

메타데이터 정보 (다중기입가능)	분야	데이터 유형 ¹⁾	구축 데이터량	원천데이터 형식 ²⁾	라벨링 형식 ³⁾	라벨링 유형 ⁴⁾
	농축수산	이미지	182,000건	JPG	JSON	세그멘테이션
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관(총괄)	가공기관	검수기관	
	자체수집	2021년	푸디웍	글로벌여성ICT 네트워크	AWASOFT, 서우엠에스	
	데이터 문의처	기관명	문의담당자명	전화번호 (유선전화번호기입)	메일주소	
		아와소프트	우영춘	070-4355-5387	wyc11218@aw asoft.co.kr	
	데이터 소개	인공지능 학습용 곤충 사육데이터 세트 구축				
	주요키워드	곤충, 동애등에, 누에, 갈색거저리, 흰점박이꽃무지				
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	지능형 곤충 사육 데이터
		Intelligent Insect Breeding Data
구축목적	1. 실제 상업용 곤충사육시설에 적용 가능한 학습용 데이터 및 AI 모델 구축 2. 곤충사육시설 최적의 환경을 위한 핵심 관리요소 확보 3. AI 기술 개발에 필수적인 공공데이터를 제공하고, 민간에서 AI 기술 발전에 따라 자생적으로 데이터를 확장, 개방하는 선순환 생태계 조성	
활용서비스	곤충영상분석모델, 곤충환경분석모델	
소개	곤충 사육 현장시스템 구축을 통한 곤충 및 감염 곤충의(동애등에, 흰점박이꽃무지, 갈색거저리, 누에)의 사육 원시 데이터의 획득하여 실제 상업용 곤충사육시설에 적용 가능한 학습용 데이터 및 AI 모델 구축 및 곤충사육시설 최적의 환경을 위한 핵심 관리요소 확보하여 AI 기술 개발에 필수적인 공공데이터를 제공하고, 민간에서 AI 기술 발전에 따라 자생적으로 데이터를 확장, 개방하는 선순환 생태계 조성	
	1. 데이터 구축 규모	
데이터셋 통계 (구축 규모 및 분포)		

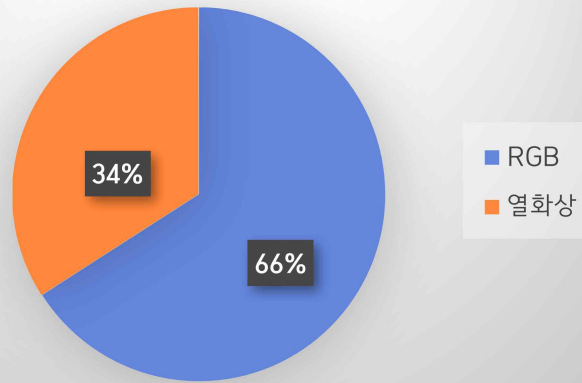
종류			획득규모	
곤충 RGB 영상)	동애등에	알	500장	곤충 길이, 성장모델 분석
		애벌레	12,000장	
		번데기	3,000장	
		성충	7,000장	
		감염곤충	7,500장	
	흰점박이 꽃무지	알	500장	
		애벌레	14,000장	
		번데기	1,000장	
		성충	7,000장	
		감염곤충	7,500장	
	갈색거저리	알	500장	
		애벌레	12,000장	
		번데기	3,000장	
		성충	7,000장	
		감염곤충	7,500장	
	누에	알	500장	
		애벌레	15,000장	
		번데기	4,000장	
		성충	3,000장	
		감염곤충	7,500장	
곤충 열화상 영상	동애등에	알	0장	곤충 길이, 성장모델 분석
		애벌레	5,000장	
		번데기	2,000장	
		성충	6,000장	
		감염곤충	2,500장	
	흰점박이 꽃무지	알	0장	
		애벌레	7,000장	
		번데기	0장	
		성충	6,000장	
		감염곤충	2,500장	
	갈색거저리	알	0장	
		애벌레	5,000장	
		번데기	2,000장	
		성충	6,000장	
		감염곤충	2,500장	
	누에	알	0장	
		애벌레	8,000장	
		번데기	3,000장	
		성충	2,000장	
		감염곤충	2,500장	
센서데이터	RGB 영상		120,000건	센서 수집 데이터
	열화상 영상		62,000건	
관리데이터	곤충길이 데이터(수동측정)		2,500개 이상	동애등에, 흰점박이 꽃무지, 갈색거저리, 누에 집적 측정을 통한 관리 데이터 구축
	곤충평균무게데이터(수동측정)		2,500개 이상	
	곤충탈피데이터		1,000개 이상	
	사료공급량 데이터		1,000개 이상	
	사료 종류 데이터		1,000개 이상	
	수분 제공 데이터		1,000개 이상	
	감염 상태 정보 데이터		1,000개 이상	

2. 데이터 분포

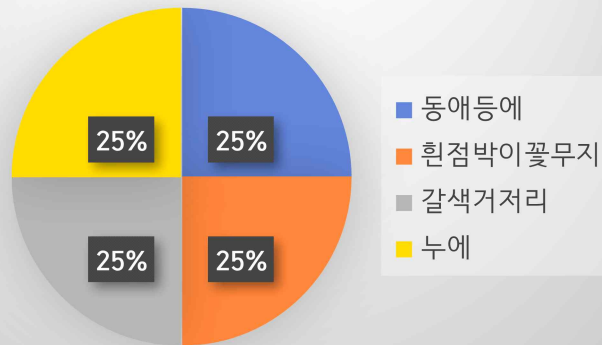
- 구축 데이터 종류 : 누에, 흰점박이 꽃무지, 갈색거저리, 동애등에 RGB, 열화상 이미지
- 구축 데이터 클래스별 비율

종류		획득규모		비율	비고
곤충 RGB 영상	동애등에	알	500장	0.27%	
		애벌레	12,000장	6.59%	
		번데기	3,000장	1.65%	
		성충	7,000장	3.85%	
		감염곤충	7,500장	4.12%	
	흰점박이 꽃무지	알	500장	0.27%	
		애벌레	14,000장	7.69%	
		번데기	1,000장	0.55%	
		성충	7,000장	3.85%	
		감염곤충	7,500장	4.12%	
	갈색거저리	알	500장	0.27%	
		애벌레	12,000장	6.59%	
		번데기	3,000장	1.65%	
		성충	7,000장	3.85%	
		감염곤충	7,500장	4.12%	
	누에	알	500장	0.27%	
		애벌레	15,000장	8.24%	
		번데기	4,000장	2.20%	
		성충	3,000장	1.65%	
		감염곤충	7,500장	4.12%	
곤충 열화 상 영상	동애등에	알	0장	0.00%	열화상 이미지로 탐지 불가
		애벌레	5,000장	2.75%	
		번데기	2,000장	1.10%	
		성충	6,000장	3.30%	
		감염곤충	2,500장	1.37%	
	흰점박이 꽃무지	알	0장	0.00%	열화상 이미지로 탐지 불가
		애벌레	7,000장	3.85%	
		번데기	0장	0.00%	
		성충	6,000장	3.30%	
		감염곤충	2,500장	1.37%	
	갈색거저리	알	0장	0.00%	열화상 이미지로 탐지 불가
		애벌레	5,000장	2.75%	
		번데기	2,000장	1.10%	
		성충	6,000장	3.30%	
		감염곤충	2,500장	1.37%	
	누에	알	0장	0.00%	열화상 이미지로 탐지 불가
		애벌레	8,000장	4.40%	
		번데기	3,000장	1.65%	
		성충	2,000장	1.10%	
		감염곤충	2,500장	1.37%	

RGB/열화상 데이터 분포



곤충별 데이터 분포



RGB_ 곤충 생육단계별 분포



열화상_ 곤충 생육단계별 분포



<데이터 분포표>

1. 라벨링데이터 구성

구분	ID	요소명(영/한)		타입	예시	필수 여부	비고
데이터셋 구성 메타정보	1	Image File Name	이미지파일명	String	FO1_CA01_INSEC_T01_20210805_01.JPG	Y	
	2	Image File Identification	이미지파일아이디	number	220394	Y	[0-999999]
	3	Image Photography Date and Time	이미지촬영일시	String	2021.08.05. 13:00:00	Y	YYYY.MM.DD HH:MM:mm
	4	Envrn. Info. MEASUREMENT Date and Time	환경정보 측정일시	String	2021.08.05. 13:00:00	Y	YYYY.MM.DD HH:MM:mm
	5	Image File Format	이미지파일포맷	String	JPG	Y	JPG
	6	Image File MAGNITUDE	이미지파일크기(사진파일크기)	number	3000	Y	[0-1,000,000,000]
	7	Equipment Information	장비정보	String	RGB CAM	Y	[RGB CAM, 열화상]
	8	Camera Info.	카메라정보	String	Canon ESO R	Y	카메라모델명
	9	Image Obtain Place Type	이미지획득장소유형	String	Huge_farm	Y	[Huge_farm, Middle_farm, Small_farm]
	10	Photography Area Name	촬영 지역명	String	청주	Y	[청주, 제주도, 강원도]
	11	Image Photography Location Information	이미지 촬영 위치 정보	String	2번 사육박스	Y	[고해상도 사육박스, 1번사육박스, 2번사육박스, 3번사육박스]
	12	Image Copyright Information	이미지저작권정보	String	FOODYWORM	Y	FOODYWORM
	13	Image Width MAGNITUDE	이미지가로크기	number	1920	Y	[0-5000]
	14	Image Height MAGNITUDE	이미지세로크기	number	1080	Y	[0-5000]
	15	Resolution	해상도	String	FHD	Y	FHD
	16	Bit	비트	number	24	Y	24

구분	ID	요소명(영/한)		타입	예시	필수 여부	비고
	17	Image Color Group Code	이미지색상군코드	String	# 8B0000	N	
	18	Length AND Breadth Rate	종횡비율	String	4:3	Y	[1:1, 3:2, 4:3, 16:9]
	19	Image Pixel	이미지화소	number	1K	Y	[0-1000k]
	20	Image Color Depth	이미지색상심도	String	sRGB	Y	sRGB
	21	ISO Speed	ISO 감도	number	3200	Y	[25-3200]
	22	White Balance	화이트 밸런스	number	5500	Y	[1K~10,000K]
	23	Shutter Speed	셔터 속도	String	f2.8 1/80	Y	f2.8 1/80
	24	F-STOP	조리개값	number	2.8	Y	[2.8, 4, 5.6, 8, 11, 16, 22, 32]
	25	Flashlight Use ALTERNATIVE	플래시사용여부	boolean	true	Y	true, false
	26	Filter Use ALTERNATIVE	필터사용여부	boolean	true	Y	true, false
	27	Focal Length	초점거리	number	50	Y	[0-100,000]
	28	Feed Supply Date and Time	사료공급일시	String	2021.12. 23.12.30. 52	Y	YYYY.MM.D D HH:MM:mm
	29	Feed Supply Amount	사료 공급량	number	100	Y	[0~1000]
	30	Feed Ingredient	사료 성분	String	밀기울	Y	x, 빵잎, 버섯배지, 곤충젤리, 밀기울,배추, 호박,무
	31	CCTV Name	CCTV명	String	cam_01	Y	
	32	Insect Name	곤충명	String	누에	Y	[동애등애, 흰점박이 꽃무지, 누에, 갈색거저리]
	33	Insect Farm Name	곤충 농장명	String	푸디원	Y	
	34	Insect Breeding Room Box Number	곤충 사육실 사육박스 번호	String	4	Y	[고해상도, 1,2,3,4]
	35	Insect Breeding Box Size	곤충 사육박스 사이즈	String	Huge_cag e	Y	[Huge_cage, Small_cage, Middle_cage]
	36	Scientific Name	학명	String	Ptecticus - tenebrife r	N	[Ptecticus_ tenebrifer, Protaetia_ brevitarsis, Silkworm, mealworm_ beetle]
	37	Image File Type	이미지파일종류(유형)	String	jpg	N	jpg
	38	Store Course	저장경로	String		N	
	39	Growth Step Name	생육단계명	String	알	Y	[알, 애벌레, 번데기, 성충]
	40	Growth Step Code MEASUREMENT	생육단계코드	String	0	Y	[0,1,2,3]
	41	Sensor INSTALLATION EXISTENCE AND NONEXISTENCE	측정센서설치유무	boolean	Y	N	[Y/N]
	42	Photographer	촬영자	String		N	
	43	THEMA	주제(테마)			N	

	구분	ID	요소명(영/한)	타입	예시	필수 여부	비고
		44	PHOTOGRAPHY YEAR	촬영년도	number	2021	N
노테이션 정보	1	Image File MAGNITUDE	이미지파일크기(사진파일크기)	number	748979	Y	[0-999999]
	2	Image File Name	이미지파일명	String	곰뱅이1.jpg	Y	
	3	Annotation ID	어노테이션아이디	number	1	Y	[0-9999]
	4	Annotation Type	어노테이션타입	String	polygon	Y	polygon
	5	Polygon X Coordinate	바운딩박스 X좌표	array	[149,139,127,121]	Y	
	6	Polygon Y Coordinate	바운딩박스 Y좌표	array	[120,124,127,130]	Y	
	7	Object Name	객체명	String	Ptecticus _tenebrifer_egg	Y	p7 Object Name 내용 참고
	8	Object Class Code	객체클래스코드	String		N	추가
실서 정보	1	Breeding Room Indoor Temperature	사육장 실내 온도	number	22	Y	0~50
	2	Breeding Room Indoor Humidity	사육장 실내 습도	number	50	Y	0~100
	3	Breeding Room Indoor Co2	사육장 실내 Co2	number	10	Y	0~100
	4	Breeding Room Indoor Surface Moisture Quantity	사육장 실내 표면 수분량	number	30	Y	0~100
	5	Breeding Room Indoor Illumination Intensity	사육장 실내 조도	number	300	Y	0 ~ 100,000
	6	Breeding Room Indoor Ammonia	사육장 실내 암모니아	number	50	Y	0~100
	7	Breeding Room Indoor Oxygen	사육장 실내 산소	number	50	Y	0~100
	8	Breeding Room External Air Velocity	사육장 외부 풍속	number	15	Y	0~100
	9	Breeding Room External Temperature	사육장 외부 온도	number	22	Y	0~50
	10	Breeding Room External Humidity	사육장 외부 습도	number	50	Y	0~100
	11	Breeding Room External Wind Direction	사육장 외부 풍향	String	N,W,E,S 중 하나	Y	[N,W,E,S]
	12	Breeding Room External Rain Alternative	사육장 외부 강우 여부	String	유/무	Y	[유/무]
	13	Insect Breeding Box Temperature	곤충 사육박스 온도	number	22	Y	0~50
	14	Insect Breeding Box Humidity	곤충 사육박스 습도	number	50	Y	0~100
	15	Insect Breeding Box Co2	곤충 사육박스 Co2	number	30	Y	0~100
	16	Insect Breeding Box Illumination Intensity	곤충 사육박스 조도	number	300	Y	0 ~ 100,000
처리 정보	1	Insect Length	곤충 길이	number	5	Y	100이하
	2	Insect Average Weight	곤충 평균 무게	number	5	Y	1000이하
	3	Insect Molting	곤충탈피정보	String	O	Y	O,X

	구분	ID	요소명(영/한)		타입	예시	필수 여부	비고
			Information					
		4	Feed Supply Amount	사료 공급량	number	2	Y	100이하
		5	Feed Type	사료 종류	String	밀기울	Y	x, 뽕잎, 버섯배지, 곤충젤리, 밀기울,배추, 호박,무
		6	Hydration Information	수분제공정보	String	O	Y	O,X
		7	Insect infection status	감염상태	String	O	Y	O,X

2 라벨링데이터 실제 예시

Meta_data
<pre> "METADATA": { "Image File Name": "F01_CA01_Insect01_20210805_01.jpg", "Image File Identification": 220394, "Image Photography Date and Time": "2021.08.05. 13:00:00", "Environ. Info. MEASUREMENT Date and Time": "2021.08.05. 13:00:00", "Image File Fprmat": ".jpg", "Image File MAGNITUDE": "3000", "Equipment Information": "RGB CAM", "Camera Info.": "Canon ESO R", "Image Obtain Place Type": "Huge_farm", "Photography Area Name": "청주", "Image Photography Location Information": "2번 사육박스", "Image Copyright Information": "FOODYWORM", "Image Width MAGNITUDE": "1920", "Image Height MAGNITUDE": "1080", "Resolution": "FHD", "Bit": "24", "Image Color Group Code": "#8B0000", "Length AND Breadth Rate": "4:3", "Image Pixel": "1K", "Image Color Depth": "sRGB", "ISO Speed": "3200", "White Balance": "5500", "Shutter Speed": "f2.8 1/80", "F-STOP": "f2.8", "Flashlight Use ALTERNATIVE": "", "Filter Use ALTERNATIVE": "", "Focal Length": "50mm", "Feed Supply Date and Time": "", "Feed Supply Amount": "", "Feed Ingredient": "", "CCTV Name": "cam_01", "Insect Name": "누에", "Insect Farm Name": "", "Insect Breeding Room Box Number": "4", "Insect Breeding Box Size": "Huge_cage", "Scientific Name": "", "Image File Type": "", "Store Course": "", "Growth Step Name": "", "Growth Step Code": "", "MEASUREMENT Sensor INSTALLATION EXISTENCE AND NONEXISTENCE": "", "Photographer": "", "THEMA": "", "PHOTOGRAPHY YEAR": "" }, </pre>

	Image_Annotation_data <pre>"ANNOTATION": { "Image File MAGNITUDE": 748979, "Image File Name": "굼벵이1.png", "Annotation Regions": [{ "Annotation ID": 1, "Annotation Type": "polygon", "Polygon X Coordinate": [149, 139, 127, 121, 149], "Polygon Y Coordinate": [120, 124, 127, 130, 120], "Object Name": "worm", "Object Class Code": "" }] }</pre>					
	sensor_data <pre>"SENSOR": { "Breeding Room Indoor Temperature": "22", "Breeding Room Indoor Humidity": "50", "Breeding Room Indoor Co2": 10, "Breeding Room Indoor Surface Moisture Quantity": 30, "Breeding Room Indoor Illumination Intensity": "300", "Breeding Room Indoor Ammonia": 50, "Breeding Room Indoor Oxygen": 50, "Breeding Room External Air Velocity": "15", "Breeding Room External Temperature": "22", "Breeding Room External Humidity": "50", "Breeding Room External Wind Direction": "N", "Breeding Room External Rain Alternative": "유", "Insect Breeding Box Temperature": "22", "Insect Breeding Box Humidity": "50", "Insect Breeding Box Co2": 30, "Insect Breeding Box Mat Temperature": "22" },</pre>					
	management_data <pre>"MANAGEMENT": { "Insect Length": "5mm", "Insect Average Weight": "5", "Insect Molting Information": 10, "Feed Supply Amount": "2", "Feed Type": "밀기울", "Hydration Information": 10, "Insect infection status": "감염" }</pre>					
데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호가입)	메일주소	담당업무
		푸디웜(주)	이유석	043-716-0282	senselys@foo dyworm.com	데이터 수집 실무책임자
	참여기관	기관명	담당업무	기관명	담당업무	
		아와소프트	정제 및 검수	글로벌여성IC 네트워크	가공	
		서우엠에스	수집 및 검수	강원도농산물 원종장	수집 및 검수	
		케이엘큐브	AI 모델	전북테크노파크	홍보	