



## Reinforcement learning-integrated evolutionary algorithm for enhanced unmanned aerial vehicle coverage path planning

Seung Chan Choi<sup>a</sup>, Yohan Lee<sup>a</sup>, Sung Won Cho<sup>a,b</sup>,\*

<sup>a</sup> Department of Industrial Engineering, Dankook University, Cheonan, Republic of Korea

<sup>b</sup> Department of Management Engineering, Dankook University, Cheonan, Republic of Korea

### ARTICLE INFO

#### Keywords:

Reinforcement learning (RL)  
Particle swarm optimization (PSO)  
Proximal policy optimization (PPO)  
Search and rescue (SAR)  
Coverage path planning (CPP)  
Unmanned aerial vehicle (UAV)

### ABSTRACT

The rapid development of unmanned aerial vehicle (UAV) technologies has led to their increased utilization across various industries. In search and rescue (SAR) missions, UAVs play a critical role in overcoming mobility constraints in search environments, particularly in time-sensitive situations such as maritime operations. To enhance the efficiency of search missions, this study addresses the Coverage Path Planning (CPP) problem for multiple UAVs in irregularly shaped search areas. We propose a novel CPP framework consisting of two main phases. In Phase 1, a reinforcement learning-integrated evolutionary algorithm is introduced for search area decomposition, aiming to minimize the area of the grid map exceeding the search area. Specifically, proximal policy optimization-based particle swarm optimization (PPO-PSO) is employed to effectively adapt to complex and irregular shapes. In Phase 2, a Mixed Integer Linear Programming (MILP) model is formulated to minimize mission completion time while ensuring collision avoidance and efficient task allocation for multiple UAVs. The proposed methodology was validated through 15 experimental scenarios, including real-world maritime environments, and demonstrated superior performance compared to existing methods in managing irregularly shaped search areas.

### 1. Introduction

With rapid development of technologies related to unmanned aerial vehicles (UAVs) such as battery efficiency, sensors, and cameras, they are being applied in various industries. UAVs have been deployed across various industries including defense [1], agriculture [2,3], construction [4], and logistics [5,6].

UAVs offer advantages in situations involving tasks that are difficult, cumbersome, or hazardous for humans to perform. UAV used when image acquisition for the purpose of conducting terrain analysis [7]. Additionally, UAVs provide easier access than humans during mobile, exploratory, and data collection processes [8]. Furthermore, UAVs enable rapid response in various situations such as disasters and emergencies [9].

Search and Rescue (SAR) missions can lead to delays in completion time due to mobility constraints in search environments. In marine environments, delays can lead to significant human casualties and economic damage [10]. UAVs are widely used for reconnaissance purposes due to their low environmental constraints [11]. Nevertheless, since UAVs necessitate human remote control, human errors can occur, and pilot fatigue can be accumulated, and thereby diminishing concentration over prolonged missions [12]. Therefore, it is important to automate the search operations of UAVs.

When it comes to UAV search operations, covering the entire search area quickly is essential, which is expressed through Coverage Path Planning (CPP). To achieve rapid mission execution, it is necessary to allocate the given area appropriately to the UAVs and derive the shortest path to accomplish the task in the shortest time possible. Since multiple UAVs are operated simultaneously, it is crucial to prevent collisions between them during mission execution.

In this study, we aim to efficiently address the CPP, a crucial part in searching target areas during SAR missions using multiple UAVs. To minimize the area of the grid map exceeding the search area, we propose a novel reinforcement learning-integrated evolutionary algorithm to determine optimal point locations for dividing the target area into cells of triangles connecting the vertices of the polygon and interior points. Ultimately, we aim to compute the shortest path, to minimize the time of search mission for UAVs.

The remainder of this paper is organized as follows. Section 2 introduces existing studies on CPP with UAVs. Section 3 describes the proposed methodology. Section 4 describes the experimental setup and the results of the experiments. Section 5 provided the conclusion of this paper.

\* Corresponding author.

E-mail address: [sungwon.cho@dankook.ac.kr](mailto:sungwon.cho@dankook.ac.kr) (S.W. Cho).

## 2. Literature review

In CPP, the goal is to generate a route that covers a given target area. CPP is used in a wide range of fields, such as cleaning robots, self-driving cars, landmine removal robots, and lawn mowers [13]. In addition, research has been conducted on CPP under numerous constraints, including the number of robots, number and shape of areas, and presence of obstacles [14]. This study focuses on a comprehensive analysis of decomposition methods used in CPP to divide the target area effectively.

Most decomposition methods in CPP are categorized as either cell-based or grid-based approaches. CPP using cell-based decompositions — including exact cellular decomposition, trapezoidal decomposition, boustrophedon decomposition, Morse-based decomposition, and slice decomposition — generally has faster computation times due to simpler path planning processes. However, cell-based methods can produce inefficient paths in irregularly shaped areas, as coverage efficiency may vary with cell size and shape, leading to performance degradation in optimization. In contrast, CPP with grid-based decompositions, such as the wavefront algorithm, spanning trees, neural networks, and hexagonal grid decomposition, divides the target area into small grid units, enabling more detailed and optimized path planning. This approach often results in higher optimization performance [13,15,16].

This study employs grid-based decomposition methods due to their superior performance in complex, irregular environments. Grid-based decomposition involves dividing the target area into a regular grid for coverage. Valente et al. [17] utilized a Sukharev grid in CPP, with nodes visited once per the coverage solution using a depth-limited search. Further, Valente et al. [18] applied a harmony search algorithm to CPP for UAVs, reducing the number of rotations by generating random harmony vectors. Nedjati et al. [19] developed a grid-based method that adapted to the resolution of UAV cameras, applying linear strength techniques within a 5-index model to achieve lower error and computational time. Modares et al. [20] optimized energy-efficient coverage for UAVs by using a mixed integer linear programming (MILP) model and Lin-Kernighan heuristics. Choi et al. [21] introduced a route optimization model with a turn penalty, resulting in less rotation and greater path efficiency. Li et al. [22] developed a hexagonal grid-based coverage method for environmental mapping, improving spatial exploration by optimizing edge configurations within the grid. Khiati et al. [23] demonstrated the efficiency of the grid-based approach in mission environments, where it outperformed the point-gathering method. Pérez-González et al. [24] processed UAV images through deep learning servers, comparing three CPP methods and identifying that boustrophedon and grid-based wavefront coverage were most efficient depending on CPP width and area type.

Zhu et al. [25] proposed an online heuristic that integrates geometric and stochastic elements, making it resilient to UAV failures and environmental changes. Heydari et al. [26] utilized reinforcement learning to surpass the limitations of heuristic and approximate algorithms. Kyriakakis et al. [27] implemented a parallel weighted greedy randomized adaptive search for CPP in convex areas. Huang et al. [28] proposed a time-efficient CPP algorithm addressing energy-limited scenarios, and Hu et al. [29] introduced cooperative multi-UAV strategies using a distributed deep Q-learning algorithm for uncertain environments. Yu and Lee [30] proposed two algorithms for multi-UAV cooperative CPP. One coordinates UAVs via multiple base stations, the other decomposes the target area into grids and plans paths independently for each grid. Wu et al. [31] proposed an integrated framework for maritime SAR missions using multiple vessels, which utilizes irregular drift prediction results of persons-in-water to construct a hierarchical environmental map and leverages a deep reinforcement learning algorithm to autonomously generate efficient, priority-oriented coverage paths.

However, despite these advances, existing decomposition methods primarily focus on polygonal or convex-shaped areas, making them

challenging to apply to irregular environments. This limitation results in inefficient coverage paths that do not take into account the characteristics of search areas.

A few studies have tackled CPP for irregular areas. Cabreira et al. [32] introduced an energy-aware grid-based CPP method that reduced computation time by using energy-dependent cost functions to optimize energy consumption. Ma and Xiong [33] proposed an improved grid method that minimizes excessive coverage and optimizes flight paths with an enhanced ant colony algorithm. Ghaddar et al. [34] proposed a partition-based approach to divide the entire area into grids in environments with a non-flying zone (NFZ). Apostolidis et al. [35] proposed a platform that distributes tasks across UAVs based on their specific sensing and operational capabilities to cover complex-shaped areas. Gong et al. [36] introduced a parallel self-adaptive ant colony optimization to solve the CPP for a single UAV considering a NFZ. Ai et al. [37] proposed an autonomous maritime CPP approach that leverages a reinforcement learning model, Q-learning, based on a markov decision process, by integrating a multi-objective reward function with a sine function-based non-linear action selection policy.

The aforementioned studies provide valuable insights into grid-based CPP for irregular shapes. However, they primarily focus on optimizing the performance of path planning algorithms, rather than examining how decomposition methods themselves impact the efficiency of coverage paths. This gap motivates further research into decomposition methods tailored specifically for irregular search areas to enhance coverage path optimization. Furthermore, most research on CPP with irregular search areas has utilized a single UAV, with only a few studies considering multiple UAVs. The literature related to the scope of this study is summarized in Table 1.

The contributions of this study are as follows: First, this study introduces a novel grid-based decomposition methodology tailored for irregularly shaped search areas. Unlike existing methods that primarily focus on polygonal or convex shapes, the proposed approach minimizes the area of the grid map exceeding the search area by determining optimal interior points within the search area, effectively adapting to the complexity of irregular shapes. Second, we propose a novel reinforcement learning-integrated evolutionary algorithm for search area decomposition. Specifically, the introduced Proximal Policy Optimization-based Particle Swarm Optimization (PPO-PSO) approach leverages the adaptability of reinforcement learning to guide the search dynamics of swarm intelligence. This integration enables more robust and flexible decomposition of complex search areas, marking a novel contribution in combining deep reinforcement learning with evolutionary optimization in the CPP domain. Third, the study extends the CPP framework to account for the simultaneous operation of multiple UAVs, emphasizing collision avoidance and optimal task allocation. This multi-UAV strategy ensures faster mission execution and reduces search times, which is critical in time-sensitive applications such as maritime SAR missions. Lastly, the proposed methodology is applicable beyond SAR missions to fields like precision agriculture, environmental mapping, and disaster response, addressing challenges in irregular terrains. Additionally, the approach can be extended to support collaborative multi-robot operations, contributing to coordinated multi-agent systems.

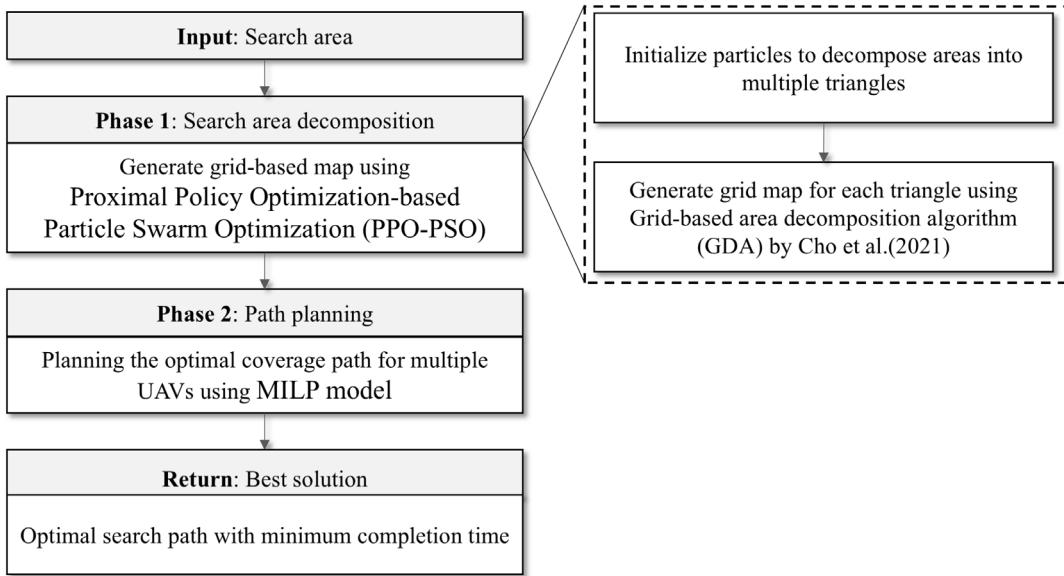
## 3. Methodology

We present a multiple UAVs CPP method using PPO-PSO for search area decomposition. The proposed method consists of two phases, as shown in Fig. 1. In Phase 1, the search area is divided into triangles by connecting a point in the region with the vertices of the area. For each triangle, a grid map is generated using a grid-based region decomposition algorithm (GDA) by Cho et al. [12]. PPO-PSO method is used to determine the optimal point within the region. In Phase 2, optimal coverage paths for multiple UAVs are generated using MILP model based on the grid map generated in the first phase.

**Table 1**  
Overview of literature on CPP with grid-based decomposition methods.

Papers	Area shape	Number of robots	Collision avoidance	Solution method <sup>a</sup>
Valente et al. [17]	Convex	Single		Heuristic
Valente et al. [18]	Convex	Multiple homogeneous		HS
Nedjati et al. [19]	Convex	Multiple homogeneous		MILP
Modares et al. [20]	Irregular	Multiple homogeneous		Heuristic
Choi et al. [21]	Convex	Multiple homogeneous		MILP
Li et al. [22]	Irregular	Single		Heuristic
Khiati et al. [23]	Irregular	Single		Heuristic
Pérez-González et al. [24]	Convex	Single		Heuristic
Zhu et al. [25]	Convex	Multiple homogeneous	✓	Heuristic
Heydari et al. [26]	Irregular	Single		RL
Kyriakakis et al. [27]	Convex	Multiple homogeneous		GRASP
Huang et al. [28]	Irregular	Single		GLNS
Hu et al. [29]	Irregular	Multiple homogeneous	✓	RL, heuristic
Yu and Lee [30]	Irregular	Multiple homogeneous	✓	Heuristic
Cabreira et al. [32]	Irregular	Single		Heuristic
Ma and Xiong [33]	Irregular	Single		ACO
Ghaddar et al. [34]	Irregular	Multiple homogeneous	✓	Heuristic
Apostolidis et al. [35]	Irregular	Multiple homogeneous	✓	SA
Gong et al. [36]	Irregular	Single		Parallel ACO
Ai et al. [37]	Irregular	Single		RL
Wu et al. [31]	Convex	Multiple homogeneous	✓	RL
Cho et al. [12]	Convex	Multiple heterogeneous	✓	MILP, heuristic
This paper	Irregular	Multiple heterogeneous	✓	MILP, RL, PSO

<sup>a</sup> Harmony search (HS), mixed integer linear programming (MILP), reinforcement learning (RL), greedy randomized adaptive search procedure (GRASP), generalized large neighborhood search (GLNS), ant colony optimization (ACO), simulated annealing (SA), particle swarm optimization (PSO).



**Fig. 1.** Research framework.

### 3.1. Phase 1: Search area decomposition

Grid-based decomposition algorithm divides a target area into grids, typically setting the grid shape as a square to consider the shape of the camera footprint [17]. When the target area with irregular shape is divided into square grids, a grid map with a size exceeding the target area is created. This excess area increases the completion time of the search mission. Therefore, in this study, we propose a PPO-PSO for search area decomposition to minimize the area of the grid map exceeding the search area.

The PPO-PSO for search area decomposition divides the search area into multiple triangles by randomly selecting a point within the search area. The grid map is then generated for each triangle using GDA proposed by Cho et al. [12]. The optimal particle is searched using PPO-PSO as an internal point where the grid map minimizes the region

exceeding the search area. The following is a detailed description using the notations in Table 2.

As shown in Fig. 2, given a search area  $P$  with vertices  $V = \{v_1, v_2, v_3, v_4, v_5\}$ , three particles  $X = \{x_1, x_2, x_3\}$  are randomly placed within the search area. Each particle is then connected to the vertices of the search area  $P$ , forming triangles  $T = \{t_1, t_2, t_3, t_4, t_5\}$  that divide the search area.

GDA is applied to each triangle. It creates a grid map that minimizes the number of nodes by generating the grid map based on each edge. For example, in the case of  $t_1$  in Fig. 2, GDA generates a grid map based on the edge from  $v_1$  to  $v_2$ , as shown in Fig. 3.

The PPO-PSO for search area decomposition evaluates the grid maps for each particle using Eq. (1). This equation identifies a particle  $x^*$  where the region  $A(x)$ , representing the area where the grid map exceeds the search area, is minimal. The particle  $x^*$  corresponds to  $l_j$ , which is determined in a single iteration and is used to update

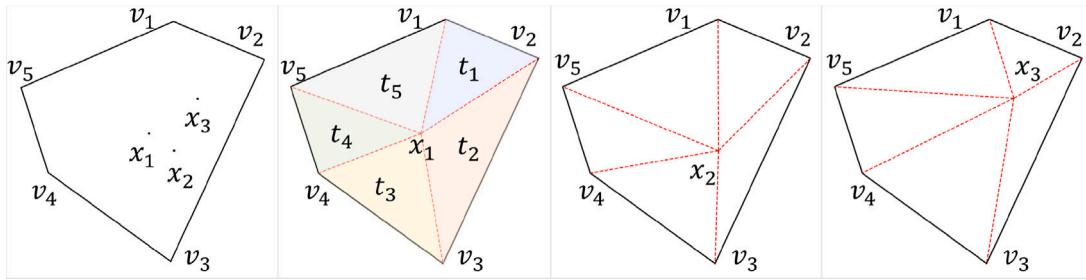


Fig. 2. Generation of triangles to divide the search area.

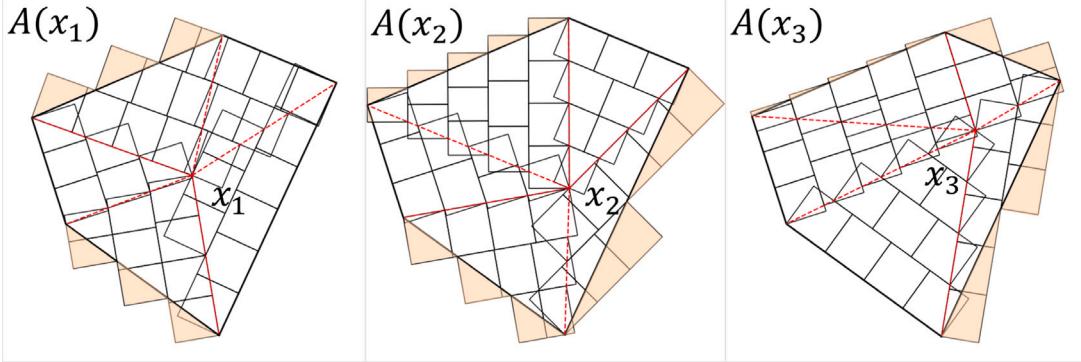


Fig. 3. Generation of grid map from each particle.

**Table 2**  
Notation of PPO-PSO for search area decomposition.

Symbol	Description
$P$	Given search area
$V = \{v_1, v_2, \dots, v_i\}$	Vertices defining the search area
$X = \{x_1, x_2, \dots, x_j\}$	Particles within the search area
$T = \{t_1, t_2, \dots, t_k\}$	Triangles that result from partitioning the search area
$A(x)$	The area where the grid map from $x$ exceeds the search area
$x^*$	Optimal dividing point within the search area
$U = \{u_1, u_2, \dots, u_j\}$	Velocity of Particles within the search area
$L = \{l_1, l_2, \dots, l_j\}$	Local best position of Particles within the search area
$g$	Global best position of Particles within the search area
$w$	Inertia weight
$c_1$	Cognitive learning rate
$c_2$	Social learning rate
$r_1, r_2$	Random numbers in the range [0,1]
$N$	Iteration number
$\pi$	Policy network
$s_i$	Current state of particle $i$
$a_i$	Action generated by policy network for particle $i$
$d_k$	Swarm diversity measure at iteration $k$
$\delta$	Convergence threshold for diversity measure
$\tau$	Maximum number of consecutive stagnation iterations

the velocities and positions of the particles. The PPO-PSO for search area decomposition iteratively adjusts particle positions, converging to a particle position  $g$  that yields the optimal grid map for the search area.

$$x^* = \arg \min_{x \in X} A(x) \quad (1)$$

The proposed PPO-PSO integrates policy-driven actions derived from the Proximal Policy Optimization (PPO) framework into the Particle Swarm Optimization (PSO) process. At each iteration, the PPO policy network observes the current state of the swarm, represented as a pixel-based image of the search area. In this image, particles are visualized as discrete points within the spatial domain. This visual representation captures the spatial distribution and clustering patterns of the particles, providing the PPO agent with comprehensive

situational awareness of the swarm's current configuration. Based on this image-based observation, the PPO policy network generates guidance actions that direct each particle toward the most appropriate partition within the search area, enhancing adaptability to irregularly shaped environments.

The PPO network architecture consists of two separate components: an Actor and a Critic, each incorporating a dedicated convolutional neural network (CNN) for feature extraction. Each CNN comprises five convolutional layers interleaved with max-pooling operations, followed by six fully connected layers. The Actor network outputs two-dimensional movement vectors via a  $\tanh$  activation function, enabling bounded continuous actions, while the Critic network produces scalar value estimates through a linear activation function. The overall architecture of the PPO network is illustrated in Fig. 4.

The proposed PPO-PSO yields a time complexity of  $O(T \times n \times v)$ , where  $T$  represents the maximum number of iterations,  $n$  denotes the number of particles in the swarm, and  $v$  corresponds to the number of vertices defining the search area polygon.

### 3.2. Phase 2: Coverage path planning

Phase 2 involves generating optimal coverage path for multiple UAVs based on the grid-based search area generated in Phase 1. The objective is to minimize the completion time of the search mission, which can be defined as the CPP problem. We utilize MILP model, as proposed by Cho et al. [12], to obtain an exact solution. This model aims to minimize the completion time while determining the optimal paths for multiple UAVs to cover the nodes. By applying this model, we ensure that the UAVs complete the search in the shortest possible time while covering all nodes.

#### 3.2.1. Assumption

To formulate the mathematical model, we set the following assumption and use the notations in Table 3.

- All UAVs start from  $i^{start}$ , which is located outside the search area.
- The adjacent nodes of  $i^{start}$  do not overlap with other nodes.

**Algorithm 1** PPO-PSO for search area decomposition

---

```

1: Input: Search area  $P$ , Policy network  $\pi$ 
2: Set initial position  $X$  and velocity  $U$ 
3: Initialize  $l_i = x_i$  and  $g$  to the best position found so far
4: for iteration = 1 to  $N$  do
5:   for each particle  $i$  do
6:     Generate triangulation  $T$  for position  $x_i$ 
7:     for each triangle  $t_k \in T$  do
8:       Apply GDA to each edge of triangle  $t_k$ 
9:       Generate grid map with minimum number of nodes
10:    end for
11:     $s_i \leftarrow$  Get current particles state
12:     $a_i \leftarrow \pi(s_i)$ 
13:     $u_i \leftarrow w_k u_i + c_1 r_1 (l_i - x_i) + c_2 r_2 (g - x_i) + a_i$ 
14:     $x_i \leftarrow x_i + u_i$ 
15:    Combine grid maps of all triangles to form complete grid for
 $x_i$ 
16:    Calculate  $A(x_i)$  as area where combined grid exceeds  $P$ 
17:    if  $A(x_i) < A(l_i)$  then
18:       $l_i \leftarrow x_i$ 
19:    end if
20:    if  $A(l_i) < A(g)$  then
21:       $g \leftarrow l_i$ 
22:    end if
23:  end for
24:  if  $d_k < \delta$  for  $\tau$  consecutive iterations then
25:    Break
26:  end if
27: end for
28: Return Best position  $g$ 

```

---

- Each UAV moves with fixed values of cruising and rotational speeds.

**3.2.2. Mathematical model**

$$\min z = \max_{\{m\}} t_m \quad (2)$$

s.t.

$$\sum_{i \in N} \sum_{j \in N} \sum_{k \in N} x_{ijkms} = 1, \quad \text{for } m \in U, s \in S \quad (3)$$

$$\sum_{j \in N} \sum_{k \in N} x_{ijkms} = 1, \quad \text{for } i = i^{start}, m \in U, s \in 1 \quad (4)$$

$$\sum_{i \in N} \sum_{j \in N} \sum_{k \in N} x_{ijkms} = 1, \quad \text{for } m \in U, s \in S \setminus \{1\} \quad (5)$$

$$\sum_{i \in N} \sum_{l \in N} (x_{jklms} - x_{jklm(s-1)}) = 0, \quad \text{for } j, k \in N, m \in U, s \in S \setminus \{1\} \quad (6)$$

$$\sum_{i \in N} \sum_{k \in N} \sum_{m \in U} \sum_{s \in S} x_{ijkms} = 1, \quad \text{for } i \in N^{IN} \quad (7)$$

$$t_m = \sum_{i \in N} \sum_{j \in N} \sum_{k \in N} \sum_{s \in S} \frac{d_{ij}}{v_m} \times x_{ijkms} \\ + \sum_{i \in N} \sum_{j \in N} \sum_{k \in N} \sum_{s \in S} \left( \frac{d_{jk}}{v_m} + \frac{\varphi_{ijk}}{w_m} \right) \times x_{ijkms}, \quad \text{for } m \in U \quad (8)$$

$$t_m \leq e_m, \quad \text{for } m \in U \quad (9)$$

$$t_m \geq 0 \quad (10)$$

**Table 3**

Notation of mathematical model.

Symbol	Description
$i$	Index of nodes within the search area, $i \in N^{IN}$
$i^{start}$	Starting node for all UAVs
$m$	Index of UAVs, where $m \in U$
$s$	Index of sequences for nodes, where $s \in S$
$N^{IN}$	Set of nodes within the search area, indexed by $i$
$N$	Set of all nodes, defined as $N = N^{IN} \cup \{i^{start}\}$
$U$	Set of UAVs, indexed by $m$
$S$	Set of sequences for nodes, indexed by $s$
$N^{ADJ}(i)$	Set of adjacent nodes for a given node $i$
$d_{ij}$	Distance between node $i$ and node $j$
$\varphi_{ijk}$	Rotation angle for the path $i \rightarrow j \rightarrow k$
$v_m$	Maximum cruising speed of UAV $m$
$w_m$	Rotational speed of UAV $m$
$e_m$	Endurance of UAV $m$ (maximum allowable operation time)
$x_{ijkms}$	1 if UAV $m$ covers the path $i \rightarrow j \rightarrow k$ at the $s$ th sequence, otherwise 0
$t_m$	Completion time of UAV $m$

$$x_{ijkms} \in \{0, 1\} \quad (11)$$

Objective function (2) aims to minimize the maximum completion time among all UAVs, thereby reducing the overall mission duration for the UAV with the longest time. Constraint (3) defines the paths for each UAV, and constraint (4) ensures that each UAV starts its mission from the designated starting node  $i^{start}$ . Constraint (5) ensures that each UAV selects exactly one path at each sequence  $s$ . Constraint (6) ensures the paths are continuous, and constraint (7) defines that all nodes are covered by the UAVs without overlap. Constraint (8) calculates the total completion time by considering both the cruising time and the rotation time of the UAVs. Constraint (9) ensures that the completion time for each UAV does not exceed its endurance limit, i.e., the maximum energy that each UAV can consume. Constraint (10) defines a non-negativity condition on the completion time, and constraint (11) defines that the decision variables are binary.

**4. Numerical experiments**

To evaluate and validate the effectiveness of our proposed methodology, we conducted comprehensive numerical experiments that demonstrate its performance in terms of computational efficiency and solution quality.

**4.1. Experimental setup**

In the experiments, two types of UAVs were used at a constant altitude during the mission, as shown in Table 4. The performance of the proposed methodology was evaluated using 15 test cases, with illustrations provided in Fig. 5. Test cases 1 and 2 were derived from Cho et al. [12]. Test cases 3 to 6 were created based on real marine environments, with special focus on the shape of the coastline in a marine area located on the west coast of South Korea, as shown in Fig. 6. Test cases 7 to 15 were randomly generated, with the former representing convex areas and the latter representing irregular areas.

The experiments are performed on a PC using an Intel Core i9-14900k 3.2 GHz processor and 32.0 GB of RAM. We run the commercial solver Gurobi 11.0.2 to solve the MILP model. All of the algorithms are programmed using Python.

**4.2. Sensitivity analysis**

In this section, we analyze the impact of key PPO hyperparameters — learning rate ( $\alpha$ ), discount factor ( $\gamma$ ), clipping range ( $\epsilon$ ), and action standard deviation ( $\sigma$ ), as defined in Section 3.1 — on the optimization performance of the proposed PPO-PSO framework. The baseline PSO parameters ( $X = 40$ ,  $N = 200$ ,  $w = 0.4$ ,  $c_1 = 0.8$ ,

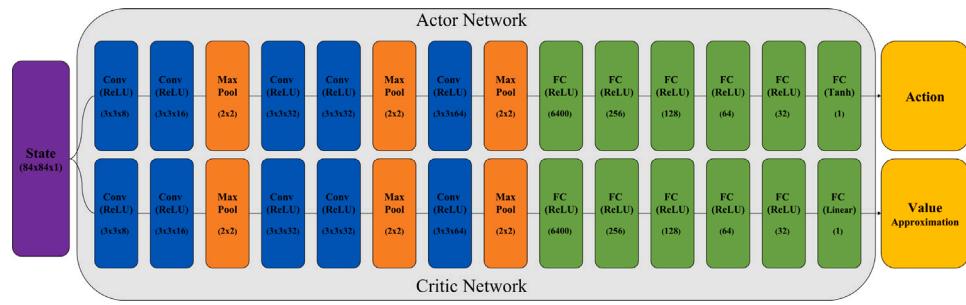


Fig. 4. PPO network architecture.

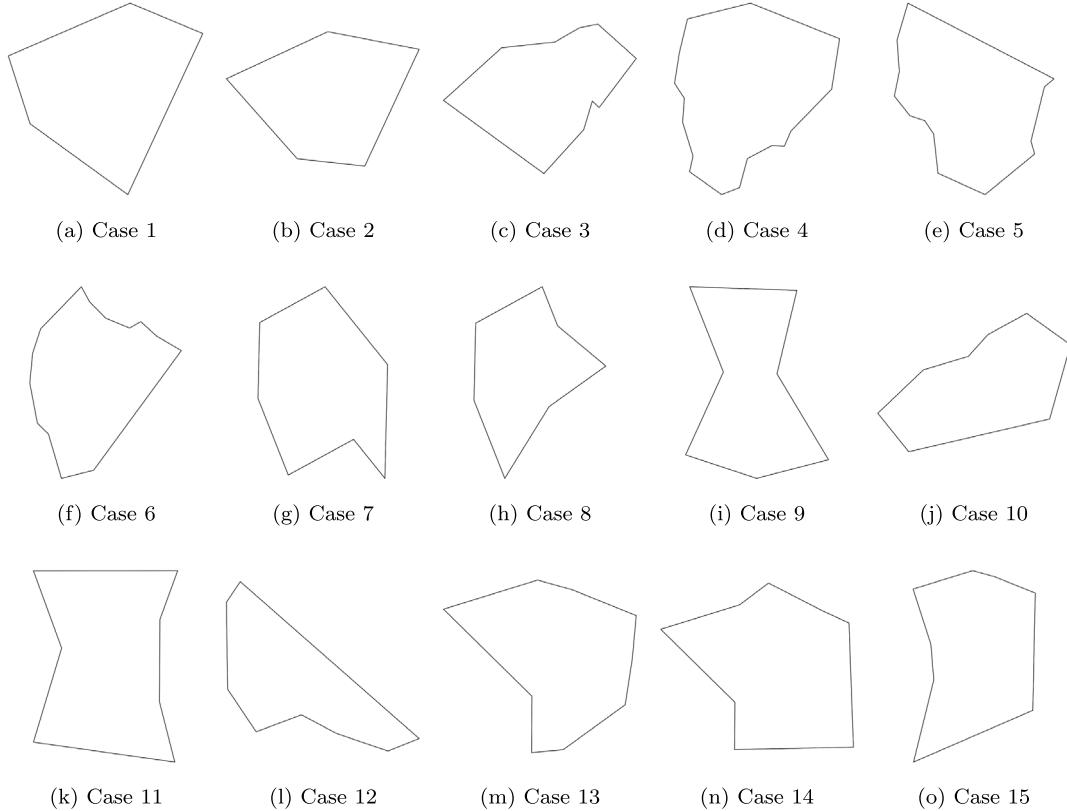


Fig. 5. Illustration of test cases.



Fig. 6. Test cases from Google Earth satellite imagery.

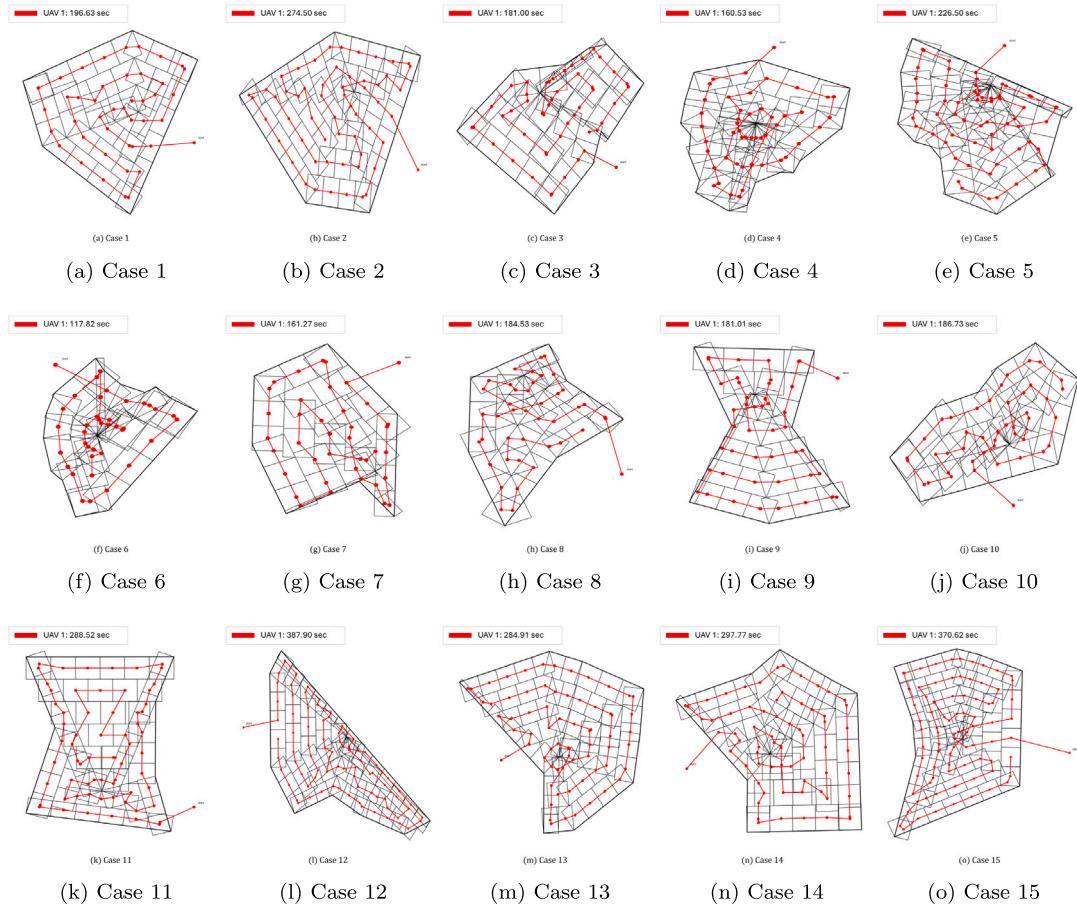


Fig. 7. Solutions for the test cases using a single UAV.

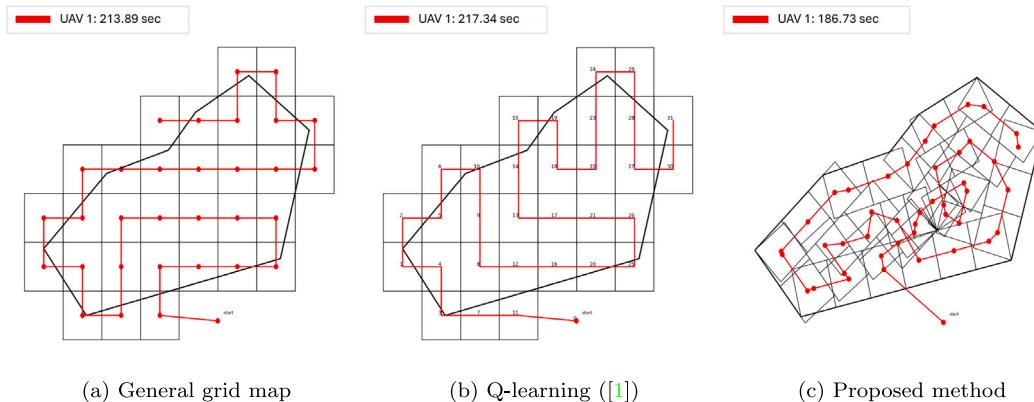


Fig. 8. Coverage path for Case 10.

**Table 4**  
Performance of UAVs.

UAV type	Velocity (m/s)	Angular velocity (rad/s)	Maximum flight time (min)	Camera FOV (deg)	Camera footprint size
1	4	0.5	30	100	20 m × 20 m
2	4	1.0	30	100	20 m × 20 m

$c_2 = 0.9$ ) were selected based on preliminary experiments. For each combination of PPO hyperparameters shown in Table 5, individual models were trained for 500 episodes.

Table 6 presents the actor and critic losses obtained from PPO-PSO models trained under various combinations of hyperparameters. While the learning rate  $\alpha$  was initially tested with values of 0.0001, 0.00025,

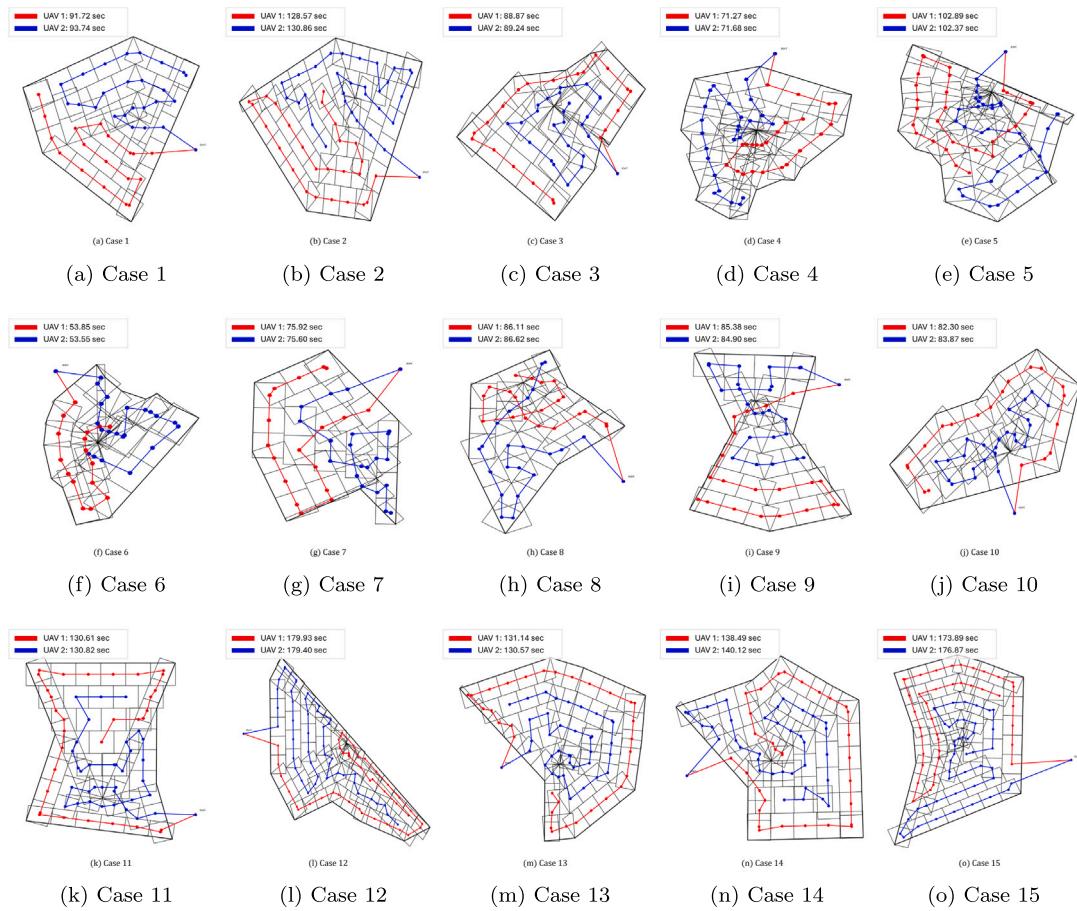


Fig. 9. Solutions for the test cases using two UAVs.

Table 5

Experimental settings for PPO hyperparameter tuning.

Hyperparameter	Experimental range
$\alpha$	0.0001, 0.00025, 0.0005
$\gamma$	0.9, 0.95, 0.99
$\epsilon$	0.1, 0.2, 0.3
$\sigma$	0.2, 0.3

and 0.0005, the models with  $\alpha$  set to 0.0001 and 0.0005 exhibited unstable behavior, resulting in diverging loss values. Therefore, the results reported in this table correspond exclusively to the stable case with  $\alpha = 0.00025$ .

The setting with  $\gamma = 0.9$ ,  $\epsilon = 0.1$ , and  $\sigma = 0.2$  achieved the best overall performance, yielding the lowest actor loss of  $-0.000783$  and a comparably low critic loss of  $0.010210$ . While another configuration with  $\gamma = 0.9$ ,  $\epsilon = 0.2$ , and  $\sigma = 0.2$  produced the lowest critic loss of  $0.010933$ , its actor loss was relatively less stable. These findings suggest that the selected optimal setting effectively balances policy learning and value estimation, contributing to the stable and efficient performance of the proposed PPO-PSO. Accordingly, this setting was adopted for all subsequent experiments.

#### 4.3. Performance comparison with existing algorithm using a single UAV

In this section, we present experimental results to verify the performance of the PPO-PSO for search area decomposition. We evaluate the performance of the proposed method using a single Type 1 UAV through a comparative analysis with the general grid map, the GDA method proposed by Cho et al. [12] and Q-learning approach proposed by Ai et al. [37].

Table 6

Comparison of PPO-PSO performance across different hyperparameter settings.

$\alpha$	$\gamma$	$\epsilon$	$\sigma$	Actor loss	Critic loss
0.00025	0.9	0.1	0.2	0.001858	0.015302
0.00025	0.9	0.2	0.2	-0.000493	<b>0.010933</b>
0.00025	0.9	0.1	0.2	<b>-0.000783</b>	0.010210
0.00025	0.99	0.2	0.2	-0.004995	0.057640
0.00025	0.95	0.2	0.1	0.000949	0.032084
0.00025	0.99	0.1	0.2	-0.003669	0.038402

Table 7 summarizes the experimental results, and Fig. 7 illustrates the coverage path generated by the proposed method. The “Phase 1” column indicates whether the corresponding method was able to generate a valid grid map, and the  $Z$  values represent the search completion time for each method. The GDA method fails to construct grid maps for irregular search areas, leading to unsuccessful planning in Cases 3 to 15. Although it produces slightly better  $Z$  values than our method in the regular-shaped Cases 1 and 2, its limited applicability restricts its usefulness. In contrast, the proposed PPO-PSO approach successfully generates feasible grid maps in all 15 cases, demonstrating robustness and generalizability across both regular and irregular search areas.

The Q-learning method consistently completes Phase 1 and produces feasible paths, but in most cases, the resulting  $Z$  values are higher than those of the proposed method, indicating less efficient search trajectories. For example, in Case 12, the proposed method achieves a  $Z_{PPO-PSO}$  of 387.9, which is notably better than the  $Z_{Q\text{-learning}}$  value of 437.15.

The general grid map method also completes Phase 1 for all cases. However, as shown in Fig. 8, it tends to generate unnecessarily large

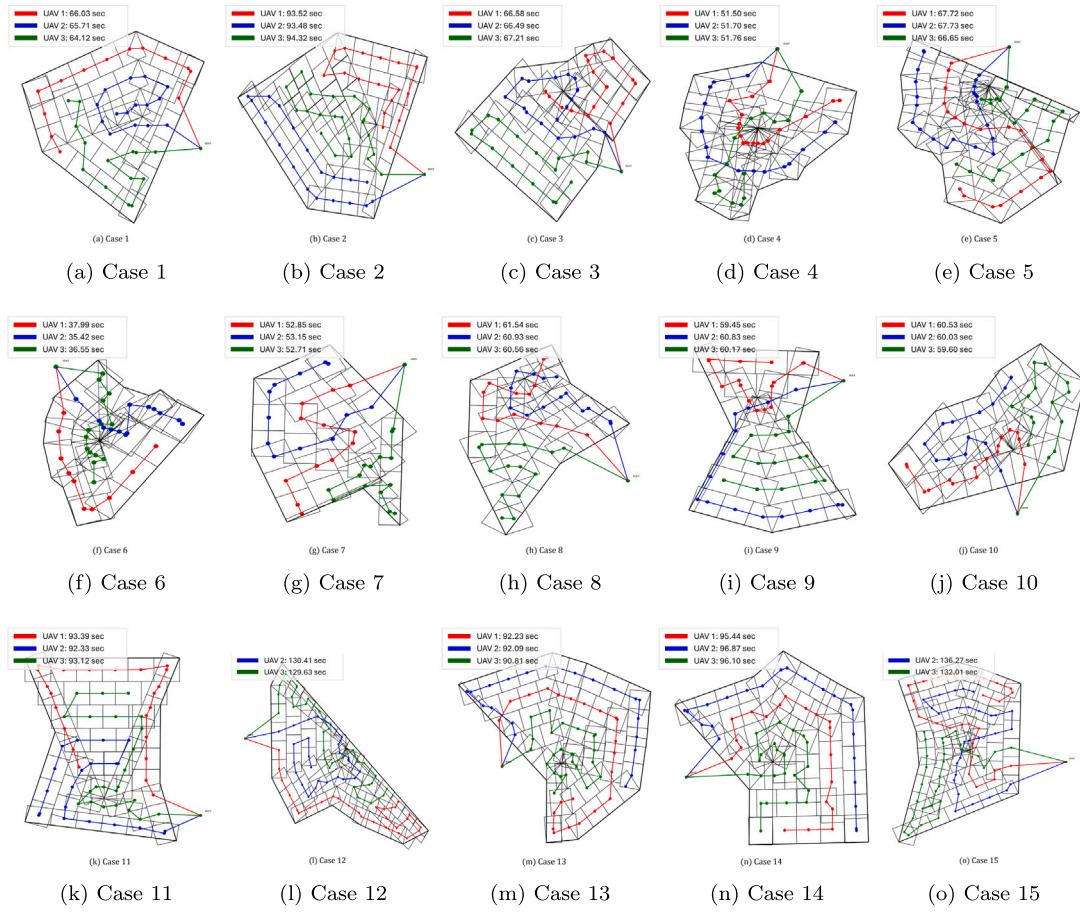


Fig. 10. Solutions for the test cases using multiple UAVs.

**Table 7**  
Experimental results of the test cases using a single UAV.

Case no.	General grid map		GDA [12]		Q-learning [37]		Our proposed method	
	Phase 1	Z	Phase 1	$Z_{GDA}$	Phase 1	$Z_{Q\text{-learning}}$	Phase 1	$Z_{PPO\text{-PSO}}$
1	O	246.81	O	189.58	O	263.71	O	196.63
2	O	-	O	262.29	O	355.3	O	274.5
3	O	252.95	X	-	O	237.88	O	181.0
4	O	169.65	X	-	O	173.37	O	160.53
5	O	228.8	X	-	O	223.09	O	226.5
6	O	120.98	X	-	O	113.73	O	117.82
7	O	164.57	X	-	O	176.88	O	161.27
8	O	-	X	-	O	234.15	O	184.53
9	O	-	X	-	O	208.15	O	181.01
10	O	213.89	X	-	O	217.34	O	186.73
11	O	275.19	X	-	O	270.98	O	288.52
12	O	-	X	-	O	437.15	O	387.9
13	O	-	X	-	O	331.09	O	284.91
14	O	-	X	-	O	333.39	O	297.77
15	O	378.58	X	-	O	396.14	O	370.62

Phase 1: Whether or not grid maps are generated in Phase 1.

grid maps, resulting in suboptimal paths. In some cases such as Cases 2, 8, 9, 12–14, this method fails to output a Z value, implying it was unable to produce a fully feasible or efficient plan despite completing Phase 1.

These results show that the proposed PPO-PSO approach balances feasibility and performance, providing a reliable decomposition strategy for search missions involving irregular search areas.

#### 4.4. Performance comparison with existing algorithm using multiple UAVs

To evaluate the scalability of the proposed method, we conducted comparative experiments using multiple UAVs. The same 15 test cases

from the single UAV experiments were applied, with two UAVs (Type 1 and Type 2) operating simultaneously to cover the search area.

As shown in Table 8, the proposed method consistently produced feasible solutions for all test cases, including irregularly shaped search areas where the GDA method failed to generate grid maps (Cases 3 to 15). While the general grid map was able to generate maps for all cases, it failed to provide feasible solutions in several cases. Additionally, in most cases, the PPO-PSO method achieved better objective values than the general grid map, indicating more efficient area coverage and shorter mission completion times.

**Table 8**  
Experimental results of the test cases using multiple UAVs.

Case no.	General grid map		GDA [12]		Our proposed method	
	Phase 1	Z	Phase 1	$Z_{GDA}$	Phase 1	$Z_{PPO-PSO}$
1	O	112.83	O	91.15	O	93.74
2	O	164.48	O	127.94	O	130.86
3	O	104.96	X	—	O	89.24
4	O	78.46	X	—	O	71.68
5	O	101.63	X	—	O	102.89
6	O	50.27	X	—	O	53.85
7	O	79.23	X	—	O	75.92
8	O	—	X	—	O	86.62
9	O	99.06	X	—	O	85.38
10	O	101.71	X	—	O	83.87
11	O	124.12	X	—	O	130.82
12	O	—	X	—	O	179.93
13	O	—	X	—	O	131.14
14	O	—	X	—	O	140.12
15	O	185.9	X	—	O	176.87

Phase 1: Whether or not grid maps are generated in Phase 1.

Overall, these results demonstrate that the proposed PPO-PSO method scales effectively to multiple UAVs and maintains robust and efficient performance across diverse search environments.

Fig. 9 illustrates the coverage paths generated by the proposed method for all test cases using two UAVs.

#### 4.5. Performance according to the number of UAVs

In this section, we evaluate the performance of the proposed method when extended to three UAVs, designating UAVs 1 and 2 as Type 1 and UAV 3 as Type 2. The proposed method generated coverage paths for all three UAVs, as illustrated in Fig. 10. These results demonstrate the capability of the proposed method to handle heterogeneous UAV types and complex coverage scenarios, ensuring efficient multi-agent coordination.

## 5. Conclusion

In this study, we proposed a novel multi-UAV CPP framework for the efficient coverage of irregularly shaped search areas, consisting of two main phases. In the first phase, a PPO-PSO for search area decomposition was developed to minimize the excess grid map area for irregular search regions. In the second phase, an MILP model was formulated to minimize the total search completion time. The effectiveness of the proposed methodology was evaluated through 15 experimental scenarios, including cases derived from real marine environments. The results demonstrated that the proposed framework outperforms existing methods in addressing irregularly shaped search areas.

This study provides a foundation for several potential future research directions. First, it is anticipated that further studies could develop methodologies that account for dynamic factors in marine environments, such as ocean currents and wave patterns. Second, research incorporating the prioritization of search areas to enhance the rapid detection of individuals in distress would be highly valuable. Finally, this framework could be extended to investigate cooperative operations between UAVs and Unmanned Surface Vehicles (USVs) in SAR missions, thereby enhancing operational efficiency and effectiveness.

## CRediT authorship contribution statement

**Seung Chan Choi:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Yohan Lee:** Writing – original draft, Investigation, Formal analysis. **Sung Won Cho:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

## Declaration of competing interest

The authors declare no conflict of interest.

## Acknowledgments

The authors are grateful to the anonymous reviewers for reading the manuscript carefully and providing constructive comments which greatly helped to improve this paper. This research was supported by Korea Institute of Marine Science & Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries, Korea (RS-2024-00407024, International Joint R&D and Collaborative Demonstration of VDES).

## Data availability

Data will be made available on request.

## References

- [1] Kozena Cyprian Aleksander, Military use of unmanned aerial vehicles—a historical study, *Saf. Déf.* 4 (2018) 17–21.
- [2] Praveen Kumar Reddy Maddikunta, Saqib Hakak, Mamoun Alazab, Sweta Bhattacharya, Thippa Reddy Gadekallu, Wazir Zada Khan, Quoc-Viet Pham, Unmanned aerial vehicles in smart agriculture: Applications, requirements, and challenges, *IEEE Sens. J.* 21 (16) (2021) 17608–17619.
- [3] Wouter H. Maes, Kathy Steppé, Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture, *Trends Plant Sci.* 24 (2) (2019) 152–164.
- [4] Yan Li, Chunlu Liu, Applications of multirotor drone technologies in construction management, *Int. J. Constr. Manag.* 19 (5) (2019) 401–412.
- [5] Robin Kellermann, Tobias Biehle, Liliann Fischer, Drones for parcel and passenger transportation: A literature review, *Transp. Res. Interdiscip. Perspect.* 4 (2020) 100088.
- [6] Wonhee Lee, Sung Won Cho, Reinforcement learning approach for outbound container stacking in container terminals, *Comput. Ind. Eng.* 204 (2025) 111069.
- [7] Marion Jaud, Sophie Passot, Réjanne Le Bivic, Christophe Delacourt, Philippe Grandjean, Nicolas Le Dantec, Assessing the accuracy of high resolution digital surface models computed by PhotoScan® and MicMac® in sub-optimal survey conditions, *Remote. Sens.* 8 (6) (2016) 465.
- [8] Alena Otto, Niels Agatz, James Campbell, Bruce Golden, Erwin Pesch, Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey, *Networks* 72 (4) (2018) 411–458.
- [9] Agoston Restas, Drone applications for supporting disaster management, *World J. Eng. Technol.* 3 (3) (2015) 316–321.
- [10] Abhilasha Nanda, Sung Won Cho, Hyeopwoo Lee, Jin Hyoung Park, KOLOM-VERSE: Korea open large-scale image dataset for object detection in the maritime universe, *IEEE Trans. Intell. Transp. Syst.* 25 (12) (2024) 20832–20840.
- [11] Håkon Hagen Helgesen, Frederik Stendahl Leira, Torleiv H Bryne, Sigurd M Albrektsen, Tor Arne Johansen, Real-time georeferencing of thermal images using small fixed-wing UAVs in maritime environments, *ISPRS J. Photogramm. Remote Sens.* 154 (2019) 84–97.

- [12] Sung Won Cho, Hyun Ji Park, Hanseob Lee, David Hyunchul Shim, Sun-Young Kim, Coverage path planning for multiple unmanned aerial vehicles in maritime search and rescue operations, *Comput. Ind. Eng.* 161 (2021) 107612.
- [13] Enric Galceran, Marc Carreras, A survey on coverage path planning for robotics, *Robot. Auton. Syst.* 61 (12) (2013) 1258–1276.
- [14] Krishan Kumar, Neeraj Kumar, Region coverage-aware path planning for unmanned aerial vehicles: A systematic review, *Phys. Commun.* 59 (2023) 102073.
- [15] Tauá M Cabreira, Lisane B Brisolara, Ferreira Jr Paulo R, Survey on coverage path planning with unmanned aerial vehicles, *Drones* 3 (1) (2019) 4.
- [16] Georgios Fevgas, Thomas Lagkas, Vasileios Argyriou, Panagiotis Sarigiannidis, Coverage path planning methods focusing on energy efficient and cooperative strategies for unmanned aerial vehicles, *Sensors* 22 (3) (2022) 1235.
- [17] Joao Valente, David Sanz, Jaime Del Cerro, Antonio Barrientos, Miguel Ángel de Frutos, Near-optimal coverage trajectories for image mosaicing using a mini quad-rotor over irregular-shaped fields, *Precis. Agric.* 14 (2013) 115–132.
- [18] João Valente, Jaime Del Cerro, Antonio Barrientos, David Sanz, Aerial coverage optimization in precision agriculture management: A musical harmony inspired approach, *Comput. Electron. Agric.* 99 (2013) 153–159.
- [19] Arman Nedjati, Gokhan Izbirak, Bela Vizvari, Jamal Arkat, Complete coverage path planning for a multi-UAV response system in post-earthquake assessment, *Robotics* 5 (4) (2016) 26.
- [20] Jalil Modares, Farshad Ghanei, Nicholas Mastronarde, Karthik Dantu, Ub-anc planner: Energy efficient coverage path planning with multiple drones, in: 2017 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2017, pp. 6182–6189.
- [21] Younghoon Choi, Youngjun Choi, Simon Briceno, Dimitri N Mavris, Coverage path planning for a UAS imagery mission using column generation with a turn penalty, in: 2018 International Conference on Unmanned Aircraft Systems, ICUAS, IEEE, 2018, pp. 1109–1117.
- [22] Teng Li, Chaoqun Wang, Clarence W. de Silva, et al., Coverage sampling planner for uav-enabled environmental exploration and field mapping, in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2019, pp. 2509–2516.
- [23] Wassim Khiati, Younes Moumen, Ali El Habchi, Ilham Zerrouk, Jamal Berrich, Toumi Bouchentouf, Grid based approach (GBA): A new approach based on the grid-clustering algorithm to solve a CPP type problem for air surveillance using UAVs, in: 2020 Fourth International Conference on Intelligent Computing in Data Sciences, ICDS, IEEE, 2020, pp. 1–5.
- [24] Andrés Pérez-González, Nelson Benítez-Montoya, Álvaro Jaramillo-Duque, Juan Bernardo Cano-Quintero, Coverage path planning with semantic segmentation for UAV in PV plants, *Appl. Sci.* 11 (24) (2021) 12093.
- [25] Kai Zhu, Bin Han, Tao Zhang, Multi-UAV distributed collaborative coverage for target search using heuristic strategy, *Guid. Navig. Control.* 1 (01) (2021) 2150002.
- [26] Javad Heydari, Olimpiya Saha, Viswanath Ganapathy, Reinforcement learning-based coverage path planning with implicit cellular decomposition, 2021, arXiv preprint [arXiv:2110.09018](https://arxiv.org/abs/2110.09018).
- [27] Nikolaos A Kyriakakis, Magdalene Marinaki, Nikolaos Matsatsinis, Yannis Marinakis, A cumulative unmanned aerial vehicle routing problem approach for humanitarian coverage path planning, *European J. Oper. Res.* 300 (3) (2022) 992–1004.
- [28] Yanxi Huang, Jiankang Xu, Mengting Shi, Liang Liu, Time-efficient coverage path planning for energy-constrained UAV, *Wirel. Commun. Mob. Comput.* 2022 (1) (2022) 5905809.
- [29] Wenjian Hu, Yao Yu, Shumei Liu, Changyang She, Lei Guo, Branka Vucetic, Yonghui Li, Multi-UAV coverage path planning: A distributed online cooperation method, *IEEE Trans. Veh. Technol.* 72 (9) (2023) 11727–11740.
- [30] Yang Yu, Sanghwan Lee, Multi-UAV coverage path assignment algorithm considering flight time and energy consumption, *IEEE Access* 12 (2024) 26150–26162.
- [31] Jie Wu, Liang Cheng, Sensen Chu, Yanjie Song, An autonomous coverage path planning algorithm for maritime search and rescue of persons-in-water based on deep reinforcement learning, *Ocean Eng.* 291 (2024) 116403.
- [32] Tauá M Cabreira, Paulo R Ferreira, Carmelo Di Franco, Giorgio C Buttazzo, Grid-based coverage path planning with minimum energy over irregular-shaped areas with UAVs, in: 2019 International Conference on Unmanned Aircraft Systems, ICUAS, IEEE, 2019, pp. 758–767.
- [33] Fang Ma, Feng Xiong, Research on path planning of plant protection UAV based on grid method and improved ant colony algorithm, in: IOP Conference Series: Materials Science and Engineering, 612, IOP Publishing, 2019.
- [34] Alia Ghaddar, Ahmad Merei, Enrico Natalizio, PPS: Energy-aware grid-based coverage path planning for UAVs using area partitioning in the presence of NFZs, *Sensors* 20 (13) (2020) 3742.
- [35] Savvas D Apostolidis, Pavlos Ch Kapoutsis, Athanasios Ch Kapoutsis, Elias B Kosmatopoulos, Cooperative multi-UAV coverage mission planning platform for remote sensing applications, *Auton. Robots* 46 (2) (2022) 373–400.
- [36] Yiguang Gong, Kai Chen, Tianyu Niu, Yunping Liu, Grid-based coverage path planning with NFZ avoidance for UAV using parallel self-adaptive ant colony optimization algorithm in cloud IoT, *J. Cloud Comput.* 11 (1) (2022) 29.
- [37] Bo Ai, Maoxin Jia, Hanwen Xu, Jiangling Xu, Zhen Wen, Benshuai Li, Dan Zhang, Coverage path planning for maritime search and rescue using reinforcement learning, *Ocean Eng.* 241 (2021) 110098.