

YOLOv8 기반 군사용 가상훈련체계의 훈련자 캐릭터 검출 방법

박영제¹, 한재혁², 김미혜³

¹주젠시스템 부장

²충북대학교 컴퓨터공학과 박사과정

³충북대학교 컴퓨터공학과 교수

redo@naver.com, haraisi2@gmail.com, mhkim@cbnu.ac.kr

Training Participant Character Detection Method for YOLOv8-based Military Virtual Training System

Yong-Jae Park¹, Jae-Hyeok Han², Mi-Hye Kim³

¹Zenstem Co., Ltd., Deputy General Manager

^{2,3}Dept. of Computer Engineering, Chungbuk National University

요 약

실제 전투와 유사한 군사 훈련을 수행하기 위해서는 훈련 공간 확보, 악천후 극복, 실 사격 훈련, 민간인 통제 등 다양한 제약이 있다. 이러한 제약을 극복하기 위해 과학화 훈련이 도입되었으며, 현대전의 양상이 대규모 전투에서 소규모 교전으로 전환되면서 가상 훈련 시스템이 주목을 받고 있다. 가상 현실에서 적을 감지하기 위해 광선투사방식이 사용되지만, 이 방법은 인간의 시각 지각능력을 넘어서기 때문에 현실적인 훈련을 시뮬레이션 하는 데 한계가 있다. 본 논문은 가상 환경 내 가상자율군(Computer Generated Forces)이 현실적인 적 시뮬레이션을 달성하기 위하여 이미지 기반의 적 검출을 적용하여, 광선투사방식에 비해 인간 시각 지각에 더 가까운 결과를 얻었다.

1. 서론

1980년대부터 세계각국을 비롯한 우리나라는 다양한 과학화훈련 체계를 발전시켜 왔으며 걸프전쟁 등 여러 전쟁에서 그 효과를 확인하였다. 미 해군은 베트남전쟁에서 실기동 훈련을 실전과 유사하게 실시하는 탐전 조종학교를 통해 극적으로 격추율을 향상시켰으며, 걸프전쟁에서 미 육군은 과병 이전 중동의 가상지형을 먼저 체험하기위한 SIMNET 에서의 훈련을 통해 승리할 수 있었다. 4차 산업혁명 시대가 도래한 최근에는 과학화훈련도 증강현실, 가상현실, 인공지능, 빅데이터 등과 같은 기술을 활용하는 과학화 훈련체계로의 전환이 급속하게 이루어지고 있다.

과학화훈련체계는 실기동 훈련 체계(live training system), 위게임훈련 체계(constructive training system)로 구분할 수 있으며[1], 각각의 훈련 체계는 대표적으로 KCTC(Korea Combat Training Center), 시뮬레이터, BCTP(Battle Command Training Program) 등이 있다. 이를 연동하여 운영하는 체계 개념을 LVC(Live, Virtual, Constructive) 훈련체계로 정의하며, LVC 훈련체계

전술게임체계를 포함하여 LVCG(Live, Virtual, Constructive, Game) 훈련체계로 정의하기도 한다[2].

현재까지 개발된 가상훈련체계에서 가상자율군은 탐지를 위한 기술로 광선투사(raycasting) 방식을 사용하고 있어 정확한 탐지가 가능하다. 광선투사 방식은 임의의 비가시적인 광선을 투사해 광선에 닿는 표면을 파악하여 대상을 검출하는 기술이다[3]. 3차원 이미지를 실시간으로 생성하기에는 컴퓨터의 성능이 낮아 제한이 있었던 시기에 등장한 개념으로 3차원 공간을 이미지로 만들어내는 것에서 출발되었으며, 이후 게임과 시뮬레이터 등에서 해당 프로그램 내 인공지능 캐릭터가 적을 탐지하는 기법으로도 활용되고 있으며 군사용 가상훈련체계의 가상자율군도 동일한 기법을 활용하고 있다.

본 논문은 기존 가상훈련체계 내 인공지능 적군의 탐지 방식인 광선투사 방식의 문제점인 시각적으로 확인할 수 없는 적을 탐지하는 상황을 해결하고자 기존의 광선투사 방식을 사용하여 사격이 가능한 상태의 대상을 100% 정확도로 검출한 뒤 검출된 대상의

이미지를 추출하여 이미지 기반의 검출 모델을 통한 검출 결과를 더하여 인간의 시각에 의한 검출과 유사한 결과를 도출하는 탐지 방식을 구현하여 탐지 결과가 인간의 시각적 판단 결과와 유사한 결과를 도출할 수 있게 함으로써, 군사용 가상훈련체계의 색적 기능 사실성을 향상시켜 사실적 전투행동 모의를 가능하게 하고자 한다.

2. 관련연구

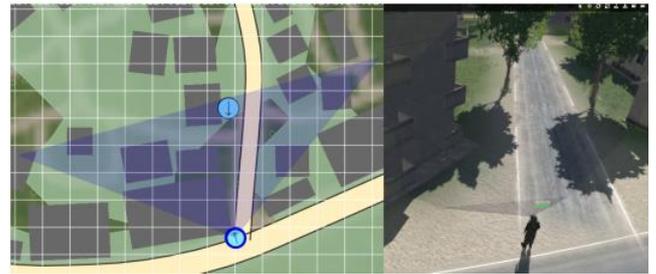
군사용 가상훈련체계는 게임 분야에서 사용되는 대상 검출 기술인 광선투사 방식을 사용한다. 검출 시작 지점으로부터 측정 범위에 포함된 모든 목표 오브젝트들을 향해 가상의 선, 면, 공간 등을 그리고 각각의 요소들이 목표하는 대상까지 도달하는 상황을 모의하여 장애물에 의한 방해 여부를 확인하는 방식이다. 광선투사방식은 과거 컴퓨터의 성능이 삼차원 그래픽을 구현하기 어려운 상황에서 등장한 그래픽 구현 기술이었으나 컴퓨터의 성능이 올라감에 따라 가상환경 내에서의 대상검출을 위한 기술로 사용되고 있으며 가상환경의 특성에 따라 100%의 검출성능을 가지게 된다. 대표적 상용 게임 엔진인 Unreal 시리즈, Unity3D 와 군사용 시뮬레이션 엔진 VBS 시리즈 등에서도 대상검출을 위한 기술로 광선투사 방식을 제공하고 있다.

최근 컴퓨터 비전 분야의 발전으로 딥러닝 기반의 객체검출이 활용되고 있다. 딥러닝 기반의 객체검출은 영상이나 비디오에서 객체 인스턴스를 찾기 위한 컴퓨터 비전 기법으로 인간의 시각인지에 의한 객체 검출 과정과 유사하다. 객체검출 알고리즘은 일반적으로 딥러닝 또는 머신러닝을 활용하여 검출결과를 생성한다. 최근 가장 널리 활용되는 알고리즘으로는 YOLO(You Only Look Once)가 있다. YOLO 는 이미지 내의 객체를 찾고, 찾은 객체를 구분하기 위하여 입력된 이미지를 일정한 그리드로 나눈 후 신경망을 통과시켜 이미지 안에서 객체의 위치와 객체의 종류를 함께 찾는 알고리즘이다. 2015 년 Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi 에 의해 소개된 후 현재 v8 까지 출시되었고 특징적인 장점으로 빠른 속도를 보유하고 있으며 이를 기반으로 실시간 의사결정이 필요한 분야에서 많이 사용되고 있다[4]. 최근에 발표된 YOLO v8 은 이전 버전과 비교하여 속도, 정확성 등 가시적인 성능이 향상되었고 기본적으로 다양한 객체탐지, 추적, 인스턴스 분할, 이미지 분류, 포즈 추정 등을 수행할 수 있다.

3. 인간과 유사한 판단을 위한 이미지 기반 훈련자 캐릭터 검출

현재 군사용 가상훈련체계에서 활용되는 가상자율군은 시나리오, 또는 규칙에 기반한 시뮬레이션이 적용되고, 이는 모의에 필요한 완전한 삼차원 가상전장 정보를 바탕으로 전투에 참여하는 개체의 사격, 방어, 계획, 이동, 탐지 등의 행위를 제어하는 수준으로 가

상훈련체계에 적용되어 군사 훈련에 이용되고 있다.



(그림 1) 가상훈련체계 광선투사 구현 예.

그림 1 과 같이 투사된 광선의 적중 여부로써 대상을 검출하는 광선투사 방식은 적중대상의 재질, 조명 상태 등과 무관하게 작동된다. 이 때문에 야간이나 암전된 실내에서의 상황, 복장 및 자연물 등으로 위장한 상태의 대상, 나뭇잎 및 반투명 장애물에 가로막혀 대상의 신체 일부만 보이는 상황 등 인간이 시각적으로 인지할 수 없는 대상까지 검출한다. 특히 가상환경 구현에 있어 적은 리소스를 사용하여 사실적인 렌더링 이미지를 만들기 위해 투명한 영역이 적용된 재질을 사용하는 경우 가상자율군은 인간이 시각적으로 판단하였을 때 인지할 수 없는 대상도 탐지하는 결과를 보인다. 이는 위장, 은폐를 이용한 은밀 기동을 무시하고 대상을 탐지하여 사실적 훈련에 제약이 발생한다. 현재 절대다수의 가상훈련체계, 시뮬레이터, 게임 등에 사용되는 광선투사방식 검출은 가상환경의 특성에 따라 완벽한 검출성능을 보유하고 있다. 그러나 인간이 인지할 수 없는 상황의 대상에 대해서도 정확하게 검출을 수행한다는 중대한 문제가 있다. 대표적인 사례로 인간의 시력으로는 식별 불가능한 거리의 대상을 검출하는 경우, 식별이 불가능한 수준의 어두운 환경 내의 대상을 검출하는 경우, 덩불이나 나뭇잎같이 불투명한 장애물에 가려진 대상을 검출하는 경우가 그것이다. 이는 광선투사 검출 방식의 한계로 눈에 비치는 시각정보를 통해 대상을 인지해내는 인간의 대상식별 방식이 NPC(Non Player Character)나 가상자율군에게도 필요하다는 것을 역설하고 있다.

본 논문에서는 이미지 기반 훈련자 캐릭터 검출 시스템을 제안하고 이 시스템을 사용하는 군사용 가상훈련체계에서의 가상자율군이 훈련자를 인간과 유사한 수준으로 인식하게 하고자 한다. 이를 위하여 국군이 사용하는 군사용 가상훈련체계를 대상으로 적용이 가능한 구조를 갖추어야 하고, 일반적인 인간의 인지 이후 반응속도인 3 초 이내의 성능을 보유한 이미지 기반의 객체검출 모델을 구현한다. 이러한 객체검출 모델 구현을 위해 그림 2 와 같이 기존의 군사용 가상훈련체계 내 훈련자 캐릭터의 이미지를 바탕으로 기준 사격거리인 50m, 100m, 150m, 200m, 250m 거리에 위치한 훈련자 캐릭터로 데이터셋을 구축한 뒤, 구축된 데이터셋을 바탕으로 속도에 장점을 가진 YOLO v8 을 학습시켜 훈련자 캐릭터 검출모델을 개발한다. 그 뒤 검증을 위해 기존의 체계와 비교를 위

해 기준 사격거리에 위치한 훈련자 캐릭터를 낮, 밤, 은폐 여부, 선 자세, 앉은 자세, 엎드린 자세에서 검출한 후 이미지를 생성하여 검출모델의 입력 데이터로 사용해 개발된 검출모델과 기존 검출모델의 결과를 비교한다.



(그림 2) 이미지 기반 훈련자 캐릭터 검출 시스템 구현.

군사용 가상훈련체계의 캐릭터 이미지 데이터 셋을 구축하기 위하여 가상훈련체계의 훈련자 캐릭터 1종을 대상으로 앉기, 서기, 엎드리기 동작 데이터를 제작하였고 각 자세별 소총 조준 동작, 전/후/좌/우 4방향으로의 이동 동작을 구현하였다. 가상환경의 기상환경조건으로 시간적 요소인 낮, 밤을 구분하여 각각 환경의 밝기를 구현하였고 수풀, 나무 등 자연 지물과 차량 등 인공 지물을 배치하여 훈련자 캐릭터의 은폐 상황을 조성하였다. 마지막으로 대상과의 거리는 실거리사격의 기준거리인 50m, 100m, 150m, 200m, 250m를 각각 기준으로 설정하여 대상과의 거리별 검출 상황을 조성하였다.

생성된 데이터 셋은 학습 데이터 70%, 검증 데이터 20%, 시험 데이터 10%로 구분하였고 필요시 데이터의 확보가 용이한 가상환경의 특성상 데이터 증강은 수행하지 않았다.

개발지에서의 거리별, 자세별 훈련자 캐릭터 이미지를 30장씩 생성한 후 모의훈련체계 사용자 30명을 대상으로 육안식별 가능여부를 조사하여 표 1과 같은 결과를 얻었다. 육안식별 일부 가능한 경우 훈련자 캐릭터의 방향이 횡방향으로 엎드린 이미지에 한하여 식별 가능한 것으로 조사되었다. 조사 결과에 따라 200m 이상의 움직임이 없는 대상은 자세와 무관하게 육안식별이 불가능한 것으로 판단되었기 때문에 해당 이미지 데이터는 학습 대상에서 제외하였다.

훈련자 은폐상태에서의 거리별, 자세별 훈련자 캐릭터 이미지를 30장씩 생성한 후 동일 대상자에게 육안 식별 가능여부를 조사한 결과 은폐한 상태의 움직임이 없는 대상은 자세, 거리와 무관하게 육안식별이 불가능한 것으로 조사되었다.

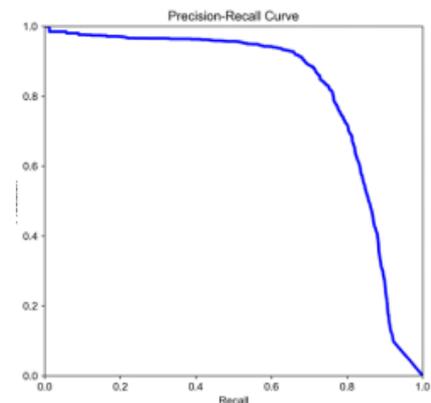
마지막으로 야간 상황에서의 거리별, 자세별 훈련자 캐릭터 이미지를 30장씩 생성한 후 동일 대상자에게 육안 식별 가능여부를 조사한 결과 야간환경의 경우 50m 거리, 선 자세의 캐릭터를 일부 육안식별이 가능하나 나머지 경우 육안식별이 불가능한 것으로 조사되었다.

거리	캐릭터 자세	데이터 수	육안식별여부
50m	선 자세	30	가능
	앉은 자세	30	가능
	엎드린 자세	30	가능
100m	선 자세	30	가능
	앉은 자세	30	가능
	엎드린 자세	30	일부 가능
150m	선 자세	30	가능
	앉은 자세	30	가능
	엎드린 자세	30	불가
200m	선 자세	30	불가
	앉은 자세	30	불가
	엎드린 자세	30	불가
250m	선 자세	30	불가
	앉은 자세	30	불가
	엎드린 자세	30	불가

<표 1> 개발지 데이터셋 구성

본 연구의 목표는 인간의 인지능력과 유사한 능력의 객체검출 성능을 확보하는 것으로써, 위의 조사 결과를 바탕으로 확인된 육안식별 불가 항목은 학습 데이터에서 제외하였다.

캐릭터 검출 모델을 개발하기 위하여 YOLOv8을 사용하였다. 모델 개발환경은 Colab Pro+를 사용하였으며 Colab Pro+는 A100 GPU, 52GB RAM의 환경을 제공한다. YOLOv8은 n 모델과 l 모델을 사용하였고 근소한 차이로 l 모델의 신뢰도가 우세하였다. 속도의 측면에서 n 모델의 경우 검출에 평균 75ms 정도의 시간이 소요되었고, l 모델의 경우 140ms 정도의 시간이 소요되어 수치적으로는 차이가 있었으나 훈련에서 요구하는 반응속도는 3초 이내이기 때문에 동등한 결과로 판단하였다. 이와 같은 결과에 따라 l 모델을 바탕으로 하이퍼 파라미터를 조정하여 모델을 학습하였다. 학습조건은 최종적으로 epoch: 100, batch size: 10, imgsz: (1920, 1080), weightsyolov8l.pt, cfgyolov8l.yaml을 사용하였다.



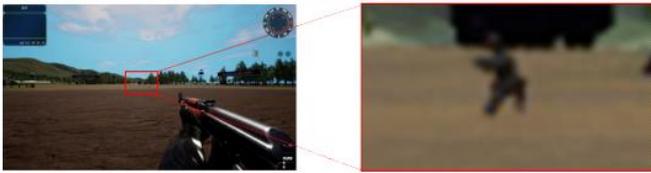
(그림 3) Precision-Recall Curve.

학습결과 주요지표는 Precision: 0.872, Recall: 0.781, mAP50: 0.843, mAP50-95: 0.52로 학습되었고 PR Curve는 그림 3과 같이 도출되었다. 테스트셋의 육안확인이 불가능했던 항목들을 포함하여 검출하는 경우 검출율이 20% 이내의 결과가 도출되어 일반적인 검출모

델로써는 저조한 성능으로 판단할 수 있었으나 인간의 시각에 의한 판단결과와 유사한 결과를 도출하기 위한 검출모델이라는 특수성에는 오히려 부합하는 결과라 해석할 수 있다.

4. 실험결과 및 분석

구축된 군사용 가상훈련체계에서 광선투사 방식을 적용한 가상자율군은 그림 4 와 같이 거리, 환경 밝기와 무관하게 훈련자 캐릭터를 검출하였고 투명 재질이 적용된 나무, 수풀 등의 장애물에도 방해받지 않고 훈련자 캐릭터를 검출하였다.

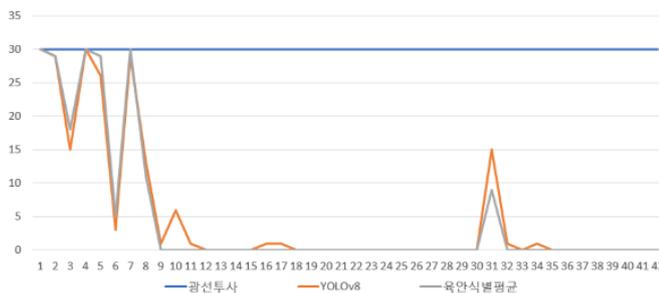


(그림 4) 광선투사방식 250m 거리 대상 검출결과.

YOLO v8 을 사용하여 개발한 캐릭터 객체 검출모델은 가상자율군 시점에서 바라본 훈련자 캐릭터 이미지를 대상으로한 검출 결과 거리, 환경 밝기에 따라 육안으로 식별할 수 없는 경우, 훈련자가 앞거나 옆드린 경우, 투명 재질이 적용된 나무, 수풀 등의 장애물에 훈련자 캐릭터의 일부가 가려진 경우에는 그림 5 와 같이 훈련자 캐릭터를 검출하지 않았다.



(그림 5) 객체검출 모델 50m 거리 대상 검출결과.



(그림 6) 객체검출 결과 비교.

광선투사방식의 결과와는 대조적으로 본 연구를 통하여 개발된 이미지 기반 검출모델은 그림 6 과 같이 육안식별 가능 여부와 유사한 검출결과를 보인다. 세부적으로는 훈련자 캐릭터의 자세와 각도에 따라 비교적 검출이 쉬운 상황에서도 검출하지 못하는 경우가 있었으나 이는 실제 훈련 상황을 가정하였을 때 전후의 동작에서 훈련자를 검출할 것으로 추정할 수

있다. 또한 육안으로 확인이 어려운 데이터를 대상으로 훈련자 캐릭터를 검출할 경우가 있었으나 전후의 동작에서 도출된 검출결과를 바탕으로 가상자율군이 인간과 유사한 검출결과를 얻을 수 있다. 이러한 검출결과에 기반한 행동분기를 적용하여 가상자율군의 사실적 행동을 유도할 수 있다.

5. 결론

광선투사 방식의 검출 결과를 육안식별 가능 여부와 비교하여 보았을 때 가상환경 내 150m 이상 거리에서 시야의 제한 또는 훈련자 캐릭터의 자세에 의하여 육안식별이 어려운 상태의 훈련자 캐릭터를 검출하였다. 이는 광선투사 방식의 검출방식이 인간이 시각적으로 검출하기 어려운 상태의 캐릭터를 검출한다는 것을 의미한다.

이미지 기반 검출모델의 경우 육안식별이 어려운 상태의 훈련자 캐릭터를 대체로 검출하지 못하였으며 검출된 결과에서도 신뢰도 점수를 제공하여 인간이 시각적으로 검출하기 어려운 상태의 훈련자 캐릭터는 낮은 신뢰도 점수를 부여하였다. 이에 따라 광선투사 방식의 검출결과와 비교하여 이미지 기반의 검출모델을 이용하였을 때 인간의 시각적 검출 결과와 유사한 결과를 도출할 수 있음이 증명되었다.

본 연구는 정지된 이미지를 기반으로 검출하여 움직이는 대상에 대한 검출결과의 사실성이 제한된다. 향후 실시간 영상 기반의 검출구조를 갖는 모델을 개발하여 사실성을 높이고, 신뢰도 점수에 기반하여 가상자율군의 훈련수준을 정의하고 이에 따른 사실적인 반응을 연동하고자 한다.

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역진흥화 혁신인재양성사업임(IITP-2024-2020-0-01462)

참고문헌

- [1] 강주석(2018), 가상환경에서 CNN 을 활용한 자율주행 및 표적탐지, 한밭대학교 대학원 박사학위 논문
- [2] 박상근, “육군 과학화훈련체계의 한계와 발전방안,” 『한국군사학논총』, 제 12 집 1 권 (2023), p.135.
- [3] Roth, Scott D. (February 1982), “Ray Casting for Modeling Solids”, 《Computer Graphics and Image Processing》 18 (2): 109- 144, doi:10.1016/0146-664X(82)90169-1
- [4] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).