

## C3D 기반의 광학 흐름을 결합한 CCTV에서의 이상 탐지

박슬기<sup>o</sup>, 홍명덕<sup>\*</sup>, 조근식<sup>\*</sup>

<sup>o</sup>인하대학교 컴퓨터공학과,

<sup>\*</sup>인하대학교 컴퓨터공학과

e-mail: seolgy111@gmail.com<sup>o</sup>, hongmyungduk@gmail.com<sup>\*</sup>, gsjo@inha.ac.kr<sup>\*</sup>

## Anomaly Detection with C3D-based Optical Flow in CCTV

SeulGi Park<sup>o</sup>, MyungDuk Hong<sup>\*</sup>, GeunSik Jo<sup>\*</sup>

<sup>o</sup>Dept. of Computer Engineering, Inha University,

<sup>\*</sup>Dept. of Computer Engineering, Inha University

### ● 요약 ●

기존 CCTV 비디오에서 딥러닝 기반의 이상 탐지 연구는 객체의 행동 값만을 이용하여 이상을 탐지하였기 때문에, 시간 흐름에 따른 정보가 축소되는 문제점이 있었다. 그러나 CCTV 비디오에서의 이상의 원인은 다양한 요소와 시계열 분석에 따른 정보로 이루어져 있어 시간 정보를 유지하면서 다양한 특징 값을 사용한 모델을 설계할 필요가 있다. 따라서 본 논문에서는 C3D에 광학 흐름을 결합한 새로운 앙상블 모델을 제안한다. 실험 결과 본 논문에서 제안하는 모델이 75.83의 AUC를 얻어 기존에 연구되었던 행동 값만을 사용한 모델보다 높은 정확도를 달성하였다. 또한 이상 탐지 모델 설계 시 객체의 행동에 다양한 측면을 고려할 수 있는 여러 특징 값과 시계열 분석에 따른 정보를 사용하는 것이 적절하다는 결론을 도출하였다.

**키워드:** 이상 탐지(Anomaly Detection), 광학 흐름(Optical Flow), 딥러닝(Deep Learning)

## I. Introduction

이상 탐지는 일상생활에서 예상되는 다른 패턴을 보이는 객체나 데이터를 찾는 것을 말한다. 최근 이상 탐지는 불법 트래픽 탐지와 같은 다양한 분야에서 기존의 정확도를 높이거나 새로운 문제를 해결하는 방법으로 활발하게 사용되고 있다[1]. 기존의 이상 탐지 연구들은 데이터셋의 한계로 규칙기반 학습을 사용하는 경우가 많았으나, 빅 데이터로 이상 탐지를 위한 데이터셋이 생성되고, 딥러닝 기술의 발전으로 이상 탐지를 위한 딥러닝 기반의 여러 방법론이 제안되고 있다[2]. 본 논문에서는 효과적인 이상 탐지를 위해 C3D(Convolutional 3D) 모델에 광학 흐름(Optical Flow)을 결합한 새로운 딥러닝 기반의 이상 탐지 방법론을 제안한다.

## II. Preliminaries

### 1. Background

C3D는 딥러닝 컨볼루션 네트워크로, 공간의 특징만을 인식하는 2D 컨볼루션 네트워크를 보완하여 시간의 특징을 추가로 인식할 수 있는 3D 컨볼루션 네트워크이다. 이러한 C3D를 이용하여 객체의 행동 값을 추출하기 위해 Sport1M 데이터로 미리 학습된 가중치를 이용하여, 객체의 행동 값에 따른 시공간의 특징을 추출한다.

광학 흐름은 두 개의 연속된 비디오 프레임 사이에서 객체의 가시적

인 동작 패턴을 추출하여 객체의 움직임을 추출하는 방법이다. 이를 이용하여 모든 프레임에서 움직임이 높은 프레임에 가중치를 주는 방법으로 이상을 탐지한다.

### 2. Related Works

기존의 CCTV 감시카메라 비디오 데이터셋인 UCF-Crime 데이터셋으로 객체의 행동 값만을 이용한 이상 탐지 연구[3]는 C3D 모델의 행동 값 추출[4]에서 시간 흐름에 따른 정보가 축소되는 문제가 있다. 그러나 CCTV 비디오에서의 이상의 원인은 다양한 요소와 시계열 분석에 따른 정보로 이루어져 있기 때문에 시간 정보를 유지하면서 다양한 특징 값을 사용한 모델을 설계할 필요가 있다. 따라서 본 연구에서는 효율적인 이상 탐지를 위해 C3D 모델에 광학 흐름을 결합한 앙상블 모델을 제안하고자 한다. 이는 기존 연구의 시간 흐름에 따른 정보 축소 문제를 보완하고, 다양한 특징 값을 이용하여 이상을 탐지하는 모델이다.

### III. The Proposed Model

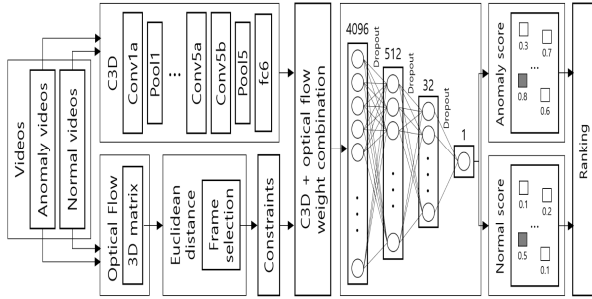


Fig. 1. Model Architecture

먼저, 전체 모델의 구성은 그림 1과 같다. 이상 비디오와 정상 비디오를 각 C3D와 광학 흐름의 입력 값으로 사용한다. 입력된 각 이상 비디오와 정상 비디오는 C3D 모델에서 객체의 행동 값으로 추출되고, 동시에 광학 흐름에서는 모든 프레임의 변화량을 계산하여 가장 큰 변화 값을 가지는 프레임을 추출한다.

C3D에서는 모든 프레임을 사용하지 않고, 16프레임 당 1프레임씩 추출하여 사용하기 때문에 시간 흐름에 대한 정보가 축소된다. 이를 보완하기 위해 광학 흐름에서는 모든 프레임을 입력 값으로 변화 값을 추출하여, 시간 흐름에 대한 정보 축소 문제를 보완한다. 추출된 변화 값은 3차원 행렬로 유클리디안 거리를 이용하여 매 프레임에서의 이동 값을 계산한다.

계산된 변화 값을 이용하여, 전체 프레임에서 객체의 이동이 많은 프레임을 추출하고, 추출된 프레임을 기준으로 가중치를 부여한다.

가중치 부여 단계에서 변화 값이 큰 프레임 기준으로 가중치를 부여하면, 카메라의 시점이 변경되는 부분이나 샷 경계의 부분에서는 해당 프레임의 모든 포인트가 이동하기 때문에 이상 값이 크게 나타날 수밖에 없다. 이를 제약하기 위해 광학 흐름의 출력인 3차원 벡터의 전체 포인트(240x360, 76,800)에서 90%(69,120) 이상 포인트가 이동한다면, 시점 변화로 판단하여 해당 프레임에서는 가중치를 부여하지 않는다.

이후 C3D와 광학 흐름의 값을 결합하여 4096개의 행동 값에 따른 벡터를 드롭아웃을 통해 1개의 이상 값으로 출력한다. 출력된 이상 값은 해당 비디오에서 가장 큰 이상 값을 가진 부분으로 탐지한다.

### IV. Experiments

#### 1. Dataset

실험에는 UCF-Crime[3] 데이터셋을 사용하였다. UCF-Crime 데이터셋은 총 13개의 범죄 비디오와 정상 비디오로 구성되며, 레이블은 비디오 레벨에서의 레이블만 존재하기 때문에 효율적인 이상 탐지를 위해 각 범주의 분류 문제가 아닌, 이상 비디오와 정상 비디오를 구분하는 바이너리 문제로 실험하였다.

### 2. Experiment Result and Comparison

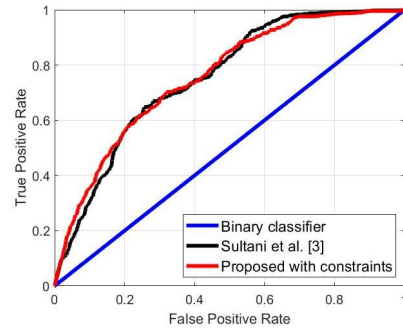


Fig. 2. ROC Curve

그림 2는 UCF-Crime 데이터셋에 대해 바이너리 분류와 행동 값만을 이용한 이상 탐지 모델[3], 본 논문에서 제안하는 모델 간에 ROC(Receiver Operating Characteristic) curve를 비교한 결과이다. ROC curve는 False positive rate와 True Positive Rate 사이의 관계를 나타내는 곡선이다.

Table 1. AUC Value

Method	AUC
Binary classifier	50.00
Sultani et al. [3]	75.41
Proposed w/o constraints	74.28
Proposed w constraints	<b>75.83</b>

표 1은 UCF-Crime 데이터셋에 대해 본 논문에서 제안하는 모델과 행동 값만을 이용한 이상 탐지 모델[3] 간에 AUC(Area Under the Curve) 정확도를 비교한 결과이다. 제약조건이 없는 방법은 시점 변화와 샷 경계와 같은 이상이 아닌 부분은 제외하지 않은 모델로 74.28의 AUC 값을 나타내며, 제약조건이 있는 방법은 시점 변화와 샷 경계 프레임에 가중치를 주지 않아, 이상이 아닌 부분은 제외한 모델로 75.83의 AUC 값으로 가장 높은 성능을 달성하였다. 구체적으로 폭발과 같은 빠르게 움직이는 객체가 있는 CCTV 비디오의 경우, 행동 값만을 사용한 모델[3]과 비교하여 4.69% 탐지율이 상승하였다.

### V. Conclusion

본 연구에서는 C3D 모델에 광학 흐름을 결합하여 효율적으로 이상을 탐지할 수 있는 앙상블 모델을 구성하였다. 실험 결과 행동 값만을 사용한 모델보다 높은 75.83의 AUC를 달성하여 다양한 요소와 시계열 분석에 따른 정보의 값을 동시에 이용하도록 모델을 구성하는 것이 이상 탐지 정확도 향상에 도움이 되는 것을 확인하였다. 향후 효율적인 이상 탐지를 위해 제약조건들의 세부 내용을 수정하여 추가 연구할 계획이며, 여러 범주에 따른 다양한 방법을 적용하면 보다 높은 이상 탐지가 가능할 것으로 예상되어, 이를 추가 연구할 계획이다.

## ACKNOWLEDGMENT

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT연구센터지원사업의 연구결과로 수행되었음" (IITP-2017-0-01642)

## REFERENCES

- [1] R. Chalapathy, and S. Chawla, "Deep learning for anomaly detection: A survey." arXiv preprint arXiv:1901.03407, 2019.
- [2] RT. Ionescu, FS. Khan, MI. Georgescu, and L. Shao, "Object-centric auto-encoders and dummy anomalies for abnormal event detection in video." CVPR, 2019.
- [3] W. Sultani, C. Chen, and M. Shah. "Real-world anomaly detection in surveillance videos." arXiv preprint arXiv:1801.04264, 2018.
- [4] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri. "Learning spatiotemporal features with 3d convolutional networks." ICCV, 2015.