

中图法分类号: U675.7 文献标识码: A 文章编号: 1006-8961(2026)05-1394-31

论文引用格式: Wang Y Y, Wu H, Qing Y H, Zhang W D, Shen L Q and Xu X. 2026. Progress in deep learning-driven intelligent perception and decision-making technologies for unmanned surface vehicles. Journal of Image and Graphics, 31(5): 1394-1424(王曰英, 吴浩, 庆雨豪, 张卫东, 沈礼权, 徐昕. 2026. 深度学习驱动的海上无人船智能感知与决策技术进展. 中国图象图形学报, 31(5): 1394-1424)[DOI: 10.11834/jig.250439]

## 深度学习驱动的海上无人船智能感知与决策技术进展

王曰英<sup>1\*</sup>, 吴浩<sup>1</sup>, 庆雨豪<sup>1</sup>, 张卫东<sup>2</sup>, 沈礼权<sup>3</sup>, 徐昕<sup>4</sup>

1. 上海大学机电工程与自动化学院, 上海 200444; 2. 河南科技学院信息工程学院, 新乡 453003;
3. 上海大学通信与信息工程学院, 上海 200444; 4. 国防科技大学智能科学学院, 长沙 410073

**摘要:** 近年来智能无人系统技术持续推动海上无人水面艇(unmanned surface vehicle, USV)感知与决策能力的提升, 涌现出诸多面向复杂海洋环境的创新研究与实践部署。本文首先系统梳理了USV的发展历程与体系架构, 分析其在船体设计、动力系统、通信控制与多传感器集成等方面的演进特征; 进而围绕智能感知这一核心环节, 重点综述了深度学习模型及多模态传感器融合在海上目标检测、障碍物识别、海况感知与多目标跟踪等任务中的应用进展, 结合典型海事视觉数据集探讨了算法在跨域泛化、实时性与环境鲁棒性方面面临的挑战; 进一步, 本文总结了基于感知的导航、制导与控制方法, 以及多船协同与群体智能在复杂动态海域中的研究现状与应用前景; 最后, 从恶劣海况下的感知稳健性、多模态融合机制、实时安全决策与分布式协同等角度, 展望了海上无人船智能技术未来发展的关键问题与研究方向。

**关键词:** 海上无人船(USV); 复杂海洋环境; 智能感知; 深度学习; 多传感器融合; 自主决策; 群体智能

### Progress in deep learning-driven intelligent perception and decision-making technologies for unmanned surface vehicles

Wang Yueying<sup>1\*</sup>, Wu Hao<sup>1</sup>, Qing Yuhao<sup>1</sup>, Zhang Weidong<sup>2</sup>, Shen Liqun<sup>3</sup>, Xu Xin<sup>4</sup>

1. School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China;
2. School of Information Engineering, Henan Institute of Science and Technology, Xinxiang 453003, China;
3. School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China;
4. College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China

**Abstract:** Unmanned surface vehicles (USVs) are rapidly emerging as one of the most promising unmanned maritime platforms, offering high autonomy, low operational costs, and long endurance, and they have demonstrated significant potential in missions such as hydrographic surveying, environmental monitoring, port security, search and rescue, and military reconnaissance. Despite these advantages, the ocean remains one of the most challenging and unpredictable operational environments owing to its dynamic and unstructured nature, where sudden illumination changes, wave-induced reflections, fog, rain, strong winds, adverse weather conditions, and limited communication links all introduce severe uncertainties for the perception and decision-making capabilities of USVs. Against this background, this paper presents a comprehensive

收稿日期: 2025-09-12; 修回日期: 2026-01-04; 预印本日期: 2026-01-11

\* 通信作者: 王曰英 yueyingwang@shu.edu.cn

基金项目: 国家自然科学基金项目(62421004, U24A20279, 62473243)

Supported by: National Natural Science Foundation of China(62421004, U24A20279, 62473243)

survey of intelligent perception and decision-making technologies for USVs under complex maritime environments, with a particular emphasis on deep learning, multimodal sensor fusion, and cooperative swarm intelligence. The development history and overall system architecture of USVs are first reviewed, highlighting the evolution of hull platforms, onboard sensors, navigation and control modules, and communication subsystems and clarifying the technological foundations that have shaped their current capabilities. In terms of perception, recent advances are categorized in accordance with algorithmic paradigms such as convolutional neural networks (CNNs), transformer-based architectures, graph neural networks (GNNs), and deep reinforcement learning (DRL). Their roles in object detection, obstacle recognition, semantic scene understanding, sea-state perception, and multitarget tracking are evaluated. Multimodal sensing and fusion, including the integration of optical cameras, LiDAR, millimeter-wave radar, sonar, and inertial navigation units, are further analyzed, with an emphasis on feature- and decision-level fusion strategies that enhance robustness under adverse conditions. Representative maritime vision datasets, such as MODD, MaStr1325, SMD, and SeaDronesSee, are introduced to reveal gaps in data diversity, annotation quality, and generalization capacity. In the decision-making domain, methods that transform perception outputs into safe and efficient navigation actions, ranging from traditional graph search and sampling-based planners to emerging artificial intelligence approaches such as imitation and reinforcement learning, are examined. Particular attention is given to the incorporation of environmental disturbances including wind, waves, and currents into dynamic planning, as well as the integration of international maritime rules such as COLREGs into collision avoidance strategies. Beyond individual vehicles, this survey extends to multi-USV cooperative frameworks that enable distributed task allocation, formation control, and swarm intelligence for collective perception and decision-making in large-scale missions. The analysis reveals that deep learning models have significantly improved perception accuracy in unstructured maritime scenarios compared with traditional rule-based or handcrafted feature methods, with CNNs excelling at feature extraction, transformers showing advantages in capturing global contextual dependencies in water-sky backgrounds, and GNNs and DRL extending perception into temporal and interaction-aware domains. Nonetheless, real-world deployment remains constrained by generalization bottlenecks, as models trained on specific datasets often fail under unseen weather, lighting, or sea conditions. Multimodal sensor fusion proves indispensable, given that cameras provide rich semantic information but degrade under poor visibility, radar and LiDAR offer reliable distance estimation but face issues of low resolution or cost, and sonar extends sensing underwater but poses challenges of alignment with surface sensors. While deep learning-based fusion techniques show promise, spatiotemporal misalignment under dynamic vessel motion continues to hinder robust performance. On the decision-making side, reinforcement learning approaches outperform classical planners in dynamic environments, especially for real-time obstacle avoidance and long-horizon trajectory optimization. However, their interpretability limitations and high sample complexity remain problematic for safety-critical applications, motivating research on hybrid frameworks that integrate domain knowledge such as COLREGs rules into learning processes. Multi-USV collaboration is recognized as a frontier for scaling maritime autonomy, with swarms offering distributed sensing, information sharing, and coordinated mission execution, yet unresolved challenges in communication latency, distributed consensus, and scalability restrict practical adoption. Overall, the findings of this survey highlight persistent gaps in dataset availability, computational efficiency on edge hardware, interpretability of deep models, and standardization of benchmarks, all of which restrict the path toward widespread operational deployment. In light of these findings, several future research directions are proposed. They include enhancing perception robustness under extreme conditions via domain adaptation and physics-informed learning; advancing cross-modal and domain fusion methods to integrate surface, subsurface, and aerial data; designing lightweight real-time models with embedded safety constraints and explainability for onboard decision-making; exploring distributed collaborative intelligence for scalable swarm autonomy; and establishing standardized open benchmarks for fair evaluation and industrial adoption. By systematically reviewing current progress, challenges, and prospects, this survey provides a structured reference for researchers and technical insights for practitioners, aiming to accelerate the development of intelligent, safe, and resilient USVs capable of autonomous operation in complex and uncertain ocean environments.

**Key words:** unmanned surface vehicle (USV); complex maritime environment; intelligent perception; deep learning; multimodal sensor fusion; autonomous decision-making; swarm intelligence

## 0 引言

随着海洋资源开发、海上安全保障和海洋科学研究需求的日益增长,人类对海洋的探索和利用正在向更深、更远、更智能的方向发展。海洋作为地球表面最大的生态系统,其复杂性和不可预测性给传统海洋作业带来了巨大挑战。在这一背景下,无人化、智能化的海洋装备技术成为突破海洋环境限制、提升海洋作业效率的关键途径。

海上无人船(unmanned surface vehicle, USV)又称自主水面艇(autonomous surface vehicle, ASV),作为无人海洋载体(unmanned marine vehicle, UMV)的一种重要类型(Ab Rahman 等, 2017),近年来在科研、国防与商业领域的关注度持续攀升。这类平台凭借无需人员登乘、长续航、低运营成本以及在高风险海域执行任务的能力,正在成为未来海洋作业体系的重要组成部分(Liu 等, 2016)。其应用范围覆盖了海洋测绘与资源勘探、海上安全(Manley, 2008)与执法、环境监测与灾害预警、航道测深与维护、通信中继以及军事情报侦察等多个方面(Wibisono 等, 2023)。例如,在水文与地质勘探中,USV 可搭载多波束测深仪、侧扫声呐及水质传感器进行自动化大面积测绘(Zolich 等, 2019);在海上交通管理中,USV 可执行自主巡逻和目标识别任务,提升海域态势感知能力(Manley 等, 2004);在军事应用中,其低可探测性与可编队作业能力可有效扩展舰队作战半径(Wu 等, 2024)。

USV 技术的快速发展得益于多学科交叉的推动。从船舶与海洋工程提供的船型设计与动力推进技术,到自动控制与人工智能提供的导航、制导与控制(navigation, guidance, and control, NGC)方法,再到传感器与通信技术支持的感知与数据交换,USV 已逐渐从单一功能向多任务、多传感器融合、高自主性的智能系统演进。在这一过程中,系统设计者不仅要关注水动力性能、能源利用与航行稳定性,还需要在软硬件架构上满足实时数据处理、环境建模与自主决策等复杂需求。

然而,复杂的海洋环境仍然是制约 USV 可靠运行的重要因素。复杂海洋环境是指由多种动态因素相互作用形成的具有高度不确定性、时变性和多尺度特征的海域环境,主要包括:海况变化(波浪高度、

周期、方向)、气象条件(风速、降雨、能见度、温度)、水文特征(海流、潮汐、盐度)、地理环境(水深、海底地形、海岸线)以及人为干扰(船舶交通、海洋工程、渔业活动)等要素。如图 1 所示,首先,海况的变化会引起船体姿态剧烈波动,影响传感器稳定性与定位精度;其次,恶劣气象(如强风、降雨、雾霾)会导致光学与毫米波传感器性能下降,甚至造成关键任务中断(Ahmed 等, 2023)。再者,远海或极地等特殊区域的通信条件受限,可能出现长时间链路中断或带宽不足,从而影响远程操控与协同任务。此外,海洋环境的多变性和不可预测性还会对任务规划与风险评估提出更高要求,如突发漂浮物、动态目标与水下地形变化等。

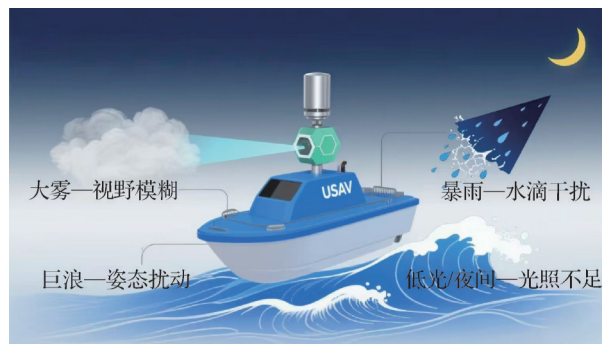


图1 复杂海洋环境图

Fig. 1 Complex marine environment diagram

在应对这些挑战的过程中,人工智能、先进传感器、通信与导航技术的融合发挥了重要作用。人工智能(尤其是深度学习)在目标检测、语义分割、路径规划与行为预测等任务中展现了强大的特征提取与决策能力。多传感器融合技术可以结合光学、声学、雷达与惯性导航数据,实现更稳健的环境感知与状态估计。通信与导航方面,高精度全球导航卫星系统(global navigation satellite system, GNSS)、差分定位(real time kinematic-GPS, RTK-GPS)、惯性导航系统(inertial navigation system, INS)以及新一代低轨卫星通信为 USV 在远距离、高动态与弱信号环境下的定位和数据传输提供了可靠支撑。

特别是近几年,海事视觉数据集与开源基准的出现为智能感知算法的开发与评估提供了良好条件。例如,SeaShips 数据集针对船舶检测任务提供了高分辨率图像与精确标注,MODS (maritime obstacle detection and segmentation dataset)数据集结

合检测与障碍物分割任务以适配USV场景,WSODD (weakly supervised object detection in the dark dataset)数据集则探索弱监督检测在海事感知中的潜力。这些数据集推动了深度学习模型在实际海洋场景下的泛化与鲁棒性研究,也为USV感知系统的工程落地提供了实验基础(Xue等,2021)。

尽管USV在多个应用领域展现了巨大的潜力,但其在智能感知、协同作业及决策规划方面仍面临显著瓶颈(Lu等,2022)。首先,感知层面的不确定性问题尤为突出。现有视觉和传感器系统在高海况、低光照、雨雾天气下的鲁棒性依然不足,容易出现目标漏检、误检和跟踪丢失等问题。复杂海洋背景中,波浪、反光和漂浮物等干扰因素会造成传统特征提取方法失效,即使是深度学习模型在跨域场景中也存在性能显著下降的风险。其次,在协同作业方面,多船编队及异构平台的任务分配与信息共享效率仍有待提高。海洋环境下通信链路的时延与带宽限制会直接影响协同感知和联合决策的实时性,而现有的分布式控制与多智能体系统(multi-agent system, MAS)算法在动态组网、抗通信中断等方面的研究尚不完善(Wu等,2024)。

决策规划同样存在挑战。在不确定性和高动态性的环境中,传统基于规则或确定性模型的路径规划与避碰算法难以适应实时变化的海况与任务需求。此外,不同海域的航行规则、避碰标准和安全约束具有差异性,导致算法的可迁移性和适用性受到限制(Hashali等,2024)。

针对上述问题,深度学习、传感器融合及多船协作被普遍视为潜在突破口。深度卷积网络、Transformer、图神经网络(graph neural network, GNN)等模型在多模态特征提取与任务决策中表现出良好的适应性与泛化能力,能够从大规模、多源数据中学习复杂的时空模式。传感器融合则通过光学、雷达、声呐及惯性数据的互补性提高感知精度和鲁棒性,尤其是在单一传感器失效或数据丢失情况下(Guan等,2023)。在协同控制方面,结合分布式深度强化学习(deep reinforcement learning, DRL)与群体智能优化方法,可在动态通信网络中实现多平台的高效协调与任务分解(Zheng等,2024)。从更广的视角看,这些研究不仅推动了无人船的智能化发展,还对海洋工程、海洋资源开发及海上安全保障等战略领域具有深远意义。一方面,提升USV在极端条件下

的自主作业能力,将显著增强人类对远海、极地和深海的探索与利用效率;另一方面,可大规模部署的智能无人船集群,将成为未来智慧海洋和智能航运的重要基础设施。

鉴于当前USV研究在应用推广中面临的挑战,本文旨在整合无人海洋载体(UMV)的发展历程、视觉数据集与深度学习方法、导航—制导—控制系统的关键技术及多平台协同作业策略,形成一份面向复杂海洋环境的系统性综述。具体如下:

1)USV发展历程与体系架构。在回顾USV从军事起源到商业与科研拓展的演进历史基础上,分类讨论了船型、任务类型及系统组成(如船体动力、通信控制、传感器和NGC(navigation, guidance and control)软件),并明确了复杂环境中USV运行所需的关键环节与任务导向。

2)感知需求与任务场景。围绕海面目标检测与识别、海况感知与障碍物识别、多目标跟踪与动态环境建模等核心任务,分析了传感器类型(自感、外感、多融合)的特性与挑战,揭示了多模态数据在提升鲁棒性方面的作用。

3)视觉数据集与感知算法基础。梳理了USV视觉数据集的现状、发展趋势与任务类别(如检测、分割、深度估计),并对比了传统计算机视觉方法的局限性与深度学习在检测、分割、跟踪中的优势,强调数据集在算法泛化与评测中的驱动作用。

4)深度学习在USV控制系统中的应用。聚焦导航系统(环境感知、状态估计、多传感器融合定位)、制导与路径规划(全局/局部规划、数据驱动趋势、规则约束避碰、“学习+规则”融合),并总结未来趋势。

通过上述结构,本文回顾了基于深度学习的USV在智能感知与决策领域的技术现状,系统整合了USV从发展历程到多模态感知、深度学习模型应用以及导航—决策—控制的全链路分析,强调深度学习“学习+规则”融合范式在避碰与协同作业中的工程化价值,同时,深度剖析了视觉数据集的现状与不足,提出基于深度学习的任务驱动评测与改进方向;此外,还扩展到多平台群体智能与跨域泛化挑战,提供更全面的未来展望,包括标准化基准建设和物理信息学习的潜力,从而为复杂海洋环境下的自主USV体系构建提供独特的技术参考与启示。

## 1 海上无人船发展历程与体系架构

### 1.1 无人船发展简史

海上无人船(USV)作为无人海洋系统的重要组成部分,能够使用自身控制系统进行导航并执行任务。USV具有相对较小尺寸和成本效益高的特点,广泛应用于大型海洋表面区域。由于人类在执行任务时几乎独立工作且协调性较差,USV越来越多地被部署在需要长时间巡逻的危险区域任务中(Jung等,2018;Breivik,2010;Breivik等,2008;Roberts和Sutton,2006)。在USV发展方面,多个USV项目通过公共和私人规划在各个领域和目的下进行开发(Marichal等,2001;Calderón等,2014),USV正在科学研究、环境任务、海洋勘探、军事应用以及其他应用(如交通、通信中继和燃料供应)中商业化使用(Liu等,2016)。

#### 1.1.1 军事起源与二战应用

USV最早出现在第二次世界大战期间,但直到20世纪90年代,其使用才真正扩展到执行国家项目(Corfield和Young,2006)。第一艘USV是1944年在加拿大开发的“Comox”(Agarwala,2022)。Comox是为诺曼底行动开发的,代表飞机执行烟雾幕操作,负责支持扫雷和火箭艇摧毁;虽然实验成功,但未实际使用。大约在同一时期,美国开发并演示了配备扫雷火箭的USV,即“Porcupine”、“Bob-Sled”和“Woo-fus 120”,用于清除区域内的水雷和障碍物。1946年,Able和Baker开始使用无人机船收集原子弹爆炸后的辐射样本。20世纪60年代,USV被开发出来并用做导弹发射训练的目标(Batu等,2013;Dağlı等,2013)。即使在今天,许多目标USV仍在服役,包括海上动力目标(seaborne powered targets,SEPARs)、高速机动海上目标(high-speed maneuverable surface target,HSMST)、移动船舶目标(mobile ship targets,MSTs)、QST-33和QST-35/35A。1950年代后,USV开始用于布雷目的。20世纪60年代末,更大的扫雷无人机USV也被开发并部署在越南,用于排雷。到20世纪90年代,开发了R/C Dyads、Moss和先进轻型影响扫雷系统(advanced lightweight influence sweep system,ALISS)用于排雷;最近,开发了复杂的USV排雷系统,即Brown、Palmer和Brizzola。

直到20世纪90年代末,USV才开始越来越多地

用于各种目的,包括信息/监控/侦察、港口监视和海洋调查。自主搜索和水文测量载具(autonomous search and hydrographic vehicles,ASHs)和Roboski最初开发为船载部署水面目标(ship deployable surface targets,SDSTs),即作为船舶自卫训练的喷气滑雪型目标。它们目前用于例如侦查等的困难任务(Agarwala,2022;Othman,2015)。

#### 1.1.2 商业与科研拓展

随着技术的不断进步,USV在商业领域的应用日益广泛。许多国家通过公共和私人规划在各个领域和目的下进行USV开发,USV正在科学研究、环境任务、海洋应用中被商业化使用。这种商业化应用推动了USV技术的快速发展和成本降低,使其能够为更多用户提供服务。USV在海洋科学研究中发挥着越来越重要的作用。海洋研究团队和私营公司使用USV进行海洋勘探,如伍兹霍尔海洋研究所(Woods Hole oceanographic institution,WHOI)开发的“Mesobot”(Yoerger等,2018)用于跟踪中上层带生物。美国在USV海洋勘探方面发挥先锋作用,海洋高级机器人(MAR)开发的“波浪自适应模块化船舶(WAM-V)”(Pandey和Hasegawa,2015)具有双体船船体设计,用于减轻波浪冲击和振动,为海洋科学研究提供了重要工具。2024年1月,Robosys自动化公司推出了VOYAGER AI自动驾驶系统,使USV能够成为完全自主的船只,提供水文和海洋调查服务。一些企业如L3Harris Technologies、Textron Inc.和Lockheed Martin Corporation正在通过产品发布和合作伙伴关系等各种举措来扩大市场,这些举措开发创新功能并扩展USV的行业应用空间。

这些发展表明,USV在商业和科研领域的应用正在快速扩展,技术创新和政策支持为这一领域的持续增长提供了坚实基础。

#### 1.1.3 遥控到自主化、多平台协作

USV技术经历了从遥控操作到自主化的重要演进。早期USV主要依靠遥控操作,操作员需要实时控制载具的每一个动作。随着人工智能、传感器技术和控制技术的发展和USV逐渐具备了自主感知、决策和控制能力,能够在复杂海洋环境中独立执行任务。多平台协作是USV技术发展的重要方向。USV可以与无人水下航行器(unmanned underwater vehicles,UUV)、无人机( unmanned aerial vehicles,UAV)等其他无人平台协同作业,形成异构平台协作

系统。这种协作模式能够充分发挥各平台的优势,实现更复杂的任务目标。例如,USV可以承担水面监视和通信中继任务,而UUV则执行水下探测和作业任务,UAV提供空中侦察和通信支持。群体智能是USV多平台协作的重要技术支撑。通过分布式决策和任务分配算法,多个USV可以形成协同作业群体,共同完成复杂的海洋任务。这种群体智能方法能够提高系统的鲁棒性和效率,为海洋应用提供更强大的技术支撑。随着技术的不断进步,USV将朝着更高层次的自主化和智能化方向发展,多平台协作将成为USV应用的重要特征,群体智能和分布式控制技术将为USV的协同作业提供更先进的技术支撑。通过系统性的技术发展和应用拓展,USV技术已经从最初的军事应用发展成为一个涵盖军事、商业以及科研等多个领域的综合性技术体系,为海洋智能装备的发展奠定了坚实基础,为海洋强国建设提供了重要技术支撑。

## 1.2 无人船的分类与任务需求

### 1.2.1 USV船型与尺寸分类

每种类型的USV都可以根据其大小、功能或独特特征用于特定应用,如图2所示。例如,海军USV需要远程控制和高速操作。此外,还需要重型有效载荷和能耗限制。USV海洋学研究需要自主导航、节能和低速稳定测量。商用监视USV通常配备先进的视觉监控传感器、高性能控制器以及与岸上管理中心的高效通信系统。

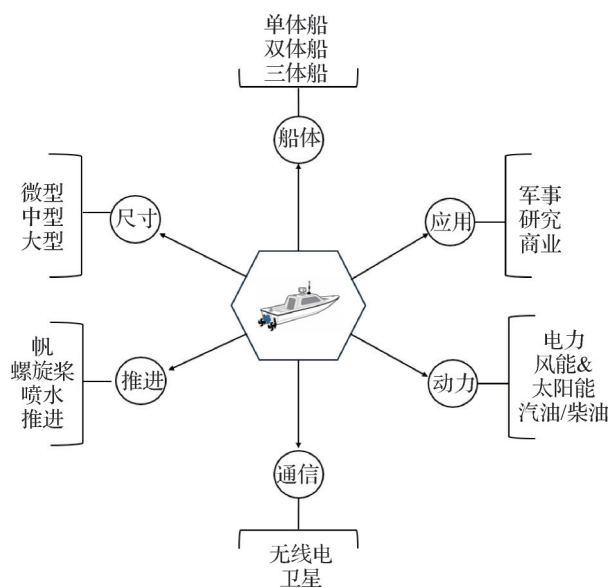


图2 无人船分类

Fig. 2 Unmanned surface vessel classification

美国海军在2007年定义了高效开发USV的标准,称为USV Master Plan(Yan等,2010)。该标准根据USV船型或尺寸对仅限于海洋船舶的USV进行分类。按船型可分为半潜式(semi-submersible, SS)型、常规滑行船体型、水翼艇型和其他船型。

半潜式USV平均长度7 m,速度25节,比传统船舶类型受海况影响较小,稳定且广泛用于向作业区域供应设备的运输,SS形态下难以被探测,便于武装作业,但设计过程比其他船舶类型更复杂。常规规划船体型具有各种船体形状(V型、改进V型、M型),在20节以下速度时具有高作业效率,有效载荷大,设计不复杂,生产成本低,但运输稳定性差,在海况影响下容易发生横摇和砰击现象。水翼艇型具有最强的海况适应能力和良好的稳定性,速度可达40节,作业效率突出,但由于速度特性不适合拖拽,发射和回收操作复杂,生产成本高。

按尺寸分类,USV可分为X级(小型,~3 m)、港口级(7 m)、潜水员级(7 m SS)和舰队级(11 m)。X级主要用于支援特种作战部队(special operations forces, SOF),属于面向具体任务的低成本、可消耗型平台,通常针对特定用途专门设计制造,不采用标准化模块化设计,主要执行海上拦截作战支援和‘低端’情报、监视与侦察(intelligence, surveillance, and reconnaissance, ISR)任务。港口级是大多数海军舰艇上用于海洋安全的船只大小,执行最基本的任务,具有ISR/火炮有效载荷、反水雷(mine countermeasures, MCM)投送、水面战(surface warfare, SUW)和SOF支持。潜水员级在高海上比其他USV更可靠地工作,用于MCM搜索、海洋盾牌(anti-submarine warfare, ASW)和特殊任务支持。舰队级具有平面或半平面船体,由于使用而耐用,用于MCM扫雷、海洋盾牌、SUW、火炮和鱼雷、“高端”SOF和高功率电子战(Toh, 2017; Graham, 2008)。

### 1.2.2 任务类型

USV的主要应用包括MCM、反恐、ISR、ASW和海洋勘探(Yan等,2010;谢伟等,2019)。在MCM应用中,无人平台可以在不需要载人平台进入疑似有雷区的情况下运行(Yang等,2021),舰队可以通过寻找或清除无雷区域来运行,缩短航行时间,USV被大型舰队用于快速建立安全作业区域、运输路线和运输通道。使用USV发现的水雷通过载具上的扫雷器进行中和,需要在海军系统内快速运行,独立于

其他作战能力。

在反恐应用中,USV集成到更广泛的安全网络中,为港口和河岸安全提供保护,对抗隐蔽和非传统威胁。USV在延长期内远程执行昼夜港口或河流监控,如果在潜在威胁平台调查中确定平台构成威胁,USV负责清除它。它还提供远程探测、询问和与港口监视、商船和海军舰艇的潜在威胁交战。

在ISR应用中,监控应该能够在海上或海岸成功执行,符合指挥官和操作员的一般指示,无论情况危险如何(Winstead, 2018)。USV配备前向红外激光测距仪用于探测和跟踪附近目标,以及可以昼夜运行的安装式摄像头。应保持USV和控制站之间的持续连接以通过USV进行监控。ISR能力可以帮助识别海上交通中的可疑行为、消除危险并保护海洋设施和船舶。此外,USV完成的任务将为作战海洋单位提供增强的能力和灵活性,因为指挥和控制中心人员可以远程操作。

在ASW应用中,有效部署并维持应对日益增长的潜艇威胁的反潜作战能力至关重要(Heo等, 2017)。ASW可根据发射端与接收端的相对配置方式开展作业,主要包括单基地、双基地和多基地等模式:单基地模式中,发射端和接收端部署在同一艘USV上;双基地和多基地模式中,发射端和接收端

分别部署在不同平台或多艘USV上。

这允许USV基于其他可用资产和能力使用特定USV补充和扩展现有ASW能力。此外,USV可以高效地代替人类在危险情况下执行任务,如海上救援和防灾引航支持、海洋领土保护、海洋环境勘探和鱼雷作业支持。

### 1.3 系统组成架构

#### 1.3.1 船体与动力系统

USV的船体设计直接影响其在海洋环境中的稳定性和性能表现。根据不同的应用需求,USV可以采用多种船体结构,包括单体船、双体船、三体船等。单体船设计简单,成本较低,适合浅水作业;双体船具有更好的稳定性和载荷能力,适合深水作业;三体船则在高速航行时表现出色,适合需要快速响应的任务场景。动力系统是USV的核心组成部分,决定了载具的续航能力和作业效率。现代USV主要采用电力推进系统,包括电池组、电机和推进器等组件。电池组通常采用锂离子电池或燃料电池,提供稳定的电力输出;电机负责将电能转换为机械能,驱动推进器工作;推进器则根据具体需求选择螺旋桨或喷水推进器。此外,部分USV还配备太阳能板或风力发电装置,实现可再生能源的利用,延长作业时间。典型的USV架构如图3所示。

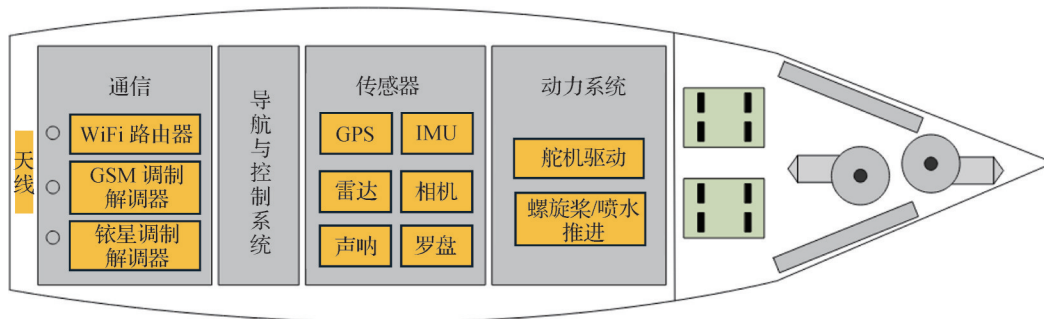


图3 典型的USV架构

Fig. 3 The architecture of a typical USV

#### 1.3.2 通信与控制系统

通信系统是USV与地面控制站或其他平台进行信息交换的关键环节(Lyu等, 2025)。现代USV通常配备多种通信方式,包括无线电通信、卫星通信和网络通信等。无线电通信适用于近距离作业,具有延迟低、成本低的优势;卫星通信适用于远距离作业,确保在远离海岸的区域仍能保持通信联系;网络通信则支持数据传输和远程控制功能。控制系统是USV实现自主作业的核心,包括硬件控制系统和软

件控制系统两个层面。硬件控制系统负责执行具体的控制指令,包括舵机、推进器控制器等执行机构;软件控制系统则负责决策制定和任务规划,包括路径规划算法、避障算法和任务调度算法等。通过硬件和软件的有机结合,USV能够实现复杂的自主作业任务。

#### 1.3.3 传感器系统

传感器系统为USV提供环境感知和状态监测能力,是自主作业的重要基础(Sotelo-Torres等,

2023)。现代USV通常配备多种传感器,包括定位传感器、环境传感器和任务传感器等。定位传感器主要包括GPS(global positioning system)接收机、惯性测量单元(inertial measurement unit, IMU)和罗盘等,用于确定载具的位置、姿态和运动状态;环境传感器包括雷达、声纳和摄像头等,用于感知周围环境的变化和障碍物;任务传感器则根据具体应用需求配置,如水质监测传感器、气象传感器等。

传感器数据的融合处理是USV智能化的关键。通过多传感器数据融合算法,USV能够获得更准确的环境信息和状态估计,为决策制定提供可靠的数据支撑。此外,传感器系统还需要具备故障检测和容错能力,确保在部分传感器失效的情况下仍能维持基本功能。

#### 1.3.4 NGC软件结构

NGC系统是USV自主作业的核心软件架构,负责实现载具的自主导航和精确控制(Chen, 2025)。NGC系统通常采用分层架构设计,包括任务规划层、路径规划层、制导层和控制层等。任务规划层负责制定整体任务目标和约束条件;路径规划层根据任务要求生成最优路径;制导层负责跟踪规划路径;控制层则执行具体的控制指令。

NGC系统的软件架构需要具备模块化、可扩展和可维护等特性。通过模块化设计,系统各功能模块可以独立开发和测试,提高开发效率;通过可扩展设计,系统能够适应不同的应用需求和硬件配置;通过可维护设计,系统能够快速定位和修复问题,提高系统可靠性。此外,NGC系统还需要支持在线学习和自适应控制,能够根据环境变化和任务需求调整控制策略,提高系统的适应性和鲁棒性。

通过船体与动力系统、通信与控制系统、传感器系统和NGC软件结构的有机结合,USV形成了完整的自主作业系统。这种系统架构不仅保证了载具的可靠性和安全性,还为未来的技术升级和功能扩展提供了良好的基础。随着人工智能、大数据等技术的不断发展,USV的系统架构将进一步完善,为海洋智能装备的发展提供更强大的技术支撑。

## 2 深度学习驱动的USV智能感知

许多研究已经聚焦于深度学习的进展(LeCun等, 2015; Bengio等, 2021),以及深度学习在机器人

(Sünderhauf等, 2018)、自动驾驶汽车(Grigorescu等, 2020)、无人机(Carrio等, 2017)和无人水下航行器中的应用。因此,为避免不必要的重复,本文简要介绍按不同学习方法(即监督学习、无监督学习和强化学习方法)分类的深度学习模型及其应用,其中,主要介绍与USV强相关的模型。

### 2.1 深度学习方法及应用

#### 2.1.1 监督学习

监督深度学习方法是使用标记良好的数据进行训练,学习将输入映射到输出的函数。在USV应用中,监督学习方法面临海洋环境数据标注困难、跨域泛化能力不足等挑战,但通过合理的网络设计和训练策略,仍能在导航、制导和控制任务中发挥重要作用。

1)卷积神经网络。在视觉感知任务中,卷积神经网络(convolutional neural network, CNN)是最具代表性的监督学习模型。通过局部感受野、参数共享与层次化特征提取机制, CNN能够在复杂海面背景下实现目标检测、图像分割和多目标跟踪。与海洋视觉任务的物理特性相结合,可以看出: CNN的局部卷积机制有助于提取具有显著几何边界的局部结构信息(如船体轮廓、桅杆、浪尖等),因此在光照稳定、视距较短的场景下具有优异性能。其参数共享机制有效减少模型参数量,能够在计算资源受限的船载平台上实现高效运行。基于CNN的基本思想,设计了先进的架构解决计算机视觉领域的3个基本问题:图像分类、目标检测和图像分割。一些流行的架构简要介绍如下,用于通过图像、视频或在海洋环境中收集的数据感知周围环境。

图像分类旨在预测给定图像的标签。基于传统CNN,提出了几种具有更复杂结构的改进架构,包括LeNet-5、AlexNet、VGG(Visual Geometry Group)、GoogLeNet、Inception Net、ResNet(residual network)等(Zhao等, 2024)。作为ILSVRC(imagenet large scale visual recognition challenge)的代表性模型,这些CNN架构实现了高分类精度,其中一些甚至超过了ImageNet数据库上的人类水平性能(Krizhevsky等, 2017)。这些网络在许多应用中广泛用作特征提取的基本模型,或作为骨干网络应用于目标检测和分割任务。在USV环境中,这些分类网络可用于识别摄像头或遥感图像中的典型目标类别(如船只类型、漂浮垃圾或辅助导航标志)

(1) CNN 目标检测。自主载具在通过传感器感知周围环境时必须识别和定位目标。一般而言,有两种类型的目标检测框架。其一,两阶段算法首先提取可能包含目标的候选区域,然后确定提议区域是否包含目标。第一阶段使用选择性搜索或区域提议网络(region proposal network, RPN)等方法生成感兴趣区域(region of interest, ROI)。然后,利用另一个网络(如 ResNet)对提议区域进行分类。流行的两阶段目标检测方法包括 R-CNN(region-based convolutional neural network)、Fast R-CNN(fast region-based convolutional neural network)(Girshick, 2015)、Faster R-CNN(Ren 等, 2017)和 Mask R-CNN(mask region-based convolutional neural network)。其二,单阶段算法直接对目标进行分类,并在一步中预测图像的边界框。更具体地说,输入图像,同时学习类别概率和边界框坐标。典型框架包括 YOLO(you only look once)、SSD(single shot multibox detector)和 RetinaNet(Jiao 和 Abdullah, 2024)。在 USV 实际部署中,研究者常对 YOLO 的颈部网络进行改造:通过引入多尺度融合与时序特征聚合模块,减弱浪涌导致的小目标抖动影响;采用可变形卷积(deformable Conv)和注意力模块增强特征对齐;并结合量化与剪枝策略减少船载计算负担。这些改进显著提升了模型在真实海况中的检测鲁棒性。

(2) CNN 分割。在 USV 自主导航中,准确理解水面环境至关重要。图像分割技术能够为 USV 提供像素级的场景理解,这对于识别航道边界、障碍物、其他船只和危险区域具有重要价值(Zhang 等, 2026)。全卷积网络(fully convolutional network, FCN)通过将高级特征图上采样为与原始图像相同大小的热图,使 USV 能够精确识别水面上的每个像素类别。金字塔注意力网络(pyramid attention network, PAN)通过特征金字塔注意力(feature pyramid attention, FPA)和全局注意力上采样(global attention upsample, GAU)模块,让 USV 在复杂的水面环境中更好地理解全局上下文信息,这对于在繁忙航道中做出安全导航决策至关重要。U-Net 的 U 型架构通过收缩路径捕获上下文信息和扩展路径实现精确定位,特别适合 USV 需要同时理解大范围水域环境和精确定位小型障碍物的需求。最新分割架构如 SegFormer(Xie 等, 2021)、Mask2Former(Cheng 等, 2022)、InternImage(Wang 等, 2023a)和 SAM(seg-

ment anything model)(Kirillov 等, 2023)等,为 USV 提供了更强大的场景理解能力,使其能够在各种天气和水面条件下准确识别和分割目标,这对于 USV 的安全自主导航和避障系统至关重要。

2) 循环神经网络。循环神经网络(recurrent neural network, RNN)在 USV 应用中具有特殊价值,因为它能够利用历史信息预测未来行为,这对于 USV 的轨迹预测和动态避障至关重要。RNN 的时序建模能力强,能够学习时间序列数据中的长期依赖关系,特别适合处理 USV 的轨迹预测任务,通过隐藏层传递历史信息,使 USV 能够学习复杂的时间模式,如潮汐变化、季节性航道变化等,同时相比其他深度学习方法,RNN 的计算复杂度相对较低,适合 USV 的实时决策需求。然而,传统 RNN 存在梯度消失和爆炸问题,在训练期间难以收敛,严重影响模型性能。此外,其序列计算特性使其无法并行化,在 USV 的实时性要求下可能成为瓶颈,即使使用长短期记忆网络(long short-term memory, LSTM)和门控循环单元(gated recurrent unit, GRU),在处理超长期依赖关系时仍存在困难。LSTM 通过专门设计的记忆网络(包括单元、输入门、输出门和遗忘门)解决了梯度消失问题,使 USV 能够更好地记忆长期依赖关系,GRU 作为 LSTM 的改进版本,具有更少的训练参数,在 USV 的实时决策系统中能够提供更快的推理速度,但需要注意的是,这些改进方法在处理复杂海洋环境下的多目标跟踪任务时,仍存在计算复杂度过高的问题。因此,RNN 特别适合处理轨迹预测、动态避障等时序任务,但不适合处理需要全局信息的海况评估任务。

3) Transformer。近年来,Transformer 结构在视觉领域取得了显著进展,并逐步拓展至 USV 感知与决策任务。Transformer 以自注意力机制为核心,最初在自然语言处理中取得突破,随后被广泛应用于视觉表征学习与序列建模任务(Vaswani 等, 2017)。相较于依赖局部感受野进行特征提取的 CNN,Transformer 更擅长建模长距离依赖关系和全局上下文信息,因此在复杂海天背景、小目标检测以及动态场景感知等 USV 任务中展现出较大潜力。

作为对比,CNN 的技术基础在于局部感受野与层次化特征提取,即通过  $3 \times 3$ 、 $5 \times 5$  等固定尺寸卷积核对局部像素块进行滑动计算,并通过多层堆叠完成从船体边缘等低层纹理到完整船只轮廓等高层

语义的特征抽象。该机制虽然能够有效捕捉局部细节,但在模糊海天背景下存在以下适配局限。(1)局部像素关联范围受限。在雾天等低对比度场景中,海天交界线纹理极弱,卷积核仅能覆盖局部像素关系,难以建立局部区域与全局场景之间的有效关联,容易将波浪反光、水雾颗粒等局部噪声误判为小型船只。(2)特征在多层传递过程中可能出现全局信息衰减。目标与水天线的相对位置、多目标之间的空间关系等关键信息,在层间传播过程中容易被削弱,尤其当目标面积占比较小时,全局表达能力下降更为明显,进而影响远距离小目标的识别效果(Guo等,2023)。(3)CNN对动态模糊场景的适应能力有限。面对浪涌引起的目标模糊,仅依赖单帧图像进行静态特征提取,难以区分目标真实运动与模糊噪声(Feng等,2022)。

相比下Transformer的优势在于基于自注意力机制的全局建模能力,通过计算全图像素间的注意力权重,直接建立像素与目标、水天交界线等环境区域的长距离关联,适配复杂海天背景。在应对水天模糊方面,自注意力可动态分配权重(如像素块与交界线的距离、与已知船只特征的相似性),即便在雾天或灰白色调背景中,也能通过全局像素关联锁定目标(Carion等,2020;Bovcon和Kristan,2022)。在小目标检测方面,Transformer的多头自注意力可并行捕捉局部细节与全局位置信息,一组注意力头聚焦船体边缘等细节特征,另一组建模目标与水天线距离等全局信息,提升小目标检测能力(Guo等,2023;Feng等,2022)。在应对动态模糊方面,Transformer可借助跨帧自注意力(如Video Swin Transformer的时序窗口设计)将多帧像素纳入全局运算,构建目标运动轨迹,在浪涌环境中提升跟踪鲁棒性(Luiten等,2021;Lin等,2024)。因此,Transformer架构通过自注意力机制实现全局感知与时序建模,有效克服了CNN在模糊、动态复杂场景下的局限,提升了小目标检测性能及动态环境中的鲁棒性与精度。

### 2.1.2 无监督与自监督学习

在海上无人船感知任务中,获取大规模高质量标注数据往往十分困难。由于海洋环境复杂多变,人工标注成本高昂且容易出现主观误差,这使得传统监督学习方法在实际应用中受到限制。为此,无监督学习与自监督学习逐渐受到研究者关注。这类方法能够在缺乏标注的条件下自动挖掘数据的潜在

结构,或利用对比、预测等预设任务进行特征学习,从而提升模型的泛化能力和数据利用效率。

1)生成对抗网络。生成对抗网络(generative adversarial network, GAN)是为数据生成设计的对抗学习框架,具有生成模型G(生成与训练数据相似的样本)和判别模型D(判别样本来自训练数据还是G)。G的训练过程是最小化真实数据分布和生成数据分布之间的Jensen-Shannon(JS)散度。D的训练过程是尽可能区分真实数据和生成数据(Goodfellow等,2014)。此外,提出了深度卷积GAN(deep convolutional generative adversarial network, DCGAN)来提供基于实验的最优超参数集和训练技巧,以提高GAN训练过程的稳定性(Radford等,2016)。为了从结构上进一步改善GAN的稳定性并消除模型崩溃和梯度消失等问题,GAN已广泛用于通过生成假图像扩充数据集。如果训练数据集不足,这对USV导航模块非常有用。最新发展包括StyleGAN系列能够生成高质量、可控的人脸图像,在图像生成质量方面取得了重大突破(Karras等,2019,2020,2021);BigGAN大规模GAN模型,在ImageNet等大型数据集上训练,生成质量显著提升(Brock等,2019)。无监督与自监督学习在解决数据不足、标注困难以及域适应等问题方面具有重要意义。它们不仅能够提升模型在复杂环境下的鲁棒性,还为构建大规模海洋感知模型提供了可能。

### 2.1.3 深度强化学习

在无人船智能决策与控制领域,强化学习(reinforcement learning, RL)为解决复杂环境下的自主航行问题提供了新的思路。不同于监督学习依赖大规模标注样本,强化学习通过智能体与环境的交互不断试错,并基于奖励信号优化策略,能够在动态、不确定的海洋场景中逐渐学习出有效的感知—决策—执行链条(Sutton和Barto,2018),特别适合处理海洋环境中数据标注困难的问题,在路径规划、避碰控制、编队协同以及复杂任务调度中展现出广阔的应用前景。然而,强化学习在USV应用中面临训练效率低、奖励函数设计困难、仿真与现实差距、安全性问题等挑战,强化学习需要大量交互数据,在USV这种高风险、高成本的平台上训练效率极低,如何设计合适的奖励函数以引导USV学习安全、高效的航行策略是一个关键挑战,强化学习通常依赖仿真训练,仿真与现实之间的差距会导致策略迁移困难,强

化学习在训练过程中可能出现不安全的行为,在USV这种安全关键的应用中风险较高。深度强化学习(deep reinforcement learning, DRL)通过将深度神经网络作为函数逼近器,能够直接从高维感知数据中学习最优策略,但需要注意的是,DRL在USV应用中仍存在计算复杂度过高、训练不稳定等问题。因此,强化学习特别适合处理路径规划、避碰控制等决策任务,但不适合处理需要高精度控制的精确导航任务。

早期的强化学习方法由于依赖有限的状态空间和显式策略更新,难以应对连续、高维的感知与控制任务。深度强化学习(deep reinforcement learning, DRL)的提出有效解决了这一瓶颈。通过将深度神经网络作为函数逼近器,DRL能够直接从高维感知数据中学习最优策略。典型方法如深度Q网络(deep Q-network, DQN)适用于离散动作空间,已应用于船舶离散操舵控制与避障决策(Mnih等, 2015);而深度确定性策略梯度(deep deterministic policy gradient, DDPG)(Lillicrap等, 2016)、近端策略优化(proximal policy optimization, PPO)(Schulman等, 2017)等则能够处理连续动作空间,更契合USV在速度、航向与舵角上的连续控制需求。相关研究表明,基于DRL的自主避碰方法在动态环境中可以超越传统基于规则或优化的路径规划方法,表现出更强的适应性与实时性。

强化学习在仿真环境中已取得显著成果,但其在实际海上部署中仍面临若干挑战。首先,训练过程需要大量交互数据,而真实海上实验的成本和风险过高,这使得大部分研究依赖于仿真平台。然而,仿真与现实之间的差异(sim-to-real gap)会导致训练好的策略在实船中表现下降。深度网络带来的高计算开销与不稳定性问题,也制约了其在资源受限平台上的实时部署。为缓解这些问题,近年来的研究尝试通过迁移学习、元学习以及领域自适应技术缩小仿真与现实之间的差距,同时利用模型压缩、轻量化网络和分布式计算提升算法的可部署性。

上述深度学习模型为USV智能感知提供了技术基础,但其性能验证与泛化能力提升需依托高质量海事数据集,具体数据集特征及基于数据的感知算法对比分析,将在第3节展开。本节余下篇幅将简单介绍USV智能感知的感知任务与对应技术。

## 2.2 感知需求与任务场景

### 2.2.1 海面目标检测与识别

海面目标检测与识别是USV感知系统的首要任务,直接关系到航行安全和任务执行效果。海洋环境中存在多种类型的目标,包括其他船只、浮标、障碍物以及海洋生物等,这些目标具有不同的运动特性和威胁程度。

1)光学图像处理技术。光学图像是USV感知系统的重要数据源,能够提供丰富的视觉信息。然而,海洋环境中的光学图像处理面临着独特的挑战,包括图像质量受时间和云层覆盖影响、高分辨率数据难以实时处理、港口场景中船舶目标分割与区分难度较大、目标船只形状细长且旋转任意等。

在复杂背景下检测船只时,提取能够更好地区分船只与周围湍流的特征表示非常重要。研究人员尝试整合高低层特征进行船只检测。基于Mask R-CNN, Nie等人(2020)添加了自下而上的路径,将低层特征传播到顶层。Yang等人(2018)采用了密集特征金字塔网络(dense FPN),通过密集连接和特征融合实现不同层特征图之间的信息交互。对于多尺度船只检测, Li等人(2021)构建了两个回归分支,基于不同的CNN特征独立预测边界框的中心点、宽度、高度和方向。Zhang等人(2021b)通过将船只检测任务转换为二元语义分割任务,提出了一种无锚点旋转船只检测方法。

2)雷达图像处理技术。雷达图像由机载、岸基传感器或卫星传感器捕获,通常面临分辨率低、像素少、多尺度船只密集聚集、陆地类似散射物体导致高误报率等挑战。虽然合成孔径雷达(synthetic aperture radar, SAR)能够在任何天气条件下全天候工作,但SAR图像通常比光学遥感图像具有更低的分辨率和更少的像素。

为了获得更多语义信息, Cheng等人(2021)融合了从雷达和光学图像中提取的特征以识别水面上小物体。Zhao等人(2020)提出了一种两阶段检测方法,采用并组合感受野块和卷积块注意力模块,增强局部特征与其全局依赖的关系。缺乏用于训练的大量真实数据是SAR图像中船只和其他物体检测和分类的主要问题之一(Ward等, 2018)。

3)状态估计与行为识别。基于DNN的方法可以在存在非线性动力学和不同采样频率的高维传感器数据的情况下估计USV的运动。Chen等人

(2020)基于视频数据分4个步骤识别船只行为:基于YOLO的船只特征提取、边界框生成、基于几何理论的位置识别以及行为分析。Zhang等人(2021a)利用LSTM捕获每个频率尺度上船只运动的内在规律,随后采用注意力机制进一步提高预测的准确性。

### 2.2.2 海况感知与障碍物识别

在自主航行过程中,航行器不仅需要感知水面目标的位置与类别,还必须对海况进行实时监测,并识别潜在障碍物,以确保航行安全和任务顺利完成。海况感知主要涵盖海浪高度、方向与周期,洋流速度与流向,以及风速、风向等气象条件,这些信息对于路径规划、避碰决策和动力控制至关重要(Zhang等,2024)。在实际应用中,这些数据可通过船载传感器(如波浪计、气象站、声呐、雷达、激光雷达等)与远程观测平台(如卫星、无人机)协同获取(Briguglio和Crupi,2024)。

障碍物识别是海况感知的重要组成部分,其目标是在不同气象条件和海面状态下,准确识别可能影响航行安全的固定或移动物体,包括浮标、渔网、漂浮物、冰块以及其他船只等(Dima等,2004)。早期方法多依赖传统图像处理与几何建模技术,例如利用边缘检测、颜色分割以及形状匹配等方式,从视觉图像中提取疑似障碍物(Liu等,2021)。然而,在强反射、浪花以及低光照等复杂环境中,这类方法容易产生较高的误检率与漏检率。

为克服上述不足,近年来的研究趋势是引入深

度学习与多模态传感器融合技术(Balemans等,2024;Malvadkar等,2025;Zuo等,2025)。在视觉域中,卷积神经网络(CNN)与基于注意力机制的架构广泛应用于障碍物识别任务,它们能够在复杂背景中自动学习区分障碍物与海面杂波的特征(Yan等,2024)。同时,结合雷达、激光雷达等主动感知设备获取的距离与速度信息,可以在感知系统中引入几何约束,从而有效抑制海浪、白沫等干扰带来的误判(Ding等,2023;Cui等,2021)。

多模态融合的实现路径主要有两类:特征级融合与决策级融合(Huang等,2020)。前者在特征提取阶段将来自不同传感器的数据映射到统一特征空间,并通过深度网络进行联合建模,以提升障碍物识别的鲁棒性与泛化能力;后者则在各模态独立完成识别任务后(El-Din等,2024),通过贝叶斯推理、证据理论或加权投票等方式综合不同模态的识别结果,从而提高整体决策的准确性,如图4所示。

此外,在恶劣天气条件下,如大雾、暴雨及风浪等,单一模态感知的性能会显著下降(Wang等,2024b;Panchal等,2025)。针对这一问题,部分研究提出基于生成模型或物理先验的图像增强方法,在输入阶段对视觉数据进行去雾、去雨或波动抑制处理(Li等,2026);刘春晓等人(2025)提出了一种解耦合的三阶段增强网络,通过分层处理亮度与饱和度,有效提升了非均匀雾霾环境下的图像清晰度,这对

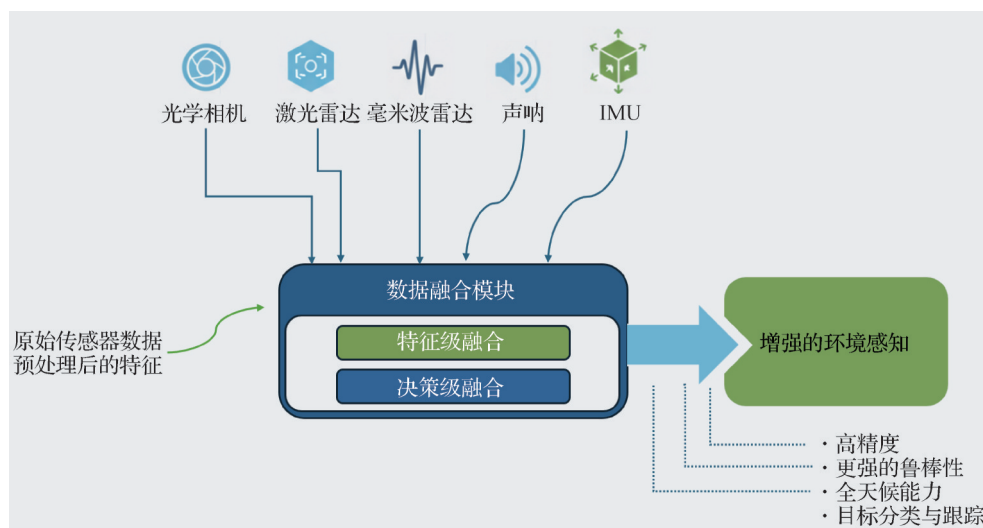


图4 多传感器融合示意图

Fig. 4 Multi-sensor fusion diagram

USV 的海上视觉增强具有重要参考价值。另一些工作则通过时序信息建模,将历史观测数据与当前观测结果结合,利用卡尔曼滤波、粒子滤波等技术对障碍物位置和海况参数进行动态估计,从而提升系统的稳健性(Lei等,2024)。

### 2.2.3 多目标跟踪与动态环境建模

多目标跟踪(multi-object tracking, MOT)与动态环境建模是自主水面航行器实现长期稳定运行和安全避碰的核心环节,其目的是在复杂、多变的海上场景中,持续追踪多个动态目标的位置、速度与轨迹,同时构建环境的时空表示,为路径规划和任务分配提供实时数据支撑(Liu等,2019)。与单目标跟踪相比,MOT需要同时应对目标间的相互遮挡、航迹交叉以及环境噪声等问题,这在波浪起伏、光照变化及目标运动模式多样的海面环境中尤为突出(Luiten等,2021)。

在实现方法上,传统多目标跟踪(MOT)主要依赖检测—关联(tracking-by-detection)框架,即先通过目标检测器在每帧图像中定位目标,再利用匈牙利算法(Lin等,2021)、联合概率数据关联(joint probabilistic data association, JPDA)(Wen等,2020)或多假设跟踪(multiple hypothesis tracking, MHT)(Groves, 2013)等数据关联方法建立跨帧匹配。然而,这类方法对检测结果质量高度依赖,在低信噪比条件下容易产生身份(identity, ID)切换和轨迹丢失。

近年来,深度学习的发展推动了端到端跟踪器的应用,这类方法通过共享特征提取网络,实现检测与跟踪的联合优化,从而在实时性与鲁棒性之间取得平衡(Krizhevsky等,2017)。部分研究引入 ReID(re-identification)特征,以增强在短时遮挡或外观相似情况下的身份保持能力(Zheng等,2015);另一些方法利用时序建模网络(如LSTM、Transformer)编码目标历史轨迹,从而提升在复杂机动模式下的预测精度。

与MOT聚焦于离散目标不同,动态环境建模旨在构建一个对周边环境连续、全面的时空表示。这一过程通常依赖于多传感器信息融合技术,将来自雷达、激光雷达(light detection and ranging, LiDAR)、摄像头等多种传感器的数据进行整合,以同时获取环境的静态结构(如岸线、岛屿、固定障碍物)与动态要素(如其他船只、浮标、漂浮物)(Liu等,2019)。

在USV系统中,基于栅格地图(grid map)或点云

地图的建模方法得到了广泛应用(Danescu等,2011)。其中,占据栅格地图(occupancy grid mapping, OGM)是一种主流技术。通过将环境划分为离散的网格单元,并估计每个单元被障碍物占据的概率,OGM能够有效地表示环境并处理传感器数据的不确定性。为了捕捉环境的动态变化,研究人员进一步发展了动态占据栅格地图(dynamic occupancy grid maps, DOGM)。该方法通过对传感器数据进行时序叠加,并利用贝叶斯滤波理论(如卡尔曼滤波或更先进的贝叶斯占据滤波器)对每个网格单元的状态(如占据概率和速度)进行递归估计,从而生成能够反映环境演化趋势的动态地图(Coué等,2006)。

此外,随着深度学习技术的发展,部分前沿工作开始探索利用深度生成模型进行环境预测。这些模型能够从历史传感器数据中学习环境变化的复杂模式,并在有限观测条件下推断甚至生成未来一段时间内场景的可能状态。例如,基于深度学习的预测性动态占据网格(predictive dynamic occupancy grid)技术,能够预测未来环境中障碍物的位置和形态(Jang等,2024)。这种预测能力为USV实现预防性避碰与高级航迹优化提供了宝贵的先验信息,是实现更高层次自主性的关键。

## 2.3 基于智能感知的传感器类别

### 2.3.1 自感传感器

自感传感器是USV感知与导航系统的基础,其特点是通过内部测量直接获取平台自身状态参数。全球定位系统(GPS)及全球导航卫星系统(global navigation satellite system, GNSS)可在开阔水域提供米级至分米级的定位精度(Groves, 2013),并且在差分GPS(differential global positioning system, DGPS)或实时动态定位(real-time kinematic, RTK)技术支持下,可实现厘米级定位(Groves, 2013)。然而,GNSS在城市港口、高纬度地区或强干扰环境下易受到遮挡和多径效应的影响(Zhang等,2024)。

惯性测量单元(IMU)通过加速度计与陀螺仪获取平台的加速度与角速度信息,可提供高频率的姿态与运动状态估计。在短时间内,IMU具有较高的稳定性,但长期运行会积累漂移误差,因此常与GNSS进行互补融合,构建惯性导航系统(INS),以实现连续稳定的定位。

### 2.3.2 外感传感器(相机、LiDAR、雷达、声呐等)

外感传感器通过观测外部环境为USV提供感知信息,是目标检测、障碍物识别、海况评估等任务的关键(Zhong等,2021;Han等,2018)。相机具备高分辨率与纹理细节捕捉能力,适合在良好光照条件下执行视觉任务,但对光照变化与天气敏感(Sazzadul Alam等,2025)。激光雷达(LiDAR)能够精确测量物体的距离与形状,且不受光照影响,但其在强雾和雨雪中性能下降(Zhong等,2021)。

毫米波雷达凭借其全天候工作能力和远距离探测范围,成为USV感知系统的重要组成部分(Lee等,2023)。尤其在能见度较低的环境中,如雾、雨、雪,雷达能够保持稳定的性能,提供可靠的目标检测和跟踪。其独特的多普勒效应测量能力使其能够精确测量目标的径向速度。然而,毫米波雷达的角分辨率相对较低,这可能导致对密集障碍物的区分能力不足(Yao等,2024)。最近的研究开始利用4D雷达技术来克服传统雷达的数据稀疏性和测量噪声问题,并结合相机数据进行融合,以提升检测精度(Baumann等,2024)。

声呐传感器在水下探测中发挥着不可替代的作用(Fan等,2019)。它通过发射声波并在水下传播,然后接收反射回来的声波来探测水下目标,获取目标的距离和方位信息。这对于USV进行水下地形测绘、水下障碍物检测和水下目标识别至关重要。但是,声呐的分辨率和刷新率通常低于光学或雷达传感器,尤其是在复杂水下环境中,声波的衰减和多径效应会进一步限制其性能。

## 3 视觉数据集与感知算法基础

### 3.1 USV视觉数据集现状

#### 3.1.1 数据集发展趋势

近年来,伴随深度学习在计算机视觉领域的广泛应用,海上无人船(USV)视觉数据集的规模与多样性持续增长。早期(2015—2018年)的数据集数量有限,主要依赖人工采集的实船视频或实验室环境下的模拟数据,任务多集中于目标检测和简单的语义分割。从2019年起,随着海洋探测、港口管理和自主航行研究的加速,数据集规模显著扩大,涵盖多种环境条件(平静海面、波浪、水雾、夜间)与复杂任务(实例分割、深度估计、多目标跟踪等)(Zhang

等,2015)。进入2020年代后,部分数据集开始强调多模态信息的同步采集,例如同时提供RGB视频、红外影像、激光雷达点云和雷达信号,从而为多传感器融合算法的研究奠定了基础(Patino等,2021;Shao等,2025)。图5展示了2025年9月之前汇编的USV视觉数据集。

#### 3.1.2 任务类别(检测、分割、分类、深度估计等)

现有USV视觉数据集按任务类型大致可分为目标检测、图像分割、目标分类、深度估计以及多目标跟踪等(Bakht等,2025)。目标检测数据集通常提供边界框标注,用于定位海面上的船只、漂浮物等对象;分割数据集则包含像素级标注,能够区分船体、海面、天空以及背景杂波等区域;分类任务主要面向船型识别、渔具类型识别等细粒度类别;深度估计任务数据集则提供稠密或稀疏深度信息,用于三维环境建模和避碰决策(Lin等,2024)。

#### 3.1.3 典型公开数据集特点

典型的USV视觉数据集如MODD(maritime obstacle detection dataset)、Singapore Maritime Dataset、SeaShips等,覆盖不同的采集环境与任务类型。MODD数据集专注于动态障碍物检测,包含在不同天气与光照条件下采集的视频序列(Bovcon和Kristan,2022);Singapore Maritime Dataset具有多种视角(船载、岸基)与长时间序列的特点,适合多目标跟踪与交通流量分析;SeaShips数据集则规模较大,船只类别丰富,适合深度学习模型的训练与迁移学习研究。此外,新一代数据集正逐渐引入恶劣天气、夜间航行及多模态同步采集,以解决现实应用中数据匮乏的问题(Su等,2023)。

### 3.2 数据集在算法发展中的作用

数据集的规模与多样性直接影响算法的泛化能力与鲁棒性。丰富的训练样本能够帮助模型学习更加稳健的特征表示,从而在不同海域、天气和光照条件下保持稳定性能(Fan等,2023)。例如,一项全面的综述研究指出,现有海事数据集在场景覆盖度上的局限性是制约模型泛化能力的主要瓶颈之一,增加来自不同地理位置、季节和一天中不同时间的样本,对于提升模型的鲁棒性至关重要(Žust等,2023)。

此外,跨域数据集(例如不同港口或不同季节采集的数据)能够减少模型在域外环境下的性能退化,相关研究通过在跨场景数据集上进行测试,验证了

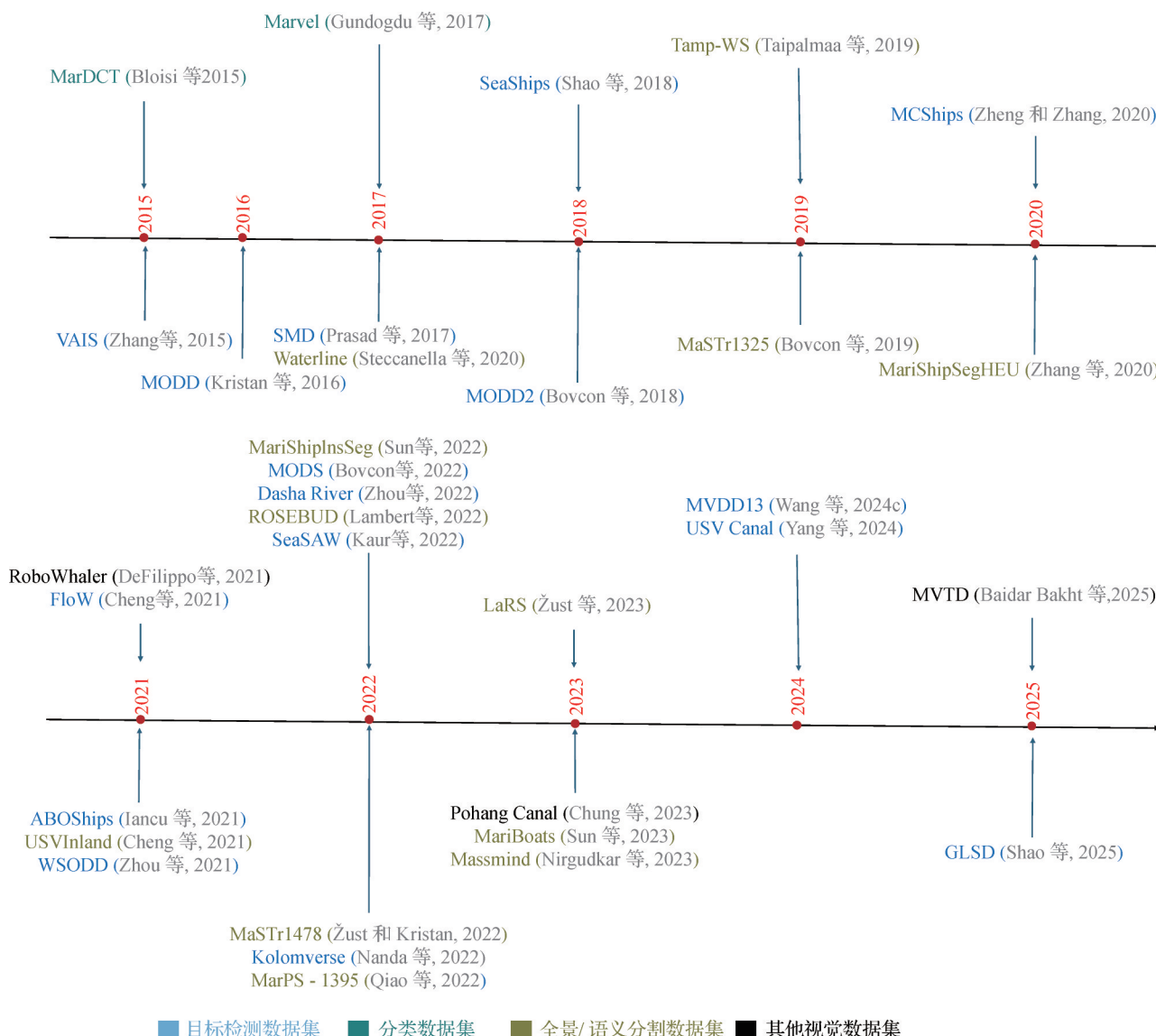


图5 USV 视觉数据集的时间顺序概述(2015—2025年)

Fig. 5 Chronological overview of vision datasets for USV (2015—2025年)

模型在迁移到未见过的环境中时,拥有更多样化训练数据的模型表现出更小的性能衰减(Bovcon 和 Kristan, 2022)。部分研究表明,在多任务联合数据集上训练的深度学习可以同时提升检测、分割和跟踪等任务的性能,这种跨任务迁移得益于共享的底层视觉特征。例如,有研究者提出了一种多任务学习框架,通过共享编码器同时处理检测和分割任务,实验证明这种联合训练方式比单任务训练在两项任务上均取得了性能增益(Li 等, 2021)。

尽管已有多个高质量数据集,但夜间、恶劣天气(如大雾、暴雨、风浪)以及多模态(红外+可见光、雷达+视觉)场景下的公开数据仍严重不足,近年来如“USVTrack”这类包含4D雷达与摄像头数据的专

用数据集的出现,正是为了填补这一空白(Yao 等, 2025)。数据缺口导致模型在实际应用中容易出现性能骤降,尤其在低能见度与复杂海况下更为明显,例如在利用海上无人船(USV)对潜艇尾迹进行遥感等复杂感知任务中,环境因素对性能的影响尤为突出。未来的发展方向包括建立更大规模、更高多样性、覆盖多气候与多任务的综合性数据集,以及推动跨模态标注与自监督学习数据集的建设,例如推动更多类似 USVTrack 的雷达—视觉融合数据集的开发与应用。

### 3.3 USV 目标检测的方法

#### 3.3.1 传统计算机视觉局限性

早期 USV 视觉感知系统大多依赖人工特征提

取与基于规则的分类器。例如,边缘检测(Canny、Sobel)、颜色阈值分割、背景差分以及基于运动的光流法等广泛应用于船舶检测与障碍物识别。这些方法的优势在于计算开销低、实现简单、对小规模数据依赖较少。然而,它们在多变海况下的鲁棒性有限:海面波浪反射、目标与背景颜色接近、低光照及恶劣天气都会导致特征提取失败。此外,传统方法在多目标跟踪、动态背景抑制等任务上往往需要额外的手工调参,缺乏可扩展性与自适应性。

### 3.3.2 深度学习在检测、分割、跟踪的优势

在USV视觉感知任务中,深度学习与传统方法的性能差异及适用场景可通过技术特性与海洋环境的适配性进一步明确:传统方法依赖人工设计特征(如边缘检测、颜色分割),虽然计算开销低、对数据量需求小,适用于计算资源有限且环境稳定的特定场景,但难以应对海洋环境的高动态性(如浪涌、光照突变)与目标多样性(如多尺度船只、未知障碍物),在复杂海况下易出现特征失效与误检漏检问题。

相比之下,深度学习通过端到端特征学习框架,能从原始数据中自动提取多层次、判别性强的特征表示,在多任务中展现显著优势:在目标检测领域,卷积神经网络(CNN)架构(如Faster R-CNN、YOLO、SSD)已在SMD、SeaShips等海事数据集上实现高性能,其中单阶段的YOLO系列表现突出——YOLOv5在SeaDronesSee/SMD数据集上mAP约0.943,YOLOv8进一步提升至0.952,多阶段的Faster R-CNN、Mask R-CNN在SMD、SeaShips上mAP分别达到0.71、0.852(Zuo等,2025);在分割任务中,FCN、U-Net及其变体可精确分离船体与海面边界,即使在水面反光、波浪干扰下仍保持高精度;多目标跟踪任务中,深度特征结合LSTM、Transformer等时序建模框架,能有效减少ID切换、提升轨迹连续性;跨模态感知场景下,通过视觉与雷达数据的特征级融合,还可增强低能见度、恶劣天气下的检测跟踪鲁棒性。

这些深度学习方法的优越性能多基于可控场景与充分标注样本,在真实海况中仍存在与模型结构强相关的瓶颈:其一,面对尺度强变化目标(如近距大型船与远距小目标),YOLO系列依赖的预设锚框易超出海面目标尺度分布范围,小目标特征易被海浪杂波掩盖,而Faster R-CNN虽然通过RPN提升小

目标检出率,但两阶段推理导致实时性下降,难以适配USV高速航行的在线检测需求(Zuo等,2025);其二,在船体剧烈摇摆与光照突变场景,CNN的局部感受野特性易引发目标边界框偏移,Transformer检测模型虽然能凭借全局注意力抑制背景扰动,但其自注意力机制计算量高,在嵌入式GPU上难以稳定满足 $\leq 200$  ms的在线时延约束(Zuo等,2025;Carion等,2020);其三,针对未知类别障碍物(如漂浮集装箱、油污团)存在“开集”识别难题,监督学习模型对固有类别依赖度高,易出现漏检误检,而自监督与开放集学习方法虽然有改进,却因缺乏海事专属预训练任务与域移适配策略,泛化能力仍显不足(Bovcon和Kristan,2022;Guo等,2023)。

因此,USV目标检测技术的优化需摆脱“基准数据集依赖”,围绕海洋环境动态特性(浪涌导致的目标位移、光照随机性)与模型结构局限(锚框适配性、推理效率、开集泛化能力)展开针对性设计;随着硬件加速与模型压缩技术的发展,深度学习在USV感知系统中的应用范围有望进一步扩大,逐步在主流感知任务中取代传统方法,但其工程化落地仍需突破真实海况下的鲁棒性与实时性瓶颈。

### 3.3.3 USV目标检测的方法分类和性能评估

目标检测技术在USV视觉感知中主要分为单阶段和多阶段两类方法。单阶段方法以YOLO系列为代表,从早期的YOLOv2-v4(Lee等,2018;Li等,2021;Wang等,2020)到最新的YOLOv5-v8、YOLOX等(Kim等,2022;Ding等,2023;Ruiz-Ponce等,2023),在USV数据集中应用最为广泛。此外,SSD(Soloviev等,2020;Li等,2021)和FCOS(He等,2022)等单阶段检测器也被用于USV目标检测任务。虽然此类单阶段检测算法在通用数据集上表现优异,但面对海面小目标时仍存在漏检问题。针对类似的小目标检测难点,国内学者也提出了注意力机制与SSD结合的轻量化改进方案(贾可心等,2022),有效提升了复杂背景下的检测精度。近年来,研究者开始关注模型改进,如Zhang等人(2023)通过多尺度特征融合改进YOLOv7-Tiny,Zhao等人(2024)在YOLOv8s中集成卷积块注意力和自注意力机制,Cai等人(2024)提出FE-YOLO架构,结合通道注意力机制和GhostNet的Ghostconv进行网络层聚合。Wang等人(2024a)提出MDD-ShipNet架构,采用极化自注意力和加权双向特征金字塔网络进行

检测,并集成CNN去雾滤波器。Guo等人(2023)将注意力和Transformer机制融入目标检测,在SMD数据集上取得了更好的检测效果。

多阶段方法主要包括Faster R-CNN(Soloviev等,2020)、Mask R-CNN(Nie等,2020)、FPN(Soloviev等,2020)等,但在USV领域应用相对较少。WaSR-T(Žust和Kristan,2022)引入时间上下文进行水面环境检测。为量化评估深度学习模型在USV场景下的检测性能,现有研究普遍采用平均精度(average precision, AP)、平均精度均值(mean average precision, mAP)、精确率(precision)、召回率(recall)以及F1分数(F1-score)等国际通用指标(Everingham等,2010)。其中,mAP能够综合反映模型在多类别检测中的整体性能,而F1分数则平衡了精确率与召回率,常用于评估模型在漏检与误检之间的权衡能力。尽管这些方法在SMD、SeaShips等USV数据集上表现优异(多数超过80%),但其泛化能力仍需更广泛的USV视觉数据集中验证。当前USV目标检测技术发展趋势显示,单阶段方法比多阶段方法更受青睐,同时序列模型和基于注意力的视觉Transformer也逐渐被探索用于USV目标检测任务。

图6展示了USV目标检测中深度学习技术的发展历程。从图中可以看出,海上无人船目标检测模型的发展趋势明显倾向于单阶段方法,而非多阶段方法。近年来,研究者开始探索序列模型和基于注意力的视觉Transformer在USV目标检测任务中的应用。此外,当前USV领域深度学习技术的发展主要基于单阶段和多阶段方法的通用架构。用于一般目标检测的常见架构包括YOLO、FCOS(fully convolutional one-stage)、SSD(单阶段)以及Faster R-CNN和Mask R-CNN(多阶段)。这些方法通常以SMD和SeaShips等知名USV视觉数据集作为性能基准。尽管这些方法在基准数据集上表现出色,许多方法的性能指标超过80%,但其泛化能力尚未在更广泛的USV视觉数据集中得到充分验证。这引发了一个重要问题:当前开发的目标检测方法能否有效推广到USV领域中的其他目标检测数据集?

基于现有数据集训练的感知算法已实现复杂海况下的目标检测与环境建模,而如何将这些感知结果转化为安全、高效的导航控制指令,以及深度学习在USV控制系统中的落地路径,将在第4节详细阐述。

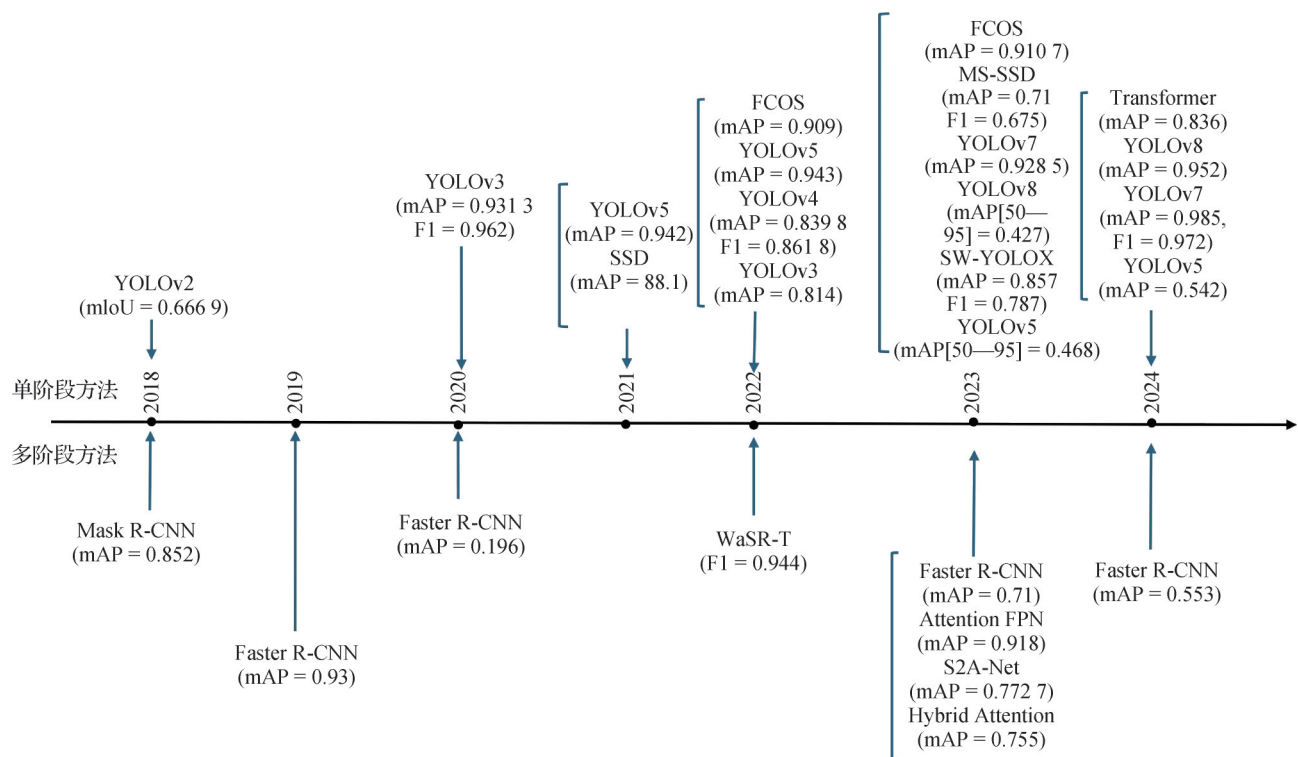


图6 目标检测的深度学习技术趋势及其在无人艇场景中的性能

Fig. 6 Trends of deep learning techniques for object detection and its performance in USVs scenarios

## 4 深度学习在USV控制系统中的应用

### 4.1 导航系统中的应用

#### 4.1.1 环境感知与状态估计

在USV的导航系统中,核心目标是基于采集到的数据向制导与控制模块提供环境信息与自身状态信息(Zheng等,2013)。典型做法是:舰载与岸基/空基传感器协同,既覆盖近场的岸线、船只与漂浮物,又补充远场的港区与航道态势(如AIS(automatic identification system)(Tu等,2018)、SAR图像(Jin等,2020)以及来自其他平台的视觉数据),形成面向导航任务的态势感知数据面。在此框架下,环境感知负责测量波浪、海流、风与天气等外部因素,以及实时适应包含大量静/动态障碍的复杂海面场景;状态估计则在噪声与不确定性条件下重建位置、姿态、速度与加速度等系统状态,并结合先验模型持续校正。

然而,海事环境中的高维多源数据(光学/红外图像、雷达图、点云等)在时间与空间分辨率、格式与几何对齐方面差异显著,给一致性建模与在线处理带来挑战,这就需要压缩、降维与融合等处理链条来提升数据质量与可用性。同时,基于GNSS与IMU的测量常含有噪声与误差,可能诱发状态估计失败;近年来相机(Chen等,2018)、雷达与LiDAR以及远程SAR的引入显著改善了估计精度,但风、浪、天气与通信条件仍会在开阔海域场景中持续施压,使得高可靠状态估计成为长期难点。从方法论演进看,深度学习在环境感知中按数据源类型开展:对光学/红外图像,需处理昼夜与云层敏感、分辨率高而实时性受限等问题;对雷达图像与点云,则强调抗恶劣天气能力与几何结构表达。这一“按源建模”的思路与传统滤波/估计链配合,使导航系统能够在复杂海况中更稳定地向下游制导与控制提供状态与环境先验。

#### 4.1.2 多传感器融合定位

就近场导航而言,自感(proprioceptive)与外感(exteroceptive)传感器的体系化组合,是维持鲁棒定位与态势理解的现实路径:前者(GPS/GNSS、IMU/INS)追踪平台位置、运动与里程计;后者(相机、LiDAR、雷达、超声)感知周边环境、风浪天气与外部障碍。在系统层面,导航感知—融合—定位通常遵

循“多模态数据的时间同步、空间配准与统一表征,以及特征层融合或决策层融合”的处理流程,以提升系统在低能见度、海雾和强反射等复杂场景下的鲁棒性。具体到USV视觉任务,相机—雷达融合已被用于目标检测、可通行区域分割、水线分割与雷达点云分割等多任务联合;在模型结构上,序列雷达帧的特征时序融合、图像引导的雷达特征、以及跨模态的Transformer交互均已用于提升检测与分割的综合表现。但在真实海事环境中仍面临若干瓶颈:跨模态标定误差与视角失配导致空间错配、雷达点云稀疏与低分辨率限制精细分割、小目标(远距或弱反射)检测依然困难、以及在拥挤水域中海杂波与多船反射引发的噪声敏感性与动态杂波滤除不稳等问题。

在USV相机—LiDAR—雷达三模态融合定位系统中,不同融合策略针对融合链路中不同层级的问题进行设计,其适用场景存在显著差异,简单的特征拼接往往难以充分利用不同模态的优势。鉴于此,提出了一种基于动态特征选择的融合网络,能够在红外与可见光特征之间进行自适应筛选,这种机制对于提升USV在夜间或复杂光照下的检测鲁棒性尤为重要。概率数据关联(probabilistic data association, PDA)方法聚焦于数据关联层的关键挑战,主要解决海杂波干扰下的目标匹配问题。在海洋环境中,雷达传感器易受海浪反射影响而产生大量虚假点云,相机传感器在雾、雨等恶劣天气条件下可能因水汽遮挡导致目标检测结果出现断连或缺失,LiDAR点云在浪涌环境下同样面临噪声干扰。PDA方法通过计算各传感器观测数据与目标航迹的关联概率,建立概率模型描述观测与目标之间的匹配关系,从而有效筛选并剔除杂波数据,保留高置信度的目标关联。该方法在近岸高海况环境或港口密集船舶场景中表现优异,能够有效应对海洋环境特有的干扰源(Zhong等,2021)。

区域提议网络(region proposal network, RPN)与度量学习方法则主要作用于特征融合层,针对跨模态特征不一致这一核心问题进行优化。相机、LiDAR和雷达3种传感器从本质上提取的是不同维度的特征信息:相机主要捕获目标的纹理特征、颜色特征和表观特征;LiDAR提供精确的几何特征、形状特征和空间结构信息;雷达则侧重于距离特征、速度特征和反射强度信息。这种固有的模态差异导致不同传感器提取的特征在特征空间中分布不一致,直

接进行特征融合往往难以获得理想的融合效果。RPN方法通过生成跨模态统一的目标候选区域,为后续的特征对齐和融合提供空间约束。度量学习技术则通过设计合适的距离度量函数和损失函数(如对比损失函数、三元组损失函数),将不同模态的特征映射到统一的特征空间,使得同一目标的跨模态特征在该空间中具有较高的相似度,而不同目标的特征则具有较低的相似度。这种特征空间的对齐和统一显著提升了多模态融合的精度和鲁棒性,该策略更适用于对目标定位精度要求较高的场景,如狭窄航道避障、精确对接任务等(Baumann等,2024)。

小样本学习(few-shot learning)方法则主要针对数据稀缺场景进行设计,解决远海、极地等特殊海域难以采集大规模标注数据的问题。与近岸、港口等常规作业区域不同,远海、极地等特殊环境下的USV作业数据采集成本高昂,且标注工作面临更多挑战,导致可用的高质量标注数据极其有限。小样本学习方法通过迁移学习策略,利用在近海等常规环境中已标注的大规模多模态数据作为预训练样本,使模型学习到多模态数据的一般性特征表示和融合规律。在此基础上,再利用小样本学习算法(如元学习算法、原型网络等)对模型进行快速适配,仅需少量远海样本即可实现融合模型在新环境下的有效迁移。该方法避免了因数据不足导致的融合精度显著下降问题,主要用于远海资源勘探、极地科学考察等特殊任务场景(Agrawal等,2023)。

总体而言,上述3种融合策略分别从数据关联层、特征融合层和数据适应层完善融合链路,针对不同层级的技术瓶颈提供了相应的解决方案。在实际应用中,需要根据具体的任务场景、环境条件和性能要求选择适配的融合方案,而非简单地叠加使用多种策略。

## 4.2 制导与路径规划

在USV的制导环节,路径规划可分为全局与局部两个层面:全局路径规划侧重于在静态或长时间尺度的环境知识(如历史航迹、静态障碍/岸线地图)上,从起点到终点生成一条(近)最优的无碰撞路径;局部路径规划则面向动态与不确定环境,在执行过程中根据实时传感与交通互动做出调整与避碰决策。这一两级框架已成为海上自主航行任务的通用共识:前者在任务前或缓变环境下给出宏观航线,后者在拥挤水域、复杂港内或航道等场景中承担避碰

与微调职责。此外,真实任务往往受地理、尺度、运动学与动力学等多重约束制约(如岸线/礁岩、船体几何尺寸、速度/加速度范围、惯性力矩等),这些约束在不同空间尺度的规划中权重不同:远洋跨港航行主要受地理约束主导;中尺度(如航道/峡道)需显式考虑船体形状;近岸精确操纵(如靠泊)则必须同时纳入运动学与动力学约束(Zhou等,2020)。这些定义、层级与约束,为后续的数据驱动与学习型方法提供了问题刻画的基础框架。

### 4.2.1 全局路径规划与数据驱动趋势

传统的确定性/启发式搜索(如可保证存在解时的全局搜索或近似最优的启发式法)仍是全局规划的基石。但随着AIS等大规模历史数据的可用性提升,如何利用海量轨迹学习“经验航路”、“能效最优路径”与“多任务最优访问序列”成为新课题:已有工作用深度网络在冰区将冰浓度—船速—能效指标的关系学习后,联合速度控制求得能效最优航线(Zhang等,2019);用DNN/LSTM对未来轨迹进行多步预测,为全局规划提供先验(Gao等,2021);在多水质监测点任务中,将任务抽象为旅行商问题,利用自组织映射等DNN学习站点位置与最优访问序列的关系,以降低总成本。这些研究展示了深度学习在处理非线性环境—性能关系、从历史轨迹抽取模式以及多目标调度上的优势,是对传统全局规划的功能性扩展。

### 4.2.2 局部路径规划与规则约束避碰。

在拥挤的港内或狭窄航道环境中,USV需要遵循国际海上避碰规则(international regulations for preventing collisions at sea, COLREGs),并与周边本船(own ship, OS)和目标船(target ship, TS)进行实时交互决策。典型的船舶相遇场景可分为对遇、直航、让路与追越4类。在生成局部避碰路径之前,如何在多船相遇的动态环境中正确识别相遇情形并实时做出符合COLREGs规则的决策是核心难点(Zhao和Roh,2019)。传统基于模型的避碰方法在多目标与强动态环境下常面临计算复杂度高、预定义架构依赖过度简化的风险评估与船舶动力学假设、控制律难以在新情形下自适应扩展等瓶颈问题。面向这些挑战,深度强化学习(DRL)与监督式深度神经网络(DNN)成为两条重要的技术路径。DRL方法通过设计奖励函数在未知环境中通过交互学习实现安全航行与规则遵循的避碰策略(Shen等,2019);DNN

类方法则通过输入结构设计适配多船相遇场景,并利用长短期记忆网络(LSTM)或序列条件生成对抗网络(seq-cGAN)等架构,学习类似人类驾驶员经验的避碰模式(Gao和Shi,2020)。

对于DRL方法而言,奖励函数的设计是连接USV避碰任务的复杂约束与学习目标的关键桥梁。在USV这一特定应用场景下,奖励函数需要综合考虑多个关键要素:1)安全距离约束是奖励函数设计的核心要素之一,需要实时监测USV与障碍物、目标船之间的距离,当距离小于安全阈值时给予负奖励以惩罚潜在碰撞风险,反之则给予正奖励鼓励保持安全距离;2)航向偏差控制需要纳入奖励机制,包括与期望航向的偏差以及为规避碰撞而进行的航向调整成本,既要鼓励USV朝向目标航点,又要平衡避碰所需的临时航向改变;3)能量消耗优化应考虑推进系统的能耗特性,频繁的速度变化、大幅度的航向调整都会导致能量消耗增加,奖励函数需要在此类能量效率与避碰紧迫性之间建立权衡机制;4)COLREGs规则违反惩罚是USV避碰任务区别于一般机器人导航的关键特征,奖励函数必须将COLREGs中的规则条款(如相遇场景下的右行原则、交叉相遇下的让路优先级、追越场景下的安全超越约束等)转化为可量化的奖惩信号,对违反规则的行为施加严厉惩罚,对合规行为给予正面激励。此外,路径跟踪误差、完成任务进度等任务级目标也需要纳入奖励函数设计,形成多目标协同的奖励机制。

将COLREGs规则编码为DRL奖励函数的核心逻辑,是将规则中的对遇、追越、交叉相遇等离散场景转化为可量化的连续奖惩信号,实现规则遵循性与避碰性能的协同优化。Shen等人(2019)在USV避碰DRL框架中,将COLREGs规则拆解为安全会遇距离、行动优先级、航向航速调整约束3类指标。当目标船处于对遇场景时,USV保持右行且与目标船保持充足安全距离给予正奖励,违反右行规则或距离过近则给予负奖励;在追越场景中,若USV从目标船正后方狭窄扇区超越则触发惩罚信号,从左侧安全扇区超越且保持稳定速度差则给予正向奖励。Gao和Shi(2020)进一步引入规则满足度系数,将COLREGs中避免横越他船前方、及时采取避碰行动等较为模糊的规则条款转化为连续值奖励,该系数在完全合规时达到最大值,触发违规阈值时降至最小值,并与路径跟踪误差奖励进行加权融合。这

类编码策略在多船相遇仿真场景中实现了COLREGs合规率和避碰成功率的显著提升,并在真实港域AIS数据迁移测试中验证了其有效性(Shen等,2019;Gao和Shi,2020)。

然而,奖励函数设计在USV避碰任务中面临显著的奖励塑造(reward shaping)难题,这些难题主要体现在以下几个方面:将COLREGs规则转化为可计算的奖励信号往往依赖人工经验和领域专家知识,例如及时采取避碰行动的时间阈值需要根据海况、能见度等环境因素进行调整,缺乏自适应机制导致在不同环境条件下性能不稳定。

另外,当多个COLREGs规则在特定场景下产生冲突时(如交叉相遇场景中同时需要考虑让路优先级与避碰紧迫性),奖励函数可能出现梯度消失或梯度冲突,导致模型收敛速度显著下降,难以学习到同时满足多个规则约束的最优策略。此外,将多个关键要素(安全距离、航向偏差、能量消耗、规则遵循等)全部纳入奖励函数设计,会导致奖励信号维度高、训练样本需求大,相比简化的无规则约束模型,需要更长的训练周期才能达到稳定性能,增加了实际应用的难度。

不同奖励项之间的权重比例对最终学习策略具有决定性影响,但在USV避碰这一安全关键应用中,如何科学地分配权重缺乏理论指导,往往需要通过大量试错实验确定,且权重选择对泛化性能的影响机制尚不明确。

这些奖励塑造难题严重制约了DRL方法在USV避碰任务中的实际应用效果,需要研究更加智能的奖励函数设计方法,如基于逆强化学习的奖励函数自动学习、基于课程学习的渐进式奖励设计、以及结合领域知识的混合奖励架构等,在保证安全性和规则遵循性的前提下提升训练效率和策略质量。

#### 4.2.3 “学习+规则”的融合实践。

在规则遵循层面,学习方法用于补充与增强基于规则/势场/搜索的传统框架:一方面,已有工作将A\*、人工势场(artificial potential field, APF)与DQN结合,在局部动态环境中实现符合COLREGs的避碰与路径修正;通过势场改造动作空间与奖励设计,提高DRL在全局一局部耦合任务中的性能(Liu等,2017)。另一方面,在编队/队形保持与多USV协同中,受限A\*与Leader-Follower等策略被用于兼顾队

形稳定与路径可行性(包含“基于USV不同机动响应时间”的约束 $A^*$ 与势场耦合),以降低跟踪误差与实现多平台协同导航(Zheng和Sun,2017)。上述“规则可解释性+学习适应性”的结合,成为工程上从可验证性走向灵活智能的常见落地路径。

#### 4.2.4 开放问题与发展方向。

在多USV协同决策任务中,图神经网络(GNN)通过图结构建模智能体间的交互关系,为分布式协同控制提供了一种数据驱动的解决方案。GNN将系统中的每个USV抽象为图节点,将USV间的通信链路和任务协作关系抽象为边,构建了一个动态的拓扑结构(Kipf和Welling,2017)。边权重可根据通信信号强度、任务依赖度等动态因素进行自适应调整,以反映智能体间交互强度的实时变化。通过图卷积操作,GNN能够聚合邻域智能体的状态信息(包括位置、任务进度、环境感知数据等),实现对智能体间动态交互关系的量化建模(Battaglia等,2018)。该方法特别适用于海洋环境中USV位置实时变化、通信链路时断时续的动态场景,能够有效捕捉多智能体间的关联依赖与协作逻辑(Zhang等,2021c)。

与传统共识算法相比,GNN在拓扑适应性和学习能力方面展现出显著优势。传统共识算法(如基于一致性协议的分布式控制方法)依赖预设的固定通信拓扑结构,在通信链路频繁切换的复杂环境中易出现收敛速度慢、抗干扰能力弱等问题(Olfati-Saber等,2007)。GNN方法无需预设固定拓扑,能够从交互数据中自主学习关联特征,降低了对人工规则设计的依赖性(Jiang和Lu,2021)。在分布式任务分配、动态编队调整等复杂协同任务中,GNN展现出更强的自适应能力和泛化性能(Sunehag等,2018)。在通信拓扑动态变化的场景下,基于GNN的协同策略较传统共识算法在收敛速度和鲁棒性方面具有明显提升(Wang等,2023b)。

然而,GNN方法在多USV协同应用中仍面临计算复杂度和数据需求的挑战。图卷积操作的计算复杂度随USV数量呈指数级增长,在USV规模较大时,相比轻量级传统共识算法存在实时性不足的问题(Gilmer等,2017)。此外,GNN需要大量多智能体交互样本进行训练,在多USV协同的小样本场景下泛化性能显著下降,难以直接适配远海等数据采集困难的任务环境(Foerster等,2018)。为缓解上述问

题,研究提出了轻量级图注意力网络、迁移学习等技术方案,在保证协同性能的同时降低计算复杂度,提升模型在资源受限条件下的适用性(Veličković等,2018;Yang等,2022)。

### 4.3 感知与决策的协同与反馈

随着USV系统日益复杂化,传统“感知—决策”分离式架构的局限性愈发明显。感知模块的误差与延迟会直接威胁到决策系统的安全性和最优性。因此,研究二者的协同与反馈机制,乃至实现一体化设计,已成为领域内的重要趋势(Yao等,2024;Chen等,2022)。

#### 4.3.1 基于感知不确定性的风险评估与安全决策

传统的决策控制策略往往基于一个理想假设,即感知输入是准确且可靠的。然而,深度学习感知模型在面对恶劣海况、低能见度或罕见目标时,其输出必然带有不确定性(Feng等,2021)。将这种不确定性纳入决策框架是提升USV系统鲁棒性和安全性的关键。

量化感知的不确定性是前提,主流方法包括利用贝叶斯神经网络(Bayesian neural network, BNN)估计模型参数的后验分布,或通过蒙特卡洛Dropout(Monte Carlo dropout, MC-dropout)在推理时进行多次随机前向传播以获取预测方差(Gal和Ghahramani,2016;Kendall和Gal,2017)。这些不确定性度量随后集成到决策框架中,例如作为模型预测控制(model predictive control, MPC)或强化学习(RL)中的风险成本项(Michelmore等,2020),指导系统在感知不确定性高时采取更保守的策略。

例如,在目标检测任务中,模型会输出每个检测框的类别和位置置信度。决策系统应利用这些置信度作为风险评估的量化依据:当关键障碍物的检测置信度较低时,决策模块应触发更保守、更谨慎的行为,例如立即减速、扩大安全距离或进入紧急预警模式(Gu等,2022)。这种将“感知不确定性”直接转化为“决策风险成本”的机制,有效地弥补了感知与决策之间的信息鸿沟,是实现安全关键型自主航行的重要趋势。

#### 4.3.2 决策需求驱动下的预测性感知任务

决策控制任务对感知系统提出了超越当前时刻的预测性要求。在动态避碰和协同航行场景中,USV的有效决策不能仅依赖于目标当前的航向和速度,而必须预判其他动态物体的短期未来行为,即意

图识别与轨迹预测(Rasouli和Tsotsos,2020)。

意图识别旨在从观测到的运动模式、环境信息及历史数据中推断出目标的潜在航行目的(例如,是否进行捕鱼作业、转向或保持航向)。在此基础上,轨迹预测则结合运动学和动力学模型,利用如图神经网络(GNN)或长短期记忆网络(LSTM)等深度模型,生成目标未来数秒内最可能的运动路径(Ai-Molegi等,2022)。决策模块随后可基于这些预测信息,采用模型预测控制(MPC)或基于预测状态的强化学习等方法,进行主动、平滑、远距离的优化路径规划和避碰操作(Williams等,2018),从而避免被动反应,极大地提高了USV在复杂交通流中的导航效率和安全性。

#### 4.3.3 感知—决策一体化的端到端策略

作为感知与决策耦合的极致形式,端到端的自主航行策略(end-to-end navigation)近年来受到了广泛关注(Chen等,2022;Shah等,2023)。该方法试图构建一个单一的深度神经网络,直接将原始传感器数据(如摄像头图像、LiDAR点云)映射到船舶的控制指令(如期望的航向和速度),从而绕过了传统架构中显式的环境建模、路径规划等中间步骤。

早期的端到端研究(如ALVINN项目(Pomerleau,1989))验证了这一思想的可行性。近年来,得益于深度强化学习(DRL)的飞速发展,端到端模型能够在模拟环境中通过海量的交互训练,学习到从感知到控制的复杂映射关系(Qiao等,2022)。这种方法的优势在于减少了模块化设计中因人工特征提取和接口定义而导致的信息损失,理论上能实现全局最优。然而,端到端策略也面临着可解释性差、数据依赖性强(尤其是需要海量真实场景的标注数据)以及在复杂或罕见场景下泛化能力不足等严峻挑战,目前在安全关键的USV应用中仍处于探索阶段。

## 5 结 语

本文系统综述了基于深度学习的面向复杂海洋环境的无人船智能感知与决策关键技术,深入分析了海洋无人系统在高动态、不确定环境中实现可靠自主作业所面临的核心挑战。通过对国内外相关领域研究的梳理,本文回顾了海上无人船的发展历程与体系架构,从军事起源到多领域应用拓展,明确了

其在智能海洋装备中的重要地位,并对其船体平台、动力系统、通信控制、传感器配置及NGC软件结构进行了全面总结,为理解USV系统的复杂性与技术演进提供了结构化框架。在智能感知技术方面,本文详细分析了深度学习模型(如CNN、Transformer、GNN、DRL等)在海上目标检测、障碍物识别、海况评估与多目标跟踪任务中的应用现状,突出了多模态传感器融合在提升感知鲁棒性方面的关键作用。结合现有海事视觉数据集(如MODD、MaSTr1325、SeaDronesSee等),本文指出当前算法在跨域泛化、恶劣天气适应性及实时推理能力方面仍存在显著瓶颈。在决策与控制方面,本文探讨了感知输出如何驱动决策与控制,分析了结合学习与规则的导航与避碰方法,以及多USV协同与群体智能在复杂任务中的潜力,强调了其在提升系统效能和应对通信受限场景中的独特优势。

尽管取得了一些进展,当前技术仍面临若干挑战:

1)数据质量与多样性不足。现有海事数据集多集中于良好天气和常见场景,极端条件(如大雾、巨浪、夜间)下的样本较为稀缺,限制了模型的泛化能力和稳健性。

2)多模态融合技术瓶颈。多传感器融合技术在实际应用中受到时空标定偏差、模态异质性和融合机制适应性不足等问题的制约。

3)仿真与现实差距。许多学习方法(如DRL)依赖仿真训练,仿真与现实间的差距导致策略迁移困难,且模型的可解释性与实时部署能力有待提高。

4)协同决策研究不足。面向多船协同的分布式决策与通信可靠性研究尚处于初步阶段,缺乏系统化的理论框架与标准验证体系。

基于本文讨论,未来的研究可从以下几个方向展开:构建涵盖风浪流耦合多海况参数、USV运动状态及多模态感知原始数据,并强化雾、涌浪等非理想场景标注的专属动态海洋数据集,解决数据与实际作业脱节问题;研发嵌入海洋物理先验(如基于PINN融合海浪预测模型实现运动补偿)的轻量化感知—决策模型,量化感知精度、推理延迟与车载电池续航的耦合关系,建立多目标优化准则以平衡能耗与性能;设计COLREGs导向的“感知—决策”闭环融合体系,模态失效时动态调整融合权重并配套硬件冗余,同时构建结合符号推理与深度学习的可解

释混合决策框架及安全验证指标;突破 USV 仿真—实物迁移瓶颈,搭建“风浪流—船体动力学”耦合高保真仿真环境,通过域自适应等迁移学习优化,建立“仿真预训练—实船微调—海域验证”范式。通过系统性的技术综述,本文为复杂海洋环境下的自主 USV 体系构建提供了全面的技术参考与启示,随着人工智能、传感器技术和通信技术的不断发展,USV 智能感知与决策技术将在海洋资源开发、海上安全保障和海洋科学研究等领域发挥越来越重要的作用。

## 参考文献 (References)

- Ab Rahman A, Hamid U Z A and Chin T A. 2017. Emerging technologies with disruptive effects: a review. *Perintis eJournal*, 7 (2) : 111-128
- Ahmed M, Bakht A B, Hassan T, Akram W, Humais A, Seneviratne L, He S M, Lin D F, and Hussain I. 2023. Vision-based autonomous navigation for unmanned surface vessel in extreme marine conditions [EB/OL]. [2025-09-01]. <https://arxiv.org/pdf/2308.04283.pdf>
- Al-Molegi A, Jabreel M and Martínez-Ballesté A. 2022. MoveSense: space-time trajectory prediction using deep neural networks. *Neural Computing and Applications*, 34 (13) : 10797-10816 [DOI: 10.1007/s00521-022-07005-7]
- Agarwala N. 2022. Integrating UUVs for naval applications. *Maritime Technology and Research*, 4 (3) : #254470 [DOI: 10.33175/mtr.2022.254470]
- Agrawal S, Bhandari S, Doycheva K and Elger G. 2023. Static multitarget-based autocalibration of RGB cameras, 3-D radar, and 3-D lidar sensors. *IEEE Sensors Journal*, 23 (18) : 21493-21505 [DOI: 10.1109/jsen.2023.3300957]
- Bakht A B, Din M U, Javed S and Hussain I. 2025. MVTD: a benchmark dataset for maritime visual object tracking [EB/OL]. [2025-09-01]. <https://arxiv.org/pdf/2506.02866.pdf>
- Balemans N, Anwar A, Steckel J and Mercelis S. 2024. Multi-modal sensor fusion in latent embedding space for robust autonomous navigation//Proceedings of 2024 IEEE SENSORS. Kobe, Japan; IEEE: 1-4 [DOI: 10.1109/sensors60989.2024.10785240]
- Bati S, Oğul H A, Karaçizmeli C and Tükel D. 2013. Human-computer interface for Doğuş Unmanned Sea Vehicle//Robotic Sailing 2012: Proceedings of the 5th International Robotic Sailing Conference. Berlin; Springer: 61-71 [DOI: 10.1007/978-3-642-33084-1\_6]
- Battaglia P W, Hamrick J B, Bapst V, Sanchez-Gonzalez A, Zambaldi V, Malinowski M, et al. 2018. Relational inductive biases, deep learning, and graph networks [EB/OL]. [2025-09-01]. <https://arxiv.org/pdf/1806.01261.pdf>
- Baumann N, Baumgartner M, Ghignone E, Kühne J, Fischer T, Yang Y H, et al. 2024. CR3DT: Camera-RADAR fusion for 3D detection and tracking//2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Abu Dhabi, United Arab Emirates: IEEE: 4926-4933 [DOI: 10.1109/IROS58592.2024.10801848]
- Bengio Y, LeCun Y and Hinton G. 2021. Deep learning for AI. *Communications of the ACM*, 64(7) : 58-65 [DOI: 10.1145/3448250]
- Bloisi D D, Iocchi L, Pennisi A and Tombolini L. 2015. ARGOS-Venice boat classification//Proceedings of 2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). Karlsruhe, Germany: IEEE: 1-6 [DOI: 10.1109/AVSS.2015.7301727]
- Bovcon B and Kristan M. 2022. WaSR—a water segmentation and refinement maritime obstacle detection network. *IEEE Transactions on Cybernetics*, 52 (12) : 12661-12674 [DOI: 10.1109/TCYB.2021.3085856]
- Bovcon B, Mandeljc R, Pers J and Kristan M. 2018. Stereo obstacle detection for unmanned surface vehicles by IMU-assisted semantic segmentation. *Robotics and Autonomous Systems*, 104: 1-13 [DOI: 10.1016/j.robot.2018.02.017]
- Bovcon B, Muhovic J, Pers J and Kristan M. 2019. The MaSTr1325 dataset for training deep USV obstacle detection models//Proceedings of 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Macau, China: IEEE: 3431-3438 [DOI: 10.1109/IROS40897.2019.8967909]
- Bovcon B, Muhovic J, Vranac D, Mozetic D, Pers J and Kristan M. 2022. MODS—A USV-oriented object detection and obstacle segmentation benchmark. *IEEE Transactions on Intelligent Transportation Systems*, 23 (8) : 13403-13418 [DOI: 10.1109/TITS.2021.3124192]
- Breivik M. 2010. *Topics in Guided Motion Control of Marine Vehicles*. Trondheim, Norway: NTNU
- Breivik M, Hovstein V E and Fossen T I. 2008. Straight-line target tracking for unmanned surface vehicles. *Modeling, Identification and Control*, 29(4) : 131-149 [DOI: 10.4173/mic.2008.4.2]
- Brock A, Donahue J and Simonyan K. 2019. Large scale GAN training for high fidelity natural image synthesis. *International Conference on Learning Representations (ICLR)*. [s.l.]: [s.n.]
- Briguglio G and Crupi V. 2024. Review on sensors for sustainable and safe maritime mobility. *Journal of Marine Science and Engineering*, 12(2) : #353 [DOI: 10.3390/jmse12020353]
- Cai S W, Meng H and Wu J B. 2024. FE-YOLO: YOLO ship detection algorithm based on feature fusion and feature enhancement. *Journal of Real-Time Image Processing*, 21 (3) : #61 [DOI: 10.1007/s11554-024-01445-5]
- Calderón A J, González I and Calderón M. 2014. Management of a PEM electrolyzer in hybrid renewable energy systems//Matía F, Marichal

- G N and Jiménez E, eds. *Fuzzy Modeling and Control: Theory and Applications*. Amsterdam: Atlantis Press; 217-234 [DOI: 10.2991/978-94-6239-082-9\_12]
- Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, and Zagoryko S. 2020. End-to-end object detection with transformers//European Conference on Computer Vision (ECCV). Cham: Springer; 213-229 [DOI: 10.1007/978-3-030-58452-8\_13]
- Carrio A, Sampedro C, Rodriguez-Ramos A and Campoy P. 2017. A review of deep learning methods and applications for unmanned aerial vehicles. *Journal of Sensors*, 2017(1): #3296874 [DOI: 10.1155/2017/3296874]
- Chen L, Wu P, Chitta K, Jaeger B, Geiger A and Hong K. 2022. End-to-end autonomous driving: challenges and frontiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46 (1) : 204-227
- Chen X Q, Qi L, Yang Y S, Luo Q, Postolache O, Tang J J, et al. 2020. Video-based detection infrastructure enhancement for automated ship recognition and behavior analysis. *Journal of Advanced Transportation*, 2020 (1) : #7194342 [DOI: 10.1155/2020/7194342]
- Chen Y, Chen X D, Zhu J J, Lin F and Chen B M. 2018. Development of an autonomous unmanned surface vehicle with object detection using deep learning//IECON 2018—44th Annual Conference of the IEEE Industrial Electronics Society. Washington, USA: IEEE; 5636-5641 [DOI: 10.1109/IECON.2018.8591129]
- Chen Z T. 2025. Research on autonomous navigation and control of unmanned surface vehicles based on reinforcement learning algorithms and multi-objective optimization models//Proceedings of the 4th International Conference on Computer, Artificial Intelligence and Control Engineering. Hefei, China: ACM; 365-372 [DOI: 10.1145/3727648.3727709]
- Cheng B, Misra I, Schwing A G, Kirillov A and Girdhar R. 2022. Masked-attention mask transformer for universal image segmentation//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE; 1290-1299 [DOI: 10.1109/CVPR52688.2022.00135]
- Cheng Y, Jiang M, Zhu J and Liu Y. 2021. Are we ready for unmanned surface vehicles in inland waterways? The USVInland multisensor dataset and benchmark. *IEEE Robotics and Automation Letters*, 6(2): 3964-3970 [DOI: 10.1109/LRA.2021.3067271]
- Cheng Y W, Xu H and Liu Y M. 2021. Robust small object detection on the water surface through fusion of camera and millimeter wave radar//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, Canada: IEEE; 15243-15252 [DOI: 10.1109/ICCV48922.2021.01498]
- Cheng Y, Zhu J, Jiang M, Fu J, Pang C, Wang P, et al. 2021. FloW: A dataset and benchmark for floating waste detection in inland waters//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, Canada: IEEE; 10953-10962 [DOI: 10.1109/ICCV48922.2021.01077]
- Chung D, Kim J, Lee C and Kim J. 2023. Pohang canal dataset: A multimodal maritime dataset for autonomous navigation in restricted waters. *The International Journal of Robotics Research*, 42 (12) : 1104-1114 [DOI: 10.1177/02783649231191145]
- Corfield S J and Young J M. 2006. Unmanned surface vehicles—game changing technology for naval operations//Roberts G N and Sutton R N, eds. *Advances in Unmanned Marine Vehicles*. London, UK: IET; 311-328 [DOI: 10.1049/PBCE069E\_ch15]
- Coué C, Pradalier C, Laugier C, Fraichard T, and Bessière P. 2006. Bayesian occupancy filtering for multitarget tracking: an automotive application. *The International Journal of Robotics Research*, 25(1): 19-30 [DOI: 10.1177/0278364906061158]
- Cui Y D, Chen R, Chu W B, Chen L, Tian D, Li Y and Cao D P. 2021. Deep learning for image and point cloud fusion in autonomous driving: a review. *IEEE Transactions on Intelligent Vehicles*, 6(4): 722-739. [DOI: 10.1109/TIV.2021.3071291]
- Dağlı E, Tükel D and Civan C. 2013. Doğuş-USV unmanned sea vehicle: obstacle localization with stereo vision and path planning//Proceedings of the 11th International Conference (ETAI 2013). Ohrid, Macedonia: Society for ETAI
- Danescu R, Oniga F and Nedevschi S. 2011. Modeling and tracking the driving environment with a particle-based occupancy grid. *IEEE Transactions on Intelligent Transportation Systems*, 12(4): 1331-1342. [DOI: 10.1109/TITS.2011.2158097]
- DeFilippo M, Sacarny M and Robinette P. 2021. RoboWhaler: A robotic vessel for marine autonomy and dataset collection//OCEANS 2021: San Diego - Porto. San Diego, USA / Porto, Portugal: IEEE; 1-7 [DOI: 10.23919/OCEANS44145.2021.9705871]
- Dima C S, Vandapel N and Hebert M. 2004. Classifier fusion for outdoor obstacle detection//IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. New Orleans, USA: IEEE; 665-671 [DOI: 10.1109/robot.2004.1307225]
- Ding S H, Xu Y, Zhang Q X, Yu J, Sun T, Ni J, et al. 2023. Obstacle detection method based on RSU and vehicle camera fusion. *Sensors*, 23(10): #4920 [DOI: 10.3390/s23104920]
- El-Din D M, Hassanein A E and Hassanien E E. 2024. An adaptive and late multifusion framework in contextual representation based on evidential deep learning and Dempster-Shafer theory. *Knowledge and Information Systems*, 66 (11) : 6881-6932 [DOI: 10.1007/s10115-024-02150-2]
- Everingham M, Van Gool L, Williams C K I, Winn J and Zisserman A. 2010. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88(2): 303-338
- Fan S, Li Y, and He Y. 2019. Underwater target detection based on sonar imagery: a review. *IEEE Access*, 7: 132455-132468
- Fan Z X, Zhu Y Z, He Y L, Sun Q, Liu H Y and He J. 2023. Deep

- learning on monocular object pose detection and tracking: a comprehensive overview. *ACM Computing Surveys*, 55 (4) : #81 [DOI: 10.1145/3524496]
- Feng D, Rosenbaum L and Dietmayer K. 2021. A review and comparative study on probabilistic object detection in autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, 22 (5) : 3161-3180 [DOI: 10.1109/TITS.2020.2970766]
- Feng M, Li J and Zhang Y. 2022. Dynamic scene image deblurring and object detection for unmanned surface vehicles based on convolutional neural networks. *Ocean Engineering*, 250: #111054 [DOI: 10.1016/j.oceaneng.2022.111054]
- Foerster J, Farquhar G, Afouras T, Nardelli N and Whiteson S. 2018. Counterfactual multi-agent policy gradients//*Proceedings of the 32nd AAAI Conference on Artificial Intelligence*. New Orleans, USA: AAAI Press: 2974-2982.
- Gal Y and Ghahramani Z. 2016. Dropout as a Bayesian approximation: representing model uncertainty in deep learning//*International Conference on Machine Learning (ICML)*. New York, USA: PMLR: 1050-1059
- Gao D W, Zhu Y S, Zhang J F, He Y K, Yan K and Yan B R. 2021. A novel MP-LSTM method for ship trajectory prediction based on AIS data. *Ocean Engineering*, 228: #108956 [DOI: 10.1016/j.oceaneng.2021.108956]
- Gao M and Shi G Y. 2020. Ship collision avoidance anthropomorphic decision-making for structured learning based on AIS with Seq-CGAN. *Ocean Engineering*, 217: #107922 [DOI: 10.1016/j.oceaneng.2020.107922]
- Gilmer J, Schoenholz S S, Riley P F, Vinyals O and Dahl G E. 2017. Neural message passing for quantum chemistry//*International Conference on Machine Learning (ICML)*. Sydney, Australia: PMLR: 1263-1272
- Grishick R. 2015. Fast R-CNN//*Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile: IEEE: 1440-1448 [DOI: 10.1109/ICCV.2015.169]
- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, and Bengio Y. 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems*, #27
- Graham M M. 2008. Unmanned surface vehicles: An operational commander's tool for maritime security. Newport, USA: Naval War College Joint Military Operations Dept
- Grigorescu S, Trasnea B, Cocias T and Macesanu G. 2020. A survey of deep learning techniques for autonomous driving. *Journal of Field Robotics*, 37(3) : 362-386 [DOI: 10.1002/rob.21918]
- Groves P D. 2013. Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems. 2nd ed. Boston, USA: Artech House
- Guo M H, Lu C Z, Liu Z N, Cheng M M and Hu S M. 2023. Visual attention network. *Computational Visual Media*, 9 (4) : 733-752 [DOI: 10.1007/s41095-023-0364-2]
- Gu Y, Zhang X, Wang H and Cao D. 2022. Risk-aware decision making for autonomous vehicles based on perception uncertainty. *IEEE Transactions on Intelligent Transportation Systems*, 23(9) : 15462-15474.
- Guan R, Yao S, Liu R W and Yue Y. 2023. Multi-sensor fusion for robust environment perception of unmanned surface vehicles: a review and new perspectives. *IEEE Sensors Journal*, 23 (14) : 15234-15248 [DOI: 10.1109/JSEN.2023.3281023]
- Gundogdu E, Solmaz B, Yucesoy V and Koc A. 2017. MARVEL: A large-scale image dataset for maritime vessels//*Computer Vision - ACCV 2016 Workshops*. Cham: Springer: 165-180 [DOI: 10.1007/978-3-319-54193-8\_11]
- Han J, Liao Y P, Zhang J Y, Wang S F and Li S X. 2018. Target fusion detection of LiDAR and camera based on the improved YOLO algorithm. *Mathematics*, 6(10) : #213 [DOI: 10.3390/math6100213]
- Hashali S D, Yang S and Xiang X. 2024. Route planning algorithms for unmanned surface vehicles (USVs) : a comprehensive analysis. *Journal of Marine Science and Engineering*, 12 (3) : #382 [DOI: 10.3390/jmse12030382]
- Heo J, Kim J and Kwon Y. 2017. Analysis of design directions for unmanned surface vehicles (USVs). *Journal of Computer and Communications*, 5(7) : 92-100 [DOI: 10.4236/jcc.2017.57010]
- Huang Y, Wang J, Jiang C P and Zhang Y. 2020. A review of multi-sensor fusion in autonomous driving. *IEEE Access*, 8: 151516-151534 [DOI: 10.1109/ACCESS.2020.3016423]
- Iancu B, Soloviev V, Zelioli L and Lilius J. 2021. ABOships—An inshore and offshore maritime vessel detection dataset with precise annotations. *Remote Sensing*, 13(5) : #988 [DOI: 10.3390/rs13050988]
- Jang C, Kim J and Sunwoo M. 2024. Deep learning-based predictive dynamic occupancy grid mapping for autonomous driving. *IEEE Transactions on Intelligent Vehicles*, 9(2) : 3120-3132.
- Jiao L and Abdullah A M. 2024. A survey of deep learning-based object detection in maritime environment. *Ocean Engineering*, 293: #116645
- Jia K X, Ma Z H, Zhu R and Li Y G. 2022. Attention-mechanism-based light single shot multiBox detector modelling improvement for small object detection on the sea surface. *Journal of Image and Graphics*, 27(4) : 1161-1175 (贾可心, 马正华, 朱蓉, 李永刚. 2022. 注意力机制改进轻量SSD模型的海面小目标检测. *中国图象图形学报*, 27(4) : 1161-1175) [DOI: 10.11834/jig.200517]
- Jin K, Chen Y L, Xu B, Yin J J, Wang X S and Yang J. 2020. A patch-to-pixel convolutional neural network for small ship detection with PolSAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 58 (9) : 6623-6638 [DOI: 10.1109/TGRS. 2020. 2978268]
- Jiang J and Lu Z. 2021. Learning attentional communication for multi-agent cooperation. *Advances in Neural Information Processing Sys-*

- tems (NeurIPS), 31: 7254-7264
- Jung D W, Hong S M, Lee J H, Cho H J, Choi H S and Vu M T. 2018. A study on unmanned surface vehicle combined with remotely operated vehicle system. *Proceedings of Engineering and Technology Innovation*, 9: 17-24
- Karras T, Laine S and Aila T. 2019. A style-based generator architecture for generative adversarial networks//*Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, USA: IEEE: 4401-4410. [DOI: 10.1109/CVPR.2019.00453]
- Karras T, Laine S, Aittala M, Hellsten J, Lehtinen J and Aila T. 2020. Analyzing and improving the image quality of stylegan//*Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA: IEEE: 8110-8119 [DOI: 10.1109/CVPR42600.2020.00813]
- Karras T, Aittala M, Laine S, Härkönen E, Hellsten J, Lehtinen J, et al. 2021. Alias-free generative adversarial networks. *Advances in Neural Information Processing Systems*, 34: 852-863.
- Kaur P, Aziz A, Jain D, Patel H, Hirokawa J, Townsend L, et al. 2022. Sea situational awareness (SeaSAW) dataset//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. New Orleans, USA: IEEE: 2579-2587 [DOI: 10.1109/CVPRW56347.2022.00290]
- Kendall A and Gal Y. 2017. What uncertainties do we need in Bayesian deep learning for computer vision? *Advances in Neural Information Processing Systems (NeurIPS)*: #30
- Kim J H, Kim N, Park Y W, and Won C S. 2022. Object detection and classification based on YOLO-V5 with improved maritime dataset. *Journal of Marine Science and Engineering*, 10(3): #377 [DOI: 10.3390/jmse10030377]
- Kipf T N and Welling M. 2017. Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations (ICLR)*. [s.l.]: [s.n.]
- Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al. 2023. Segment anything//*Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris, France: IEEE: 4015-4026 [DOI: 10.1109/ICCV51070.2023.00371]
- Kristan M, Sulic Kenk V, Kovacic S and Pers J. 2016. Fast image-based obstacle detection from unmanned surface vehicles. *IEEE Transactions on Cybernetics*, 46(3): 641-654 [DOI: 10.1109/TCYB.2015.2412251]
- Krizhevsky A, Sutskever I and Hinton G E. 2017. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6): 84-90 [DOI: 10.1145/3065386]
- Lambert R, Chavez-Galaviz J, Li J and Mahmoudian N. 2022. ROSEBUD: A deep fluvial segmentation dataset for monocular vision-based river navigation and obstacle avoidance. *Sensors*, 22(13): #4681 [DOI: 10.3390/s22134681]
- LeCun Y, Bengio Y and Hinton G. 2015. Deep learning. *Nature*, 521(7553): 436-444 [DOI: 10.1038/nature14539]
- Lee C, Schätzle S, Lang S A and Oksanen T. 2023. Design considerations of a perception system in functional safety operated and highly automated mobile machines. *Smart Agricultural Technology*, 6: #100346 [DOI: 10.1016/j.atech.2023.100346]
- Lee J, Hwang J, Kim H and Park C H. 2018. Vision-based object detection for unmanned surface vehicles (USVs)//*2018 18th International Conference on Control, Automation and Systems (ICCAS)*. PyeongChang, Korea: IEEE: 1-4
- Lei L, Huang J and Zhou Y. 2024. Multimodal fusion-based spatiotemporal incremental learning for ocean environment perception under sparse observation. *Information Fusion*, 108: #102360 [DOI: 10.1016/j.inffus.2024.102360]
- Li L H, Zhou Z Q, Wang B, Miao L J and Zong H. 2021. A novel CNN-based method for accurate ship detection in HR optical remote sensing images via rotated bounding box. *IEEE Transactions on Geoscience and Remote Sensing*, 59(1): 686-699 [DOI: 10.1109/TGRS.2020.2995477]
- Lillicrap T P, Hunt J J, Pritzel A, Heess N, Erez T, Tassa Y, et al. 2016. Continuous control with deep reinforcement learning. *International Conference on Learning Representations (ICLR)*. [s.l.]: [s.n.]
- Li X L, Liu W Y, Li X S, Zhou F Q, Li H F and Nie F P. 2026. All-weather multi-modality image fusion: unified framework and 100k benchmark. *Information Fusion*, 131: #104130 [DOI: 10.1016/j.inffus.2026.104130]
- Lin J H, Zhu L, Shen J X, Fu H Z, Zhang Q and Wang L S. 2024. ViD-SOD-100: a new dataset and a baseline model for RGB-d video salient object detection. *International Journal of Computer Vision*, 132(11): 5173-5191 [DOI: 10.1007/s11263-024-02051-5]
- Lin X W, Guo Y A and Wang J Q. 2021. Global correlation network: end-to-end joint multi-object detection and tracking [EB/OL]. [2025-09-01]. <https://arxiv.org/pdf/2103.12511.pdf>
- Liu C, Chen C L P, Zou Z J and Li T S. 2017. Adaptive NN-DSC control design for path following of underactuated surface vessels with input saturation. *Neurocomputing*, 267: 466-474 [DOI: 10.1016/j.neucom.2017.06.042]
- Liu C X, Hu P J, Li S C, Wang C H and Ling Y. 2025. Decoupled triple-stage enhancement network for non-homogeneous image dehazing. *Journal of Image and Graphics*, 30(1): 83-94 (刘春晓, 胡鹏靖, 厉世昌, 王成骅, 凌云. 2025. 面向非均质图像去雾的解耦合三阶段增强网络. *中国图象图形学报*, 30(1): 83-94) [DOI: 10.11834/jig.240069]
- Liu W W, Liu Y C and Bucknall R. 2019. A robust localization method for unmanned surface vehicle (USV) navigation using fuzzy adaptive kalman filtering. *IEEE Access*, 7: 46071-46083 [DOI: 10.1109/ACCESS.2019.2909151]

- Liu W W, Liu Y C, Gunawan B A and Bucknall R. 2021. Practical moving target detection in maritime environments using fuzzy multi-sensor data fusion. *International Journal of Fuzzy Systems*, 23(6) : 1860-1878 [DOI: 10.1007/s40815-020-00963-1]
- Liu Z X, Zhang Y M, Yu X and Yuan C. 2016. Unmanned surface vehicles: An overview of developments and challenges. *Annual Reviews in Control*, 41: 71-93 [DOI: 10.1016/j.arcontrol.2016.04.018]
- Lu Z, Li B and Yan J. 2022. Research on unmanned surface vessel perception algorithm based on multi-sensor fusion//2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC). Qingdao, China: IEEE: 493-498 [DOI: 10.1109/ICFTIC57696.2022.10075187]
- Luiten J, Osep A, Dendorfer P, Torr P, Geiger A, Leal-Taixé L, et al. 2021. HOTA: a higher order metric for evaluating multi-object tracking. *International Journal of Computer Vision*, 129(2) : 548-578 [DOI: 10.1007/s11263-020-01375-2]
- Lyu Z C, Wang X Y, Wang G, Xing X F, Lyu C L and Yu F. 2025. Unmanned surface vessels in marine surveillance and management: advances in communication, navigation, control, and data-driven research. *Journal of Marine Science and Engineering*, 13(5) : 969 [DOI: 10.3390/jmse13050969]
- Malvadkar P, Bhosale A and Handore S. 2025. AI-driven multi-sensor fusion for autonomous robotic navigation. *Journal of Information Systems Engineering and Management*, 10(24s) : 246-257 [DOI: 10.52783/jisem.v10i24s.3893]
- Manley J E. 2008. Unmanned surface vehicles, 15 years of development//Proceedings of OCEANS 2008. Quebec City, Canada: IEEE: 1-4 [DOI: 10.1109/OCEANS.2008.5152052]
- Manley J E, Mineart G M and Sheridan A E. 2004. Underwater marine domain awareness for first responders: a low cost approach//Oceans '04 MTS/IEEE Techno-Ocean '04. Kobe, Japan: IEEE: 1695-1700 [DOI: 10.1109/oceans.2004.1406379]
- Marichal G N, Acosta L, Moreno L, Méndez J A, Rodrigo J J and Sigut M. 2001. Obstacle avoidance for a mobile robot: a neuro-fuzzy approach. *Fuzzy Sets and Systems*, 124(2) : 171-179 [DOI: 10.1016/S0165-0114(00)00095-6]
- Michelmore R, Wicker M, Laurenti L, Cardelli L, Kwiatkowska M, and Poloczek M. 2020. Uncertainty quantification with statistical guarantees in end-to-end autonomous driving control//2020 IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE: 7344-7350
- Mnih V, Kavukcuoglu K, Silver D, Rusu A A, Veness J, Bellemare M G, et al. 2015. Human-level control through deep reinforcement learning. *Nature*, 518 (7540) : 529-533 [DOI: 10.1038/nature14236]
- Nanda A, Cho S W, Lee H and Park J H. 2022. KOLOMVERSE: Korea open large-scale image dataset for object detection in the maritime universe [EB/OL]. [2025-09-01].  
<https://arxiv.org/pdf/2206.09885.pdf>
- Nie X, Duan M Y, Ding H X, Hu B L and Wong E K. 2020. Attention mask R-CNN for ship detection and segmentation from remote sensing images. *IEEE Access*, 8: 9325-9334 [DOI: 10.1109/access.2020.2964540]
- Nirgudkar S, DeFilippo M, Sacarny M, Benjamin M and Robinette P. 2023. MassMIND: Massachusetts Maritime INfrared Dataset. *The International Journal of Robotics Research*, 42 (1-2) : 21-32 [DOI: 10.1177/02783649231153020]
- Olfati-Saber R, Fax J A, and Murray R M. 2007. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1) : 215-233 [DOI: 10.1109/JPROC.2006.887293]
- Othman E H. 2015. A review on current design of unmanned surface vehicles (USVs). *Journal of Advanced Review on Scientific Research*, 16(1) : 12-17
- Panchal K, Korat A S and Pathak S R. 2025. Sensor fusion using machine learning for robust object detection in adverse weather conditions for self-driving cars. *International Journal of Computational and Experimental Science and Engineering*, 11(3) : #3589 [DOI: 10.22399/ijcesen.3589]
- Pandey J and Hasegawa K. 2015. Study on manoeuvrability and control of an autonomous Wave Adaptive Modular Vessel (WAM-V) for ocean observation//Proceedings of the 2015 International Association of Institutes of Navigation World Congress (IAIN). Prague, Czech Republic: IEEE: 1-7 [DOI: 10.1109/IAIN.2015.7352248]
- Patino L, Cane T and Ferryman J. 2021. A comprehensive maritime benchmark dataset for detection, tracking and threat recognition//2021 17th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). Washington, USA: IEEE: 1-8 [DOI: 10.1109/avss52988.2021.9663739]
- Pomerleau D A. 1989. ALVINN: an autonomous land vehicle in a neural network. *Advances in Neural Information Processing Systems (NIPS)*, #1
- Prasad D K, Rajan D, Rachmawati L, Rajabally E and Quek C. 2017. Video processing from electro-optical sensors for object detection and tracking in a maritime environment: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 18(8) : 1993-2016 [DOI: 10.1109/TITS.2016.2634580]
- Qiao D, Liu G, Li W, Lyu T and Zhang J. 2022. Automated full scene parsing for marine ASVs using monocular vision. *Journal of Intelligent & Robotic Systems*, 104(2) : #37 [DOI: 10.1007/s10846-021-01543-7]
- Qiao L, Wang Y and Zhang W. 2022. Path planning and obstacle avoidance of unmanned surface vehicles based on deep reinforcement learning. *IEEE Access*, 10: 12345-12356
- Radford A, Metz L and Chintala S. 2016. Unsupervised representation learning with deep convolutional generative adversarial networks.

- International Conference on Learning Representations (ICLR). [s.l.]: [s.n.]
- Rasouli A and Tsotsos J K. 2020. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. *IEEE Transactions on Intelligent Transportation Systems*, 21(3): 900-918
- Ren S, He K, Girshick R and Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1137-1149 [DOI: 10.1109/TPAMI.2016.2577031]
- Roberts G N and Sutton R. 2006. *Advances in Unmanned Marine Vehicles*. Stevenage, UK: IET [DOI: 10.1049/PBCE069E]
- Ruiz-Ponce P, Ortiz-Perez D, Garcia-Rodriguez J, and Kiefer B. 2023. POSEIDON: A data augmentation tool for small object detection datasets in maritime environments. *Sensors*, 23(7): #3691 [DOI: 10.3390/s23073691]
- Sazzadul Alam A K M, Hei X and Zhang Y P. 2025. In-progress: Enhancing traffic signal perception for connected and autonomous vehicles (CAVs) via multi-sensor fusion of camera, LiDAR, radar, and SPaT data//2025 IEEE Security and Privacy Workshops (SPW). San Francisco, USA: IEEE: 361-363 [DOI: 10.1109/spw67851.2025.00052]
- Schulman J, Wolski F, Dhariwal P, Radford A and Klimov O. 2017. Proximal policy optimization algorithms [EB/OL]. [2025-09-01]. <https://arxiv.org/pdf/1707.06347.pdf>
- Shao Z F, Wu W J, Wang Z Y, Du W and Li C Y. 2018. SeaShips: A large-scale precisely annotated dataset for ship detection. *IEEE Transactions on Multimedia*, 20(10): 2593-2604 [DOI: 10.1109/TMM.2018.2865686]
- Shao Z F, Wang Y, Wang J M, Deng L B, Huang X, Lu T, et al. 2025. GLSD: a global large-scale ship database with baseline evaluations. *Geo-Spatial Information Science*, 28(4): 1566-1580 [DOI: 10.1080/10095020.2024.2416896]
- Shah D, Osiński B, Ichter B and Levine S. 2023. LM-Nav: Robotic navigation with large pre-trained models of language, vision, and action//Proceedings of the 6th Conference on Robot Learning (CoRL). Auckland, New Zealand: PMLR: 492-504.
- Shen H, Hashimoto H, Mathews A, Tamba T A and Shimizu E. 2019. Automatic collision avoidance of multiple unmanned surface vehicles based on deep reinforcement learning. *Navigation*, 66(3): 261-278 [DOI: 10.1002/navi.303]
- Soloviev V, Farahnakian F, Zelioli L and Heikkonen J. 2020. Comparing CNN-based object detectors on two novel maritime datasets//2020 International Conference on System, Computation, Automation and Networking (ICSCAN). Pondicherry, India: IEEE: 1-6 [DOI: 10.1109/ICSCAN49426.2020.9262410]
- Sotelo-Torres F, Alvarez L V and Roberts R C. 2023. An unmanned surface vehicle (USV): development of an autonomous boat with a sensor integration system for bathymetric surveys. *Sensors*, 23(9): 4420 [DOI: 10.3390/s23094420]
- Steccanella L, Bloisi D D, Castellini A and Farinelli A. 2020. Waterline and obstacle detection in images from low-cost autonomous boats for environmental monitoring. *Robotics and Autonomous Systems*, 124: #103346 [DOI: 10.1016/j.robot.2019.103346]
- Su L, Chen Y S, Song H and Li W Y. 2023. A survey of maritime vision datasets. *Multimedia Tools and Applications*, 82(19): 28873-28893 [DOI: 10.1007/s11042-023-14756-9]
- Sunehag P, Lever G, Gruslys A, Czarnecki W M, Zambaldi V, Jaderberg M, et al. 2018. Value-decomposition networks for cooperative multi-agent learning//Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS). Stockholm, Sweden: IFAAMAS: 2085-2087
- Sun Y, Su L, Luo Y K, Meng H, Li W Y, Zhang Z, et al. 2022. Global Mask R-CNN for marine ship instance segmentation. *Neurocomputing*, 480: 257-270 [DOI: 10.1016/j.neucom.2022.01.017]
- Sun Z, Meng C, Huang T, Zhang Z and Chang S, et al. 2023. Marine ship instance segmentation by deep neural networks using a global and local attention mechanism. *PLOS ONE*, 18(2): #0279248 [DOI: 10.1371/journal.pone.0279248]
- Sünderhauf N, Brock O, Scheirer W, Hadsell R, Fox D, Leitner J, et al. 2018. The limits and potentials of deep learning for robotics. *The International Journal of Robotics Research*, 37(4/5): 405-420 [DOI: 10.1177/0278364918770733]
- Sutton R S and Barto A G. 2018. *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge, MA, USA: MIT Press
- Taipalmaa J, Passalis N, Zhang H, Gabbouj M and Raitoharju J. 2019. High-resolution water segmentation for autonomous unmanned surface vehicles: A novel dataset and evaluation//Proceedings of 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP). Pittsburgh, USA: IEEE: 1-6 [DOI: 10.1109/MLSP.2019.8918694]
- Toh Y J. 2017. *Development of a vision-based situational awareness capability for unmanned surface vessels*. Monterey, USA: Naval Postgraduate School
- Tu E M, Zhang G H, Rachmawati L, Rajabally E and Huang G B. 2018. Exploiting AIS data for intelligent maritime navigation: a comprehensive survey from data to methodology. *IEEE Transactions on Intelligent Transportation Systems*, 19(5): 1559-1582 [DOI: 10.1109/TITS.2017.2724551]
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*, #30
- Veličković P, Cucurull G, Casanova A, Romero A, Liò P and Bengio Y. 2018. Graph attention networks. *International Conference on Learning Representations (ICLR)*, [s.l.]: [s.n.]
- Wang C, Jiang Z, Li L and Liu Y. 2020. Ship detection for visual maritime surveillance from unmanned surface vehicles. *IEEE Transac-*

- tions on Geoscience and Remote Sensing, 58(10): 7311-7325
- Wang N, Wang Y, Feng Y and Wei Y. 2024a. MDD-ShipNet: Math-data integrated defogging for fog-occlusion ship detection. *IEEE Transactions on Intelligent Transportation Systems*, 25 (10) : 15040-15052 [DOI: 10.1109/TITS.2024.3391850]
- Wang S Y, Xie X L, Li M J, Wang M F, Yang J M, Li Z M, et al. 2024b. An adaptive multimodal fusion 3D object detection algorithm for unmanned systems in adverse weather. *Electronics*, 13(23) : #4706 [DOI: 10.3390/electronics13234706]
- Wang N, Wang Y Y, Wei Y, Han B and Feng Y. 2024c. Marine vessel detection dataset and benchmark for unmanned surface vehicles. *Applied Ocean Research*, 142: #103835 [DOI: 10.1016/j.apor.2023.103835]
- Wang W, Dai J, Chen Z, Huang Z, Li Z, Zhu X, et al. 2023a. Intern-Image: exploring large-scale vision foundation models with deformable convolutions//*Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, Canada: IEEE: 14408-14419 [DOI: 10.1109/CVPR52729.2023.01384]
- Wang Y, Li Y and He Y. 2023b. Dynamic graph neural networks for multi-agent cooperative control of unmanned surface vehicles. *Ocean Engineering*, 281: #114865
- Ward C M, Harguess J and Hilton C. 2018. Ship classification from overhead imagery using synthetic data and domain adaptation//*OCEANS 2018 MTS/IEEE Charleston*. Charleston, USA: IEEE: 1-5 [DOI: 10.1109/OCEANS.2018.8604662]
- Wen L Y, Du D W, Cai Z W, Lei Z, Chang M C, Qi H G, et al. 2020. UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking. *Computer Vision and Image Understanding*, 193: #102907 [DOI: 10.1016/j.cviu.2020.102907]
- Wibisono A, Piran M J, Song H K and Lee B M. 2023. A survey on unmanned underwater vehicles: challenges, enabling technologies, and future research directions. *Sensors*, 23 (17) : #7321 [DOI: 10.3390/s23177321]
- Williams G, Drews P, Goldfain B, Reh J M and Theodorou E A. 2018. Information-theoretic model predictive control: theory and applications to autonomous driving. *IEEE Transactions on Robotics*, 34(6): 1603-1622 [DOI: 10.1109/TRO.2018.2865891]
- Winstead P J. 2018. *Implementation of unmanned surface vehicles in the distributed maritime operations concept*. Newport, USA: Naval War College
- Wu H Z, Wang M, Wang J S and Wang G Q. 2024. Distributed information fusion based trajectory tracking for USV and UAV clusters via multi-agent deep learning approach. *Aerospace Systems*, 7 (2) : 193-207 [DOI: 10.1007/s42401-024-00275-4]
- Xie E, Wang W, Yu Z, Anandkumar A, Alvarez J M, and Luo P. 2021. SegFormer: simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*, 34: 12077-12090
- Xie W, Yang M and Gong J B. 2019. Underwater attack-defense confrontation system and its future development. *Strategic Study of CAE*, 21(6): 71-79 (谢伟, 杨萌, 龚俊斌. 2019. 水下攻防对抗体系及其未来发展. *中国工程科学*, 21(6): 71-79) [DOI: 10.15302/J-SSCAE-2019.06.014]
- Xue H, Chen X, Zhang R, Wu P, Li X, and Liu Y. 2021. Deep learning-based maritime environment segmentation for unmanned surface vehicles using superpixel algorithms. *Journal of Marine Science and Engineering*, 9(12): #1329 [DOI: 10.3390/jmse9121329]
- Yan C H, Yan S F, Yao T Y, Yu Y, Pan G, Liu L, et al. 2024. A lightweight network based on multi-scale asymmetric convolutional neural networks with attention mechanism for ship-radiated noise classification. *Journal of Marine Science and Engineering*, 12(1): #130 [DOI: 10.3390/jmse12010130]
- Yan R J, Pang S, Sun H B and Pang Y J. 2010. Development and missions of unmanned surface vehicle. *Journal of Marine Science and Application*, 9(4): 451-457 [DOI: 10.1007/s11804-010-1033-2]
- Yang H, Zhao X and Zhang Y. 2022. Lightweight graph attention networks for multi-USV cooperative navigation in resource-constrained environments. *IEEE Transactions on Intelligent Vehicles*, 7 (4) : 952-964
- Yang Q S, Yin Y, Chen S and Liu Y. 2021. Autonomous exploration and navigation of mine countermeasures USV in complex unknown environment//*Proceedings of the 2021 33rd Chinese Control and Decision Conference (CCDC)*. Kunming, China: IEEE: 4373-4377 [DOI: 10.1109/CCDC52312.2021.9602457]
- Yang X, Sun H, Fu K, Yang J R, Sun X, Yan M L, et al. 2018. Automatic ship detection in remote sensing images from Google earth of complex scenes based on multiscale rotation dense feature pyramid networks. *Remote Sensing*, 10 (1) : #132 [DOI: 10.3390/rs10010132]
- Yang X, She H, Lou M, Ye H, Guan J and Li J. 2024. A joint ship detection and waterway segmentation method for environment-aware of USVs in canal waterways. *IEEE Transactions on Automation Science and Engineering*, 22: 2132-2144 [DOI: 10.1109/TASE.2024.3375300]
- Yao S L, Guan R W, Huang X Y, Li Z X, Sha X Y, Yue Y, et al. 2024. Radar-camera fusion for object detection and semantic segmentation in autonomous driving: a comprehensive review. *IEEE Transactions on Intelligent Vehicles*, 9(1): 2094-2128 [DOI: 10.1109/TIV.2023.3307157]
- Yao S L, Guan R W, Ni Y, Xu S, Yue Y, Zhu X H, et al. 2025. USVTrack: USV-based 4D radar-camera tracking dataset for autonomous driving in inland waterways//*IEEE/RSJ International Conference on Intelligent Robots and Systems*. Hangzhou, China: IEEE: 19016-19023
- Yoerger D R, Curran M, Fujii J, German C R, Gomez-Ibanez D, Govindarajan A F, et al. 2018. Mesobot: an autonomous underwater vehicle for tracking and sampling midwater targets//*Proceedings of*

- the 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV). Porto, Portugal: IEEE: 1-7 [DOI: 10.1109/AUV.2018.8729822]
- Zheng C, Mao Z and Zhao X. 2024. Priority experience replay-based multi-agent reinforcement learning for unmanned surface vehicle cooperative target hunting. *Ocean Engineering*, 295: #116854 [DOI: 10.1016/j.oceaneng.2024.116854]
- Zhang C, Zhang D, Zhang M Y and Mao W G. 2019. Data-driven ship energy efficiency analysis and optimization model for route planning in ice-covered Arctic waters. *Ocean Engineering*, 186: #106071 [DOI: 10.1016/j.oceaneng.2019.05.053]
- Zhang D Y, Chu X M, Liu C G, He Z B, Zhang P L and Wu W X. 2024. A review on motion prediction for intelligent ship navigation. *Journal of Marine Science and Engineering*, 12(1): #107 [DOI: 10.3390/jmse12010107]
- Zhang L, Anwar A, Steckel J and Mercelis S. 2026. Scene-adaptive semantic segmentation guided by multi-level boundary-semantic-reinforcement for unmanned surface vessels. *Measurement*, 257: #118655 [DOI: 10.1016/j.measurement.2025.118655]
- Zhang L, Du X, Zhang R and Zhang J. 2023. A lightweight detection algorithm for unmanned surface vehicles based on multi-scale feature fusion. *Journal of Marine Science and Engineering*, 11(7): #1392 [DOI: 10.3390/jmse11071392]
- Zhang M M, Choi J, Daniilidis K, Wolf M T and Kanan C. 2015. VAIS: A dataset for recognizing maritime imagery in the visible and infrared spectrums//2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Boston, USA: IEEE: 10-16 [DOI: 10.1109/CVPRW.2015.7301291]
- Zhang T, Zheng X Q and Liu M X. 2021a. Multiscale attention-based LSTM for ship motion prediction. *Ocean Engineering*, 230: #109066 [DOI: 10.1016/j.oceaneng.2021.109066]
- Zhang X R, Wang G C, Zhu P, Zhang T Y, Li C and Jiao L C. 2021b. GRS-Det: an anchor-free rotation ship detector based on Gaussian-mask in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4): 3518-3531 [DOI: 10.1109/TGRS.2020.3018106]
- Zhang Y, Li Y, and He Y. 2021c. Distributed consensus of USVs under heterogeneous UAV-USV multi-agent systems cooperative control scheme. *Journal of Marine Science and Engineering*, 9(11): #1314 [DOI: 10.3390/jmse9111314]
- Zheng Y and Zhang S. 2020. Mcships: A large-scale ship dataset for detection and fine-grained categorization in the wild//Proceedings of 2020 IEEE International Conference on Multimedia and Expo (ICME). London, United Kingdom: IEEE: 1-6 [DOI: 10.1109/ICME46284.2020.9102907]
- Zhao L M and Roh M I. 2019. COLREGs-compliant multiship collision avoidance based on deep reinforcement learning. *Ocean Engineering*, 191: #106436 [DOI: 10.1016/j.oceaneng.2019.106436]
- Zhao X, Wang L M, Zhang Y F, Han X M, Deveci M and Parmar M. 2024. A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 57(4): #99 [DOI: 10.1007/s10462-024-10721-6]
- Zhao Y, Zhao L J, Xiong B L and Kuang G Y. 2020. Attention receptive pyramid network for ship detection in SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 2738-2756 [DOI: 10.1109/JSTARS.2020.2997081]
- Zheng H R, Negenborn R R and Lodewijks G. 2013. Survey of approaches for improving the intelligence of marine surface vehicles//16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013). The Hague, Netherlands: IEEE: 1217-1223 [DOI: 10.1109/ITSC.2013.6728398]
- Zheng Z and Sun L. 2017. Path following control for a marine surface vessel with asymmetric saturation actuators. *Robotics and Autonomous Systems*, 97: 83-91 [DOI: 10.1016/j.robot.2017.08.005]
- Zheng L, Shen L Y, Tian L, Wang S J, Wang J D and Tian Q. 2015. Scalable person re-identification: a benchmark//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE: 1116-1124 [DOI: 10.1109/ICCV.2015.133]
- Zhong H Z, Wang H, Wu Z R, Zhang C, Zheng Y W and Tang T. 2021. A survey of LiDAR and camera fusion enhancement. *Procedia Computer Science*, 183: 579-588 [DOI: 10.1016/j.procs.2021.02.100]
- Zhou C H, Gu S D, Wen Y Q, Du Z, Xiao C S, Huang L, et al. 2020. The review unmanned surface vehicle path planning: based on multi-modality constraint. *Ocean Engineering*, 200: #107043 [DOI: 10.1016/j.oceaneng.2020.107043]
- Zhou R, Gao Y, Wu P, Zhao X, Dou W, Sun C, Zhong Y and Wang Y. 2022. Collision-free waterway segmentation for inland unmanned surface vehicles. *IEEE Transactions on Instrumentation and Measurement*, 71: 1-16 [DOI: 10.1109/TIM.2022.3165803]
- Zhou Z, Li M, Yu H, Sun Z and Qu Y. 2021. An image-based benchmark dataset and a novel object detector for water surface object detection. *Frontiers in Neurorobotics*, 15: #723336 [DOI: 10.3389/fnbot.2021.723336]
- Zolich A, Palma D, Kansanen K, Fjortoft K, Sousa J, Johansson K H, et al. 2019. Survey on communication and networks for autonomous marine systems. *Journal of Intelligent and Robotic Systems*, 95(3/4): 789-813 [DOI: 10.1007/s10846-018-0833-5]
- Zuo Q, Li C Z H, Fan Y L, Kang W S, Yang T, Feng N N, et al. 2025. Advanced multi-modal sensor fusion architectures for robust autonomous driving systems//2025 IEEE 5th International Conference on Electronic Technology, Communication and Information (ICETCI). Changchun, China: IEEE: 1556-1560 [DOI: 10.1109/icetci64844.2025.11084124]
- Žust L and Kristan M. 2022. Temporal context for robust maritime obstacle detection//2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Kyoto, Japan: IEEE: 6340-6346 [DOI: 10.1109/IROS47612.2022.9982043]

Žust L, Perš J and Kristan M. 2023. LaRS: a diverse panoptic maritime obstacle detection dataset and benchmark//Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE: 20304-20314 [DOI: 10.1109/ICCV51070.2023.01857]

### 作者简介

王曰英,男,研究员,主要研究方向为无人艇/水下机器人智能控制理论与技术和海洋智能无人系统。

E-mail:yueyingwang@shu.edu.cn

吴浩,男,硕士研究生,主要研究方向为图像处理。

E-mail:2544557363@shu.edu.cn

庆雨豪,男,博士研究生,主要研究方向为图像增强、目标检测、语义分割和多目标跟踪。E-mail:yuhao\_fly@163.com

张卫东,男,副教授,主要研究方向为深度学习、底层视觉感知和遥感影像解译。E-mail:zwd\_wd@163.com

沈礼权,男,研究员,主要研究方向为图像处理。

E-mail:jsslq@shu.edu.cn

徐昕,男,教授,主要研究方向为智能控制、强化学习、机器人以及自主车辆。E-mail:xinxu@nudt.edu.cn