

# 이상행동 인식 영상 데이터셋 EDA 보고서

데이터 출처: AI Hub - 사람 이상행동 인식 영상 데이터셋

프레임: 3fps | 길이: 60초 (180프레임) | 형식: MP4 + XML

## 1. 데이터셋 개요

### 1.1 데이터셋 목적

해당 데이터셋은 실제 편의점/매장 내에서 발생하는 **\*\*이상행동(전도, 파손, 흡연, 유기, 절도, 폭행 등)\*\***을 인식하고, 관련 인공지능 모델 학습을 위한 원천 및 라벨링 데이터를 포함

AI Hub에서 제공하는 사람 이상행동 인식 영상 데이터셋을 기반으로 하며, 학습용(Training)과 검증용(Validation)으로 구성

### 1.2 폴더 및 파일 구성

- 데이터셋 전체 구조

```
[ROOT]
├── 238-2.실내(편의점, 매장) 사람 이상행동 데이터/
├── 01-1.정식개방데이터/
│   ├── Training/
│   │   ├── 01.원천데이터/
│   │   │   ├── TS_03.이상행동_07.전도/
│   │   │   ├── TS_03.이상행동_08.파손/
│   │   │   └── ... (이상행동별 분류)
│   │   ├── 02.라벨링데이터/
│   │   │   ├── TL_03.이상행동_07.전도/
│   │   │   ├── TL_03.이상행동_08.파손/
│   │   │   └── ... (이상행동별 XML)
│   │   └── Validation/
│   │       ├── 01.원천데이터/
│   │       │   ├── VS_03.이상행동_07.전도/
│   │       │   ├── VS_03.이상행동_08.파손/
│   │       │   └── ...
│   │       ├── 02.라벨링데이터/
│   │       │   ├── VL_03.이상행동_07.전도/
│   │       │   ├── VL_03.이상행동_08.파손/
│   │       │   └── ...
```

- TS\_ , VS\_ : Training / Validation의 **Source (원천 영상)**

- **TL**, **VL**: Training / Validation의 **Label** (라벨링 데이터)

### 1.3 원천 데이터 특징

항목	내용
영상 길이	1분 (60초)
프레임 수	180프레임 (3fps)
해상도	RGB, 정규화된 해상도
포맷	<b>.mp4</b> (ZIP으로 압축되어 제공)
촬영 환경	실제 편의점/매장 내 CCTV 기반
행동 유형	전도, 파손, 흡연, 유기, 절도, 폭행, 방화, 교통약자 (이상행동 8종)

\*\* 행동 유형 중 <방화>는 데이터 품질의 문제때문에 <교통 약자> 는 주제와 맞지 않는다고 판단하여 프로젝트에 사용할 데이터셋에서 제외하였기 때문에 최종적으로는 그 둘을 제외한 총 6종의 이상 행동이 행동 유형에 포함되었음.

### 1.3 라벨링 데이터 (XML) 구조 및 정보

- 주요 정보

항목	설명
<b>VideoName</b>	라벨이 적용된 영상 파일명
<b>BehaviorType</b>	이상행동 종류 (예: 전도, 절도 등)
<b>StartFrame</b> / <b>EndFrame</b>	행동이 시작되고 끝나는 프레임 번호 (0부터 시작)
<b>PersonID</b>	이상행동을 수행한 인물 식별 ID
<b>Skeleton</b>	이상행동을 수행한 인물의 2D 관절 좌표 정보

- XML 예시 구조

```

...
<track id="0" label="smoke_start" source="manual">
  <box frame="22" outside="0" occluded="0" keyframe="1" xtl="999.11" ytl="88.67" xbr="1"
    <attribute name="age group">young people</attribute>
    <attribute name="gender">male</attribute>
    <attribute name="ID">1</attribute>
  </box>
  <box frame="23" outside="1" occluded="0" keyframe="1" xtl="999.11" ytl="88.67" xbr="11
    <attribute name="age group">young people</attribute>
    <attribute name="gender">male</attribute>
    <attribute name="ID">1</attribute>
  </box>
</track>
<track id="1" label="smoke_end" source="manual">
  <box frame="168" outside="0" occluded="0" keyframe="1" xtl="943.01" ytl="43.12" xbr="
  <box frame="169" outside="1" occluded="0" keyframe="1" xtl="943.01" ytl="43.12" xbr="1

```

</track>

...

\*\* 이 데이터셋은 프레임 기반 이상행동 구간 라벨링과 함께 인물 식별 ID와 스켈레톤(관절) 정보까지 포함하고 있지만, 프로젝트를 위한 모델 학습 시 위와 같이 label 이 (이상행동)\_start 인 프레임부터 (이상행동)\_end 인 프레임 정보만 사용할 것.

## 2. 데이터 분포 분석

### 2.1 이상행동 유형별 구성

데이터셋에는 총 **6가지 이상행동 유형**이 존재하며, 각 유형은 다음과 같이 정의

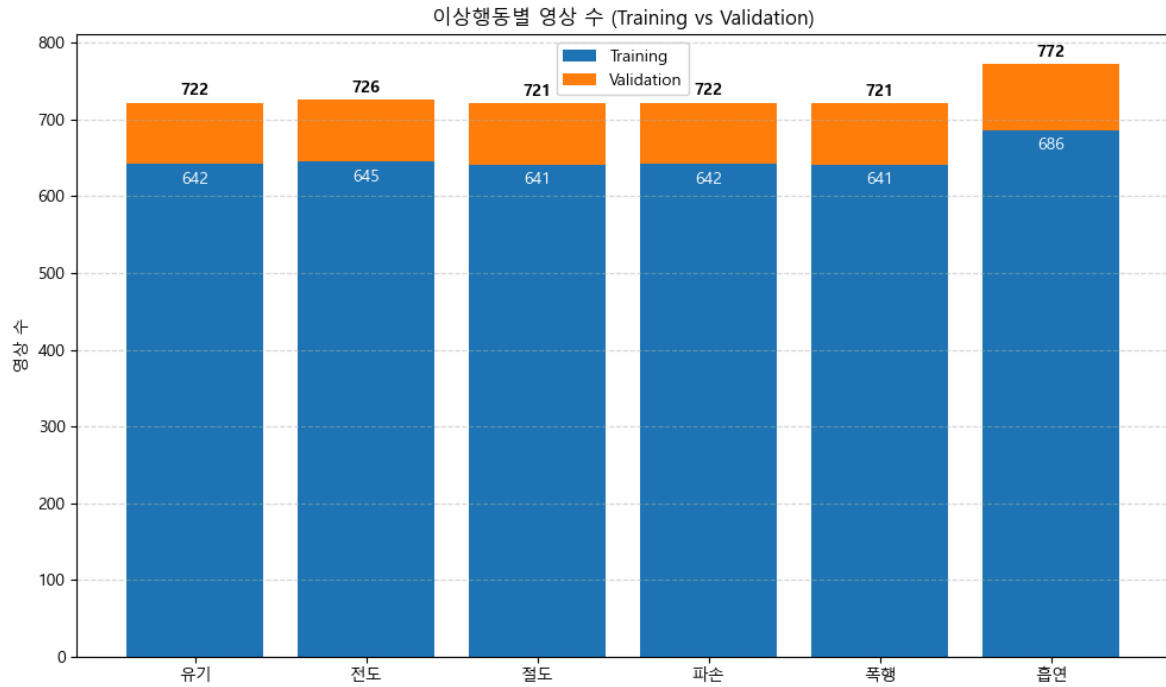
이상행동 유형	설명 예시
전도	사람이 넘어진 상황
파손	물건을 부순 행위
흡연	실내에서의 흡연 장면
유기	유아, 동물 등의 방치
절도	상품을 훔치는 행위
폭행	사람 간 물리적 폭력

각 행동은 원천데이터와 라벨링데이터로 나뉘며, 각각 Training과 Validation 용도로 분리

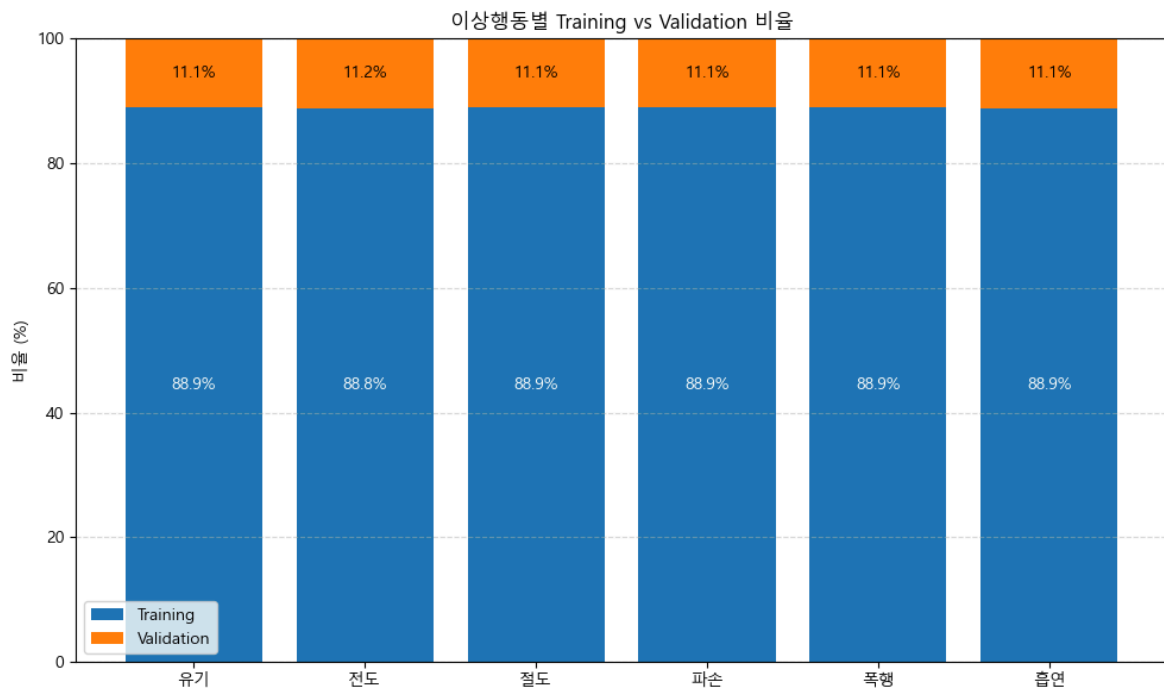
### 2.2 이상행동별 영상 수량 분석

이상행동	Training (원천)	Training (라벨)	Validation (원천)	Validation (라벨)	총합
전도	645	645	81	81	1,452
파손	642	642	80	80	1,444
흡연	686	686	86	86	1,544
유기	642	642	80	80	1,444
절도	641	641	80	80	1,442
폭행	641	641	80	80	1,442
합계	3,897	3,897	487	487	8,768

### 2.3 데이터 분포 시각화



## 2.4 Training vs Validation 비율



## 2.5 데이터 불균형 여부

이상행동 유형 간 영상 수 차이가 크지 않아, 전체적으로 **균형 잡힌 분포**를 이루고 있음

# 3. 영상 메타데이터 분석

## 3.1 해상도, FPS, 길이

```

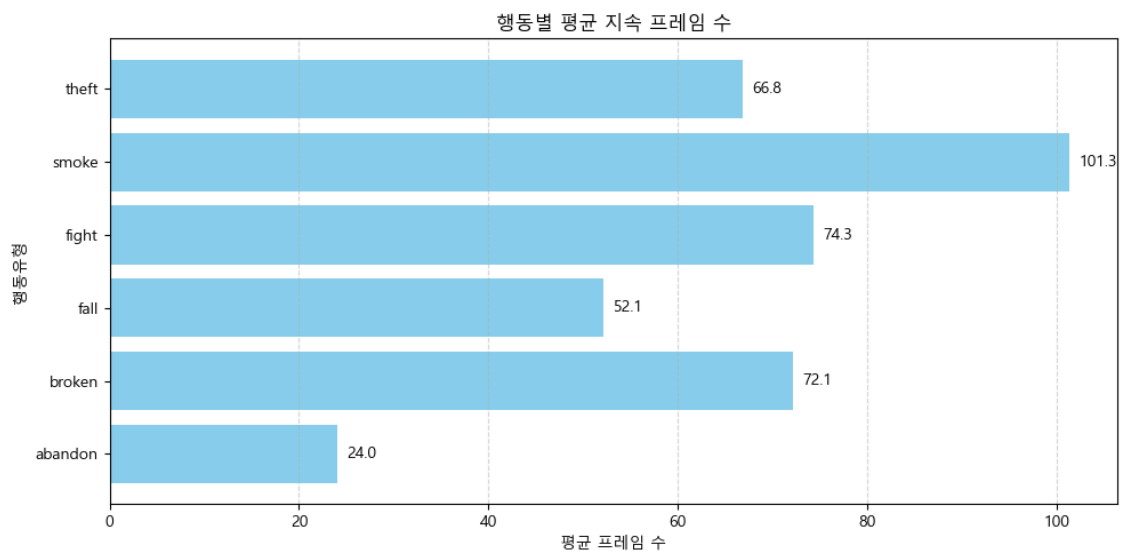
영상 정보 수집 중...
✓ 평균 해상도: 1920.0 x 1080.0
✓ 평균 FPS: 3.0
✓ 프레임 수 분포:
  프레임수
180      3766
181       618
Name: count, dtype: int64

```

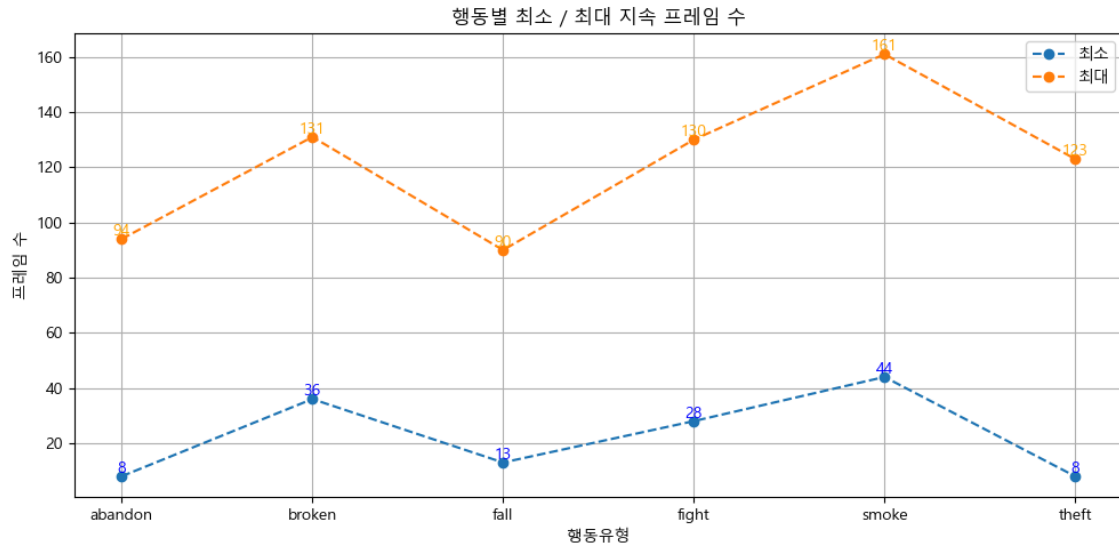
## 4. 라벨링 데이터(XML) 분석

### 4.1 이상행동 지속 시간 분석

- 행동별 평균 지속 프레임 수



- 행동별 최소/최대 지속 시간



- 전체 프레임 중 이상행동이 차지하는 비율



## 7. 종합 특이점 요약

### 7.1 데이터 불균형 여부

- 각 이상행동 유형별 영상 수는 거의 균일하게 분포되어 있음  
→ Training과 Validation 모두 640~680개 사이로 분포 (예: 유기 642개, 절도 641개 등).
- 하지만 실제 라벨링된 **이상행동 지속 프레임 수** 기준으로는 편차 존재  
→ 예: smoke(흡연)은 평균 101프레임, abandon(유기)은 평균 24프레임으로 **약 4배 차이** 발생

#### 🔍 결론:

- 샘플 수 기준 데이터는 균형 있음
- 이상행동 지속시간 기준으로는 편차가 존재 → 일부 행동(예: 흡연, 절도)은 비교적 오랜 시간 지속됨

- 이상행동 지속시간 편차
- 행동 발생 위치 편향성

## 7.2 이상행동 지속시간 편차

- **최소 지속 프레임:** 8프레임 (abandon, theft 등)
- **최대 지속 프레임:** 161프레임 (smoke)  
→편차 매우 큼 (8 ~ 161 프레임)
- 특히 smoke는 지속시간 평균, 최대값 모두 가장 높은 수치를 기록

### 결론:

- 이상행동의 지속시간은 **행동 유형에 따라 매우 상이함**
- 따라서 학습 시 일정 프레임 수로 clip을 자를 경우, **pad 또는 truncate 전략 고려 필요**