

## ●○ 안전 영역

# 이상행동 CCTV 영상



## ●○ 개요: 이상행동 CCTV 영상 데이터셋이란?

지금은 데이터를 기반으로 AI 기술을 활용하여 새로운 제품과 서비스를 창출하는 경제 시대이다. AI와 데이터는 전체 사업에서의 혁신 성장을 가속화 할 수 있는 중요 요소이다. AI와 데이터는 4차 산업혁명 시대의 중요한 생산요소이며, 미래의 경쟁은 대규모 데이터 확보와 AI 역량 강화가 핵심일 것이다. 경쟁력을 갖추기 위해서는 민·관 협동하여 대규모 데이터를 구축하고 개방하는 것이 필요하다. 미국, 유럽 등 AI 선도국에서는 정부와 민간 협업과 투자로 대규모 데이터를 구축하여 공개하고 있다. 국내 중소·벤처기업들은 AI 학습용 원천 데이터의 확보가 어렵고, 데이터의 자체 구축과 가공에 많은 시간과 비용을 소요하고 있다. 따라서 AI 선도국과 기술 격차를 줄이기 위한 노력이 필요하다. 데이터, AI 분야별 육성전략과 융합촉진을 위한 정책을 통해 다양한 영역에서 데이터와 AI 발전을 도모해야 한다.

그 중 CCTV 데이터는 안면인식, 의료 등과 더불어 규제가 특히 심한 부문이다. 개인정보보호법 및 시행령, 시행규칙과 지방자치단체의 CCTV 설치 및 운영 규정, 공공기관 CCTV 관리 가이드라인 등에 의거하여 CCTV 화상 정보의 수집·처리 관련 행위가 엄격하게 제한된다. 정보·공공 주도 하에 공개 가능한 CCTV 영상 데이터의 구축 및 공개·유통은 민간기업의 AI 기술 발전을 위해 큰 의의를 지닌다.

공개된 해외 CCTV 영상은 총격, 폭발, 상점털이 등의 이상행동이 주를 이뤄 한국적 환경에 적용하기 어려움이 있다. 따라서 실제 지자체 CCTV 환경과 유사하고 한국 생활에 밀접한 데이터가 필요하고, 정부 차원에서 관리하는 5대 중대범죄 중심의 이상행동 정의 및 그에 따른 AI 학습용 CCTV 영상 데이터가 필요하다.

본 데이터셋은 CCTV 또는 보안 카메라로부터 입력되는 동영상에서 이상행동을 탐지하는 지능형 영상 인식 AI의 학습 개발에 활용하기 위한 데이터셋을 구축하고 시범 모델을 개발하여 데이터셋을 검증한 뒤 공개하였으며, 추후 인공지능 CCTV 또는 보안 카메라에 사용되는 인공지능 모델 학습을 위한 최적의 이상행동 CCTV 데이터셋으로 활용되는 것을 목표로 한다.

## ●○ 데이터셋의 구성

본 데이터셋은 데이터 수집, 데이터 정제, 데이터 가공, 데이터셋 검수, 데이터 활용의 5단계를 거쳐 구축하였으며, 12종의 이상행동, 실내/실외, 장소, 시간대, 계절, 날씨, 이상행동을 취한 인원, 성별/연령대를 고려하여 최대한 현실과 근접한 영상 제작을 위해 실제 CCTV 영상을 참고하여 전문 촬영 업체와 협업하여 사람 동작과 이상행동을 재현하여 직접 제작하였다.

또한 구축한 데이터셋은 공개의 목적을 가지고 있으므로, 촬영 대상자에게 반드시 개인정보 활용 동의서 작성 등의 법적 절차를 거쳐 촬영된 데이터를 공개하는데 있어, 법적 의무사항을 준수할 수 있도록 절차와 동의서 등에 대한 법률 자문을 받아 진행하였다.

### | 이상행동 CCTV 영상 데이터셋 구성 |

데이터 종류	포함 내용	제공 방식
이상행동 CCTV 영상	12종의 이상행동이 담긴 제작된 CCTV 영상 (8436 개)	MP4 포맷 파일
정보 XML 파일	이상행동에 대한 레이블링 정보를 포함하고 있는 각 영상과 쌍을 이루는 데이터 (8436 개)	XML 포맷 파일

(주)마인즈랩

(ver. 1.0, 2019.08)

## 개인 식별 데이터 공개에 대한 동 의 서

(이상행동 CCTV 학습용 데이터)

과제명 : 이상행동 CCTV 영상 AI 데이터 구축

1. 나는 본 과제를 통해 얻어진 본인의 데이터를 과제 결과 공유의 목적으로 공개하는 것에 대한 설명을 들었습니다.
2. 나는 이 촬영을 통해 얻어진 본인 식별 데이터를 연구목적으로 공개하는 것에 대하여 자발적으로 동의합니다.
3. 나는 이 촬영에서 얻어진 나에 대한 정보를 현행 법률과 연구윤리심의위원회 규정이 허용하는 범위 내에서 과제 수행기관이 수집하고 처리하는데 동의합니다.
4. 나는 과제 수행기관이나 위임 받은 대리인이 과제를 진행하거나 결과 관리를 하는 경우와 과제 수행 기관에서 실태 조사를 하는 경우에는 비밀로 유지되는 나의 개인 신상 정보를 직접적으로 열람하는 것에 동의합니다.
5. 나는 언제라도 이 촬영의 참여를 철회할 수 있고 이러한 결정이 나에게 어떠한 해도 되지 않을 것이라는 것을 압니다.
6. 나의 서명은 이 동의서의 사본을 받았다는 것을 뜻하며 촬영 참여가 끝날 때까지 사본을 보관하겠습니다.

촬영 대상자      성명:                      서명:                      서명일:

법정대리인      성명:                      서명:                      서명일:  
(필요시)      촬영 대상자와의 관계:

입회인              성명:                      서명:                      서명일:  
(필요시)

## ●○ 데이터셋의 설계 기준과 분포

데이터셋을 설계할 때 최대한 현실적인 데이터셋을 구축하기 위하여, 동영상에서의 동작 인식을 다루는 여러 가지 논문과 오픈 데이터셋을 분석하고 동작의 후보들을 정리, 동영상에서 이상행동 인식을 다루는 여러 가지 논문을 분석하고 이상행동의 후보들을 정리, 동작과 이상행동을 객관적 및 실제로 정의하기 위해, 참여기관인 수원시에서 관리하고 있는 CCTV 관제부서에서 정의하는 이상행동을 정리하여, '동작(Action)' 사람의 움직임 등을 구체적으로 정의하는 것, '이상행동(Abnormal event)'를 여러 가지 동작들이 모여서 이루어지는 행동 가운데 정상이 아닌 행동으로 정의되는 행동으로 정의하였다.

### | 이상행동 리스트 |

No	이상행동	출처	객체	객체수	이상행동 내 동작									
1	폭행	논문, 수원시	사람	1인 이상	휘두르기	차기	밀치기	당기기	던지기	쓰러짐	위협하기	찌르기		
2	싸움	논문, 수원시, KISA	사람	2인 이상	휘두르기	차기	밀치기	당기기	던지기	쓰러짐	위협하기	찌르기		
3	절도	논문, 수원시	사람	1인 이상	줍기	운반하기	당기기							
4	기물파손	논문, 수원시	사람, 기물	1인 이상	휘두르기	차기	밀치기	당기기	던지기					
5	살인	논문, 수원시, KISA	사람	1인 이상	쓰러짐	비틀거림								
6	배회	논문, KISA	사람	1인 이상	맴돌기	가다써다 반복								
7	침입	수원시, KISA	사람	1인 이상	월담	당기기	걸기							
8	투기	수원시	사람, 쓰레기, 가방	1인 이상	던지기	떨어뜨림	운반하기							
9	강도	논문	사람	1인 이상	휘두르기	차기	밀치기	당기기	던지기	위협하기	찌르기	쓰러짐	빼앗기	
10	데이트 폭력 및 추행	수원시	사람	1인 이상	휘두르기	차기	밀치기	당기기	던지기	안기	만지기	위협하기	찌르기	쓰러짐
11	납치	논문, 수원시	사람	1인 이상	휘두르기	차기	밀치기	당기기	안기	위협하기	찌르기	쓰러짐		
12	주취행동	수원시	사람	1인 이상	비틀거림	주저앉음								

## | 데이터셋 구성 개요 |

	실내		실외		합계	
	영상 개수	영상 시간	영상 개수	영상 시간	영상 개수	영상 시간
01. 폭행(Assault)	518	44:37:20	395	33:28:21	913	78:05:41
02. 싸움(Fight)	480	41:51:37	694	58:03:08	1174	99:54:45
03. 절도(Burglary)	270	23:12:18	569	46:21:06	839	69:33:24
04. 기물파손(Vandalism)	54	4:44:41	436	36:44:05	490	41:28:46
05. 실신(Swoon)	432	44:37:20	480	39:48:56	912	84:26:16
06. 배회(Wander)	161	14:10:09	484	41:14:41	645	55:24:50
07. 침입(Trespass)	0	0:00:00	259	22:03:15	259	22:03:15
08. 투기(Dump)	216	18:27:33	407	31:39:38	623	50:07:11
09. 강도(Robbery)	216	18:40:29	148	12:12:26	364	30:52:55
10. 데이트폭력및추행(Datefight)	378	32:28:07	315	25:53:38	693	58:21:45
11. 납치(Kidnap)	104	9:05:34	158	13:18:34	262	22:24:08
12. 주취행동(Drunken)	594	50:31:30	668	53:49:07	1262	104:20:37
총 합계	3423	302:26:38	5013	414:36:55	8436	717:03:33

이상행동과 동작에 대한 주요 특징은 다음과 같다.

- 동작(Action): 사람의 움직임 등을 구체적으로 정의하는 것
- 이상행동(Abnormal event) : 여러가지 동작들이 모여서 이루어지는 행동 가운데 정상이 아닌 행동

구분	내용(종류)
동작(Action)	<ul style="list-style-type: none"> <li>• 이상행동을 구성하는 세부이상동작 20가지               <ul style="list-style-type: none"> <li>- 휘두르기, 차기, 밀치기, 당기기, 던지기, 위협하기, 찌르기, 쓰러짐, 줍기, 운반하기, 비틀거림, 맴돌기, 가다 서다 반복, 월담, 떨어뜨림, 빼앗기, 안기, 만지기, 주저앉음</li> </ul> </li> </ul>
이상행동(Abnormal event)	<ul style="list-style-type: none"> <li>• 이상행동 12가지               <ul style="list-style-type: none"> <li>- 폭행, 싸움, 절도, 기물파손, 실신, 배회, 침입, 투기, 납치, 강도, 데이트폭력 및 추행, 주취행동</li> </ul> </li> </ul>

## ●○ 데이터 구조

데이터셋에 따른 항목과 해당 값은 아래 테이블과 같다.

분류	이상행동 CCTV 영상	레이블링 파일
수량	8436건(약 717시간)	8436건
레이블링항목(정보 XML 파일 항목 설명)		
level1	level2	level3
folder : 상위 folder 명 (event끼리 folder를 묶으므로 event명이 folder명이 됨)		
filename : xml 파일과 매칭되는 원본 영상 파일 이름		
source	database : Database 버전 명	
	annotation : annotation 구조 명	
size	width : xml 파일에 대응되는 영상의 width 값	
	height : xml 파일에 대응되는 영상의 height 값	
	depth : xml 파일에 대응되는 영상의 channels값	
header	duration : 영상 길이	
	fps : 초당 프레임 수	
	frames : 영상의 총 프레임 수	
	location : 영상 속 장소	
	weather : 영상 속 날씨	
	time : 시간대	
	population : (이상행동을 한)인원수	
	character : 성별, 연령대	
event	eventname : 이상행동 명	
	starttime : 이상행동 발생 시각	
	duration : 이상행동 지속 시간	
object : xml파일과 대응되는 영상 속의 object 정보에 대한 tag	objectname : object명	
	position : object의 위치정보, object를 명시하기 위한.	frame : object가 등장한 프레임(위치를 알려주기 위한)
		keypoint : 위 frame에서 object의 위치(가슴/머리 부근)
	action	actionname : action 명
		frame : action이 발생하는 시작 프레임과 끝 프레임을 리스트로 나타냄

●○ 데이터 예시



| 이상행동 CCTV 영상 데이터 예시 |

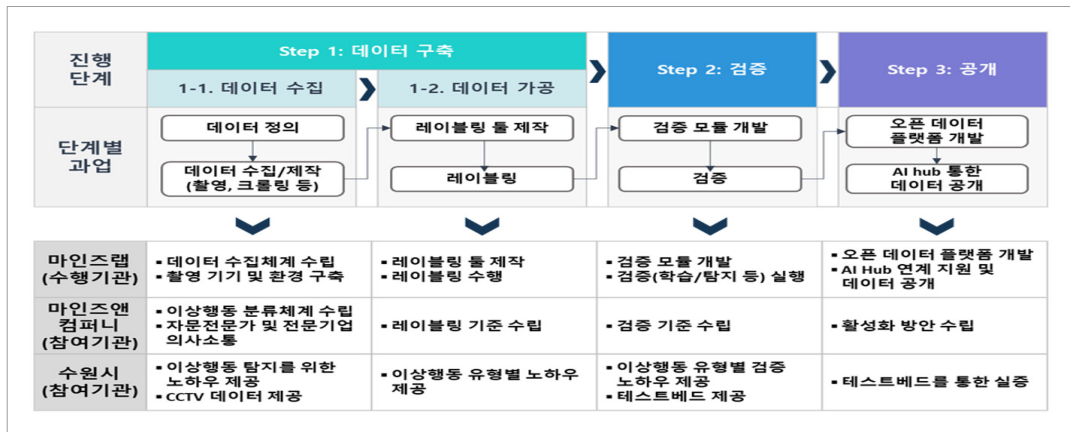
```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<annotation>
  <folder>swoon</folder>
  <filename>100-1_cam01_swoon01_place02_day_summer.mp4</filename>
  <source>
    <database>NIA2019 Database v1</database>
    <annotation>NIA2019</annotation>
  </source>
  <size>
    <width>3840</width>
    <height>2160</height>
    <depth>3</depth>
  </size>
  <header>
    <duration>00:05:14.9</duration>
    <fps>30</fps>
    <frames>9448</frames>
    <inout>OUT</inout>
    <location>PLACE02</location>
    <season>SUMMER</season>
    <weather>SUNNY</weather>
    <time>DAY</time>
    <population>1</population>
    <character>M20</character>
  </header>
  <event>
    <eventname>swoon</eventname>
    <starttime>00:02:30.6</starttime>
    <duration>00:01:26.4</duration>
  </event>
  <object>
    <objectname>person_1</objectname>
    <position>
      <keyframe>4518</keyframe>
      <keypoint>
        <x>1975</x>
        <y>1030</y>
      </keypoint>
    </position>
    <action>
      <actionname>falldown</actionname>
      <frame>
        <start>4518</start>
        <end>4675</end>
      </frame>
      <frame>
        <start>6705</start>
        <end>6790</end>
      </frame>
      <frame>
        <start>6790</start>
        <end>7107</end>
      </frame>
    </action>
  </object>
</annotation>
```

| XML 포맷 파일 예시 |



## ●○ 데이터 구축 과정

데이터 구축은 이상행동 CCTV 영상 데이터 설계 및 분류, 데이터 수집 및 정제/가공, 이상행동 검출 모델 개발 및 실증 테스트베드 구축, 테스트베드를 통한 데이터셋 검증 과정을 통해 이루어졌다.



| 데이터 수집/정제/가공/검증/활용 단계 |

LB

장소/시간/날씨

실내/실외

실외

장소

유흥가톨릭

시간대

밤

계절

여름

날씨

맑음

인원/성별/연령대

인원수

3

성별/연령대

M20,M20,M20

이상행동

이상행동

폭행

작업 파일명: 16-3\_cam01\_assault01\_place02\_night\_summer.mp4

최종작업자: checker17

작업완료: YES

검수완료: YES

00:03:37.5 / 00:05:00.3

- 0.1

Play

+ 0.1

이상행동 (Event) 시작지점 저장

00:03:23.5

6105

이상행동 (Event) 종료지점 저장

00:04:46.8

8604

이상행동 (Event) 구간보기

이상동작(Action) 오브젝트

오브젝트 추가

ObjectName	KeyTime	KeyPoints X	KeyPoints Y	Action Buttons
person_3	6303	2194	999	키 지정, 확인, 오브젝트 삭제, 액션 추가
ActionName	쓰러짐			액션 삭제, 프레임 추가
액션 시작 등록	6303	액션 끝 등록	6359	액션 삭제, 프레임 시작으로, 프레임 끝으로, 프레임 시작으로, 프레임 삭제
액션 시작 등록	6793	액션 끝 등록	6859	

| 효율적인 데이터 정제/가공을 위한 웹 기반 툴 |

사람마다 한가지 동작을 다양하게 인식할 수 있기에 일관적인 레이블링 데이터를 확보하기 위하여 레이블링 작업자를 위한 가이드라인을 작성하여, 웹 기반의 레이블링 툴을 사용하여 데이터셋을 구축.

### 3. 이상행동 및 세부이상행동 xml 태깅 가이드

① 이상행동을 하는 사람(주체)의 액션만 태깅합니다.

② 객체(주체에 의해 당하는 사람)를 태깅하는 이상행동(event)은 폭행, 데이트폭행 및 추행, 납치, 강도 총 4가지이며 쓰러짐의 액션만 태깅합니다.

③ 명확한 액션이며, 이상행동(event)에 해당되는 세부이상행동(action)일 경우에만 태깅합니다.

(예 : 주취행동에서 쓰러지는 행동을 '주저앉음'으로 태깅하지 않습니다.)

④ 이상행동의 시작지점과 끝지점을 동일하게 맞춥니다.

예) 휘두르기 : 이상행동 시작(주먹을 올리기 전 차렷자세)→주먹 휘두르기→이상행동 끝(주먹을 휘두른 후 다시 차렷자세)

- 같은 동작 반복하는 사이에 다른 동작이 들머갈 경우 (예 : 주먹질 + 차기 + 주먹질)로 태깅합니다.

- 같은 동작을 반복하는 사이에 툴에 적혀있지 않는 행동이 들머갈 경우 끊어서 태깅합니다.

⑤ 오브젝트가 같은 동작을 짧은 시간 안에 연속적으로 반복할 경우 1번만 태깅합니다.

- 단, 휘두르는 여러 번의 액션 사이에 틈이 생길 때, 액션의 시간보다 틈이 길면 끊어서 태깅합니다.

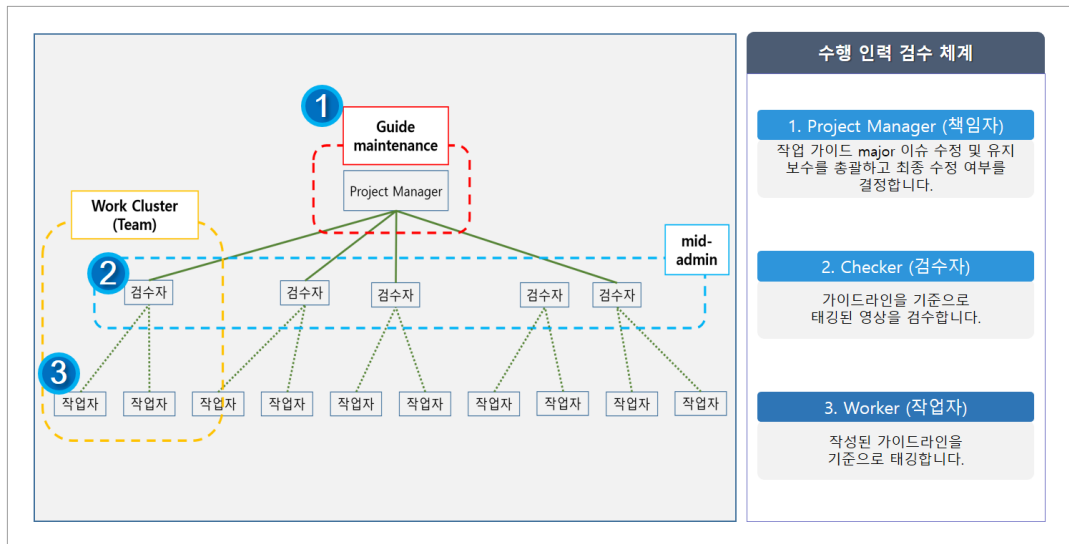
ex) 휘두르기(3초) - 휴식(3초) - 휘두르기(3초) 일 경우 휘두르기를 한 번으로 태깅합니다.

ex) 휘두르기(3초) - 휴식(5초) - 휘두르기(3초) 일 경우, 중간에 쉬는 부분(5초)을 끊고 두 번 태깅합니다.

### | 레이블링 작업자 가이드라인 예시 |

## ●○ 검수와 품질 확보

대량의 데이터를 높은 품질로 생성하기 위하여 검수 프로세스의 정립은 데이터셋 구축에 매우 중요한 의미를 갖는다. 이 데이터셋에서는 2단계 검수 체계를 구축했는데, 가장 하위 레벨에는 작업자들이 작업한 결과물을 가이드라인에서 제시한 형식에 맞는지 확인하는 검수자가 있었고, 이들이 검수한 결과물에 대해서 재검수를 하고 데이터셋을 전체적으로 들여다보며 데이터셋의 밸런스나 가이드라인의 적절성을 제시해주는 관리자는 마인즈랩의 직원으로 학습 경험이 풍부한 인력을 배치하여 최종적인 데이터셋의 품질을 담보할 수 있었다.



| 고품질 데이터셋 구축을 위한 품질 검수 체계 |

## ●○ 데이터 구축 담당자

수행기관(주관) : (주)마인즈랩

(전화: 031-625-4349, 이메일: pworks@mindsrab.ai)