

Q러닝 기반 지형중심의 지상 전장환경 경로계획 기법

Q-learning-Based Path Planning with Terrain-Aware Concealment in Ground Battlefield Environments

이지선^{*,1)} . 황현호²⁾

Jisun LEE^{*,1)} . Hyunho Hwang²⁾

[초 록]

전장 환경에서의 경로계획은 높은 불확실성과 제한된 환경 정보, 그리고 전술적 생존성 확보라는 요구로 인해 일반적인 경로 탐색 문제와 본질적으로 다른 특성을 지닌다. 기존의 경로계획 연구는 주로 장애물 회피와 최단 거리 탐색에 초점을 두어 왔으나, 이러한 접근은 지상 전장환경에서 요구되는 은폐와 엄폐를 충분히 반영하지 못하는 한계가 있다.

본 논문에서는 지형 특성을 전술적 자원으로 활용하여 은폐와 엄폐가 가능한 경로를 탐색하는 Q-learning 기반 경로계획 기법을 제안한다. 전장 환경은 다양한 지형 요소를 포함하는 이차원 격자 지도로 모델링되며, 에이전트는 이동 거리와 탐색 효율성뿐만 아니라 지형의 은폐·엄폐 효과를 고려한 보상 함수를 통해 학습을 수행한다. 제안한 기법은 장애물을 단순히 회피하는 대상이 아닌, 전술적 이동을 지원하는 요소로 활용함으로써 전장 환경의 특수성을 반영한다.

서로 다른 목표 지점을 갖는 다수의 실험 시나리오를 통해 제안한 기법의 성능을 검증한 결과, 복잡한 전장 환경에서도 높은 성공률과 안정적인 경로 탐색 성능을 확인하였다. 특히 경로 길이가 증가하더라도 은폐·엄폐 지형을 우선적으로 활용하는 전략적 경로가 학습됨을 확인하였다. 본 연구는 지상 전장환경에서 자율 이동체 및 군사용 로봇의 생존성과 임무 수행 능력을 향상시키기 위한 실질적인 경로계획 프레임워크를 제시한다.

[ABSTRACT]

Path planning in ground battlefield environments requires more than finding the shortest path, as safety and survivability are critical under uncertain and hostile conditions. Conventional path planning methods mainly focus on obstacle avoidance, which is insufficient for battlefield scenarios where terrain can be used for concealment and cover.

This paper proposes a Q-learning-based path planning method that exploits terrain features to enable safe and tactical movement in ground battlefield environments. The environment is modeled as a two-dimensional grid map with various terrain types. A reward function is designed to consider both path efficiency and terrain-based concealment and cover, allowing the agent to learn strategic paths rather than purely shortest routes.

Experimental results from multiple target scenarios show that the proposed method achieves high success rates and stable performance, even in complex environments. The results demonstrate that terrain-aware reinforcement learning is effective for path planning in battlefield environments and can support autonomous ground vehicles and military robotic systems.

Key Words : Path planning(경로계획), Q-learning(Q러닝), Reinforcement learning(강화학습), Ground battlefield environments(지상 전장환경), Concealment and concealment(은폐 및 엄폐)

1. 서론

일반적으로 경로 탐색(path planning)은 시작점에서 목표 지점까지 이동하는 과정에서 장애물과의 충돌을 회피하면서 이동 거리 및 소요 시간을 최소화하는 경로를 탐색하는 문제로 정의된다. 이러한 경로 탐색 기법은 환경 정보가 완전히 알려진 경우부터 부분적으로만 알려지거나 전혀 알려지지 않은 경우까지 다양한 조건을 대상으로 연구되어왔다. 그러나 환경에 대한 정보가 제한되거나 불확실성이 증가할수록 경로 탐색의

난이도는 급격히 상승하며, 기존의 최적화 중심 접근 방식만으로는 효과적인 해결이 어려워진다.

1) 한국국방기술학회(The Korean Institute of Defense Technology) 2) 한양사이버대학교 국방융합학과(Hanyang Cyber University, Korea)

* Corresponding author, E-mail: jisun0319@gmail.com
Copyright © The Korean Institute of Defense Technology
Received : December 23, 2025 Revised : December 30, 2025
Accepted : December 30, 2025

특히 전장 환경에서의 경로 탐색은 일반적인 민간 환경과 본질적으로 다른 특성을 지닌다. 전장에서는 적군의 위치와 위협요소를 사전에 정확히 파악하기 어렵기 때문에, 단순히 목적지까지의 최단 경로나 최단 시간 경로를 선택하는 것은 오히려 생존성과 임무 성공률을 저하시킬 수 있다. 따라서 은폐와 엄폐를 고려한 안전한 이동 경로의 확보가 무엇보다 중요하다. 목적지에 신속히 도달하는 것뿐만 아니라, 무기의 연료 소모, 병사의 체력 유지, 보급 자원의 효율적 사용 등 전력 유지 요소 또한 경로 계획 단계에서 함께 고려되어야 한다. 이러한 이유로 전장 환경에서는 때로는 더 길고 느린 경로라 하더라도, 적의 탐지를 회피하고 전력을 보존할 수 있는 전략적 경로가 선택될 필요가 있다. 결과적으로 전장 환경에서의 경로 탐색은 전술적 안전성과 임무 수행 효율성을 동시에 만족해야 하는 복합적인 문제라 할 수 있다.

지상 전장 환경에서의 경로 탐색은 지형의 경사, 단차, 요철 등으로 인해 더욱 복잡한 문제로 확장된다. 특히 로봇이나 전차와 같은 지상 이동체가 비정형 지형을 통과해야 하는 경우, 실행 가능한 경로를 효율적으로 계획하는 것은 매우 어려운 과제이다. 숲, 하천, 웅덩이 등 다양한 자연 지형이 존재하고, 지도 정보가 불완전하거나 존재하지 않는 전장 환경에서는 조건의 불확실성이 더욱 커지며, 기존의 장애물 회피 중심 경로 탐색 기법은 한계를 드러낸다.

기존의 경로 탐색 연구에서는 장애물을 반드시 회피해야 할 요소로 간주해 왔다. 그러나 전장 환경에서는 이러한 접근이 항상 최선의 선택이 아닐 수 있다. 예를 들어 전차가 적 기지로 이동하는 상황을 가정할 때, 도로를 따라 이동하면 빠르게 이동할 수 있지만 적에게 노출될 위험이 커진다. 반면, 필요에 따라 건물이나 구조물을 활용하거나, 경우에 따라 파괴하며 이동하는 전략적 선택이 가능하다. 즉, 전장 환경에서의 경로 탐색은 장애물을 단순히 회피 대상이 아닌, 은폐와 엄폐를 제공하는 전략적 요소로 활용해야 한다는 특수성을 지닌다.[1]

이에 본 논문에서는 지상 전장 환경에서 은폐와 엄폐를 가능하게 하는 새로운 경로 탐색 기법을 제안한다. 제안하는 접근 방식은 장애물을 회피하는 기존의 경로 탐색 개념에서 벗어나, 장애물을 은폐와 엄폐의 수단으로 활용하여 경로를 탐색하는데 초점을 둔다. 또한 숲, 강, 웅덩이 등 다양한 지형 요소를 함께 고려함으로써, 복잡하고 불확실한 지상 전장 환경에서도 안전하고 전략적인 이동이 가능하도록 설계하였다. 본 연구는 지상 전장 환경의 특수성을 반영한 경로 탐색 방법을 제시함으로써, 향후 군사 로봇 및 자율 전투체계의 생존성과 임무 수행 능력 향상에 기여하고자 한다.

2. Q러닝

Q러닝은 강화학습 기법 중 하나로, 환경에 대한 사전 모델이 존재하지 않는 상황에서도 에이전트가 시행착오를 통해 최적의 행동 정책을 학습할 수 있는 대표적인 모델 프리 학습 알고리즘이다. Q러닝의 목표는 주어진 상태(state)에서 특정 행동(action)을 수행했을 때 얻을 수 있는 누적 보상의 기댓값을 나타내는 상태-행동 가치 함수(state-action value

function), 즉 Q-함수를 추정하는 것이다.

에이전트는 현재 상태에서 행동을 선택하고, 그 결과로 다음 상태와 보상을 관측한다. 이러한 상호작용을 반복함으로써 Q러닝은 벨만 최적 방정식에 기반하여 Q-값을 점진적으로 갱신하며, 최종적으로 최적 정책을 수렴한다. Q-함수의 갱신은 다음과 같은 식으로 표현된다.

$$Q(s,a) \leftarrow Q(s,a) + \alpha [r + \gamma \max_{a'} Q(s',a') - Q(s,a)] \quad (1)$$

여기서 s 와 a 는 각각 현재 상태와 선택한 행동을 의미하며, r 은 해당 행동에 대한 보상, s' 는 다음 상태를 나타낸다. α 는 학습률(learning rate)로서 새로운 정보가 기존 Q-값에 반영되는 정도를 조절하고, γ 는 할인율(discount factor)로 미래 보상의 중요도를 결정한다.

Q러닝의 중요한 특징 중 하나는 탐험(exploration)과 활용(exploitation) 간의 균형을 통해 학습을 수행한다는 점이다. 일반적으로 ϵ -greedy 전략과 같은 정책을 사용하여, 일정 확률로 임의의 행동을 선택함으로써 미지의 상태 공간을 탐색하고, 나머지 확률로는 현재까지 학습된 Q-값을 기반으로 최적 행동을 선택한다. 이를 통해 에이전트는 국소 최적해에 머무르지 않고 전역 최적 정책을 학습할 수 있다.[2,3]

Q러닝은 환경 모델이 명확하지 않거나 동적으로 변화하는 문제에 효과적으로 적용될 수 있어, 로봇 제어, 게임 AI, 자율주행, 경로 계획과 같은 분야에서 폭넓게 활용되어왔다. 특히 경로 계획 문제에서는 각 위치를 상태로, 이동 방향을 행동으로 정의하고, 목표 지점 도달, 충돌 회피, 이동 비용 등을 보상 함수로 설계함으로써 에이전트가 최적의 이동 경로를 학습하도록 유도할 수 있다.

전장 환경과 같이 불확실성이 높고, 사전 정보가 제한적인 상황에서는 Q러닝의 모델 프리 특성이 큰 장점으로 작용한다. 에이전트는 완전한 환경 지도를 요구하지 않으며, 반복적인 학습을 통해 은폐와 엄폐, 위험 회피, 이동 효율성 등 다양한 전술적 요소를 보상 함수에 반영할 수 있다. 이러한 특성으로 인해 Q러닝은 지상 전장 환경에서의 경로 탐색 문제를 해결하기 위한 유효한 학습 프레임워크로 활용될 수 있다.

3. 문제정의

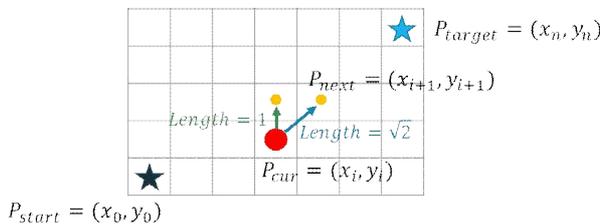
본 연구의 목적은 지형 조건을 고려하여 지상 전장 환경에서 출발 지점으로부터 목표 지점까지 이동하는 최적의 경로(path)를 탐색하는 데 있다. 특히 단순한 최단 경로 탐색이 아닌, 지형지물을 은폐와 엄폐의 수단으로 활용할 수 있는 경로 계획을 제안하는 것을 목표로 한다. 이를 통해 전장 환경에서의 생존성과 전술적 이동 효율성을 동시에 향상시키고자 한다. 전장 환경은 시작 지점과 목표 지점, 그리고 다양한 지형지물을 포함하는 이차원 격자 지도로 모델링하였다. 지도는 총 50×50 크기의 격자로 구성되며, 각 격자 셀은 하나의 지형 유형을 나타낸다. 표 1은 지도상에 표현된 지형의 종류와 이에 대응하는 코드 값을 나타낸다.

3.1 경로 길이 정의

본 연구에서 경로 길이는 출발 지점(start point)에서 목표 지점(target point)까지 이동한 전체 거리의 합으로 정의한다. 경로 길이는 각 상태 간 이동 거리의 누적 합으로 계산되며, 이는 다음 식 (2)과 같이 표현된다.

$$L_{path} = \sum_{i=0}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \quad (2)$$

여기서 $i = 0, 1, \dots, n$ 은 경로상의 상태 인덱스를 의미한다. $i = 0$ 일 때의 위치는 출발 지점으로, $P_{start} = (x_0, y_0)$ 이며, $i = n$ 일때의 위치는 목표지점으로 $P_{target} = (x_n, y_n)$ 이다. 또한, 현재 상태 위치는 $P_{cur} = (x_i, y_i)$, 다음 상태 위치는 $P_{next} = (x_{i+1}, y_{i+1})$ 로 정의 된다. 그림 *은 이러한 경로 길이 계산 개념을 시각적으로 나타낸다. [4]



3.2 상태와 행동 정의

본 연구에서는 강화학습 기반 경로 탐색을 위해 전장 환경을 상태(state)와 행동(action)의 집합으로 정의한다. 상태는 에이전트가 위치한 격자 좌표를 의미하며, 각 상태는 이차원 좌표로 표현된다. 즉, 상태 s_i 는 다음과 같이 정의된다.

$$s_i = (x_i, y_i)$$

여기서 (x_i, y_i) 는 50×50 격자 지도 상에서 에이전트의 현재 위치를 나타낸다. 각 상태는 해당 위치의 지형 정보(평지, 숲, 건물, 강 등)를 포함하며, 이를 통해 지형 특성에 따른 보상 설계가 가능하도록 하였다.

행동(action)은 에이전트가 현재 상태에서 다음 상태로 이동하기 위해 선택할 수 있는 이동 방향을 의미한다. 본 연구에서는 상, 하, 좌, 우 및 대각선 방향을 포함한 총 8개의 행동 집합을 정의하였다. 행동 집합 A 는 다음과 같이 표현된다.

$$A = \{a_1, a_2, \dots, a_8\}$$

단, 선택한 이동 방향에 장애물(파괴 불가 건물, 강 등)이 존재하는 경우 해당 행동은 불가능한 행동으로 간주되며, 에이전트는 가능한 행동 집합 내에서만 다음 행동을 선택할 수 있다. 이를 통해 물리적으로 실행 불가능한 경로가 학습되는 것을 방지한다.

3.3 보상함수 정의

전장 환경에서의 경로 탐색은 단순한 최단 경로 문제가 아니라, 은폐와 엄폐를 고려한 전술적 이동을 요구한다. 이에 본

연구에서는 보상 함수(reward function)를 설계할 때 이동 효율성뿐만 아니라 지형의 전술적 가치가 반영되도록 구성하였다.

에이전트가 목표 지점에 도달한 경우에는 높은 양의 보상을 부여하여 목표 도달을 학습의 최우선 목표로 설정하였다. 반면, 이동 과정에서 불리한 지형(예: 적에게 노출되기 쉬운 개활지, 이동 비용이 큰 지형)을 통과하는 경우에는 음의 보상을 부여하였다. 반대로 숲이나 건물과 같이 은폐 및 엄폐가 가능한 지형을 통과할 경우에는 상대적으로 높은 보상을 부여하여, 에이전트가 자연스럽게 안전한 경로를 학습하도록 유도하였다. 이를 일반화하여, 보상 함수 $r(s,a)$ 는 다음 요소들을 종합적으로 고려하여 정의된다.

- (1) 목표 지점 도달 여부, (2) 이동에 따른 거리, (3) 지형 유형에 따른 은폐·엄폐 효과, (4) 이동 불가능 영역 진입 시 패널티와 같은 보상 설계를 통해 에이전트는 단순히 짧은 경로가 아닌, 전장 환경에서 생존성과 전술적 우위를 확보할 수 있는 경로를 학습하게 된다.

4. 실험 결과 및 분석

본 절에서는 제안한 Q러닝 기반 은폐·엄폐 경로계획 기법의 성능을 검증하기 위해, 전장 환경에서 요구되는 이동 거리, 목표 방향, 지형 복잡도 차이를 반영한 세 가지 실험 시나리오를 구성하고 그 결과를 분석한다. 모든 실험에서 출발 지점은 (0,0)(0,0)으로 동일하게 설정하였으며, 각 시나리오별로 목표 지점의 위치를 달리하여 경로 길이, 탐색 난이도, 은폐·엄폐 지형 활용 양상이 서로 다르게 나타나도록 설계하였다.

시나리오 1(Target (28,29))은 중간 거리 이동을 요구하는 환경으로, 비교적 짧은 이동 거리와 균형적인 수평·수직 이동이 혼합된 조건을 갖는다. 이는 제안한 기법이 기본적인 전장 환경에서 안정적으로 수렴하는지를 검증하기 위한 기준 시나리오로 설정하였다.

시나리오 2(Target (38,48))는 세 시나리오 중 가장 장거리 이동이 요구되며, 이동 과정에서 다양한 지형을 광범위하게 통과해야 하는 조건을 갖는다. 이 시나리오는 경로 길이 증가 및 탐색 공간 확장에 따른 학습 안정성, 성공률 변화, 그리고 은폐·엄폐 기반 경로 선택 전략의 유효성을 평가하기 위해 구성되었다. 특히 복잡한 지형 분포로 인해 탐색 난이도가 가장 높은 환경에 해당한다.

시나리오 3(Target (45,16))는 장거리 이동 조건이지만, 상대적으로 수평 이동 비중이 높은 목표 지점을 설정한 환경이다. 이를 통해 목표 방향과 지형 분포에 따라 은폐·엄폐 지형 활용 방식이 어떻게 달라지는지, 그리고 동일한 장거리 조건에서도 탐색 효율성과 성공률이 어떻게 변화하는지를 분석하고자 하였다.

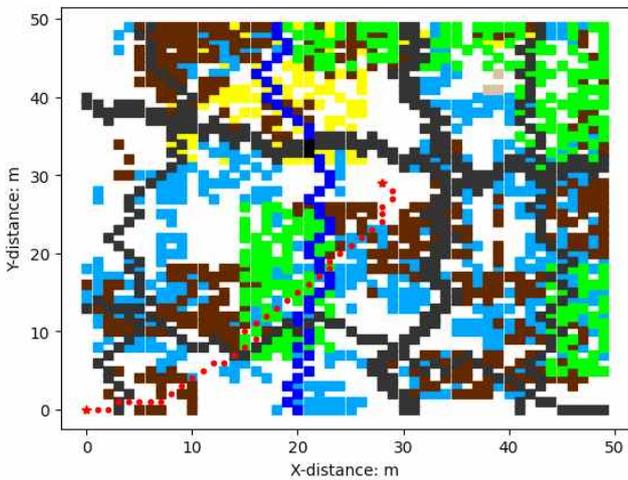
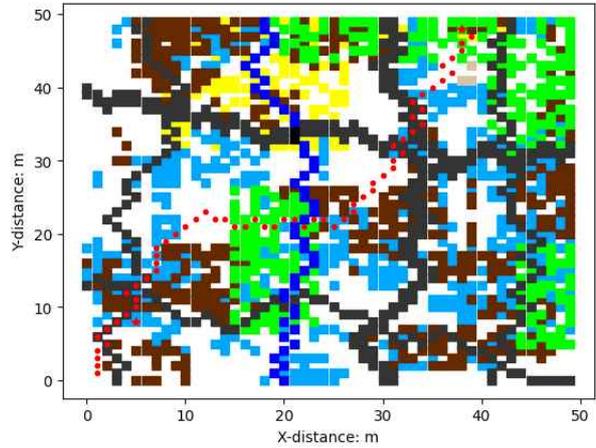
이와 같이 세 가지 시나리오는 단순히 목표 지점을 임의로 변경한 것이 아니라, 이동 거리(중·장거리), 목표 방향(대각·수평), 지형 복잡도 차이를 단계적으로 반영하도록 설계되었다. 이를 통해 제안한 경로계획 기법이 다양한 전장 환경 조건에서도 일관된 성능과 전략적 경로 선택 능력을 보이는지를 종합적으

로 검증하고자 하였다.

4.1 시나리오 1: Target (28, 29)

첫 번째 시나리오는 비교적 중간 거리의 목표 지점 (28,29)을 설정한 환경이다. 실험 결과, 평균 경로 길이는 46.113, 평균 탐색 시간은 8.374, 평균 성공률은 97.677%로 나타났다. 표준편차는 각각 1.85, 1.12, 0.74로 매우 낮아, 반복 실험 간 결과의 일관성이 높음을 확인할 수 있다.

이는 비교적 단순한 환경에서도 제안한 경로계획 기법이 안정적으로 수렴하며, 은폐·엄폐 지형을 활용한 전략적 경로 선택이 경로 품질과 성공률 측면에서 효과적으로 작동함을 의미한다. 특히 높은 성공률은 목표 지점 도달뿐만 아니라 장애물 회피 및 안전한 이동이 동시에 달성되었음을 보여준다.



4.2 시나리오 2: Target (38, 48)

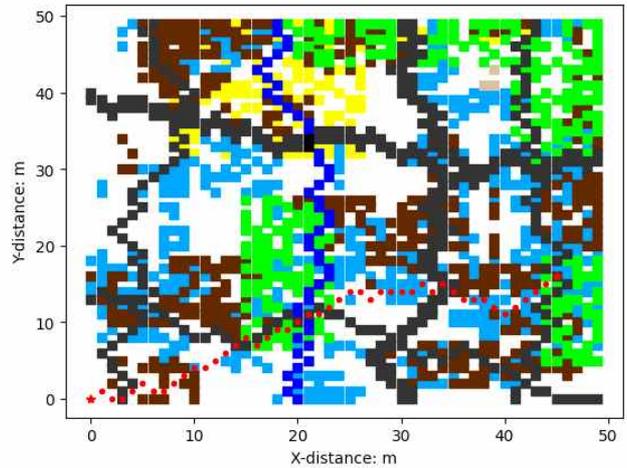
두 번째 시나리오는 상대적으로 장거리 이동이 요구되는 목표 지점 (38,48)을 설정한 경우로, 전반적으로 가장 복잡한 환경에 해당한다. 이 경우 평균 경로 길이는 67.768, 평균 탐색 시간은 26.245, 평균 성공률은 85.934%로 나타났다.

경로 길이와 탐색 시간이 증가함에 따라 성공률이 감소하는 경향을 보였으며, 특히 탐색 시간의 표준편차(7.47)가 크게 나타났다. 이는 장거리 이동 과정에서 은폐·엄폐 지형을 우선적으로 선택하는 전략이 경로의 다양성을 증가시키고, 일부 실험에서는 더 많은 탐색이 요구되었기 때문으로 해석된다. 그럼에도 불구하고 약 86% 수준의 평균 성공률을 유지한 것은, 불확실성과 위험 요소가 높은 환경에서도 제안한 기법이 일정 수준 이상의 안정적인 성능을 보장함을 의미한다.

4.3 시나리오 3: Target (45, 16)

세 번째 시나리오는 장거리이지만 상대적으로 수평 이동 비중이 큰 목표 지점 (45,16)을 설정하였다. 실험 결과, 평균 경로 길이는 57.473, 평균 탐색 시간은 11.489, 평균 성공률은 95.429%로 측정되었다.

특히 탐색 시간과 성공률 모두에서 시나리오 2에 비해 크게 개선된 결과를 보였으며, 이는 지형 분포와 목표 방향에 따라 은폐·엄폐 경로를 보다 효율적으로 활용할 수 있었기 때문으로 판단된다. 성공률의 표준편차 또한 1.56으로 낮아, 제안 기법의 학습 안정성과 재현성이 우수함을 확인할 수 있다.



5.4 종합 분석

세 가지 실험 시나리오의 결과를 종합적으로 분석한 결과, 제안한 Q-learning 기반 경로계획 기법은 전장 환경에서 요구되는 안전성, 안정성, 그리고 전략적 이동 측면에서 효과적인 성능을 보였다. 목표 지점까지의 거리가 증가하고 환경이 복잡해질수록 경로 길이와 탐색 시간은 증가하는 경향을 보였으나, 은폐·엄폐를 고려한 보상 설계를 통해 비교적 높은 성공률을 유지할 수 있었다.

특히 단순한 최단 경로 탐색이 아닌, 위험 노출을 최소화하는 경로를 학습하도록 유도함으로써 일부 시나리오에서는 경로 길이가 증가하더라도 성공률과 안정성이 확보되는 특징을 확인하였다. 이는 전장 환경에서 요구되는 전술적 이동 특성을 잘 반영한 결과로, 제안 기법이 실제 지상 전투 환경 및 자율 지상 플랫폼에 적용 가능성이 있음을 시사한다.

5. 결론

본 논문에서는 지상 전장 환경의 특수성을 반영한 은폐·엄폐 기반 경로계획 기법을 제안하고, 이를 Q-learning 기반 강화 학습 프레임워크로 구현하였다. 기존의 경로 탐색 연구가 장애물을 반드시 회피해야 할 요소로 간주한 반면, 본 연구는 지형 지물을 전술적 자원으로 활용하여 생존성과 임무 성공률을 동시에 고려한 경로 탐색을 수행한다는 점에서 차별성을 갖는다. 제안한 기법은 전장 환경을 50×50 이차원 격자 지도로 모델링하고, 다양한 지형 요소를 포함한 상태-행동 공간을 정의하였다. 또한 보상 함수 설계를 통해 이동 효율성뿐만 아니라 은폐 및 엄폐가 가능한 지형 선택을 강화학습 과정에 반영함으로써, 에이전트가 단순 최단 경로가 아닌 전략적 경로를 학습하도록 유도하였다.

실험 결과, 서로 다른 목표 지점을 갖는 세 가지 시나리오 모두에서 제안 기법은 안정적인 성능을 보였다. 비교적 단거리 환경에서는 평균 97% 이상의 높은 성공률과 낮은 분산을 기록하였으며, 장거리 및 복잡한 환경에서도 약 86% 이상의 성공률을 유지함으로써 불확실성이 높은 전장 환경에서도 효과적으로 경로를 탐색할 수 있음을 확인하였다. 특히 경로 길이와 탐색 시간이 증가하는 경우에도 은폐·엄폐 지형을 활용한 경로 선택을 통해 전술적 안전성을 확보할 수 있음을 실험적으로 입증하였다.

본 연구의 결과는 지상 전투 환경에서 자율 이동체, 전차, 군사용 로봇 등의 경로 계획에 있어 단순한 이동 최적화가 아닌 전술적 의사결정을 포함한 경로 탐색이 필요함을 시사한다. 제안한 접근 방식은 실제 전장 환경에서의 위험 노출 최소화, 전력 보존, 임무 지속성 향상에 기여할 수 있을 것으로 기대된다.

다만 본 연구에서는 정적인 지도 환경을 기반으로 실험을 수행하였으며, 적의 이동이나 위협 요소의 실시간 변화는 고려하지 못했다는 한계가 있다. 향후 연구에서는 동적 전장 환경을 반영한 강화학습 모델 확장, 심층 강화학습(DQN, A3C 등)을 활용한 대규모 상태 공간 처리, 그리고 적 위협 확률 모델을 결합한 경로계획 기법으로의 확장이 필요하다. 이를 통해 보다 현실적이고 고도화된 전장 환경에서도 적용 가능한 자율 경로 탐색 체계를 구축할 수 있을 것으로 기대된다.

References

- [1] 김민경. (2024). Ray RLlib 기반 QMIX 와 RND 를 이용한 희소 보상 전장 환경에서의 멀티에이전트 강화학습 협업. 한국컴퓨터정보학회논문지, 29(1), 11-19.
- [2] 김호원, & 이원창. (2020). Q-learning을 이용한 이동 로봇의 실시간 경로 계획. 전기전자학회논문지, 24(4), 71-77.
- [3] 용성중, 박효경, 유연휘, & 문일영. (2022). 효율적인 경로 선택을 위한 Q-Learning 정책 및 보상 설계. 한국향행학회논문지, 26(2), 72-77.
- [4] Lee, J., & Seo, Y. (2024). Q-learning based on

strategic artificial potential field for path planning enabling concealment and cover in ground battlefield environments. Applied Intelligence, 54(13), 7170-7200.