

복합환경정보 및 증강 기법을 활용한 작물 병해 영상 분류시스템 설계 및 구현

김태현*, 백정현*, 임동혁*, 김만중*, 박성진*, 이정호^o

Design and Implementation of Crop Disease Image Classification System Using Complex Environmental Information and Data Augmentation

Tae-Hyun Kim*, Jeong-Hyun Baek*, Dong-Hyeok Im*,
 Man-Jung Kim*, Seong-Jin Park*, Jeong-Ho Lee^o

요약

본 논문은 표준 기반의 인공지능 복합환경제어시스템을 이용하여 온실 내부 환경정보와 작물 영상 데이터를 자동으로 취득하였다. 또한 취득된 영상 데이터와 환경정보를 이용해 병해 발생 예측 정보와 일사량, 광투과율, 산란광 등의 광 환경 정보를 활용하여, 온실 실증 환경에서 쉽게 검출하기 힘들었던 작물 영상을 분석하였다. 작물 병해 영상 분석 시 영상 변환 증강기법과 파라미터 미세 조절 등 전처리 기술 적용을 통해 기존 실증 환경에서 2020년도에 92.5%였던 병해 검출율을 95.2%로 향상시켰다. 이는 비교 대상이었던 keras를 활용한 딥러닝 병해 예측 모델 검출 정확도 최대치인 89%와 비교했을 경우, 6.2%p 이상 향상됐고, 실험실에서 조절된 환경이 아닌 외부 환경에서 취득한 unseen data를 이용한 것을 감안할 경우(Practical outcome), 매우 높은 병해 분류 정확도를 보여준다고 할 수 있다.

키워드 : 스마트농업, 작물병해, 전처리, 지능형, 환경정보

Key words : smart agriculture, crop disease, preprocessing, intelligence, environmental information

ABSTRACT

This paper uses a KS standard-based artificial intelligence complex environmental control system. In the automatically acquired crop image data, the environmental information from inside the greenhouse is used to predict the occurrence of diseases. In addition, by using light environment information such as insolation, light transmittance, and scattered light, crop diseases that were difficult to detect in the greenhouse demonstration environment were solved by applying pre-processing techniques such as the image conversion enhancement technique and fine parameter adjustment. As a result, the disease detection rate, which was 92.5% in 2020 in the

※ 본 연구는 (재)스마트팜연구개발사업단의 스마트팜다부처패키지혁신기술개발사업의 지원을 받아 연구되었음(과제번호: 421005-04)

• First Author : Division of Smart Farm Development, Department of Agricultural Engineering, National Institute of Agricultural Sciences, thkim8205@korea.kr, 정희원

o Corresponding Author : Division of Smart Farm Development, Department of Agricultural Engineering, National Institute of Agricultural Sciences, ljh09150915@korea.kr, 정희원

* Division of Smart Farm Development, Department of Agricultural Engineering, National Institute of Agricultural Sciences, butterfly@korea.kr, 정희원; imdh1004@korea.kr; kmj0403@korea.kr; psj8475@korea.kr, 정희원

논문번호 : 202307-012-0-SE, Received July 17, 2023; Revised September 21, 2023; Accepted October 24, 2023

existing demonstration environment, was raised to 95.2%. This was improved by more than 6.2%p when compared with the 89% maximum detection accuracy of the deep learning disease prediction model using keras. In other words, considering that the unseen data acquired from an external environment rather than an environment controlled by a laboratory was used (Practical outcome), it can be said that the disease classification accuracy is very high.

I. 서 론

최근 인공지능 모델의 발전에 힘입어 온실 내의 작물의 생체정보와 환경 정보를 융합하여 병해 진단과 작물 관리에 적용하여 온실 운영 및 작물 생산의 효율성 증대와 관련한 연구가 많이 진행되고 있다. 이 중 온실 내의 작물 생산 관련 농업 분야의 경우 질병 검출 및 분류^[1], 최적 환경 조건을 찾아내는 작물의 표현체 분석^[2], 재배 현황 분석을 위한 환경 정보 메타데이터 생성^[3] 등을 머신러닝 기법을 이용하여 수행함으로써 실시간으로 환경과 작물의 상호작용을 통한 피드백과 보상 관계 등을 통해 생산력과 이윤의 증대를 추구하고 있다. 병해에 의한 농작물 피해를 최소화하기 위한 작업의 자동화를 위해서는 작물 영상 및 환경 정보를 자동으로 취득하는 시스템과 분류 작업을 수행할 수 있는 진단 모델이 있어야 한다. 현재 영상 분류 작업에는 CNN(Convolution Neural Network)^[4] 모델이 좋은 성능을 나타내어 널리 사용되고 있다. 작물의 질병을 분류하는 문제를 사람이 파악하여 해결한다고 가정했을 때, 질병의 병징이 나타난 위치 정보와 더불어 해당 병징이 나타난 환경 정보를 활용하는 것이 좀 더 명확하게 분류하는데 도움이 될 수 있다. 예를 들어 영상의 경우 CNN 분석을 할 경우 GAP(Global Average Pooling)을 수행함으로써 특징 맵의 평균값을 구하는데 이 경우 공간 정보의 평균값을 사용하기 때문에 원본(RAW) 영상의 병징의 경계가 얼마나 명확하게 구분 되느냐에 따라 병징의 검출 성능을 많이 좌우하게 된다^[5]. 이를 위해 애초에 학습시킬 때 되도록 해당 병징의 많은 원본(RAW) 영상을 확보하려고 하고 이를 학습하는 모델이 다양한 학습을 할 수 있도록 데이터 증강을 통해 모델의 검출 및 분류 정확도를 높이려고 하지만 실증을 해보면 모델을 학습시킨 사진의 온실 환경과 그 외의 온실의 환경에 따라 같은 병해라도 검출 정확도에는 많은 차이가 나타나는 것이 사실이다. 본 논문에서는 이러한 부분에 착안하여 환경에 따른 영상 차이를 극복하기 위한 방안으로 이동식 영상 장치를 통해 일정시간마다 실시간으로 취득한 영상데이터와 연동되는 인공지능 기반의 복합환경제어 시스템을 이용해 작물의 영상을 실시간으로 변형, 증강

및 결과 피드백을 통해 작물의 병해 영상 분류 정확도를 높이는 시스템을 설계하였다.

II. 본 론

인공신경망 연구를 통해 객체 인식 기술이 많이 발전했으며, 특히 영상 인식을 다루는 분야에서 높은 활용성과 정확도를 보여준다. 본 연구에서 다루는 인공지능 모델은 맥락정보에 기반하여 올바른 레이블로 추정하는 것을 기본으로 하고 있다. 이를 위해 AlexNet, VGG, GoogleNet 다양한 CNN 아키텍처가 사용될 수 있다. 본 연구에서 다루고 있는 CNN은 다음에서 자세히 설명하겠지만 기본적으로 VGG16 아키텍처를 기반으로 하고 있으며, 컨볼루션 레이어와 서브 샘플링 레이어 영역을 활용한 특징 추출기가 있다. CNN의 컨볼루션은 하나의 뉴런이 인접한 층 뉴런의 일부에 연결되어 있다. 이러한 컨볼루션 레이어는 일반적으로 다중 특징 평면이 있고 각 특징면은 직사각형으로 배열된 뉴런의 수로 구성된다. 같은 특징면을 사용하는 뉴런들은 같은 가중치를 가지게 되고 이는 본 논문에서도 동일하다.

컨볼루션 커널은 일반적으로 무작위 분수 행렬의 형태로 초기화된다. 네트워크 훈련 동안, 컨볼루션 커널은 합리적인 가중치를 얻는 법을 배울 것이다.

가중치는 과적합의 위험을 줄이면서 네트워크의 계층 간 연결의 감소를 위해 존재한다. 또한 컨볼루션 및 하위 샘플링은 모델 복잡성을 크게 단순화하고 모델의 매개 변수를 줄이는데 일조한다.

본 연구에서 사용된 시스템 내 데이터 검출의 기본 구조는 입력 이미지를 다중 차원으로 나누고 학습 파라미터와 보정된 하이퍼 파라미터를 활용해 다중 차원 내에서 검출된 결과를 크게 분류 클래스와 검출 위치로 표현하도록 구현되어 있다. 그림 1에서 보이는 Faster R-CNN(Recursive Convolution Neural Network) 모델은 CNN 백본(backbone), ROI 풀링 레이어(pooling layer) 및 완전 연결 레이어(fully connected layer)로 구성되며, Region Proposal Network(RPN)는 CNN 백본에 이미지가 입력되어 실행되며, CNN 백본에서 출력되는 특징 맵(feature map)의 모든 지점에 대해 네트

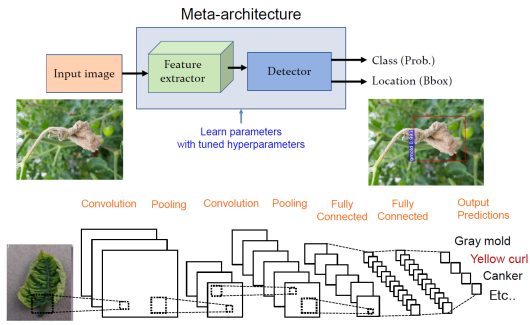


그림 1. 검출 기본 구조와 다중 검출 처리
Fig. 1. Detection basic structure and multiple detection processing

워크는 해당 위치에 입력 이미지에 개체가 있는지를 학습하고 크기를 추정해야 한다.

다중 검출을 위해 각 레이어는 가중치 내에서 조절되며 최종적으로 5개의 클래스를 분류하고 나머지는 무분류 데이터로 처리하는 구조를 가지고 있다.

또한 검출 완료 후 처리되는 바운딩 박스를 통해 병해 위치를 추정할 수 있는데, 실제 병해 위치와 추정되는 병해 위치 차이가 발생할 수 있고 해당 성능 평가를 위해 IoU(Intersection-Over Union) 기법을 통해 정량적으로 성능을 검증할 수 있다^{8,11}.

제어 클래스는 객체 검출이 시스템의 우선순위는 아니지만, 학습 중에 잠재적 장애에 대한 특징과 정보를 제공할 수 있는 클래스이며, 반면에 대상 클래스는 객체 검출 목표의 일부 클래스가 포함된다. 본 연구 수행을 위한 핵심 접근 방법은 전체 훈련 데이터 셋의 전체 카테고리(controlled)에서 모델을 학습시켜 주요 카테고리(target)의 성능을 점진적으로 향상하는 것이다^{9,10}. 객체 검출 위치의 정확도는 그림 2와 같이 IoU 인식 모델 평가지수로 실제 병징의 위치와 검출된 위치와의 차이를 가지고 검출 정확도를 평가할 수 있다.

또한 일반적으로 데이터 수의 불균등(imbalanced)은 학습에 좋지 않은 영향을 줄 수 있다¹². 따라서 데이터 셋의 구성함에 있어서 데이터 증강은 데이터 수가 최대한 클래스와 균등한 정도가 될 수 있도록 클래스별로

(IoU) metric

$$IoU(A, B) = \left| \frac{A \cap B}{A \cup B} \right|$$

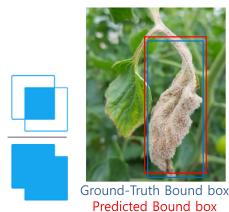


그림 2. 검출 성능 평가 지표
Fig. 2. Recognition model evaluation index

표 1. 대상 클래스와 제어 클래스에서의 주석 바운딩 박스 수
Table 1. Annotation bounding Box of Target class and control class

No.	Class	Previous dataset	New dataset
1	Leaf mold	7,389	7,178
2	Gray mold	1,010	523
3	Canker	1,657	618
4	Powdery mildew	464	6,277
5	Yellow curl	8,564	12,918
6	제어 클래스*	35,662	32,858
	TOTAL	54,746	60,372

증가량을 다르게 적용해야 한다. 전복대와 수행했던 이전 연구⁶에서 병해 진단 개선을 위해 추가 학습 데이터를 사용했고, 미분류 데이터의 정보를 신규 데이터로 추가하기 위해 특징의 분포를 파악하여 새로운 클래스로 만들거나 이전 클래스로 포함시켰다.

Ⅲ. 실시간 증강 데이터를 활용한 토마토 병해 분류기의 설계 및 구현

본 연구에서 사용된 병해영상 수집 시스템은 농촌진흥청 스마트팜개발과에서 개발된 작물 영상 취득 장치를 활용하였으며, 기본 구조는 그림 5와 같다. 영상 취득 장치는 작물의 병해를 인식하는데 사용되는 PTZ 지원 RGB 카메라와 높낮이 조절 가능한 리프트, 광측정 센서, 온습도 센서, 온실 내 자율이동을 지원하는 모바일 플랫폼으로 이루어져 있다.

본 논문의 기본 딥러닝 구조는 VGG-16 특징 추출기(feature extractor)의 Faster R-CNN 구조를 사용한다.

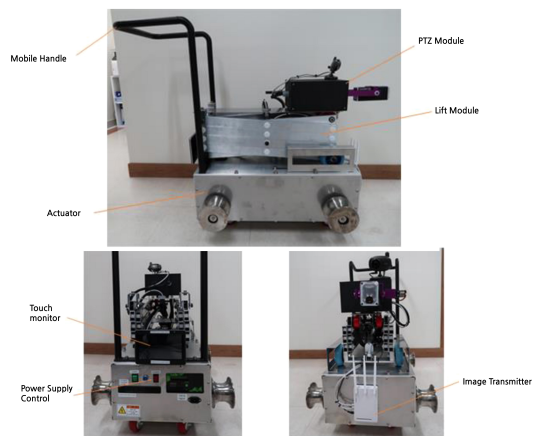


그림 3. 영상 정보 수집 장치 세부
Fig. 3. Video information collection system

이 Faster R-CNN은 CNN 백본, ROI 풀링 계층 (pooling layer) 및 완전 연결 레이어로 구성되며, 분류 및 바운딩 박스 회귀(regression)를 위한 두 개의 브랜치가 있다.

그림 4에서 보듯이, RPN은 CNN 백본에 이미지가 입력되어 실행되며, CNN백본에서 출력되는 특징 맵의 모든 지점에 대해 네트워크는 해당 위치의 입력 이미지에 개체가 있는지를 학습하고 크기를 추정해야 한다. RPN의 바운딩 박스 제안은 백본 특징 맵에서 ROI 풀링 계층에 의해 특징을 풀링하는데 사용되며, ROI 풀링 레이어는 기본적으로 a) 백본 특징 맵의 제안에 해당하는 영역을 선택하고, b) 이 해당 영역을 고정된 수의 하위 윈도우로 나누기, c) 고정된 크기 출력을 제공하기 위해 하위 윈도우를 통해 맥스 풀링(max pooling)을 수행한다.

이후 완전 연결 된 두 개의 레이어를 통과하면 특징은 Sibling Layer의 분류 및 회귀 분기로 공급되며, 분류 및 회귀 손실 함수를 사용하여 end-to-end 방식으로 학습이 수행된다. 아래 그림 5은 Faster R-CNN의 Meta-Architecture를 보여주고 있다.

현재 구현된 모델은 토마토 병해에 대해 케양병(Canker), 잎곰팡이(Leafmold), 잿빛곰팡이(Graymold), 흰가루병(Powdery mildew), 황화잎말림 바이러스(Yellow curl)에 대해 검출이 가능하다.

본 연구에서 객체검출 모드의 실험은 다음과 같이

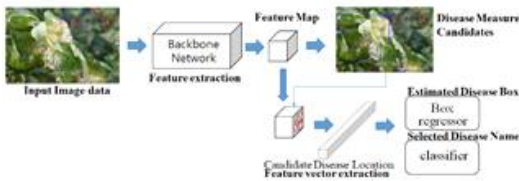


그림 4. 객체 검출 추론 과정
Fig. 4. Object detection inference process

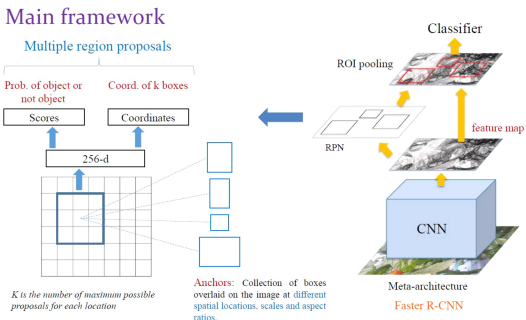


그림 5. Faster R-CNN의 메타 아키텍처
Fig. 5. Faster R-CNN's Meta-Architecture

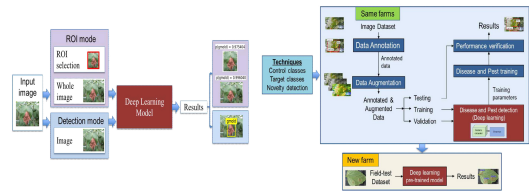


그림 6. 토마토 병해 진단 모델
Fig. 6. Tomato disease diagnosis model

수행하였다. 먼저 기존 data set의 바운딩 박스와 라벨을 수정한 후 baseline data set으로 사용하여, 전체 모델을 훈련한 후 성능을 평가하였다. 다음 그림은 baseline data set의 검출 결과를 보여준다.

본 논문에서는 병해 데이터를 수집할 때 온실 정보를 함께 수집하여 온실 정보를 토대로 영상의 병해 데이터를 변형, 증강시켜 분류하는 시스템을 고안하였다. 제안된 설계안은 병해 영상 자동 취득 장치와 연동되는 인공지능 모델이 탑재된 표준 기반[15]의 복합환경제어 시스템을 활용하여 실시간으로 영상과 환경 데이터를 취득하여 분석하는 시스템이며, 이를 위해 Faster RCNN 기반 병해 영상 분류기와 작물 영상 취득 장치, 아두이노 레퍼런스 보드를 활용하여 센서노드와 인공지능 기반 복합환경제어시스템을 올린 시범장치를 구성하였다.

또한 본 연구에 사용되는 전처리 이미지를 위한 클라우드 스마트팜 시스템과 병해진단 서비스의 통합DB 적용을 위한 스마트팜 시스템의 DB에 대한 개체 관계 모델링을 수정 개발 진행하여 향후 스마트팜 시스템 데이터와 병해진단 전처리 데이터 및 신규 병해 클래스와의 유기적인 서비스 구성이 가능하도록 설계를 진행하였다. 병해진단 관련 정보의 DB 통합은 스마트팜의 생육 및 환경정보 그리고 경영체 및 시설정보와 같이 병해 발생과 밀접한 연관성이 있는 정보들의 개체 관계 모델링의 수정 개발을 진행하였으며 농가 시설 및 환경을 고려한 병해진단 정보 / 결과 저장 가능한 구조로 설계되었으며 그에 대한 ERD는 아래 그림과 같다.

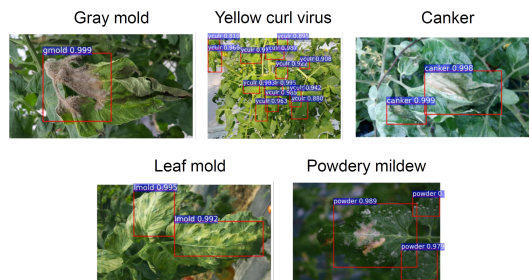


그림 7. Baseline Dataset 의 예시
Fig. 7. Example of Baseline Dataset

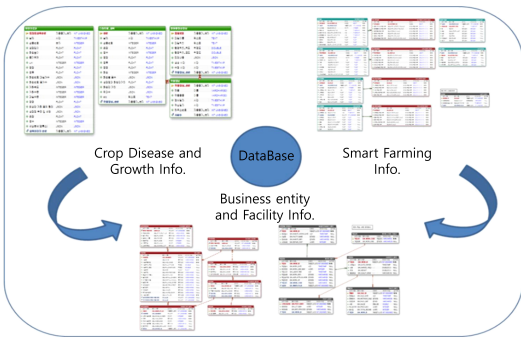


그림 8. 병해 관련 정보 통합 ERD 구성
Fig. 8. Disease-related information ERD configuration

병해 영상 장치를 인식 하기 위해 표준 기반 복합환경 제어시스템에 현행 인공지능 모델을 탑재해서 온실 환경 정보와 병해영상 데이터를 처리하여 데이터 변형, 증강 및 피드백을 통해 미검출 데이터 검출 및 병해 진단 분류 정확도를 향상시킬 수 있는 시스템을 고안하였다. 병해 영상 장치를 인식 하기 위해 복합환경제어시스템에 현행 인공지능 모델을 탑재해서 온실 환경 정보와 병해영상 데이터를 처리하여 데이터 변형, 증강 및 피드백을 통해 미검출 데이터 검출 및 병해 진단 분류 정확도를 향상시킬 수 있는 시스템을 고안하였다. 본 연구 설계에서 고려 된 복합환경제어시스템은 장치 간 호환을 위한 KS X 3267^[13]과 KS X 3286^[14]을 준수하는 표준 기반 인터페이스를 채택해 영상 장치를 PnP 방식으로 인식하고 python 코드로 구현 된 CNN 모델을 탑재하여 환경 수집 및 영상 분석이 가능한 오픈소스 기반 시스템으로 아두이노에 4채널 릴레이 모듈과 센서노드 및 영상 장치를 탑재하고 결과 분석 및 피드백 결과를 이용한 아두이노 제어는 현장의 라즈베리파이 에 구축 된 서버를 통해 진행되는 식으로 엣지컴퓨팅 기반으로 환경을 조성하였다. 기본적으로 영상을 처리 하기 위해서는 알 수 없는 새로운 병증이 들어왔을 때의 분류체계를 만들고 unknown classification에 속하는 비분류 영상데이터를 증강하여 만들었다.

본 연구의 검출과 분류 단계를 설명하면 다음과 같다.

- 이 연구는 5 가지 질병 (기준 데이터 세트)을 인식하는데 초점을 맞춤
- 딥러닝 모델은 위 5가지 질병을 인식하도록 설계됨. 그러나 Test 중 입력 이미지가 개발된 모델의 특징과 일치하지 않으면 "unknown"으로 인식됨
- "unknown" 데이터로 모델에 추정된 경우, 시스템은 도메인 전문가의 지원으로 이를 확인한 후 새로운 질

- 병에 적응할 수 있음. 이후 새로운 질병에 대응하는 더 많은 데이터가 수집되어야 함
- 확장된 데이터 집합을 빌드하기 위해 새 클래스를 기존 데이터 집합에 추가할 수 있고 딥 러닝 모델은 확장된 데이터 집합을 사용하여 학습을 수행함
- 이 부분은 기존의 딥러닝 모델에서 미세 조정된(fine tuning) 하이퍼 파라미터를 사용하여 매개 변수를 배울 수 있어야 함
- 새로운 unknown 데이터가 입력될 때마다 위 절차를 반복 수행하여 기존 모델의 확장을 도모함. 시스템은 기존 데이터 집합의 질병을 인식하며 새로운 질병은 후방으로 추가됨

상기 장치를 이용해 취득한 영상을 보정하기 위한 광 데이터는 일사량 (w/m^2) 기준으로 측정하였으며, 날씨에 따른 광환경 분석을 위해 일사량 추정식을 활용하였다. 지표면으로 입사되는 일사량의 분석은 대기권 밖 일사량 I_0 로부터 산출가능하다. 다음의 식 (1)에서 I_0 는 위도, 일적위, 시간각에 관한 함수이며 대기권 통과 전

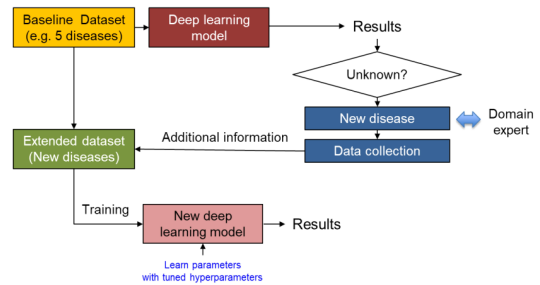


그림 9. 새로운 병증을 처리하기 위한 전략의 도식
Fig. 9. Schematic of strategies for dealing with new conditions



그림 10. 영상수집장치
Fig. 10. Image analysis device

일사량이다. 식(2)에서 계수 K_T 는 구름의 양에 관한 계수이며 대기권 통과 전 일사량 I_0 과 지표에 도달하는 일사량 I 의 비로 다음의 식으로 나타낼 수 있다⁷⁾.

$$I_0 = \frac{12 \times 3600}{\pi} G_{sc} (1 + 0.033 \cos \frac{360n}{365}) \times [\cos \phi \cos \sigma (\sin \omega_2 - \sin \omega_1) + \frac{\pi(\omega_2 - \omega_1)}{180} \times \sin \phi \sin \sigma] \quad (1)$$

- ϕ : 해당 지역의 위도 ($^\circ$, degree)
- σ : 일적위 ($\sigma = 23.45 \sin(360 \times \frac{284+n}{365})$)
- n : Julian date
- ω_1, ω_2 : 시간각 ($15^\circ/h$)
- G_{sc} : 태양상수 (1367 Wh/m^2)
- I_0 : 대기권 통과 전 일사량 (MJ/m^2)

I_0 는 위도, 일적위, 시간각을 이용한 함수 계산으로 대기권 통과 전 일사량을 추정하는 식

$$K_T = \frac{I}{I_0} \quad (2)$$

K_T 는 청명도를 나타내는 식으로 I_0 와 I 의 비로 나타낼 수 있다.

$$\frac{I_d}{I} = \begin{cases} 1.0 - 0.09K_T & 0 < K_T \leq 0.22 \\ 0.9511 - 0.1604K_T \frac{2}{T} & 0.22 < K_T \leq 0.8 \\ + 4.388K_T^2 & \\ - 16.638K_T^3 + 12.336K_T^4 & \\ 0.165 & K_T > 0.8 \end{cases} \quad (3)$$

- I : 해당지역에서 측정된 일사량 (MJ/m^2)으로 지표에 도달하는 실제 일사량값 (센서를 통해 측정됨)
- I_d : 천공 일사량 (MJ/m^2)
- K_T : Clearness

위 수식을 통해 얻어진 일사량 추정 정보에 온실 내에서 수집된 광강도 측정 정보, 직사광과 산란광의 비율, 계절 및 작물별 광포화점 정보를 반영, 병이 잘 발생하는 환경 여부를 인지하여 영상 증강 처리 여부 및 방법을 결정하여 인식률을 향상시키고자 하였다.

영상 데이터는 기본적으로 RAW를 기준으로 분석

표 2. 전처리를 위한 환경정보
Table 2. Environmental Information for Preprocessing

환경 정보	기준	판별
온실 광투과율	70%	$\pm 10 \%$
작물 광포화점	토마토: $1,400 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$	투과율 \times 일사량
계절	여름, 겨울	광포화점 증폭여부 + 계절 정보
직사광, 산란광 비율	피복재질 50%	계절, 일사량, 광투과율 + 직사, 산란비율 고려

후 10초 이내로 검출이 안될 경우 또는 검출 결과 피드백에서 유사도가 50% 이하로 나타날 경우 광 환경에 따른 일사량 추정식 및 환경 정보를 토대로 3단계(청천공, 부분 담천공, 담천공) 광 환경 정보와 병 발생 확률 정보를 이용하여, 영상 밝기 조절 및 90, 180, 270도 회전, 상하, 좌우 반전 등 영상 당 최대 8개의 증강 데이터를 활용해 검출 또는 피드백 데이터를 기준으로 10% 이상 유사도 향상이 있는지를 검증할 수 있도록 설계하였다.



그림 11. 데이터 증강 기법
Fig. 11. Data argumentation method



그림 12. 병해진단 검증 설계
Fig. 12. Design of disease diagnosis verification

IV. 실험결과

본 연구의 인공지능 기본 모델은 전문가가 검증한 1만장 이상의 토마토 병해 영상 RAW 데이터를 활용하여 8만 9천장의 증강 데이터를 포함해 병해단단 엔진을 만들었고 실증 환경에서의 병해 엔진 성능 향상을 위해 5천장 이상의 현장 RAW 데이터를 추가하여, 실제 온실 환경에서 높은 정확도의 분석이 가능하도록 점진적으로 성능을 향상시켰다. 또한 검증용 현장 데이터는 별도로 2천장 이상 추가로 수집하였다.

본 연구에 쓰인 학습용 데이터는 국립원예특작과학원 보유 영상데이터와 전북대학교에서 추가로 수집한 이미지 데이터를 포함하였다. 데이터의 주석은 충남농업기술원 토마토연구소의 도움을 얻어 정확한 병명을 부여하여 학습 데이터를 준비하였다. 그림 13와 표 3은 국립원예특작과학원 토마토 병해 이미지 수집 및 제공 내용을 보여주고 있다.



그림 13. 신규 토마토 병해 사진(일부)
Fig. 13. New disease image of tomato

표 3. 신규 토마토 병해 이미지 내역
Table 3. Descript of new tomato Disease image samples

병해명	매수	비중
케양병	507	16.0%
앞곰팡이병	921	29.1%
젓빛곰팡이	865	27.3%
토마토황화잎말림바이러스	409	12.9%
흰가루병	468	14.8%

4.1 토마토 병해영상 분류 시스템 현장데이터 활용 검증

학습에 의해 얻어진 인공지능 병해판독 추론엔진의 실증을 위해 새로운 환경과 조건에서 시스템의 성능을 검증하는 것을 목표로 새로운 데이터 셋을 구성하였다. 학습모델을 실증환경에서 검증하기 위한 7곳의 새로운 농장에서 획득되어 제공된 데이터를 사용하여 테스트를 수행하였다. 수집된 데이터 수는 표 6과 같다.

현장 테스트에는 5개의 클래스의 이미지(yellow curl, leaf mold, canker, powdery mildew)가 제공되었으며, canker, yellow curl virus의 경우, 병 발생이 잘 일어나지 않고 발병 후에는 보통 바로 전부 제거하므로 일반 농가에서 발생하는 것을 바로 현장에서 검증하기에는 어려움이 있어 부득이 기존의 학습에 사용되지 않은 unlabeled 데이터 집합의 일부를 사용하였다.

표 4. 수집 된 병해 영상 수
Table 4. Number of newly acquired data

질병종류/과일	영상 수
Tomato Yellow curl virus	139
Tomato Canker	522
Tomato Graymold	596
Tomato Leafmold	691
Tomato Powdery Mildew	431
계	2,379

4.2 검증 시험 수행 및 결과

진술한 바와 같이 변화된 환경 적응 여부를 실증 시험에 의해 확인하기 위해 기존 데이터 추출 장소가 아닌 다른 곳 또는 학습에 사용하지 않은 데이터를 사용 시험을 수행하였다. 그림은 실증 데이터의 실험 결과를 예로써 보여주고 있다.

본 논문에 쓰인 데이터셋은 기존 데이터를 보강하고 증강하여 불균형도를 개선하여 unseen data(신규 데이터)에 대한 성능을 검증하기위해 k-교차검증(k-fold cross validation)(k=5)을 수행했으며, 실증 테스트 수행 결과 평균 약 95.175%의 만족스러운 결과를 달성함으로써 전처리 하기 전 데이터 편별 시 92.5% 수준의 정확도를 상회하는 결과를 얻을 수 있었다.

기존 비교 대상이던 Keras RCNN 모델의 경우, 그림 같이 90%를 겨우 상회하는 분류 성능을 보여줬으며, 2020년에 증강 데이터와 Unknown data를 활용한 향상된 모델에서도 92.5% 정도의 분류 정확도를 보여줬다. 그러나 일사와 온실 환경 데이터를 활용해 변형, 증강 및 파라미터를 조절한 모델에서는 95.2%의 평균 분류



그림 14. 실증 데이터의 실험 결과 예시(궤양병과 황화잎말림바이러스는 기존의 학습 데이터 셋을 일부 사용)
 Fig. 14. Example of experimental results of empirical data (Canker and Yellow curl use some of the existing training data sets)

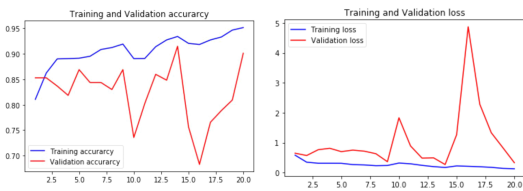


그림 15. Keras RCNN 모델 분류 정확도
 Fig. 15. Keras RCNN model classification accuracy

정확도를 보여줬다. 이는 실험실 레벨이 아닌 일반적인 온실에서 수집한 영상 데이터를 활용하여 판별한 것으로 실증 환경에서의 분류 정확도임을 감안했을 시, 매우 높은 수준이라고 볼 수 있다.

본 연구를 통해서 기존 진단 대상 클래스에서 새로운 실증 데이터를 병해 분류를 위한 추론 시스템에 적용할 때 환경 정보를 복합으로 적용해 데이터를 전처리하는 방식을 통해 병해 인식 정확도가 상승됨을 확인하였다.

V. 결론

본 연구는 기존의 병해 진단 엔진을 점진적 학습 연구로 외부 광 환경 정보와 온실 내 환경 정보 활용 병 발생 예측 정보를 조합한 전처리 기술 연구를 통해 실제 온실 환경 내에서 발생하는 병해 증상을 실증 단계에서 95.2% 수준으로 판별함으로써 실험실 수준에서 머무른 작물 병해 진단 기술을 현장에서도 실시간 전처리 기술을 통해 높은 정확도로 판별 가능하다는 것을 보여주었다. 온실 내 병해 진단 같이 주변 환경이 민감하게 판별에 영향을 미치는 경우, 점진적 학습을 통해 판별 정확도를 올리는 것은 한계가 있으며, 특정 클래스 학습 시 오히려 기존 병해 진단엔진에 영향을 미쳐 판별 정확도의 저하가 발생할 수 있으나, 본 연구의 경우, 점진적 학습 및 세부 파라미터의 조절을 통해 실증 단계에서

어느 정도 검증된 수준까지 판별 정확도를 끌어올린 후, 엔진에 영향이 미치지 않도록 환경 정보를 활용한 전처리 데이터 활용으로 매우 높은 수준의 실증 판별 정확도를 얻을 수 있었다. 향후 연구 시에는 온실 내 정보를 조합하여, 전처리 단계를 간소화하거나 다중 병검출을 예측하고 검출 수준을 낮춘 상태에서 ROI를 통해 정보를 습득한 뒤 부위별로 전처리를 다르게 하여 다중 판별 정확도를 좀 더 끌어올릴 수 있도록 추가 연구를 진행할 예정이다.

References

- [1] A. F. Fuentes, S. Yoon, J. Lee, and D. S. Park, "High-performance deep neural network based tomato plant diseases and pests diagnosis system with refinement filter bank," *Frontiers in Plant Sci.*, vol. 9, Aug. 2018. (<https://doi.org/10.3389/fpls.2018.01162>)
- [2] F. Fiorani and U. Schurr, "Future scenarios for plant phenotyping," *Annu. Rev. Plant Biology*, vol. 64, pp. 267-291, 2013. (<https://doi.org/10.1146/annurev-arplant-050312-120137>)
- [3] P. L. Suarez, A. D. Sappa, and B. X. Vintimilla, "Learning image vegetation index through a conditional generative adversarial network," 2017 *IEEE ETCM*, pp. 1-6, Oct. 2017. (<https://doi.org/10.1109/ETCM.2017.8247538>)
- [4] C. Szegedy, et al., "Inception-v4, inceptionresnet and the impact of residual connections on learning," *Thirty-first AAAI Conf. Artificial Intell.*, pp. 4278-4284, San Francisco, USA, Feb. 2017. (<https://dx.doi.org/10.14801/jkiit.2020.18.1.1>)
- [5] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv.1312.4400*, 2013. (<https://doi.org/10.48550/arXiv.1312.4400>)
- [6] J. H. Lee, T. H. Kim, J. H. Park, and J. Y. Lee, "Enhancing the UI for web-based tomato disease diagnosis installment of the system for expert's application," *Enhancement and Industrialization of the 1st Generation Smart Farm Report*, IPET, Jul. 2020.

- [7] S. T. Jang and S. J. Jang, "Exploration of a light shelf system for multi-layered vegetable cultivation," *J. Korea Inst. Ecological Architecture and Environ.*, vol. 13, no. 2, pp. 61-66, 2013.
(<https://doi.org/10.12813/kieae.2013.13.2.061>)
- [8] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Sci.*, vol. 7, 1419, 2016.
(<https://doi.org/10.3389/fpls.2016.01419>)
- [9] A. Fuentes, S. Yoon, and D. S. Park, "Deep learning-based phenotyping system with global description of plant anomalies and symptoms," *Frontiers in Plant Science*, vol. 10, 2019.
(<https://doi.org/10.3389/fpls.2019.01321>)
- [10] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, 2022.
(<https://doi.org/10.3390/s17092022>)
- [11] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Netw. and Learn. Syst.*, vol. 30, no. 11, pp. 3212-3232, 2019.
(<https://doi.org/10.1109/TNNLS.2018.2876865>)
- [12] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. CVPR*, pp. 2117-2125, 2017.
(<https://doi.org/10.48550/arXiv.1612.03144>)
- [13] National Radio Research Agency, *KS X 3267 RS485 MODBUS interface between sensor/ actuator node and greenhouse controller in smart greenhouse*, 2022.1.11.
(from <https://www.rra.go.kr>)
- [14] National Radio Research Agency, *KS X 3286 RS485/ MODBUS-based smart greenhouse node/device registration procedures and description specification*, 2022.1.11.
(from <https://www.rra.go.kr>)

김 태 현 (Tae-Hyun Kim)



2011년 2월 : 세종대학교 컴퓨터 공학과 박사수료
2019년 1월 : (주)아이앤씨테크 놀로지 책임연구원
2019년 2월~현재 : 농촌진흥청 국립농업과학원 연구사
<관심분야> HCI, 인공지능, 스마트팜

[ORCID:0000-0002-4478-667X]

백 정 현 (Jeong-Hyun Baek)



2016년 2월 : 군산대학교 컴퓨터 정보공학 박사
2020년 1월 : 농촌진흥청 국립농업과학원 전문연구원
2020년 2월~현재 : 농촌진흥청 국립농업과학원 농업연구사
<관심분야> 스마트팜, 클라우드 컴퓨팅, 데이터분석

[ORCID:0000-0002-5867-2171]

임 동 혁 (Dong-Hyeok Im)



1999년 2월 : 성균관대학교 생물기전공학과 석사
2005년 2월 : 성균관대학교 바이오메카트로닉스학과 박사
2016년 4월~현재 : 농촌진흥청 국립농업과학원 연구관

<관심분야> 객체지향시스템, 빅데이터, 스마트팜

[ORCID:0000-0003-1204-8274]

김 만 중 (Man-Jung Kim)



2022년 2월 : 전북대학교 기계시스템공학과 박사
2023년 1월 : 농촌진흥청 국립농업과학원 전문연구원
2023년 2월~현재 : 농촌진흥청 국립농업과학원 연구사
<관심분야> 농업 로봇, 농업자동화, 자동화시스템

[ORCID:0000-0003-0600-489X]

이 정 호 (Jeong-Ho Lee)



2015년 2월 : 성균관대학교 생명공학과 석사
2022년 8월 : 성균관대학교 바이오메카트로닉스 박사
2022년 9월~현재 : 농촌진흥청 국립농업과학원 전문연구원
<관심분야> 영상처리, 인공지능, 스마트팜

[ORCID:0009-0003-4088-7064]

박 성 진 (Seong-Jin Park)



2018년 8월 : 순천대학교 산업기계공학과 학사
2020년 2월 : 순천대학교 산업기계공학과 석사
2022년 4월~현재 : 농촌진흥청 국립농업과학원 전문연구원

<관심분야> 스마트팜, 자동화 온실, 농업로봇

[ORCID:0000-0001-8808-6686]