

메타데이터 정보 (다중기입가능)	분야	데이터 유형 ¹⁾	구축데이터량	원천데이터 형식 ²⁾	라벨링 형식 ³⁾	라벨링 유형 ⁴⁾
	재난안전환경	이미지	5만	tif	json	폴리곤 세그멘테이션
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관 (총괄)	가공기관	검수기관	
	서울시 제공	2021년	서울디지털재단	엘티메트릭(주), 인크루트	공간정보산업진흥원 (주)다비오	
	데이터 문의처	기관명	담당자명	전화번호 (유선전화번호기입)	메일주소	담당분야
		서울디지털재단 (주관기관)	박지혜	02-570-4644	jhpark@sdf.seoul.kr	과제실무책임
		다비오	임해웅	070-7777-2489	bryanrhim@dbeco.com	AI 모델학습
		인크루트	유지애	02-2186-9096	yuja@incruit.com	저작도구 /라벨링데이터
		엘티메트릭	이상준	02-855-5724	ls79j@ltmetric.com	원천데이터 제작
		공간정보산업진흥원	홍송표	070-4139-1553	sp.hong@spacen.or.kr	품질관리
	데이터 소개	도시에서의 건축물 변화를 탐지하는 AI 기술 개발을 위한 항공영상 기반의 변화 건물 이미지 데이터				
	주요키워드	건물 변화, 변화탐지, 항공사진, U-Net, 변화탐지모델				
	카테고리 정의서		별첨			

1) 텍스트, 오디오, 이미지, 비디오,

2) txt, jpg,.....

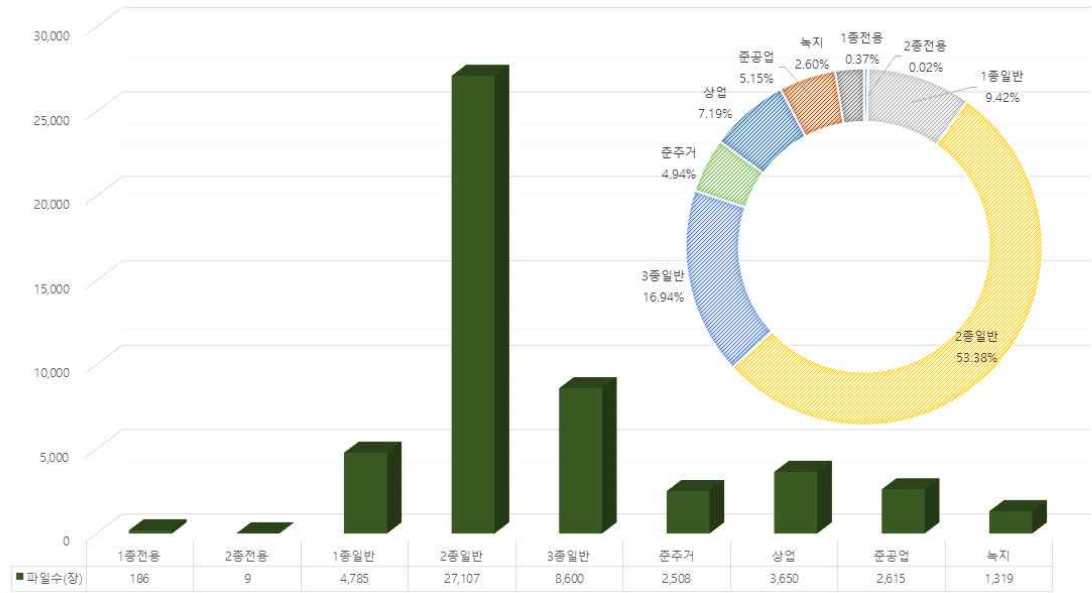
3) json, csv,.....

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

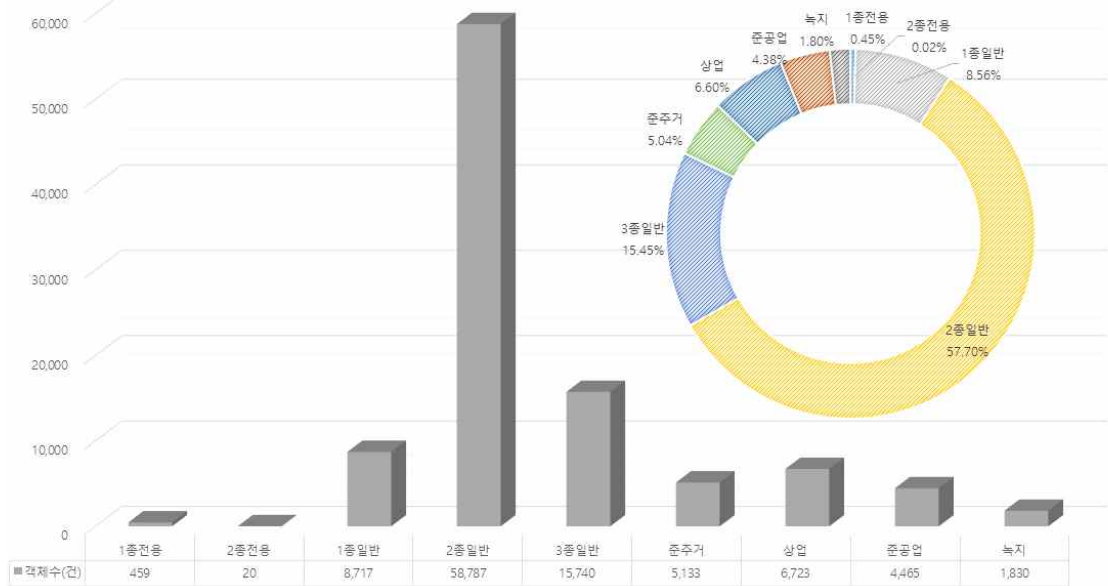
5) 4대 언론기사, 자체 수집,.....

데이터셋명	국문영문	항공영상 도시건물 변화탐지 데이터셋																																																																						
	문	City Building Change Detection using Aerial imagery Dataset																																																																						
구축목적	도시 안전을 위협하는 불법건축물의 적발과정이 단축되도록 기존 사람의 육안식별 대신 인공지능이 변화한 건물을 찾아내는 알고리즘 개발에 필요한 학습용 데이터 구축																																																																							
활용서비스	건물변화 모니터링 서비스																																																																							
소개	<p>서울시 항공사진을 정제·가공하여 전후비교 형태의 이미지로 제작하고 과거대비 변화가 있는 건물을 식별·라벨링한 데이터셋으로, 기존 객체인식용 데이터셋과 달리 건물의 변화를 탐지해내는 인공지능 알고리즘 개발을 위해 제작되었음</p> <p><항공영상 도시건물 변화탐지 데이터셋 구축 프로세스></p>																																																																							
	<p>1. 데이터 구축 규모</p> <table> <tr> <th>구분</th><th>내용</th><th>파일포맷</th><th>구축량(장)</th></tr> <tr> <td>원천데이터</td><td>항공사진을 정제하여 건축물의 변화여부 판단이 가능한 비교쌍 형태(N년,N+1년)의 이미지</td><td>.tif</td><td>50,779</td></tr> <tr> <td>라벨링데이터</td><td>원천데이터내 변화한 건물을 식별하여 폴리곤 세그멘테이션한 정보(변화유형, 갱신유형 등)</td><td>.json</td><td>50,779</td></tr> </table> <p>2. 데이터 분포</p> <table> <tr> <th rowspan="2">용도지역 구분 (하위폴더 단위)</th><th colspan="2">구축파일</th><th colspan="2">변화건물 객체</th></tr> <tr> <th>개수(장)</th><th>비중(%)</th><th>개수(건)</th><th>비중(%)</th></tr> <tr> <td>제1종전용주거지역</td><td>186</td><td>0.37%</td><td>459</td><td>0.45%</td></tr> <tr> <td>제2종전용주거지역</td><td>9</td><td>0.02%</td><td>20</td><td>0.02%</td></tr> <tr> <td>제1종일반주거지역</td><td>4,785</td><td>9.42%</td><td>8,717</td><td>8.56%</td></tr> <tr> <td>제2종일반주거지역</td><td>27,107</td><td>53.38%</td><td>58,787</td><td>57.70%</td></tr> <tr> <td>제3종일반주거지역</td><td>8,600</td><td>16.94%</td><td>15,740</td><td>15.45%</td></tr> <tr> <td>준주거지역</td><td>2,508</td><td>4.94%</td><td>5,133</td><td>5.04%</td></tr> <tr> <td>상업지역</td><td>3,650</td><td>7.19%</td><td>6,723</td><td>6.60%</td></tr> <tr> <td>준공업지역</td><td>2,615</td><td>5.15%</td><td>4,465</td><td>4.38%</td></tr> <tr> <td>녹지지역</td><td>1,319</td><td>2.60%</td><td>1,830</td><td>1.80%</td></tr> <tr> <td>합 계</td><td>50,779</td><td>100.00%</td><td>101,874</td><td>100.00%</td></tr> </table> <p>※ 구축 이미지 1장 당 평균 2개의 변화건물 객체 포함</p>		구분	내용	파일포맷	구축량(장)	원천데이터	항공사진을 정제하여 건축물의 변화여부 판단이 가능한 비교쌍 형태(N년,N+1년)의 이미지	.tif	50,779	라벨링데이터	원천데이터내 변화한 건물을 식별하여 폴리곤 세그멘테이션한 정보(변화유형, 갱신유형 등)	.json	50,779	용도지역 구분 (하위폴더 단위)	구축파일		변화건물 객체		개수(장)	비중(%)	개수(건)	비중(%)	제1종전용주거지역	186	0.37%	459	0.45%	제2종전용주거지역	9	0.02%	20	0.02%	제1종일반주거지역	4,785	9.42%	8,717	8.56%	제2종일반주거지역	27,107	53.38%	58,787	57.70%	제3종일반주거지역	8,600	16.94%	15,740	15.45%	준주거지역	2,508	4.94%	5,133	5.04%	상업지역	3,650	7.19%	6,723	6.60%	준공업지역	2,615	5.15%	4,465	4.38%	녹지지역	1,319	2.60%	1,830	1.80%	합 계	50,779	100.00%	101,874
구분	내용	파일포맷	구축량(장)																																																																					
원천데이터	항공사진을 정제하여 건축물의 변화여부 판단이 가능한 비교쌍 형태(N년,N+1년)의 이미지	.tif	50,779																																																																					
라벨링데이터	원천데이터내 변화한 건물을 식별하여 폴리곤 세그멘테이션한 정보(변화유형, 갱신유형 등)	.json	50,779																																																																					
용도지역 구분 (하위폴더 단위)	구축파일		변화건물 객체																																																																					
	개수(장)	비중(%)	개수(건)	비중(%)																																																																				
제1종전용주거지역	186	0.37%	459	0.45%																																																																				
제2종전용주거지역	9	0.02%	20	0.02%																																																																				
제1종일반주거지역	4,785	9.42%	8,717	8.56%																																																																				
제2종일반주거지역	27,107	53.38%	58,787	57.70%																																																																				
제3종일반주거지역	8,600	16.94%	15,740	15.45%																																																																				
준주거지역	2,508	4.94%	5,133	5.04%																																																																				
상업지역	3,650	7.19%	6,723	6.60%																																																																				
준공업지역	2,615	5.15%	4,465	4.38%																																																																				
녹지지역	1,319	2.60%	1,830	1.80%																																																																				
합 계	50,779	100.00%	101,874	100.00%																																																																				

<용도지역별 구축파일 개수 및 비중>



<용도지역별 변화건물 객체 개수 및 비중>



1. 대표도면

데이터셋 구성

원천 데이터	라벨링 데이터
<p>※ 예시로 어노테이션된 폴리곤을 빨간선 실선 표시 (실제 원천데이터(이미지) 상 표시되지 않으며, 저작도구(뷰어)로 원천/라벨링 결과를 결합하여 보면 위와 같이 표시됨)</p>	<pre> { "info": { "url": "http://www.khri.go.kr/...", "version": "1.0", "description": "KHRI Building Data", "copyright": "KHRI" }, "images": { "url": "http://www.khri.go.kr/...", "width": 60, "height": 60 }, "annotations": { "polygons": [{ "id": 1, "x": 10, "y": 10, "order": 1 }, { "id": 1, "x": 50, "y": 10, "order": 2 }, { "id": 1, "x": 50, "y": 50, "order": 3 }, { "id": 1, "x": 10, "y": 50, "order": 4 }, { "id": 1, "x": 10, "y": 10, "order": 5 }] } } </pre>

※ 원천데이터에서 한쪽 이미지의 실면적은 라벨링 가공 편의를 위해 최적화 재단된 수치임 (가로 60m*세로 60m)

2. 라벨링 데이터 구성

– 메타데이터는 <info>, <image>, <annotation>의 3개 분야로 구성됨

– <info>

- 구축된 항공영상 도시건물 변화탐지 데이터셋에 대한 공통 정보. 원천/라벨링 데이터 구축 완료 후 일괄적으로 부여한 값임

– <image>

- 원천데이터(이미지) 파일에 대한 정보로써, 이미지식별자(파일명), 이미지파일확장자(포맷), 파일생성일자, 가로/세로크기(픽셀), 수평/수직해상도(DPI), 비트수준이 포함됨. 해당정보는 원시데이터인 항공영상을 원천데이터로 정제/가공하는 과정에서 추출하여 저장한 값임

– <annotation>

- 원천데이터(이미지)내 변화한 건물을 식별하여 폴리곤 세그멘테이션한 정보로써 폴리곤 식별자, 폴리곤 객체타입, 이미지내 폴리곤내 위치(이미지내 좌표), 변화건물(폴리곤) 외형, 건물 변화유형, 건물 갱신유형, 변화건물 주소가 포함됨

※ 이미지쌍에서 변화한 건물은 양쪽 이미지가 아닌 한쪽에만 라벨링 되어있으며, 이에 대한 위치기준은 세부내용은 구축 가이드라인 문서에서 설명함

No	항목		길이	타입	필수 여부	예시
	한글명	영문명				
1	데이터셋 정보	info	-	-	-	-
1-1	데이터셋명	info.name	18	string	Y	항공영상 도시건물 변화탐지 데이터
1-2	데이터셋 식별자	info.id	-	string	Y	TIF_CHANGE_safe_01
1-3	데이터셋 생성일자	info.date_created	19	string	Y	2021-04-09 09:00:00
1-4	이미지파일 폴더 경로	info.src_path	-	string	Y	/dataSet/change
1-5	라벨링파일 폴더 경로	info.label_path	-	string	Y	/dataSet/change
1-6	데이터셋 카테고리	info.category	1	int	Y	0 : 제1종 전용주거지역 1 : 제2종 전용주거지역 2 : 제1종 일반주거지역 3 : 제2종 일반주거지역 4 : 제3종 일반주거지역 5 : 준주거지역 6 : 상업지역 7 : 준공업지역(공업지역) 8 : 녹지지역
1-7	데이터셋 타입	info.type	1	int	Y	1 : 이미지
2	이미지 정보	images	-	-	-	-
2-1	이미지 식별자	images.id	-	string	Y	2019_DBG_1LB_000005.tif
2-2	이미지 파일 확장자	images.type	3	string	Y	TIF
2-3	이미지 생성 일자	images.data_captured	19	string	Y	2021-04-09 20:50:15
2-4	이미지 파일용량(KB)	images.size	4	int	Y	1713
2-5	이미지 가로크기(픽셀)	images.width	4	int	Y	1508
2-6	이미지 세로크기(픽셀)	images.height	3	int	Y	754
2-7	이미지 수평해상도(DPI)	images.h_resolution	3	int	Y	150
2-8	이미지 수직해상도(DPI)	images.v_resolution	3	int	Y	150
2-9	이미지 비트수준(bit)	images.bit	2	int	Y	24
3	어노테이션 정보	annotations	-	-	-	-
3-1	폴리곤 식별자	polygon.id	10	string	Y	PLG_000001
3-2	폴리곤 객체타입	polygon.name	8	string	Y	building
3-3	이미지내 폴리곤내 위치	polygon.points	-	array	Y	[(100.440, 105.123),..., (160.333, 0.000)]
3-4	변화건물(폴리곤) 외형	polygon.shape	1	int	Y	0 : 전체가 온전한 건물 1 : 일부가 잘린 건물
3-5	건물 변화유형	polygon.shift	-	array	Y	0 : 신 축 1 : 소 멸 2 : 갱 신
3-6	건물 갱신유형	polygon.update	-	array	-	0 : 총 변화 1 : 철골구조 완성 2 : 설치·구조물 추가 3 : 옥상 색상 변화 4 : 외곽선 변화
3-7	변화건물 주소*	polygon.address	-	string	-	서울특별시 도봉구 쌍문1동 512-21

* 구축대상 : 19-20년 원천데이터

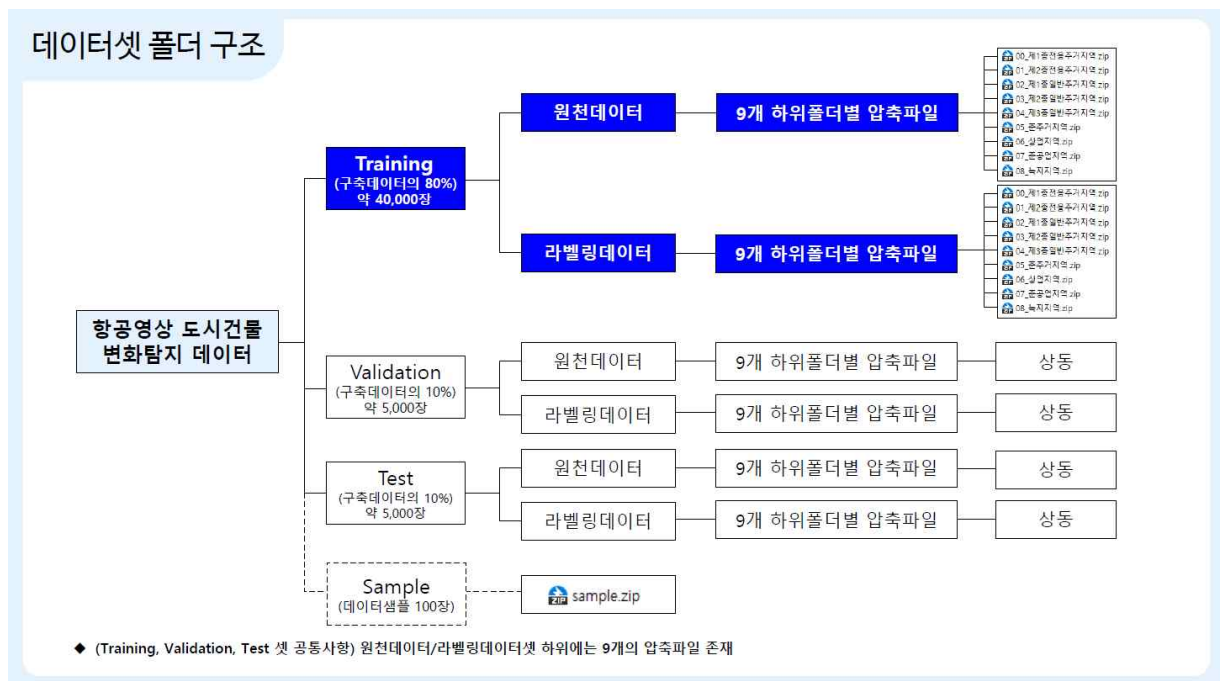
3. 라벨링 데이터 실제예시

```
{
  "info": {
    "info.name": "항공영상 도시건물 변화탐지 데이터",
    "info.id": "TIFF_CHANGE_safe_01",
    "info.date_created": "2021-04-09 09:00:00",
    "info.src_path": "/dataSet/change",
    "info.label_path": "/dataSet/change",
    "info.category": 2,
    "info.type": 1
  },
  "images": {
    "images.id": "2019_DBG_1LB_000005.tif",
    "images.type": "TIF",
    "images.data_captured": "2021-08-26 17:49:24",
    "images.size": 1703,
    "images.width": 1508,
    "images.height": 754,
    "images.h_resolution": 150,
    "images.v_resolution": 150,
    "images.bit": 24
  },
  "annotations": [
    {
      "polygon.id": "PLG_000000",
      "polygon.name": "building",
      "polygon.points": [
        [
          1056.5897341067248,
          316.18122145441106
        ],
        [
          1190.713823174877,
          327.9108389457345
        ],
        [
          1178.4142342210291,
          450.48534173006465
        ],
        [
          1117.5019841638768,
          444.62053298440287
        ],
        [
          1115.7449000276129,
          467.4932870924837
        ]
      ]
    }
  ]
}
```

1041.9473663045246,
462.2149592213881

```
    ],  
    "polygon.shape": 0,  
    "polygon.shift": 2,  
    "polygon.updates": [  
        3  
    ],  
    "polygon.address": "서울특별시 도봉구 쌍문1동 512-21"  
  }  
]  
}
```

4. 업로드 데이터 폴더구조



데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)	메일주소	담당업무
		서울디지털재단	박지혜	02-570-4644	jhpark@sdf.seoul.kr	실무 책임
	참여기관	기관명	담당업무			
		엘티메트릭(주)	원시데이터 정제, 원천 데이터 제작			
		공간정보산업진흥원	품질 관리			
		인크루트	라벨링 데이터 제작, 제작도구 개발			
		(주)다비오	AI 알고리즘 개발			
		서울특별시 (수요기관)	원시데이터 제공			