

스켈레톤 정보를 활용한 공공 보안용 CCTV의 이상행동 학습용 데이터 적합성 검증 연구

A Study on Verification of Data Suitability for Learning Abnormal Behavior Using Skeleton Information of CCTV for Public Security.

김경수¹, 이세림¹, 이지상¹, 이양선¹ · 손윤식²

KyoungSoo Kim, SeRim Lee, JiSang Lee, YangSun Lee and Yunsik Son

¹서경대학교 컴퓨터공학과

E-mail: {tictac49, leeserim3287, 2017305064, yslee}@skuniv.ac.kr

²동국대학교 컴퓨터공학과

E-mail: sonbug@dongguk.edu

요 약

인공지능은 인간의 지능이 가지는 기능을 갖춘 컴퓨터 시스템을 말한다. 최근에는 컴퓨터뿐만 아니라 인간의 삶에서도 빠질 수 없는 것으로 발전해오고 있다. 이에 더불어 한국지능정보사회진흥원은 인공지능 개발에 중요한 요소 중 하나인 학습용 데이터의 품질을 중요하다 인식하여 인공지능 학습데이터 구축사업을 추진하고 있다.

본 논문은 LSTM과 유사하나 구성원간 상호 종속성을 모델링하고 영상 속 인물의 스켈레톤 정보를 사용하는 MPED-RNN 모델을 이용해 다양한 평가지표로 공공 보안용 CCTV에서 이상행동 발생 시 이상행동을 검출하기 위한 데이터가 학습용 데이터로 적합한지 검증해보았다.

키워드 : 인공지능, 학습용 데이터, 이상행동, 스켈레톤 정보, MPED-RNN

1. 서 론

인공지능은 인간의 지능이 가지는 기능을 갖춘 컴퓨터 시스템을 말한다. 최근에는 컴퓨터뿐만 아니라 인간의 삶에서도 빠질 수 없는 것으로 발전해오고 있다. 이에 더불어 한국지능정보사회진흥원은 인공지능 개발에 중요한 요소 중 하나인 학습용 데이터의 품질을 중요하다 인식하여 인공지능 학습데이터 구축사업을 추진하고 있다.

본 논문은 LSTM과 유사하나 구성원간 상호 종속성을 모델링하고 영상 속 인물의 스켈레톤 정보를 사용하는 MPED-RNN 모델을 이용해 다양한 평가지표로 공공 보안용 CCTV에서 이상행동 발생 시 이상행동을 검출하기 위한 데이터가 학습용 데이터로 적합한지 검증해보았다.

2. 관련 연구

2.1 스켈레톤 정보

스켈레톤 정보란 딥러닝 기반 포즈 인식으로 사람을 인식하여 사람의 움직이는 관절부위를 탐지하여 미리 정해진 skeleton을 분할 하여 나타내는 것을 말한다. [1].

이 성과는 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행되는 연구임(No.2018R1A5A7023490).

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2020R1A2C1013296).

2.2 MPED-RNN 모델

MPED-RNN 모델은 스켈레톤 데이터를 입력으로 받아 학습 시킨 후 평가하는 오토인코더 기반의 이상행동 탐지 모델이다. 오토 인코더와 같이 인코더-디코더로 구성 되어 있으며 스켈레톤 궤적의 시공간 패턴을 특징으로 한다.

MPED-RNN 모델은 입력인 스켈레톤 정보를 프레임 내에서 바운딩 박스의 모양, 크기 및 변형이 거의 없는 큰 움직임에 대한 정보인 global과 스켈레톤 움직임의 내부 변형과 같은 미세한 움직임에 대한 정보인 local로 분해하여 사용한다.

3. 스켈레톤 정보를 활용한 학습용 데이터 검증

그림 1은 본 논문에서 학습용 데이터를 검증하기 위한 시스템 구성도이다.

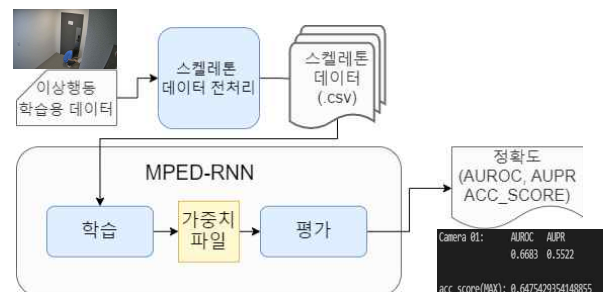


그림 2. 학습용 데이터 검증 시스템

3.1 스켈레톤 정보 전처리

MPED-RNN 모델의 학습과 평가 과정에서 영상에서 추출한 인물별 스켈레톤 정보가 필요하다. 해당 스켈레톤 정보를 이상행동 학습용 데이터의 영상에서 추출하였다. 추출한 파일은 인물별로 만들어지며 인물별 각 영상에서의 프레임 번호와 프레임별 스켈레톤 좌표를 가지고 있다.

추출한 스켈레톤 정보를 입력 데이터와 같은 형태로 만들기 위하여 영상의 각 인물별로 파일을 나누고 파일마다 인물이 등장한 프레임 번호와 해당 프레임에서 관측된 관절 17개의 좌표를 csv 파일 형태로 저장하였다. 그림 2는 제작한 스켈레톤 정보 파일을 나타낸다.

frame	joint_1	joint_2	joint_3	joint_4
1	1272	274	1369	318
2	1278	277	1366	318
3	1272	277	1366	318
4	1266	274	1366	318
5	1260	271	1366	319
6	1260	271	1363	318
7	1260	274	1363	318

그림 4. 스켈레톤 정보 파일
Fig 2. Skeleton data file

3.2 학습

제작한 스켈레톤 정보 파일을 사용하여 학습하였다. 총 4종류의 이상행동을 각각 80개의 영상으로 학습하였다. 그림 3은 학습에 사용된 파일 중 일부이다.

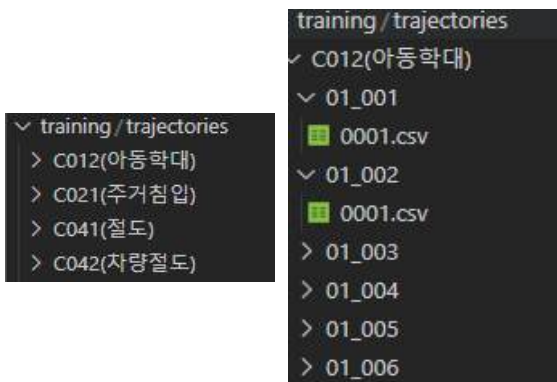


그림 3. 학습에 사용된 파일 일부

3.3 평가

이상행동 분류의 정확도 평가를 위해서 MPED-RNN은 스켈레톤 정보와 프레임별 이상행동 유무를 0과 1로 표현한 이진 파일인 프레임 레벨 마스크(frame level mask) 파일을 사용한다. 해당 파일은 학습용 데이터와 함께 제공된 주석 파일에서 이상행동이 발생하는 프레임들을 사용하여 제작하였다.

중간 결과는 비지도 학습의 모델이기 때문에 프레임별 이상행동 유무의 예측값이 나오게 된다. 해당 예측값을 이용하여 임의의 분류점을 기준으로 TPR과 FPR을 계산하여 AUROC를 출력하였다.

프레임별 예측값을 가지고 정상으로 예측한 프레임 중에 실제 정상인 비율인 정밀도와 실제 정상인 프레임 중에 정상으로 예측한 비율인 재현율을 가지고 정밀-재현 그래프의 아래 면적인 AUPR을 출력하였다.

AUROC 출력에서 사용하는 분류점을 조절하여 정확도가 가장 높아지는 구간을 확인한 후 해당 최적분류점에서의 정확도를 출력하였다 [2].

4. 실험결과 및 분석

4.1 이상행동별 평가 결과

그림 4는 각각 아동학대, 주거침입, 절도, 차량절도 학습 데이터의 평가 결과를 나타낸다.

Camera 01:	AUROC	AUPR
	0.5222	0.3682
acc_score(MAX):	0.3029017615908354	
Camera 02:	AUROC	AUPR
	0.1889	0.3387
acc_score(MAX):	0.5334833833500056	
Camera 03:	AUROC	AUPR
	0.5821	0.5789
acc_score(MAX):	0.5739511897159675	
Camera 04:	AUROC	AUPR
	0.0768	0.2652
acc_score(MAX):	0.39348408584378836	

그림 4. 이상행동별 평가 결과

학습 영상의 총 프레임 수 중에 이상행동 프레임의 수가 50% 넘는 주거침입과 차량절도 이상행동의 경우 평가 결과가 매우 낮게 나오는 것을 확인하였다.

4.2 종합 정확도

그림 5는 전체 학습 데이터로 학습을 하여 정확도를 확인한 결과이다.

Camera 01:	AUROC	AUPR
	0.6683	0.5522
acc_score(MAX):	0.6475429354148855	

그림 5. 이상행동 종합 평가 결과

AUROC는 약 0.67, AUPR은 약 0.55, 최적분류점 정확도는 약 0.65가 출력된 것을 확인하였다.

5. 결 론

본 논문은 스켈레톤 정보를 사용하는 MPED-RNN 모델을 이용해 이상행동 학습용 데이터가 학습용 데이터로 적합한지 검증해보았다.

본 논문에서 검증한 데이터가 이상행동 프레임 수가 영상 전체 프레임 수의 50% 이상인 경우가 많아 비지도 학습 기반의 모델에서는 높은 정확도를 출력하진 않지만, 이상행동의 프레임과 아닌 프레임이 명확하게 구분되기 때문에 지도 학습 기반의 모델에서는 높은 정확도를 출력할 수 있다고 판단된다.

참 고 문 헌

- [1] loner의 학습노트, 딥러닝 기반 포즈 인식(skeleton 탐지), <https://wiserloner.tistory.com/1103>
- [2] 홍종선, 전해선, 신혜수, "부분 AUC와 VUS를 최대화하는 선형결합 스코어에 대한 최적분류점 구간," 한국데이터정보과학회지, 제30권 4호, pp. 759-770, 2019.