

YOLO 기반 학교폭력 감지 시스템

신찬휘^o, 문미경^{*}

^o동서대학교 소프트웨어학과,

^{*}동서대학교 소프트웨어학과

e-mail: 20181548@g.dongseo.ac.kr^o, mkmoon@dongseo.ac.kr^{*}

YOLO-based School Violence Detection System

Chanhwi Shin^o, Mikyeong Moon^{*}

^oDept. of Software, Dongseo University,

^{*}Dept. of Software, Dongseo University

● 요약 ●

학교폭력은 교육 환경에서 심각한 문제이다. 피해자에게 심리적 고통과 육체적 상해를 입히고 학교 내 안전과 안정성을 위협한다. 이에 많은 교육기관과 정부 기관이 학교폭력 예방과 대처를 위한 다양한 방안을 제시하고 있지만, 여전히 어려운 문제이다. 최근에는 인공지능 기술을 활용하여 학교폭력 방지와 대처에 관한 연구가 이루어지고 있다. 본 연구에서는 YOLOv5(You Only Look Once version 5) 딥러닝 알고리즘을 활용하여 학교 내부에서 발생하는 폭력 행위를 실시간으로 탐지하는 모델을 제안한다. 이 모델은 CCTV와 같은 영상 데이터를 입력으로 받아들이며 학교 내부에서 발생하는 폭력 행위를 실시간으로 식별하는 것을 목표로 한다.

키워드: YOLOv5(You Only Look Once version 5), Backbone Network(백본 네트워크), Focal Loss(손실 함수), Feature Pyramid(특징 피라미드)

I. Introduction

학교폭력은 교육 환경에서 가장 심각한 문제 중 하나이다. 학교폭력은 피해자에게 심리적 고통과 육체적 상해를 입히며, 학교 내부의 안전과 안정성을 위협한다. 또한, 학생들의 학업 성취도와 학교생활 만족도를 낮추며, 학습과 성장에 대한 장애요인이 된다. 이러한 문제에 대응하기 위해 많은 교육기관과 정부 기관이 학교폭력 예방과 대처를 위한 다양한 방안을 제시하고 있다. 교육부에서 초·중·고 학생들을 대상으로 실시한 학교폭력 실태조사에서 학교폭력 중 신체 폭력의 비율이 2021년 12.4%, 2022년 14.6%로 언어폭력에 이어 두 번째 높은 비율을 차지했다.

최근에는 인공지능 기술이 학교폭력 방지와 대처에 활용될 수 있다는 연구들이 진행되고 있다. YOLOv5(You Only Look Once version 5)는 Backbone Network와 Focal Loss를 통하여 객체 탐지에 높은 정확도와 빠른 속도를 보여준다. 또한 Feature Pyramid를 사용하여 다양한 크기의 객체를 탐지한다. 본 논문은 YOLOv5[1] 모델을 사용하여 영상 데이터를 입력받아 폭력 행위를 실시간으로 탐지할 수 있는 모델을 제안한다.

II. Preliminaries

연구 [2]은 하키 영상을 사용하여 폭력을 분석하고 탐지하는 방법을 연구하였다. 이는 특정 환경과 상황에 한정된 폭력 유형을 다룰 수밖에 없는 한계가 있다. 그에 반해 본 연구는 AI-Hub의 이상행동 CCTV 데이터 셋을 활용함으로써 다양한 현실적인 폭력 사례를 포함한다. 이상행동 CCTV 데이터 셋은 실제 일상 환경, 공공장소에서 촬영된 영상을 기반으로 구축되었으며, 다양한 폭력행위를 포함하고 있다. 펀칭, 발차기, 눕히고 폭행하는 등의 폭력 행동은 실제 사회에서 발생할 수 있는 형태이므로, 이를 통해 본 연구는 이전 연구들과 비교하여 보다 더 현실적인 폭력 사례를 포착하고 분석하는 데 있어서 더 나은 탐지 성능을 제공한다.

연구 [3]에서는 폭력행위를 객체의 너비와 높이 변화량 및 변화 횟수를 기준으로 판단한다. 만약 일정 횟수 이상의 변화가 20초 내에 측정된다면 폭력행위로 간주한다. 본 연구는 이러한 메트릭 측정이 아닌 실시간으로 폭력 상황 객체를 검출하여 즉각적으로 감지한다. 이를 통해 폭력 객체를 실시간으로 검출하여 클라이언트에게 즉각적인 시각화 결과를 제공하며, 실시간 CCTV 데이터를 이용한 모니터링이 가능하다.

III. The Proposed Scheme

본 연구에서는 CCTV 등 영상 감시 장치로부터 전송되는 데이터를 전처리하고, Flask 웹 서버와 YOLOv5 모델을 결합한 서버에서 객체 탐지를 수행한다. 또한 감지된 객체의 경계 박스를 저장 및 실시간으로 모니터링에 적용하며, 인식된 내용은 관련 부서에도 알림으로 전송한다.

1. DataSet

본 연구에서는 YOLOv5 모델을 사용하여 AI-Hub의 이상행동 CCTV 데이터를 선별하여 학습시켰다. 이 데이터 셋은 다양한 이상행동 상황을 담고 있으며, 그중에서도 폭력 상황을 감지하기 위해 Fig 1과 같이 폭행 직전의 동작을 취하는 장면과 Fig 2와 같은 폭행 직후의 장면을 선별하여 학습시켰다. 이를 통해 모델은 폭력 상황에서 객체를 더욱 정확하게 감지할 수 있게 되었다.



Fig. 1. 폭행 직전 이미지



Fig. 2. 폭행 직후 이미지

2. Violence detection using YOLOv5

본 논문에서는 YOLOv5를 활용하여 학습 모델을 구현하였다. 학습은 총 200 epochs를 수행하였으며, batch의 크기는 2로 설정했다. 입력 이미지는 416x416 크기로 조정했다. 이와 같은 설정을 통해 200개의 이미지에 대한 학습이 수행되었으며, Fig 3과 같이 실시간 영상 데이터에서 폭력 객체를 91%로 감지하는 높은 정확도를 보여준다. 감지되는 폭력의 종류는 머리채 잡기, 주먹질, 발차기, 눕히고 때리기가 해당된다.



Fig. 3. 실시간 영상 속 폭력 검출

IV. Conclusions

본 논문에서는 YOLOv5 모델을 활용한 폭력 감지 시스템을 연구하였다. 학습 데이터로 사용한 이상행동 CCTV 영상 속 폭력은 높은 검출률을 보였다. 또한 Web-Cam 혹은 휴대폰 카메라를 통한 유사 폭력 행동에서도 감지하는 모습을 보여주었다. 향후 계획은 검출률을 높이기 위해 더 많은 상황에서의 폭력 데이터를 학습시키고, 유사 폭력 행동을 검출하는 방향으로 이어갈 것이다.

ACKNOWLEDGEMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음(2019-0-01817)

REFERENCES

- [1] Jocher, G. YOLOv5 by Ultralytics (Version 7.0) [Computersoftware].<https://doi.org/10.5281/zenodo.3908559>, 2020.
- [2] N.Y.Lee, C.W.Jo, K.Y.Lee. Real-Time Violence Detection System using YOLO. Proceedings of the Korean Institute of Electronics and Electrical Engineers Conference, pp.754-755, 2020.
- [3] Y.B.Shim, H.J.P. A Study on a Violence Recognition System with CCTV. Journal of Digital Contents Society, Vol. 16, No. 1, pp.25-32, 2015.