

21년도 인공지능 학습용 데이터 구축 가이드라인

< 고품질 과수작물 통합 데이터 >

인공지능 데이터 구축	사업 총괄	 GDS CONSULTING global data service		
	데이터 설계	 GDS CONSULTING global data service		
	데이터 수집 및 정제	 GDS CONSULTING global data service	 제주대학교 JEJU NATIONAL UNIVERSITY	
	데이터 가공	 GDS CONSULTING global data service		
	데이터 검수	 DATA WAY		
	클라우드 소싱	 GDS CONSULTING global data service	 DATA WAY	 와이비에스에듀
	AI모델 개발	 DSTI DATA SCIENCE TECHNOLOGY INSTITUTE		
	홍보	 JTP 제주테크노파크 JEJU TECHNOPARK		
가이드라인 작성	지디에스컨설팅그룹	박경채		
	데이터웨이	박준식		
가이드라인 버전	ver 1.0 ('22. 1. 15)			

목 차

1. 데이터 명세 정보 1

1.1 데이터 정보 요약 1

1.2 데이터 포맷 1

1.3 어노테이션 포맷 5

1.4 데이터 구성 7

1.5 데이터 통계 9

1.6 원시데이터 특성 11

1.7 기타정보 11

2. 데이터 구축 가이드 12

2.1 데이터 구축 개요 12

2.2 문제정의 13

2.3 수집·정제 14

2.4 어노테이션/라벨링 25

2.5 검수 29

2.6 활용 35

1. 데이터 명세 정보

1.1 데이터 정보 요약

데이터 이름	고품질 과수작물 통합 데이터	
활용 분야	과수작물(감귤,키위) 병충해 방제 및 품질향상을 위한 인공지능 학습	
데이터 요약	고품질 과수작물에 대한 최적의 생육조건 및 수확시기를 인공지능 기술을 통해 예측하고, 병충해 방제 및 최적의 생육 상태를 유지하기 위한 정보를 제공하며 이를 위해 고품질 과수작물 생육 및 환경 통합 데이터 구축	
데이터 출처	병충해 이미지와 초분광 이미지는 제주도 소재 22개 감귤, 키위 농가에서 크라우드 워커 인력이 직접 수집하여 촬영한 데이터 이며, 환경데이터는 제주대학교에서 보유한 자동기상관측장비(AWS) 20곳으로부터 기상 정보를 수집하고, 기상청 데이터는 공개된 데이터를 활용	
데이터 이력	배포버전	V 1.0
	개정이력	
	작성자/ 배포자	(주)지디에스컨설팅그룹/㈜지디에스컨설팅그룹

1.2 데이터 포맷

□ 구축데이터 요약

구분		데이터 설명	데이터 포맷
고품질 과수작물 통합 데이터	이미지	- 과수작물(감귤/키위) 병충해 이미지 촬영 및 Annotation - 초분광 이미지 촬영 및 Annotation	JPG, HDR, JSON
	환경 데이터	환경데이터 총 200,000건 이상	JSON

□ 구축데이터 명세

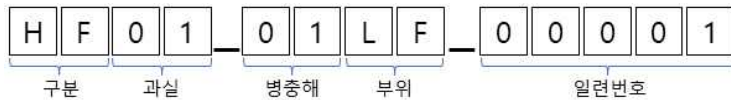
항목	내용
이미지 데이터	- 과수작물 과실 병충해 이미지
	- 과수작물 초분광 이미지
기상환경 데이터	- 장비로부터 수집된 환경 데이터 - 수치모델 데이터와 결합
메타데이터	- 이미지 파일 정보 - 어노테이션 정보 - 환경 정보

□ 파일 포맷

- 어노테이션 포맷은 JSON 파일 형태의 메타데이터로 구성되어 특정 소프트웨어에 독립적으로 구축 데이터셋을 활용할 수 있도록 함
- 과수작물의 원천 이미지 데이터는 다양한 연구에 활용도를 높이기 위하여 JPG 포맷으로 저장하여 제공

□ 파일명명 규칙

- 파일아이디는 총 13자리이며, 타이틀, 과실분류, 병충해분류, 부위분류, 일련번호 등의 5개의 카테고리로 구분하고 4_4_5 구조로 "_"를 삽입



구분	과실분류	병충해		부위	일련번호
HF (고품질 과수작물)	01 : 감귤 02 : 키위	00 : 정상 01 : 감귤_궤양병 02 : 감귤_굴응애 03 : 감귤_진딧물	04 : 키위_점무늬병 06 : 키위_총채벌레 07 : 키위_과실무름병	FT : 과실 LF : 잎	000001, 000002, ...

<파일아이디 체계>

[파일명명규칙]

1) 병충해 데이터 명명 규칙

□ 과수/병충해/부위별 구분 : 파일명으로 구별

- 자리수 : 1 2 3 4 5 6 7 8 9 A B C D E F 0
- 파일명 : H F 0 0 _ 0 0 X X _ 0 0 0 0 0 0

□ 과수별 구분 : 자릿수 위치 - 3, 4(데이터 형태 : 숫자)

- 자리수 : 1 2 3 4 5 6 7 8 9 A B C D E F 0
- 감 귤 : H F 0 1 _ 0 0 X X _ 0 0 0 0 0 0
- 키 위 : H F 0 2 _ 0 0 X X _ 0 0 0 0 0 0

□ 병충해별 구분 : 자릿수 위치 - 6, 7(데이터 형태 : 숫자)

- 자리수 : 1 2 3 4 5 6 7 8 9 A B C D E F 0
- 정 상 : H F 0 1 _ 0 0 X X _ 0 0 0 0 0 0
- 궤양병 : H F 0 2 _ 0 1 X X _ 0 0 0 0 0 0

정상(00), 궤양병(01), 굴응애(02), 진딧물(03), 점무늬병(04), 총채벌레(06), 과실무름병(07)

□ 부위별 구분 : 자릿수 위치 - 8, 9(데이터 형태 : 알파벳)

- 자리수 : 1
- 과 실 : H
- 앞 : H

2) 초분광 데이터 명명 규칙

□ 초분광 영상 구분 : 자릿수 위치 - B, C (데이터 형태 : 알파벳)

- 자리수 :
- 초분광(감귤) :
- 초분광(키위) :

3) 기상환경 데이터 명명 규칙

□ 기상환경 농장코드 구분 : 자릿수 위치 - 6, 7, 8 (데이터 형태 : 알파벳&숫자)

날짜 구분 : 자릿수 위치 - A~1A (데이터 형태 : 날짜)

- 자리수 :
-
기상환경(감귤농장)
:
-
기상환경(키위농장)
:

□ 적용형식 및 예시

파일종류	형식
병충해영상 원시이미지	FileD.jpg
Annotation MetaData	FileD.json
초분광영상 원시이미지	FileD.dat FileD.hdr
초분광영상 라벨링데이터	FileD.json
환경기상정보 MetaData	FileD.csv

1.3 어노테이션 포맷

□ 클래스 설계

- 최상위 대분류는 작물명(감귤, 키위)
- 중분류는 병충해명
예) 정상, 궤양병, 굴응애, 진딧물, 점무늬병, 과실무름병, 총채벌레
- 세부 클래스 체계는 과수명/병충해명으로 구성함

대분류/중분류 (작물코드)	중분류 (병충해코드)	CLASSCODE
감귤(01)	정상 (00)	감귤_정상
	궤양병 (01)	감귤_궤양병
	굴응애 (02)	감귤_굴응애
	진딧물 (03)	감귤_진딧물
키위(02)	정상 (00)	키위_정상
	점무늬병 (04)	키위_점무늬병
	총채벌레 (06)	키위_총채벌레
	과실무름병 (07)	키위_과실무름병

□ 데이터 설계 - 병충해 데이터

구분	속성명	타입	필수여부	설명	범위	비고
1	Info 기본 정보					
1-1	IMAGE_FILE_NM	string	Y	파일명	HF01_00FT_000001~ HF02_07FT_999999, HF01_00LF_000001~ HF02_07LF_999999	
1-2	IMAGE_FILE_TY	string	Y	종류	JPEG	
1-3	STRE_COURS	string	Y	저장폴더 ¹⁾	W1.TrainingW원천 데이터W01.감귤~0 2.키위 W1.TrainingW라벨 링데이터W01.감귤 ~02.키위	
1-4	RSOLTN	string	Y	해상도	병충해영상 : (1920,1080) or (1440,1080)	
1-5	IMAGE_POTOGRF_DT	string	Y	날짜, 시간정보	YYYY:MM:DD HH24:MI:SS	
1-6	CMRA_INFO	string	Y	카메라정보	"Apple","LGE","sam	

구분	속성명	
1-7	F_STOP	
1-8	FILM_SPD	
1-9	FSLT_USE_AT	
1-10	LCINFO	
1-11	IMAGE_OBTAIN_PLA CE_TY	
1-12	GRWH_STEP_CODE	
1-13	DBYHS	
1-14	OCPRD	
1-15	SPCIES_NM	
2	Annotations & 라벨링정보	
2-1	ANTN_ID	
2-2	ANTN_TY	
2-3	OBJECT_CLASS_COD E	
2-4	ANTN_PT	
3	Environment 환경정보 ³⁾	
3-1	ENVRN_INFO_FILE_N M	
3-2	MSRG_DATETM	

구분	속성명	타입	필수여부	설명	범위	비고
3-3	SOLRAD_QY	number	N	일사량	0~2000	
3-4	AFR	number	N	강우	0~600	
3-5	TP	number	N	온도	-30~70	
3-6	HD	number	N	습도	0~105 ⁴⁾	
3-7	SOIL_MITR	number	N	토양수분	0~115 ⁶⁾	

□ 데이터 설계 - 초분광 데이터

구분	속성명	타입	필수여부	설명	범위	비고
1	Info 기본 정보					
1-1	IMAGE_FILE_NM	string	Y		HF01_00FT_HS_000001~HF02_00FT_HS_999999	
1-2	IMAGE_FILE_TY	string	Y	종류	DAT,HDR	
1-3	STRE_COURS	string	N	저장폴더 ⁵⁾	W1.TrainingW원천데이터W03.초분광_감굴~04.초분광_키워 W1.TrainingW라벨링데이터W03.초분광_감굴~04.초분광_키워	
1-4	RSOLTN	string	N	해상도		
1-5	IMAGE_POTOGRF_DT	string	Y	날짜, 시간정보	YYYY:MM:DD HH24:MI:SS	
1-6	CMRA_INFO	string	N	카메라정보		
1-	F_STOP	string	N	조리개값		

1) 파일명/저장폴더 구조 : <별첨4. 파일명명 규칙, 데이터 디렉토리 규칙> 참조

2) 생육단계(9단계) : 논의발달(1), 잎의발달(2), 가지의발달(3), 화총의발달(4), 개화(5), 과실비대(6), 과실성숙(7), 노화기(8), 휴면기(9), 생육Data 획득 대상에만 기록함

3) 환경기상정보 : 측정센서 없음/고장/오작동 및 비대상(선별장 채집의 경우) 등의 사유로 데이터 일부(또는 전체)가 미기록(공백) 될 수 있음.

4) 습도, 토양수분값은 포화상태인 경우 센서가 100을 초과하는 포화값으로 측정될 수 있음.

5) 파일명/저장폴더 구조 : <별첨4. 파일명명 규칙, 데이터 디렉토리 규칙> 참조

6) 생육단계(9단계) : 논의발달(1), 잎의발달(2), 가지의발달(3), 화총의발달(4), 개화(5), 과실비대(6), 과실성숙(7).

구분	속성명	
7		
1-8	FILM_SPD	
1-9	FLSLT_USE_AT	
1-10	LCINFO	
1-11	IMAGE_OBTAIN_PLACE_TY	
1-12	GRWH_STEP_CODE	
1-13	DBYHS	
1-14	OCPRD	
1-15	SPCIES_NM	
2	Annotations & 라벨링정보	
2-1	ANTN_ID	
2-2	ANTN_TY	
2-3	OBJECT_CLASS_CODE	
2-4	BNDBX	
3	Environment 환경정보 ⁷⁾	
3 - 1	ENVRN_INFO_FILE_NM	
3 - 2	MSRG_DATETM	
3 - 3	SOLRAD_QY	
3 - 4	AFR	
3 - 5	TP	
3 - 6	HD	
3 - 7	SOIL_MITR	

구분	속성명	타입	필수여부	설명	범위	비고
4	Hyperspectral 초분광정보					
4-1	HS_NON_BRX	number	N	당도	0~100	
4-2	HS_NON_DM	number	N	건물율	0~100	
4-3	HS_BRX	number	N	정밀검사 당도	0~150	
4-4	HS_DM	number	N	정밀검사 건물율	0~100	
4-5	HS_CHROMA	number	N	정밀검사 색도	50~150	
4-6	HS_HN	number	N	정밀검사 경도	0~20	
4-7	HS_ACIDITY	number	N	정밀검사 산도	0~50	

1.4 데이터 구성

□ 데이터 저장 구조

- 어노테이션 정보는 JSON 파일 형태의 구조체 형식으로 각 대상 객체의 클래스별로 어노테이션 정보를 저장
- 각 이미지 데이터와 일대일 매핑이 될수 있도록 파일명 형식을 갖추어 저장
- 저장 디렉토리는 각 종별로 분류하고 종별로 분류된 저장공간에 이미지, 어노테이션 메타데이터를 구조체 형태로 저장

□ 데이터 저장 디렉토리

1단계	2단계	3단계	파일명
/Data	/라벨링데이터	/01.감귤	HF01_00FT_000001.json ~ HF01_03FT_999999.json
		/02.키위	HF02_00FT_000001.json ~ HF02_07FT_999999.json
		/03.초분광_감귤	HF01_00FT_HS_000001.json ~ HF01_00FT_HS_999999.json

노화기(8), 휴면기(9), 생육Data 획득 대상에만 기록함

7) 환경기상정보 : 측정센서 없음/고장/오작동 및 비대상(선별장 채집의 경우) 등의 사유로 데이터 일부(또는 전체)가 미기록(공백) 될 수 있음.

	/원천데이터	

1.5 데이터 통계

1.5.1 데이터 구축 규모

□ 데이터셋 구축량

구분		데이터 구축 내용		구축 수량 (장)
		데이터 설명	데이터 포맷	
고품질 과수작물 통합 데이터	이미지	- 과수작물(감귤/키위) 병충해 이미지 촬영 및 Annotation - 초분광 이미지 촬영 및 Annotation	JPG, HDR, JSON	40,000장 100,000세트
	환경 데이터	환경데이터 총 200,000건 이상	JSON	200,000건

□ 데이터셋 상세

1단계	2단계	3단계	수량	내용
1.Dataset	1_원천데이터	01.감귤	20,000	○ 이미지 파일(jpg) -정상(HF01_00) : 5,000 -궤양병(HF01_01) : 11,268 -응애(HF01_02) : 1,814 -진딧물(HF01_03) : 1,918
		02.키위	20,000	○ 이미지 파일(jpg) -정상(HF01_00) : 5,000 -점무늬병(HF02_04) : 7,678 -충채벌레(HF02_06) : 5,585 -과실무름병(HF02_07) : 1,737
		03.초분광_감귤	100,000	○ 초분광 영상 파일(dat, hdr)
		04.초분광_키위	100,000	
	2_라벨링데이터	01.감귤	20,000	○ 라벨 데이터(json) -수량은 상동
		02.키위	20,000	
		03.초분광_감귤	50,000	○ 초분광 라벨 데이터(json) -수량은 상동
		04.초분광_키위	50,000	

1.5.2 데이터 분포

□ 병충해 이미지

- 해년마다 병충해 발생 양상이 달라지기 때문에 수집단계에서 병충해 종류와 비율은 수집 결과의 분포확인함.
- AI학습모델 개발 시에는 정상군과의 비교가 필요하기 때문에 정상데이터도 25% 수집
- 알고리즘 학습에 필요한 데이터 수량 확보를 위하여 클래스당 최소 수량 이상이 되도록 수집

○ 감귤 병충해 유형별 분포

분류
정상
궤양병
굴응애
진딧물
소계

○ 키위 병충해 유형별 분포

분류
정상
과실무름병
점무늬병
충채벌레
소계

□ 초분광 이미지

- 감귤 - 생육시기별 분포 확인
 - 비대 단계와 성숙 단계별 확보 후 데이터 분포확인
- 키위 - 품종별 분포 확인
 - 그린키위는 방제기술의 발전으로 인해 거의 병충해가 발생하지 않았고, 레드키위와 골드키위에서 주로 발생
 - 레드키위와 골드키위로 데이터 분포를 확인

1.6 원시데이터 특성

1.6.1 대상분류

과수작물 병충해 이미지 및 초분광 이미지는 실제 형상을 촬영한 실제에 해당

1.6.2 제약조건

- 수집 이미지 촬영을 위한 촬영 방법 및 기준 제시(상세내용 - 2.3.6 수집·정제 기준)

1.6.3 속성

과수작물 병충해 이미지 : 이미지크기 1920*1080 또는 1440*1080

1.7 기타정보

1.7.1 포괄성

- 지역적인 다양성 분포를 고려하여 제주도에 다양한 위치를 산정하여 도내 22개 감귤,키위 농가에서 대상 과수작물의 과실과 잎을 수집하여 촬영

1.7.2 독립성

- 직접 촬영한 데이터로 라이선스 없음
- 협력 농가에 대한 법적 문제발생을 방지하기 위하여, 농가들과 협약서를 체결하고, 개인정보 수집 및 활용 동의를 득함

1.7.3 유의사항

- 개인정보보호를 위한 비식별화를 위하여 개인정보 해당하는 정보(농가명 및 위치정보)를 제거하고 배포

1.7.4 관련 연구

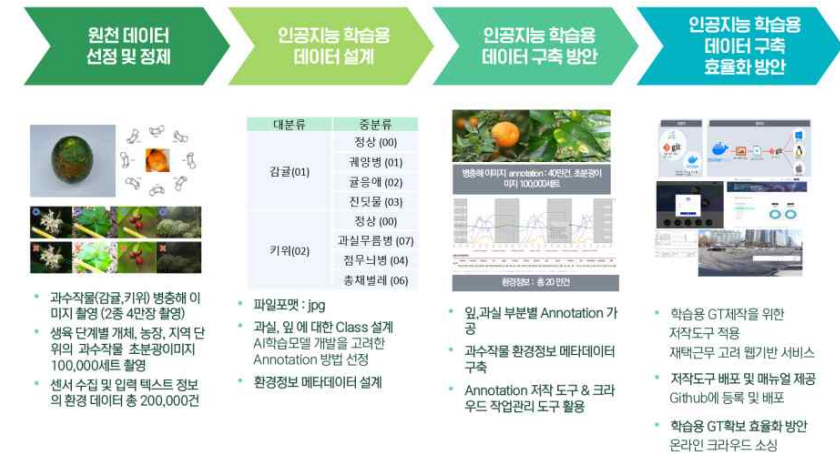
- Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014)
- He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016
- Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017
- Liang, Qiaokang, et al. "PD2SE-Net: Computer-assisted plant disease diagnosis and severity estimation network." Computers and electronics in agriculture 157 (2019): 518-529.

2. 데이터 구축 가이드

2.1 데이터 구축 개요

□ 데이터 구축 절차

- 본 사업의 데이터 구축 절차는 단계별로 구성하며, 각 단계별 세부 절차 내역을 정의함



<과수작물 AI 학습용 데이터 구축 프로세스>

□ 데이터 획득 정제

- 지역적인 다양성 분포를 고려하여 제주도에 전체적인 위치를 산정하여 도내 22개 감귤,키위 농가에서 대상 병충해 과실 수집
- 감귤은 정상, 궤양병, 굴응애, 진딧물 병충해 이미지를 키위는 정상, 점무늬병, 총재벌레, 과실무름병 병충해 이미지 데이터를 수집 획득함
- 환경데이터는 제주도에 20곳으로부터 기상정보를 수집하고, 기상청 데이터는 공개된 데이터를 활용

□ 데이터 가공(어노테이션/라벨링)

- 객체 클래스 및 메타데이터 설계를 통하여 객체 클래스를 분류하고, 코드화 설계함
- 정제 데이터에 학습대상 객체를 어노테이션하고, 객체 분류 및 속성 정보를 입력함

□ 데이터 품질검증

- 정제된 데이터를 대상으로 정제 불완전처리 여부를 점검함
- 어노테이션/레이블 완료 데이터를 대상으로 어노테이션 대상 객체가 누락되거나 잘못된 분류/속성이 입력되었는지 여부를 점검함
- 최종적으로 외부기관(TTA)에 의한 검수를 수행함

2.2 문제정의

2.2.1 임무 정의

- 감귤/키위 과수작물에 대한 병충해를 인식하여 방제 계획을 수립할 수 있는 학습용 데이터 구축
- AI를 통한 과수 품질을 예측하는 과학영농 활성화를 통해 현실적인 문제들을 극복하고 적극적으로 대응할 수 있는 데이터셋 구축
- 과수작물(감귤, 키위) 병충해 이미지를 AI학습을 통해서 분류 판별 할수 있는 AI 모델 개발 및 응용 서비스 개발

2.2.2 데이터 구축 유의사항

- 데이터의 적합성 측면
 - 데이터 획득 대상, 획득방법이 법·제도를 저촉하거나 또는 사회 윤리에 어긋나지 않도록 함
 - 개인정보 및 사생활 보호가 필요한 항목 획득 시, 개인정보보호법 등에 따라 적절한 법적, 기술적 절차를 거친 데이터를 활용하며, 그렇지 않은 데이터는 정제 과정에서 처리될 수 있도록 함
 - 법적 절차 : 개인정보 활용 동의, 초상 활용 동의, 명예훼손 가능성 여부 검토 등
 - 기술적 절차 : 데이터 유형별로 적용할 수 있는 익명처리 기법 적용
 - 협력 농가에 대한 법적 문제발생을 방지하기 위하여, 농가들과 협약서를 체결하고, 개인정보 수집 및 활용 동의를 득함
- 데이터의 무결성측면
 - 작업 환경에서 데이터베이스의 트랜잭션 보장으로 1차적 데이터 무결성 확보
 - 각 이미지 데이터 작업 시에 다수의 작업자가 같은 이미지 작업하는 것을 방지
- 데이터 보안 측면
 - 클라우드 소싱 인력은 비식별화 처리된 원시이미지에 대하여 어노테이션 작업을 진행하도록 작업을 배분하며, 어노테이션 도구를 통하여 작업을 진행함
 - 회원가입 및 인증 후 시스템 내 데이터 접근 허용
 - 작업자 구분별로 접근 권한을 차등 부여하여 관리 할당된 데이터에만 제한적으로 접근 허용
 - 개별 서버에 대한 방화벽 설정으로 서버 데이터로의 직접적인 접근을 막고, 권한이 부여된 특정 URL을 통해서만 접근을 허용함

2.3 수집·정제

2.3.1 원시데이터 선정

□ 병충해 선정

- 어떤 과실이든 해마다 유행하는 병충해가 달라질 수가 있기 때문에 감귤과 키위 각각 병충해 전문가를 섭외하여 병충해를 선정
- 병충해의 특성상 여러 가지 병충해가 고르게 유행하기 보다는 한가지 병충해가 dominant할 경우 다른 병충해의 발생이 억제되는 현상이 있음
- 금년의 경우, 때 이른 장마와 9월에도 비가 많이 내렸던 날씨 영향으로 감귤의 경우 점무늬병의 발생이 많았음
- 키위 병충해 역시 점무늬병의 발생이 우세하였으며, 총채벌레 데이터도 확보하기 용이하였으나, 세 번째로 유행하는 병충해를 찾기가 어려웠음

과실	병충해명
감귤	궤양병
	꿀응애
	진딧물

과실	병충해명	특징
		- 일반적으로 5월 하순부터 발생하기 시작하여 봄순에서는 5월 하순~6월 상순에 그리고 여름순에서는 7월 하순~8월 상순에 밀도가 가장 높는데, 전반적으로 조팝나무 진딧물보다 발생량이 다소 적으며, 특히 여름순에서의 발생이 더 적음 - 식물체에 달라붙어서 즙액을 빨아먹고 기생
키위	점무늬병	- 키위 잎에서 가장 많이 나타나는 병 중의 하나이나 피해는 크지 않음 - 6월 이후부터 가을까지 계속해서 발생하며, 주로 강우가 많은 여름철에 많이 발생 - 처음에는 작은 회갈색의 점무늬가 형성되고, 점차 진전되면서 원형 내지 불규칙한 병반으로 확대 - 오래된 병반상에는 소립점(병자각)이 형성되며, 탄저병과 흡사하여 구분하기가 매우 어려움
	충채벌레	- 암컷성충의 크기는 약 1.4 ~ 1.7mm정도이며, 몸색은 황색에서 갈색으로 다양하고, 더듬이 첫마디는 황색, 둘째 마디는 갈색, 3 ~ 5번째 마디는 황갈색이고 나머지는 갈색 - 수컷성충의 크기는 약 1.0 ~ 1.2mm정도이며 황색. 성충의 날개는 2쌍으로 되어 있으나 잘 발달하지 못하였고, 충채모양을 하고 있어서 잘 날지 못하므로 툭툭 튀듯이 이동 - 피해는 수관의 하부보다는 중앙부 이상의 과실에서 피해가 많은 경향이며, 과실이 피해를 받으면 초기에는 가해받은 부분이 구름 모양의 백색반점이 형성되고, 시간이 경과함에 따라 피해부위가 갈변되어 상품가치를 떨어뜨림
	과실무름병	- 병든 부위에서 병자각 형태로 월동하여 다음해 1차전염원이 됨 - 병자각에서 월동된 병포자는 비바람에 의하여 전반, 과경을 침입하여 갈변시키고, 점차 진전되면 생육 중 또는 저장 중에 연부증상을 일으키는데, 특히 저장중의 피해가 큼 - 담색의 병반으로 나타나고, 주위는 불분명한 농녹색을 띠며, 진전되면 무른 상태로 썩음

※ 출처 : 국가농작물병충해관리시스템

2.3.2 병충해 데이터 수집

□ 병충해 데이터 촬영

- 병충해에 걸린 개체를 채취하여 A4용지 등의 흰색 배경으로 실내 촬영을 진행함
- 일부 병충해는 채취하지 않고 매달린 상태로 촬영을 진행
- 채취하여 촬영하는 경우, 운반 및 보관 이슈를 고려하여 현장에서 직사광선을 차단시키는 촬영세트를 구성하여 실내촬영의 효과를 얻을 수 있었음



<과수작물 병충해 이미지 현장촬영 모습>

- 데이터 수집매뉴얼 작성 : 피사체 전체가 나오도록 정확한 화면 구도와 비율에 맞게 촬영하고, 일관된 이미지 품질 및 수량 범위를 위해 취득 품종별로 높이, 각도, 거리, 조명 등을 매뉴얼로 작성
- 스마트폰에 '오픈 카메라'라는 카메라 전용앱을 설치하여 촬영함으로써, 카메라 설정의 일관성을 확보함 (해상도 1920 X 1080, 종횡비율 4:3, 지오태킹 등)

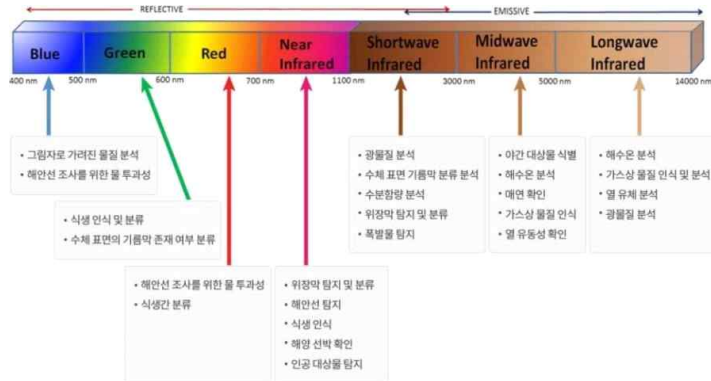


<오픈카메라 앱을 통해 촬영 설정 일관성 확보>

2.3.3 초분광 데이터 수집

□ 초분광 데이터 개요

- 초분광은 수백 개의 분광채널을 통해서 대상물의 특성을 나타내는 연속적인 스펙트럼(spectrum)을 수집하는 센서를 이용한 촬영기법을 의미하며, 각각의 화소는 대상체의 연속적인 분광스펙트럼을 가지고 있음



<초분광 데이터 대역별 활용범위>

- 과수작물에 대한 데이터의 활용목표는 크게 생산량 예측과 품질 예측 2가지로 나눌 수 있음
- 초분광 드론 촬영을 통해 과수원 전체에 대한 생산량 예측을 시도한 사례는 있으나, 이는 초분광 데이터 특정 밴드 데이터로부터 과실을 count하는 방법으로 농민들이 농장을 운영하는 데 크게 도움이 되는 방법은 아님
- 본 과제에서는 개별 과실에 대한 초분광 데이터를 기반으로 품질을 예측하는 것으로 활용 목적을 설정함
- 이를 위해 감귤과 키위 과실에 대해 초분광이미지를 촬영하고 당도 등의 과실 품질을 측정하였음

□ 초분광 카메라 Specification

- 본 과제에서는 총 3대의 초분광 카메라를 사용하였으며, 이중 2대는 본 과제 예산으로 구매하였고, 일정만회를 위해 추가로 1대를 임대하여 활용
- 본 과제 활용목적인 과실 품질 예측에 활용될 수 있도록 400~1,000nm 대역의 카메라를 사용하였음

항 목
Wavelength band
Spectral resolution FWHM
Spatial Sampling
Spectral bands
F/number
Peak SNR
FOV
Viewfinder camera

<초분광카메라 Specification>



<본 과제에서 사용된 초분광카메라>

□ 초분광데이터 촬영



<초분광 데이터 수집 프로세스>

- 초분광카메라는 빛에 매우 민감하기 때문에 현장촬영을 통해서는 품질예측에 활용할 수 있는 고품질의 데이터를 확보하기는 어려워 전량 실내 촬영을 진행

- 농장주에 의해 과실을 획득한 후 클라우드워크 수거팀이 농장을 방문하여 수거하며, 수거 시점에 열매별로 농장, 품종, 획득일시 등을 라벨링
- 클라우드워커 촬영팀에 의해 초분광 촬영을 진행한 후, 비파괴검사와 파괴검사의 방법을 병행하여 과실의 품질을 측정
- 측정된 품질은 각 이미지에 속성으로 라벨링



<초분광데이터 촬영 현장사진>

2.3.4 기상환경 데이터 수집

□ 기상데이터 수집

- 제주대학교에서 보유한 자동기상관측장비(AWS) 20곳으로부터 기상정보를 수집하고, 기상청 데이터는 공개된 데이터를 활용



<제주지역 자동기상관측장비(AWS) : 초록색 제주대, 파란색 기상청>



<제주지역에 설치된 자동기상관측장비(AWS) 현장사진>

□ 기상데이터 수치모델링

- 수집된 AWS 데이터는 FNL/GFS/SST 등의 글로벌 기상자료와 병합하여 WRF 수치

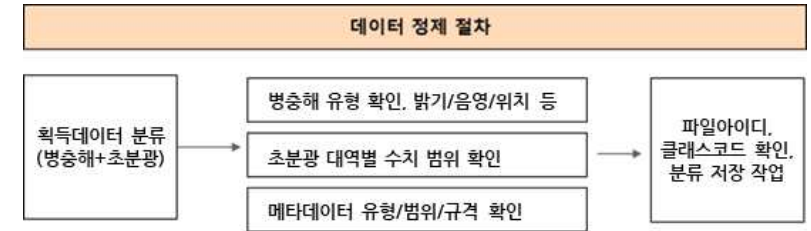
모델링 프로세싱을 거쳐 최종 기상환경 데이터를 구축

※ WRF 모델(Weather Research and Forecasting Model, 중규모 기상 수치 모델)은 미국 국립대기환경연구소를 중심으로 개발된 수치모형으로서 자료동화, 예보 연구, 모델간 결합, 실시간 예보 등에 활용

- 기상데이터 수치모델링은 단기간에 강력한 서버 리소스를 필요로 하는 작업이며, 본 과제에서는 NBP 클라우드 기반으로 진행됨
- WRF 수치모델링에 사용된 데이터는 다음과 같음
 - . AWS : 제주 지역 AWS로부터 수집된 데이터
 - . FNL : 미국 국립대기환경연구소의 Final Global Data Assimilation System 자료
 - . GFS : Global Forecast System 데이터
 - . SST : Sea Surface Temperature (해수면 온도) 데이터
- 1단계 WPS Processing
 - . 초기자료 검토 (AWS데이터, FNL, GFS, SST)
 - . 지형자료를 이용하여 해당 실험 영역의 지형, 지표 정보 생산
 - . 중간단계 파일 생성 : 입력자료 구축을 위한 자료 디코딩
- 2단계 Real Data Initialization
 - . 연직내삽 : 중간단계파일(결과물) 활용
 - . 모델레벨로 변환 -> WRF 초기자료& 경계자료 생산
- 3단계 WRF Processing
 - . 메인 모델링 과정
 - . nc파일형태 -> netCDF 라이브러리 활용
 - . 분석및 Final Data 산출

2.3.5 수집·정제 절차

□ 데이터 정제 절차



- 어노테이션 단계에 들어가기 전에 학습용 데이터로 적합한 데이터를 선별하고 처리하는 정제 프로세스를 획득데이터 유형별로 수행
- 과수작물 데이터는 병충해이미지와 초분광데이터로 구성되어 있어서 데이터 정제 작업도 각 유형에 맞게 진행
- 병충해 데이터는 병충해 유형이 맞는지 확인하고 밝기, 음영, 위치, 해상도 등이 규격에 맞게 수집되었는지 확인
- 또한 과실, 잎의 이미지를 분류하여 정제하고 해당 객체의 차지하는 비율과 이미지 내 객체 외 여백이 과다하게 많은 경우 등을 점검
- 초분광 데이터는 각 대역별로 수치의 이상값 등을 확인
- 기타 AI 학습용 데이터의 정확도의 대부분은 획득데이터의 품질에 좌우되는 만큼 고품질의 원천데이터를 확보할 수 있도록 필요한 정제 작업을 수행

2.3.6 수집·정제 기준

○ 수집 방법 및 기준

촬영 구분	설명
촬영 방법	▪ 촬영 장비 : 스마트폰(또는 초분광카메라) + 삼각대 + 배경지 + 빛가리개 + 조명 ▪ 촬영 대상을 명확하게 보여주도록 촬영 ▪ 촬영이미지 해상도 - 병충해 이미지 : FHD규격(1920x1080, 1440x1080) (※ 학습용 영상은 필요에 의해 크롭, 축소하여 학습용 input data 규격에 맞도록 최적화 함.) ▪ 피사체 전체가 나오도록 정확한 화면 구도와 비율에 맞게 촬영 ▪ 동일한 피사체를 촬영 각도를 다르게 여러장 촬영 ▪ 일관된 이미지 품질 및 수량 범위를 위해 취득 품종별로 촬영

촬영 구분	설명	비고
	영각도, 조명, 빗가림 등의 촬영기준이 표준화된 카메라 앱 활용	
촬영 환경	- 실내/실외 : 병충해 영상은 판별을 위해 대상을 채집하여 실내 촬영을 기본으로 하고 채취가 어려운 부분은 현장 촬영, 초분광 영상은 채집 후 실내 촬영함 - 위치정보(GPS) 활성화 - 부분 과노출을 방지하기 위한 플래시 사용 금지 및 빗가림 판 사용(단, 필요한 경우 인공 조명 사용) - 채집된 객체는 깨끗한 배경에서 촬영 - 피사체 흔들림, 음영 혹은 반사를 최소화하여 촬영 - 객체가 카메라앱의 가이드박스를 벗어나지 않도록 중앙에 배치하여 촬영	

○ 이미지 데이터 촬영 항목 및 획득 방법

항목 구분	항목별 획득(촬영) 방법	비고
열매	- 열매를 현장 또는 채집 후 촬영 - 화면중앙에 열매 전체가 포함되는 구도로 촬영 - 대상 객체의 최대 크기가 화면을 벗어나지 않도록 촬영 - 대상 객체의 최소 크기는 이미지 세로폭의 30% 이상이 되도록 촬영 - 대상 객체 정면, 상하좌우 45도 각도에서 여러장 촬영(구도는 동일) - 해충이 있는 경우 포함하여 촬영* - 구도 : 객체가 중앙에 배치되는 구도로 촬영 - 모양 : 객체의 일부분, 잘림 등으로 객체 고유의 모양(원형, 타원형 등)이 상실되지 않도록 선정 및 촬영	병충해 학습모델 데이터
잎	- 잎을 현장 또는 채집 후 촬영 - 화면중앙에 잎 전체가 포함되는 구도로 촬영 - 대상 객체의 최대 크기가 화면을 벗어나지 않도록 촬영 - 대상 객체의 최소 크기는 이미지 세로폭의 30% 이상이 되도록 촬영 - 대상 객체 정면, 상하좌우 45도 각도에서 여러장 촬영(구도는 동일) - 앞면, 뒷면 각각 동일한 방법으로 촬영 - 해충이 있는 경우 포함하여 촬영* - 구도 : 객체가 중앙에 배치되는 구도로 촬영 - 모양 : 객체의 일부분, 둥글게 말림 등으로 고유의 모양(유선형, 부채꼴 등)이 상실되지 않도록 객체 선정 및 촬영할 것	병충해 학습모델 데이터

(※ 충해의 경우 농약, 해충제거 등 방제 활동에 따라 해충을 확보할 수 없거나 병징 부위에 해충이 군집을 이루어 붙어 있는 경우가 있으므로 해충 유무와 상관없이 촬영함.)

○ 정제 방법 및 기준

구분	적합
밝기	
	- 객체 전체의 밝기가 균등함 - 객체를 식별하기에 적절한 밝기(색농도, 색상, 질감이 실물과 유사한 정도)
음영	
	- 객체 주변 배경의 그림자가 거의 없음 - 객체와 주변 배경을 구별할 수 있는 윤곽선(색상, 색 농도 등)이 명확
위치	
	- 객체가 이미지의 중앙에 위치함 - 검수도구(SW)의 가이드 박스의 중앙에 있는 경우

구분	적합	부적합
초점		
	- 객체에 초점이 맞춤 - 객체 표면의 질감이 식별됨	- 객체에 초점이 어긋남 - 객체 표면의 질감을 식별할 수 없음

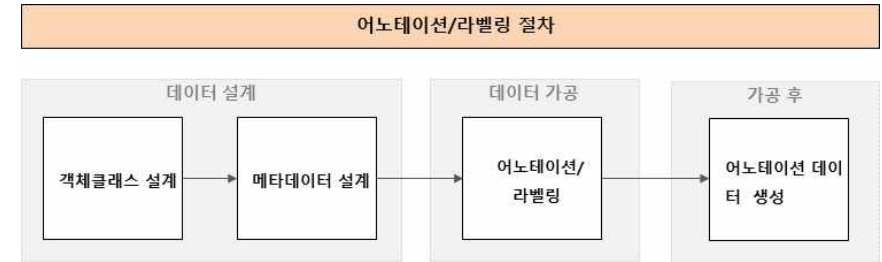
2.4 어노테이션/라벨링

2.4.1 어노테이션/라벨링 절차

□ 가공 데이터 규모

고품질 과수작물 통합 데이터	○ 과수작물 이미지 데이터 40,000장 - 어노테이션/라벨링 가공작업 수행
	○ 초분광 이미지 총 100,000세트 - 라벨링 가공 작업 수행
	○ 환경데이터 총 200,000건 이상 - 메타데이터 규격에 맞게 변환

□ 데이터 가공 절차

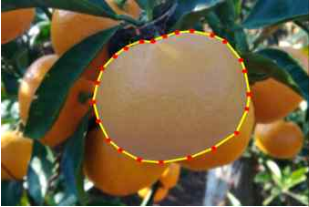
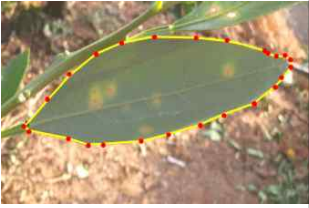
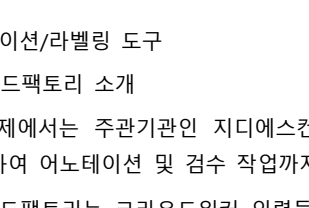


- 본 과제의 데이터 가공은 병충해 이미지에 대한 어노테이션과 라벨링으로 분류됨
- 어노테이션은 병충해 객체를 포함하는 도형을 그리는 작업이며, 라벨링은 해당 이미지 또는 도형에 대해 속성을 입력하는 작업을 말함
- 데이터 가공을 통해 이미지와 메타데이터를 생성하여 정해진 규격에 맞게 저장

□ 병충해 데이터 어노테이션

- 어노테이션은 AI모델 학습목표(Classification)에 맞게 학습용 데이터셋 활용 전 이미지의 대상 객체와 배경의 분리가 필요한 경우(현장 촬영)에 한하여 구현함
- 채집하여 하얀 배경으로 촬영한 이미지의 경우는 AI학습에 그대로 활용할 수 있기 때문에 별도의 어노테이션을 하지 않음
- 현장사진의 경우 배경과 대상 객체의 구분이 필요하기 때문에 해당 객체를 폴리곤으로 구획하고 해당 어노테이션 정보를 라벨링
- 어노테이션 도형 타입은 Polygon으로 하며, 추후 데이터의 활용도를 높이기 위하여 Fine Segmentation 방식으로 진행
- 어노테이션은 병진이 아니라 과실/잎 객체를 대상으로 하며, 해당 객체를 완전하게 포함하도록 구획

2.4.2 어노테이션/라벨링 기준

분류	이미지 예시	어노테이션 방법
채집촬영		- 채집 대상은 어노테이션 구현하지 않음 - 라벨 데이터(작물명, 정상, 병충해명, 부위명 등) 메타데이터에 정의한 값 기록함
현장촬영		【 Annotation 기준 】 - 병충해 객체를 대상으로 Polygon 도형으로 구현 - 대상 객체 전체를 포함하도록 Fine Segmentation 방식으로 구현
		【 Annotation 작업 방법 】 - 병진이 아닌 대상 객체 이미지 영역 전체가 포함되도록 작업 - 폴리곤 어노테이션으로 작업 - 라벨 클래스는 작물명, 부위명, 정상, 병충해명(3종)으로 구별함
		【 Annotation 작업 시 주의사항 】 객체 이미지 영역 내부를 침범하여 일부가 어노테이션 되는 경우 오류

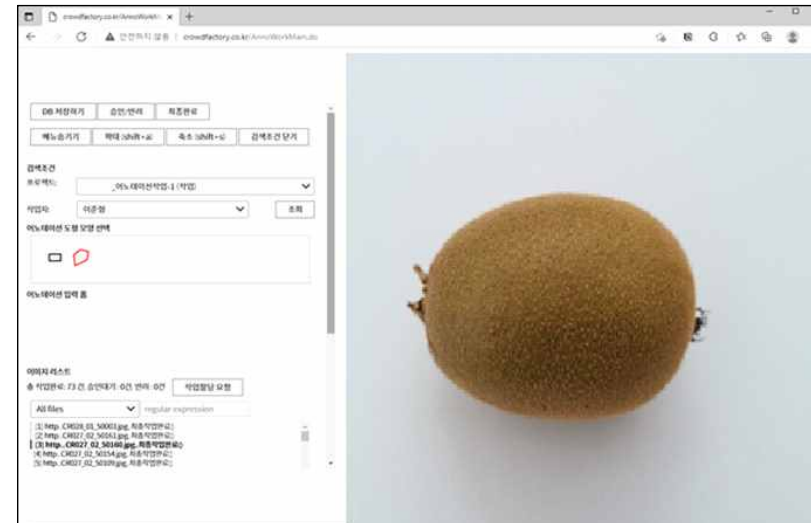
2.4.3 어노테이션/라벨링 도구

□ 클라우드팩토리 소개

- 본 과제에서는 주관기관인 지디에스컨설팅그룹이 보유한 어노테이션 통합솔루션을 사용하여 어노테이션 및 검수 작업까지 수행
- 클라우드팩토리는 클라우드워커 인력들이 재택근무로 활용할 수 있도록 웹기반의 서비스 제공
- VGG 오픈소스SW를 기반으로 개발된 클라우드팩토리는 기존 인공지능 데이터셋 구축 사업에서 활용하여 검증된 바 있음

□ 주요 기능

- Annotation 속성 파일을 저장 또는 로딩
- 주석을 만들 이미지를 불러오는 기능
- 객체 종류 등 속성 확인 기능
- Annotation 좌표는 DB에 저장되며, 최종적으로 JSON, CSV, XML 형태로 Export
- 영상이미지 전환 기능

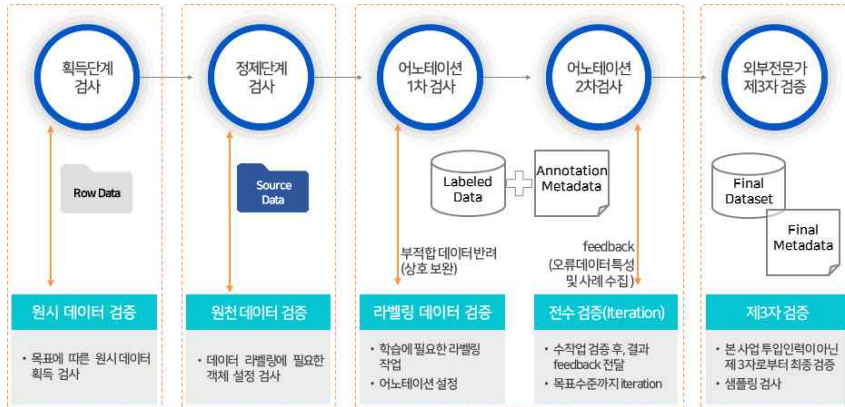


<클라우드팩토리 화면>

2.5 검수

2.5.1 검수 절차

□ 데이터 검수 절차



□ 어노테이션 데이터 검수절차

구분	세부 내용
테스트 결과 검수	- 관리자의 어노테이션 테스트 결과 확인 및 합격/불합격 통보 * 불합격 이유(미분류, 과분류, 오분류 등) 통보
균질성 교육 및 재교육	(합격) 어노테이션 결과의 균질성 확보를 위한 분류 결과 교육 (불합격) 분류항목, 분류기준 및 방법 등 재교육
교차검수 실시	작업자 A와 B의 어노테이션 결과 교차 검수
어노테이션 수정	작업자 교차검수에 따른 수정

- 작업자의 숙련도 부족에 따른 오류 발생을 최소화하고자 어노테이션 교육 및 테스트를 진행하며, 테스트 결과 검수를 통해 합격된 인원이 어노테이션을 실시
- 합격된 인원은 작업자간의 균질성 확보를 위해 테스트 분류결과에 대한 교육을 실시
- 테스트 결과에 따라 불합격된 인원은 분류항목, 분류기준 및 방법에 대해 재교육을 실시하며, 어노테이션 테스트 및 결과 검수 과정을 반복
- 어노테이션 결과에 대해 작업자 2인 이상을 1개 조로 편성하여, 작업자 간의 교차검수를 통해 검수 실시
- 수정사항 확인 시, 해당 작업자에게 오류사항에 대해 FeedBack 하며 어노테이션 수정 실시

- 필요 시, 어노테이션 실시 인원도 분류항목, 분류기준 및 방법에 대한 재교육 실시

□ 정제 데이터 검수 (검수작업팀)

- 정제된 데이터에 대한 검수를 진행
- 정제 데이터 검수 후 오류를 입력하고 데이터 정제팀에 수정 보완 요청 (어노테이션 검수결과와 종합)
- 품질검증팀에서는 어노테이션/라벨링 검수와 동시 실시

□ 가공(어노테이션) 데이터 검수 (검수작업팀)

- 최종 가공(어노테이션) 완료된 AI학습용 데이터에 대해 전수 검사 실시
- 전수검사를 통해 정확도가 목표치에 도달하지 못할 시에는 해당기간에 가공한 데이터를 가공팀에 재송신하여 보완작업을 수행함
- 검증용 View 도구를 이용 육안 검수를 진행
- 검증결과는 가공(Annotation)팀과 공유하여 상호 feedback을 주고 받으면서 목표품질수준에 도달할 때까지 iterative하게 반복
- 데이터 획득·정제팀은 데이터 획득·정제된 이미지 데이터를 검수할 수 있도록 송신
- 메타정보에 대해서는 JSON 자동 검증 프로그램으로 확인하여 Rule대로 작성되었는지 검수함
- 검수 오류내역은 다시 데이터 가공팀에 송신하고 데이터 가공팀은 오류 데이터에 대하여 보완 작업 수행함
- 오류 데이터 수신 → 어노테이션 팀에서 수정 반영 → 데이터 검수팀에 데이터 재송신, 검수팀에서 재검수 실시
- 검수를 통과하지 못한 데이터의 처리 방법
 - 검수를 통과하지 못한 데이터는 어노테이션 팀에 전달하여 어노테이션 작업을 재실시

□ 외부전문가에 의한 제3자 품질검증

- 본 과제에 참여하지 않는 외부 전문가에게 제3자 품질검증을 의뢰하여 진행
- 품질검수팀 최종 검수 완료 후 외부 전문가에 의한 전수 검수 수행

2.5.2 검수 기준

□품질검증 지표 및 목표

품질특성	항목명	측정 지표	정량 목표	지표 및 목표 설정 근거
다양성	클래스 (감귤 부위) 별 라벨 분포	비율	분포확인	일부 병충해가 우점종이 되면 경쟁 병충해가 감소하는 현상 발생으로 병충해가 균등하게 발생하지 않음
	클래스 (키위 부위) 별 라벨 분포	비율	분포확인	일부 병충해가 우점종이 되면 경쟁 병충해가 감소하는 현상 발생으로 병충해가 균등하게 발생하지 않음
	클래스 (병충해) 별 라벨 분포	비율	분포확인	일부 병충해가 우점종이 되면 경쟁 병충해가 감소하는 현상 발생으로 병충해가 균등하게 발생하지 않음
	감귤 초분광 생육 단계 분포	비율	분포확인	데이터구축방법 및 검증계획서
	키위 초분광 품종 분포	비율	분포확인	데이터구축방법 및 검증계획서
구문 정확성	구문 정확도	정확도(%)	99.9% 이상	TTA 권고사항
의미 정확성	폴리곤 라벨 정확도	정밀도 재현율 F1-score	95% 이상	데이터구축방법 및 검증계획서
	분류 라벨 정확도	정밀도 재현율 F1-score	95% 이상	초기데이터 검수 결과 Precision: 92%, Recall: 100%, F-1 Score : 95.83%, Accuracy: 91.16%
유효성	객체 분류	F1-score	85% 이상	본 과제는 농업현장에서 구축된 데이터셋으로 학습하므로 실질적으로 실현 가능한 85%를 목표로 설정함

□최종데이터셋

분류	과실	클래스
병충해 이미지	감귤	감귤_정상
		감귤_궤양병
		감귤_굴응애
		감귤_진딧물
	키위	키위_정상

분류	과실	클래스	이미지
			
		키위_점 무늬병	
		키위_총 채벌레	
		키위_과 실무름 병	

분류	과실	클래스
초분 광 이미 지	감귤	감귤_초 분광
	키위	키위_초 분광

2.6 활용

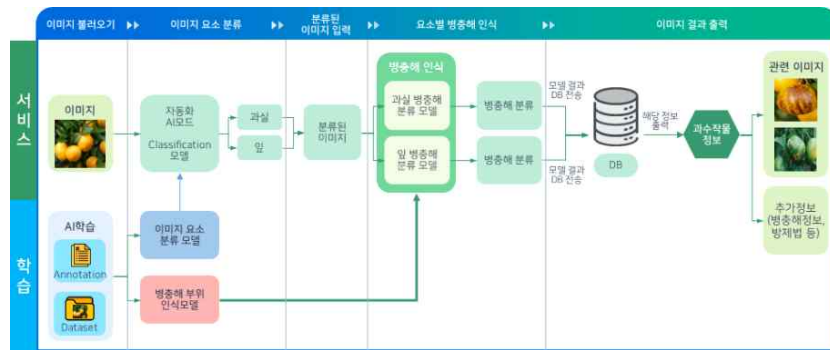
2.6.1 활용 모델

2.6.1.1 모델 학습

□ 모델 적합성 검토

- Classification 모델에는 기존 연구결과⁸⁾에 따라 아래 표와 같이 3개의 모델에 대하여 선정 기준과 절차에 따라 학습모델 선정을 진행함

Model	VGGNet	ResNet	DenseNet
Accuracy	74.5%	83.2%	78.44%

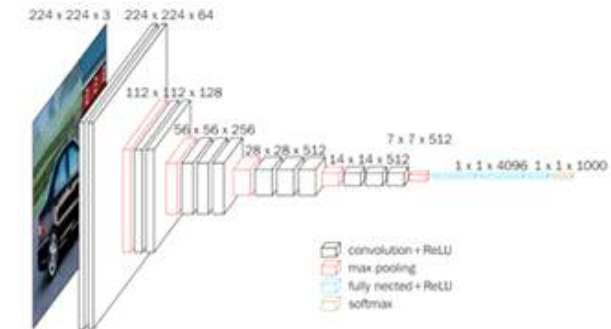


<AI모델 개발 프로세스>

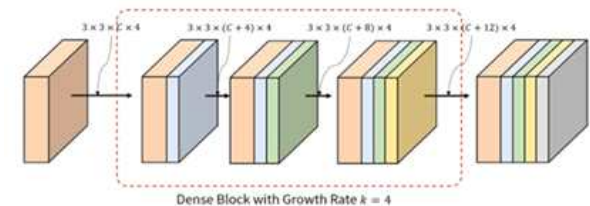
- 과수작물 병충해 데이터 비교 사례 논문을 통해 실현 가능한 목표 설정
- 논문⁹⁾에서는 ResNet50을 바탕으로 구성된 PD2SE-Net 모델을 사용하여 사과, 배, 후추 등 9종의 서로 다른 식물의 질병 심각도 추정, 종 인식 및 질병 분류 학습 모델
- 과적합 방지를 위한 모델 검증 및 대상 병충해 변경
- 과적합 (Overfitting) 방지를 위해 모델 학습시 후보 학습셋 데이터를 모델 구축용 (training dataset)과 모델 검증용 데이터셋(validation dataset)으로 분리
- 두 데이터에 대해 동일한 모형을 적용하고 두 모델의 설명력(결정계수 : R^2)기준 10% 이상 차이나는 경우 병충해 대상에 대한 변경 검토
- 병충해 대상 변경 및 선정은 NIA와 협의 후 변경
- 인공지능 학습 모델 선정

- 대규모 ImageNet 데이터셋에서 우수한 성능을 보인 VGGNet, ResNet, Densenet 모델을 Classification 인공지능 학습 모델로 선정함
- VGGNet, ResNet 모델은 ImageNet 데이터셋에 대해 높은 정확도를 보이며, 국제 ILSVRC 대회에서 이미지 분류 오류율이 각 각 7.3%, 3.57% 로 낮은 결과로 좋은 성능을 나타냄
- Inception V3 모델의 경우 높은 정확도를 보이나 소모하는 컴퓨팅 파워가 크므로 Densenet 모델이 고품질 과수작물 병충해 분류 작업에 더 적합할 것으로 예상함

□ 적용 AI모델 구조도



VGG Net 구조도

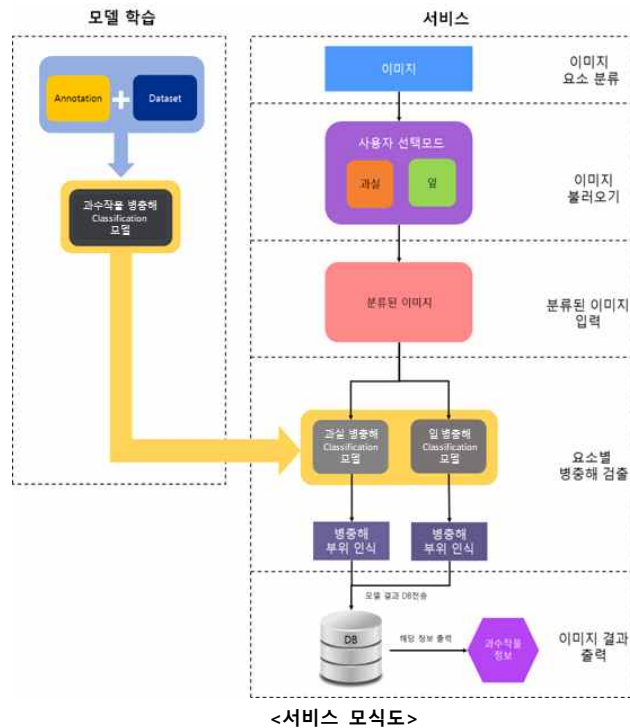


DenseNet구조도

2.6.1.2 서비스 활용 시나리오

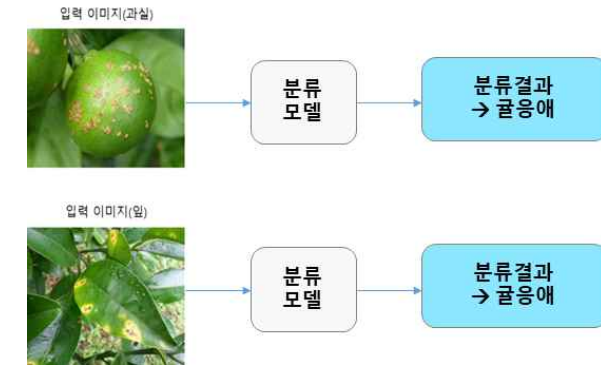
□ 과수작물 병충해 식별 사전 AI 서비스 개요

- 병충해를 자동으로 식별하는 시스템에서는 병해의 종류에 따라 과실, 잎에 나타나는 색깔, 무늬 등 시각적 특성이 다양하므로 판별에 어려움이 있음
- 병충해를 식별하는 시스템에 관해 본 서비스는 과수작물의 병충해 정보를 얻기 위하여 과실, 잎의 다양한 종류의 병충해 이미지들이 포함된 학습데이터를 활용한 분류 및 분류 모델을 개발함
- 본 서비스는 과수작물 이미지 내에서 AI 모델 개발을 통하여 병충해를 식별한 후, 해당 병충해의 종류 및 방제법 등의 다양한 정보를 출력해주는 서비스 제공



□ 분류(Classification) 모델

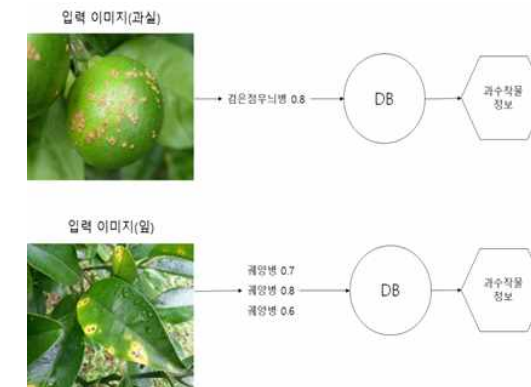
- 본 서비스에서는 사용자 선택에 의해 요소(과실, 잎)가 분류된 이미지를 AI 모델의 입력이미지로 사용함
- “과실”또는 “잎”이미지를 입력하면 어떤 병충해인지를 판별하는 분류(Classification) 모델을 구축
- 분류모델 학습 과정에서는 어노테이션(annotation)된 학습용 데이터셋을 사용



<과수 작물 병충해 분류 모델 예시>

- 분류(Classification) 모델의 병충해 추론 결과를 출력
- 위 그림과 같이 이미지를 입력하면 분류모델을 통해 어떤 병충해인지를 분류하여 결과를 제시

□ 관련 정보 출력

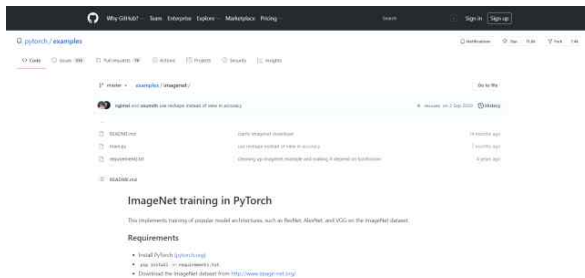


<과수작물 병충해 정보 출력 예시>

- 최종 인식 결과에 해당되는 병명ID를 DB에 전송
- 병명 ID를 통하여 DB내의 해당 병충해의 정보와 방제법 등과 같은 관련된 정보를 출력
- AI모델을 통해 과수작물의 병명 이외의 다양한 정보를 제공하는 과수작물 병충해 식별 사전을 개발하여 사용자의 접근성을 확대

2.6.2 데이터 제공

- 개발 인공지능 모델의 공개
 - 개발 인공지능 모델의 공개를 통하여 관련 연구 기술 개발의 확장을 유도
 - 소스 공개를 통하여 서비스의 개선 사항에 대한 피드백 및 성능 향상
 - 개발 소스 코드 공유를 통하여 유사 프로젝트 창출 유도
 - 서비스 공개를 통하여 유사 응용기술 개발 기대
 - AI 학습 데이터 활용을 통한 창의적 서비스 개발 및 부가가치 창출
 - 사용자들의 공개 코드 사용을 통하여 구축 데이터셋의 활용률 증가 기대
 - 사용자의 기술적 참여를 통한 서비스 신뢰도를 향상
- 모델 공개 및 기술 지원
 - 공유 플랫폼(Git hub)을 통하여 모델 공유 및 기술 지원
 - 개발 인공지능 모델의 구조 공유
 - 딥러닝 학습이 완료된 모델의 가중치 파일 공유
 - 공개 모델 및 가중치 파일의 사용 방법 공유
 - 본 사업에서 수집한 인공지능 학습 데이터셋의 활용 방법 공유



<Git-hub를 통한 소스 공개>

- 세부 공개 방법
 - 서비스 코드 및 활용 예제를 오픈소스 형태로 Git-hub를 통한 공개
 - 서비스 코드 사용 튜토리얼 및 예제의 문서화 작업
 - AI 모델의 학습(Train) 및 테스트(Test) 코드 이분화와 사용 설명 공개
 - 사전 학습된 AI 모델의 가중치 파일의 사용목적별 분류와 세부 사용 방법 문서화



<AI 모델 공개 도식도>

2.6.3 접근 및 활용 권한

- AI Hub 의 운영방침에 따름