection[C]// 17th European Conference on Computer Vision. Tel Aviv: Springer, 2022.

(责任编辑: 曹晓霖)