

도심 디지털 가상환경에서 UAV를 활용한 심층 강화학습 기반 차량 검출 및 추적 시스템

박채원, 임지혜, 이승준, 유상조
인하대학교

cwp2002@inha.edu, wlgP12211827@inha.edu, pank147@inha.edu, sjyoo@inha.ac.kr

Deep Reinforcement Learning–Based Vehicle Detection and Tracking Using UAVs in Urban Digital Environments

Chae–Won Park, Ji–Hye Lim, Seung–Jun Lee and Sang–jo Yoo
Inha Univ.

요 약

본 연구는 무인항공기(UAV)가 복잡한 도시 환경에서 차량을 실시간으로 검출하고 안정적으로 추적할 수 있는 시스템을 제안한다. 이를 위해 합성곱 신경망(CNN)을 이용한 객체 검출 모델과 심층 Q-네트워크(DQN)를 결합한 프레임워크를 구성하였다. 객체 검출 모델은 Self-Attention 모듈과 Heatmap-Regression 병렬 구조를 통해 차량의 중심 위치와 Bounding Box 를 정밀하게 추정하며, DQN 기반 추적기는 차량의 이동 패턴을 학습하고 다양한 환경 변화에 대응할 수 있도록 설계되었다. 제안된 시스템은 Unity ML–Agents 를 활용한 도시 시뮬레이션 환경에서 평가되었으며, 터널, 고가도로, 다양한 날씨 조건 등 복잡한 환경에서도 높은 탐지율과 안정적인 추적 성능을 입증하였다.

I. 서론

무인항공기(UAV, Unmanned Aerial Vehicle)는 물류, 재난 구조, 농업, 감시, 군사 작전 등 다양한 산업 분야에서 활용도가 높아지고 있으며, 특히 자율 비행과 실시간 객체 추적 기술은 핵심 요소로 주목받고 있다[1–4]. UAV가 차량이나 사람과 같은 이동 객체를 지속적으로 감지하고 추적하는 능력은 정보 수집과 임무 수행의 효율성을 크게 좌우한다. 그러나 기존의 규칙 기반(Rule-based) 객체 탐지 시스템은 환경 변화에 유연하게 대응하지 못하는 한계를 지니며, 특히 차량이 비선형 경로를 따라 이동하거나 터널 및 장애물로 가려지는 상황에서 추적 성능이 현저히 저하된다. 이에 본 연구는 합성곱 신경망(CNN, Convolutional Neural Network)을 활용하여 객체를 정확히 검출하고, 심층 강화학습(Deep Reinforcement Learning)의 일종인 심층 Q 네트워크(DQN, Deep Q Network) 알고리즘을 활용하여 UAV가 차량의 이동 패턴을 학습하고 복잡한 환경에서도 적용할 수 있는 방법을 제안한다. 본 연구의 목표는 객체 검출과 추적 기술을 결합한 접근법을 통해 안정적인 실시간 객체 검출과 이동 예측이 가능한 UAV 추적 기술을 구현함으로써, 다양한 산업 분야에서의 활용성을 극대화하는 데 있다.

II. 심층 강화학습을 이용한 차량 검출 및 추적 방법 제안

본 장에서는 UAV 기반 차량 검출 및 추적을 위해, 객체 검출 모델과 DQN 알고리즘을 통합한 방법론을 제안한다. 제안하는 전체 시스템 아키텍처는 그림 1과 같이 구성한다.

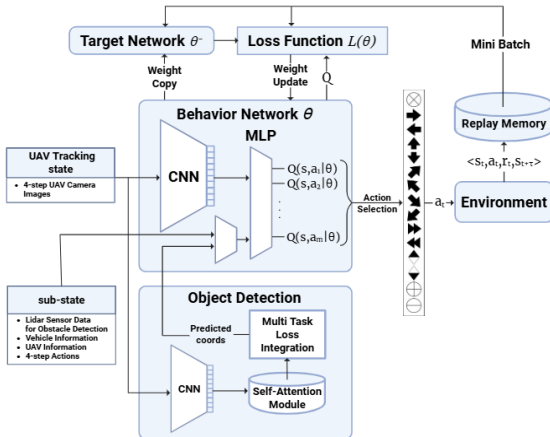


그림 1. 제안하는 차량 검출 및 추적 모델의 전체 구조

2-1. 차량 검출 모델

제안된 객체 검출 모델은 CNN 기반 특징 추출기, Self-Attention 모듈, 그리고 다중 분기(Multi-Branch) 구조로 구성된다. 입력 이미지는 단단계 CNN 계층을 통해 공간적 특징을 추

출하고, Self-Attention 모듈을 통해 영상 내 다양한 객체 간의 장거리 관계를 학습한다. 이를 통해 차량이 부분적으로 가려지거나 복잡한 배경과 혼재된 상황에서도 강인한 특징 표현이 가능하다. 차량의 중심 위치는 Heatmap 분기를 통해 검출되며, Sigmoid 기반의 출력과 Binary Cross-Entropy 손실 함수를 통해 학습된다. Heatmap은 식 (1)과 같은 2차원 Gaussian 분포를 따른다.

$$H(x, y) = \exp\left(-\frac{(x - x^c)^2 + (y - y^c)^2}{2\sigma^2}\right) \quad (1)$$

(x^c, y^c) 는 객체 중심 좌표를 의미하며, σ 는 분산을 나타낸다. 중심에서 멀어질수록 $H(x, y)$ 값은 0에 수렴하고, 중심에 가까울수록 1에 가까운 값을 갖는다. 이를 통해 모델은 중심 위치를 빠르고 정확하게 수렴하도록 학습된다. 한편, 회귀(Regression) 분기는 Self-Attention을 거친 전역 특징 벡터를 입력으로 받아 차량의 Bounding Box 좌표를 직접 회귀한다. 최종적으로 중심 예측과 Bounding Box 회귀를 동시에 수행하는 다중 과제 학습(Multi-Task Learning) 구조를 통해 정확한 객체 검출 성능을 달성한다.

2-2. Deep Q-Network 기반 차량 추적 시스템

DQN 알고리즘을 기반으로 UAV의 차량 추적 시스템을 구성한다. 그림 1에서 환경(Environment)은 도시의 도로 환경을 의미한다. UAV의 카메라와 라이다 센서를 통해 건물, 도로, 차량 그리고 날씨 등 도심 상황에 대한 데이터를 수집한다. 객체 검출 모델에서 얻은 차량의 예측 정보와 도심 상황 정보를 통합하여 상태(s_t)로 입력한다. Behavior Network는 상태 정보를 바탕으로 Q값을 계산하고, 그 중 가장 큰 값을 가지는 행동(a_t)을 선택하여 차량을 추적한다.

Q 값은 식(2)와 같이 정의된다. r_t 는 시점 t 에서 수행한 행동에 대한 보상이고, γ 는 미래 시점에 대한 보상을 조절하는 할인 인자이다. UAV가 행동하여 변화한 환경 상태(s_{t+1})를 획득하고 보상(r_t)을 계산한다. 한 스텝 동안 얻은 $\langle s_t, a_t, r_t, s_{t+1} \rangle$ 튜플을 Replay Memory에 저장하고, 무작위로 추출된 Mini Batch는 경사 하강법(Gradient Descent) 기반 학습을 수행한다.

$$Q(s_t, a_t) = r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) \quad (2)$$

손실 함수는 식 (3)과 같이 정의된다. 손실은 Target Network θ^- 로부터 계산된 목표 Q값과 현재 Q값 간의 오차를 기반으로 하며, 이를 최소화하도록 θ^- 는 주기적으로 업데이트 된다.

$$Loss = \sum_{j=1}^{mb} \left[\left(r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1} | \theta^-) \right) - Q(s_t, a_t | \theta) \right]^2 \quad (3)$$

2-3. Behavior Network 구조

그림 2는 세 가지 요소로 구성된 Behavior Network 구조를 나타낸다. UAV의 카메라로부터 입력된 연속된 프레임은 CNN 계층을 통해 시각적 특징을 추출한다. 추출된 특징은 완전 연결 계층(Fully Connected Layer)로 전달되어 고차원적 특징

표현으로 변환된다. 이 과정에서 차량 이동 패턴, 환경적 요소, UAV 행동 이력 등이 통합되어 보다 정교한 표현 학습이 이루어진다. 최종적으로, 네트워크는 최대 Q 값을 갖는 행동을 선택한다. UAV의 제어 행동은 총 15개의 조작으로 구성된다.

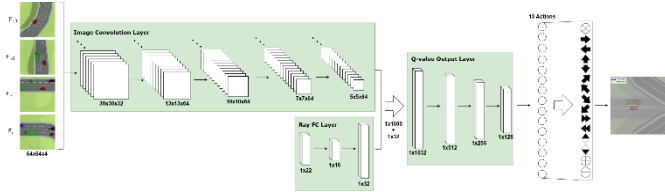


그림 2. Behavior Network 구조

2-4. 상태, 행동, 보상 정의

차량 추적 시스템에서의 상태(state), 행동(action), 보상(reward)은 다음과 같이 정의된다.

- 상태(state): 본 연구에서는 UAV와 차량의 상태를 시각적 정보와 수치 정보를 결합하여 정의한다. 차량 상태를 포함하는 연속적인 이미지 프레임은 입력 받아 시각적 정보를 추출한다. 이는 시계열 기반의 데이터를 대신한다. 라이다 센서를 통한 장애물 감지 정보, UAV의 직진 위치, 수평 속도, 행동, UAV와 차량 간의 유클리드 거리, 상태 각도, 차량 유무, 그리고 객체 검출 모델의 차량 예측 정보를 포함한다. 이는 UAV가 환경 변화에 유연하게 대응하며 최적의 행동을 선택할 수 있도록 유도한다.
- 행동(action): 본 연구에서 정의하는 UAV의 행동은 차량 추적을 위한 이동 조작을 의미한다. 행동은 정지, 수평 이동, 고도 조절, 속도 조절, 해상도 조절의 다섯 가지 범주로 구분되며, 총 15개의 이산적 행동으로 구성된다. UAV는 차량을 추적하기 위해 호버링, 상·하·좌·우 및 대각선 방향의 수평 이동을 수행하며, 고가도로 및 터널과 같은 구조물에 효과적으로 대응하기 위해 고도 조절을 포함한다. 또한, 차량의 속도 변화에 대응하기 위해 가속 및 감속 행동이 정의되며, 정밀한 추적을 위해 줌 인/아웃 방식의 해상도 조절도 포함된다.
- 보상(reward): 본 연구에서의 보상 함수는 도로 환경에서 UAV의 차량 추적 성능을 최적화하는 것을 목적으로 설계되었다. 전체 보상 값은 식 (4)과 같이 정의되며, 차량 검출 여부, UAV와 차량 간 거리, UAV의 속도조절(가감속) 정보를 기반으로 구성된다.

$$R_{total} = R_{detection} + R_{distance} + R_{acceleration} \quad (4)$$

$R_{detection}$ 은 이전 상태와 현재 상태를 고려하여 차등적으로 부여되는 보상이다. $R_{distance}$ 은 UAV가 차량과 적정 거리를 유지하는 경우 최대 보상이 주어지며, 거리 차이에 비례하여 선형적으로 감소하도록 설계되었다. 또한 $R_{acceleration}$ 은 UAV가 차량의 속도 변화에 부드럽게 대응할 수 있도록 유도하는 보상이다.

III. 모의 실험 결과

본 논문에서 사용한 모의 실험은 Unity ML-Agent 환경에서 구축하였다. 도심환경을 기반으로 고가도로, 터널 및 건물을 배치하고, 약 20 대의 차량을 생성하였다. 또한 4 가지의 날씨 조건을 구현하여 실제 도심 도로 상황을 고려한 실험을 수행하였다. 실험 결과로 얻어진 보상 값의 평균(Score)은 그림 2 를 통해 나타내어 추적 모델의 성능을 분석하였다.

표 1. 시뮬레이션 파라미터

Parameter	Value	Parameter	Value
Learning rate	2×10^{-5}	Learning rate	5×10^{-4}
Discount factor (γ)	0.95	Heatmap Gaussian	2
Replay Memory capacity	15000	Self Attention heads / key	8 / 32
Batch size	32	Batch size	64
Target Network update	1000	Conv Layers	5

표 2. 객체 검출 모델에 따른 성능 비교

Model (Heatmap : Regression)	Average MAE	Average Error (pixels)
1:9	0.01772	5.03 px
3:7	0.01663	4.26 px
5:5	0.01706	4.37 px

표 2는 차량의 중심 좌표를 예측하는 데 있어 회귀 기반의 직접적인 좌표 예측 뿐만 아니라, 확률적 공간 정보를 제공하기 위한 Heatmap 예측을 병렬적으로 학습하는 멀티태스킹 구조의 성능을 나타낸다. 특히, 차량 중심이 존재할 가능성이 높은 위치를 확률 분포 형태로 제공함으로써, 단일 회귀 기반 접근 방식보다 더 안정적인 추론이 가능하다는 점에서 주목할 만하다. 실험 결과, Heatmap 보조 모델은 회귀 단독 모델에 비해 평균 중심 오차에서 일관되게 더 낮은 값을 기록하였다. 특히 회귀의 비중이 상대적으로 높은 설정에서 가장 우수한 성능을 보였으며, 이는 해당 모델이 대략적인 위치 정보를 제공해 회귀의 초기 추정을 유도하고, 이후 정밀한 위치 보정을 수행할 수 있게 하는 구조적 이점 때문으로 해석된다. 이러한 구조는 노이즈에 민감하고 예측 위치의 불안정성이 존재하는 회귀 단독 모델의 단점을 보완할 수 있다. 반면, Heatmap만을 사용하는 경우에는 정밀한 좌표 추출을 위한 명확한 기준이 부족하다는 한계가 존재한다. 따라서 본 모델은 두 방식을 병행함으로써 정확도와 안정성 측면에서 균형을 이룬다.

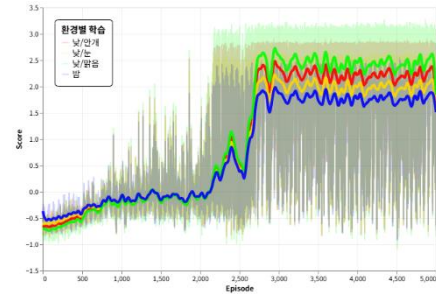


그림 3. 환경별 학습 결과

그림 3 은 각 환경에서의 UAV 차량 추적 결과를 나타낸다. 약 2800 에피소드 시점부터 보상 평균값이 최대 값에 수렴하는 양상을 보이며, 이는 UAV가 충분한 탐색(Exploration)을 통해 환경에 적응하고 최적에 가까운 추적 정책을 학습했음을 의미한다. 환경 조건이 복잡해질수록 평균 보상은 상대적으로 낮은 수준을 유지하지만, 이는 비전 기반 인식 특성상 자연스러운 현상으로 해석될 수 있다. 낮/밤을 환경과 유사한 수렴 경향을 보였으며, 고가도로 및 터널이 존재하는 복잡한 환경에서도 안정적인 추적 성능을 유지하는 것으로 나타났다. 이러한 결과는 제안 시스템이 다양한 환경 조건에서도 효과적으로 추적 성능을 발휘함을 보여준다.

IV. 결론

본 논문에서는 차량 검출 모델과 심층 강화학습(DQN) 기반의 차량 추적 시스템을 통합하여, UAV가 차량을 실시간으로 검출하고 지속적으로 추적할 수 있는 시스템을 제안한다. 제안된 CNN 기반 차량 검출기는 Self-Attention 모듈과 Heatmap + Regression 기반의 멀티태스킹 구조를 통해 복잡한 환경에서도 높은 검출 성능을 보였다. 특히, 차량이 부분적으로 가려지거나 복잡한 배경 조건에서도 Bounding Box와 중심 좌표를 안정적으로 예측할 수 있음을 실험을 통해 확인하였다. 전체 시스템의 성능 평가 결과, 기존의 Rule-based 방식에 비해 차량 탐지율과 거리 유지 측면에서 우수한 성능을 보였으며, 복잡한 교통 환경에서도 강인한 차량 추적 성능을 유지하였다. 이는 DQN 기반 접근 방식이 고정된 규칙에 의존하지 않고 다양한 상황에 효과적으로 적응하며 이동 객체를 추적할 수 있음을 입증한다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획지원-대학ICT연구센터(ITRC)의 지원을 받아 수행된 연구임 (RS-2021-II212052)

참고 문헌

- [1] J. Wang et al., "OmniTracker: Unifying Visual Object Tracking by Tracking-With-Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 47, no. 4, pp. 3159–3174, April 2025.
- [2] K. Parkhodov, J. -H. Park, S. -H. Lee and K. -R. Kwon, "Virtual Simulation based Visual Object Tracking via Deep Reinforcement Learning," International Conference on Information Science and Communications Technologies (ICISCT), 2022.
- [3] 이준환, "지능형 관제시스템을 위한 딥러닝 기반의 다중 객체 분류 및 추적에 관한 연구", 스마트미디어저널, vol.12, no.5, pp. 73–80, 2023
- [4] D. Barrientos, M. Medina, B.Fernandes, P. Barros, "The Use of Reinforcement Learning Algorithms in Object Tracking: A Systematic Literature Review," Neurocomputing, vol. 596, 2024.