

C3D 모델을 활용한 이상행동 분류 및 영상 구간 분석



전은성¹. 이양선¹. 손윤식²

¹서경대학교 컴퓨터공학과

Email: {jjjes758, yslee}@skuniv.ac.

kr

²동국대학교 컴퓨터공학과

E-mail: sonbug@dongguk.edu



목 차

1. 서론
2. 관련 연구
3. C3D 모델을 활용한 이상행동 분류 및 영상 구간 분석
4. 실험 결과
5. 결론

1. 서론

● 연구 배경

- 이상행동이란 통계적 기준, 사회적 규범 그리고 인간행동 유형에서 벗어나는 행동을 뜻함
- 요즘 소위 ‘묻지마 범죄’ 등 이러한 이상행동인 범죄 행위로 인해 사람들이 불안해하고 있음
- 최근 CCTV, 카메라 사용이 증가함에 따라 머신러닝, 딥러닝을 활용한 HAR(Human Action Recognition) 기반의 이상 탐지 연구로 이상행동을 탐지하고 선조치할 수 있음

● 연구 목표

- AI 영상감시 분야의 이상행동 탐지(Anomaly Detection) 과정 중 시공간적 특징을 잃지 않게 되는 3D Convolution Network를 포함하는 C3D 모델을 사용하여 이상행동의 전조부터 발생을 탐지
- 탐지과정 중 영상을 16프레임의 조각으로 나누어 이상행동의 세부 분야를 구분하며 영상의 구간을 분석
- 결론적으로 국내 보안 및 안전 전문 민간회사의 관련 기술 수준을 향상하고, 지능형 CCTV의 상용화로 국민의 범죄 두려움을 낮추고 사회 전반의 안전에 대한 신뢰를 높일 것으로 판단됨

2. 관련 연구

2.1 이상행동 탐지(Anomaly Detection)

2.2 C3D 모델

2.1 이상행동 탐지(Anomaly Detection)

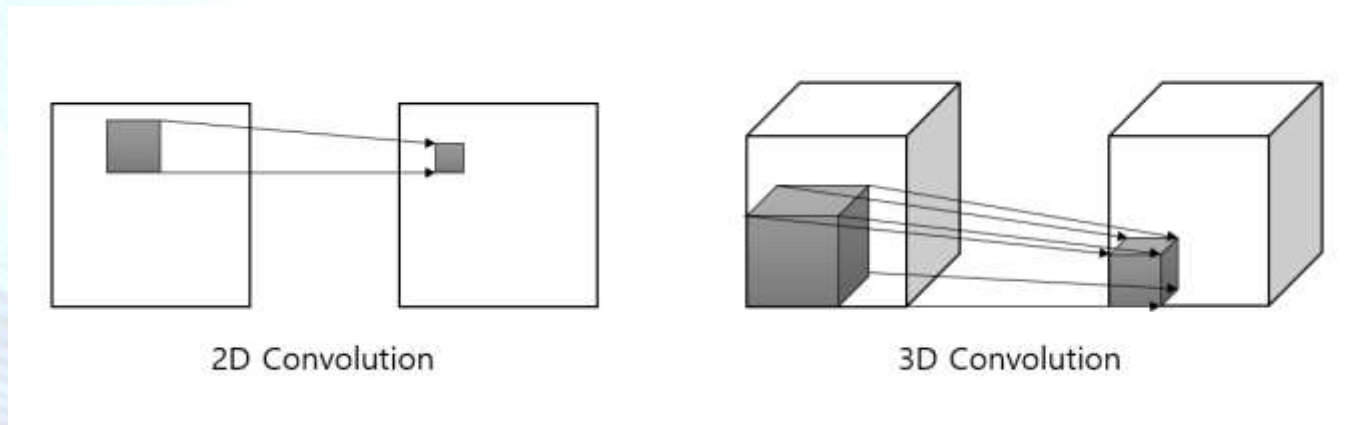
- 통계적 기준, 사회적 규범 그리고 인간행동 유형에서 벗어나는 행동을 딥러닝을 사용하는 신경망 모델을 이용하여 CCTV, 카메라 영상을 분석하여 영상 속 사람의 행동이 정상행동인지 이상행동인지 판별하는 것을 뜻함

1	0: 아동학대 (신체학대)
2	1: 절도 (문 앞)
3	2: 침입 (문 앞)
4	3: 절도 (주차장)

실험에 사용된 이상행동 분야

2.2 C3D 모델

- 3D Convolution Network를 대량의 영상 데이터에 지도학습 방식으로 훈련한 모델

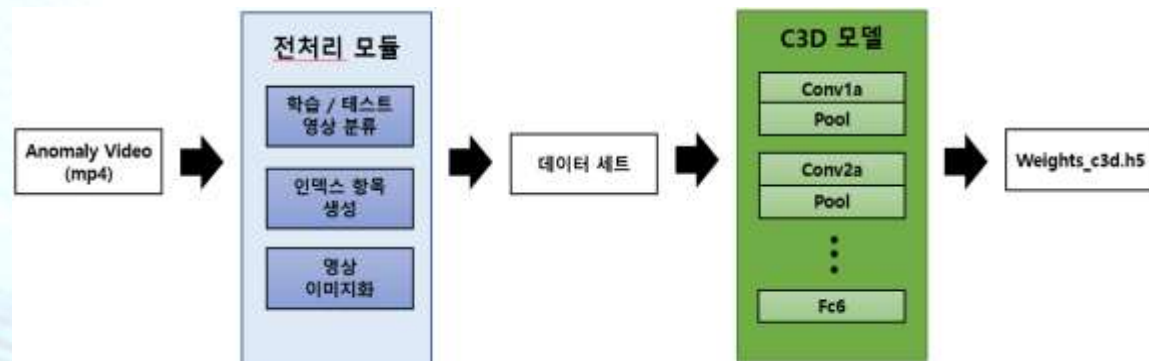


2D Convolution 과 3D Convolution

- 여러 이미지나 영상을 2D Convolution Network로 처리하게 되면 시공간적인 특징을 잃기 때문에 시공간적 특징을 잃지 않게 되는 3D Convolution Network를 사용

3. C3D 모델을 활용한 이상행동 분류 및 영상 구간 분석

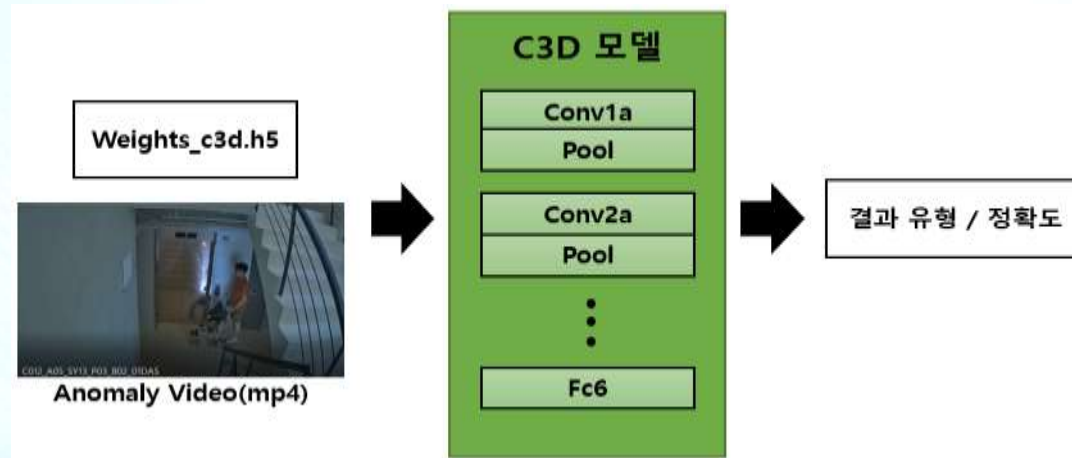
- 3D Convolution Networks를 사용하는 C3D 모델은 시공간적 특징을 잃지 않기 때문에 대량의 영상 데이터를 학습하고 검증할 때 효율적



C3D 모델 학습 과정

- 데이터 세트 영상들을 학습/테스트 영상으로 분류
- 각 영상에 해당하는 범죄유형, 의심 행동, 전조현상의 인덱스 항목을 생성 및 데이터 세트 구축
- C3D 모델을 통해 학습하여 가중치 파일을 생성

3. C3D 모델을 활용한 이상행동 분류 및 영상 구간 분석



C3D 모델 테스트 과정

- 학습 결과로 생성된 **weight 파일**과 **이상행동 영상**을 입력으로 사용
- 영상을 **16프레임 조각**으로 나누어 범죄, 의심 행동, 그리고 전조현상을 평가하여 유사한 유형에 점수를 부여하여 결과 유형과 정확도를 판별

3. C3D 모델을 활용한 이상행동 분류 및 영상 구간 분석

3.1 학습데이터 구성

3.2 전처리 과정

3.3 이상행동 학습

3.4 이상행동 분류 및 영상 구간 분석

3.1 학습데이터 구성

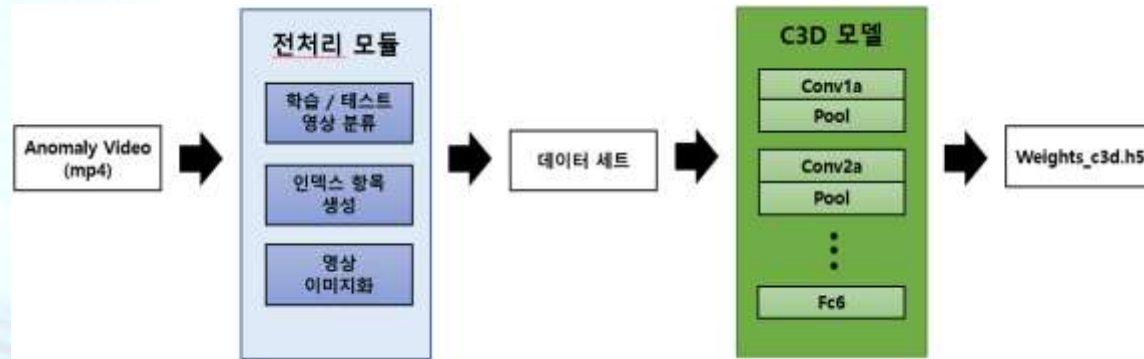
<인공지능 학습데이터 이상행동 시나리오: 4가지 범죄유형(아동학대, 주거침입, 폭행/강도, 절도)>

No.	범죄유형	전조현상	세부 분류	이상행동(행위)	구축 비율
1	아동학대	1,2,3	방임	아이 혼자 있음	7%
2		7,8	신체학대	성인이 아이를 세게 밀침	7%
3		5,8	신체학대	성인이 아이를 발로 참	7%
4		6,8	신체학대	성인이 물체로 아이를 때림	7%
5		13	방임	유모차 방치	3%
6		13	신체학대	유모차를 세게 밀침	4%
7		13	신체학대	유모차를 발로 참	4%
8		13	신체학대	유모차를 세게 올려내리침	4%
9	주거침입	15,16,17	문	서성거림	2%
10		15,16,17	문	도어락 오픈 시도	2%
11		15,16,17	문	문을 발로 참	2%
12		15,16,17	문	문 안쪽을 보려고 시도	2%
13		15,16,17	문	문을 두드림	2%
14	폭행/강도	18	위협물체	칼로 위협	9%
15		19	위협물체	방망이로 위협	9%
16		21,22,23	위협행동	주먹으로 침	9%
17		21,22,23	위협행동	뺨때리기	9%
18		21,22,23	위협행동	발차기	9%
19	절도	25,26,27	문 앞	택배 훔쳐가기	1%
20		28,29,30	주차장	차량 앞 서성거림	1%

E2ON 영상 데이터 세트 구성

- 이상행동 영상으로 E2ON 영상 데이터 세트를 사용
- E2ON 영상 데이터는 아동학대, 주거침입, 폭행/강도, 절도의 4가지 범죄유형과 29가지 세부 유형으로 구분

3.2 전처리 과정



C3D 모델 전처리 과정

- 모델에 적용하기 위한 데이터 세트 구축을 위해 전처리 과정을 진행

3.2.1 학습·테스트 영상 분류

- 학습을 위한 학습 영상과 학습된 결과를 바탕으로 테스트를 진행할 테스트 영상이 필요
- 학습 영상과 테스트 영상의 비율을 8:2로 분류를 하고 리스트 파일을 생성
- 분류 방식은 영상을 읽어오면서 1~10의 난수를 부여하고 각 영상에 부여된 수가 2를 초과하면 학습영상으로, 반대의 경우는 테스트 영상으로 분류

```
1 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 1 0
2 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 17 0
3 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 33 0
4 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 49 0
5 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 65 0
6 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 81 0
7 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 97 0
8 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 113 0
9 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 129 0
10 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 145 0
11 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 161 0
12 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 177 0
13 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 193 0
14 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 209 0
15 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 225 0
16 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 241 0
17 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 257 0
18 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 273 0
19 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 289 0
20 main_train_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B01_01DAS 305 0
```

학습 영상 리스트

```
1 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 1 0
2 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 17 0
3 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 33 0
4 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 49 0
5 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 65 0
6 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 81 0
7 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 97 0
8 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 113 0
9 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 129 0
10 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 145 0
11 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 161 0
12 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 177 0
13 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 193 0
14 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 209 0
15 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 225 0
16 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 241 0
17 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 257 0
18 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 273 0
19 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 289 0
20 main_test_img/ChildAbuse(PhysicalAbuse)/C012_A05_SY13_P03_B02_01DAS 305 0
```

테스트 영상 리스트

3.2.2 인덱스 항목 생성

- 영상에는 4가지 범죄유형과 29가지 세부 유형이 존재
- 동영상의 제목에서 범죄 및 세부 유형을 추출하여 **인덱스 항목을 생성**

```
1 0: ChildAbuse(PhysicalAbuse) 1 0: A01: Child is alone.
2 1: HomeInvasion(InfrontDoor) 2 1: A05: Stroller left unattended.
3 2: Theft(InfrontDoor) 3 2: A06: Push the stroller hard.
4 3: Theft(Park) 4 3: A07: Hold the stroller with your feet.
5 | 5 4: A08: Pulled the stroller down hard.
6 6 5: A14: Adult throws a child.
7 6: A16: Overturn the stroller.
8 7: A17: Pacing around.
9 8: A18: Trying to open the door lock.
10 9: A19: kick the door.
11 10: A20: Trying to look inside the door.
12 11: A21: Knocked on the door.
13 12: A29: Hitting with tools.
14 13: A30: Stealing packages.
15 14: A31: Hanging around in front of a car.
16 15: N0: Normal behavior.
17 16: N1: Normal behavior.
18 17: SY13: A kid is in a stroller.
19 18: SY14: A child is walking around.
20 19: SY15: A person is pacing around the door.
21 20: SY16: A person is standing around the door.
22 21: SY17: A person is sitting around the door.
23 22: SY25: A person is pacing in front of the package.
24 23: SY26: A person is standing around the delivery.
25 24: SY28: A person standing in front of a car.
26 25: SY29: A person is standing around a car.
27 26: SY30: A person is sitting around a car.
28 27: SY31: A person is leaning against the door.
29 28: SY32: A person is leaning against a wall (or pole).
```

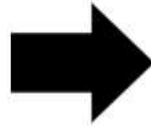
범죄유형 • 세부 유형 인덱스

3.2.3 영상 이미지화

- 영상을 읽으면서 모든 프레임별 이미지를 저장하여 데이터 세트를 구축



Anomaly Video(mp4)



영상 이미지화

3.3 이상행동 학습

- 영상의 프레임별 이미지들을 16프레임씩 분할을 하고 이미지의 특징을 추출
- 생성된 인덱스와 대응하여 C3D 모델에서 학습

epoch	loss	acc	val_loss	val_acc
0	2.2488326258794347	0.964267106634248	0.2009035155977834	0.9895969855832241
1	0.22770594007065736	0.9821378951024662	0.22307701216300224	0.9867463958060289
2	0.2220583382705385	0.9835576589093435	0.23871958106951507	0.9801769331585846
3	0.21014006002034608	0.9852596387634596	0.22562760266220272	0.985353866317169
4	0.14538798414250967	0.9950503647099688	0.13364187187440743	0.992300131061599
5	0.12592618663036065	0.9949982632858632	0.13118829240297428	0.9921690694626475
6	0.1207662018011452	0.9950720736366794	0.13351111710071564	0.9909239842726081
7	0.11850334703197658	0.9951328586314693	0.12503713977719666	0.9914318479685452
8	0.11140460658711085	0.9965699895797152	0.11907216504821615	0.992840760157274
9	0.10940409067056792	0.9968044459881903	0.12161130472763806	0.9931847968545217

범죄유형 정확도와 손실

epoch	loss	acc	val_loss	val_acc
0	4.386001282269946	0.3343728222996516	2.181888134298342	0.5307816197738627
1	2.1082534808481195	0.5970426829268293	2.037021577405265	0.6415330002629503
2	2.0088979222051773	0.6555357142857143	2.0261223289234964	0.659265711280568
3	1.9708151309498512	0.6757012195121951	1.9818540001241527	0.665182093084407
4	1.3779431475247241	0.8043728222996516	1.3393841424867288	0.7808966605311596
5	1.20472930814331	0.822630662020906	1.29103415566559	0.7877991059689718
6	1.1678720084964607	0.8304311846689896	1.2886243817870817	0.7850874309755457
7	1.153011773729158	0.8341419860627177	1.3544625628944325	0.7683243491980016
8	1.0268331209417003	0.8725783972125436	1.1798956914697858	0.8247436234551669
9	0.9891397418959216	0.8791289198606271	1.1477968731648729	0.8271101761767026

세부유형 정확도와 손실

3.4 이상행동 분류 및 영상 구간 분석

- 가중치가 저장된 weight 파일과 학습 • 테스트 영상 분류에서 나온 테스트 영상을 활용하여 이상행동 분류 및 영상 구간 분석 실시
- 테스트 영상을 16프레임으로 분할 하여 생성된 이미지가 어떤 분야에 가장 유사한지에 대한 점수를 매긴 후 분류

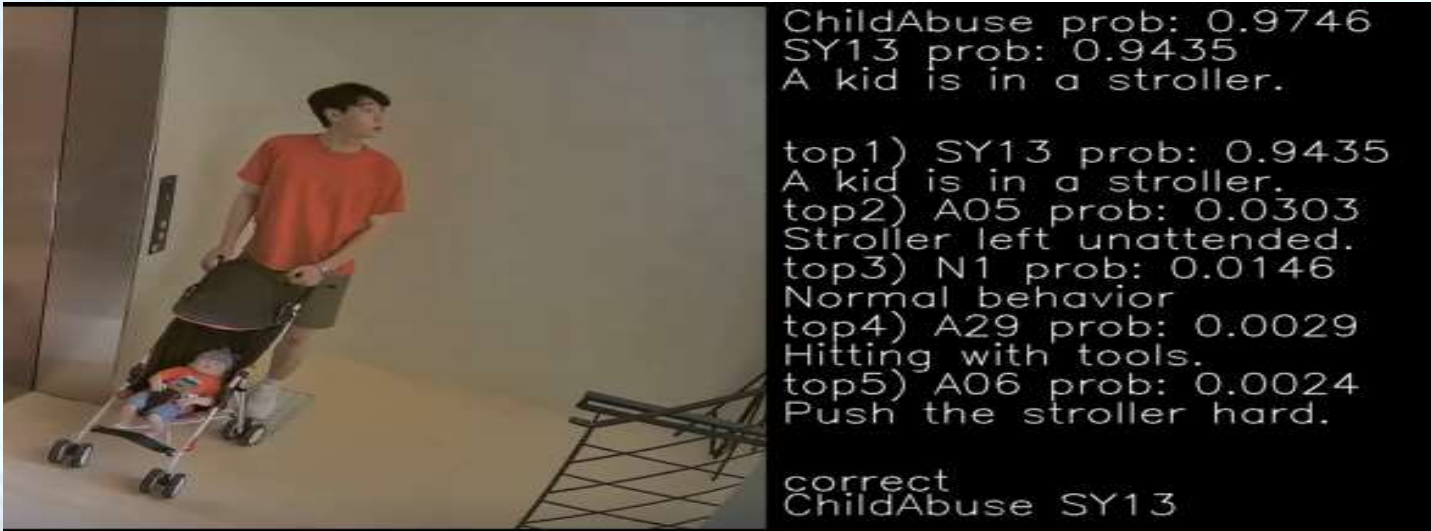
Invasion prob: 0.9993	이상행동 분류
A19 prob: 0.9682	
kick the door.	영상 구간 분석


```
top1) A19 prob: 0.9682
kick the door.
top2) N1 prob: 0.0069
Normal behavior
top3) A18 prob: 0.0049
Trying to open the door lock.
top4) SY17 prob: 0.0046
A person is sitting around the door
top5) A17 prob: 0.0042
Pacing around.

correct
Invasion A19
```

분류 테스트 결과 예시

4. 실험 결과



아동학대 영상 테스트 결과

- 이상행동 동영상을 16프레임의 조각으로 나눈 후 각각 범죄유형, 의심 행동, 전조현상에 점수를 매긴 후 분류를 하는 방식으로 진행
- 아동학대(신체학대) 영상의 테스트 결과로 영상의 모든 프레임 중 한 프레임을 캡처한 결과
- 우측 상단 부분에는 범죄 분야인 아동학대, 그 하단에는 세부 유형인 SY13이 각각 정확도가 0.9746, 0.9435로 분류되었으며 영상 구간 분석 결과인 “A kid is in a stroller.”가 출력

5. 결론

- 이상행동이란 통계적 기준, 사회적 규범 그리고 인간행동 유형에서 벗어나는 행동을 뜻함
- 최근 ‘묻지마 범죄’ 등 이러한 이상행동인 범죄 행위로 인해 사람들이 불안해하고 있음
- 최근 CCTV, 카메라 사용이 증가함에 따라 머신러닝, 딥러닝을 활용한 HAR(Human Action Recognition) 기반의 이상 탐지 연구로 이상행동을 탐지하고 선조치하기 위해 노력중
- C3D 모델로 이상행동 영상을 탐지한 결과, 범죄유형과 전조현상을 구분하는데 상당히 높은 정확도를 보임
- 국내 보안 및 안전 전문 민간회사의 관련 기술 수준을 향상
- 지능형 CCTV의 상용화로 국민의 범죄 두려움을 낮추고 사회 전반의 안전에 대한 신뢰를 높일 것으로 판단

감사합니다