

	분야	데이터 유형	구축 데이터량	원천데이터 형식	라벨링 형식	라벨링 유형
메타데이터 정보 (다중기입가능)	NIA 기입	이미지	<ul style="list-style-type: none"> o 101. 정밀농업 노지작물 데이터 <ul style="list-style-type: none"> - 학습용 데이터 672,000개 이상 - 토양 및 환경센서 데이터 144,000set 이상 - 적외선 데이터 128,000개 이상 - 다분광 데이터 32,000개 이상 o 102. 정밀농업 농기계 잡초 인식 데이터 <ul style="list-style-type: none"> - 학습용 데이터 358,000개 이상 - 토양 및 환경센서 데이터 120,000set 이상 - 적외선 데이터 89,000개 이상 	jpg	json	세그멘테이션 (이미지/동영상)
	데이터 출처	데이터 구축년도	구축기관(총괄)	가공기관	검수기관	
	자체 수집	2021년	전라북도	(주)업데이터	(주)캠틱종합기술원 디에스엔전주(주)	
	데이터 문의처	기관명	문의담당자명	전화번호 (유선전화번호기입)	메일주소	
		(주)업데이터	정진우	031-754-9985	jw16@updatergroup.com	
	데이터 소개	<ul style="list-style-type: none"> o 101. 정밀농업 노지작물 데이터 <ul style="list-style-type: none"> - 정밀농업에 필요한 노지작물 생육 데이터를 확보하는 것을 목적으로, 작물 10종에 대해 지상에서 촬영하는 2D RGB 이미지 데이터, 다분광 드론 촬영 데이터, 적외선 촬영 온도값, 토양 및 환경 센서 데이터를 통합적으로 융합한 학습데이터 구축 o 102. 정밀농업 농기계 잡초 인식 데이터 <ul style="list-style-type: none"> - 농업 노동력 부족 문제로 잡초를 인식·탐지하여 제거하는 농기계 개발을 위해, 잡초 14종에 대해 지상에서 촬영하는 2D RGB 이미지 데이터, 적외선 촬영 온도값, 토양 및 환경 센서 데이터를 통합적으로 융합한 학습데이터를 구축 				
	주요키워드	정밀농업, 노지작물, 생육, 잡초, 탐지, 학습데이터				
카테고리 정의서	첨부의 카테고리 정의서 엑셀파일에 데이터카테고리 작성하여 제출					

데이터셋명	국문영문	101. 정밀농업 노지작물 데이터
		101. Precision agricultural Outdoor Crop Data

구축목적 주요 작물10종을 대상으로 2D RGB 데이터(환경 데이터/적외선/다분광 데이터 연계된 메타데이터 포함)를 구축하여 정밀농업에 활용될 수 있는 AI 학습데이터 구축

학습모델 개발

- 인공지능 기반 노지작물 탐지 및 분류 모델
- 인공지능 기반 노지작물 생육 분류 예측 모델

소개

- 정밀농업에 필요한 노지작물 생육 데이터를 확보하는 것을 목적으로, 작물 10종에 대해 지상에서 촬영하는 2D RGB 이미지 데이터, 다분광 드론 촬영 데이터, 적외선 촬영 온도값, 토양 및 환경 센서 데이터를 통합적으로 융합한 학습데이터셋

드론	
데이터수집도구	다분광
수집주기	주 1회(작물 당)
수집방법	고도별 2단계 촬영
기타	촬영시의 대기 기상환경 및 비활정보 추가



토양/대기 환경			
성장환경 수집	온도	습도	일사량
토양환경 수집	토양 온도	토양수분	전기전도도
수집주기	10분 주기		
기타	재배 장소 및 시간 연계취득		

생육정보		
데이터획득내용	2D	적외선
수집주기	주 5회(작물 당)	
수집방법	방향과 각도, 대상에 따라 촬영(작물 불 상이)	
기타	센서 데이터와 연계	

드론 + 토양 + 대기환경 + 생육정보

1. 학습용 데이터-2D RGB 이미지 데이터-구축 규모

품종	생육단계	객체 촬영 범위		합	비율
		전체	부분		
벼(조생종)	3	4,050	16,841	20,891	2.92%
	1	3,721	10,267	13,988	1.96%
벼(중만생종)	2	1,653	3,583	5,236	0.73%
	3	7,915	39,415	47,330	6.62%
콩(풍산)	1	8,013	56,869	64,882	9.07%
	2	750	5,911	6,661	0.93%
	3	1,605	18,964	20,569	2.88%
콩(대풍2호)	1	3,762	41,473	45,235	6.32%
	2	2,078	9,064	11,142	1.56%
	3	4,317	25,413	29,730	4.16%
옥수수(일미찰)	1	2,139	26,414	28,553	3.99%
	2	1,318	6,401	7,719	1.08%
	3	3,483	14,618	18,101	2.53%
옥수수(광평옥)	1	7,087	73,549	80,636	11.27%
	2	1,220	6,268	7,488	1.05%
	3	4,413	17,490	21,903	3.06%
감자	1	1,154	35,035	36,189	5.06%
	2	5,757	13,990	19,747	2.76%
	3	2,175	11,681	13,856	1.94%

품종	생육단계	객체 촬영 범위		합	비율
		전체	부분		
고구마	1	1,557	45,387	46,944	6.56%
	2	1,180	16,361	17,541	2.45%
	3	1,660	3,242	4,902	0.69%
참깨	1	34,338	733	35,071	4.90%
	2	2,406	2,001	4,407	0.62%
	3	2,459	1,883	4,342	0.61%
들깨	1	9,411	29,375	38,786	5.42%
	2	2,610	50,287	52,897	7.40%
	3	3,134	7,298	10,432	1.46%
총합계		125,365	589,813	715,178	100%

※총 구축 목표인 672,000개의 106%인 715,178개 학습데이터 구축

2. 학습용 데이터 품종별 분포



3. 학습용 데이터 이외 추가 데이터 구축 규모

추가 데이터 종류	구축 목표	실제 구축 규모
토양 및 환경센서데이터	144,000 set	158,200 set
적외선 데이터	128,000 개	133,911 개
다분광 데이터	32,000 개	83,448 개

데이터셋 구성

1. 객체 클래스 정의

- 객체의 품종명

품종명	품종코드	품종명	품종코드
벼(조생종)	AA	옥수수(광평옥)	CB
벼(중만생종)	AB	감자	DA
콩(풍산)	BA	고구마	EA
콩(대풍2호)	BB	참깨	FA
옥수수(일미찰)	CA	들깨	GA

- 객체의 품종별 생육단계

작물 품종	생육단계		
	01	02	03
벼 (조생종, 중만생종)	- 열매가 맺기 전까지 단계	다음 중 하나에 해당하는 경우 - 열매가 맺힌 후 고개를 90도 이상 숙이기 전까지 - 열매가 맺힌 후 열매 또는 잎이 변색되기 전까지	다음 중 하나에 해당하는 경우 - 열매가 익어가며 고개를 90도 이상 숙인 후 수확까지 - 열매 또는 잎이 변색되기 시작한 후 수확까지
콩 (풍산, 대풍2호)	- 발아 후 곁가지가 생성되기 전까지 단계	- 곁가지가 생성되어 잎 또는 열매가 황변되기 전까지 단계	- 잎 또는 열매가 황변되기 시작한 후 수확까지
옥수수 (일미찰, 광평옥)	- 꽃봉오리(꽃대)나 열매가 생기기 전까지 단계	- 꽃봉오리(꽃대)가 생성된 후 열매*가 맺히기 전까지 단계 * 열매는 옥수수 수염이 보이는 경우 맺힌 걸로 판단	- 열매*가 맺힌 후 수확까지 * 수확기에 이르러 옥수수 잎 및 열매가 말라가면서 수염이 빠진 경우를 포함
감자	- 전체 객체가 멸칭 (또는 봉분) 내부에 존재할 때까지 단계 * 객체가 비닐 포장(또는 봉분) 끝 라인을 벗어나지 않는 경우 멸칭(또는 봉분) 내부에 존재하는 것으로 판단 * 외부 요인으로 인해 일부가 갈변되더라도 전체가 멸칭 (또는 봉분) 내부에 존재하는 경우 1단계로 분류	- 전체 객체가 멸칭(또는 봉분)을 벗어난 시점부터 잎의 갈변이 시작되기 전까지	- 잎의 갈변이 시작된 시점부터 수확까지
고구마	- 객체의 뿌리에서 나오는 줄기가 육안으로 확인 가능할 때까지 단계	- 전체 객체가 멸칭 내부*에 존재할 때까지 단계 * 감자의 멸칭 내부 판단 기준 준용	- 전체 객체가 멸칭을 벗어나 무성한 단계
참깨	- 꽃이 피기 전까지 단계 * 흰색 꽃잎이 보이는 경우 꽃이 핀 것으로 판단	다음 중 하나에 해당하는 경우 - 꽃이 피어있고 열매가 맺히기 전까지 단계 - 꽃이 피어있고 열매가 맺혀 있는 단계	- 꽃이 지고 열매가 맺혀있는 단계
들깨	- 발아 후 꽃봉오리가 생성되기 전까지	- 꽃봉오리가 생성된 이후 꽃봉오리 또는 잎의 변색이 시작되기 전까지	- 꽃봉오리 또는 잎이 변색되는 시점부터 수확까지

2. 라벨링데이터 구성

구분	항목명	설명	타입	필수여부	범위
1. info-기본 정보					
1-1	description	데이터셋 이름	string	Y	crop_data
1-2	url	데이터셋 제작자 url	string	Y	https://www.jeonbuk.go.kr/
1-3	version	제작버전	string	Y	v.1.0
1-4	year	제작년도	number	Y	2021
1-5	type	데이터셋 타입	string	Y	jpg
1-6	img_path	이미지데이터 폴더 경로	string	Y	/품종명/img
1-7	label_path	라벨링데이터 폴더 경로	string	Y	/품종명/json

구분	항목명	설명	타입	필수여부	범위
2. collection - 수집 정보					
2-1	obj_num	객체 번호(위치)	string	Y	001~999
2-2	tod_attribute	적외선 촬영 장비 속성	string		FLIR C3X
2-3	tod_temper	객체 표면 온도	number		-20 ~ 300
2-4	spec_way	다분광 촬영 방식	string		drone
2-5	spec_attribute	다분광 촬영 장비 속성	string		Micasense RedEdge MX Dual
2-6	spec_ndvi	정규 식생 지수	number		0~9
2-7	soil_ec	토양 전기전도도(EC)	number	Y	0~10
2-8	soil_temper	토양 온도	number	Y	0~60
2-9	soil_humidity	토양 습도	number	Y	0~100
2-10	soil_potential	토양 수분장력	number	Y	0 ~ -1000000
2-11	temperature	온도	number	Y	-40 ~ 124
2-12	humidity	습도	number	Y	0~100
2-13	sunshine	일사량	number	Y	0~2000
2-14	sunrise_time	일출시간	string	Y	04:00~09:59
2-15	sunset_time	일몰시간	string	Y	16:00~22:59
2-16	img_attribute	촬영 장비 속성	string	Y	SM-G965N 등
2-17	img_time	촬영일시	string	Y	YYYY-MM-DD
2-18	img_part	작물 촬영 부위	string	Y	[W, X, Y]
2-19	img_dist	촬영 거리(cm)	number		1~1000
2-20	img_angle	촬영 각도	number		0~360
3. licenses - 저작권 정보					
3-1	licenses_id	라이선스 고유 번호	number		1
3-2	licenses_name	라이선스 이름	string		crop01
4. images - 이미지 데이터 정보					
4-1	img_file_name	원천데이터 이름(파일명)	string	Y	예)20210803_RGB_AA_1_X_01_07.jpg
4-2	img_height	세로	number	Y	1080~4800
4-3	img_width	가로	number	Y	1080~4800
5. annotations - 어노테이션 정보					
5-1	segmentation	segmentation 좌표	string	Y	예)[2.14474345,114.3444 656]
5-2	area	어노테이션 면적	number	Y	0~ 23040000
5-3	bbox	bounding box 정보	string	Y	예)[0.0, 0.0, 440.0, 273.0]
5-4	crops_kind	작물 품종명	string	Y	[A, B, C, D, E, F, G]
5-5	crops_detail	작물 세부 품종명	string	Y	[AA, AB, BA, BB, CA, CB, DA, EA, FA, GA]
5-6	crops_step	객체 생육단계	string	Y	01~03

3 라벨링 데이터 실제예시

```

JSON
{
  "info": {
    "description": "crop_data",
    "url": "https://www.jeonbuk.go.kr/",
    "version": "v.1.0",
    "year": 2021,
    "type": "jpg",
    "img_path": "/감자/img",
    "label_path": "/감자/json"
  }
}

```

```

collection
├── obj_num : "006"
├── tod_attribute : null
├── tod_temper : null
├── spec_way : null
├── spec_attribute : null
├── spec_ndvi : null
├── soil_ec : 1.14
├── soil_temper : 21.96
├── soil_humidty : 55.44
├── soil_potential : -0.1
├── temperature : 21.7
├── humidity : 91.01
├── sunshine : 95
├── sunrise_time : "05:56"
├── sunset_time : "18:49"
├── img_attribute : "Galaxy Note20 5G"
├── img_time : "2021-09-06"
├── img_part : "W"
├── img_dist : 100
├── img_angle : 0
├── licenses
│   ├── licenses_id : 1
│   └── licenses_name : "crop01"
├── images
│   ├── img_file_name : "20210906_RGB_DA_006_1_00_06.jpg"
│   ├── img_height : 4000
│   └── img_width : 3000
├── annotations
│   ├── segmentation
│   │   └── area : 1066367
│   └── bbox
│       ├── crops_kind : "D"
│       ├── crops_detail : "DA"
│       └── crops_step : "01"

```

데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)	메일주소	담당업무
		전라북도	조광희	063-280-2790	jokwhi@korea.kr	업무 총괄
	참여기관	기관명	담당업무	기관명	담당업무	
		전라북도농업 기술원	데이터 검증	(주)디에스엔전주	데이터 전처리	
		국립농업과학원	데이터 수집, 검증	(주)이레아이에스	데이터 수집	
		완주군	잡초 데이터 수집 지원 및 지문	(유)호정솔루션	데이터 수집	
		캠틱종합기술원	클라우드소싱 인력 관리	(주)엔에이치 네트웍스	수집데이터 검사	
(주)업데이터	데이터 가공, 모델 개발					

데이터셋명	국문영문	102. 정밀농업 농기계 잡초인식 데이터
		102. Precision agricultural Weeds Data

구축목적	잡초를 인식 탐지하기 위한 학습 데이터로써 향후 관련 농기계에 적용할 수 있는 데이터 구축
------	--

학습모델 개발	- 인공지능 기반 잡초 탐지 및 분류 모델 개발
---------	----------------------------

- 농업 노동력 부족 문제로 잡초를 인식·탐지하여 제거하는 농기계 개발을 위해, 잡초 14종에 대해 지상에서 촬영하는 2D RGB 이미지 데이터, 적외선 촬영 온도값, 토양 및 환경 센서 데이터를 통합적으로 융합한 학습 데이터셋

토양/대기 환경			
성장 환경 수집	온도	습도	일사량
토양 환경 수집	토양 온도	토양 습도	토양 수분 장력
수집 주기	10분 주기		
기타	재배 장소 및 시간 연계 취득		

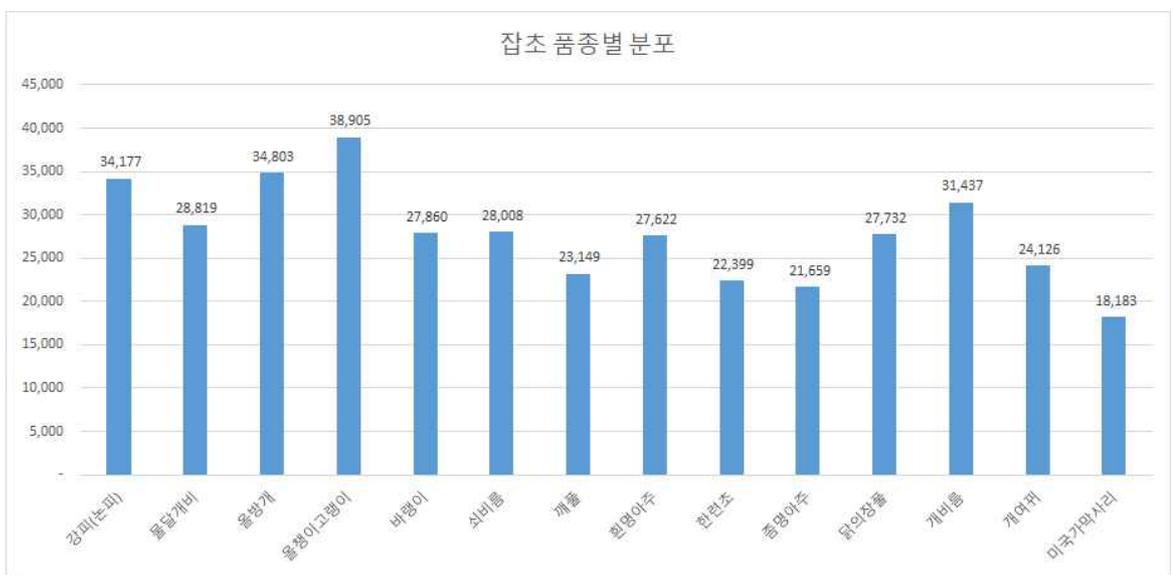


생육정보		
데이터 획득 내용	2D	적외선
수집 주기	주 5회(작물 당)	
수집 방법	방향과 각도 대상에 따라 촬영(작물 별 상이)	

1. 학습용 데이터-2D RGB 이미지 데이터-구축 규모

품종	구축 수량	비율	품종	구축 수량	비율
강피(논피)	34,177	9%	흰명아주	27,622	7%
물달개비	28,819	7%	한련초	22,399	6%
올방개	34,803	9%	좀명아주	21,659	6%
올챙이고랭이	38,905	10%	닭의장풀	27,732	7%
바랭이	27,860	7%	개비름	31,437	8%
쇠비름	28,008	7%	개여뀌	24,126	6%
깨풀	23,149	6%	미국가막사리	18,183	5%
총 구축 목표	358,000 개		실제 구축 규모	388,879 개	

2. 학습용 데이터 품종별 분포



데이터셋
통계
(구축 규모
및 분포)

3. 학습용 데이터 이외 추가 데이터 구축 규모

추가 데이터 종류	구축 목표	실제 구축 규모
토양 및 환경센서데이터	120,000 set	170,205 set
적외선 데이터	89,000 개	107,579 개

1. 객체 클래스 정의

- 객체의 품종명

품종명	품종코드	품종명	품종코드
강피(논피)	A	흰명아주	H
물달개비	B	한련초	I
올방개	C	좀명아주	J
올챙이고랭이	D	닭의장풀	K
바랭이	E	개비름	L
쇠비름	F	개여뀌	M
깨풀	G	미국가막사리	N

2. 라벨링데이터 구성

구분	항목명	설명	타입	필수여부	범위
1. info-기본 정보					
1-1	description	데이터셋 이름	string	Y	weed_detection
1-2	url	데이터셋 제작자 url	string	Y	https://www.jeonbuk.go.kr/
1-3	version	제작버전	string	Y	v.1.0
1-4	year	제작년도	number	Y	2021
1-5	type	데이터셋 타입	string	Y	jpg
1-6	img_path	이미지데이터 폴더 경로	string	Y	/품종명/img
1-7	label_path	라벨링데이터 폴더 경로	string	Y	/품종명/json
2. collection - 수집 정보					
2-1	obj_num	객체 번호(위치)	string	Y	001~999
2-2	tod_attribute	적외선 촬영 장비 속성	string		FLIR C3X
2-3	tod_temper	객체 표면 온도	number		-20 ~ 300
2-4	soil_ec	토양 전기전도도(EC)	number	Y	0~10
2-5	soil_temper	토양 온도	number	Y	0~60
2-6	soil_humidity	토양 습도	number	Y	0~100
2-7	soil_potential	토양 수분장력	number	Y	0 ~ -1000000
2-8	temperature	온도	number	Y	-40 ~ 124
2-9	humidity	습도	number	Y	0~100
2-10	sunshine	일사량	number	Y	0~2000
2-11	sunrise_time	일출시간	string	Y	04:00~09:59
2-12	sunset_time	일몰시간	string	Y	16:00~22:59
2-13	img_attribute	촬영 장비 속성	string	Y	SM-G965N 등
2-14	img_time	촬영일시	string	Y	YYYY-MM-DD
2-15	img_part	잡초 촬영 범위	string	Y	[whole, part]
2-16	img_dist	촬영 거리(cm)	number		1~1000

데이터셋
구성

2-17	img_angle	촬영 각도	number		0~360
3. licenses – 저작권 정보					
3-1	licenses_id	라이선스 고유 번호	number		1
3-2	licenses_name	라이선스 이름	string		weed01
4. images – 이미지 데이터 정보					
4-1	img_file_name	원천데이터 이름(파일명)	string	Y	예:20210803_RGB_A_0_001_05.jpg
4-2	img_height	세로	number	Y	1080~4800
4-3	img_width	가로	number	Y	1080~4800
5. annotations – 어노테이션 정보					
5-1	segmentation	segmentation 좌표	string	Y	예: [2.14474345,114.3444656]
5-2	area	어노테이션 면적	number	Y	0~ 23040000
5-3	bbox	bounding box 정보	string	Y	예: [0.0, 0.0, 440.0, 273.0]
5-4	weeds_kind	잡초명	string	Y	[A, B, C, D, E, F, G, H, I, J, K, L, M, N]

3 라벨링데이터 실제예시

```

JSON
├── info
│   ├── description : "weed_detection"
│   ├── url : "https://www.jeonbuk.go.kr/"
│   ├── version : "v.1.0"
│   ├── year : 2021
│   ├── type : "jpg"
│   ├── img_path : "/강피/img"
│   └── label_path : "/강피/json"
├── collection
│   ├── obj_num : "012"
│   ├── tod_attribute : null
│   ├── tod_temper : null
│   ├── soil_ec : 1.03
│   ├── soil_temper : 22.62
│   ├── soil_humidty : 75.84
│   ├── soil_potential : -0.2
│   ├── temperature : 24.61
│   ├── humidity : 76.93
│   ├── sunshine : 78
│   ├── sunrise_time : "05:39"
│   ├── sunset_time : "19:17"
│   ├── img_attribute : "Galaxy Note20 5G"
│   ├── img_time : "2021-08-17"
│   ├── img_part : "whole"
│   ├── img_dist : 25
│   └── img_angle : 45

```

```

{
  licenses: {
    licenses_id: 1,
    licenses_name: "weed01"
  },
  images: {
    img_file_name: "20210817_RGB_A_012_45_13.jpg",
    img_height: 1836,
    img_width: 1836
  },
  annotations: {
    segmentation: {
      area: 115037
    },
    bbox: {
      weeds_kind: "A"
    }
  }
}

```

데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)	메일주소	담당업무
		전라북도	조광희	063-280-2790	jokwhi@korea.kr	업무 총괄
	참여기관	기관명	담당업무	기관명	담당업무	
		전라북도농업 기술원	데이터 검증	(주)디에스엔전주	데이터 전처리	
		국립농업과학원	데이터 수집, 검증	(주)이레아이에스	데이터 수집	
		완주군	잡초 데이터 수집 자원 및 자문	(유)호정솔루션	데이터 수집	
		캠텍종합기술원	클라우드소싱 인력 관리	(주)엔에이치 네트웍스	수집데이터 검사	
	(주)업데이터	데이터 가공, 모델 개발				