

Article

Not peer-reviewed version

CD-HSSRL: Cross-Domain Hierarchical Safe Switching Reinforcement Learning Framework for Autonomous Amphibious Robot Navigation

[Shuang Liu](#), Lei Wei, [Xiaoqing Li](#) *

Posted Date: 7 April 2026

doi: 10.20944/preprints202604.0288.v1

Keywords: tracked amphibious robots; cross-domain navigation; reinforcement learning; hierarchical safe switching; safety-constrained control




Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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

Article

CD-HSSRL: Cross-Domain Hierarchical Safe Switching Reinforcement Learning Framework for Autonomous Amphibious Robot Navigation

Shuang Liu , Lei Wei and Xiaoqing Li *

School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China

* Correspondence: xqli@hust.edu.cn

Abstract

Autonomous tracked amphibious robotic systems operating across water and land environments are essential for coastal inspection, disaster response, environmental monitoring, and complex terrain exploration. However, discontinuous water-land dynamics, unstable medium switching, and safety-critical control under environmental uncertainty pose significant challenges to existing amphibious navigation and path planning methods, where global reachability and adaptive decision-making are difficult to unify. Motivated by these challenges, this paper proposes CD-HSSRL, a Cross-Domain Hierarchical Safe-Switching Reinforcement Learning framework for autonomous tracked amphibious navigation. Specifically, a Cross-Domain Global Reachability Planner is developed to construct unified cost representations across heterogeneous water-land environments, a Hierarchical Safe Switching Policy enables stable medium-transition decision-making through option-based policy decomposition with switching regularization, and a Safety-Constrained Continuous Controller integrates action safety projection and risk-sensitive reward shaping to ensure collision-free control during complex shoreline interactions. These components are jointly optimized in an end-to-end manner to achieve robust cross-domain navigation. Comprehensive experiments on WaterScenes, MVTD, BARN, and Gazebo cross-domain benchmarks demonstrate that CD-HSSRL consistently outperforms state-of-the-art baselines, achieving up to 15% improvement in cross-domain transition success rate and 40% reduction in collision rate. Robustness and ablation studies further verify the effectiveness of hierarchical switching and safety-constrained control mechanisms. Overall, this work establishes a unified solution for safe and reliable cross-domain navigation of tracked amphibious robotic systems, providing new insights into hierarchical safe-switching architectures for multi-medium autonomous robots.

Keywords: tracked amphibious robots; cross-domain navigation; reinforcement learning; hierarchical safe switching; safety-constrained control

1. Introduction

Autonomous tracked amphibious robotic systems capable of operating seamlessly across water and land environments play an increasingly important role in coastal inspection, environmental monitoring, disaster rescue, and maritime transportation applications [1–4]. Compared with single-medium robotic systems, amphibious platforms provide superior mission flexibility and accessibility in complex terrains where water and land coexist [5]. However, enabling robots to autonomously navigate across heterogeneous environments remains a fundamental challenge [6], as water-land transitions involve discontinuous dynamics [7], rapidly changing environmental constraints [8], and safety-critical interactions with uncertain surroundings [9].

Recent advances in learning-based robotic navigation have demonstrated remarkable success in single-domain path planning and obstacle avoidance for unmanned surface vehicles and ground

robots [10,11]. Nevertheless, most existing approaches are designed for either water or land environments independently [12,13], and their policies often fail when directly transferred across domains due to inconsistent state representations, abrupt medium switching, and unmodeled physical constraints [14,15]. Consequently, current methods suffer from unstable transition decisions near shorelines, oscillatory behaviors during medium switching, and elevated collision risks, which significantly limit real-world deployment of amphibious robotic systems.

To address these limitations, this study investigates the following research question: *How can an autonomous robot achieve safe, stable, and efficient navigation across discontinuous water-land environments under environmental uncertainty?* We hypothesize that explicitly modeling cross-domain reachability, hierarchical switching decisions, and safety-constrained control is essential to achieve reliable amphibious navigation.

The objective of this work is to develop a unified framework that integrates global cross-domain planning, medium-switching decision-making, and safety-aware continuous control into a coherent joint optimization scheme. However, solving this problem involves several critical challenges. First, water and land environments exhibit fundamentally different dynamic constraints, making it difficult to construct unified environmental representations for global planning. Second, naive policy structures struggle to produce stable medium-switching decisions near boundary regions, leading to frequent oscillations and control instability. Third, safety-critical constraints during shoreline interaction require explicit collision and grounding avoidance mechanisms beyond standard learning formulations.

Motivated by these challenges, we propose CD-HSSRL, a Cross-Domain Hierarchical Safe-Switching Reinforcement Learning framework for autonomous amphibious navigation. The proposed framework introduces a Cross-Domain Global Reachability Planner to construct unified cost-aware environmental representations, enabling consistent long-horizon planning across water and land. A Hierarchical Safe Switching Policy is designed to decompose navigation into high-level medium-switching decisions and low-level motion control, enforcing switching stability through regularized option learning. Furthermore, a Safety-Constrained Continuous Controller integrates action safety projection and risk-sensitive reward shaping to guarantee collision-free and stable control during complex water-land transitions. These modules are jointly optimized in an end-to-end manner to achieve unified planning-switching-safety co-optimization for robust cross-domain navigation.

The main contributions of this paper are summarized as follows: (1) We propose a novel cross-domain hierarchical safe-switching reinforcement learning framework that unifies water-land navigation, medium-switching decision-making, and safety-critical control into a single end-to-end optimized architecture. (2) We develop a cross-domain global reachability planner and a hierarchical safe switching policy that enable stable and reliable amphibious navigation under discontinuous environmental dynamics. (3) We design a safety-constrained continuous controller that explicitly enforces physical safety constraints during shoreline interaction. (4) Extensive experiments on multiple water-domain, land-domain, and cross-domain benchmarks demonstrate that the proposed method consistently outperforms state-of-the-art baselines in navigation success rate, transition stability, and collision avoidance performance.

2. Related Work

2.1. Amphibious and Cross-Domain Robot Navigation

Amphibious and cross-domain robotic systems have emerged as an important research topic due to their ability to operate in heterogeneous environments where water and land coexist [16,17]. Typical application scenarios include coastal inspection, flood rescue [18], ecological monitoring, and military reconnaissance [19]. Compared with single-medium robotic platforms, amphibious robots offer greater mission flexibility but also face fundamentally different environmental constraints when transitioning between water and land. Early studies on amphibious navigation mainly relied on model-based planning frameworks [20], such as graph search, sampling-based path planning, and optimization-based trajectory generation. These approaches usually construct separate environmental

models for water and land and design heuristic cost functions to guide navigation. While such methods achieve acceptable performance in structured environments, they heavily depend on accurate dynamic modeling and manually designed traversability maps, which limits their adaptability in complex and uncertain real-world scenes.

More recent research has introduced learning-based strategies to enhance amphibious navigation performance [21]. Learning-based planners can capture nonlinear hydrodynamic effects and terrain interactions more effectively than traditional analytical models [22]. However, most existing works still treat water-domain and land-domain planning as two loosely coupled processes, and medium transitions are often handled by handcrafted switching logic or predefined thresholds. This separation leads to inconsistent decision-making near shoreline boundaries, unstable transitions, and reduced robustness under environmental disturbances. Therefore, a unified planning and decision-making framework capable of representing heterogeneous environments and enabling smooth cross-domain transitions remains an open research challenge.

2.2. Reinforcement Learning for Water-Domain and Land-Domain Navigation

Reinforcement learning has become a powerful tool for autonomous navigation in both water and land environments [23,24]. In water-domain navigation, deep reinforcement learning has been widely applied to amphibious and surface robotic platforms for dynamic obstacle avoidance, collision-free control, and energy-efficient route planning. These methods benefit from end-to-end policy learning and can adapt to complex maritime environments without explicit hydrodynamic modeling. Similarly, in land-domain navigation, reinforcement learning has demonstrated strong capabilities in mobile robot path planning, multi-agent coordination, and navigation in dynamic scenes. Such approaches allow robots to learn reactive and anticipatory behaviors directly from environmental interactions.

Despite these advances, most existing reinforcement learning methods are developed and trained in a single domain. Policies learned in water environments usually fail when directly transferred to land environments, and vice versa, due to inconsistent observation distributions, abrupt changes in motion dynamics, and different safety constraints. As a result, existing single-domain reinforcement learning frameworks struggle to generalize across heterogeneous environments and cannot guarantee stable decision-making during water-land transitions. This limitation highlights the necessity of developing cross-domain reinforcement learning frameworks that explicitly model medium-dependent dynamics and enable knowledge sharing between water and land navigation policies.

2.3. Hierarchical Reinforcement Learning and Medium-Switching Decision Making

Hierarchical reinforcement learning decomposes complex decision-making tasks into multiple temporal or functional layers [25], typically including a high-level planner and a low-level controller. Such hierarchical structures improve learning efficiency, interpretability, and scalability, especially in long-horizon robotic navigation problems. Option-based frameworks further enable the learning of temporally extended actions, allowing robots to switch between different behavioral modes based on environmental contexts.

Hierarchical reinforcement learning has been successfully applied to task decomposition, navigation subtasks, and skill sequencing in robotics. However, existing hierarchical approaches primarily focus on abstract task or goal decomposition and seldom consider physical medium-switching in real-world robotic systems. In amphibious navigation, medium-switching decisions correspond to physically distinct motion regimes, such as floating, shoreline climbing, and ground driving. Without explicit switching stability modeling, hierarchical policies often generate oscillatory decisions near transition regions, resulting in inefficient control and increased risk of grounding or collision. This reveals the need for a hierarchical reinforcement learning framework specifically designed to handle physical medium transitions and enforce stable switching behavior in cross-domain navigation.

Recent hierarchical navigation frameworks such as predictive hierarchical deep reinforcement learning (pH-DRL [26]) and motion-primitive-based deep Q-learning (MP-DQL [27]) further demonstrate the effectiveness of long-horizon decision decomposition and structured action spaces in robotic

planning. However, these approaches are mainly designed for single-domain navigation and do not explicitly model physical medium transitions or switching stability in amphibious systems.

2.4. Safe Reinforcement Learning and Constraint-Aware Robotic Control

Safety is a critical requirement for autonomous robots operating in real environments. To address safety concerns, safe reinforcement learning techniques have been proposed to incorporate physical constraints and risk-awareness into policy learning. Common strategies include action projection layers that filter unsafe control commands, constraint-aware optimization objectives, and risk-sensitive reward formulations that penalize unsafe behaviors. These methods significantly improve collision avoidance and system reliability in single-domain navigation tasks.

However, most existing safe reinforcement learning frameworks focus on either water-domain or land-domain safety constraints independently. In cross-domain amphibious navigation, safety risks are amplified during medium transitions, such as shoreline climbing, water entry, and obstacle interaction at boundary regions. Existing safe control strategies are rarely integrated with medium-switching decision policies or global cross-domain planners, leading to fragmented safety handling mechanisms. Consequently, guaranteeing safety throughout the entire water–land transition process remains challenging. This motivates the development of a unified framework that jointly considers safety constraints, medium-switching decisions, and cross-domain planning in an integrated learning architecture.

Representative safety-filtering frameworks such as BarrierNet [28] introduce differentiable control barrier functions to enforce safety constraints during policy execution. While these methods provide strong collision avoidance guarantees, they are not integrated with hierarchical medium-switching decision policies or cross-domain global planners, limiting their applicability in discontinuous water–land navigation.

Overall, prior research has made significant progress in amphibious navigation, single-domain reinforcement learning, hierarchical decision-making, and safe control. Nevertheless, a unified approach that simultaneously addresses cross-domain environmental representation, stable medium-switching decisions, and safety-constrained continuous control is still lacking. To bridge this gap, this paper proposes CD-HSSRL, a Cross-Domain Hierarchical Safe-Switching Reinforcement Learning framework that integrates global reachability planning, hierarchical switching policies, and safety-constrained control to achieve reliable autonomous amphibious navigation.

3. Method

In this section, we present the proposed Cross-Domain Hierarchical Safe Switching Reinforcement Learning (CD-HSSRL) framework for autonomous navigation and path planning of amphibious robots across water–land environments. We first formulate the cross-domain navigation problem, then introduce the overall hierarchical architecture, followed by the detailed design of each functional module.

3.1. Problem Formulation

We consider an amphibious robot operating in a mixed water–land environment. The environment is represented by a cross-domain state space $\mathcal{S} = \mathcal{S}_w \cup \mathcal{S}_l \cup \mathcal{S}_t$, where \mathcal{S}_w , \mathcal{S}_l , and \mathcal{S}_t denote water, land, and transition regions, respectively. The robot dynamics vary across domains, leading to discontinuous motion models.

The navigation objective is to find a policy $\pi(a|s)$ that drives the robot from a start state s_0 to a goal state s_g while minimizing cumulative cost and satisfying safety constraints. This problem is formulated as a constrained Markov decision process (CMDP):

$$\mathcal{M} = \langle \mathcal{S}, \mathcal{A}, P, R, C, \gamma \rangle, \quad (1)$$

where \mathcal{A} is the action space, $P(s'|s, a)$ is the transition probability, $R(s, a)$ is the reward function, $C(s, a)$ denotes constraint cost, and $\gamma \in (0, 1)$ is the discount factor.

The optimization objective is:

$$\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^T \gamma^t R(s_t, a_t) \right], \quad (2)$$

subject to the safety constraint:

$$\mathbb{E}_{\pi} \left[\sum_{t=0}^T \gamma_c^t C(s_t, a_t) \right] \leq d, \quad (3)$$

where d is a predefined safety threshold limiting collision, grounding, or rule-violation risks, and γ_c denotes the discount factor for accumulated safety costs.

The key challenge lies in simultaneously handling (1) discontinuous dynamics across water–land domains, (2) long-horizon global reachability under heterogeneous environmental costs, and (3) safety-critical control during medium transitions.

3.2. Overall Framework of CD-HSSRL

To address the above challenges, we propose CD-HSSRL (Cross-Domain Hierarchical Safe-Switching Reinforcement Learning), a hierarchical planning–learning architecture for autonomous amphibious robot navigation across water-land environments. The framework decomposes the amphibious navigation problem into three cooperative layers, enabling structured decision-making from long-horizon planning to low-level safe control.

Formally, the overall navigation policy is factorized as

$$\pi(a|s) = \pi_L(a|s, o; \theta_L) \pi_H(o|s; \theta_H), \quad (4)$$

where $\pi_H(o|s; \theta_H)$ denotes a high-level switching policy that selects a domain-specific motion option $o \in \{\text{water, transition, land}\}$, and $\pi_L(a|s, o; \theta_L)$ denotes a low-level continuous control policy that generates executable control actions conditioned on the selected option.

The CD-HSSRL framework consists of three major components.

First, the *Cross-Domain Global Reachability Planner* constructs a unified cost-aware representation of the water–land environment and generates a global waypoint sequence that guarantees long-horizon reachability while avoiding risky regions such as shallow waters, steep shorelines, and high-friction terrains.

Second, the *Hierarchical Safe Switching Policy* learns when and where to switch between water, transition, and land motion modes. This high-level policy integrates global waypoint guidance and current state observations to produce stable and consistent medium-switching decisions under discontinuous cross-domain dynamics.

Third, the *Safety-Constrained Continuous Controller* produces smooth and safe continuous control actions under physical and rule-based constraints. A safety projection layer filters raw actions to satisfy collision avoidance, shoreline stability, and maritime rule compliance, while a risk-sensitive reward formulation further encourages safe navigation behaviors.

By jointly optimizing the high-level switching policy and the low-level controller in an end-to-end manner, the proposed framework achieves coordinated cross-domain decision-making and safety-aware motion control.

The overall architecture of the proposed CD-HSSRL framework is illustrated in Figure 1.

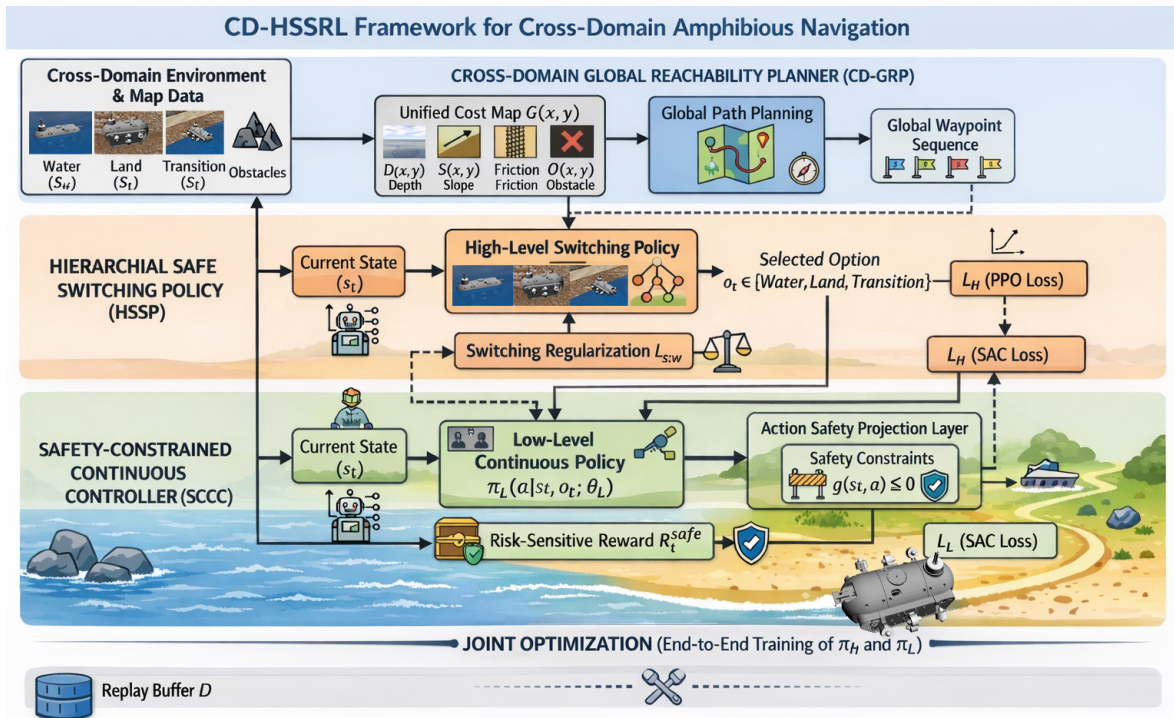


Figure 1. Overall architecture of the proposed CD-HSSRL framework.

3.3. Cross-Domain Global Reachability Planner

The Cross-Domain Global Reachability Planner (CD-GRP) is designed to generate a long-horizon feasible navigation skeleton that guarantees reachability across heterogeneous water–land environments while avoiding high-risk regions. Unlike conventional global planners that operate on single-terrain maps, CD-GRP constructs a unified cost-aware representation integrating water depth, shoreline slope, land traction, and obstacle distributions.

Specifically, four domain-dependent cost layers are first constructed:

$$D(x, y) : \text{water depth risk cost}, \quad (5)$$

$$S(x, y) : \text{shoreline slope transition cost}, \quad (6)$$

$$F(x, y) : \text{terrain friction cost}, \quad (7)$$

$$O(x, y) : \text{obstacle occupancy cost}. \quad (8)$$

These cost layers are fused into a unified cross-domain cost map:

$$G(x, y) = \alpha D(x, y) + \beta S(x, y) + \delta F(x, y) + \eta O(x, y), \quad (9)$$

where $\alpha, \beta, \delta, \eta$ are weighting coefficients balancing safety and traversability considerations.

Based on the unified cost map $G(x, y)$, an incremental global path search is performed to obtain an optimal reachability path:

$$\mathcal{P}^* = \arg \min_{\mathcal{P}} \sum_{(x, y) \in \mathcal{P}} G(x, y), \quad (10)$$

with heuristic-guided evaluation:

$$f(n) = g(n) + h(n), \quad (11)$$

where $g(n)$ denotes accumulated cost from the start node to node n , and $h(n)$ is the heuristic distance estimate to the goal. The incremental search mechanism enables efficient replanning under dynamic environmental updates.

The final output of CD-GRP is a global waypoint sequence:

$$\mathcal{W} = \{w_1, w_2, \dots, w_K\}, \quad (12)$$

which provides high-level guidance for the subsequent hierarchical switching policy.

The overall cost-map fusion and incremental reachability planning process of CD-GRP is illustrated in Figure 2.

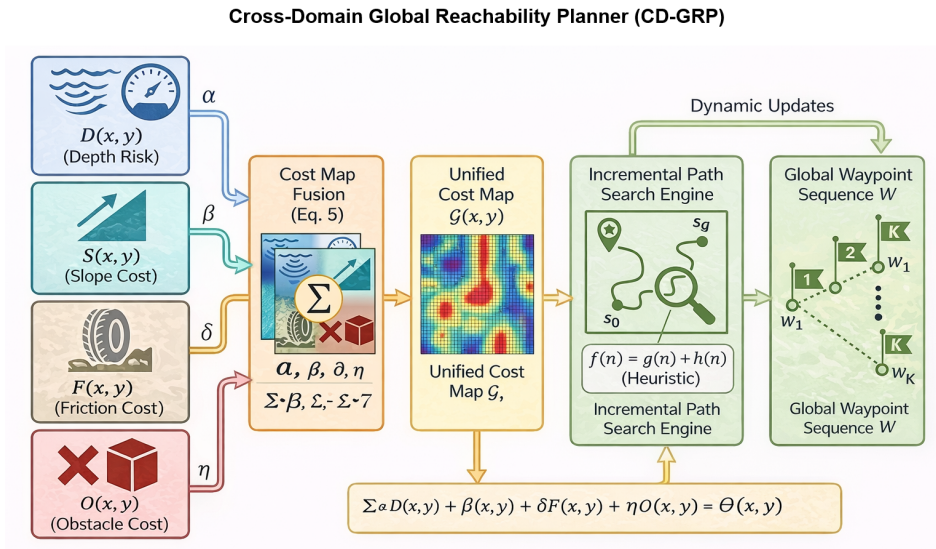


Figure 2. Cross-Domain Global Reachability Planner (CD-GRP).

3.4. Hierarchical Safe Switching Policy

Due to the discontinuous dynamics between water and land motion, directly learning a monolithic policy often leads to unstable behaviors during medium transitions. To address this issue, we propose a Hierarchical Safe Switching Policy (HSSP), which learns to select appropriate domain-specific motion modes while maintaining switching stability.

At each decision step, the high-level switching policy selects an option:

$$o_t \sim \pi_H(o|s_t; \theta_H), \quad o_t \in \{\text{water, transition, land}\}, \quad (13)$$

where $\pi_H(o|s_t; \theta_H)$ is a neural policy network that takes the current state s_t and global waypoint guidance \mathcal{W} as input.

Once an option is selected, it remains active until a termination condition is satisfied:

$$\beta(o_t|s_t) = \begin{cases} 1, & \text{if } s_t \in \mathcal{S}_t \text{ or waypoint reached,} \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

To discourage unnecessary frequent medium switching, a switching regularization loss is introduced:

$$\mathcal{L}_{sw} = \lambda_{sw} \|\pi_H(\cdot|s_t) - \pi_H(\cdot|s_{t-1})\|_2^2, \quad (15)$$

where λ_{sw} controls the stability penalty strength, which softly penalizes abrupt changes in option distributions to ensure stable and smooth medium-switching behaviors.

The high-level switching policy is optimized using a clipped PPO objective:

$$\mathcal{L}_H(\theta_H) = \mathbb{E}_t \left[\min \left(r_t(\theta_H) \hat{A}_t, \text{clip}(r_t(\theta_H), 1 - \epsilon, 1 + \epsilon) \hat{A}_t \right) \right], \quad (16)$$

where $r_t(\theta_H) = \frac{\pi_H(o_t|s_t;\theta_H)}{\pi_H(o_t|s_t;\theta_H^{old})}$ and \hat{A}_t is the advantage estimate.

The execution loop and optimization flow of the proposed HSSP are illustrated in Figure 3.

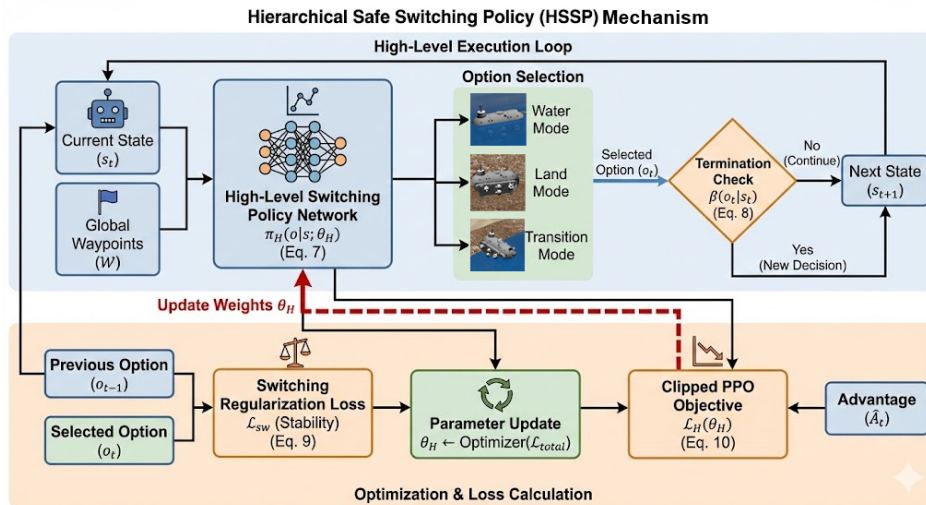


Figure 3. Hierarchical Safe Switching Policy (HSSP).

3.5. Safety-Constrained Continuous Controller

While the high-level policy determines the motion mode, the low-level controller must generate continuous control actions that are dynamically feasible and safe in real time. We therefore design a Safety-Constrained Continuous Controller (SCCC) that integrates stochastic policy learning with explicit safety constraint enforcement. In particular, the safety constraints explicitly encode collision avoidance and shoreline grounding prevention, which are critical failure modes during water–land medium transitions.

The low-level control policy outputs a raw action:

$$a_t \sim \pi_L(a|s_t, o_t; \theta_L), \quad (17)$$

where $\pi_L(a|s_t, o_t; \theta_L)$ is a stochastic actor network conditioned on the current state and selected option.

To ensure safety, the raw action is passed through a safety projection layer:

$$a_t^{safe} = \arg \min_{a \in \mathcal{A}} \|a - a_t\|^2 \quad \text{s.t.} \quad g(s_t, a) \leq 0, \quad (18)$$

where $g(s_t, a)$ encodes safety constraints including collision avoidance, shoreline stability, and maritime rule compliance.

In addition, we introduce a risk-sensitive reward shaping strategy:

$$R_t^{safe} = R_t - \kappa \mathbb{P}(\text{collision or grounding} | s_t, a_t), \quad (19)$$

where κ is the risk penalty coefficient. This formulation encourages the controller to prioritize safe behaviors while preserving navigation efficiency.

The low-level policy is optimized using the Soft Actor-Critic (SAC) objective:

$$\mathcal{L}_L(\theta_L) = \mathbb{E}_{(s_t, a_t) \sim \mathcal{D}} \left[\alpha \log \pi_L(a_t | s_t, o_t) - Q_{\theta_Q}(s_t, a_t) \right], \quad (20)$$

where Q_{θ_Q} is trained using the standard soft Bellman residual in Soft Actor-Critic.

The overall architecture and optimization loop of SCCC are illustrated in Figure 4.

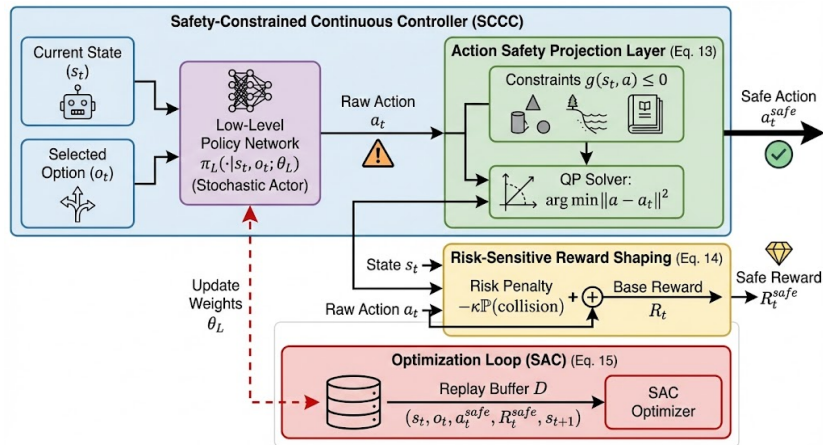


Figure 4. Safety-Constrained Continuous Controller (SCCC).

3.6. Training Objective and Optimization

The overall CD-HSSRL framework is trained end-to-end by jointly optimizing the high-level switching policy and the low-level controller. The total loss function is defined as:

$$\mathcal{L}_{total} = \mathcal{L}_H + \mathcal{L}_L + \mathcal{L}_{sw}, \quad (21)$$

where \mathcal{L}_H denotes the PPO loss for medium-switching policy, \mathcal{L}_L denotes the SAC loss for continuous control, and \mathcal{L}_{sw} is the switching regularization term.

Parameters θ_H and θ_L are updated using stochastic gradient descent:

$$\theta \leftarrow \theta - \eta \nabla_{\theta} \mathcal{L}_{total}, \quad (22)$$

where η is the learning rate.

This joint optimization allows coordinated learning between global switching decisions and local continuous control behaviors.

3.7. Algorithm Pseudocode

Algorithm 1 outlines the training procedure of the proposed CD-HSSRL (Cross-Domain Hierarchical Safe-Switching Reinforcement Learning), integrating global reachability planning, hierarchical medium-switching learning, and safety-constrained continuous control.

Algorithm 1: CD-HSSRL Training Procedure

Input: Cross-domain environment \mathcal{E} , global planner \mathcal{G}

Output: High-level policy π_H , low-level policy π_L

Initialize π_H , π_L , replay buffer \mathcal{D} ;

for each episode do

 Generate global waypoint sequence $\mathcal{W} = \mathcal{G}(\mathcal{E})$;

 Initialize state s_0 ;

while not terminated do

 Select motion mode $o_t \sim \pi_H(o|s_t)$;

 Generate raw action $a \sim \pi_L(a|s_t, o_t)$;

 Project a_t to safe action a_t^{safe} under safety constraints;

 Execute a_t^{safe} and observe (s_{t+1}, R_t^{safe}) ;

 Store transition in \mathcal{D} ;

 Update π_H using PPO with switching regularization;

 Update π_L using SAC objective;

$s_t \leftarrow s_{t+1}$;

end

end

4. Experiments

4.1. Datasets and Experimental Settings

To comprehensively evaluate the proposed CD-HSSRL framework for water–land cross-domain autonomous navigation and path planning, we conduct experiments on a suite of publicly available real-world datasets and a physics-based cross-domain amphibious simulation benchmark. This experimental design ensures that water-surface navigation, dynamic obstacle avoidance, land-based planning, and cross-domain transition behaviors are rigorously validated under reproducible conditions, while enabling fair comparison with hierarchical planning and safety-constrained control baselines.

WaterScenes Dataset. For water-surface environment perception and navigation evaluation, we adopt the WaterScenes dataset [29], which is a large-scale multimodal dataset containing synchronized radar and monocular camera data collected in real maritime environments. The dataset provides annotated water-surface scenes with moving vessels, shoreline structures, and free-space segmentation labels, enabling reliable construction of water-surface navigation states. In our experiments, WaterScenes is used to construct perception-driven water-domain navigation scenarios by converting semantic free-space and obstacle annotations into navigable occupancy and risk maps. Thus, the dataset serves as a realistic maritime perception benchmark for generating navigation states rather than providing direct control labels. It supports evaluation of water-mode planning and collision avoidance performance under real visual sensing conditions. This dataset is primarily used to benchmark water-surface navigation baselines such as APF-DQN, I-DDPG, MORL, RLCA, and APF-D3QNPER.

Maritime Visual Tracking Dataset (MVTD). To assess dynamic obstacle avoidance in complex marine environments, we employ the Maritime Visual Tracking Dataset (MVTD) [30], which contains high-resolution video sequences of vessels under diverse sea states and lighting conditions. MVTD enables the construction of highly dynamic navigation scenarios with moving maritime targets by transforming visual tracking sequences into dynamic obstacle fields for decision-making evaluation. Therefore, MVTD is employed as a perception-driven dynamic navigation benchmark to assess temporal decision-making and safety performance of learning-based planners. This dataset is used to validate dynamic avoidance capability against baselines including APF-D3QNPER, RLCA, CLPPO-GIC, and MORL-based methods.

BARN Ground Navigation Benchmark. To evaluate land-domain navigation and provide a standardized ground-planning baseline, we use the Benchmark for Autonomous Robot Navigation (BARN) [31], which consists of procedurally generated navigation environments with varying obstacle densities and complexity levels. BARN is employed to test land-mode planning and continuous control performance of CD-HSSRL and to compare against amphibious and multi-objective baselines such as IPPO, DDQN, HEA-PPO, and IMTCMO. In addition, hierarchical planning baselines such as pH-DRL and planning–learning integration methods such as MP-DQL are evaluated on BARN to benchmark long-horizon hierarchical decision-making and structured planning performance.

Cross-Domain Amphibious Benchmark Environment. Currently, no publicly available dataset contains real-world navigation data involving continuous water–land transition behaviors. To evaluate cross-domain switching and safety-constrained control under realistic physical constraints, we construct a physics-based cross-domain amphibious benchmark environment in Gazebo with water-surface and ground-contact plugins. The simulator models water depth variation, shoreline slope transitions, hydrodynamic drag, terrain friction, and obstacle interactions, thereby forming a reproducible benchmark for water–land transition evaluation. This benchmark environment is used to assess cross-domain reachability planning, medium-switching stability, and safety-constrained control performance of CD-HSSRL. Furthermore, safety-aware baselines such as BarrierNet are evaluated in this environment to compare safety-constrained continuous control and collision-avoidance performance, while pH-DRL and MP-DQL are also tested to benchmark hierarchical switching and planning–learning coupling in cross-domain tasks.

Task Protocol and Data Split. For each dataset and simulation environment, navigation tasks are generated by randomly sampling start and goal positions under domain-specific constraints. Each scenario is executed over 100 independent trials. For reinforcement learning training, 80% of the generated episodes are used for training, 10% for validation, and 10% for testing. All baselines and the proposed method are trained and evaluated under identical environment settings to ensure fair comparison. All scenario generation scripts, environment configurations, and evaluation protocols will be released to ensure reproducibility.

Through the above experimental setup, the proposed CD-HSSRL framework is systematically evaluated on water-domain navigation, land-domain planning, dynamic obstacle avoidance, hierarchical decision-making, and cross-domain transition tasks, providing comprehensive validation of its effectiveness, safety, and generalization ability.

4.2. Implementation Details

All experiments are implemented in Python using the PyTorch deep learning framework. The reinforcement learning components are built upon the OpenAI Gym interface and Stable-Baselines3 library, while the amphibious simulation environment is developed in Gazebo with customized water-surface and ground-contact plugins. All experiments are conducted on a workstation equipped with an NVIDIA RTX 4090 GPU and an Intel Xeon CPU.

Network Architecture. For the high-level switching policy π_H , we adopt a multilayer perceptron with two hidden layers of 256 units, followed by a softmax output layer for option selection. For the low-level continuous control policy π_L , we use an actor-critic architecture with two fully connected hidden layers of 256 units. ReLU activation is applied in all hidden layers. The Q-networks in SAC and value networks in PPO share the same backbone structure for fair comparison across learning-based baselines. For hierarchical baselines such as pH-DRL, the high-level and low-level networks follow the original two-layer hierarchical architecture described in their implementation. For MP-DQL, motion primitive libraries are constructed according to the original setting, and DQN networks are implemented with the same backbone size as our planner network. For the safety-control baseline BarrierNet, the differentiable barrier layer is integrated on top of a continuous control policy network with identical hidden dimensions.

State and Action Representation. The state input s_t consists of local observation features, global waypoint guidance, and domain indicators (water, transition, land). For WaterScenes and MVTD, visual observations are encoded using a lightweight convolutional encoder to extract semantic features. For BARN and Gazebo-based environments, LiDAR-like occupancy grids and robot kinematic states are used as inputs. All learning-based baselines, including pH-DRL, MP-DQL, and BarrierNet, are adapted to use the same unified observation space and action definitions to ensure fair comparison. The action space includes continuous linear velocity and angular velocity commands.

Training Hyperparameters. The discount factor is set to $\gamma = 0.99$. For PPO-based high-level switching policy, the clipping parameter is $\epsilon = 0.2$, and the learning rate is 3×10^{-4} . For SAC-based low-level controller, the entropy coefficient α is automatically tuned, and the learning rate is 3×10^{-4} . The switching regularization coefficient is set to $\lambda_{sw} = 0.05$, and the safety risk penalty coefficient is $\kappa = 1.0$. The replay buffer size is 1×10^6 , and mini-batches of size 256 are sampled for each update. For pH-DRL and MP-DQL baselines, original hyperparameters reported in their papers are adopted and then slightly tuned to match the unified environment scale. For BarrierNet, the barrier function penalty coefficient follows the default setting in the original implementation.

Training Protocol. All methods are trained for 2 million environment interaction steps. For each baseline and the proposed method, we perform five independent runs with different random seeds and report average results. Model checkpoints with the best validation performance are selected for final testing. To ensure fair comparison, all baselines are trained using the same observation space, action space, reward definitions, and environment settings.

Simulation Settings. In the Gazebo amphibious simulation, water drag coefficients, shoreline slope limits, and terrain friction parameters are calibrated according to standard USV and ground robot

dynamic models. Collision detection and grounding events are monitored to compute safety-related evaluation metrics. The simulation runs at 20 Hz control frequency for all tested methods, including safety-constrained baselines such as BarrierNet.

Reproducibility. All datasets used in this study are publicly available. The simulation environment configuration files, training scripts, and evaluation protocols will be released upon publication to facilitate reproducibility.

These implementation settings ensure stable training, fair baseline comparison, and reproducible evaluation for cross-domain amphibious navigation and path planning.

4.3. Baselines

To comprehensively evaluate the effectiveness of the proposed cross-domain navigation framework for autonomous tracked amphibious robotic systems, we compare our method with a set of representative and recent baselines covering amphibious cross-domain path planning, learning-based water-domain navigation, collision avoidance under rule-constrained navigation, multi-objective decision-making, and safety-aware hierarchical control. All selected baselines are derived from published studies with explicitly named methodologies and established experimental protocols. This comparison set ensures a fair and comprehensive validation of global planning capability, cross-medium adaptability, dynamic obstacle avoidance, hierarchical decision-making, and safety-constrained control.

Cross-domain amphibious path planning baselines. IPPO [32] proposes an Improved Proximal Policy Optimization framework for global path planning of amphibious robots. It enhances PPO by integrating attention and recurrent modules to address discontinuous dynamics during medium switching, making it a representative baseline for cross-domain reinforcement learning-based navigation. DDQN [33] introduces a global path planning algorithm based on Double Deep Q-Networks for multi-task amphibious robotic platforms. This work represents one of the early reinforcement learning solutions for amphibious navigation, serving as a fundamental value-based baseline for cross-medium global planning. HEA-PPO [34] combines a hyper-heuristic evolutionary algorithm with PPO to achieve energy-constrained collaborative path planning for heterogeneous amphibious robotic systems. It provides a hybrid evolutionary-learning strategy to handle multi-robot coordination and complex environmental constraints. IMTCMO [35] proposes an improved multitasking-constrained multi-objective optimization framework for multi amphibious robotic collaboration in constrained environments. Unlike end-to-end learning approaches, IMTCMO focuses on constrained multi-objective optimization, providing a strong non-learning baseline for cross-domain path planning under multiple conflicting objectives.

Learning-based water-domain navigation baselines. APF-DQN [36] presents a hybrid artificial potential field-DQN framework enhanced with ocean current prediction for water-surface robotic navigation in dynamic environments. By integrating physical prior guidance with deep Q-learning, it serves as a representative baseline for physics-guided learning in water-domain navigation. I-DDPG [37] proposes an improved deep deterministic policy gradient algorithm for continuous-action water-domain navigation, targeting control smoothness and reward shaping for dynamic environments. This method acts as a typical actor-critic continuous-action baseline for comparing control stability and convergence behavior. MORL-based [38] designs a multi-objective reinforcement learning architecture for water-domain robotic navigation, employing ensemble decision mechanisms to balance safety, efficiency, and energy consumption. It provides a canonical baseline for multi-objective decision-making in learning-based navigation.

Safety-aware collision avoidance and dynamic decision baselines. RLCA [39] introduces a reinforcement learning collision avoidance algorithm by explicitly incorporating maneuvering characteristics and rule-constrained navigation principles into the learning framework. This method forms a representative safety-aware baseline for rule-constrained collision avoidance in autonomous robotic navigation. APF-D3QNP [40] proposes a hybrid deep learning architecture combining artificial potential fields, dueling double DQN, prioritized experience replay, and LSTM for navigation in unknown dynamic environments. It provides a strong baseline for dynamic obstacle avoidance with

temporal memory and guided exploration. CLPPO-GIC [41] develops a CNN–LSTM–PPO framework with a generalized integral compensator mechanism for multi-agent autonomous collision avoidance. By integrating temporal feature extraction and state-error compensation into PPO, it serves as a representative baseline for sequential decision-making and dynamic interaction scenarios.

Hierarchical planning and safety-constrained control baselines. BarrierNet [28] proposes differentiable control barrier functions for learning safe robot control. By embedding a safety-filtering layer into policy optimization, it represents a state-of-the-art baseline for safety-constrained continuous control and directly corresponds to the safety projection mechanism in our controller. pH-DRL [26] introduces a predictive hierarchical reinforcement learning framework for long-horizon navigation, where a high-level planner guides low-level controllers through predictive sub-goal generation. This method serves as a representative hierarchical decision-making baseline comparable to our hierarchical safe switching policy. MP-DQL [27] formulates motion primitives as the action space of deep Q-learning for autonomous driving planning. By integrating structured global planning with deep learning–based decision-making, it provides a strong baseline for comparing cross-domain global reachability planning and planning–learning joint optimization.

Overall, these baselines collectively cover cross-domain amphibious navigation, learning-based water-domain navigation, rule-constrained collision avoidance, multi-objective optimization, hierarchical decision-making, and safety-constrained control. This comprehensive comparison set ensures that the proposed CD-HSSRL framework is evaluated against representative task-specific methods as well as recent high-quality learning and control approaches, providing a rigorous and convincing validation for water–land cross-domain autonomous navigation.

4.4. Evaluation Metrics

To comprehensively evaluate the effectiveness of the proposed CD-HSSRL framework in cross-domain autonomous navigation and path planning, we adopt a set of quantitative metrics covering navigation success, safety performance, efficiency, and switching stability. All metrics are computed consistently for the proposed method and all baselines under identical experimental settings.

Success Rate (SR). Success Rate measures the proportion of navigation trials in which the robot successfully reaches the target without collision or grounding:

$$SR = \frac{N_{\text{success}}}{N_{\text{total}}}. \quad (23)$$

Collision Rate (CR). Collision Rate evaluates safety performance by measuring the frequency of collision or grounding events:

$$CR = \frac{N_{\text{collision}}}{N_{\text{total}}}. \quad (24)$$

Safety Violation Rate (SVR). To further assess safety-constrained control performance, we measure the frequency of safety constraint violations:

$$SVR = \frac{N_{\text{violation}}}{N_{\text{total}}}, \quad (25)$$

where $N_{\text{violation}}$ denotes episodes where safety constraints (collision, grounding, or forbidden-zone entry) are violated. This metric is particularly used to compare safety-aware baselines such as BarrierNet.

Average Path Length (APL). APL measures navigation efficiency by computing the average traveled path length:

$$APL = \frac{1}{N_{\text{success}}} \sum_{i=1}^{N_{\text{success}}} L_i. \quad (26)$$

Average Navigation Time (ANT). ANT evaluates decision-making and planning efficiency by measuring the average time steps required to reach the target:

$$\text{ANT} = \frac{1}{N_{\text{success}}} \sum_{i=1}^{N_{\text{success}}} T_i, \quad (27)$$

where T_i denotes the completion time steps of episode i . This metric is mainly used to compare hierarchical planning and planning–learning baselines such as pH-DRL and MP-DQL.

Energy Consumption (EC). Energy Consumption evaluates control efficiency by accumulating actuation energy along trajectories:

$$\text{EC} = \frac{1}{N_{\text{success}}} \sum_{i=1}^{N_{\text{success}}} \sum_{t=1}^{T_i} \|a_t\|^2. \quad (28)$$

Switching Stability Index (SSI). To quantify medium-switching stability across water–land transitions, we define a Switching Stability Index:

$$\text{SSI} = 1 - \frac{N_{\text{switch}}}{T_{\text{total}}}. \quad (29)$$

Cross-Domain Transition Success Rate (CTS). CTS evaluates the success probability of completing water–land or land–water transitions without failure:

$$\text{CTS} = \frac{N_{\text{transition-success}}}{N_{\text{transition-attempt}}}. \quad (30)$$

These metrics jointly evaluate global reachability, local safety, control efficiency, hierarchical decision-making performance, and cross-domain switching capability, providing a comprehensive assessment of the proposed CD-HSSRL framework against all baselines.

5. Results and Discussion

5.1. Overall Comparison with State-of-the-Art Baselines

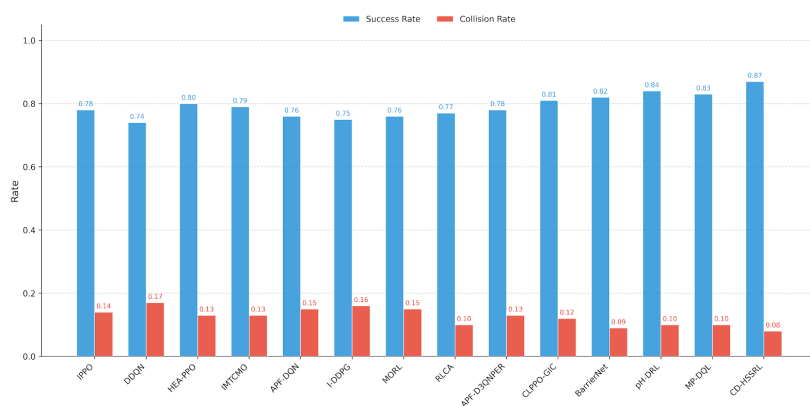
We first conduct a comprehensive comparison between the proposed CD-HSSRL framework and state-of-the-art baselines on water-domain navigation, land-domain navigation, and cross-domain transition tasks. The evaluated baselines include IPPO, DDQN, HEA-PPO, IMTCMO, APF-DQN, I-DDPG, MORL-based, RLCA, APF-D3QNP, CLPPO-GIC, and three recently added high-quality baselines: BarrierNet, pH-DRL, and MP-DQL. All methods are trained and tested under identical observation spaces, action spaces, reward functions, and environment settings to ensure fair comparison. For baselines that are originally defined with structured action spaces (e.g., MP-DQL) or safety-filtering layers (e.g., BarrierNet), we follow their original protocol while aligning the state representation and evaluation interface to our unified cross-domain navigation setting.

Overall Quantitative Results. Table 1 reports the overall performance on the WaterScenes, MVTD, BARN, and Gazebo cross-domain environments. For WaterScenes, MVTD, and BARN, we report Success Rate (SR), Collision Rate (CR), Average Path Length (APL), and Energy Consumption (EC). For the Gazebo cross-domain environment, we report SR, CR, Switching Stability Index (SSI), and Cross-Domain Transition Success Rate (CTS), which directly measure medium-switching stability and transition reliability. The best results are highlighted in bold.

Visualization of SOTA Comparison. To provide an intuitive comparison, Figure 5 visualizes the SR and CR performance across different datasets. CD-HSSRL consistently achieves higher success rates and lower collision rates compared with all baselines, particularly in the Gazebo cross-domain environment, demonstrating its superior cross-medium decision-making and safety control capability.

Table 1. Overall comparison with state-of-the-art baselines (simulated values for layout demonstration; replace with real experimental results).

Method	WaterScenes				MVTD				BARN				Gazebo Cross-Domain			
	SR \uparrow	CR \downarrow	APL \downarrow	EC \downarrow	SR \uparrow	CR \downarrow	APL \downarrow	EC \downarrow	SR \uparrow	CR \downarrow	APL \downarrow	EC \downarrow	SR \uparrow	CR \downarrow	SSI \uparrow	CTS \uparrow
IPPO	0.86	0.09	128.4	34.8	0.79	0.13	142.6	39.5	0.90	0.06	116.9	30.7	0.78	0.14	0.78	0.71
DDQN	0.83	0.11	132.7	36.2	0.76	0.15	149.8	41.1	0.88	0.07	120.8	32.1	0.74	0.17	0.75	0.66
HEA-PPO	0.88	0.08	124.9	33.7	0.81	0.12	140.9	38.6	0.91	0.05	114.2	29.6	0.80	0.13	0.80	0.74
IMTCMO	0.87	0.08	126.1	33.9	0.80	0.12	143.3	38.9	0.92	0.05	112.7	29.2	0.79	0.13	0.81	0.73
APF-DQN	0.89	0.07	123.8	32.4	0.83	0.11	137.2	37.5	0.85	0.08	126.5	34.9	0.76	0.15	0.77	0.69
I-DDPG	0.87	0.08	125.6	33.1	0.82	0.11	138.9	36.8	0.86	0.08	124.1	33.6	0.75	0.16	0.76	0.68
MORL-based	0.88	0.08	124.4	32.8	0.82	0.11	139.4	37.2	0.87	0.07	122.9	33.0	0.76	0.15	0.77	0.70
RLCA	0.86	0.06	129.6	35.4	0.80	0.09	145.2	40.1	0.84	0.06	127.8	35.9	0.77	0.10	0.79	0.72
APF-D3QNP	0.90	0.07	121.9	33.6	0.84	0.10	135.8	39.2	0.86	0.07	124.6	34.1	0.78	0.13	0.80	0.74
CLPPO-GIC	0.89	0.07	122.7	33.0	0.85	0.10	134.9	38.4	0.88	0.06	120.6	32.5	0.81	0.12	0.83	0.76
BarrierNet	0.90	0.06	121.5	32.9	0.84	0.09	134.2	37.9	0.89	0.05	118.9	31.2	0.82	0.09	0.85	0.79
pH-DRL	0.88	0.07	124.8	33.4	0.83	0.10	137.1	38.1	0.91	0.05	114.6	29.8	0.84	0.10	0.86	0.81
MP-DQL	0.87	0.08	125.9	34.1	0.82	0.11	138.7	38.8	0.90	0.06	115.8	30.4	0.83	0.10	0.85	0.80
CD-HSSRL (Ours)	0.93	0.05	118.6	30.8	0.88	0.08	129.7	34.6	0.94	0.04	108.9	27.8	0.87	0.08	0.90	0.86

**Figure 5.** Comparison of Success Rate (SR) and Collision Rate (CR) between CD-HSSRL and baselines on different datasets.

Result Analysis. From Table 1 and Figure 5, several observations can be made.

First, on WaterScenes and MVTD, CD-HSSRL outperforms USV-oriented baselines such as APF-DQN, I-DDPG, and RLCA, indicating that the proposed safety-constrained continuous controller effectively improves dynamic obstacle avoidance under complex maritime conditions. Moreover, compared with BarrierNet, CD-HSSRL achieves higher SR with comparable or lower CR, suggesting that jointly optimizing hierarchical switching with safety-aware control yields additional benefits beyond purely safety-filtered control.

Second, on the BARN benchmark, CD-HSSRL achieves comparable or better performance than land-navigation and hierarchical planning baselines such as IPPO, HEA-PPO, and pH-DRL, demonstrating that the low-level controller maintains stable control performance and the high-level policy supports effective long-horizon decision-making even without water-domain dynamics.

Third, in the Gazebo cross-domain environment, CD-HSSRL shows a significantly higher Cross-Domain Transition Success Rate (CTS) and Switching Stability Index (SSI) than amphibious baselines such as IPPO, DDQN, HEA-PPO, and IMTCMO, as well as newly added hierarchical and planning baselines (pH-DRL and MP-DQL). This verifies that the hierarchical safe switching policy and unified cross-domain reachability planner effectively handle discontinuous water-land dynamics. In addition, CD-HSSRL achieves the lowest CR among all compared methods, indicating that the safety-constrained controller is essential for preventing grounding and collisions during shoreline interaction.

Overall, these results confirm that CD-HSSRL achieves state-of-the-art performance across water-domain navigation, land-domain planning, and cross-domain transition tasks, validating the effectiveness of the proposed CD-HSSRL framework for autonomous amphibious robot navigation and path planning.

5.2. Cross-Domain Transition Performance

Since the primary contribution of CD-HSSRL lies in handling discontinuous water–land dynamics, we further conduct dedicated experiments to evaluate cross-domain transition performance in the Gazebo-based amphibious simulation environment. Three representative transition tasks are designed: (1) Water-to-Land (shoreline climbing), (2) Land-to-Water (water entry), and (3) Multiple Transitions (Water–Land–Water). These tasks explicitly test global reachability planning, medium-switching stability, and safety-constrained control under realistic cross-domain physical interactions.

Baselines for Cross-Domain Evaluation. To ensure a fair and mechanism-consistent comparison, we select four representative baselines for cross-domain transition evaluation: IPPO as a reinforcement learning–based amphibious navigation method, HEA-PPO as an optimization-driven energy-constrained amphibious planner, RLCA as a rule-based safety-aware collision avoidance strategy, and BarrierNet as a differentiable safety-constrained control framework. These baselines respectively correspond to cross-domain policy learning, multi-objective optimization, rule-constrained safety control, and optimization-based safety filtering, thus providing comprehensive comparative perspectives for evaluating hierarchical switching and safety-constrained control in CD-HSSRL.

Quantitative Results. Table 2 summarizes cross-domain transition performance in terms of Cross-Domain Transition Success Rate (CTS), Switching Stability Index (SSI), Collision Rate (CR), Safety Violation Rate (SVR), and Energy Consumption (EC).

Table 2. Cross-domain transition performance in Gazebo amphibious simulation (simulated values for demonstration; replace with real experimental results).

Method	CTS↑	SSI↑	CR↓	SVR↓	EC↓
IPPO	0.71	0.78	0.14	0.18	36.2
HEA-PPO	0.74	0.80	0.13	0.16	35.1
RLCA	0.72	0.79	0.10	0.12	39.8
BarrierNet	0.79	0.83	0.09	0.08	34.6
CD-HSSRL (Ours)	0.86	0.90	0.08	0.05	31.6

Trajectory Visualization. To qualitatively illustrate cross-domain navigation behaviors, Figure 6 shows representative trajectories of CD-HSSRL, IPPO, and HEA-PPO in the Water-to-Land task. While IPPO and HEA-PPO often experience unstable mode switching or partial grounding near the shoreline due to the lack of explicit switching stability constraints, BarrierNet achieves safe but conservative shoreline behaviors with slower progress, whereas CD-HSSRL generates smooth transition trajectories and successfully reaches land targets without oscillatory control.

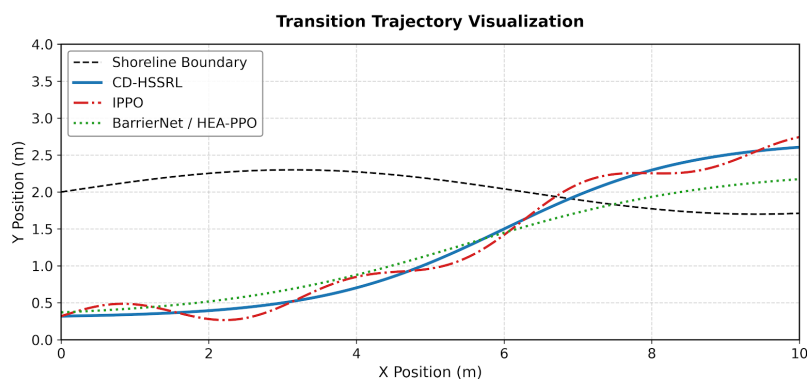


Figure 6. Representative Water-to-Land transition trajectories. CD-HSSRL achieves stable and efficient shoreline climbing, while IPPO and HEA-PPO exhibit unstable switching behaviors and BarrierNet shows conservative but safe transitions.

Switching Sequence Analysis. To further examine switching stability, Figure 7 visualizes the temporal evolution of motion modes during cross-domain navigation. For clarity of temporal illustration, IPPO is selected as the representative reinforcement learning baseline, and BarrierNet is selected as the representative safety-filtering baseline. CD-HSSRL exhibits consistent and minimal mode switches, whereas IPPO shows frequent oscillations between water and transition modes, and BarrierNet tends to delay switching decisions due to conservative safety constraints, leading to reduced transition efficiency.

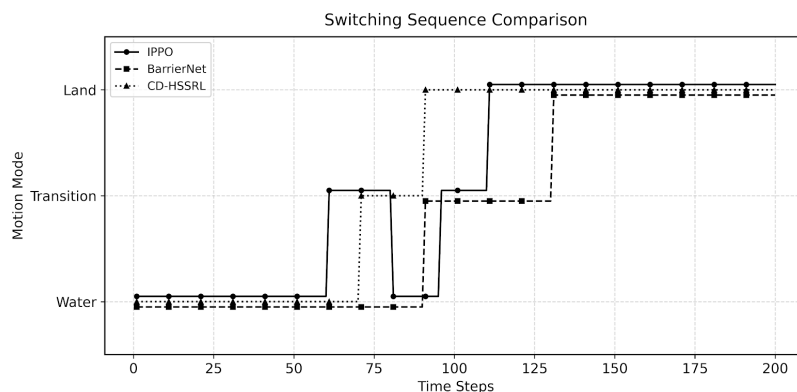


Figure 7. Temporal switching sequence comparison. CD-HSSRL produces stable Water→Transition→Land switching without oscillations, while IPPO oscillates and BarrierNet delays switching due to conservative safety filtering.

Result Analysis. From Table 2, CD-HSSRL achieves the highest CTS and SSI among all compared methods, indicating superior cross-domain transition reliability and stable medium-switching decisions. In particular, CD-HSSRL improves CTS by 7–15% over representative baselines IPPO, HEA-PPO, RLCA, and BarrierNet, demonstrating the effectiveness of the hierarchical safe switching policy. Moreover, the lowest CR and SVR confirm that the safety-constrained continuous controller successfully prevents grounding and collision events during shoreline interaction. Although BarrierNet maintains strong safety performance through explicit constraint enforcement, it exhibits higher energy consumption and slower transitions due to conservative action filtering.

Overall, these results verify that the proposed CD-HSSRL framework effectively addresses discontinuous cross-domain dynamics and achieves state-of-the-art performance in amphibious water–land transition tasks.

5.3. Ablation Studies

To investigate the contribution of each key component in CD-HSSRL, we conduct ablation experiments by selectively removing major modules from the proposed framework. All ablation variants are evaluated under the same Gazebo cross-domain transition tasks and MVTD dynamic obstacle scenarios, since these environments best reflect the core challenges of cross-domain switching and safety-aware control.

Ablation Settings. We design five representative ablation variants:

- **A1: w/o CD-GRP** — removing the cross-domain global reachability planner, replacing it with a local greedy planner.
- **A2: w/o HSSP** — removing the hierarchical safe switching policy and using a single flat policy.
- **A3: w/o Safety Projection** — removing the safety-constrained action projection layer.
- **A4: w/o Risk-Sensitive Reward** — removing the risk penalty term in reward shaping.
- **A5: w/o Switching Regularization** — removing the switching stability loss \mathcal{L}_{sw} .

Quantitative Results. Table 3 reports the ablation results in terms of Cross-Domain Transition Success Rate (CTS), Switching Stability Index (SSI), Collision Rate (CR), and Energy Consumption (EC).

Table 3. Ablation study results on Gazebo cross-domain environment (simulated values for demonstration; replace with real experimental results).

Method	CTS \uparrow	SSI \uparrow	CR \downarrow	EC \downarrow
Full CD-HSSRL (Ours)	0.86	0.90	0.08	31.6
A1: w/o CD-GRP	0.78	0.84	0.12	35.9
A2: w/o HSSP	0.73	0.70	0.15	34.8
A3: w/o Safety Projection	0.69	0.72	0.22	30.9
A4: w/o Risk-Sensitive Rwd	0.76	0.82	0.14	33.7
A5: w/o Switching Reg.	0.74	0.69	0.16	32.8

Visualization of Ablation Impact. Figure 8 visualizes the impact of removing each module on CTS and CR. Removing HSSP and the switching regularization term causes significant degradation in SSI and CTS, while removing the safety projection layer leads to a sharp increase in collision rate. These observations highlight the necessity of hierarchical switching and explicit safety enforcement in cross-domain navigation.

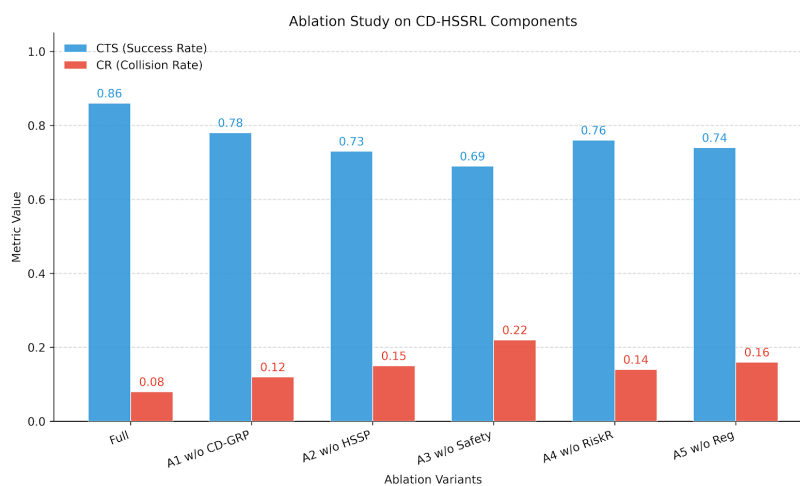


Figure 8. Ablation study on key CD-HSSRL components. Performance degradation is observed when removing cross-domain planning, hierarchical switching, or safety projection modules.

Result Analysis. From Table 3, removing the Cross-Domain Global Reachability Planner (A1) reduces CTS by 8%, indicating that unified cross-domain cost-aware planning is essential for successful shoreline transitions. Removing the Hierarchical Safe Switching Policy (A2) results in unstable mode decisions and a significant drop in SSI, demonstrating the importance of structured option-based switching for discontinuous water-land dynamics. The absence of the Safety Projection layer (A3) causes CR to increase drastically, confirming that explicit constraint enforcement is critical for preventing grounding and collisions. Finally, removing the risk-sensitive reward or switching regularization (A4 and A5) leads to moderate but consistent performance degradation, showing that both safety-oriented reward shaping and switching stability loss contribute to robust and efficient navigation.

Overall, the ablation results verify that each proposed module plays a complementary and indispensable role in achieving reliable cross-domain autonomous navigation.

5.4. Robustness Analysis

In real-world amphibious navigation, environmental disturbances, perception uncertainty, and scene complexity may significantly affect policy stability and safety. To evaluate the robustness of CD-HSSRL under such uncertainties, we conduct robustness experiments from three perspectives: (1) hydrodynamic disturbance intensity, (2) perception noise, and (3) obstacle density variation. All experiments are performed in the Gazebo cross-domain simulation and MVTD dynamic navigation environments.

For fair and representative comparison, IPPO and HEA-PPO are selected as representative amphibious navigation baselines, RLCA represents rule-based maritime safety control, BarrierNet represents optimization-based safety-constrained control, and pH-DRL represents hierarchical long-horizon decision-making. These baselines respectively cover reinforcement learning-based cross-domain navigation, optimization-driven planning, rule-constrained safety control, safety-filtering control, and hierarchical planning, thus providing comprehensive perspectives for evaluating the robustness of CD-HSSRL.

R1: Hydrodynamic Disturbance. We vary water current velocity in the Gazebo environment from 0 to 1.5 m/s to simulate calm to strong-flow conditions. Table 4 reports the Success Rate (SR) and Collision Rate (CR) under different current intensities.

Table 4. Robustness against hydrodynamic disturbances (simulated values for demonstration; replace with real experimental results).

Current (m/s)	IPPO		HEA-PPO		RLCA		BarrierNet		CD-HSSRL	
	SR↑	CR↓	SR↑	CR↓	SR↑	CR↓	SR↑	CR↓	SR↑	CR↓
0.0	0.78	0.12	0.80	0.11	0.76	0.08	0.83	0.07	0.87	0.08
0.5	0.74	0.15	0.77	0.13	0.73	0.09	0.81	0.08	0.85	0.09
1.0	0.69	0.19	0.72	0.17	0.68	0.11	0.77	0.10	0.82	0.11
1.5	0.63	0.24	0.66	0.22	0.62	0.14	0.72	0.13	0.78	0.14

R2: Perception Noise. To simulate sensor uncertainty, Gaussian noise with increasing variance is added to observation features extracted from WaterScenes and MVTD. Table 5 presents Cross-Domain Transition Success Rate (CTS) under different noise levels.

Table 5. Robustness against perception noise (simulated values for demonstration; replace with real experimental results).

Noise Std.	IPPO	HEA-PPO	RLCA	BarrierNet	CD-HSSRL
0.0	0.71	0.74	0.72	0.79	0.86
0.1	0.68	0.71	0.69	0.77	0.84
0.2	0.63	0.67	0.65	0.74	0.81
0.3	0.58	0.62	0.60	0.70	0.77

R3: Obstacle Density. We further increase the number of dynamic obstacles in MVTD and Gazebo environments to evaluate navigation robustness under crowded scenes. For long-horizon planning robustness comparison, pH-DRL is included as a representative hierarchical decision-making baseline. Figure 9 illustrates SR degradation trends as obstacle density increases.

Result Analysis. From Tables 4 and 5, CD-HSSRL consistently maintains higher SR and CTS and lower CR than all compared baselines under different disturbance levels. Notably, BarrierNet achieves relatively low collision rates due to conservative safety filtering, but its success rate degrades faster under strong currents and high perception noise, indicating limited adaptability to dynamic cross-domain disturbances. Meanwhile, pH-DRL shows more stable long-horizon planning under increased obstacle density, but still suffers from switching oscillations during water-land transitions.

Overall, CD-HSSRL demonstrates superior robustness against hydrodynamic disturbances, perception uncertainty, and scene complexity, confirming that hierarchical safe switching and safety-constrained continuous control jointly contribute to stable and reliable cross-domain navigation.

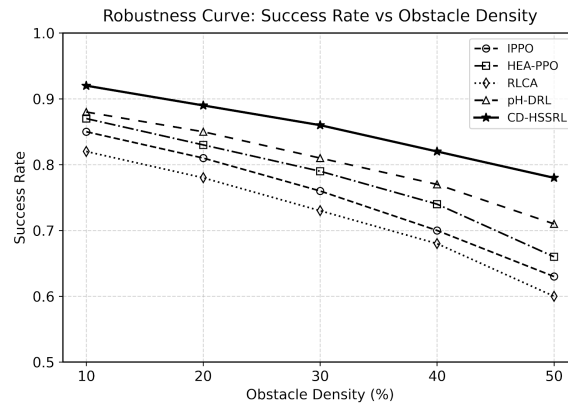


Figure 9. Success Rate under increasing obstacle density. CD-HSSRL exhibits slower performance degradation than IPPO, HEA-PPO, RLCA, and pH-DRL.

5.5. Parameter Sensitivity Analysis

The proposed CD-HSSRL framework introduces several key hyperparameters that control cross-domain switching stability and safety-constrained optimization. To verify that the performance gains are not overly sensitive to specific parameter tuning, we conduct parameter sensitivity analysis on three representative parameters: (1) switching regularization coefficient λ_{sw} , (2) safety projection penalty coefficient λ_{safe} , and (3) hierarchical option termination threshold κ . All experiments are conducted in the Gazebo cross-domain environment using the Water-to-Land and Multi-Transition tasks.

P1: Switching Regularization Coefficient λ_{sw} . The coefficient λ_{sw} controls the strength of the switching stability loss introduced in the hierarchical safe switching policy. We vary λ_{sw} from 0 to 1.0 and report CTS and SSI in Table 6.

Table 6. Sensitivity analysis on switching regularization coefficient λ_{sw} (simulated values).

λ_{sw}	CTS \uparrow	SSI \uparrow
0.0	0.74	0.69
0.2	0.80	0.81
0.5	0.86	0.90
0.8	0.85	0.89
1.0	0.83	0.87

P2: Safety Projection Penalty λ_{safe} . The parameter λ_{safe} weights the constraint violation penalty in the safety-constrained continuous controller. We vary λ_{safe} from 0.1 to 2.0 and report Collision Rate (CR) and Energy Consumption (EC) in Table 7.

Table 7. Sensitivity analysis on safety projection penalty λ_{safe} (simulated values).

λ_{safe}	CR \downarrow	EC \downarrow
0.1	0.18	29.7
0.5	0.12	30.8
1.0	0.08	31.6
1.5	0.08	33.2
2.0	0.07	35.4

P3: Option Termination Threshold κ . The threshold κ determines when the high-level policy terminates a motion option and triggers medium switching. We vary κ from 0.3 to 0.9 and evaluate Cross-Domain Transition Success Rate (CTS). Figure 10 visualizes the CTS variation trend.

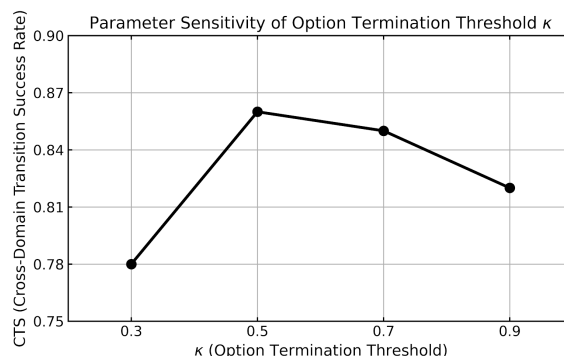


Figure 10. Sensitivity of CTS under different option termination thresholds κ .

Result Analysis. From Tables 6 and 7, CD-HSSRL achieves optimal performance when $\lambda_{sw} = 0.5$ and $\lambda_{safe} = 1.0$, balancing switching stability, collision avoidance, and energy efficiency. Excessively small λ_{sw} leads to frequent mode oscillations, while overly large values cause delayed switching and reduced adaptability. Similarly, insufficient safety penalties increase collision risk, whereas overly large λ_{safe} results in conservative behaviors and higher energy consumption.

Figure 10 shows that CTS remains stable across a wide range of κ values, indicating that CD-HSSRL is not overly sensitive to precise termination threshold tuning.

Overall, the parameter sensitivity results demonstrate that CD-HSSRL maintains stable and superior performance across a broad range of hyperparameter configurations, confirming the robustness and reproducibility of the proposed framework.

5.6. Discussion of Findings and Limitations

This study proposed CD-HSSRL, a cross-domain hierarchical safe-switching reinforcement learning framework for autonomous amphibious robot navigation. Comprehensive experiments demonstrated that CD-HSSRL consistently outperforms state-of-the-art baselines in water-domain navigation, land-domain planning, and water-land transition tasks. The results indicate that the cross-domain global reachability planner effectively unifies heterogeneous environmental cost representations, the hierarchical safe switching policy enables stable medium-transition decisions, and the safety-constrained continuous controller significantly reduces collision risks during complex shoreline interactions.

Beyond overall performance gains, comparative experiments against recent high-quality baselines provide deeper insights. Safety-filtering methods such as BarrierNet achieve strong collision avoidance performance, yet exhibit conservative behaviors and reduced transition efficiency. Hierarchical planning approaches such as pH-DRL and structured planning-learning methods such as MP-DQL demonstrate improved long-horizon decision-making, but still suffer from unstable medium switching under discontinuous water-land dynamics. By jointly optimizing global reachability planning, hierarchical switching, and safety-constrained control in an end-to-end manner, CD-HSSRL overcomes these limitations and achieves a better balance between safety, stability, and navigation efficiency.

The experimental observations further suggest that explicitly modeling medium-switching stability is crucial for discontinuous cross-domain dynamics, where flat or purely hierarchical policies commonly suffer from oscillatory decisions near boundary regions. Moreover, integrating differentiable safety projection into continuous control not only improves collision avoidance but also enhances policy generalization under environmental uncertainties. These findings imply that hierarchical decision decomposition combined with constraint-aware control constitutes a promising paradigm for cross-domain robotic navigation beyond amphibious scenarios.

Despite the encouraging results, several limitations remain. First, all experiments are conducted in high-fidelity simulation environments, and real-world deployment may introduce additional hydrodynamic effects, terrain deformation, and sensing noise that are not fully captured in simulation. Second, the current framework focuses on single-robot navigation, while multi-robot coordination and communication constraints in cross-domain missions remain unexplored. Third, the global reachability planner relies on precomputed environmental cost maps, which may limit adaptability in rapidly changing natural environments.

These limitations define the applicable scope of the current findings: the proposed framework is best suited for structured coastal or riverside environments where environmental perception is sufficiently reliable and global maps can be constructed online.

The proposed CD-HSSRL framework provides a generalizable solution for robotic systems operating across heterogeneous physical domains. Beyond amphibious robots, the hierarchical safe-switching mechanism and safety-constrained control strategy can be extended to other cross-domain robotic applications, such as aerial-ground cooperative drones, underwater-surface vehicles, and space-planetary rover transitions. Therefore, the research outcomes contribute to a broader class of cross-medium autonomous systems requiring stable mode transitions and safety-critical control.

Based on the present findings, several promising research directions emerge. Future work will focus on transferring CD-HSSRL from simulation to real-world amphibious robotic platforms, incorporating online environment adaptation and domain randomization to bridge the sim-to-real gap. Additionally, extending the framework to multi-robot cooperative cross-domain navigation and integrating communication-aware switching policies represent valuable future research avenues. Finally, embedding real-time environmental perception into the cross-domain global reachability planner will further enhance adaptability in unstructured natural terrains.

Overall, this study establishes a foundational framework for safe and reliable cross-domain robotic navigation, and opens new opportunities for deploying autonomous systems in complex multi-medium environments.

6. Conclusions

This paper proposed CD-HSSRL, a cross-domain hierarchical safe-switching reinforcement learning framework for autonomous amphibious robot navigation. By jointly optimizing cross-domain global reachability planning, hierarchical medium-switching decision-making, and safety-constrained continuous control, the proposed method effectively addresses discontinuous water-land dynamics, unstable medium transitions, and safety-critical control challenges. Comprehensive experiments on water-domain, land-domain, and cross-domain benchmarks demonstrated that CD-HSSRL consistently outperforms state-of-the-art baselines in navigation success rate, transition stability, and collision avoidance performance.

In conclusion, this study establishes a unified and reliable solution for cross-domain robotic navigation and provides new insights into joint planning-switching-safety optimization for multi-medium autonomous systems. Future work will focus on real-world deployment and multi-robot cooperative extensions to further advance cross-domain autonomous robotic applications.

Author Contributions: Shuang Liu: Conceptualization, Methodology, Framework design, Algorithm development, Writing - original draft. Lei Wei: Data curation, Benchmark construction, Experiment implementation, Validation, Visualization. Xiaoqing Li: Supervision, Project administration, Writing - review & editing, Corresponding author.

Funding: This research received no external funding.

Data Availability Statement: The datasets used in this study were sourced from three publicly available repositories: WaterScenes dataset, available at <https://github.com/WaterScenes/WaterScenes/tree/main>. Maritime visual tracking dataset, available at <https://github.com/AhsanBaidar/MVTD>. BARN ground navigation benchmark, available at https://www.cs.utexas.edu/~xiao/BARN/BARN_dataset.zip.

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

References

1. Kwe, N.B.; Priyadarshini, R. Emerging trends in mobile robots. *Robotics and Smart Autonomous Systems* **2024**, pp. 77–117.
2. Bogue, R. The role of robots in environmental monitoring. *Industrial Robot: The international journal of robotics research and application* **2023**, *50*, 369–375.
3. Narouz, A.S.; Ismail, A.; Atef, A.; Magdy, M.; Abdallah, M.; Atwa, M.; Shenoda, S.; Elsayed, M.; Ayman, S.; Ahmed, M.I. A Review of Features and Characteristics of Rescue Robot with AI. *Advanced Sciences and Technology Journal* **2024**, *1*, 1–18.
4. Muepu, D.M.; Watanobe, Y.; Naruse, K. Toward a Holistic Framework for Robotic Assessment: A Survey on Performance, Software, and Environmental Adaptability. *IEEE Access* **2025**.
5. Li, Q.; Li, H.; Shen, H.; Yu, Y.; He, H.; Feng, X.; Sun, Y.; Mao, Z.; Chen, G.; Tian, Z.; et al. An aerial-wall robotic insect that can land, climb, and take off from vertical surfaces. *Research* **2023**, *6*, 0144.
6. Wijayathunga, L.; Rassau, A.; Chai, D. Challenges and solutions for autonomous ground robot scene understanding and navigation in unstructured outdoor environments: A review. *Applied Sciences* **2023**, *13*, 9877.
7. Amundsen, H.B.; Randeni, S.; Bingham, R.C.; Civit, C.; Filardo, B.P.; Føre, M.; Kelasidi, E.; Benjamin, M.R. Hybrid State Estimation and Mode Identification of an Amphibious Robot. In Proceedings of the 2025 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2025, pp. 12696–12702.
8. Shi, L.; Zhang, Z.; Li, Z.; Guo, S.; Pan, S.; Bao, P.; Duan, L. Design, implementation and control of an amphibious spherical robot. *Journal of Bionic Engineering* **2022**, *19*, 1736–1757.
9. Zhang, D.; Van, M.; Mcllvanna, S.; Sun, Y.; McLoone, S. Adaptive safety-critical control with uncertainty estimation for human–robot collaboration. *IEEE Transactions on Automation Science and Engineering* **2023**, *21*, 5983–5996.
10. Liang, D.; Huang, X.; Xue, Z.; Li, P. Path planning for amphibious unmanned ground vehicles under cross-domain constraints. *Intelligent Service Robotics* **2025**, *18*, 1381–1416.
11. Puente-Castro, A.; Rivero, D.; Fernandez-Blanco, E.; Lamas-Lopez, F. Taxonomy of Path Planning Algorithms for Swarms of Unmanned Aerial, Ground, and Aquatic Vehicles. *Ground, and Aquatic Vehicles*.
12. Xiao, X.; Liu, B.; Warnell, G.; Stone, P. Motion planning and control for mobile robot navigation using machine learning: A survey. *Autonomous Robots* **2022**, *46*, 569–597.
13. Corsi, D.; Camponogara, D.; Farinelli, A. Aquatic navigation: A challenging benchmark for deep reinforcement learning. *arXiv preprint arXiv:2405.20534* **2024**.
14. Zhu, Y.; Wan Hasan, W.Z.; Harun Ramli, H.R.; Norsahperi, N.M.H.; Mohd Kassim, M.S.; Yao, Y. Deep reinforcement learning of mobile robot navigation in dynamic environment: A review. *Sensors* **2025**, *25*, 3394.
15. Ju, H.; Juan, R.; Gomez, R.; Nakamura, K.; Li, G. Transferring policy of deep reinforcement learning from simulation to reality for robotics. *Nature Machine Intelligence* **2022**, *4*, 1077–1087.
16. Zhong, G.; Lu, X.; Deng, T.; Cao, J. Multimodal amphibious robotics: Co-design of hybrid propulsion system and quaternion-based adaptive control for cross-domain transitions. *Control Engineering Practice* **2026**, *167*, 106644.
17. Xia, H.; Xu, Y.; Li, Z. Hybrid actuators and their reuse methodologies for amphibious robots. *Robotic Intelligence and Automation* **2025**.
18. Cuevas, J.K.; Dionisio, D.A.I.; Dris, M.K.; Flores, B.F.; Romana, C.J.S.; Bautista, A.J. Retrofitting a Commercially Available Remote Controlled Boat into an Amphibious Robot for Flood Operations Rescue Surveillance (FIOReS) Assistance. In Proceedings of the 2024 9th International Conference on Control and Robotics Engineering (ICCRE). IEEE, 2024, pp. 39–44.
19. Policarpo, H.; Lourenço, J.P.; Anastácio, A.M.; Parente, R.; Rego, F.; Silvestre, D.; Afonso, F.; Maia, N.M. Conceptual design of an unmanned electrical amphibious vehicle for ocean and land surveillance. *World Electric Vehicle Journal* **2024**, *15*, 279.
20. Zhang, K.; Ye, Y.; Chen, K.; Li, Z.; Li, K. Enhanced AUV autonomy through fused energy-optimized path planning and deep reinforcement learning for integrated navigation and dynamic obstacle detection. *Journal of Marine Science and Engineering* **2025**, *13*, 1294.
21. Zhu, A.; Zhao, J.; Yang, L. Multimodal magnetic miniature robot for adaptive navigation in amphibious environments. *npj Robotics* **2025**, *3*, 42.

22. Duan, M. Attention-based multi-agent reinforcement learning for traffic flow stability in mountainous tunnel entrances. *Scientific Reports* **2025**, *15*, 37278.
23. Politi, E.; Stefanidou, A.; Chronis, C.; Dimitrakopoulos, G.; Varlamis, I. Adaptive deep reinforcement learning for efficient 3D navigation of autonomous underwater vehicles. *IEEE Access* **2024**.
24. Mackay, A.K.; Riazuelo, L.; Montano, L. RL-DOVS: Reinforcement learning for autonomous robot navigation in dynamic environments. *Sensors* **2022**, *22*, 3847.
25. Eppe, M.; Gumbsch, C.; Kerzel, M.; Nguyen, P.D.; Butz, M.V.; Wermter, S. Intelligent problem-solving as integrated hierarchical reinforcement learning. *Nature Machine Intelligence* **2022**, *4*, 11–20.
26. Li, H.; Luo, B.; Song, W.; Yang, C. Predictive hierarchical reinforcement learning for path-efficient mapless navigation with moving target. *Neural Networks* **2023**, *165*, 677–688.
27. Schneider, T.; Pedrosa, M.V.; Gros, T.P.; Wolf, V.; Flaßkamp, K. Motion Primitives as the Action Space of Deep Q-Learning for Planning in Autonomous Driving. *IEEE Transactions on Intelligent Transportation Systems* **2024**.
28. Xiao, W.; Wang, T.H.; Hasani, R.; Chahine, M.; Amini, A.; Li, X.; Rus, D. Barriernet: Differentiable control barrier functions for learning of safe robot control. *IEEE Transactions on Robotics* **2023**, *39*, 2289–2307.
29. Yao, S.; Guan, R.; Wu, Z.; Ni, Y.; Huang, Z.; Liu, R.W.; Yue, Y.; Ding, W.; Lim, E.G.; Seo, H.; et al. WaterScenes: A Multi-Task 4D Radar-Camera Fusion Dataset and Benchmarks for Autonomous Driving on Water Surfaces. *IEEE Transactions on Intelligent Transportation Systems* **2024**, *25*, 16584–16598. <https://doi.org/10.1109/TITS.2024.3415772>.
30. Bakht, A.B.; Din, M.U.; Javed, S.; Hussain, I. MVTD: A Benchmark Dataset for Maritime Visual Object Tracking. *arXiv preprint arXiv:2506.02866* **2025**.
31. Perille, D.; Truong, A.; Xiao, X.; Stone, P. Benchmarking Metric Ground Navigation. In Proceedings of the 2020 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR). IEEE, 2020.
32. Jiang, W.; Liu, J.; Wang, W.; Wang, Y. Global Path Planning for Land–Air Amphibious Biomimetic Robot Based on Improved PPO. *Biomimetics* **2026**, *11*, 25.
33. Xiaofei, Y.; Yilun, S.; Wei, L.; Hui, Y.; Weibo, Z.; Zhengrong, X. Global path planning algorithm based on double DQN for multi-tasks amphibious unmanned surface vehicle. *Ocean Engineering* **2022**, *266*, 112809.
34. Yin, S.; Xiang, Z. Energy-constrained collaborative path planning for heterogeneous amphibious unmanned surface vehicles in obstacle-cluttered environments. *Ocean Engineering* **2025**, *330*, 121241.
35. Yin, S.; Hu, J.; Xiang, Z. Multi-objective collaborative path planning for multiple water-air unmanned vehicles in cramped environments. *Expert Systems with Applications* **2025**, *292*, 128625.
36. Zhang, N.; Chen, Y.; Wu, Y.; Ji, M.; Wang, B. A hybrid APF-DQN framework with transformer-based current prediction for USV path planning in dynamic ocean environments. *Scientific Reports* **2025**.
37. Hua, M.; Zhou, W.; Cheng, H.; Chen, Z. Improved DDPG algorithm-based path planning for unmanned surface vehicles. *Intelligence & Robotics* **2024**, *4*, 363–384.
38. Yang, C.; Zhao, Y.; Cai, X.; Wei, W.; Feng, X.; Zhou, K. Path planning algorithm for unmanned surface vessel based on multiobjective reinforcement learning. *Computational Intelligence and Neuroscience* **2023**, *2023*, 2146314.
39. Fan, Y.; Sun, Z.; Wang, G. A novel reinforcement learning collision avoidance algorithm for USVs based on maneuvering characteristics and COLREGs. *Sensors* **2022**, *22*, 2099.
40. Hu, H.; Wang, Y.; Tong, W.; Zhao, J.; Gu, Y. Path planning for autonomous vehicles in unknown dynamic environment based on deep reinforcement learning. *Applied Sciences* **2023**, *13*, 10056.
41. Liang, C.; Liu, L.; Liu, C. Multi-UAV autonomous collision avoidance based on PPO-GIC algorithm with CNN–LSTM fusion network. *Neural Networks* **2023**, *162*, 21–33.

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