

보안 과제( ), 일반 과제( O ) / 공개( O ), 비공개( )발간등록번호( O )  
스마트팜 다부처 패키지 혁신기술개발사업 2024년도 최종보고서

발간등록번호

11-1543000-004728-01

# 인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발

2024. 05. 28.

주관연구기관 / 국립순천대학교 산학협력단  
공동연구기관 / 지에스아이티엠  
국립순천대학교 산학협력단  
홍양영농조합법인  
위탁연구기관 / 데이터스퀘어

농림축산식품부  
과학기술정보통신부  
농촌진흥청  
(전문기관)농림식품기술기획평가원  
스마트팜연구개발사업단

제출문

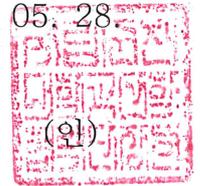
## 제 출 문

농림축산식품부 장관 · 과학기술정보통신부 장관 · 농촌진흥청장 귀하

본 보고서를 “인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발”(개발기간 : 2021. 04. 07. ~ 2023. 12. 31.)과제의 최종보고서로 제출합니다.

2024. 05. 28.

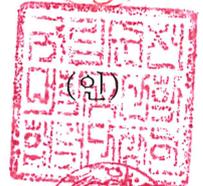
주관연구기관명 : 국립순천대학교 산학협력단 (대표자) 여 현



공동연구기관명 : 지에스아이티엠 (대표자) 변 재 철



국립순천대학교 산학협력단 (대표자) 여 현



홍양영농조합법인 (대표자) 김 중 울



위탁연구기관명 : 데이터스퀘어 (대표자) 심 형 용



주관연구책임자 : 신 창 선

공동연구책임자 : 양 지 현

정 세 훈

안 채 홍

위탁연구책임자 : 심 형 용

「국가연구개발혁신법」 제17조에 따라 보고서 열람에 동의 합니다.

<b>최종보고서</b>										보안등급			
										일반[√], 보안[ ]			
중앙행정기관명		농림축산식품부 과학기술정보통신부 농촌진흥청			사업명		사업명			스마트팜 다부처패키지 혁신기술개발사업			
전문기관명 (해당 시 작성)		농림식품기술기획평가원 (재)스마트팜연구개발사업단					내역사업명 (해당 시 작성)			차세대 융합·원천기술 연구사업			
공고번호		농림축산식품부 공고 제 농축 2021-45호			총괄연구개발 식별번호 (해당 시 작성)					-			
					연구개발과제번호					421028-03			
기술 분류	국가과학기술 표준분류		1순위 EE1001	50%	2순위 LB0805	30%	3순위 EE0108	20%					
	농림식품과학기술분류		1순위 CA0301	60%	2순위 AA0203	30%	3순위 CA0399	10%					
총괄연구개발명 (해당 시 작성)		국문		인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발									
		영문		Development of Big-data Platform for Autonomous Smart Greenhouse Production Environment Management Based on Artificial Intelligence									
연구개발과제명		국문		인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발									
		영문		Development of Big-data Platform for Autonomous Smart Greenhouse Production Environment Management Based on Artificial Intelligence									
주관연구개발기관		기관명		순천대학교 산학협력단		사업자등록번호		416-82-14326					
		주소		(57922)전남 순천시 중앙로 255		법인등록번호		201371-0002165					
연구책임자		성명		신창선		직위		교수					
		연락처		직장전화		휴대전화							
				전자우편		국가연구자번호							
연구개발기간		전체		2021. 04. 07. - 2023. 12. 31.(2년 9개월)									
		단계 (해당 시 작성)		1단계		2021. 04. 07. - 2022. 12. 31.(21개월)							
				2단계		2023. 01. 01. - 2023. 12. 31.(12개월)							
연구개발비 (단위: 천원)		정부지원 연구개발비		기관부담 연구개발비		그 외 기관 등의 지원금				합계		연구개발비 외 지원금	
		현금		현금		현물		현금		현물		합계	
총계		2,000,000		51,776		369,034		-		-		2,051,776	
1단계		1년차		600,000		13,934		105,716		-		613,934	
		2년차		700,000		19,621		137,959		-		719,621	
2단계		1년차		700,000		18,221		125,359		-		718,221	
공동연구개발기관 등 (해당 시 작성)		기관명		책임자		직위		휴대전화		전자우편		비고	
												역할	
												기관유형	
공동연구개발기관		지에스아이티엠		양지현		이사						공동 중견기업	
		순천대학교 산학협력단		정세훈		교수						공동 대학	
		홍양명농조합법인		안채홍		실장						공동 중소기업	
위탁연구개발기관		데이터스퀘어		심형용		대표이사						위탁 중소기업	
연구개발기관 외 기관		-		-		-		-		-		-	
연구개발담당자 실무담당자		성명		이명배		직위		책임연구원					
		연락처		직장전화		휴대전화							
				전자우편		국가연구자번호							

이 최종보고서에 기재된 내용이 사실임을 확인하며, 만약 사실이 아닌 경우 관련 법령 및 규정에 따라 제재처분 등의 불이익도 감수하겠습니다.

2024년 2월 29일

연구책임자: 신 창 선



주관연구개발기관의 장: 여 현



공동연구개발기관의 장: 변 재 철



공동연구개발기관의 장: 여 현



공동연구개발기관의 장: 김 종 울



위탁연구개발기관의 장: 심 형 용



농림축산식품부장관 과학기술정보통신부장관 농촌진흥청장 농림식품기술기획평가원장 귀하

## < 요약 문 >

※ 요약문은 5쪽 이내로 작성합니다.

사업명	스마트팜 다부처패키지 혁신기술개발사업		총괄연구개발 식별번호 (해당 시 작성)		-		
내역사업명 (해당 시 작성)	차세대 융합·원천기술 연구사업		연구개발과제번호		421028-03		
기술 분류	국가과학기술 표준분류	1순위 EE1001	50%	2순위 LB0805	30%	3순위 EE0108	20%
	농림식품 과학기술분류	1순위 CA0301	60%	2순위 AA0203	30%	3순위 CA0399	10%
총괄연구개발명 (해당 시 작성)	인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발						
연구개발과제명	인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발						
전체 연구개발기간	2021. 04. 01 - 2023. 12. 31 ( 2년 9개월)						
총 연구개발비	총 2,420,810 천원 (정부지원연구개발비: 2,000,000 천원, 기관부담연구개발비 :420,810천원, 지방자치단체지원연구개발비: 천원, 그 외 지원연구개발비: 천원)						
연구개발단계	기초[ ] 응용[ ] 개발[✓] 기타(위 3가지에 해당되지 않는 경우)[ ]		기술성숙도 (해당 시 작성)		착수시점 기준 ( 3 ) 종료시점 목표( 5 )		
연구개발과제 유형 (해당 시 작성)	-						
연구개발과제 특성 (해당 시 작성)	-						
연구개발 목표 및 내용	최종 목표		인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발 및 실증				
	전체 내용		스마트 온실(비닐온실, 유리온실, 수직농장)의 미처리 데이터를 포함한 빅데이터의 통합 수집 및 분석을 통해 최적 성장환경, 성장관리 및 의사결정 지원이 가능한 인공지능 학습 기술 기반 빅데이터 플랫폼 개발				
	1단계	목표	온실 미처리 데이터 자동 가공 및 처리를 통한 표준 데이터 생성 기술과 자동화된 인공지능 학습 기술 기반의 스마트온실 빅데이터 플랫폼 개발				
		내용	<ul style="list-style-type: none"> <li>○ 클라우드 기반 스마트 온실 빅데이터 시스템 설계 및 구축</li> <li>○ 실시간 데이터 시각화 및 분석 기능을 위한 웹기반 Toolbox 개발</li> <li>○ 데이터 표준화 및 통합데이터베이스 설계 및 구축</li> <li>○ 미처리 온실 작물 영상 데이터 자동 가공을 위한 인공지능 네트워크 개발</li> <li>○ 미처리 온실 작물 텍스트 자동 가공을 위한 인공지능 네트워크 개발</li> <li>○ 온실의 영상 데이터 생성을 위한 자율형 인공지능 생성 모델 개발</li> </ul>				

			○ 인공지능 최적 모델 자동 탐색과 생성 및 최적화 인공 네트워크 개발
			○ 스마트 온실 데이터 수집 및 실증 기술 설계 및 사이트 구축
2단계[ ]	목표		스마트 온실 빅데이터 분석 플랫폼(a-BRAIN) 실증 및 고도화
	내용		○ 실시간 데이터 시각화 및 분석 기능을 위한 웹기반 Toolbox 개발 및 실증 ○ 클라우드 기반 스마트 온실 빅데이터 플랫폼 실증 ○ 대상 작물을 고려한 데이터 표준화 제정 및 통합 데이터 베이스 최적화 ○ 미처리 온실 작물 텍스트 자동 가공을 위한 인공지능 네트워크 실증 및 고도화 ○ 온실 데이터 생성 및 파라미터 최적화 인공지능 네트워크 기술 실증 및 고도화 ○ 스마트 온실 데이터 수집 및 실증 사이트 운영 ○ 고품 스마트팜 혁신밸리 과제류 및 업체류 실증 단지 활용 실증

연구개발성과	<p><b>■ 연구개발 내용 및 성과</b></p> <ul style="list-style-type: none"> <li>○ 클라우드 기반 스마트 온실 빅데이터 시스템 설계 및 구축 <ul style="list-style-type: none"> <li>▪ 온실의 센서와 제어기에서 생성되는 데이터와 연동 인터페이스를 표준화한 내역을 기반으로 온실과의 연동 인터페이스를 설계하고 데이터 저장소 구조 설계</li> <li>▪ 스마트 온실의 정형, 비정형 데이터를 수집할 수 있는 서비스 API 설계 및 구축</li> <li>▪ Data Lake 아키텍처 및 ETL 설계 <ul style="list-style-type: none"> <li>- 모든 데이터는 Data Lake에 저장하고 메타데이터 추출, 모니터링 항목 및 상관 분석 항목에 맞도록 데이터 가공할 수 있는 아키텍처 설계 및 구축</li> <li>- Data Lake에 저장되는 파일 포맷은 csv, excel, parquet, zip, png, bmp 등이며, 온실 환경 및 생장 데이터 등의 정형데이터는 표준화 통합 작업 후 처리</li> <li>- 시계열 기반 데이터와 연관 이미지/동영상 정보 융합/가공을 통해 데이터 전처리 수행</li> </ul> </li> </ul> </li> <li>○ 실시간 데이터 시각화 및 분석 기능을 위한 웹기반 Toolbox 개발 및 실증 <ul style="list-style-type: none"> <li>▪ BIRT, Apache Zeppelin, Metabase, knime등의 오픈 소스 기반 BI Tool 기술 검토</li> <li>▪ 온실 환경 모니터링 <ul style="list-style-type: none"> <li>- 모니터링 항목 : 온도(냉/난방 제어 사항) 습도, CO2, 배양, 측정위치, 온실 외부 환경(광량, 온도, 습도 등 수집 가능한 데이터)</li> <li>- 데이터 현황/이력/상관관계 비교/조회/검색을 통한 시각화 처리 : 각 센싱 위치별 온실 환경과 제어 내역 조회, 기간 설정 기능을 제공하고 기간별 온실 환경 및 제어 내역을 비교/조회 기능 제공</li> </ul> </li> </ul> </li> </ul>
--------	---

- 분석 기능
  - 데이터 분석을 위한 시각화 툴 구축
  - 데이터 수집 현황과 데이터의 트렌드 분석을 위한 데이터마트 및 대시보드 개발
  - 작물의 특성에 따라 필터링 조건을 조정 가능하며, 데이터 리스트와 차트 및 Pivot 기능을 제공하여 분석 편의성 제공
- 시각화 기능
  - 수집 데이터 시각화를 위한 그리드/차트 출력
  - 데이터 분석용 집계 테이블, 필수 데이터 생성
  - 작기별/작물별 분석 그리드/차트 생성
- 클라우드 기반 스마트 온실 빅데이터 플랫폼 실증
  - 다수의 농가에서 데이터를 업로드/분석할 수 있도록 권한 기능 부여
  - 농업전문가의 의견을 반영하여 분석기능 확장
- 데이터 표준화 및 통합데이터베이스 설계 및 구축
  - “반밀폐 스마트 온실 대상 데이터 항목 표준”을 개발
    - 최첨단 반밀폐형 유리온실이 많이 확산됨에 따라 해당 온실의 특성을 감안한 “반밀폐 스마트 온실 대상 데이터 항목 표준”을 개발하고 이를 농업기술 실용화재단에 제안
  - 스마트 온실 관련 데이터 추가·수정에 따른 데이터베이스 스키마 보완
    - 영농일지·이미지 획득장소 추가, 제어정보(제어데이터, 장치유형 및 값) 수정 등 데이터 추가·수정이 발생함에 따른 데이터베이스 스키마 보완
  - 본 사업에서 수집가능한 데이터를 기준으로 데이터베이스 스키마 재설계 및 데이터베이스 구축
    - 전체적인 데이터베이스 스키마 대비 본 컨소시엄을 통해 수집할 수 있는 데이터 항목을 고려하여 데이터베이스 스키마 재설계 및 데이터베이스 구축
- 대상 작물을 고려한 데이터 표준화 제정 및 통합 데이터베이스 최적화
  - 반밀폐 스마트 온실 대상 데이터 항목 표준 초안 작성 및 제출
    - 스마트팜 ICT 융합 표준화 포럼을 통해 표준 신규 제정이 아닌 선행표준인 SPS-X KOAT-0009-7470(시설원에 분야 스마트팜 수집 데이터 규격)의 개정으로 변경
  - 시설원에 분야 스마트팜 수집 데이터 규격 개정 진행
    - 선행표준 제정 기관인 농림수산물교육문화정보원과 협의를 통해 표준 개정 완료
    - 2023년 KoAT 단체표준(안) 제정 및 개정 관련하여 예고고시를 농진원 홈페이지를 통해 진행
- 미처리 온실 작물 영상 데이터 자동 가공을 위한 인공지능 네트워크 개발
  - 데이터 스케일링 기법을 활용한 데이터 정제 및 전처리 기술/AutoEncoder를 활용한 특징 추출 기술 개발
    - 데이터 스케일링 기법을 활용한 데이터 정제 및 전처리
  - Mask R-CNN 기반 Segmentation 및 YOLO기반 Object Detection 모델 개발
    - YOLOv3,4 기반의 작물 생육단계 객체 탐지 모델 개발 및 성능평가
  - E-HRNet을 활용한 Semantic Segmentation 모델 개발
    - Squeeze 및 Excitation을 활용한 E-HRNet 기반의 Semantic Segmentation 모델 개발

- 미처리 온실 작물 텍스트 자동 가공을 위한 인공지능 네트워크 개발
  - 영농일지 OCR 인공지능을 위한 텍스트 영역 추출 및 텍스트 인식 모델 개발
    - VGG, ResNet, GRCL 기반 특징 추출 모듈 개발
    - BiLSTM 기반 Sequence Modeling 모듈 개발
  - 농업 텍스트 정보 저장 및 활용을 위한 데이터 전처리 및 가시화 기능 개발
    - 정규표현식을 활용한 특정패턴에 대한 텍스트 전처리 및 매핑 모듈 개발
    - 텍스트 정보 및 BBox 정보를 활용한 OCR 인식 시각화 모듈 개발
  - 사전학습된 모델을 활용한 전이학습기반 NLP 주석 툴 개발
    - 한국어 인식 성능 및 학습 속도 향상을 위한 전이학습 모델 개발
- 미처리 온실 작물 텍스트 자동 가공을 위한 인공지능 네트워크 실증 및 고도화
  - 온실 작물 텍스트 데이터 검증을 위한 실증
    - 영농일지 OCR 프로그램 실증을 위한 a-BRAIN 플랫폼 적용
  - 온실 작물 데이터 가공을 위한 인공지능 네트워크 실증
    - 온실 작물 데이터 가공 인공지능 네트워크 실증을 위한 객체 탐지 및 분할 모델(Mask R-CNN) 결합 및 고도화
    - 객체 탐지 및 분할 인공지능 네트워크 실증을 위한 a-BRAIN 플랫폼 적용
  - 인공지능 네트워크 기반 작물 생장/생육/품질 예측 시스템 개발
    - 작물 영상 및 시계열 데이터 활용 Multimodal 작물 분류 알고리즘 개발
    - Few-shot 및 Zero-shot 기반 품질 측정 알고리즘 개발
- 온실의 영상 데이터 생성을 위한 자율형 인공지능 생성 모델 개발
  - 과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 개발
    - 대상 작물 과채류 3종(파프리카, 토마토, 고추), 엽채류 3종(이자벨, 오비레드, 페얼리) 총 6종의 데이터 명세 정의
    - 테스트베드(휴양영농조합법인, 테라영농조합법인, 알로하채소농장)의 기본 환경데이터 백업
  - 기존에 공개되어 병충해 데이터 세트와 기 확보되어 있는 데이터를 이용하여 인공지능 생성모델(GAN-Generative Adversarial Network) 개발
    - 기존 보유중인 병해충 이미지 데이터 분석 및 GAN 기반 이미지 생성 프로세스 정의
  - 생육 데이터(예, 과실)를 이용한 데이터 생성 모델 개발
    - 공개데이터 제공 플랫폼(AI hub, kaggle 등)에서 제공하는 생육 데이터를 기반으로 하는 생육 단계 분류 체계 정의
  - 생성된 데이터를 이용한 추론 모델 재학습 기술 개발
    - 단계별 작물, 병해 데이터 이미지 생성 및 결합 모델 설계
    - 작물별 병해 이미지 생성 모델 개발을 위한 전이학습 수행
- 인공지능 최적 모델 자동 탐색과 생성 및 최적화 인공지능 네트워크 개발
  - 자동 학습 시스템을 통해 데이터에 맞는 최적 학습 방법 선택 기술 개발
    - 의사결정트리 기반 자동 머신러닝 모델 기술 개발
    - 이미지 데이터 분류를 위한 최적 학습법 탐색
  - 학습된 인공지능 모델간의 중요도(가중치) 선택 인공지능 네트워크 개발
    - 입력데이터 분류 모델 설계를 위해 원시데이터로부터 추출한 데이터셋을 사용한 상관관계 분석

- 최적의 다계층 뉴럴 네트워크 모델을 위한 Pretrained Model간 정확도 비교 및 모델 선정
- 온실 데이터 생성 및 파라미터 최적화 인공지능 네트워크 기술 실증 및 고도화
  - 과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 실증
    - 기 선정되어있던 실증사이트 외 고흥 스마트팜 혁신밸리 실증단지 선정으로 빅데이터 수집 사이트 추가 구축 완료
  - 표준화 기반 작물 병충해 데이터셋 인공지능 생성모델 실증 및 고도화
    - 공개 데이터를 활용한 병행 이미지 생성모델을 개발하여 a-BRAIN 플랫폼에 기능 탑재
    - 실데이터에 생성된 가상 데이터를 추가하여 Inception v3 모델을 활용한 Disease Detection 모델 개발
- 스마트 온실 데이터 수집 및 실증 기술 설계 및 사이트 구축
  - 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 구축
    - 과채류 3종, 엽채류 6종에 대한 현장 데이터 수집
    - 작물 재배일지 작성 및 이슈사항 정리
  - 커넥티드 팜 데이터 수집 및 실증 사이트 구축
    - 과채류 생육데이터 수집을 위한 API 설치
    - 온실 생육 의사 결정 지원을 위한 양액 배액량 및 성분 측정 설비 구축
- 스마트 온실 데이터 수집 및 실증 사이트 운영
  - 무선통신 기반 커넥티드 팜 실증사이트 운영 및 스마트팜 표준화 데이터 수집
    - 과채류 2종 환경 및 제어데이터 API 연계를 통한 DB 적재
    - 과채류 1종 및 엽채류 6종 환경 데이터 csv파일을 활용한 DB 수동 적재
    - 실증 사이트를 대상으로 한 a-BRAIN 플랫폼 개발에 따른 시범 적용 및 운영
  - 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 사이트 구축
    - 고흥 스마트팜 혁신밸리 과채류 및 엽채류 실증 단지 활용 실증
    - 고흥 스마트팜 혁신밸리 실증단지 선정에 따른 실증단지 및 실증작물(과채류 1종) 추가
  - 과채류 대상 스마트 온실 빅데이터 수집 사이트 구축
    - 고흥 스마트팜 혁신밸리 실증단지 선정(온실 8번 : 1,296m<sup>2</sup>(392))
    - 고흥 스마트팜 혁신밸리 실증단지 계약 완료(계약일 : 2023.11.20.)
    - 과채류 1종 정식 완료(정식일 : 2023.11.24, 실증작물 : 딸기(죽향, 설향 2종))

## ■ 정량적 연구개발성과

- 특허출원 8건 달성 : 계획 4건 대비 200% 달성(PCT 출원 3건)
- 특허등록 2건 달성 : 계획 4건 대비 50% 달성
  - ※ 2023.12.31. 이후 특허 2건 추가 등록 완료 : 총 4건 달성
- SW저작권 3건 달성 : 계획 3건 대비 100% 달성
- 기술이전 3건 달성 : 계획 1건 대비 300% 달성
- 기술료 8백만원 달성 : 계획 5백만원 대비 160% 달성
- 제품화 2건 달성 : 계획 1건 대비 200% 달성
- 매출액 103백만원 달성 : 계획 120백만원 대비 85.8% 달성

	<ul style="list-style-type: none"> <li>○ 고용창출 3건 달성 : 계획 2건 대비 150% 달성</li> <li>○ 표준화 1건 달성 : 계획 1건 대비 100% 달성</li> <li>○ 논문(SCI) 5건 달성</li> <li>○ 논문(비SCI) 5건 달성</li> <li>○ 논문평균 IF 3.0 달성</li> <li>○ 학술발표 15건 달성 : 계획 6건 대비 250% 달성</li> <li>○ 교육지도 18건 달성</li> <li>○ 인력양성 4건 달성</li> <li>○ 정책활용 1건 달성 : 계획 1건 대비 100% 달성</li> <li>○ 홍보전시 2건 달성</li> <li>○ 기타실적(정책활용 외) 3건 달성 : 계획 3건 대비 100% 달성</li> </ul>
<p>연구개발성과 활용계획 및 기대 효과</p>	<p><b>■ 연구개발 성과의 수요처</b></p> <ul style="list-style-type: none"> <li>○ 2021년 기준 전국 시설재배 면적은 83,823ha, 시설재배 농가는 146,000가구로 2019년 126,000가구 대비 증가 추세를 보이며, 이중 자동화 비닐하우스(단동, 연동)는 35%, 자동화 유리 온실은 0.2%로 구성</li> <li>○ 정부 주도의 청년농 육성과 귀농 인구 증가에 따라 자동화 재배시설의 신설 및 전환 농가와 함께 관련 기술의 수요도 증가할 것으로 예상</li> <li>○ 스마트농업 생산관리에 최적화된 빅데이터 플랫폼 구축 기술 및 클라우드 기반의 서비스 상품화를 통한 국내외 농가 및 관련 업체를 대상으로 한 데이터 기반 서비스 사업 가능</li> <li>○ 스마트 온실 빅데이터 분석 플랫폼 목표 시장 : 정보화 수립 희망 농가</li> <li>○ 생산관리 의사결정 지원 시스템 목표 시장 : 스마트팜 도입 희망 농가</li> </ul> <p><b>■ 연구개발 성과의 활용</b></p> <ul style="list-style-type: none"> <li>○ 스마트 온실 농가에서 생산된 작물별 데이터 분석 결과를 통해 각 작물의 생산 환경 관리 및 생산성 향상을 위한 업무·운영 매뉴얼의 기초자료로 활용 가능</li> <li>○ 동 연구개발과제를 통해 도출된 과채류·엽채류 작물의 데이터 표준화를 이용하여 최적의 온실 환경관리용 빅데이터 플랫폼 구축 시 데이터베이스 설계 및 구축에 활용 가능</li> <li>○ 데이터 표준화를 통해 온실 생육, 환경, 영농 등에 대한 정보를 규격화하여 데이터의 품질 및 해당 정보의 정확성을 제고하고 스마트 온실 생산 환경 관리를 위해 일관성 있는 분석자료 구축에 활용 가능</li> <li>○ 클라우드를 기반으로 하는 플랫폼 기술은 높은 접근성과 확장성을 제공하여 농가에서 하드웨어를 직접 구매하고 관리할 필요 없이 IT 자원을 빠르고 유연하게 확보할 수 있어, H/W 관리에 필요한 리소스 비용을 줄임으로서 응용개발에 집중할 수 있음</li> <li>○ API, 수동 업로드 등 다양한 설비의 데이터 제공 방식에 대응하는 데이터 연계 기</li> </ul>

능을 적용하여 농가 규모 및 데이터 수집 시스템에 상관없이 플랫폼 활용 가능하며, 외부 공유 데이터와의 연계를 통한 농가 수요에 따른 서비스 확장에도 용이함

- 온실 생육, 환경, 영농 등에 대한 정보를 수집하여 작물 생장에 관여하는 VPD, Entalphy 등을 계산 및 시각화하는 기능을 탑재하여 보다 정밀한 환경 관리와 제어 가능하도록 하여 에너지 효율을 높이고 작물 생산 품질 향상에 기여 가능

**■ 경제적 파급효과**

- 온실 관련 환경·생육·영농경영 등 데이터와 인공지능 기술을 연계하여 자율형 스마트 온실을 위한 생산 환경 관리용 빅데이터 플랫폼을 개발 및 보급하여 스마트 온실의 경쟁력 제고에 기여, 영농업자에게 스마트 온실 환경의 최적화 분석 결과를 제시함으로써 생산량 증대와 품질 향상을 통한 매출 증가 효과 발생
- 스마트 온실 운영의 리스크 요소를 감소시킴으로써 스마트 온실 운영자의 재무 건정성을 확보 가능

연구개발성과의 비공개여부 및 사유

연구개발성과의 등록·기탁 건수	논문	특허	보고서 원문	연구 시설·장비	기술 요약 정보	소프트 웨어	표준	생명자원		화합물	신품종	
								생명 정보	생물 자원		정보	실물
10	5	1	0	0	3	1	0	0	0	0	0	0
연구시설·장비 종합정보시스템 등록 현황	구입 기관	연구시설·장비명	규격 (모델명)	수량	구입 연월일	구입가격 (천원)	구입처 (전화)	비고 (설치장소)	ZEUS 등록번호			
국문핵심어 (5개 이내)	빅데이터		인공지능		개발도구		자율 복합환경제어		클라우드 서버			
영문핵심어 (5개 이내)	Big-data		Artificial Itelligence		Toolbox		Autonomous Climate		Cloud Server			

## <목차>

1. 연구개발과제의 개요 .....	1
1) 연구개발과제의 최종 목표 .....	1
2) 단계별 연구 목표 .....	2
3) 단계 및 연차별 연구개발 내용 .....	6
2. 연구개발과제의 수행 과정 및 수행 내용 .....	28
1) 세부별 연구 수행 결과 .....	28
2) 연구 수행 세부 내용 .....	46
3. 연구개발과제의 수행 결과 및 목표 달성 정도 .....	201
1) 연구수행 결과 .....	201
2) 목표 달성 수준 .....	227
4. 목표 미달 시 원인분석 .....	227
1) 목표 미달 원인(사유) 자체분석 내용 .....	227
2) 자체 보완활동 .....	227
3) 연구개발 과정의 성실성 .....	227
5. 연구개발성과의 관련 분야에 대한 기여 정도 .....	227
6. 연구개발성과의 관리 및 활용 계획 .....	229
[별첨 자료] .....	233

# 1. 연구개발과제의 개요

## 1) 연구개발과제의 최종 목표

### 인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발

R  
F  
P

#### 표준화 데이터 기반 스마트 온실 빅데이터 분석 플랫폼 개발

- 스마트 온실 빅데이터 수집, 저장, 전송, 표준화 및 클라우드 서버에 인텔리전스 통합 데이터베이스 구축
- 클라우드 서버의 스마트 온실 빅데이터를 활용하기 위한 분석 파이프라인 개발
- 시각화 인터페이스 개발 및 웹기반 Toolbox 개발 및 실증

#### 미처리(raw) 데이터 자동 가공 및 인공지능 모델 자동 학습 기술 개발 및 실증

- 미처리 온실 작물 영상 빅데이터 자동 가공을 위한 인공지능 머신러닝 네트워크 및 SW 개발
- 미처리 온실 작물 텍스트 빅데이터 전처리 자동화 인공지능 머신러닝 네트워크 개발
- 온실의 영상 빅데이터 생성을 위한 자율형 인공지능 생성 모델 개발
- 인공지능 최적 모델 구현을 위한 빅데이터 자동 탐색과 전처리, 파라미터 최적화를 위한 인공지능 자율 학습기 개발 및 실증

기술개발목표

#### 농업의 디지털화 (Agriculture Digitalization)

- 스마트 온실 미처리 데이터 자동 가공 및 처리
- 표준 데이터 및 자동화된 인공지능 학습 기술 기반 스마트 온실 빅데이터 플랫폼

#### 클라우드 기반 빅데이터 플랫폼 (cloud based Big-Data Platform)

- 클라우드 기반 스마트 온실 빅데이터 시스템
- 실시간 데이터 시각화 및 분석 기능을 위한 웹기반 Toolbox 개발

#### 미처리 데이터 자동화 처리 (Raw Data automated process)

- 미처리 온실 작물 영상 데이터 자동 가공을 위한 인공지능 네트워크 기술
- 미처리 온실 작물 텍스트 자동 가공을 위한 인공지능 네트워크 기술

#### 자율형 인공지능 (Autonomous artificial Intelligence)

- 온실의 영상 데이터 생성을 위한 자율형 인공지능 생성 모델
- 인공지능 최적 모델 자동 탐색과 생성 파라미터 최적화 인공지능 기술

#### 영농 데이터 통합 데이터 베이스 (Integrated database)

- 데이터 표준화 및 통합 데이터 베이스

#### 커넥티드 팜 실증 (field demonstration at coNnected farm)

- 스마트 온실 데이터 수집 및 실증 기술
- 5G 통신 기반 커넥티드 팜 및 실증 사이트 운영 및 표준화 데이터 수집



1  
차  
년  
도

- 온실 작물에 대한 데이터 생성 모델 개발
- 기확보된 데이터를 활용한 데이터 생성 모델 개발

- 데이터 가공 기술을 이용한 데이터 주요 정보 추출 기술 개발
- 객체 polygon 탐색 및 추론 기술 개발

- 클라우드 기반 스마트 온실 빅데이터 시스템 설계 및 인프라 구축

- 커넥티드 팜 데이터 수집 및 실증 사이트 그룹 선정
- 스마트 온실 빅데이터 수집 환경 설계

- 대상 작물별 데이터 명세 정의 및 표준 데이터 범위 선정
- 데이터 분석 및 인공지능 모델 구축을 위한 데이터 전처리 기술 고안

2  
차  
년  
도

- 자동 학습 시스템을 통해 데이터에 맞는 최적 학습 방법 선택 기술 개발
- 학습된 인공지능 모델간의 가중치 선택 인공지능 네트워크 개발

- 텍스트 추출을 위한 OCR 인공지능 네트워크 개발
- 텍스트 정보 전처리 및 가시화 기능 개발
- 전이학습을 활용한 NLP 주석 툴 개발

- 빅데이터 플랫폼 구축
- 분석 파이프라인 구축
- 실시간 데이터 시각화 및 분석 Tool Box

- 커넥티드 팜 데이터 수집 및 실증 사이트 구축
- 스마트 온실 빅데이터 수집 환경 구축

- 데이터 표준화 개발 및 제안
- 데이터 전처리 기술 개발
- 최적화된 데이터베이스 구축

3  
차  
년  
도

- 과채류 및 엽채류 온실 작물에 대한 데이터 생성 모델 실증
- 표준화 기반 작물 병충해 데이터셋 인공지능 생성모델 실증

- 인공지능네트워크 기반 생장/생육/품질 예측 시스템 개발
- 작물 텍스트 데이터 검증을 위한 실증
- 데이터 가공을 위한 인공지능 네트워크 실증

- 클라우드 기반 스마트 온실 빅데이터 플랫폼 실증
- 실시간 온실 데이터 시각화 및 분석 Tool Box 실증

- 5G 기반 커넥티드 팜 및 실증 사이트 운영 및 스마트팜 표준화 데이터 수집
- 고품 스마트팜 혁신밸리 실증 단지 선정 및 실증

- 작물별 데이터 명세를 고려한 데이터 표준 제정
- 데이터베이스 갱신 및 최적화
- 빅데이터 플랫폼 구조 조사 및 데이터 전송 기능 구현

## 2) 단계별 연구 목표

### 인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발

- 스마트 온실(비닐온실, 유리온실, 수직농장)의 미처리 데이터를 포함한 빅데이터의 통합 수집 및 분석을 통해 최적 생산환경, 생산관리 및 의사결정 지원이 가능한 인공지능 학습 기술 기반 빅데이터 플랫폼 개발

1단계(2021~2022) 연구개발 목표	온실 미처리 데이터 자동 가공 및 처리를 통한 표준 데이터 생성 기술과 자동화된 인공지능 학습 기술 기반의 스마트온실 빅데이터 분석 플랫폼 개발
---------------------------	--

#### ■ 1차년도(2021) 연구 개발 목표

세부 연구개발 목표	세부 연구개발 내용 및 범위	수행기관
클라우드 기반 스마트 온실 빅데이터 시스템 설계 및 구축	<ul style="list-style-type: none"> <li>클라우드 기반 스마트 온실 빅데이터 수집/저장/관리/분석할 수 있는 빅데이터 플랫폼 설계</li> <li>빅데이터 플랫폼에 필요한 하드웨어 및 소프트웨어 인프라 구축</li> <li>스마트온실 데이터 상관관계 분석 파이프라인 설계</li> <li>생육-환경-영농 데이터 사이의 상관관계 분석 설계</li> <li>오믹스 데이터를 확보할 수 있는 작물의 경우, 생명 정보와의 상관관계를 파악하여 더욱 정밀한 의사결정 및 관리지원 달성</li> </ul>	GS ITM
데이터 표준화 및 통합 데이터베이스 설계	<ul style="list-style-type: none"> <li>대상 작물별 데이터 명세 정의 및 표준데이터 범위 선정</li> <li>데이터 분석 및 인공지능 모델 구축에 활용하기 위한 온실 생육·환경·영농 정보의 데이터 전처리(가공) 기술 고안</li> <li>데이터 분석 및 인공지능 모델 구축에 활용할 수 있는 최적화 된 데이터베이스 설계</li> </ul>	데이터스퀘어
미처리 온실 작물 영상 데이터 자동 가공을 위한 인공지능 네트워크 개발	<ul style="list-style-type: none"> <li>기본 전처리 기술(PCA 등)과 Auto Encoder 등의 데이터 가공 기술을 이용하여 데이터 주요 정보 추출 기술 개발</li> <li>사전 학습된 인공지능 모델로 탐색(Detection), 분할(Segmentation)을 1차로 하고 2차적으로 보정하는 기술 개발</li> <li>픽셀 레벨의 작물 영상의 주석을 만들 때(Polygon) 수작업을 최소화 할 수 있는 polygon 포인트들의 추론 기술 개발</li> </ul>	순천대학교 (공동)

<p>온실의 영상 데이터 생성을 위한 자율형 인공지능 생성 모델 개발</p>	<ul style="list-style-type: none"> <li>▪ 과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 개발</li> <li>▪ 기존에 공개되어 병충해 데이터 세트와 기 확보되어 있는 데이터를 이용하여 인공지능 생성모델 (GAN-Generative Adversarial Network) 개발</li> <li>▪ 생육 데이터(예, 과실)를 이용한 데이터 생성 모델 개발</li> <li>▪ 생성된 데이터를 이용한 추론 모델 재학습 기술 개발</li> </ul>	<p>순천대학교 (주관)</p>
<p>스마트 온실 데이터 수집 및 실증 기술 설계</p>	<ul style="list-style-type: none"> <li>▪ 커넥티드 팜 데이터 수집 및 실증 사이트 그룹 선정</li> <li>▪ 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 설계               <ul style="list-style-type: none"> <li>- 과채류 : 3종 ( 파프리카, 토마토, 고추)</li> <li>- 엽채류 : 3종 ( 고수, 바질, 로메인상추)</li> </ul> </li> </ul>	<p>홍양 영농조합법인</p>

■ 2차년도(2022) 연구 개발 목표

세부 연구개발 목표	세부 연구개발 내용 및 범위	수행기관
<p>클라우드 기반 스마트 온실 빅데이터 시스템 구축</p>	<ul style="list-style-type: none"> <li>▪ 클라우드 기반 스마트 온실 데이터 수집/적재/분석 개발</li> <li>▪ 스마트온실 데이터 상관관계 분석 파이프라인 개발</li> <li>▪ 생육-환경-영농 데이터 사이의 상관관계 분석 개발</li> </ul>	<p>GS ITM</p>
<p>실시간 데이터 시각화 및 분석 기능을 위한 웹기반 Toolbox 개발</p>	<ul style="list-style-type: none"> <li>▪ 스마트 온실 데이터 조회 및 다운로드, 실시간 모니터링, 온실 간 Head-to-Head 비교 기능 개발</li> <li>▪ 환경-생육-영농 정보를 연결하고 동시에 분석·활용할 수 있는 실시간 시각화 인터페이스 개발</li> <li>▪ 커넥티드 팜 데이터를 활용한 개별 생육-환경 -영농 인자분석 도구 개발</li> </ul>	<p>GS ITM</p>
<p>미처리 온실 작물 텍스트 자동 가공을 위한 인공지능 네트워크 개발</p>	<ul style="list-style-type: none"> <li>▪ 기존에 축적된 스캔 파일의 텍스트를 추출하기 위한 OCR 인공지능 네트워크 개발</li> <li>▪ 농업 텍스트 정보 전처리 및 가시화 기능 개발</li> <li>▪ 전이학습 모델을 이용한 자동 Natural Language Processing(NLP) 주석 툴 개발</li> </ul>	<p>순천대학교 (공동)</p>
<p>인공지능 최적 모델 자동 탐색과 생성 파라미터 최적화 인공지능 네트워크 개발</p>	<ul style="list-style-type: none"> <li>▪ 자동 학습 시스템을 통해 데이터에 맞는 최적 학습 방법 선택 기술 개발</li> <li>▪ 학습된 인공지능 모델간의 중요도(가중치) 선택 인공지능 네트워크 개발</li> </ul>	<p>순천대학교 (주관)</p>

데이터 표준화 개발·제안 및 통합 데이터베이스 구축	<ul style="list-style-type: none"> <li>▪ 대상 작물별 데이터 명세를 고려한 데이터 표준화 개발 및 제안</li> <li>▪ 데이터 분석 및 인공지능 모델 구축에 활용하기 위한 온실 생육·환경·영농 정보의 데이터 전처리(가공) 기술 개발</li> <li>▪ 대상 작물별 데이터 분석 및 인공지능 모델 구축에 활용할 수 있는 최적화 된 데이터베이스 구축</li> </ul>	데이터스퀘어
스마트 온실 데이터 수집 실증 사이트 구축	<ul style="list-style-type: none"> <li>▪ 커넥티드 팜 데이터 수집 및 실증 사이트 구축</li> <li>▪ 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 구축 <ul style="list-style-type: none"> <li>- 과채류 : 3종 ( 파프리카, 토마토, 고추)</li> <li>- 엽채류 : 3종 ( 고수, 바질, 로메인상추)</li> </ul> </li> </ul>	홍양 영농조합법인

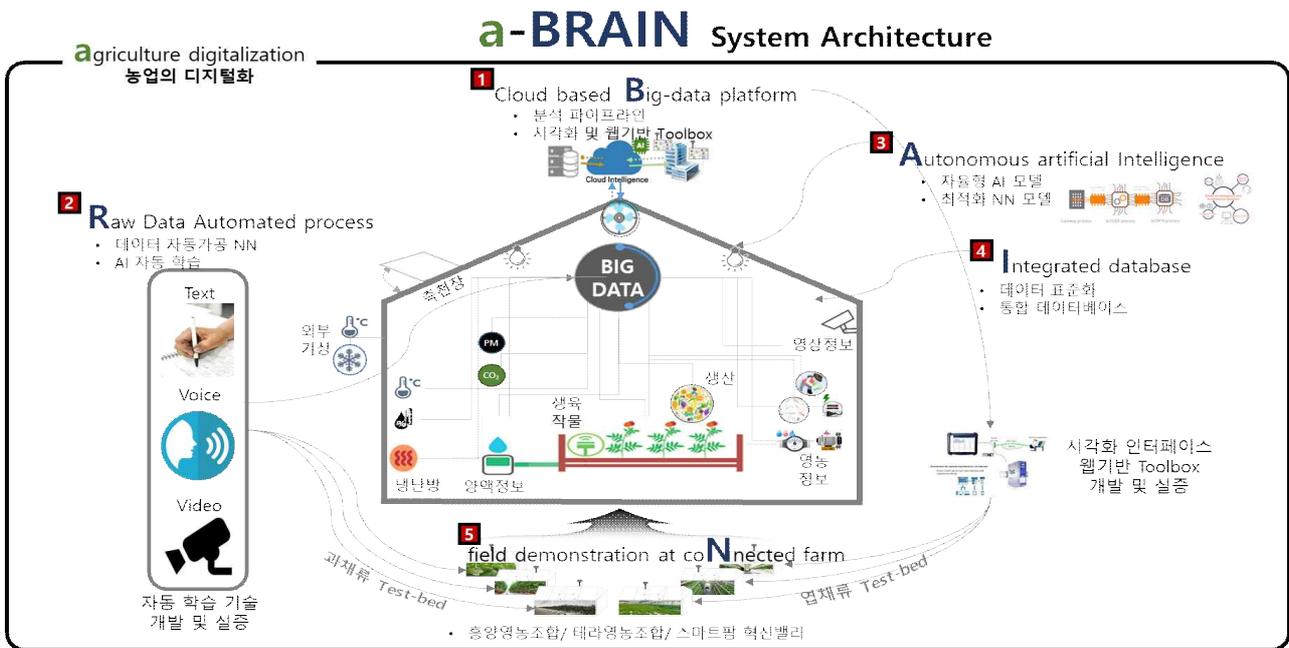
<b>2단계(2023) 연구개발 목표</b>	스마트온실 빅데이터 분석 플랫폼(a-BRAIN) 실증 및 고도화
------------------------------	-------------------------------------

### ■ 3차년도(2023) 연구 개발 목표

세부 연구개발 목표	세부 연구개발 내용 및 범위	수행기관
실시간 데이터 시각화 및 분석 기능을 위한 웹기반 Toolbox 개발 및 실증	<ul style="list-style-type: none"> <li>▪ 분석 파이프라인의 분석 결과의 시각화 실증</li> <li>▪ 영농 전문가와 시각화 요소의 적절성 평가/보완</li> </ul>	GS ITM
클라우드 기반 스마트 온실 빅데이터 플랫폼 실증	<ul style="list-style-type: none"> <li>▪ 표준화에 입각한 데이터의 수집/저장/연동에 대한 실증</li> <li>▪ 지능형 모델의 학습/배포 실증</li> </ul>	
데이터 표준화 제정 및 통합 데이터베이스 최적화	<ul style="list-style-type: none"> <li>▪ 대상 작물별 데이터 명세를 고려한 데이터 표준 제정</li> <li>▪ 실증을 통한 데이터베이스 갱신 및 데이터베이스 최적화</li> <li>▪ 스마트 팜 빅데이터 플랫폼 구조 조사 및 데이터 전송 기능 설계·구현</li> </ul>	데이터스퀘어
미처리 온실 작물 텍스트자동 가공을 위한 인공지능 네트워크 실증 및 고도화	<ul style="list-style-type: none"> <li>▪ 온실 작물 텍스트 데이터 검증을 위한 실증</li> <li>▪ 온실 작물 데이터 가공을 위한 인공지능 네트워크 실증</li> <li>▪ 인공지능 네트워크 기반 작물 성장/생육/품질 예측 시스템 개발</li> </ul>	순천대학교 (공동)
온실 데이터 생성 및 파라미터 최적화 인공지능 네트워크 기술 실증 및 고도화	<ul style="list-style-type: none"> <li>▪ 과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 실증</li> <li>▪ 표준화 기반 작물 병충해 데이터셋 인공지능 생성 모델 실증 및 고도화</li> </ul>	순천대학교 (주관)

스마트 온실 데이터 수집 및 실증 사이트 운영	<ul style="list-style-type: none"> <li>▪ 무선 통신 기반 커넥티드 팜 및 실증 사이트 운영 및 스파트팜 표준화 데이터 수집</li> <li>▪ 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 사이트 구축 <ul style="list-style-type: none"> <li>- 과채류 : 3종 ( 파프리카, 토마토, 고추)</li> <li>- 엽채류 : 3종 ( 고수, 바질, 로메인상추)</li> </ul> </li> </ul>	<p style="text-align: center;">홍양 영농조합법인</p>
	<ul style="list-style-type: none"> <li>▪ 과채류 대상 스마트 온실 빅데이터 수집 사이트 구축 <ul style="list-style-type: none"> <li>- 과채류 : 1종(딸기)</li> </ul> </li> </ul>	<p style="text-align: center;">전남스마트팜 혁신밸리 (연구협력기관)</p>

### 3) 단계 및 연차별 연구개발 내용



#### **a** agriculture Digitalization – 농업의 디지털화

- 스마트 온실 미처리 데이터 자동 가공 및 처리
- 표준 데이터 및 자동화된 인공지능 학습 기술 기반 스마트 온실 빅데이터 플랫폼

#### **B** cloud based Big-data Platform – 클라우드 기반 빅데이터 플랫폼

- 클라우드 기반 스마트 온실 빅데이터 시스템
- 실시간 데이터 시각화 및 분석 기능을 위한 웹기반 Toolbox 개발

#### **R** Raw Data automated process – 미처리 데이터 자동화 처리

- 미처리 온실 작물 영상 데이터 자동 가공을 위한 인공지능 네트워크 기술
- 미처리 온실 작물 텍스트 자동 가공을 위한 인공지능 네트워크 기술

#### **A** Autonomous artificial Intelligence – 자율형 인공지능

- 온실의 영상 데이터 생성을 위한 자율형 인공지능 생성 모델
- 인공지능 최적 모델 자동 탐색과 생성 파라미터 최적화 인공지능 기술

#### **I** Integrated database – 영농 데이터 통합 데이터 베이스

- 데이터 표준화 및 통합 데이터 베이스

#### **N** field demonstration at coNnected farm – 커넥티드 팜 실증

- 스마트 온실 데이터 수집 및 실증 기술
- 무선 통신 기반 커넥티드 팜 및 실증 사이트 운영 및 표준화 데이터 수집

## [1단계]

### (1) 1차년도 : 인공지능 학습 기술 기반의 스마트온실 빅데이터 분석 플랫폼 설계

#### ■ 주관연구기관 - 순천대학교

##### ○ 과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 개발

- 실증 대상 작물에 대한 기존 데이터 셋 기반 데이터 명세 정의

- 기존 작물별 스마트팜 데이터의 전처리 과정을 통한 데이터 정규화 수행
- 작물 분류별 공통 요소와 작물별 특성 요소에 대한 범주를 분류하여 데이터 공통 명세를 정의

- 작물 분류, 작물별 공통 데이터 명세에 따른 분석 데이터 생성 모델 개발

- 작물 분류 (과채류, 엽채류)별 스마트 온실 공통 데이터 요소 분류 및 정의
- 작물별 스마트 온실 공통 데이터 요소 분류 및 정의
- 요소별 공통 데이터 분류를 기반한 분석 데이터 생성 모델 설계

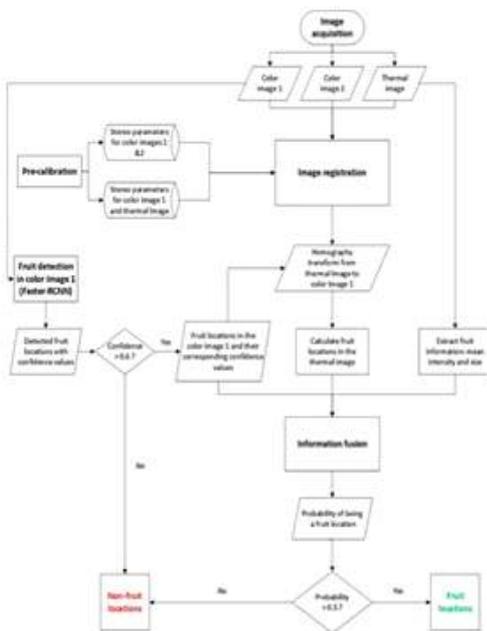
##### ○ 기존에 공개되어 병충해 데이터 세트와 기 확보되어 있는 데이터를 이용하여 인공지능 생성모델(GAN-Generative Adversarial Network) 개발

- GAN 기반의 데이터 모델 구축

- GAN 모델 구성은 생성(generator) 모델과 분별(discriminator) 모델, 결합된 GAN모델 구성
- 분별자 모델은 실제 데이터와 가상 데이터를 구별할 수 있도록 별도로 훈련시키고, GAN 모델에서는 생성하는 데이터가 분별자 모델을 통과했을 때 True에 가까워지도록 생성자 모델만 학습시킴

##### ○ 생육 데이터(예, 과실)를 이용한 데이터 생성 모델 개발

- 과채류 생육 데이터 기반 생육 단계별 데이터 분류 체계 정의



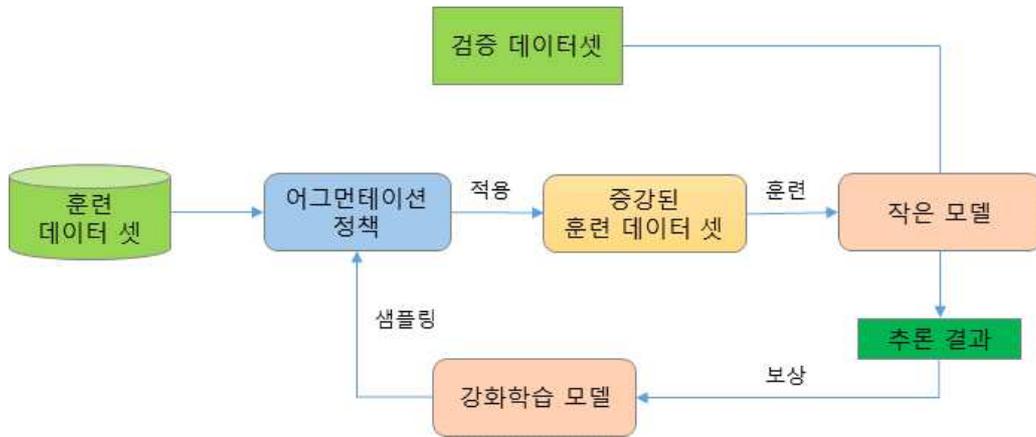
<이미징 시스템 기반 작물 정보 분류 알고리즘>



<데이터 추론 워크 플로우>

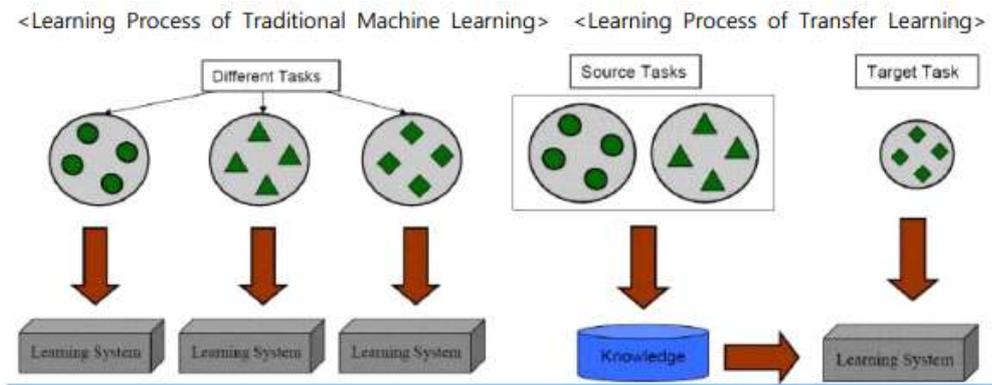
○ 생성된 데이터를 이용한 추론 모델 재학습 기술 개발

- GAN 데이터 기반 추론데이터를 활용한 재학습 데이터 분류를 통해 대상 작물에 해당하는 특정요소에 대한 재학습 수행 기술 개발



< 데이터 추론 모델 예시 >

- 기 생성된 학습 모델의 재사용 및 학습용 데이터 확보를 위한 전이학습 모델 개발



< 일반 머신 러닝과 전이학습 모델 비교 >

- 딥러닝 은닉계층 중 하위 계층이 상위 계층보다 상대적으로 기초적인 지식을 제공하는 관계로 전이 학습을 통해 지식의 재활용 가능
- 하위계층 지식의 재활용을 위한 파라미터 재정의의 통한 상위 계층 재학습 수행 기술 연구

**공동연구기관 1 - GS ITM**

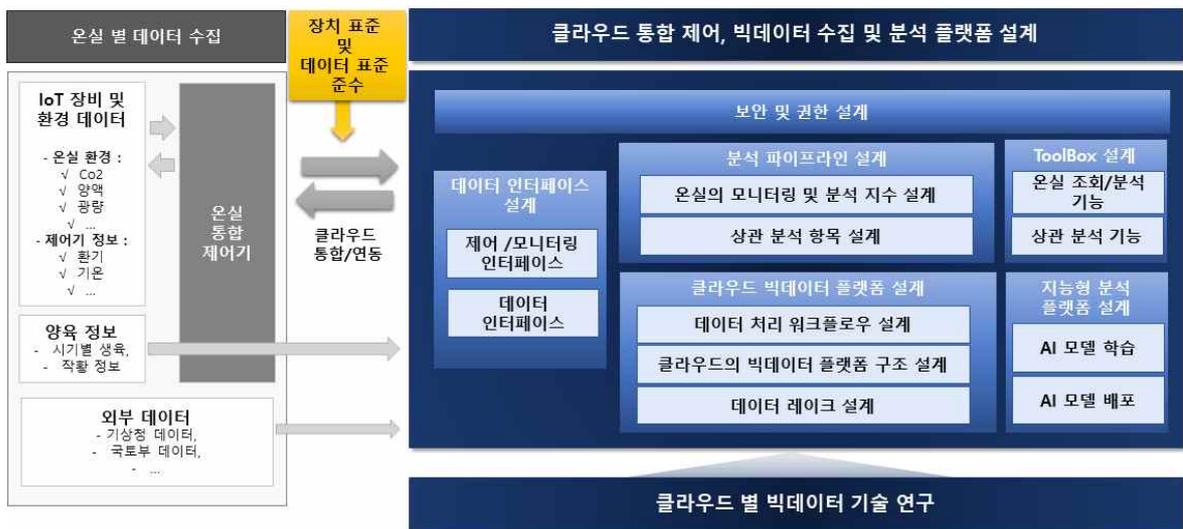
○ 클라우드 기반 스마트 온실 빅데이터 시스템 설계 및 인프라 구축

- 클라우드 기반 스마트 온실 빅데이터 시스템 설계

- 온실의 센서와 제어기에서 생성되는 데이터와 연동 인터페이스를 표준화한 내역을 기반으로 온실과의 연동 인터페이스를 설계하고 데이터 저장소 구조를 설계
- 오픈 소스에서 제공하는 빅데이터 플랫폼 기술을 조사, 연구하여 최적으로 오픈 소스 프레임 워크로 빅데이터 플랫폼을 구성함. 그리고 여러 퍼블릭 클라우드에서 제공하는 빅데이터 플랫폼과 인공지능 기반 분석 기술을 연구하여 각 클라우드 별 최적의 빅데이터 플랫폼을 설계
- 퍼블릭 클라우드에서는 대체로 오픈소스를 기반으로 시스템을 구축하거나, 오픈 소스 프레임 워크를 최적화하여 별도의 기능을 제공함. 따라서 오픈 소스 기반으로 빅데이터 플랫폼을 구

축할 경우 여러 클라우드 시스템에 적용할 수 있는 유연성과 이식성을 확보 가능

- 클라우드 기반 온실 빅데이터 시스템의 경우 다수의 농장에서 수집되는 데이터를 통합 관리하고 분석하며 조회할 수 있게 해야 하므로, 데이터의 접근 권한을 제어하여 데이터의 보안을 확보할 수 있도록 설계 필요
- 온실의 실시간 모니터링을 위해 온실의 환경, 제어기의 상태 데이터를 실시간으로 수집하고 웹 기반으로 적절하게 표현해 줄 수 있도록 설계
- 온실의 환경과 생육, 영농 인자 등을 상관 분석하기 위해 분석 파이프라인을 설계함. 이때 생육 및 작황을 기준으로 상관성 있는 시기별 온실 환경 정보, 날씨와 기온 등 외부 환경 정보, 모종과 면적, 밀집도, 지역 등의 영농 정보를 상관 분석하여 각 요소가 작물의 생육과 작황에 어떤 영향을 주는지 분석할 수 있도록 기능 개발



< 클라우드 기반 스마트 온실 빅데이터 시스템 설계 내역 >

- 클라우드 기반 스마트 온실 빅데이터 시스템 인프라 구축

- 1차년도에서는 오픈 소스와 클라우드의 빅데이터 플랫폼 기술을 연구하고 설계한 내역을 기반으로 클라우드 기반 빅데이터 플랫폼의 하드웨어/소프트웨어 인프라 구축 및 테스트
- 구축 시스템
  - ✓ 데이터 수집을 위한 메시지 전송 플랫폼
  - ✓ RestAPI를 기반으로 데이터를 송수신할 수 있는 오픈 소스 WAS 서버
  - ✓ 동영상이나 이미지 등 대용량의 데이터를 전송할 수 있는 서버 (ex. sftp 등)
  - ✓ 데이터 레이크 시스템 : 하둡의 HDFS, AWS의 S3 등 원천 데이터를 분산 환경에서 저장할 수 있는 데이터 레이크 시스템과 데이터의 메타데이터를 추출할 수 있는 카탈로깅 시스템
  - ✓ 빅데이터 쿼리 시스템 : 데이터레이크에 저장된 원천 데이터를 가공하여 쿼리 기반으로 조회할 수 있는 시스템 (하둡의 Impala, Hive 등). SQL과 NoSQL 기반으로 검색하거나 실시간 쿼리를 위한 별도의 DBMS 구축
  - ✓ 컨테이너 기반 분석 시스템 : 인공 지능 기반 학습 모듈과 분석 파이프라인 모듈을 docker와 kubernetes로 컨테이너화하여 관리할 수 있는 시스템 구축. 이때 각 모듈의 컨테이너 이미지와 필요 시스템 사양(메모리, CPU, GPU 사양 등)을 지정하여 수행할 수 있도록 구성
  - ✓ 실시간 모니터링 및 분석 시각화를 위한 Tool Box : 분석 Tool Box를 구성하기 위한 웹 및 WAS 서버와 차트, 플롯 등 시각화 모듈 구성
  - ✓ 워크플로우 엔진 : 빅데이터 처리 혹은 인공 지능 학습 모델 생성, 배포를 위한 워크플로우 엔진 (ex.

Apache Airflow, 하둡의 Oozie 등)

- 빅데이터 플랫폼 구성에 관련된 오픈 소스 프레임워크와 기술 요소들로 이 프레임워크 중 본 과제에 필요한 기술 요소를 선별하여 구축

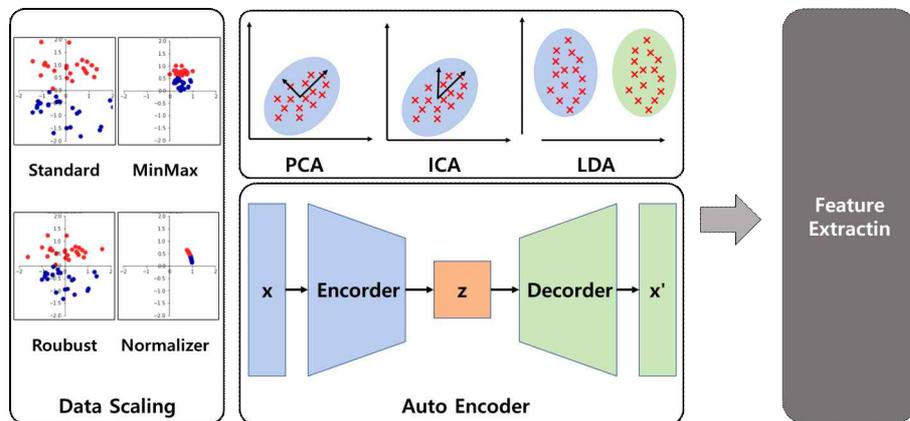


< 빅데이터 플랫폼에 대한 퍼블릭 클라우드 기술 제공 예시(AWS 예시) >

- 구성된 빅데이터 플랫폼을 기반으로 2차 년도에 데이터 레이크 및 데이터 저장소에 데이터 저장 구조를 구축하며 데이터 처리 시스템을 구축하고 인공 지능 학습 모델 및 분석 모델의 학습과 예측 시스템을 구축할 계획

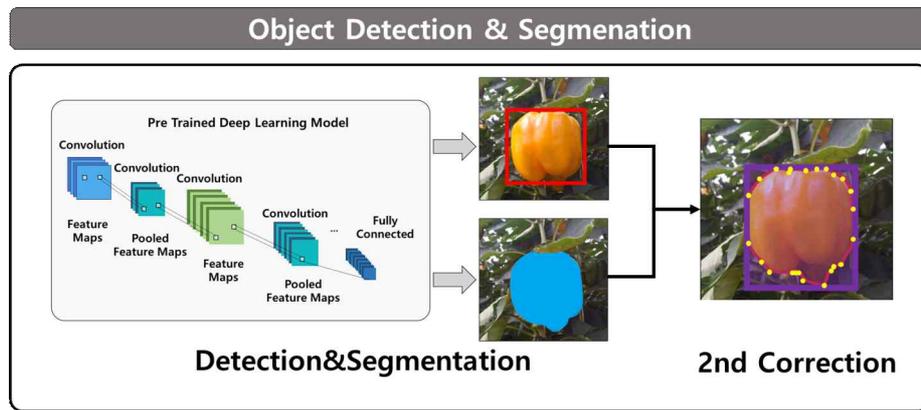
## ■ 공동연구기관 2 - 순천대학교(공동)

- 기본 전처리 기술(PCA 등)과 Auto Encoder 등의 데이터 가공 기술을 이용하여 데이터 주요 정보 추출 기술 개발
  - 데이터 스케일링 기법을 활용한 데이터 정제(Standard, MinMax, Roubust, Normalizer)
  - 기본 전처리 기술(PCA(Principal Component Analysis), ICA(Independent Component Analysis), LDA(Latent Dirichlet Allocation))을 활용한 주요 정보 추출
  - Auto Encoder를 통한 데이터 가공 및 정보 추출



<Auto Encoder 처리 과정>

- 사전 학습된 인공지능 모델로 탐색(Detection), 분할(Segmentation)을 1차로 하고 2차적으로 보정하는 기술 개발
  - 객체 탐지 및 분할을 위한 사전 학습된 인공지능 모델 기반 객체 탐지 및 분할 기술 개발
  - 2차적 보정을 통한 영상 객체 탐지 범위 세분화 및 객체 분할 정밀화 기술 개발
- 픽셀 레벨의 작물 영상의 주석을 만들 때(Polygon) 수작업을 최소화 할 수 있는 polygon 포인트들의 추론 기술 개발
  - 객체 탐지 및 분할 기술을 활용한 주석 예측 및 Polygon 포인트 추론 기술 개발



<객체 탐지 및 분할 과정>

### ■ 공동연구기관 3 - 흥양영농조합법인

- 커넥티드 팜 데이터 수집 및 실증 사이트 그룹 선정
  - 과채류
    - 흥양영농조합법인
    - 테라영농조합법인
  - 엽채류
    - 알로하채소농장
- 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 설계
  - 과채류 : 2종 (파프리카, 토마토)
    - 주요 수집 데이터 set : 외부기상, 광량, 내부 온도, 습도, CO2농도, 양액공급량, 공급시간, EC, PH 등
  - 엽채류 : 1종 (고수)
    - 주요 수집 데이터 set : 외부기상, 내부온도, 습도, 양액 EC, PH, 양액공급시간 등
    - 데이터수집 환경 구축



<실증사이트 현황>

■ 위탁연구기관 - 데이터스퀘어

○ 대상 작물별 데이터 명세 정의 및 표준데이터 범위 선정

- 한국정보통신기술협회, 농업기술실용화재단, 농촌진흥청 등 기존 스마트팜 관련 데이터 표준화 현황\* 조사 및 분석
  - (KSX3268) 스마트온실 구동기 메타데이터, (KSX3269) 스마트온실 센서 메타데이터, (TTAK.K0-10.0843) 시설 원예 생육 진단 메타데이터, (TTAK.K0-10.1092) 농장 빅데이터 서비스 제공자와 온실 관제 시스템 간의 인터페이스, (SPS-X FACT-0006-일련번호:2021) 스마트온실 과채류 생육정보 메타데이터 등
- 스마트 온실 생산환경 관리를 위한 대상 작물별 측정 항목 조사 및 데이터 명세 정의

<생육데이터 측정 항목·명칭·단위 예시(과채류 중심)>

측정 항목	표준단어	표준단어(영문)	측정단위	비고
엽장, 엽길이, 잎의길이	잎길이	LEF_LT	cm	국화, 딸기, 방울토마토, 오이, 참외, 토마토
엽폭, 잎의너비	잎너비	LEF_BT	cm	국화, 딸기, 방울토마토, 오이, 참외, 토마토
잎수, 엽수, 잎의갯수	잎개수	LEF_CUNT	개	국화, 딸기, 방울토마토, 오이, 토마토
초장	초장	TOT_LT	cm	국화, 딸기, 방울토마토, 참외, 토마토, 파프리카
생장길이	생장길이	GRWT_LT	cm	국화, 딸기, 방울토마토, 참외, 토마토, 파프리카
줄기굵기, 줄기직경	줄기굵기	STEM_THCK	mm	방울토마토, 오이, 토마토, 파프리카
착과화방	착과화방	FRTST_FCLU	화방	방울토마토, 토마토, 파프리카
화방당착과수	착과수	FRTST_CUNT	개	딸기, 방울토마토, 토마토, 파프리카

※ 출처: 농촌진흥청

- 기존 스마트팜 관련 표준화와 대상 작물별 데이터 명세를 고려한 데이터 표준화 설계

○ 데이터 분석 및 인공지능 모델 구축에 활용하기 위한 온실 생육·환경·영농 정보의 데이터 전처리(가공) 기술 고안

- 온실 생육·환경·영농 정보 등에 대한 각 데이터 특성 및 구조 분석

- 데이터의 누락값·불일치·오류·숫자 및 날짜 형식의 일관성 문제 등 불완전하고 모순된 데이터를 정제(Cleaning)하는 데이터의 무결성 검증 로직 고안
- 온실 생육·환경·영농 정보 관련 다양한 파일·자료를 데이터 분석·인공지능 활용이 용이하도록 데이터 통합 방안 조사·분석

○ 데이터 분석 및 인공지능 모델 구축에 활용할 수 있는 최적화 된 데이터베이스 설계

- 대상 작물별 수집된 온실 생육·환경·영농 정보(원천 데이터) 등을 기반으로 온실 생산 환경 관리용 데이터베이스 구축을 위한 데이터 현황 분석
- 대상 작물별 온실 생육·환경·영농정보 관련 데이터의 현황 분석을 통한 주요 데이터 항목 발굴
- 작물별 온실 생육·환경·영농 등 주요 영역(category)을 고려하여 효과적인 온실 생산환경 관리를 지원할 수 있는 데이터베이스 논리적·물리적 스키마(Schema) 설계

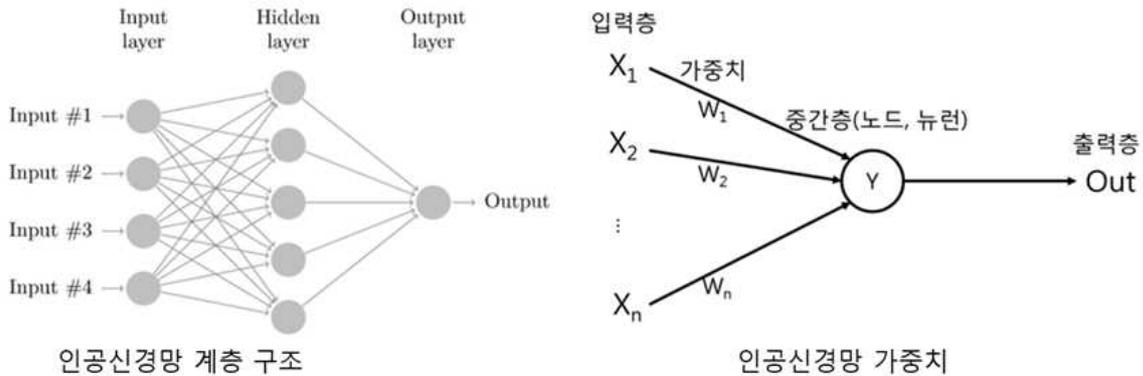
<스마트 온실 관리를 위한 주요 데이터 현황(예시)>

구분	주요 데이터
생육	(토마토) 초장, 생장길이, 엽수, 엽장, 엽폭, 줄기굵기, 화방높이, 개화군, 착과군, 수확군, 착과수, 수확과수, 수확과중 등 (파프리카) 초장, 생장길이, 엽수, 엽장, 엽폭, 줄기굵기, 화방높이, 개화마디, 착과마디, 착과수, 열매수, 수확마디, 수확과수, 과폭, 과고, 무게 등 (딸기) 초장, 엽수, 엽장, 엽폭, 엽병장, 관부직경, 화방출퇴기, 개화기, 꽃수, 착과수, 액아 발생시기, 과중, 과장, 과폭, 당도, 산도 등
환경	(온실외부) 일사량, 온도, 상대습도, 풍향, 풍속, 강우 등 (온실내부) 일사량, 온도, 상대습도, CO2 농도, 풍속 등 (토양환경) 근권 온도·수분, 양액 수온, 양액 pH·EC, 배지 pH·EC 등 (시설환경) 시설 유형·면적, 피복겹수, 수평커튼 겹수, 천창 유형, 측창 유형, 차광막·보호커튼 유무, 제어기 유형, 양액기 유형, 난방기 유형 등
영농	(농작업관리) 농업인ID, 시설ID, 농작업명, 농작업유형, 농작업내용, 비용 등 (인력관리) 농업인ID, 시설ID, 인력활용 일시·유형·내용, 인력비용, 인원수 등 (유지관리) 농업인ID, 시설ID, 인력활용 일시·유형·내용, 비용 등 (원자재관리) 농업인ID, 시설ID, 원자재 항목·구입일시·비용 등 (양액관리) 농업인ID, 시설ID, 양액 항목·구입일시·무게·비용 등 (방제관리) 농업인ID, 시설ID, 방제 일시·작업내용·비용 등 (생산관리) 농업인ID, 시설ID, 작물명, 생산 일시, 생산량 등 (출하관리) 농업인ID, 시설ID, 작물명, 출하 일시, 출하량, 출하 단위 당 가격, 매출 등

(2) 2차년도 : 인공지능 학습 기술 기반의 스마트온실 빅데이터 분석 플랫폼 구축

■ 주관연구기관 - 순천대학교

- 자동 학습 시스템을 통해 데이터에 맞는 최적 학습 방법 선택 기술 개발
  - 표준 데이터셋 명세를 기반한 스마트 농업 자동 머신러닝 모델 기술 개발
  - 작물 데이터 분류별(작물 분류별, 작물별) 자동 학습 모델 개발
- 학습된 인공지능 모델간의 중요도(가중치) 선택 인공지능 네트워크 개발

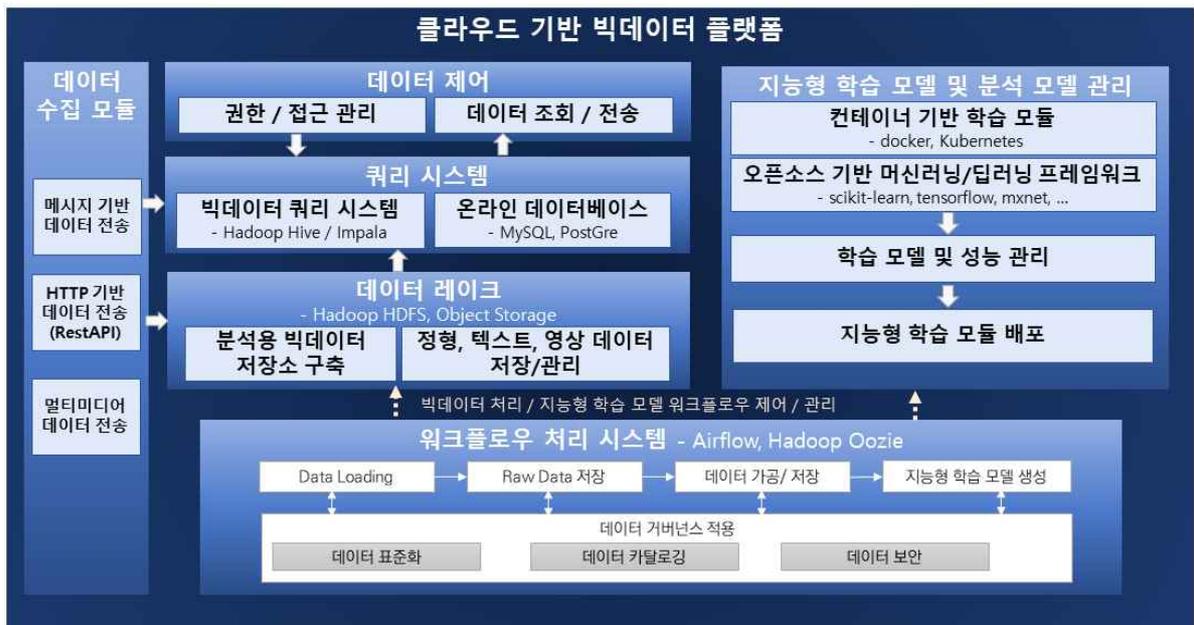


<인공신경망 계층 구조 및 가중치>

- 원시 입력 데이터 분류 모델 및 의미 분석 모델 설계
  - 가중치 분류(Classification)를위하여 원시 데이터로부터 데이터셋 추출
- 가중치 선택 인공지능 네트워크 개발을 위한 다계층 뉴럴 네트워크 모델 개발

■ 공동연구기관 1 - GS ITM

- 빅데이터 플랫폼 구축
  - 표준화된 데이터에 기반한 빅데이터 저장 구조 및 처리 모듈을 개발

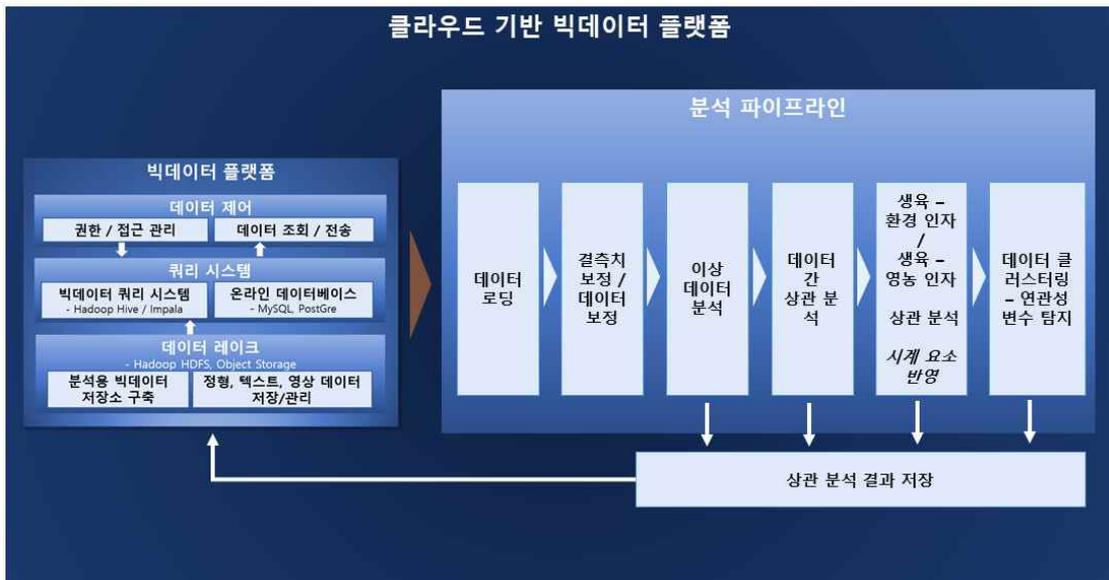


<클라우드 기반 빅데이터 플랫폼 구조도>

- 데이터 수집 모듈 : 스마트 온실에서 생성되는 데이터를 각각의 온실 제어기의 연동 인터페이스와 데이터 스펙에 따라 데이터를 수집할 수 있는 인터페이스를 개발함. 데이터 및 제어기의 연동 인터페이스가 표준화 될 경우 이 표준화 된 내역에 따라 표준화 인터페이스 제공
- 데이터레이크 : 데이터 수집 모듈에서 수집된 원천 데이터를 저장함. 각 제어기에서 제공되는 정형 데이터 뿐만 아니라 텍스트, 이미지, 영상 데이터까지 저장함. 데이터 처리의 워크플로우에 따라 데이터를 카탈로깅하여 메타데이터를 생성 관리하며 본 연구에 연구하는 표준에 맞는지 점검하며, 표준에 맞도록 변환하는 작업 수행
- 쿼리 시스템 : 데이터레이크에 저장된 데이터를 집계/통계하여 실시간 모니터링 시각화 및 분석 Tool Box에서 조회할 수 있도록 함. 데이터 쿼리 시스템에 온실 데이터, 생육 데이터, 영농 인자 데이터의 표준화된 내역을 반영하여 표준에 따른 데이터 집계 및 조회가 가능
- 데이터 제어 : 사용자별, 역할별 데이터의 접근 권한을 설정하고 이를 기반으로 각 모듈이 데이터에 대한 접근 권한이 있는지 확인하고 제어
- 인공지능 학습 모델 관리 : 미처리 데이터의 인공지능 모델, 분석 파이프라인 모듈이 별도의 컨테이너로 구성, 구동될 수 있도록 각 모듈을 컨테이너로 제작하고, 제작된 컨테이너를 관리함. 데이터 분석 량에 따라 각 분석 모듈 컨테이너를 자연스럽게 증감시킬 수 있도록 구현

○ 분석 파이프라인 구축

- 온실의 환경-생육-영농 인자의 상관 분석을 위한 분석 파이프라인 모듈을 개발하고 이를 분석 모듈의 컨테이너에 넣어서 관리

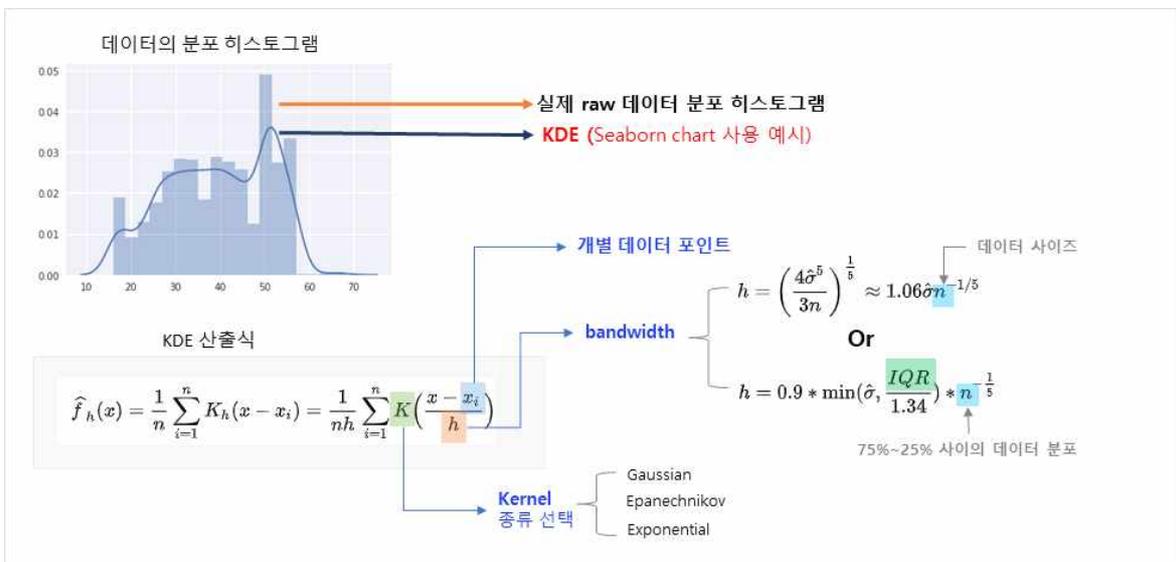


<상관 분석을 위한 분석 파이프라인>

- 위 그림과 같이 분석 파이프라인에서는 데이터레이크 혹은 쿼리 시스템으로부터 데이터를 로딩하여 상관 분석을 진행하고 각 단계의 상관 분석 결과를 저장하여 분석 Tool Box에서 조회/분석 가능

- 분석 파이프라인의 분석 단계

- 데이터 로딩 : 데이터레이크 혹은 쿼리시스템으로부터 상관 분석에 필요한 데이터를 조회하여 분석 데이터를 생성
- 데이터 전처리 : 로딩한 데이터에 결측치 데이터가 있는지, 보정할 데이터가 있는지 확인하여 결측치 데이터를 보정(삭제, 평균값 사용, 이전 값 사용 등)하고, 데이터의 분포를 확인하여 보정 할 수치가 있는지 확인하고 보정함. 데이터 자체 통계 정보(전체 분포도, 평균, 분산 등)를 분석
- 이상 데이터 분석 : 데이터에 Outlier가 있는지 확인하고 Outlier 데이터의 처리 방법을 결정함. 예를 들어 Outlier 데이터를 보정하여 다른 데이터들과의 분포에 맞출 수도 있고, Outlier 데이터만을 별도로 분석할 수 있음. Outlier 데이터가 필요없을 경우에는 삭제를 하고 분석을 진행함. Outlier 데이터들도 분석 Tool Box에서 조회할 수 있도록 하여 온실의 환경, 생육, 영농 인자의 데이터 중 독특하게 생성되는 패턴을 조회 및 분석
- PCA와 같은 주성분 분석 알고리즘이나 KDE와 같은 알고리즘을 통해 이상 데이터를 분석



< KDE 정의 >

- 데이터간 상관 분석 : Correlation 분석 및 OLS(Ordinary Least Square) 분석을 통해 데이터 간의 상관 관계를 분석하고 상관 관계가 유효한 변수들을 추출

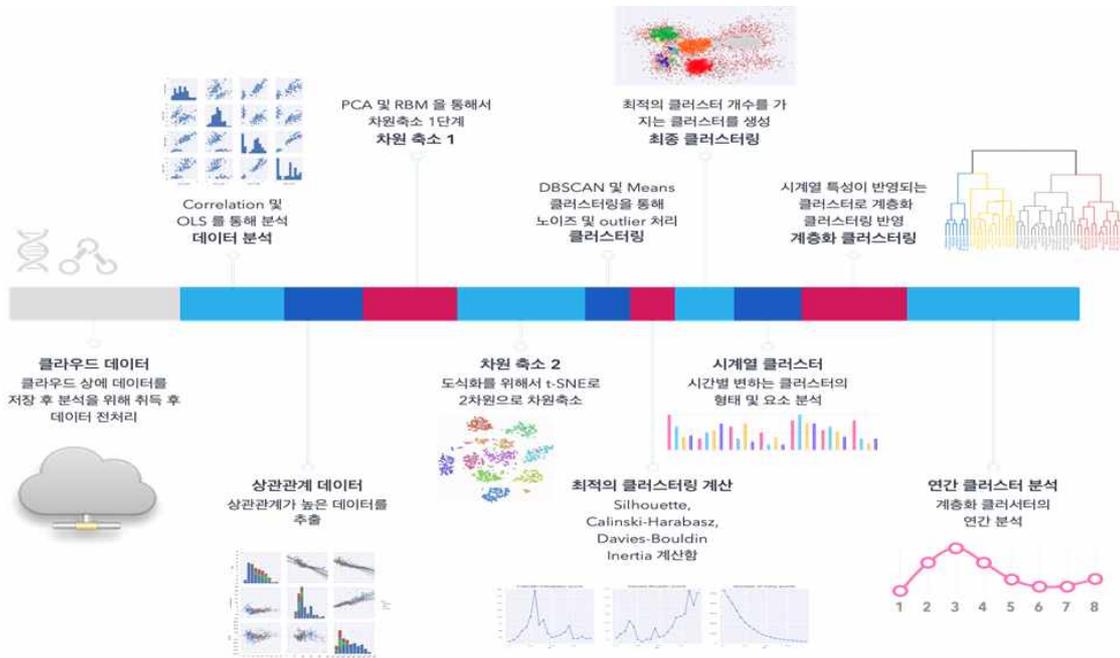
수식	기본 설명	동일한 의미	자세한 설명
$\sum_{i=1}^n (y_i - y')^2$	<b>Sum of Squares Total (SST)</b>	<b>Total Sum of Squares (TSS)</b>	데이터 베이스 전체의 가변성
$\sum_{i=1}^n (y_i - y^{\wedge})^2$	<b>Sum of Squares Regression (SSR)</b>	<b>Explained Sum of Squares (ESS)</b>	선형 회귀 분석으로 추정된 가변성
$\sum_{i=1}^n e_i^2$	<b>Sum of Squares Error (SSE)</b>	<b>Residual Sum of Squares (RSS)</b>	에러 항목

OLS 정의	OLS 정의	수식
	<b>OLS</b> = <b>Ordinary Least Squares</b> = <b>Ordinary Least SSE</b>	$\sum_{i=1}^n (y_i - y')^2 = \sum_{i=1}^n (y_i - y^{\wedge})^2 + \sum_{i=1}^n e_i^2$

- 분석 의미
- 데이터 베이스 전체의 가변성은 선형 회귀 분석으로 설명이 된 가변성과 설명이 안된 가변성의 합
  - 선형 회귀 분석으로 설명이 안된 에러인 SSE 부분이 0에 가까울 수록 선형 회귀 분석이 잘 된 것으로 판별

< OLS 정의 >

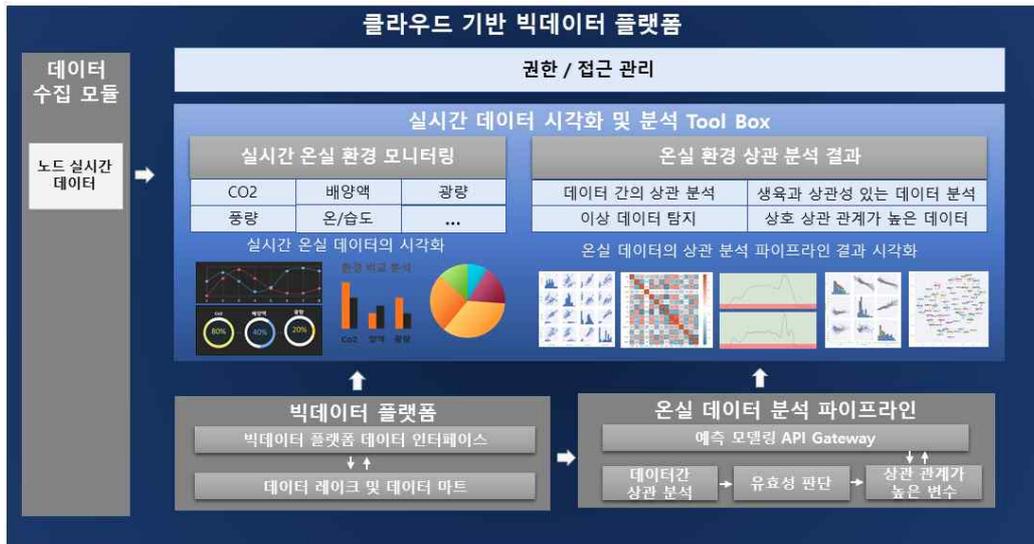
- 생육 기준 환경, 영농인자 상관 분석 : 온실 작물의 생육과 상관 관계가 있는 환경 정보나 영농 인자를 분석하기 위해 OLS 방법론을 활용하여 분석하고, 유용한 변수를 탐지함. 상관 관계가 유효한 데이터의 분포와 데이터 간의 연관성을 파악하기 위해 데이터의 클러스터링을 진행함. 이 때 데이터들은 모두 시계열 데이터이므로 시계열 요소를 반영하여 클러스터링을 진행
- 최적의 데이터 차원 축소 : 데이터 차원 변환 에러인 Kullback Leibler Divergence를 최소화 시키는 차원 축소 알고리즘 Restricted Boltzmann Machine(RBM)와, t-distributed Stochastic Neighbor Embedding(t-SNE)의 최적의 하이퍼 파라미터를 선택
- 최적의 클러스터링 : 비지도 학습인 클러스터링에서 최적의 클러스터 개수를 판별하기 위해 Silhouette, Calinski-Harabasz, Davies-Bouldin score 및 inertia 방법론을 사용하여 최적의 클러스터링 개수를 계산 후 이를 반영
- 연관성 변수 탐지 : 데이터들 간의 시계열 분포를 클러스터링함으로써 데이터들 사이에 유사한 패턴을 보이는 데이터를 연관성 있는 변수로 탐지하며, 탐지된 군집들을 평가하고 조정



<상관 분석 파이프라인 구성 예시>

○ 실시간 데이터 시각화 및 분석 Tool Box

- 온실의 환경 정보의 실시간 시각화를 위한 Tool Box와 상관 분석의 파이프라인 결과를 시각화로 표현하기 위한 Tool Box를 개발

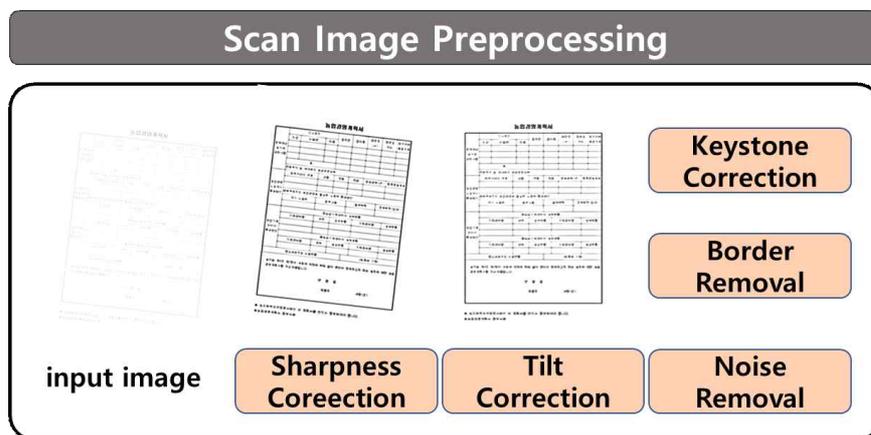


<스마트 온실 실시간 데이터 시각화 및 분석 Tool Box 구성>

- 시각화 모듈 및 분석 Tool Box는 웹기반으로 개발하며, HTML5와 차트 및 플롯, 표 구성을 위한 자바스크립트 모듈을 사용
- 실시간 온실 환경을 위해서는 각 온실의 제어기와 직접 연동하여 데이터를 수집하고 표현해 주며, 과거 이력 데이터를 시각화하여 보여 줄 경우에는 빅데이터 플랫폼에 저장되고, 집계된 데이터를 기반으로 표현
- 분석 Tool Box에서는 상관 분석을 위한 분석 파이프라인 과정에서 생성되는 분석 결과 데이터를 시각화하여 표현해 주며, Correlation Matrix, 회귀 분석 Chart, HeatMap Chart, KDE Chart, Scatter Chart 등 다양한 시각화 기능을 제공하여 데이터의 분포와 상관성 판단

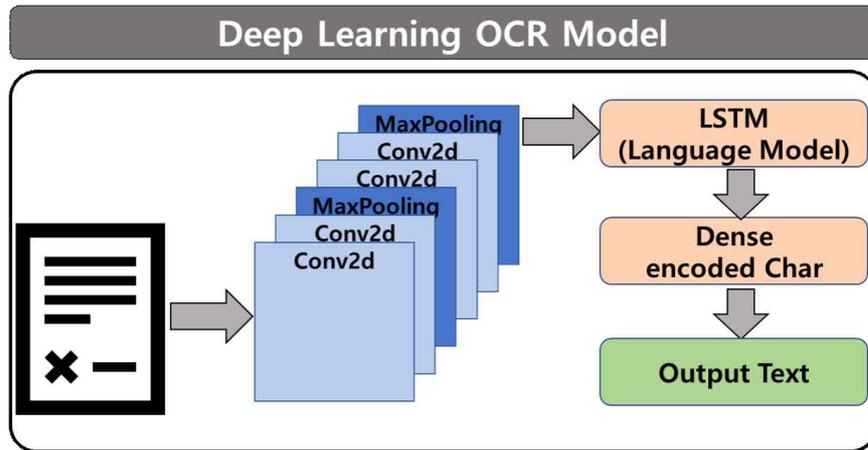
■ 공동연구기관 2 - 순천대학교(공동)

- 기존에 축적된 스캔 파일의 텍스트를 추출하기 위한 OCR 인공지능 네트워크 개발
  - 스캔 파일 이미지 전처리(잡티 제거, 기울기 보정, 여백 제거) 모듈 개발



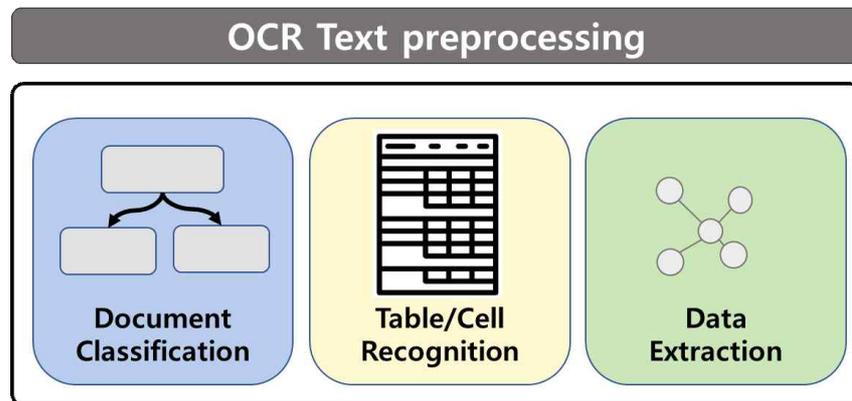
<스캔 이미지 전처리 과정>

- 딥러닝 기반 스캔 파일 서식 판별/인식 및 추출 인공지능 네트워크 개발



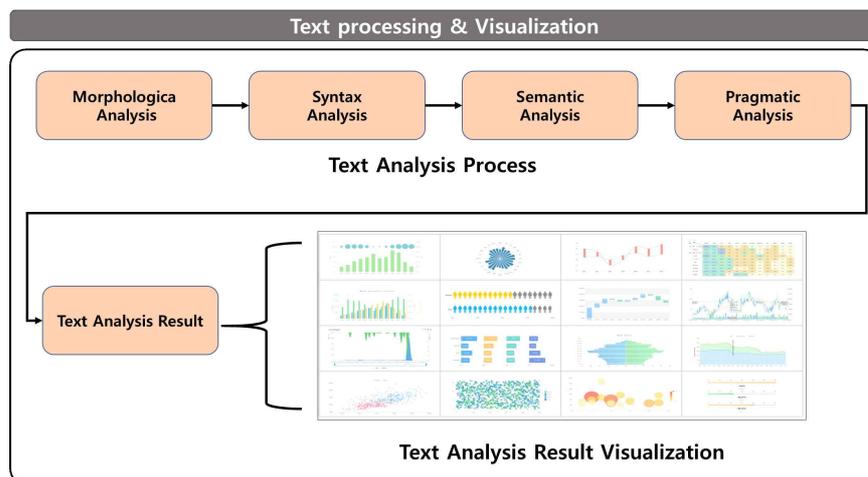
<OCR 모델 처리 과정>

- 농업 텍스트 정보 전처리 및 가시화 기능 개발
  - 농업 텍스트 정보 전처리 모듈 개발(문서 분류, 테이블 인식, 데이터 추출 등)



<OCR 모델 텍스트 처리 과정>

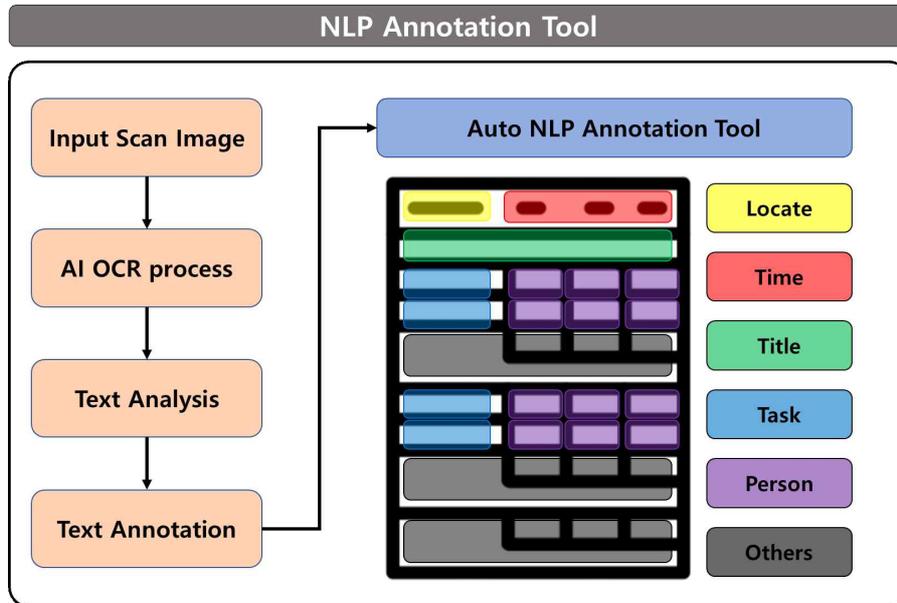
- 농업 정보 텍스트 분석(형태소, 구문, 의미, 화용 분석) 및 가시화 기능 개발



<텍스트 분석 및 시각화 과정>

- 전이학습 모델을 이용한 자동 Natural Language Processing(NLP) 주석 툴 개발
  - 인공지능 전이학습 모델을 활용한 자동화된 NLP 주석 툴 개발

- 스캔 이미지 입력 시 OCR AI Network 및 Text Analysis 과정을 거쳐 텍스트/테이블 컬럼 별 자동 주석 처리



<자연어 처리 과정>

### ■ 공동연구기관 3 - 흥양영농조합법인

#### ○ 커넥티드 팜 데이터 수집 및 실증 사이트 운영

##### - 과채류

- 흥양영농조합법인
- 테라영농조합법인

##### - 엽채류

- 알로하채소농장

#### ○ 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 구축

##### - 과채류 : 3종(파프리카, 토마토, 고추)

- 주요 수집 데이터 set : 외부기상, 광량, 내부 온도, 습도, CO2농도, 양액공급량, 공급시간, EC, PH 등

##### - 엽채류 : 3종(고수, 바질, 로메인 상추)

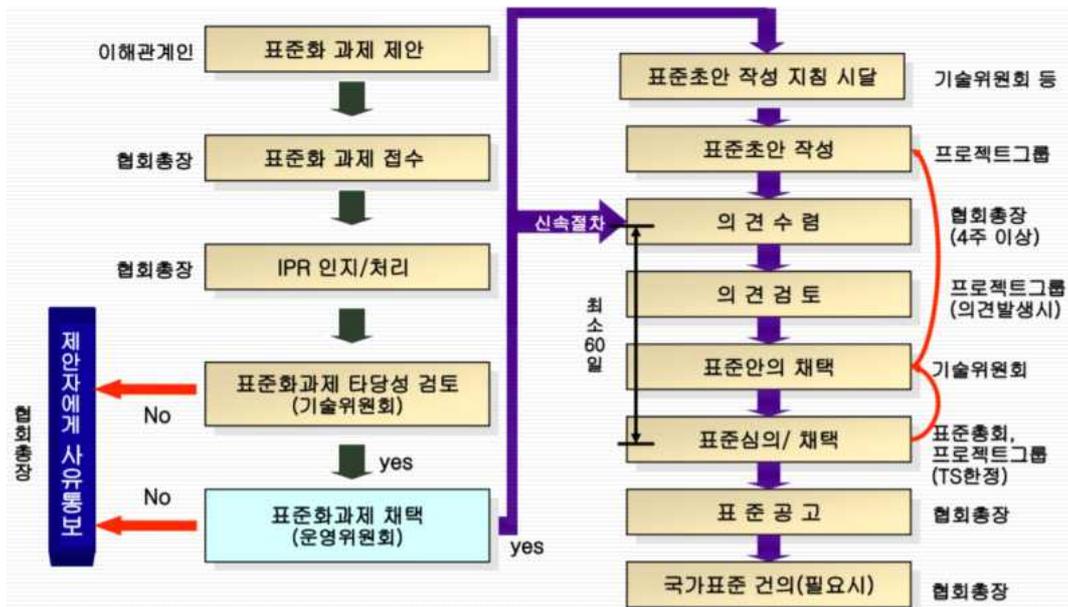
- 주요 수집 데이터 set : 외부기상, 내부온도, 습도, 양액 EC, PH, 양액공급시간 등 [데이터 수집]



■ 위탁연구기관 - 데이터스퀘어

○ 대상 작물별 데이터 명세를 고려한 데이터 표준화 개발 및 제안

- 기존 스마트팜 관련 표준화와 대상 작물별 데이터 명세를 고려한 스마트 생산환경 관리용 데이터 표준화 개발
- 국내 KAMICO(한국농기계공업협동조합) 또는 TTA PG426(한국정보통신기술협회)을 통한 단체표준에 스마트 온실 생산환경 관리용 데이터 표준화 제안
- 단체표준의 표준화과제 타당성 검토를 통한 스마트 생산환경 관리용 데이터 표준화 과제 추진



<한국정보통신기술협회(TTA) 정보통신단체표준 표준화 추진체계 및 절차 예시>

○ 데이터 분석 및 인공지능 모델 구축에 활용하기 위한 온실 생육·환경·영농 정보의 데이터 전처리(가공) 기술 개발

- 데이터의 불일치, 일관성 문제 등 불완전하고 모순된 데이터를 정제(Cleaning)하는 데이터의 무결성 검증 로직 개발

< 데이터 무결성 >

무결성	내용	비고
개체 무결성	▪ 기본키는 중복되거나 누락될 수 없는 특성(기본키는 유일성 보장)이 있음	Null값이 허용되지 않음
참조 무결성	▪ 외래키의 참조값은 다른 개체의 기본키 또는 NULL 값을 가짐	-
속성 무결성	▪ 속성값은 지정된 데이터 형식을 반드시 만족하는 값만 가짐	정의된 값만 입력 (Number, Character, Date 등)
도메인 무결성	▪ 특정 속성값은 미리 정의된 범위 내에 포함됨	허용된 범위 값만 입력
키 무결성	▪ 한 릴레이션에 동일한 키 값을 가진 튜플은 허용되지 않음	중복 배제

- 온실 생육·환경·영농 관련 수집 데이터를 통한 전처리 기술 시험 및 검증
- 대상 작물별 데이터 분석 및 인공지능 모델 구축에 활용할 수 있는 최적화 된 데이터베이스 구축
  - 농업인정보, 시설정보, 센서정보, 시설제어기 정보, 환경 정보, 작물 생육정보, 경영관리 정보 등 작물별 주요 온실 환경 정보 분석
  - 작물별 온실 생육·환경·영농 관련 데이터베이스 스키마(Schema)를 고려한 데이터베이스 구축
  - 데이터 전처리 기술을 통한 온실 생육·환경·영농 관련 데이터베이스 입력 및 저장

## [2단계]

### (3) 3차년도 : 스마트온실 빅데이터 분석 플랫폼(a-BRAIN) 실증 및 고도화

#### ■ 주관연구기관 - 순천대학교

○ 과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 실증



홍양 영농조합법인

전남 스마트팜 혁신밸리

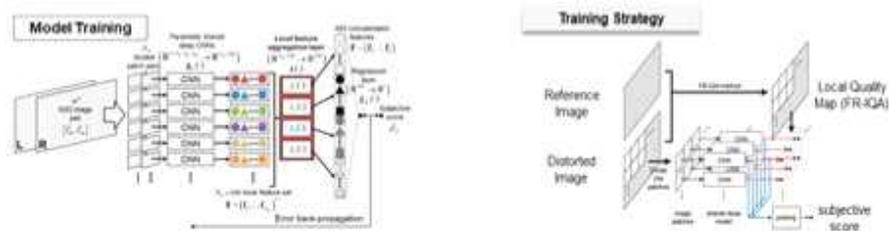
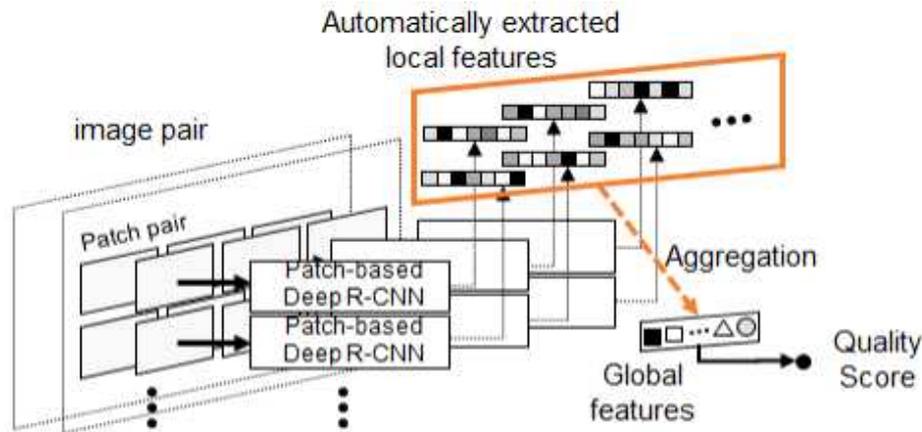
<인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 실증 사이트>

- 전남 스마트팜 혁신 밸리 TEST-BED 대상 온실 데이터 생성 모델 실증
- 무선 통신 통신기반 스마트 온실 빅데이터 플랫폼 실증
  - 스마트영농 실증 단지 내 실증 사이트 활용 협약
    - ✓ 구축 장비 : 통합관제센터 및 빅데이터시스템 등
  - 작물별 기 확보된 실증 사이트를 활용한 작물 데이터 생성 모델 실증 수행

< 실증 사이트 목록 >

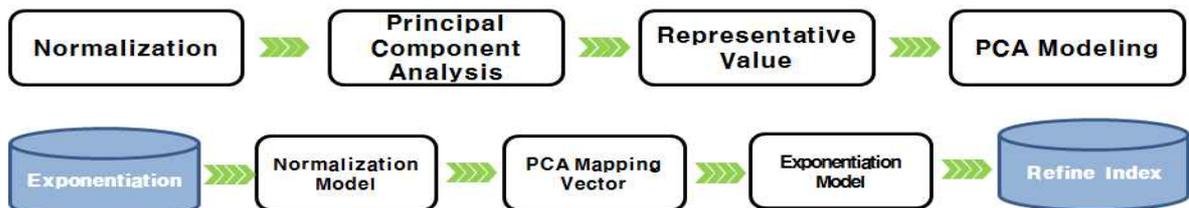
구분	농장명	소재지	품목구분	실증재배 작목	실증 가능년도	수집 가능 데이터
1	홍양영농 조합법인	고흥군 고흥만로	과채류	파프리카	전체 연구 기간	외부기상, 광량, 내부 온도, 습도, CO2농도, 양액공급량, 공급시간, EC, PH 등
2	테라영농 조합법인	고흥군 고흥만로	과채류	토마토	전체 연구 기간	외부기상, 광량, 내부 온도, 습도, CO2농도, 양액공급량, 공급시간, EC, PH 등
3	ALOH4 채소농장	고흥군 포두면 외산길	엽채류	로메인상추 바질 고추	전체 연구 기간	외부기상, 내부온도, 습도, 양액 EC, PH, 양액공급시간등
4	홍양영농 조합법인	고흥군 고흥만로	과채류	고추	2021년 8월 이후	홍양, 테라 와 외부기상은 공유하며 내부 환경 센서는 추후 보완 예정

- 표준화 기반 작물 병충해 데이터셋 인공지능 생성모델 실증 및 고도화
  - 작물별 실증 사이트 활용 e-BRAIN 및 로컬 사이트 수집 데이터 활용 작물 병해 데이터셋 생성 모델 실증 수행
  - 인공지능 모델 고도화를 위한 실데이터와 가상 데이터를 활용한 추론모델 재학습 수행
- 인공지능 네트워크 기반 작물 성장/생육/품질 예측 시스템 개발
  - 스마트 온실 작물의 영상 이미지 데이터를 활용한 작물 성장/생육/품질 분류 알고리즘 개발
  - No-reference Product Quality Assessment(NR-PQA) for Agricultural 알고리즘 개발



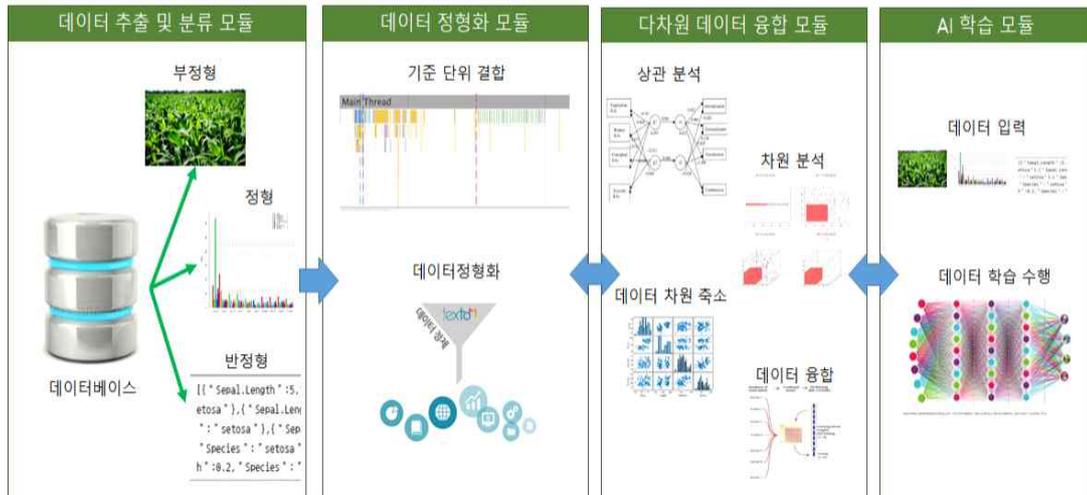
<Automatically extracted local features>

- 온실 작물 텍스트 데이터 검증을 위한 실증
  - 데이터 정제(결측치, 잡음 데이터 처리, 데이터 불일치, Field Overloading)
  - 데이터 통합(메타데이터, 상관성 분석, 데이터 충돌탐지, 의미적 이질성 해소)
  - 데이터 변환(Aggregation, Smoothing, Generalization, Normalization)
  - 데이터 축소(데이터큐브집계, 속성 subset 선택, 차원 축소, 수량 축소, 표본채취)
  - 데이터 이산화(히스토그램 분석, 그룹화 분석, 엔트로피식 이산화, 세분식 그룹화)



<텍스트 데이터 검증 모델 개발 절차>

○ 온실 작물 데이터 가공을 위한 인공지능 네트워크 실증



< 인공지능 기반 작물 데이터 수집/활용 및 예측을 위한 데이터 실증 구성도 >

■ 공동연구기관 1 - GS ITM

○ 클라우드 기반 스마트 온실 빅데이터 플랫폼 실증

- 데이터 개방 API 개발 : 클라우드 기반 스마트 온실의 빅데이터 플랫폼에서 분석 결과나 집계 및 통계된 데이터를 공공기관이나 전남 스마트팜 실증 단지 등에 오픈 API로 제공하여 데이터의 유효성을 실증
- 실시간으로 들어오는 온실의 데이터를 저장, 분석되는지를 테스트하고 정상 처리되는지를 실증
- 클라우드 기반 빅데이터 플랫폼의 성능을 실증함. 초기에 실증 과정에서 데이터가 적어서 빅데이터 플랫폼의 성능을 확인하지 못할 경우, 기존 데이터를 복사 및 변환하여 빅데이터 규모의 데이터로 임의 증가하여 성능을 실증
- 온실 작물의 생육/작황 상태를 기준으로 온실 환경 변수와 영농 인자의 유효한 상관 변수와 변수 간의 상관 관계를 실제 온실의 데이터를 적용하여 분석하고 이 결과를 영농 전문가와 협업하여 실제 상관성을 실증
- 클라우드 빅데이터 플랫폼에서 제공하는 인공지능 학습 모델의 학습 과정과 학습 후 모델의 배포, 배포된 모델에서 예측 결과가 도출되도록 테스트를 진행하고 실증함. 학습이나 예측 결과 도출 시 시스템의 성능을 테스트하여 시스템적으로 보완해야 하는 사항을 도출하고 이를 고도화

○ 실시간 온실 데이터 시각화 및 분석 Tool Box 실증

- 스마트 온실과 직접 연결하여 실시간으로 온실의 데이터가 시각화되는지 실증
- 스마트 온실의 데이터가 적합하고 유용하게 표현되는지, 시각화의 기능이 스마트 온실을 운영하는 데 적합한지 영농 전문가와 함께 실증함. 실증한 결과를 보완하여 스마트 온실 운영에 필요하고 유용하도록 보완
- 분석 파이프라인에서 도출되는 결과 시각화하여 보여주는 Tool Box를 영농 전문가와 같이 분석하여 상관 관계가 유효한지 실증하고 영농 전문가의 의견을 반영하여 보완

- 상관 분석을 보여주는 분석 Tool Box의 기능이 데이터 분석가들과 실증하여 정확한 데이터의 표현인지, 데이터 분석에서 유효한지 확인하고 데이터 분석가들의 의견을 반영하여 보완

### ■ 공동연구기관 3 - 흥양영농조합법인

#### ○ 무선 통신 기반 커넥티드 팜 및 실증 사이트 운영 및 스마트팜 표준화 데이터 수집

- 과채류
  - 흥양영농조합법인
  - 테라영농조합법인

- 엽채류
  - 알로하채소농장

#### ○ 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집

- 과채류 : 4종 (파프리카, 토마토, 고추, 딸기)
  - 주요 수집 데이터 set : 외부기상, 광량, 내부 온도, 습도, CO2농도, 양액공급량, 공급시간, EC, PH 등
- 엽채류 : 3종 (고수, 바질, 로메인상추)
  - 주요 수집 데이터 set : 외부기상, 내부온도, 습도, 양액 EC, PH, 양액공급시간 등

#### ○ 고품 스마트팜 혁신밸리 과채류 실증 단지 선정 및 실증

- 데이터수집



<실증사이트 현황>

### ■ 위탁연구기관 - 데이터스퀘어

#### ○ 대상 작물별 데이터 명세를 고려한 데이터 표준 제정

- 데이터 표준화에 대한 국내 KAMICO(한국농기계공업협동조합) 또는 TTA PG426(한국정보통신기술협회) 단체표준의 표준 초안 작성
- 데이터 표준화에 대한 KAMICO 또는 TTA PG426 단체표준 기술위원회 검토의견 반영 및 표준 제정

#### ○ 실증을 통한 데이터베이스 갱신 및 데이터베이스 최적화

- 실증을 통한 농업인정보, 시설정보, 센서정보, 시설제어기 정보, 환경 정보, 작물 생육정

보, 경영관리정보 등 작물별 주요 온실 환경 정보의 데이터베이스 갱신

- 실증 결과를 기반으로 데이터베이스 스키마의 적절성, 데이터 정제 및 무결성 검증 등을 통한 데이터베이스 구조·기능 보완 및 최적화

○ 스마트 팜 빅데이터 플랫폼 구조 조사 및 데이터 전송 기능 설계·구현

- 동 사업의 데이터 공유를 위한 스마트팜 빅데이터 플랫폼 구조 조사

- 스마트팜 빅데이터 플랫폼의 데이터 공유 기능(Open API 등)을 통한 온실 환경 데이터 전송 기능 설계 및 구현

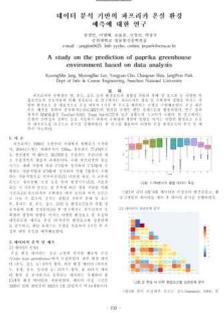
## 2. 연구개발과제의 수행 과정 및 수행 내용

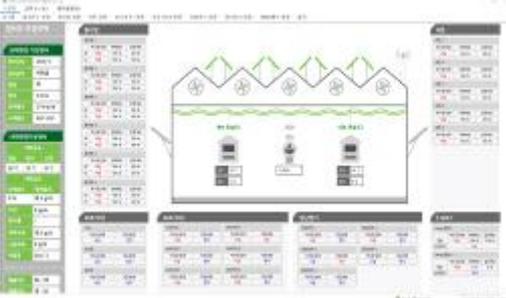
### 1) 세부별 연구 수행 결과

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
순천대학교	온실의 영상 데이터를 위한 자율형 인공지능 생성 모델 개발	과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 개발	<p>실증 대상 작물에 대한 기존 데이터 셋 기반 데이터 명세 정의</p> <p>작물 분류, 작물별 공통 데이터 명세에 따른 분석 데이터 생성 모델 개발</p>	<ul style="list-style-type: none"> <li>- 대상 작물 과채류 3종 (파프리카, 토마토, 고추), 엽채류 3종(이자벨, 페얼리, 스텔릭스) 이외에 과채류 1종(딸기), 엽채류 3종(로메인, 고수, 바질) 총 10종의 작물에 대한 데이터 명세 정의</li> <li>- 스마트 팜 데이터 표준 조사 및 공통 데이터 명세</li> <li>- 테스트베드(홍양영농조합법인, 테라영농조합법인, 알로하채소농장, 고흥 스마트팜 혁신밸리)의 기본 환경데이터 백업</li> <li>- 획득 데이터 기반 작물 별 공통 데이터 명세에 따른 데이터 생성 모델 개발</li> </ul>
		기존에 공개되어 병충해 데이터 세트와 기 확보되어 있는 데이터를 이용하여 인공지능 생성모델(GAN-Generative Adversarial Network) 개발	GAN 기반의 데이터 모델 구축	<ul style="list-style-type: none"> <li>- 기존 보유중인 병해충 이미지 데이터 분석 및 GAN 기반 이미지 생성 프로세스 정의</li> <li>- 데이터 모델 구축을 위한 병해 이미지 Data Augmentation 모델 구현</li> </ul> 
		생육 데이터(예, 과실)를 이용한 데이터 생성 모델 개발	과채류 생육 데이터 기반 생육 단계별 데이터 분류 체계 정의	<ul style="list-style-type: none"> <li>- 공개데이터 제공 플랫폼(AI hub, kaggle 등)에서 제공하는 생육 데이터를 활용하여 모델 학습 데이터로 활용</li> <li>- 공개 데이터를 기반으로 생육 단계 분류 체계 정의</li> <li>- 학술발표 2건, 비SCI 논문 2건</li> <li>· 스마트팜 환경데이터 기반 파프리카 생육 예측 모델 연구</li> <li>· 온도 변화에 따른 파프리카 성장 예측 모델을 위한 기계학습 알고리즘 비교 연구</li> <li>· 머신러닝 기반 시설재배 딸기 생산량 예측 연구</li> <li>· A Crop Growth Prediction Model Using Energy Data Based on Machine</li> </ul>



세부구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
인공지능 최적 모델 자동 탐색과 생성 파라미터 최적화 인공지능 네트워크 개발	자동 학습 시스템을 통해 데이터에 맞는 최적 학습 방법 선택 기술 개발	표준 데이터셋 명세를 기반한 스마트 농업 자동 머신러닝 모델 기술 개발	<p>표준 데이터셋 명세를 기반한 스마트 농업 자동 머신러닝 모델 기술 개발</p>	<p>- 의사결정트리 기반 자동 머신러닝 모델 기술 개발</p> <ul style="list-style-type: none"> <li>· 데이터 접근 방식에 따라 머신러닝 모델의 분류도 다양한 형태를 취함</li> <li>· 데이터셋 매개변수에 따른 의사결정트리 기반의 자동 머신러닝 모델 분류 진행</li> </ul> <p>- 학술발표 1건</p> <ul style="list-style-type: none"> <li>· Prediction of Greenhouse Air Vapor Pressure Deficit Based on Machine Learning</li> </ul> <p style="text-align: center;">Prediction of Greenhouse Air Vapor Pressure Deficit Based on Machine Learning</p> 
			<p>작물 데이터 분류별(작물 분류별, 작물별) 자동 학습 모델 개발</p>	<p>- 이미지 데이터 분류를 위한 최적 학습법 탐색</p> <ul style="list-style-type: none"> <li>· YOLO를 활용한 이미지 분류 모델 개발</li> <li>· CNN을 기반으로 한 Pretrained Model간 loss, auc 비교</li> <li>· 최적의 Pretrained Model 선정하여 작물데이터 분류 학습 모델에 적용</li> </ul> <p>- 학술발표 2건</p> <ul style="list-style-type: none"> <li>· YOLO를 이용한 유해조수 침입 감지 모델</li> <li>· Harmful Tide Intrusion Detection Model Using YOLO</li> </ul> <p style="text-align: center;">YOLO를 이용한 유해조수 침입 감지 모델</p> <p style="text-align: center;">Harmful Tide Intrusion Detection Model Using YOLO</p> 

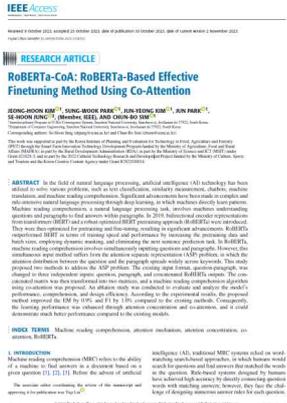
세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
			<p>원시 입력 데이터 분류 모델 및 의미 분석 모델 설계</p>	<ul style="list-style-type: none"> <li>- 가중치 분류를 위해 원시 데이터로부터 데이터셋 추출</li> <li>- 입력데이터 분류 모델 설계를 위해 원시데이터로부터 추출한 데이터셋을 사용한 상관관계 분석</li> <li>- 학술발표 1건</li> <li>· 데이터 분석 기반의 파프리카 온실 환경 예측에 대한 연구</li> </ul> 
		<p>학습된 인공지능 모델간의 중요도(가중치) 선택 인공지능 네트워크 개발</p>	<p>가중치 선택 인공지능 네트워크 개발을 위한 다계층 뉴럴 네트워크 모델 개발</p>	<ul style="list-style-type: none"> <li>- 최적의 다계층 뉴럴 네트워크 모델을 위한 Pretrained Model간 정확도 비교 및 모델 선정</li> <li>- 학술발표 1건, 비SCI 논문 1건</li> <li>· 호박잎 병해 진단 시스템을 위한 합성곱 신경망 기반 이미지 분류 연구</li> <li>· 딥러닝기반 토마토 병해 진단 서비스 연구</li> </ul>  

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
		과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 실증	전남 스마트팜 혁신 밸리 TEST-BED 대상 온실 데이터 생성 모델 실증	<ul style="list-style-type: none"> <li>- 실증농가 8번에 딸기(설향, 죽향) 정식 완료 및 빅데이터 수집을 위한 시스템 구비 완료</li> <li>· MAGMA 3.0</li> </ul>  <ul style="list-style-type: none"> <li>- 환경제어기, 양액기, 유통팬, FCU, 튜브레일, 행잉커터, 수평스크린, 천창 스크린 개폐장치, CO2 공급장치, 살균기 등 장비 구축 완료</li> </ul>
	온실 데이터 생성 및 파라미터 최적화 인공지능 네트워크 기술 실증 및 고도화	표준화 기반 작물 병충해 데이터셋 인공지능 생성모델 실증 및 고도화	작물별 실증 사이트 활용 e-BRAIN 및 로컬 사이트 수집 데이터 활용 작물 병해 데이터셋 생성 모델 실증 수행	<ul style="list-style-type: none"> <li>- Grad-CAM 기능을 통한 모델 성능 향상 방안 연구</li> <li>- epoch 변화에 따른 FID를 비교하여 FID값이 가장 낮은 최적 Epoch 탐색</li> <li>- 공개 데이터를 활용한 병해 이미지 생성 모델을 개발하여 a-BRAIN 플랫폼에 기능 탑재</li> <li>- 자체평가보고서 1건 완료</li> <li>· 자율형 영상 빅데이터 생성모델 적합도</li> <li>- PCT출원 1건</li> <li>· GAN을 이용한 작물 이미지 생성 및 작물 병해충 분류 모델 학습 시스템</li> <li>- 학술발표 1건</li> <li>· Grad-CAM을 활용한 식물 병해 이미지 객체 분류 분석</li> </ul> 
			인공지능 모델 고도화를 위한 실데이터와 가상 데이터를 활용한	<ul style="list-style-type: none"> <li>- EDA(Exploratory Data Analysis)를 통한 작물 데이터 분석</li> <li>- CycleGAN으로 생성된 병해 작물 이미지 추가 및 질병별 데이터 불균형</li> </ul>

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
			추론모델 재학습 수행	해결을 위한 가중치 설정 - 실데이터에 생성된 가상데이터를 추가하여 재학습을 통한 Inception v3 모델을 활용한 Disease Detection 모델 개발 - 학술발표 1건 · A Study on Honey Bee Object Detection Using DETR 
순천대학교 (공동)	미처리 온실 작물 관련 인공지능 네트워크 기술 개발	기본 전처리 기술(PCA 등)과 Auto Encoder 등의 데이터 가공 기술을 이용하여 데이터 주요 정보 추출 기술 개발	데이터 스케일링 기법을 활용한 데이터 정제 및 전처리 기술/AutoEncoder를 활용한 특징 추출 기술 개발	- 데이터 스케일링 기법을 활용한 데이터 정제 및 전처리 - AutoEncoder 기반의 데이터 처리 및 특징 추출 모델 개발 - SCI 급 논문 1건 · AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals 
		사전 학습된 인공지능 모델로 탐색(Detection), 분할(Segmentation)을 1차로 하고 2차적으로 보정하는 기술 개발	Mask R-CNN 기반 Segmentation 및 YOLO기반 Object Detection 모델 개발	- MSCOCO 데이터셋으로 사전학습된 MASK RCNN 모델을 활용한 Sementation 모델 설계 - YOLOv3,4 기반의 작물 생육단계 객체 탐지 모델 개발 및 성능평가 - SCI 급 논문 1건 · A Novel on Conditional Min Pooling and Restructured Convolutional Neural Network

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
				 <p>- 학술대회 1건</p> <ul style="list-style-type: none"> <li>· 이미지 컨투어 및 딥러닝을 이용한 엽면적 지수추정 모델 설계</li> </ul>  <p>- 특허 출원 1건</p> <ul style="list-style-type: none"> <li>· 딥러닝 기반의 객체 탐지 모델을 이용한 과수 생육 단계 판별 시스템</li> </ul> 
		픽셀 레벨의 작물 영상의 주석을 만들 때(Polygon) 수작업을 최소화 할 수 있는 polygon 포인트들의 추론 기술 개발	E-HRNet을 활용한 Semantic Segmentation 모델 개발	<ul style="list-style-type: none"> <li>- Squeeze 및 Excitation을 활용한 E-HRNet 기반의 Semantic Segmentation 모델 개발</li> <li>- 기존 벤치마크 데이터셋을 활용한 Semantic Segmentation 모델 성능 평가</li> </ul>
미처리 온실 작물 텍스트 자동 가공을 위한 인공지능	기존에 축적된 스캔 파일의 텍스트를 추출하기 위한 OCR 인공지능	영농일지 OCR 인공지능을 위한 텍스트 영역 추출 및 텍스트 인식 모델 개발		<ul style="list-style-type: none"> <li>- 흥양영농조합법인 영농일지 데이터 수집</li> <li>- VGG, ResNet, GRCL 기반 특징 추출 모듈 개발</li> <li>- BiLSTM 기반 Sequence Modeling</li> </ul>

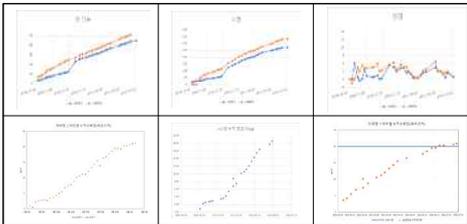
세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
	네트워크 개발	네트워크 개발		<p>모듈 개발</p> <ul style="list-style-type: none"> <li>- 학술대회 1건, 특허출원 1건</li> <li>- 영농일지 디지털화를 위한 딥러닝 기반 문자 인식 모델 설계</li> <li>- 인공지능 기반의 광학적 문자 판독 모델을 활용한 스마트팜 영농데이터 변환 시스템</li> </ul> 
		농업 텍스트 정보 전처리 및 가시화 기능 개발	농업 텍스트 정보 저장 및 활용을 위한 데이터 전처리 및 가시화 기능 개발	<ul style="list-style-type: none"> <li>- 정규표현식을 활용한 특정패턴에 대한 텍스트 전처리 및 매핑 모듈 개발</li> <li>- 텍스트 정보 및 BBox 정보를 활용한 OCR 인식 시각화 모듈 개발</li> <li>- KCI 급 1건</li> <li>- I-QANet: 그래프 컨볼루션 네트워크를 활용한 향상된 기계독해</li> </ul> 
	전이학습 모델을 이용한 자동 NLP 주석 툴 개발	사전학습된 모델을 활용한 전이학습기반 NLP 주석 툴 개발	<ul style="list-style-type: none"> <li>- 한국어 인식 성능 및 학습 속도 향상을 위한 전이학습 모델 개발</li> </ul>	
	미처리 온실 작물 텍스트자동 가공을 위한 인공지능 네트워크 실증 및 고도화	데이터 정제(결측치, 잡음 데이터처리, 데이터불일치, Field Overloading) 데이터 통합(메타데이터, 상관성 분석, 데이터 충돌탐지, 의미적 이질성 해소)	<ul style="list-style-type: none"> <li>- 영농일지 OCR 프로그램 실증을 위한 a-BRAIN 플랫폼 적용</li> <li>- 저작권등록 1건</li> <li>- 영농일지디지털화프로그램</li> </ul>	

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
			<p>데이터 변환(Aggregation, Smoothing, Generalization, Normalization)</p> <p>데이터 축소(데이터큐브집계, 속성 subset 선택, 차원 축소, 수량 축소, 표본채취)</p> <p>데이터 이산화(히스토그램 분석, 그룹화 분석, 엔트로피식 이산화, 세분식 그룹화)</p>	 <p>– SCIE급 1건</p> 
		<p>온실 작물 데이터 가공을 위한 인공지능 네트워크 실증</p>		<p>– 온실 작물 데이터 가공을 위한 객체 탐지 및 분할 모델 결합 및 고도화</p> <p>– 객체 탐지 및 분할 인공지능 네트워크 실증을 위한 a-BRAIN 플랫폼 적용</p> <p>– 학술대회 1건</p> <ul style="list-style-type: none"> <li>· Mask R-CNN 기반 토마토 객체 분할 및 생육단계 판별 모델에 관한 연구</li> </ul>  <p>– SCIE급 1건</p> <ul style="list-style-type: none"> <li>· E-HRNet, Enhanced Semantic Segmentation Using Squeeze and Excitation</li> </ul>

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
				 <p>- 특허등록 1건 · 딥러닝 기반의 객체 탐지 모델을 이용한 과수 생육 단계 판별 시스템</p> 
		<p>인공지능 네트워크 기반 작물 생장/생육/품질 예측 시스템 개발</p>	<p>스마트 온실 작물의 영상 이미지 데이터를 활용한 작물 생장/생육/품질 분류 알고리즘 개발</p> <p>No-reference Product Quality Assessment (NR-PQA) for Agricultural 알고리즘 개발</p>	<p>- 작물 영상 및 시계열 데이터를 활용한 Multimodal 작물 생장/생육/ 품질 분류 알고리즘 설계 및 개발</p> <p>- Multimodal 융합 방법 연구 및 성능평가 수행</p> <p>- Few-shot 및 Zero-shot 기반 품질 측정 알고리즘 설계 및 개발</p> <p>- prototypical networks 기반 모델 구현 및 성능평가 수행</p>
GSITM	클라우드 기반 스마트 온실 빅데이터 시스템 설계 및 구축	클라우드 기반 스마트 온실 빅데이터 수집/저장/관리/분석할 수 있는 빅데이터 플랫폼 설계	빅데이터 플랫폼 구조 설계 및 저장소 설계	<p>- 빅데이터 플랫폼 구조 설계 완료</p> <p>- 빅데이터 플랫폼 하드웨어/소프트웨어 인프라 구축 완료</p> <p>- 아마존(AWS) EC2 리눅스 서버 구축</p>

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
		소프트웨어 인프라 구축		<ul style="list-style-type: none"> <li>및 네트워크 구성</li> <li>- 아마존(AWS) RDS 구성(MariaDB)/환경 설정 및 튜닝</li> </ul>
		스마트온실 데이터 상관관계 분석 파이프라인 설계	수집-분석을 위한 파이프라인 구성 및 자동화 설계	- Airflow를 활용한 프라이프라인 설계 및 Dags 관리 규칙 정의
		생육-환경-영농 데이터 사이의 상관관계 분석 설계	작물과 환경 정보의 상관 분석 방안 설계	<ul style="list-style-type: none"> <li>- 다양한 환경 데이터들에 대한 다변량 상관 분석</li> <li>- 환경데이터와 영농데이터의 선형적 관계 분석 방법으로 Kendal, Spearman, Person 등 활용</li> </ul>
		오믹스 데이터를 확보할 수 있는 작물의 경우, 생명 정보와의 상관관계를 파악하여 더욱 정밀한 의사결정 및 관리지원 달성	확보된 작물에 대한 분석을 위한 데이터 자동 정비	<ul style="list-style-type: none"> <li>- 선형데이터 여부 체크</li> <li>- 데이터 사이의 연관성 정의</li> <li>- 에러 항목 데이터 처리</li> <li>- 데이터 전처리(autocorrelation 항목 제거 및 상호 선형적 특징 큰 값 제거</li> <li>- VIF(Variance Inflation Factor) 상한 10 이상 값 제거</li> </ul>
클라우드 기반 스마트 온실 빅데이터 시스템 구축	클라우드 기반 스마트 온실 데이터 수집/적재/분석 개발	표준화된 데이터 기반 빅데이터 저장 구조 및 처리 모듈을 개발	<ul style="list-style-type: none"> <li>- 환경정보 ==&gt; RDBMS 적재 (Redshift, MySQL)</li> <li>- 생육정보 파일 데이터 ==&gt; AWS S3 Bucket 기반 DataLake 적재</li> <li>- 이미지 데이터 ==&gt; AWS S3 Bucket 기반 DataLake 적재</li> <li>- API를 통한 데이터 수집과 기존 적재된 데이터는 Bulk 형태로 이관</li> <li>- AWS S3 Bucket 기반의 DataLake에 데이터 적재</li> </ul>	
	스마트온실 데이터 상관관계 분석 파이프라인 개발	온실의 환경-생육-영농 인자의 상관 분석을 위한 분석 파이프라인 모듈을 개발하고 이를 분석 모듈의 컨테이너에 넣어서 관리	- 데이터 상관관계 분석을 위한 ETL(Extraction/Transaction/Loading) 프로세스를 개발하고 관리를 위한 파이프라인 환경 구축	

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
		생육-환경-영농 데이터 사이의 상관관계 분석 개발	의사결정에 영향을 주는 데이터간의 유효한 상관관계를 분석하고 결과를 저장	- 데이터간 상관 분석 : Correlation 분석 및 OLS(Ordinary Least Square) 분석을 통해 데이터 간의 상관 관계를 분석하고 상관 관계가 유효한 변수들을 추출
실시간 데이터 시각화 및 분석 기능을 위한 웹기반 Toolbox 개발	스마트 온실 데이터 조회 및 다운로드, 실시간 모니터링, 온실 간 Head-to-Head 비교 기능 개발		실시간으로 수집된 데이터를 조회/검색/다운로드/모니터링을 위한 웹기반의 Portal 개발	- 정기 스케줄러를 통해 수집된 데이터 데이터 조회/다운로드 제공 웹기반 포털 개발
		온실의 환경 정보의 실시간 시각화를 위한 Tool Box와 상관 분석의 파이프라인 결과를 시각화로 표현하기 위한 Tool Box를 개발	- 온실 환경 정보 시각화 대시보드 개발 - 웹기반 Portal을 위한 서버 구축 및 사용자 UI를 개발 - 수집/조회/분석된 데이터는 엑셀(또는 PDF)로 다운로드가 가능	
	환경-생육-영농 정보를 연결하고 동시에 분석·활용할 수 있는 실시간 시각화 인터페이스 개발	전문기술자만 해석 가능한 상관관계 분석결과를 보기 쉬운 방법(Chart, Plot 등)으로 변환하는 시각화 인터페이스를 개발	- 오픈소스 기반의 툴을 활용한 대시보드와 데이터 테이블 개발 - 웹기반 Portal은 HTML5, Chart, plot, Table을 이용하여 데이터를 표현하며, 이를 위해 JavaScript Module을 사용 - 사용자 분석에 활용할 수 있는 환경을 고려하여 Tool Box 구축	
	커넥티드 팜 데이터를 활용한 개별 생육-환경-영농 인자분석 도구 개발	개발된 모델링과 커넥티드 팜에서 수집된 데이터 연계를 위한 프로세스 개발 상관분석 및 모델링 결과로 추출된 결과 데이터를 다양한 방법으로 시각화	- 상관 분석을 위한 분석 파이프라인 과정에서 생성되는 분석 결과 데이터를 시각화하여 표현해 주며, Correlation Matrix, 회귀 분석 Chart, HeatMap Chart, KDE Chart, Scatter Chart 등 다양한 시각화 기능을 제공하여 데이터의 분포와 상관성 판단 가능	
실시간 데이터	분석 파이프라인의	고흥혁신 밸리 내의 스마트 온실과 연동	- 고흥 스마트팜 혁신밸리의 딸기, 토마토 데이터 연동	

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
	시각화 및 분석 기능을 위한 웹기반 Toolbox 개발 및 실증	분석 결과의 시각화 실증		- 실제 플랫폼에 데이터가 적재되도록 파이프라인 개발
		영농 전문가와 시각화 요소의 적절성 평가/보완	데이터 연동 결과, 시각화 화면에 대한 적정성 평가	- 실제 데이터를 화면에 표출 - 적재된 데이터만 표출하는 것으로는 분석이 부족하여 데이터 분석이 용이하도록 작기별 / 작물별 분석 차트/데이터 마트 개발
	클라우드 기반 스마트 온실 빅데이터 플랫폼 실증	표준화에 입각한 데이터의 수집/저장/연동에 대한 실증	추가되는 데이터 연동에 필요한 데이터 컬럼 자동 생성	- 데이터연동 확장성을 위해 적재 공간 자동 생성 - 추가 데이터(혁신밸리 데이터) 연동 시, 표준 데이터 기반으로 데이터 마트 생성 완료
		지능형 모델의 학습/배포 실증	인공지능 모델 플랫폼 연동	- 3개 모델을 플랫폼에 탑재 - 실제 데이터를 업로드 후 모델 결과 확인 기능 완료
	스마트 온실 데이터 수집 및 실증 기술 설계	커넥티드 팜 데이터 수집 및 실증 사이트 그룹 선정	과채류(토마토, 파프리카) 실증 사이트 데이터 수집 및 재배 일지 작성	- 토마토 및 파프리카 재배 일지 작성 - 재배 일지 중 이슈 사항 정리 
		과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 설계	과채류(토마토, 파프리카) 실증 사이트 생육 현황	- 작물 생육 데이터 2020.09.23. ~ 2021.07.15. 7일 단위로 수기로 입력 
커넥티드 팜 데이터 수집 및 실증 사이트 구축		실증 사이트를 통한 실시간 데이터 수집 환경 구축	- 과채류 생육데이터 수집을 위한 API 설치 - 온실 생육 의사 결정 지원을 위한 양액 배액량 및 성분 측정 설비 구축 · SW 등록 1건	

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
				 <p>- 알로하 채소농장 업체류 생육데이터 수집용 영상수집 환경 구축</p>
		과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 구축	과채류 3종 (파프리카, 토마토 고추) 엽채류 6종 (고수, 바질, 로메인 상추, 페얼리, 이자벨, 스텔릭스) 빅데이터 수집 현황	<p>- 과채류 3종, 엽채류 6종에 대한 현장 데이터 수집</p> <ul style="list-style-type: none"> <li>· 파프리카 : 2020. 4.~현재 (실시간 데이터 수집 중), 8개 블록의 Blockout screen 등 27개 항목(세부 항목 약 300개)의 데이터 수집</li> <li>· 토마토 : 2020. 4.~현재 (실시간 데이터 수집 중), 8개 블록의 Blockout screen 등 27개 항목(세부 항목 약 300개)</li> <li>· 고추 : 2020.01.~2022.06.(30개월)의 환경데이터 수집. 22개의 세부 항목</li> <li>· 엽채류 6종(고수, 바질, 로메인 상추, 페얼리, 이자벨, 스텔릭스) : 종류별 2개월의 환경데이터 수집. 13개의 세부 항목</li> </ul> <p>※ 엽채류의 경우 재배주기가 짧고 종류가 다양하여 데이터 수집기간이 2개월 이내임</p>
	스마트 온실 데이터 수집 및 실증 사이트 운영	무선 통신 기반 커넥티드 팜 및 실증 사이트 운영 및 스마트팜 표준화 데이터 수집	실증사이트 운영 및 스마트팜 표준화 데이터 수집 진행	<p>- 실증사이트 표준화 데이터 수집 현황</p> <ul style="list-style-type: none"> <li>· 과채류 2종 환경 및 제어데이터 API 연계를 통한 DB 적재</li> <li>· 과채류 1종 및 엽채류 6종 환경 데이터 csv파일을 활용한 DB 수동 적재</li> </ul> <p>- 스마트팜 R&amp;D 빅데이터 플랫폼에서 정의한 데이터 표준을 기반으로 데이터 정제</p> <p>- 실증 사이트를 대상으로 한 a-BRAIN 플랫폼 개발에 따른 시범 적용 및 운영</p>
		과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 사이트 구축	과채류 및 엽채류 실증 사이트 및 대상작물 확대	<p>- 고흥 스마트팜 혁신밸리 실증단지 선정에 따른 실증단지 및 실증작물(과채류 1종) 추가고흥 스마트팜 혁신밸리 과채류 및 엽채류 실증 단지 활용 실증</p> <p>- 흥양영농조합법인(파프리카), 테라영농</p>

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과						
		과채류 대상 스마트 온실 빅데이터 수집 사이트 구축	고흥 스마트팜 혁신밸리 실증단지 작물 재배 및 데이터 수집 수행 중	<p>조합법인(토마토), 알로하 채소농장(고수, 바질, 로메인상추, 페얼리, 이자벨, 스텔릭스), 고흥 스마트팜 혁신밸리 실증단지(딸기) 총 4개 실증단지에서 10종 데이터 수집</p>						
				<p>- 기 선정되어있던 실증사이트 외 고흥 스마트팜 혁신밸리 실증단지 선정으로 빅데이터 수집 사이트 추가 구축 완료</p> <ul style="list-style-type: none"> <li>· 온실8번(유리온실, 1,296m<sup>2</sup>(392))</li> <li>· 계약일 : 2023.11.20.</li> <li>· 이용 기간 : 2023.11.~ 2026.11.(36개월)</li> </ul> <p>- 과채류 1종 정식 완료</p> <ul style="list-style-type: none"> <li>· 정식일 : 2023.11.24.</li> <li>· 실증작물 : 딸기(죽향, 설향 2종)</li> </ul> <p>[별지 제2호] 이용계약서  <b>전남 고흥 스마트팜 혁신밸리 실증단지 이용계약서</b>  전남 고흥 스마트팜 혁신밸리 실증단지 이용계약서를 위해 실증단지 운영자(전남 알로하농업)와 조합법인(알로하)은 실증단지 운영자(알로하)와 조합법인(알로하)의 "이용계약"을 다음과 같이 체결하고, 실증단지(온실 8번) 부지 재배(이하 "이용계약")는 다음과 같이 체결을 하였다.</p> <p><b>당해 일</b>  <b>운영기간</b> : 전남테크노파크 원 장 오 지 영  <b>당해 일</b>  <b>주 관 기 관</b> : 순천대학교 산학협력단 원 장 서 행  <b>주 관 기 관</b> : KINARA 연구개발팀, 김현아</p> <p><b>제1조(목적)</b> 본 계약은 운영자(이하 "스마트팜 혁신밸리 실증단지"라 함 "실증단지") 이용 과 산 실증단지 이용에 관한 계약사항 규정을 목적으로 한다.</p> <p><b>제2조(정의)</b> ① 운영기간은 당해 일 실증단지 운영을 위해 노력하고 이용자는 실증단지의 일정한 운영을 위해 노력한다.  ② 운영자(이하 "운영자")는 실증단지 운영을 위해 노력한다. 본 계약에서 운영자가 맡은 사항은 "전남 고흥 스마트팜 혁신밸리 실증단지 운영규칙" (이하 "운영규칙") 및 "전남 고흥 스마트팜 혁신밸리 실증단지 운영규칙" (이하 "운영규칙")에 의해서 운영되는 것으로 간주한다.  ③ 실증단지 및 운영자(이하 "운영자")는 본 계약의 운영목적에 따라 운영한다. 본 계약이 맺어 있는 동안 "운영규칙", "운영지침"의 범위내에서 상호 존중을 통한 협의를 위해 일방적행위를 수행할 수 없다.</p> <p><b>제3조(이용 및 계약기간)</b> ① 본 계약에 의한 이용 및 계약기간은 다음 표와 같다.</p> <table border="1"> <tr> <td>이용 기간</td> <td>2023. 11. 20. ~ 2026. 11. 19. (36개월)</td> </tr> <tr> <td>이용 범위/이용</td> <td>온실 8번 운영계약서상</td> </tr> <tr> <td>이용료(월/일)</td> <td>0원/0원</td> </tr> <tr> <td>주요시설(월/일)</td> <td>0원/0원</td> </tr> <tr> <td>합계(월/일)</td> <td>0원/0원</td> </tr> </table> <p>* 운영기간은 실제 운영일  ② 운영기간은 이용기간 종료 3개월 전에 이용자에게 이용기간 만료 사실을 서면통지하여야 한다.</p>	이용 기간	2023. 11. 20. ~ 2026. 11. 19. (36개월)	이용 범위/이용	온실 8번 운영계약서상	이용료(월/일)	0원/0원
이용 기간	2023. 11. 20. ~ 2026. 11. 19. (36개월)									
이용 범위/이용	온실 8번 운영계약서상									
이용료(월/일)	0원/0원									
주요시설(월/일)	0원/0원									
합계(월/일)	0원/0원									
데이터 스캐어	데이터 표준화 및 통합 데이터베이스 설계	대상 작물별 데이터 명세 정의 및 표준데이터 범위 선정	기존 스마트팜 관련 데이터 표준화 현황 조사 및 분석	<p>- 기존 스마트팜 관련 표준화와 본 사업의 전문가 의견을 고려한 데이터 표준화 설계(작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보)</p>						
			스마트 온실 생산환경 관리를 위한 대상 작물별 측정 항목 조사 및 데이터 명세 정의							
			온실 생육·환경·영농 정보 등에 대한 각 데이터 특성 및 구조 분석							
데이터 분석 및 인공지능 모델 구축에 활용하기 위한 온실 생육·환경·영농 정보의 불완전하고 모순된 데이터를 정제(Cleaning)하는	<p>- 수집 데이터에 대한 정제가 요구되는 항목 도출 및 데이터 무결성 검증 로직 설계(데이터 구조·입력값 범위 관련 무결성 검증)</p>									

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
		데이터 전처리(가공) 기술 고안	데이터의 무결성 검증 로직 고안	
		데이터 분석 및 인공지능 모델 구축에 활용할 수 있는 최적화 된 데이터베이스 설계	온실 생산환경 관리용 데이터베이스 구축을 위한 데이터 현황 분석	- 스마트 온실 생산환경 관리용 개체관계 모델(ER 다이어그램) 도출 및 해당 모델 기반 데이터베이스 스키마 설계
	효과적인 온실 생산환경 관리를 지원할 수 있는 데이터베이스 설계			
	데이터 표준화 개발·제안 및 통합 데이터베이스 구축	대상 작물별 데이터 명세를 고려한 데이터 표준화 개발 및 제안	반밀폐 온실의 특성을 고려한 데이터 표준화 추진	- 반밀폐 스마트 온실 대상 데이터 항목 표준 작성 및 해당 표준을 농업기술실용화재단에 제안 - 데이터 표준화 (실용화재단 단체 표준) 추진 중
		데이터 분석 및 인공지능 모델 구축에 활용하기 위한 온실 생육·환경·영 농 정보의 데이터 전처리(가공) 기술 개발	데이터의 불일치, 일관성 문제 등 불완전하고 모순된 데이터를 정제(Cleaning)하는 데이터의 무결성 검증 로직 개발	- 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보로 구성된 데이터 항목에 대해 수집된 데이터의 신뢰성 및 정확성을 지원하는 검사 기준 및 방법 기반 데이터 무결성 검증 로직 개발
		대상 작물별 데이터 분석 및 인공지능 모델 구축에 활용할 수 있는 최적화 된 데이터베이스 구축	작물별 온실 생육·환경·영농 관련 데이터베이스 스키마(Schema)를 고려한 데이터베이스 설계 수정 및 데이터베이스 구축	- 1차년도 데이터베이스 스키마의 추가·수정 추진(영농일지, 이미지 획득장소, 제어정보(제어데이터, 장치유형 및 값) 등 수정 보완) - 본 사업에서 수집가능한 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보를 기반으로 데이터베이스 재설계 및 개체관계 모델 수정 및 데이터베이스 구축 - 흉양 및 알로하 농장 데이터 준실시간 환경 데이터 수집 및 생육 정보는 일괄 등록 - 이미지, 동영상 등 반정형/비정형 데이터는 클라우드 저장소에 원본을 저장하고, 관련 메타 데이터는 RDBMS에 저장

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
	데이터 표준화 제정 및 통합 데이터베이스 최적화	대상 작물별 데이터 명세를 고려한 데이터 표준 제정	데이터 표준화를 위한 국내 단체표준기관 제출용 표준 초안 작성 및 제출	<ul style="list-style-type: none"> <li>- 반밀폐 스마트 온실 대상 데이터 항목 표준 초안 작성 및 제출</li> <li>- 공통데이터 항목은 기본적으로 선행표준인 SPS-X KOAT-0009-7470(시설원에 분야 스마트팜 수집 데이터 규격)을 인용하여 용어, 정의, 단위를 통일</li> <li>- 기존 선행표준에 반밀폐 온실에 사용되는 용어 및 장비 정보를 추가한 상세 메타데이터를 정의</li> <li>- “반밀폐 스마트 온실 대상 수집 데이터 규격” 표준을 한국농업기술진흥원(농업기술실용화재단)으로 제안</li> </ul>
			데이터 표준화를 위한 단체표준 위원회 및 포럼 참가를 통한 검토의견을 반영한 표준 제정(개정)	<ul style="list-style-type: none"> <li>- 스마트팜 ICT 융합 표준화 포럼(제 7차 정기총회)를 통해 선행표준인 SPS-X KOAT-0009-7470(시설원에 분야 스마트팜 수집 데이터 규격)에 반밀폐 특화 정보의 추가하여 개정으로 진행할 것을 권고 받음</li> <li>- 선행표준 제정기관인 농림수산식품교육문화교육원과 협의를 통해 시설원에 분야 스마트팜 수집 데이터 규격 표준에 반밀폐 온실에 사용되는 용어 및 장비 정보를 추가하여 자체심의, 전문가 심의 등 진행</li> <li>- 2023년 KoAT 단체표준(안) 제정 및 개정 관련하여 예고고시를 농진원 홈페이지를 통해 진행</li> <li>- 2023.12.29.일 부터 2024.01.30.일 까지 총 30일간 진행</li> </ul> <div style="text-align: center;">  <p>시설원에 분야 스마트팜 수집 데이터 규격 SPS X KOAT 0009-7470:2024</p> <p>한국농업기술진흥원 2024년 1월 1일 개정</p> </div>
		실증을 통한 데이터베이스 갱신 및 데이터베이스	실증을 통한 농업인정보, 시설정보, 센서정보, 시설제어기 정보,	<ul style="list-style-type: none"> <li>- 과채류·엽채류 생육 및 반밀폐 온실 특성을 고려한 데이터스키마 추가·보완</li> <li>- 실제 수집 데이터 및 데이터스키마 간</li> </ul>

세부 구분	세부과제명	세부연구목표	연구개발 수행내용	연구결과
		최적화	환경 정보, 작물 생육정보, 경영관리정보 등 작물별 주요 온실 환경 정보의 데이터베이스 갱신	매핑 테이블을 작성하여 작물환경정보, 시설정보, 제어정보 등을 고려한 작물별 주요 온실 환경 정보 및 생육정보의 데이터베이스 구축
			실증 결과를 기반으로 데이터베이스 스키마의 적절성, 데이터 정제 및 무결성 검증 등을 통한 데이터베이스 구조·기능 보완 및 최적화	- 데이터스키마(생육정보 등) 수정으로 인한 품질관리(데이터 정제 등) 기능 보완 - 보완된 품질관리 도구를 통한 실증 데이터 검증 및 최적화 수행
		스마트 팜 빅데이터 플랫폼 구조 조사 및 데이터 전송 기능 설계·구현	동 사업의 데이터 공유를 위한 스마트 팜 빅데이터 플랫폼 구조 조사	- 스마트팜연구개발사업단의 "스마트팜 R&D 빅데이터 플랫폼 연구데이터 제출 및 작성안내서"를 통해 데이터 공유를 위한 스마트 팜 빅데이터 플랫폼 구조 조사 - 동 작성안내서의 연구데이터 분류·파일명 생성규칙·데이터셋 정의·비정형 파일 제출 등 세부 연구데이터 제출 내용 검토
			스마트 팜 빅데이터 플랫폼의 데이터 공유 기능(Open API 등)을 통한 온실 환경 데이터 전송 기능 설계 및 구현	- 동 작성안내서의 연구데이터 분류·파일명 생성규칙·데이터셋 정의·비정형 파일 제출 등 세부 기준을 고려하여 정형 및 비정형 데이터 업로드

## 2) 연구 수행 세부 내용

### [순천대학교\_주관]

#### [1단계(2021)]

#### □ 과채류 및 엽채류 각 3종 온실 작물에 대한 데이터 생성 모델 개발

##### ○ 실증 대상 작물에 대한 기존 데이터 셋 기반 데이터 명세 정의

##### • 흥양영농조합 토마토 및 파프리카 온실

- 흥양영농조합의 경우 토마토는 반밀폐온실에서 재배되고 있으며, 파프리카는 비닐온실에서 재배 중
- 온실 데이터 수집을 위한 시스템은 네덜란드 HORTIMAX 사의 SYNOPTA제어 프로그램 구축
- 네덜란드 온실시공업체인 쿠보의 설계에 의해 설치되어 있으며, 생육환경 데이터 및 영농일지 작성 진행



#### <흥양영농조합법인 현황>

#### <흥양영농조합법인 데이터 수집 항목 리스트>

구분	변수	단위
온실외부 기상환경	일사량	W·m <sup>-2</sup>
	기온	℃
	상대습도	%
	풍향	E/W/S/N
	풍속	m·s <sup>-1</sup>
온실내부 기상환경	기온	℃
	습구온도	℃
	상대습도	%
	절대습도	g/kg
	습도부족분	g/kg
	CO <sub>2</sub> 농도	ppm
	이슬점	℃
온실 공기의 증기압 결핍	kPa	

구분	변수	단위
토양(근권)환경	엔탈피	kJ/kg
	양액 수온	℃
	양액 pH	pH
	양액 EC	dS·m <sup>-1</sup>
제어 환경	난방 설정 온도	℃
	냉방 설정 온도	℃
	센서 작동 상태	active/inactive
	양액통 상태	full/empty
	팬 작동 상태	on/off
생육 환경	초장	cm
	생장길이	cm
	잎길이	cm
	잎폭	cm
	잎수	장
	줄기굵기	cm
	화발높이	cm
	개화군	개
	착과군	개
	열매수	개
평균과중	g/개	

• 알로하 농장 엽채류 온실 데이터 수집

- 알로하 농장의 경우 비닐온실 환경에서 엽채류들을 아쿠아포닉스 방식으로 재배
- 온실 데이터 수집을 위한 시스템은 국내 그린씨에스사의 마그마 프로그램 구축
- 양액 데이터 및 기타 장치 제어는 직접 개발한 IoT 장비를 구축
- 작물의 육묘단계에서부터 관리를 하며, 육묘는 습도 90%이상의 환경에서 정식 때까지 2~6일 육성
- 아쿠아포닉스 방식으로 재배하여 작물 뿌리에 분단위로 물을 뿌려 주기 때문에 여름철 따로 냉방을 진행하지 않음
- 작물을 출하하기까지 걸리는 기한이 30일 이내로 굉장히 짧아 따로 생육데이터 수집은 하지 않음



<알로하 채소농장 현황>

**<알로하채소농장 데이터 수집 항목 리스트>**

구분	변수	단위
온실외부 기상환경	일사량	W·m-2
	적산일사량	J/cm-2
	온도	℃
	상대습도	%
	풍향	E/W/S/N
	풍속	m·s-1
	감우	비움(1)/비안움(0)
	일출 시간	시/분
	일몰 시간	시/분
온실내부 기상환경	온도	℃
	습도	%
	포화수분	g/m-2
	절대습도	g/m-2
	수분부족	g/m-2
	이슬점	℃
토양(근권)환경	양액 수온	℃
	양액 pH	pH
	양액 EC	dS·m-1
제어 환경	난방 설정 온도	℃
	환기설정	℃
	2중 환기설정	℃

○ 작물 분류, 작물별 공통 데이터 명세에 따른 분석 데이터 생성 모델 개발

- 흥양 및 알로하 공통 데이터 명세

- 기존 데이터의 명세를 기반으로 아래의 표의 내용과 같이 공통 데이터를 추출

**<흥양 및 알로하 공통 데이터 항목 리스트>**

구분	변수	단위
온실외부 기상환경	일사량	W·m-2
	온도	℃
	상대습도	%
	풍향	E/W/S/N
	풍속	m·s-1
온실내부 기상환경	온도	℃
	습도	%
	절대습도	g/m-2
	수분부족	g/m-2
	이슬점	℃
토양(근권)환경	양액 수온	℃
	양액 pH	pH
	양액 EC	dS·m-1
제어 환경	난방 설정 온도	℃

- 작물별로 특정데이터를 추가로 수집하는 경우는 없었으며, 시스템을 구축하는 회사에 따라 구축된 센서에 따른 차이 발생

- 두 온실 모두 기본적인 온실 환경데이터 및 양액데이터를 수집하고 있었으며, 과채류의 경우

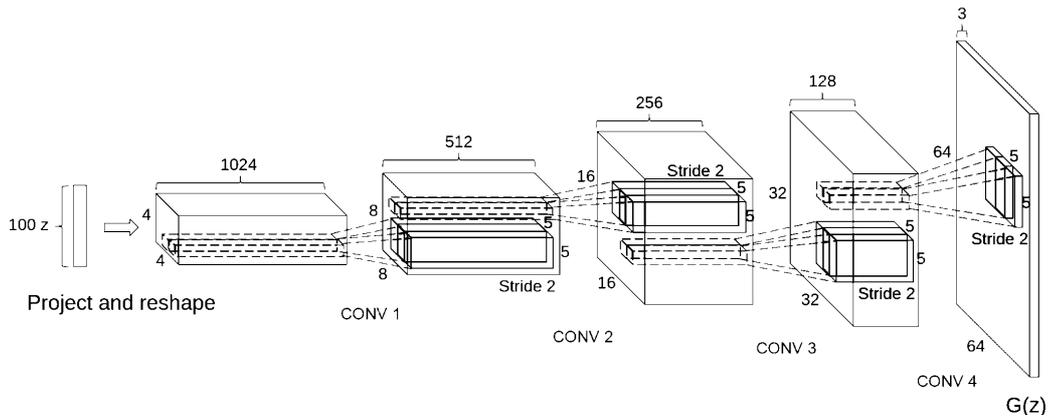
재배기간이 1년으로 1주일 단위로 수동으로 생육데이터를 수집하고 있었지만 엽채류의 경우 재배기간이 30일로 짧아 따로 수집하지 않는 차이점

- 제어 환경의 경우 냉/난방시설을 통해 온도를 제어하는 과채류와 다르게 엽채류의 경우 난방시설만 갖추고 있으며 냉방의 경우 환기시설을 기반으로 35℃ 만 넘지 않도록 관리하며, 육묘 시기에만 별도의 재배 공간에서 온도 19℃ 이내 습도 90%의 냉장 상태에서 정식까지 관리

□ 기존에 공개되어 병충해 데이터 세트와 기 확보되어 있는 데이터를 이용하여 인공지능 생성모델 (GAN-Generative Adversarial Network) 개발

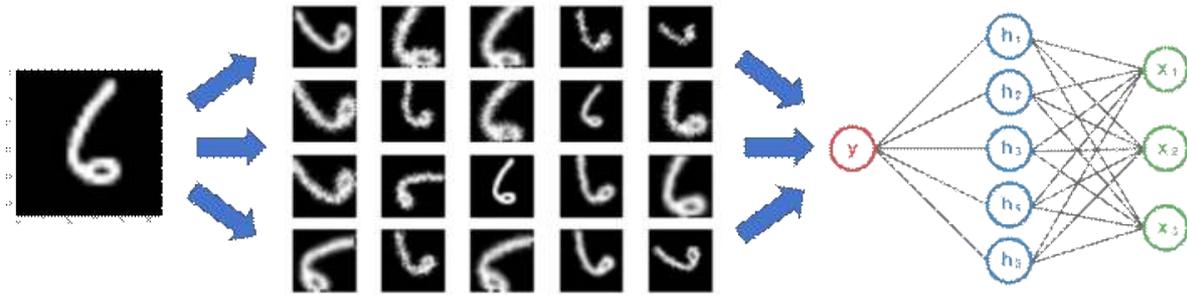
○ GAN(Generative Adversarial Network : 생성적 적대 신경망) 모델 개발

- GAN 기반 초기 이미지 생성 테스트 결과 이미지 생성 이미지와 기존 이미지간의 매칭비율이 떨어지는 문제 발생
  - 고해상도 이미지 생성 한계와 학습 불안정 문제 발생
- DCGAN 구조적 특징 분석
  - Max-Pooling Layer 를 없애고, Strided Convolution이나 fractional-strided convolution을 사용하여 feature-map 크기를 조절
  - Batch normalization을 적용
  - Fully connected hidden layer를 제거하기
  - Generator의 출력단의 활성화함수로 Tanh를 사용하고, 나머지 layer는 ReLU를 사용
  - Discriminator의 활성화함수로 LeakyReLU를 사용
- GAN 구조적 단점 해결을 위해 Convolution 구조를 GAN에 사용한 DCGAN(Deep Convolution GAN)을 기반하여 이미지 생성 필요



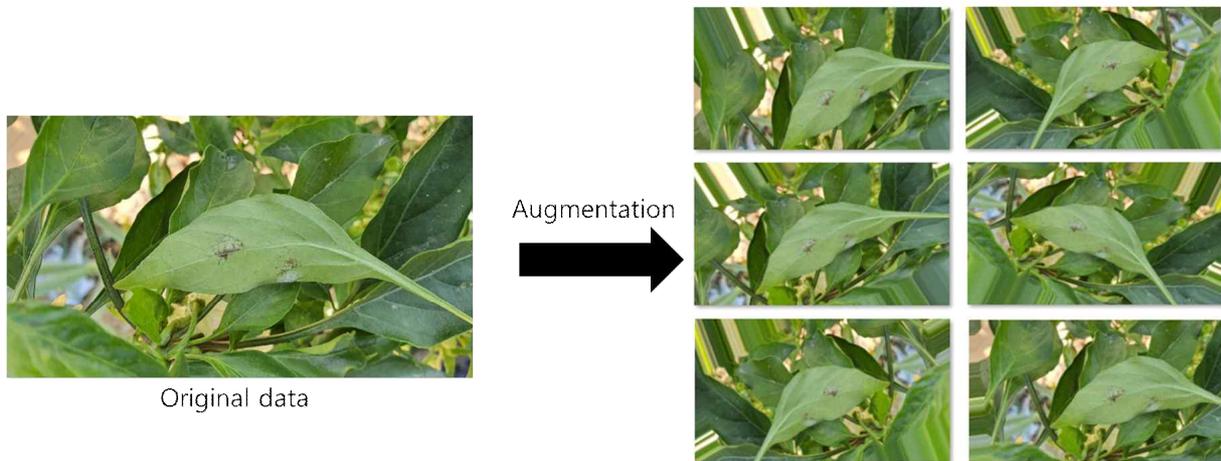
<DCGAN Generator Architecture>

- augmentation 매개 변수를 사용하여 1장의 이미지 데이터를 변환하여 여러장의 이미지로 증강 생성하여 학습용 데이터 확보



<Data Augmentation 개념>

- augmentation을 통해 훈련 데이터 셋의 이미지 양을 증가시켜 재학습 모델의 정확도 향상
- rotation\_range로 임의 회전의 각도 범위를 제어
- width\_shift\_range, height\_shift\_range로 수평 및 수직 이동
- shear\_range로 이미지를 시계 반대 방향의 각도를 제어
- 1-zoom\_range, 1 + zoom\_range로 이미지를 확대·축소 제어



<이미지 증강 예시>

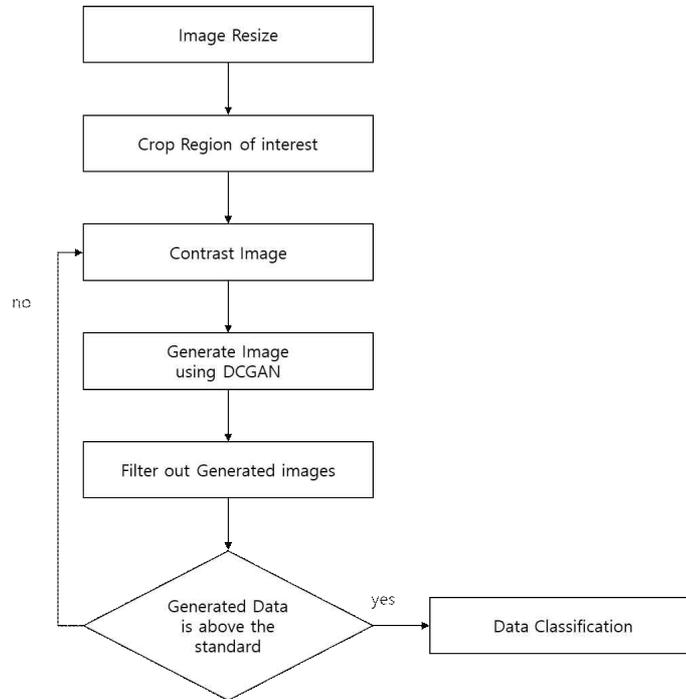
□ 생육 데이터(예, 과실)를 이용한 데이터 생성 모델 개발

○ 과채류 데이터 기반 생육 단계별 데이터 분류 체계 정의

- 과채류 정상 이미지 및 병해 이미지 데이터 분류
  - 과실, 잎, 줄기에 대한 공유 데이터 수집 및 재분류
  - 과채류 관련 병해별 공유 샘플 이미지 수집 및 분류
- 이미지 색상 분석을 통한 원천 이미지 종류에 따른 변위 도출
  - 1차년도 파프리카 과실 대상 색상 범위 도출

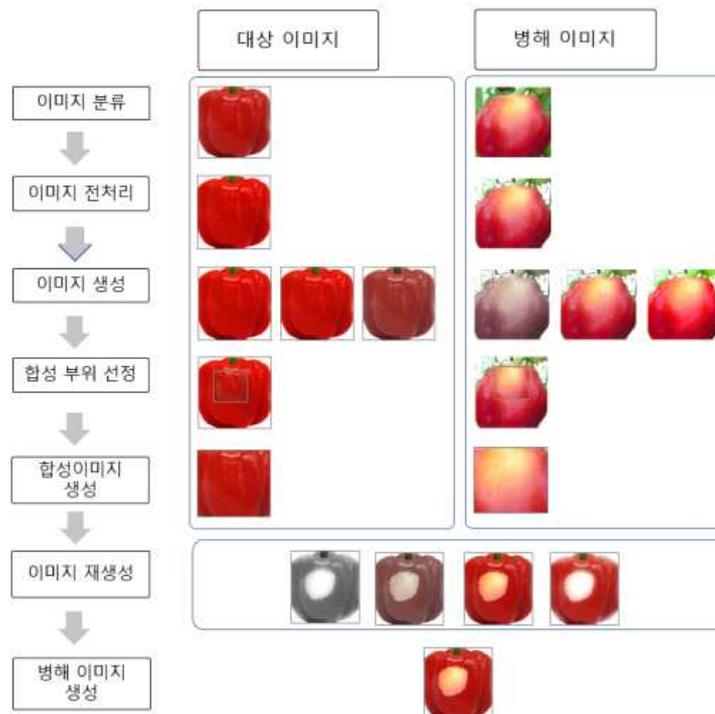
○ 이미지 생성 프로세스 정의

- Image Generation Flow using DCGAN



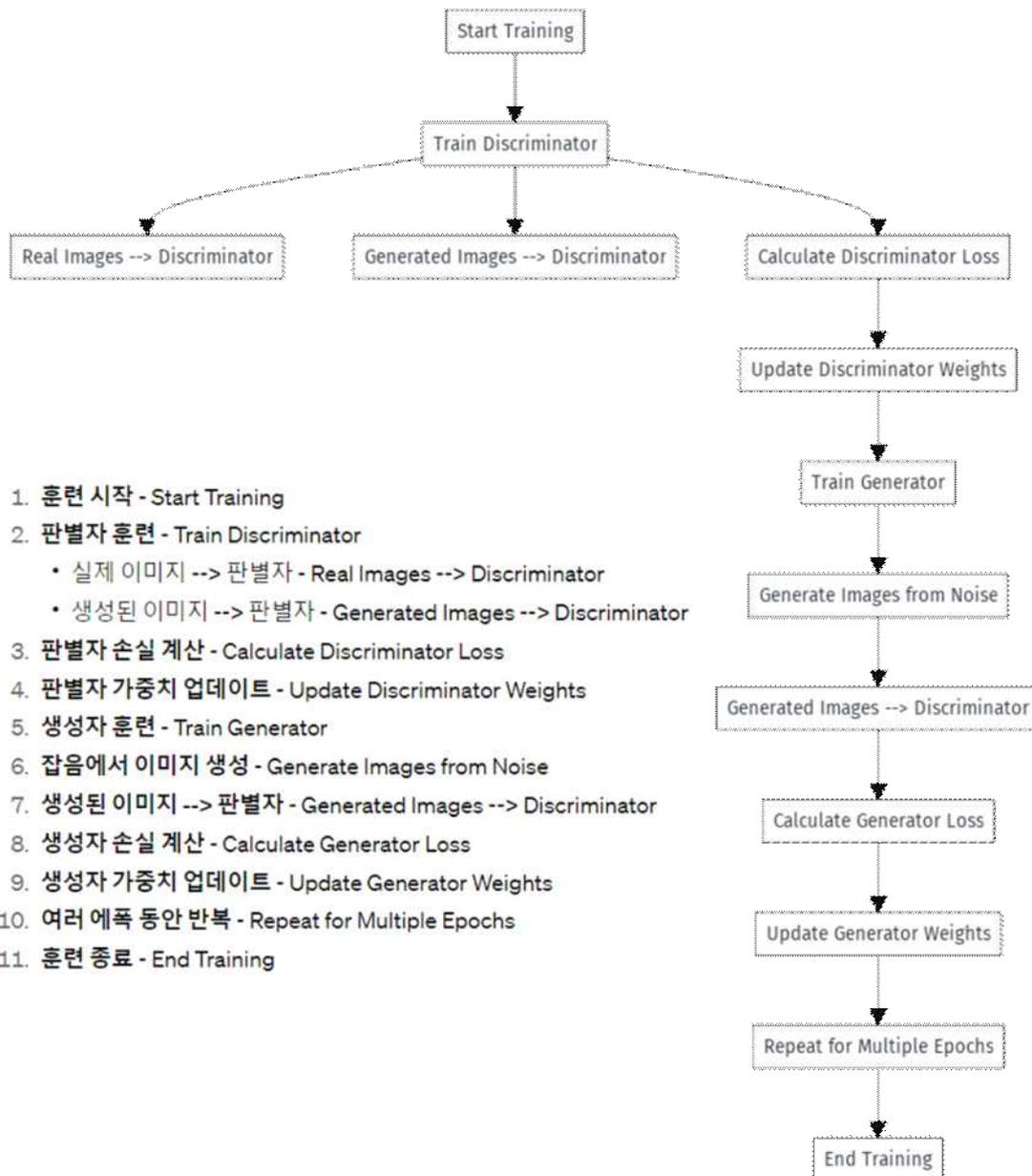
<Image Generation Flow using DCGAN>

○ 병해 이미지 생성 프로세스 정의



<병해 이미지 생성 프로세스>

- DCGAN을 이용해 대량 복제 이미지 생성 후 원본 이미지간의 비교 분석을 통해 일치도가 80% 이상인 이미지만 선별하여 재학습 시킴으로서 작물 부위별 병해의 기초 이미지 확보



<GAN 훈련과정 플로우차트>

- 생성자, 판별자 신경망 구성 및 연결모델 생성

```

# 생성자 신경망을 구성
def build_generator():
    model = tf.keras.Sequential()
    model.add(tf.keras.layers.Dense(128 * 8 * 8, input_dim=n_noise))
    model.add(tf.keras.layers.BatchNormalization())
    model.add(tf.keras.layers.LeakyReLU(alpha=0.2))
    model.add(tf.keras.layers.Reshape((8, 8, 128)))
    model.add(tf.keras.layers.Conv2DTranspose(64, (4, 4), strides=(2, 2), padding="same"))
    model.add(tf.keras.layers.BatchNormalization())
    model.add(tf.keras.layers.LeakyReLU(alpha=0.2))
  
```

```

model.add(tf.keras.layers.Conv2DTranspose(3, (4, 4), strides=(2, 2), padding="same",
activation="tanh"))
return model

# 판별자 신경망을 구성
def build_discriminator():
    model = tf.keras.Sequential()
    model.add(tf.keras.layers.Conv2D(64, (4, 4), strides=(2, 2), padding="same",
input_shape=(32, 32, 3)))
    model.add(tf.keras.layers.LeakyReLU(alpha=0.2))
    model.add(tf.keras.layers.Dropout(0.3))
    model.add(tf.keras.layers.Conv2D(128, (4, 4), strides=(2, 2), padding="same"))
    model.add(tf.keras.layers.LeakyReLU(alpha=0.2))
    model.add(tf.keras.layers.Dropout(0.3))
    model.add(tf.keras.layers.Flatten())
    model.add(tf.keras.layers.Dense(1, activation="sigmoid"))
    return model

# 생성자와 판별자를 생성
generator = build_generator()
discriminator = build_discriminator()

# 생성자의 출력을 판별자의 입력으로 연결하는 모델을 생성
gan_input = tf.keras.layers.Input(shape=(n_noise,))
x = generator(gan_input)
gan_output = discriminator(x)
gan = tf.keras.models.Model(gan_input, gan_output)

# 손실 함수와 최적화 알고리즘을 설정
discriminator.compile(loss="binary_crossentropy",
optimizer=tf.keras.optimizers.Adam(lr=learning_rate))
discriminator.trainable = False
gan.compile(loss="binary_crossentropy",
optimizer=tf.keras.optimizers.Adam(lr=learning_rate))

```

- 잡음 벡터를 입력받아 생성자로부터 이미지를 생성 -> 판별자로부터 결과 피드백 -> 잡음 벡터와 판결 결과 연결 모델 생성
- 손실함수와 최적화 알고리즘을 적용하여 생성 이미지 선별
- 판별자 및 생성자-판별자 연결모델 : 이진 교차 엔트로피 손실 함수와 Adam 최적화 알고리즘 적용

□ 생성된 데이터를 이용한 추론 모델 재학습 기술 개발

○ GAN 데이터 기반 추론데이터를 활용한 재학습 데이터 분류를 통해 대상 작물에 해당하는 특정요소에 대한 재학습 수행 기술 개발

- 단계별 작물, 병해 데이터 이미지 생성 및 결합 모델 설계
- 이미지 생성 모델 학습을 위한 작물 이미지 수집
  - 공공기관에서 수집/배포하고 있는 작물 이미지 데이터 셋을 활용



<정상 이미지>

<병해 이미지>

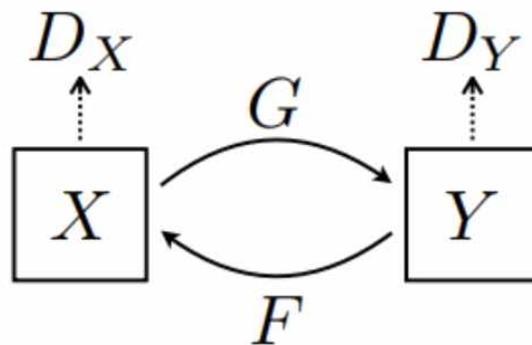
- 이미지 데이터 정보
  - 사이즈 : 3024 x 4032
  - 정상 이미지 : 13,000장
  - 병해 이미지 : 1,967장
- 어노테이션
  - 이미지 파일명
  - 사이즈
  - B\_BOX
  - 병해 종류
- 이미지 생성 모델 학습용 데이터셋 생성
  - 이미지의 RGB 값을 모두 0~1 사이 값으로 치환
  - 보다 빠른 모델 학습을 위해 원본이미지 resize 진행 3024x4032 => 512x512
  - 정상 이미지 30% 활용
  - 작업 환경 구축

<병해 데이터 생성을 위한 작업환경>

개발 환경	
운영체제	Window10 pro 64비트
프로세서	Intel® Core™ i7-9700k
메모리	64GB
VGA	NVIDIA GeForce GTX 4090
Tools	python 3.9, cuda 12.3

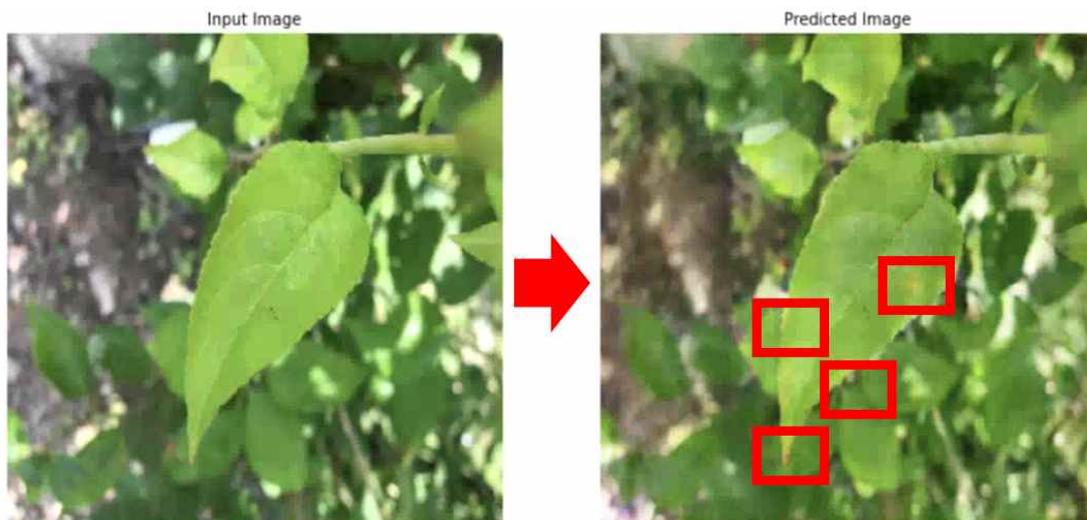
- 이미지 CycleGAN 모델 학습

- CycleGAN은 image to image translation() 기법을 이용한 모델로서, input 이미지와 output 이미지를 mapping하는 것을 목표로 하는 생성 모델의 한 분야
- CycleGAN은 배치 정규화 대신 인스턴스 정규화를 사용하며, 2개의 생성기(G 및 F)와 2개의 판별자(D<sub>X</sub> 및 D<sub>Y</sub>)를 훈련
  - 생성기 G는 이미지 X를 이미지 Y로 변환하는 방법을 학습
  - 생성기 F는 이미지 Y를 이미지 X로 변환하는 방법을 학습
  - 판별자 D<sub>X</sub>는 이미지 X와 생성된 이미지 X(F(Y))를 구별하는 방법을 학습
  - 판별자 D<sub>Y</sub>는 이미지 Y와 생성된 이미지 Y(G(X))를 구별하는 방법을 학습



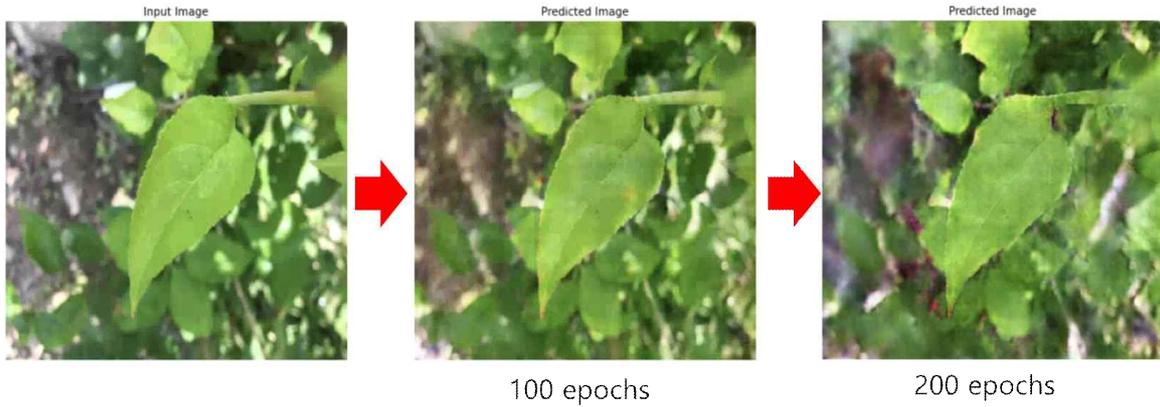
<GAN의 기본 구조>

- unpaired image 데이터를 가지고 image to image translation를 하는 것이 cycle GAN의 목적
- 정상 이미지 픽셀을 병해 픽셀로 변환하여 새로운 병해 이미지 데이터 셋을 생성
- 이미지 데이터 셋을 넣어 학습 진행
- 100 epochs 학습에서부터 유의미한 이미지 생성



<INPUT 이미지와 OUTPUT 이미지 비교>

- 100 epochs 단위로 추가 학습 진행하여 생성된 병해 이미지 데이터 확인
  - 변화과정을 보기 위하여 동일한 test 이미지 데이터 사용
  - 각 epochs 별 체크포인트 저장을 통한 전이학습용 파라미터 관리



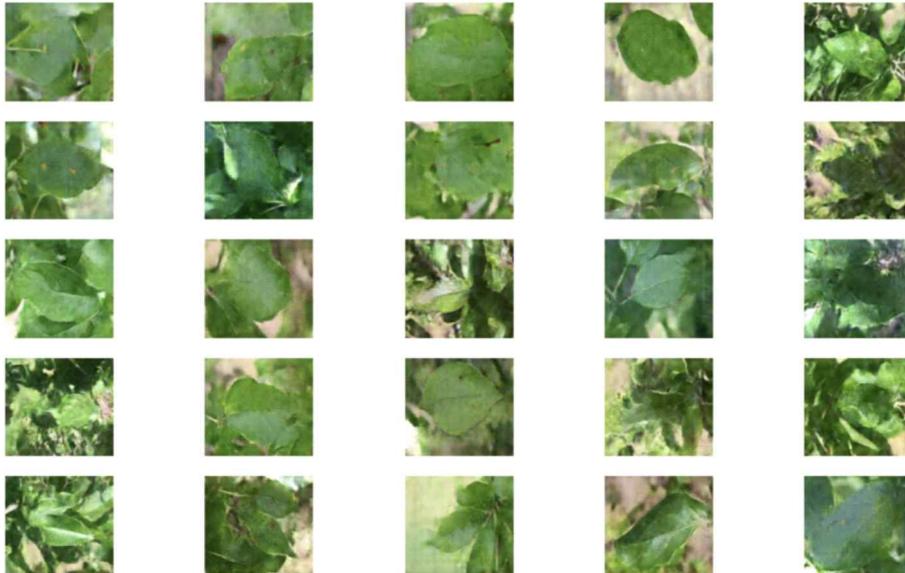
<epoch 변화에 따른 질병 생성 과정>

- 생성된 모델별 Frechet Inception distance(FID)를 통한 생성 이미지 적합도 검증
- Frechet Inception Distance(FID)는 생성된 이미지의 분포와 원래 이미지의 분포가 어느정도 비슷한지 측정하는 지표

$$\|m - m_w\|_2^2 + \text{Tr}(C + C_w - 2(CC_w)^{1/2})$$

<FID 계산 공식>

- 생성 모델을 통한 병해 이미지 데이터 생성



<생성된 병해 이미지>

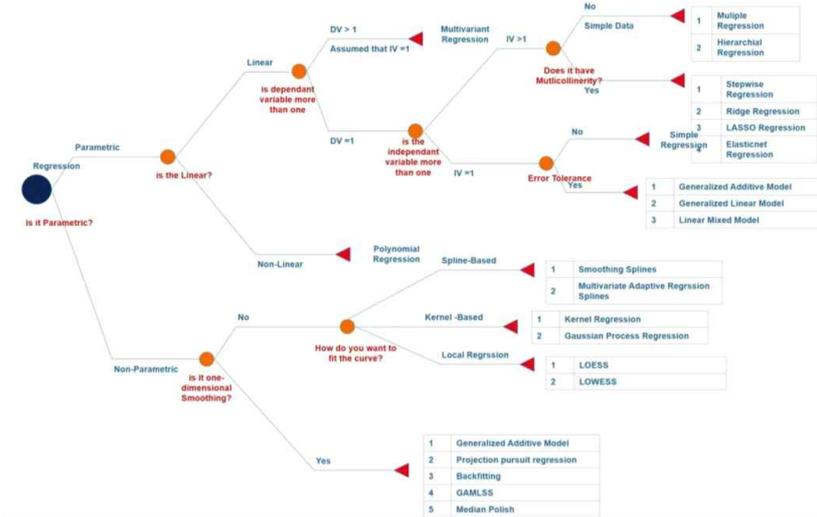
- 작물/병해 병합 기초 이미지 생성 테스트 수행
- 기 생성된 학습 모델의 재사용 및 학습용 데이터 확보를 위한 전이학습 모델 개발
  - epochs별 학습 파라미터값 체크포인트 생성
  - 추가 작물 모델 개발 생성시 기존 파라미터값을 넣어 전이학습 함으로써 학습 시간 절감
  - 체크포인트 관리 프로세스 도출

[1단계(2022)]

□ 자동 학습 시스템을 통해 데이터에 맞는 최적 학습 방법 선택 기술 개발

○ 표준 데이터셋 명세를 기반한 스마트 농업 자동 머신러닝 모델 기술 개발

- 온실에서 수집되어지는 데이터 명세들에 따라 사용되어지는 머신러닝 모델들도 다양하게 사용
- 머신러닝 모델들에 사용되는 매개변수에 따라 모수/비모수 모델들로 구분되며, 함수형태의 모수 모델들은 매개변수가 고정되어 상관관계가 분명한 데이터들의 분석에 용이
- 비모수 모델의 경우 함수의 형태를 가정하지 않고 진행하여 데이터들 간의 특성을 모를 때 사용



- 위 그림의 모듈에서 생성된 데이터를 이전 데이터와 통합하여 트리를 재평가
- 전체 데이터 세트에 대해 엔트로피가 다시 계산되고 재할당하여 엔트로피 계산과 함께 Information Gain도 다시 계산되어 데이터 셋에 추가
- 엔트로피와 정보 이득 값으로 데이터에 맞는 머신러닝 알고리즘의 순위 생성

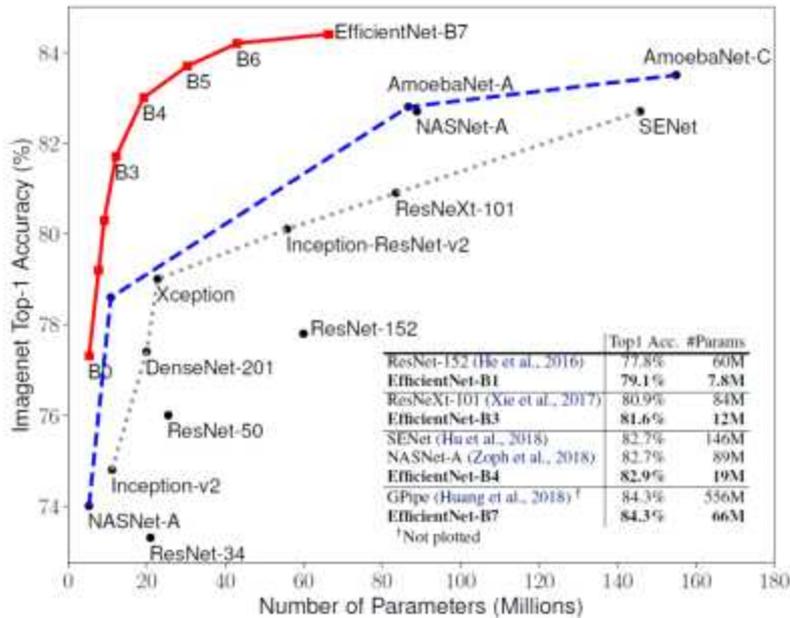
□ 학습된 인공지능 모델간의 중요도(가중치) 선택 인공지능 네트워크 개발

- 원시 입력 데이터 분류 모델 및 의미 분석 모델 설계
  - 가중치 분류(Classification)를 위하여 원시 데이터로부터 데이터셋 추출
  - 입력데이터 분류 모델 설계를 위해 원시 데이터로부터 추출한 공통 데이터셋을 사용한 상관관계 분석
  - 상관관계 분석을 기반으로 한 데이터 가중치 선택 모델 개발



<상관관계 분석 시각화>

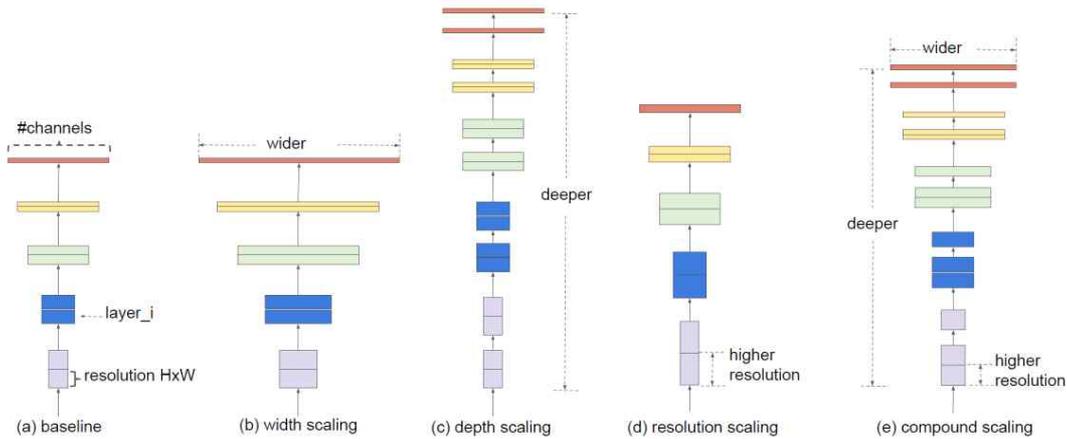
- 가중치 선택 인공지능 네트워크 개발을 위한 다계층 뉴럴 네트워크 모델 개발
  - 다계층 뉴럴 네트워크의 구조를 설계하고, 각 계층의 뉴런 수와 연결 가중치를 결정하는 방법 고안
  - 가중치 선택 알고리즘 개발을 통해 네트워크의 가중치를 자동으로 선택하기 위한 최적의 가중치 조합 탐색
    - ImageNet 데이터셋을 활용한 가중치 선택 알고리즘을 적용하여 최적의 가중치 조합 탐색



<ImageNet 모델별 정확도>

- 작물 이미지 분류 모델 선정을 위한 탐색 매개변수 도출
  - 이미지 데이터 분류를 위한 최적 학습법 탐색
    - YOLO를 활용한 이미지 분류 모델 개발(학술발표 2건)
    - YOLO를 이용한 유해조수 침입 감지 모델
    - Harmful Tide Intrusion Detection Model Using YOLO
  - CNN을 기반으로 한 Pretrained Model간 loss, auc 비교

- ResNet50 : 마이크로소프트에서 개발한 알고리즘으로 VggNet을 베이스로 했으며, 지름길 (Shortcut)을 연결해서 잔차(Residual)를 최소가 되게 학습한 것. ResNet-18 부터 ResNet-200 까지 망의 깊이(depth)를 늘려 성능 향상
- Xception : eXtreme Inception의 약자로 기존 Inception 모델이 채널, 공간, correlation을 분리한 것을 depthwise separable convolution으로 강화한 모델, Inception V3과 비슷한 수의 파라미터를 가지면서 image classification에서 더 좋은 성능을 이끌어 냄
- EfficientNet : ConvNet 구조의 복잡도에 영향을 미치는 요소들을 분석하여 이를 최적화하는 방식으로 성능 개선을 접근한 모델



<모델 스케일링 구조 비교>

- 최적의 다계층 뉴럴 네트워크 모델을 위한 Pretrained Model 정확도 비교
- 다음과 같이 설정값 지정
  - IMAGE\_SIZE = 320, 512
  - BATCH\_SIZE = 8
  - Learning Rate Scheduler = ReduceLROnPlateau
  - Learning Rate = 0.0001
  - epochs = 10
  - metrics = ROC-AUC

<모델별 정확도 결과 비교>

구분	ResNet50	Xception	EfficientNetb3	EfficientNetb7
time	<b>87s</b>	77s	105s	<b>268s</b>
loss	<b>0.1521</b>	0.6454	0.0909	<b>0.0623</b>
AUC	<b>0.9936</b>	0.9302	0.9975	<b>0.9988</b>

- 인공지능 모델간의 중요도(가중치) 선택을 위한 변수값으로 학습 및 분석 속도 변수로 time 값을 정확도 관련 변수로 loss, AUC 값을 선정하여 모델 선택
  - loss : 손실함수(loss function)는 머신러닝에서 모델이 나타내는 확률 분포와 데이터가 따르는 실제 확률 분포 사이의 차이를 나타내는 함수로 값이 0에 가까울수록 모델의 정확도가 높고, 반대로 0에서 멀어질수록 모델의 정확도가 낮음
  - auc : 정확도(accuracy)는 실제 데이터에서 예측 데이터가 얼마나 같은지를 판단하는 지표
  - 딥러닝파라미터(batch size, epochs)별 ResNet50, Xception등 기존 딥러닝모델 학습 및 time, loss, AUC 값 도출
  - 사용자 의도에 따른 결과 도출 값 선택을 통한 모델 선정
  - 정확도 : loss > AUC > time 값을 우선순위로 두어 모델 선정

- 학습 및 분석 속도 : time > AUC > loss 값을 우선순위로 두어 모델 선정

## [2단계(2023)]

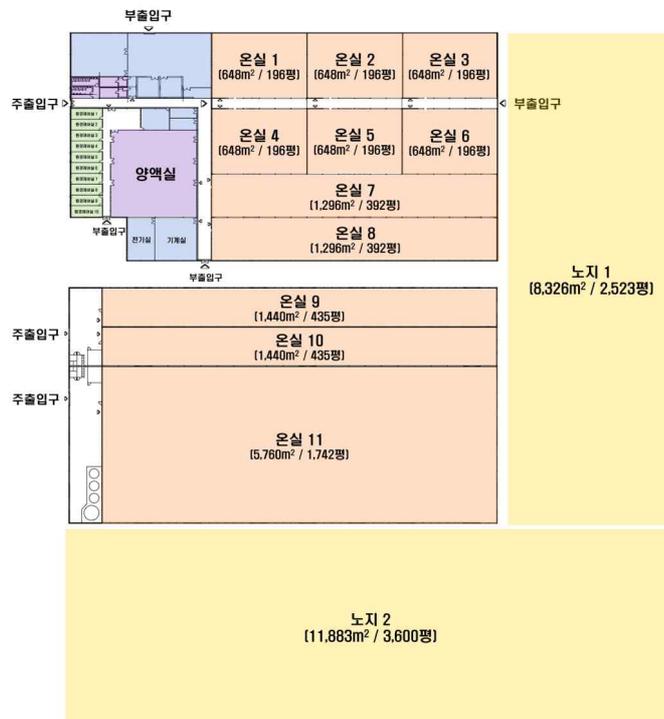
### □ 과제류 및 업체류 각 3종 온실 작물에 대한 데이터 생성 모델 실증

#### ○ 전남 스마트팜 혁신밸리 TEST-BED 대상 온실 데이터 생성 모델 실증

- 스마트팜 관련 기술을 육성을 위해 전남 고흥에서 ‘스마트팜 혁신밸리 실증단지’의 운영하여 전남 고흥 스마트팜 혁신밸리 실증단지 입주 기업·기관을 모집

#### - 모집 내용

- 모집 공고 : 2023.05.16.
- 모집 기간 : 2023.05.16.~2023.05.25.
- 입주 시설 : 유리온실 8구역, 반밀폐형온실 3구역, 노지 2구역



<고흥 스마트팜 혁신밸리 실증단지 구역 배치도>

- 총 13개 구역 중 유리온실 8번(1,296㎡(392)) 신청
- 유리온실 8번에 최종 선정되었으며 전남테크노파크와의 계약 체결을 통해 빅데이터 수집 사이트 추가 완료
  - 면적(평) : (유리온실, 1,296㎡(392))
  - 계약일 : 2023.11.20.
  - 이용 기간 : 2023.11.~ 2026.11.(36개월)

[별지 제8호] 이용계약서

**전남 고흥 스마트팜 혁신밸리 실증단지 이용계약서**

전남 고흥 스마트팜 혁신밸리 실증단지 이용계약을 위해 실증단지 운영기관인 전남테크노파크(이하 "운영기관")와 실증단지이용자 **순천대학교 산학협력단(Grand ICT 연구센터)**(이하 "이용자")는 다음과 같이 계약을 체결한다.

**임대인**  
운영기관 : 전남테크노파크 원장 오익현

**임차인**  
주관기관 : **순천대학교 산학협력단** 단장 여현  
(Grand ICT 연구센터) (센터장)

제1조(목적) 본 계약은 운영기관과 스마트팜 혁신밸리 실증단지(이하 "실증단지") 이용자 간 실증단지 이용에 관한 제반사항 규정을 목적으로 한다.

제2조(원칙) ① 운영기관은 양질의 실증결과를 위해 노력하고 이용자는 실증단지의 원활한 운영을 위해 적극 협조한다.  
② 실증단지의 이용에 관한 사항은 본 계약에 따른다. 또한, 본 계약에서 규정하지 않은 사항은 「전남 고흥 스마트팜 실증단지 운영규칙」(이하 "운영규칙") 및 「전남 고흥 스마트팜 실증단지 운영지침」(이하 "운영지침")에 따르며 본 계약의 일부로 간주한다.  
③ 실증 진행 및 수행과 관련한 모든 사항은 실증계획서에 따라 진행한다. 단, 변동사항이 있는 경우 "운영규칙", "운영지침"의 범위내에서 상호 문서를 통한 합의 하에 실증계획서를 수정할 수 있다.

제3조(이용 및 계약기간) ① 본 계약에 의한 이용 및 계약기간은 아래 표와 같다.

이용 기간	2023. 11. 20. - 2026. 11. 19. (36개월)
시설 번호, 이름	온실 8, 환경제어실
이용료(원/월)	876,900원
부가세(원/월)	87,690원
합계(원/월)	964,590원

\* 관리비는 실비 청구함.  
② 운영기관은 이용기간 만료 3개월 전에 이용자에게 이용기간 만료 사실을 서면통지하여야 한다.

위 법령 및 운영규칙·지침을 준용한다.  
② 제1항에도 불구하고 해석상 이견이 생겼을 경우 상위기관의 해석에 따른다.  
제21조(기타사항) 본 계약서에 기재되지 아니한 사항은 운영기관의 제반 규정 및 관리지침과 통상의 관련 계약에 준한다.  
제22조(계약의 효력) ① 본 계약서는 2부를 작성하여 운영기관과 이용자가 각각 서명 날인하고, 각 1부씩 보관한다.  
② 운영기관의 장이 변경되어도 본 계약의 효력은 유지된다.

2023년 11월 20일  
**임대인**  
운영기관 : 전남테크노파크 원장 오익현 (인)  
**임차인**  
주관기관 : **순천대학교 산학협력단** 단장 여현 (인)  
(Grand ICT 연구센터) (센터장)

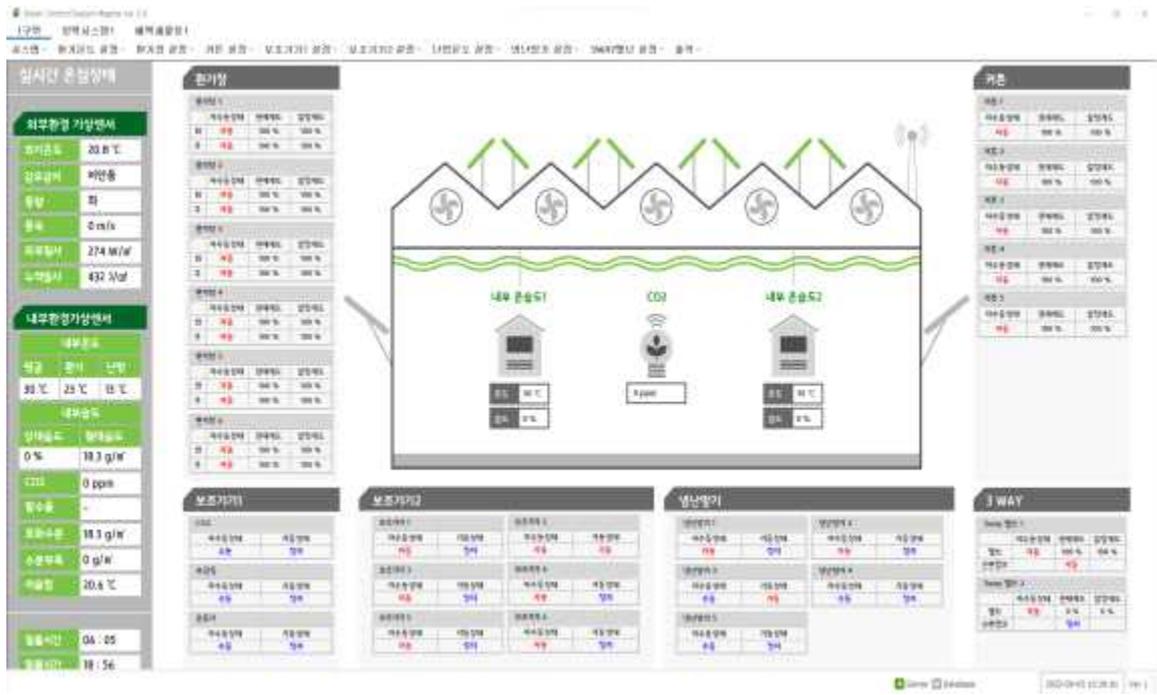
첨부 : 1. 인감증명서(개인사업자인 경우 대표자 개인인감), 1부.  
2. 사업자등록증, 사본 1부.  
3. 실증계획서 1부.  
\* 기존 실증계획서 양식 참조, 참여기관별 수행 업무 작성

**<고흥 스마트팜 혁신밸리 실증단지 이용계약서>**



**<고흥 스마트팜 혁신밸리 실증단지 8번 유리온실>**

- 데이터 수집을 위한 통합 제어기, 데이터 수집용 센서, 관수펌프 등 제어장비, 통합제어 PC 및 통합제어용 프로그램(MAGMA 3.0) 설치 완료
- MAGAM 3.0을 통해 환경 및 제어 데이터가 수집되고 있으며 통합제어용 프로그램을 통해 수집되고 있는 데이터를 기간 및 데이터 간격을 설정하여 CSV파일로 제공받고 있음



<고흥 스마트팜 혁신밸리 실증단지 통합 제어 시스템(MAGMA 3.0) 제어 화면>

- 환경 제어기, 양액기, 유동팬, FCU, 튜브레일, 행잉커터, 수평스크린, 천창, 스크린 개폐장치, CO2 공급장치, 살균기 등 장비 구축 완료



유동팬, 튜브레일, 행잉커터, 천창 등



양액기

<고흥 스마트팜 혁신밸리 실증단지 구축 장비>



<통합 제어기>



<양액 제어기>

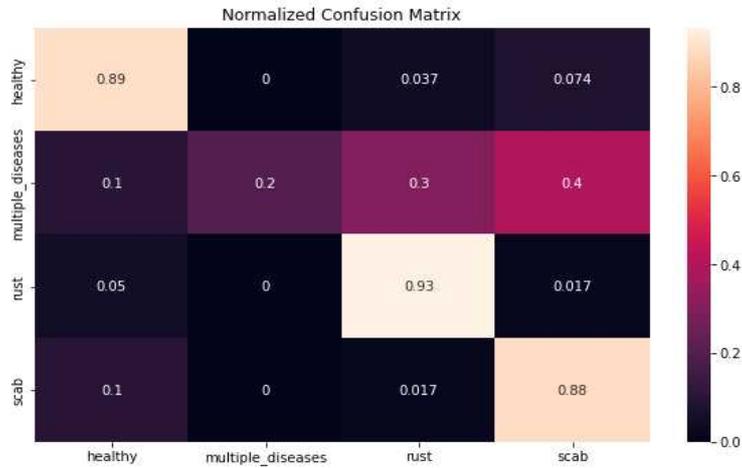
<고흥 스마트팜 혁신밸리 실증단지 장비 현황>

□ 표준화 기반 작물 병충해 데이터셋 인공지능 생성모델 실증 및 고도화

○ 작물별 실증 사이트 활용 e-BRAIN 및 로컬 사이트 수집 데이터 활용 작물 병해 데이터셋 생성 모델 실증 수행

• Grad-CAM 기능을 통한 모델 성능 향상 방안 연구

- Grad-CAM(Gradient-weighted Class Activation Mapping)은 딥러닝 모델의 학습된 가중치를 사용하여 입력 이미지에서 모델의 예측에 어떤 부분이 주요하게 기여했는지 시각화하는 기법
- CNN(Convolutional Neural Network)과 같은 컨볼루션 기반의 신경망에서 활성화된 특징 맵을 이용하여 클래스에 대한 기여를 시각적으로 확인 가능
- 이미지 데이터는 총 1,821장으로 healthy(건강) 516장, rust(녹병, 곰팡이병) 622장, scab(붉은곰팡이병) 592장, multiple\_diseases(다발성 질병) 91장의 4가지로 구성
- epoch를 10으로 설정하여 모델 성능을 확인해본 결과 Loss=0.43177, Accuracy=86.34%의 성능을 보임

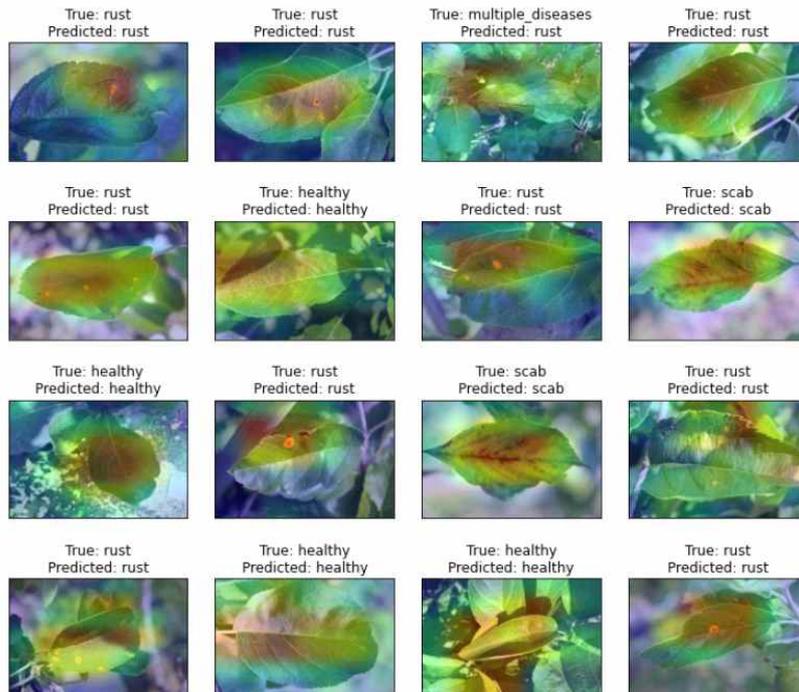


<Grad-CAM을 통해 분석한 병해데이터 성능 히트맵>

- rust, healthy, scab의 순서로 높은 분류 정확도를 보였으며, multiple\_diseases의 경우 약 0.2의 낮은 성능을 보임
- accuracy 뿐만 아니라 다양한 방법으로 모델의 성능을 평가해본 결과 rust의 recall=0.98, f1-score=0.92, support=60으로 가장 높은 성능을 보였고, multiple\_diseases는 recall=0.20, f1-score=0.33, support=10의 매우 낮은 성능을 보임

<병해별 구분 정확도 비교>

구분	recall	f1-score	support
healthy	0.89	0.86	54
<b>multiple_diseases</b>	<b>0.20</b>	<b>0.33</b>	<b>10</b>
<b>rust</b>	<b>0.93</b>	<b>0.92</b>	<b>60</b>
scab	0.83	0.87	59



<Grad-CAM을 적용한 병해 판별 원인 분석>

- 공개 데이터를 활용한 병해 이미지 생성 모델을 개발하여 a-BRAIN 플랫폼에 탑재
  - 1,925장의 작물 이미지를 확보하여 epoch 변화에 따라 생성된 병해 이미지의 FID를 비교하여 FID값이 가장 낮은 최적 Epoch 탐색
    - Discriminators와 Generators의 손실을 계산하고, 그래디언트를 역전파하여 모델을 최적화
    - 학습 변화도를 확인하기 위해 동일한 test 이미지를 가지고 1epochs 마다 변화를 확인할 수 있도록 코드 작성



<각 EPOCH별 생성된 병해 이미지>

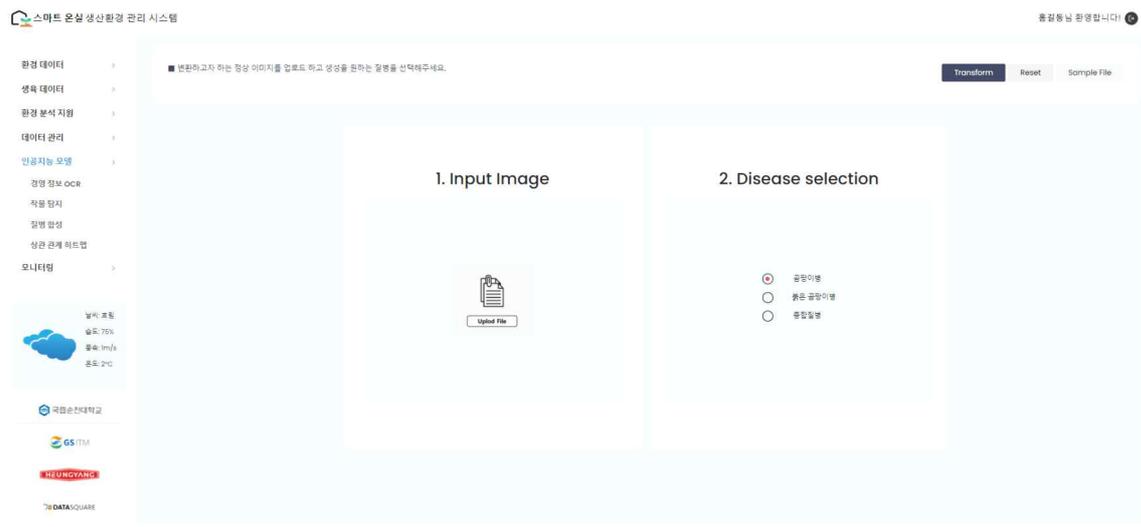
- 300epochs 이후 모델 평균 FID score의 변화가 미비하여 400epochs로 훈련 종료

<EPOCH별 IS와 FID 비교>

구분	1st IS		2nd IS		FID
	Mean	Standard Deviation	Mean	Standard Deviation	
100epochs	2.2637	0.1113	1.7251	0.1668	213.3931
200epochs	2.2613	0.1421	1.6587	0.1759	<b>204.7196</b>
300epochs	2.2667	0.0946	1.7092	0.1698	211.4339
400epochs	2.2619	0.1414	1.6287	0.1453	212.1680

- 정상-실제 병해 이미지의 평균 FID는 약 223, 실제 병해 이미지(train, test)의 평균 FID는 약 207
- 해당 포인트를 기준으로 epochs별 FID를 비교했을 때 200epochs의 생성 병해 이미지에서 약 204의 낮은 FID를 보임

- 200epochs 데이터셋은 IS가 높지만 FID가 낮은 결과값을 보였으며, 이 결과는 생성된 이미지가 품질은 높지만 분포의 차이가 존재한다는 것을 나타냄
- 200epochs 생성 데이터셋을 기준으로 생성된 병해 이미지의 FID가 실제 병해 병해이미지 FID 보다 높은 수준의 FID를 나타냄
- 실증단지에서 수집한 과채류 작물 잎 이미지를 활용한 모델 실증
- 잎의 형태가 확실한 작물을 대상으로 모델 실증을 위한 이미지 데이터 수집
- 공개 데이터를 활용한 병해 이미지 생성 모델을 개발하여 a-BRAIN 플랫폼에 기능 탑재



<a-BRAIN 플랫폼 병해 데이터셋 생성 모델 적용 화면>

- 곰팡이병(rust), 붉은 곰팡이병(scab), 종합질병(multy disease) 총 3가지 질병을 선정하여 정상 작물 이미지를 선택한 질병이미지로 변경
- 충분한 학습데이터 확보가 가능한 질병의 경우 모델 추가 구현이 가능함

○ 인공지능 모델 고도화를 위한 실데이터와 가상 데이터를 활용한 추론모델 재학습 수행

- CycleGAN으로 생성된 병해 작물 이미지 추가
- 곰팡이병(rust), 붉은 곰팡이병(scab), 종합질병(multy disease) 총 3가지 질병 데이터를 CycleGAN을 사용하여 가상 이미지 데이터 생성



<rust>



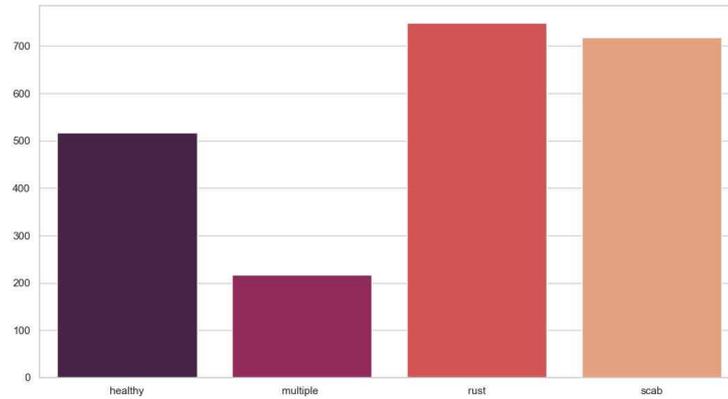
<scab>



<multiple disease>

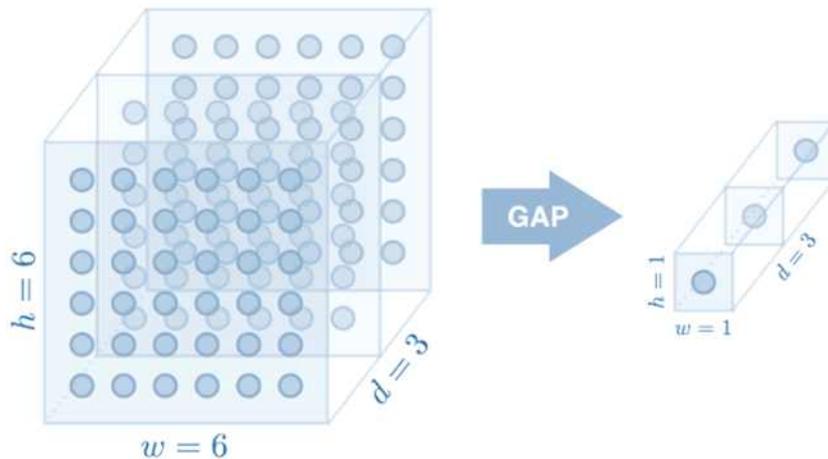
<CycleGAN을 통해 생성된 병해 작물 이미지>

- EDA(Exploratory Data Analysis)를 통한 작물 데이터 분석
- Validation\_split=0.1, training\_split=0.9로 설정



### <병해 데이터 구성 시각화>

- 질병별 데이터 불균형 해결을 위한 가중치 설정
  - $weight\_for\_0 = (1 / \text{각 클래스의 샘플 수가 저장된 리스트[리스트 번호]}) * (\text{total} / \text{클래스 수})$  공식을 사용하여 각 클래스의 가중치 계산
  - class\_weight 기능을 사용하여 상대적으로 샘플 수가 적은 클래스(multiple disease)에 높은 가중치를 적용하여 데이터 불균형 해결
- Inception v3 모델을 적용한 Disease Detection 모델 개발
  - Inception V3 모델은 GoogLeNet 아키텍처에서 파생된 모델로 복잡한 구조와 다양한 크기의 convolutional 필터를 사용하여 효율적으로 다양한 이미지 특징 학습 가능
  - Fully Connected Layer 대신 Global Average Pooling을 사용하여 출력을 생성



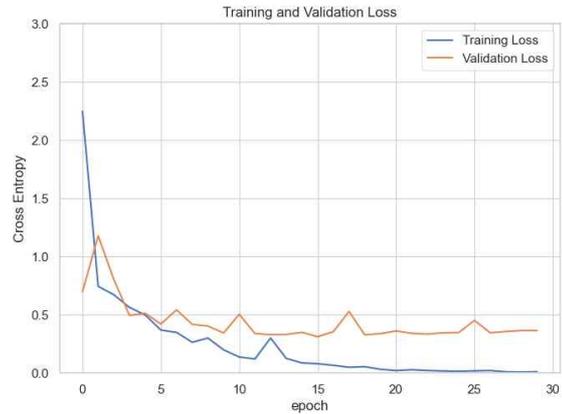
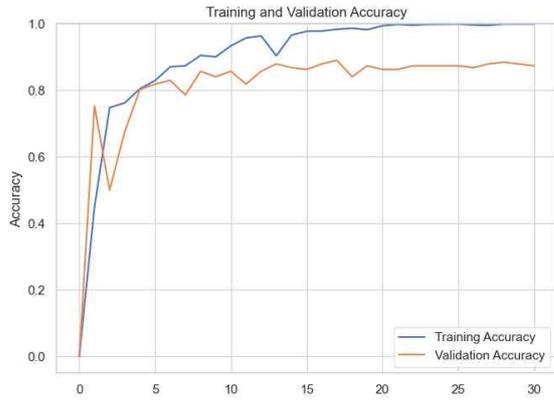
### <Global Average Pooling>

- 생성된 가상 병해이미지를 활용한 추론모델 재학습 수행
  - 기존 데이터를 기반으로 최대 epoch를 20으로 설정
  - 기본 epoch를 30으로 설정하여 최적의 성능을 보이는 epoch를 탐색
  - 최적 epoch에서의 loss, accuracy를 기반으로 생성된 가상 병해이미지를 활용한 추론모델의 성능 비교

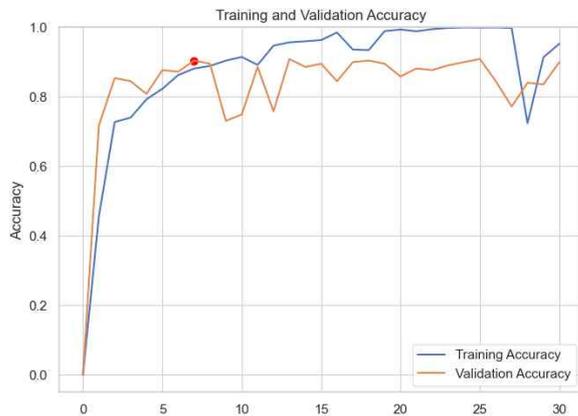
### <생성된 병해 데이터를 활용한 추론모델 재학습 결과>

구분	최적epoch	loss	accuracy	val_loss	val_accuracy
기존 모델	16	0.3119	0.8791	0.3646	0.8736
생성 데이터 추가 모델	7	<b>0.3385</b>	<b>0.9041</b>	<b>0.3403</b>	<b>0.8995</b>

- 약 350장의 생성 병해 이미지를 추가하여 병해 진단 모델을 구현한 결과 loss=0.0266, accuracy=0.0250, val\_loss=0.0243, val\_accuracy=0.0259로 성능 향상



<기존 데이터의 분석 모델 성능 시각화>



<생성 이미지 추가 데이터의 분석 모델 성능 시각화>

## [순천대학교\_공동]

### [1단계(2021)]

□ 기본 전처리 기술(PCA 등)과 Auto Encoder 등의 데이터 가공 기술을 이용하여 데이터 주요 정보 추출 기술 개발

○ 데이터 스케일링 기법을 활용한 데이터 정제

- Standard, MinMax, Roubust, Normalizer기법을 이용한 데이터 정제

- standardScaler : 데이터 특성들의 평균을 0, 분산을 1로 스케일링

-> 데이터 이상치에 민감하며, 분류 모델에 적용할 특성에 적용

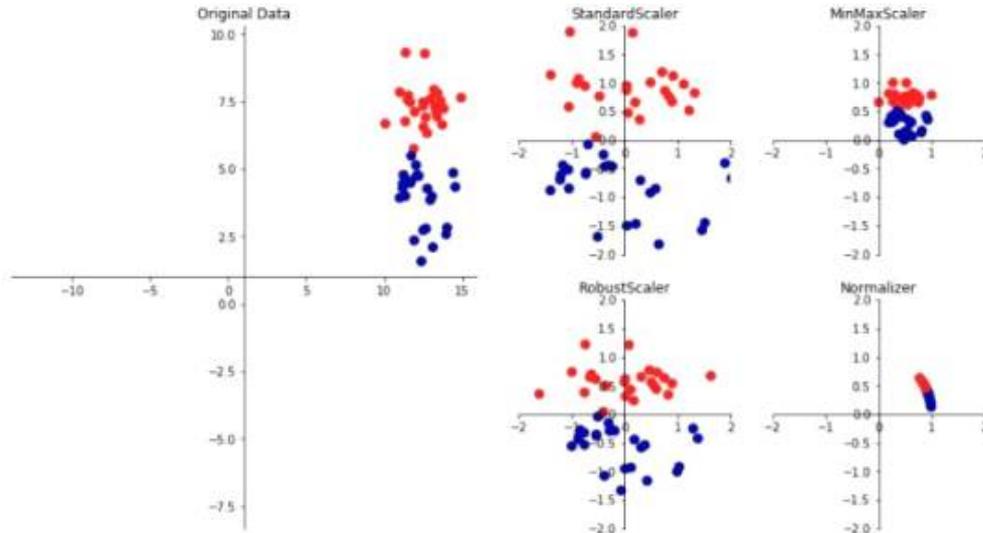
- MinMaxScaler : 데이터 특성들을 특정 범주로 스케일링([0,1] - 최솟값 : 0, 최댓값 : 1)

-> 데이터 이상치에 민감하며, 회귀분석을 진행할 특성에 적용

- RoubustScaler : 중간 값과 사분위 값을 사용하여 스케일링

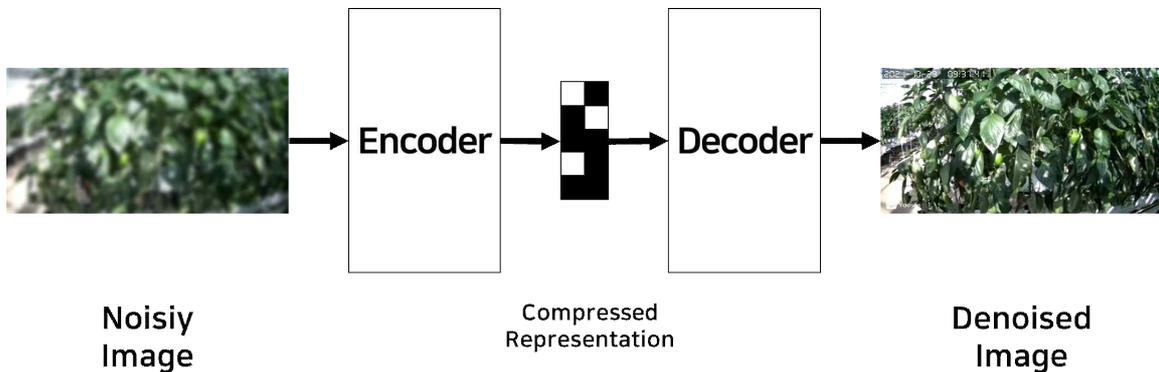
-> 이상치의 영향을 최소화할 수 있는 스케일링 방법

- Normalizer : 한 행의 모든 특성들 사이의 거리가 1이 되도록 스케일링  
-> 모델 내 학습 벡터에 적용할 때 주로 사용(동일 단위의 데이터에 적용)
- 각 스케일링 기법에 대해 모듈화 진행
- 실증 농가에서 수집되는 데이터의 활용 용도에 맞게 데이터 스케일링 기법을 적용하여 데이터 처리 진행 예정  
-> 작물별, 주기별, 활용 용도별, 환경데이터별 스케일링 기법 선정 예정



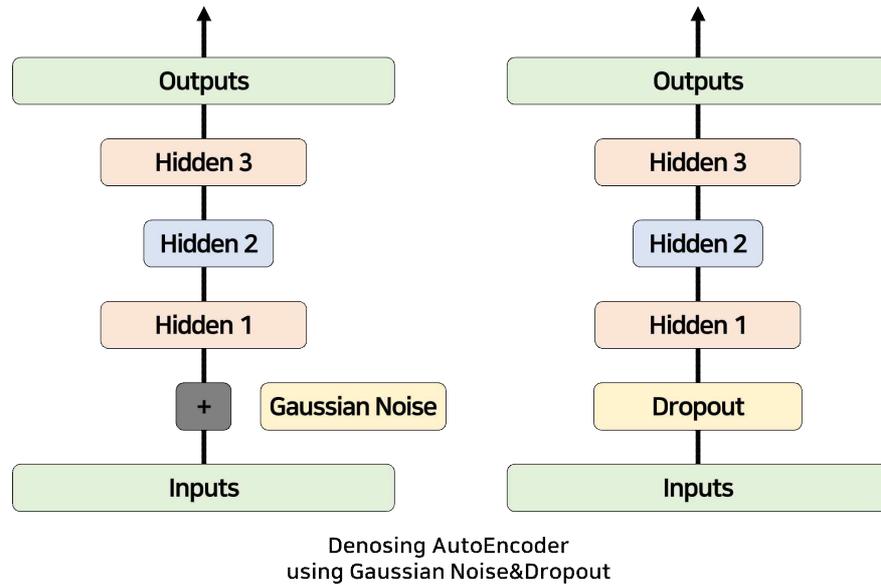
<데이터 스케일링 종류>

- 기본 전처리 기술(PCA(Principal Component Analysis), ICA(Independent Component Analysis), LDA(Latent Dirichlet Allocation))을 활용한 주요 정보 추출
- Denoising Auto Encoder를 통해 수집 영상에서의 잡음 제거
  - training data에 노이즈를 추가하여 인코더에 넣어 학습된 결과가 노이즈를 붙이기 전 데이터와의 error를 최소화하도록 학습
  - 카메라 이동 시 및 사람의 움직임으로 인한 초점이 맞지 않는 이미지를 Denoising Auto Encoder 기반 모델을 활용하여 잡음을 제거
  - 추가로 습기 등으로 인한 잡음, 실증 농가에 날아다니는 벌 등으로 인한 잡음을 제거하는 용도로 활용이 가능



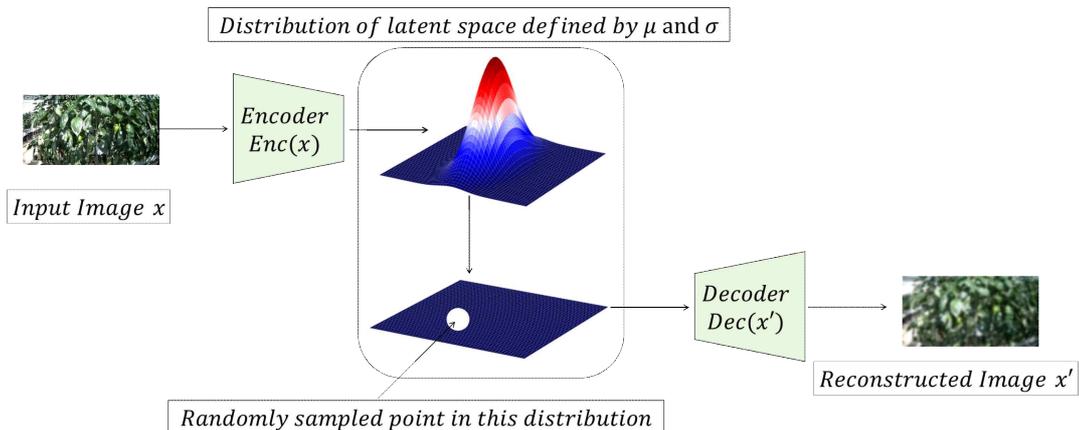
<Denoising Auto Encoder 과정>

- 노이즈가 있는 데이터의 경우 데이터의 양이 많지 않기 때문에 원본 이미지에 가우시안 노이즈를 추가하거나, 드롭아웃 등을 활용하여 입력 데이터에 노이즈를 발생시키는 모듈을 구성
- 드롭아웃은 사전에 정의한 확률로 이미지의 픽셀 중 일부 개수를 삭제하여 노이즈를 추가한 것과 같은 효과를 낼 수 있음



<Denoising Auto Encoder의 노이즈 제거>

- Variational Auto Encoder를 통한 데이터 생성
  - 작물의 이미지 데이터를 수집하는데 있어 작물의 생장에 걸리는 시간이 있어 많은 양의 데이터를 확보하기 어려움
  - 이에 입력 이미지와 유사한 새로운 데이터를 생성하는 Variational Auto Encoder 모델을 개발
  - Variational Auto Encoder는 latent vector  $z$ 의 분포  $p(z)$ 는 가우시안 분포를 따른다는 가정하에 최대 우도 추정을 통해 실제와 유사한 분포를 근사하도록 구성

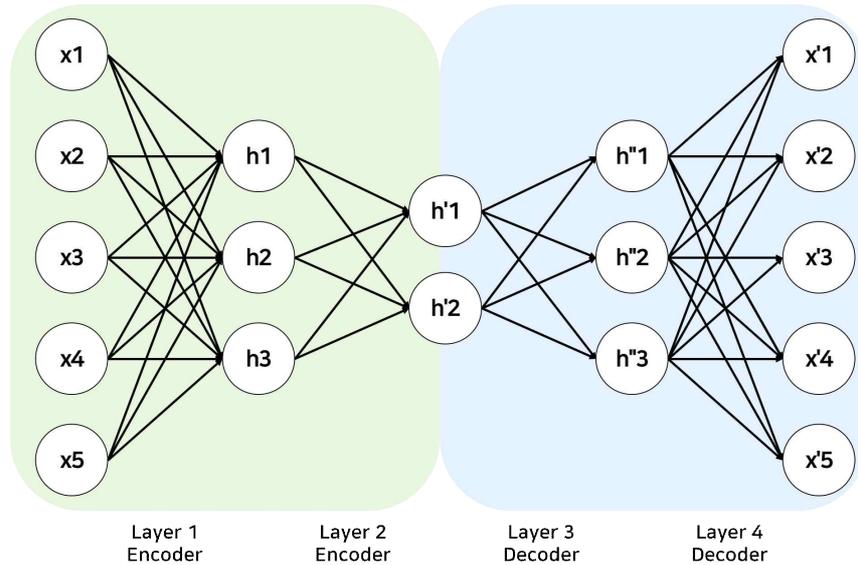


<Variational Auto Encoder 구조>

- Stacked Auto Encoder 모델을 통해 영상 특징 추출 및 영상 생성
  - Auto Encoder를 이용해 더 깊은 층을 쌓은 형태로, 입력 데이터의 추상적인 특성들을

fine-tuning하여 특징들을 학습시켜 이미지를 생성하는 방법

- 기존 Auto Encoder와 다르게 Deep Belief Network 구조를 사용
- DBN에는 Fully connected boltzmann machine을 변형한 Restricted Boltzmann Machine을 도입하였고, Gibbs sampling을 활용한 Contrastive Divergence를 활용하여 Maximum Likelihood Estimation 문제를 해결하여 성능을 향상시킴

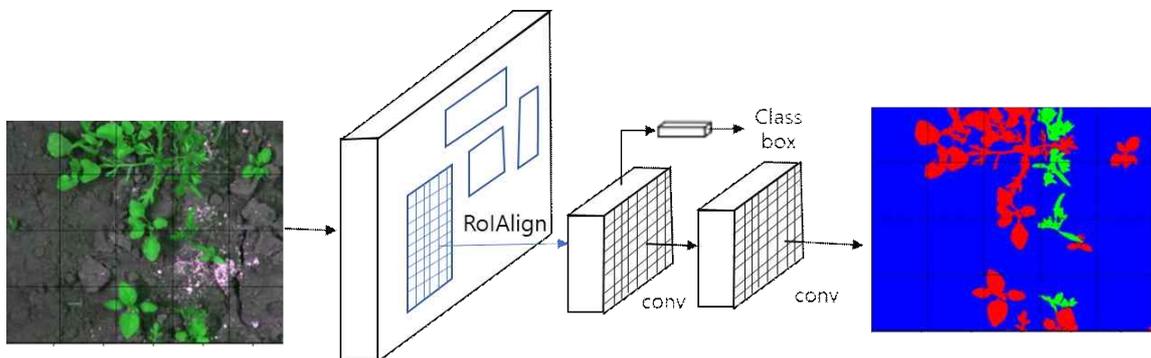


<Stacked Auto Encoder 구조>

□ 사전 학습된 인공지능 모델로 탐색(Detection), 분할(Segmentation)을 1차로 하고 2차적으로 보정하는 기술 개발

○ Mask R-CNN을 이용한 객체 탐지 및 세그멘테이션 모델 개발

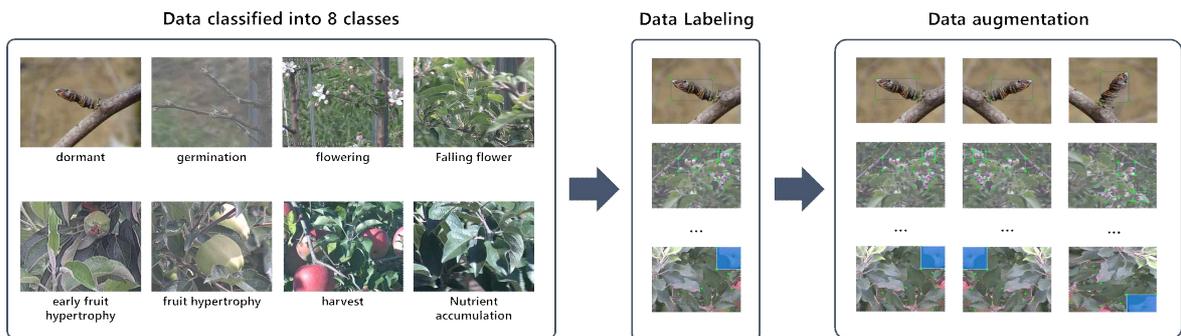
- 작물의 엽면적지수 (Leaf Area Index, LAI)를 추정하기 위해 작물의 잎, 줄기, 과수 등을 분할할 수 있는 기술이 필요
- 일반적인 데이터 집합을 이용해 학습된 인공지능 모델을 활용하여 적은 데이터로도 빠르게 학습하기 위해 사전 학습 방법을 활용
- 객체 탐지 및 분할 알고리즘은 MS-COCO 데이터셋으로 사전 학습된 Mask R-CNN 모델을 활용하여 객체 분할 모델을 설계
- 실증 농가에서 과채류의 영상 데이터 수집 중이며, 학습 데이터로 활용하기 위해 라벨링 작업을 진행



<Mask R-CNN 구조>

○ YOLO 기반 생육단계 인식 모델 개발

- 생육단계 인식 모델을 위한 데이터 수집, 라벨링, 증강
  - 8개 생육단계에 대한 데이터 수집 및 라벨링 진행  
(휴면기(dormant), 발아기(germination), 개화기(flowering), 낙화기(falling flower), 과실비대초기(early fruit hypertrophy), 과실비대기(fruit hypertrophy), 수확기(harvest), 양분축적기(nutrient accumulation))
  - 부족한 데이터 보안을 위한 데이터 증강 기법 사용 - Flip(반전), Rotation(회전) 활용  
(각 클래스별 500장, 총 4,000장 이미지 -> 11,135장으로 증강)
  - 학습/검증/테스트용 데이터셋 분할(학습 7 : 검증 2 : 테스트 1)



<생육단계 인식 모델 개발 과정>

- YOLO 모델을 이용한 생육단계 인식 모델 훈련 및 평가
  - 빠른 속도가 장점인 1-stage 기반의 객체 탐지 모델 YOLO를 활용하여 생육단계 인식 모델을 개발
  - YOLOv3, YOLOv3-tiny, YOLOv4, resized\_YOLOv4를 사용하여 생육단계 인식 모델을 비교 평가
  - Train 7,789장, Validation 2,236장, Test 1,110장(총 11,315장)을 사용하여 훈련 및 평가 진행
  - Batch Size 64, Subdivision 32, Learning Rate 0.0013, Max Batches 16,000, IoU Threshold 0.5로 파라미터 설정
  - 각 모델 별 클래스 AP 비교 결과 resized\_YOLO모델이 각 클래스별 AP에서 가장 좋은 성능을 보여주었으나, YOLO 모델의 전체 성능 비교 표를 보았을 때 전체적인 성능은 YOLOv3와 큰 차이를 보여주지 못함
  - 속도면에서는 YOLOv3보다 늦은 학습시간을 보여주오 YOLOv3모델이 적합한 모델이라 판단
  - 각 모델 별 클래스 AP 비교 결과

<각 모델 별 클래스 AP 비교 결과>

구분	YOLOv3	YOLOv3-Tiny	YOLOv4	resized YOLOv4
dormant	44.04%	20.31%	41.75%	44.79%
early_fruit_hypertrophy	72.58%	50.53%	75.26%	79.02%
falling_flower	46.38%	19.94%	46.66%	45.00%
flowering	65.01%	43.14%	57.21%	67.91%
fruit_hypertrophy	90.33%	82.67%	88.24%	91.44%
germination	63.96%	29.03%	52.60%	64.66%
harvest	94.95%	87.59%	92.08%	94.83%
nutrient_accumulation	79.29%	77.45%	72.94%	81.05%

- 모델 전체 성능 비교 결과

<모델 전체 성능 비교 결과>

구분	YOLOv3	YOLOv3-Tiny	YOLOv4	resized YOLOv4
Precision	0.68	0.43	0.86	0.64
Recall	0.63	0.45	0.35	0.61
mAP	0.695667	0.513359	0.658419	0.710881
F1-Score	0.65	0.44	0.49	0.65

- YOLOv3 객체인식 결과

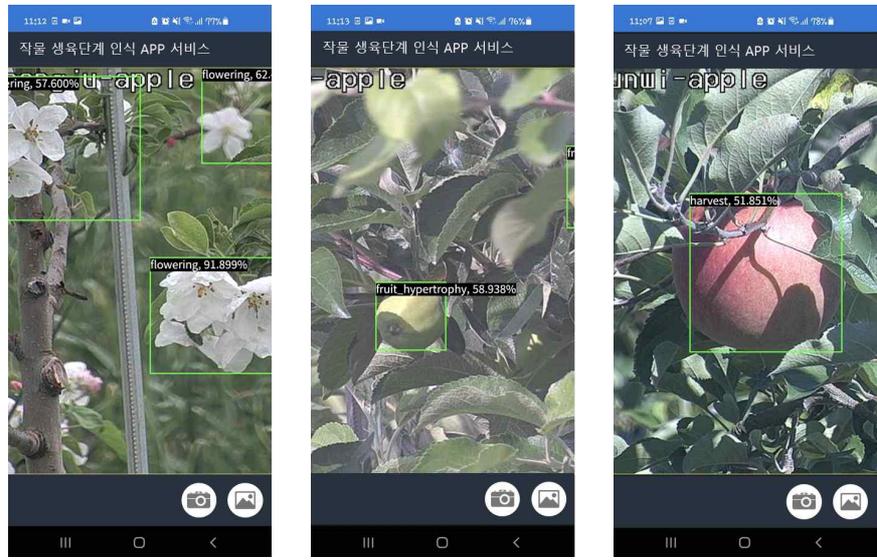
<YOLOv3 객체인식 결과>

클래스	객체인식 결과 이미지	클래스	객체인식 결과 이미지
dormant		fruit_hypertrophy	
early_fruit_hypertrophy		germination	
falling_flower		harvest	
flowering		nutrient_accumulation	

클래스	객체인식 결과 이미지	클래스	객체인식 결과 이미지
dormant		fruit_hypertrophy	

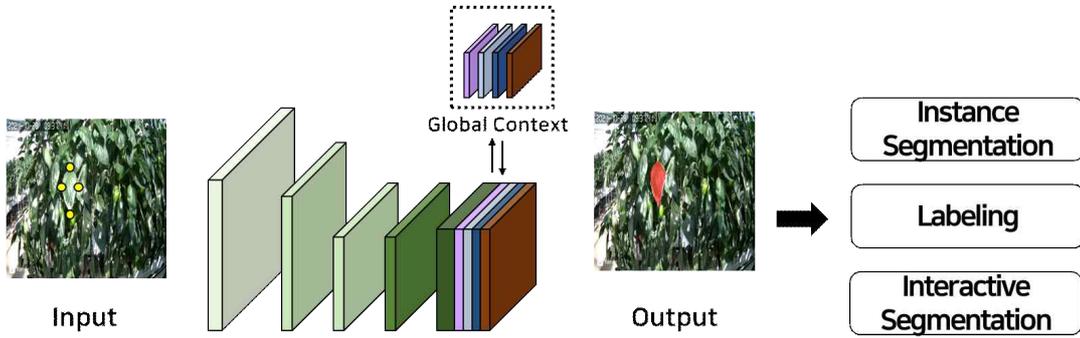
early_fruit_hypertrophy		germination	
falling_flower		harvest	
flowering		nutrient_accumulation	

- 모바일&웹 기반 서비스 제공 가능



<모바일&웹 기반 서비스>

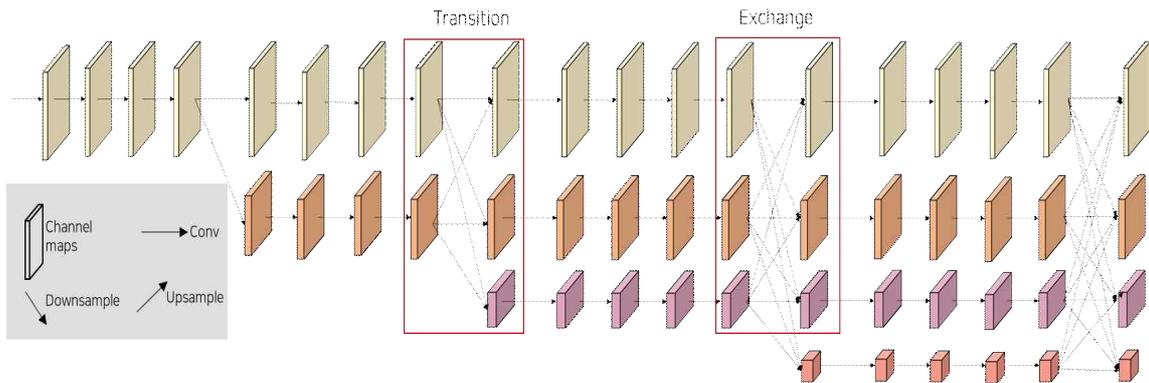
- 픽셀 레벨의 작물 영상의 주석을 만들 때(Polygon) 수작업을 최소화 할 수 있는 polygon 포인트들의 추론 기술 개발
  - Deep Extreme Cut(DEXTR)을 활용한 작물 영역 추론 기술 설계
    - 이미지에서 개체를 분할하여 라벨링할 객체의 4개의 키 포인트(상, 하, 좌, 우)를 선택
    - 각 포인트에 중심을 둔 가우스를 포함한 CNN의 입력에 추가 채널을 두어 포인트들과 일치하는 객체 세그멘테이션을 학습
    - PASCAL VOC, MS-COCO 데이터셋을 사전학습하여 활용
    - Instance segmentation, Labeling, Interactive Segmentation를 적용



<Deep Extreme Cut(DEXTR) 구조>

○ E-HRNet을 활용한 Semantic Segmentation 모델 개발

- HRNet을 활용한 Semantic Segmentation 모델 설계
  - HRNet은 고해상도의 특징맵들을 유지하면서 그와 평행하게 고해상도보다 작은 크기를 가진 특징맵들을 분기된 가지로 유지
  - 각기 다른 해상도로 이루어진 가지 안에 있는 특징맵을 서로 융합하여 새로운 특징맵을 생성하여 서로 다른 해상도의 정보를 교환해 더 풍부한 정보를 포함
  - 이렇게 생성된 특징맵을 토대로 업샘플링 과정을 보다 정밀하게 하여 픽셀 단위의 클래스와 위치를 판별
  - HRNet은 4단계 구조로 이루어져있으며, 1단계는 ResNet-50과 같이 너비 64채널의 bottleneck구조이고, 2,3,4 단계는 각각 변환 유닛(Transition Unit)과 교환 유닛(Exchange Unit)으로 구성
  - 변환 유닛은 서로 다른 가지의 특징맵을 융합하여 새로운 특징맵을 만들며, 교환 유닛은 생성된 다른 가지의 특징맵의 정보를 교환하여 특징을 융합

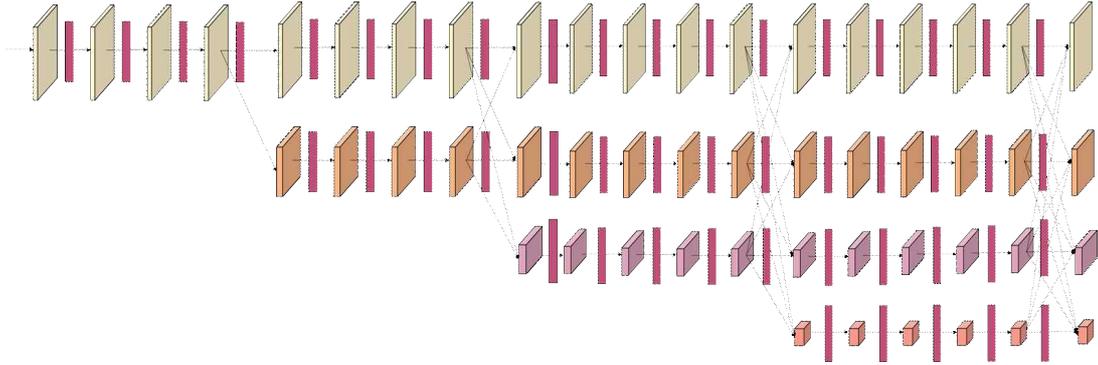


<HRNet 구조>

- Squeeze 및 Excitation을 이용한 E-HRNet 모델 설계 및 개발
  - 기존 HRNet은 평행하는 가지들 사이의 특징을 융합하여 특징을 효율적으로 추출하지만 다운샘플링 과정에서 여전히 정보 손실이 발생
  - Squeeze 및 Excitation을 이용한 E-HRNet 모델의 경우 기존 HRNet의 Convolution Block의 끝마다 어텐션 모듈을 추가함으로써 객체 영역 내 전역적 문맥 정보가 재조정 될 수 있도록 함
  - 어텐션 모듈을 추가할 때 파라미터 수 증가는 필연적이기 때문에 연산 부하가 적은 SE Block을 사용하여 부하를 최소화하도록 함
  - Convolution Layer 307개, Batch Normalization 306개, ReLU Layer 269개, Bottleneck Layer 4개, Basicblock 104개, High Resolution Module 8개로 이루어진 기존 HRNet 모델에 SE

Block을 추가하여 ReLU Layer가 71개 추가되었으며, 각 채널을 하나의 특징으로 압축하기 위한 GAP Layer는 108개 추가하였음.

- 또한, Fully Connected Layer 206개, Sigmoid Layer 108개, SE Block 108개가 추가되었다.
  - 기존 파라미터 수는 65.8M에서 0.4M 증가해 66.2M이 되어 1% 미만으로 증가했다.
- GFLOPs(GPU Floating point Operations Per Second)는 0.0004로 미비하게 증가



<Squeeze 및 Excitation을 이용한 E-HRNet 모델>

<Attention Module Pseudocode>

---

**Algorithm 1** : Attention Module Pseudocode, Torch-like

---

```

1 # input features with shape [N, C, H, W] = x
2 # N : batch size, C : channels, H : height, W : width
3 # SEBlock = Attention module
4 # reduction ratio = 16
5 def SEBlock (x)
6     squeeze = torch.nn.AdaptiveAvgPool2d(x)
7     excitation = torch.nn.Fully_connected(squeeze, out_channel / ratio)
8     excitation = torch.nn.ReLU(excitation)
9     excitation = torch.nn.Fully_connected(excitation / ratio, out_channel)
10    excitation = torch.nn.Sigmoid(excitation)
11    scale = x * excitation
12    return scale
13
14 def Convolution Basicblock (x)
15    residual = x
16    out = torch.nn.Convolution(x)
17    out = torch.nn.BatchNorm2d(out)
18    out = torch.nn.ReLU(out)
19    out = torch.nn.Convolution(out)
20    out = torch.nn.BatchNorm2d(out)
21    out = torch.nn.SEBlock(out)
22    out += residual
23    out = ReLU(out)
24    return out

```

---

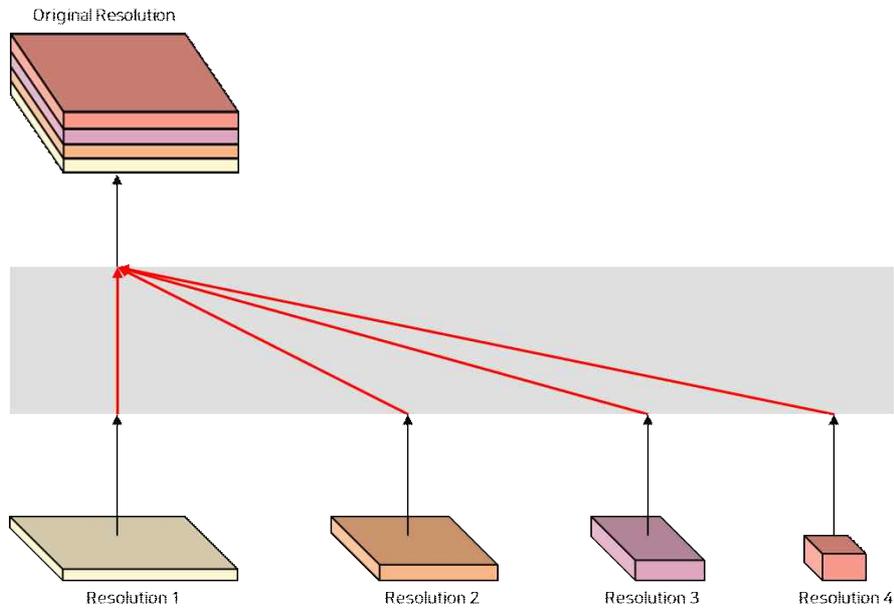


그림 179

- 벤치마크 데이터셋을 활용한 모델 성능평가
  - Semantic Segmentation의 경우 수집 데이터에 라벨링에 대한 시간이 많이 소요되어, 기존 벤치마크 데이터셋을 활용하여 모델에 대한 성능평가 수행
  - Scene Parsing 데이터셋인 Cityspcaes와 Human Parsing 데이터셋인 LIP, 일반적인 이미지 데이터셋인 PASCAL Context를 사용 비교평가 수행
  - Cityspcaes를 활용한 비교평가

Method	#Params[M]	GFLOPs	mIoU	MeanACC
HRNetV2-W18	15.4	7.774	63.6	72.9
HRNetV2-W48	65.8	174.043	79.4	<b>87.1</b>
Proposed Method	66.2	174.044	<b>79.7</b>	87.0

Method	backbone	mIoU
UNet++	ResNet-101	75.5
DeepLabv3	Dilated-ResNet-101	78.5
DeepLabv3+	Dilated-Xception-71	79.6
Proposed Method	HRNetV2-W48	<b>79.7</b>

- LIP를 활용한 비교평가

Method	#Params[M]	GFLOPs	mIoU	MeanACC
HRNetV2-W18	15.4	3.3798	13.4	19.1
HRNetV2-W48	65.8	75.9817	52.6	65.4
Proposed Method	66.2	75.9821	<b>53.0</b>	<b>65.7</b>

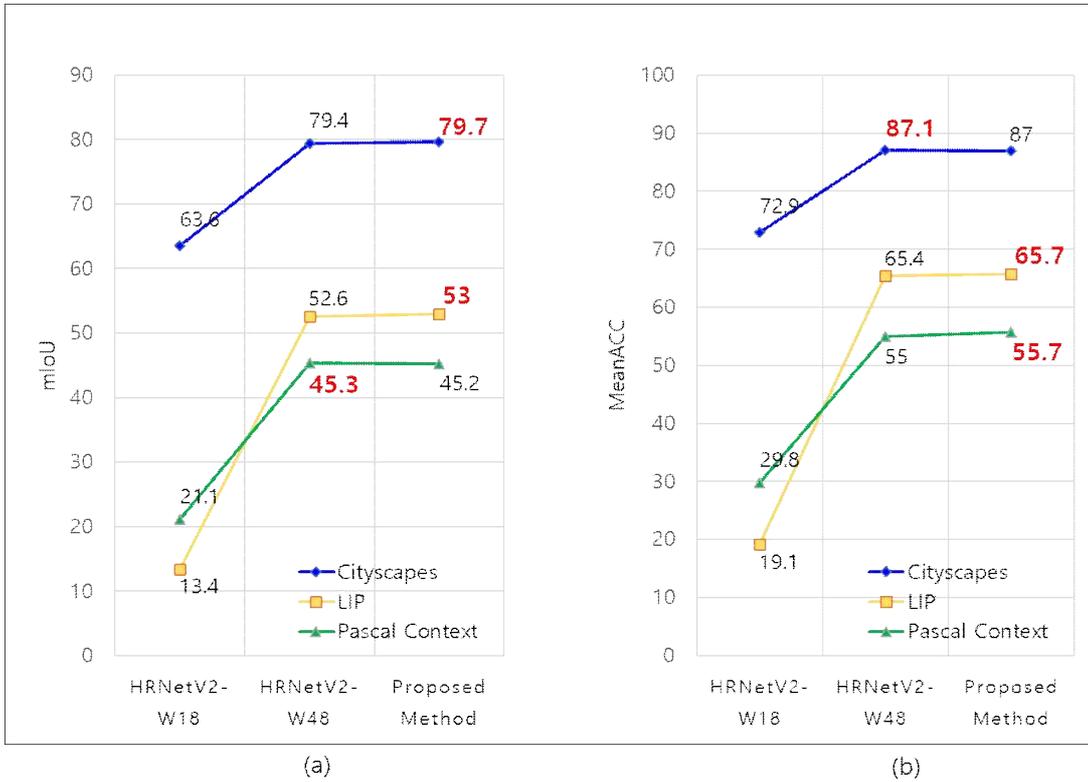
Method	backbone	extra	mIoU
Attention+SSL	VGG-16	Pose	44.7
DeepLabv3+	Dilated-ResNet-101	-	44.8
MMAN	Dilated-ResNet-101	-	46.8
SS-NAN	ResNet-101	Pose	47.9
MuLA	Hourglass	Pose	49.3
JPPNet	Dilated-ResNet-101	Pose	51.3
Proposed Method	HRNetV2-W48	-	<b>53.0</b>

- PASCAL Context를 활용한 비교평가

Method	#Params[M]	GFLOPs	mIoU	Mean_ACC
HRNetV2-W18	15.5	3.5484	21.1	29.8
HRNetV2-W48	65.8	76.8800	<b>45.3</b>	55.0
Proposed Method	66.3	76.8804	45.2	<b>55.7</b>

Method	backbone	mIoU
FCN-8s	VGG-16	35.1
BoxSup	-	40.5
HO_CRF	-	41.3
Piecewise	VGG-16	43.3
Proposed Method	HRNetV2-W48	45.2

- HRNet기반 모델 성능 비교 결과



<HRNet기반 모델 성능 비교 결과>

[1단계(2022)]

□ 기존에 축적된 스캔 파일의 텍스트를 추출하기 위한 OCR 인공지능 네트워크 개발

○ 영농일지 데이터 수집

• 흥양영농조합 영농일지 데이터 수집 진행

- 특정 형태의 영농일지 활용 데이터 수기 작성(작성일자, 작성자, 온실평균기온, 외부평균기온, 누적광량, 작업시간, 작업내용, 작업면적, 작업인원, 특기사항, 내일 작업사항 등 자유롭게 작성)

영 농 작 업 일 지 흥양영농조합법인				영 농 작 업 일 지 흥양영농조합법인				영 농 작 업 일 지 흥양영농조합법인				영 농 작 업 일 지 흥양영농조합법인				영 농 작 업 일 지 흥양영농조합법인			
작성일자	작성일자	작성일자	작성일자																
작성일자	작성일자	작성일자	작성일자																
작성일자	작성일자	작성일자	작성일자																
작성일자	작성일자	작성일자	작성일자																
작성일자	작성일자	작성일자	작성일자																

- 작성일자 : 2021.09.01. ~ 2022.04.01. (242건)

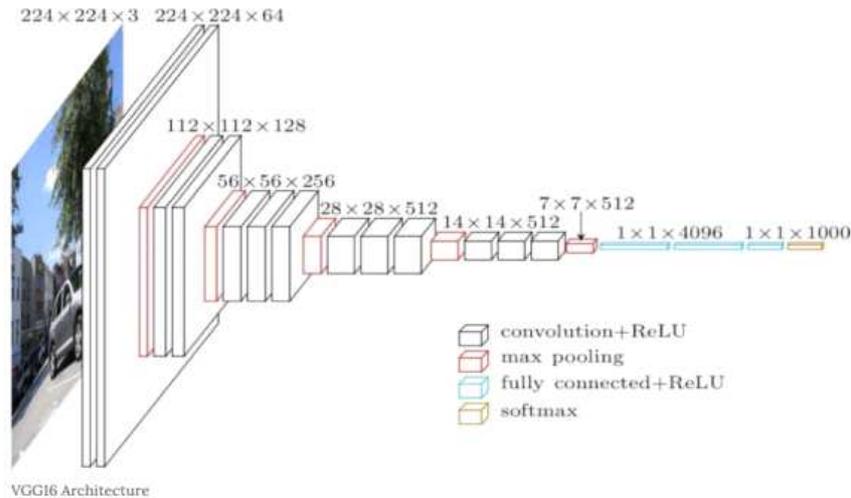
○ 텍스트 영역 추출을 위한 영역 검출 모델 설계 및 개발

• 텍스트 영역 추출을 위한 Feature Extraction 모듈 개발

- VGG, ResNet, GRCL 기반의 Feature Extractor를 통한 이미지 특징 추출

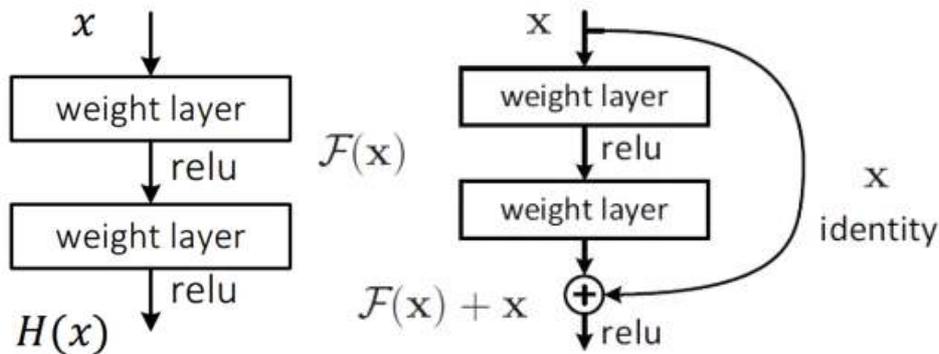
- VGG Feature Extractor

- : VGGNet은 Network의 깊이가 어떤 영향을 주는지 연구를 하기 위해 설계된 네트워크
- : 구조적인 측면에서 굉장히 간단하데 성능은 뛰어나 변형을 시켜가면서 테스트하기에 좋은 모델
- : VGGNet-16은 Layer 수가 16개이며, 망이 깊어지면 출력단에서 하나의 픽셀이 담고 있는 정보의 양이 많아지며 복잡한 문제를 해결 가능
- : 필터의 경우 3x3으로 strid는 1로 고정, MaxPooling은 2x2에 strid를 2로 하여 진행



- ResNet Feature Extractor

- : ResNet은 Residual Neural Network로,네트워크의 출력값이  $x$ 가 되도록  $H(x)-x$ 를 최소화하는 방향으로 학습을 진행
- :  $F(x) = H(x)-x$ 를 잔차라고 하며, 이 잔차를 학습하는 것인 Residual Learning이 적용
- : 또한 네트워크는 0이 되도록 학습시키고 마지막에  $x$ 를 더해서  $H(x)$ 가  $x$ 가 되도록 학습을 하여 gradient vanishing 문제를 해결함

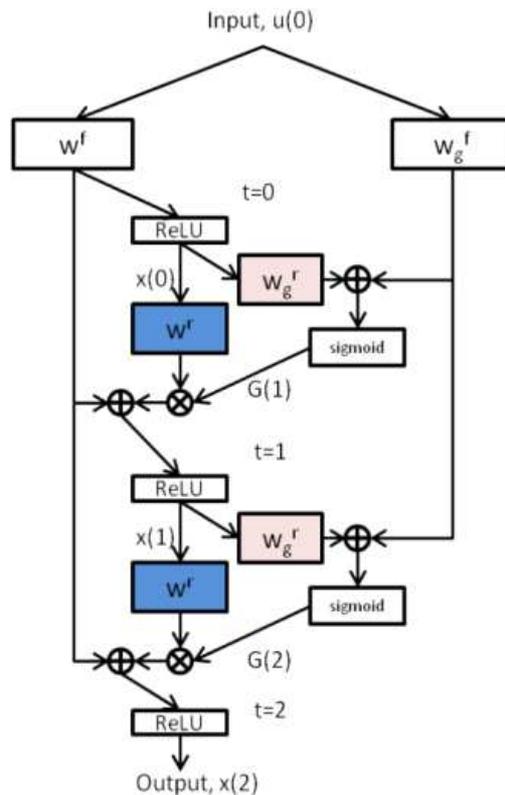


- : 최종적으로 이미지에서는  $H(x)=x$  가 되도록 학습 -> 네트워크의 output  $F(x)$ 는 0이 되도록 학습
- >  $F(x)+x = H(x) = x$ 가 되도록 학습시켜, 미분을 해도 미분값이 최소 1이상 -> 모든 레이어에서 gradient가  $1+F'(x)$  이므로 gradient vanishing 현상을 해결
- : ResNet은 네트워크의 깊이에 따라 ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-152 등 다양한 구조로 변형이 가능

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

- GRCL Feature Extractor

- : Gated Recurrent Convolution Neural Network로 recurrent convolution 레이어에 있는 컨텍스트 모듈을 제어하기 위해 RCNN에 gate가 더해진 모델
- : gate는 컨텍스트 모듈 제어뿐만 아니라 CNN의 정보와 RNN의 정보를 균형 있게 제어하는 역할을 함

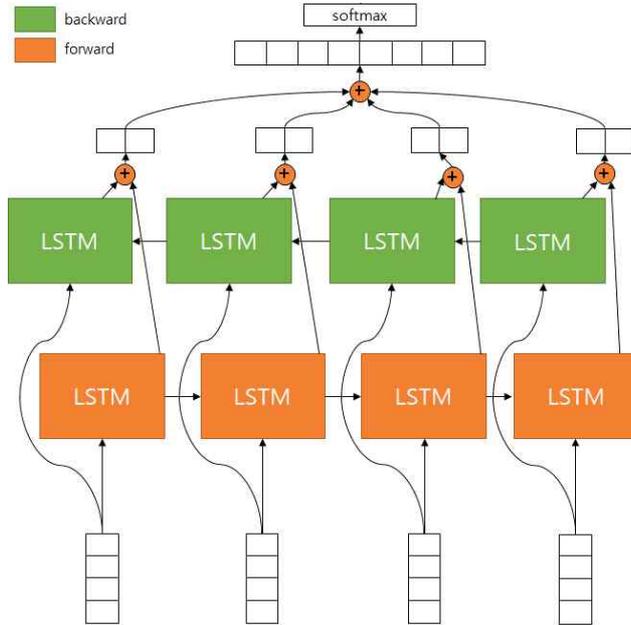


○ 텍스트 인식을 위한 sequence modeling 모델 설계 및 개발

- Bidirectional LSTM을 활용한 sequence modeling
  - BiLSTM은 양방향성을 갖는 LSTM으로 Forward 정방향 LSTM과 backward 역방향 LSTM이 있음
  - : Tokenizing과 Embedding 과정이 완료된 Sequence 입력이 들어오면 forward방향의 LSTM과

backward방향의 LSTM에 각각 입력

- : forward 방향 LSTM의 t번째 Cell과 backward 방향 LSTM의 N-t번째 Cell의 은닉값을 Concat
- : 모든 Concat 결과물을 가지고 최종 출력을 뽑게됨



○ 텍스트 인식 모델 실행 결과

영농작업일지  
중앙영농조합법인

작성일자	2021년 9월 16일 (수요일)	작업자	윤 경배
작성시간	23.5	작업내용	3농장 경작 완료

image_path	predicted_labels	confidence score
./temp\0.jpg	영농작업일일지	0.3309
./temp\1.jpg	중앙영농조합법인	0.6887
./temp\2.jpg	작성일자	0.9955
./temp\3.jpg	2021년9	0.6733
./temp\4.jpg	월15일	0.3523
./temp\5.jpg	(수요일)	0.9455
./temp\6.jpg	작성자	0.9956
./temp\7.jpg	윤 경배	0.3369
./temp\8.jpg	날씨사항	0.8427
./temp\9.jpg	은실명군기온	0.3578

영농작업일지  
중앙영농조합법인

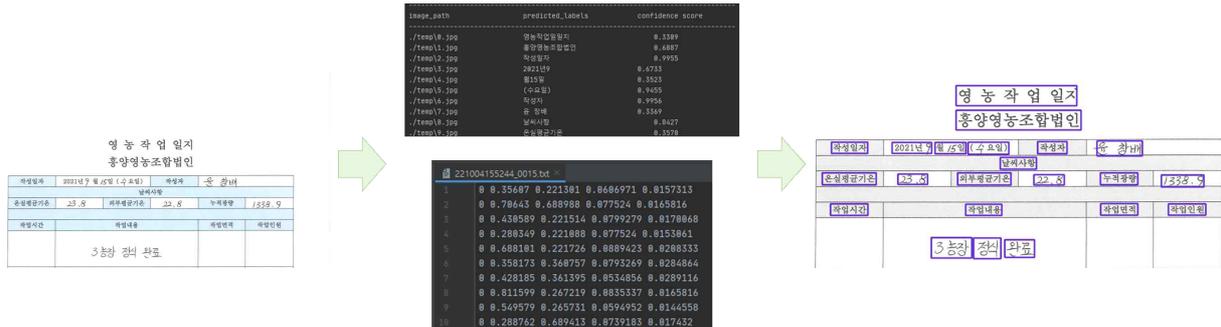
작성일자	2021년 9월 17일 (목요일)	작업자	윤 경배
작성시간	26.1	작업내용	1농장 B동 원형정립
		작업내용	마중알 경수기 파쇄 연결 후 가동
		작업내용	1농장 인력 조성
		작업내용	육묘장 피시드 연결 라인 복귀 (포피 대비)

image_path	predicted_labels	confidence score
./temp\0.jpg	영농작업일일	0.6023
./temp\1.jpg	중앙영농조합법인	0.5700
./temp\2.jpg	작성일자	0.9963
./temp\3.jpg	2021년	0.9987
./temp\4.jpg	요일일	0.0551
./temp\5.jpg	작성자	0.9954
./temp\6.jpg	윤	0.4986
./temp\7.jpg	경배	0.5068
./temp\8.jpg	날씨사항	0.7868
./temp\9.jpg	은실명군기온	0.4727
./temp\10.jpg	26,	0.6709
./temp\11.jpg	1	0.8989
./temp\12.jpg	외부명군기온	0.4678
./temp\13.jpg	23.1	0.8251
./temp\14.jpg	누적광량	0.9880
./temp\15.jpg	2037.77	0.3567
./temp\16.jpg	작성시간	0.9965
./temp\17.jpg	작업내용	0.9949

□ 농업 텍스트 정보 전처리 및 가시화 기능 개발

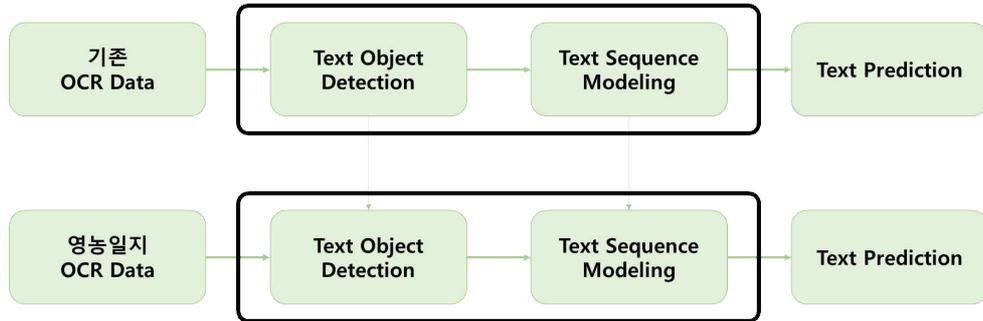
- OCR 텍스트 인식 결과 및 정규표현식을 활용한 텍스트 정보 전처리
  - 날짜정보(년, 월, 일), 농장정보(농장명 정보), 환경정보(기온, 온도, 광량 등에 대한 텍스트) 등 특정패턴을 가진 텍스트에 대한 전처리 및 매핑 진행
  - 특정 패턴 외에 일반적인 농업 텍스트 정보의 경우 하나의 텍스트로 연결하여 저장

- 한국어에 대한 OCR 인식 성능이 불안정하여 N-gram 방식의 유사도 비교 필요
- 텍스트 정보 및 BBox 정보를 활용한 가시화 기능
- 텍스트 인식 결과에 대한 텍스트 정보와 BBox 정보를 활용 가시화 기능 개발
- 텍스트 인식 결과 및 BBox 정보 추출
- 텍스트 정보 전처리 결과 특정 패턴의 주요 정보를 바탕으로 BBox 가시화 기능 개발



□ 전이학습 모델을 이용한 자동 Natural Language Processing(NLP) 주석 툴 개발

- 기존 학습된 모델을 활용한 NLP 주석 툴 개발
- 한국어 자연어 처리의 경우 다양한 글자 조합으로 인해 글자 인식의 성능이 떨어지며, 학습을 위해서는 대량의 데이터와 오랜 학습 시간이 필요함
- 그러나, 기존 학습된 모델을 활용해 전이학습을 할 경우, 학습에 소요되는 시간과 데이터 수가 급격히 줄어들 수 있음
- 이를 위해 기존 한국어 데이터가 학습된 모델을 활용 NLP 주석 툴 개발



□ 미처리 데이터 활용 방안

- 미처리 영상 데이터 가공 처리를 통해 생육, 생장, 품질 데이터 분석에 필요한 데이터를 확보 가능
- 엔드 유저 관점에서 영상 데이터 학습을 통해 생육, 생장, 품질 등 이상상황이 발생 시 알림 서비스 등에 활용이 가능할 것으로 판단
- 미처리 텍스트 데이터 가공 처리를 통해 수기 작성된 영농일지를 디지털화하며, 이를 빅데이터 분석에 활용
- 엔드 유저 관점에서 영농일지 내용을 검색하여 기타 작업에 대한 자세한 사항을 파악하는데 도움을 줄 수 있음

[2단계(2023)]

□ 온실 작물 텍스트 데이터 검증을 위한 실증

- 영농일지 OCR 프로그램 실증을 위한 a-BRAIN 플랫폼 적용

- 영농일지 OCR 프로그램 기능별 모듈화 진행

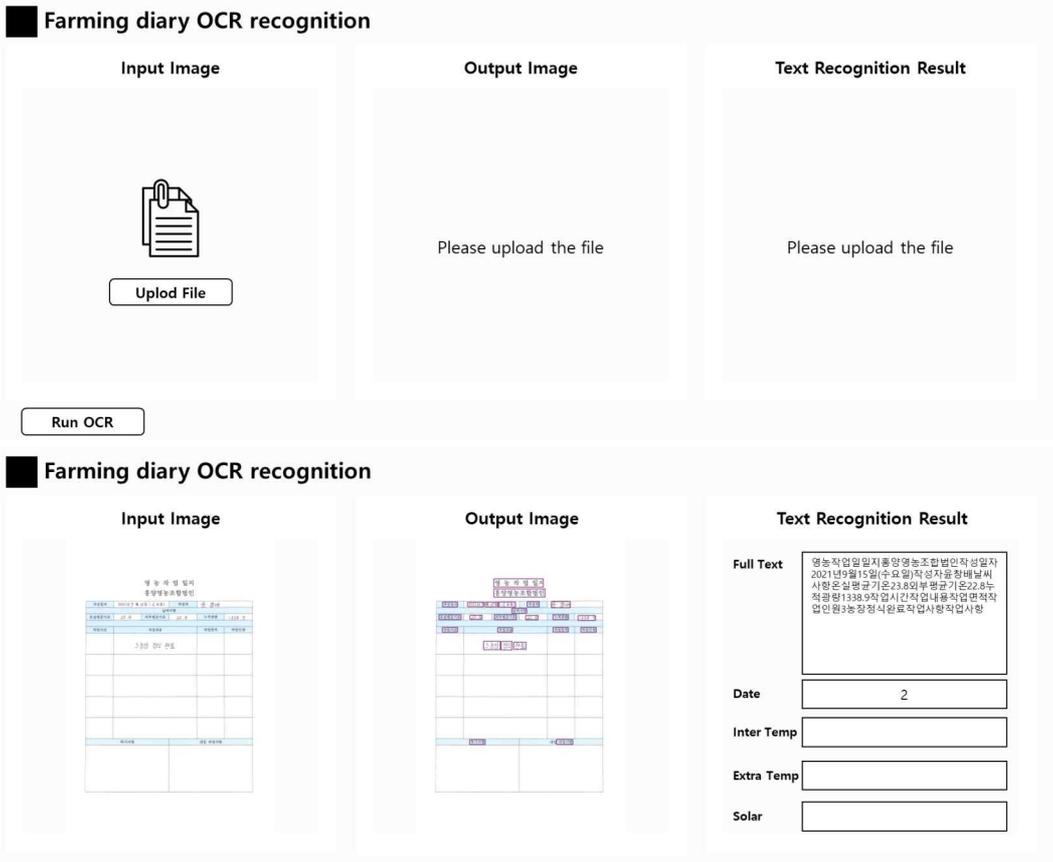
- Data Loading and Resize, Normalize
- 이미지 크기 및 기울기 보정
- 글자 영역 객체 탐지 기능
- 탐지 영역 이미지 추출
- OCR 인식 및 모델 실행

- 영농일지 OCR 인식 페이지 및 데이터베이스 설계

번호	컬럼명	데이터형식	설명
1	ID	string(255)	파일명(Primary Key)
2	full_text	TEXT	전체 인식된 텍스트
3	date	string(20)	추출된 날짜
4	temp_avg	float	추출된 온실 평균 기온
5	temp_avg_out	float	추출된 외부 평균 기온
6	light	float	추출된 누적 광량
7	etc	TEXT	기타 나머지 텍스트
8	original_image	string(255)	원본 이미지 저장 경로
9	detect_image	string(255)	텍스트 바운딩 박스가 포함된 이미지 저장 경로
10	detect_bbox_info	string(255)	바운딩 박스에 대한 텍스트파일 경로
11	cutting_image	string(255)	탐지된 글자 커팅 이미지 경로
12	cutting_image_count	int	탐지된 글자 커팅 이미지 개수
13	text_recognition_confidence_score	string(255)	텍스트 인식에 대한 신뢰 점수에 대한 텍스트 파일 경로

- 영농일지 OCR 프로그램 연동을 위한 데이터베이스 및 S3 저장소 입출력 코드 수정

- 영농일지 OCR 프로그램 a-BRAIN 플랫폼 기능 탑재



<a-BRAIN 플랫폼 영농일지 OCR 프로그램 적용 화면>

- 스캔한 문서 또는 이미지를 업로드하여 OCR 프로그램 실행
- 문자 영역 인식 결과 및 문자 인식 결과(Full Text, Date, Temp, Solar 등 특정 데이터 추출)를 출력
- 손글씨 인식 성능 고도화 및 의미 추출 등 고도화 필요

□ 온실 작물 데이터 가공을 위한 인공지능 네트워크 실증

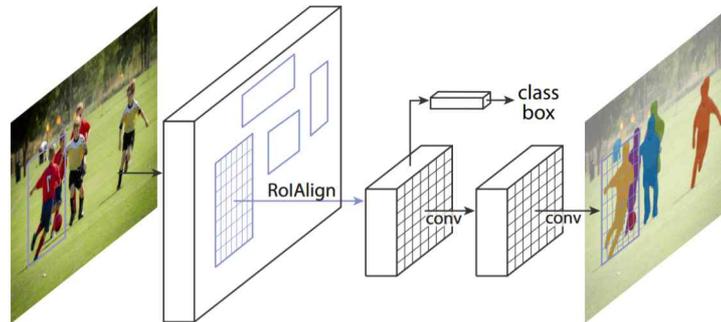
○ 온실 작물 데이터 가공 인공지능 네트워크 실증을 위한 객체 탐지 및 분할 모델 결합 및 고도화

- Mask R-CNN을 활용한 객체 탐지 및 분할 모델 고도화
  - 토마토 데이터를 활용한 객체 탐지 및 분할 모델 학습
    - 토마토 데이터세트 개요(작물 숙성도 및 크기에 따라 6 클래스로 분류)
      - fully\_ripend - 90%이상 붉게 변한 경우
      - half\_ripend - 30~89%가 붉게 변한 경우
      - green - 0~30%k 붉게 변한 경우

Class	Train	Test	Total
b_fully_ripened	343	77	420
b_half_ripened	506	130	636
b_green	1,511	343	1,854
l_fully_ripened	963	288	1,251
l_half_ripened	805	215	1,020
l_green	3,627	967	4,594

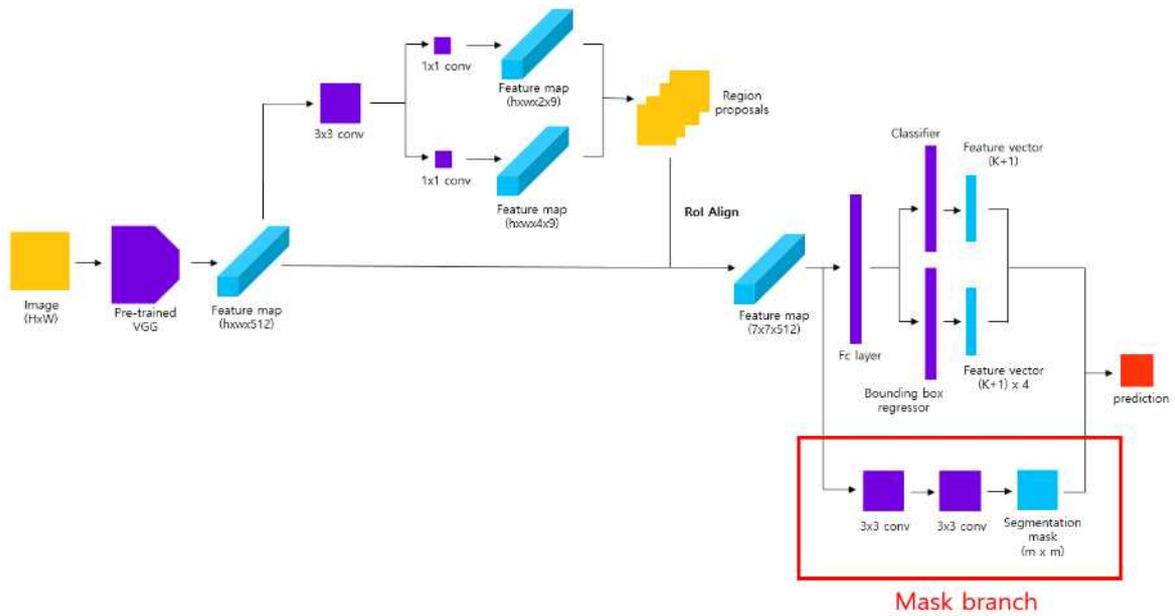
- Mask R-CNN

- ResNet-50 및 FPN(Feature Pyramid Networks) 기반 Mask R-CNN 설계 및 개발



<Mask R-CNN>

- Mask R-CNN은 classification branch, bbox regression branch, mask branch로 구성
- classification branch - RPN에서 얻은 RoI에 대하여 객체의 class를 예측
- bbox regression branch - bounding box에 대한 regression을 수행
- mask branch : segmetation mask를 예측(객체의 공간적 특징, 위치를 보존하는 RoIAlign Layer 추가)



<Mask R-CNN Configuration>

- RoIAlign 방법을 적용하여 RoI 위치 추정
  - RoI projection을 통해 Feature Map을 활용 다음 출력 Feature Map의 크기에 맞게 projection된 feature map을 분할
  - 분할된 하나의 cell에서 4개의 sample point 지정
  - Bilinear interpolation을 적용하여 4개의 sampled location 값을 연산 (2차원 좌표 상에서 두 좌표가 주어졌을 때 중간에 위치한 값을 추정하는 방법)

$$P \approx \frac{y_2 - y}{y_2 - y_1} \left( \frac{x_2 - x}{x_2 - x_1} Q_{11} + \frac{x - x_1}{x_2 - x_1} Q_{21} \right) + \frac{y - y_1}{y_2 - y_1} \left( \frac{x_2 - x}{x_2 - x_1} Q_{12} + \frac{x - x_1}{x_2 - x_1} Q_{22} \right)$$

- 모든 cell에 대하여 sampled location 값을 연산
- 하나의 cell에 있는 4개의 sampling point에 대하여 max pooling을 수행
- Mask R-CNN 모델 학습 : 이미지 전처리
  - 원본 이미지의 width, height 값 조정 짧은 길이에 맞춰 resize 수행 (default 값 - target size=800, maximum size = 1333)
  - ResNet-FPN 기반 feature pyramid 생성
  - RPN을 활용한 feature pyramid 별 Resion proposal 출력
  - top-N RoI 선택(anchor N = 12,000, NMS threshold = 0.7, anchor box N = 2,000)
  - ROIAlign을 활용한 feature map 출력(7x7 feature map)
  - Fast R-CNN 기반 Classification 및 Bounding box regression 수행
  - Mask branch를 사용한 segmentation 수행
  - Multi task Loss 기반 학습 수행

$$L = L_{cls} + L_{box} + L_{mask}$$

- Mask R-CNN 모델 훈련 및 평가 : 모델 학습 스케줄

```
# training schedule for 1x
train_cfg = dict(type='EpochBasedTrainLoop', max_epochs=12, val_interval=1)
val_cfg = dict(type='ValLoop')
test_cfg = dict(type='TestLoop')

# learning rate
param_scheduler = [
    dict(
        type='LinearLR', start_factor=0.001, by_epoch=False, begin=0, end=500),
    dict(
        type='MultiStepLR',
        begin=0,
        end=12,
        by_epoch=True,
        milestones=[8, 11],
        gamma=0.1)
]

# optimizer
optim_wrapper = dict(
    type='OptimWrapper',
    optimizer=dict(type='SGD', lr=0.02, momentum=0.9, weight_decay=0.0001))

auto_scale_lr = dict(enable=False, base_batch_size=16)
```

- 성능평가는 detection 문제에서 사용되는 mAP(mean Average Precision)와 segmentation 문제에서 사용되는 mIoU(mean Intersection-over-Union)을 사용
- mAP는 각각의 class에 대한 AP의 평균을 의미하며, mIoU는 각각의 class에 대한 IoU의 평균을 의미

- 모델 전체 성능 비교결과

model	detection	segmentation
mAP	0.644	0.6560
mAP_50	0.8210	0.8170
mAP_75	0.7340	0.7330
mAP_s	0.0	0.0
mAP_m	0.1450	0.1270
mAP_l	0.6790	0.6930

- Mask R-CNN 객체인식 및 분할 결과

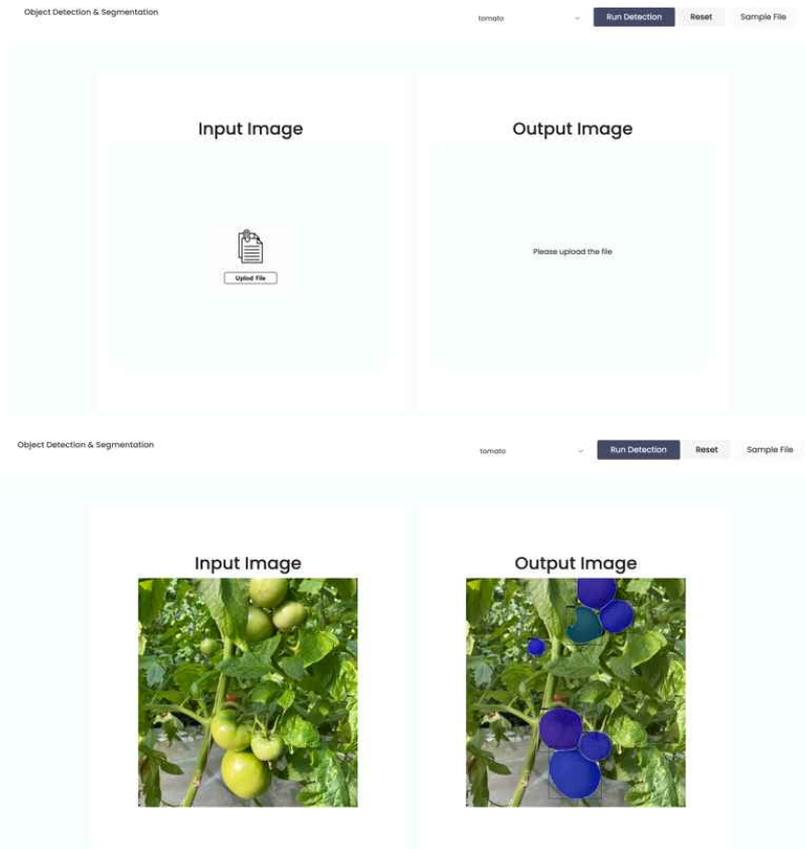


- 객체 탐지 및 분할 인공지능 네트워크 실증을 위한 a-BRAIN 플랫폼 적용
  - 페이지 및 데이터베이스 설계

번호	컬럼명	데이터형식	설명
1	c_id	int	crop ID
2	c_name	string(255)	작물 이름(작물 선택 리스트)
3	c_model	string(255)	작물별 detection model weight 경로(gpu 서버)
4	c_config	string(255)	작물별 detection model config 경로(gpu 서버)

번호	컬럼명	데이터형식	설명
1	ID	int	detection ID
2	original_image	string(255)	원본 이미지 저장 경로(S3 저장소)
3	detect_image	string(255)	결과 이미지 저장 경로(S3 저장소)
4	detect_json	string(255)	결과 json저장 경로(S3 저장소)
5	c_id	int	crop_id

- 웹 페이지 데모 구현을 위한 Inference 코드 작성
- 객체 탐지 및 분할 모델 연동을 위한 데이터베이스 및 S3 저장소 입출력 코드 수정
- 객체 탐지 및 분할 모델 a-BRAIN 플랫폼 기능 탑재

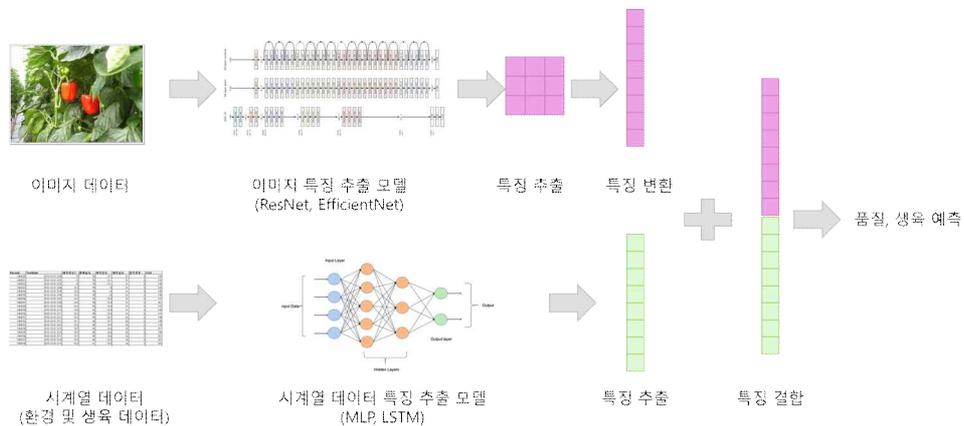


<a-BRAIN 플랫폼 객체 탐지 및 분할 모델 적용 화면>

- 작물 영상을 업로드하고 작물선택란에서 해당 작물을 선택 시 학습된 모델을 통해 객체 탐지 및 분할 실행
- 작물 별 학습데이터 확보가 가능한 경우 다양한 작물의 객체 탐지 및 분할이 가능하며, 이를 활용한 다양한 서비스 활용 가능

□ 인공지능 네트워크 기반 작물 성장/생육/품질 예측 시스템 개발

- 스마트 온실 작물의 이미지 데이터를 활용한 작물 성장/생육/품질 분류 알고리즘 개발
- 작물 영상 및 시계열 데이터를 활용한 Multimodal 작물 성장/생육/품질 분류 알고리즘 개발

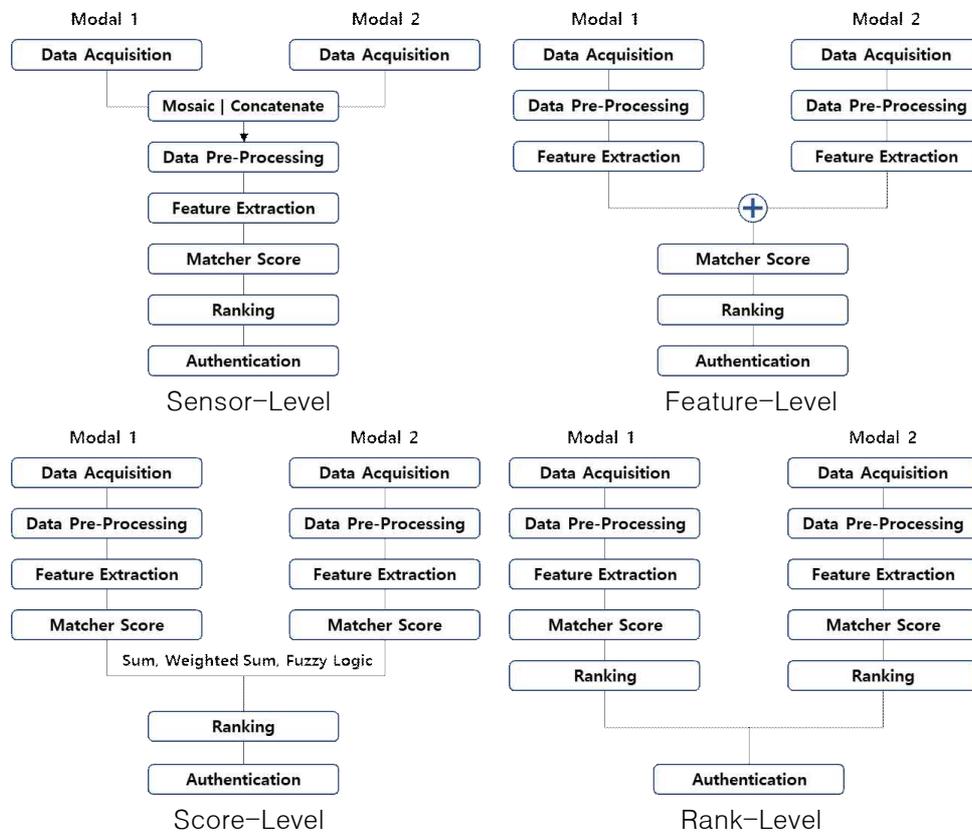


<Multimodal 작물 성장/생육/품질 분류 알고리즘>

- 생장, 생육, 품질 분류를 위한 ResNet, MLP(Multi Layer Perceptron) 기반 Multimodal 인공지능 모델 개발
- Multimodal 모델은 예측 성능을 높이기 위해 작물 영상 뿐만 아니라 온실 환경 및 재배와 관련된 수치 데이터를 입력으로 받아 각 데이터 유형별 특징을 추출하고 생장, 생육, 품질 예측에 활용
- 작물 영상은 전체 영상 중 작물의 위치에 해당하는 Bounding box의 좌표를 활용하여 이미지에서 배경을 제외하여 노이즈를 제거
- 데이터 증강, 정규화 등 데이터 전처리 등을 통한 일반화 성능 조정
- ResNet 기반 작물 영상 데이터 특징 추출 및 MLP 기반 시계열 데이터 특징 추출
- 추출한 두 특징 정보를 결합하여 최종 생장, 생육, 품질을 학습 및 예측

- Multimodal 융합 방법

- Multimodal 융합 방법은 ‘Sensor-Level’, ‘Feature-Level’, ‘Score-Level’, ‘Rank-Level’ 융합 방법 등이 있음



<Multimodal 융합 방법>

- ‘Sensor-Level’ 융합은 ‘Data-Level’ 융합이라고도 하며, 일반적으로 데이터 수집 후 결합, 특징 추출에 앞서 원시 데이터에서 직접 융합을 수행, 영상과 이미지 간의 데이터 융합이기 때문에 적합하지 않음
- ‘Feature-Level’ 융합은 원시 데이터에서 특징 추출 모델을 통해 추출된 특징맵을 융합하는 방법으로 이미지의 질감 및 구조적 특징, 시계열 데이터의 시간적 특징을 융합
- 각 특성에서 추출된 특징맵을 단순 연결(concatenation) 또는 Convolution Layer로 구성된 Fusion Block을 통해 융합

<Fusion Block 예시>

Layer	Input Channels	Output Channels	Kernel Size	Padding	Stride
Conv1	2×a	a	1	0	1
Conv2	a	a/2	3	1	1
Conv3	a/2	a	3	1	1

- ‘Score-Level’ 융합은 Matcher에서 생성된 매치 점수를 결합하는 방법

$$M_i = \frac{M_i - \min_{mci}}{\max_{si} - \min_{si}}$$

$$F_s = \sum_i^N (M_{m1i} + M_{m2i})$$

- ‘Rank-Level’ 융합은 매치 점수 비교 후, 생성된 순위를 결합하는 방법으로 생성된 순위를 결합할 때 Borda Count, Logistic Regression, Highest Rank Method 등을 사용해 순위를 결합
- Borda Count 방법은 각 특성별로 가능한 모든 클래스에 우선순위를 부여하는 투표 방식으로 다수의 투표자들이 각각 후보자들을 순위로 평가하고, 순위에 따라 점수를 부여한 뒤 이를 합산하여 순위를 도출

- Multimodal 융합 방법 성능평가

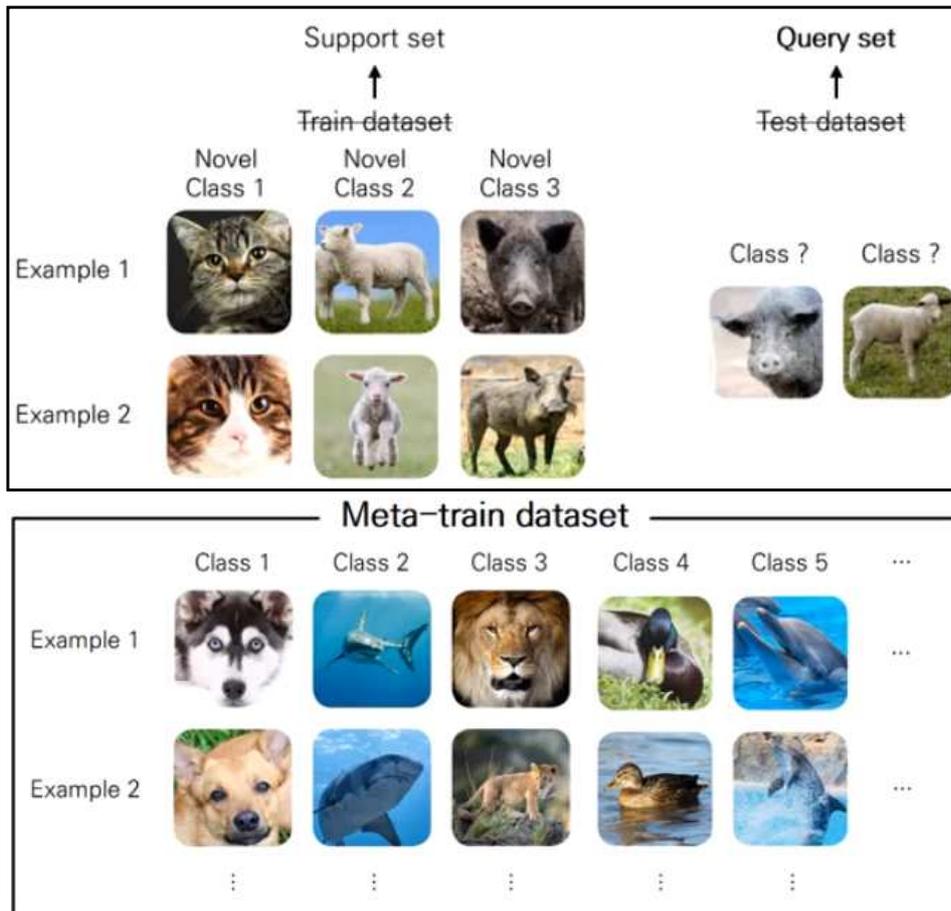
<성능평가 결과>

Target	Fusion Strategy	Model	Precision	Recall	f1-score	Accuracy
생장	Feature	ResNet-50, MLP	76.12	74.8	75.33	75.15
		<b>EfficientNet-B4, MLP</b>	<b>79.13</b>	<b>76.56</b>	<b>78.29</b>	<b>78.21</b>
	Score	ResNet-50, MLP	74.62	72.43	74.05	73.87
		EfficientNet-B4, MLP	76.8	75.09	76.04	77.27
	Rank	ResNet-50, MLP	74.56	73.5	73.1	74.06
		EfficientNet-B4, MLP	78.2	74.21	77.25	75.06
생육	Feature	<b>ResNet-50, MLP</b>	<b>85.79</b>	<b>87.25</b>	<b>87.97</b>	<b>88.61</b>
		EfficientNet-B4, MLP	84.97	81.62	83.21	80.9
	Score	ResNet-50, MLP	86.66	85.99	86.88	84.31
		EfficientNet-B4, MLP	83.54	79.39	82.03	81.96
	Rank	ResNet-50, MLP	86.33	85	85.61	83.38
		EfficientNet-B4, MLP	82.77	80.67	81.24	79.9
품질	Feature	ResNet-50, MLP	64.17	59.57	60.87	52.18
		<b>EfficientNet-B4, MLP</b>	<b>61.58</b>	<b>59.01</b>	<b>61.84</b>	<b>59</b>
	Score	ResNet-50, MLP	62.12	57.13	61.57	50.81
		EfficientNet-B4, MLP	59.18	59.81	58.49	55.79
	Rank	ResNet-50, MLP	64.42	57.32	61.89	50.31
		EfficientNet-B4, MLP	61.52	60.36	58.99	57.74

- Sensor-Level 융합 방법을 제외한 세 가지 융합 방법에 대해 성능평가를 비교했으며, 이미지의 경우 ResNet-50 및 EfficentNet-B4를 이용해 특징을 추출, 시계열 데이터의 경우 MLP를 이용해 특징을 추출
- 생장, 생육 분류 모두 Feature-Level 융합 방법이 전체적으로 비교적 우수한 수준의 성능을 보였으나, 특징 추출 모델에 따라, 다른 융합 방법이 우수한 경우도 발생
- 생장, 생육 분류 모델의 경우 일정 수준의 성능을 달성하였으나, 품질 분류의 경우 Accuracy가 60%이하로 좋지 않은 성능을 보였음 품질 분류 모델에 대한 보완 필요

○ No-reference Product Quality Assessment(NR-PQA) for agriculture 알고리즘 개발

- Few-shot, Zero-shot 기반 품질 측정 알고리즘 설계 및 개발

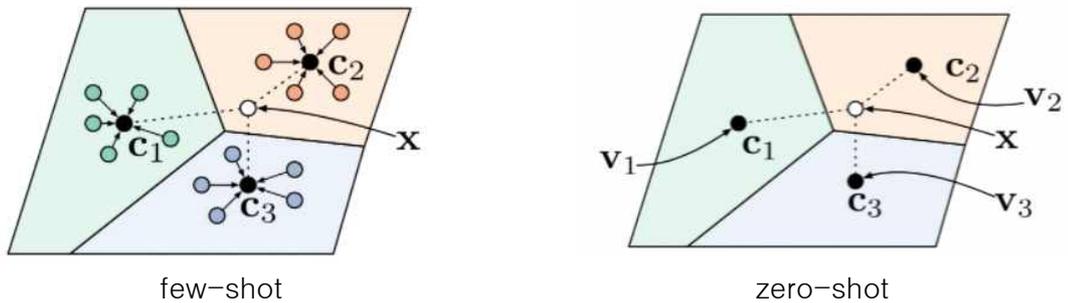


<Few-shot learning dataset>

- Few-shot Learning은 일반적으로 10개 미만의 한정된 데이터로 학습하는 방법론
- 지정된 훈련 데이터가 부족하거나, 효과적인 representation 학습이 필요할 경우, 학습하지 않은 새로운 클래스로 일반화를 필요로 할 때 활용 가능
- 메트릭 기반 방법, 모델 기반 방법, 최적화 기반 방법으로 구분 가능
- few-shot 기반 모델 학습 방법
  - n개의 클래스 별로 k개의 데이터로 구성이 되며, support set(학습 데이터)과 query set(테스트 데이터)으로 구분
  - Episode 기반 학습 수행

$$\theta = \underset{\theta}{\operatorname{argmax}} \underbrace{E_{L \sim T}}_4 \left[ \underbrace{E_{S \sim L, B \sim L}}_1 \left[ \underbrace{\sum_{(x,y) \in B}}_2 \log P_{\theta}(y|x, S) \right] \right] \underbrace{\quad}_3 \underbrace{\quad}_4$$

- ① Task sampling - 모든 class 중 학습시킬 N개의 class 샘플링 수행
  - ② Support set sampling - N개 class 별 t개의 examples 샘플링 수행
  - ③ Query set sampling - N개 class 별 u개의 examples 샘플링 수행
  - ④ Query set의 class를 가장 잘 맞추는 파라미터 학습
- prototypical networks 모델기반 few-shot, zero-shot 모델 구현
- 신경망을 사용하여 임베딩 공간에 대한 입력의 비선형 매핑을 학습
  - 임베딩 공간에 설정된 support set의 평균으로 class의 prototype을 생성
  - Prototypical networks는 각 class의 prototype 역할을 하기 위해 meta-data를 공유 공간에 임베딩하는 것을 학습
  - Classification을 수행할 때 임베딩 된 query point에서 가장 가까운 class prototype을 찾음
  - 각 class의 prototype을 만들고 Euclidean distance를 이용해 query point와의 거리를 계산



<Prototypical networks>

- 모델 성능평가
- 기존 prototypical networks의 특징 추출 모델은 GoogleNet과 기존 활용 중인 ResNet, EfficientNet을 활용하여 모델 구현 및 평가
  - few-shot learning의 경우 작물의 품질을 5단계로 분류하여 분류별 1장의 이미지를 사용하는 5-way 1-shot과, 5장의 이미지를 사용하는 5-way 5-shot, 품질에 대한 이미지가 없는 5-way 0-shot 평가
- 모델 성능평가 결과

<성능평가 결과>

Model	Feature Extraction	Accuracy		
		5-way 1-shot	5-way 5-shot	5-way 0-shot
prototypical networks	GoogleNet	32.5	41.6	25.9
	ResNet-50	64.2	68.9	56.8
	EfficientNet-B4	68.4	72.1	61.2

- 기존 GoogleNet을 이용한 prototypical networks는 매우 낮은 성능을 보였으나, ResNet-50, EfficientNet-B4를 이용한 prototypical networks는 성능이 낮지만 유의미한 성능을 보임
- 기존 특징 추출 모델 및 few-shot learning 모델을 개선할 경우 다양한 방면에 활용 가능할 것으로 판단

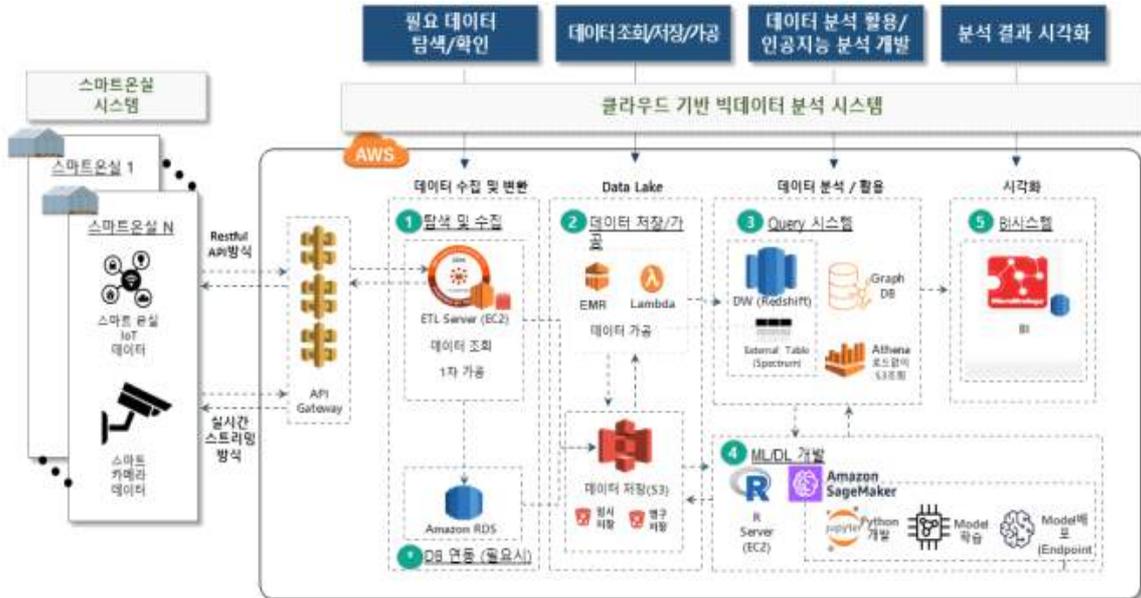
[GS ITM]

[1단계(2021)]

□ 클라우드 기반 스마트 온실 빅데이터 시스템 설계 및 인프라 구축

○ 클라우드 기반 스마트 온실 빅데이터 시스템 설계

- 온실의 센서와 제어기에서 생성되는 데이터와 연동 인터페이스를 표준화한 내역을 기반으로 온실과의 연동 인터페이스를 설계하고 데이터 저장소 구조 설계
- 본 빅데이터 시스템 아키텍처는 크게 4개의 영역으로 구분하여 설계 및 기술 연구 추진
  - 스마트온실의 다양한 데이터를 수집 및 서빙을 위한 API 영역
  - 수집된 데이터를 가공/저장/분석을 위한 Data Lake 영역
  - 수집 가공된 데이터를 사용자가 쉽게 분석하고, 조회할 수 있는 시각화(BI) 영역



<스마트온실 클라우드 기반 빅데이터 아키텍처>

- 작물 데이터를 가공/정제/분석용 AI 모델들의 소스/배포 관리를 위한 MLOps 영역

• 스마트온실 데이터 수집/서빙 API 설계

- 정형데이터, 텍스트 데이터, 이미지 데이터, 동영상 데이터에 대한 저장/통신 프로토콜 설계

- API 통신 방식: AWS API Gateway 적용, HTTP방식 (REST API), 파일 전송(sftp) 제공
- 정형 데이터, 텍스트 데이터, 이미지 데이터 수집
  - : HTTP 기반의 REST API 방식과 AWS S3 직접 전송 할 수 있는 sftp방식 제공
- 동영상 데이터 수집
  - : 스트리밍 서버 기반의 실시간 영상 수집 (AWS CloudFront + AWS Kinesis Video Streams)을 사용하여, AI 분석 모델의 실시간 학습 및 처리가 가능하도록 데이터를 제공
  - : 기 저장된 영상의 경우는 Workflow 엔진인 Airflow를 사용하여 배치 처리
- 부가적인 데이터 수집 방법 : 각 데이터별 topic 정의 후 Kafka에 저장하는 통신 API 제공
- 프로토콜 설계 : 인증 (ID기반 인증-역활과 권한 설계), 데이터 전송 포맷 설계 (Json, 파일 전송)



< AWS Kinesis Video Streams 동작원리 >

< AWS Kinesis Data Streams 동작원리 >

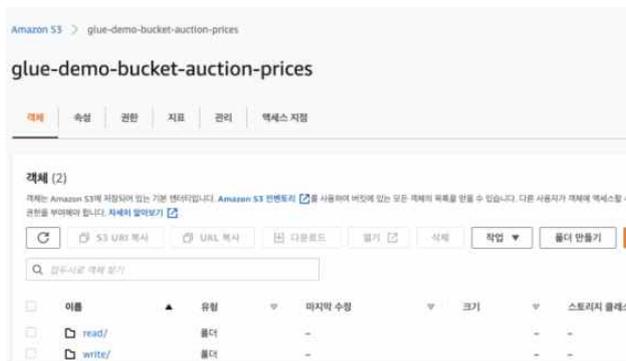
• Data Lake 아키텍처 및 ETL 설계

- 모든 데이터는 Data Lake에 저장하고, 메타데이터 추출(카탈로그), 모니터링 항목 및 상관 분석 항목에 맞도록 데이터 가공하도록 아키텍처 설계를 연구함
- Data Lake에는 원천 데이터, 가공 데이터, 최종 데이터를 모두 보관하게 하고, 이를 기준으로 데이터 분석에 직접 사용



<Data Lake 아키텍처 및 ETL 설계 절차>

- Data Lake에 저장되는 파일 포맷은 CSV, Excel, Parquet, ZIP 형태 지원
- 스마트폰으로부터 수집된 데이터에 대한 Data Lake 카탈로그는 AWS Glue 서비스를 사용하여 데이터를 분류 저장하고, 크롤러 정의를 통해 다양한 데이터 소스로부터 데이터를 수집하는 형태로 구성함
- ※ AWS Glue란? ETL 서비스로 AWS에 있는 RDS, Dynamodb, S3, Redshift 등에 있는 데이터를 불러오는 기능을 하는 AWS 서비스



<S3로부터 AWS Glue를 활용하여 데이터를 읽어드리는 화면>

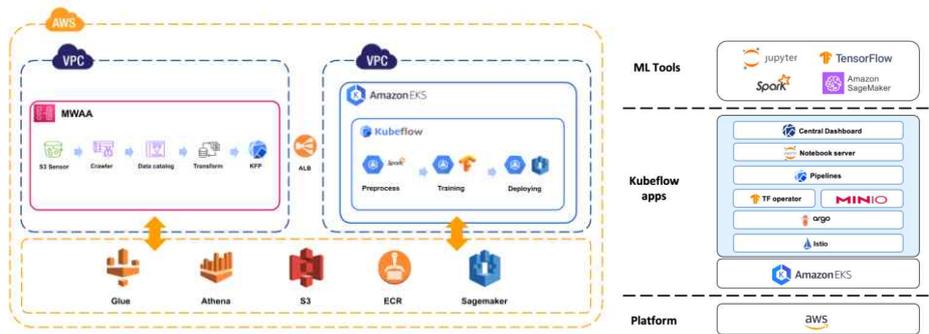
- ※ 크롤러란? 크롤러는 데이터 스토어에 연결하여, 우선순위가 지정된 분류자의 목록을 기준으로 데이터의 스미카를 결정한 다음, 데이터 카탈로그에 메타데이터 테이블을 생성하는 역할

이름	일정	상태	로그	마지막 런타임	중간 런타임	테이 뎀
glue-demo-crawler		Ready		0초	0초	0

<크롤러 설정화면>

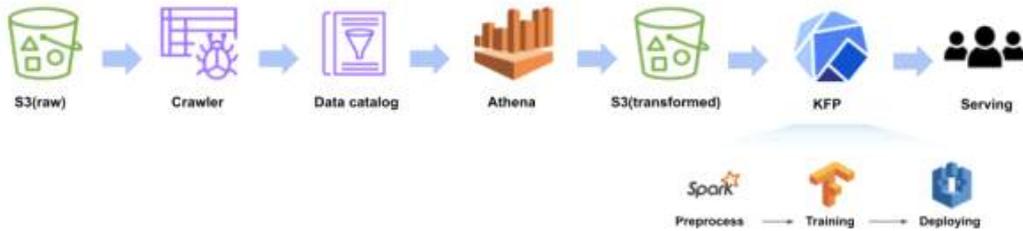
- ETL 및 데이터 가공 설계
  - 온실 환경 및 생산 데이터 등의 정형 데이터는 표준화 후 raw 데이터 형태로 처리
  - 데이터 가공: 시계열 기반 및 연관 이미지/동영상 정보 융합/가공(시간별, 생육 주기별 등)
- 실시간 모니터링에 사용할 시스템: AWS Redshift(선정), mariaDB, MongoDB 검증

• 스마트 온실 인공지능 모델 소스관리/배포를 위한 MLOps 설계 및 관련 기술 연구



<MLOps의 전체 시스템 구조>

- MLOps의 전체 아키텍처는 아래와 같으며, 이번 아키텍처 검증을 위해서 사용한 AI모델은 AWS SageMaker를 사용하여 검증을 추진함
- MLOps의 전체 파이프라인 흐름도는 아래와 같은 형태로 구현
  - ① S3 에 데이터 입력
  - ② 데이터 크롤링
  - ③ 데이터 1차 가공 후 S3에 외부테이블 형태로 저장
  - ④ ML pipeline (전처리, 학습, 배포) 실행



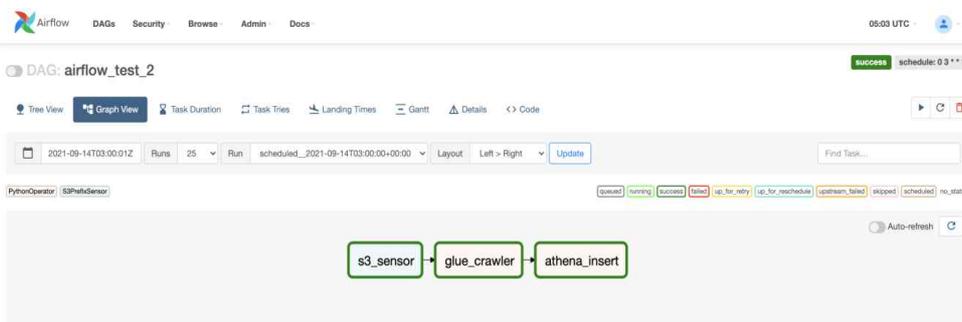
<MLOps 전체 파이프라인 flow>

- MLOps 전체 워크플로우 관리를 위해 Airflow와 KubeFlow(ML 워크플로우 엔진)에 대한 기술 연구를 진행했으며, 이 2개의 기술 중 Cloud 내 워크플로우 관리를 위한 Airflow 엔진 기반의 전체 아키텍처는 아래와 같으며, 해당 아키텍처에서는 자동 워크플로우 관리, 보안, 자동 패치, 모니터링 등 다양한 기능을 제공하도록 설계
  - 자동 Airflow 설정 - Amazon MWSA 환경을 생성할 때 Apache Airflow 버전을 선택하여 Apache Airflow를 빠르게 설정

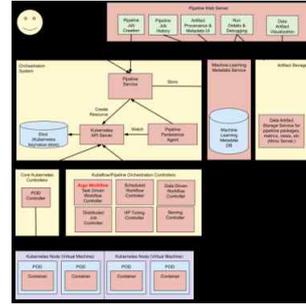
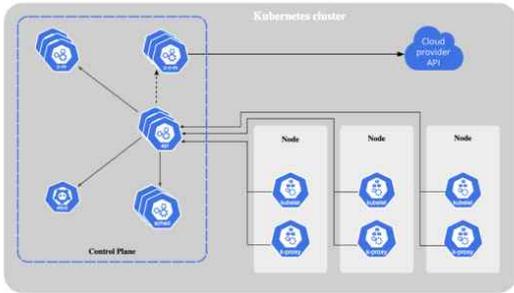
- 자동 스케일링 - 환경에서 실행되는 최소 및 최대 작업자 수를 설정하여 Apache Airflow 작업을 자동으로 확장하며, Amazon MWAA는 환경의 작업자를 모니터링하고 자동 크기 조정 구성 요소를 사용하여 사용자가 정의한 최대 작업자 수에 도달할 때까지 수요를 충족하기 위해 작업자 추가 가능
- 기본 제공 인증 - AWS Identity and Access Management(IAM)에서 액세스 제어 정책을 정의하여 Apache Airflow 웹 서버에 대한 역할 기반 인증 및 권한 부여를 활성화
- 기본 제공 보안 - Apache Airflow 작업자 및 스케줄러는 Amazon MWAA의 Amazon VPC에서 실행 . 또한 데이터는 AWS Key Management Service를 사용하여 자동으로 암호화되므로 기본적으로 안전한 환경 보장
- 공개 또는 비공개 액세스 모드 - 개인 또는 공개 액세스 모드를 사용하여 Apache Airflow 웹 서버에 액세스하며, 공용 네트워크 액세스 모드는 인터넷을 통해 액세스할 수 있는 Apache Airflow 웹 서버용 VPC 엔드포인트를 사용함. 비공개 네트워크 액세스 모드는 VPC에서 액세스할 수 있는 Apache Airflow 웹 서버용 VPC 엔드포인트를 사용
- 간소화된 업그레이드 및 패치 - Amazon MWAA는 새로운 버전의 Apache Airflow를 주기적으로 제공하며, 이러한 버전의 이미지는 Amazon MWAA 팀에서 업데이트 및 패치
- 워크플로우 모니터링 - Amazon CloudWatch에서 Apache Airflow 로그 및 Apache Airflow 지표를 확인하여 추가 도구 없이 Apache Airflow 작업 지연 또는 워크플로우 오류를 식별하며, Amazon MWAA는 환경 지표를 자동으로 전송하고 활성화된 경우 Apache Airflow 로그를 CloudWatch 전송이 가능한 형태로 운영
- AWS 통합 - Amazon MWAA는 Amazon Athena, AWS Batch, Amazon CloudWatch, Amazon DynamoDB, AWS DataSync, Amazon EMR, AWS Fargate, Amazon EKS, Amazon Kinesis Data Firehose, AWS Glue, AWS Lambda, Amazon Redshift, Amazon과의 오픈 소스 통합을 지원
- 작업자 플릿 - Amazon MWAA는 AWS Fargate의 Amazon ECS를 사용하여 요청 시 작업자 플릿을 확장하고 스케줄러 종단을 줄이기 위해 컨테이너를 사용할 수 있는 지원을 제공함. Amazon ECS 컨테이너에서 작업을 호출하는 운영자와 Kubernetes 클러스터에서 포드를 생성하고 실행하는 Kubernetes 운영자 지원
- AWS에서의 airflow를 통한 데이터 워크플로우 전처리와 실제 airflow 엔진 기반의 데이터 전처리 화면을 모듈 형태로 구현



<데이터 처리 워크플로우 단계 - AWS(클라우드)에서 데이터 전처리 하는 과정>



<에어플로우로 워크플로우를 구현한 화면>



<Kubeflow 구성도>

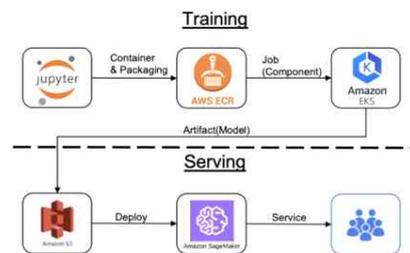
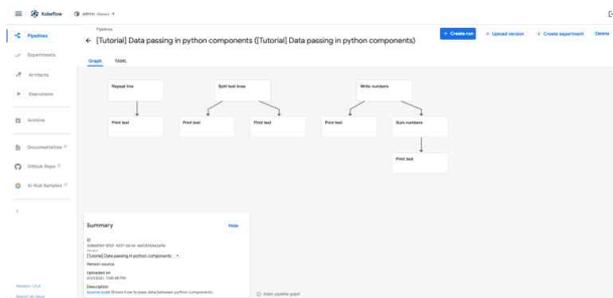
<Kubeflow 파이프라인 구성 아키텍처>

- MLOps 내에서의 AI 모델의 개발 및 배포를 위한 워크플로우 관리 엔진인 Kubeflow에 대한 기술 연구를 진행하였으며, 아래부터는 kubeflow 워크플로우 엔진과 해당 엔진 관리를 위한 kubernetes 기반의 전체 아키텍처를 구성
- Kubeflow Pipelines는 kubernetes 용 컨테이너 네이티브 워크 플로 엔진인 Argo Workflows를 기반으로 구성했으며, 일반적으로 높은 수준에서의 파이프 라인 구성

<Kubeflow Pipelines>

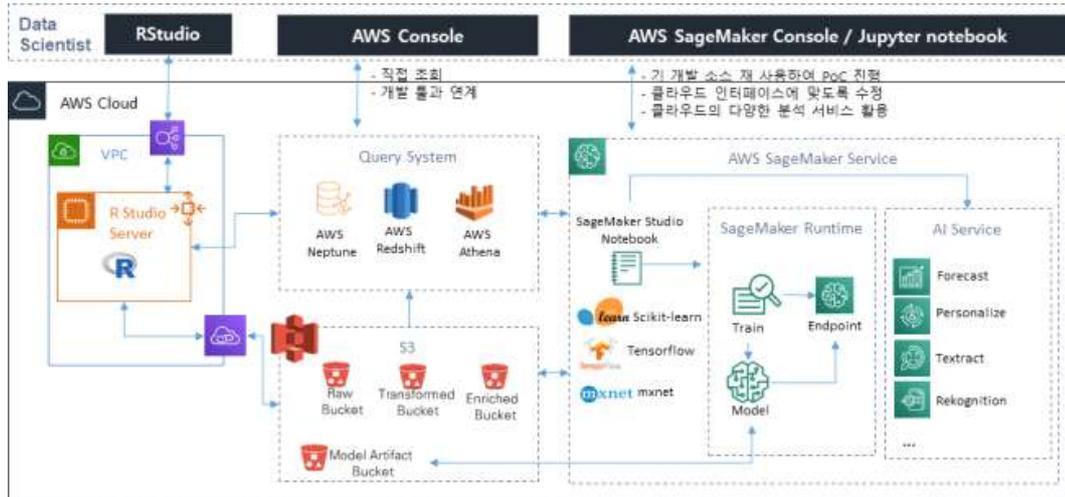
Python SDK	Kubeflow Pipelines 도메인 별 언어(DSL)를 사용하여 컴포넌트를 생성하거나 파이프라인을 지정
DSL 컴파일러	DSL 컴파일러는 파이프라인의 Python 코드를 YAML파일로 변환
파이프 라인 서비스	파이프라인 서비스를 호출하여 YAML파일에 정의된 파이프라인을 생성
Kubernetes 리소스	파이프라인 서비스는 Kubernetes API 서버를 호출하여 파이프라인을 실행하는 데 필요한 Kubernetes 리소스를 생성
오케스트레이션 컨트롤러	일련의 오케스트레이션 컨트롤러가 파이프라인을 완료하는데 필요한 컨테이너를 실행하며, 컨테이너는 가상 머신의 Kubernetes 포드 내에서 실행이 됨
아티팩트 저장소	Pod는 두 가지 종류의 데이터를 저장
메타데이터	실험, 작업, 파이프라인 실행 및 메트릭 데이터는 MySQL 데이터베이스에 저장
아티팩트	파이프 라인 패키지,보기 및 대규모 지표는 Minio 서버 또는 Cloud Storage 와 같은 아티팩트 저장소에 아티팩트를 저장

- Kubeflow 기반의 파이프라인은 아래와 같이 구성했으며, 실제 동작방식은 아래와 같음



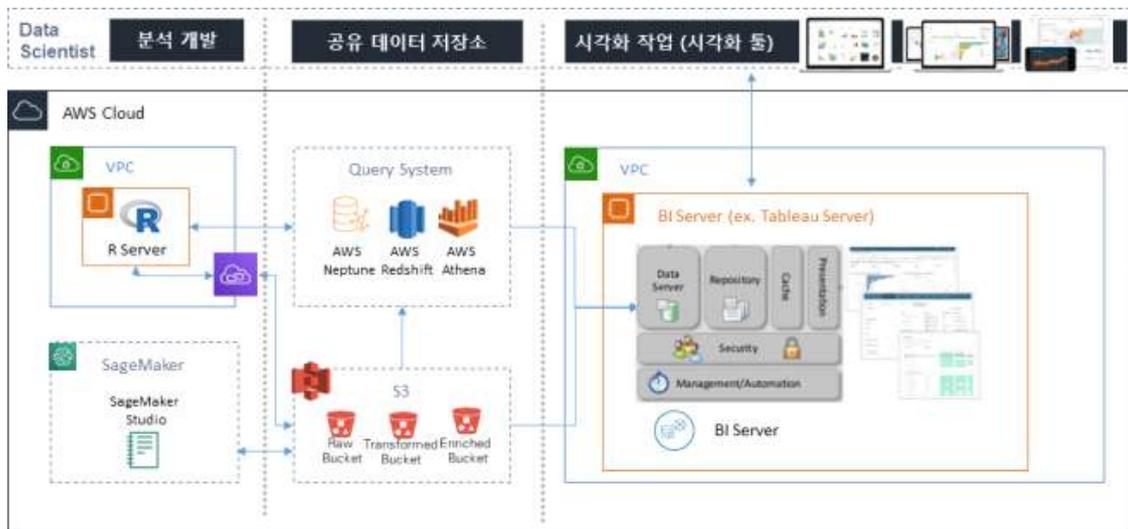
<Kubeflow 엔진으로 구성한 파이프라인 화면>

- Jupyter 노트북에서 데이터 처리 및 파이프라인을 구현 한 후 컴파일. Sagemaker 에서 제공하는 이미지를 사용하기 때문에 별도의 이미지 빌드 불가
  - 컴파일된 파이프라인을 Kubeflow 상에 업로드
  - 파이프라인을 실행. 이 작업은 Kubernetes(EKSs)환경에서 진행. 학습 과정에서 나오는 모델 및 산출물들은 S3 의 지정된 경로에 저장
  - S3에 저장된 모델을 sagemaker 를 사용하여 서빙
  - 생성된 엔드포인트에 테스트 데이터를 전송하고 결과값 확인
- 향후 Data Lake에 모아진 스마트온실 데이터간의 다양한 상관 관계 분석과 AI 분석 모델 수립을 위한 분석 환경은 아래와 같이 설계 진행



<분석환경 설계도>

- 스마트온실 모니터링 및 분석 Tool Box 정책 기획 및 관련 기술 연구
  - 오픈소스 기반 BI Tool 기술 검토: BIRT, Apache Zeppelin, MetaBase, Knime
  - 온실 환경 모니터링: 온실 환경, 온실 제어 데이터 모니터링
    - 모니터링 항목
      - : 온도 (냉/난방 제어 사항), 습도, Co2, 배양 - 측정 위치,
    - 온실 외부 환경 (광량, 온도, 습도 등 수집 가능한 데이터)
      - : 모니터링 주기: 최소 10분 ~ 최대 60분
    - 데이터 현황/이력/상관관계 비교/조회/검색을 통한 시각화
      - : 현황 조회 - 각 센싱 위치별 온실 환경 과 제어 내역 조회
      - : 이력 조회 - 기간을 설정 기능 제공, 기간별 온실 환경 및 제어 내역 비교/조회
      - : 상관 관계 조회 - 기간별, 시계열, 온실 환경+제어 내역+작물의 성장 정보 등의
      - : 복합적 상관관계 표기 (수치 데이터와 영상/이미지 데이터 혼합 조회)
  - 생육 모니터링: 작물의 생육 상태에 대한 데이터 모니터링
    - 모니터링 항목 : 엽장, 두께, 키 (기타 작물별 생육 측정 항목-데이터스퀘어 정의 항목)
    - 모니터링 주기: 실증 작물별 조절이 가능하도록 Admin 기능 제공
    - 생육 데이터 현황 조회 기능 제공: 이미지/영상 데이터 조회, 조회 조건: 작물+기간별
- 초기 데이터 시각화는 BI툴을 활용하여 시각화를 진행하고, 2차년도에 뷰잉할 데이터 확정 후 AWS의 EC2기반으로 별도 서버를 구성하여 별도 개발 추진 계획

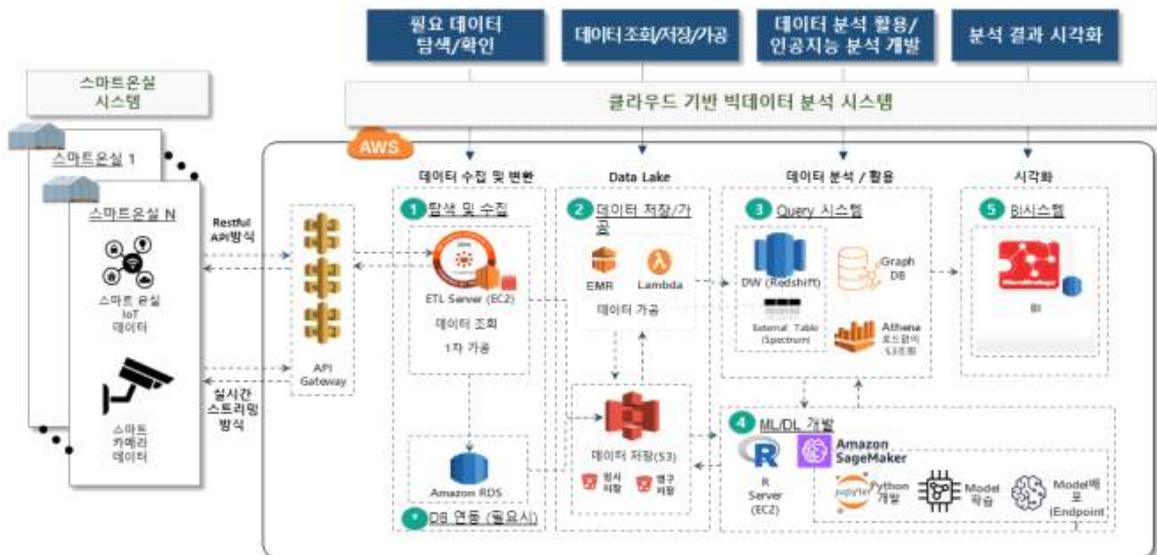


<1차년도 구성된 스마트온실 시각화 Toolbox 아키텍처>

□ 클라우드 기반 스마트 온실 빅데이터 시스템 설계 및 인프라 구축

○ 클라우드 기반 스마트 온실 빅데이터 시스템 설계

- 온실의 센서와 제어기에서 생성되는 데이터와 연동 인터페이스를 표준화한 내역을 기반으로 온실과의 연동 인터페이스를 설계하고 데이터 저장소 구조 설계
- 본 빅데이터 시스템 아키텍처는 크게 3개의 영역으로 구분하여 설계 및 기술 연구 추진
  - 스마트온실의 다양한 데이터를 수집 및 서빙을 위한 API 영역
  - 수집된 데이터를 가공/저장/분석을 위한 Data Lake 영역
  - 수집 가공된 데이터를 사용자가 쉽게 분석하고, 조회할 수 있는 시각화(BI) 영역



<스마트온실 클라우드 기반 빅데이터 아키텍처>

• 스마트온실 데이터 수집/서빙 API 설계

: 정형데이터, 텍스트 데이터, 이미지 데이터, 동영상 데이터에 대한 저장/통신 프로토콜 설계

- API 통신 방식: AWS API Gateway 적용, HTTP방식 (REST API), 파일 전송(sftp) 제공
- 정형 데이터, 텍스트 데이터, 이미지 데이터 수집

- : HTTP 기반의 REST API 방식과 AWS S3 직접 전송 할 수 있는 sftp방식 제공
- 동영상 데이터 수집
  - : 스트리밍 서버 기반의 실시간 영상 수집 (AWS CloudFront + AWS Kinesis Video Streams)을 사용하여, AI 분석 모델의 실시간 학습 및 처리가 가능하도록 데이터를 제공
  - : 기 저장된 영상의 경우는 Workflow 엔진인 Airflow를 사용하여 배치 처리
- 부가적인 데이터 수집 방법 : 각 데이터별 topic 정의 후 Kafka에 저장하는 통신 API 제공
- 프로토콜 설계 : 인증 (ID기반 인증-역활과 권한 설계), 데이터 전송 포맷 설계 (Json, 파일 전송)

• Data Lake 아키텍처 및 ETL 설계

- 모든 데이터는 Data Lake에 저장하고, 메타데이터 추출(카탈로깅), 모니터링 항목 및 상관 분석 항목에 맞도록 데이터 가공하도록 아키텍처 설계를 연구함
- Data Lake에는 원천 데이터, 가공 데이터, 최종 데이터를 모두 보관하게 하고, 이를 기준으로 데이터 분석에 직접 사용

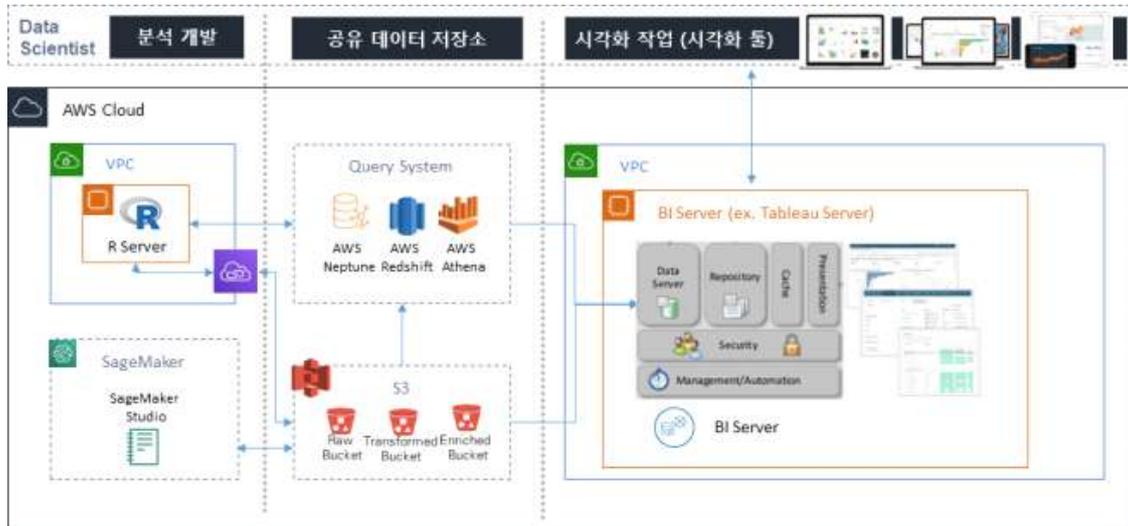


- Data Lake에 저장되는 파일 포맷은 CSV, Excel, Parquet, ZIP 형태 지원
- ETL 및 데이터 가공 설계
  - : 온실 환경 및 생장 데이터 등의 정형 데이터는 표준화 후 raw 데이터 형태로 처리
  - : 데이터 가공: 시계열 기반 및 연관 이미지/동영상 정보 융합/가공(시간별, 생육 주기별 등)
- 실시간 모니터링에 사용할 시스템: AWS Redshift(선정), mariaDB, MongoDB 검증

• 스마트온실 모니터링 및 분석 Tool Box 정책 기획 및 관련 기술 연구

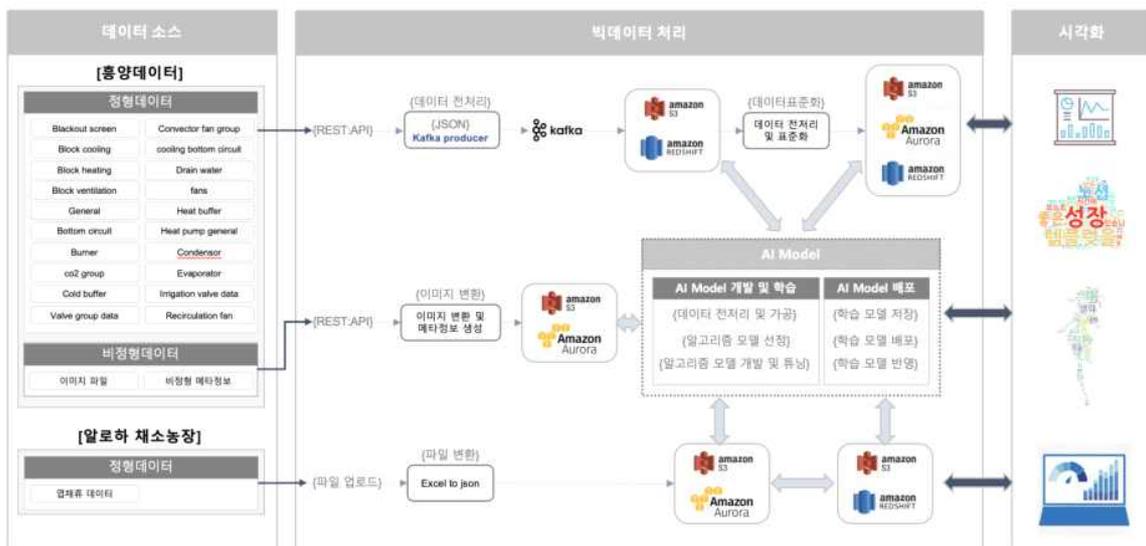
- 오픈소스 기반 BI Tool 기술 검토: BIRT, Apache Zeppelin, MetaBase, Knime
- 온실 환경 모니터링: 온실 환경, 온실 제어 데이터 모니터링
  - 모니터링 항목
    - : 온도 (냉/난방 제어 사항), 습도, Co2, 배양 - 측정 위치, 온실 외부 환경 (광량, 온도, 습도 등 수집 가능한 데이터)
    - : 모니터링 주기: 최소 10분 ~ 최대 60분
  - 데이터 현황/이력/상관관계 비교/조회/검색을 통한 시각화
    - : 현황 조회 - 각 센싱 위치별 온실 환경 과 제어 내역 조회
    - : 이력 조회 - 기간을 설정 기능 제공, 기간별 온실 환경 및 제어 내역 비교/조회
    - : 상관 관계 조회 - 기간별, 시계열, 온실 환경+제어 내역+작물의 성장 정보 등의
    - : 복합적 상관관계 표기 (수치 데이터와 영상/이미지 데이터 혼합 조회)
- 생육 모니터링: 작물의 생육 상태에 대한 데이터 모니터링
  - 모니터링 항목 : 엽장, 두께, 키 (기타 작물별 생육 측정 항목-데이터스퀘어 정의 항목)
  - 모니터링 주기: 실증 작물별 조절이 가능하도록 Admin 기능 제공
  - 생육 데이터 현황 조회 기능 제공: 이미지/영상 데이터 조회, 조회 조건: 작물+기간별

- 초기 데이터 시각화는 BI툴을 활용하여 시각화를 진행하고, 2차년도에 뷰잉할 데이터 확정 후 AWS의 EC2기반으로 별도 서버를 구성하여 별도 개발 추진 계획



<1차년도 구성한 스마트온실 시각화 Toolbox 아키텍처>

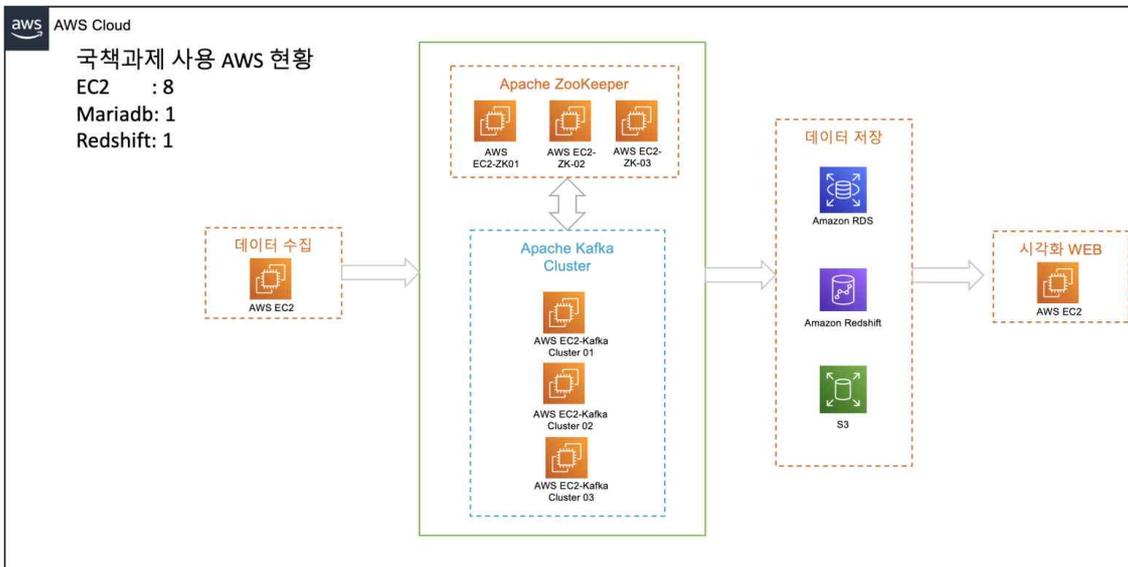
- 빅데이터 플랫폼의 구성
  - 흥양영농조합과 알로하 채소농장으로부터 데이터를 수집
  - 데이터 수집은 REST API 등 기술을 기반으로 설계하였으며, 데이터 관리를 위하여 전처리와 표준화를 수행
  - 데이터의 종류에 따라 AWS S3, Redshift, MySQL에 자동으로 분기되어 적재되고, 시각화를 위한 집계 테이블은 별도로 관리
  - Multi Cloud 지원 서비스와 Edge Cloud 기술을 활용하여 워크로드의 분산과 Cloud 서비스 장애에 대해 유연하게 대처할 수 있는 아키텍처를 설계 및 구현



<빅데이터 플랫폼 구성도>

- AWS Cloud 구조
  - 빅데이터 플랫폼을 위해서 AWS S3, EC2, MariaDB, Redshift 서비스를 활용

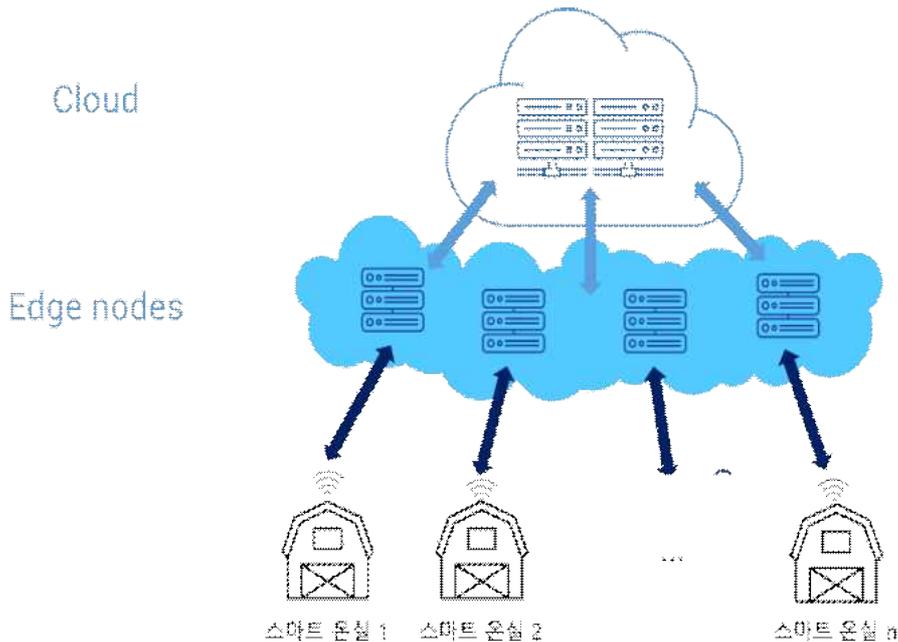
- EC2는 8종, MariaDB, Redshift는 각 1종으로 설정



<AWS Cloud 아키텍처>

□ 멀티 클라우드 기반 빅데이터 플랫폼 설계 및 구축

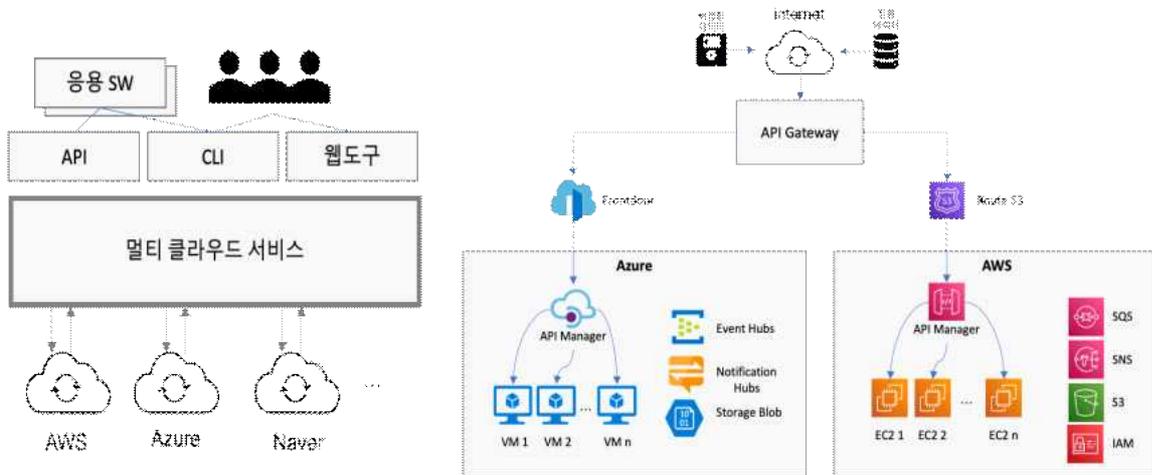
- 멀티 클라우드 환경은 특정 애플리케이션 또는 서비스를 제공하는 클라우드 플랫폼을 최소 두 개 이상 사용하는 환경을 의미
- 여러 퍼블릭 클라우드 제공 업체를 기반으로 실행되는 서비스와 애플리케이션을 온프레미스 운영과 결합하여, 목표에 따라 퍼블릭, 프라이빗, 엣지 클라우드로 구성 가능



<Multi Cloud/Edge Cloud 개념도>

- 에지 클라우드 기술을 활용하여 다수의 스마트 온실에서 수집한 데이터들을 관리
- 중앙 집중형 클라우드에는 대용량/고성능 처리가 필요한 워크로드에 집중하고, 초저지연/초연결이 필요한 워크로드 수행을 에지 노드에서 수행하여 상호 보완적 모델 구현

- 이질적인 운영 방식의 클라우드 인프라들을 동일한 방식, 동일한 인터페이스로 활용 가능하도록 연동



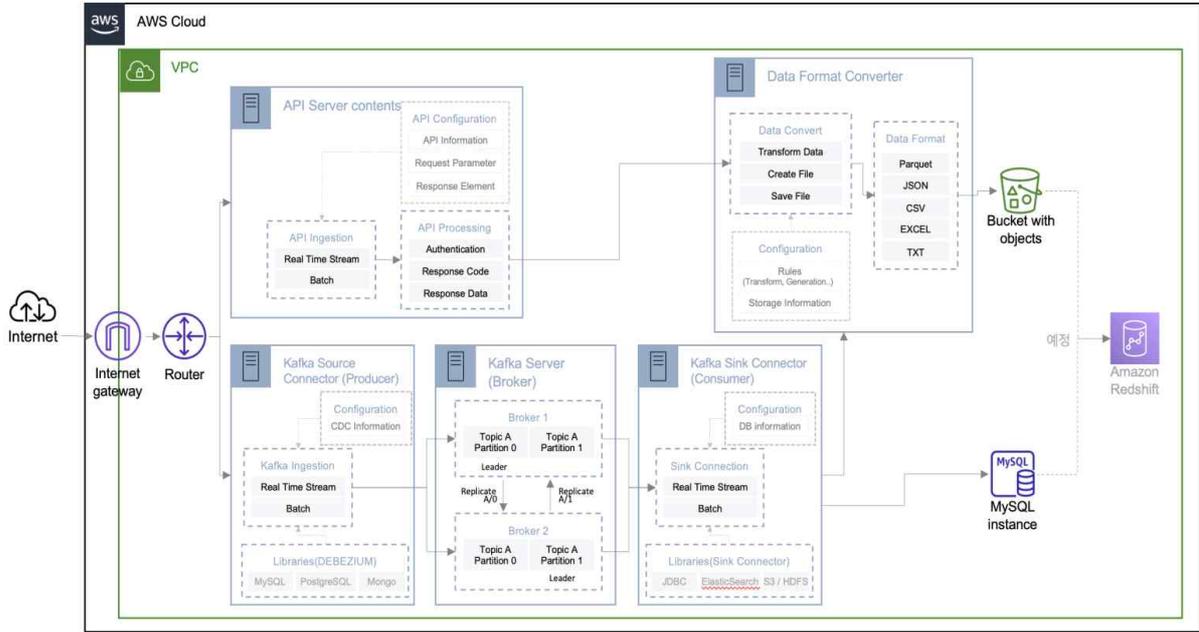
- 인프라 연동을 위해서는 1) 공통방법으로 연동 및 제어를 위한 단일 API, 2) 클라우드 연동 드라이버, 3) 자원 공통 라이프 사이클 관리 기능 필요
- 다수의 클라우드 상에서 구동 되는 자원들 (예: 가상머신 등)을 통합 배치, 운영, 관리하는 기능을 제공해야 하며 자동 설치, 라이프 사이클 관리, 로드밸런싱 등의 제공이 필요
- 또한 다양한 애플리케이션의 상태 및 성능 정보를 수집, 저장, 관리 기능을 통해 서로 다른 모니터링 매트릭을 일관된 방식의 모니터링 기능을 제공
- 내부 서비스들의 복잡한 관계성을 사용자로부터 격리시키고, 사용 편의성을 극대화함으로써 비전문가들도 쉽게 시스템을 사용할 수 있는 환경 제공
- 제공되는 기능으로는 1) 개방형 API: 업계 표준인 REST API 응용 프로그래밍용 인터페이스, 2) CLI: 터미널 환경의 편리한 시스템 활용을 위한 명령어 인터페이스, 웹환경의 편리한 시스템 활용을 위한 GUI 인터페이스 등 필요

### [1단계(2022)]

#### □ 클라우드 기반 스마트 온실 데이터 수집/적재/분석 개발

##### ○ 수집기능

- 농가에 설치된 스마트팜 장비에서 제공하는 REST API를 활용하여 준 실시간 데이터 수집
- Apache Kafka Streaming 서비스 기반 데이터 수집 자동화



<스마트팜 빅데이터 플랫폼 파이프라인 구조>

- 과채류(토마토, 파프리카, 고추)에 대해 매 20분마다 온실 환경 데이터 수집
- 수집 종류 및 API 종류(API 27종, 데이터 약 500여종)
- API 전달 프로토콜 : http 표준 웹 프로토콜
- 전달 주소 : [http://112.187.38.94:61480/syncustom.json?name="API\\_Name"](http://112.187.38.94:61480/syncustom.json?name=)
- 사용예 : [http://112.187.38.94:61480/syncustom.json?name=\\_ws\\_scnu\\_blackuot\\_screen](http://112.187.38.94:61480/syncustom.json?name=_ws_scnu_blackuot_screen)

<농가의 스마트팜 기기에서 제공하는 REST API 목록>

No.	API 이름	No.	API 이름
1	_ws_scnu_blackout_screen	16	_ws_scnu_Sensor_unit
2	_ws_scnu_Block_cooling	17	_ws_scnu_Shading_screen
3	_ws_scnu_Block_heating	18	_ws_scnu_Soil_heating
4	_ws_scnu_Block_ventilation	19	_ws_scnu_Thermal_screen
5	_ws_scnu_Bottom_circuit	20	_ws_scnu_Top_circuit
6	_ws_scnu_Burner	21	_ws_scnu_Valve_group_data
7	_ws_scnu_co2_group	22	_ws_scnu_Ventilation_leeward_side
8	_ws_scnu_Cold_buffer	23	_ws_scnu_Ventilation_side1
9	_ws_scnu_Convector_fan_group	24	_ws_scnu_Ventilation_side2
10	_ws_scnu_cooling_bottom_circuit	25	_ws_scnu_Ventilation_windward_side
11	_ws_scnu_General	26	_ws_scnu_Water_tanks
12	_ws_scnu_Heat_buffer	27	_ws_scnu>Weater_stations
13	_ws_scnu_Heat_pump_general		
14	_ws_scnu_Irrigation_valve_data		
15	_ws_scnu_Recirculation_fan_group		

- REST API 전달 데이터 명세

API명	_ws_scnu_blackout_screen
번호	항목명
1	blackout_screen
2	control_based_on_block
3	air_temperature
4	rh
5	hd
6	calculated_outside_temperature_for_apply_gap
7	calculated_rh
8	calculated_hd
9	calculated_screen_position
10	screen_position
11	dark_duration
12	gap_temperature
13	temperature_influence
14	humidity_influence
15	remaining_delay_time_form_last_stage_open_screen
16	override_on_screen
17	screen_status

API명	_ws_scnu_Block_cooling
번호	항목명
1	block_cooling
2	air_temperature
3	rh
4	alarm_limit_for_rh_too_low
5	hd
6	alarm_limit_for_hd_too_high
7	co2_concentration
8	calculated_rh_cooling
9	calculated_hd_cooling
10	pre_set_cooling_setpoint
11	radiation_intensity_influence_on_cooling_setpoint
12	radiation_sum
13	radiation_sum_influence_on_cooling_setpoint
14	humidity_influence_on_cooling_setpoint
15	corrected_cooling_setpoint
16	remaining_influence_on_cooling_setpoint
17	calculated_cooling_setppoint
18	alarm_limit_for_air_temperature_too_high
19	override_on_cooling
20	cooling_status

API명	_ws_scnu_Block_heating
번호	항목명
1	block_heating
2	air_temperature
3	rh
4	alarm_limit_for_rh_too_low
5	hd
6	alarm_limit_for_hd_too_high
7	co2_concentration
8	calculated_rh_ch
9	calculated_hd_ch
10	pre_set_cooling_setpoint
11	outside_temperature_influence_on_heating_setpoint
12	wind_speed_influence_on_heating_setpoint
13	radiation_intensity_influence_on_heating_setpoint
14	radiation_sum
15	radiation_sum_influence_on_heating_setpoint
16	humidity_influence_on_heating_setpoint
17	tempeature_sum_influence_on_heating_setpoint
18	corrected_heating_setpoint
19	remaining_influence_on_heating_setpoint
20	calculated_heating_setpoint_ch
21	alarm_limit_for_air_temperature_too_low

API명	_ws_scnu_Block_ventilation
번호	항목명
1	block_ventilation
2	air_temperature
3	rh
4	alarm_limit_for_rh_too_low
5	hd
6	alarm_limit_for_hd_too_high
7	co2_concentration
8	calculated_rh_ventilation
9	calculated_hd_ventilation
10	pre_set_ventilation_setpoint
11	outside_temperature_influence_on_ventilation_setpoint
12	wind_speed_influence_on_ventilation_setpoint
13	radiation_influence_on_ventilation_setpoint
14	radiation_sum
15	radiation_sum_influence_on_ventilation_setpoint
16	humidity_influence_on_ventilation_setpoint
17	cooling_influence_on_ventilation_setpoint
18	corrected_ventilation_setpoint
19	remaining_influence_on_ventilation_setpoint
20	calculated_ventilation_setpoint
21	alarm_limit_for_air_temperature_too_high
22	thermal_inversion_status

API명	_ws_scn_u_Bottom_circuit
번호	항목명
1	bottom_circuit
2	control_based_on_block
3	air_temperature
4	rh
5	hd
6	calculated_cooling_setpoint
7	calculated_rh
8	calculated_hd
9	calculated_pipe_temperature
10	alarm_limit_for_above_calculated_pipe_temperature
11	pipe_temperature
12	alarm_limit_for_pipe_temperature_too_low
13	alarm_limit_for_pipe_temperature_too_high
14	minimum_pipe_temperature
15	maximum_pipe_temperature
16	radiation_intensity_influence_on_minimum_pipe_temperature
17	humidity_influence_on_minimum_pipe_temperature
18	humidity_influence_on_maximum_pipe_temperature
19	calculated_circulation_pump_status
20	circulation_pump_status
21	override_on_minimum_pipe_temperature
22	override_on_maximum_pipe_temperature
23	override_on_circuit

API명	_ws_scn_u_co2_group
번호	항목명
1	co2_group
2	pre_set_co2_concentration
3	radiation_intensity_influence_on_co2_concentration_trajectory_1
4	radiation_intensity_influence_on_co2_concentration_trajectory_2
5	vent_position_influence_on_co2_concentration
6	wind_speed_influence_on_co2_concentration
7	humidity_influence_on_co2_concentration
8	minimum_co2_concentration
9	alarm_limit_for_co2_concentration_too_low
10	maximum_co2_concentration
11	alarm_limit_for_co2_concentration_too_high
12	calculated_co2_concentration
13	co2_concentration
14	override_on_co2_group
15	co2_group_status

API명	_ws_scnu_Burner
번호	항목명
1	burner
2	boiler_type
3	calculated_boiler_output
4	boiler_output
5	calculated_burner_position
6	burner_position
7	calculated_burner_status
8	boiler_status

API명	_ws_scnu_Cold_buffer
번호	항목명
1	cold_buffer
2	segment_temperature_01
3	segment_temperature_02
4	segment_temperature_03
5	segment_temperature_04
6	segment_temperature_05
7	segment_temperature_06
8	segment_temperature_07
9	segment_temperature_08
10	segment_temperature_09
11	segment_temperature_10
12	segment_temperature_11
13	calculated_average_segment_temperature_at_which_buffer_0
14	pre_set_average_segment_temperature_buffer_100
15	fill_level
16	buffer_status
17	buffer_status

API명	_ws_scnu_Convector_fan_group
번호	항목명
1	convector_fan_group
2	control_based_on_block
3	type_of_fan_group
4	air_temperature
5	rh
6	hd
7	calculated_rh
8	calculated_hd
9	activated_based_on
10	override_on_fan_group
11	fan_group_status
12	status_of_fan_group_communication
13	fill_level
14	buffer_status
15	buffer_status

API명	_ws_scnu_cooling_bottom_circuit
번호	항목명
1	Bottom circuit
2	Control based on block
3	air_temperature
4	rh
5	hd
6	calculated_cooling_setpoint
7	calculated_rh
8	calculated_hd
9	calculated_pipe_temperature
10	Alarm limit for above calculated pipe temperature
11	Pipe temperature
12	Alarm limit for pipe temperature too low
13	Alarm limit for pipe temperature too high
14	maximum_pipe_temperature
15	radiation_intensity_influence_on_minimum_pipe_temperature
16	humidity_influence_on_minimum_pipe_temperature
17	humidity_influence_on_maximum_pipe_temperature
18	calculated_circulation_pump_status
19	circulation_pump_status

API명	_ws_scnu_irrigation_valve_data
번호	항목명
1	irrigation_valve_data
2	date_of_irrigation_cycle
3	start_time
4	activated_based_on
5	irrigation_phase
6	calculated_ec
7	measured_ec
8	lowest_ec
9	highest_ec
10	average_ec
11	calculated_ph
12	measured_ph
13	lowest_ph
14	highest_ph
15	average_ph
16	measured_supply_ec
17	irrigation_duration
18	irrigation
19	irrigation_volume
20	irrigation_dripper
21	cumulative_irrigation
22	cumulative_irrigation_dripper
23	link_irrigation_valve_to_irrigation_group

API명	_ws_scnu_General
번호	항목명
1	general
2	boiler_type
3	calculated_boiler_temperature
4	boiler_temperature
5	alarm_limit_for_boiler_temperature_too_low
6	alarm_limit_for_boiler_temperature_too_high
7	calculated_boiler_output
8	boiler_output
9	calculated_burner_position
10	burner_position
11	calculated_isolation_valve_position
12	isolation_valve_position
13	calculated_burner_status
14	calculated_boiler_status
15	boiler_status

API명	_ws_scnu_Shading_screen
번호	항목명
1	shading_screen
2	control_based_on_block
3	control_based_on_block_radiation_sensor
4	air_temperature
5	rh
6	hd
7	radiation_intensity
8	calculated_rh
9	calculated_hd
10	calculated_screen_position
11	screen_position
12	radiation_intensity_for_close_to_screen_position_1
13	radiation_intensity_for_open_screen_position_1
14	radiation_intensity_for_close_to_screen_position_2
15	radiation_intensity_for_open_screen_position_2
16	radiation_intensity_for_close_to_screen_position_3
17	radiation_intensity_for_open_screen_position_3
18	radiation_intensity_for_close_to_screen_position_4
19	radiation_intensity_for_open_screen_position_4
20	gap_temperature
21	screen_position_active
22	temperature_influence
23	humidity_influence
24	override_on_screen
25	screen_status

API명	_ws_scnu_Heat_buffer
번호	항목명
1	heat_buffer
2	supply_temperature
3	return_temperature
4	segment_temperature_01
5	segment_temperature_02
6	segment_temperature_03
7	segment_temperature_04
8	segment_temperature_05
9	segment_temperature_06
10	segment_temperature_07
11	segment_temperature_08
12	segment_temperature_09
13	segment_temperature_10
14	segment_temperature_11
15	calculated_average_segment_temperature_at_which_buffer_0%
16	pre_set_average_segment_temperature_buffer_100%
17	fill_level
18	buffer_status

API명	_ws_scnu_Heat_pump_general
번호	항목명
1	heat_pump_general
2	calculated_system_mode
3	system_position
4	heat_output_setpoint
5	cold_output_setpoint
6	operating_time
7	output
8	output_at_integrator_position
9	calculated_heat_pump_status
10	heat_pump_status
11	status_of_heat_pump_control_cabinet

API명	_ws_scnu_Recirculation_fan_group
번호	항목명
1	recirculation_fan_group
2	control_based_on_block
3	type_of_fan_group
4	air_temperature
5	rh
6	hd
7	calculated_rh
8	calculated_hd
9	activated_based_on
10	override_on_fan_group
11	fan_group_status
12	status_of_fan_group_communication

API명	_ws_scn_u_Sensor_unit
번호	항목명
1	sensor_unit
2	control_based_on_block
3	air_temperature
4	wet_bulb_temperature
5	dew_point_temperature
6	rh
7	hd
8	vapour_pressure_deficit_of_greenhouse_air
9	absolute_humidity
10	enthalpy
11	co2_concentration
12	status_of_sensor_unit_fan
13	water_container_status
14	sensor_unit_status
15	status_of_sensor_unit_communication

API명	_ws_scn_u_Soil_heating
번호	항목명
1	soil_heating
2	control_based_on_block
3	air_temperature
4	rh
5	hd
6	heat_manifold_temperature_for_decrease_of_heating_setpoint
7	calculated_heating_setpoint_ch
8	calculated_rh
9	calculated_hd
10	calculated_pipe_temperature
11	alarm_limit_for_below_calculated_pipe_temperature
12	pipe_temperature
13	alarm_limit_for_pipe_temperature_too_low
14	alarm_limit_for_pipe_temperature_too_high
15	minimum_pipe_temperature
16	maximum_pipe_temperature
17	radiation_intensity_influence_on_minimum_pipe_temperature
18	humidity_influence_on_minimum_pipe_temperature
19	humidity_influence_on_maximum_pipe_temperature
20	dew_point_temperature_influence_on_minimum_pipe_temperature
21	calculated_circulation_pump_status
22	circulation_pump_status
23	override_on_minimum_pipe_temperature
24	override_on_maximum_pipe_temperature
25	override_on_circuit

API명	_ws_scnu_Thermal_screen
번호	항목명
1	thermal_screen
2	control_based_on_block
3	air_temperature
4	rh
5	hd
6	radiation_intensity
7	calculated_outside_temperature_for_apply_gap
8	wind_speed_influence_on_maximum_outside_temperature_for_close_screen
9	radiation_influence_on_maximum_outside_temperature_for_close_screen
10	calculated_maximum_outside_temperature_for_close_screen
11	calculated_rh
12	calculated_hd
13	calculated_screen_position
14	screen_position
15	radiation_intensity_for_close_screen
16	radiation_intensity_for_open_screen
17	highest_pipe_temperature
18	gap_temperature
19	temperature_influence
20	humidity_influence
21	remaining_delay_time_for_last_stage_open_screen
22	override_on_screen
23	screen_status

API명	_ws_scnu_Water_tanks
번호	항목명
1	water_tanks
2	control_based_on_supply_group
3	control_based_on_return_group
4	supply_group_status
5	return_group_status
6	water_tank_supply_valve_position
7	alarm_status
8	level
9	alarm_limit_for_minimum_level
10	alarm_limit_for_maximum_level
11	volume
12	alarm_limit_for_open_irrigation_valve

API명	_ws_scn_u_Top_circuit
번호	항목명
1	top_circuit
2	control_based_on_block
3	air_temperature
4	rh
5	hd
6	heat_manifold_temperature_for_decrease_of_heating_setpoint
7	calculated_heating_setpoint_ch
8	calculated_rh
9	calculated_hd
10	calculated_pipe_temperature
11	alarm_limit_for_below_calculated_pipe_temperature
12	pipe_temperature
13	alarm_limit_for_pipe_temperature_too_low
14	alarm_limit_for_pipe_temperature_too_high
15	minimum_pipe_temperature
16	maximum_pipe_temperature
17	radiation_intensity_influence_on_minimum_pipe_temperature
18	humidity_influence_on_minimum_pipe_temperature
19	humidity_influence_on_maximum_pipe_temperature
20	dew_point_temperature_influence_on_minimum_pipe_temperature
21	calculated_circulation_pump_status
22	circulation_pump_status
23	override_on_minimum_pipe_temperature
24	override_on_maximum_pipe_temperature
25	override_on_circuit

API명	_ws_scn_u_Valve_group_data
번호	항목명
1	valve_group_data
2	date_of_irrigation_cycle
3	start_time
4	activated_based_on
5	irrigation_phase
6	calculated_supply_ec
7	measured_supply_ec
8	average_ec
9	average_ph
10	irrigation_duration
11	irrigation
12	irrigation_dripper
13	cumulative_irrigation
14	cumulative_irrigation_dripper
15	drain_water

API명	_ws_scnu_Ventilation_leeward_side
번호	항목명
1	ventilation_leeward_side
2	control_based_on_block
3	air_temperature
4	rh
5	hd
6	absolute_humidity
7	calculated_rh
8	calculated_hd
9	calculated_ventilation_setpoint_leeward_side
10	p_band_at_minimum_temperature_difference_leeward_side
11	temperature_influence_on_p_band_leeward_side
12	outside_temperature_increase_of_p_band_leeward_side
13	wind_speed_influence_on_p_band_leeward_side
14	p_band_leeward_side
15	humidity_influence_on_leeward_side
16	cooling_influence_on_maximum_vent_position_leeward_side
17	wind_speed_influence_on_leeward_side
18	rain_influence_on_maximum_vent_position_leeward_side
19	dry_wind_influence_on_leeward_side
20	temperature_i_influence_on_leeward_side
21	minimum_vent_position_leeward_side
22	maximum_vent_position_leeward_side
23	override_on_leeward_side
24	calculated_vent_position_leeward_side
25	remaining_open_in_stages_leeward_side
26	vent_position_leeward_side

API명	_ws_scnu_Ventilation_side1
번호	항목명
1	ventilation_side_1
2	status_side_1
3	vent_position_side_1
4	status_of_vent_position_control_side_1
5	status_of_vent_position_sensor_side_1
6	calculated_vent_position_side_1
7	pre_set_fixed_vent_position_side_1
8	directly_measured_vent_position_side_1
9	vent_position_side_1_achieved
10	alarm_status_side_1

API명	_ws_scnv_Ventilation_side2
번호	항목명
1	ventilation_side_2
2	status_side_2
3	vent_position_side_2
4	status_of_vent_position_control_side_2
5	status_of_vent_position_sensor_side_2
6	calculated_vent_position_side_2
7	pre_set_fixed_vent_position_side_2
8	directly_measured_vent_position_side_2
9	vent_position_side_2_achieved
10	alarm_status_side_2

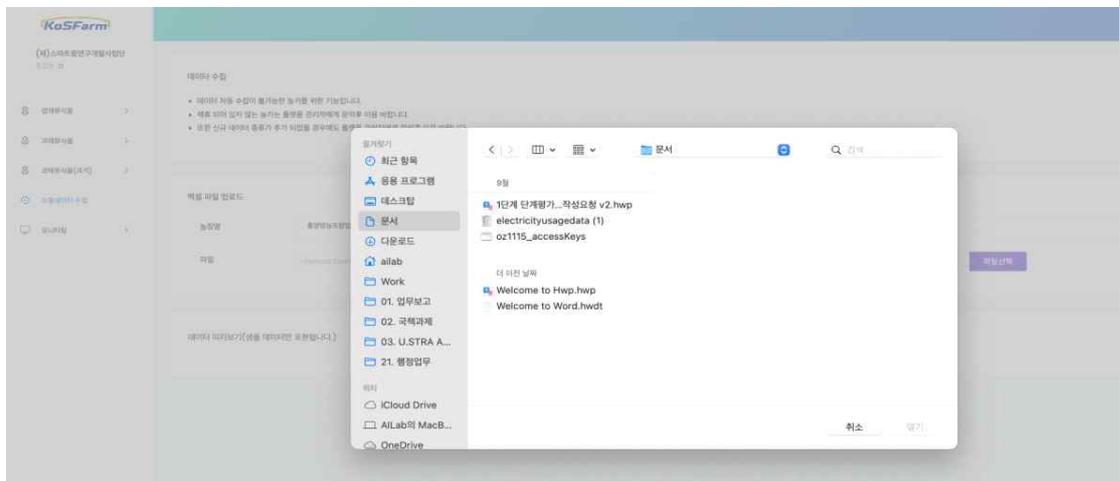
API명	_ws_scnv_Ventilation_windward_side
번호	항목명
1	ventilation_windward_side
2	control_based_on_block
3	air_temperature
4	rh
5	hd
6	absolute_humidity
7	calculated_rh
8	calculated_hd
9	calculated_ventilation_setpoint_windward_side
10	p_band_at_minimum_temperature_difference_windward_side
11	temperature_influence_on_p_band_windward_side
12	outside_temperature_increase_of_p_band_windward_side
13	wind_speed_influence_on_p_band_windward_side
14	p_band_windward_side
15	lag_windward_side
16	humidity_influence_on_windward_side
17	cooling_influence_on_maximum_vent_position_windward_side
18	wind_speed_influence_on_windward_side
19	rain_influence_on_maximum_vent_position_windward_side
20	dry_wind_influence_on_windward_side
21	temperature_i_influence_on_windward_side
22	minimum_vent_position_windward_side
23	maximum_vent_position_windward_side
24	override_on_windward_side
25	calculated_vent_position_windward_side
26	remaining_open_in_stages_windward_side
27	vent_position_windward_side

API명	_ws_scnu_Weater_stations
번호	항목명
1	weather_stations
2	outside_temperature
3	unsmoothed_wind_speed
4	wind_speed
5	wind_direction
6	cardinal_wind_direction
7	unsmoothed_radiation_intensity
8	radiation_intensity
9	standard_radiation_intensity
10	radiation_sum
11	rain
12	rh
13	hd
14	dew_point_temperature
15	absolute_humidity
16	enthalpy
17	barometric_pressure
18	weather_station_status
19	status_of_weather_station_communication

- 실시간으로 수집하지 못하는 데이터를 위한 수동 데이터 업로드 기능 개발
  - 농가에 설치된 스마트팜 장비에서 다운로드 받은 엑셀 데이터를 사용자가 직접 업로드
  - Bootstrap Framework 기술 기반으로 데이터 업로드 기능 개발
  - 수집 종류는 업체류(이자벨, 페얼리, 스텔릭스) 해당



<수동 데이터 업로드 프로세스>



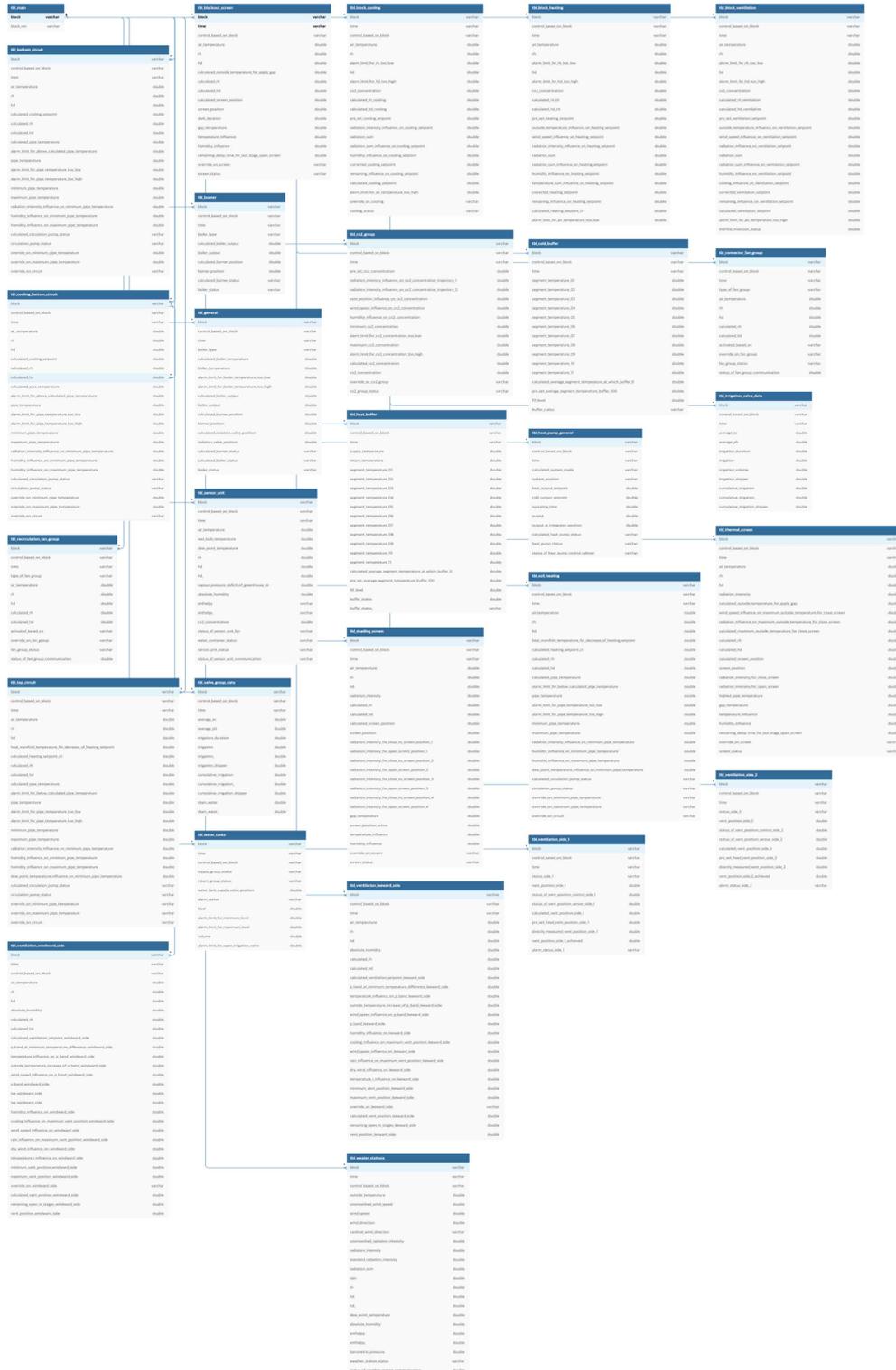
<빅데이터 플랫폼의 수동 데이터 업로드 기능>

- 적재기능

- 정형 데이터(준실시간, 수동데이터 포함) :

- 데이터 관리를 위한 DB 구조 설계

- AWS Cloud의 Redshift 서비스에 정형 데이터 적재 (총 28개 테이블, 600여 종 필드 관리)



<빅데이터 플랫폼 데이터 저장소 ERD>

- 테이블 정의서 작성 : 27개 DB테이블 정의

No.	테이블 이름	No.	테이블 이름
1	tbl_blackout_screen	15	tbl_Recirculation_fan_group
2	tbl_Block_cooling	16	tbl_Sensor_unit
3	tbl_Block_heating	17	tbl_Shading_screen
4	tbl_Block_ventilation	18	tbl_Soil_heating
5	tbl_Bottom_circuit	19	tbl_Thermal_screen
6	tbl_Burner	20	tbl_Top_circuit
7	tbl_co2_group	21	tbl_Valve_group_data
8	tbl_Cold_buffer	22	tbl_Ventilation_leeward_side
9	tbl_Convector_fan_group	23	tbl_Ventilation_side1
10	tbl_cooling_bottom_circuit	24	tbl_Ventilation_side2
11	tbl_General	25	tbl_Ventilation_windward_side
12	tbl_Heat_buffer	26	tbl_Water_tanks
13	tbl_Heat_pump_general	27	tbl_Weater_stations
14	tbl_Irrigation_valve_data		



테이블 정의서

인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발 사업

2022. 09. 26

Copyright © GSITM. All rights reserved.

No.	테이블명	테이블 타입	PK 여부	생성 필드명	데이터 설명	데이터 범위
1	tbl_blackout_screen	tbl_blackout_screen	Y	blackout	블랙아웃 화면	0 ~ 1000 (°C)
2	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
3	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
4	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
5	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
6	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
7	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
8	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
9	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
10	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
11	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
12	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
13	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
14	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
15	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
16	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
17	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
18	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
19	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
20	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
21	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
22	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
23	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
24	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
25	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
26	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
27	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)

No.	테이블명	테이블 타입	PK 여부	생성 필드명	데이터 설명	데이터 범위
1	tbl_blackout_screen	tbl_blackout_screen	Y	blackout	블랙아웃 화면	0 ~ 1000 (°C)
2	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
3	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
4	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
5	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
6	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
7	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
8	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
9	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
10	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
11	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
12	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
13	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
14	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
15	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
16	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
17	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
18	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
19	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
20	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
21	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
22	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
23	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
24	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
25	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
26	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)
27	tbl_blackout_screen	tbl_blackout_screen	N	blackout	블랙아웃 화면	0 ~ 1000 (°C)

- 실제 구축한 정형 데이터 저장소 화면

테이블명	Location	Input format	Output format	Tablespace	Row Count	Estimate	Partiti
tbl_aggregation	s3://smart-farm/SF_v1_0001/Ingestion/data/api-json/00001/aggregation	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_aloh4	s3://smart-farm/manual_upload/00002/00001	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_blackout_screen	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/blackout screen	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_block_cooling	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/block cooling	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_block_heating	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/block heating	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_block_ventilation	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/block ventilation	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_bottom_circuit	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/bottom circuit	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_burner	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/burner	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_co2_group	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/co2 group	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_cold_buffer	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/cold buffer	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_convactor_fan_group	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/convactor fan group	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_cooling_bottom_circuit	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/cooling bottom circuit	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_general	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/general	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_heat_buffer	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/heat buffer	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_heat_pump_general	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/heat pump general	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_irrigation_valve_data	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/irrigation valve data	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_recirculation_fan_group	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/recirculation fan group	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_sensor_unit	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/sensor unit	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_shading_screen	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/shading screen	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_soil_heating	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/soil heating	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_thermal_screen	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/thermal screen	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_top_circuit	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/top circuit	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_valve_group_data	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/valve group data	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_ventilation_leeward_side	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/ventilation leeward side	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_ventilation_side_1	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/ventilation side 1	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_ventilation_side_2	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/ventilation side 2	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_ventilation_windward_side	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/ventilation windward side	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	
tbl_water_tanks	s3://smart-farm/SF_v1_0001/Ingestion/data/json/00001/water tanks	org.apache.hadoop.org.apache.ig	org.apache.ig	org.apache.ig	0	[ ]	

<데이터 관리를 위해 구축 한 Redshift를 Deaver Tool로 접속한 화면>

- BI 집계를 위해 MariaDB 운영 및 집계 테이블에 데이터 적재

테이블명	Engine	Auto Increment	Data Length	Description
api_keys	InnoDB	3	16K	
tb_image_meta	InnoDB	1,709	384K	
tbl_aggregation	InnoDB	637,109	89M	
tbl_aggregation11	InnoDB	1,005	16K	
tbl_aloh4	InnoDB	0	272K	
user	InnoDB	4	16K	

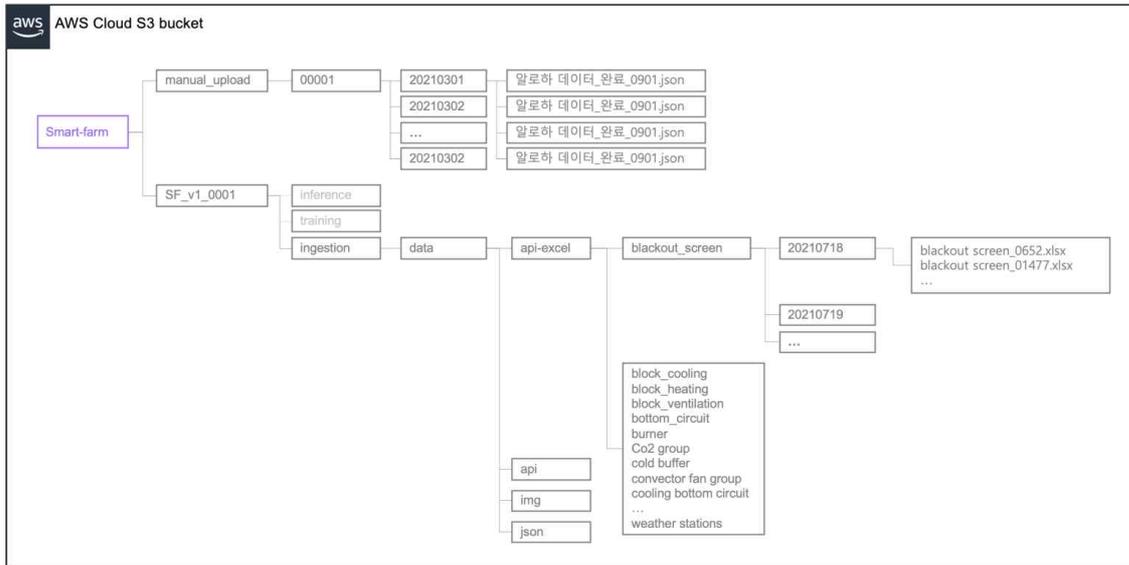
<대시보드를 위한 MariaDB 데이터 구조 및 테이블 리스트>

- 대시보드를 위한 테이블 구조는 53개의 컬럼으로 구성되어 있으며 약 70만개 데이터가 존재

rowid	block	check_time	blackout	air_temperature	rh	hd	screen_status	pipe_temperature	alarm_limit_for_pipe_temperat
1	Block 01	2022-08-22 13:40:02	Block 01 Shade	33.6616	54	15.9396	0	0	
2	Block 02	2022-08-22 13:40:02	Block 02 Shade	35.4855	46	20.8998	0	0	
3	Block 03	2022-08-22 13:40:02	Block 03 Shade	34.5474	44	20.3575	0	0	
4	Block 04	2022-08-22 13:40:02	Block 04 Shade	34.7723	54	16.9211	0	0	
5	Block 05	2022-08-22 13:40:02	Block 05 Shade	38.1948	79	9.58569	0	0	
6	Block 06	2022-08-22 13:40:02	Block 06 Shade	37.8552	82	8.19586	0	0	
7	Block 07	2022-08-22 13:40:02	Block 07 Shade	39.4002	67	16.3735	0	0	
8	Block 08	2022-08-22 13:40:02	Block 08 Shade	39.1732	53	22.5408	0	0	
9	Block 09	2022-08-22 13:40:02	Block 09 Shade	38.19	49	23.193	0	0	
10	Block 10	2022-08-22 13:40:02	Block 10 Shade	39.6061	49	25.3066	0	0	
11	Block 11	2022-08-22 13:40:02	Block 11 Shade	38.5249	25	34.3559	0	0	
12	Block 12	2022-08-22 13:40:02	Block 12 Shade	35.6574	54	17.8271	0	0	
13	Block 01	2022-08-22 14:00:02	Block 01 Shade	33.7695	52	16.6327	0	0	
14	Block 02	2022-08-22 14:00:02	Block 02 Shade	35.3243	45	20.9431	0	0	
15	Block 03	2022-08-22 14:00:02	Block 03 Shade	34.7235	43	21.1428	0	0	
16	Block 04	2022-08-22 14:00:02	Block 04 Shade	34.9579	53	17.683	0	0	
17	Block 05	2022-08-22 14:00:02	Block 05 Shade	39.133	78	10.7877	0	0	
18	Block 06	2022-08-22 14:00:02	Block 06 Shade	38.8649	81	8.97007	0	0	

<대시보드를 위한 집계 데이터>

- 비정형 데이터(json, 이미지, 동영상 데이터)
  - 정형 데이터 외에는 모두 AWS S3에 적재하며, API를 통해 수집한 데이터는 json형식으로 저장하고, 이미지/동영상 데이터는 원본 그대로 저장하며 연관된 메타 데이터는 MariaDB에 저장



<비정형 데이터 저장 구조>

- 이미지/동영상 관련 메타 데이터는 비정형 데이터 관리를 위해 필수적으로 필요
- 메타 데이터는 17개 필드로 관리

컬럼명	#	Data Type	Not Null	Auto Increment	Key	인덱스	Extra	Expression	Comment
created_at	1	datetime	[V]	[ ]					
updated_at	2	datetime	[V]	[ ]					
id	3	int	[V]	[V]	PRI		auto_increment		
camera_id	4	varchar(50)	[V]	[ ]	MUL				
width	5	int	[ ]	[ ]					
height	6	int	[ ]	[ ]					
channel	7	int	[ ]	[ ]					
shooting_date	8	varchar(10)	[ ]	[ ]					
shooting_time	9	varchar(50)	[ ]	[ ]					
save_path	10	varchar(200)	[ ]	[ ]					
camera_installation_heigh	11	int	[ ]	[ ]					
camera_installation_angle	12	varchar(100)	[ ]	[ ]					
camera_installation_locat	13	varchar(100)	[ ]	[ ]					
farm_name	14	varchar(100)	[ ]	[ ]					
farm_area	15	varchar(100)	[ ]	[ ]					
target_crop	16	varchar(50)	[ ]	[ ]					
growth_stage	17	varchar(50)	[ ]	[ ]					

<비정형 데이터 관리를 위한 메타 데이터 구조>

	created_at	updated_at	id	camera_id	width	height	channel	shooting_date	shooting_time	save_path
1	2022-05-31 08:55:30	2022-05-31 08:55:31	3	24345571	608	360	3	20220530	144539	SF_v1_0001/Ingestion/data/im
2	2022-06-01 23:55:43	2022-06-01 23:55:43	4	24345571	608	360	3	20220530	144539	SF_v1_0001/Ingestion/data/im
3	2022-06-02 00:05:35	2022-06-02 00:05:35	5	24345571	608	360	3	20220530	144539	SF_v1_0001/Ingestion/data/im
4	2022-06-02 00:06:02	2022-06-02 00:06:02	6	24345571	608	360	3	20220530	144539	SF_v1_0001/Ingestion/data/im
5	2022-06-02 00:07:14	2022-06-02 00:07:14	7	24345571	608	360	3	20220530	144539	SF_v1_0001/Ingestion/data/im
6	2022-06-02 00:07:34	2022-06-02 00:07:35	8	24345571	608	360	3	20220530	144539	SF_v1_0001/Ingestion/data/im
7	2022-06-02 00:08:58	2022-06-02 00:08:59	9	24345571	608	360	3	20220530	144539	SF_v1_0001/Ingestion/data/im
8	2022-06-02 00:09:00	2022-06-02 00:09:00	10	24345571	608	360	3	20220530	155829	SF_v1_0001/Ingestion/data/im
9	2022-06-02 00:09:01	2022-06-02 00:09:01	11	24345571	608	360	3	20220530	160516	SF_v1_0001/Ingestion/data/im
10	2022-06-02 00:09:03	2022-06-02 00:09:03	12	24345571	608	360	3	20220530	161744	SF_v1_0001/Ingestion/data/im
11	2022-06-02 00:09:04	2022-06-02 00:09:04	13	24345571	608	360	3	20220601	110936	SF_v1_0001/Ingestion/data/im
12	2022-06-02 00:09:05	2022-06-02 00:09:05	14	24345914	608	360	3	20220530	144541	SF_v1_0001/Ingestion/data/im
13	2022-06-02 00:09:07	2022-06-02 00:09:07	15	24346205	608	360	3	20220530	144543	SF_v1_0001/Ingestion/data/im
14	2022-06-02 00:09:08	2022-06-02 00:09:08	16	24346205	608	360	3	20220530	155832	SF_v1_0001/Ingestion/data/im
15	2022-06-02 00:09:09	2022-06-02 00:09:10	17	24346205	608	360	3	20220530	160518	SF_v1_0001/Ingestion/data/im
16	2022-06-02 00:09:11	2022-06-02 00:09:11	18	24346205	608	360	3	20220530	161122	SF_v1_0001/Ingestion/data/im
17	2022-06-02 00:09:12	2022-06-02 00:09:12	19	24346205	608	360	3	20220530	161746	SF_v1_0001/Ingestion/data/im
18	2022-06-02 00:09:14	2022-06-02 00:09:14	20	24346205	608	360	3	20220530	162048	SF_v1_0001/Ingestion/data/im
19	2022-06-02 00:09:15	2022-06-02 00:09:15	21	24346205	608	360	3	20220531	082259	SF_v1_0001/Ingestion/data/im
20	2022-06-02 00:09:17	2022-06-02 00:09:17	22	24346205	608	360	3	20220601	002522	SF_v1_0001/Ingestion/data/im
21	2022-06-02 00:09:18	2022-06-02 00:09:18	23	24346205	608	360	3	20220601	145852	SF_v1_0001/Ingestion/data/im
22	2022-06-02 00:09:19	2022-06-02 00:09:20	24	24346205	608	360	3	20220601	230001	SF_v1_0001/Ingestion/data/im
23	2022-06-02 00:09:21	2022-06-02 00:09:21	25	24345524	608	360	3	20220530	050820	SF_v1_0001/Ingestion/data/im
24	2022-06-02 00:09:22	2022-06-02 00:09:22	26	24345524	608	360	3	20220530	144537	SF_v1_0001/Ingestion/data/im
25	2022-06-02 00:09:23	2022-06-02 00:09:24	27	24345524	608	360	3	20220530	152440	SF_v1_0001/Ingestion/data/im
26	2022-06-02 00:09:25	2022-06-02 00:09:25	28	24345524	608	360	3	20220530	152444	SF_v1_0001/Ingestion/data/im
27	2022-06-02 00:09:26	2022-06-02 00:09:26	29	24345524	608	360	3	20220530	155825	SF_v1_0001/Ingestion/data/im
28	2022-06-02 00:09:28	2022-06-02 00:09:28	30	24345524	608	360	3	20220530	160514	SF_v1_0001/Ingestion/data/im
29	2022-06-02 00:09:29	2022-06-02 00:09:29	31	24345524	608	360	3	20220530	161118	SF_v1_0001/Ingestion/data/im

<MariaDB에서 관리자하고 있는 비정형 데이터 메타 데이터>

- 준 실시간 수집 데이터는 20분 간격으로 json 파일을 생성하고 저장
- 작물별로 블록을 구분하고 일자를 구분하는 구조로 생성하여 데이터 관리 체계 마련

Amazon S3 > 버킷 > smart-farm > SF\_v1\_0001 / Ingestion / data / api-json / 00001 / aggregation / block 01 / 2022-07-15/

2022-07-15/ S3 URI 복사

객체 속성

객체 (46)

객체는 Amazon S3에 저장되어 있는 기본 엔티티입니다. Amazon S3 인벤토리 링크를 사용하여 버킷에 있는 모든 객체의 목록을 얻을 수 있습니다. 다른 사용자가 객체에 액세스할 수 있게 하려면 명시적으로 권한을 부여해야 합니다. 자세히 알아보기

🔄
S3 URI 복사
URL 복사
다운로드
열기
삭제
작업
폴더 만들기
업로드

🔍 점두사로 객체 찾기

<input type="checkbox"/>	이름	유형	마지막 수정	크기	스토리지 클래스
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 08:50:01.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 09:00:02.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 09:20:02.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 09:40:01.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 10:00:02.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 10:20:01.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 10:40:02.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 11:00:02.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 11:20:02.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 11:40:02.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 12:00:02.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.7KB	Standard
<input type="checkbox"/>	aggregation-Block 01-2022-07-15 12:20:03.json	json	2022. 7. 20, pm 3:29:06 PM KST	7.6KB	Standard

<실제 적재된 비정형 데이터>

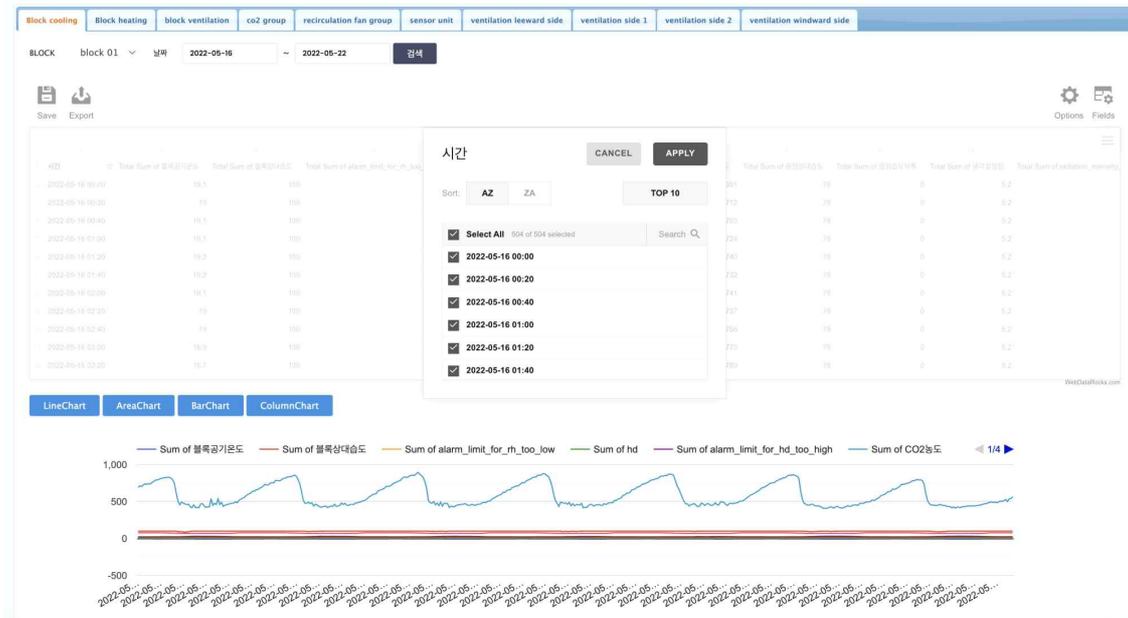
- 수동 업로드 데이터는 업로드 한 일자별로 데이터 관리



<실제 적재된 비정형 데이터(수동 업로드 데이터)>

○ 분석기능

- 데이터 분석을 위한 시각화 툴 구축
- 데이터 수집 현황과 데이터의 트렌드 분석을 위한 데이터마트 및 대시보드 개발
- 작물의 특성에 따라 필터링 조건을 조정 가능하며, 데이터 리스트와 차트 및 Pivot 기능을 제공하여 분석의 편의성 제공



<분석을 위한 툴박스 화면>

□ 스마트온실 데이터 상관관계 분석 파이프라인 개발

○ Python Library를 활용하여 스마트 온실의 각 센서별 데이터 간의 Scatter Plot, Covariance, Correlation Coefficient 분석 개발

• Scatter Plot : 두 변수간의 관계를 표현하는 기법

- 각 데이터 포인트들을 흩 뿌려놓은 형태, 크게 양의 상관관계(한 변수가 증가할 때 나머지 변수도 같이 증가) 또는 음의 상관관계(한 변수가 증가할 때 나머지 변수는 감소) 그리고 무상관(두 변수 간 상관성이 없음)을 표현
- 데이터 포인트 각각을 표시해주기 때문에 해상도 높은 정보를 표현 가능

- 모든 포인트를 표현하는 것이 단점이 될 수 있음
- Covariance : 두 변수 사이에 어떤 상관관계가 있는지 분석
  - Covariance를 이용하면 두 변수의 관계를 객관적으로 수량화(quantify) 가능
  - 두 짝지어진 변수가 동시에 변하는 정도를 나타내는 지수로, 어떤 방향으로 움직이는지 측정
  - 두 변수의 선형적인 관계(lear relationship)을 나타내는 지수
  - 추정치가 양수가 되면 두 변수가 서로 정적 상관을 보이고 있다고 해석하고 음수가 되면 두 변수가 서로 부적 상관을 가지고 있다고 해석
  - 0에 가까운 값을 갖는다면 두 변수는 서로 상관이 없다는 의미
- Correlation Coefficient : 이변량(두 변수) 상관 측정치이며 변수의 단위로부터 자유롭다는 특징
  - 상관계수 중 상관계수의 경우, -1에서 1사이에서 움직이며, 상관계수의 부호는 두 변수의 관계 방향을 표현
  - 상관계수가 양수이면 두 변수는 정적 상관이 있다는 의미이며, 음수이면 부적 상관이 있음을 의미
  - 상관계수의 절대적 크기는 관계의 강도를 의미
- 데이터 분석이 가능하도록 데이터 저장소로부터 데이터를 조회/다운로드 기능 구현
- Apache Airflow 기술 기반 파이프라인 자동화 개발

□ 생육-환경-영농 데이터 사이의 상관관계 분석 개발

- 상관관계 관계 분석은 생육 데이터와 환경 데이터 등 2가지 이상의 데이터들 조합으로 어떠한 연관성이 있는지 분석
- 데이터 분석을 위해 작물의 생육 데이터(이미지)와 메타 데이터를 축적하고 있으며, 스마트 온실의 환경 데이터를 수집하는 과정 진행
- 정교한 데이터 분석을 위해서는 지속적으로 데이터를 축적하고, 데이터의 범위를 확장해야 하며 이를 위한 인프라의 자동 Scale Out 및 모니터링 환경 구축 진행



<빅데이터 플랫폼에서 저장/관리하는 작물 생육 이미지 예시>

	created_at	updated_at	id	camera_id	width	height	channel	shooting_date	shooting_time	save_path
1	2022-10-01 00:01:15	2022-10-01 00:01:15	1,353	24313849	608	360	3	20220930	233918	SF_v1_0001/Ingestion/data/img/

<상기 이미지와 연관된 메타 데이터>

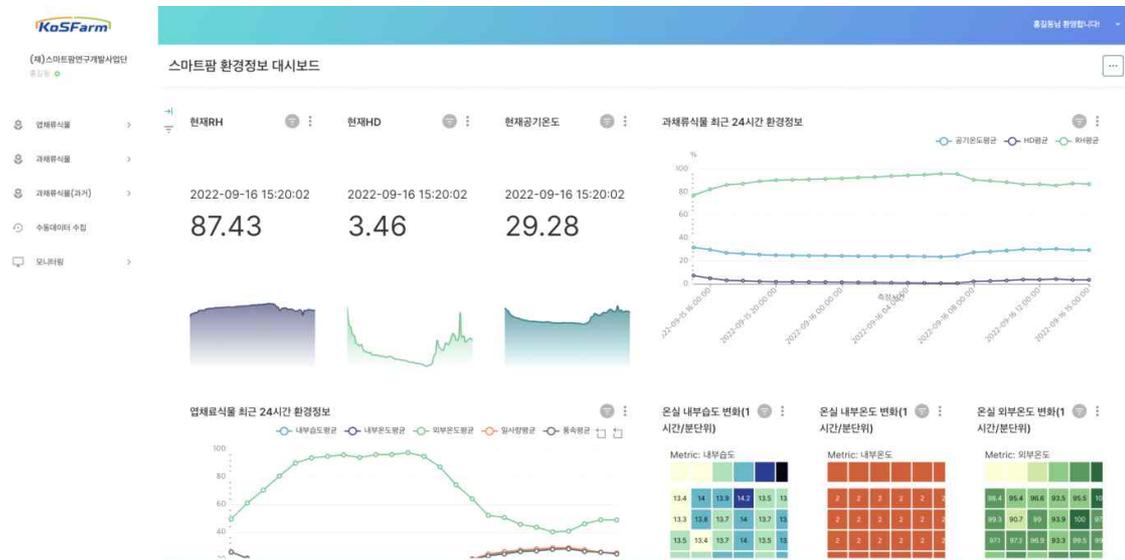
이름	유형	마지막 수정	크기	스토리지 클래스
<input checked="" type="checkbox"/> 20221001_083918.bmp	bmp	2022. 10. 1. am 9:01:16 AM KST	185.5KB	Standard
<input type="checkbox"/> 20221001_093141.bmp	bmp	2022. 10. 1. am 9:52:38 AM KST	224.6KB	Standard
<input type="checkbox"/> 20221001_104000.bmp	bmp	2022. 10. 1. am 11:01:17 AM KST	180.8KB	Standard
<input type="checkbox"/> 20221001_124041.bmp	bmp	2022. 10. 1. pm 1:01:19 PM KST	180.6KB	Standard
<input type="checkbox"/> 20221001_144123.bmp	bmp	2022. 10. 1. pm 3:01:20 PM KST	178.3KB	Standard
<input type="checkbox"/> 20221001_153439.bmp	bmp	2022. 10. 1. pm 3:52:40 PM KST	201.1KB	Standard
<input type="checkbox"/> 20221001_164204.bmp	bmp	2022. 10. 1. pm 5:01:22 PM KST	203.4KB	Standard

<상기 이미지가 실제 저장되고 관리되고 있는 저장소>

□ 스마트 온실 데이터 조회 및 다운로드, 실시간 모니터링, 온실 간 Head-to-Head 비교 기능 개발

○ 데이터 조회/다운로드/데이터 검증(데이터 비교)

- 수집된 데이터를 조회/검증할 수 있는 시각화 플랫폼 구축
- 준 실시간 수집 및 수동 업로드 데이터와 S3에 적재된 비정형 데이터 Viewer 개발
- 기간별 데이터 조회 기능을 포함하고 있으며, 필요한 컬럼들만 선택 가능한 옵션 기능 개발
- 시각화 플랫폼의 화면에 조회된 데이터 리스트 다운로드 기능 개발



<시각화 플랫폼 메인 화면>

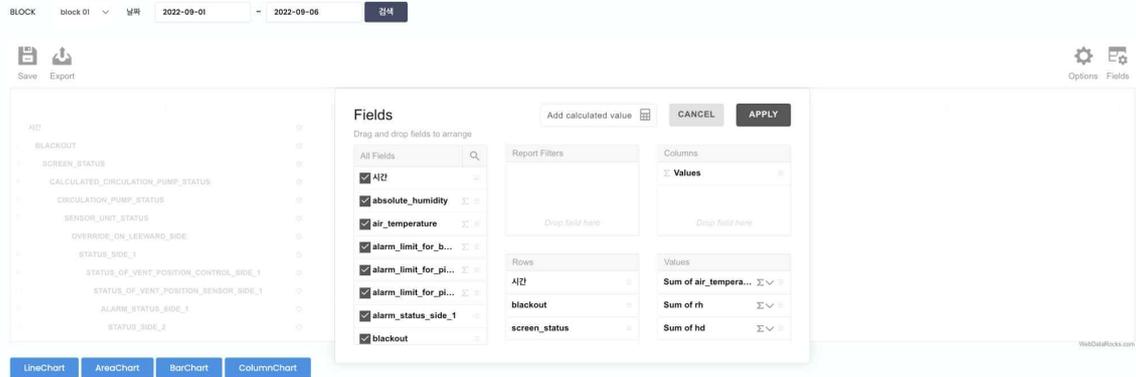
날짜: 2022-05-13 | 2022-05-19 | 검색

Save | Export | Options | Fields

이차별	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 시간	4,456,081	4,932	792,508	172,803.4	49,407.9	48,572.6	50,242.1	46,154.3	0	33,524.5	1,859.4				
2 (blank)	3,000	2	0	94.7	13.2	13.2	13.2	13.6	0	14	0				
4 (blank)	3,000	2	0	96.8	13.4	13.3	13.6	13.6	0	14	0				
5 (blank)	3,000	2	0	88.4	13.5	13	14	13.6	0	14	0.1				
6 (blank)	3,000	2	0	89.8	13.5	12.7	14.3	13.5	0	14	0.3				
7 (blank)	3,000	2	0	99.5	13.2	12.9	13.6	13.5	0	14	0.1				
8 (blank)	3,000	2	0	99.2	13.4	13	13.8	13.4	0	14	0.3				
9 (blank)	3,000	2	0	86.5	13.4	12.9	13.9	13.6	0	14	0.7				
10 (blank)	3,000	2	0	100	13.3	12.6	14	14.1	0	14	0.3				
11 (blank)	3,000	2	0	100	13.2	12.6	13.7	14.1	0	14	0.6				
12 (blank)	3,000	2	0	99.7	13.1	12.5	13.6	13.9	0	14	0				

WotDataPaaS.com

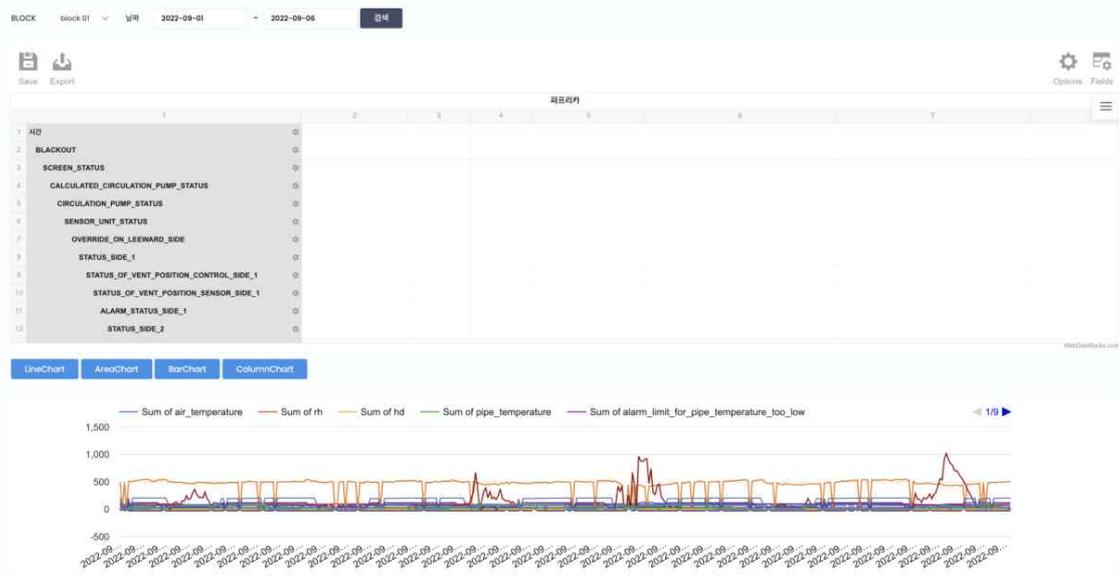
<조회 기간 설정 후 데이터 조회 화면>



### <컬럼 필터링 옵션>

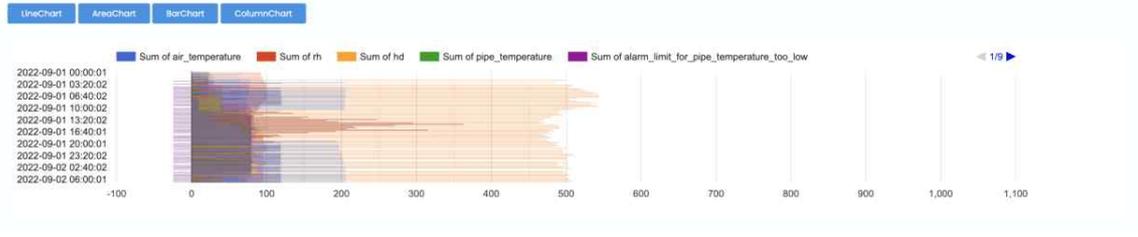


### <Pivot 기능>

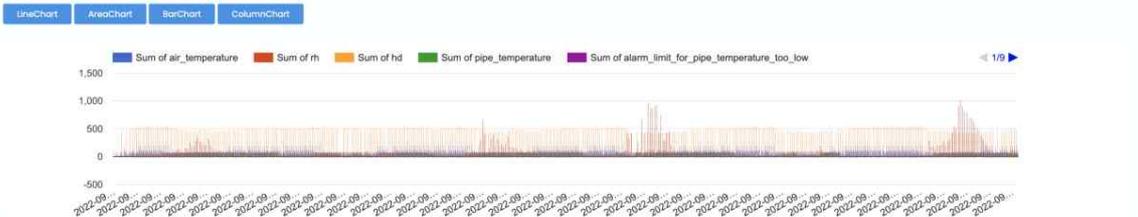


### <데이터 차트 기능>

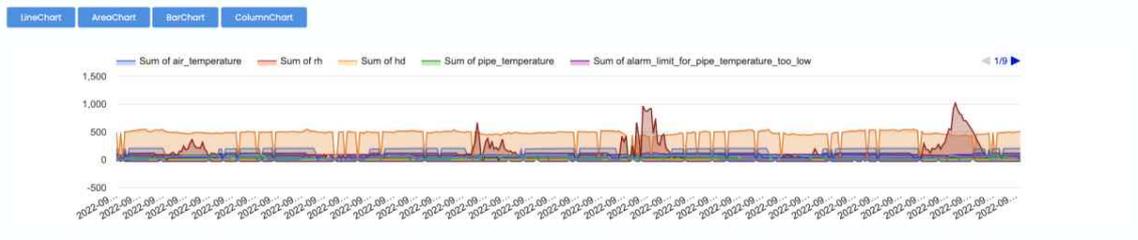
- 조회 데이터는 그리드(테이블) 형식과 차트 형식으로 제공하며, 사용자가 선택 가능하도록 구성



<바차트 기능>



<컬럼 차트 기능>



<영역 차트 기능>

○ 실시간 모니터링

- 데이터의 수집 현황 모니터링
  - - Apache Kafka를 통해 수집되는 데이터의 현황과 Kafka 수집 현황에 대한 모니터링



<데이터 수집을 위한 Kafka 대시보드>



## <데이터 수집을 위한 Kafka Broker 모니터링 현황>

Topic Name	Total Partitions	Out of sync replicas	Replication Factor	Number of messages	Size
IN_...consumer_offsets	50	0	1	609237	35KB
IN_...schemas	1	0	3	139	219KB
ws_scmu_Block_cooling	1	0	3	0	0Bytes
ws_scmu_Block_heating	1	0	3	0	0Bytes
ws_scmu_Block_ventilation	1	0	3	0	0Bytes
ws_scmu_Bottom_circuit	1	0	3	0	0Bytes
ws_scmu_Burner	1	0	3	0	0Bytes
ws_scmu_Cold_buffer	1	0	3	0	0Bytes
ws_scmu_Convector_fan_group	1	0	3	0	0Bytes
ws_scmu_General	1	0	3	0	0Bytes
ws_scmu_Heat_buffer	1	0	3	0	0Bytes
ws_scmu_Heat_pump-general	1	0	3	0	0Bytes
ws_scmu_Irrigation_valve_data	1	0	3	0	0Bytes
ws_scmu_Recirculation_fan_group	1	0	3	0	0Bytes
ws_scmu_Sensor_unit	1	0	3	0	0Bytes
ws_scmu_Shading_screen	1	0	3	0	0Bytes
ws_scmu_Soil_heating	1	0	3	0	0Bytes

## <데이터 수집을 위한 Kafka Topic List 모니터링 현황>

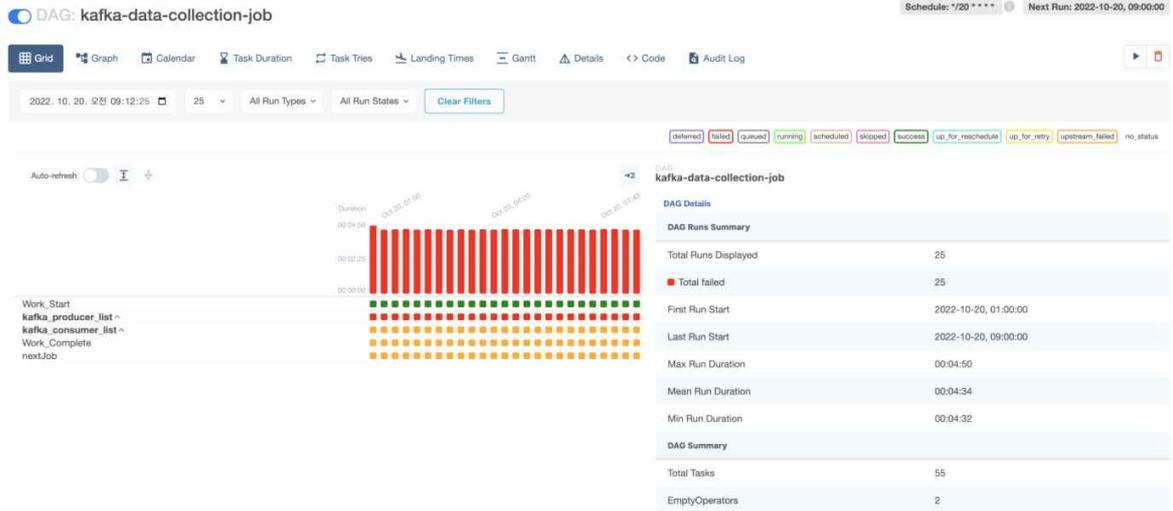
Schema Name	Version	Compatibility
blackout_screen	2	FULL
block_cooling	4	FULL
block_heating	1	FULL
block_ventilation	1	FULL
bottom_circuit	2	FULL
burner	1	FULL
co2_group	2	FULL
cold_buffer	1	FULL
convector_fan_group	1	FULL
general	1	FULL
heat_buffer	1	FULL
heat_pump_general	1	FULL
irrigation_valve_data	1	FULL
recirculation_fan_group	1	FULL
sensor_unit	1	FULL
shading_screen	1	FULL

## <데이터 수집을 위한 Kafka Schema Registry 모니터링 현황>

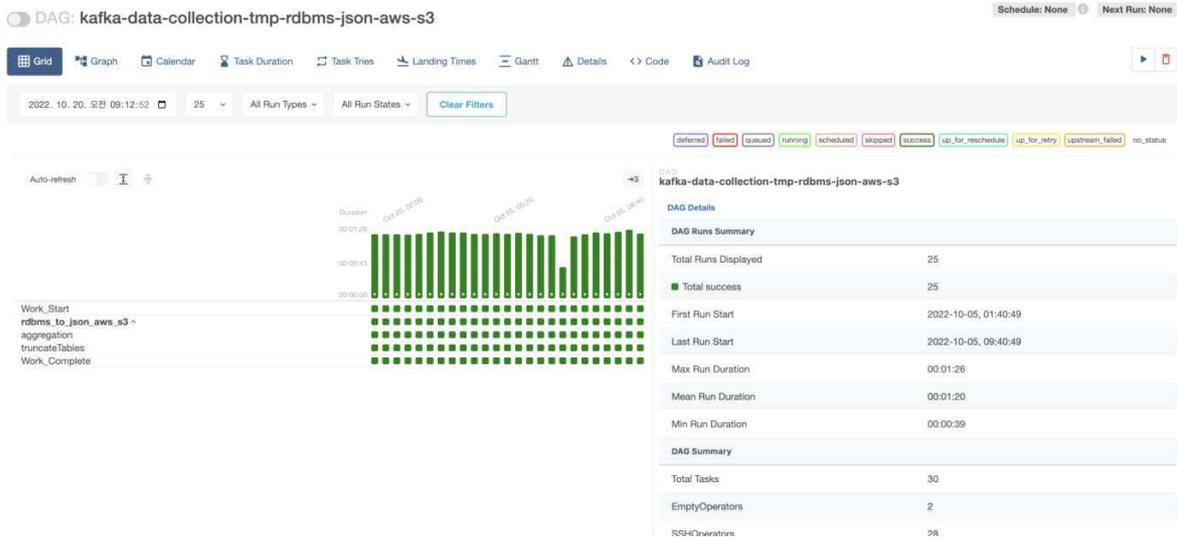
- 수집 자동화 기능 모니터링
  - 준실시간적 데이터를 수집하기 위한 Apache Airflow의 자동화 기능을 모니터링
  - 총 3종의 자동화에 대한 모니터링 수행

DAG	Owner	Runs	Schedule	Last Run	Next Run	Recent Tasks	Actions	Links
kafka-data-collection-job	airflow	1000	*/20 * * * *	2022-10-20, 08:40:00	2022-10-20, 09:00:00	1000	▶ 🗑	...
kafka-data-collection-tmp-rdbms-json-aws-s3	airflow	1	None	2022-10-05, 09:40:47		1	▶ 🗑	...
redShift-append-partition	airflow	1	0 01 * * * *	2022-10-20, 01:00:58	2022-10-20, 01:00:59	1	▶ 🗑	...

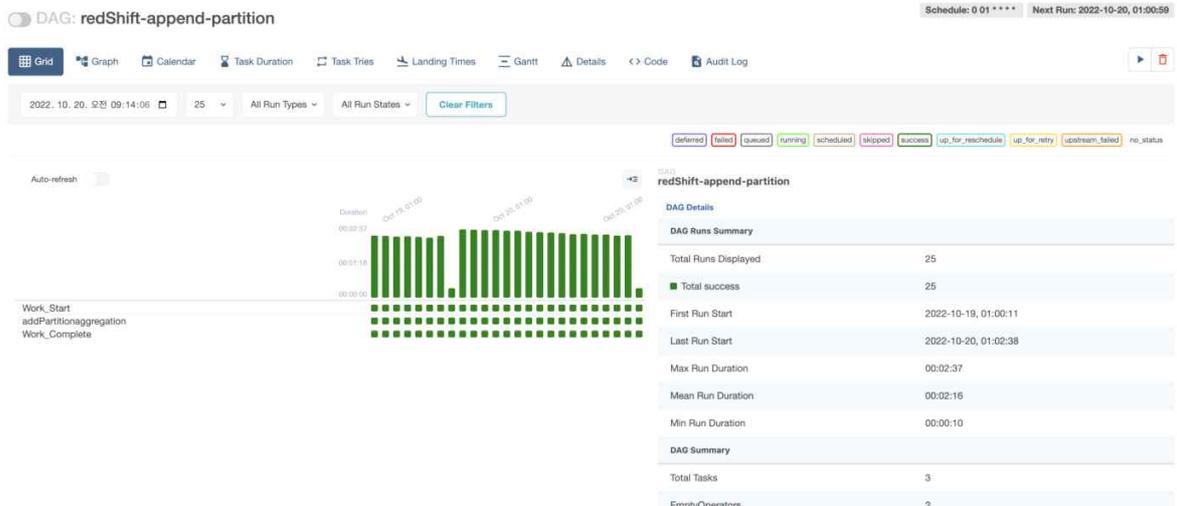
## <자동화를 위한 Airflow 대시보드>



<데이터 수집 자동화 기능 모니터링 현황>



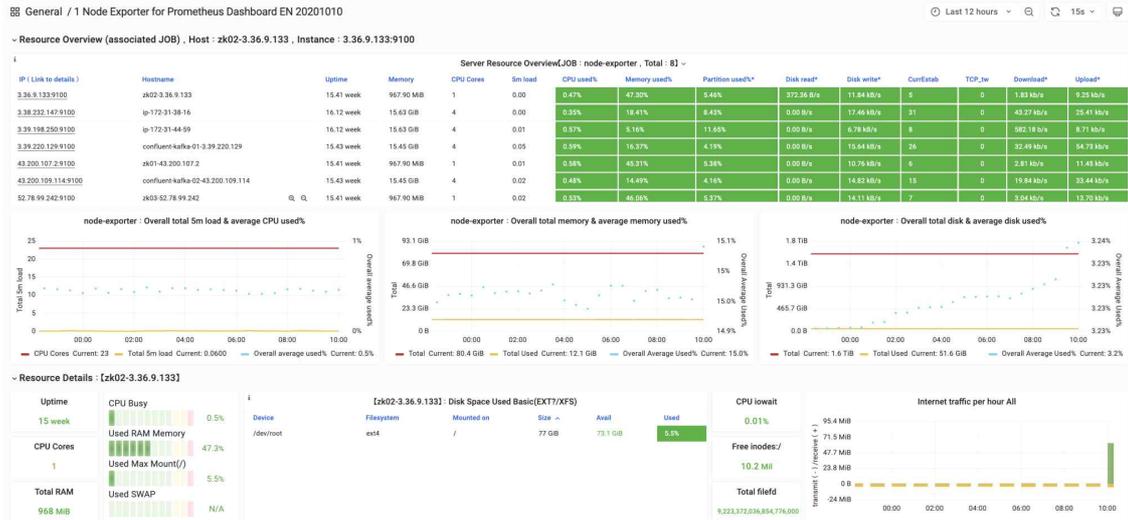
<비정형 데이터 적재 자동화 기능 모니터링 현황>



<데이터 저장소 파티셔닝 자동화 기능 모니터링 현황>

- 시스템 모니터링

- 빅데이터 플랫폼과 시각화 플랫폼을 운영하기 위한 클라우드 자원 모니터링
- 스마트 온실 플랫폼은 AWS S3, EC2, Redshit 등 다양한 자원을 활용
- EC2의 CPU, Memory, Network, Traffic, Disk Space, I/O 등을 실시간적으로 모니터링하고 관리



<AWS 자원 모니터링 현황>

□ 환경-생육-영농 정보를 연결하고 동시에 분석·활용할 수 있는 실시간 시각화 인터페이스 개발

○ 실시간 시각화 인터페이스

- WebDataRocks를 활용하여 실시간 분석 가능한 BI 개발
- 데이터 Pivot 기능을 제공
- Pivo 결과를 활용한 차트 기능을 제공하여 시각화

Block cooling | Block heating | Block ventilation | co2 group | recirculation fan group | sensor unit | ventilation leeward side | ventilation side 1 | ventilation side 2 | ventilation windward side

BLOCK block 01 날짜 2022-05-16 ~ 2022-05-22 검색

Save Export Options Fields

시간	블록공기온도	블록상대습도	ALARM_LIMIT_FOR_RH_TOO_LOW	HD	ALARM_LIMIT_FOR_HD_TOO_HIGH	블록공기온도	블록상대습도	RADIATION_INTENSITY_INFLUENCE_ON_COOLING_SETPOINT	RADIATION...
11:708.1	50.128	0	60.3	0	0	0	0	0	0
2022-05-16 00:00	19.1	100	0	0	0	0	0	0	0
2022-05-16 00:20	19	100	0	0	0	0	0	0	0
2022-05-16 00:40	19.1	100	0	0	0	0	0	0	0
2022-05-16 01:00	19.1	100	0	0	0	0	0	0	0
2022-05-16 01:20	19.2	100	0	0	0	0	0	0	0
2022-05-16 01:40	19.2	100	0	0	0	0	0	0	0
2022-05-16 02:00	19.1	100	0	0	0	0	0	0	0
2022-05-16 02:20	19	100	0	0	0	0	0	0	0
2022-05-16 02:40	19	100	0	0	0	0	0	0	0
2022-05-16 03:00	18.9	100	0	0	0	0	0	0	0

Fields

Select and arrange columns

- Select All
- 시간
- 블록공기온도
- 블록상대습도
- alarm\_limit\_for\_rh\_too\_low
- hd
- alarm\_limit\_for\_hd\_too\_high
- CO2농도
- 용량상대습도

<데이터 분석을 위한 필터링 옵션 기능>

Block cooling | Block heating | Block ventilation | co2 group | recirculation fan group | sensor unit | ventilation leeward side | ventilation side 1 | ventilation side 2 | ventilation windward side

BLOCK block 01 날짜 2022-05-16 ~ 2022-05-22 검색

Save Export Options Fields

1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	시간	블록공기온도	블록상대습도	ALARM_LIMIT_FOR_RH_TOO_LOW	HD	ALARM_LIMIT_FOR_HD_TOO_HIGH							
2		11 708.1	50.128	0	60.3	0							
3	2022-05-16 00:00	19.1	100	0	0	0							
4	2022-05-16 00:20	19	100	0	0	0							
5	2022-05-16 00:40	19.1	100	0	0	0							
6	2022-05-16 01:00	19.1	100	0	0	0							
7	2022-05-16 01:20	19.2	100	0	0	0							
8	2022-05-16 01:40	19.2	100	0	0	0							
9	2022-05-16 02:00	19.1	100	0	0	0							
10	2022-05-16 02:20	19	100	0	0	0							
11	2022-05-16 02:40	19	100	0	0	0							
12	2022-05-16 03:00	18.9	100	0	0	0							

<옵션 적용 후 데이터 조회 현황>

Block cooling | Block heating | Block ventilation | co2 group | recirculation fan group | sensor unit | ventilation leeward side | ventilation side 1 | ventilation side 2 | ventilator

BLOCK block 01 날짜 2022-05-16 ~ 2022-05-22 검색

Save Export Options Fields

1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	시간	블록공기온도	블록상대습도	ALARM_LIMIT_FOR_RH_TOO_LOW	HD	ALARM_LIMIT_FOR_HD_TOO_HIGH							
2		11 708.1	50.128	0	60.3	0							
3	2022-05-16 00:00	19.1	100	0	0	0							
4	2022-05-16 00:20	19	100	0	0	0							
5	2022-05-16 00:40	19.1	100	0	0	0							
6	2022-05-16 01:00	19.1	100	0	0	0							
7	2022-05-16 01:20	19.2	100	0	0	0							
8	2022-05-16 01:40	19.2	100	0	0	0							
9	2022-05-16 02:00	19.1	100	0	0	0							
10	2022-05-16 02:20	19	100	0	0	0							
11	2022-05-16 02:40	19	100	0	0	0							
12	2022-05-16 03:00	18.9	100	0	0	0							
13	2022-05-16 03:20	18.7	100	0	0	0							
14	2022-05-16 03:40	18.7	100	0	0	0							
15	2022-05-16 04:00	18.8	100	0	0	0							

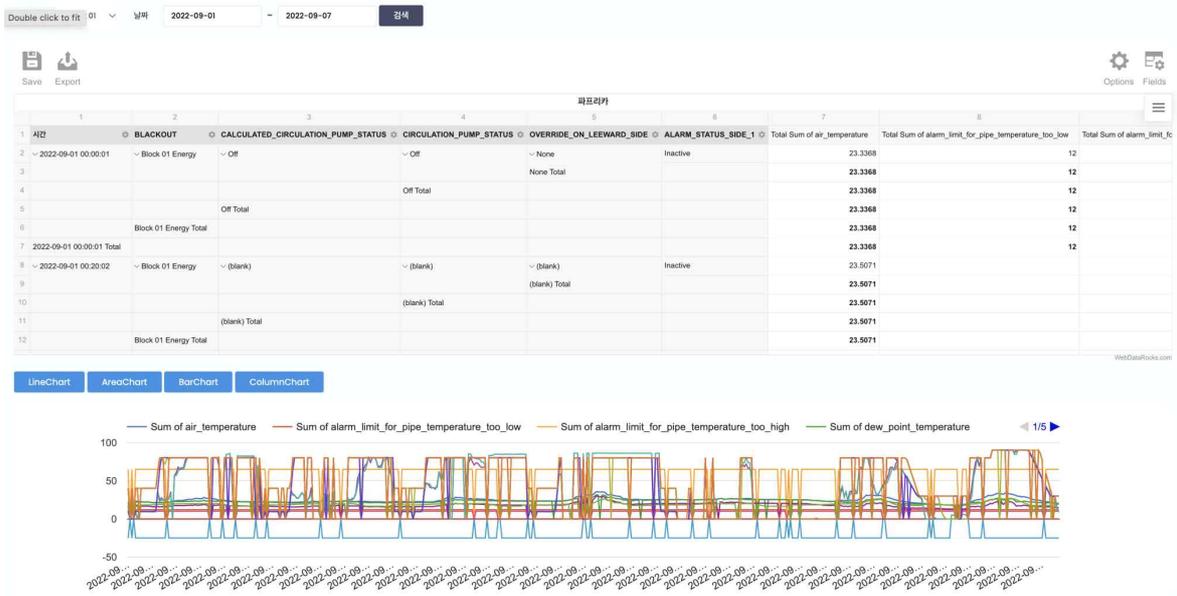
블록공기온도

Sort: AZ ZA TOP 10

Select All 131 of 131 selected Search Q

- 17.8
- 18
- 18.3
- 18.4
- 18.5
- 18.6

<데이터 분석을 위한 랭킹 옵션 기능>



### <데이터 분석을 위한 Pivot 및 차트 기능>

#### □ 커넥티드 팜 데이터를 활용한 개별 생육-환경 -영농 인자분석 도구 개발

- 인자모형 추정작업
  - 주성분 방법(Pricipal Component Method)
  - 최우추정법(Maximum Likelihood Method)
  - 최소 잔차법(Minimum residual)
- 인자모형을 기반으로 변수들간의 관련성 분석

#### [2단계(2023)]

#### □ 분석 파이프라인의 분석 결과의 시각화 실증

- 스마트팜 빅데이터 플랫폼 분석 항목
  - 스마트팜 온실에서 수집된 데이터 종료
    - 농가에 따라 API로 수집된 데이터 또는 수동으로 업로드한 데이터에 대한 분석 수행
    - 총 27종의 API 및 수동으로 수집된 데이터를 기반으로 농업에 필요한 분석 자료를 제공 (스마트팜 온실의 환경 데이터. 예:온도, 습도, CO2 등)
  - 데이터 분석 결과
    - 작기별 작물 비교 분석 : 전년 대비 금년에 수집된 데이터를 비교할 수 있는 차트 제공
    - 농가 작물 분석 : 기준 데이터와 농가의 작물 데이터를 비교할 수 있는 차트 제공
    - 일별 환경 데이터를 비교할 수 있는 데이터와 농업 분석에 필수적인 수식 데이터를 자동 생성
  - 상관관계 분석
    - 스마트팜 온실의 환경데이터와 생육데이터의 관계성을 분석하고 필수적인 변수들을 도출
    - 상관관계 분석 결과는 히트맵 형식으로 정리하여 표현

○ 분석 결과에 대한 시각화 데이터 실증

- 수집된 데이터 기반 데이터 분석

- 수집 데이터의 현황을 제공하기 위해 1~2차에 걸쳐 집계 테이블 생성
- 1차 집계테이블: 27종 API에서 수집되는 450여 종의 데이터 중 분석에 반드시 필요한 60여종을 별도로 집계하여 수집
- > API가 수집되어 원천테이블에 적재 후, 집계테이블에도 별도로 적재
- 집계 결과는 다음과 같다.

block	check_time	backout	air_temperatu	rh	hd	screen_status	pipe_temperal	alarm_limit	for_alarm_limit	for_maximum_pip	humidity_inflx	calculated_circ	circulation_pu	wet_bulb	benj	dew_point	ten_vapour	presure	absolute_hum	enthalpy	status	of_sens_vapor	contain	sensor	unit	at_radiation	inter_alarm	limit	for	minimum	pi
Block 06	2023-12-16	Block 06 Shw	14.5413	87	1.4098	Thermal	0	5	20	0	0	0	0	Off	13.222	12.3289	0.22549	8.91118	37.1624	0	Full	Active	4	(NULL)	(NULL)						

<1차 집계된 테이블의 내용>

- 2차 집계되는 테이블은 분석에 필요한 수식을 반영하여 새로운 계산값을 도출하여 적재
- 계산된 수식 값은 차트와 그리드를 통해 웹화면에서 조회하도록 구성
- 수식값은 매일 오전 7시에 배치를 통해 생성

이슬점 온도	$es = 6.112 * \exp(17.67 * T) / (T + 243.5)$ $e = es * (RH / 100.0)$ $Td = \log(e / 6.112) * 243.5 / (17.67 - \log(e / 6.112))$	T = temperature in de c es = saturation vapor pressure in ml e = vapor pressure in mb RH = relative humidity in percent Td = dew point in deg C
질대 습도	$e = 6.112 * \exp(17.67 * Td) / (Td + 243.5)$ $q = (0.622 * e) / (p - (0.378 * e))$ $q = q * 1000$	p = surface pressure in mb q = specific humidity in kg/kg.
수분부족분		
증기압 부족분	$SVP = 0.6108 * \exp(17.27 * T / (T + 237.3))$ $AVP = SVP * (RH / 100)$ $VPD = SVP - AVP$	
Enthalpy	$air\ enthalpy = 1.007 * T - 0.026$ $Water\ vapor\ enthalpy = q * (2501 + 1.84 * T)$ $Enthalpy = air\ enthalpy + water\ vapor\ entalpy$	

<데이터 생성 수식>



선택한다.

③ OLS(Ordinary Least Square) 분석

하나의 종속변수와 하나 또는 여러 독립 변수 간의 선형 관계를 추정하는 방법

① 모델설정은 종속 변수와 잠재적인 독립 변수들을 포함하는 선형 모델을 설정한다.

② OLS추정은 최소제곱법을 사용하여 모델 파라미터(계수)를 추정하고,

③ 계수의 통계적 유의성 검증은 각 독립 변수의 계수에 대한 t 통계량과 p-value 값을 계산하여 통계적 유의성을 검증한다

④ 모델의 적합도 평가는 결정 계수를 평가하여 모델의 전반적인 적합도를 확인하며,

⑤ 변수 선택은 통계적으로 유의미한 계수를 가진 독립 변수를 채택하여 높은 p-value를 가진 변수는 제외한다.

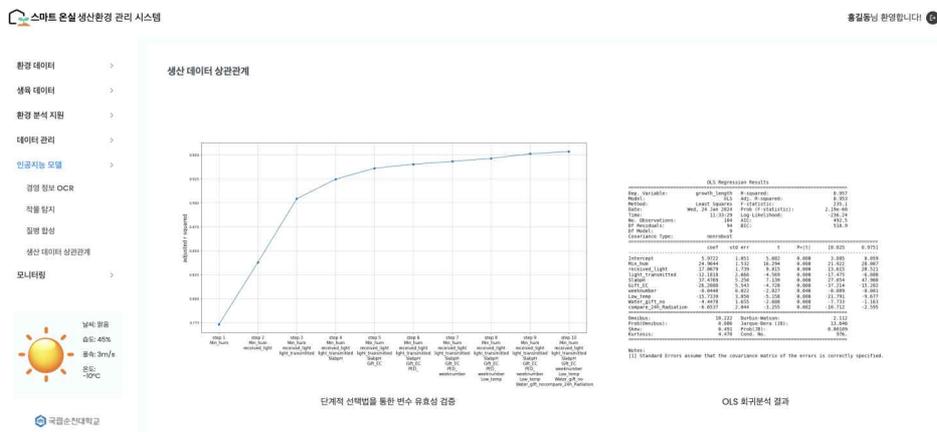
④ 단계적 선택법을 이용한 유효한 변수 추출

단계적 선택법은 모델에 변수를 추가하거나 제거함으로써 최적의 변수 조합을 찾는 방법

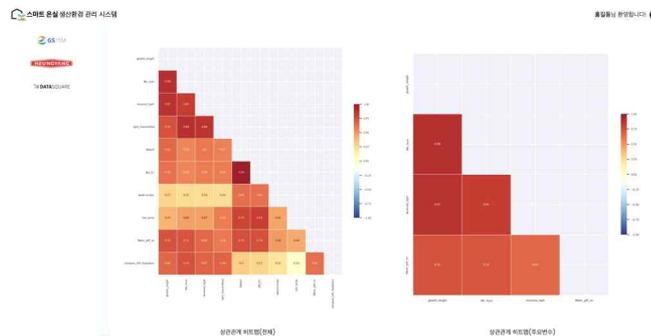
① 전진 선택법은 모델에 아무 변수도 포함하지 않은 상태에서 시작하여 통계적으로 유의미한 변수를 하나씩 추가한다.

② 후진 제거법은 모든 변수를 포함한 모델에서 시작하여 가장 유의미하지 않은 변수를 하나씩 제거한다

③ 단계적 선택법은 전진 선택법과 후진 제거법을 결합한 방법으로 변수를 추가하는 동시에 기존 변수의 유의성을 재평가하여 필요없는 변수를 제거한다.



<상관분석 결과>



<상관분석 히트맵>

□ 영농 전문가와 시각화 요소의 적절성 평가/보완

○ 시각화 요소의 적절성

- 수집되는 데이터를 기반으로 다양한 농업에 관련된 데이터를 시각화로 표현
- 필수 데이터들 : 수집된 원천데이터, 계산이 수반되는 필수 데이터, 경영데이터, 작물 이미지 등
- 원천데이터를 기반으로 그래프와 테이블 형식으로 표현하고 있는지, 분석을 효율적으로 할 수 있는지를 검토하고 보완

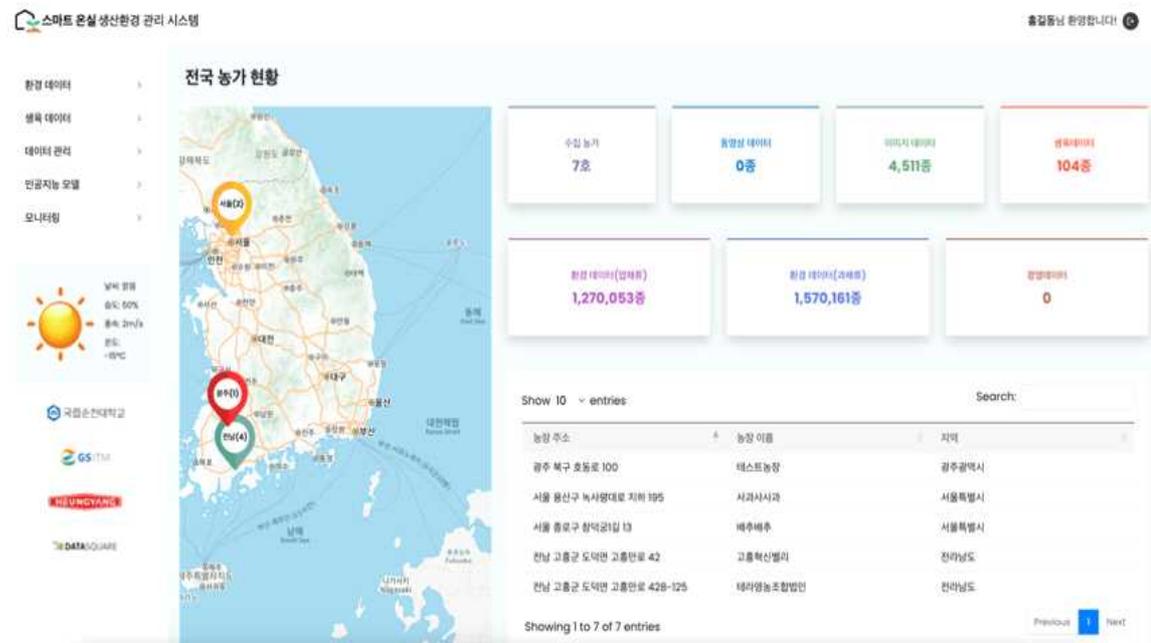
○ 빅데이터 플랫폼의 시각화 기능

- 빅데이터 플랫폼에서 관리되는 데이터들을 분석용 데이터로 구성하여 시각화
- 주요 기능으로는 원천데이터 조회(정형/비정형데이터) / 분석용 데이터 조회 / 권한 별 데이터 조회 / 권한 관리 / 모니터링 등을 제공

• 원천 데이터 조회 검색

- 대시보드 :농가 집계 및 데이터 집계를 메인 화면으로 구성되어 있으며, 농가와 수집되는 데이터의 현황을 한눈에 볼수 있는 카드형태로 제공

- ① 수집되는 데이터 현황 표현(농가 수, 비정형데이터 수, 온실환경데이터 수 등)
- ② 수집되는 농가별/작물별 메뉴 자동 생성
- ③ 기상청 API를 통해 날씨 표현
- ④ 지도 기반 농가의 위치 표시



<시각화 플랫폼 대시보드>

- 환경 데이터
  - 홍양영농조합법인 >
    - 테스트(수동)
    - 고추
    - 토마토
    - 파프리카
  - 알로하채소농장 >
    - 패밀리
    - 이자별
    - 스텔릭스
    - 바질
    - 로메인 상추
    - 고수
  - 테라영농조합법인 >
    - 토마토



<데이터 적재 현황>

<작물 메뉴>

- 업로드된 정형 데이터 조회
  - 사용자가 수동으로 업로드 한 작물의 정형 데이터 조회
  - csv 형태 그대로 데이터를 표출하며, 엑셀/csv/html 형식으로 저장 가능

스마트 온실 생산환경 관리 시스템 홍일동님 환영합니다

평균 14.69 25.15 18.60 -0.49 9.64 4.57 20.20 17.25 88.08 84.88 389.25 507.43 4

원본 테이블

시간	날씨	실내온도	습구온도	건구온도	상대습도	절대습도	수분부족량	CO2	VPD	Enthalpy	외부온도	상대습도	풍향	풍속	일사량
2023-12-27 00:00:00	2023-12-27	16.87	15.22	14.25	84.00	10.12	1.90	663.00	0.30	42.57	1.92	84.00	0.28	309.00	0.00
2023-12-27 00:20:00	2023-12-27	16.96	15.40	14.50	85.00	10.29	1.78	659.00	0.28	43.09	2.00	85.00	0.15	28.00	0.00
2023-12-27 00:40:00	2023-12-27	17.06	15.43	14.48	85.00	10.28	1.88	667.00	0.30	43.16	1.72	85.00	0.14	108.00	4.00
2023-12-27 01:00:00	2023-12-27	17.06	15.37	14.39	84.00	10.21	1.93	671.00	0.31	43.00	1.72	84.00	0.34	43.00	4.00
2023-12-27 01:20:00	2023-12-27	16.91	15.15	14.12	84.00	10.03	2.01	671.00	0.31	42.38	1.60	84.00	0.14	349.00	0.00
2023-12-27 01:40:00	2023-12-27	16.81	15.05	13.99	83.00	9.95	2.01	683.00	0.32	42.10	1.19	83.00	0.29	48.00	0.00
2023-12-27 02:00:00	2023-12-27	16.74	15.01	13.97	84.00	9.94	1.97	691.00	0.31	41.98	1.47	84.00	0.22	349.00	0.00
2023-12-27 02:20:00	2023-12-27	16.70	14.98	13.94	84.00	9.92	1.96	692.00	0.31	41.89	1.04	84.00	0.25	43.00	0.00
2023-12-27 02:40:00	2023-12-27	16.65	14.83	13.72	83.00	9.78	2.06	688.00	0.33	41.49	0.72	83.00	0.29	13.00	0.00
2023-12-27 03:00:00	2023-12-27	16.50	14.60	13.43	82.00	9.59	2.19	691.00	0.34	40.84	0.51	82.00	0.61	71.00	0.00
2023-12-27 03:20:00	2023-12-27	16.30	14.45	13.30	82.00	9.51	2.06	687.00	0.33	40.44	1.15	82.00	0.56	185.00	0.00
2023-12-27 03:40:00	2023-12-27	16.05	14.27	13.16	83.00	9.42	1.98	684.00	0.31	39.96	0.11	83.00	0.33	347.00	0.00
2023-12-27 04:00:00	2023-12-27	15.91	14.06	12.88	82.00	9.25	2.01	699.00	0.32	39.38	-0.13	82.00	0.15	346.00	4.00
2023-12-27 04:20:00	2023-12-27	15.80	13.92	12.72	82.00	9.15	2.06	695.00	0.32	39.01	0.13	82.00	0.43	347.00	4.00
2023-12-27 04:40:00	2023-12-27	15.84	13.99	12.81	82.00	9.20	2.03	704.00	0.32	39.18	0.35	82.00	0.31	346.00	0.00

<데이터 조회 화면>

- 이미지, 영농일지 등 비정형 데이터 조회
  - 사용자가 수동으로 업로드 한 비정형 데이터 조회
  - 비정형 데이터는 데이터레이크(AWS S3)에 적재, (농가 / 날짜 / 파일목록)하여 관리하고 있으며, 시각화 플랫폼에서는 뷰어를 제공하여 비정형 데이터 출력

<비정형 데이터 뷰어>



<적재된 이미지 샘플>

<비정형 데이터(작물 이미지) 데이터레이크>

영농작업일지  
홍양영농조합법인

작성일자	2023년 7월 / 일 (수요일)	작성자	속성
작성필요기준	2023.5	작성필요기준	2023.6
작성시각	작성내용	작성일자	작성인원
	1농장 A를 정리		
	3농장 캔 미수리		
	3농장 인력 조정		그림
	4농장 격물 교체		대우기술사님
특기사항	내일 작업사항		

<비정형 데이터(영농일지) 데이터레이크>

<적재된 영농일지 샘플>

- 분석용 데이터 검색/조회

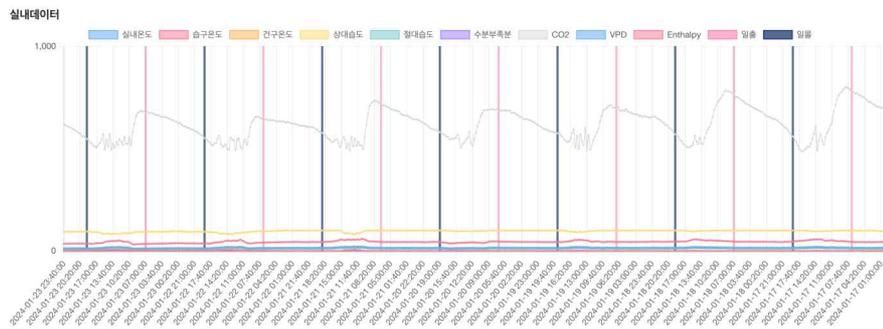
- 스마트 온실의 환경 데이터 분석을 위한 시각화 툴 제공
- 분석 차트는 수집데이터에서 스마트 온실에 대한 분석을 위해 필요한 수식을 생성하고 이를 기

반으로 가공/요약된 데이터를 표현

- 매일의 실내/실외/급액 등의 데이터를 한눈에 비교하고 트렌드를 분석하여 이상 상황을 감지하고, 이에 따라 더 좋은 온실 환경을 구성하기 위한 기반 마련
- 분석 데이터는 신래/실외/급액에 대한 데이터를 기반으로 제공하고, 월평균 데이터와 RTR 데이터를 생성하여 제공

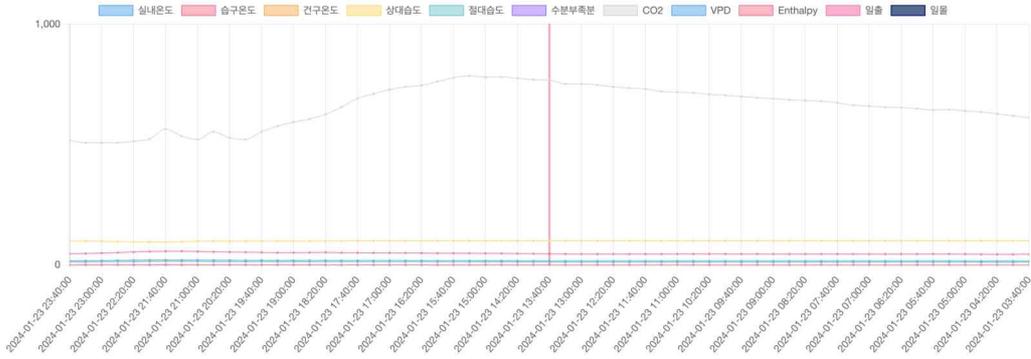


- 실내 데이터로는 실내온도, 습구온도, 건구온도, 상대습도, 절대습도, 수분부족분, CO2, VPD, Enthalpy를 제공하고 있으며, 매일 일출/일몰 기준으로 비교 분석할 수 있는 차트 제공
- 일주일 데이터를 기준으로 하루 7분할 데이터를 표출하며, 확대/축소 기능 제공



<차트 기본 - 일주일 요약 데이터>

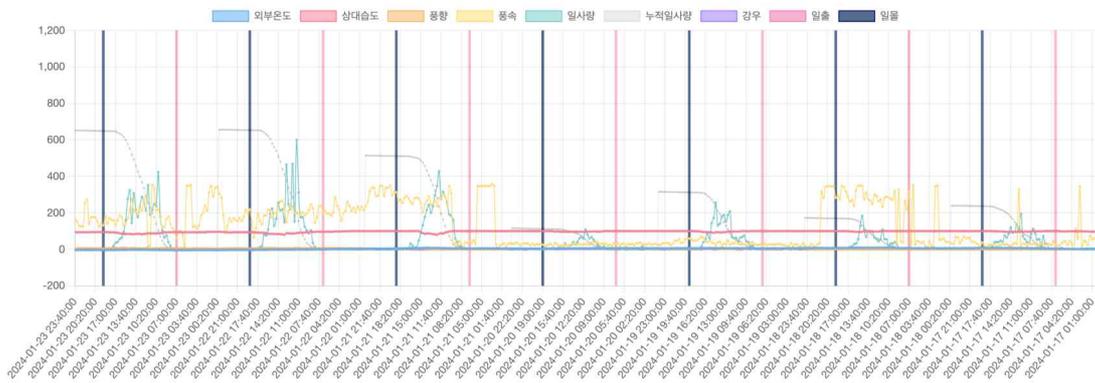
실내데이터



<차트 확대 - 최대 확대 시 매 20분 간격의 데이터 확인 가능>

- 실외데이터는 외부온도, 상대습도, 풍향, 풍속, 일사량, 누적일사량, 강우 데이터를 제공하며 일주일 데이터를 기초로 제공한다.

실외데이터



<실외 데이터 차트>

- 급액/배액 데이터는 수분함량, 급액EC, 급액pH, 급액량, 급액수, 급액시작, 급액종료, 배액EC, 배액pH, 배액량, 배액시작, 배액 후 누적량을 제공하며 데이터의 종류에 따라 취사 선택을 한다.

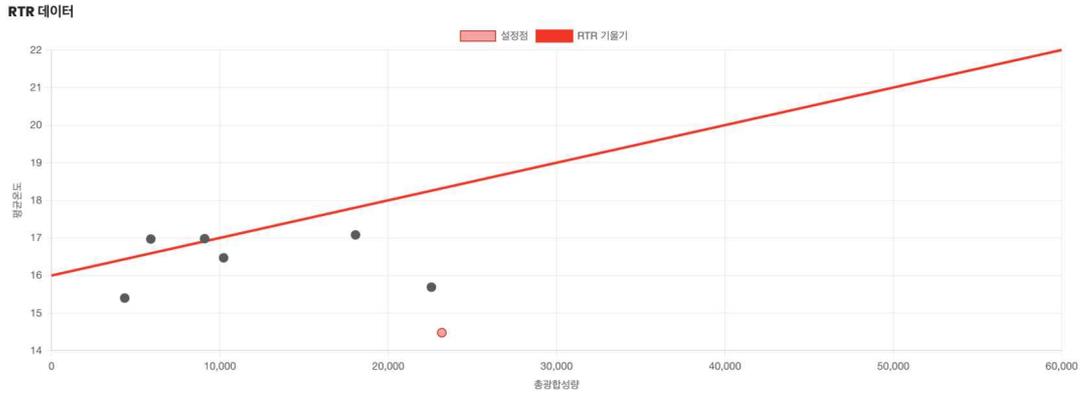
급액/배액 데이터



<급액/배액 데이터 차트>

- RTR 그래프

- 식물의 온도가 어떻게 변화하는지, 그 변화가 식물의 성장이나 건강에 어떤 영향을 미치는지를 시각적으로 표현
- 다양한 센서를 통해 실시간으로 수집하는 데이터(온도, 습도, 빛의 강도, 이산화탄소 농도 등)가 식물의 성장에 영향을 미치므로 이러한 요소들을 적절히 조절하는 것은 매우 중요
- RTR 그래프는 이러한 환경 조절의 일부로써 스마트 온실 관리에 중요한 도움을 줄 수 있음



<RTR 그래프>

- 평균 데이터 비교표

- 주요 변수들의 일평균 데이터를 계산하여 생산
- 평균 데이터는 내부온도(최저/최고/평균), 외부온도(최저/최고/평균), 주간평균온도, 야간평균온도, 주간평균습도, 야간평균습도, 누적광량, 최고일사량, 최저Co2, 최고Co2, DIF를 산출

2023-10-23 ~ 2024-01-23	내부최저온도	내부최고온도	내부평균온도	외부최저온도	외부최고온도	외부평균온도	주간평균온도	야간평균온도	주간평균습도	야간평균습도	누적광량	최고일사량	최저Co2	최고Co2
2024-01-23	11.59	19.91	14.48	-5.36	-2.19	-4.11	15.95	13.33	88.30	94.44	321.98	424.00	495.00	684.00
2024-01-22	13.25	21.00	15.69	-4.50	2.82	-1.56	16.86	14.70	89.52	96.44	313.38	600.00	494.00	657.00
2024-01-21	15.33	21.85	17.08	2.98	8.44	5.70	18.73	15.69	94.42	100.00	250.78	428.00	492.00	734.00
2024-01-20	13.14	16.73	15.40	4.30	6.99	5.82	14.99	15.75	97.76	100.00	60.37	108.00	505.00	694.00
2024-01-19	15.38	20.07	16.47	5.26	7.18	6.42	17.34	15.73	96.94	99.97	141.95	256.00	494.00	714.00
2024-01-18	15.91	20.31	16.97	6.29	9.99	8.03	17.99	16.12	98.76	100.00	81.90	184.00	507.00	785.00
2024-01-17	15.40	20.44	16.98	0.16	8.25	5.60	18.21	15.94	98.15	99.05	126.30	196.00	486.00	800.00
2024-01-16	14.17	25.20	17.63	-3.70	7.34	1.87	20.17	15.49	83.58	95.95	560.45	620.00	494.00	723.00
2024-01-15	15.02	25.80	18.07	-1.27	5.78	1.95	20.83	15.74	84.42	98.21	622.69	604.00	460.00	714.00
2024-01-14	15.41	26.00	18.25	-3.31	14.21	5.33	20.71	16.17	84.00	98.21	509.55	548.00	474.00	806.00
2024-01-13	15.36	25.36	18.18	-2.06	8.63	3.25	20.97	15.81	83.55	96.72	580.01	584.00	463.00	764.00
2024-01-12	15.62	25.42	18.23	2.62	10.26	6.37	20.83	16.03	86.24	99.74	505.33	536.00	468.00	751.00
2024-01-11	15.03	26.18	18.09	-1.20	6.79	2.25	20.71	15.86	85.64	99.21	580.76	580.00	450.00	764.00
2024-01-10	14.90	25.43	17.64	0.47	7.34	3.43	19.95	15.69	90.45	99.79	442.44	544.00	469.00	726.00
2024-01-09	15.12	18.91	16.56	-3.75	3.21	0.62	17.60	15.68	98.85	98.26	125.45	168.00	503.00	793.00
평균	14.71	22.57	17.05	-0.20	7.00	3.40	18.78	15.58	90.71	98.40	348.22	425.33	483.60	740.60

<일 평균 데이터>

- 컨설팅을 위한 분석 자료 제공

- 수집 데이터를 적극 활용하여 스마트 온실 가이드를 구축하기 위해 분석 시각화 툴을 개발하였다
  - ① 가이드 데이터
    - 작물에 대한 표준 데이터를 설정하고, 표준 데이터와 수집 데이터를 비교 분석할 수 있는 기능을 제공
    - 각 농가 별 작물의 온실 데이터에 대한 그래프와 그리드를 제공하여, 표준과의 차이점을 비교하고 개선점을 도출할 수 있는 기능을 제공함으로써, 농가의 작물을 더욱 잘 성장할 수 있는 컨설팅에 활용 가능

가이드

2024-01-18 ~ 2024-01-24	내부최저온도	내부최고온도	내부평균온도	외부최저온도	외부최고온도	외부평균온도	주간평균온도	야간평균온도	주간평균습도	야간평균습도	주간평균습도	이산화탄소	최고CO2	DIF
2024-01-23	10.84	19.46	14.39	-5.36	-2.19	-4.11	15.29	13.62	89.45	83.00	321.98	424.00	0	0.59
2024-01-22	11.86	19.50	15.02	-4.50	2.82	-1.56	15.69	14.45	90.33	88.00	313.38	600.00	0	0.07
2024-01-21	13.06	20.74	15.80	2.98	8.44	5.70	17.45	14.40	95.82	98.00	250.78	438.00	0	1.64
2024-01-20	12.84	17.10	14.65	4.30	6.99	5.82	15.20	14.18	97.73	100.00	60.37	108.00	0	1.33
2024-01-19	13.49	19.17	15.42	5.26	7.18	6.42	17.05	14.04	97.73	100.00	141.95	256.00	0	2.71
2024-01-18	13.52	17.96	15.29	6.29	9.99	8.03	16.61	14.16	99.48	100.00	81.90	184.00	0	2.85
평균	12.60	18.99	15.09	1.50	5.54	3.38	16.21	14.14	95.09	94.83	195.06	334.67	0.00	1.53

<토마토 작물의 표준데이터>

태라링농조합법인 | 1월 | 3월 | 7월 | 15일 | 1개월 | 3개월 | 2024-01-18 | 2024-01-24 | 검색

가이드

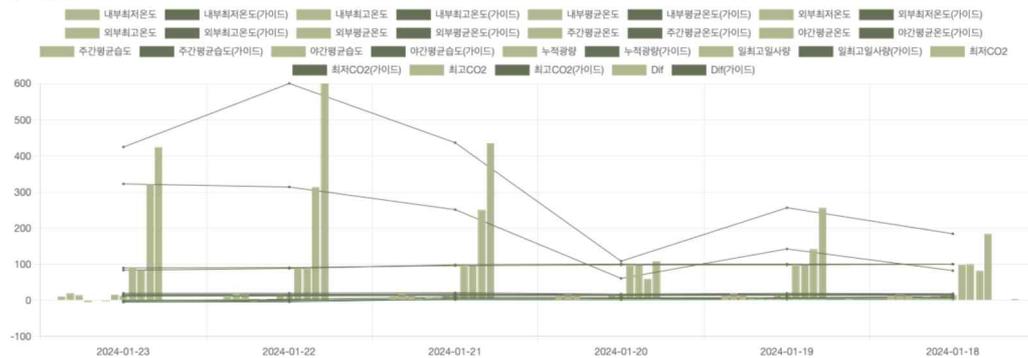
2024-01-18 ~ 2024-01-24	내부최저온도	내부최고온도	내부평균온도	외부최저온도	외부최고온도	외부평균온도	주간평균온도	야간평균온도	주간평균습도	야간평균습도	주간평균습도	이산화탄소	최고CO2	DIF
2024-01-23	10.84	19.46	14.39	-5.36	-2.19	-4.11	15.29	13.62	89.45	83.00	321.98	424.00	0	0.59
2024-01-22	11.86	19.50	15.02	-4.50	2.82	-1.56	15.69	14.45	90.33	88.00	313.38	600.00	0	0.07
2024-01-21	13.06	20.74	15.80	2.98	8.44	5.70	17.45	14.40	95.82	98.00	250.78	438.00	0	1.64
2024-01-20	12.84	17.10	14.65	4.30	6.99	5.82	15.20	14.18	97.73	100.00	60.37	108.00	0	1.33
2024-01-19	13.49	19.17	15.42	5.26	7.18	6.42	17.05	14.04	97.73	100.00	141.95	256.00	0	2.71
2024-01-18	13.52	17.96	15.29	6.29	9.99	8.03	16.61	14.16	99.48	100.00	81.90	184.00	0	2.85
평균	12.60	18.99	15.09	1.50	5.54	3.38	16.21	14.14	95.09	94.83	195.06	334.67	0.00	1.53

선택농가

2024-01-18 ~ 2024-01-24	내부최저온도	내부최고온도	내부평균온도	외부최저온도	외부최고온도	외부평균온도	주간평균온도	야간평균온도	주간평균습도	야간평균습도	주간평균습도	이산화탄소	최고CO2	DIF
2024-01-23	10.84	19.46	14.39	-5.36	-2.19	-4.11	15.29	13.62	89.45	83.00	321.98	424.00	0	0.59
2024-01-22	11.86	19.50	15.02	-4.50	2.82	-1.56	15.69	14.45	90.33	88.00	313.38	600.00	0	0.07
2024-01-21	13.06	20.74	15.80	2.98	8.44	5.70	17.45	14.40	95.82	98.00	250.78	438.00	0	1.64
2024-01-20	12.84	17.10	14.65	4.30	6.99	5.82	15.20	14.18	97.73	100.00	60.37	108.00	0	1.33
2024-01-19	13.49	19.17	15.42	5.26	7.18	6.42	17.05	14.04	97.73	100.00	141.95	256.00	0	2.71
2024-01-18	13.52	17.96	15.29	6.29	9.99	8.03	16.61	14.16	99.48	100.00	81.90	184.00	0	2.85
평균	12.60	18.99	15.09	1.50	5.54	3.38	16.21	14.14	95.09	94.83	195.06	334.67	0.00	1.53

<토마토 작물의 표준 데이터와 농가의 작물 데이터>

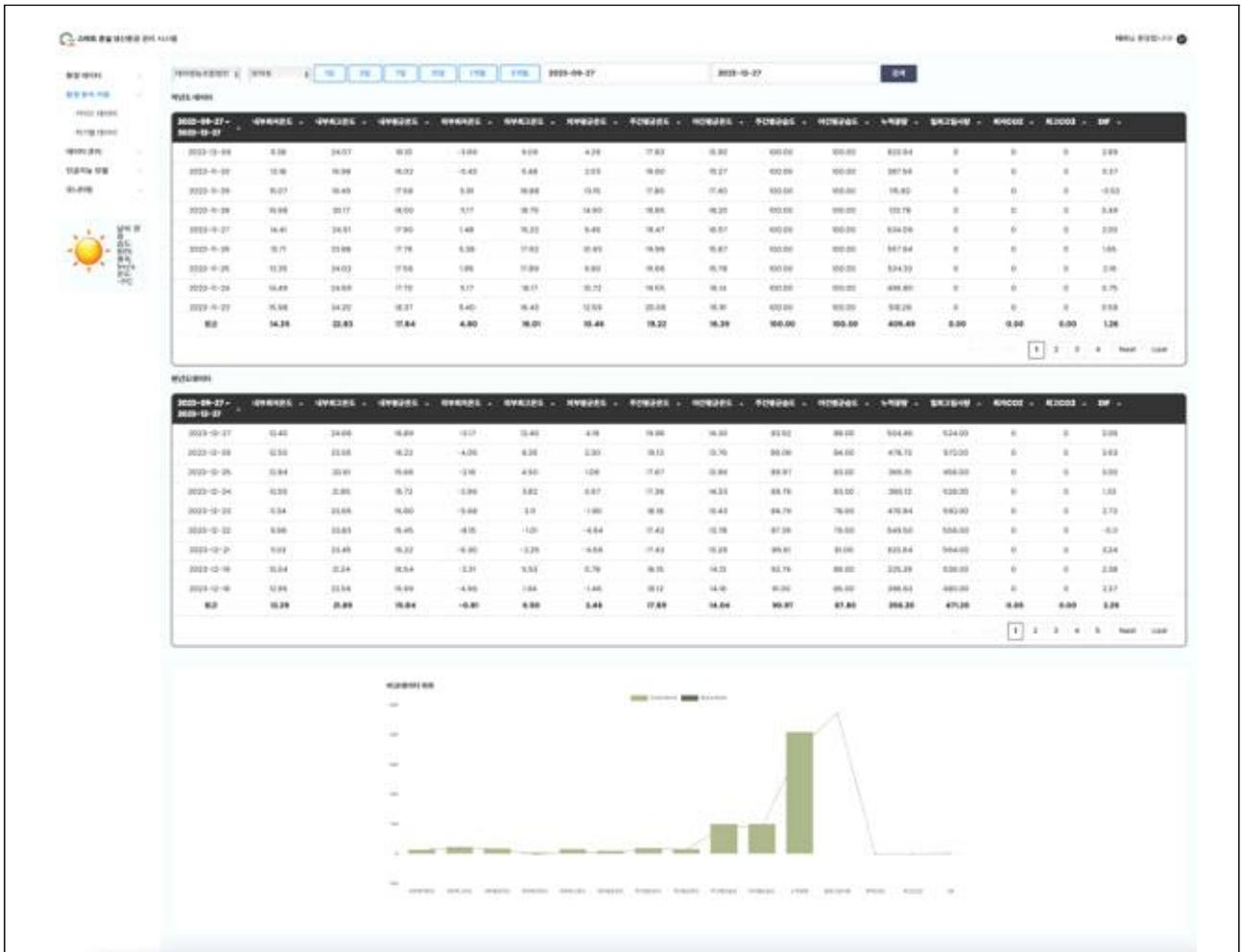
비교데이터 차트



<토마토 작물의 표준데이터와 농가 데이터 비교 그래프>

② 작기별 데이터

- 생물의 생장의 중요 변수인 환경 변수들을 기간별로 비교할 수 있는 기능을 제공
- 금년과 지난해의 변수를 비교하여 생물 성장에 미치는 영향도를 비교



<작기별 데이터 비교>

- 데이터 다운로드 기능(스마트팜 R&D 빅데이터 플랫폼)
  - 스마트팜 빅데이터 플랫폼에서 관리하고 있는 데이터를 타 플랫폼과 연계하기 위하여 자동으로 매핑한 집계 데이터를 제공
  - API로 수집되는 데이터는 총 29종의 API를 연계하고 있으며, 각 API의 데이터를 연계하기 위한 키를 설정하고 하나의 집합체로 조합
  - 데이터는 총 16개의 block으로 구성되어 있으므로 다운로드 효율을 위하여 block별로 다운로드 할 수 있도록 구성
  - 다운로드 가능한 데이터는 1백만개('22년7월 ~ 현재), 550여종 컬럼으로 구성
  - 이 중 타 플랫폼과의 연계를 위한 중요 컬럼(100여종)으로 구성된 데이터셋을 제공



데이터 미리보기(샘플 데이터만 표현됩니다.)

데이터 업로드

온도	온도 방향	화면 상태	상대 습도	습도 결빙	매니폴드 온도	난방값	상대 습도	파이프 온도	최소 파이프 온도	최대 파이프 온도	증기값	필트 상태
4.32	0	2.798525629	89.83	0.004599002	11.53866004	0	0	0.084528507	0	0	0	0
4.36	0	2.863135865	89.99	0.004620179	11.59213349	0	0	0.083430125	0	0	0	0
4.43	0	2.90778465	89.83	0.004634665	11.62907523	0	0	0.085180131	0	0	0	0
4.45	0	3.033415833	90.51	0.004676408	11.7339879	0	0	0.079596046	0	0	0	0
4.4	0	3.088169494	91.18	0.004694618	11.77924643	0	0	0.073718053	0	0	0	0
4.5	0	3.174699212	91.1	0.004723523	11.85264196	0	0	0.074909121	0	0	0	0
4.5	0	3.126503466	90.79	0.004707404	11.81219409	0	0	0.077518316	0	0	0	0
4.46	0	3.152187162	91.21	0.004715988	11.83338653	0	0	0.073776508	0	0	0	0
4.6	0	3.310876008	91.34	0.004769333	11.96847067	0	0	0.073400592	0	0	0	0
4.57	0	3.257926808	91.19	0.004751474	11.92339184	0	0	0.074515522	0	0	0	0

### <업로드 데이터 미리보기>

② 각 농가에서 업로드된 데이터는 다음과 같이 업로드된 테이블명/적재된 데이터 수/ 업로드 일자 확인 가능

업로드된 데이터는 자동으로 작물명을 왼쪽 메뉴로 생성하며, 이를 통하여 데이터를 조회 가능



### <업로드 데이터 목록과 자동으로 생성된 작측 메뉴>



### <수동으로 업로드 데이터 조회 화면>

#### • 데이터 접속 권한 부여

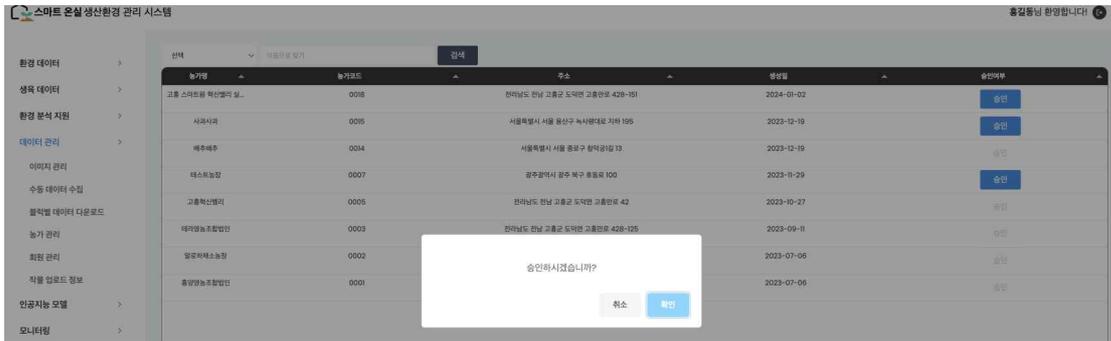
- 각 농장의 담당자는 농장의 데이터에 접속하거나 업로드 할 수 있는 권한을 선 취득해야 하며, '농장가입 -> 승인 -> 회원가입 -> 승인' 의 4단계로 권한 획득.

#### ① 농장가입/승인

농장명, 위치, 담당자 연락처를 입력한다.



<농장 가입 화면>



<농장 승인(시스템 관리자) 화면>

② 회원가입/승인

로그인 화면에서 회원가입 버튼을 클릭하여 진입하며, 회원 정보와 농가를 선택



<회원가입 화면>

시스템 관리자는 회원 목록을 확인할 수 있으며, 신규로 등록한 회원의 승인처리를 수행한다.

회원명	농가명	이메일	생성일	승인여부
강강한	홍양영농조합법인	korea123@gmail.com	2023-12-18	승인
강영현	홍양영농조합법인	korea123@naver.com	2023-12-19	승인
조승연	홍양영농조합법인	farm1234@naver.com	2023-12-19	승인
홍양	홍양영농조합법인	farm1@test.com	2023-12-18	승인
테라	테라영농조합법인	farm3@test.com	2023-12-18	승인
강강한	홍양영농조합법인	abc@naver.com	2023-12-18	승인
김도원	달로하채소농장	one@naver.com	2023-12-15	승인
이영배	홍양영농조합법인	imb@scnu.ac.kr	2023-10-11	승인
오지영	시스템운영	oat115@giltn.com	2023-07-06	승인
윤거문	시스템운영	gayun526@korea.kr	2023-07-06	승인
장강민	시스템운영	jagnm0925@scnu.ac.kr	2023-07-06	승인

<회원 목록>

<회원 승인 처리>

- 시스템 운영을 위한 모니터링 기능
  - 시스템 모니터링: 데이터 파이프라인의 성능을 모니터링(예:Prometheus, Grafana)
  - 유지보수: 오류 수정, 성능 최적화 등을 정기적으로 수행
- 데이터 수집 모니터링 기능
  - Grafana를 이용하여 Apache kafka를 모니터링하는 것은 kafka 클러스터의 성능과 상태를 시각화하고 이해하기 위한 중요한 과정
    - ① 메트릭 수집 : Kafka의 성능지표를 수집하기 위해 Prometheus, JMX Exporter 또는 Kafka 자체의 메트릭 수집 시스템과 같은 도구를 사용할 수 있으며, JMX Exporter는 Java Management Extension(JMX)를 통해Kafka에서 메트릭을 수집하여 Prometheus 형식으로 변환하는 역할
    - ② Prometheus 설정 : Prometheus 를 사용하는 경우, 'prometheus.yml' 설정 파일에 Kafka를 타겟으로 추가하여 메트릭을 수집하도록 설정
    - ③ Grafana 대시보드 설정 : Grafana에 로그인한 후 데이터 소스로 Prometheus를 추가. 이후 Kafka 모니터링을 위한 대시보드를 새로 생성하거나, 이미 존재하는 Kafka 대시보드 템플릿을 가져옴
    - ④ 대시보드 커스터마이징 : Kafka에 관련된 다양한 패널을 대시보드에 추가. 예를 들면, 메시지 처리량, 파티션별 레이턴시, 브로커 상태, 컨슈머 그룹의 오프셋 등을 시각화
    - ⑤ 알림 설정 : Kafka 클러스터의 특정 조건을 모니터링 하고, 이상 징후가 발생할 경우 알림을 설정. 예를 들어, 메시지 큐의 길이가 임계값을 초과하거나 브로커가 다운된 경우 등에 대한 알림 발생

항목	설명
Topics	Kafka에서 데이터 스트림을 토픽이라는 카테고리 분류. 토픽은 메시지를 게시(publish) 하거나 구독(subscribe)하는데 사용되는 이름 토픽은 하나 이상의 메시지를 포함할 수 있으며, 이러한 메시지는 토픽내에서 순차적으로 관리
Partitions	토픽은 여러 파티션으로 나누어짐. 각 파티션은 메시지의 순서를 유지하는 독립적인 시퀀스로, 병렬 처리와 확장성을 가능하게 함. 프로듀서는 메시지를 특정 파티션에 게시하며, 컨슈머는 하나 이상의 파티션에서 메시지를 읽음
Replicas	고가용성과 내구성을 위해, 각 파티션은 복제(replica) 될수 있음. 복제는 주 파티션의 정확한 복사본으로 브로커간에 분산
In Sync Replicas	동기화된 복제본은 주 파티션의 최신 상태를 유지하고 있는 복제본 집합을 의미. 주 브로커 장애가 발생하면 ISR 중 하나가 새로운 주 브로커로 승격 됨
Under Replicated Patitions	부족 복제 파티션은 주 파티션에 비해 동기화되지 않은 복제본을 가진 파티션을 의미함. 이는 복제본이 주 파티션과 동일한 데이터를 가지고 있지 않음을 의미하며, 보통 복제본이 충분히 빠르게 데이터를 복사하지 못했을 때 발생함
Partitions at minimum ISR	최소 ISR을 가진 파티션은 설정된 최소 ISR 크기에 도달한 파티션을 의미. 이는 파티션의 복제본이 최소한으로 요구되는 복제본수와 동일
Partitions under minimum ISR	최소 ISR 미만 파티션은 최소 ISR 크기보다 적은 복제본을 가지고 있는 파티션을 말한다. 이는 데이터 손실 위험을 증가시키며, 일반적으로 시스템 문제가 있을 알려주는 신호이다.
Partitions not on preferred node	신호 노드에 있지 않은 파티션은 파티션의 주 복제본이 선호하는 노드에 위치해 있는 않은 경우를 의미
Messages in per second	토픽으로 들어오는 메시지의 비율. 즉 초당 몇개의 메시지가 kafka 시스템에 의해 받아들여지고 있는지 표현
Messages comsumed per second	컨슈머에 의해 처리되는 메시지의 비율. 이는 초당 몇 개의 메시지가 컨슈머 그룹에 의해 소비되는지 표현
Lag by Consumer Group	특정 컨슈머 그룹의 lag 상태를 표현. 레그는 컨슈머가 아직 읽지 않은 메시지의 수를 의미하며, 데이터 처리의 지연을 나타내는 지표
Consumer Group Offsets	컨슈머 그룹 오프셋은 컨슈머 그룹이 마지막으로 읽은 메시지의 위치를 표현. 각 파티션에 대한 컨슈머의 진행 상태를 표현
Consumer Group Lag	컨슈머 그룹 레그는 특정 컨슈머 그룹이 처리해야 할 메시지의 양을 나타낸다. 이는 토픽의 최신 오프셋 차이를 통해 계산
Number of Partitions	특정 토픽이 가지고 있는 파티션의 총 수. 파티션 수는 토픽의 병렬 처리 능력과 관련
Latest Offsets	최신 오프셋은 각 파티션에서 가장 최근에 쓰여진 메시지의 오프셋을 의미. 토픽의 현재 길이를 나타내는 지표
Oldest Offsets	가장 오래된 오프셋은 각 파티션에서 가장 오래전에 쓰여진 메시지의 오프셋을 의미함. 데이터의 보존 기간과 관련이 있으며, 이 오프셋 이전의 데이터는 삭제될 수 있음.



<데이터 수집을 위한 Kafka 모니터링>

Consumer Group Offsets				Consumer Group Lag			
Time	consumergroup	partition	topic	Time	consumergroup	partition	topic
2023-12-20 09:44:16	weater_stations_group	0.00	ws_scnu_weater_stations	2023-12-20 09:44:16	valve_group_data_group	0.00	ws_scnu_irrigation_valve_data
2023-12-20 09:44:16	water_tanks_group	0.00	ws_scnu_water_tanks	2023-12-20 09:44:16	weater_stations_group	0.00	ws_scnu_irrigation_valve_data
2023-12-20 09:44:16	ventilation_windward_side_group_001	0.00	ws_scnu_ventilation_windward_side	2023-12-20 09:44:16	weater_stations_group	0.00	ws_scnu_weater_stations
2023-12-20	valve_group_data_group	0.00	ws_scnu_valve_group_data	2023-12-20	water_tanks_group	0.00	ws_scnu_water_tanks

Number of Partitions			Latest Offsets			Oldest Offsets		
Time	topic	Partitions	Time	partition	Offset	Time	partition	Offset
2023-12-20 09:44:29	ws_scnu_weater_stations	1	2023-12-20 09:44:29	0.00	ws_scnu_Thermal_screen	2023-12-20 09:44:29	0.00	ws_scnu_Thermal_screen
2023-12-20 09:44:29	ws_scnu_water_tanks	1	2023-12-20 09:44:29	0.00	ws_scnu_blackout_screen	2023-12-20 09:44:29	0.00	ws_scnu_blackout_screen
2023-12-20 09:44:29	ws_scnu_ventilation_windward_side	1	2023-12-20 09:44:29	0.00	ws_scnu_Soil_heating	2023-12-20 09:44:29	0.00	ws_scnu_Soil_heating
2023-12-20 09:44:29	ws_scnu_valve_group_data	1	2023-12-20 09:44:29	0.00	ws_scnu_Shading_screen	2023-12-20 09:44:29	0.00	ws_scnu_Shading_screen
2023-12-20 09:44:29	ws_scnu_irrigation_valve_data	1	2023-12-20	0.00		2023-12-20	0.00	

<데이터 수집 현황 상세>

- <시스템 모니터링 기능>
  - 데이터 수집을 하는 3개의 Node 중 첫번째 Node에 대한 모니터링
  - node exporter : 서버 시스템(cpu, loadAvg, memory) 및 네트워크 정보를 수집하여 메트릭(metric) API로 제공. 프로메테우스(Prometheus)에서 노드 익스포터 매트릭 API(endpoint)를 target에 등록하여 스크랩 가능

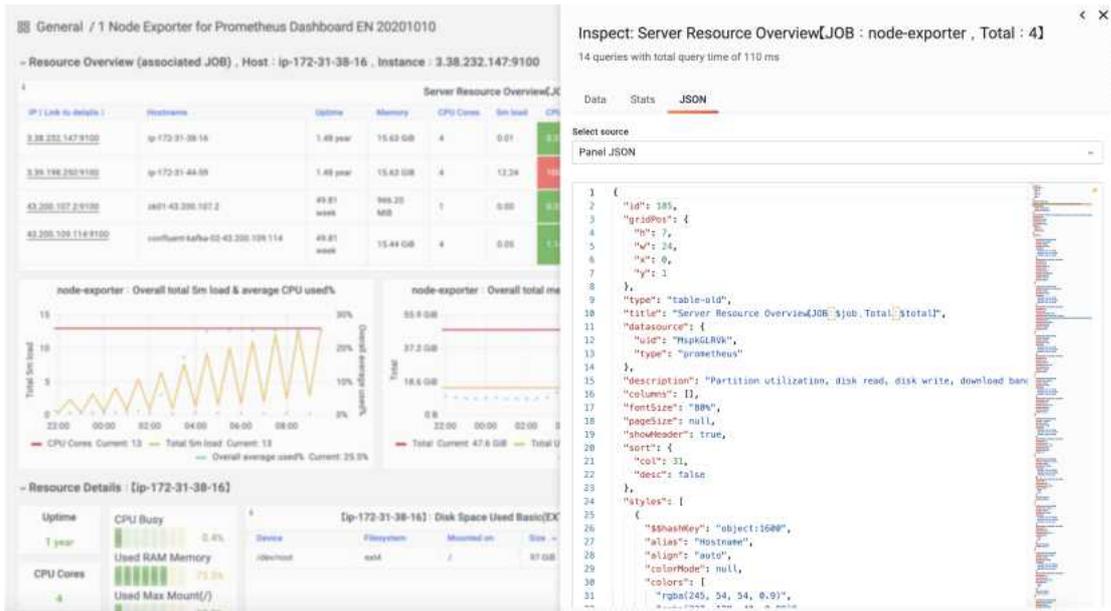
항목	설명
Node-exporter:Overall total 5m load & average CPU used%	시스템 평균 로드를 나타내며, 5분 동안의 CPU 부하와 사용률 표시
Node-exporter:Overall total memory & average memory used%	시스템 평균 로드를 나타내며, 전체 메모리 와 사용률 표시
Node-exprter:Overall total disk & average disk used%	시스템 평균 로드를 나타내며, 전체 디스크와 사용률 표시



<인프라 모니터링>



<인프라 모니터링 상세 예시 1>



<인프라 모니터링 상세 예시 2>



<인프라 모니터링 상세 예시 3>

□ 표준화에 입각한 데이터의 수집/저장/연동에 대한 실증

- 빅데이터 플랫폼 파이프라인 구축 프로세스
  - 파이프라인 실증을 위해 소스 API 데이터와 수집 데이터를 비교
  - 총 27종의 API 연동
- 빅데이터 플랫폼 파이프라인 구축에 대한 실증
  - 파이프라인 실증을 위해 소스 API 데이터와 수집 데이터를 비교
  - 데이터 처리 및 변환에 대한 실증
    - 이상치 제거, 누락된 값 처리 등을 수행하고, 필요에 따라 데이터를 적절한 형식으로 변환
    - 처리 도구 : Apache Spark, Apache Flink 등을 사용하여 대용량 데이터를 빠르게 처리
    - 데이터 이상치 제거 및 누락 값 처리 예시

[http://121.180.10.69:61480/syncustom.json?name=\\_ws\\_scnu\\_blackout\\_screen](http://121.180.10.69:61480/syncustom.json?name=_ws_scnu_blackout_screen)

```
{
  "name": "_ws_scnu_blackout_screen",
  "range": {
    "col1": 1,
    "row1": 1,
    "col2": 34,
    "row2": 18
  },
  "rows": [
    [
      "Blackout screen",
      "Block 01 Shade",
      "Block 01 Energy",
      "Block 02 Shade",
      ...
      null
    ],
    [
      "Control based on block",
      {
        "type": "hortimax.synopta.textid",
        "id": 2,
        "value": "Block 01"
      },
      {
        "type": "hortimax.synopta.textid",
        "id": 2,
        "value": "Block 01"
      }
    ]
  ]
}
```

```

    ...
    null
  ],
  [
    "Air temperature [◆C]",
    20.3532,
    20.3532,
    19.9884,
    19.9884,
    19.7833,
    19.7833,
    .
    ...
  ],
  [
    "RH [%]",
    92,
    92,
    96,
    96,
    85,
    85,
    94,
    ...
    null
  ],
  ...
],
"ts": "2023-12-15T14:35:49+09:00",
"pt": 0
}

```

- 상기 수신 데이터를 DB에 저장 시, null to zero로 치환하여 저장

calculated_outside_temperature_for_apply_gap	calculated_rh	calculated_hd
13	0	0
75	0	0
75	0	0
0	0	0
0	0	0
0	0	0
0	0	0
55	0	0
55	0	0
27	0	0
27	0	0
16	0	0

- 데이터 저장에 대한 실증

- 데이터 웨어하우스/레이크 : 처리된 데이터는 Hadoop, Amazon S3, Google BigQuery 등에 저장
- 분산 데이터베이스 : Cassandra, MongoDB와 같은 NoSQL 데이터베이스를 사용할 수 있음
- 본 과제에서는 데이터의 안전한 관리를 위해 정형데이터를 json형식으로 데이터 레이크에 저장/운영하고 있으며,
- 데이터 분석을 위해 데이터마트를 생성하여 시각화로 표현하고 있음
- 데이터 분석 시각화를 위한 데이터마트는 매일 아침 7시에 자동 배치로 생성됨
- <예시-1. API 수집 데이터 : blackout\_screen> : S3>blackout\_screen>block01>{date}>{json}

이름	유형	마지막 수정	크기	스토리지 클래스
2023-12-03/	폴더	-	-	-
2023-12-04/	폴더	-	-	-
2023-12-05/	폴더	-	-	-
2023-12-06/	폴더	-	-	-
2023-12-07/	폴더	-	-	-
2023-12-08/	폴더	-	-	-
2023-12-09/	폴더	-	-	-
2023-12-10/	폴더	-	-	-
2023-12-11/	폴더	-	-	-
2023-12-12/	폴더	-	-	-
2023-12-13/	폴더	-	-	-
2023-12-14/	폴더	-	-	-
2023-12-15/	폴더	-	-	-
2023-12-16/	폴더	-	-	-
2023-12-17/	폴더	-	-	-
2023-12-18/	폴더	-	-	-
2023-12-19/	폴더	-	-	-

<데이터 저장소 - 날짜별 폴더>

이름	유형	마지막 수정	크기	스토리지 클래스
blackout_screen-Block01-2023-12-1900:00:19.json	json	2023. 12. 19. am 12:01:14 AM KST	517.0B	Standard
blackout_screen-Block01-2023-12-1900:02:02.json	json	2023. 12. 19. am 12:21:02 AM KST	519.0B	Standard
blackout_screen-Block01-2023-12-1900:40:02.json	json	2023. 12. 19. am 12:41:05 AM KST	519.0B	Standard
blackout_screen-Block01-2023-12-1901:00:02.json	json	2023. 12. 19. am 1:01:01 AM KST	519.0B	Standard
blackout_screen-Block01-2023-12-1901:20:02.json	json	2023. 12. 19. am 1:21:03 AM KST	519.0B	Standard
blackout_screen-Block01-2023-12-1901:40:02.json	json	2023. 12. 19. am 1:41:00 AM KST	519.0B	Standard
blackout_screen-Block01-2023-12-1901:59:02.json	json	2023. 12. 19. am 2:01:04 AM KST	519.0B	Standard

<저장된 데이터 파일 목록>

<저장된 데이터 파일 상세>

```

blackout_screen-Block01-2023-12-1900_00_19.json
{"time": "2023-12-19 00:00:19", "blackout": "Block 01 Energy", "block": "Block 01", "air_temperature": "18.2997", "rh": "76", "hd": "3.196",
"calculated_outside_temperature_for_apply_gap": "0", "calculated_rh": "0", "calculated_hd": "0", "calculated_screen_position": "100",
"screen_position": "100", "dark_duration": "0", "gap_temperature": "0", "temperature_influence": "0", "humidity_influence": "0",
"remaining_delay_time_for_last_stage_open_screen": "0", "override_on_screen": "None", "screen_status": "Thermal"}

```

<저장된 데이터 파일>

```
select * from real_time.tb_ws_scnu_blackout_screen twsbs order by time desc
```

time	blackout	block	air_temperature	rh	hd	calculated_outside_temperature_for_apply_gap	calculated_rh	calculated_hd
2023-12-19 18:00:03.000	Block 05 Shade	Block 05	13.976	100	0	0	0	0
2023-12-19 18:00:03.000	Block 03 Energy	Block 03	18.418	76	3.245	0	0	0
2023-12-19 18:00:03.000	Block 01 Energy	Block 01	18.989	87	1.829	0	0	0
2023-12-19 18:00:03.000	Block 06 Energy	Block 06	14.567	100	0	0	0	0
2023-12-19 18:00:03.000	Block 07 Shade	Block 07	14.881	80	2.095	0	0	0
2023-12-19 18:00:03.000	Block 02 Shade	Block 02	19.371	87	1.849	0	0	0
2023-12-19 18:00:03.000	Block 01 Shade	Block 01	18.989	87	1.829	0	0	0
2023-12-19 18:00:03.000	Block 05 Energy	Block 05	13.976	100	0	0	0	0
2023-12-19 18:00:03.000	Block 08 Shade	Block 08	15.897	100	0	0	0	0
2023-12-19 18:00:03.000	Block 02 Energy	Block 02	19.371	87	1.849	0	0	0
2023-12-19 18:00:03.000	Block 06 Shade	Block 06	14.567	100	0	0	0	0
2023-12-19 18:00:03.000	Block 04 Shade	Block 04	18.948	88	1.657	0	0	0
2023-12-19 18:00:03.000	Block 09 Energy	Block 09	18.545	90	1.367	0	0	0
2023-12-19 18:00:03.000	Block 09 Shade	Block 09	18.545	90	1.367	0	0	0
2023-12-19 18:00:03.000	Block 03 Shade	Block 03	18.418	76	3.245	0	0	0
2023-12-19 18:00:03.000	Block 04 Energy	Block 04	18.948	88	1.657	0	0	0
2023-12-19 18:00:03.000	Block 07 Energy	Block 07	14.881	80	2.095	0	0	0
2023-12-19 18:00:03.000	Block 08 Energy	Block 08	15.897	100	0	0	0	0
2023-12-19 17:40:03.000	Block 01 Energy	Block 01	19.313	89	1.621	0	0	0
2023-12-19 17:40:03.000	Block 08 Energy	Block 08	15.206	100	0	0	0	0
2023-12-19 17:40:03.000	Block 08 Shade	Block 08	15.206	100	0	0	0	0
2023-12-19 17:40:03.000	Block 04 Shade	Block 04	19.115	91	1.251	0	0	0

<RDBMS에서 관리하는 정제 데이터>

- 데이터 수동 수집
- API로 수집하지 못하는 데이터를 위해 파일 업로드 기능 제공

- 시각화 플랫폼에서 정형 데이터를 업로드하는 경우, Database에 테이블이 자동으로 생성되며, 작물에 대한 링크, 테이블 매핑, 메타 정보가 함께 생성되도록 구성됨

id	tbl_nm	col_nm_eng	col_nm_kor	col_id	col_data_type	col_not_null	col_auto_incr	col_key	col_default	col_sensor	col_descriptio	col_data_rang	create_dt	use_yn	tbl_nm	tbl_description
2.061	tbl_image_meta	id	일련번호	1	int	TRUE	TRUE	PRI	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.062	tbl_image_meta	created_at	생성일	2	datetime	TRUE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.063	tbl_image_meta	updated_at	수정일	3	datetime	TRUE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.064	tbl_image_meta	camera_id	카메라 아이디	4	varchar(50)	TRUE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.065	tbl_image_meta	width	넓이	5	int	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.066	tbl_image_meta	height	높이	6	int	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.067	tbl_image_meta	channel	채널	7	int	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.068	tbl_image_meta	shooting_date	촬영일	8	varchar(10)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.069	tbl_image_meta	shooting_time	촬영시간	9	varchar(50)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.070	tbl_image_meta	save_path	저장소	10	varchar(200)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.071	tbl_image_meta	camera_installation_height	카메라 설치 높이	11	int	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.072	tbl_image_meta	camera_installation_angle	카메라 설치 각도	12	varchar(100)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.073	tbl_image_meta	camera_installation_location	카메라 설치 위치	13	varchar(100)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.074	tbl_image_meta	farm_name	농장명	14	varchar(100)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.075	tbl_image_meta	farm_area	농장 넓이	15	varchar(100)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.076	tbl_image_meta	target_crop	타겟	16	varchar(50)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.077	tbl_image_meta	growth_stage	생장 상태	17	varchar(50)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	비정형데이터 메타 정보
2.078	tbl_aloh4	date	날짜	1	varchar(10)	TRUE	FALSE	PRI	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.079	tbl_aloh4	time	시간	2	datetime	TRUE	FALSE	PRI	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.080	tbl_aloh4	ventilation_temperature	환기 온도	3	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.081	tbl_aloh4	internal_temperature	내부 온도	4	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.082	tbl_aloh4	heating_temperature	난방 온도	5	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.083	tbl_aloh4	outside_temperature	외부 온도	6	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.084	tbl_aloh4	internal_humidity	내부 습도	7	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.085	tbl_aloh4	wind_speed	풍속	8	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.086	tbl_aloh4	insolation	일사량	9	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.087	tbl_aloh4	accumulated_insolation	적사량	10	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.088	tbl_aloh4	rain	강우	11	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.089	tbl_aloh4	internal_temperature_1	내부 온도1	12	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.090	tbl_aloh4	internal_temperature_2	내부 온도2	13	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)				2023-07-27	Y	smartfarm	일로와 농장 데이터(과거분)
2.091	tbl_chili	date	날짜	1	varchar(10)	TRUE	FALSE	PRI	(NULL)				2023-07-27	Y	smartfarm	고추 데이터 (과거분)
2.092	tbl_chili	time	시간	2	datetime	TRUE	FALSE	PRI	(NULL)	시간			2023-07-27	Y	smartfarm	고추 데이터 (과거분)
2.093	tbl_chili	temperature_1	온도 1	3	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)	온도 1			2023-07-27	Y	smartfarm	고추 데이터 (과거분)
2.094	tbl_chili	temperature_2	온도 2	4	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)	온도 2			2023-07-27	Y	smartfarm	고추 데이터 (과거분)
2.095	tbl_chili	temperature_3	온도 3	5	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)	온도 3			2023-07-27	Y	smartfarm	고추 데이터 (과거분)
2.096	tbl_chili	temperature_4	온도 4	6	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)	온도 4			2023-07-27	Y	smartfarm	고추 데이터 (과거분)
2.097	tbl_chili	humidity_1	습도 1	7	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)	습도 1			2023-07-27	Y	smartfarm	고추 데이터 (과거분)
2.098	tbl_chili	humidity_2	습도 2	8	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)	습도 2			2023-07-27	Y	smartfarm	고추 데이터 (과거분)
2.099	tbl_chili	humidity_3	습도 3	9	decimal(10,5)	FALSE	FALSE	(NULL)	(NULL)	습도 3			2023-07-27	Y	smartfarm	고추 데이터 (과거분)

### <업로드 파일-자동생성테이블-자동생성컬럼 매핑 테이블>

codeid	data_type	data_type_nm	corp_type	corp_type_nm	corp_nm	corp_code	corp_code_nm	farm_code	mart_yn	menu_path	use_yn	create_dt	tbl_nm	where_condition
41	1	환경 데이터	1	과채류	1	파프리카	1	block 01	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 01'
42	1	환경 데이터	1	과채류	1	파프리카	2	block 02	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 02'
43	1	환경 데이터	1	과채류	1	파프리카	3	block 03	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 03'
44	1	환경 데이터	1	과채류	1	파프리카	4	block 04	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 04'
45	1	환경 데이터	1	과채류	2	파프리카	1	block 09	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 09'
46	1	환경 데이터	1	과채류	2	파프리카	2	block 10	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 10'
47	1	환경 데이터	1	과채류	2	파프리카	3	block 11	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 11'
48	1	환경 데이터	1	과채류	2	파프리카	4	block 12	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 12'
49	1	환경 데이터	1	과채류	3	토마토	1	block 05	3	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 05'
50	1	환경 데이터	1	과채류	3	토마토	2	block 06	3	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 06'
51	1	환경 데이터	1	과채류	3	토마토	3	block 07	3	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 07'
52	1	환경 데이터	1	과채류	3	토마토	4	block 08	3	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 08'
53	1	환경 데이터	1	과채류	4	토마토	1	block 13	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 13'
54	1	환경 데이터	1	과채류	4	토마토	2	block 14	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 14'
55	1	환경 데이터	1	과채류	4	토마토	3	block 15	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 15'
56	1	환경 데이터	1	과채류	4	토마토	4	block 16	1	Y	/fruitvegetab	2023-07-27	tbl_aggregati	where block = 'block 16'
57	1	환경 데이터	2	엽채류	5	고추	1	May.21	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2021-05%'
58	1	환경 데이터	2	엽채류	5	고추	2	Jun.21	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2021-06%'
59	1	환경 데이터	2	엽채류	6	로메인 상추	1	Jul.21	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2021-07%'
60	1	환경 데이터	2	엽채류	6	로메인 상추	2	Aug.21	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2021-08%'
61	1	환경 데이터	2	엽채류	7	비질	1	Sep.21	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2021-09%'
62	1	환경 데이터	2	엽채류	7	비질	2	Oct.21	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2021-10%'
63	1	환경 데이터	2	엽채류	8	패얼리	1	Nov.21	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2021-11%'
64	1	환경 데이터	2	엽채류	8	패얼리	2	Dec.21	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2021-12%'
65	1	환경 데이터	2	엽채류	9	이자벨	1	Jan.22	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2022-01%'
66	1	환경 데이터	2	엽채류	9	이자벨	2	Feb.22	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2022-02%'
67	1	환경 데이터	2	엽채류	10	스텔리스	1	Mar.22	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2022-03%'
68	1	환경 데이터	2	엽채류	10	스텔리스	2	Apr.22	2	Y	/leafvegetab	2023-07-27	tbl_aloh4	where date like '2022-04%'
69	1	환경 데이터	2	엽채류	11	고추	1		1	Y	/fruitvegetab	2023-07-27	tbl_chili	
84	1	환경 데이터	1	과채류	1	텍스트	(NULL)	(NULL)	1	N	/datacollect	2023-12-26	tbl_datacollec	(NULL)

### <업로드 파일의 작물 코드 데이터>

#### - 집계 데이터 생성

- 1차 집계 테이블(tbl\_aggregation) : 27종의 API연계 데이터 중 필수적으로 필요한 데이터만 선별하여 집계
- 2차 집계 테이블(tbl\_mart) : 수식으로 계산되어야 할 항목을 생성하고, 그래프로 표현하기 위한 항목만 선별하여 제공
- 필수 데이터 : 이슬점 온도/절대 습도/증기압 부족분/엔탈피

#### ① 이슬점 온도

$$e_s = 6.112 * \exp((17.67 * T)/(T + 243.5));$$

$$e = e_s * (RH/100.0);$$

$$\text{이슬점 온도 } T_d = \log(e/6.112) * 243.5 / (17.67 - \log(e/6.112));$$

- T=temperature in Deg C/es=saturation vapor pressure in mb/e=vapor pressure in mb/RH=Relative Humidity in percent/Td=dew point in deg C

② 절대습도

$$e = 6.112 \cdot \exp\left(\frac{17.67 \cdot T_d}{T_d + 243.5}\right);$$

$$q = \frac{0.622 \cdot e}{p - (0.378 \cdot e)}; \quad 1013.25$$

- e=vapor pressure in mb/Td=dew point in deg C/p=surface pressure in mb/q=specific humidity in kg/kg

③ 증기압부족분

- 식물의 성장과 생산에 중요한 요소이며, 공기가 포화되었을 때 포함할 수 있는 수분량과 공기 중의 수분량 사이의 차이를 계산. 공기중의 포화 수증기 압력은 온도에 의해 정해진다.

$$SVP = 0.6108 \cdot \exp\left(\frac{17.27T}{T + 237.3}\right) [kPa]$$

- 현재 온도에서 공기 중의 수증기 압력 활성 증기압이라고 하며 온 상대 습도에 의해 결정

$$AVP = SVP \cdot \frac{RH}{100} \quad VPD = SVP - AVP$$

- VPD는 SVP와 AVP의 차이로 계산

④ 엔탈피

- 공기의 열 용량 즉 엔탈피는 두가지로 구성
- 온도계로 측정하는 분자의 운동 에너지(건조 공기의 엔탈피와 증발된 수분이 포함되어 있는 잠열)
- 모든 공기는 작은 양일지라도 수분을 포함하고 있기 때문에 공기의 엔탈피를 계산하는 경우 두 가지 요인을 모두 고려

$$\text{Air enthalpy} = 1.007 \cdot \text{Temperature} - 0.026 [kJ/kg]$$

- 수증기의 엔탈피는 공기중의 수분의 양과 공기의 온도를 이용하여 구한다.

$$\text{Water vapor enthalpy} = \text{water content of air} \times (2501 + 1.84 \cdot \text{temperature})$$

- 전체 엔탈피는 두가지 요소를 더해서 구한다

$$\text{Enthalpy} = \text{Air enthalpy} + \text{Water vapor enthalpy}$$

- 수식 생성을 위한 정리 작업

- ① 쿼리 생성을 위한 순차적인 수식 정의

이슬점 온도	$es = 6.112 \cdot \exp((17.67 \cdot T)/(T+243.5))$ $e = es \cdot (RH/100.0)$ $Td = \log(e/6.112) \cdot 243.5 / (17.67 - \log(e/6.112))$	T = temperature in de c es. = saturation vapor pressure in mb e = vapor pressure in mb
절대 습도	$e = 6.112 \cdot \exp(17.67 \cdot Td) / (Td + 243.5)$ $q = (0.622 \cdot e) / (p - (0.378 \cdot e))$ $q = q \cdot 1000$	RH = relative humidity in percent Td = dew point in deg C p = surface pressure in mb q = specific humidity in kg/kg.
증기압 부족분	$SVP = 0.6108 \cdot \exp(17.27 \cdot T / (T+237.3))$ AVP = SVP * (RH / 100) VPD = SVP - AVP	
Enthalpy	air enthalpy = 1.007 * T - 0.026 Water vapor enthalpy = q * (2501 + 1.84 * T) Enthalpy = air enthalpy + water vapor entalpy	

<수식 정의>

② 관련 테이블과 컬럼 정의

차트 항목	알로라	고수	표면인상수	태일	배일리	미차별	스텔릭스	종양	파프리카	토마토 (홍양)	토마토 (태라)	교주	말기	
실내	실내온도	internal_temperature						air temperature	block1-4, 9-12	block 13-16	block 5-8	temperature	tp	gh_indoor_env_info
	실내습도	internal_humidity						humidity				humidity_1	hd	gh_indoor_env_info
	이슬점온도	dew_point_temperature						dew point temperature						
	수분부족분(air)	absolute_humidity						absolute humidity						
	수분부족분(water)	humidity_deficit						humidity deficit						
	VPD	vpd						vapour pressure deficit of greenhouse air	tbi_ws_sclu_block_cooling			vpd	vpd	gh_indoor_env_info
	Enthalpy	enthalpy						enthalpy				enthalpy		
실외	외부온도	outside_temperature						outside temperature	tb_ws_sclu_weater_stations			temperature	tp	gh_outdoor_env_info
	외부습도	outside_humidity						humidity				humidity_1	hd	gh_outdoor_env_info
	풍향	wind_direction						wind direction	tb_ws_sclu_weater_stations			wind_direction	wd	gh_outdoor_env_info
	풍속	wind_speed						wind speed	tb_ws_sclu_weater_stations			wind_speed	ws	gh_outdoor_env_info
	일사량	irradiation						radiation intensity	tb_ws_sclu_weater_stations			irradiation	ir	gh_outdoor_env_info
	누적 일사량	accumulated_irradiation						radiation sum	tb_ws_sclu_weater_stations			irradiation_sum	is	gh_outdoor_env_info
	강수량	rain						rain	tb_ws_sclu_weater_stations			rain	rc	gh_outdoor_env_info
급액/배액	배액 수분량	measured_water_out						measured water out	tb_ws_sclu_irrigation_valve_data			ec_1	ec	gh_indoor_env_info
	급액 EC	measured_ec_in						measured ec in	tb_ws_sclu_irrigation_valve_data			ec_1	ec	gh_indoor_env_info
	급액 pH	measured_ph_in						measured ph in	tb_ws_sclu_irrigation_valve_data			ph_1	ph	gh_indoor_env_info
	급액 온도	measured_temperature_in						measured temperature in	tb_ws_sclu_irrigation_valve_data			temperature_in	ti	gh_indoor_env_info
	급액 수	measured_flow_in						measured flow in	tb_ws_sclu_irrigation_valve_data			flow_in	fi	gh_indoor_env_info
	급액 시작	start_time						start time	tb_ws_sclu_irrigation_valve_data			start_time	st	gh_indoor_env_info
	급액 종료	end_time						end time	tb_ws_sclu_irrigation_valve_data			end_time	et	gh_indoor_env_info
	배액 EC	measured_ec_out						measured ec out				ec_2	ec	gh_indoor_env_info
	배액 pH	measured_ph_out						measured ph out				ph_2	ph	gh_indoor_env_info
	배액 온도	measured_temperature_out						measured temperature out				temperature_out	to	gh_indoor_env_info
	배액 수	measured_flow_out						measured flow out				flow_out	fo	gh_indoor_env_info
	배액 시작	start_time_out						start time out				start_time_out	sto	gh_indoor_env_info
	배액 후 누적일사량	accumulated_irradiation_out						accumulated irradiation out				irradiation_sum_out	is_o	gh_indoor_env_info

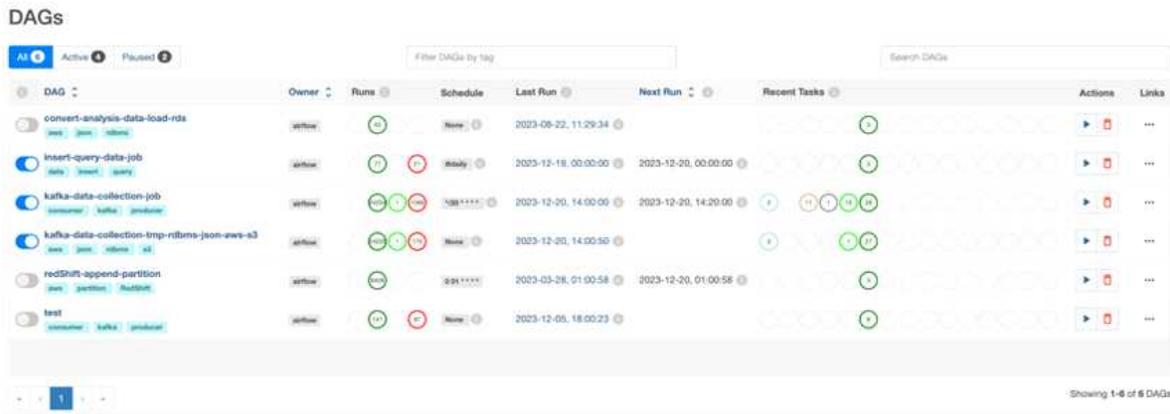
<수식 계산을 위한 필수 변수 및 테이블 정보>

③ 쿼리 수행 후 데이터 적재

④ 시각화 플랫폼 차트로 표현

- 데이터 수집을 위한 자동화 배치 기능

· 본 과제에서는 데이터 수집 및 데이터 마트 구축을 위한 자동화 시스템으로 Airflow를 활용



<Airflow 활용 화면>

· 총 6개의 DAG 중에서 3개가 실행 중에 있으며, 각 다음과 같은 역할을 수행

항목	설명
Kafka-data-collection-job	API 데이터를 수집하고 테이블로 저장하는 스케줄러
Kafka-data-collection-tmp-rdbms-json-aws-s3	수집된 데이터를 1차 집계하여 aggregation 테이블로 저장하는 스케줄러





- ② 모델 개발 및 훈련
- ③ 모델 검증
- ④ 모델 패키징
- ⑤ CI/CD 통합
- ⑥ 모델 배포

○ 지능형 모델 탑재 및 운영

- 본 플랫폼에서 스마트 온실 데이터 분석을 위해 총 3개의 모델을 제공

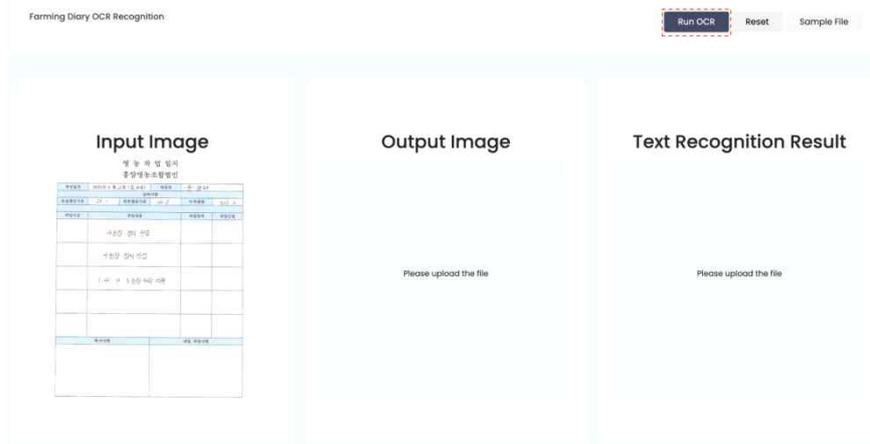
- 경영일지 OCR 인식

① 데이터 업로드

· 원본 이미지를 업로드 하면, 데이터레이크에 적재가 되며, 웹 화면에 미리보기로 표현



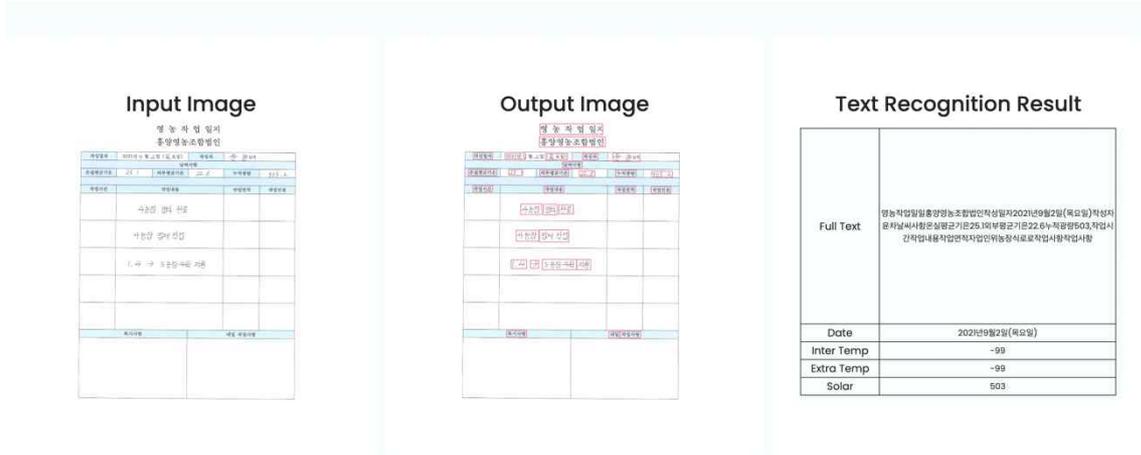
<이미지 업로드 화면>



<미리보기 화면>

② 인식 모델 실행 및 결과물 저장

· 경영일지에서 텍스트 추출이 완료되면 결과 이미지를 생성하고 데이터레이크에 저장이 되며, 추출된 텍스트는 RDBMS로 저장



<추출 결과물>

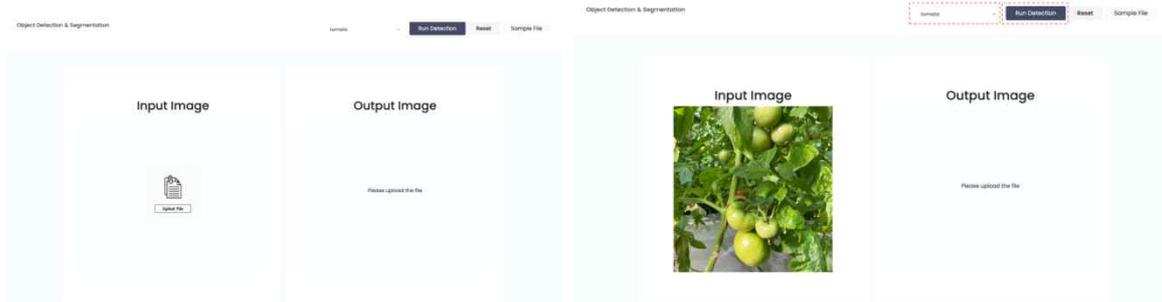
ocr_id	full_text	date	temp_avg	temp_avg_out	light	etc
1	영농직업일일홍양영농조합법인작성일자2021년월1일(수요일)작성자윤창혜배은(	[NULL]	27	25	727	영농직업일일홍양영농조합법인2021년월1일(수요일)작성자윤창혜배은(
2	영농직업일일홍양영농조합법인작성일자2021년월1일(수요일)작성자윤창혜배은(	[NULL]	27	25	727	영농직업일일홍양영농조합법인2021년월1일(수요일)작성자윤창혜배은(
3	영농직업일일홍양영농조합법인작성일자2021년월1일(수요일)작성자윤창혜배은(	[NULL]	27	25	727	영농직업일일홍양영농조합법인2021년월1일(수요일)작성자윤창혜배은(
4	영농직업일일홍양영농조합법인작성일자2021년월1일(수요일)작성자윤창혜배은(	[NULL]	27	25	727	영농직업일일홍양영농조합법인2021년월1일(수요일)작성자윤창혜배은(
5	영농직업일일홍양영농조합법인작성일자2021년월1일(수요일)작성자윤창혜배은(	[NULL]	27	25	727	영농직업일일홍양영농조합법인2021년월1일(수요일)작성자윤창혜배은(
6	영농직업일일홍양영농조합법인작성일자2021년9월15일(수요일)작성자윤창혜배은(	2021년9월15일(수요일)	23	22	1,338	영농직업일일홍양영농조합법인2021년9월15일(수요일)작성자윤창혜
7	영농직업일일홍양영농조합법인작성일자2021년9월2일(목요일)작성자윤창혜배은(	2021년9월2일(목요일)	-99	-99	503	영농직업일일홍양영농조합법인2021년9월2일(목요일)작성자윤창혜배은(
8	영농직업일일홍양영농조합법인작성일자2021년월1일(수요일)작성자윤창혜배은(	[NULL]	27	25	727	영농직업일일홍양영농조합법인2021년월1일(수요일)작성자윤창혜배은(
9	영농직업일일홍양영농조합법인작성일자2021년월1일(수요일)작성자윤창혜배은(	[NULL]	27	25	727	영농직업일일홍양영농조합법인2021년월1일(수요일)작성자윤창혜배은(
10	영농직업일일홍양영농조합법인작성일자2021년월1일(수요일)작성자윤창혜배은(	[NULL]	-99	23	-99	영농직업일일홍양영농조합법인2021년월1일(수요일)작성자윤창혜배은(

<추출 결과물>

• 작물 탐지 모델

① 데이터 업로드

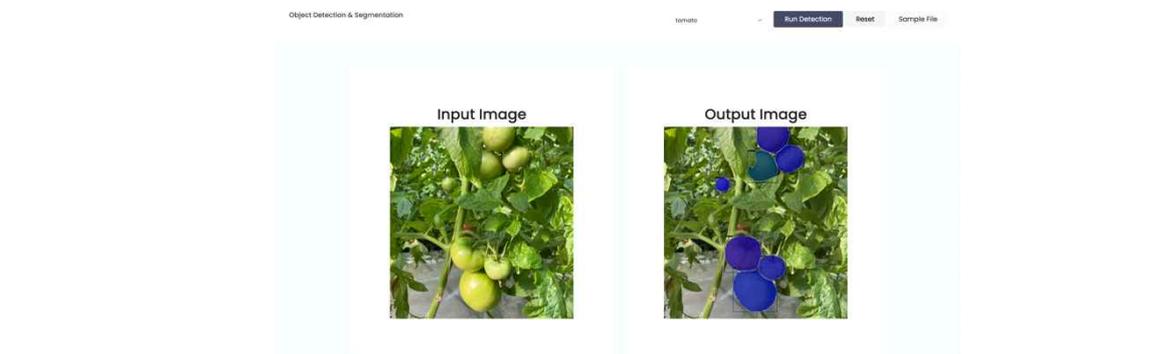
· 원본 이미지를 업로드 하면, 데이터레이크에 적재가 되며, 웹 화면에 미리보기로 표현



<이미지 업로드 화면>

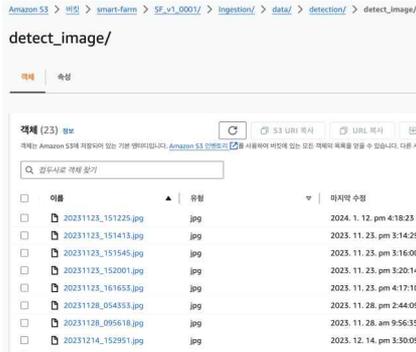
<미리보기 화면>

② 탐지 모델 실행



③ 탐지 결과물

- 작물 탐지가 완료되면 결과 이미지를 생성하고 데이터레이크에 저장이 되며, 저장경로가 메타 테이블에 저장



<결과 이미지 저장 데이터레이크>

<합성 결과 이미지>

detect_id	original_image	detect_image	detect_json	code_id
1	SF_v1_0001/Ingestion/data/detection/original_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_json/IMG_0983	1
2	SF_v1_0001/Ingestion/data/detection/original_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_json/IMG_0983	1
3	SF_v1_0001/Ingestion/data/detection/original_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_json/IMG_0986	1
4	SF_v1_0001/Ingestion/data/detection/original_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_json/IMG_0987	1
5	SF_v1_0001/Ingestion/data/detection/original_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_json/IMG_0988	1
6	SF_v1_0001/Ingestion/data/detection/original_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_json/IMG_0991	1
7	SF_v1_0001/Ingestion/data/detection/original_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_image/IMG_09	SF_v1_0001/Ingestion/data/detection/detect_json/IMG_0991	1
8	SF_v1_0001/Ingestion/data/detection/original_image/IMG_09	[NULL]	[NULL]	1
9	SF_v1_0001/Ingestion/data/detection/original_image/IMG_10	SF_v1_0001/Ingestion/data/detection/detect_image/IMG_10	SF_v1_0001/Ingestion/data/detection/detect_json/IMG_1013	1
10	SF_v1_0001/Ingestion/data/detection/original_image/202311	SF_v1_0001/Ingestion/data/detection/detect_image/202311	SF_v1_0001/Ingestion/data/detection/detect_json/20231123	1
11	SF_v1_0001/Ingestion/data/detection/original_image/202311	SF_v1_0001/Ingestion/data/detection/detect_image/202311	SF_v1_0001/Ingestion/data/detection/detect_json/20231123	1
12	SF_v1_0001/Ingestion/data/detection/original_image/202311	SF_v1_0001/Ingestion/data/detection/detect_image/202311	SF_v1_0001/Ingestion/data/detection/detect_json/20231123	1
13	SF_v1_0001/Ingestion/data/detection/original_image/202311	SF_v1_0001/Ingestion/data/detection/detect_image/202311	SF_v1_0001/Ingestion/data/detection/detect_json/20231123	1
14	SF_v1_0001/Ingestion/data/detection/original_image/202311	SF_v1_0001/Ingestion/data/detection/detect_image/202311	SF_v1_0001/Ingestion/data/detection/detect_json/20231123	1
15	SF_v1_0001/Ingestion/data/detection/original_image/202311	SF_v1_0001/Ingestion/data/detection/detect_image/202311	SF_v1_0001/Ingestion/data/detection/detect_json/20231128	1
16	SF_v1_0001/Ingestion/data/detection/original_image/202311	SF_v1_0001/Ingestion/data/detection/detect_image/202311	SF_v1_0001/Ingestion/data/detection/detect_json/20231123	1
17	SF_v1_0001/Ingestion/data/detection/original_image/202312	[NULL]	[NULL]	

<결과물 저장 경로를 관리하는 메타 테이블>

· 병충해 합성 모델

① 데이터 업로드

- 원본 이미지를 업로드 하면, 데이터레이크에 적재가 되며, 웹 화면에 미리보기로 표현



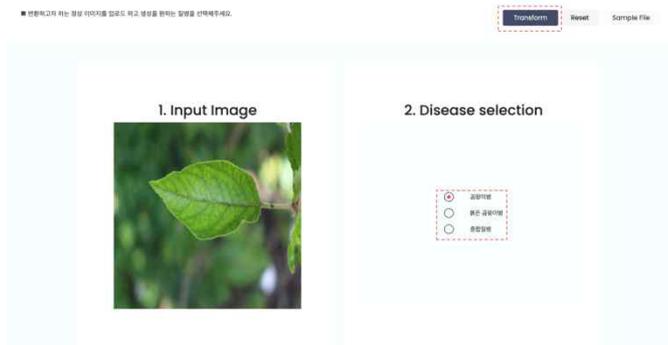
<이미지 업로드 화면>



<미리보기 화면>

② 합성 모델 실행

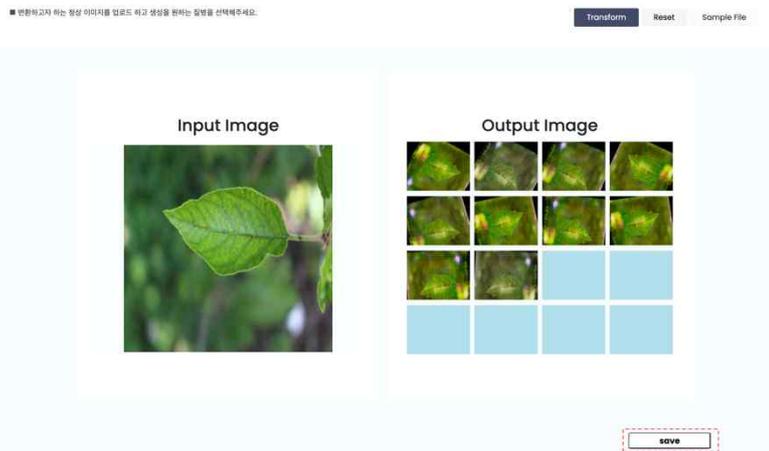
- 합성하고자 하는 질병을 선택하여 모델을 실행



<합성하고자 하는 질병 선택>

③ 합성 결과 출력

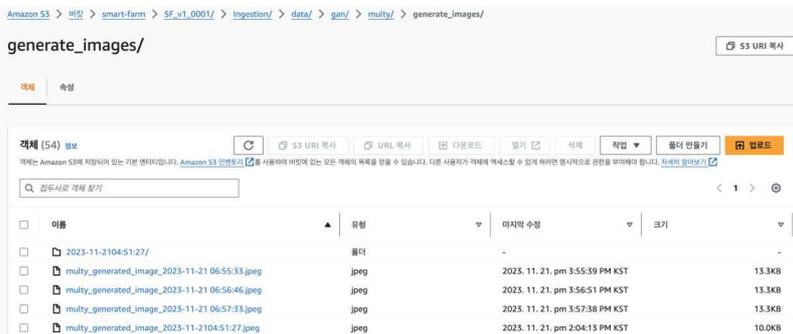
- 합성을 완료하면 결과 이미지는 데이터레이크의 결과 폴더에 저장이 되며, 그 경로는 메타 데이터로 RDBMS에 저장



<합성 결과>

gan_id	original_image	generate_image	aug_images_folder
1	SF_v1_0001/Ingestion/data/gan/original_images/Test_57.jpg	SF_v1_0001/Ingestion/data/gan/muly/generate_images/muly_generat	SF_v1_0001/Ingestion/data/gan/muly/aug_images/2023-11-21 05:
2	SF_v1_0001/Ingestion/data/gan/original_images/Test_10.jpg	SF_v1_0001/Ingestion/data/gan/muly/generate_images/muly_generat	SF_v1_0001/Ingestion/data/gan/muly/aug_images/2023-11-21 06:
3	SF_v1_0001/Ingestion/data/gan/original_images/Test_10.jpg	SF_v1_0001/Ingestion/data/gan/rust/generate_images/rust_generat	SF_v1_0001/Ingestion/data/gan/rust/aug_images/2023-11-23 04:3
4	SF_v1_0001/Ingestion/data/gan/original_images/Test_10.jpg	SF_v1_0001/Ingestion/data/gan/scab/generate_images/scab_generat	SF_v1_0001/Ingestion/data/gan/scab/aug_images/2023-11-23 04:
5	SF_v1_0001/Ingestion/data/gan/original_images/Test_57.jpg	SF_v1_0001/Ingestion/data/gan/scab/generate_images/scab_generat	SF_v1_0001/Ingestion/data/gan/scab/aug_images/2023-11-27 01:
6	SF_v1_0001/Ingestion/data/gan/original_images/Test_57.jpg	SF_v1_0001/Ingestion/data/gan/scab/generate_images/scab_generat	SF_v1_0001/Ingestion/data/gan/scab/aug_images/2023-11-27 01:
7	SF_v1_0001/Ingestion/data/gan/original_images/Test_10.jpg	SF_v1_0001/Ingestion/data/gan/scab/generate_images/scab_generat	SF_v1_0001/Ingestion/data/gan/scab/aug_images/2023-11-27 01:
8	SF_v1_0001/Ingestion/data/gan/original_images/Test_165.jpg	SF_v1_0001/Ingestion/data/gan/scab/generate_images/scab_generat	SF_v1_0001/Ingestion/data/gan/scab/aug_images/2023-11-27 01:
9	SF_v1_0001/Ingestion/data/gan/original_images/Test_233.jpg	SF_v1_0001/Ingestion/data/gan/scab/generate_images/scab_generat	SF_v1_0001/Ingestion/data/gan/scab/aug_images/2023-11-28 00:
10	SF_v1_0001/Ingestion/data/gan/original_images/Test_452.jpg	SF_v1_0001/Ingestion/data/gan/scab/generate_images/scab_generat	SF_v1_0001/Ingestion/data/gan/scab/aug_images/2023-11-28 00:
11	SF_v1_0001/Ingestion/data/gan/original_images/Test_671.jpg	SF_v1_0001/Ingestion/data/gan/scab/generate_images/scab_generat	SF_v1_0001/Ingestion/data/gan/scab/aug_images/2023-11-28 00:
13	SF_v1_0001/Ingestion/data/gan/original_images/20231127/20231127_3!	[NULL]	[NULL]
14	SF_v1_0001/Ingestion/data/gan/original_images/20231127/20231127_!	SF_v1_0001/Ingestion/data/gan/rust/generate_images/rust_generat	SF_v1_0001/Ingestion/data/gan/rust/aug_images/2023-11-27 06:2
15	SF_v1_0001/Ingestion/data/gan/original_images/20231127/20231127_!	SF_v1_0001/Ingestion/data/gan/rust/generate_images/rust_generat	SF_v1_0001/Ingestion/data/gan/rust/aug_images/2023-11-27 06:0
16	SF_v1_0001/Ingestion/data/gan/original_images/20231127/20231127_!	SF_v1_0001/Ingestion/data/gan/rust/generate_images/rust_generat	SF_v1_0001/Ingestion/data/gan/rust/aug_images/2023-11-27 06:0

<데이터 저장 경로에 대한 메타 데이터>



<결과물 저장 데이터레이크>



<생성 이미지>

○ 인공지능 운영/배포

- 3종의 인공지능 모델은 MLOps 프로세스에 따라 9개의 과정으로 관리

① 데이터 분석 및 데이터 준비

- OCR 인식 모델을 위해 한글데이터, 형태소 모음, 구글 테스터 데이터를 수집하고 dvc로 관리
- 작물 탐지를 위해 laboro-tomato 데이터를 수집하여 dvc로 관리
- 병충해 합성을 위해 일반 식물 데이터와 3가지 병충해 식물 데이터를 수집하고 dvc로 관리

② 모델 훈련

- 머신러닝 프레임워크 pytorch를 통해 모델 훈련, 데이터, 모델 아키텍처, 하이퍼파라미터 튜닝 등을 통해 훈련

③ 모델 직렬화

- Pytorch 모델의 .pt, .pth 파일 형식으로 직렬화를 수행하고 저장

④ 컨테이너화

- 직렬화된 모델은 kubernetes 컨테이너에 배치
- 이 컨테이너에는 모델 뿐만 아니라 모델을 실행하는데 필요한 모든 의존성, 라이브러리 및 환경이 포함되면, 모델이 다양한 컴퓨팅 환경에서 일관되게 실행되도록 보장

⑤ Kubeflow 컴포넌트 생성

- Kubeflow에서 모델은 일반적으로 컴포넌트로 패키징
- Kubeflow 컴포넌트는 ML 워크플로우의 한 단계를 수행하는 자체 포함된 코드 세트이며, 이 단계에서 컴포넌트의 입력, 출력 및 사용할 컨테이너 이미지를 정의하는 컴포넌트 사양을 작성

⑥ 파이프라인 구축

- 컴포넌트를 생성 한 후에는 Kubeflow 파이프라인을 구축한다. 파이프라인은 ML 워크플로우의 단계들이 순차적으로 연결

⑦ 모델 배포

- 축된 파이프라인을 통해 모델을 배포

⑧ 모니터링 및 로깅

- Prometheus/Grafana를 통해 모니터링 및 로깅을 수행

⑨ 버전 관리 및 모델 레지스트리

- 위의 과정들을 기반으로 버전을 관리

□ 빅데이터 플랫폼 제품화

○ 빅데이터 플랫폼 제품화

- 빅데이터 플랫폼의 활용성 검증을 위해 제품화 2건을 구성
- 본 연구에서 구축한 빅데이터 플랫폼 API연동을 통해 자동으로 데이터를 수집하고 시각화하는 기능을 기반으로 하여 데이터 분석을 위한 다양한 기능을 제공
- 빅데이터 플랫폼의 농가의 환경과 활용도에 따라 커스터마이징 가능하며, 인터넷망과 폐쇄망에서 데이터를 관리하는 플랫폼으로 이중화

○ 인터넷 망에서의 빅데이터 플랫폼 제품화

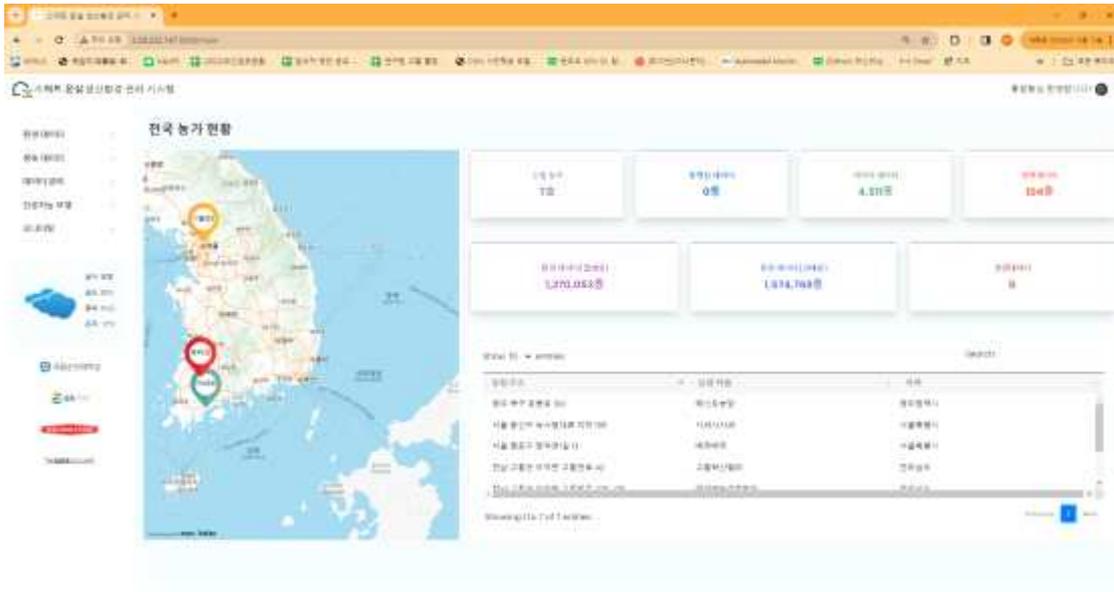
- 농가에서 관리하는 관리 시스템의 API를 통하여 데이터를 수집하고 개별적으로 관리하는 파일들을 업로드 할 수 있는 기능을 제공
- 수집된 데이터들을 활용하여 비교 분석할 수 있는 작기별/작물별 분석 차트와 상관 분석 기능 제공
- 데이터 기반의 인공지능 모델 제공으로 분석 효과 극대화
- 실시간 데이터 수집을 위한 인프라와 서비스 모니터링 기능 제공

○ 폐쇄망에서의 빅데이터 플랫폼 제품화

- 폐쇄망 내에 연계되어 있는 농가의 데이터를 수집/관리
- 폐쇄망 내부 스케줄러가 동작하며, 타 시스템과 연동을 통해 데이터를 수집하며 외부의 불필요한 데이터는 마이그레이션 됨
- 모니터링/인공지능 모델 등 무거운 리소스를 차지하는 서비스를 개선

○ 차별성 및 기대 효과

- 본 빅데이터 플랫폼은 Public Cloud 기반으로 설계/구축되었으며, 활용성 증대를 위해 다양한 접속 환경에서 데이터를 수집하거나 전송할 수 있는 사용자 접근성에 중점을 두고 설계
- 또한 Private 한 환경에서도 활용이 가능할 수 있도록 불필요한 보안 및 관리 도구 연계 부분을 삭제하여 비교적 낮은 시스템 리소스로 이중 설계되었으며, 기존 CDN방식의 연계 라이브러리 및 패키지 등을 로컬 인프라 설치형으로 변경하였으며, 내부 백그라운드 서비스를 폐쇄망 형으로 이중 구조로 변경
- 외부 API 연계를 위한 다양한 모니터링 서비스와 리소스를 소요하고 있으나, 폐쇄망 내 빅데이터 플랫폼은 데이터 임의 복제/반출 불가능한 데이터 취급 보안 정책에 따라 데이터베이스를 일부 공유
- 데이터 스토리지에 대한 별도 제원 요구 및 데이터 수집 속도와 분석을 위한 전처리 요소를 최적화
- 2건의 제품화를 통해 수집되는 환경 및 제어 데이터를 수집/분석/시각화하여 생산환경 관리를 위한 컨설팅 자료로 활용 가능하며, 생산성 향상을 위한 업무 운영 매뉴얼의 기초자료로 활용이 가능
- 데이터 표준화를 통해 온실 생육, 환경, 영농 등에 대한 정보를 규격화하여 데이터의 품질 및 해당 정보의 정확성을 제고하고 스마트 온실 생산환경 관리를 위해 일관성 있는 분석 자료 구축에 활용 가능할 것으로 판단 됨



< 인터넷 망에서의 빅데이터 플랫폼 제품화 >

The screenshot displays a web interface with two main data tables. The top table is titled '산채농장' and the bottom table is titled '산채농장' (likely a typo for '산채농장'). Both tables show columns for dates, various agricultural metrics, and numerical values. The interface includes a navigation menu on the left and a search bar at the top.

< 폐쇄망에서의 빅데이터 플랫폼 제품화 >

## [홍양영농조합]

### [1단계(2021)]

#### □ 커넥티드 팜 데이터 수집 및 실증 사이트 그룹 선정

##### ○ 과채류(토마토, 파프리카) 실증 사이트 데이터 수집 및 재배 일지 작성

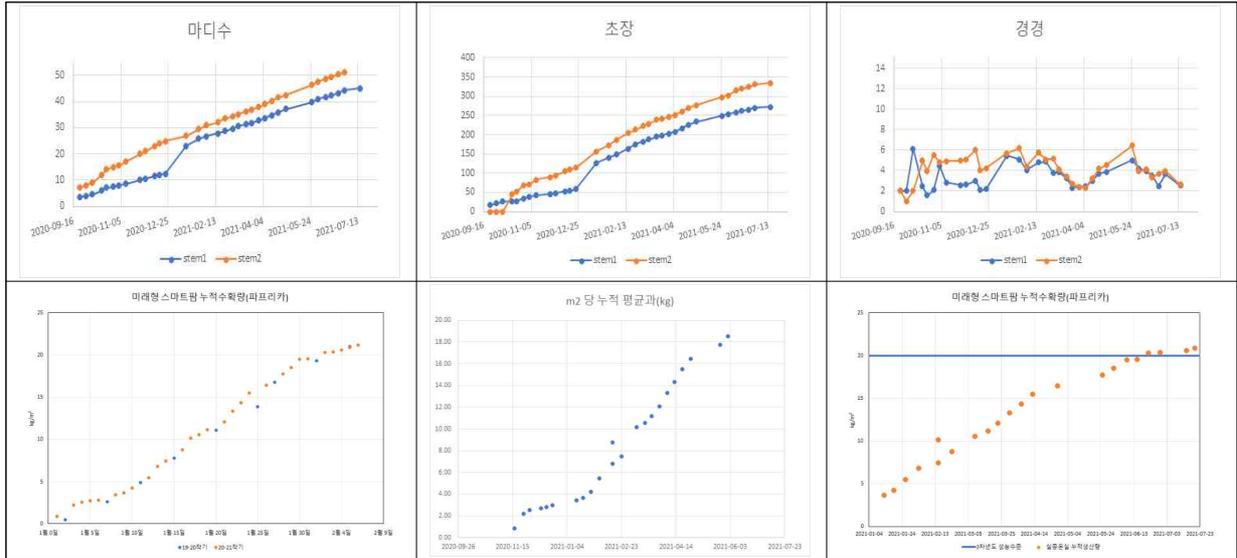
- 토마토(반밀폐온실) 실증 사이트 재배 이슈
  - 21년 06월 10일 토마토(품종 : 레드 250) 파종 시작
  - 21년 06월 23일 이식 -> 07월 02일 가정 -> 07월 20일 정식 시작
  - 재배 33주차 부터 상부 성장점과 결순이 보라 빛을 띄며 생식생장 기질이 보이기 시작



- 1~ 3번까지 착과된 모습 보이나 4번은 아직 착과 안됨
- 3번 화방의 경우 7개 중 1~2개 착과 수준임
- 나머지 꽃들은 톤장애로 추측되어 시들어버린 것이 20% 보임
- 레드칸 샘플과 레드250의 1~3번 착과수준 비교해보면 레드칸이 더 많음
- 과일에 갈라짐 증상이 조금 보임, 꼭지부분이 갈변하는 현상 발생
- 원인으로 천창 차광제도포로 인한 광 투과율 40% 수준 저하 및 차광 스크린도 일부 사용함
- 또한, 생식 및 착과유도를 위한 관수량 제한 등으로 상부 세력이 약해진 것으로 파악
- 해결책으로 점진적으로 온실 24시간 평균온도를 하락시키고, 관수를 충분히 공급하며 관수 EC를 서서히 3.0까지 하락 시킴
- 재배 37주차 상부 세력이 서서히 잡히기 시작하며 온실 일평균 22도 수준으로 가을 날씨에 접어듦
- 7화방 개화가 완료 되었고, 8화방에 꽃이 피기 시작하며 야간 냉방을 통해 최대한 온도를 낮추어

세력을 높이기 위한 노력 진행

- 과채류(토마토, 파프리카) 실증 사이트 생육 현황
  - 2020.09.23. ~ 2021.07.15. 일간 7일 단위로 수기로 입력
  - 마디수, 초장, 경경, 생장점, 개화수, 착과수, 수확수, 수확무게, 수확일 등 수집



□ 채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 설계

- 온실 환경데이터 수집 및 작물 생육 상태 확인을 위한 이미지 수집용 카메라 설치
  - 순천대학교와의 연구 협업을 통해 현재 시스템의 데이터를 한달 주기로 순천대학교 서버로 백업 될 수 있도록 작업 진행
  - 기존 데이터의 경우 별도의 하드에 백업되어 있어, 확인을 위해선 온실에 구축되어 있는 SYNOPTA 프로그램에 연동하여 진행해야 하는 문제점이 있어 순천대와 같이 데이터를 보다 빨리 확보 할 수 있는 방안 모색 중
  - 추가로 수집이 필요한 배액데이터(배액 EC, pH, 배액온도 등) 확보를 위해 순천대학교에 데이터 수집 시스템 설계 요청 및 공동 연구 진행
  - 순천대학교(공동)와 연구 협업을 통해 카메라 위치, 화질, 설치 장소 등을 고려하여 작물 생육 상태 수집을 위한 카메라 설치 진행

[1단계(2022)]

□ 커넥티드 팜 데이터 수집 및 실증 사이트 구축

○ 실증 사이트를 통한 실시간 데이터 수집 환경 구축

- 과채류 생육데이터 수집을 위한 API 설치
- 온실 생육 의사 결정 지원을 위한 양액 배액량 및 성분 측정 설비 구축
- 알로하 채소농장 엽채류 생육데이터 수집용 영상수집 환경 구축
  - 로보뷰P3

CPU	화소	해상도	영상코딩	화각	렌즈	무선규격
Hisilicon	FULL HD 300만 화소	2304×1296	H265	상하 90° 좌우 355°	2.8~12mm	WiFi 802.11 b/g/n

□ 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 환경 구축

○ 과채류 3종(파프리카, 토마토, 고추), 엽채류 6종(고수, 바질, 로메인상추, 페얼리, 이자벨, 스텔릭스) 빅데이터 수집 현황

- 흥양영농조합, 테라영농조합법인, 알로하채소농장 3개의 사이트에서 총 9개 작목의 환경·제어 데이터를 수집
  - 데이터 형태 : JPEG, CSV, JSON, QRP
  - 수집기간 : 2020.06.~현재(실시간 데이터 수집 진행 중)
  - 데이터 수집 용량 : 1.39GB(2022.11.기준)
- 과채류 2종(파프리카, 토마토)
  - 수집처 : 흥양영농조합법인, 테라영농조합법인
  - 수집기간 : 2020.06.~현재(실시간 데이터 수집 진행 중)
  - 데이터 제공 형태 : CSV 및 JSON

수집 데이터 항목		
항목	소스	요약
Sensor unit	환경 센서	온실 내부 환경 데이터(온도, 습도, 이슬점, 엔탈피등)
Weaterstations	외부 환경 센서	외부 환경 데이터(외부 온도, 풍속, 풍향, 감우등)
CO2 group	CO2 센서	CO2 현황 및 제어 관련 데이터(CO2농도, 설정값, 환기에 따른 CO2 영향값등)
General	보일러	버너 및 soil 보일러의 전반적인 데이터(보일러 온도, 밸브 개방율, 보일러 상태)
Block heating	시스템	온실 냉/난방을 위한 설정값및 영향 데이터
Block cooling		(냉/난방 설정점, 외부 복사에 의한 냉/난방 설정점, 블록별냉/난방 설정값, 현재 냉/난방 상태 등)
Block ventilation	시스템	환기제어를 위한 설정값및 영향 데이터 (환기 설정점, 외부 온도에 의한 영향, 바람에 의한 영향 등)
Top circuit	냉/온수 파이프	냉/온수 공급 파이프 상태 데이터
Bottom circuit		(파이프 온도, 습도로 인한 파이프 온도 영향, 오버라이드활성 상태 등)
cooling bottom circuit		
Burner	보일러	보일러 상태 데이터(보일러 종류, 보일러 출력, 밸브 개방율, 보일러 상태 등)
Soil heating	시스템	배지 난방을 위한 설정 데이터(난방 설정값, 공급 펌프 상태, 난방에 따른 온도 계산 값 등)
Heat buffer	축열조	냉/온수 축열조데이터(수위별수온, 수위, 축열조상태, 저장 열량값 등)
Cold buffer		

Drain water	양액기	양액관련 데이터(EC, pH, 누적 관개량, 누적 배수량 등)
Irrigation valve data		양액공급 관련 데이터(공급 주기, 공급 시작 시간, 관개 밸브 상태, 최대/평균/최소 EC, 최대/평균/최소 pH 등)
Heat pump general	히트 펌프	히트펌프 데이터(시스템 상태, 냉/난방 출력량, 작동 시간 등)
Water tanks	물탱크	물탱크 상태 데이터(공급/반환 상태, 밸프개방률, 수위, 용량 등)
Blackout screen	스크린	암전 상태에서의 스크린 작동 데이터(스크린 제어 상태, 암전 지속 시간, 암전에 따른 온도 및 습도 차이 등)
Shading screen		음영 상태에서의 스크린 작동 데이터(복사열에 따른 스크린 제어 상태, 음영에 의한 온도 영향, 음영 스크린에 의한 온도 차이 등)
Thermal screen		열에 따른 스크린 작동 데이터(복사열에 따른 스크린 개방율, 과도한 복사열 제거를 위한 온도/습도 차이 계산값등)
Ventilation leeward side	환기 시스템	환기 시스템 위치별상태 데이터 (통풍구/환기구상태, 바람 방향에 따른 온도 영향, 환기 제어 상태 등)
Ventilation side 1		
Ventilation side 2		
Ventilation windward side	환기 팬	환기용 팬 그룹 상태 데이터(팬 그룹 별 활성화 상태, 비활성 상태 이유, 제어되는 팬 그룹 등)
Convector fan group		
Recirculation fan group	순환 팬	순환용 팬 그룹 상태 데이터(팬 그룹 별 활성화 상태, 비활성 상태 이유, 제어되는 팬 그룹 등)

• 과채류 1종(고추)

- 수집처 : 흥양영농조합법인(외부데이터)
- 수집기간 : 2020.01.~2022.06.
- 데이터 제공 형태 : CSV

수집 데이터 항목		
항목	소스	요약
온도	온도 센서(5개)	5개 구역별 온실 내부 온도 상태 데이터
습도	습도센서(5개)	5개 구역별 온실 내부 습도 상태 데이터
CO2	CO2(5개)	5개 구역별 온실 내부 CO2 상태 데이터
일사량	일사량 센서	일사량 데이터
풍속	풍속 센서	풍속 데이터
EC	양액기	공급 양액EC, pH 데이터
pH		

• 엽채류 6종(고수, 바질, 로메인상추, 이자벨, 스텔릭스, 페얼리)

- 수집처 : 알로하채소농장
- 수집기간 : 2021.05.~2022.04.(작물별 2개월의 데이터)

- 데이터 제공 형태 : QRP

수집 데이터 항목		
항목	소스	요약
내부온도	온도 센서(2개)	2개 구역별 온실 내부 온도 상태 데이터
내부습도	습도 센서(2개)	2개 구역별 온실 내부 습도 상태 데이터
외부 온도	외부 환경 센서	외부 온도 데이터
풍속		풍속 데이터
일사량		일사량 데이터
적산일사량		일일 누적 일사량 데이터
감우	감우 센서	감우상태 데이터
난방온도	시스템	난방 및 환기를 위한 온도 셋팅 데이터
환기온도		

[2단계(2023)]

□ 무선 통신 기반 커넥티드 팜 및 실증 사이트 운영 및 스마트팜 표준화 데이터 수집

○ 실증사이트 운영 및 스마트팜 표준화 데이터 수집 진행

- 스마트팜 R&D 빅데이터 플랫폼에서 정의한 데이터 표준을 기반으로 데이터 정제
- 스마트팜연구개발사업단에서 개발한 빅데이터 R&D 플랫폼 활용
  - 정형데이터 표준화를 위해 단체표준 및 국가표준, 2차년도 연구데이터 항목을 토대로 스마트팜 표준화 데이터 설정
  - 과채류 4종(파프리카, 토마토, 고추, 딸기), 엽채류 6종(고수, 바질, 로메인상추, 이자벨, 페얼리, 스텔릭스) 데이터를 스마트팜 R&D 빅데이터 플랫폼에서 정의한 스마트팜 표준화 데이터 수집 규칙에 맞춰 데이터 정제하여 플랫폼 업로드

구분	작목	데이터 항목
과채류	파프리카, 토마토	조사일시, 실험구, 내부온도, 습구온도, 노점, 내부상대습도, 수분부족분(중량), 수증기압차, 내부절대습도(중량), 질량당 엔탈피, 평균누적내부일사량, 난방기_난방온도, 내부CO2, 냉방기_저습경보설정값, 천창_개도율, 천창_개도율_1, 천창_개도율_2, 환기구_작동여부, CO2발생기_설정농도, 최저내부CO2, 최고내부CO2, CO2발생기_과다경보설정값, CO2발생기_부족경보설정값, CO2발생기_작동여부, 보온커튼(덮개)_개도율, 보온커튼(덮개)_작동방식, 일사량, 차광커튼_개도율, 보온커튼(덮개)_단힘광량, 보온커튼(덮개)_열림광량, 단열커튼(천정)_작동방식, 단열커튼(천정)_개도율, 환기구_환기온도, 보일러_온도, 보일러_저온경보설정값, 보일러_고온경보설정값, 보일러_작동여부, 외부온도, 외부풍속, 외부풍속_1, 외부풍향, 외부일사량, 전천일사량, 감우, 외부습도, 외부이슬점온도, 외부절대습도(중량), 양액시스템(관수모터)_최종작동일,

		양액시스템(관수모터)_공급상직, 최저EC, 최고EC, 평균급액EC, 최저pH, 최고pH, 평균급액pH, 급액시간, 급액량, 급액량(Dripper기준), 양액원수조_원수높이, 양액원수조_원수높이_최저, 양액원수조_원수높이_최고, 양액원수조_용량
	고추	조사일사, 조사시간, 내부온도, 내부온도_1, 내부온도_2, 내부온도_3, 내부온도_4, 내부상대습도, 내부상대습도_1, 내부상대습도_2, 내부상대습도_3, 내부CO2, 일사량, 외부풍속, 토양(배지)EC, 토양(배지)pH
	딸기	조사일시, 외부온도, 외부풍향, 외부풍속, 감우, 외부일사량, 외부누적일사량, 일사량, 내부누적일사량, 내부온도, 내부온도_1, 계산_평균내부온도, 내부상대습도, 평균내부상대습도, 내부CO2, 양액원수조_공급온도, 양액원수조_공급온도_1, 수분부족분, 포화수분, 내부절대습도, 외부이슬점온도
엽채류	고수, 바질, 로메인상추, 이자벨, 페어리, 스텔릭스	조사일자, 조사시간, 내부온도, 내부온도_1, 내부온도_2, 내부상대습도, 외부온도, 외부풍속, 일사량, 내부누적일사량, 감우, 환기구_환기온도, 난방기_난방온도

□ 과채류 및 엽채류 대상 스마트 온실 빅데이터 수집 사이트 구축

- 과채류 및 엽채류 실증사이트 및 대상작물 확대
- 실증사이트 3개 → 4개, 실증작물 6종 → 10종

구분	당초	변경	비고
농장	홍양영농조합법인 테라영농조합법인 알로하 채소농장	홍양영농조합법인 테라영농조합법인 알로하 채소농장 <b>고흥 스마트팜 혁신밸리 실증단지</b>	1개 증가
작물	토마토 파프리카 고추 고수 바질 로메인상추	토마토 파프리카 고추 고수 바질 로메인상추 <b>페얼리 이자벨 스텔릭스 딸기</b>	4종 증가

□ 과채류 대상 스마트 온실 빅데이터 수집 사이트 구축

- 고흥 스마트팜 혁신밸리 실증단지 작물 재비 및 데이터 수집 진행 중
- 과채류 1종 정식 완료

- 정식일 : 2023.11.24.
- 실증작물 : 딸기(죽향, 설향 2종)



<정식 1주차>



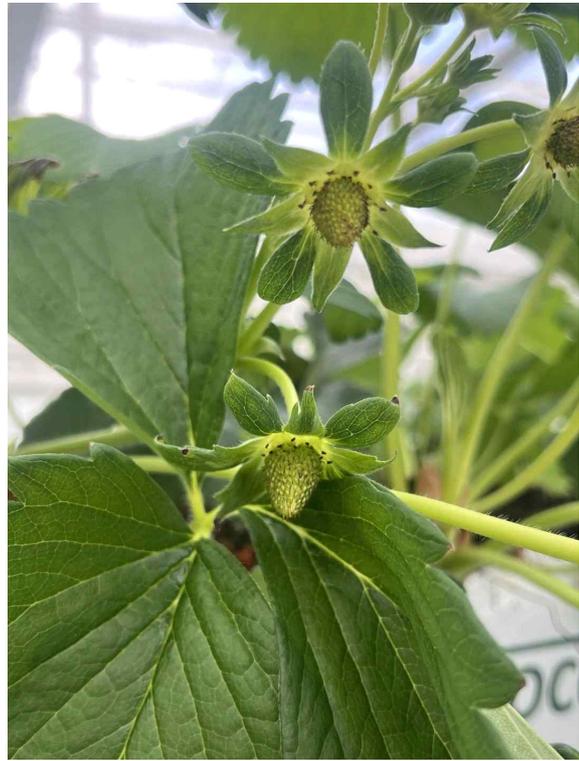
<정식 2주차>



<정식 4주차>



<정식 7주차>



<정식 9주차>

· 데이터 수집 항목

구분	데이터 수집 간격	데이터 수집 항목
항목	1분 10분 30분	외부온도, 풍향, 풍속, 감우, 외부일사, 누적일사(외부), 내부일사, 내부온도(1), 내부온도(2), 평균(내부온도(1+2)), 내부습도(1), 내부습도(2), 평균(내부습도(1+2)), CO2농도, 공급온도, 공급온도, 수분부족분, 포화수분, 절대습도, 이슬점

[데이터스퀘어]

[1단계(2021)]

□ 대상 작물별 데이터 명세 정의 및 표준데이터 범위 선정

- 인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼을 개발하기 위해 기존 스마트팜 데이터 표준을 조사·분석을 수행
  - 이를 위해 한국정보통신기술협회, 농업기술실용화재단, 농촌진흥청 등에서 제정된 기존 스마트팜 관련 데이터 표준화 조사하여 온실 작물에 대한 데이터 항목 검토·분석을 수행
    - (KSX3269) 스마트온실 센서 메타데이터, (TTAK.KO-10.0843) 시설 원예 생육 진단 메타데이터, (SPS-X FACT-0006-XXXX) 스마트온실 과채류 생육정보 메타데이터, (SPS-X FACT-0009-XXXX) 시설원예분야 스마트팜 수집 데이터, (농촌진흥청) 스마트 팜 확산 활용

시설원에 스마트팜 데이터 표준, 시설원에 스마트팜 측정 빅데이터 수집 기준 등

<시설원예분야 스마트팜 환경데이터 측정항목 정의(농촌진흥청)>

구분	변수	단위	범위
온실외부 기상환경	일사량	W·m <sup>-2</sup>	0 ~ 3000
	기온	℃	-30.0 ~ 100.0
	상대습도	%	0 ~ 100
	풍향	E/W/S/N	0 ~ 360
	풍속	m·s <sup>-1</sup>	0 ~ 50.0
	강우	Y/N	0/1
온실내부 기상환경	일사량*	W·m <sup>-2</sup>	0 ~ 1000
	기온	℃	-30.0 ~ 100.0
	상대습도	%	0 ~ 100
	CO <sub>2</sub> 농도	ppm	0 ~ 3000
	풍속	m·s <sup>-1</sup>	0 ~ 50.0
	냉방과 난방관 온도	℃	0 ~ 100.0
토양(근권)환 경	근권 온도	℃	-30.0 ~ 100.0
	근권수분	%	0 ~ 100
	근권 유효수분	Mpa or %	0 ~ 100
	양액 수온	℃	0 ~ 100
	양액 pH	pH	0.0 ~ 7.0
	양액 EC	dS·m <sup>-1</sup>	0.0 ~ 10.0
	토양(배지) pH	pH	0.0 ~ 7.0
	토양(배지) EC	dS·m <sup>-1</sup>	0.0 ~ 10.0

※ 출처: 농촌진흥청, '스마트 팜 확산 활용 시설원에 스마트팜 데이터 표준안 제안'

<시설원예 환경데이터 측정항목 정의(한국정보통신기술협회)>

번호	항목명	정의
1	작물 식별자	시설 원예에서 재배되는 작물을 정의
2	광량	일정한 시간동안 광도를 측정한 값의 합계를 나타내며 klx/hr로 표시
3	광질	광이 작물(作物)의 성장(生長)에 미치는 영향은 광파장에 기인하는 광선으로 390~760nm의 파장을 가진 가시광선(可視光線)이며 광파장에 따라 그 질이 다름
4	일장	하루 낮의 길이를 나타냄
5	광도	광원에서 비추는 면에 입사하는 단위면적 및 단위시간당 광에너지를 말하는데 그 단위는 Lux임.
6	CO <sub>2</sub> concentration (이산화탄소 농도)	대기 중에 있는 유효한 CO <sub>2</sub> 의 양을 나타내며 단위는 ppm임
7	온도	작물을 재배하는 동안의 온실 대기 중의 온도를 나타냄
8	이슬점 온도	일정량의 수증기를 함유한 공기가 차츰 냉각하여 포화 상태가 되고 수증기가 응축하여 물방울이 되기 시작하는 온도를 나타냄
9	상대 습도	최대 수증기량(포화 수증기량)에 대비 상대적인 수증기량을 퍼센트(%)로 나타냄
10	주위온도	작물 주위의 대기 온도를 나타냄

11	주위습도	작물 주위의 대기 습도를 나타냄
12	엽습윤	식물체 잎의 습도를 나타냄
13	토양온도	토양의 온도, 지표 및 토심별로 측정된 값
14	토양습도	토양공극내의 습도인데 토양 수분장력이 15기압 이하일 때 토양습도는 99% 이상임
15	토양산도	토양의 산도를 pH로 표시함
16	토양전기전도도	토양에 용존하는 염류의 농도를 나타내며 단위는 dS/m임
17	토양 수분장력	토양의 흡인력을 압력으로 표시함
18	양액온도	식물의 양액 재배 시 공급하는 양액의 온도를 나타냄
19	양액 전기전도도	양액에 용존하는 염류의 농도를 나타내며 단위는 dS/m임
20	양액 산도	양액의 산도를 pH로 표시함
21	폐액 산도	폐액의 산도를 pH로 표시함
22	폐액 전기전도도	폐액에 용존한 염류의 농도를 나타내며, 단위는 dS/m임

※ 출처: 한국정보통신기술협회, '시설원에 생육진단 메타 데이터', 2015.12

<시설원에분야 스마트팜 생육데이터 측정항목 - 완숙토마토(농촌진흥청)>

측정항목	측정기준	측정주기	단위
초장	지표면에서 생장점까지 길이	1주	cm
생장길이	지난주 생장점에서 금주 생장점 길이	1주	cm
엽수	개화 화방 아래의 완전히 전개한 엽 수	1주	개/주
엽장	맨 위에 개화 화방에서 3번째 해당하는 엽 길이 (착과된 바로 위의 엽)	1주	cm
엽폭	맨 위에 개화 화방에서 3번째 해당하는 엽의 폭	1주	cm
줄기굵기	개화화방(생장점 첫 화방) 위치에 바로 아래의 줄기 굵기	1주	cm
화방높이	개화화방에서 생장점까지의 거리	1주	cm
개화군	꽃이 핀 개수를 점수로 측정	1주	점
착과군	수정된 열매가 달려있는 것을 점수로 측정	1주	점
수확군	수확하는 열매를 점수로 측정	1주	점
착과수	착과수의 수정된 상태에서의 개수	1주	개
수확과수	수확한 열매 개수	1주	개
수확과중 (평균과중)	수확한 열매 개수의 무게 측정	1주	g

※ 출처: 농촌진흥청, '스마트 팜 확산 활용 시설원에 스마트팜 데이터 표준안 제안'

<시설원에분야 스마트팜 생육데이터 측정항목 - 파프리카(농촌진흥청)>

측정항목	측정기준	측정주기	단위
초장	지표면에서 생장점까지 길이	1주	cm

측정항목	측정기준	측정주기	단위
생장길이	지난주 생장점에서 금주 생장점 길이	1주	cm
엽수 (줄기마다 조사)	개화화방 기준 아래 완전 전개된 엽수	1주	개/주
엽장 (줄기마다 조사)	착과된 파프리카의 바로 위 엽장 혹은 맨위에개화화방에서 3번째 아래있는 엽길이	1주	cm
엽폭 (줄기마다 조사)	착과된 파프리카의 바로 위 엽장 혹은 맨 위에 개화화방에서 3번째 아래 있는 엽폭	1주	cm
줄기굵기 (줄기마다 조사)	개화화방(생장점 첫 화방) 바로 아래의 줄기굵기(낙화된 경우 추정하여 측정)	1주	cm
화방높이 (줄기마다 조사)	생장점에서 개화화방까지의 거리 (낙화된 경우 추정하여 계산)	1주	cm
개화마디 (개화화방)	현재 꽃이 피는 위치 (마디)	1주	마디
착과마디 (착과화방)	과일이 보이고(씨방이 보이면) 착과로 간주	1주	마디
착과수	화방의 착과에서 수정된 상태(열매가 맺은 상태)에서의 개수	1주	개
열매수	열매가 달려 있는 개수(착과제외)	1주	개
수확마디 (수확화방)	파프리카를 수확한 위치	1주	마디
수확과수	수확한 열매 개수	1주	개
과폭	착과된 파프리카의 과폭	1주	cm
과고	착과된 파프리카의 높이	1주	cm
무게	수확한 열매의 무게	1회	g

※ 출처: 농촌진흥청, '스마트 팜 확산 활용 시설원에 스마트팜 데이터 표준안 제안'

<토마토 생육조사 항목 및 조사방법(농정원)>

조사항목	단위	조사방법	비고
생장길이	cm	집게유인 : 전체 길이 측정 파스칼(혹크)유인 : 지난주 생장점에서 금주 생장점 길이	- (변경전) 전체 길이 측정
화방높이	cm	생장점에서 제일 위에 개화 화방*(꽃, 열매)까지의 길이	
줄기굵기	mm	개화화방(생장점 첫 화방) 위치에 바로 아래의 줄기 굵기	- (변경전) 지난주 생장점 위치 굵기 측정 - 개화화방 아래 2cm 정도에서 가장 넓은 곳(타원형임)
엽장	cm	개화 화방 바로 아래 잎의 길이를 측정	
엽폭	cm	개화 화방 바로 아래 잎의 가장 폭이 큰 부분 폭 길이 측정	- 잎의 가장 넓은 곳으로 한쪽만재서 2배 가능, 최대한 넓게 펼쳐서 길이를 잴
엽수	개	개화 화방 아래의 완전히 전개한 엽 수(측지 제거)	- (변경전) •3cm이상 잎의 수 측정 •화방이 달려있는 부분 제외 •측지는 제외하되, 3cm이상이면 잎수에 포함

개화군	점	꽃이 핀 개수를 점수로 측정	- (예)1화방에서 5개중에서 1개 개화:1/5=0.2, 3화방에서 꽃 1개 개화(다섯개 기준)=2.2 - 농가에서 꽃을 몇 개를 유지해서 생산할지 확인
착과군	점	수정된 열매가 달려 있는 것을 점수로 측정	- 과일이 보인다는 의미는 꽃이 빠지면서 열매가 맺으면 착과로 생각 - (예)1화방에서 꽃이 4개중 3개가 피고 착과가 2개인 경우 0.66
착과수 (열매수)	개	열매가 달려있는 총 개수를 측정함	- 착과된 것까지 포함하여 측정
최종 화방번호	번	현재 개화 되어 있는 화방의 차수 기입	

※ 출처: 농정원, 작물별 생육조사 방법

<파프리카 생육조사 항목 및 조사방법(농정원)>

조사항목	단위	조사방법	비고
생장길이	cm	지난주 생장점에서 금주 생장점 길이	(변경전)측정 당일의 전체 길이에서 이전주차 측정 전체길이를 뺀
엽수	개	개화화방 기준 아래 완전 전개된 잎의 수 ※ 측지는 제외하고 본 줄기 앞만 셈	(변경전) • 3cm이상 되는 잎의 수를 측정함 • 화방이 달려있는 부분은 제외 • 측지는 제외하되, 3cm이상이면 잎수에 포함
엽장	cm	착과된 파프리카의 바로 위 잎의 길이 또는 맨 위 개화 화방 아래 있는 잎의 길이	(변경전) 맨 위의 개화 화방에서부터 3번째 아래에 있는 잎의 길이를 측정
엽폭	cm	착과된 파프리카의 바로 위 잎의 너비 또는 맨 위 개화 화방 아래 있는 잎의 너비	(변경전) 맨 위의 개화 화방에서부터 3번째 아래에 있는 잎의 가장 폭이 큰 부분 폭 길이 측정
줄기굵기	mm	개화화방(생장점 첫 화방) 바로 아래(2cm)의 줄기 굵기	(변경전) 줄기의 굵기 측정은 지난주 생장점 위치의 굵기를 측정
화방높이	cm	생장점에서 개화화방까지의 거리	(변경전) 생장점에서 제일 위에 개화한 화방(꽃, 열매)까지의 길이
개화마디 (개화화방)	마디	현재 꽃이 피는 위치(마디) ※ 맨 위 꽃이 핀 마디번호만 표시	(변경전) 개화군
착과마디	마디	현재 착과가 된 위치(마디)	(변경전) 착과군

조사항목	단위	조사방법	비고
(착과화방)		※ 과일로 보이고(씨방이 보이면) 착과로 생각	
착과수	개	화방의 착과에서 수정된 상태(열매가 맺은 상태)에서의 개수 ※ 착과 후 꽃이 지고 열매(미착색) 높이 3cm 이하만 셈	
열매수	개	열매가 달려있는 개 수(착과제외) ※ 본 줄기에서 나온 열매 높이 3cm 초과만 셈	
최종 화방번호	번	현재 개화 되어 있는 화방의 차수 기입	- 신규 추가

※ 출처: 농정원, 작물별 생육조사 방법

- 한국정보통신기술협회, 농업기술실용화재단, 농촌진흥청 등 기존 스마트팜 관련 데이터 표준화와 본 사업의 스마트 온실 생산환경 관리용 빅데이터 플랫폼의 특성 및 관련 전문가의 의견을 고려하여 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보로 구성된 스마트팜 데이터 표준화를 위한 데이터 항목을 정의
- 스마트 온실 생산환경 관리용 스마트팜 데이터 표준화는 전반적인 온실 생산 환경의 스마트팜을 대상으로 요구되는 데이터 항목을 고려하였으며, 각 스마트팜 농가에서는 작물종류·온실형태 등에 의한 특성에 따라 스마트팜 데이터 표준화에서 제시된 데이터 항목이 상호 상이
  - 각 스마트팜 농가에서는 자체 농가의 특성 및 환경 등을 고려하여 스마트팜 데이터 표준화에서 제시된 데이터 항목의 일부를 활용할 것으로 예상
  - 현재 제시된 스마트팜 데이터 표준화는 신 농업기술, 최신 IT 기술 등을 통해 데이터 항목이 추가 또는 변경될 수 있으며, 이를 위해 표준화에 대한 지속적인 데이터 항목 관리가 필요
  - 작물의 생육정보 측정 및 값 추출의 자동화를 위한 촬영기기가 요구되며, 일관성 있는 생육정보 획득을 위해 촬영기기의 유형, 설치위치, 촬영 각도, 촬영 거리 등 정보가 필요함에 따라 이를 시설 및 제어 정보에 추가

<데이터 표준화 - 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보 데이터 항목>

구분	데이터 항목	
기본 데이터	○ 농업인ID, 시설ID, 품종ID, 생육측정 작물ID, 장비ID, 센서ID, 카메라ID, 온실유형, 온실 경도, 온실 위도, 재배유형, 작물 생육단계, 시설명, 우편번호, 시설주소, 시설규모(동수), 시설면적, 재배면적, 천창유형, 피복재질, 온실(시설) 형태-방향, 온실(시설) 형태-길이, 온실(시설) 형태-넓이, 온실(시설) 형태-높이, 배지형태(배지종류), 배지무게(배지 부피), 환경정보 측정일시, 생육정보 측정일시	
작물환경 데이터	온실 외부	○ 일사량, 누적 일사량, 일사량(최고), 광질, 일장, 광도, 온도, 온도(야간 평균), 온도(주간평균), 온도(최고), 온도(평균), 상대습도, 풍향, 풍속, 강우, 이슬점온도
	온실 내부	○ 일사량, 누적 일사량, 광질, 일장, 광도, 광파장, 온도, 이슬점 온도, 온도(최고), 온도(최저), 온도(평균), 습도, 평균 습도, 상대 습도, 절대 습도, 주위 습도, 상대습도(야간평균), 상대습도(주간평균), 상대습도(최고), 상대습도(최저), 상대습도(평균), 수분부족분, 엽습윤, CO <sub>2</sub> 농도, 풍향, 풍속, 냉방 및 난방관 온도, 배양실(CO <sub>2</sub> ), 배양실(습도), 배양실(온도), 생육실(CO <sub>2</sub> ), 생육실(습도), 생육실(온도), 조도, 차광율, 드리퍼

		당 공급량, 광양자(퀀텀)
	근권부	○ 근권 온도, 근권 수분/습도, 근권 유효수분(수분함수율), 근권 pH, 근권 EC, 근권 수분장력, 양액 온도, 양액 pH, 양액 EC, 토양(배지) pH, 토양(배지) EC, 공급 pH, 공급 EC, 배액 pH, 배액 EC, 총 배액량, 배지 온도, 양액 공급량, CO <sub>2</sub> 설정값, 양액공급횟수, 관수 배수율, 관수 일 배액량, 관수 일회공급량, 관수량, 지온
작물생육 데이터	잎	○ 엽장, 최대 엽장, 엽폭, 최대 엽폭, 엽병장, 잎자루 길이, 엽수, 잎 마디 번호, 엽면적지수, 엽온, 엽생체중, 광합성 속도
	줄기	○ 초장, 생장길이, 줄기굵기(줄기직경, 관부직경), 제1절간장, 줄기 절간, 절간 번호, 마디수, 줄기 생체중, 줄기 건물중, 수액이동 측정(SAP), 줄기 EC, 줄기유인방법, 줄기유인수
	꽃	○ 착과화방(착과마디), 출현화방수, 화방당착과수, 착과군, 생장점과화방 거리(화방높이), 수확화방(수확마디), 개화화방(개화마디), 개화기, 개화군, 화방당소화수(개화수), 개화 위치, 화방꽃수, 암꽃수, 수꽃수, 꽃자루 길이, 꽃자루 두께, 꽃받침 길이, 화방출뢰기, 최종 화방번호
	과실	○ 과실ID, 수확여부, 수확마디, 수확수(수확과수), 수확군, 과수, 과장(과고), 과폭, 과중, 수확과중, 수확평균과중, 과실생체중, 과실건물중, 착과수, 열매수, 주당 열매수, 수확과수, 유과수, 비상품과율, 비상품과수, 비상품과중, 착과마디번호
	뿌리	○ 근생체중, 근건물중
	품질	○ 당도, 산도, 완숙비율, 비타민함량, 잔류농약함량
경영영농 정보	경영정보	○ 작기 일련번호, 경영 일련번호, 경영 관리일시, 경영 관리유형, 농장명, 조직명, 사업자명, 대표자명, 대표품목, 대표품종, 지역, 주소, 연락처, 우편번호, 사업비-총사업비, 사업비-국고, 사업비-지방비, 사업비-자부담, 고용현황-사무실근로자, 고용현황-현장근로자, 인력-활용일시, 인력-활용유형, 인력-활용내용, 인력-일급/시급, 인력-인원수, 인력-시간, 인력-급여
	영농정보	○ (기본정보) 정식일시, 작목, 품종, 재식밀도, 적엽일시, 적엽량, 적심유무, 적심시기, 적심절위, 적과량, 육묘기간, 육묘 일수 ○ (방제관리) 방제정보 일련번호, 병해유형, 병해발생일시, 발생정도, 방제일시, 방제횟수, 방제작업명, 방제작업내용, 방제비용 ○ (생산관리) 수확정보(생산정보) 일련번호, 수확 시작일자(생산일시), 수확 종료일자, 수확량(생산량) ○ (출하관리) 출하정보 일련번호, 출하일시, 출하량, 출하단위, 출하단위당 가격, 매출 ○ (작기정보) 작기 일련번호, 작기명, 재배방식, 재배면적, 식재된 총 작물의 수, m <sup>2</sup> 당 식재된 작물의 수, 배지 1개당 작물의 줄기수, 배지1개당 작물의 수, 재식밀도 ○ (농작업관리) 농작업 일련번호, 농작업 일시, 농작업명, 농작업 유형, 농작업 내용, 농작업 비용(가격) ○ (유지관리) 유지보수 일련번호, 유지보수 일시, 유지보수 유형, 유지보수 내용, 유지보수 비용 ○ (양액관리) 양액정보 일련번호, 구입일시, 유형, 이름, 무게, 비용 ○ (기초원자재 관리) 원자재정보 일련번호, 구입일시, 항목, 내용, 비용
시설정보	온실정보	○ 통합제어기-제조사, 통합제어기-유형, 통합제어기-모델, 통합제어기-설치위치, 통합제어기-수량, 양액기/관수기-제조사, 양액기/관수기-유형, 양액기/관수기-모델, 양액기/관수기-설치위치, 양액기/관수기-수량, 냉난방장치-제조사, 냉난방장치-유형, 냉난방장치-모델, 냉난방장치-설치위치, 냉난방장치-수량

	에너지 정보	○ 에너지 유형, 에너지 사용량, 에너지 관리방식
	센서정보	○ 센서유형, 센서위치, 센서정상범위, 센서정확도, 측정일시, 측정값
	구동기 정보	○ 구동기 유형, 구동기 수량, 구동기 위치, 구동기 자동유무
	촬영기기 정보	○ 촬영기기 유형, 촬영기기 위치, 촬영기기 각도, 촬영거리, 측정일시, 측정값
	기타	○ 보온자재, 난방 방법, 환기 방법, 관수 방법, UPS 설치 유무, 인터넷 유무
제어정보	천창	○ 작동시간, 작동여부, 작동상태, 개도율, 천창 동작 설정값
	측창	○ 작동시간, 작동여부, 작동상태, 개도율, 측창 동작 설정값
	단열커튼 (천장)	○ 작동시간, 개폐시간, 개도율
	단열커튼 (측면)	○ 작동시간, 개폐시간, 개도율
	보온커튼 (덮개)	○ 작동시간, 작동여부, 작동상태, 개도율, 보온커튼 동작 설정값
	차광커튼	○ 작동시간, 작동여부, 작동상태, 개도율, 차광커튼 동작 설정값
	유동팬	○ 작동시간, 작동여부, 작동상태, 유동팬 동작 설정값
	보광등	○ 작동시간, 작동여부, 작동상태, 보광등 동작 설정값
	3WAY 밸브	○ 작동시간, 작동여부, 작동상태, 공급수 온도, 3WAY밸브 동작 설정값
	배기팬	○ 작동시간, 작동여부, 작동상태, 배기팬 동작 설정값
	환기구	○ 작동시간, 작동여부, 작동상태, 환기조절
	관수작동상태(토경)	○ 작동시간, 작동여부, 작동상태, 관수 공급 설정값
	양수작동상태(수경)	○ 작동시간, 작동여부, 작동상태
	습관리장치	○ 작동구분(제습/가습), 작동시간, 작동여부, 작동상태
	난방기	○ 작동시간, 작동여부, 작동상태, 난방 설정 온도
	냉방기	○ 작동시간, 작동여부, 작동상태, 냉방 설정 온도
	냉난방기	○ 작동구분(냉방/난방), 작동시간, 작동여부, 작동상태, 냉난방 설정 온도
	훈증기	○ 작동시간, 작동여부, 작동상태, 훈증기 동작 설정값
	수막시스템	○ 작동시간, 작동여부, 작동상태
	CO <sub>2</sub> 발생기	○ 작동시간, 작동여부, 작동상태, CO <sub>2</sub> 공급 설정농도값
	양액시스템(관수모터)	○ 작동시간, 작동여부, 작동상태, 양액 공급 설정pH값, 양액 공급 설정 EC값
	유량계	○ 작동시간, 작동여부
	분무(살수장치)	○ 작동시간, 작동여부, 작동상태, 분무기 동작 설정값
	보일러	○ 작동시간, 작동여부, 작동상태

모니터링	○ CCTV, 웹캠, 녹화장비
포그분사기	○ 작동시간, 작동여부, 작동상태, 포그분사기 동작 설정값
두상살수 장치	○ 작동시간, 작동여부, 작동상태
보온덮개 개폐기	○ 작동시간, 작동여부, 작동상태
센서	○ 작동시간, 작동여부, 작동상태
활영기기	○ 작동시간, 작동여부, 작동상태

○ 흥양영농조합법인의 시설, 장비, 작물 등을 고려하여 해당 법인에 특화된 스마트 온실 생산환경 관리용 데이터 표준화를 위한 데이터 항목을 도출

- 본 사업 관련 전문가(농업+IT 전문가)와 협의를 통해 스마트 온실 생산환경 관리용 데이터 표준화에서 흥양영농조합법인의 작물종류·온실형태·설치장비 등 특성을 고려하여 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보로 구성된 데이터 항목을 선정
- 흥양영농조합법인에서 기존 장비 제거, 신규 장비 설치, 신규 작물 추가 등으로 인해 스마트 온실 생산환경 관리용 데이터 표준화가 변경될 수 있음

<데이터 표준화 -

흥양영농조합법인의 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보 데이터 항목>

구분		데이터 항목
기본 데이터		○ 농업인ID, 시설ID, 품종ID, 생육측정 작물ID, 장비ID, 센서ID, 카메라ID, 온실유형, 배지형태(배지종류), 배지무게(배지 부피), 환경정보 측정일시, 생육정보 측정일시
작물환경 데이터	온실 외부	○ 일사량, 누적 일사량, 일사량(최고), 광도, 온도, 온도(야간평균), 온도(주간평균), 온도(최고), 온도(평균), 상대습도, 풍향, 풍속, 강우
	온실 내부	○ 일사량, 누적 일사량, 광도, 온도, 이슬점 온도, 온도(최고), 온도(최저), 온도(평균), 습도, 평균 습도, 상대 습도, 절대 습도, 주위 습도, 상대습도(야간평균), 상대습도(주간평균), 상대습도(최고), 상대습도(최저), 상대습도(평균), 수분부족분, CO <sub>2</sub> 농도, 풍속, 냉방 및 난방관 온도, 생육실(CO <sub>2</sub> ), 생육실(습도), 생육실(온도), 드리퍼당 공급량
	근권부	○ 근권 pH, 근권 EC, 양액 pH, 양액 EC, 공급 pH, 공급 EC, 배액 pH, 배액 EC, 양액 공급량, CO <sub>2</sub> 설정값, 양액공급횟수, 관수 배수율, 관수 일배액량, 관수 일회공급량, 관수량
작물생육 데이터	잎	○ 엽장, 최대 엽장, 엽폭, 최대 엽폭, 잎자루 길이, 엽수, 잎 마디 번호
	줄기	○ 초장, 생장길이, 줄기굵기(줄기직경, 관부직경), 절간 번호, 마디수
	꽃	○ 착과화방(착과마디), 출현화방수, 화방당착과수, 착과군, 생장점과화방거리(화방높이), 수확화방(수확마디), 개화화방(개화마디), 개화군, 화방당 소화수(개화수), 개화 위치, 화방꽃수
	과실	○ 수확수(수확과수), 수확군, 수확평균과중, 착과수, 열매수
	품질	○ 해당 사항 없음
경영영농 정보	경영정보	○ 농장명, 조직명, 사업자명, 대표자명, 대표품목, 지역, 주소, 연락처
	영농정보	○ (기본정보) 정식일시, 작목, 품종, 재식밀도, 적시유무, 육묘기간, 육묘

		<p>일수</p> <ul style="list-style-type: none"> <li>○ (생산관리) 수확 시작일자(생산일시), 수확량(생산량)</li> <li>○ (출하관리) 출하일시, 출하량</li> <li>○ (작기정보) m<sup>2</sup>당 식재된 작물의 수, 배지 1개당 작물의 줄기수, 배지1개당 작물의 수, 재식밀도</li> <li>○ (양액관리) 구입일시, 이름, 무게</li> </ul>
시설정보	온실정보	<ul style="list-style-type: none"> <li>○ 통합제어기-제조사, 통합제어기-유형, 통합제어기-모델, 통합제어기-설치위치, 양액기/관수기-제조사, 양액기/관수기-유형, 양액기/관수기-모델, 양액기/관수기-설치위치, 양액기/관수기-수량, 냉난방장치-제조사, 냉난방장치-유형, 냉난방장치-모델, 냉난방장치-설치위치, 냉난방장치-수량</li> </ul>
	에너지정보	○ 해당 사항 없음
	센서정보	○ 센서유형, 센서위치, 센서정상범위
	구동기정보	○ 해당 사항 없음
	촬영기기정보	○ 촬영기기 유형, 촬영기기 위치, 촬영기기 각도, 촬영거리, 측정일시, 측정값
	기타	○ 해당 사항 없음
제어정보	천창	○ 개도율, 천창 동작 설정값
	측창	○ 개도율, 측창 동작 설정값
	단열커튼(천장)	○ 해당 사항 없음
	단열커튼(측면)	○ 해당 사항 없음
	보온커튼(덮개)	○ 작동시간, 작동여부, 개도율, 보온커튼 동작 설정값
	차광커튼	○ 작동시간, 작동여부, 작동상태, 개도율, 차광커튼 동작 설정값
	유동팬	○ 작동시간, 작동여부, 작동상태, 유동팬 동작 설정값
	보광등	○ 해당 사항 없음
	3WAY 밸브	○ 작동여부, 공급수 온도
	배기팬	○ 해당 사항 없음
	환기구	○ 해당 사항 없음
	관수작동상태(토경)	○ 해당 사항 없음
	양수작동상태(수경)	○ 해당 사항 없음
	습관리장치	○ 해당 사항 없음
	난방기	○ 해당 사항 없음
	냉방기	○ 해당 사항 없음
냉난방기	○ 해당 사항 없음	

훈증기	○ 작동시간, 작동여부, 작동상태, 훈증기 동작 설정값
수막시스템	○ 해당 사항 없음
CO <sub>2</sub> 발생기	○ 작동시간, 작동여부, 작동상태, CO <sub>2</sub> 공급 설정농도값
양액시스템(관수모터)	○ 작동시간, 작동여부, 작동상태, 양액 공급 설정pH값, 양액 공급 설정EC 값
유량계	○ 해당 사항 없음
분무(살수 장치)	○ 해당 사항 없음
보일러	○ 해당 사항 없음
모니터링	○ 해당 사항 없음
포그분사기	○ 해당 사항 없음
두상살수 장치	○ 해당 사항 없음
보온덮개 개폐기	○ 해당 사항 없음
센서	○ 작동시간, 작동여부, 작동상태
촬영기기	○ 작동시간, 작동여부, 작동상태

□ 데이터 분석 및 인공지능 모델 구축에 활용하기 위한 온실 환경·생육·경영영농·시설·제어 정보의 데이터 전처리(가공) 기술 고안

- 온실 환경·생육·경영영농·시설·제어 정보 등에 대한 각 데이터 특성 및 구조를 분석
  - 온실 환경·생육·경영영농·시설·제어 정보 등 온실 생산환경 관리를 위한 온실 데이터를 조사 및 분석을 수행
  - 온실 생산환경 관리 관련 데이터 표준 자료 조사·검토를 통한 주요 데이터 항목 발굴과 각 데이터 특성 및 구조를 분석
- 불완전하고 모순된 데이터를 정제(Cleaning)하기 위한 데이터의 무결성 검증 방안을 고안함
  - 온실 생산환경 관리를 위해 수집되는 데이터 현황을 조사하고 분석
  - 수집 데이터에 대한 정제가 요구되는 항목 도출 및 데이터 무결성 검증 방안을 설계함(데이터 구조·입력값 범위 관련 무결성 검증)

<데이터 무결성 예시>

무결성	내용	비고
개체 무결성	○ 기본키는 중복되거나 누락될 수 없는 특성(기본키는 유일성 보장)이 있음	Null값이 허용되지 않음
참조 무결성	○ 외래키의 참조값은 다른 개체의 기본키 또는 NULL 값을 가짐	-
속성 무결성	○ 속성값은 지정된 데이터 형식을 반드시 만족하는 값만 가짐	정의된 값만 입력(Number, Character, Date 등)
도메인 무결성	○ 특정 속성값은 미리 정의된 범위 내에 포함됨	허용된 범위 값만 입력
키 무결성	○ 한 릴레이션에 동일한 키 값을 가진 튜플은 허용되지 않음	중복 배제

- 스마트 온실 생산환경 관리용 데이터 표준화를 기반으로 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보로 구성된 데이터 항목에 대해 수집된 데이터의 신뢰성 및 정확성을 제고하기 위해 검사 기준 및 방법을 고안
- 스마트 온실 생산환경 관리용 데이터 표준화에 정의된 데이터의 형식에 대한 준수 여부를 판별할 수 있는 검사 기준 및 방법을 고안

<수집된 데이터의 형식 검사>

분류		검사기준 및 방법
검사 상세		○ 스마트 온실 생산환경 관리를 위해 수집된 데이터의 형식을 준수하는지 확인
검사 대상	산출물	○ 데이터 형식 검사를 통한 정제된 데이터
	데이터 종류	○ 텍스트
	내용	○ 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보로 구성된 데이터 항목의 값
검사방법	SW 기반 검사	○ 데이터 표준화에 정의된 데이터 형식을 준수하지 않는 데이터가 있는지 확인

- 스마트 온실 생산환경 관리용 데이터 표준화에 정의된 데이터의 속성값이 입력 유효값 범위 내에 존재하는지 판별할 수 있는 검사 기준 및 방법을 고안

[표] 수집된 데이터의 입력값 범위 검사

분류		검사기준 및 방법
검사 상세		○ 스마트 온실 생산환경 관리를 위해 수집된 데이터의 속성값이 입력 유효 값 내에 존재하는지 확인
검사 대상	산출물	○ 입력값 범위 검사를 통한 정제된 데이터
	데이터 종류	○ 텍스트
	내용	○ 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보로 구성된 데이터 항목의 값
검사방법	SW 기반 검사	○ 데이터 표준화에 정의된 데이터를 구성하는 속성값이 입력 유효값에 해당하지 않는 데이터가 있는지 확인

□ 데이터 분석 및 인공지능 모델 구축에 활용할 수 있는 최적화된 데이터베이스 설계

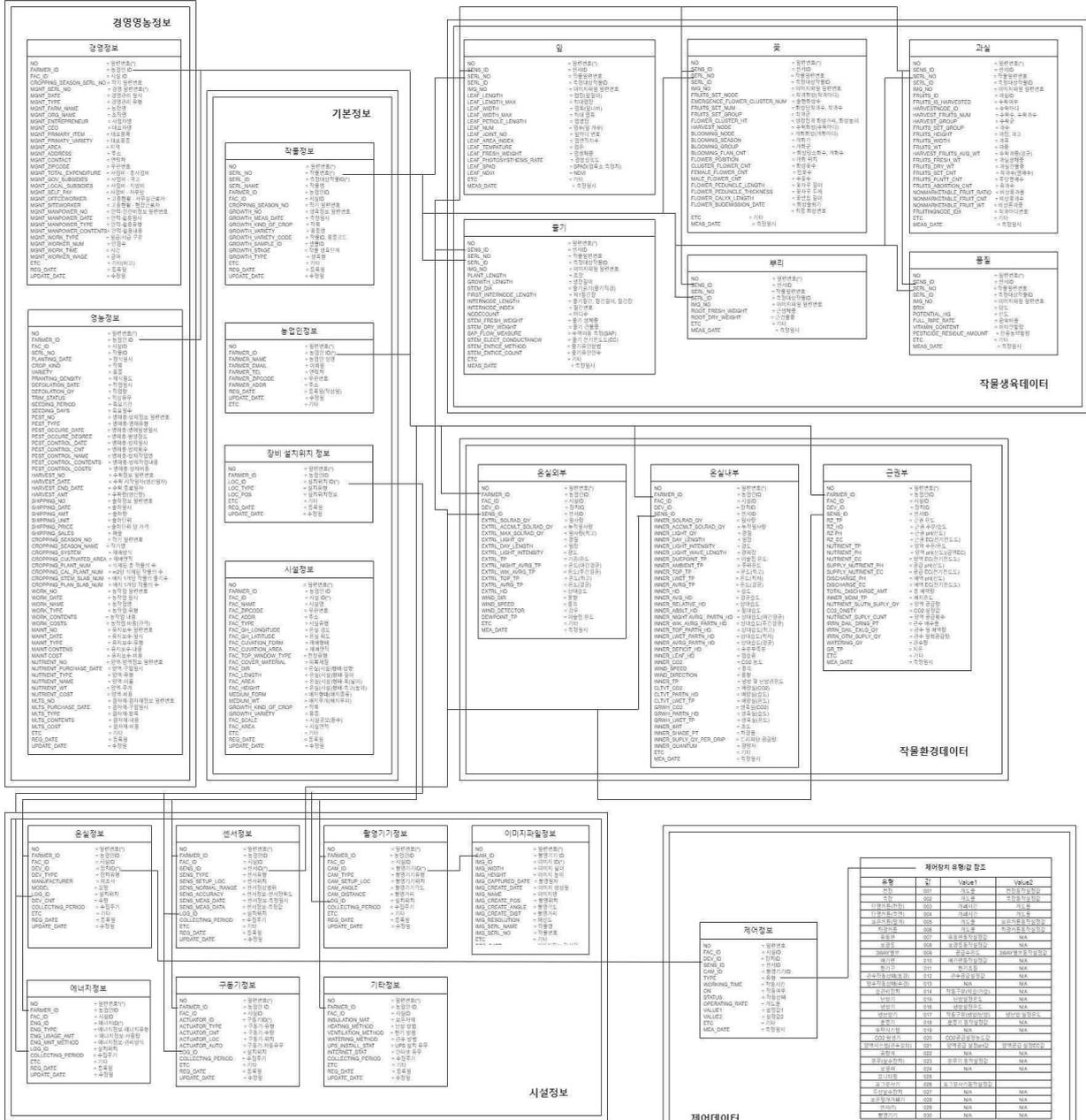
- 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보를 활용하여 인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼 개발을 효과적으로 지원할 수 있도록 개체관계 모델을 도출하고 데이터베이스를 설계
- 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보 등 온실 및 작물 관련 데이터 현황, 표준 데이터 항목 등을 조사하고 작물의 온실 생산환경 관리를 위한 데이터베이스의 요구사항

을 분석하여 온실 생산환경 관리용 데이터베이스 설계(안)을 도출

- 온실 생산환경 관리용 데이터베이스 설계(안)에서 기본정보(농업인정보, 경영정보, 영농정보, 영농정보, 작물 정보, 시설정보), 장비정보(장치정보, 센서정보, 촬영기기정보, 장비설치정보), 작물환경정보(온실 외부, 온실 내부, 근권부), 생육정보(잎, 줄기, 꽃, 과실) 등 데이터를 포함

- 도출한 DB스키마 : 별첨파일(1단계2021항목) 참조

- 온실 생산환경 관리를 위한 기본정보, 장비정보, 작물환경정보, 생육정보 등을 고려하여 온실 생산 환경 관리용 개체관계 모델(ER 다이어그램)을 도출



<온실 생산환경 관리용 개체관계 모델>

[1단계(2022)]

□ 대상 작물별 데이터 명세를 고려한 데이터 표준화 개발 및 제안

- “반밀폐 스마트 온실 대상 데이터 항목 표준”을 위한 데이터 표준화 추진(주관기관 협력)
  - 인공지능 기반 자율형 스마트 온실 생산환경 관리용 빅데이터 플랫폼을 개발하기 위해 기존 스마트팜 데이터 표준에 대한 조사·분석 수행
  - 최근 스마트팜의 경우 대규모의 최첨단 반밀폐형 유리온실이 많이 확산됨에 따라 해당 온실의 특성을 감안한 시설 및 제어 관리 표준이 요구되고 있음
  - ※ 반밀폐형 유리온실: 일정 범위까지 밀폐형을 유지하다가 특정 조건하에서 천장 및 측장이 개폐식으로 되어 있어 공기의 순환이 가능하며, 재배과정을 일부 통제할 수 있는 온실
  - 본 컨소시엄의 전문가 등을 통해 반밀폐형 온실에 특화된 온실환경·시설·제어 정보를 위주로 데이터 표준화를 추진
  - “반밀폐 스마트 온실 대상 데이터 항목 표준”을 농업기술실용화재단에 제안하였으며, 수정 및 보완 사항에 대처할 예정
  - 반밀폐형 온실의 특성을 고려한 시설 및 제어 정보관리를 통해 정밀한 작물의 생육관리를 지원하여 스마트팜 농가의 생산성 향상에 기여



농업기술실용화재단  
20XX년 XX월 XX일 제정

<반밀폐 스마트 온실 대상 데이터 항목 표준>

□ 데이터 분석 및 인공지능 모델 구축에 활용하기 위한 온실 생육·환경·영농 정보의 데이터 전처리

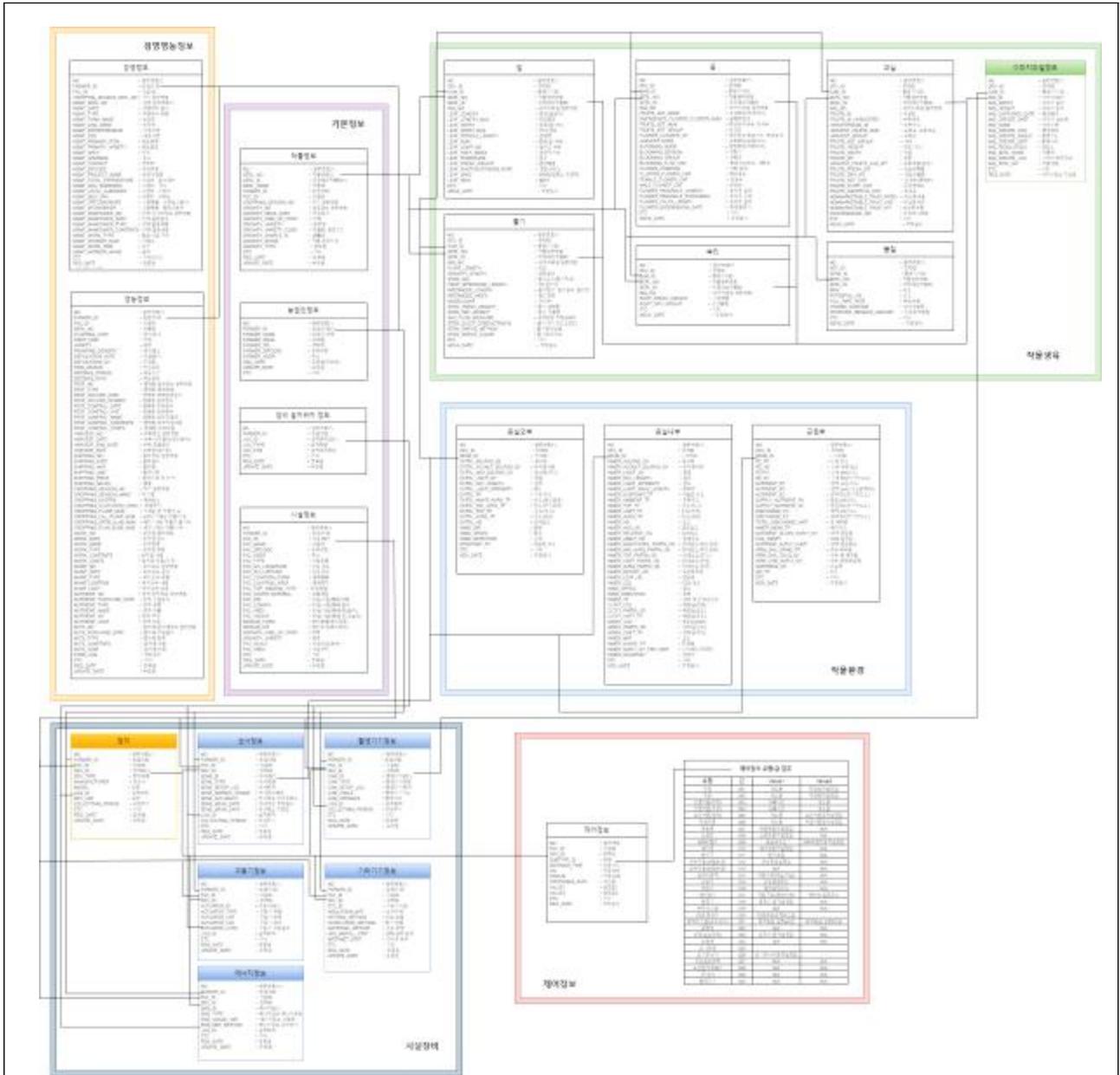
### (가공) 기술 개발

- 수집된 데이터의 형식, 데이터 속성값의 유효성 등을 점검하여 불완전하고 모순된 데이터를 정제할 수 있는 데이터 무결성 검증 로직 개발
  - 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보로 구성된 데이터 항목에 대해 수집된 데이터의 신뢰성 및 정확성을 제고하기 위한 검사 기준 및 방법 제시
  - 수집 데이터의 형식(데이터 형식 검사), 데이터 속성값의 입력 유효값(입력값 범위 검사)에 대한 준수 여부를 확인할 수 있는 SW 기반 데이터 무결성 검증 로직 개발
  - 데이터 유효성 검사는 데이터 형식을 숫자와 텍스트로 구분하고 포맷(날짜형식 등) 및 데이터 유효 범위를 설정하여 데이터의 이상여부를 일괄 검사 가능
  - 수집된 엑셀·JSON 형식의 데이터 → 설정된 레이블에 대한 유효값을 확인 → 유효성검사 내용 및 검사 결과 표시



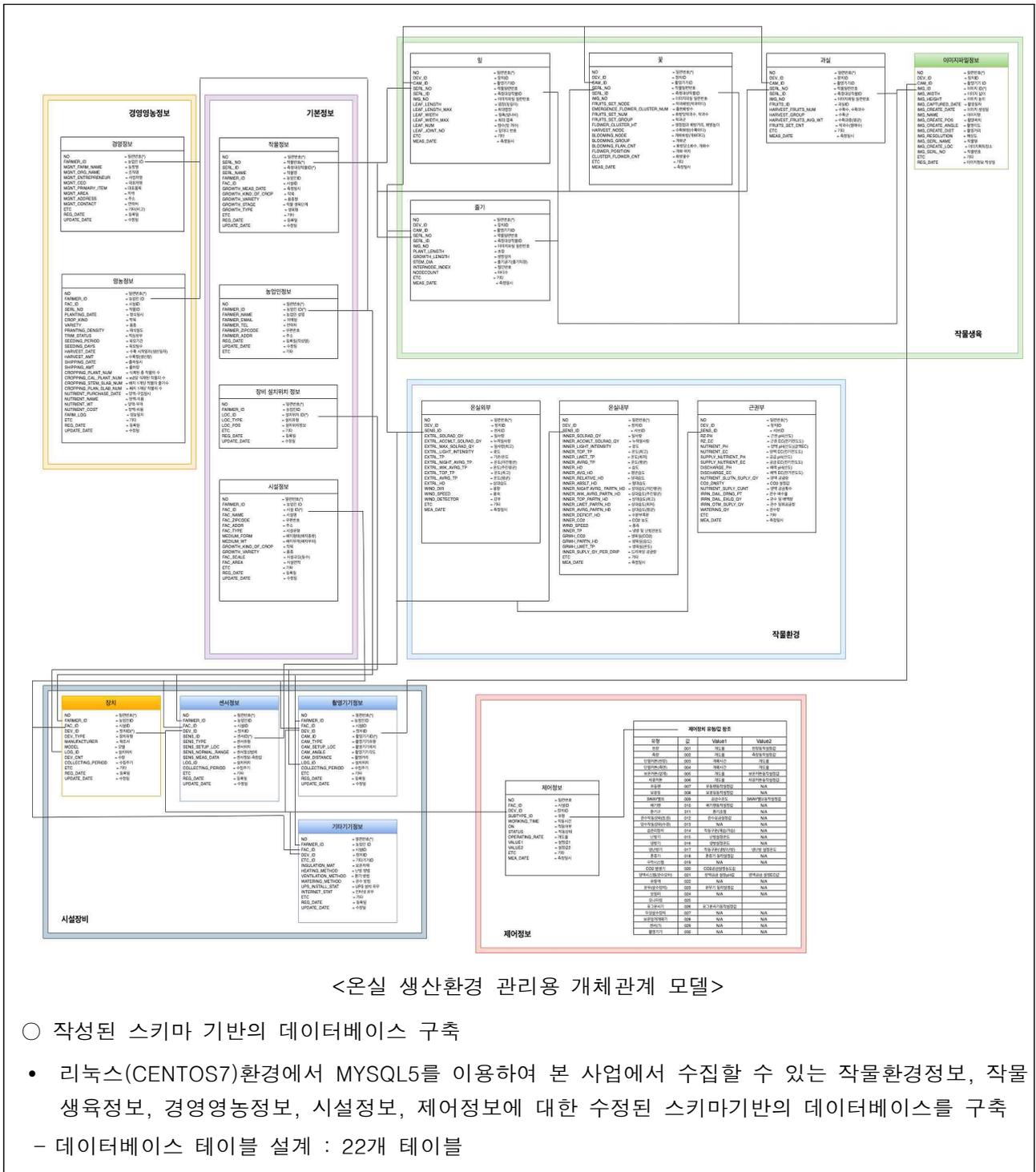
<데이터 유효성 검사화면 예시>

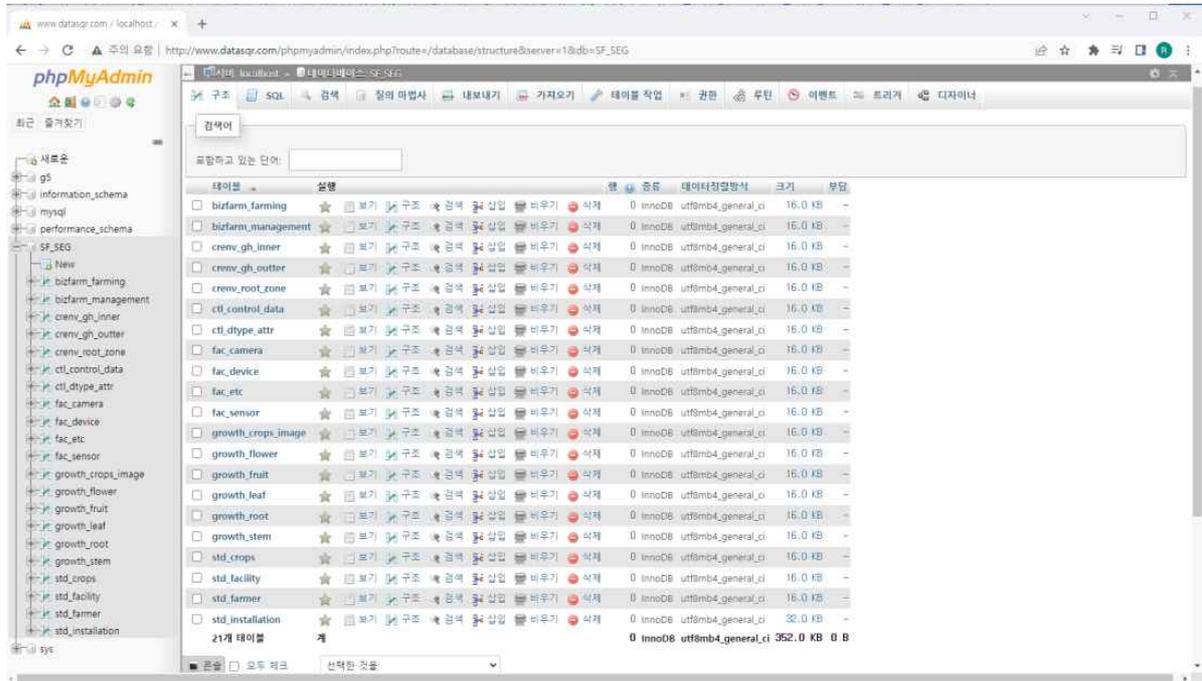
- 수집 데이터의 데이터 무결성 검사를 통해 오류 내용에 대한 통계 데이터를 보여줄 수 있는 기능 개발
- 대상 작물별 데이터 분석 및 인공지능 모델 구축에 활용할 수 있는 최적화 된 데이터베이스 구축
- 1차년도 데이터베이스 스키마의 추가·수정 추진
    - 영농일지·이미지 획득장소 추가, 제어정보(제어데이터, 장치유형 및 값) 수정 등 1차년도에 도출된 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보의 추가 및 수정이 발생함에 따라 데이터베이스 스키마 보완 추진
      - 수정된 전체 DB스키마 별첨파일(1단계2022항목) 참조
      - 수정된 데이터베이스 스키마를 고려하여 개체관계 모델(ER 다이어그램)을 보완



<온실 생산환경 관리용 개체관계 모델>

- 본 사업에서 수집할 수 있는 작물환경정보, 작물생육정보, 경영영농정보, 시설정보, 제어정보를 기반으로 데이터베이스 재설계 및 개체관계 모델 수정 및 데이터베이스 구축
  - 전체적인 데이터베이스 스키마에 비해 에너지 정보·구동기 정보·품질 데이터 테이블이 없으며, 나머지 테이블의 데이터 항목도 변경된 경우가 발생
    - 테스트베드(흥양)에 최적화 된 DB스키마 : 별첨파일(1단계2022항목) 참조
- 본 사업에서 수집이 용이한 기본정보, 장비정보, 작물환경정보, 생육정보 등을 고려하여 온실 생산환경 관리용 개체관계 모델(ER 다이어그램)을 도출





<설계변경된 스키마를 기반으로 구축된 DB테이블 목록>

No.	테이블 이름	No.	테이블 이름
1	기본정보-농업인정보 테이블	12	제어정보-제어데이터 테이블
2	기본정보-작물정보 테이블	13	작물환경-온실외부 데이터 테이블
3	기본정보-시설정보 테이블	14	작물환경-온실내부 데이터 테이블
4	경영영농정보-경영정보 테이블	15	작물환경-근권부데이터 테이블
5	경영영농정보-영농정보 테이블	16	작물생육-잎 데이터 테이블
6	시설장치-장치정보 테이블	17	작물생육-줄기 데이터 테이블
7	시설장치-에너지정보 테이블	18	작물생육-꽃 데이터 테이블
8	시설장치-센서정보 테이블	19	작물생육-과실 데이터 테이블
9	시설장치-구동기정보 테이블	20	작물생육-뿌리 데이터 테이블
10	시설장치-촬영기기정보 테이블	21	작물생육-품질 데이터 테이블
11	시설장치-기타정보 테이블	22	작물생육-이미지파일 데이터 테이블

## [2단계(2023)]

### □ 대상 작물별 데이터 명세를 고려한 데이터 표준 제정

○ 데이터 표준화를 위한 국내 표준기관 제출용 표준 초안 작성

- 반밀폐 스마트 온실 대상 데이터 항목 표준 초안 작성 및 제출

- 공통데이터 항목은 기본적으로 선행표준인 SPS-X KOAT-0009-7470(시설원에 분야 스마트팜 수집 데이터 규격)을 인용하여 용어, 정의, 단위를 통일

- 기존 선행표준에 반밀폐 온실에 특화된 정보를 추가한 상세 메타데이터를 정의

- 기존 표준 대비 반밀폐형 온실을 고려하여 센서 용어 추가(복사에너지 측정센서, 기압계, 적외선





내용이므로 표준문서에 해당 내용을 포함시키는 것은 적절하지 않다고 판단됩니다.  
실제로 스마트팜에서 환경정보 관련 XML 내용(예시)은 아래와 같다고 예상됩니다.

/\* 외부환경데이터 예제 \*/

<ENV\_GH\_EXTERNAL\_DATA>

:

<EXTRL\_DATA>

<NO>20</NO>

<FARMER\_ID>HY01</FARMER\_ID>

<FAC\_ID>F001</FAC\_ID>

<DEV\_ID>D1</DEV\_ID>

<SENS\_ID> S1</SENS\_ID>

<EXTRL\_SOLRAD\_QY>1.8</EXTRL\_SOLRAD\_QY>

<EXTRL\_ACCMLT\_SOLRAD\_QY>722.1

</EXTRL\_ACCMLT\_SOLRAD\_QY>

<EXTRL\_MAX\_SOLRAD\_QY>35.6</EXTRL\_MAX\_SOLRAD\_QY>

<EXTRL\_LIGHT\_INTENSITY>0</EXTRL\_LIGHT\_INTENSITY>

<EXTRL\_TP>7.4</EXTRL\_TP>

<EXTPRL\_NIGHT\_AVRG\_TP>7.3</EXTPRL\_NIGHT\_AVRG\_TP>

<EXTRL\_WIK\_AVRG\_TP>7.5</EXTRL\_WIK\_AVRG\_TP>

<EXTRL\_TOP\_TP>7.7</EXTRL\_TOP\_TP>

<EXTRL\_AVRG\_TP>7.2</EXTRL\_AVRG\_TP>

<EXTRL\_RH>78</EXTRL\_RH>

<EXTRL\_HD>0</EXTRL\_HD>

<WIND\_DIR>277</WIND\_DIR>

<WIND\_SPEED>2</WIND\_SPEED>

<RAIN\_DETECTOR>0</RAIN\_DETECTOR>

<ETC >-</ETC>

<MEAS\_DATE>2023-05-26 19:53:50</MEAS\_DATE>

</EXTRL\_DATA>

:

</ENV\_GH\_EXTERNAL\_DATA>

/\* 근권부환경데이터 예제 \*/

<ENV\_ROOTZONE\_DATA>

:

<ENV\_DATA>

<NO>34</NO>

<FARMER\_ID>HY01</FARMER\_ID>

<FAC\_ID>F001</FAC\_ID>

<FAC\_BLOCK>BLK001</FAC\_BLOCK>

<DEV\_ID>R1</DEV\_ID>

<SENS\_ID>RS2</SENS\_ID>

<DISCHARGE\_PH>1</DISCHARGE\_PH>

<DISCHARGE\_EC>2</DISCHARGE\_EC>

<NUTRIENT\_SLUTN\_SUPLY\_QY>64</NUTRIENT\_SLUTN\_SUPLY\_QY>

```

<CO2_DNSTY>300</CO2_DNSTY>
<NUTRIENT_SUPLY_CUNT>12</NUTRIENT_SUPLY_CUNT>
<IRRN_OTM_SUPLY_QY>708</IRRN_OTM_SUPLY_QY>
<ETC >-</ETC>
<MEAS_DATE>2023-05-28 7:10:12</MEAS_DATE>
</ENV_DATA>
:
</ENV_ROOTZONE_DATA>

```

- 선행표준인 SPS-X KOAT-0009-7470 (시설원에 분야 스마트팜 수집 데이터 규격)에 반밀폐 온실에서 사용되는 용어 및 장비정보 등을 추가하여 최종 제출 완료
- 2023년 KoAT 단체표준(안) 제정 및 개정 관련하여 예고고시를 농진원 홈페이지를 통해 진행
  - 2023.12.29.일부터 2024.01.30.일 까지 총 30일간 진행



SPSPSPSP SPS X KOAT 0009-7470

SPSPSPS

SPSPSP

SPSPS

SPSP

SPS

시설원에 분야 스마트팜 수집 데이터 규격  
SPS X KOAT 0009-7470:2023

**2023년 KoAT단체표준안 예고**

구분	내용	비고
관련근거	「단체표준 적용 및 추진 과정 규정」 제2조(단체표준안 제고)	-
제고기간	(2023.12.29. ~ 2024.01.30.)	-

**KOAT단체표준안 목록**

순번	단체표준명	구분
1	스마트온실 - 농업용 로보 - 제3부:구동 시험 방법	제정
2	스마트온실 - 농업용 로보 - 자율인식 시험방법(완속 도차도)	제정
3	스마트온실 - 농업용 로보 - 수확 시험방법(완속 도차도)	제정
4	시설원에 분야 스마트팜 수집 데이터 규격	개정

한국농업기술진흥원  
2023년 1월 1일 개정

**<시설원에 분야 스마트팜 수집 데이터 규격 개정 표준>**

**□ 실증을 통한 데이터베이스 갱신 및 데이터베이스 최적화**

- 실증을 통한 농업인정보, 시설정보, 센서정보, 시설제어기 정보, 환경 정보, 작물 생육정보, 경영관리정보 등 작물별 주요 온실 환경 정보의 데이터베이스 갱신
  - 과채류·엽채류 생육 및 반밀폐 온실 특성을 고려한 데이터스키마 수정
    - 과채류 생육의 특성을 고려한 데이터스키마 수정
      - ※ SPS-X KOAT-0006-7468\_스마트온실 과채류 생육 정보 메타데이터(가지,고추,미니오이,방울토마토,오이,토마토,파프리카,딸기,멜론,참외) 표준 참고
    - 수정된 과채류 작물생육 DB스키마 : 별첨파일(1단계2022항목) 참조
    - 엽채류 생육의 특성을 고려한 데이터스키마 수정
      - ※ SPS-X KOAT-0007-7471\_스마트온실 엽채류 생육 정보 메타데이터(상추,양상추) 표준 참고

- 수정된 업체류 작물생육 DB스키마 : 별첨파일(1단계2022항목) 참조
- 실제 수집 데이터 및 데이터스키마 간 매핑 테이블을 작성하여 작물환경정보, 시설정보, 제어정보 등을 고려한 작물별 주요 온실 환경 정보 및 생육정보의 데이터베이스 구축
- 실제 수집 데이터에서 데이터스키마의 항목별 데이터를 추출하기 위한 매핑 테이블을 작성함

<과채류 작물환경-온실외부 데이터 매핑 테이블 예시>

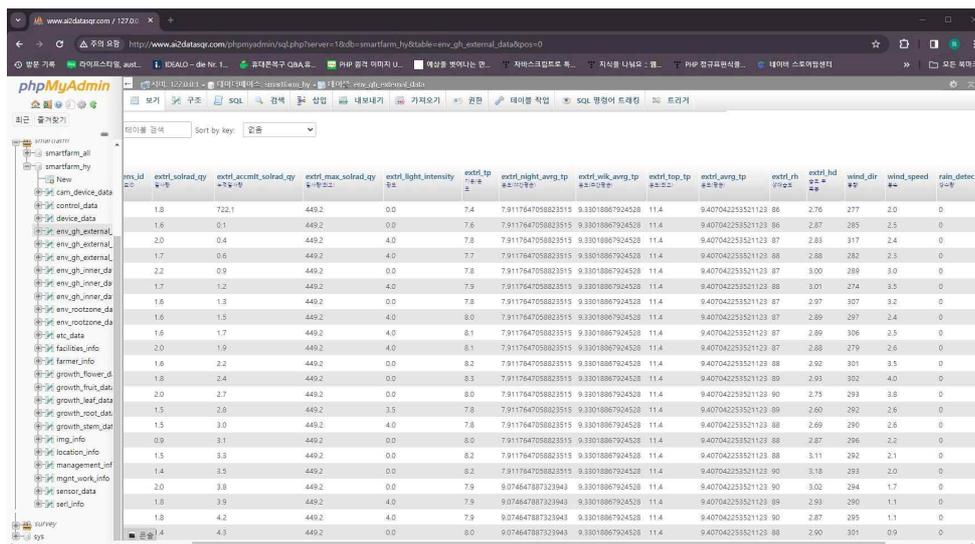
필드명	속성명	타입	키	비고
EXTRL_SOLRAD_QY	일사량	VARCHAR(20)		Weater stations 시트 - Radiation intensity [W/m2]
EXTRL_ACCMLT_SOLRAD_QY	누적일사량	VARCHAR(20)		Weater stations 시트 - Radiation sum [J/cm2]
EXTRL_MAX_SOLRAD_QY	일사량(최고)	VARCHAR(20)		계산해서 반영(1일 기준)
EXTRL_TP	기온/온도	VARCHAR(20)		Weater stations 시트 - Outside temperature
EXTRL_NIGHT_AVRG_TP	온도(야간평균)	VARCHAR(20)		계산해서 반영 (1일 기준, 0시~6시, 18시~24시)
EXTRL_WIK_AVRG_TP	온도(주간평균)	VARCHAR(20)		계산해서 반영(1일 기준, 6시~18시)
EXTRL_TOP_TP	온도(최고)	VARCHAR(20)		계산해서 반영(1일 기준)
EXTRL_AVRG_TP	온도(평균)	VARCHAR(20)		계산해서 반영(1일 기준)
EXTRL_RH	상대습도	VARCHAR(20)		Soil heating 시트 - RH [%] ※ HD : 습도 부족분(Humidity Deficit), 상대습도(RH)와 구별 필요
EXTRL_HD	습도 부족분	VARCHAR(20)		필드 신규 추가 Sensor unit 시트 - HD [g/m³] Weater stations 시트 - HD [g/kg] Weater stations 시트 - HD [g//m3] → 우선 Weater stations 시트 - HD [g/kg] 수집 필요시 다른 항목 수집
WIND_DIR	풍향	VARCHAR(20)		Weater stations 시트 - Wind direction
WIND_SPEED	풍속	VARCHAR(20)		Weater stations 시트 - Wind speed [m/s]
RAIN_DETECTOR	강우	VARCHAR(20)		Weater stations 시트 - Rain [0-1]
MEAS_DATE	측정일시	DATETIME		각 시트별 첫 번째 항목

<작물환경-온실내부 데이터 매핑 테이블 예시>

필드명	속성명	타입	키	비고
INNER_TP	온도	VARCHAR(20)		Soil Heating 시트 - Air Temperature
INNER_TOP_TP	온도(최고)	VARCHAR(20)		계산해서 반영(1일 기준)
INNER_LWET_TP	온도(최저)	VARCHAR(20)		계산해서 반영(1일 기준)
INNER_AVRG_TP	온도(평균)	VARCHAR(20)		계산해서 반영(1일 기준)
INNER_DP_TP	이슬점 온도	VARCHAR(20)		Sensor unit 시트 - Dew point temperature
INNER_Hum	습도	VARCHAR(20)		※ HD : 습도 부족분(Humidity Deficit), 상대습도(RH)와 구별

				필요
INNER_RELATIVE_Hum	상대습도	VARCHAR(20)		Sensor unit 시트 - RH [%]
INNER_ABSLT_Hum	절대습도	VARCHAR(20)		Sensor unit 시트 - Absolute humidity [g/kg]
INNER_NIGHT_AVRG_RH	상대습도(야간평균)	VARCHAR(20)		계산해서 반영 (1일 기준, 0시~6시, 18시~24시)
INNER_WIK_AVRG_RH	상대습도(주간평균)	VARCHAR(20)		계산해서 반영 (1일 기준, 6시~18시)
INNER_TOP_RH	상대습도(최고)	VARCHAR(20)		계산해서 반영 (1일 기준)
INNER_LWET_RH	상대습도(최저)	VARCHAR(20)		계산해서 반영 (1일 기준)
INNER_AVRG_RH	상대습도(평균)	VARCHAR(20)		계산해서 반영 (1일 기준)
INNER_HD	수분부족분	VARCHAR(20)		Sensor unit 시트 - HD [g/kg]
INNER_CO2	CO2농도	VARCHAR(20)		co2 group 시트 - CO2 concentration [ppm]
INNER_Cooling_TP	냉방 온도	VARCHAR(20)		Block cooling 시트 - Air temperature
INNER_Heating_TP	냉방 온도	VARCHAR(20)		Block heating 시트 - Air temperature
INNER_SUPLY_QY_PERR_DRIP	드리퍼당 공급량	VARCHAR(20)		Drain water 시트 - Irrigation (dripper) [ml/dripp]
MEAS_DATE	측정일시	DATETIME		각 시트별 첫 번째 항목

- 데이터스키마 기반 과채류(토마토, 파프리카, 고추) 및 엽채류(고수, 로메인, 바질, 피얼리, 이자벨, 스텔릭스)의 온실 환경정보 및 생육정보의 데이터베이스를 구축



<작물환경-온실외부 데이터베이스 구축사례(과채류)>

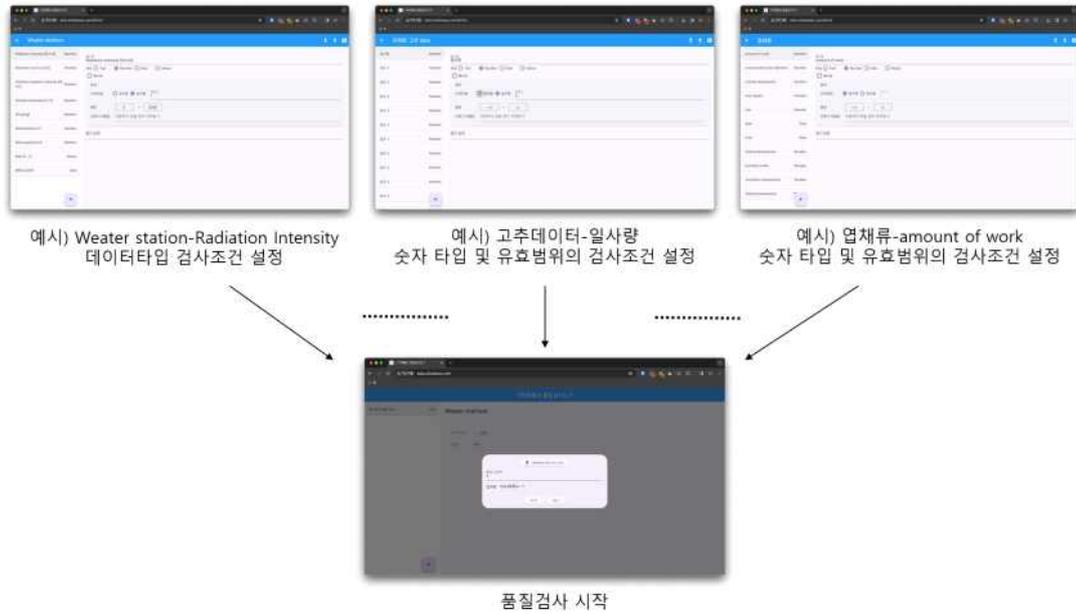
no	inner_top_tp	inner_level_tp	inner_avg_tp	inner_dp_tp	inner_hum	inner_avg_hum	inner_relative_hum	inner_abot_hum	inner_night_avg_partn_rh	inner_wik_avg_partn_rh	inner_top_partn_rh	inner_level_partn_rh	inner_avg_partn_rh	inner_hd
23.7	18.2	19.614084507942554	24.3	0.00	0.00	19.25	100	96.7792452830189	100	80.4	96.87042233521128	0.00	496	
23.7	18.2	19.614084507942554	23.9	0.00	100.0	18.78	100	96.7792452830189	100	80.4	96.87042233521128	0.00	454	
23.7	18.2	19.614084507942554	24.1	0.00	100.0	18.95	100	96.7792452830189	100	80.4	96.87042233521128	0.00	476	
23.7	18.2	19.614084507942554	24.1	0.00	100.0	18.97	100	96.7792452830189	100	80.4	96.87042233521128	0.00	492	
23.7	18.2	19.614084507942554	23.9	0.00	100.0	18.78	100	96.7792452830189	100	80.4	96.87042233521128	0.00	488	
23.7	18.2	19.614084507942554	24.0	0.00	100.0	18.82	100	96.7792452830189	100	80.4	96.87042233521128	0.00	490	
23.7	18.2	19.614084507942554	24.1	0.00	100.0	19.00	100	96.7792452830189	100	80.4	96.87042233521128	0.00	496	
23.7	18.2	19.614084507942554	24.1	0.00	100.0	19.98	100	96.7792452830189	100	80.4	96.87042233521128	0.00	501	
23.7	18.2	19.614084507942554	24.1	0.00	100.0	19.00	100	96.7792452830189	100	80.4	96.87042233521128	0.00	488	
23.7	18.2	19.614084507942554	24.1	0.00	100.0	19.02	100	96.7792452830189	100	80.4	96.87042233521128	0.00	501	
23.7	18.2	19.614084507942554	24.1	0.00	100.0	18.98	100	96.7792452830189	100	80.4	96.87042233521128	0.00	513	
23.7	18.2	19.614084507942554	24.1	0.00	100.0	18.98	100	96.7792452830189	100	80.4	96.87042233521128	0.00	513	
23.7	18.2	19.614084507942554	24.2	0.00	100.0	19.13	100	96.7792452830189	100	80.4	96.87042233521128	0.00	513	
23.7	18.2	19.614084507942554	24.3	0.00	100.0	19.16	100	96.7792452830189	100	80.4	96.87042233521128	0.00	516	
23.7	18.2	19.614084507942554	24.3	0.00	100.0	19.25	100	96.7792452830189	100	80.4	96.87042233521128	0.00	524	
23.7	18.2	19.614084507942554	24.8	0.00	100.0	19.78	100	96.7792452830189	100	80.4	96.87042233521128	0.00	505	
23.7	18.2	19.614084507942554	25.4	0.00	100.0	20.58	100	96.7792452830189	100	80.4	96.87042233521128	0.00	504	
23.7	18.2	19.614084507942554	26.4	0.00	100.0	21.89	100	96.7792452830189	100	80.4	96.87042233521128	0.00	501	
23.7	18.2	19.614084507942554	26.7	0.00	100.0	22.20	94.4169140845067	96.7792452830189	100	80.4	96.87042233521128	0.00	509	
23.7	18.2	19.614084507942554	27.5	0.55	68.1	23.35	94.4169140845067	96.7792452830189	100	80.4	96.87042233521128	0.48	488	
23.7	18.2	19.614084507942554	27.9	2.01	93.7	23.93	94.4169140845067	96.7792452830189	100	80.4	96.87042233521128	1.74	480	
23.7	18.2	19.614084507942554	28.5	3.21	90.4	24.77	94.4169140845067	96.7792452830189	100	80.4	96.87042233521128	2.75	472	
23.7	18.2	19.614084507942554	28.9	4.43	87.8	25.43	94.4169140845067	96.7792452830189	100	80.4	96.87042233521128	3.88	467	
23.7	18.2	19.614084507942554	29.0	5.32	86.4	26.02	94.4169140845067	96.7792452830189	100	80.4	96.87042233521128	4.81	444	

<작물환경-온실내부 데이터베이스 구축사례(엽채류 6종)>

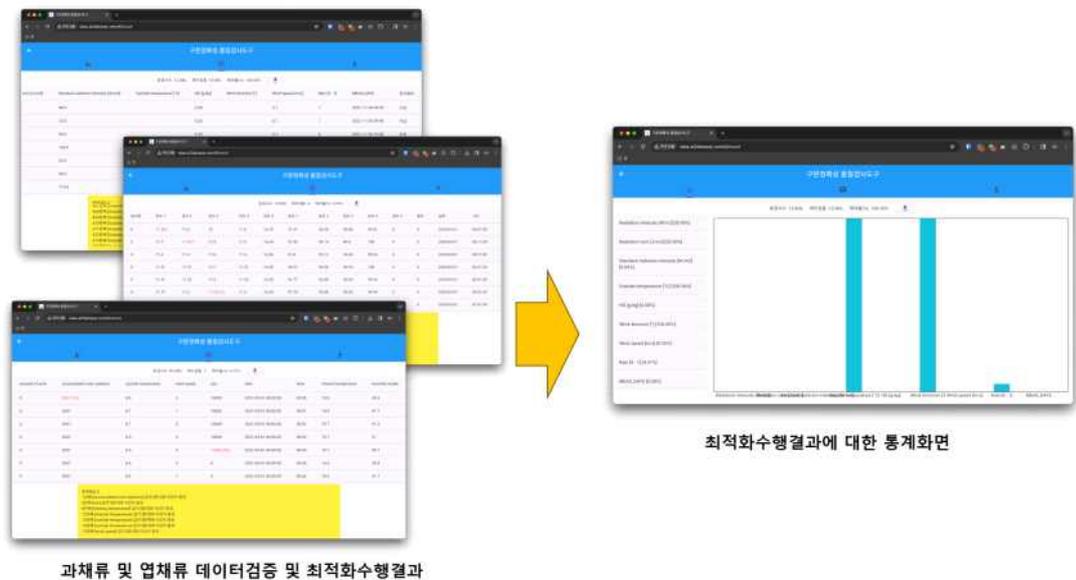
no	dev_id	cam_id	serl_no	serl_id	img_no	harvest_fruits_num	harvest_group	harvest_fruits_avg_wt	fruits_set_cnt	etc	meas_date
1	1	1	1	1	1	1	1	1	1		2021-08-11 00:00:00
2	1	1	1	1	1	1	1	5.625	1		2021-09-18 00:00:00
3	1	1	1	1	1	1	1	8.5	1		2021-09-25 00:00:00
4	1	1	1	1	1	1	1	8.25	1		2021-09-01 00:00:00
5	1	1	1	1	1	1	1	10.125	1		2021-09-08 00:00:00
6	1	1	1	1	1	1	1	15.875	1		2021-09-15 00:00:00
7	1	1	1	1	1	1.15625	1	15.125	1		2021-09-22 00:00:00
8	1	1	1	1	1	2.46875	1	15.125	1		2021-09-29 00:00:00
9	1	1	1	1	1	3.25	1	15.5	1		2021-10-06 00:00:00
10	1	1	1	1	1	3.625	1	16.375	1		2021-10-13 00:00:00
11	1	1	1	1	1	4.15625	1	17.375	1		2021-10-20 00:00:00
12	1	1	1	1	1	5.5	1	14	1		2021-10-27 00:00:00
13	1	1	1	1	1	6.25	1	17.875	1		2021-11-03 00:00:00
14	1	1	1	1	1	7.28125	1	18.625	1		2021-11-10 00:00:00
15	1	1	1	1	1	8.125	1	18.625	1		2021-11-17 00:00:00
16	1	1	1	1	1	8.78125	1	16.6666666	1		2021-11-24 00:00:00
17	1	1	1	1	1	9.625	1	19	1		2021-12-01 00:00:00
18	1	1	1	1	1	1	1	0	1		2022-08-10 00:00:00
19	1	1	1	1	1	1	1	1	1		2022-08-17 00:00:00
20	1	1	1	1	1	1	1	3.875	1		2022-08-24 00:00:00
21	1	1	1	1	1	1	1	6.25	1		2022-08-31 00:00:00
22	1	1	1	1	1	1	1	6.25	1		2022-09-07 00:00:00
23	1	1	1	1	1	1	1	8.625	1		2022-09-14 00:00:00
24	1	1	1	1	1	1	1	11.25	1		2022-09-21 00:00:00
25	1	1	1	1	1	1	1	15.25	1		2022-09-28 00:00:00

<작물생육-꽃데이터 데이터베이스 구축사례>

- 데이터베이스 스키마 변경 및 품질관리 도구 수정 · 고도화를 통한 데이터베이스 구조 · 기능 보완 및 최적화
  - 과채류 · 엽채류의 특성을 반영한 생육정보 · 환경정보 및 반밀폐 관련 제어정보 관련 데이터 추가 · 수정 · 삭제로 인해 데이터스키마가 변경됨에 따라 이를 고려한 품질관리 도구 기능을 수정 및 보완
  - 보완된 품질관리 도구를 통한 실증 데이터 검증 및 최적화 수행



<구문상태검증 조건설정(데이터타입, 유효범위등) 후 품질검사시작>



<과채류 및 업채류데이터의 품질관리 수행 결과화면>

□ 스마트 팜 빅데이터 플랫폼 구조 조사 및 데이터 전송 기능 설계·구현

○ 동 사업의 데이터 공유를 위한 스마트 팜 빅데이터 플랫폼 구조 조사

- 스마트팜연구개발사업단의 "스마트팜 R&D 빅데이터 플랫폼 연구데이터 제출 및 작성안내서"를 통해 데이터 공유를 위한 스마트 팜 빅데이터 플랫폼 구조 조사
- 동 작성안내서의 연구데이터 분류·파일명 생성규칙·데이터셋 정의·항목정보·비정형 파일 제출 등 세부 연구데이터 제출 내용 검토
  - (연구데이터 분류 - 대분류) 시설원예, 축산, 노지, 기타
  - (연구데이터 분류 - 연구분야) 생육, 병충해, 질병, 환경, 경영, 사양, 생산, 기타

- (파일명 생성규칙) 파일명명 규칙 “[세부과제번호]\_[파일명]\_[생성일자]\_[순번]”, 파일명 200자 이내, 생성일자 8자리(YYYYMMDD) 등 준수
- (데이터셋- 기본정보) 분야, 연구분야, 연구품목, 품종, 엑셀파일제목, 자료목적, 수집주기, 키워드 정보
- (데이터셋- 엑셀파일) 데이터 수집 정보의 엑셀파일 저장을 위한 규칙
- (항목정보) 엑셀파일 타이틀의 설명정보
- (첨부파일) 엑셀파일 관련 이미지·그래프·문서 등 원본 파일

○ 스마트 팜 빅데이터 플랫폼의 데이터 공유 가이드라인을 통한 온실 환경 데이터 전송 방안 마련

- "스마트팜 R&D 빅데이터 플랫폼 연구데이터 제출 및 작성안내서"의 연구데이터 분류·파일명 생성규칙·데이터셋 정의·항목정보·비정형 파일 제출 등 세부 기준을 고려하여 정형 및 비정형 데이터 업로드
- 엑셀파일 관련 이미지, 도면, 음향, 문서 등은 해당 엑셀파일을 업로드할 때 함께 첨부파일로 업로드

#### 주 의

1. 이 보고서는 농림축산식품부·과학기술정보통신부·농촌진흥청에서 시행한 “스마트팜 다부처 패키지 혁신기술개발사업”의 “과제명” 연구개발과제 최종보고서입니다.
2. 이 연구개발 내용을 발표하는 때에는 반드시 농림축산식품부·과학기술정보통신부·농촌진흥청((재)스마트팜연구개발사업단)에서 시행한 “스마트팜 다부처 패키지 혁신기술개발사업”의 “과제명” 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀 유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 안 됩니다.