technologies

MDPI

Review

AUV Intelligent Decision-Making System Empowered by Deep Learning: Evolution, Challenges and Future Prospects

Qiulin Ding [1,2], Lugang Ye [1], Hao Chen [1,2], Hongyuan Liu [1,2], Aoming Liang [3] and Weicheng Cui [1,2,*]

[1] Department of Electronic and Information Engineering, School of Engineering, Westlake University, Hangzhou 310030, China; dingqiulin@westlake.edu.cn (Q.D.); chenhao@westlake.edu.cn (H.C.); liuhongyuan@westlake.edu.cn (H.L.)
[2] Zhejiang Engineering Research Center of Micro/Nano-Photonic/Electronic System Integration, Hangzhou 310030, China
[3] Zhejiang University-Westlake University Joint Training, Zhejiang University, Hangzhou 310024, China; liangaoming@westlake.edu.cn
* Correspondence: cuiweicheng@westlake.edu.cn

Abstract

The intelligent decision-making systems of Autonomous Underwater Vehicles (AUVs) are undergoing a significant transformation, shifting from traditional control theories to data-driven paradigms. Deep learning (DL) serves as the primary driving force behind this evolution; however, its application in complex and unstructured underwater environments continues to present unique challenges. To systematically analyze the development, current obstacles, and future directions of DL-enhanced AUV decision-making systems, this paper proposes an innovative 'four-module' decomposition framework consisting of information processing, understanding, judgment, and output. This framework enables a structured review of the progression of DL technologies across each stage of the AUV decision-making information flow. To further bridge the gap between theoretical advancements and practical implementation, we introduce a task complexity–environment uncertainty four-quadrant analytical matrix, offering strategic guidance for selecting appropriate DL architectures across diverse operational scenarios. Additionally, this work identifies key challenges in the field as well as anticipates future developments to solve these challenges. This paper aims to provide researchers and engineers with a comprehensive and strategic overview to support the design and optimization of next-generation AUV decision-making architectures.

Keywords: autonomous underwater vehicle; deep learning; intelligent decision-making system; intelligent control system; deep reinforcement learning

check for updates

Academic Editor: Salvatore Brischetto

Received: 10 November 2025
Revised: 9 December 2025
Accepted: 10 December 2025
Published: 12 December 2025

Citation: Ding, Q.; Ye, L.; Chen, H.; Liu, H.; Liang, A.; Cui, W. AUV Intelligent Decision-Making System Empowered by Deep Learning: Evolution, Challenges and Future Prospects. *Technologies* **2025**, *13*, 586. https://doi.org/10.3390/technologies13120586

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Research Background

The vast ocean is rich in resources, including oil and gas, minerals, biological and renewable energy, and the development and utilization of these resources are of great significance to the sustainable development of human society. The core tools for carrying out underwater exploration tasks are underwater robots, whose development level directly decides the depth, width and efficiency of underwater exploration. Underwater robots are mainly divided into Remotely Operated Vehicles (ROVs) and Autonomous Underwater Vehicles (AUVs) based on whether they are controlled in real-time or not [1]. ROVs rely on cables and real-time control by operators. Although they have a certain degree of intuitiveness in operation, their range of motion is limited due to the existence of this cable.

In contrast, AUVs have a high degree of autonomy and can independently complete a series of complex tasks, such as navigation, obstacle avoidance, reconnaissance and sampling, based on the information sensed by the sensors they carry and the preset mission objectives, greatly expanding the depth, breadth and duration of underwater operations [2].

1.2. The Development History of AUV Intelligent Decision-Making Systems

Figure 1 shows the gradual yet impactful integration of DL into underwater robotics, typically requiring 2–7 years for technologies to transition from foundational computer science breakthroughs to specialized marine applications.

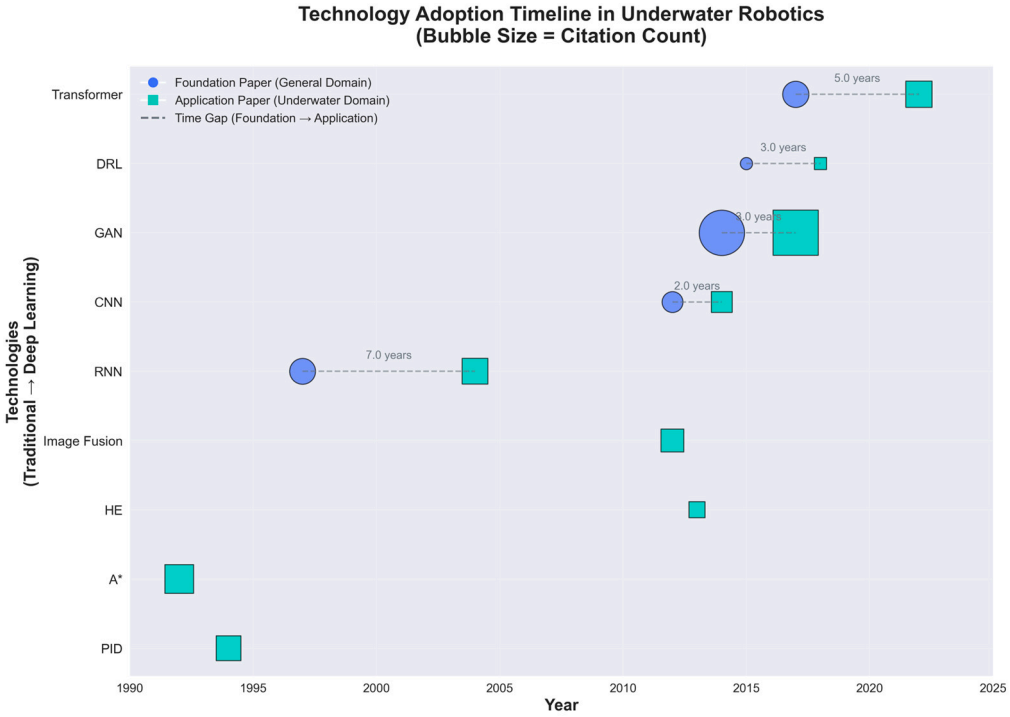

Figure 1. Technology adoption timeline in underwater robotics.

To enable AUVs to assume tasks currently dominated by ROVs, the pivotal challenge lies in advancing control system architectures to increase vehicle intelligence [3]. The underwater environment has many unique physical characteristics that pose a serious challenge to traditional control methods based on precise physical models and explicit rules [4,5]. This makes it difficult for vision-dependent underwater robots to obtain clear images and accurate visual information underwater, and they cannot use vision for precise target recognition and navigation as they do on land.

In recent years, artificial intelligence technologies represented by DL have revolutionized the decision-making systems of underwater robots with their outstanding ability to handle high-dimensional, nonlinear data (data with complex relationships not fitting a simple linear model), and have become the key technology to break through the constraints of the underwater environment and achieve the true intelligence of AUVs [6], which means that AUVs can truly possess understanding and intentionality. DL algorithms can automatically extract effective features from complex sensor data, avoiding the cumbersome manual feature extraction process in traditional methods and greatly improving the efficiency and accuracy of data processing.

Deep-learning technologies, such as CNNs, effectively overcome underwater environmental constraints by enhancing visual image quality through de-scattering and denois-

ing [7], while simultaneously improving the resolution and target recognition accuracy of sonar data.

The application of DL technology has facilitated the leap of underwater robot decision-making systems from static, model-based control logic to dynamic, data-driven autonomous learning paradigms, greatly enhancing the intelligence level of underwater robots. Through DL algorithms, underwater robots can autonomously learn and adjust behavioral strategies based on real-time perceived environmental information, achieving more flexible and efficient task execution [8].

DL is being widely applied in various modules of underwater robot decision-making systems [9]. In order to analyze this trend, this paper explicitly restricts its scope to DL algorithm in AUVs. Specifically, we focus on the evolution of intelligent decision-making systems for AUVs brought by DL development, analyzing how DL technologies are replacing or augmenting traditional control theories to explore the boundary of underwater application.

1.3. Research Gaps and Motivation

Although some progress has been made in the application of DL to decision-making systems for underwater robots, there are still some shortcomings in the current research. In terms of comprehensive analysis, the current review articles of the same type [10–14] lack a comprehensive and integrated analysis of the AUV control system, either not focusing on the particular robot form of AUV [12], or only introducing a single field separately, lacking an in-depth analysis of the AUV decision-making system as a whole [10,11]. It is difficult to meet the needs of scholars who need to build and improve AUV control systems.

On the other hand, as shown in Figure 1, the development of DL has had a profound impact on intelligent decision-making systems for underwater robots. This makes the review of control systems based on traditional model analysis no longer suitable for the current stage of development. Due to the particularity of underwater robots, their DL applications have always lagged behind those of onshore quadruped/biped robots, and their intelligent decision-making system architectures are often limited to improvements and fine-tuning of onshore solutions [14]. This pattern of development is not adapted to the special dynamics and decision-making environment of underwater robots.

As larger onshore models [15] gradually become less suitable for the computing power conditions underwater, where it is difficult to deploy large models in real-time through communication, it is foreseeable that in the future, there will be efficient architectures and standalone systems focused on underwater autonomous control systems. There is a lack of articles in the academic community that analyze and predict the nascent stage of developing autonomous control systems for underwater robots.

In terms of module relationship sorting, with the development of the multi-module consolidation wave led by end-to-end large models, how to sort out the relationships of different modules according to application requirements and thereby achieve better configuration and integration of the perception end and control end of the autonomous system has become an urgent problem to be solved. However, there is currently a lack of comprehensive analysis and sorting of different modules and their complex relationships within the profession, leaving researchers lacking effective references when designing and optimizing AUV decision-making systems.

In the study of DL architectures adapted to underwater environments, simply transplanting the DL framework of ground robots is not sufficient to solve the problem of autonomous decision-making of AUVs, and there are relatively few studies of DL architectures tailored to the characteristics of underwater environments at present. There is a

particular lack of articles focusing on the impact of DL on autonomous decision-making systems (ADSs).

1.4. The Purpose of This Article

This paper aims to address the deficiencies in the current research by systematically sorting out the role and impact of DL in the AUV intelligent decision-making system through a novel "four-module" perspective. This work fills a significant gap in the literature by providing the first systematic, multi-dimensional, and end-to-end analysis of DL within the AUV control domain. By leveraging this framework alongside a scientifically classified "four-scenario" analysis matrix, we deeply explore the specific application and technology selection of DL across each link of the decision-making chain. Our objective is to provide a comprehensive and forward-looking perspective that helps researchers at the intersection of underwater robotics and artificial intelligence to profoundly understand the intrinsic relationship between DL algorithms and target applications. By clarifying development priorities based on practical engineering needs, this paper strives to offer a solid theoretical foundation and technical reference for the design and implementation of the next generation of fully autonomous AUV decision-making systems.

1.5. Paper Structure Arrangement

This paper is divided into five sections, each of which closely revolves around the theme of DL empowering decision-making systems for underwater robots. The first section is the introduction, which elaborates on the research background, significance, existing problems and the purpose of this study. Section 2 defines the proposed 'four-module' analysis framework, detailing the division of labor across the system's information flow to lay a theoretical foundation for the subsequent exploration of DL applications. Section 3, as an overview of core technologies, reviews the technological evolution and role of DL in the four modules one-by-one, analyzes the integration trends of different modules, and elaborates in detail how DL technology plays a key role in each module and the integration development trends among different modules.

Section 4 will divide the application environment into four typical scenarios based on task complexity and environmental uncertainty, and analyze one-by-one the challenges each scenario poses to the intelligent decision-making system and the contribution of DL algorithms in addressing these challenges. Through the analysis of actual scenarios, the advantages and adaptability of DL in different application scenarios will be further revealed. Section 5 summarizes the core challenges and cutting-edge trends in the current field and looks forward to future developments, comprehensively sorting out the problems and future development directions in the field of DL-enabled decision-making systems for underwater robots at present.

2. Definition and Module Division of Intelligen Decision-Mtaking Systems

2.1. Intelligent Decision Systems: Definitions, Paradigms, and Autonomous Cores

The Intelligent Decision System is the core of achieving a high degree of autonomy in AUVs and is recognized as the brain of AUVs. The essence lies in the system's ability to integrate real-time perceptual data to autonomously analyze environmental conditions, assess mission risks, and select and generate optimal action strategies in highly complex, dynamic, and information-constrained underwater environments, in order to efficiently and robustly complete scheduled or emergent tasks.

In general, traditional AUV decision-making relies on rule-driven paradigms, expert systems, and classical control theory [16]. However, traditional control strategies often

prove inadequate in meeting stringent performance specifications [17]. The introduction of DL enables AUVs to establish complex perception–decision mappings [18] directly from high-dimensional sensor data, overcoming the reliance on precise physical models required by traditional approaches.

2.2. The Four-Module Deconstruction and DL Function of the Intelligent Decision-Making System

To precisely sort out the impact of DL technology on the autonomous decision-making process of AUVs, as Figure 2 shows, this paper uses the logic of information flow to innovatively deconstruct the intelligent decision-making system into four core modules: the information processing module, the information understanding module, the information judgment module, and the output module. This division aims to establish a clear analytical framework that enables us to closely track the specific functions and contributions of DL technology in each link of the information transformation chain.

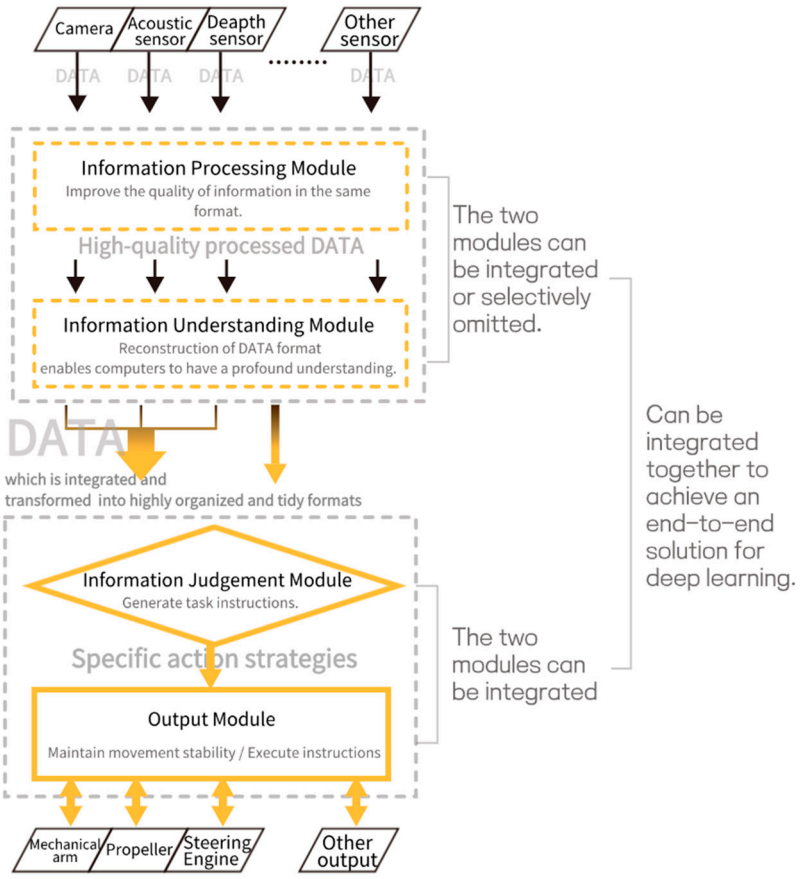

Figure 2. Information flow logic.

2.2.1. Definition of Information Processing Module

This module is the starting point of the information flow, and its core function is to receive the raw data flow from various sensors of the AUV, and use advanced algorithms to preprocess and optimize it to purify and enhance the information. For example, underwater image processing, sonar signal denoising, etc., fall within the scope of this module. It optimizes the information through technical means. This is different from the information understanding module, which improves the quality of information in the same format and essentially purifies and optimizes the same kind of information.

2.2.2. The Information Understanding Module

Building on the high-quality data provided by the information processing module, the information understanding module is responsible for converting low-level sensor data into high-level environmental cognition that computers can understand, such as semantic segmentation, SLAM mapping, etc. Its function lies in the understanding and transformation of information, that is, using DL to transform and understand processed data into high-level semantic information, to extract a small amount of easily understandable single-dimensional simple information from the complex multi-dimensional information that is difficult for the judgment module to understand. It is essentially a reconfiguration of the information format.

2.2.3. Definition of Information Judgment Module

The information judgment module is the core hub of the entire Intrusion Detection System (IDS), responsible for translating the environmental cognition and task objectives provided by the information understanding module into specific action strategies. This module is an embodiment of AUV autonomy and must have the ability to generate optimal decisions in complex and uncertain environments.

In order to better cover intelligent decision-making systems, the information judgment module in this paper is defined as generating simple tasks from information without being responsible for directly converting the tasks into motor pulse signals or rotation angles. After the task instructions are produced by the information judgment module, the output module is responsible for the specific implementation of the lower computer.

2.2.4. Definition of Output Module

As mentioned above, the output module is equivalent to the reflex center of the AUV, mainly handling real-time micro-decisions at the tactical, kinematic or dynamic levels. Its core function is to maintain stability while ensuring that instructions from the advanced planning layer can be safely, smoothly and efficiently converted into execution actions. DL is mainly used at this layer to model the real-world for predictive adjustment of actions, such as using DL for efficient obstacle-avoidance strategies and adaptive motion control to make up for the shortcomings of traditional geometric programming algorithms in dynamic environments.

2.2.5. Module Splitting and Mix

In many articles, the output module and the judgment module are mixed and defined as decision modules, and the concepts of monolithic models and split models are generated based on the degree of fusion [19]. In decision systems, whether to integrate the output module and the information judgment module together is a very different technical approach. This article will also delve deeper into this issue later on.

Many review articles thus analyze the output module and the information judgment module as a whole. This does not provide a comprehensive review of the components of an intelligent decision-making system, because there are many architectural designs that handle the design of two separate modules in a specialized hierarchical architectural design model. For example, Article [20] has a clear hierarchical structure—the AI agent is responsible for cognition and planning, and the underlying controller is responsible for execution. Combining the two modules for analysis can lead to ambiguous semantics and an inability to clearly define trends, and related scholars tend to overlook the architecture scheme of separating control when designing the architecture. So, although there are some overlaps in the examples, this article specifically summarizes the two separately.

2.3. The Flexibility of Deconstructing Intelligent Decision-Making Systems

Although this paper presents a four-module analytical framework, it must be made clear that the framework aims to provide a theoretical path for information transformation, rather than strictly limiting the engineering implementation of all intelligent decision-making systems. In fact, it is the incorporation of DL technology that has greatly facilitated the integration of different modules. Because DL has a natural end-to-end tendency, many cutting-edge intelligent decision-making systems, especially those that pursue end-to-end learning, will naturally exhibit selective omitting or deep integration of modules. For example, due to the limited computing power of AUV hardware or the pursuit of real-time decision-making, there are often systems that skip the information processing stage and directly feed the raw sensor data without denoising or enhancement into the subsequent information understanding or information judgment module. The advantage of this approach is that it avoids the delay caused by preprocessing, but at the cost of significantly increasing the robustness and generalization difficulty of subsequent modules in dealing with high-noise, low-signal-to-noise ratio data.

More aggressive fusion trends are seen in single-layer decision models represented by Vision–Language–Action (VLA) large models. These models are designed to establish a direct mapping from complex sensor inputs to action outputs [21]. In this architecture, the information processing module and the information understanding module are often embedded or implicitly integrated into a single, giant feature extraction layer of the model, thereby eliminating explicit module boundaries at the system architecture level. For example, Transformer-based VLA [22] models can directly receive the raw image and output path points or control thrust, bypassing traditional semantic middleware. While this design retains the maximum richness of input information, it also significantly increases the learning difficulty and the amount of data required for the information judgment module.

Although fusion schemes seem to be the current trend, this paper still analyzes them separately according to the four-module theory. End-to-end solutions have not yet become a consensus in the industry due to their shortcomings [23]; at the same time, there are still many improvements in DL solutions for specific modules being proposed. Separate research will make the technical route clearer and also provide more professional insights and summaries for readers who need to design AUV intelligent decision-making systems. This paper summarizes the greatest innovation points of the multi-module deeply integrated technical solutions in the academic field, and then references and analyzes them in a targeted manner based on the modules where the greatest innovation points are located.

It is also worth noting that, regardless of how the system structure is simplified or integrated, an intelligent decision-making system must contain two core and indispensable modules: the information judgment module and the output module. Without the information judgment module, the system will lose its ability to make autonomous decisions. Without an output module, decision-making strategies cannot be translated into practical actions, and ADS loses its engineering value and task-driven significance. Of course, there are many cases of integration between the output module and the information judgment module, but its presence in the overall intelligent decision-making system makes it impossible for either to be omitted.

2.4. Module Collaboration and System Integration

The four modules of an intelligent decision-making system do not work in isolation but form a complete cognitive decision-making loop through close collaboration. The information flows through the modules in an orderly manner, forming a complete chain from perception to action.

In a typical workflow, the information processing module cleans and enhances the raw sensor data to provide high-quality input for the understanding module. The understanding module semantically parses the processed data to generate a highly clean environment description. The judgment module formulates decision-making strategies based on the environment description and task requirements. The output module eventually converts the strategy into control instructions to drive the actuator to perform the corresponding actions. The execution results of each module form a closed loop through sensor feedback, enabling continuous optimization of the system.

Module interface design and data specification affect system performance. For example, the information understanding module needs to provide an environment description with an appropriate level of abstraction for the information judgment module. Excessive detail increases the computational burden, while excessive simplicity results in the loss of decision-making information. To address this, Li et al. [24] explored the temporal semantic communication paradigm. They integrated the ISC3 (Integrated Sensing, Computing, Communication, and Control) architecture, using SFE (Semantic Feature Extractor) at the transmitting end to identify the temporal series correlation of control information to adjust the information update strategy, and using SFR (Semantic Feature Reconstructor) at the receiving end to predict and reconstruct untransmitted control information, ensuring control accuracy while reducing communication overhead. It provides a new approach to solving the communication efficiency problem of real-time control systems.

The performance of intelligent decision-making systems is ultimately reflected in their ability to complete complex tasks. A well-designed system should be able to achieve optimal task performance under resource constraints, which requires collaborative optimization of modules and a well-designed overall architecture.

3. Modules for Autonomous Decision-Making Empowered by Deep Learning

3.1. Information Processing Module

3.1.1. The Evolution of Information Processing Module

Optical and acoustic information, which are the most commonly used, are often distorted underwater. As a result, the information processing module of the AUV focuses more on these two types of information restoration. As the perceptual front end of the decision-making system, the evolution of the information processing module clearly reflects the shift from relying on prior physical assumptions to embracing data-driven learning.

Before the advent of DL, the field was dominated by schemes based on physical models and traditional filtering. For example, the recovery of underwater visual information has long relied on the reverse solution of the physical process of light propagation in water, such as the classic Jaffe–McGlamery [25] model. These methods attempt to restore images by modeling effects such as forward scattering, backscattering, and absorption, but their performance is highly dependent on precise estimates of the optical parameters of the water body, which leads to the limited generalization ability of the models [26]. Parallel to this are signal-processing-based enhancement techniques such as contrast stretching [27], Retinex theory [28], and Wiener filtering for acoustic signals [29]. These methods have low computational overhead and do not rely on complex physical models, but by nature, they adjust based on the statistical characteristics of the image, often amplifying noise or losing details, and have difficulty coping with non-uniform complex degradation.

With the increase in computing power and the emergence of large datasets, DL-based processing solutions have become mainstream [30]. CNNs, with their powerful feature-extraction capabilities, were first used to construct mappings from degraded images to clear ones. Wang et al. [31] first introduced CNN into the field of underwater image

enhancement and proposed UIE-Net, the core of which is the co-optimization of color correction and defogging through dual-task joint training. The network consists of a shared feature extraction layer, S-Net, and two branches. The CC-Net outputs a three-channel attenuation coefficient to correct color distortion, and the defogging network HR-Net outputs a single-channel transmission image to enhance contrast. At the same time, pixel-scrambling strategies are used to suppress local texture interference. They randomly shuffle the pixels of the image blocks to improve convergence speed and accuracy. However, there is a problem of color oversaturation in some high-frequency regions, and the reliance on image block overlap tests results in high computational overhead.

Subsequently, the introduction of Generative Adversarial Network (GAN) improved the efficiency of information processing. WaterGAN, as introduced by Li et al. [32], generates networks through unsupervised adversarial training to generate paired training data, alleviating the scarcity problem of underwater datasets. Since then, applications of unpaired image models such as CycleGAN [33] have enabled model training without the need for strictly corresponding pairs of clear-degenerate image pairs, lowering the threshold for data acquisition. To address the problem that traditional DL methods require a large amount of paired data for underwater image enhancement, UW-CycleGAN [34] introduced the CycleGAN framework to solve this challenge. It enables the model to perform image-to-image transformation learning between unpaired degradation and clear image sets by training two generators and two discriminators, using cyclic consistency loss to ensure reversibility of image transformation and content retention, achieving high-quality underwater image enhancement without explicitly pairing data and expanding the application boundaries of DL in underwater image processing.

The same technical approach has also been successfully applied to the problem of sonar image processing. Du et al. [35] showed that even a four-layer CNN network outperforms traditional techniques in sonar information augmentation. Wang et al. [36] delved into and compared the application effects of multiple DL denoising algorithms designed for optical images on underwater sonar images. Notably, they treated images processed by different denoising algorithms as multi-frame data of the same scene and fused them using multi-frame denoising techniques, achieving good results.

While DL has achieved impressive results in denoising and information augmentation, there are also different findings. For example, Huang et al. [37] clearly pointed out the limitations of DL-based denoising algorithms in AUV applications when exploring speckle noise denoising in underwater sonar images. They noted that although DL methods typically outperform traditional methods in terms of denoising performance, their huge computational load, long training time, and strict requirements for large-scale raw image datasets make it difficult to deploy efficiently under the limited computing resources and storage space of AUV. This is indeed a huge constraint on the application of DL in current modules.

3.1.2. Applications of DL in Information Processing

Comparing DL-based schemes with traditional ones, the differences are reflected in the fundamental differences between data-driven and model-driven schemes [38]. The traditional approach attempts to solve a well-defined inverse problem through precise mathematical modeling and the inversion of physical laws. The advantage of this approach lies in its solid theoretical foundation and strong interpretability, but its effectiveness is greatly compromised when there is a huge gap between the model assumptions and the complex reality. DL solutions bypass the complex intermediate modeling process and view information processing as a high-dimensional pattern recognition and mapping problem. By learning from massive amounts of data, it summarizes the empirical rules of how to

recover from degraded patterns to clear patterns. This data-driven paradigm gives the system unprecedented generalization ability and robustness, enabling it to adaptively handle previously unseen, complex degradation scenarios [35].

Of course, the acquisition of this capability comes at a cost; the large amount of computation required by DL schemes does not meet the inherent requirements of AUV information processing models for speed and low consumption, and its black box nature makes the decision-making process lack transparency, a problem that current research is striving to solve.

3.1.3. Development Summary and Future Projections of Information Processing Module

The information understanding module aims to maximize signal fidelity and content in uncertain underwater environments.

Looking ahead, this module will follow three trends—multimodal generative sensing: generative AI fuses multi-source data (vision/acoustics) to restore signals and infer missing information [39,40]; integrating physical knowledge and data-driven methods: PINN embeds propagation models to avoid nonsensical results [41]; self-supervised/weakly supervised learning for scarce underwater labeled data.

3.2. Information Understanding Module

3.2.1. The Evolution of Information Understanding Module

As shown in Table 1, the technical evolution of the information understanding module has shifted from manual feature engineering to automatic feature learning. Original technical solutions relied heavily on human experts' insights, using descriptors like SIFT and HOG [42] combined with classifiers like SVM. While effective in structured scenes, these methods struggle to capture abstract semantic concepts in complex underwater environments.

Table 1. The relevant technologies of the information processing module Section 1.

Traditional/DL	Name	Technology	Summary of the Technical Route
Solution based on physical models and traditional filtering	Jaffe–McGlamery Model [31]	Physics-based model for underwater vision recovery	Recovers images by inversely solving the physical process of light propagation in water, modeling forward scattering, backscattering, and absorption effects.
	Contrast Stretching [33]	Signal processing enhancement technique	Enhances image contrast based on statistical properties, with low computational cost and no reliance on complex physical models.
	Retinex Theory [34]	Signal processing enhancement technique	Enhances image quality based on statistical properties, with low computational cost and no reliance on complex physical models.
	Wiener Filtering [35]	Signal processing enhancement technique (for acoustic signals)	Denoises acoustic signals based on statistical properties, with low computational cost and no reliance on complex physical models.

Table 1. *Cont.*

Traditional/DL	Name	Technology	Summary of the Technical Route
Deep-learning solution	UIE-Net [37]	CNN for underwater image enhancement	Achieves synergistic optimization of color correction and defogging through dual-task joint training, marking an early application of CNNs in this field.
	WaterGAN [38]	GAN for underwater image processing	Generates paired training data through unsupervised adversarial training, effectively addressing the scarcity of underwater datasets.
	CycleGAN [39]	Non-paired image model	Lowers data acquisition barriers by enabling model training without strictly corresponding paired clear-degraded image sets.
	UW-CycleGAN [40]	CycleGAN framework for underwater image enhancement	Enables high-quality underwater image enhancement without explicit paired data by learning image-to-image translation between unpaired degraded and clear image sets, using cyclic consistency loss.
	Multi-frame denoising technique with OPD [42]	Multi-frame denoising for underwater sonar images	Fuses multi-frame data, treated as images processed by different denoising algorithms, to achieve better denoising results in underwater sonar imaging.
	PINN (Physics-Informed Neural Networks) [43]	Neural network with embedded physical models	Embeds physical models (e.g., optical or acoustic propagation) as inductive biases into neural network structures or loss functions, enhancing model stability and interpretability by preventing physically unrealistic outcomes.

DL, especially deep convolutional neural networks, can automatically learn feature representations from the original pixels [44].

In underwater classification tasks, models focus on identifying specific species or targets. For example, MLR-VGGNet [45], based on the CNN architecture, and improved methods, based on mResNet [46], both achieved superior classification accuracy of over 96% on the Fish4Knowledge dataset; DAMNet [47] and MCANet [43] used advanced attention mechanisms to handle complex biological image classification and achieved good results.

In the field of underwater object detection and segmentation, detection and segmentation frameworks in various DL domains are widely applied, ranging from two-stage methods such as Faster R-CNN to single-stage methods such as YOLO [48]. It is notable that the YOLO algorithm has taken the mainstream position in underwater object monitoring due to its simplicity, openness and ease of deployment. A wealth of improved algorithms based on YOLO have emerged in the field [49–51]. The industry has developed more sophisticated architectures for small targets such as underwater garbage, such as FocusDet [52], which focuses on small-object monitoring, and MLDet [53], which focuses on underwater garbage monitoring.

In terms of semantic segmentation, models have evolved from classic fully convolutional networks, FCN, [54] to more complex Encoder–Decoder architectures such as MTHI-Net [55] and BCMNet [56]. A notable frontier trend is the use of the underlying models that many scholars are using to solve the problem of underwater fine segmenta-

tion, that is, to make underwater-specific improvements starting from onshore algorithms. For example, Meta's Segment Anything Model (SAM) [57] has given rise to underwater specialized variants such as Dual-SAM [58].

DL has also revolutionized underwater SLAM and 3D reconstruction techniques. Traditional visual SLAM relies heavily on geometric features such as corners in the environment and is prone to failure in underwater weakly textured scenes [59]. Traditional hand-designed features perform poorly when underwater image quality deteriorates, while DL can learn feature descriptions that are more invariant to lighting and blurring, significantly improving the accuracy and robustness of feature matching. Advanced features that are more robust to light and blur, such as those learned by SuperPoint [60], have directly led to more robust visual odometry and pose estimation networks, such as DeepVO and PoseNet. Furthermore, at the back end of SLAM, DL has also brought breakthroughs to loopback detection, a key step in eliminating cumulative errors. For example, RCNN [61] borrowed the idea of probabilistic appearance recognition to determine whether the AUV had returned to the previously passed area and achieved good results.

Further, DL has driven the development of multimodal sensor fusion and semantic SLAM. For example, S2L-SLAM [62] converts sonar data into LiDAR point clouds through DL models, enabling existing LiDAR SLAM algorithms to continue to work accurately in complex environments where traditional sensors fail, achieving dynamic selection and fusion of sensor modalities. And the Sonar-CAD for Underwater Semantic 3D Mapping method proposed by Guerneve et al. [63] not only effectively fuses visual and SONAR heterogeneous data but also gives high-level semantic information to the map through semantic segmentation and target recognition techniques. This makes the final output of the information understanding module no longer a simple geometric point cloud, but a understood three-dimensional environment model, providing an unprecedented advanced cognitive input for the subsequent information judgment module.

3.2.2. Applications of Deep Learning in Information Understanding

Traditional schemes are based on artificially designed features and are essentially based on template matching or statistical matching. DL schemes build a deep, hierarchical internal representation of the world through data-driven learning, which contains object geometry and texture information, as well as abstract categories and context relationships. As a result, DL schemes have a qualitative leap in robustness and generalization, and can handle more complex and variable scenarios. However, its powerful representation learning ability relies on massive labeled data, and the inexplicability of the decision logic limits its application in some areas with high security requirements.

3.2.3. Development Summary and Future Predictions of Information Understanding Module

As shown in Table 2, the information understanding module's core evolves from low-level data description to high-level world understanding.

Table 2. Technologies related to the information understanding module section.

Traditional/DL	Name	Technology	Summary of the Technical Route
Traditional Solution	SIFT, HOG [10]	Hand-designed feature extractor	Extracts simple, predefined features for basic environmental understanding. Relies on geometric features like corners, prone to failure in weak-texture scenes.
	Traditional Visual SLAM	Geometric feature-based SLAM	

Table 2. *Cont.*

Traditional/DL	Name	Technology	Summary of the Technical Route
Deep-Learning Solution	MLR-VGGNet [50]	CNN architecture	Combining the VGGNet backbone with multi-layer residual, asymmetric and depthwise separable convolutions to optimize fish classification and reduce model parameters.
	The method based on mResNet [51]	CNN architecture	Underwater target recognition method based on mResNet and optimized feature engineering.
	DAMNet [52]	CNN with attention mechanism	Utilizes advanced attention mechanisms for complex biological image classification.
	MCANet [53]	CNN with attention mechanism	Utilizes advanced attention mechanisms for complex biological image classification.
	Faster R-CNN [54]	Two-stage deep-learning detection framework	Widely applied for underwater object detection and segmentation.
	YOLO improved variants [10,55,56]	Deep-learning detection framework	Mainstream for underwater object monitoring due to simplicity, open-source nature, and ease of deployment.
	FocusDet [57]	Fine-grained architecture for small object detection	Specialized for monitoring small objects like underwater trash.
	MLDet [58]	Fine-grained architecture for small object detection	Specialized for monitoring underwater trash.
	MTHI-Net [60]	Encoder–Decoder architecture	By using multi-task learning to hierarchically segment images, performance is enhanced, demonstrating innovation and potential in the field of image segmentation.
	BCMNet [61]	Encoder–Decoder architecture	Through bidirectional contrastive representation learning, more effective motion representations can be extracted from multimodal data,
	Dual-SAM [63]	Specialized foundation model for segmentation	Underwater-specific variant for fine-grained segmentation based on SAM.
	SuperPoint [8]	Feature-learning network	Learns robust high-level features for improved visual odometry and pose estimation.
	RCNN [64]	Deep learning for loop closure detection	Breaks through in loop closure detection by using probabilistic appearance recognition to eliminate cumulative errors in SLAM.
	S2L-SLAM [65]	Deep-learning model for multimodal sensor fusion	Converts sonar data to LiDAR point clouds, enabling LiDAR SLAM in challenging environments and dynamic sensor selection.
	SONAR-CAD for Underwater Semantic 3D Mapping [66]	Deep learning for multimodal sensor fusion and semantic SLAM	Fuses visual and sonar data, adding high-level semantic information to maps through segmentation and object recognition.

Future research focuses on building large underwater base models via large-scale self-supervised pre-training on multi-source data and multi-source fusion to integrate visual/sonar/laser data for robust cognition, supporting AUV's long-term autonomous operation.

3.3. Information Judgment Module

The fundamental task of the information judgment module is to transform abstract task objectives into specific, safe and efficient sequences of physical actions in highly uncertain, dynamic and information-constrained underwater environments and hand them over to the output module for implementation.

Overall, the decision-making logic of the information judgment module can be roughly deconstructed into two closely coupled but functionally distinct levels: kinematic decision-making and task-driven decision-making.

It is notable that DRL and its derived framework are important directions for the application research of information judgment modules and are expected to endow AUV information judgment modules with cognitive planning capabilities [8,67]. Reinforcement learning (RL) involves the core agent interacting with the environment and performing action A in the state for reward R to learn the optimal strategy aimed at maximizing the expected long-term cumulative return [64]. However, traditional RL methods (such as Q-Learning [65]) rely on tables to store the value of state–action pairs and encounter the curse of dimensionality when facing the high-dimensional or even continuous state space of AUV sensors, making storage and computation infeasible. The introduction of DL has addressed [66] this problem. DRL is an organic combination of DL and RL, which uses high-capacity deep neural networks as function approximators [68] to distill high-dimensional, noisy raw sensor data received by the AUV into low-dimensional, information-intensive, task-relevant feature vectors. Based on this, DRL directly learns end-to-end mappings through the network, either state-to-value mappings (like DQN [68]) or state-to-action direct mappings, which is similar to Policy Gradients in RL [69].

It is foreseeable that DRL is being deeply integrated into information judgment systems in AUV intelligent decision-making systems [70,71].

3.3.1. The Evolution of Information Judgment Module

Also shown in Table 2, kinematic decisions are the physical basis on which AUVs perform all tasks. Before DL intervened, this field was dominated by traditional planning algorithms based on precise models. For instance, article [72] describes a model that utilized graph search algorithms, including the A* algorithm, for global path planning. It also combined these with the dynamic window approach or the artificial potential field method for local, real-time obstacle avoidance. These schemes are essentially a kind of deductive logic [73], that is, the optimal motion trajectory is solved through optimized computation given an exact environmental model and dynamic model. In an unstructured real marine environment, the assumption of model accuracy is broken, and the robustness of the algorithm drops sharply [74]. At the same time, building a completely real and complex decision-making model requires elaborate mathematical modeling, which further increases the difficulty of making correct decisions.

DRL provides a completely different inductive logic for this. The DRL scheme bypasses the difficulty of precisely modeling the world and instead learns mapping from perception to action directly from high-dimensional, noisy raw sensor data through massive environmental interaction trial and error [75]. The goal is to generalize an optimal strategy that maximizes long-term cumulative returns, a strategy functionally similar to a highly optimized driving intuition.

DRL has developed different mainstream algorithms for different action spaces: deep Q-networks perform well for discrete actions such as turn left rudder, turn right rudder, and go straight [75]. DCMAC [76] approximates Q values through neural networks, enabling AUV to judge the long-term value of performing each discrete action from high-dimensional perception. Algorithms based on the Actor–Critic architecture, such as Deep Deterministic Policy Gradient [77] and Proximal Policy Optimization [78], are more applicable to continuous action spaces that are more common in AUV kinematic decision-making, such as precise rudder angles or thruster speeds. They output continuous control instructions directly through an Actor network and evaluate the quality of the instructions through a Critic network [79], thus efficiently learning smooth and robust motion strategies in complex continuous spaces.

The transplantation of VLM algorithms in the field of decision-making also shows great value. OceanPlan [80] pioneered an innovative Large Language Model (LLM) task–motion planning and re-planning framework aimed at addressing the core challenge of efficient and robust navigation of AUV through natural language instructions in vast, unknown marine environments. The core lies in a hierarchical planning system that includes LLM planners, HTN mission planners and DQN motion planners, complemented by a comprehensive re-planner to address the uncertainty of the underwater environment. Likewise, Autonomous Vehicle Maneuvering [81] integrates cognitive, decision-making, path planning and control functions to achieve real-time environmental adaptive LLM-guided path planning. Yang et al. [82] also reported on a VLM-powered ASV navigation system that enhances success in dynamic marine environments through improved path planning.

This data-driven paradigm essentially shifts reliance on precise models to reliance on massive data. The core advantage lies in its ability to implicitly encode high uncertainty in unstructured environments into policy networks, evolving the kinematic decisions of AUVs from static trajectory tracking to dynamic, real-time feedback environmental adaptation [83].

3.3.2. The Evolution of Task-Driven Decision-Making Schemes

Task decision systems involve the significance of AUVs. In traditional schemes, this level is typically handled by expert systems such as finite state machines or behavior trees. For example, HUXLEY [84] is a typical implementation based on hierarchical expert systems. The system uses a modular control hierarchy and internally organizes task flows through predefined state machines and behavior trees, enabling the AUV to perform reliable behavior switching in accordance with the sequence of rules set by the engineer when it is executed.

The introduction of DRL is historic: its learning paradigm makes it naturally suitable for the complex task-driven decision-making of AUVs. It allows AUVs to break away from reliance on human expert rules, no longer requiring precise physical models, and instead be able to self-learn how to perform complex tasks directly from interactions with the environment. DRL enables AUVs to perform exploratory complex tasks. For example, DRL-guided Autonomous Exploration with Waypoint Navigation [85] enables AUVs to train in completely unknown underwater cave simulation environments, using DRL agents to start from the points of interest perceived by the environment Autonomously plan waypoints and carry out exploration. Without any prior maps, you can eventually complete full coverage exploration in complex three-dimensional cave structures.

At the same time, VLM models are gradually becoming involved in the control system to deal with more advanced complex tasks. DREAM [86] has developed a VLM-driven autonomous underwater monitoring system that integrates multimodal perception, chain-based cognitive planning, and low-level control to enable underwater robots to

conduct efficient and comprehensive long-term exploration and target monitoring without human intervention. It builds dedicated maps to provide environmental memory and uses carefully designed prompts to guide the VLM to generate humanoid navigation strategies, demonstrating outstanding efficiency and coverage in both simulation and real-world experiments.

A lot of useful work has also been produced using VLMS to optimize AUV operations. Word2Wave [87] enables real-time programming and parameter configuration of AUV tasks through natural language, for example. Its proposed W2W framework includes a novel set of language rules and command structures, a GPT-based prompt engineering module for generating training data, a sequence-to-sequence learning pipeline based on the T5-Small small language model for generating task commands from human speech or text, and a user interface for 2D task map visualization and human–computer interaction, thus reducing task programming time and enhancing the user experience; this is designed for future AUV task programming without manual operation.

In terms of multi-agent decision-making, multi-agent reinforcement learning (MARL) provides a core framework for collaborative decision-making in multi-AUV clusters, enabling them to autonomously learn task allocation, formation flight, and collaborative confrontation strategies. UW-MARL [81] uses multi-agent RL to achieve adaptive sampling of AUVs, first exploring the environment through distributed Q-learning, collecting data and calculating variance as rewards to construct the initial environment map; then, in the second stage, tasks are assigned based on priority index, allowing the vehicle to be finely reconfigured within the MARL framework. Data sharing and collaboration are achieved through a customized communication protocol, and underwater environment monitoring is completed efficiently and economically. Similarly, HA-MRAL [88] focuses on the coordination, stability, convergence speed and high winning rate of wireless data sharing, and designs intelligent game strategies for multi-AUV underwater network systems.

3.3.3. Applications of Deep Learning in Information Analysis

The value of DL in the information judgment module lies in providing a completely new solution for the decision-making logic of AUV. Conventional planning algorithms follow deductive logic and rely on precise environmental and dynamic models to find the optimal solution, showing strong vulnerability in the highly uncertain conditions of the underwater unstructured environment. DRL provides inductive logic, circumventing the difficulty of precise modeling, learning end-to-end mapping with massive interactive data, transforming reliance on precise models into reliance on massive data, and facilitating the evolution of AUV decision-making from trajectory tracking to environmental adaptation.

3.3.4. Development Summary and Future Projections of Information Judgment Module

Despite the great potential of DRL and VLM, there are three core challenges on the way to high-reliability applications. One is the issue of sample efficiency. Learning requires massive interaction data, which is costly and risky in underwater environments [89]. The second is the black box problem, where the strategy decision-making logic is hidden in neural network parameters, and stability and security are difficult to analyze and verify, which is unacceptable in critical tasks. The third issue is deployment failure. Because simulators have difficulty in precisely simulating real physics [89], strategies trained in simulation environments often experience performance degradation or even failure when deployed to physical entities.

In response to these challenges, we believe that the information judgment module will make progress and develop in several areas in the future. First, a strong combination of AUVs and VLM. According to Wang et al. [23], existing VLA models are roughly divided

into three architectures: Discrete Token VLAs, Generative Action Head VLAs, and Custom Architecture VLAs. In fact, almost all existing VLAs in the field of underwater robotics are based on Discrete tokens. The Discrete Token VLA architecture has many limitations [23]. Through Table 3, one can conduct an analysis that the subsequent combination of AUV and VLA is bound to turn to the other two for better adaptability and resolution. The second is the expected wide application of Offline RL and Sim2Real technology in the AUV field. Because AUVs are at a disadvantage in terms of sample efficiency and safe trial and error, using massive offline navigation log data for policy learning will become an important training method. Third, future research will focus on setting safe boundaries for its decision-making, quantifying behavioral uncertainty and developing explainable tools to make its black box attributes transparent to some extent.

Table 3. Some related technologies of the information judgment module.

Whether to Integrate the Two Modules	Traditional/DL	Name	Technology	Summary of the Technical Route
No	Traditional	Huxley [81]	Hierarchical expert system (state machines, behavior trees)	Organizes task flows using modular control layers with predefined state machines and behavior trees.
		A * Path Planning Approach [77]	Graph search algorithm DWA APF	Used for local real-time obstacle avoidance in traditional approaches.
	Deep Learning	DRL-Guided Autonomous Exploration with Waypoint Navigation [88]	DRL agent	Autonomously plans waypoints and performs exploration in unknown underwater cave environments without prior maps.
		Word2Wave [90]	VLM SLM	Real-time programming and parameter configuration for AUV tasks.
		DREAM [91]	VLM	The VLM-driven underwater autonomous monitoring system integrates multimodal perception, cognitive planning based on thought chains, and low-level control
		RL Adaptive Underwater Arm Control [89]	Actor–Critic structure with DNNs	Demonstrates that RL controllers can outperform MPC in fine physical interaction.
		UW-MARL [92]	Q-learning MARL	MARL with distributed Q-learning for adaptive underwater sampling, coordinating autonomous vehicles via shared Q-values.
		HA-MARL [93]	APF MAPPO	It enhances multi-AUV data sharing by integrating APF for path planning and a Tabu-Search task scheduler into MAPPO.
		UnderwaterVLA [28]	VLM VLA	The dual-brain architecture and zero-data training enable robust autonomous navigation of underwater VLA.
Yes		OceanPlan [94]	LLM planner, HTN task planner, DQN motion planner, replanner	Addresses efficient and robust AUV navigation in unknown oceans via natural language instructions.
		Autonomous Vehicle Maneuvering [81]	LLM-guided path planning	Achieves real-time environmental adaptive LLM-guided path planning by integrating cognitive, decision-making, path planning, and control functions.

3.4. Output Module

The output module is responsible for maintaining system stability and performing simple tasks assigned by the judgment module. Advanced control schemes empowered by DL do not attempt to completely subvert classical control theory, but rather, infuse it with powerful adaptive and learning capabilities through ingenious integration.

3.4.1. Evolution of Output Module

The core objective of classical control theory is to ensure the stability of the system and the precise tracking of the preset trajectory. At the motion control level, the PID controller has been widely used for a long time because of its simplicity and effectiveness [91]. But with the increasing demands for accuracy and robustness in tasks, modern robust control methods such as SMC [90] were later introduced to better suppress external disturbances such as ocean currents.

At the operational control level, for AUV with robotic arms, traditional methods mainly rely on inverse kinematics to calculate joint trajectories and track them through PID controllers in the joint space [92]. This approach shows rigidity and vulnerability when in contact with the environment or when facing an uncertain target position, making it difficult to complete fine physical interactions [92]. It can be concluded that the evolution of the output module is first reflected within traditional control theory, that is, from classical stability to robust trajectory tracking, and from fixed motion to closed-loop adaptive motion with admittance/impedance control introduced. Then comes the introduction of DL.

One core direction in dynamic stabilization is using neural networks for model identification and adaptive control. The hydrodynamic model of the AUV is highly nonlinear and time-varying, making it difficult to model precisely. In recent years, dynamic neural networks have been proposed to address the modeling uncertainty and parameter perturbation problems of AUVs, such as DNCS [93] which online-learns the unknown dynamics of the system and adjusts the control gain in real-time by constructing a parallel identifier structure based on Lyapunv stability. The advantage of this scheme is that it can achieve high-precision trajectory tracking without relying on precise hydrodynamic parameters and has strong robustness against unknown perturbations.

However, the computational load is high, and the sensitivity of neural network weight initialization may lead to a decrease in convergence speed, which may be limited in tasks with extremely high real-time requirements. Cortez et al. [95] embedded prior knowledge of fluid mechanics into the network structure to enhance the generalization ability of the model by constraining the evolution law of the state space. This approach significantly reduces training data requirements in low-speed cruise scenarios while maintaining the physical consistency of dynamic characteristics; however, it has shortcomings, such as insufficient adaptability to high-speed maneuvers or strong turbulence conditions, and the stability of the numerical solution of the differential equation is limited by the step size.

Operationally, the introduction of DL enables the output module to be more deeply integrated with the information judgment module. For example, Cimurs et al. [94] present a DRL controller based on the Actor–Critic architecture, which directly takes joint position, velocity, and target state as inputs and outputs torque instructions for precise position control. A closed-loop operation of the output module and the information judgment module is achieved. However, it is worth noting that it has not been applied in the field and still has a lot of room for development. The LLM proposed by Kim and Choi [83] is directly and deeply integrated into an algorithmic program while integrating PINN's environmental awareness network module and incorporating flow field data into the state space of the AUV to achieve iterative optimization of the AUV structure and control strategy, thereby

significantly improving the adaptability and mission performance of the AUV in complex underwater environments.

3.4.2. Applications of Deep Learning in Information Output

For kinematic stability maintenance tasks, traditional control schemes are rooted in rigorous mathematical derivations. The integration of DL, through the powerful nonlinear function approximation ability of neural networks, enables the fusion scheme to identify and compensate for unmodeled dynamics and parameter uncertainties in traditional models online. This will avoid the huge amount of work of designing algorithms manually. It would be almost impossible to design a computational formula that is so well-considered without the involvement of DL. This is, of course, is at the expense of some efficiency, which is also the advantage of the traditional control of the output module-manual algorithms that still maintain their efficiency advantage within an acceptable margin of error, which is exactly what the computationally scarce underwater robot values.

On the other hand, DL does not have a significant impact on the output module in terms of task implementation. For complex tasks, the difficulty mainly lies in understanding and breaking down the tasks. The simple tasks that are broken down can actually be done very well with the traditional approach. This does not mean that DL is of no value in the output module. In fact, integrating the output module with the information judgment module to form a unified architecture similar to the VLA model has its unique value, but there are also corresponding shortcomings, and how to make trade-offs will be analyzed in detail in the next subsection.

3.4.3. Integration and Separation of Output Modules and Information Judgment Modules

As shown in Table 4, it is notable that many of the cases cited in the introduction of the information judgment module are end-to-end solutions that combine the output module with the information judgment module. The RL strategy proposed by Kim and Choi [83] directly outputs the underlying speed and angle control instructions. In fact, it takes advantage of the easy fusion feature of DL to integrate the output module with the information judgment module as a whole. There are also many options that handle the information judgment module and the output module separately. For example, Carlucho et al. [96] reported a model which has a clear hierarchy where the RL is responsible for high-level task decision-making and the S-Surface controller is responsible for low-level real-time control. It can be seen that whether the information judgment and output modules are integrated can be used as a criterion to divide two different technical solutions. This is analyzed and explained in Section 2.4.

DL inherently has an end-to-end tendency [97]. Based on whether the algorithm can be clearly split into an information analysis end and an output end, current solutions can be roughly divided into two technical paradigms: hierarchical decision architecture and end-to-end decision architecture. A review of these two approaches constitutes the core of understanding the evolution of current intelligent decision-making technologies.

The hierarchical architecture does not fully embrace DL solutions and does not thoroughly reflect the end-to-end technical inclination of DL. It can also clearly decouple complex decision-making problems, allowing different levels of strategy to be designed, trained, and debugged independently. Because of its operational semantic sub-objectives, high-level decisions are highly interpretable, and low-level skills can be reused in different tasks after being learned, with good combinatorial generalization ability. However, the challenges are obvious: the interaction between the lower and higher levels significantly reduces resolution, and when dealing with complex problems, it is less efficient and professional than the end-to-end model.

Table 4. Output module-related technologies.

Whether to Integrate the Two Modules	Traditional/DL	Name	Technology	Summary of the Technical Route
No	Traditional	PID Controller [97]	PID	Provides simple and effective stable tracking for predefined paths.
		SMC [98]	Sliding Mode Control	Offers robust control to suppress external disturbances like sea currents.
		Inverse Kinematics + PID [79]	Inverse Kinematics, PID	Calculates and follows joint trajectories for manipulator arms.
	Deep Learning	DNCS [99]	DNN	Online learning of unknown system dynamics to adaptively adjust control gains for tracking.
		Domain-aware Control-oriented Neural Models for Autonomous Underwater Vehicles [100]	Physics-Informed Neural Network	Embeds hydrodynamic priors into the network to improve generalization.
		RL Adaptive Underwater Arm Control [89]	Actor–Critic structure with DNNs	Demonstrates that RL controllers can outperform traditional MPC for fine physical interaction.
Yes		OceanPlan [94]	LLM planner, HTN task planner, DQN motion planner, replanner	Addresses efficient and robust AUV navigation in unknown oceans via natural language instructions.
		Autonomous Vehicle Maneuvering [81]	LLM-guided path planning	Achieves real-time environmental adaptive LLM-guided path planning by integrating cognitive, decision-making, path planning, and control functions.

The theoretical charm of the end-to-end approach lies in the fact that it minimizes the injection of human prior knowledge and avoids the loss and potential suboptimal decomposition problems caused by the transmission of information between modules in a hierarchical architecture. This is precisely the drawback of the hierarchical architecture. In theory, an end-to-end model deep enough and data-rich enough could potentially discover more efficient control strategies beyond human intuition. Its technical implementation relies entirely on DRL, particularly algorithms capable of handling high-dimensional continuous input and output, such as PPO [98], SAC [79] and other RL algorithms.

However, the practical challenges of end-to-end learning are huge. The first is the astonishing sample complexity. Due to the lack of guidance and supervision from intermediate targets, agents have to blindly explore in a vast state–action space, resulting in an extremely slow learning process that requires massive amounts of interaction data, which is almost unrealistic in a real underwater environment. Next comes the serious black box problem. The entire decision-making process takes place within an unresolvable neural network, and when the system malfunctions, it is almost impossible to diagnose whether the problem lies in the perception part, the decision-making logic part, or any other link. This inexplicability greatly limits its application in safety-critical tasks. Finally, the strategies learned end-to-end are often highly overfitted to their training environment

and tasks. An end-to-end strategy trained for the pipeline tracking task may be completely unable to handle the docking task because it does not learn any transferable, modular knowledge and shows poor task generalization ability.

The choice between the two architectures is often related to the pattern of decision-making.

Applications of AUV kinematic algorithms often adopt an end-to-end approach [99,100]. Because its output is relatively straightforward, it is suitable for end-to-end deployment in all aspects. The advantage also lies in the high extraction efficiency that surpasses manual design.

For example, an end-to-end AUV navigation strategy might directly learn from the fan-shaped scan of the forward-looking sonar. It extracts implicit information about the obstacle's distance, orientation, and movement trend. Then, it maps this information directly to the thrust difference between the left and right thrusters. This process enables reactive obstacle avoidance.

But there are also different schemes, such as Underwater VLA [22], which uses a layered motion structure and achieves good results as well.

It is different in the field of task implementation. End-to-end solutions are not gaining an overwhelming advantage nowadays because of their data problems and black box issues. Scholars tend to selectively choose based on the complexity of the target operation.

3.4.4. Development Summaries and Future Projections of Out Put Module

The output module's evolution forms the physical basis for AUVs' transformation from simple active platforms to intelligent agents, core being the shift from passive instruction execution to active task completion. Future trends may include the development of AUV-specific end-to-end models (due to underwater environment specificity) and deep integration with the judgment module.

4. Application Analysis and Technology Selection of AUV Intelligent Decision-Making System Empowered by Deep Learning

The previous section systematically dissects the development of the four major modules in the context of DL. However, the vitality of technology is ultimately reflected in its ability to solve practical problems. Therefore, we will go beyond the limitations of traditional classifications based on application domains, such as scientific expedition or military defense, as such divisions tend to blur the commonalities of underlying technical requirements. Instead, this paper proposes to systematically deconstruct the application scenarios of AUVs with "task complexity" and "environmental uncertainty" as two mutually orthogonal core dimensions. The superiority of this division lies in the fact that it directly relates to two fundamental determinants of the level of intelligence required by the decision-making system: the depth of the internal decision-making logic and the difficulty of external perception adaptation, and can provide a more precise theoretical basis and practical insights for technology selection under the four-module architecture.

4.1. Division Criteria: Orthogonal Deconstruction of Task Complexity and Environmental Uncertainty

The first dimension, task complexity, measures the requirements of the task for the AUV's advanced cognition, long-term planning, and fine physical interaction capabilities. It is divided into two levels.

Simple tasks: These are tasks with a single objective, whose sequences of actions are mostly predefined, and whose interaction logic with the environment is relatively clear.

Complex tasks involve multi-objective collaboration, continuous interaction in dynamic environments, decision chains that require long-duration reasoning, or understanding and decomposing high-level abstract semantic instructions.

The second dimension, environmental uncertainty, measures the pressure exerted by the AUV operating environment on the decision-making system from the perspective of external challenges. Structured environment refers to scenarios with relatively stable and predictable environmental features and abundant prior information, such as inland lakes, aquaculture cages, ports with clear underwater structures, etc. In these environments, the changes in marine dynamics, such as light and water flow, are relatively gentle, and the degradation of perceived information is within a relatively controllable range.

The unstructured environment represents the true appearance of the ocean: dynamic and changeable, lacking prior maps, and full of unknown random factors, such as deep-sea hydrothermal vents, high-turbidity estuarine areas, and sunken ships or underwater ruins formed by human activities. These environments are often accompanied by extremely poor or even zero visibility, intense and spatiotemporal turbulence, rugged and complex seabed topography, and a large number of dynamic or static unknown obstacles, which pose challenges to the viability and environmental adaptability of AUVs.

As shown in Table 5, based on this two-dimensional orthogonal framework, typical applications of AUVs can be clearly classified into the following four quadrants. This classification goes beyond the surface application labels and points directly to the essence of the technical challenges, thus laying a solid logical foundation for subsequent targeted technical route analysis and the role positioning of DL.

Table 5. Different scenarios classified based on task difficulty and environmental characteristics.

	Simple Task	**Complex Task**
Structured environment	Scene One: Routine operations for efficiency and cost optimization	Scene Three: Pre-determined operations for high-precision physical interaction
Unstructured environment	Scene Two: Goal-oriented behavior with strong environmental robustness	Scene Four: A fully autonomous system oriented towards unknown exploration and dynamic interaction

4.2. Scene Analysis and Deep-Learning Technology Selection

4.2.1. Simple Tasks, Structured Environments

This scenario is the most mature area for the commercial application of AUV technology, with the core driving force being the replacement of high-cost, high-risk manual operations with automation. Therefore, the core design demands of the entire intelligent decision-making system focus on reliability, operational efficiency, and operational costs. Typical examples of this aspect include pipeline inspection, dam structure inspection, and cage culture monitoring. In this context, DL is not a necessary condition for achieving autonomy, but rather a powerful tool for performance optimization and process automation, whose value lies in significantly raising the performance ceiling and economic benefits of traditional technical solutions.

In the information processing module, traditional image enhancement algorithms, such as Adaptive Histogram Equalization CLAHE [101], are often able to meet the basic requirements for visual information quality.

The information understanding module is where DL exerts its core commercial value in this scenario. The information understanding module can replace human interpretation of massive video/sonar recordings for semantic segmentation and information extraction, which can greatly improve efficiency and reduce costs. For example, Musa et al. [102] developed a density-based DL model to automatically identify and calculate the number of

fish in sonar images. Additionally, a model [103] trained a high-precision object detection network through supervised learning on a large amount of labeled pipe defect data and deployed it on the edge-computing unit of the AUV to achieve automated, real-time detection and classification of abnormal conditions. The key to technology selection is to seek the best balance point between inference speed and model accuracy while ensuring as few omissions as possible, in order to fully match the limited on-board computing resources and inspection speed of the AUV.

As for the information judgment module and the output module, since the task paths are largely pre-planned, the ability to plan complex dynamic paths is not an essential need and the requirements are not high.

To sum up, in the scenario of simple tasks and a structured environment, the core contribution of DL application is to achieve efficiency improvement by automating key links in the information understanding module. Given the low demand for dynamic response and strong adaptability of the control system in this scenario, introducing redundant complex control schemes would not only bring unnecessary system complexity and verification difficulties, but also significantly increase computational overhead, contrary to the scenario's demand for low cost.

4.2.2. Simple Tasks, Unstructured Environment

Entering this section, the task objective itself may still be singular, yet the extreme uncertainty of the execution environment becomes the primary issue that determines the success or failure of the task and even the survival of the AUV itself. This environmental case, such as specific marine life tracking, deepwater hydrological profile measurement, seabed topography mapping, etc., poses challenges to system stability. Therefore, the core task of the intelligent decision-making system is to maintain basic task capabilities and ensure the system's survival. Robustness becomes the overriding technical concern.

In this scenario, the performance of the information processing module is the key point for the intervention of DL technology. Deep-sea environments are murky or completely dark, optical sensor information is of low quality, and the system is highly dependent on acoustic sensors, etc., for obtaining information. However, raw acoustic data in complex environments cannot be effectively utilized without processing [104], so the information processing module becomes the focus. For example, MHGAN [105] designs GAN-based unsupervised learning models that can learn the intrinsic manifold structure of sonar images without pairs of clean–noisy data and recover reliable seabed topography and obstacle information from a strong noise background. In addition, when visual information is partially available, powerful low-light image enhancement [106] or defogging models [107] are also key to ensuring multimodal perception redundancy.

The information understanding module also faces huge challenges. Even after processing, the perceived information remains sparse and uncertain. The task is simple, but target recognition and model training become extremely difficult. For example, continuously tracking a specific target in a dimly lit, complex coral reef background requires models with strong generalization capabilities, such as WebUOT-1M [108] and an improved LeNet-5 CNN [109]. For cadastral mapping tasks, simply outputting a three-dimensional point cloud is not enough. It is necessary to use models similar to those based on improved Deeplab [110] and Segformer [111] to achieve real-time geomorphic classification of sonar images for a smarter understanding.

High-level task planning is relatively simple in the decision-making logic system of the information judgment module. However, the output module needs to have strong anti-disturbance performance.

The core survival skill of AUVs is dynamic obstacle avoidance, which can no longer depend on traditional planners with precise maps or on DRL strategies that output end-to-end evasive actions from processed sensor raw data streams. Through trial and error training in extremely dangerous scenarios, this strategy can acquire nearly instinctive obstacle avoidance behavior beyond traditional geometric algorithms [112]. For example, Gao et al. [113] suggested using small neural networks to estimate and compensate for unknown interference torques generated by external fluid dynamics online, and by seamlessly integrating the compensation terms into robust control frameworks such as sliding mode control or model predictive control, and ensuring that the AUV is in the harshness of sea conditions, it can precisely execute various emergency avoidance or fine tracking instructions from the judgment module.

In summary, under the harsh challenge of simple tasks and unstructured environments, the core value of DL lies in ensuring the basic survivability of the system and the minimum robustness of task execution. The focus of its technical application is strategically concentrated on the information processing module that interacts directly with the external environment, the information understanding module that understands the environment, and the output module.

4.2.3. Complex Tasks, Structured Environments

The technical challenges in this scenario are in sharp contrast to those in Scenario Two. The environment itself is cooperative and predictable, and the uncertainty at the perceptual level is minimized. However, the high complexity of the task itself places demands on the logical reasoning level of the information judgment module, the ability to break down tasks, and the precision of the physical operations of the output module; for example, complex operations such as underwater plugging and unplugging docking of base stations on the seabed observation network, underwater valve operation, and multi-AUV collaborative salvage of known targets. The information processing and understanding module can usually adopt a relatively mature technical solution.

The information judgment module is the core where DL plays a decisive and irreplaceable role in this scenario. The inherent complexity of the task demands an advanced architecture that can perform temporal and semantic abstraction beyond simple reactive decision-making. Both the end-to-end model and the hierarchical model introduced earlier have made useful explorations into this. For multi-AUV collaborative tasks, the MARL framework becomes an inevitable choice for learning efficient communication protocols, role assignment, and collaborative action strategies among AUV groups in a centralized training, distributed execution paradigm to maximize overall task efficiency.

The output module is equally crucial because the ultimate success or failure of the task depends on the quality of the physical interaction. For operations involving precise contact, high demands are placed on the flexibility and precision of control. Combining deep reinforcement learning with classical admittance/impedance control theory is a promising research direction in this context [114]. Additionally, for highly complex sequences of robotic arm operations, a high-quality initial strategy can be rapidly pre-learned from remote operation teaching data of human experts via learning by demonstration (LbD) [103], and subsequently self-improved online through DRL, thereby accelerating the learning process further.

In this context, the focus of DL applications is strategically shifted to information analysis and output modules and output solutions that deeply integrate learning with classical control theory, becoming the technical core that defines the level of autonomy in this scenario.

4.2.4. Complex Tasks, Unstructured Environment

Such an environment is the ultimate test of the comprehensive capabilities of the AUV intelligent decision-making system, which integrates and amplifies the core challenges of the aforementioned scenarios. The system not only needs to complete the preset tasks but also needs to have the core capabilities of continuous learning during execution, dynamic re-planning, and coexistence with uncertainty; for example, autonomous detection and internal mapping of unknown shipwrecks in the deep sea, autonomous marine scientific research and opportunistic sample collection, and underwater post-disaster search and rescue. At present, there is no complete AUV architecture that can fully autonomously achieve the goal in this environment, so we make predictions about the development of AUV.

The information processing module must have a powerful real-time, multimodal, generative denoising preprocessing capability. In response to this environment, in addition to the solutions mentioned above, when a critical sensor fails due to environmental influences, the system should be capable of selecting generative models for missing modal normalization; for example, the architecture based on reasonable estimation of missing modal information through generative reasoning in order to maintain the continuity and integrity of cognition of the external world [115].

The deep integration of information understanding and information judgment is beneficial for responding to this scenario. The technical goal is to build an adaptive planning intelligent system. The first step is to build and maintain in real-time a dynamic external understanding model. This model should incorporate not only the geometry of the environment but also semantic information and physical dynamics. Based on such a world model, the information judgment module performs intelligent planning. Next, instead of directly learning a black box strategy from the original observation of the action, the AUV first builds a dynamic predictive model through end-to-end learning that can predict how the world model will evolve in the next moment if one takes a certain action, thereby obtaining the best sequence of actions to achieve the long-term goal.

Also, the system's ability to generate task planning is put to the test when faced with extremely vague open-ended instructions such as explore this wreckage or assess the ecological condition of the area. The decision system of the AUV, combined with the common sense knowledge base and reasoning ability of the LLM, may help to push this forward.

Finally, all decision-making processes must be risk-aware. In a situation where perception, understanding, and prediction are all uncertain, the optimal decision is not simply about maximizing the expected return, but about making careful trade-offs between return and risk. This enables the AUV to make the most reasonable choice between high-return-but-high-risk aggressive behavior and low-return-but-absolute-safety conservative behavior based on the task context and its own state.

The output module is also expected to be deeply integrated with the above modules to enhance the complexity and efficiency of actions. Form a tight closed loop with extremely high information exchange rates and ultra-low latency. The dynamics output module should exist independently as a closed-loop module to deal with sudden shocks. For example, the underlying controller needs to be able to feed back in real-time and quantify information such as its own tracking error, uncertainty of state estimation, and control margin to the upstream information judgment module. Based on these feedback, the judgment module dynamically adjusts the safety margin of its planned path and the aggressiveness of its behavior to achieve true risk sharing and system-wide collaborative optimization.

4.3. Key Points and Insights of This Chapter

Through the aforementioned systematic analysis based on the task complexity–environmental uncertainty framework, it aims to provide a higher-dimensional strategic thinking framework for future research and engineering practice in this field.

The primary point is that DL plays a significant role in dealing with different environments, but its value shifts dynamically based on the core contradictions of different scenarios. In Scenario 1, it is responsible for improving efficiency. In Scenario 2, its role is to safeguard the system. In Scenario 3, it is the core that gives the system advanced cognitive and fine operational capabilities. Ultimately, in Scenario 4, DL becomes the underlying operating system that drives the entire system to achieve general autonomous capabilities, requiring a high degree of integration and co-evolution of all modules. The evolution of this series of roles clearly indicates that the application of DL technology must identify the core problem it needs to solve in a specific scenario.

This leads to a second key insight: a precise match between the problem and the technology is a fundamental principle for achieving high-performance autonomous systems. Using the model that requires massive computing resources in Scenario 4 for the pipeline inspection task in Scenario 1 not only wastes resources but may also reduce reliability due to the complexity of the system. Instead, using the simple reactive strategy of Scenario 2 to deal with the operational tasks of Scenario 3 will not achieve the desired effect. Therefore, this framework advocates for a deep understanding of the core technical challenges represented by each quadrant and the selection of highly self-disciplined DL architectures that match them. This precise matching is not only a prerequisite for achieving technical performance optimization, but also a key to ensuring that the system can operate reliably and efficiently on resource-constrained AUV platforms.

Finally, the two-dimensional framework constructed in this section is not only a summary of the existing application scenarios but also provides a clear, nonlinear evolutionary roadmap for the development of AUV autonomous technology. The development path of a country or research institution in AUV autonomous technology can be regarded as a process of continuous expansion of its capability coverage on this two-dimensional plane. From a business perspective, starting from Scenario 1 and expanding to Scenario 2 or Scenario 3, respectively, it represents two different technological upgrade routes for enhancing environmental adaptability and task complexity. From the perspective of cutting-edge scientific exploration, Scenario 4 is undoubtedly the ultimate goal. However, the path to Scenario 4 must be based on the full resolution of the technical challenges represented by Scenarios 2 and 3. This points the way for strategic investment in research resources: continued investment in fundamental capabilities such as multimodal perception fusion, and robust control is the solid foundation that underpins the future edifice of more advanced, more general autonomous intelligence.

5. Challenges, Frontiers and Future Prospects

At this point, we have clearly mapped out a technological evolution trajectory. While each subsection of Section 3 offers a future outlook for this module, this section aims to go beyond the boundaries of existing technology for an overall comprehensive analysis, with the AUV intelligent decision-making system as the target, to envision a future of truly autonomously intelligent underwater robots driven by higher-level artificial intelligence.

5.1. Challenges and Frontiers

5.1.1. Dual Scarcity of Underwater Perception Data

The first challenge is the dual scarcity of underwater perception data in terms of both quality and quantity. DL is rooted in vast amounts of well-labeled data. However, the

nature of the underwater environment makes obtaining high-quality data costly and time-consuming [116]. This has led to current underwater datasets not only being far inferior in absolute numbers to land-based applications, but also lacking in coverage of extreme sea conditions, rare landforms, and long-term dynamic processes. It restricts the generalization ability of AUVs. A visual enhancement or target recognition model trained in clear water, shallow sea areas will experience a cliff-drop in performance when entering deep sea or high-turbidity estuaries [117]. Simulation training as an alternative also faces a huge gap from simulation to reality. Despite techniques such as domain randomization [118], current underwater simulators are still unable to precisely simulate complex optical scattering, acoustic multipath, and nonlinear fluid dynamics effects, resulting in decision-making strategies trained in simulated environments often not achieving the expected results when transplanted to real AUVs.

5.1.2. Black Box Problem

Also, there is a contradiction between the inherent black box nature of DL models, especially DRL decision models, and the absolute requirements for interpretability and verifiability in safety-critical tasks. When AUV perform high-risk tasks such as undersea critical infrastructure maintenance like Scenario 3, or deep-sea search and rescue like Scenario 4, we not only need the system to make correct decisions but also need strict validation of its behavioral boundaries before deployment. However, the decision-making logic of current mainstream end-to-end DRL strategies is buried deep within millions of neural network parameters, making it difficult to explain in human-understandable language. Once unexpected behavior occurs, we cannot make effective attributions and corrections. This situation limits the depth to which DL can be applied in the core decision-making system of AUVs.

5.1.3. Limitations on Computational Capacity

Secondly, there is a conflict between the increasing complexity of cutting-edge algorithmic models and the inherent energy, computing power, and volume constraints of AUV platforms. DL relies on massive computing resources. Because underwater communication is difficult to achieve high-speed real-time communication, cloud processing is almost impossible. Edge deployment is also difficult. As a closed system that has to operate independently for long periods without cables, AUVs have extremely limited energy supply and computing power and cannot carry powerful computing hardware. This leads to the most powerful advanced algorithms that are expected to solve other problems often being unable to be deployed in AUVs due to their huge computational overhead. Therefore, how to achieve extreme lightweight and efficient design at the algorithmic level, develop energy-efficient AI acceleration chips at the hardware level, and achieve coordinated optimization of software and hardware becomes a crucial step in determining whether cutting-edge technologies can move from the laboratory to the real ocean.

Notably, this computational overhead directly translates to significant delays in deep-learning pipelines from sensing to actuation—a key reason why classical control methods remain prevalent in AUV systems. Classical controls, with their low-latency response and deterministic timing behavior, align better with the real-time requirements of underwater operations where even minor delays can compromise mission success. To address this gap, future work should explicitly incorporate timing constraints into the design of deep-learning-based decision systems; such an explicit treatment would not only enhance the practicality of these systems but also strengthen the rigor of the research by bridging theoretical advancements with real-world operational needs.

5.1.4. Fragmentation of Applications

Finally, the fragmentation of applications poses a challenge to the academic community. Most current DL solutions are tailored for highly specific application scenarios. This customization is manifested on three levels: first, the solidification of sensor modalities, a visual recognition model designed for high-resolution optical images that cannot directly process sparse acoustic data from forward-looking sonar. The second is the rigidity of the mission target, a model trained for pipeline inspection tasks and excelling at detecting corrosion and cracks on the surface of cylinders, which cannot be transferred for identifying marine life in complex coral reef ecosystems. Finally, there is the binding of platform dynamics, a reinforcement learning control strategy trained for a particular AUV that is highly adapted to its hydrodynamic characteristics and thruster layout, which almost inevitably fails when switched to another AUV with very different dynamic characteristics. This leads to a huge amount of repetitive work and high development costs. To break through this bottleneck, both academia and industry must look to exploring and building AUV foundation models and universal decision-making frameworks that cover a wide range of underwater environments and tasks across modalities and platforms. This, in turn, requires us to build an unprecedentedly large, multimodal underwater data bank.

5.2. Frontier Technology Trends and Future Prospects

To overcome the above challenges, the academic and industrial sectors are exploring cutting-edge solutions from three dimensions: data, models, and training paradigms. These trends are reshaping the boundaries and functions of the four modules.

5.2.1. Underwater Foundation Models and Self-Supervised Learning

To address the problem of data scarcity and application fragmentation, building a large underwater base model has become an important trend, that is, to learn a general underwater scene representation model by conducting large-scale self-supervised pre-training on massive, multi-source visual, sonar unlabeled underwater data. This model can greatly enhance the fine-tuning performance of information processing and understanding tasks on small sample data, fundamentally alleviate the reliance on expensive manual labeling, and achieve a paradigm shift from one task, one model to general pre-training.

5.2.2. Physical-Data Dual-Driven and Trusted AI

To address the issues of black box and security challenges, purely data-driven methods are evolving towards a hybrid paradigm constrained by physical knowledge. PINN is used to embed optical or acoustic propagation models as inductive biases into the network, making the enhancement results more in line with physical laws. In the information judgment and output module, trustworthy AI [119] becomes the focus. The core is the development of security reinforcement learning, such as using classical robust controls as the supervision of DRL policies to ensure that the exploratory instructions of DRL are always confined within the security envelope guaranteed by traditional control theory, achieving the decoupling and unification of intelligence and security.

5.2.3. Offline Learning and Sim-to-Real Efficient Migration

To break through the interaction bottleneck and insufficient data volume of DRL, efficient data utilization paradigms are at the forefront. As predicted in 3.3.4, Offline RL is attracting a lot of attention. Meanwhile, the academic community is narrowing the simulation–reality gap by improving simulation techniques and developing domain-adaptive algorithms [120] for efficient and low-cost strategy migration from simulation to physical AUV [121].

5.3. Future Outlook: Moving Towards Integrated and Clustered Underwater Intelligence

Looking ahead, we believe that DL will drive the four modules of the AUV decision-making system from loosely coupled to deeply integrated, thereby driving the update of the specialized architecture of underwater robots and ultimately achieving the ultimate leap from individual autonomous decision-making to cluster-emergent intelligence.

5.3.1. Robot Architecture Focused on the Underwater Domain

At the same time, it cannot be ignored that land-based robots have difficulty communicating with underwater robots, which limits the development of DL applications in the underwater field. There are essential differences in fundamental characteristics between the autonomous systems of underwater AUVs and those on the ground. In terms of perception input, the underwater environment lacks stable lighting and is very prone to turbidity, and optical images have severe scattering, absorption, and color distortion, which makes it difficult to use visual input as the largest data input method as robots in the air environment [10]. The various underwater acoustic communications are characterized by low bandwidth, high latency and high noise, which makes it difficult to transmit information at high speed and in real-time like ground systems. These problems make it impossible for the mature processing models of ground systems to be directly reused underwater, and model reconstruction specifically for the fusion, enhancement and feature extraction of underwater multimodal data is required. In terms of dynamic control, the strong water-flow interference and complex fluid dynamics characteristics of the underwater environment make the motion model of the AUV highly nonlinear and uncertain, which is quite different from the relatively stable kinematic model of the ground robot, and pose higher requirements for the policy learning and generalization ability of DRL. Therefore, simply transplanting the DL framework of the ground robot is not sufficient to solve the autonomous decision-making problem of the AUV. As DL deepens in the AUV field, the differences between onshore robot control schemes and underwater robot DL control schemes are growing. In the future, there will inevitably be efficient architectures and separate systems focused on underwater autonomous control systems, rather than deploying underwater verification technologies for onshore schemes.

5.3.2. Perception–Cognition–Decision-Making Integrated Underwater Agents

In the previous text, we have discussed in depth the convergence trends of the modules of AUV intelligent decision-making systems under the influence of DL. The future AUV decision-making system will no longer be a linear series of four independent modules, but a highly integrated agent with an internal world model. AUV will no longer passively identify targets but actively generate and predict a coherent, high-fidelity underwater world within themselves. This internal world model will serve as a sandbox for the information judgment module to plan, enabling the DRL to perform efficient strategy deduction in imagination. The judgment module, which integrates the common sense reasoning ability of the VLM, will be able to understand highly abstract natural language instructions and autonomously break them down into adaptive control sequences that are executable by the output module and incorporate physical constraints.

5.3.3. From Individual Intelligence to Distributed Cluster Intelligence

The complexity and wide range of underwater tasks will eventually lead to multi-agent collaboration. As described in Section 3.3.2, MARL is at the core of achieving this goal. The future challenge will arise from autonomous decision-making for individual AUV to distributed decision-making for AUV clusters. At that time, all modules will be restructured at the cluster scale: evolving from single-point augmentation of information processing to

distributed data fusion; from individual cognition of information understanding to shared environment mapping; and from individual optimal strategies for information analysis to cluster collaborative emergence strategies. DL will endow underwater multi-AUV clusters with unprecedented collaborative capabilities, enabling them to achieve system-level performance far beyond any individual intelligence in wide-area search, collaborative confrontation, and distributed environmental monitoring.

5.3.4. Operational Paradigms of the Next Decade

To operationalize the "Four-Module" framework within the "Task-Environment" matrix proposed in Section 4, we project two distinct operational paradigms that define the future of underwater autonomy.

Semantic-to-Control Intervention (Structured Environment/Complex Task): Future AUVs will transition from trajectory tracking to semantic execution. In offshore infrastructure maintenance scenarios, the information judgment and output modules will be deeply integrated into a VLA architecture. Unlike current systems requiring explicit geometric programming for every motion, these agents will interpret abstract instructions and autonomously generate compliance control policies. By utilizing sim-to-real transfer techniques, the system will handle contact-rich manipulation tasks—such as underwater connector plugging—adjusting to hydrodynamic disturbances in real time without the latency constraints of surface teleoperation.

Open-Set Adaptive Exploration (Unstructured Environment/Complex Task): In deep-sea frontiers, the information understanding module will evolve from closed-set classification to open-set cognitive mapping driven by underwater foundation models. Upon encountering novel geological features or biological species absent from training datasets, the AUV will autonomously initiate an active perception loop. The information analysis module will dynamically restructure the mission profile, switching from coarse-grain sonar mapping to fine-grain multimodal scanning to maximize information entropy reduction. This capability transforms the AUV from a passive data logger into an active scientific agent capable of constructing high-fidelity semantic world models in situ.

6. Conclusions

The transition from model-based control to data-driven learning marks a critical inflection point in the evolution of AUV. This review has not merely cataloged the proliferation of DL algorithms, but rather dissected the underlying logic of this technological migration. By deconstructing the intelligent decision-making system into four discrete modules—processing, understanding, judgment, and output—we have clarified a chaotic landscape, demonstrating that DL is not a universal panacea, but a specialized tool that transforms how machines perceive and reason in unstructured aquatic environments.

The core contribution of this work lies in bridging the disconnect between algorithmic sophistication and engineering reality. The proposed "Task Complexity–Environment Uncertainty" analytical matrix challenges the prevailing trend of blind model stacking. We argue that the value of an intelligent system is defined not by the depth of its neural networks, but by the precision of its architectural fit. Whether deploying a lightweight reactive policy for survival in turbulent currents or a heavy cognitive planner for shipwreck exploration, the choice must be dictated by the specific constraints of the operational quadrant. This framework serves as a strategic compass for researchers to navigate the trade-offs between computational cost, robustness, and autonomy levels.

Looking beyond the horizon, the era of isolated AUVs is drawing to a close. The future belongs to embodied underwater intelligence—systems that possess internal world models, risk-aware reasoning, and the ability to generalize across missions. While data scarcity and

the "black box" nature of end-to-end learning remain formidable barriers, the emergence of underwater foundation models, physics-informed learning, and high-fidelity Sim-to-Real transfer offers a tangible path forward. We stand on the precipice of a new paradigm where AUVs evolve from pre-programmed tools into adaptive explorers, ultimately unlocking the deep ocean through collaborative, emergent cluster intelligence.

Author Contributions: Q.D.: Original draft writing, investigation, formal analysis, figure creation; L.Y.: review and editing of the manuscript; H.C.: review and editing of the manuscript; H.L.: figure creation; A.L.: review and editing of the manuscript; W.C.: conceptualization, review and editing of the manuscript, investigation, resources, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Scientific Research Funding Project of Westlake University (Grant No. WU2024A001) and start-up funding from Westlake University under grant number 041030150118.

Acknowledgments: During the preparation of this manuscript, the authors used Doubao (an AI developed by ByteDance) for providing support in searching for information and language translation from Chinese to English throughout the research and paper drafting process. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Wu, H.; Chen, Y.; Yang, Q.; Yan, B.; Yang, X. A Review of Underwater Robot Localization in Confined Spaces. *J. Mar. Sci. Eng.* **2024**, *12*, 428. [CrossRef]
2. Stefanidou, A.; Politi, E.; Chronis, C.; Dimitrakopoulos, G.; Varlamis, I. A Deep Reinforcement Learning Approach for Navigation and Control of Autonomous Underwater Vehicles in Complex Environments. In Proceedings of the 18th International Conference on Control, Automation, Robotics and Vision (ICARCV), Dubai, United Arab Emirates, 12–15 December 2024; pp. 750–755.
3. Choi, H.S.; Lee, P.-M. Development of a System Architecture for an Advanced Autonomous Underwater Vehicle, ORCA. In Proceedings of the 4th Conference of International Conference on Control, Automation and Systems, Bangkok, Thailand, 25–27 August 2004.
4. Chemori, A. Control of Complex Robotic Systems: Challenges, Design and Experiments. In Proceedings of the 22nd International Conference on Methods and Models in Automation and Robotics (MMAR), Międzyzdroje, Poland, 28–31 August 2017; pp. 622–631.
5. Li, A.; Guo, S.; Liu, M.; Yin, H. Hydrodynamic Characteristic-Based Adaptive Model Predictive Control for the Spherical Underwater Robot under Ocean Current Disturbance. *Machines* **2022**, *10*, 798. [CrossRef]
6. Kozhubaev, Y.; Belyaev, V.; Murashov, Y.; Prokofev, O. Controlling of Unmanned Underwater Vehicles Using the Dynamic Planning of Symmetric Trajectory Based on Machine Learning for Marine Resources Exploration. *Symmetry* **2023**, *15*, 1783. [CrossRef]
7. Sun, X.; Liu, L.; Dong, J. Underwater Image Enhancement with Encoding-Decoding Deep CNN Networks. In Proceedings of the IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA, USA, 4–8 August 2017; pp. 1–6.
8. Yuan, J.; Wang, H.; Zhang, H.; Lin, C.; Yu, D.; Li, C. AUV Obstacle Avoidance Planning Based on Deep Reinforcement Learning. *J. Mar. Sci. Eng.* **2021**, *9*, 1166. [CrossRef]
9. Cao, X.; Sun, C.; Yan, M. Target Search Control of AUV in Underwater Environment with Deep Reinforcement Learning. *IEEE Access* **2019**, *7*, 96549–96559. [CrossRef]
10. Singh, P.; Sehgal, A. Computer Vision, Sonar and Sensor Fusion for Autonomous Underwater Vehicle (AUV) Navigation. In Proceedings of the Conference and Planning Committee Meeting, Seville, Spain, 28–31 March 2006.
11. Yang, Y.; Xiao, Y.; Li, T. A Survey of Autonomous Underwater Vehicle Formation: Performance, Formation Control, and Communication Capability. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 815–841. [CrossRef]
12. Spaho, E.; Matsuo, K.; Barolli, L.; Xhafa, F. Robot Control Architectures: A Survey. In *Information Technology Convergence: Security, Robotics, Automations and Communication*; Springer: Dordrecht, The Netherlands, 2013. [CrossRef]

13. Baranidharan, V.; Hari Murugesh, K.; Gokulvasanth, K.; Rahul Vikash, K.; Jane Mystika, D.; Renugadevi, S. Modeling and Control Techniques for Autonomous Underwater Vehicles—A Comprehensive Review. In Proceedings of the International Conference on Recent Advances in Science and Engineering Technology (ICRASET), Mandya, India, 21–22 November 2024; pp. 1–7.
14. Sunbeam, M. Deep Learning for Visual Navigation of Underwater Robots. *arXiv* **2023**, arXiv:2310.19495. [CrossRef]
15. Wang, X.; Zhang, H.; Liu, H.; Lewis, F.L. Control Oriented Reinforcement Learning: A Survey of Recent Progress and Applications. *Int. J. Robust Nonlinear Control* **2025**. [CrossRef]
16. Yu, D.; Lin, C.; Liu, N. Logic-Optimization Behavior Tree Algorithm for Enhanced Autonomous Underwater Vehicle Cooperative Decision-Making. *Complex Intell. Syst.* **2025**, *11*, 406. [CrossRef]
17. Zhao, S.; Yuh, J. Experimental Study on Advanced Underwater Robot Control. *IEEE Trans. Robot.* **2005**, *21*, 695–703. [CrossRef]
18. Wickramasinghe, C.S.; Marino, D.L.; Manic, M. ResNet Autoencoders for Unsupervised Feature Learning from High-Dimensional Data: Deep Models Resistant to Performance Degradation. *IEEE Access* **2021**, *9*, 40511–40520. [CrossRef]
19. Shao, R.; Li, W.; Zhang, L.; Zhang, R.; Liu, Z.; Chen, R.; Nie, L. Large VLM-Based Vision-Language-Action Models for Robotic Manipulation: A Survey. *arXiv* **2025**, arXiv:2508.13073.
20. Buchholz, M.; Carlucho, I.; Grimaldi, M.; Petillot, Y.R. Distributed AI Agents for Cognitive Underwater Robot Autonomy. *arXiv* **2025**, arXiv:2507.23735. [CrossRef]
21. Skorobogatov, G.; Barrado, C.; Salamí, E. Multiple UAV Systems: A Survey. *Unmanned Syst.* **2020**, *8*, 149–169. [CrossRef]
22. Zitkovich, B.; Yu, T.; Xu, S.; Xu, P.; Xiao, T.; Xia, F.; Wu, J.; Wohlhart, P.; Welker, S.; Wahid, A.; et al. RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control. In Proceedings of the 7th Conference on Robot Learning, Atlanta, GA, USA, 2 December 2023.
23. Wang, Z.; Zhu, Y.; Yan, Y.; Tian, X.; Shao, X.; Li, M.; Li, W.; Su, G.; Cui, W.; Fan, D. UnderwaterVLA: Dual-Brain Vision-Language-Action Architecture for Autonomous Underwater Navigation. *arXiv* **2025**, arXiv:2509.22441.
24. Goyal, A.; Hadfield, H.; Yang, X.; Blukis, V.; Ramos, F. VLA-0: Building State-of-the-Art VLAs with Zero Modification. *arXiv* **2025**, arXiv:2510.13054.
25. Li, Q.; Chang, B.; Mei, W.; Chen, Z. Integrated Sensing, Computing, Communication, and Control for Time-Sequence-Based Semantic Communications. *arXiv* **2025**, arXiv:2505.03127.
26. Jaffe, J.S. Computer Modeling and the Design of Optimal Underwater Imaging Systems. *IEEE J. Ocean. Eng.* **1990**, *15*, 101–111. [CrossRef]
27. Tian, Y.; Xu, Y.; Zhou, J. Underwater Image Enhancement Method Based on Feature Fusion Neural Network. *IEEE Access* **2022**, *10*, 107536–107548. [CrossRef]
28. Luan, X.; Hou, G.; Sun, Z.; Wang, Y.; Song, D.; Wang, S. Underwater Color Image Enhancement Using Combining Schemes. *Mar. Technol. Soc. J.* **2014**, *48*, 57–62. [CrossRef]
29. Zhang, S.; Wang, T.; Dong, J.; Yu, H. Underwater Image Enhancement via Extended Multi-Scale Retinex. *Neurocomputing* **2017**, *245*, 1–9. [CrossRef]
30. Therrien, C.W.; Frack, K.L.; Ruiz Fontes, N. A Short-Time Wiener Filter for Noise Removal in Underwater Acoustic Data. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Munich, Germany, 21–24 April 1997; Volume 1, pp. 543–546.
31. Sree Vidhya, K.S.; Deepthi, P.S. A Comprehensive Analysis of Underwater Image Processing Based on Deep Learning Techniques. In Proceedings of the International Conference on Control, Communication and Computing (ICCC), Thiruvananthapuram, India, 19–21 May 2023; pp. 1–6.
32. Wang, Y.; Zhang, J.; Cao, Y.; Wang, Z. A Deep CNN Method for Underwater Image Enhancement. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Beijing, China, 17 September 2017; pp. 1382–1386.
33. Li, J.; Skinner, K.A.; Eustice, R.M.; Johnson-Roberson, M. WaterGAN: Unsupervised Generative Network to Enable Real-Time Color Correction of Monocular Underwater Images. *IEEE Robot. Autom. Lett.* **2017**, *3*, 387–394. [CrossRef]
34. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. *arXiv* **2020**, arXiv:1703.10593.
35. Du, R.; Li, W.; Chen, S.; Li, C.; Zhang, Y. Unpaired Underwater Image Enhancement Based on CycleGAN. *Information* **2022**, *13*, 1. [CrossRef]
36. Steiniger, Y.; Kraus, D.; Meisen, T. Survey on Deep Learning Based Computer Vision for Sonar Imagery. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105157. [CrossRef]
37. Wang, Z.; Xue, T.; Wang, Y.; Li, J.; Zhang, H.; Xu, Z.; Xu, G. Enhancing Object Detection Accuracy in Underwater Sonar Images through Deep Learning-Based Denoising. *arXiv* **2025**, arXiv:2503.01655. [CrossRef]
38. Huang, Y.; Li, W.; Yuan, F. Speckle Noise Reduction in Sonar Image Based on Adaptive Redundant Dictionary. *J. Mar. Sci. Eng.* **2020**, *8*, 761. [CrossRef]
39. Liu, P.; Wang, L.; He, G.; Zhao, L. A Survey on Active Deep Learning: From Model-Driven to Data-Driven. *ACM Comput. Surv.* **2022**, *54*, 1–34. [CrossRef]

40. Drap, P.; Merad, D.; Boï, J.-M.; Mahiddine, A.; Peloso, D.; Chemisky, B.; Seguin, E.; Alcala, F.; Bianchimani, O. Underwater Multimodal Survey: Merging Optical and Acoustic Data. In *Underwater Seascapes*; Musard, O., Le Dû-Blayo, L., Francour, P., Beurier, J.-P., Feunteun, E., Talassinos, L., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 221–238, ISBN 978-3-319-03439-3.
41. Bach, N.G. Underwater Image Enhancement with Physical-Based Denoising Diffusion Implicit Models. *J. Image Graph.* **2025**, *13*, 2013–2230. [CrossRef]
42. Macatangay, X.; Gabriel, S.A.; Hoseinnezhad, R.; Fowler, A.; Bab-Hadiashar, A. Machine Learning for Modeling Underwater Vehicle Dynamics: Overview and Insights. *IEEE Access* **2024**, *12*, 139486–139504. [CrossRef]
43. Qu, P.; Li, T.; Zhou, L.; Jin, S.; Liang, Z.; Zhao, W.; Zhang, W. DAMNet: Dual Attention Mechanism Deep Neural Network for Underwater Biological Image Classification. *IEEE Access* **2023**, *11*, 6000–6009. [CrossRef]
44. Arora, A.; Kumar, A. HOG and SIFT Transformation Algorithms for the Underwater Image Fusion. In Proceedings of the IEEE International Conference on Technology, Research, and Innovation for Betterment of Society (TRIBES), Raipur, India, 17–19 December 2021; pp. 1–5.
45. Pachaiyappan, P.; Chidambaram, G.; Jahid, A.; Alsharif, M.H. Enhancing Underwater Object Detection and Classification Using Advanced Imaging Techniques: A Novel Approach with Diffusion Models. *Sustainability* **2024**, *16*, 7488. [CrossRef]
46. Prasetyo, E.; Suciati, N.; Fatichah, C. Multi-Level Residual Network VGGNet for Fish Species Classification. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *24*, 5286–5295. [CrossRef]
47. Reliable Object Recognition Using Deep Transfer Learning for Marine Transportation Systems with Underwater Surveillance Semantic Scholar. Available online: https://www.semanticscholar.org/paper/Reliable-Object-Recognition-Using-Deep-Transfer-for-Moghimi-Mohanna/46ad1003d1d3ce7edfa4be357373fa918b6d2244 (accessed on 25 October 2025).
48. Li, G.; Wang, F.; Zhou, L.; Jin, S.; Xie, X.; Ding, C.; Pan, X.; Zhang, W. MCANet: Multi-Channel Attention Network with Multi-Color Space Encoder for Underwater Image Classification. *Comput. Electr. Eng.* **2023**, *108*, 108724. [CrossRef]
49. Nawarathne, U.; Kumari, H.; Kumari, H. Underwater Waste Detection Using Deep Learning A Performance Comparison of YOLOv7 to 10 and Faster RCNN. *arXiv* **2025**, arXiv:2507.18967. [CrossRef]
50. Wang, L.; Chen, L.-Z.; Peng, B.; Lin, Y.-T. Improved YOLOv5 Algorithm for Real-Time Prediction of Fish Yield in All Cage Schools. *J. Mar. Sci. Eng.* **2024**, *12*, 195. [CrossRef]
51. Lei, F.; Tang, F.; Li, S. Underwater Target Detection Algorithm Based on Improved YOLOv5. *J. Mar. Sci. Eng.* **2022**, *10*, 310. [CrossRef]
52. Chen, J.; Er, M.J. Dynamic YOLO for Small Underwater Object Detection Artificial Intelligence Review. *Artif. Intell. Rev.* **2024**, *57*, 165. [CrossRef]
53. Shi, Y.; Jia, Y.; Zhang, X. FocusDet: An Efficient Object Detector for Small Object. *Sci. Rep.* **2024**, *14*, 10697. [CrossRef] [PubMed]
54. Ma, D.; Wei, J.; Li, Y.; Zhao, F.; Chen, X.; Hu, Y.; Yu, S.; He, T.; Jin, R.; Li, Z.; et al. MLDet: Towards Efficient and Accurate Deep Learning Method for Marine Litter Detection. *Ocean Coast. Manag.* **2023**, *243*, 106765. [CrossRef]
55. Chen, J.; Tang, J.; Lin, S.; Liang, W.; Su, B.; Yan, J.; Zhou, D.; Wang, L.; Lai, Y.; Yang, B. RMP-Net: A Structural Reparameterization and Subpixel Super-Resolution-Based Marine Scene Segmentation Network. *Sec. Ocean Obs.* **2022**, *9*, 1032287. [CrossRef]
56. Lin, B.; Dong, X. A Multi-Task Segmentation and Classification Network for Remote Ship Hull Inspection. *Ocean Eng.* **2024**, *301*, 117608. [CrossRef]
57. Bidirectional Collaborative Mentoring Network for Marine Organism Detection and Beyond Semantic Scholar. Available online: https://www.semanticscholar.org/paper/Bidirectional-Collaborative-Mentoring-Network-for-Cheng-Wu/1aba94e0a51def387c99cf1350a3ebf221269612 (accessed on 25 October 2025).
58. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; et al. Segment Anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 2–6 October 2023; pp. 3992–4003. [CrossRef]
59. Zhang, P.; Yan, T.; Liu, Y.; Lu, H. Fantastic Animals and Where to Find Them: Segment Any Marine Animal with Dual SAM. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 16 June 2024; IEEE: New York, NY, USA, 2024; pp. 2578–2587.
60. Zhang, S.; Zhao, S.; An, D.; Liu, J.; Wang, H.; Feng, Y.; Li, D.; Zhao, R. Visual SLAM for Underwater Vehicles: A Survey. *Comput. Sci. Rev.* **2022**, *46*, 100510. [CrossRef]
61. DeTone, D.; Malisiewicz, T.; Rabinovich, A. SuperPoint: Self-Supervised Interest Point Detection and Description. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 18–22 June 2018; pp. 337–33712.
62. Sangari, M.S.; Thangaraj, K.; Vanitha, U.; Srikanth, N.; Sathyamoorthy, J.; Renu, K. Deep Learning-Based Object Detection in Underwater Communications System. In Proceedings of the 2nd International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichirappalli, India, 5–7 April 2023; pp. 1–6. [CrossRef]

63. Balemans, N.; Hellinckx, P.; Latré, S.; Reiter, P.; Steckel, J. S2L-SLAM: Sensor Fusion Driven SLAM Using Sonar, LiDAR and Deep Neural Networks. In Proceedings of the IEEE Sensors Conference, New York, NY, USA, 31 October–4 November 2021; pp. 1–4.
64. Yu, X.; Sun, Y.; Wang, X.; Zhang, G. End-to-End AUV Motion Planning Method Based on Soft Actor-Critic. *Sensors* **2021**, *21*, 5893. [CrossRef]
65. Sutton, R.S.; Barto, A.G. Reinforcement Learning: An Introduction. *IEEE Trans. Neural Netw.* **1998**, *9*, 1054. [CrossRef]
66. Czum, J.M. Dive into Deep Learning. *J. Am. Coll. Radiol.* **2020**, *17*, 637–638. Available online: https://www.jacr.org/article/S1546-1440(20)30146-0/abstract (accessed on 29 October 2025). [CrossRef] [PubMed]
67. Guerneve, T.; Subr, K.; Petillot, Y. Underwater 3D Structures as Semantic Landmarks in SONAR Mapping. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 614–619.
68. Kiran, B.R.; Sobh, I.; Talpaert, V.; Mannion, P.; Sallab, A.A.A.; Yogamani, S.; Pérez, P. Deep Reinforcement Learning for Autonomous Driving: A Survey. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 4909–4926. [CrossRef]
69. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M.A. Playing Atari with Deep Reinforcement Learning. *arXiv* **2013**, arXiv:1312.5602. [CrossRef]
70. Tabas, S.S.; Samadi, V. Fill-and-Spill: Deep Reinforcement Learning Policy Gradient Methods for Reservoir Operation Decision and Control. *arXiv* **2024**. [CrossRef]
71. Shi, W.; Tang, Y.; Jin, M.; Jing, L. An AUV-Assisted Data Gathering Scheme Based on Deep Reinforcement Learning for IoUT. *J. Mar. Sci. Eng.* **2023**, *11*, 2279. [CrossRef]
72. Zhang, Q.; Lin, J.; Sha, Q.; He, B.; Li, G. Deep Interactive Reinforcement Learning for Path Following of Autonomous Underwater Vehicle. *IEEE Access* **2020**, *8*, 24258–24268. [CrossRef]
73. Carroll, K.P.; McClaran, S.R.; Nelson, E.L.; Barnett, D.M.; Friesen, D.K.; William, G.N. AUV Path Planning: An A* Approach to Path Planning with Consideration of Variable Vehicle Speeds and Multiple, Overlapping, Time-Dependent Exclusion Zones. In Proceedings of the Symposium on Autonomous Underwater Vehicle Technology, Washington, DC, USA, 2 June 1992; pp. 79–84.
74. Li, J.; Zhang, Z. AUV Local Path Planning Based on Fusion of Improved DWA and RRT Algorithms. In Proceedings of the IEEE International Conference on Mechatronics and Automation (ICMA), Harbin, China, 6–9 August 2023; pp. 935–941.
75. Cao, X.; Chen, L.; Guo, L.; Han, W. AUV Global Security Path Planning Based on a Potential Field Bio-Inspired Neural Network in Underwater Environment. *Intell. Autom. Soft Comput.* **2021**, *27*, 1002. [CrossRef]
76. Liu, T.; Hu, Y.; Xu, H. Deep Reinforcement Learning for Vectored Thruster Autonomous Underwater Vehicle Control. *Complexity* **2021**, *2021*, 6649625. [CrossRef]
77. Andriotis, C.P.; Papakonstantinou, K.G. Managing Engineering Systems with Large State and Action Spaces through Deep Reinforcement Learning. *Reliab. Eng. Syst. Saf.* **2019**, *191*, 106483. [CrossRef]
78. Lillicrap, T.P.; Hunt, J.J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous Control with Deep Reinforcement Learning. *arXiv* **2019**, arXiv:1509.02971.
79. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal Policy Optimization Algorithms. *arXiv* **2017**, arXiv:1707.06347. [CrossRef]
80. Yan, J.; Zhang, L.; Yang, X.; Chen, C.; Guan, X. Communication-Aware Motion Planning of AUV in Obstacle-Dense Environment: A Binocular Vision-Based Deep Learning Method. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 14927–14943. [CrossRef]
81. Rahmati, M.; Nadeem, M.; Sadhu, V.; Pompili, D. UW-MARL: Multi-Agent Reinforcement Learning for Underwater Adaptive Sampling Using Autonomous Vehicles. In Proceedings of the International Conference on Underwater Networks & Systems, Atlanta, GA, USA, 23–25 October 2019; pp. 1–5. [CrossRef]
82. Yang, R.; Zhang, F.; Hou, M. OceanPlan: Hierarchical Planning and Replanning for Natural Language AUV Piloting in Large-Scale Unexplored Ocean Environments. In Proceedings of the 18th International Conference on Underwater Networks & Systems, Sibenik, Croatia, 27 January 2025; Association for Computing Machinery: New York, NY, USA, 2025; pp. 1–5.
83. Kim, T.-Y.; Choi, W.-S. Autonomous Vehicle Maneuvering Using Vision–LLM Models for Marine Surface Vehicles. *J. Mar. Sci. Eng.* **2025**, *13*, 1553. [CrossRef]
84. Ding, Y.; Xu, J.; Xie, G.; Zhang, S.; Li, Y. Make Your AUV Adaptive: An Environment-Aware Reinforcement Learning Framework for Underwater Tasks. *arXiv* **2025**, arXiv:2506.15082. [CrossRef]
85. Goldberg, D. Huxley: A Flexible Robot Control Architecture for Autonomous Underwater Vehicles. In Proceedings of the IEEE OCEANS Conference, Santander, Spain, 6–9 June 2011; pp. 1–10.
86. Wu, Z.; Modi, A.; Mavrogiannis, A.; Joshi, K.; Chopra, N.; Aloimonos, Y.; Karapetyan, N.; Rekleitis, I.; Lin, X. DREAM: Domain-Aware Reasoning for Efficient Autonomous Underwater Monitoring. *arXiv* **2025**, arXiv:2509.13666. [CrossRef]
87. Chen, R.; Blow, D.; Abdullah, A.; Islam, M.J. Word2Wave: Language Driven Mission Programming for Efficient Subsea Deployments of Marine Robots. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Atlanta, GA, USA, 19–23 May 2025; pp. 4107–4114.

88. Li, Z.; Du, J.; Jiang, C.; Mi, W.; Ren, Y. HA-MARL: Heuristic and APF Assisted Multi-Agent Reinforcement Learning for Wireless Data Sharing in AUV Swarms. In Proceedings of the ICC 2024—IEEE International Conference on Communications, Denver, CO, USA, 9–13 June 2024; pp. 5401–5406. [CrossRef]
89. Tong, R.; Feng, Y.; Wang, J.; Wu, Z.; Tan, M.; Yu, J. A Survey on Reinforcement Learning Methods in Bionic Underwater Robots. *Biomimetics* **2023**, *8*, 168. [CrossRef]
90. Utkin, V. Sliding Mode Control. *Control Syst. Robot. Autom.* **1999**, *12*, 1–319.
91. Lotfi, F.; Virji, K.; Dudek, N.; Dudek, G. A Comparison of RL-Based and PID Controllers for 6-DOF Swimming Robots: Hybrid Underwater Object Tracking. *arXiv* **2024**, arXiv:2401.16618. [CrossRef]
92. Alessandro, R.; Roberto, C.; Costanzi, R.; Francesco, F.; Enrico, M. A Dynamic Manipulation Strategy for an Intervention: Autonomous Underwater Vehicle. *Adv. Robot. Autom.* **2015**, *4*, 1–16.
93. Muñoz, F.; Cervantes-Rojas, J.S.; Valdovinos, J.M.; Sandre-Hernández, O.; Salazar, S.; Romero, H. Dynamic Neural Network-Based Adaptive Tracking Control for an Autonomous Underwater Vehicle Subject to Modeling and Parametric Uncertainties. *Appl. Sci.* **2021**, *11*, 2797. [CrossRef]
94. Cimurs, R.; Suh, I.H.; Lee, J.H. Goal-Driven Autonomous Exploration Through Deep Reinforcement Learning. *IEEE Robot. Autom. Lett.* **2022**, *7*, 730–737. [CrossRef]
95. Cortez, W.S.; Vasisht, S.; Tuor, A.; Drgoňa, J.; Vrabie, D. Domain-Aware Control-Oriented Neural Models for Autonomous Underwater Vehicles. *arXiv* **2022**, arXiv:2208.07333. [CrossRef]
96. Carlucho, I.; De Paula, M.; Barbalata, C.; Acosta, G.G. A Reinforcement Learning Control Approach for Underwater Manipulation under Position and Torque Constraints. In Proceedings of the Global Oceans 2020: Singapore–U.S. Gulf Coast, Biloxi, MS, USA, 5–30 October 2020; pp. 1–7.
97. Xie, G.; Xu, J.; Ding, Y.; Zhang, Z.; Zhang, S.; Li, Y. Never Too Prim to Swim: An LLM-Enhanced RL-Based Adaptive S-Surface Controller for AUV under Extreme Sea Conditions. *arXiv* **2025**, arXiv:2503.00527.
98. Battaglia, P.; Hamrick, J.B.; Bapst, V.; Sanchez-Gonzalez, A.; Zambaldi, V.; Malinowski, M.; Tacchetti, A.; Raposo, D.; Santoro, A.; Faulkner, R.; et al. Relational Inductive Biases, Deep Learning, and Graph Networks. *arXiv* **2018**, arXiv:1806.01261. [CrossRef]
99. Haarnoja, T.; Zhou, A.; Hartikainen, K.; Tucker, G.; Ha, S.; Tan, J.; Kumar, V.; Zhu, H.; Gupta, A.; Abbeel, P.; et al. Soft Actor-Critic Algorithms and Applications. *arXiv* **2019**, arXiv:1812.05905. [CrossRef]
100. Lyu, X.; Sun, Y.; Wang, L.; Tan, J.; Zhang, L. End-to-End AUV Local Motion Planning Method Based on Deep Reinforcement Learning. *J. Mar. Sci. Eng.* **2023**, *11*, 1796. [CrossRef]
101. Carlucho, I.; De Paula, M.; Wang, S.; Menna, B.V.; Petillot, Y.R.; Acosta, G.G. AUV Position Tracking Control Using End-to-End Deep Reinforcement Learning. In Proceedings of the OCEANS 2018 MTS/IEEE Charleston, Charleston, SC, USA, 22–25 October 2018; pp. 1–8.
102. Musa, P.; Rafi, F.A.; Lamsani, M. A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) Methods to Help the Application of Face Recognition. In Proceedings of the 3rd International Conference on Informatics and Computing (ICIC), Palembang, Indonesia, 17–18 October 2018; pp. 1–6.
103. Tarling, P.; Cantor, M.; Clapés, A.; Escalera, S. Deep Learning with Self-Supervision and Uncertainty Regularization to Count Fish in Underwater Images. *PLoS ONE* **2022**, *17*, e0267759. [CrossRef]
104. Li, X.; Li, X.; Han, B.; Wang, S.; Chen, K. Application of EfficientNet and YOLOv5 Model in Submarine Pipeline Inspection and a New Decision-Making System. *Water* **2023**, *15*, 3386. [CrossRef]
105. Karimanzira, D.; Renkewitz, H.; Shea, D.; Albiez, J. Object Detection in Sonar Images. *Electronics* **2020**, *9*, 1180. [CrossRef]
106. Ma, Z.; Li, S.; Ding, J.; Zou, B. MHGAN: A Multi-Headed Generative Adversarial Network for Underwater Sonar Image Super-Resolution. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–16. [CrossRef]
107. Sudevan, V.; Zayer, F.; Kausar, R.; Javed, S.; Karki, H.; De Masi, G.; Dias, J. snnTrans-DHZ: A Lightweight Spiking Neural Network Architecture for Underwater Image Dehazing. *arXiv* **2025**. [CrossRef]
108. Yu, H.; Li, X.; Feng, Y.; Han, S. Underwater Vision Enhancement Based on GAN with Dehazing Evaluation. *Appl. Intell.* **2022**, *53*, 5664–5680. [CrossRef]
109. Zhang, C.; Liu, L.; Huang, G.; Wen, H.; Zhou, X.; Wang, Y. WebUOT-1M: Advancing Deep Underwater Object Tracking with A Million-Scale Benchmark. *arXiv* **2024**. [CrossRef]
110. Wang, M.; Qiu, B.; Zhu, Z.; Ma, L.; Zhou, C. Passive Tracking of Underwater Acoustic Targets Based on Multi-Beam LOFAR and Deep Learning. *PLoS ONE* **2022**, *17*, e0273898. [CrossRef]
111. Liu, F.; Fang, M. Semantic Segmentation of Underwater Images Based on Improved Deeplab. *J. Mar. Sci. Eng.* **2020**, *8*, 188. [CrossRef]
112. Chen, B.; Zhao, W.; Zhang, Q.; Li, M.; Qi, M.; Tang, Y. Semantic Segmentation of Underwater Images Based on the Improved SegFormer. *Front. Mar. Sci.* **2025**, *12*, 1522160. [CrossRef]
113. Gao, W.; Han, M.; Wang, Z.; Deng, L.; Wang, H.; Ren, J. Research on Method of Collision Avoidance Planning for UUV Based on Deep Reinforcement Learning. *J. Mar. Sci. Eng.* **2023**, *11*, 2245. [CrossRef]

114. Vu, Q.V.; Dinh, T.A.; Nguyen, T.V.; Tran, H.V.; Le, H.X.; Pham, H.V.; Kim, T.D.; Nguyen, L. An Adaptive Hierarchical Sliding Mode Controller for Autonomous Underwater Vehicles. *Electronics* **2021**, *10*, 2316. [CrossRef]
115. Peng, G.; Chen, C.L.P.; Yang, C. Neural Networks Enhanced Optimal Admittance Control of Robot–Environment Interaction Using Reinforcement Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 4551–4561. [CrossRef] [PubMed]
116. Kebaili, A.; Lapuyade-Lahorgue, J.; Vera, P.; Ruan, S. AMM-Diff: Adaptive Multi-Modality Diffusion Network for Missing Modality Imputation. In Proceedings of the IEEE 22nd International Symposium on Biomedical Imaging (ISBI), Houston, TX, USA, 14–17 April 2025; pp. 1–4.
117. Ochal, M.; Vazquez, J.; Petillot, Y.; Wang, S. A Comparison of Few-Shot Learning Methods for Underwater Optical and Sonar Image Classification. In Proceedings of the Global Oceans 2020: Singapore–U.S. Gulf Coast, Biloxi, MS, USA, 5–30 October 2020; pp. 1–10. [CrossRef]
118. Wang, J.; Li, Q.; Fang, Z.; Zhou, X.; Tang, Z.; Han, Y.; Ma, Z. YOLOv6-ESG: A Lightweight Seafood Detection Method. *J. Mar. Sci. Eng.* **2023**, *11*, 1623. [CrossRef]
119. Tobin, J.; Fong, R.; Ray, A.; Schneider, J.; Zaremba, W.; Abbeel, P. Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 23–30. [CrossRef]
120. Li, B.; Qi, P.; Liu, B.; Di, S.; Liu, J.; Pei, J.; Yi, J.; Zhou, B. Trustworthy AI: From Principles to Practices. *ACM Comput. Surv.* **2023**, *55*, 1–46. [CrossRef]
121. Chaffre, T.; Wheare, J.; Lammas, A.; Santos, P.; Le Chenadec, G.; Sammut, K.; Clement, B. Sim-to-Real Transfer of Adaptive Control Parameters for AUV Stabilisation under Current Disturbance. *Int. J. Robot. Res.* **2025**, *44*, 407–430. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.