Journal of *Marine Science and Engineering*

MDPI

Article

Multi Autonomous Underwater Vehicle (AUV) Distributed Collaborative Search Method Based on a Fuzzy Clustering Map and Policy Iteration

Kaiqian Cai [1,2,3], Guocheng Zhang [1,3,*], Yushan Sun [1,3], Guoli Ding [1,3] and Fengchi Xu [4]

[1] College of Shipbuilding Engineering, Harbin Engineering University, No. 145, Nantong Street, Nangang District, Harbin 150001, China; caikaiqian@hrbeu.edu.cn (K.C.); sunyushan@hrbeu.edu.cn (Y.S.)
[2] Nanhai Institute of Harbin Engineering University, Yazhou Bay Science and Technology City Yazhou District, Sanya 572025, China
[3] National Key Laboratory of Autonomous Marine Vehicle Technology, No. 145, Nantong Street, Nangang District, Harbin 150001, China
[4] CSSC Systems Engineering Research Institute, No.1 Fengxian East Road, Yongfeng Science and Technology Industry Base, Haidian District, Beijing 100094, China; 18110076047@163.com
* Correspondence: zhangguocheng168@126.com

Abstract: Collaborative search with multiple Autonomous Underwater Vehicles (AUVs) significantly enhances search efficiency compared to the use of a single AUV, facilitating the rapid completion of extensive search tasks. However, challenges arise in underwater environments characterized by weak communication and dynamic complexities. In large marine areas, the limited endurance of a single AUV makes it impossible to cover the entire area, necessitating a collaborative approach using multiple AUVs. Addressing the limited prior information available in uncertain marine environments, this paper proposes a map-construction method using fuzzy clustering based on regional importance. Furthermore, a collaborative search method for large marine areas has been designed using a policy-iteration-based reinforcement learning algorithm. Through continuous sensing and interaction during the marine search process, multiple AUVs constantly update the map of regional importance and iteratively optimize the collaborative search strategy to achieve higher search gains. Simulation results confirm the effective utilization of limited information in uncertain environments and demonstrate enhanced search gains in collaborative scenarios.

Keywords: multiple AUVs; collaborative search; fuzzy clustering; reinforcement learning

Citation: Cai, K.; Zhang, G.; Sun, Y.; Ding, G.; Xu, F. Multi Autonomous Underwater Vehicle (AUV) Distributed Collaborative Search Method Based on a Fuzzy Clustering Map and Policy Iteration. *J. Mar. Sci. Eng.* **2024**, *12*, 1521. https://doi.org/10.3390/jmse12091521

Academic Editor: Marco Cococcioni

Received: 22 July 2024
Revised: 18 August 2024
Accepted: 29 August 2024
Published: 2 September 2024

1. Introduction

The ocean, which occupies more than 70% of the Earth's surface, is not only a repository of biodiversity but also exerts a profound influence on the global climate, food chain, and human life. Therefore, comprehensive scientific research and the study of the oceans is of paramount importance. In recent years, Autonomous Underwater Vehicles (AUVs) have emerged as a crucial instrument in oceanographic research. In comparison to traditional manned submersibles, AUVs offer several advantages, including low cost and flexible operation. They are capable of carrying a variety of sensing devices, such as sonar, cameras, and chemical analyzers, which enables them to adapt to a diverse range of oceanographic research tasks, including seafloor topographic mapping, biodiversity observation, and pollutant monitoring. In comparison to single AUVs, AUV clusters are capable of covering a larger area of the mission zone in the form of formations. This enables the completion of a wide range of patrol and surveillance tasks in a more expeditious manner, while also markedly enhancing the detection capabilities of underwater vehicles. This latter point represents a significant area of current research interest.

Collaboration between AUVs holds significant potential value in a variety of marine tasks, including large-scale maritime search, environmental monitoring, resource explo-

ration, and military missions. This potential value encompasses improvements in task efficiency, the optimization of resource management, enhanced adaptability to complex environments, and system scalability. By carefully designing and optimizing collaboration strategies, the overall effectiveness of multi-AUV systems can be greatly enhanced, particularly in complex marine environments where uncertainty and time sensitivity are critical factors. Among these, the main focus of multi-AUVs cooperative search research is on how multi-AUVs in known or unknown environments can effectively acquire target information in formation or other collaborative formats, so both task allocation and path planning need to be considered in the multi-AUV search problem [1].

In terms of task allocation, a novel consensus-based adaptive optimization auction (CAOA) algorithm is proposed for the task allocation of multiple AUVs [2,3], which greatly reduces the computational load while improving the system revenue, but its performance degrades significantly in communication-constrained environments. Task allocation methods based on clustering algorithms include Euclidean distance clustering, K-means clustering, and fuzzy clustering [4,5], which demonstrates a degree of dominance on robustness and computational efficiency. However, the key challenge of clustering-based task allocation is to identify the optimal number of tasks for each cluster, especially in an environment of uncertainty. Intelligent optimization-based task allocation, on the other hand, makes use of optimization techniques such as meta-heuristic algorithms to deal with the task allocation problem and is particularly suitable for situations where the number of tasks is large and the environment is complex [6]. Common intelligent swarm optimization methods include genetic algorithm (GA) [7,8], particle swarm optimization (PSO) [9–11], and self-organizing map (SOM). Among them, a decentralized task allocation method based on a decentralized genetic algorithm was proposed [8] for multi-agent cooperative search. The advantages of the PSO algorithm are simple implementation, few adjustment parameters, and fast convergence. The modified PSO algorithm for the optimization of rescue task allocation with uncertain time constraints was proposed [9], which focuses on the problem of robot rescue task allocation. Sun et al. proposed a novel game allocation method for solving the task allocation problem in underwater multi-AUV dynamic confrontations [12]. The proposed method addresses incomplete information using a multi-objective evaluation model and interval ranking, improving the convergence speed, accuracy, and global optimization ability by increasing population diversity and iterative effectiveness, ensuring real-time decision-making in complex game scenarios. Zhu et al. combined the improved SOM neural network with the Glasius Bio-inspired Neural Network (GBNN) approach to propose an integrated algorithm for multi-AUV dynamic task assignment and path planning that can solve the multi-AUV task assignment problem with good performance, where each AUV can be assigned a specific task in a time-varying ocean environment [13]. Ma et al. proposed a collaborative hunting algorithm based on the Bionic Neural Network (BNWN) algorithm [14], improving the global search efficiency under incomplete navigation map conditions and embedding a real-time reassignment method.

In terms of path planning, the path planning algorithm based on the geometric model search belongs to the category of discrete optimal planning. It is a traditional path-planning algorithm with a simple implementation process and mature technology. However, a high level of model building is required by the algorithm, as it is closely related to the final path planning results. Rajnarayan D G et al. proposed a multiple UUV search strategy based on a segmentation strategy [15], where the search area is segmented into different regions of uniform size and assigned to individual AUVs to detect the segmented region based on a boundary algorithm to improve the search efficiency. Furthermore, Zhang et al. proposed a 3D path planning method [16]. Cai et al. proposed an area partitioning method to allocate the task to multiple AUVs and maintain the possible target area as a whole, which can generate stable solutions to reduce the segmentation of target areas [17]. The scholars above have primarily investigated the potential of partitioned search to reduce the time required for search operations. Their findings offer insights and methodologies that can be applied to the research topics addressed in this paper. However, the collaborative

search problem they have studied primarily focuses on the rapid and efficient acquisition of target information. In contrast, the large sea area search problem addressed in this paper also requires consideration of the ability to cover the target area multiple times. The path planning issues in complex dynamic environments can be solved by intelligent algorithms and are suitable for path planning for AUVs. The main path planning strategies are based on intelligent algorithms such as particle swarm optimization (PSO), ant colony optimization (ACO), genetic algorithms (GAs), differential evolution (DE), and artificial neural networks (ANNs). Zhang et al. designed a novel prediction-based path evaluator to evaluate the fitness of possible paths and conducted multiple simulation experiments to verify the effectiveness and superiority of the path planner [18]. However, the dynamic obstacles were not considered.

In addition, reinforcement learning (RL) is known for its efficiency and practicality in single-agent planning, but it faces numerous challenges when applied to multi-agent scenarios [19]. In multi-agent reinforcement learning (MRL), a multilayer fully connected neural network is used for value function approximation, which solves large-scale or continuous space problems. However, in the complex dynamic environment affected by wind and wave currents and seabed topography, it is easy to fall into local optima and overfitting under partially observed environments because each agent lacks the information that plays a key role in decision-making beyond the observation field. In particular, the weak communication characteristics of the marine environment, where both perception and communication between AUVs are highly noisy, make it highly uncertain if the strategies of each AUV are used as communication messages. This poses a great challenge for AUV task execution, and traditional reinforcement learning methods, such as Markov Decision Process (MDP)-based methods, although successful in many applications, still face challenges in dealing with uncertainty and ambiguity.

Whereas fuzzy reinforcement learning as an intelligent control strategy combines the principles of fuzzy logic and reinforcement learning, fuzzy logic provides a way to deal with imprecise and uncertain information, while reinforcement learning focuses on how to optimize decision-making strategies based on interaction with the environment. By combining the two, fuzzy reinforcement learning is able to learn how to make optimal or near-optimal decisions in complex environments without full knowledge of the environment model [20]. This has better results in dealing with decision-making processes with high uncertainty and complexity and also provides new ideas for multi-AUV search tasks. Fathinezhad et al. proposed a supervised fuzzy Sarsa learning algorithm to find the optimal action for each fuzzy rule through the supervised learning of fuzzy rules based on the Sarsa algorithm, which was was validated by the E-puck robot's movement in the E-puck robot in an obstacle environment to verify that the method has the advantages of short learning time and fewer failures [21]. A method of RL with an interpretable fuzzy system (IFS) based on a neural fuzzy actor–critic (RLIFS-NFAC) framework was proposed, and the learned IFS has also been successfully applied to control a real wall-following robot in unknown environments [22]. Fu et al. focused on the policy-learning process in scenarios with dynamic competitors that evolve dynamically with MARL and a competitive automultiagent learner with fuzzy feedback (CALF), showing that CALF significantly promotes team competitiveness in adversarial competitions [23]. The two-stream fused fuzzy deep neural network (2s-FDNN) was proposed to reduce the uncertainty and noise of information in the communication channel and to improve the robustness and generalization under partially observed environments [24].

However, the current research on fuzzy reinforcement learning in the field of AUV is still relatively limited, especially in the scenario of multi-AUV collaborative search; understanding how to train effectively with the help of fuzzy reinforcement learning under the limited a priori information in order to improve the search efficiency and accuracy is an urgent problem to be solved. Therefore, this paper aims to propose a multi-AUV cooperative search method for dynamic and complex environments in the ocean based on the fuzzy reinforcement learning method, constructing a fuzzy regional importance model

based on grey system theory, clustering to generate an importance map of the mission sea area based on limited a priori information, and then updating the map as a guide to quickly determine the AUV cooperative search strategy within the map updating cycle through the strategy-iteration algorithm of reinforcement learning in order to adapt to the dynamic environment of the ocean and improve the search benefit under the requirements of low communication.

2. Problem Formulation

Compared to a single AUV, a collaborative search with multiple AUVs can significantly expand the search area and greatly enhance search efficiency. However, it also faces several challenges. Taking the typical task scenario of patrolling a large marine area with multiple AUVs to detect potential targets as an example, the main difficulties and limitations include the following.

(1) Communication Constraints: Due to the unique properties of the marine environment as a communication medium, long-distance underwater communication relies on acoustic signals. However, acoustic communication has limitations such as low bandwidth, uncertain latency, and errors in transmission, leading to incomplete or inconsistent information among members of an AUV fleet.
(2) Perception Constraints: Limited by sensor accuracy and the interference and dynamic nature of the underwater environment, the positioning and navigation of AUVs under water are not always accurate, and misjudgments regarding targets and the environment are possible.
(3) Energy Consumption Constraints: Although expanding the AUV fleet to perform periodic patrols of large marine areas is feasible, the energy resources of individual AUVs within the cluster are limited. Therefore, each AUV can only patrol a region near its starting point.
(4) Individual Loss: The larger the AUV cluster, the more likely it is to encounter losses of individual AUVs. The fleet must adjust the task sequence and re-plan movements in real time according to environmental changes to respond quickly to external variations, enhancing task efficiency and flexibility.

Therefore, collaborative search in large marine areas by multiple AUVs is a complex issue with many research directions worth exploring. Considering all influencing factors comprehensively is challenging. To focus the research while maintaining scientific rigor, this paper selects the P324 underwater vehicle, designed by T-SEA Marine Technology Co., Ltd. (Zhang Jiagang, China), as the research subject. Public images of this vehicle are shown Figure 1. In the large marine area patrol scenario addressed in this paper, a mother ship deploys multiple AUVs into the marine area. Each AUV patrols the clearly bounded large marine area based on prior knowledge and environmental perception. In this scenario, the AUVs operate using a distributed structure, where each AUV interacts only with its neighboring AUVs. The confidence level of information exchange ranges from 0 to 1; within the neighboring range, the confidence level is 1, and outside the neighboring range, it is 0. The size of the neighboring range is determined by the parameters of the communication equipment installed on the AUVs. This paper analyzes scenarios where there are no AUVs nearby, one neighboring AUV, and two neighboring AUVs.

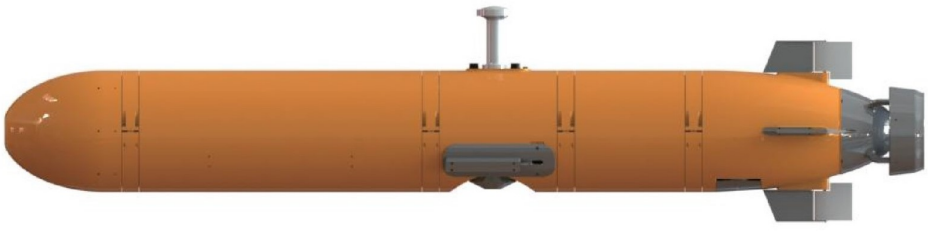


Figure 1. P324.

Given that the research focus of this paper is on designing effective strategies for multi-AUV search in large marine areas, the multi-AUV collaborative search problem discussed herein is simplified with the following assumptions:

(1) Hydroacoustic Communications: Although short-term instability may occur in acoustic communication, it is assumed that within a patrol cycle, tasks, environmental data, and state information updates can be progressively completed through interactions between neighboring AUVs. That is, information may not be synchronized between AUVs that are far apart, but this is consistent among neighboring AUVs. AUVs within the effective range of acoustic communication are defined as neighboring AUVs, capable of direct information exchange without the need for relay through other AUVs.
(2) Detection Range: It is assumed that each AUV's detection range is confined to a single grid width and can fully cover the sea area represented by that grid with satisfactory accuracy for the mission requirements. Detection outside this range is not considered, with confidence in the detection data ranging from 0 to 1, where 1 is within the range and 0 is outside.
(3) AUV Deployment and Energy: AUVs are deployed by a surface mother ship, and each AUV's energy is quantified to support movement across 1000 grid spaces.
(4) Loss of Units: It is assumed that any lost units can be promptly replenished by the mother ship, so the progress of the search mission is not affected by individual losses, and any impacts can be mitigated through periodic updates of the map and adjustments to the task sequence.

In the search task scenario within a large marine area, AUV clusters are distributed in a designated area and continuously and repeatedly patrol and monitor the area in a self-organized manner to gradually grasp the information of the target area. In this process, repeated searches by multiple AUVs for the sea area will lead to a reduction in search yield, which is the concern of this paper. This paper adopts a distributed decision-making approach, allowing each AUV to evaluate the importance of marine areas based on existing map information. Based on the assessed importance, optimal search/patrol strategies are sought using the policy iteration algorithm. Subsequent motion planning for the efficient patrolling of the marine area is then carried out according to the search strategies. The grid map's area importance model discussed in this paper is constructed based on prior information and grey system theory. The importance evaluation provides a fuzzy assessment of the marine environment under uncertain conditions. Although there are deviations and disturbances compared to the actual marine characteristics, the model can be refined by updating the map through perception information and interaction processes during the multi-AUV patrol. This process continues as the patrol progresses, allowing the model to gradually converge to the actual characteristics of the marine area. In the distributed collaborative search by multiple AUVs, the discovery of a target in a particular area increases the importance of that area and its surroundings. Conversely, repeated patrols of the same area without finding a target reduce its importance due to a deeper understanding of the area. Employing a distributed structure, the AUV cluster lacks a centralized information aggregation and distribution center. Thus, information among multiple AUVs is exchanged via acoustic communication among nearby AUVs. As the search mission progresses, interactions form among all AUVs involved in the collaborative search, ultimately updating the grid map for the AUV cluster to obtain more definitive global information. However, the progress of map updates within the AUV cluster may vary, and significant map discrepancies might exist between two distant AUVs. Since these AUVs are sufficiently far apart, task conflicts are nearly nonexistent, and their search missions minimally affect each other. Ultimately, as each member of the AUV cluster updates the map, the cluster's understanding of the grid map and area importance in the patrol marine area will tend to converge. This continues until other disturbances occur, such as the discovery of new targets or sudden changes in the marine environment, prompting the cluster to repeat the process of information perception and interaction until a new consensus is reached. The architecture designed in this paper is shown in Figure 2.

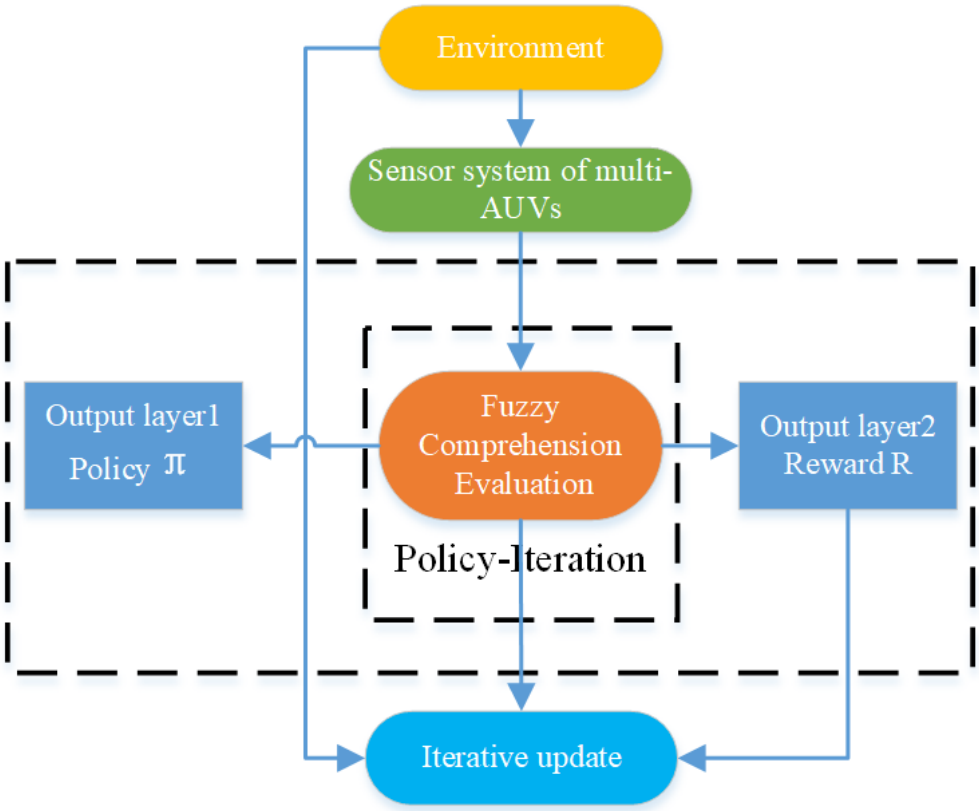


Figure 2. Fuzzy decision-making architecture based on multi-AUVs perception and interaction.

3. Methods

The large-scale maritime cooperative search task scenario discussed in this paper involves a situation where the boundaries of the maritime area are known, but the specific locations or trajectories of the targets are uncertain. This represents a typical case of an uncertain system characterized by incomplete information and inaccurate data. For the study of uncertain systems, four general theories exist: stochastic system theory, fuzzy system theory, grey system theory, and the theory of unascertained systems. Among these, the grey system theory, introduced in 1982 by the Chinese scholar Professor Deng Julong, offers a novel approach to addressing uncertainty in situations with limited data and insufficient information. This theory focuses on systems where "some information is known, while some information is unknown", making it particularly suitable for evaluating regional importance and clustering in the search process explored in this paper [25].

Regional Importance in Collaborative Search primarily indicates the probability of discovering targets in the patrolled area. The likelihood of target appearances or discoveries in different marine areas under uncertain conditions relates to the location distribution of the area, familiarity with the area (e.g., number of detection attempts and patrol frequencies), and other prior knowledge (such as historical tracks of targets). Considering that during the search process, the target primarily adopts stealthy and covert maneuvers to gather more maritime information, frequent depth changes are only necessary when evasive maneuvers are required. Additionally, the multi-AUV cooperative search mission involves continuous and repetitive coverage of the maritime area, making it crucial to avoid excessive energy consumption due to frequent maneuvers. Therefore, greater emphasis is placed on endurance, indicating that depth-fixed searches are advantageous for the execution of the mission. Thus, simplifying the task planning for this scenario into a two-dimensional problem is both reasonable and effective. Consequently, it is stipulated that AUVs primarily move within the same plane, requiring no frequent movement out of this plane, and the initial map depth is set to zero and updated only when a change in AUV depth is needed.

For two-dimensional problems, employing the grid method for optimizing multi-AUV search strategies is simple and effective and offers good real-time performance. In a grid map, multi-AUVs do not require frequent communication interactions during the search of large maritime areas. It is only necessary for neighboring AUVs to interact before deciding to enter the next grid after searching the current one in order to avoid task conflicts in the same grid. To support the resolution of search strategies based on the grid method, grid clustering is required. Considering the varying importance of different grids in a patrol scenario, this paper designs an importance evaluation and clustering method based on grey system theory, enabling the reasonable division of the patrol area by clustering the grids.

3.1. Construction of a Set of Indicators for the Evaluation of Regional Importance

In the regional importance model presented in this paper, the prior knowledge of the collaborative search task, such as historical tracks of targets and no-sail zones within the marine area, is incorporated into a grid map for visual analysis as designated danger and focus areas. Due to the fuzziness of sonar detection (where detected targets might be false images), the number of times a target is detected is initially set to zero, reflecting the lack of precise prior information. This necessitates the improvement of the model through posterior knowledge acquired during continuous patrols. Consequently, various types of distances have been selected as indicators for evaluating regional importance. The importance of each grid σ is controlled by three indicators—distance from the port, marine environment, and human attention level—with membership functions designed based on expert knowledge.

Distance from Port S_O: represented by the distance between the grid position and the center of the port.

Marine Environment S_H: characterized by the distance between the grid position and challenging navigational areas such as reefs and shallow waters.

Human Attention Level S_K: indicated by the distance between the grid position and areas of significant human interest.

As shown in Equation (1) through (4), x, y represent the horizontal and vertical coordinates, respectively. H (boundary) and K (boundary) denote the boundaries of challenging navigational areas and areas of significant human interest, respectively.

$$\sigma = f(S_O, S_H, S_K) \tag{1}$$

$$S_O = \| x - y \| \tag{2}$$

$$S_H = \min\{\| (x, y) - \mathrm{H(boundary)} \|\} \tag{3}$$

$$S_K = \min\{\| (x, y) - \mathrm{K(boundary)} \|\} \tag{4}$$

3.2. Regional Importance Assessment and Clustering Method Based on Grey System Theory

3.2.1. Construction of the Whitening Weight Function

The whitening weight function of a grey number reflects the degree of information subjectively held about that grey number. This function is used to describe a grey number's "preference" for different values within its range. Generally, a whitening weight function is designed by researchers based on available information, and there is no fixed formula for its construction. However, the starting and ending points of the curve should have meaningful interpretations.

Common whitening weight functions are shown in Figure 3. The functional forms are as follows:

(a) Typical Whitening Weight Function
The general form of the function's inflection point and piecewise function expression are shown in Equations (5) and (6).

(b) Lower Bound Measure Whitening Weight Function

The general form of the function's inflection point and piecewise function expression are shown in Equations (7) and (8).

(c) Moderate Measure Whitening Weight Function
The general form of the function's inflection point and piecewise function expression are shown in Equations (9) and (10).

(d) Upper Bound Measure Whitening Weight Function
The general form of the function's inflection point and piecewise function expression are shown in Equations (11) and (12).

$$f_j^k[x_j^k(1), x_j^k(2), x_j^k(3), x_j^k(4)] \tag{5}$$

$$f_j^k(x) = \begin{cases} 0, x \notin [x_j^k(1), x_j^k(4)] \\ \frac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, x \in [x_j^k(1), x_j^k(2)] \\ 1, x \in [x_j^k(2), x_j^k(3)] \\ \frac{x_j^k(4) - x}{x_j^k(4) - x_j^k(3)}, x \in [x_j^k(3), x_j^k(4)] \end{cases} \tag{6}$$

$$f_j^k[-, -, x_j^k(3), x_j^k(4)] \tag{7}$$

$$f_j^k(x) = \begin{cases} 0, x \notin [0, x_j^k(4)] \\ 1, x \in [0, x_j^k(3)] \\ \frac{x_j^k(4) - x}{x_j^k(4) - x_j^k(3)}, x \in [x_j^k(3), x_j^k(4)] \end{cases} \tag{8}$$

$$f_j^k[x_j^k(1), x_j^k(2), -, x_j^k(4)] \tag{9}$$

$$f_j^k(x) = \begin{cases} 0, x \notin [x_j^k(1), x_j^k(4)] \\ \frac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, x \in [x_j^k(1), x_j^k(2)] \\ \frac{x_j^k(4) - x}{x_j^k(4) - x_j^k(2)}, x \in [x_j^k(2), x_j^k(4)] \end{cases} \tag{10}$$

$$f_j^k[x_j^k(1), x_j^k(2), -, -] \tag{11}$$

$$f_j^k(x) = \begin{cases} 0, x \in [0, x_j^k(1)] \\ \frac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, x \in [x_j^k(1), x_j^k(2)] \\ 1, x \in [x_j^k(2), \infty] \end{cases} \tag{12}$$

The whitening weight functions for the model were designed using the Grey Variable Weight Clustering method. The design references the triangular membership function evaluation model from reference [25] and the whitening weight function construction method described in reference [26]. The specific constructions are shown in Table 1.

Table 1. Index set and importance grading of grey levels.

K = 4	K = 3	K = 2	K = 1	K = 0
[0, 0, 0, 0]	[0, 0, 0, 0]	[30, 40, 85, 95]	[-, -, 30, 40]	[85, 95, -, -]
[8, 10, -,-]	[5, 8, -, 10]	[3, 5, -, 8]	[1, 3, -, 5]	[-, -, 1, 3]
[-, -, 5, 10]	[5, 10, -, 20]	[10, 20, -, 30]	[20, 30, -, 40]	[30, 40, -, -]

In the table, the indicators are briefly denoted as S_O S_H S_K, with the set of importance evaluation levels W = Extremely High (K = 4), High (K = 4), Moderate (K = 2), Low (K = 1),

Extremely Low (K = 0). The numbers in brackets represent the coordinates of the inflection points. A total of 15 whitening weight functions have been constructed to evaluate the importance of each grid in three indicator dimensions. Taking the indicator S_O as an example, the meanings of some inflection points in the whitening weight functions are explained as follows:

- K = 4 (Extremely High Importance) and K = 3 (High Importance) are both arbitrarily set to 0

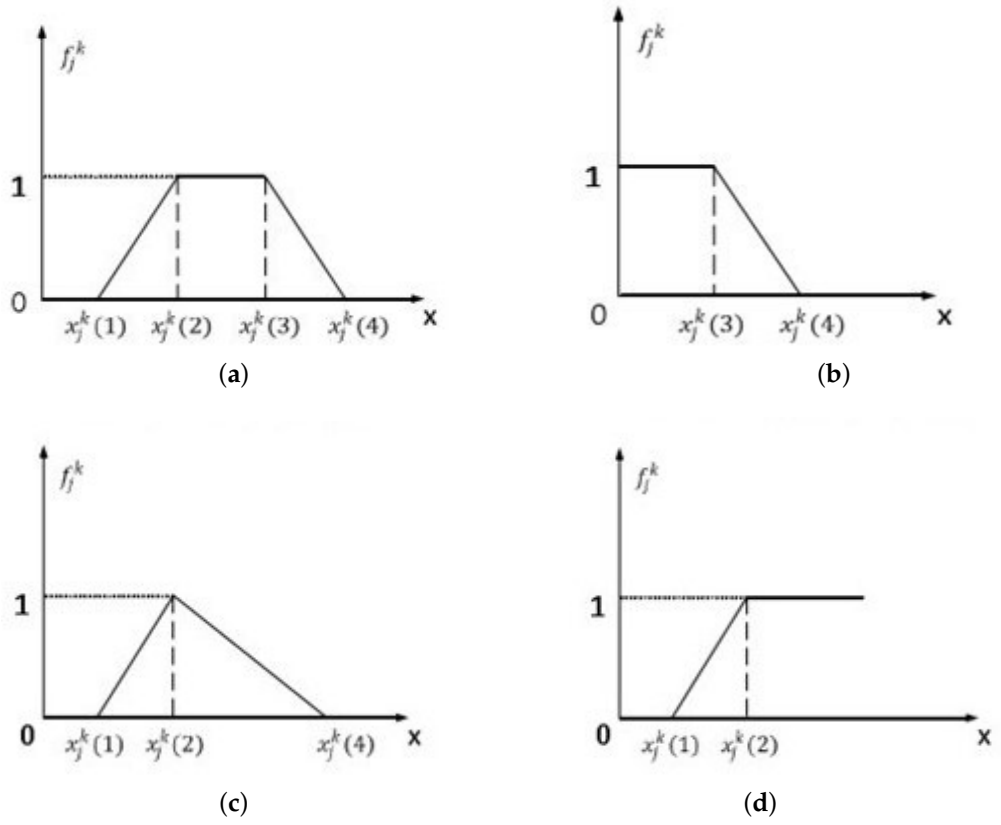


Figure 3. Common whitening weight functions. (**a**) Typical whitening weight function. (**b**) Lower bound measure whitening weight function. (**c**) Moderate measure whitening weight function. (**d**) Upper bound measure whitening weight function

The reason is that it is difficult to judge the importance based solely on the distance from the port, and at most, areas where targets are unlikely or less likely to appear can be estimated (see K = 1 and K = 0). Therefore, the membership values for these levels are arbitrarily set to 0.

- K = 1, the Lower Bound Measure Whitening Weight Function is used

This represents that the distance from the port is less than 30. This value should be classified under low importance (this journey is less than one-third of the maximum range, suggesting it is only likely if staying at the mother port or in an accidental situation). A deviation of 10 units on the right side means that when the distance from the port exceeds 40, it should no longer be classified under low importance.

- K = 0 (Extremely Low Importance), the Upper Bound Measure Whitening Weight Function is used

A value of 85 (Inflection Point 1) represents the maximum possible distance from the port, which is calculated by multiplying the *departure time* by the *maximum speed*. When greater than this value, the membership rapidly increases to 1, indicating a definite extremely low-importance area (for this indicator). With a deviation of 10 units, at 95

(Inflection Point 4) and beyond, the importance score of these grids should be recognized as 1, while within 85 it is 0, i.e., it should not be considered an extremely low-importance area.

- K = 2, the Typical Whitening Weight Function is used

Indicates that from Inflection Point 1 to Inflection Point 2, and from Inflection Point 2 to Inflection Point 4, the membership probability ascends on the left and descends on the right.

By the complement of K = 4 and K = 5 under the universe, the completeness of the importance evaluation under the indicator is ensured. Values of 40 (Inflection Point 2) and 85 (Inflection Point 3) mean that areas between 40 and 85 absolutely belong to moderate importance; i.e., under normal circumstances, most areas considering only the distance from the port, except for those that are exceptionally far (greater than the maximum range) or exceptionally close (less than one-third of the maximum range), are considered of moderate importance, which aligns with conventional understanding in such scenarios. Additionally, a deviation of 10 units is set from Inflection Points 2 and 3 to the left and right as Inflection Points 1 and 4, respectively. When less than 40 and more than 85, the membership rapidly decreases until it is 0 at less than 30 and more than 95; i.e., it is impossible to be considered a moderate-importance area (for this indicator).

3.2.2. Evaluation of Regional Significance

The critical value for subclass k of indicator j is defined as λ_j^k. Therefore, the critical values for the four types of whitening weight functions are shown in Equation (13) through (16). Thus, the weight for subclass k of indicator j is shownin Equation (17).

$$\lambda_j^k = \frac{x_j^k(2) + x_j^k(3)}{2} \tag{13}$$

$$\lambda_j^k = x_j^k(3) \tag{14}$$

$$\lambda_j^k = x_j^k(2) \tag{15}$$

$$\lambda_j^k = x_j^k(2) \tag{16}$$

$$\eta_j^k = \frac{\lambda_j^k}{\sum_{j=1}^{m} \lambda_j^k} \tag{17}$$

Then, we calculate $\lambda_j^k (1 \le j \le 3, 1 \le k \le 5)$ and $\eta_j^k (1 \le j \le 3, 1 \le k \le 5)$.

Finally, we compute the clustering coefficient matrix σ to obtain a clustering map of regional importance. Calculating as shown in Equations (18) and (19), if $\max\limits_{1 \le k \le s} \left\{ \sigma_i^k \right\} = \sigma_j^{k*}$, object i is said to belong to grey class k^*, determining the importance of the grid. By assigning different colors to grids of different importance, the importance clustering of the target marine area is completed.

$$\sigma = \begin{bmatrix} \sigma_1^1 & \sigma_1^2 & \cdots & \sigma_1^s \\ \sigma_2^1 & \sigma_2^2 & \cdots & \sigma_2^s \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_n^1 & \sigma_n^2 & \cdots & \sigma_n^s \end{bmatrix} \tag{18}$$

$$\sigma_j^k = \sum_{j=1}^{m} f_j^k(x_{ij}) \eta_j^k \tag{19}$$

$$\max_{1 \le k \le s} \left\{ \sigma_i^k \right\} = \sigma_j^{k*} \tag{20}$$

In this paper, two scenarios are considered as examples, focusing on either a fixed marine area or a target navigation route. The task area is set as a 100 × 100 grid map, with the wharf coordinates set at [50, 0]. For the scenario with a fixed marine area, the designated danger area d_A is set at [(20, 20), (20, 40), (35, 40), (35, 20)], and the focus area f_A is at [(60, 20), (75, 20), (75, 55), (60, 55)]. For the target navigation route scenario, the danger area d_A is set at [(10, 10), (10, 20), (20, 20), (20, 10)], and the focus area f_A is at [(60, 30), (80, 40), (80, 41), (60, 31)]. The marine area's importance division obtained using this clustering method is shown in Figure 4.

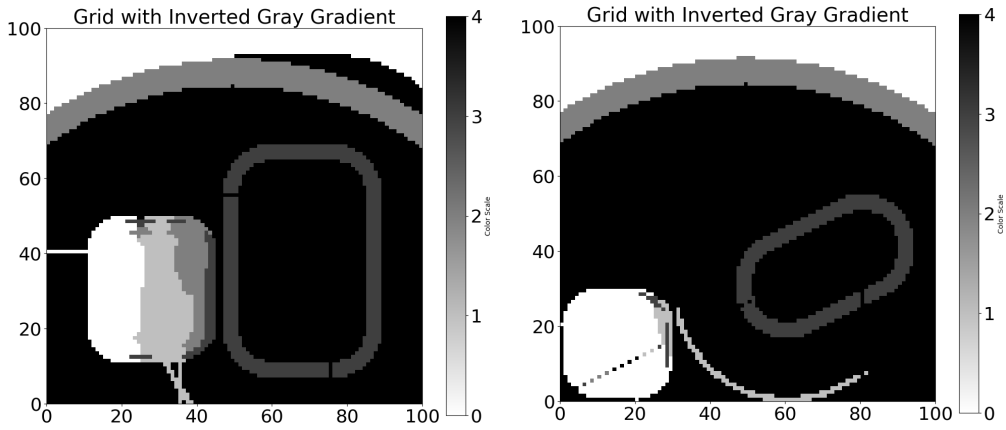


Figure 4. Grey clustering diagrams for marine area importance under the two scenarios.

In the diagram, different marine areas' importance is represented by a gradient of colors, where darker colors indicate higher importance. The importance distribution ranges from 0 to 4. Using this clustering method allows for effective segmentation of the task area based on varying levels of importance.

3.3. Multi-AUVs Cooperative Search Strategy Based on Fuzzy Reinforcement Learning

Traditional reinforcement learning methods employ Markov Decision Processes (MDPs) as the mathematical model, constructing state and action spaces to define objective functions and state/action rewards based on the analysis of the task scenario. Through continuous iterative calculations, these methods aim to determine the optimal state values or the best action strategies within the task environment. Given the fuzziness and uncertainty inherent in the regional importance model constructed in this paper, it is reasonable to integrate fuzzy logic with traditional reinforcement learning to address the multi-AUV collaborative search problem and achieve optimal search gains.

The search reward R represents the cumulative reward value obtained by an AUV while searching the grid map, indicating the effectiveness of the AUV's search operations under the weight of importance. As shown in Equations (21) and (22), k is the importance level, j is the current number of searches for a grid, i is the step number in the search, z is the AUV index, x, and y is the row and column numbers of the search grid used to identify specific grids.

$$\mathrm{R} = \sum_{i=1}^{1000} r_{iz}, z = 1, 2, 3, \ldots \tag{21}$$

$$r_{iz} = k_{xy} + 1 - j_{iz} \tag{22}$$

3.3.1. Initial Patrol Strategy π_0

AUVs start their search from different points and search one grid at each step. The reward obtained is denoted by r, and r follows the formula provided in Equation (22), indicating that the reward diminishes as a grid is repeatedly searched. The next step action

is denoted by a and $a = \{0, 1, 2, 3, 4\}$, representing staying in place, moving left, moving up, moving right, and moving down, respectively.

The initial patrol strategy π_0 employs a random search strategy where AUVs start from designated grids and randomly choose directions to move in, advancing one grid at each step and receiving the grid's importance as a reward. No reward is calculated upon touching a boundary, and another direction is chosen to continue until the search reaches the specified step length. Finally, the total search rewards for all AUVs are calculated.

3.3.2. Policy Evaluation

As shown in Equation (23), π_k denotes the policy, j represents the number of iterations, r_k represents the immediate reward under the policy π_k, γ is the discount factor, $V_k^{(j)}$ represents the state–value matrix obtained in the j-th iteration under the policy π_k, p_{π_k} represents the complete probability distribution matrix obtainable for $V_k^{(j)}$ under the policy π_k. Equation (23) is iterated until $\left\| \mathrm{v}_{\pi k}^{(j+1)} - \mathrm{v}_{\pi k}^{(j)} \right\| \leq 0.0001$, and then the optimal state value under the policy can be regarded as similar to Equation (24).

$$\mathrm{v}_{\pi k}^{(j+1)} = r_{\pi k} + \gamma P_{\pi k} \mathrm{v}_{\pi k}^{(j)},\ j = 0, 1, 2, \ldots \tag{23}$$

$$\mathrm{v}_{\pi k}^{*} = \mathrm{v}_{\pi k}^{(j)} \tag{24}$$

3.3.3. Policy Improvement

As shown in Equation (26), the policy is updated at step K + 1 with the optimal state value, and the specific formula is expanded as Equations (27)–(29).

$$\mathrm{v}_{\pi k}^{*} = \mathrm{v}_{\pi k}^{(j)} \tag{25}$$

$$\pi_{k+1} = \arg\max_{\pi}(r_{\pi} + \gamma P_{\pi k} \mathrm{v}_{\pi k}^{*}) \tag{26}$$

$$q_{\pi k}(s, a) = \sum_{r} p(r \mid s, a) r + \gamma \sum_{s'} p(s' \mid s, a) v_{\pi k}^{*}(s') \tag{27}$$

$$a_k^{*}(s) = \arg\max_{a} q_{\pi k}(s, a) \tag{28}$$

$$\pi_{k+1}(a|s) = \begin{cases} 1, a = a^{*} \\ 0, a \neq a^{*} \end{cases} \tag{29}$$

$q_{\pi k}(s, a)$ represents the reward obtained by taking action a in state s and continuing according to policy π_k until the end of the episode in the equation. The right side of the equation is the expansion of the total probability formula.

$a_k^{*}(s)$ denotes the action, which leads to the maximum reward $q_{\pi k}(s, a)$; π_{k+1} is the updated policy, and by continuously iterating and updating, the optimal policy π^{*} can be achieved.

Some of the pseudo-code for the strategy iteration algorithm is shown in Algorithm 1.

Algorithm 1 Policy Iteration for Grid Environment

Initialization:
Initialize policy π with equal probability for each action
Set $V(s) = 0$ for all states s
Set $\gamma = 0.9$, $\theta = 0.00001$
Policy Evaluation:
repeat
    Set $\Delta = 0$
    for each state s **do**
        $v \leftarrow V(s)$
        $V(s) \leftarrow \sum_a \pi(a \mid s) \sum_{s'} P(s' \mid s, a)[R(s, a, s') + \gamma V(s')]$
        $\Delta \leftarrow \max(\Delta, |v - V(s)|)$
    end for
until $\Delta < \theta$
Policy Improvement:
Policy stable $\leftarrow$ true
for each state s **do**
    old_action $\leftarrow \arg\max_a \pi(a \mid s)$
    for each action a **do**
        $Q(s, a) \leftarrow \sum_{s'} P(s' \mid s, a)[R(s, a, s') + \gamma V(s')]$
    end for
    $\pi(s) \leftarrow \arg\max_a Q(s, a)$
    if old_action $\neq \pi(s)$ **then**
        Policy stable $\leftarrow$ false
    end if
end for
if policy stable **then**
    exit
end if
Repeat:
Perform Policy Evaluation and Policy Improvement until policy is stable
End:
Return the optimal policy π and value function V

4. Simulation Results and Discussion

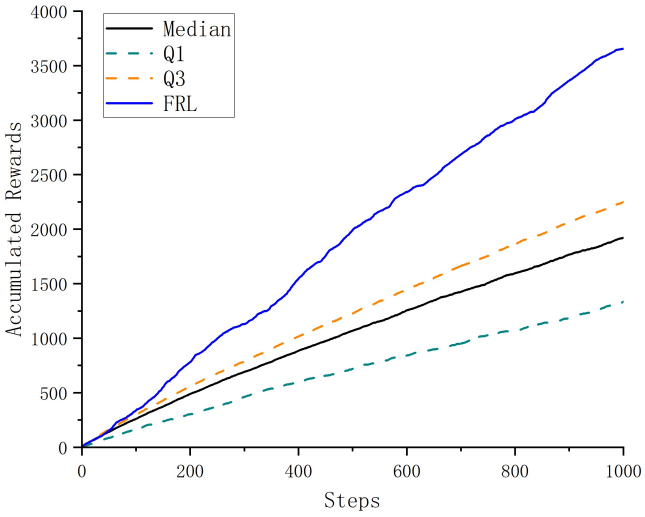
This paper presents a case study focusing on a fixed marine area task region, employing the second operational scenario to validate the effectiveness of strategy optimization following map updates. The study involves dividing a 10 km × 10 km area into a 100 × 100 grid, with each grid measuring 100 m × 100 m. It is stipulated that each AUV's detection range fully covers a grid as it passes through. Due to energy consumption limitations, each AUV can only travel through 1000 grids. Each grid, depending on its importance, provides varying search yields (with initial reward values equivalent to grid importance). During patrols, grids that are searched repeatedly have their importance decay within the map update cycle, with a decay discount factor set at 0.9. The study compares the search gains from a random search strategy and an optimized search strategy developed through policy iteration based on reinforcement learning, without considering map changes within one map update cycle. Both strategies involve AUVs traveling 1000 steps and calculating the obtained search rewards.

Given the distributed search architecture designed in this paper, each AUV's decision-making on search strategies only considers the influence of nearby AUVs. The simulation results will illustrate search strategies and rewards under scenarios with 0, 1, and 2 nearby AUVs. The figures demonstrate the effectiveness of the area search in terms of search rewards R, which represents the effectiveness of the marine area search.

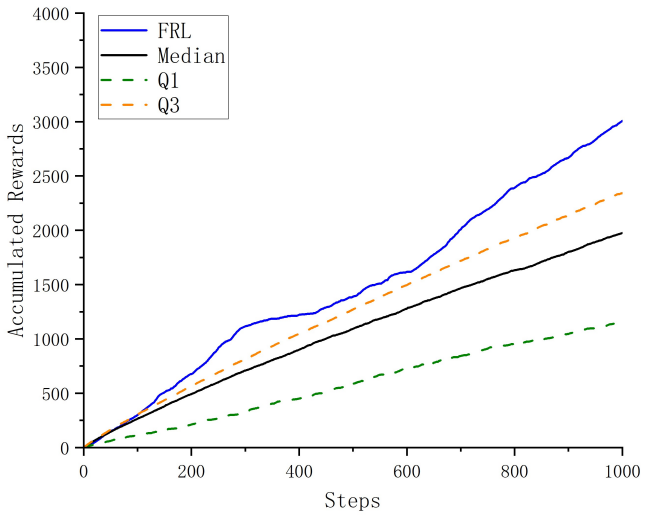
As shown in Figures 5–7, the median and quartile traces for random behavior in the figure represent the search reward curves obtained from 1000 random search strategies

performed by the target AUV and its neighboring AUVs (if there are any). These curves indicate the level of cooperative search efficiency under the initial strategy. The blue line marked as FRL represents the search reward curves obtained by the target AUV and its neighboring AUVs (if there are any) using a search strategy based on fuzzy reinforcement learning, reflecting the cooperative search efficiency of the target AUV and its neighboring AUVs. As can be observed, these curves are generally significantly superior to the initial random strategy, and the advantage of the cooperative search strategy based on fuzzy reinforcement learning (FRL) becomes more pronounced as the number of AUVs increases.

Average search rewards under different numbers of AUVs for random search strategy and cooperative search strategy based on fuzzy reinforcement learning are detailed in Table 2. The average search rewards decrease due to the increased number of nearby AUVs, which leads to more frequent grid re-searches and thus faster reward decay. Similarly, the search strategy based on fuzzy reinforcement learning shows a significant advantage in all three cases.



[Condition 1]



[Condition 2]

Figure 5. Comparison of rewards for search strategies based on fuzzy reinforcement learning and random search (1 AUV).

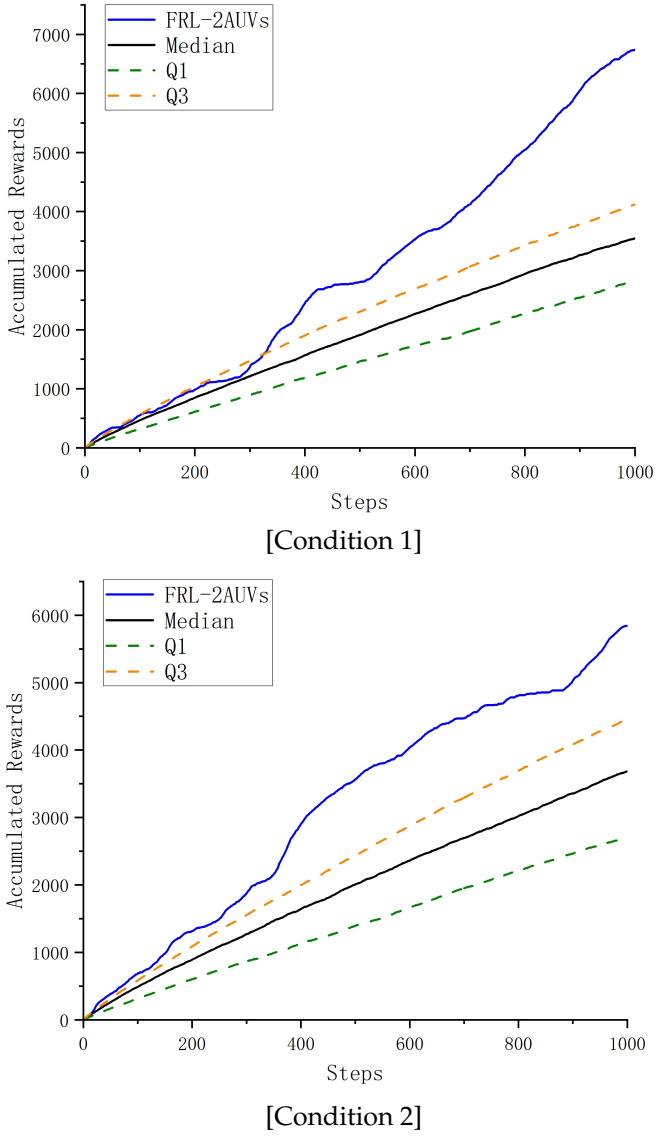


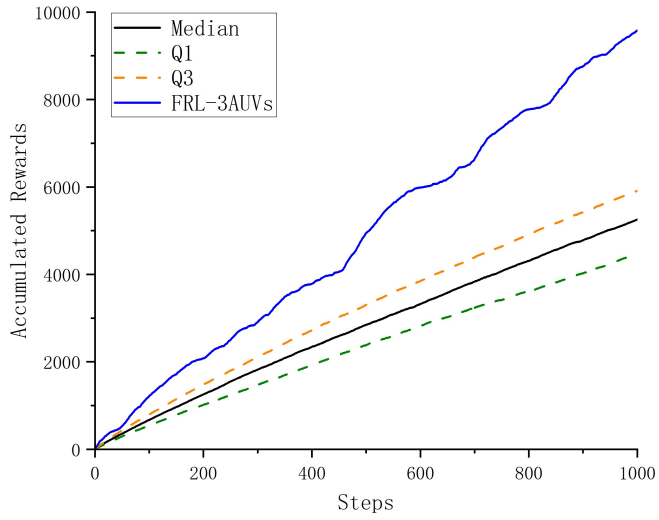
Figure 6. Comparison of rewards for search strategies based on fuzzy reinforcement learning and random search (2 AUVs).

To validate the stability of the method, this paper adopts the standard shown in Equation (30), which is a more stringent criterion for determining the convergence of strategy optimization. The search gain obtained by continuing the iteration of the strategy is considered to be no longer increasing, and the iterations are terminated at the point when $delta \leq 0.00001$. The value of *delta* here represents the difference in rewards between the strategy after iteration and the strategy before iteration. When this difference is less than a certain threshold, we can consider that the current strategy to be unlikely to be further optimized through continued iteration, indicating convergence.

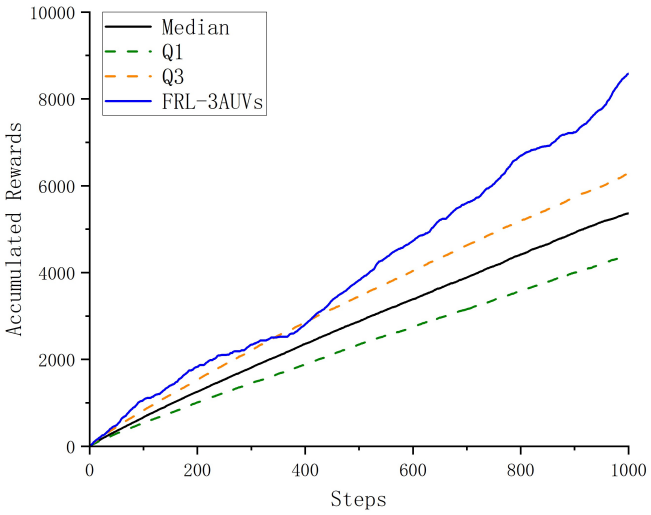
In the fuzzy reinforcement-learning-based strategy iteration used in this paper, each iteration sequence is designed to include 100 iterations. After 100 iterations, the current iteration will end and proceed to the next iteration round, regardless of whether convergence has been achieved. The 100-iteration design is intended to prevent slow convergence and excessive computation time. Each new iteration round inherits part of the strategy from the previous iteration, which helps to achieve rapid convergence over multiple iterations.The convergence behavior during the policy iteration process are shown as Figures 8 and 9. It can be observed that in each iteration cycle, the method consistently achieves convergence within 100 iterations, and the speed of convergence increases with each subsequent cycle. It

can be considered that this method has a relatively high computational efficiency, making it suitable for cooperative search scenarios with time-sensitive requirements.

$$\left\| \mathrm{v}_{\pi k}^{(j+1)} - \mathrm{v}_{\pi k}^{(j)} \right\| \leq 0.00001 \tag{30}$$



[Condition 1]



[Condition 2]

Figure 7. Comparison of rewards for search strategies based on fuzzy reinforcement learning and random search (3 AUVs).

Table 2. Comparison of average search rewards based on FRL and random search.

Number of AUV	Condition 1		Condition 2	
	FRL Search Rewards	Random Search Rewards	FRL Search Rewards	Random Search Rewards
1	3653	1770.64	3007	1825.52
2	3370	1744.54	2920	1760.94
3	3194	1649.85	2862	1727.9

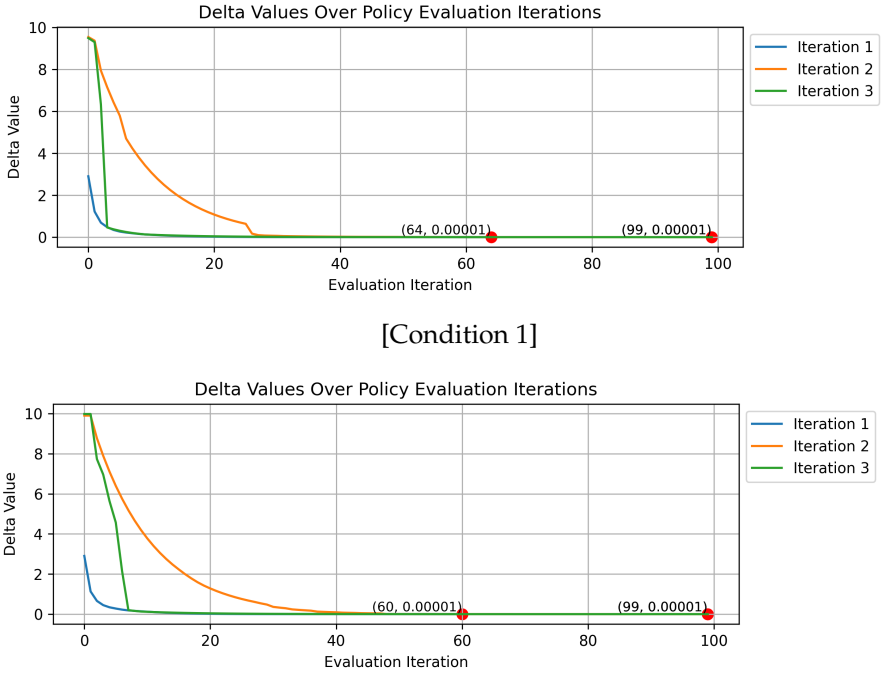


Figure 8. The delta change curve for the three iterations (100 evaluation iterations per round of improvement iterations).

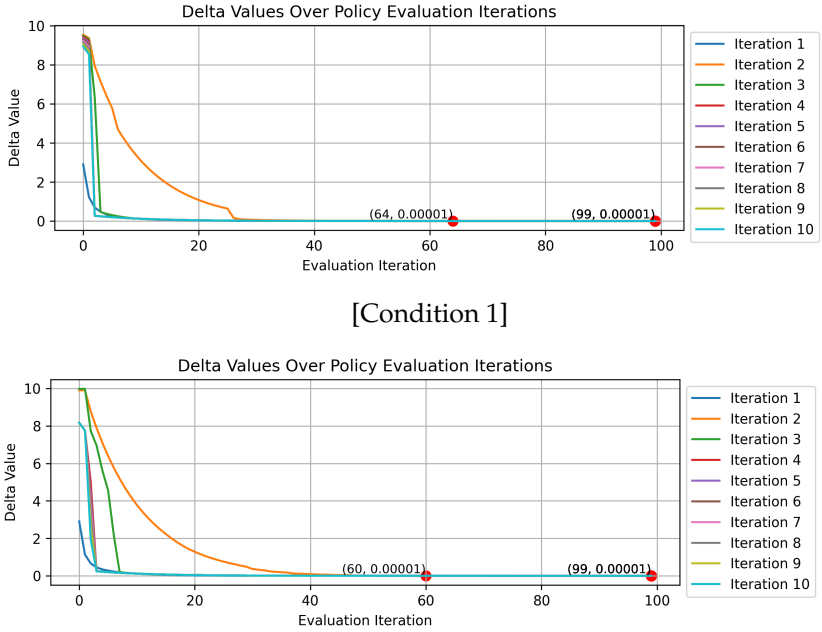


Figure 9. The delta changecurve for ten iterations (100 evaluation iterations per round of improvement iterations).

Additionally, the search gain curves for 1000-step searches using the Breadth-First Search (BFS) and Depth-First Search (DFS) algorithms in two different scenarios were analyzed. The simulation results are as follows, with the search gains detailed in Table 3. It is evident that traditional DFS and BFS algorithms exhibit significant redundant searches as the number of AUVs increases, leading to a rapid decline in search gains. In contrast,

the search strategy based on fuzzy reinforcement learning demonstrates a clear advantage in multi-AUV cooperative search, as shown in Figures 10–13.

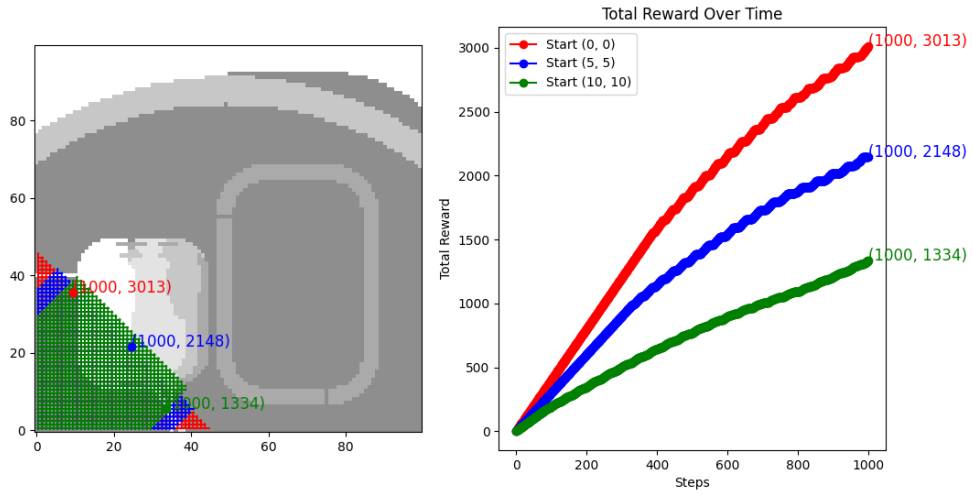


Figure 10. Searching 1000 steps with BFS under condition 1.

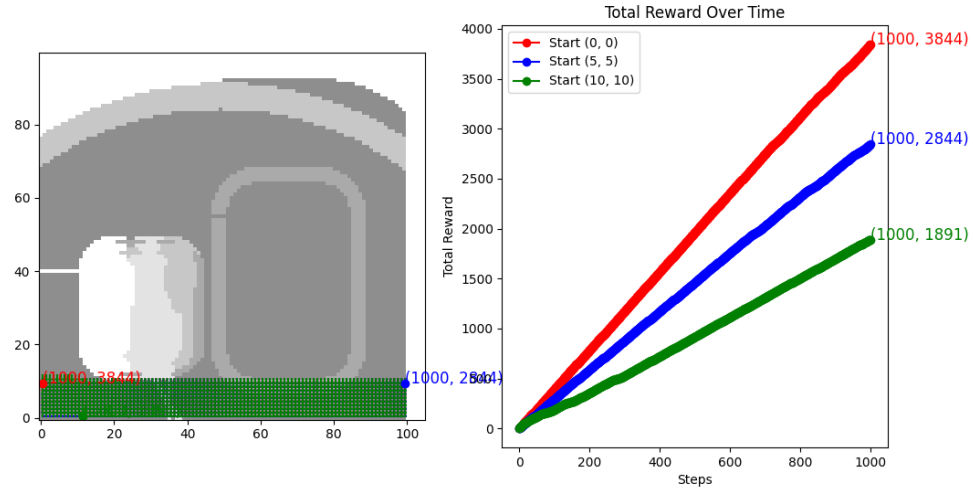


Figure 11. Searching 1000 steps with DFS under condition 1.

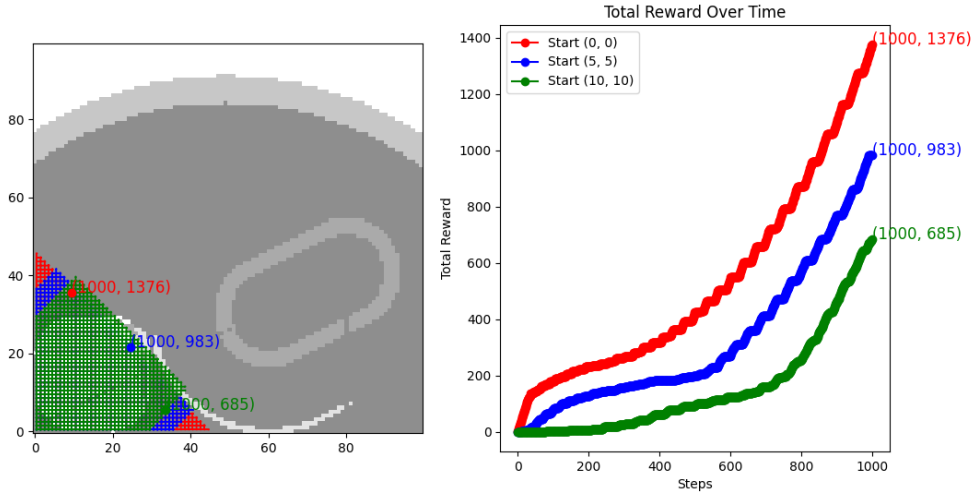


Figure 12. Searching 1000 steps with BFS under condition 2.

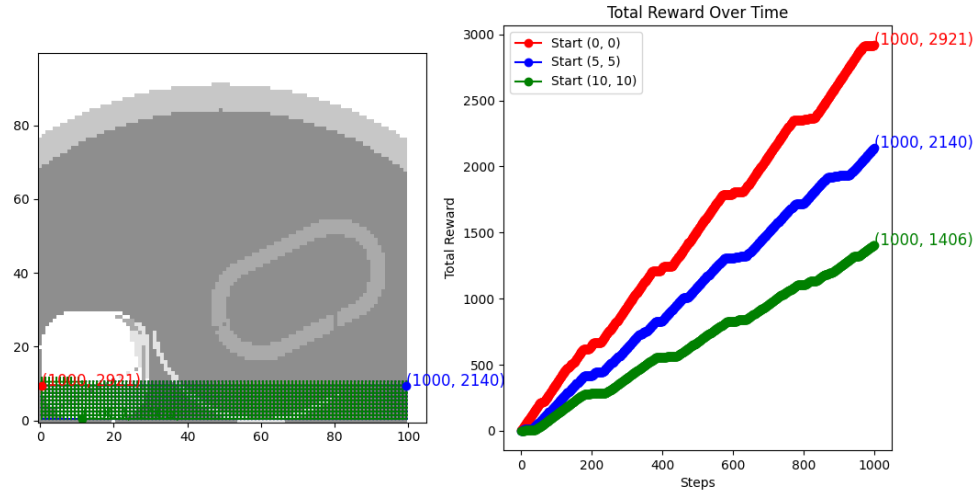


Figure 13. Searching 1000 steps with DFS under condition 2.

Table 3. Search rewards from 1000 steps with different strategies.

	Condition 1			Condition 2		
Number of AUV	**DFS**	**BFS**	**FRL**	**DFS**	**BFS**	**FRL**
1	3844	3013	3653	2921	1376	3007
2	6688	5161	6740	5061	2359	5840
3	8579	6495	9582	6467	3044	8586

5. Conclusions

The area importance model of the grid map in this paper is constructed based on prior information and grey system theory. The evaluation of importance serves as a fuzzy assessment of the marine area under uncertain conditions. By adopting a two-dimensional gridding method for the task area, the model is greatly simplified. Utilizing the policy iteration algorithm from reinforcement learning, the system optimizes search strategies for rapidly identifying high-importance areas within the marine environment, providing quick solutions, fast convergence, and substantial search gains. This approach offers valuable insights into the collaborative search/patrol of multiple AUVs in fuzzy marine environments.

Here are the limitations and shortcomings:

(1) Dimensional Scalability: This study primarily explores two-dimensional area division based on importance, which should also be applicable to three-dimensional division and task planning. However, the addition of the vertical dimension could significantly increase the number of grid areas, and the necessity of this in practical applications needs to be assessed. Typically, AUVs do not frequently adjust their depth during underwater operations, so it can be assumed that AUVs patrol within a fixed depth grid, closely approximating the two-dimensional scenario described in this paper. Only in specific areas would small-scale three-dimensional division and task planning be conducted, making this method feasible and beneficial.

(2) Model Design Simplification: The design of the area importance model cleverly uses the Euclidean distance between region locations and various points (starting points, no-sail zones, focus areas) as different dimensional indicators for importance evaluation. This approach allows for a succinct and efficient construction of the state space based solely on location. However, the importance of an area in practical situations involves more dimensions than just location (though location is almost always a critical factor), possibly including hydrological environment characteristics, the degree of area familiarity (substituted by patrol frequency in this study), and the

impact of incidental events. Therefore, the construction of the state space should be more complex and dynamic, requiring deeper research.

Author Contributions: Conceptualization, G.Z. and Y.S.; Methodology, K.C., G.Z. and Y.S.; Software, Y.S., G.D. and F.X.; Formal analysis, K.C. and G.D.; Investigation, K.C.; Data curation, K.C. and F.X.; Writing—original draft, K.C. and G.D.; Writing – review & editing, K.C., G.Z., Y.S. and F.X.; Visualization, K.C. and G.D.; Supervision, G.Z., Y.S. and F.X.; Project administration, G.Z. and Y.S.; Funding acquisition, G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 52371310.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, L.; Zhu, D.; Pang, W.; Zhang, Y. A survey of underwater search for multi-target using Multi-AUV: Task allocation, path planning, and formation control. *Ocean Eng.* **2023**, *278*, 114393 . [CrossRef]
2. Wang, Y.; Li, H.; Yao, Y. An adaptive distributed auction algorithm and its application to multi-auv task assignment. *Sci. China Technol. Sci.* **2023**, *66*, 1235–1244. [CrossRef]
3. Li, H.; Chen, M. Task allocation and path planning problems of multi-auv system based on auction-dynamic neural network. In Proceedings of the 2023 35th Chinese Control and Decision Conference (CCDC), Yichang, China, 20–22 May 2023; IEEE: New York, NY, USA, 2023; pp. 2945–2949.
4. Wang, L.; Dou, J.; Ji, Z.; Sun, Y.; Liu, M.; Wang, S. Region division strategy for multi-auv charging scheme in underwater rechargeable sensor networks. In Proceedings of the 2023 8th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA), Chengdu, China, 26–28 April 2023; IEEE: New York, NY, USA, 2023; pp. 364–368.
5. Ghassemi, P.; Chowdhury, S. Decentralized task allocation in multi-robot systems via bipartite graph matching augmented with fuzzy clustering. In Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Quebec City, QC, Canada, 26–29 August 2018; American Society of Mechanical Engineers: New York, NY, USA, 2018; Volume 51753, p. V02AT03A014.
6. Seenu, N.; Ramya, K.C.R.M.M.; Janardhanan, M.N. Review on state-of-the-art dynamic task allocation strategies for multiple-robot systems. *Ind. Robot. Int. J. Robot. Res. Appl.* **2020**, *47*, 929–942.
7. Bänziger, T.; Kunz, A.; Wegener, K. Optimizing human–robot task allocation using a simulation tool based on standardized work descriptions. *J. Intell. Manuf.* **2020**, *31*, 1635–1648. [CrossRef]
8. Patel, R.; Rudnick-Cohen, E.; Azarm, S.; Otte, M.; Xu, H.; Herrmann, J.W. Decentralized task allocation in multi-agent systems using a decentralized genetic algorithm. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; IEEE: New York, NY, USA, 2020; pp. 3770–3776.
9. Geng, N.; Chen, Z.; Nguyen, Q.A.; Gong, D. Particle swarm optimization algorithm for the optimization of rescue task allocation with uncertain time constraints. *Complex Intell. Syst.* **2021**, *7*, 873–890. [CrossRef]
10. Chen, Z.; Zhang, D.; Wang, C.; Sha, Q. Hybrid form of differential evolutionary and gray wolf algorithm for multi-auv task allocation in target search. *Electronics* **2023**, *12*, 4575. [CrossRef]
11. Zhang, J.; Ning, X.; Ma, S. An improved particle swarm optimization based on age factor for multi-auv cooperative planning. *Ocean Eng.* **2023**, *287*, 115753. [CrossRef]
12. Sun, B.; Zeng, Y.; Su, Z. Task allocation in multi-auv dynamic game based on interval ranking under uncertain information. *Ocean Eng.* **2023**, *288*, 116057. [CrossRef]
13. Zhu, D.; Zhou, B.; Yang, S.X. A novel algorithm of multi-auvs task assignment and path planning based on biologically inspired neural network map. *IEEE Trans. Intell. Veh.* **2020**, *6*, 333–342. [CrossRef]
14. Ma, X.; Yang, J. Collaborative planning algorithm for incomplete navigation graphs. *Ocean Eng.* **2023**, *280*, 114464. [CrossRef]
15. Rajnarayan, D.G.; Ghose, D. Multiple agent team theoretic decision-making for searching unknown environments. In Proceedings of the 42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475), Maui, HI, USA, 9–12 December 2003; IEEE: New York, NY, USA, 2003; Volume 3, pp. 2543–2548.
16. Zhang, H.; Cheng, Z. The method based on dijkstra of three-dimensional path planning. In Proceedings of the 2020 Chinese Automation Congress (CAC), Shanghai, China, 6–8 November 2020; IEEE: New York, NY, USA, 2020; pp. 1698–1701.
17. Cai, C.; Chen, J.; Ayub, M.S.; Liu, F. A task allocation method for multi-auv search and rescue with possible target area. *J. Mar. Sci. Eng.* **2023**, *11*, 804. [CrossRef]

18. Zhang, J.; Liu, M.; Zhang, S.; Zheng, R. Auv path planning based on differential evolution with environment prediction. *J. Intell. Robot. Syst.* **2022**, *104*, 23. [CrossRef]
19. Wang, Z.; Sui, Y.; Qin, H.; Lu, H. State super sampling soft actor–critic algorithm for multi-auv hunting in 3D underwater environment. *J. Mar. Sci. Eng.* **2023**, *11*, 1257. [CrossRef]
20. Jouffe, L. Fuzzy inference system learning by reinforcement methods. *IEEE Trans. Syst. Man, Cybern. Part C (Appl. Rev.)* **1998**, *28*, 338–355. [CrossRef]
21. Fathinezhad, F.; Derhami, V.; Rezaeian, M. Supervised fuzzy reinforcement learning for robot navigation. *Appl. Soft Comput.* **2016**, *40*, 33–41. [CrossRef]
22. Juang, C.-F.; You, Z.-B. Reinforcement learning of an interpretable fuzzy system through a neural fuzzy actor–critic framework for mobile robot control. *IEEE Trans. Fuzzy Syst.* **2024**, *32*, 3655–3668. [CrossRef]
23. Fu, Q.; Pu, Z.; Pan, Y.; Qiu, T.; Yi, J. Fuzzy feedback multi-agent reinforcement learning for adversarial dynamic multi-team competitions. *IEEE Trans. Fuzzy Syst.* **2024**, *32*, 2811–2824. [CrossRef]
24. Fang, B.; Zheng, C.; Wang, H.; Yu, T. Two-stream fused fuzzy deep neural network for multiagent learning. *IEEE Trans. Fuzzy Syst.* **2022**, *31*, 511–520. [CrossRef]
25. Deng, J. Introduction to grey system theory. *J. Grey Syst.* **1989**, *1*, 1–24.
26. Liu, S.; Forrest, J.; Yang, Y. A brief introduction to grey systems theory. In Proceedings of the 2011 IEEE International Conference on Grey Systems and Intelligent Services,Nanjing, China, 15–18 September 2011; IEEE: New York, NY, USA, 2011; pp. 1–9.