

메타데이터 정보 (다중기입가능)	분야	데이터 유형 ¹⁾	구축 데이터량	원천데이터 형식 ²⁾	라벨링 형식 ³⁾	라벨링 유형 ⁴⁾
	영상이미지	비디오	330시간	mp4	json	바운딩박스, 세그멘테이션
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관(총괄)	가공기관	검수기관	
	자체 수집	2021년	(주)디아이솔루션	(주)일주지앤에스	(주)디아이솔루션	
	데이터 문의처	기관명	문의담당자명	전화번호 (유선전화번호기입)	메일주소	
		(주)디아이솔루션	이광수	051-717-2503	leegs@di-solution.co.kr	
	데이터 소개	총 7곳의 상권으로부터 CCTV 화각을 고려하여 수집된 데이터셋으로, 촬영 시간 및 촬영 장소의 다양성을 추구하며 비식별화를 통해 개인정보 문제를 해결한 학습용 데이터셋 구축				
	주요키워드	유동인구, 인공지능, 학습 데이터셋, Person Attribute Recognition				
카테고리 정의서		첨부의 카테고리 정의서 엑셀파일에 데이터카테고리 작성하여 제출(예시참고)				

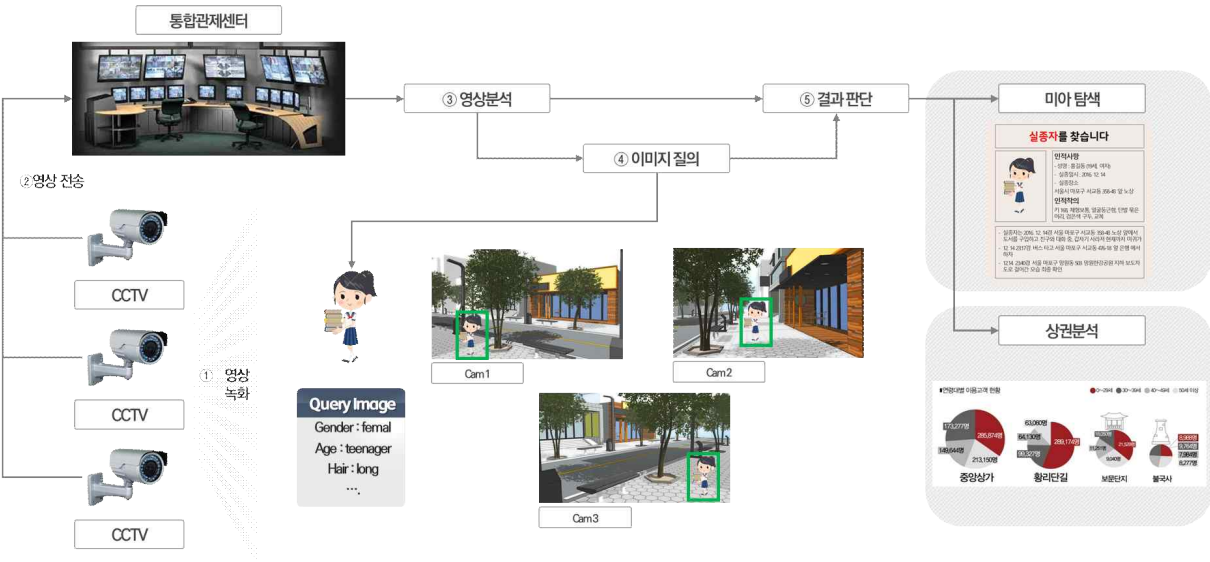
1) 텍스트, 오디오, 이미지, 비디오,

2) txt, jpg,.....

3) json, csv,.....

4) 내용요약(텍스트), 번역(자연어), 질의응답(자연어), 바운딩박스(이미지/동영상), 키포인트(이미지/동영상), 세그멘테이션(이미지/동영상), 전자(음성)

5) 4대 언론기사, 자체 수집,.....

데이터셋명	국문영문	유동인구 분석을 위한 CCTV 영상 데이터
		CCTV video data for foot traffic analysis)
구축목적	상권으로부터 유동인구의 특징 분석 인공지능 학습을 위한 데이터셋	
활용서비스	특정 인상착의 인물 또는 그룹 탐색 서비스, 유동인구 분석을 통한 버스 노선 등의 도시개혁 정책 개정에 활용, 기존의 값비싼 통신사 유동인구 데이터 대체 등에 활용 가능 	
소개	총 7곳의 상권으로부터 CCTV 화각을 고려하여 수집된 데이터셋으로, 촬영 시간 및 촬영 장소의 다양성을 추구하며 비식별화를 통해 개인정보 문제를 해결한 학습용 데이터셋 구축 지역 상권의 유동인구 및 상권특성 분석을 위한 CCTV 영상 데이터 필요성 코로나로 인해 변화하는 상권 상권분석의 중요성 대두 실제 데이터 활용성의 한계 비싼 데이터 가격 데이터셋 구축 상권의 CCTV 정보 상권내 점포의 정보 상권내 유동인구의 정보 데이터 수집 데이터 정제 데이터 가공 데이터 검수 데이터셋 공개 서비스 활성화 트렌드 분석 서비스 상권 분석 서비스 유동인구 분석 서비스	

데이터셋 통계
(구축 규모 및
분포)

1. 데이터 구축 규모

- 총 영상 수 : 6,600개 (330시간 분량)
- 영상 정보 : 3fps, FHD(1920x1080) 해상도의 비식별화 처리된 3분 클립

2. 데이터 분포

구분	분류	항목	시간	비율	구분	분류	항목	시간	비율
1	장소	자갈치시장	49.5H	15%	1	시간대	새벽(0~9시)	65.1H	19.7%
2		서면영광도서 앞	33H	10%	2		오전(9~12시)	81.8H	24.8%
3		사직구장 주변 상권	66H	20%	3		오후(12~16시)	55.4H	16.8%
4		중구 중앙동 주변	66H	20%	4		저녁(16~20시)	65.2H	19.8%
5		동래역 근처	16.5H	5%	5		야간(20~24시)	62.5H	18.9%
6		사상 시외버스 터미널	66H	20%					
7		연산동 막걸리 골목	33H	10%					

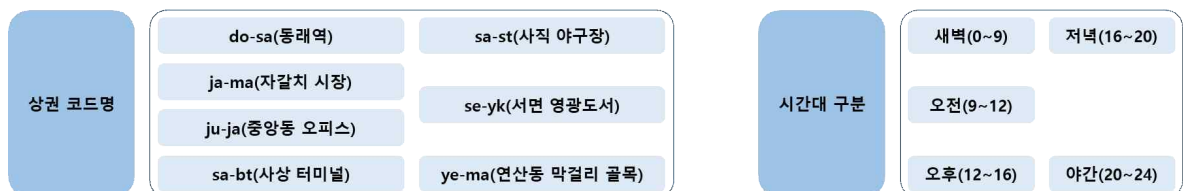
구분	분류	항목	수량	비율
1	성별 분포	남자	53,723	63%
2		여자	31,524	37%
3	연령대 분포	미취학아동	1,015	1%
4		청소년	1,393	2%
5		청년	34,343	40%
6		중장년	48,496	57%
7	상의 유형 분포	상의_긴팔	1,826,233	28%
8		상의_반팔	4,434,147	68%
9		상의_민소매	50,187	1%
10		상의_원피스	209,241	3%
11	상의 색상 분포	상의_빨간색	242,019	4%
12		상의_주황색	86,841	1%
13		상의_노란색	191,406	3%
14		상의_초록색	215,466	3%
15		상의_파란색	1,016,220	16%
16		상의_보라색	84,739	1%
17		상의_분홍색	205,707	3%
18		상의_갈색	204,656	3%
19		상의_흰색	1,157,862	18%
20		상의_회색	916,885	14%
21		상의_검은색	2,198,007	34%
22	하의 유형 분포	하의_긴바지	5,017,387	77%
23		하의_반바지	1,165,135	18%
24		하의_치마	127,446	2%
25		하의_해당없음	209,840	3%
26		하의_빨간색	48,573	0.7%
27	하의 색상 분포	하의_주황색	14,395	0.2%
28		하의_노란색	57,746	0.9%
29		하의_초록색	73,911	1.1%
30		하의_파란색	798,183	12.2%
31		하의_보라색	15,029	0.2%
32		하의_분홍색	15,499	0.2%
33		하의_갈색	206,173	3.2%
34		하의_흰색	274,255	4.2%
35		하의_회색	938,482	14.4%
36		하의_검은색	3,868,460	59.3%
37		하의_해당 없음	209,102	3.2%
38	소지품 유형 분포	캐리어	5,205	0.1%
39		우산	303,552	4.7%
40		가방	1,081,869	16.6%
41		모자	813,428	12.5%
42		안경	119,146	1.8%
43	애완동물 동행 여부 분포	해당 없음	4,196,608	64.4%
44		동행하지 않음	6,500,015	99.7%
45		동행	19,793	0.3%

1. 대표도면



```
{
  "info": {
    "year": "2021",
    "version": "1.1",
    "date_created": "2021-08-02 08:36:00",
    "day": "mon",
    "weather": "sunny",
    "username": "iljoo"
  },
  "video": {
    "file_name": "2021-08-02_08-36-00_mon_sunny_out_ju-ja_C0065.mp4",
    "resolution": [
      1920,
      1080,
      3
    ],
    "fps": 3,
    "total_frame": 540,
    "play_time": "00:03:00",
    "in_out": "out",
    "location": "junggu-jungangdong-around",
    "location_type": "office",
    "total_person": 6
  },
  "annotations": [
    {
      "id": "person_0",
      "frame": 27,
      "bbox": [
        1335.197265625,
        884.59765625,
        1582.0544128417969,
        1078.996046875
      ],
      "direction": "NW",
      "top_type": "short_sleeve",
      "top_color": "white",
      "bottom_type": "long_pants",
      "bottom_color": "black",
      "accessories": "none",
      "pet": 0
    },
    {
      "id": "person_0",
      "frame": 28,
      "bbox": [
        1320.84765625,
        841.75,
        1519.0,
        1080.0
      ],
      "direction": "NW",
      "top_type": "short_sleeve",
      "top_color": "white",
      "bottom_type": "long_pants",
      "bottom_color": "black",
      "accessories": "none",
      "pet": 0
    }
  ]
}
```

2. 라벨링데이터 구성



3 라벨링데이터 실제예시

분류	속성명	속성 설명	데이터 타입	필수 여부	예시
info (CCTV)	year	촬영연도	int	O	2021
	version	버전정보	String	O	1.1
	date_created	촬영일자	String	O	2021/01/01 15:00:00
	day	촬영요일	String	O	fri
	weather	날씨	String	X	sunny
	username	사용자 ID	String	O	admin
video (CCTV)	file_name	파일명	String	O	sample.mp4
	resolution	비디오 해상도 및 컬러	Array[int]	O	[1920, 1080, 3]
	fps	촬영 프레임	int	O	3

		total_frame		영상 총프레임	int	O	540
		play_time		영상 시간	String	O	00:03:00
		in_out		실내 · 외	String	O	out
		location		촬영장소	String	O	seomyeon_yk_fro nt
		location_type		촬영장소 상권 유형	String	O	downtown
		total_person		영상내 유동인구수	int	O	1
	annotati ons	id		객체 ID	String	O	person_1
		frame_no		현재 프레임	int	O	25
		bbox		바운딩 박스	Array[int]	O	[100, 200, 100, 200]
		direction		이동방향	String	O	N
		top_type		상의 유형	String	O	long_sleeve
		top_color		상의 색상	String	O	white
		bottom_type		하의 유형	String	O	long_pants
		bottom_color		하의 색상	String	O	blue
		accessories		소지품	String	O	backpack
		pet		애완동물 동행 여부	int	O	0
	objects	actor	actor_id	연기자 ID	String	X	actor_0
		id		객체 ID	String	X	car_1
		frame_no		현재 프레임	int	X	25
		bbox		바운딩 박스	Array[int]	X	[100, 200, 100, 200]
		direction		이동방향	String	X	N
		state		상태	String	X	parking
	events	id		객체 ID	String	X	person_1
		start_frame		이벤트 시작 프레임	int	X	45
		end_frame		이벤트 끝 프레임	int	X	70
		action		액션 유형	String	X	store_in
		action_position		액션 위치	Array[int]	X	store_2
	categori es	person	supercategory	객체 유형	String	O	person
			id	객체 ID	String	O	person_1
			gender	성별	String	O	male
			age	연령대	String	O	adult
			cam_in	객체 입장 프레임	Array[int]	O	[15]
			cam_in_direc tion	객체 입장 방향	Array[String]	O	[N]
			cam_out	객체 퇴장 프레임	Array[int]	O	[500]
			cam_out_direc tion	객체 퇴장 방향	Array[String]	O	[E]
	backgro und	store	supercategory	객체 유형	String	O	store
			id	객체 ID	String	O	store_2
			empty	공실 여부	int	X	0
			category	객체 유형	String	X	restaurant
			segmentation	객체 위치좌표	Array[int][int]	X	[100,200], [200,300], [300,400], [400,500], [500,600]
			parking	주차장 유무	int	X	1
			toilet	화장실 유무	int	X	0
			floor	층수	int	X	2
			elevator	엘리베이터 유무	int	X	1
		car	supercategory	객체 유형	String	O	car
			id	객체 ID	String	O	car_3

			cam_in	객체 입장 프레임	Array[int]	X	[20]
			cam_in_direction	객체 입장 방향	Array[String]	X	[N]
			cam_out	객체 퇴장 프레임	Array[int]	X	[90]
			cam_out_direction	객체 퇴장 방향	Array[String]	X	[E]
데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)	메일주소	담당업무	
		(주)디아이솔루션	이광수	051-717-2503	leegs@di-solution.co.kr	총괄 책임	
	참여기관	기관명	담당업무	기관명	담당업무		
		(주)일주지앤에스	데이터 가공				
		(주)나래솔루션	데이터 수집				