

메타데이터 정보 (다중기입가능)	분야	데이터 유형 ¹⁾	구축 데이터량	원천데이터 형식 ²⁾	라벨링 형식 ³⁾	라벨링 유형 ⁴⁾
	농축수산	이미지	100,000	jpg	json	bbox
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관(총괄)	가공기관	검수기관	
	자체구축	2021년	(주)씨씨미디어 서비스	(주)유비엔	(주)LK2	
	데이터 문의처	기관명	문의담자명	전화번호 (유선전화번호기입)	메일주소	
		(주)씨씨미디어 서비스	김기수	02-501-6371	kskim@ccmedia.co.kr	
	데이터 소개	수직농장 과 스마트팜 온실 환경에서 재배되는 엽채류(상추,케일,겨자채,근대)의 생육이미지와 재배환경데이터를 수집하여 정의된 포맷으로 데이터셋 구축				
	주요키워드	엽채류, 상추,근대, 겨자채, 케일 ,재배환경, 생육지표				
카테고리 정의서		첨부의 카테고리 정의서 엑셀파일에 데이터카테고리 작성하여 제출				

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)

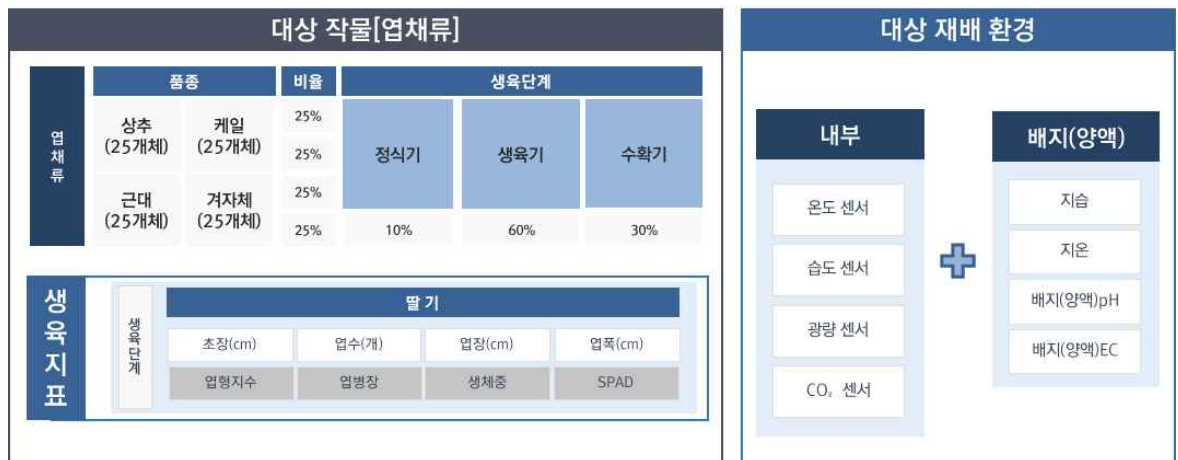
5) 4대 언론기사, 자체 수집,.....

데이터셋명	국문영문	수직농장통합데이터(엽채류)																														
		Vertical Farming Leafy Vegetables AI Training Dataset																														
구축목적	수직농장재배 엽채류에 대한 생육이미지 및 재배환경데이터를 활용한 AI 데이터 셋 구축																															
활용서비스	엽채류 데이터셋을 이용한 농업 분야 AI 서비스 분야 활성화																															
소개	수직농장 및 스마트 팜에서 재배되는 엽채류의 생육이미지를 일정 시간 간격으로 수집하고, 이미지 수집 시간대에 재배환경 센서데이터를 함께 수집하여 이미지와 재배환경데이터를 활용하여 재배농가에 혁신적인 AI 기술적용 및 응용 서비스 개발을 위한 수직농장 엽채류 데이터셋을 구축 함. 데이터 획득 데이터 수집 딸기 총 100,000장 1차 정제 딸기,엽채류 IOT 센서(측정)데이터 200,000개 텍스트 데이터 생육 정보 현장 조사 수집 데이터 정제·가공 클라우드 소싱 2차 정제 어노테이션 검수 데이터 셋 저장 데이터 셋 생성 (image + json) AI 응용서비스 개발 신기술 개발 → AI 모델 개발 알고리즘 구현 반복 검증																															
	1. 데이터 구축 규모 엽채류 이미지 데이터 10만건, 재배환경데이터 20만건 <table><tr><th>데이터종류</th><th>데이터형태</th><th>규모</th><th>어노테이션규모</th><th>결과물 규모</th></tr><tr><td>이미지</td><td>엽채류 이미지</td><td>10 만건</td><td>10만건</td><td>겨자채 37,751건 근대 38,738건</td></tr><tr><td>텍스트</td><td>재배환경 센서데이터</td><td>20 만건</td><td>해당없음</td><td>상추 46,751건 케일 39,716건</td></tr></table> 2. 데이터 분포 엽채류의 작물 (상추,케일,겨자채,근대), 생육단계(정식,생육기,수확기) 클래스에 따라 수집 ※ 작물 특성상 생육상황, 재배상황에 따라 생육단계의 비율 변동 (생육정지 현상) <table><tr><th rowspan="5">엽채류</th><th>작물</th><th colspan="3">생육단계</th></tr><tr><td></td><th>정식기</th><th>생육기</th><th>수확기</th></tr><tr><td>상추</td><td rowspan="4">10%</td><td rowspan="4">60%</td><td rowspan="4">30%</td></tr><tr><td>케일</td></tr><tr><td>겨자채</td></tr><tr><td>근대</td></tr></table>		데이터종류	데이터형태	규모	어노테이션규모	결과물 규모	이미지	엽채류 이미지	10 만건	10만건	겨자채 37,751건 근대 38,738건	텍스트	재배환경 센서데이터	20 만건	해당없음	상추 46,751건 케일 39,716건	엽채류	작물	생육단계				정식기	생육기	수확기	상추	10%	60%	30%	케일	겨자채
데이터종류	데이터형태	규모	어노테이션규모	결과물 규모																												
이미지	엽채류 이미지	10 만건	10만건	겨자채 37,751건 근대 38,738건																												
텍스트	재배환경 센서데이터	20 만건	해당없음	상추 46,751건 케일 39,716건																												
엽채류	작물	생육단계																														
		정식기	생육기	수확기																												
	상추	10%	60%	30%																												
	케일																															
	겨자채																															
근대																																
데이터셋 통계 (구축 규모 및 분포)																																

1. 대표도면

제목				
수직농장 통합데이터 업채류				
유형	• 이미지			
대상명	대상	• 업채류 재배 이미지	표시형식	• jpg
		• 재배환경데이터		• json
		• 생육지표 데이터		• json
품종	• 업채류			
수집내용	• 업채류 재배성장 이미지 100,000 장 • 재배환경 데이터 200,000장 • 생육단계에 따른 생육지표 데이터			
데이터 구성				

1. 이미지 데이터 : 업채류의 생육이미지를 1시간 간격으로 수집 (결측 존재)
2. 재배환경데이터 : 생성된 수집된 업채류 이미지에 대하여 시간 기준으로 재배환경데이터 2개 SET 구성(결측 존재)
3. 생육지표 : 1주 간격으로 생육지표 수집 (작물의 성장 및 농장환경에 따라 변동가능)
딸기 생육 image 와 재배환경데이터 , 생육지표 데이터를 AI 데이터셋 모음으로 구축 함.
4. 데이터 납품 형상



생성요약

업채류 생육이미지는 이미지 셋 으로 재배환경데이터 및 생육지표는 JSON 파일 형태로 이미지+ JSON 의 데이터셋 을 생성하여 제공.

데이터셋 구성

2. 라벨링데이터 구성

구분	한글명	영문명	설명	타입	필수여부
1	이미지정보	images	이미지정보	Object	Y
1-1	이미지식별자	images.id	이미지식별자	Number	Y
1-2	농장코드	Farm_id	농장코드	String	Y
1-3	작물개체코드	Crops_id	작물개체코드	String	Y
1-4	작물명	Crops	작물명	String	Y
1-5	품종명	Kind_Type	품종명	String	Y
1-6	파일경로	file_path	파일경로	String	Y
1-7	이미지의 이름	fname	이미지의 이름	String	Y
1-8	이미지의 종류	fext	이미지의 종류	String	Y
1-9	가로 넓이	width	가로 넓이	String	Y
1-10	세로 넓이	height	세로 넓이	String	Y
1-11	이미지 촬영일	create_date	이미지 촬영일	String	Y
1-12	이미지 변환저장일	captured_date	이미지 변환저장일	String	Y
1-13	생육단계	growth_stage	생육단계	String	Y
1-14	어노테이션 개체	leaf	어노테이션 개체	String	
1-15	어노테이션 개체	plant_body	어노테이션 개체	String	
2	재배환경정보	environment	재배환경정보	Array	Y
2-1	농장코드	Farm_id	농장코드	String	Y
2-2	수집일자	Receive_date1	수집일자	String	Y
2-3		id		String	
2-4	내부 온도	ti_value	내부 온도	String	
2-5	내부 습도	hi_value	내부 습도	String	
2-6	내부 CO2	ci_value	내부 CO2	String	
2-7	광량	ir_value	광량	String	
2-8	양액온도	tl_value	양액온도	String	
2-9	양액 EC	ei_value	양액 EC	String	
2-10	양액 PH	pl_value	양액 PH	String	
3	생육지표조사	Growth_index	생육지표조사	Object	Y
3-1	작물개체코드	Crops_id	작물개체코드	String	Y
3-2	측정일	measured_date	측정일	String	Y
3-3	초장	stem_length	초장	String	
3-4	엽수	leaf_cnt	엽수	String	
3-5	엽폭	leaf_width	엽폭	String	
3-6	엽장	leaf_length	엽장	String	
3-7	착(과중)	fr_weight	착(과중)	String	
3-8	엽병장	petiole_length	엽병장	String	
3-9	엽형지수	leaf_index	엽형지수	String	
3-10	엽록소 함량	SPAD	엽록소 함량	String	
4	기타정보	ETC_infor	기타정보	Object	

4-1	작물개체코드	Corps_id	작물개체코드	String	
4-2	수집일	Create_date	수집일	String	
4-3	정보	inform	정보	String	
5	어노테이션정보	annotations[]	어노테이션정보	Array	Y
5-1	영역	annotations[].area	영역	Number (float)	Y
5-2	bbox 위치정보	annotations[].bbox[]	bbox 위치정보	Number (float)	Y
5-3	어노테이션 아이디(툴 생성)	annotations[].id	어노테이션 아이디(툴 생성)	Number (int)	Y
5-4	어노테이션 툴 생성	annotations[].isCrowd	어노테이션 툴 생성	Number (int)	Y
5-5	어노테이션 이미지 아이디(툴생성)	annotations[].image_id	어노테이션 이미지 아이디(툴생성)	Number (int)	Y
6	라이선스	licenses[]	라이선스	Array	Y
6-1	라이선스 아이디	licenses[].id	라이선스 아이디	Number (int)	
6-2	라이선스 주체	licenses[].name	라이선스 주체	String	
6-3	라이선스 관련정보 url	licneses[].url	라이선스 관련정보 url	String	
6-4	카테고리	categories[]	카테고리	Array	Y
6-5	카테고리id	categories[].id	카테고리id	Number (int)	
6-6	카테고리이름	categories[].name	카테고리이름	String	
6-7	상위 카테고리	categories[].supercategory	상위 카테고리	String	

3 라벨링 데이터 실제예시

• 라벨 데이터

- 작물 정보 (작물명, 품종명)
- 생육단계 및 생육지표 데이터
- 재배환경센서 데이터
- 어노테이션 결과 정보(BBOX)

```
{ "images": {
  "image_id": 146500,
  "farm_id": "AIF001",
  "crops_id": "C11_L01_01",
  "crops": "상추",
  "kind_type": "로메인",
  "file_path": "/cropsImages/L/L01/AIF001/C11",
  "fname": "C11_L01_01_001_210903170621.jpg",
  "fext": ".jpg",
  "width": "1057",
  "height": "1187",
  "create_date": "2021-10-07 14:17:09",
  "date_captured": "2021-09-03 17:06:21",
  "growth_stage": "정식기",
  "leaf": "1",
  "plant_body": "0"
},
"envrionments": [
```

```

{
  "farm_id": "AIF001",
  "receive_date": "2021-08-29 00:00:02",
  "id": 0,
  "ti_value": "21.0",
  "hi_value": "57",
  "ci_value": "346",
  "ir_value": "77",
  "tl_value": "21.8",
  "ei_value": null,
  "pl_value": null
},
{
  "farm_id": "AIF001",
  "receive_date": "2021-08-29 00:05:01",
  "id": 1,
  "ti_value": "22.0",
  "hi_value": "58",
  "ci_value": "346",
  "ir_value": "9",
  "tl_value": "21.7",
  "ei_value": null,
  "pl_value": null
}
],
"growth_index": {
  "corps_id": "C11_L01_01",
  "measured_date": null,
  "stem_length": null,
  "leaf_cnt": null,
  "leaf_width": null,
  "leaf_length": null,
  "stem_thick": null,
  "fr_weight": null,
  "petiole_length": null,
  "leaf_index": null,
  "spad": null
},
"etc_infor": {
  "corps_id": "C11_L01_01",
  "create_date": "",
  "inform": ""
},
"annotations": [
  {
    "area": 39397.63,
    "bbox": [
      346.31,
      518.35,
      285.78,
      137.86
    ]
  }
],

```

	<pre> "isCrowd": 0, "id": 0, "image_id": 146500, "category_id": 0 }], "licenses": [{ "id": 1, "name": "CCMediaService", "url": "" }], "categories": [{ "id": 0, "name": "윙", "supercategory": "none" }] } </pre>					
데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)	메일주소	담당업 무
		(주)씨씨미디어서비스	김기수	02-501-6371	kskim@ccmedia.co.kr	총괄
	참여기관	기관명	담당업무	기관명	담당업무	
		(주)유비엔	김지성	(주)유비엔	어노테이션 및 라벨링	
		경상남도 농업기술원	김우일	경상남도 농업기 술원	작물 전문	
		경상국립대학교	김현태	경상국립대학교	작물 전문	