

국방 AI 플랫폼 기반 Vision 변화 탐지에 관한 연구

박병훈¹, 서범석¹, 이명규¹, 황영준¹, 김기환¹
¹티쓰리큐㈜

warmpark@t3q.com, tjqjatjr98@gmail.com, mkliving31@gmail.com, hyjun0103@t3q.com, tensor@t3q.com

Research on Vision Change Detection Based on Defense AI Platform

Byeong-Hoon Park¹, Beom-Seok Seo¹, Myeong-Gyu Lee¹, Yeong-Jun Hwang¹, Ki-Hwan Kim¹
¹T3Q

요약

본 연구는 국방 AI 플랫폼 기반 변화 탐지 서비스를 제안한다. 국방 AI 플랫폼의 특성을 반영하여 객체의 출현, 소실, 이동을 정량화하고, 이를 통해 장면의 동적 변화를 평가하는 변화율을 정의한 후, 해당 변화율을 적용한 서비스를 설계하였다. 변화 탐지 서비스는 학습 없이 다양한 환경에서 실시간 적용이 가능하며, 사건의 중요도를 효율적으로 분석할 수 있다. YOLOv10 탐지 모델과 추적 모델을 사용한 실험을 통해 변화율 기반 변화 탐지 서비스의 실시간 처리 가능성을 검증하였으며, 복잡한 상황에서도 효과적인 변화 탐지가 가능함을 확인하였다.

1. 서론

인공지능 기반 CCTV 시스템을 국방 분야에 적용하려는 다양한 연구와 시도가 활발히 이루어지고 있다.[1] 이 시스템은 다수의 객체가 존재하는 복잡한 상황에서 효과적으로 작동해야 하며, 특정 환경에서 학습된 데이터에 의존하지 않고 시시각각 변화하는 전장 상황에서 성능을 유지할 수 있어야 한다.[2][3]

본 논문에서는 이러한 요구를 충족하기 위해 변화율을 정의하고 이를 활용한 서비스를 제안한다. 변화율은 객체의 출현, 소실, 이동과 같은 동적 변화를 실시간으로 정량화하여 복잡한 환경에서도 사건의 중요성을 신속하고 정확하게 평가할 수 있는 지표이다. 본 연구에서는 YOLOv10[4] 객체 탐지 모델과 추적 모델을 사용해 변화율 기반 변화 탐지 서비스를 설계하였고, 이를 통해 실시간 서비스 구현 가능성을 확인하였다.

제안한 변화율 기반 변화 탐지 서비스의 결과는 기존 사건 기반 시스템 및 이상 감지 시스템으로 탐지한 결과와 비교하여 성능을 검증하였다. 이 서비스는 국방 AI 플랫폼에서 실시간으로 효율적인 변화 탐지를 가능하게 하여 지휘관의 신속하고 정확한 의사결정을 돕는데 기여할 것으로 기대된다.

2. 사전 연구

2.1 픽셀 기반 변화 탐지(Motion Detection)

픽셀 기반 변화 탐지 기법은 이미지의 픽셀 차이를

이용하여 발생한 변화를 분석한다. 구현이 간단하며 실시간 처리가 가능하나 조명 변화, 작은 움직임에 민감하게 반응하여 오탐지율이 높다. 특히 배경이 복잡하거나 조명 변화가 잦은 환경에서 성능이 저하되며, 객체의 정확한 위치나 크기 정보를 제공하지 못한다.[5]

2.2 장면 기반 변화 탐지(Segmentation)

장면 기반 변화 탐지 기법은 이미지 내의 특정 장면이나 영역에서 발생한 변화를 분석한다. 이미지 내의 전체 장면의 구성을 분석하여 이전 프레임과의 차이를 비교함으로써 변화를 감지한다. 이미지의 전반적인 구성을 고려하므로 단순한 픽셀 차이보다 더 높은 수준의 분석을 제공하나, 고해상도 이미지 처리에 많은 계산 자원이 필요하며 변화의 원인을 구체적으로 파악하기 어려운 경우가 있다.[5]

2.3 객체 기반 변화 탐지(Object Detection)

객체 기반 변화 탐지 기법은 객체 인식 및 추적 기술을 활용하여 영상에서 발생하는 변화를 분석한다. 객체의 위치, 크기, 이동 경로 등을 추적하여 변화 패턴을 기반으로 장면의 변화 정도를 평가한다. AI 모델을 통해 객체를 식별하고 추적하여 보다 정확한 변화 탐지를 가능하게 한다.[5] 지능형 CCTV 시스템에서 사건 기반 감지 시스템 또는 이상 감지 시스템으로 사용된다.

2.3.1 사건 기반 감지 시스템

사건 기반 감지 시스템(이하 사건 기반 시스템)은 객체의 출현, 소실, 이동과 같은 특정 이벤트를 탐지할 때마다 경고를 생성하는 방식으로 작동한다. 이 시스템은 보안 및 감시 목적으로 널리 사용되며, 사건이 발생할 때 즉각적으로 반응한다는 장점이 있다. 그러나 객체의 모든 작은 변화에도 반응하여 불필요한 경고가 빈번하게 발생할 수 있으며, 복잡한 환경에서는 과도한 알림이 발생해 운영 효율성을 저해할 가능성이 있다.[2]

2.3.2 이상 감지 시스템

이상 감지 시스템은 정상적인 행동 패턴을 모델링하고, 이를 벗어나는 비정상적인 상황을 감지하는 방식이다. 머신 러닝 알고리즘을 기반으로 학습된 데이터를 통해 정상과 비정상을 구분하며, 주로 보안 시스템에서 이상 상황 탐지에 사용된다. 하지만 새로운 환경에서는 학습되지 않은 패턴에 대해 적응하지 못하는 한계가 있으며, 학습에 많은 데이터와 시간이 소요된다는 단점이 존재한다. 특히 변화하는 환경에서는 재 학습이 필요해 실시간 대응이 어렵다.[3]

3. 변화율 정의 및 제안 서비스

변화율은 영상 내에서 객체의 출현, 소실, 이동을 기반으로 장면의 동적 변화를 정량적으로 평가하는 지표이다. 변화율은 특정 시간 동안 영상에서 발생한 주요 변화의 크기와 빈도를 수치화하여, 시스템이 이를 기반으로 변화의 중요성을 평가하고, 그 결과를 지휘관의 의사결정에 활용할 수 있도록 지원한다.

3.1 탐지 객체 정의

탐지 모델의 오차를 고려하여, 각 프레임의 추론 결과를 그대로 사용하지 않고 일정 구간 동안의 추론 결과를 종합해 변화를 판단하였다. 이 구간을 윈도우(window)로 정의하고, 각 윈도우 내에서 일정 비율 이상 탐지된 객체만을 실제로 탐지된 객체로 간주하였다. 본 연구에서는 50% 이상 탐지된 객체를 실제 탐지된 객체로 판단하였다. 또한, 윈도우 간 탐지 결과를 비교할 때는 각 윈도우의 간격을 일정한 보폭(stride)으로 설정해 효율적인 변화 탐지를 구현하였다.

3.2 변화율 정의

본 논문에서 제안하는 변화율은 기존의 시간적 변화를 나타내는 일반적인 변화율 개념과는 구별된다. 본 연구에서는 객체 탐지에서 식별되는 세 가지 요소를 각각의 지표로 활용하여 변화율을 정의하였다. 또한, 특정 클래스의 비중이 높아 소수 클래스 객체의 정보가 손실되는 것을 방지하기 위해 클래스별로 변화율을 계산한 후 이를 평균하여 최종 변화를 산출하였다.

3.2.1 출현 지표(c_a)

출현 지표는 현재 윈도우에서 탐지된 객체 중 이전 윈도우에서 탐지되지 않았던 객체의 비율을 나타낸다.

t 번째 윈도우의 출현 지표는 식 (1)과 같이 계산한다.

3.2.2 소실 지표(c_d)

소실 지표는 이전 윈도우에서 탐지된 객체 중 현재 윈도우에 탐지되지 않은 객체의 비율을 나타낸다. t 번째 윈도우의 소실 지표는 식 (2)에 따라 계산하며, 여기서 분모 $|W_{t-1,i}|$ 는 t-1 번째 윈도우에서 해당 클래스의 전체 객체 수를 의미한다. 이는 소실된 비율을 측정하기 위해 이전 윈도우에서 얼마나 많은 객체가 존재했는지를 기준으로 반영한 것이다.

3.2.3 이동 지표(c_m)

이동 지표는 현 윈도우 내에서 탐지된 객체 중 움직이고 있는 객체의 비율을 나타낸다. 이때 움직이고 있는 객체는 탐지된 객체 중 객체의 평균속도가 기준 속도 이상인 객체로 정의하였다. t 번째 윈도우에서 이동 지표는 식 (3)과 같이 계산한다. 본 연구에서 기준 속도 0.5 를 사용한다.

3.2.4 변화율(c)

변화율은 출현, 소실, 이동 지표의 가중평균으로 정의한다. 본 연구에서는 출현 가중치(w_a) 0.4, 소실 가중치(w_d) 0.4, 이동 가중치(w_m) 0.2 를 사용한다. t 번째 윈도우에서 변화율은 식 (4)과 같이 계산한다.

W_t : t 번째 윈도우에서 탐지된 객체들의 집합
 N_a : t 번째 윈도우에서 새롭게 출현한 객체들의 수
 N_d : t 번째 윈도우에서 소실된 객체들의 수
 N_m : t 번째 윈도우에서 기준 속도 이상 객체들의 수

$$N_a = |W_t - W_{t-1}|, \quad N_d = |W_{t-1} - W_t|$$

$$N_m = |\{o \in W_t \mid v(o) > v_{thr}\}|, \quad v_{thr} \text{은 기준 속도}$$

$$c_a(t) = \frac{1}{N_{cls}} \sum_{i \in cls} \frac{N_{a,i}}{|W_{t,i}|} \dots\dots\dots (1)$$

$$c_d(t) = \frac{1}{N_{cls}} \sum_{i \in cls} \frac{N_{d,i}}{|W_{t-1,i}|} \dots\dots\dots (2)$$

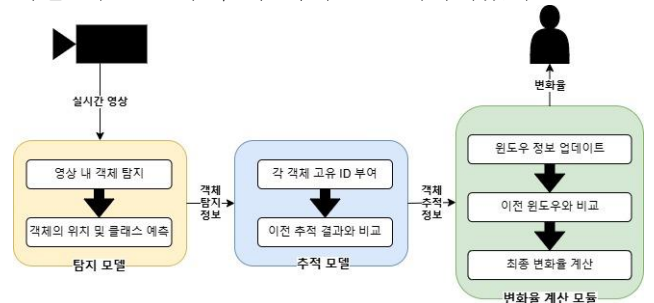
$$c_m(t) = \frac{1}{N_{cls}} \sum_{i \in cls} \frac{N_{m,i}}{|W_{t,i}|} \dots\dots\dots (3)$$

$$w_a + w_d + w_m = 1$$

$$c(t) = w_a \cdot c_a(t) + w_d \cdot c_d(t) + w_m \cdot c_m(t) \dots\dots\dots (4)$$

3.3 제안 서비스

3.2 에서 정의한 변화율을 적용하여 (그림 1)과 같이 변화율 기반 변화 탐지 서비스를 제작하였다.



(그림 1) 변화율 기반 서비스의 구조도

4. 실험 방법

4.1 사용 모델

국방 AI 플랫폼에서 변화 탐지를 수행하려면, 실시간으로 작동 가능하며 다양한 환경에서 우수한 성능을 발휘하는 모델이 필요하다. 이를 고려하여 추론 속도가 빠르고 성능이 뛰어난 YOLOv10[4] 모델 중 가장 큰 네트워크 구조와 많은 학습 파라미터를 가진 YOLOv10x 를 객체 탐지 모델로 채택하였다. 추적 모델은 속도가 비교적 빠르며 범용적으로 안정적인 성능을 제공하는 ByteTrack[6], BoT-SORT[7], GHOST[8], General Track[9]을 사용하였다.

추적 모델	HOTA	IDF1	MOTA
ByteTrack	61.3	75.2	77.8
BoT-SORT	62.6	76.3	77.7
GHOST	61.2	75.2	73.7
GeneralTrack	61.4	74.0	77.2

<표 1> MOT20 데이터[10]에서의 추적 모델 성능[7][9]

4.2 실험 설계

탐지 모델과 추적 모델을 Python 기반의 PyTorch 프레임워크로 통합하여 변화율 기반 변화 탐지 서비스를 구현하고, 이 서비스의 적합성을 평가하기 위해 세 가지 실험을 설계한다.

4.2.1 실시간 서비스 가능성 분석

변화율 기반 변화 탐지 서비스의 실시간 작동 가능성을 평가하기 위해 각 추적 모델의 프레임 당 처리 시간을 측정한다. 이때 국방 AI 플랫폼을 고려하여 실험은 AMD Ryzen 7 5800H, 32GB RAM, NVIDIA GeForce RTX 3050 Laptop GPU (VRAM 4GB) 사양의 노트북을 사용하여 진행한다.

4.2.2 사건 기반 시스템과의 비교

소수 클래스의 단일 객체들이 출현, 이동, 소실하는 상황의 Case 1 과 다수 클래스의 다수 객체가 출현, 이동, 소실하는 역동적인 상황의 Case 2 에서 사건 기반 서비스와 변화율 기반 변화 탐지 서비스를 비교한다. 실험에는 1920x1080 FHD, 30 프레임으로 직접 촬영한 영상을 사용한다.

4.2.3 이상 감지 시스템과의 비교

다양한 객체가 배경에 존재하는 상황에서 새로운 클래스의 객체가 출현하는 이상 현상 데이터를 대상으로 이상 감지 시스템과 변화율 기반 변화 탐지 서비스를 비교한다. 사용된 데이터는 이상 탐지의 Base Dataset 인 ShanghaiTech 영상 데이터[11]이다.

5. 실험 결과

5.1 실시간 서비스 가능성 분석

실험 결과 추적 모델의 종류에 따라 프레임당 처리 시간이 51 ~ 103ms 측정되었다. 이는 약 10~20 FPS 의 영상 데이터를 실시간으로 처리할 수 있음을 의미한다. 이 결과를 통해 본 서비스의 실시간 적용 가능성을 확인하였다.

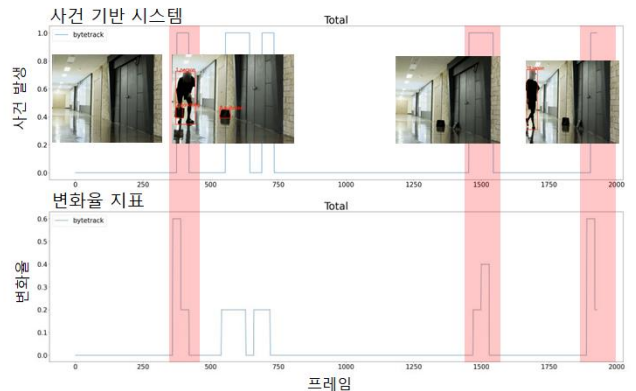
추적 모델	처리 시간 (ms/frame)
ByteTrack	50.92
BoT-SORT	58
GHOST	71.33
GeneralTrack	102.57

<표 2> 추적 모델별 변화율 알고리즘 처리시간

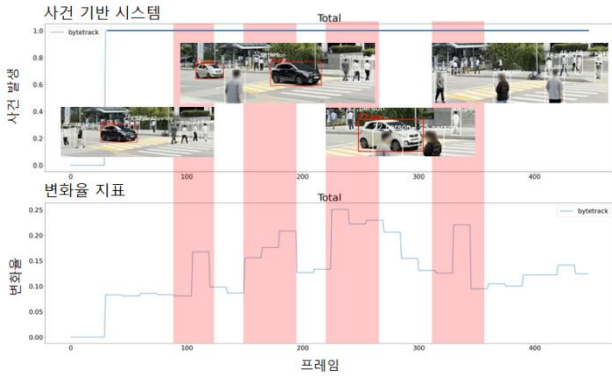
5.2 사건 기반 시스템과의 비교

Case 1 의 경우 윈도우 크기는 60 frame, 보폭은 45, 추적 모델은 ByteTrack[6]을 사용하여 변화율을 계산하였다. 실험 결과(그림 2)와 같이 사건 기반 시스템과 변화율 기반 서비스 모두 소수의 객체가 나타나고 사라지는 구간(ground truth)을 효과적으로 검출하였다.

Case 2 의 경우 윈도우 크기는 30 frame, 보폭은 15, 추적 모델은 ByteTrack[6]을 사용하여 변화율을 계산하였다. 실험 결과(그림 3), 사건 기반 서비스 그래프에서는 유의미한 변동을 확인 할 수 없었는데, 그 이유는 사건 기반 서비스가 모든 순간을 사건으로 간주했기 때문이다. 이러한 상황에서는 지휘관이 우선적으로 확인해야 할 구간을 판단하는 데 어려움이 발생 할 수 있다. 반면, 변화율 기반 변화 탐지 서비스는 소수 클래스의 출현 및 소실이 발생하는 구간(ground truth)에서 상대적으로 높은 변화율과 급격한 변화율 변동이 나타나는 것을 확인 할 수 있었다. 이를 통해 지휘관은 우선적으로 확인해야 할 중요한 구간을 파악할 수 있다.



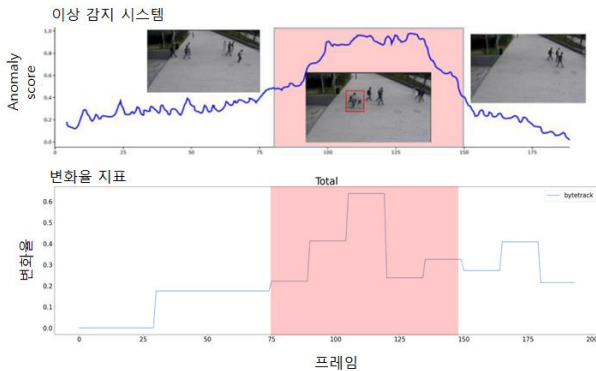
(그림 2) Case 1 에서 사건 기반 시스템과 변화율 기반 변화 탐지 서비스 비교



(그림 3) Case 2 에서 사건 기반 시스템과 변화율 기반 서비스 비교

5.3 이상 감지 시스템과의 비교

윈도우 크기는 30 frame, 보폭은 15, 추적 모델은 ByteTrack[6]으로 실험한 결과, 해당 데이터로 학습된 이상 감지 시스템[3]과 학습되지 않은 변화율 기반 변화 탐지 서비스 모두 이상 현상이 발생하는 구간(ground truth)을 (그림 4)와 같이 효과적으로 감지하였다.



(그림 4) 이상 현상 발생 상황에서 이상 감지 시스템과 변화율 기반 변화 탐지 서비스 비교

6. 결론 및 고찰

본 논문은 객체 기반 변화 탐지에서 변화를 정량적으로 평가하고 분석할 수 있도록 변화율을 새롭게 정의하고, 이를 적용한 서비스를 제시하였다. 변화율 기반 변화 탐지 서비스가 국방 AI 플랫폼에 적합한지 검증하기 위해 실시간 처리 가능성을 실험으로 확인하였으며, 기존의 사건 기반 시스템과 비교한 결과, 변화율 기반 서비스가 다양한 상황에서 지휘관에게 더 효율적인 정보를 제공할 수 있음을 확인하였다. 또한, 학습되지 않은 데이터에서도 변화율 기반 변화 탐지 서비스가 학습된 이상 감지 시스템과 유사하게 이상 현상이 발생하는 구간을 감지하는 것을 확인하였다. 이러한 실험 결과를 통해 변화율 기반 변화 탐지 서비스가 국방 AI 플랫폼에 적합함을 검증할 수 있었다.

다만, 본 논문에서는 각 추적 모델별 성능 평가를 진행하지 않았다. 실제 서비스에서 사용될 데이터를 더 확보하고, 성능 평가를 통해 서비스에 가장 적합한 추적 모델을 선정함으로써 서비스의 성능을 더욱 향상시킬 수 있을 것이다. 또한, 실시간 서비스 가능성 평가에서 FPS 가 높은 영상에서는 실시간 처리가 원활하지 않았는데, 이는 모델에 영상을 입력하고 변화율 계산 결과를 시각화하는 과정에서 병목 현상이 발생했기 때문이다. 이러한 병목 구간을 최적화한다면 더 높은 FPS 에서도 실시간 처리가 가능할 것으로 기대된다.

사사(Acknowledgement)

이 논문은 국방신속특기기술연구원의 지원을 받아 수행된 연구임(No. UC200019D)

참고문헌

- [1] 오동한, "국방 분야에서 인공지능 기술 활용실태와 XAI 기술 활용방향 제시", *Journal of Digital Contents Society*, 23, 5, 943-951, 2022.
- [2] Karbalaie A., Abtahi F., Sjöström M., "Event detection in surveillance videos: a review", *Multimed Tools Appl*, 81, 35463–35501, 2022.
- [3] Ali, Manal Mostafa, "Real-time video anomaly detection for smart surveillance", *IET Image Processing*, 17, 5, 1375-1388, 2023
- [4] A. Wang et al., "YOLOv10: Real-Time End-to-End Object Detection", arXiv:2405.14458, 2024.
- [5] Guangliang Cheng; Yunmeng Huang; Xiangtai Li; Shuchang Lyu; Zhaoyang Xu; Qi Zhao; Shiming Xiang, "Change Detection Methods for Remote Sensing in the Last Decade: A Comprehensive Review", arXiv: 2305.05813, 2023.
- [6] Y. Zhang et al., "ByteTrack: Multi-Object Tracking by Associating Every Detection Box", arXiv:2110.06864, 2022.
- [7] N. Aharon, R. Orfaig, and B.-Z. Bobrovsky, "BoT-SORT: Robust Associations Multi-Pedestrian Tracking", arXiv:2206.14651. 2022.
- [8] J. Seidenschwarz, G. Brasó, V. C. Serrano, I. Elezi, and L. Leal-Taixé, "Simple Cues Lead to a Strong Multi-Object Tracker", 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 2023, 13813-13823.
- [9] Z. Qin, L. Wang, S. Zhou, P. Fu, G. Hua, and W. Tang, "Towards Generalizable Multi-Object Tracking", 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2024, 18995-19004.
- [10] Dendorfer, P., Rezatofighi, H., Milan, A., Shi, J., Cremers, D., Reid, I., Roth, S., Schindler, K., & Leal-Taixé, L., "MOT20: A benchmark for multi object tracking in crowded scenes", arXiv:2003.09003, 2020.
- [11] W. Liu, W. Luo, D. Lian, and S. Gao, "Future Frame Prediction for Anomaly Detection - A New Baseline", 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, 6536-6545.