

메타데이터 정보 (다중기입가능)	분야	데이터 유형 ¹⁾	구축 데이터량	원천데이터 형식 ²⁾	라벨링 형식 ³⁾	라벨링 유형 ⁴⁾
	<i>NIA 기입</i>	이미지	350,621	jpg	json	-바운딩박스(이미지) -세그멘테이션(이미지)
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관(총괄)	가공기관	검수기관	
	자체수집	<i>2021년</i>	유클리드소프트	유클리드소프트	유클리드소프트	
	데이터 문의처	기관명	문의담당자명	전화번호 (유선전화번호기입)	메일주소	
		유클리드소프트	황준원	042-488-6589	jwhwang@euc lidsoft.co.kr	
	데이터 소개	스마트팜에서 재배되는 버섯 품종 인식, 수확기 판단, 병충해 판단을 위한 AI기술 개발을 위한 데이터				
	주요키워드	버섯, 스마트팜, 버섯환경, 버섯 데이터				
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)

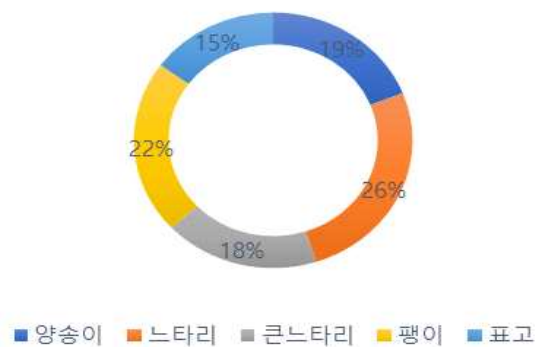
5) 4대 언론기사, 자체 수집,.....

데이터셋명	국문영문	스마트팜 통합 데이터(버섯)
		AI Data Of Smart Farm (mushroom)
구축목적	본 인공지능 학습 데이터셋 구축은 스마트팜 버섯 농업 산업의 AI 기술 선순환 생태계 조성을 목표로 함	
활용서비스	<p>스마트팜에서 재배되는 버섯 품종 인식과 버섯의 수확기 판단, 병해 판단 가능</p> <ul style="list-style-type: none"> - 소비자 수요 속성 맞춤형 작물 생산 <p>이미지데이터, 사물인터넷(IoT), 인공지능(AI) 등 실시간으로 생성·축적되는 빅데이터를 활용한 정밀 농업으로 버섯산업이 충분히 성장 가능할 것으로 예상</p> <ul style="list-style-type: none"> - 대규모 플랫폼 구축 <p>각 버섯별, 실증농가별, 생산품질별로 다양하게 수집된 데이터를 중심으로 공통구조에 기반한 데이터 연계를 추진, 중소생산농가 통합형 또는 대기업과의 상생형 플랫폼 등을 구축하여 작물의 특화성이 보다 다양해질 것으로 예상</p>	
소개	<p>양송이, 느타리, 큰느타리, 팽이, 표고 버섯 총 5품종의 전체 생육주기와 환경요소에 따른 촬영 자동화 장치를 통해 이미지 데이터를 확보. 버섯 생육주기별 데이터를 확보·구축하고 이를 통해 수확기 버섯 상태를 구분하여 생육의 편의성과 효율성을 개선. 버섯 관련 인공지능 데이터가 전무한 상황에 맞춰 수집된 데이터를 공개하여 현실 환경 개선을 하고자 함</p> <div> <div> <p>수집</p>  <p>버섯 (농장 : 지역)</p> <p>느타리 (채인 농장: 경기 화성) 큰느타리(해남 농장: 전남 해남) 팽이 (연우 농장: 충북 음성) 표고 (오름 농장: 충남 천안) 양송이 (정해평 농장: 충남 부여) 병충해(국립원예특작과학원: 충북 음성)</p> </div> <div> <p>정제</p>  <p>환경정보 데이터 매칭 영상 데이터1, 영상 데이터2, 영상 데이터3 영상 데이터1, 영상 데이터2, 영상 데이터3 이미지 데이터 변환 이미지 데이터 전달</p> </div> <div> <p>가공</p>  <p>오픈 플랫폼 : LabelOn 버섯 개체 라벨링 가공 검수(승인, 반려, 작업불가) Bounding box DATA Polygon DATA 가공된 데이터 전달</p> </div> <div> <p>모델링</p>  <p>모델링 설계 가공된 데이터 학습 테스트 모델 적용 목표수치 도출</p> <p>모델링 개발</p> <ol style="list-style-type: none"> 1. 버섯 품종 인식 2. 수확기 판단 3. 병충해 판단 </div> </div>	
데이터셋 통계 (구축 규모 및 분포)	<p>1. 데이터 구축 규모</p> <ul style="list-style-type: none"> - 원천데이터 35만 621건 (양송이 7만 9백 건, 느타리 8만 6천 460건, 큰느타리 6만 2천 419건, 팽이 7만 7천 50건, 표고 5만 3천 792건)을 활용하여 각각 Polygon Segmentation 5만 4백 건, Bounding box 27만 9천 721건으로 가공 하여 총 35만 621건의 데이터 도출 - 5품종의 버섯(양송이, 느타리, 큰느타리, 팽이, 표고)으로 분류하였으며, 생장별 3단계 (배양단계, 생육단계, 병해충)구분하여 데이터 구축 - 이미지 데이터와 환경데이터를 1:1 매칭하여 데이터 구축 	

데이터 종류	품종	CLASS	수량	데이터 형태
버섯	양송이	생육단계	50,400	Polygon
		병해충	20,500	
	느타리	배양단계	4,960	Bounding Box
		생육단계	59,500	
		병해충	22,000	
	큰느타리	배양단계	2,719	
		생육단계	43,700	
		병해충	16,000	
	팽이	배양단계	4,050	
		생육단계	53,000	
		병해충	20,000	
	표고	배양단계	1,792	
		생육단계	38,100	
		병해충	13,900	
	합계	350,621 건		

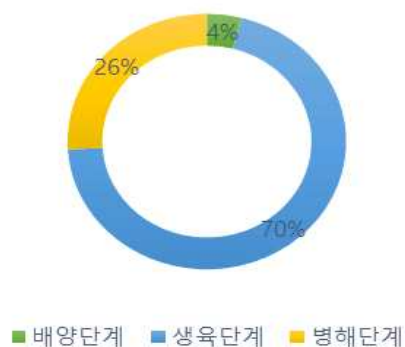
2. 데이터 분포

버섯 품종 별 구축 비율



품종	수량
양송이	70,900
느타리	86,460
큰느타리	62,419
팽이	77,050
표고	53,792
합계	350,621

버섯 주기별 비율



주기	수량
배양	13,521
생육	244,700
병해	92,400
합계	350,621

데이터셋 구성

1. 대표도면

Bounding Box



Bounding Box는 버섯, 배지, 병목을 포함하는 것을 기준으로 각 버섯의 특성에 따라 각기 다른 상세 기준 설정

▶ 총 300,221건 구축

Polygon Segmentation



Segmentation은 버섯의 외곽선을 따라 라벨링을 하는 것을 기준으로 상세 작업 기준 설정

▶ 총 50,400건 구축

양송이 생육

가공 : 저작도구 화면

검수 : 저작도구 화면



2. 라벨링데이터 구성

분류	구분	항목명	설명	범위	타입	필수 여부	비고
파일 정보	1-1	DATASET_NAME	데이터셋명	1000	String	Y	
	1-2	DATASET_DETAIL	데이터셋상세설명	1000	String	N	
	1-3	VERSION	버전	3	String	N	
	1-4	LICENSE	라이선스	100	String	N	
	1-5	CREATE_DATE_TIME	파일 생성 일자	20	String	Y	
	1-6	CONTRIBUTOR	기여자	200	String	N	
	1-7	URL	URL	200	String	N	
	1-8	CATEGORY_NAME	카테고리 명	20	String	Y	
이미지 파일 정보	2-1	IMAGE_URL	이미지 URL	100	String	Y	
	2-2	IMAGE_FILE_NAME	이미지 파일명	10	String	Y	
	2-3	WIDTH	이미지 가로	1072~1080	NUMBER	Y	
	2-4	HEIGHT	이미지 세로	1920	NUMBER	Y	
	2-5	ANNOTATION_COUNT	이미지당 라벨 개수	0~100	NUMBER	Y	

어노테이션 정보	3-1	ID	어노테이션 식별자	0~4294967295	NUMBER	Y	
	3-2	BOUNDING_BOX_X_COORDINATE	바운딩박스 X좌표	0~4294967295	NUMBER	N	
	3-3	BOUNDING_BOX_Y_COORDINATE	바운딩박스 Y좌표	0~4294967295	NUMBER	N	
	3-4	BOUNDING_BOX_WIDTH	바운딩박스 가로	0~4294967295	NUMBER	N	
	3-5	BOUNDING_BOX_HEIGHT	바운딩박스 높이	0~4294967295	NUMBER	N	
	3-6	SEGMENTATION	세그멘테이션		array	N	NUMBER ARRAY
	3-7	SEGMENTATION_AREA_TOTAL	세그멘테이션영역합	0~4294967295	NUMBER	N	
	3-8	CROWDSOURCING_OPERATION_ALTERNATIVE	크라우드소싱 작업여부		Boolean	N	
메타	4-1	DBYHS_SPCHCKN	병해충구분	10	String	N	-세균갈색무늬병 -세균성검은썩음병 -흰곰팡이 -숨털곰팡이 -푸른곰팡이
	4-2	DBYHS_NORMALITY_ALTERNATIVE	정상여부		Boolean	N	
	4-3	IP_CAMERA_ID	IP카메라 아이디	0~100	NUMBER	N	
	4-4	WIND_SPEED	풍속	0~5	NUMBER	N	
	4-5	AIR_VELOCITY	풍속	0~5	NUMBER	N	
	4-6	TEMPERATURE	온도	0~100	NUMBER	N	
	4-7	HUMIDITY	습도	0~100	NUMBER	N	
	4-8	ILLUMINATION_INTENSITY	조도	-60~100	NUMBER	N	
	4-9	CARBON_DIOXIDE	이산화탄소(CO2)	-20~7600	NUMBER	N	
	4-10	GUIDELINE	가이드라인		array	N	NUMBER ARRAY
	4-11	IMAGE_CREATE_DATE	이미지 생성 날짜	YYYY-MM-DD	String	N	
	4-12	IMAGE_CREATE_TIME	이미지 생성 시간	HH:mm:ss	String	N	
	4-13	IMAGE_CREATE_DAY_OF_WEEK	이미지 생성 요일	20	String	N	
	4-14	STIPE_LENGTH	대길이	0	NUMBER	N	
	4-15	STIPE_THICKNESS	대두께	0	NUMBER	N	
	4-16	PILEUS_DIAMETER	갓직경	0	NUMBER	N	
	4-17	PILEUS_THICKNESS	갓두께	0	NUMBER	N	
	4-18	GROSS_WEIGHT	총중량	0	NUMBER	N	

3 라벨링 데이터 실제예시

```
{
  "INFO": {
    "DATASET_NAME": "양송이 병해",
    "DATASET_DETAIL": "(스마트팜 통합데이터_버섯)",
    "VERSION": "1.0",
    "LICENSE": "",
    "CREATE_DATE_TIME": "2021-12-06 16:25:55",
    "CONTRIBUTOR": "",
    "URL": "https://www.labelon.kr",
    "CATEGORY_NAME": "양송이"
  },
  "IMAGE": {
    "IMAGE_URL": "https://images.labelon.kr/2021/11/30/5b853f816c80457ea6dc8172ab059f88.jpg",
    "IMAGE_FILE_NAME": "양송이_생육실1_8_16344202.jpg",
    "WIDTH": 1080,
    "HEIGHT": 1920,
    "ANNOTATION_COUNT": 1
  },
  "ANNOTATION_INFO": [
    {
      "ID": 65225628,
      "BOUNDING_BOX_X_COORDINATE": 204,
      "BOUNDING_BOX_Y_COORDINATE": 1261,
      "BOUNDING_BOX_WIDTH": 295,
      "BOUNDING_BOX_HEIGHT": 342,
      "SEGMENTATION": null,
      "SEGMENTATION_AREA_TOTAL": null,
      "CROWDSOURCING_OPERATION_ALTERNATIVE": true
    }
  ],
  "META": {
    "DBYHS_SPCHECKN": "푸른곰팡이병",
    "DBYHS_NORMALITY_ALTERNATIVE": false,
    "IP_CAMERA_ID": 8,
    "WIND_SPEED": 0.0,
    "AIR_VELOCITY": 0.0,
    "TEMPERATURE": 18.2,
    "HUMIDITY": 97.6,
    "ILLUMINATION_INTENSITY": 0.0,
    "CARBON_DIOXIDE": 1156.0,
    "GUIDELINE": null,
    "IMAGE_CREATE_DATE": "2021-11-28",
    "IMAGE_CREATE_TIME": "15:08:11",
    "IMAGE_CREATE_DAY_OF_WEEK": "Sunday",
    "STIPE_LENGTH": null,
    "STIPE_THICKNESS": null,
    "PILEUS_DIAMETER": null,
    "PILEUS_THICKNESS": null,
    "GROSS_WEIGHT": null
  }
}
```

[JSON 어노테이션 데이터]

데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)	메일주소	담당업무
		유클리드소프	황준원	042-488-65	jwhwang@euc	데이터 구축

		트		89	lidsoft.co.kr	사업 총괄 데이터 수집, 가공, 검증
	참여기관	기관명	담당업무	기관명	담당업무	
		엠에이치소프트	데이터 정제, AI 모델 개발	전북대학교 산학협력단	데이터, AI모델 설계	