

메타데이터 정보 (다중기입가능)	분야	데이터 유형 ¹⁾	구축 데이터량	원천데이터 형식 ²⁾	라벨링 형식 ³⁾	라벨링 유형 ⁴⁾
	농축수산	이미지	600,716장	jpg. pcd	json	polygon(이미지) bbox(이미지) Cuboid(이미지)
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관(총괄)	가공기관	검수기관	
	자체 수집	2021년	(주)대동 한국자동차연구원 국립농업과학원	(주)테스트웍스	(주)인터마인즈	
	데이터 문의처	기관명	문의담당자명	전화번호 (유선전화번호기 입)	메일주소	
		(주)테스트웍스	서길원	02-423-5178	gwseo@testworks.co.kr	
	데이터 소개	농기계 자율주행 시스템 개발을 위한 인공지능 학습용 이미지(RGB/ NIR/ Lidar) 데이터 구축				
	주요키워드	인공지능, 정밀농업, 농기계, 자율주행				
카테고리 정의서		첨부의 카테고리 정의서 엑셀파일에 데이터카테고리 작성하여 제출(예시참고)				

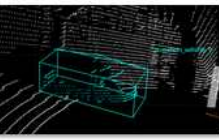
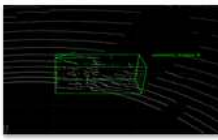
1) 텍스트, 오디오, 이미지, 비디오,

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	61번 정밀농업 농기계 자율주행 데이터																				
		No61. Precision Agricultural Autonomous Driving Data																				
구축목적	농기계 자율주행 시스템 개발을 위한 인공지능 학습용 이미지(RGB/ NIR/ Lidar) 데이터 구축																					
활용서비스	AI모델개발 예시: 농기계 자율주행을 위한 농작지 인식 모델 / 위험물 인식 모델																					
소개	인공지능 학습용 데이터 라벨링 절차 Raw-data 입수/확득 1. Raw-data 분석 2. 학습 데이터 상계 3. 에노데이터선 정제 및 데이터셋 구축 가이드라인 검수 결과를 바탕으로 에노데이터선 정제 및 가이드라인 업데이트 전처리를 위한 데이터 품질 및 입력 데이터 전처리 1단계 Annotation (자동화 실행) 2단계 annotation (수동 보정 작업) 검수 인공지능 학습 데이터셋 구축 데이터셋 Release (협의된 주기로는 데이터셋 구축 분량에 따라 시행) 학습용 데이터 라벨링 구축 인프라 환경 라벨링 Cloud (MS Azure) Storage Annotation Tool Blackolive 라벨링 데이터 Crowd Sourcing aiworks 수집 데이터 Crowd worker Crowd worker Crowd worker 2D RGB/NIR 데이터 라벨링 예시  <2D바운딩박스, 폴리곤 세그멘테이션> LiDAR 데이터 라벨링 예시  <3D바운딩박스(Cuboid)>																					
	데이터셋 통계 (구축 규모 및 분포)	1. 데이터 구축 규모 <table><tr><th>데이터 종류</th><th>가공 방식</th><th>원천데이터 및 라벨 형식</th><th>결과물 규모</th></tr><tr><td>RGB</td><td>Box Polygon</td><td>jpg / json</td><td>266,905장</td></tr><tr><td>NIR</td><td>Box Polygon</td><td>jpg / json</td><td>253,500장</td></tr><tr><td>LiDAR</td><td>Cuboid</td><td>pcd / json</td><td>80,311장</td></tr><tr><td colspan="3">총계</td><td>600,716장</td></tr></table> 2. 데이터 분포 ○ 라벨링 유형별 데이터셋 통계		데이터 종류	가공 방식	원천데이터 및 라벨 형식	결과물 규모	RGB	Box Polygon	jpg / json	266,905장	NIR	Box Polygon	jpg / json	253,500장	LiDAR	Cuboid	pcd / json	80,311장	총계		
데이터 종류	가공 방식	원천데이터 및 라벨 형식	결과물 규모																			
RGB	Box Polygon	jpg / json	266,905장																			
NIR	Box Polygon	jpg / json	253,500장																			
LiDAR	Cuboid	pcd / json	80,311장																			
총계			600,716장																			

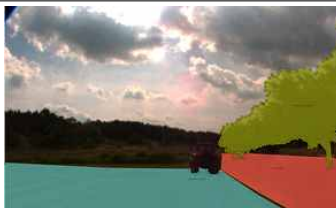
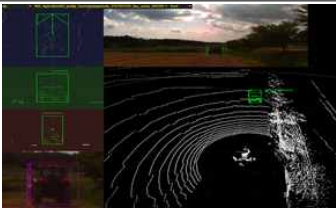
라벨링 유형	이미지 종류	라벨링 수량(장)	총 객체 수(개)	평균 이미지 당 객체(개)
Polygon	RGB	134,281	686,346	5.11
	NIR	121,287	808,194	6.66
Box	RGB	132,351	701,347	5.29
	NIR	132,213	681,017	5.15
Cuboid	LiDAR	80,311	211,724	2.63
합계		600,716	2,808,775	4.68

○ 클래스별 객체 다양성 통계

클래스	객체 명	라벨 명	라벨링 방식	구분	수량(개)	비율
논	경운전	paddy_before_dri ving	polygon	주행가능영역	108,135	3.85%
논	경운후	paddy_after_drivi ng	polygon	주행가능영역	141,186	5.03%
논	물논	paddy_water	polygon	주행가능영역	14,635	0.52%
논	논두렁	paddy_edge	polygon	주행불가능영역	104,154	3.71%
논	작물(벼)	paddy_rice	polygon	주행불가능영역	147,356	5.25%
밭	고랑	field_furrow	polygon	주행가능영역	68,745	2.45%
밭	두둑	field_levee	polygon	주행불가능영역	289,411	10.30%
밭	작물(밭)	field_corps	polygon	주행불가능영역	7,916	0.28%
과수원	농로	orchard_road	polygon	주행불가능영역	11,900	0.42%
과수원	과수	orchard_tree	polygon	주행불가능영역	21,972	0.78%
공통	농로	common_road	polygon	주행가능영역	175,310	6.24%
공통	베일	common_vail	box	정적위험개체	629,264	22.40%
			cuboid		40,363	1.44%
공통	차량	common_vehicle	box	동적위험개체	59,053	2.10%
			cuboid		5,471	0.19%
공통	농기계	common_tractor	box	동적위험개체	43,793	1.56%
			cuboid		13,164	0.47%
공통	돌	common_rocks	box	정적위험개체	29,294	1.04%
			cuboid		6,001	0.21%
공통	나무	coommon_tree	polygon	주행불가능영역	132,963	4.73%
공통	사람	common_person	box	동적위험개체	295,307	10.51%
			cuboid		84,782	3.02%
공통	전봇대	common_pole	box	정적위험개체	316,657	11.27%
			cuboid		61,943	2.21%
합계					2,808,775	100%

1. 대표도면

데이터셋 구성

원천데이터	RGB	NIR	LiDAR
가공 예시			

2. 라벨링 데이터 구성

- box 라벨링 데이터 구조

구분	항목명	타입	필수여부	설명	범위	비고
1	license	array	N	저작권		
2	info[]	array	Y	일반 정보		
	2-1 info[].contributor	string	Y	기여자		
	2-2 info[].date_created	string	Y	데이터 생성일		
	2-3 info[].description	string	Y	데이터 설명		
	2-4 info[].url	string	N	이미지 주소		
	2-5 info[].version	number	N	가공 버전		
	2-6 info[].year	number	N	가공 년도		
3	images[]	array	Y	이미지 정보		
	3-1 images[].id	number	Y	이미지 고유 ID		
	3-2 images[].file_name	string	Y	이미지 파일명		
	3-3 images[].width	number	Y	이미지 너비	0~1920	1920
	3-4 images[].height	number	Y	이미지 높이	0~1080	1080
	3-5 images[].latitude	number	N	위도		32.12345
	3-6 images[].longitude	number	N	경도		127.12355
	3-7 images[].time	string	Y	촬영시간		hh:mm:ss
	3-8 images[].heading	number	N	IMU Heading	0~360	
	3-9 images[].pitch	number	N	IMU Pitch	-999~999	
	3-10 images[].roll	number	N	IMU Roll	-999~999	
4	categories[]	array	Y	객체 정보		
	4-1 categories[].id	number	Y	객체 고유 ID		
	4-2 categories[].name	enum	Y	객체 이름		
	4-3 categories[].supercategory	string	N	객체 상위 카테고리		
5	annotations[]	array	Y	어노테이션 정보		
	5-1 annotations[].id	number	Y	고유 ID		
	5-2 annotations[].image_id	string	Y	이미지 고유 ID		
	5-3 annotations[].category_id	string	Y	객체 고유 ID		
	5-4 annotations[].bbox	array	Y	어노테이션 좌표		[x1, y1, width, height]
	5-5 annotations[].iscrowd	number	N	객체 군집 유무		
	5-6 annotations[].area	number	Y	객체 면적		
	5-7 annotations[].size	enum	Y	객체 크기 분류	"S", "M", "L"	"S", "M", "L"
	5-8 annotations[].occluded	number	Y	가림 및 잘림 여부		0, 1

- cuboid 라벨링 데이터 구조

구분	항목명	타입	필수여부	설명	범위	비고
1	mode	string	N	가공모드		
2	images[]	array	Y	이미지 정보		
2-1	images[].image_id	number	Y	이미지 고유 ID		
2-2	images[].name	string	Y	이미지 파일명		
2-3	images[].latitude	number	N	위도		32.12345
2-4	images[].longitude	number	N	경도		127.12355
2-5	images[].time	string	Y	촬영시간		hh:mm:ss
2-6	images[].heading	number	N	IMU heading	0~360	nnn
2-7	images[].pitch	number	N	IMU pitch	-999~999	nnn
2-8	images[].roll	number	N	IMU roll	-999~999	nnn
2-9	images[].objects[]	array	Y	객체 정보		
2-9-1	images[].objects[].position	number	Y	라벨링 데이터 원점	0~999	x, y, z
2-9-2	images[].objects[].scale	number	Y	객체 크기	0~999	x, y, z
2-9-3	images[].objects[].rotation	number	Y	객체 회전각	0~999	x, y, z
2-9-4	images[].objects[].area	number	Y	객체 면적		
2-9-5	images[].objects[].size	enum	Y	객체 크기 분포	"S", "M", "L"	"S", "M", "L"
2-9-6	images[].objects[].obj_type	enum	Y	객체 명		트랙터
2-9-7	images[].objects[].obj_id	number	N	객체 분류 코드		null

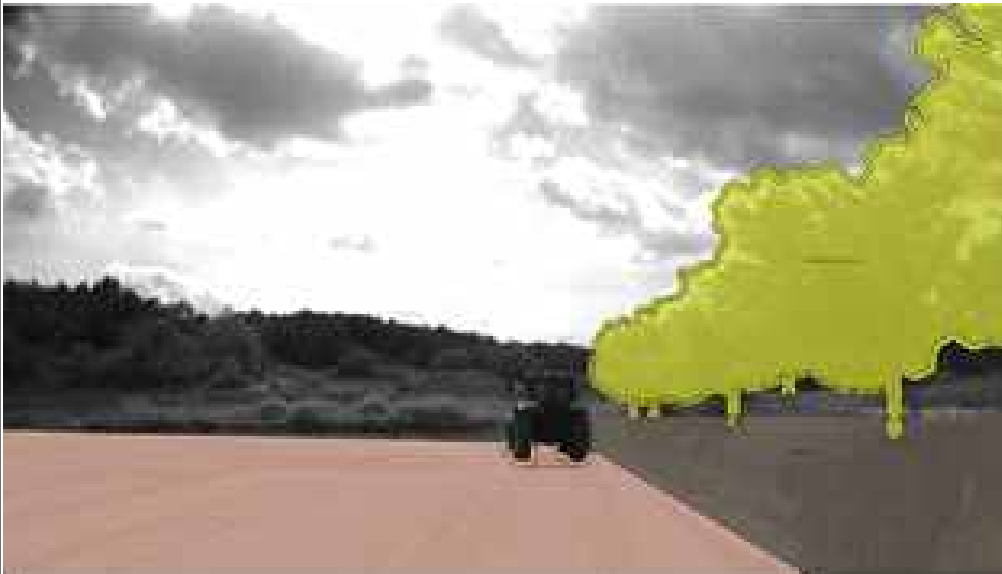
- polygon 라벨링 데이터 구조

구분	항목명	타입	필수여부	설명	범위	비고
1	mode	string	N	가공모드		
2	images[]	array	Y	이미지 정보		

2-1	images[].image_id	number	Y	이미지 고유 ID		
2-2	images[].name	string	Y	이미지 파일명		
2-3	images[].width	string	Y	이미지 너비	0~1920	1920
2-4	images[].height	string	Y	이미지 높이	0~1080	1080
2-5	images[].latitude	number	N	위도		32.12345
2-6	images[].longitude	number	N	경도		127.12355
2-7	images[].time	string	Y	촬영시간		hh:mm:ss
2-8	images[].heading	number	N	IMU heading	0~360	nnn
2-9	images[].pitch	number	N	IMU pitch	-999~999	nnn
2-10	images[].roll	number	N	IMU roll	-999~999	nnn
2-11	images[].objects[]	array	Y	객체 정보		
2-11-1	images[].objects[].label	enum	Y	객체 명		농로
2-11-2	images[].objects[].type	string	Y	가공 타입		폴리곤
2-11-3	images[].objects[].position[]	array	Y	여노테이션 좌표		[x1,y1,x2,y2...xn,yn]
2-11-4	images[].objects[].area	number	Y	객체 면적		
2-11-5	images[].objects[].size	enum	Y	객체 크기 분포	"S", "M", "L"	"S", "M", "L"
2-12	images[].occluded	number	Y	가림 및 잘림 여부		0: 기본 1: 가림
2-13	images[].z_order	number	N	객체 원근 순서		
2-14	images[].attributes[]	array	N	라벨 속성		

3 라벨링 데이터 실제예시

- polygon 라벨링 데이터 예시

구분	세부내용
라벨링 이미지 (JPG)	
여노테이션	{

	결과 (JSON)	<pre> "mode": "annotation", "images": [{ "image_id": "0", "name": "NIA_AgricultureAD_paddy_Gyeongsangnamdo_2107291234_day_sunny_000300.jpg", "width": "1920", "height": "1080", "latitude": 36.12345, "longitude": 127.12345, "time": 174106, "heading": 128, "pitch": 004, "roll": 359 "objects": [{ "label": "common_road", "type": "polygon", "position": [[523.4, 1080.0, 685.8, 981.3, 866.2, 867.0, 932.4, 824.8, 963.85, 807.04, 988.5, 807.2, 1002.8, 807.5, 1018.3, 807.8, 1031.1, 808.2, 1140.71, 845.06, 1308.3, 905.26, 1558.8, 997.9, 1813.1, 1080.0]], "occluded": 0, "z_order": 1, "attributes": [] }] } </pre>
--	--------------	---

	<pre> } } }</pre>

- box 라벨링 데이터 예시

구분	세부내용
라벨링 이미지 (JPG)	
어노테이션 결과 (JSON)	<pre> { "licenses": [], "info": { "contributor": "Testworks", "date_created": "2021-07-21 15:26:28", "description": "NIA Agricultural Autonomous Driving Data Annotations", "url": "https://board.blackolive.co.kr/", "version": 1.0, "year": 2021 }, "images": [{ "id": 1, "name": "NIA_AgricultureAD_paddy_Gyeongsangnamdo_2107291234_day_sunny_000300.jpg", "width": 1920, "height": 1080, "latitude": 36.12345, "longitude": 127.12345, "time": 174106, "Roll": -002, "Pitch": 004, "Yaw": 359 }] },]</pre>

```

"categories": [
  {
    "id": 1,
    "name": "road",
    "supercategory": "common"
  },
  {
    "id": 2,
    "name": "vail",
    "supercategory": "common"
  },
  {
    "id": 3,
    "name": "vehicle",
    "supercategory": "common"
  },
  {
    "id": 4,
    "name": "tractor",
    "supercategory": "common"
  },
  {
    "id": 5,
    "name": "rocks",
    "supercategory": "common"
  },
  {
    "id": 6,
    "name": "tree",
    "supercategory": "common"
  },
  {
    "id": 7,
    "name": "person",
    "supercategory": "common"
  },
  {
    "id": 8,
    "name": "pole",
    "supercategory": "common"
  },
  {
    "id": 9,
    "name": "road",
    "supercategory": "orchard"
  },
  {
    "id": 10,
    "name": "tree",
    "supercategory": "orchard"
  }
]

```

		<pre> "id": 11, "name": "apple", "supercategory": "orchard" }, { "id": 12, "name": "tree_suport", "supercategory": "orchard" }, { "id": 13, "name": "furrow", "supercategory": "filed" }, { "id": 14, "name": "levee", "supercategory": "filed" }, { "id": 15, "name": "bean", "supercategory": "filed" }, { "id": 16, "name": "garlic", "supercategory": "filed" }, { "id": 17, "name": "onion", "supercategory": "filed" }, { "id": 18, "name": "befor_driving", "supercategory": "paddy" }, { "id": 19, "name": "after_driving", "supercategory": "paddy" }, { "id": 20, "name": "water", "supercategory": "paddy" }, { "id": 21, "name": "edge", </pre>
--	--	---

	<pre> "supercategory": "paddy" }, { "id": 22, "name": "rice", "supercategory": "paddy" }], "annotations": [{ "id": 1, "image_id": 1, "category_id": 1, "bbox": [x1,y1,width,height] "iscrowd": 0, "area": 95468.98, "occluded": 0 }] }] } } </pre>
--	--

- cuboid 라벨링 데이터 예시

구분	세부내용
라벨링 이미지 (PCD)	

	어노테이션 결과 (JSON) <pre> { "mode": "annotation", "images": [{ "image_id": "0", "name": "NIA_AgricultureAD_paddy_Gyeongsangnamdo_2107291234_day_sunny_000300.pcd", "width": "1024", "height": "32", "points": "32768", "latitude": 36.12345, "longitude": 127.12345, "time": 174106, "heading": -002, "pitch": 004, "roll": 359 }], "objects": [{ "position": { "x": 17.436717337784955, "y": 0.18087162307816662, "z": 0.23347608639275436 }, "scale": { "x": 2.2772472174421154, "y": 1.4666516540942975, "z": 2.3923818745906233 }, "rotation": { "x": 0, "y": 0, "z": 1.7540558982543013 }, "obj_type": "tractor_common", "obj_id": "" }] } </pre>
--	---

데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)	메일주소	담당업무
		(주)대동	감병우	02-3470-7351	bwgam@daedong.co.kr	사업관리, 수집
	참여기관	기관명	담당업무	기관명	담당업무	
		유엔젤(주)	AI모델개발	한국자동차연구원	수집	
		(주)인터마인즈	검수	국립농업과학원	수집	
		(주)테스트웍스	정제/가공			