

# 딥러닝을 활용한 이미지 기반 온실 토마토 잎 질병 탐지 기법에 대한 연구

이철원, 안형태\*

한국전자통신연구원, \*한국전자통신연구원

e10000won@etri.re.kr, \*anten@etri.re.kr

## A Study on Image-based Plant Tomato Leaf Disease Detection Method Using Deep Learning

Lee Cheol Won, Ahn Hyeon Tea\*

Electronics and Telecommunications Research Institute(ETRI),

\*Electronics and Telecommunications Research Institute(ETRI).

### 요 약

본 논문에서는 딥러닝을 활용하여 온실 토마토의 잎 이미지에서 질병을 탐지 할 수 있는 기법을 제안하였다. 이미지 기반 온실 토마토 잎 질병 탐지 기술은 이미지 프로세싱을 통한 전처리 기법과 질병 탐지를 위한 딥러닝 기법으로 설계하였다. 온실 토마토 이미지의 상태를 조정하여 잡음을 줄이기 위해 전역 고정 이진화 기법과 지역 가변 이진화 기법을 사용하여 배경과 잎을 분할하였다. 그리고 8종의 온실 토마토에서 발생하는 질병을 탐지하기 위해서 딥러닝 기술인 컨볼루션 뉴럴 네트워크(Convolutional Neural Networks)를 활용하였다. 실험은 Plantvillage 데이터셋을 통해 질병 탐지 정확도에 대해 검증하였다. 최종적으로 제안한 기법은 토마토 잎의 8가지 질병 분류 정확도를 Recall과 F1-score 96퍼센트를 달성하였다.

### I. 서 론

이미지 프로세싱 기법은 객체의 분류, 추출 그리고 인식이 대표적인 연구 분야이다. 최근 딥러닝의 등장으로 이미지 프로세싱 기법의 객체 분류 정확도가 사람보다 높아지면서, 자율주행 자동차, 스마트 팩토리, 스마트 팜과 같이 다양한 분야에서 적용되고 있다 [1].

온실에서 작물을 재배하는 것은 시설 내에서 인위적으로 재배 환경을 조절하여 기후의 영향을 덜 받으면서 생산량을 올릴 수 있다는 장점을 가진다. 하지만 이렇게 인위적인 환경에서 자라는 작물들은 각종 질병에 취약하기 때문에 온도나 습도의 조절에 따라 질병이 쉽게 발생하고 확산 될 수 있다. 이로 인해 온실 환경에서 질병이 걸린 작물에 대한 탐지는 매우 중요하다.

현재는 질병이 걸린 작물을 찾기 위해서 농부들이 온실을 순회하며 경험에 따라 작물에 생기는 질병을 판정한다. 하지만 이러한 방법은 온실의 규모가 커질수록 사람이 직접 모든 작물들의 잎을 확인해야 한다는 어려움과 사람에 따라 질병에 대한 판단 기준이 다르다는 문제가 있다. 따라서 온실 내부의 작물 질병을 정확하게 탐지 할 수 있는 기법이 필요하다.

최근 이미지 프로세싱 기법들을 활용하여 작물의 질병을 탐지하는 다양한 연구들이 수행되어지고 있다. Anand R은 K-Means Clustering 기법을 활용하여 작물의 잎에서 CLS, TMV와 CR과 같은 질병들의 면적과 분포를 색을 기반으로 계산하여 탐지하는 연구를 수행하였다 [2]. R.Meena Prakash는 GLCM 기법을 사용하여 작물의 잎을 분할하고, SVM을 통해 분할된 작물의 잎과 질병을 학습하고 분류하는 연구를 수행하였다 [3]. Widiyanto S는 CNN을 활용하여 토마토 잎 이미지에서 5종의 질병을 분류할 수 있는 연구를 수행하였다. 이 외에도 여러 이미지 프로세싱을 통한 잎 질병 탐지 기술이 연구되었지만, 여전히 분류 정확도가 떨어진다.

본 논문에서 우리는 이러한 문제를 해결하기 위해서 토마토 잎에서 질병

을 정확하게 탐지 할 수 있는 새로운 기법을 제안한다. 기존 정확도가 떨어지는 문제를 해결하기 위해서 전역 고정 이진화 기법과 지역 가변 이진화 기법을 활용하여 잎의 상태에 대한 중요 정보들을 추출하며, 이 추출된 정보들만을 컨볼루션 뉴럴 네트워크(Convolutional Neural Networks)를 통해 학습하여 정상 상태와 7가지의 질병 상태를 분류하고 탐지하는 것으로 정확도를 향상시킨다.

2절에서는 딥러닝을 활용한 토마토 탐지 기법에 대해 다루며, 3절은 성능 평가를 위한 실험으로 구성된다. 마지막으로 4절은 결론으로 마무리 짓는다.

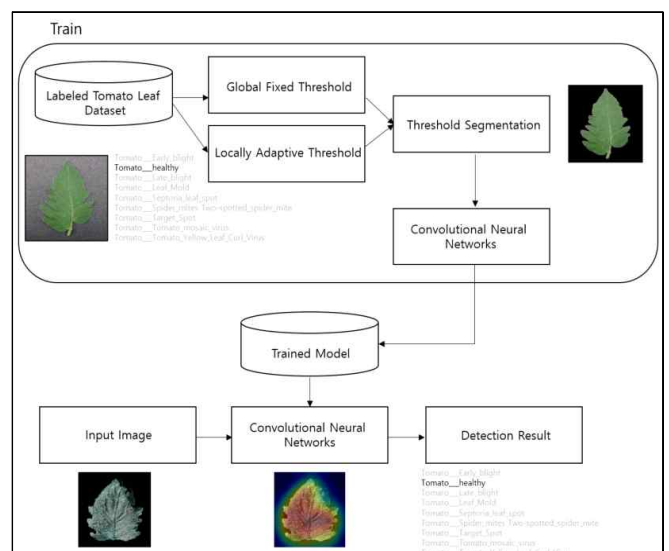


그림 1. 딥러닝을 활용한 이미지 기반 토마토 잎 질병 탐지 기법의 구조.

## II. 딥러닝을 활용한 이미지 기반 토마토 잎 질병 탐지 기법

제한한 딥러닝을 활용한 이미지 기반 토마토 잎 질병 탐지 기법은 그림 1과 같이 학습을 위한 Train 부분과 탐지를 위한 부분으로 구성되어 있다. 기본적인 아이디어는 학습하는 데이터셋의 배경에서 발생하는 잡음을 제거하여 순수하게 토마토 잎과 발생한 질병의 특징 자체를 컨볼루션 뉴럴 네트워크가 학습하도록 하는 것이다. 실제로 배경이 있는 상태로 학습을 시킬 경우 그림 2의 특징 히트맵과 같이 배경 부분도 학습하는 경우가 발생한다. 그림 2의 오른쪽 이미지는 특징 히트맵으로 붉은색 부분일수록 더 중요한 특징을 의미하며, 푸른색일수록 덜 중요한 특징을 의미한다. 이렇게 잘못된 학습된 특징들이 늘어날수록 정확도는 떨어지게 된다.

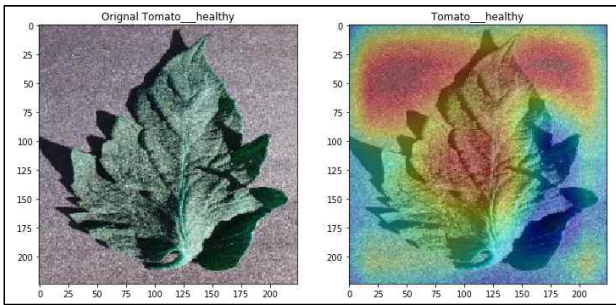


그림 2. 토마토 잎 특징 히트맵. (좌) 학습된 이미지, (우) 특징 히트맵.

따라서 이러한 잡음을 제거하기 위해 우리는 전역 고정 이진화 (Global Fixed Threshold) 기법과 지역 가변 이진화 (Locally Adaptive Threshold) 기법을 동시에 사용한다. 전역 고정 이진화 기법은 파라미터로 주어진 임계값을 중심으로 0과 1, 즉 흑과 백으로 이진화 하는 것을 통해 잡음을 필터 할 수 있다. 하지만 그림 3 (a)의 경우 질병이 걸린 부분과 연결된 부분의 색이 다르기 때문에 이 기법만으로는 (b)와 같이 배경과 질병 걸린 부분이 함께 필터링 된다. 따라서 픽셀 위치마다 다른 임계값을 지정할 수 있는 가변 지역 이진화 기법을 추가하여 (c)와 같이 토마토 잎을 분할 한다.

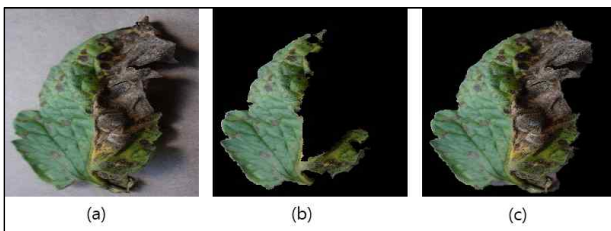


그림 3. 잎 분할 기법에 따른 차이. (a) 오리지널 이미지, (b) 전역 고정 이진화 기법을 통한 분할, (c) 지역 가변이진화와 전역고정 이진화 기법을 통한 분할.

이렇게 분할된 토마토 잎 데이터셋을 컨볼루션 뉴럴 네트워크를 통해 학습한다. 학습에 사용한 네트워크는 VGG 모델이며, 13개의 Conv레이어와 5개의 Pooling레이어 그리고 한 개의 FC레이어로 구성되어 있다. 최종적인 Softmax 출력은 Healthy, Early Blight, Late Blight, Leaf Mold, Septoria Spot, Spider Mites, Target Spot, Mosaic Virus, Yellow Curl Virus로 총 9가지의 잎 상태를 분류한다.

## III. 실험

실험에는 PlantVillage의 Tomato Leaf Disease 데이터셋을 사용하였다 [5]. 실험은 PlantVillage 데이터셋을 일반적인 컨볼루션 뉴럴 네트워크로 학습시킨 것과 제안한 기법을 통해 학습시켜 비교하였다.

일반적인 컨볼루션 뉴럴 네트워크는 Recall 및 F1-score가 각각 0.89, 0.75의 결과를 얻었다. 반면에 제안한 기법을 통해 분류한 결과는 Recall과 F1-score 모두 0.96까지 향상된 결과를 보인다. Table 1은 제안한 기법의 각 분류에 따른 성능 지표이다.

Table 1. 딥러닝을 활용한 토마토 잎 질병 분류 기법 성능

	Precision	Recall	F1-score
Early_blight	1.00	0.96	0.99
Healthy	0.98	0.99	0.98
Late_blight	1.00	0.95	0.96
Leaf_Mold	1.00	0.96	0.94
Septoria_spot	0.98	0.97	0.94
Spider_mites	0.99	0.98	0.98
Target_spot	0.97	0.97	0.93
Mosaic_virus	0.97	0.94	0.98
Yellow_curl_virus	1.00	0.95	0.96
Average	0.98	0.96	0.96

## III. 결론

본 논문에서는 전역 고정 이진화 기법과 지역 가변 이진화 기법을 통해 데이터셋의 잡음을 제거하는 방법을 통해 컨볼루션 뉴럴 네트워크의 토마토 잎 질병 분류 정확도를 향상시켰으며, 실험을 통해 성능을 검증하였다.

실제 온실에서 본 기법을 활용하기 위해서는 온실에서 수집되는 토마토 잎의 군집에서 잎을 분할하여 본 기법을 통해 실시간으로 분류하는 추가적인 연구가 필요하다.

## ACKNOWLEDGMENT

본 연구 논문은 한국전자통신연구원 연구운영지원사업의 일환으로 수행되었음. [20ZD1150, 인공지능기반 스마트팜 통합 솔루션 기술 개발]

## 참 고 문 헌

- [1] 이철원, 나대영, and 전홍석. "라즈베리파이를 활용한 비전기반 야생조류 침입 탐지 및 퇴치 시스템의 구현.", 한국컴퓨터정보학회 학술발표논문집, 26.2, pp. 506-507, Jul 2018
- [2] Anand R, Veni S, Aravinth J. "An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method.", In 2016 international conference on recent trends in information technology (ICRTIT), 8, pp. 1-6, Apr 2016.
- [3] Prakash RM, Saraswathy GP, Ramalakshmi G, Mangaleswari KH, Kaviya T. "Detection of leaf diseases and classification using digital image processing.", In 2017 international conference on innovations in information, embedded and communication systems (ICIIECS), 0, pp. 1-4, Mar 2017.
- [4] Widiyanto S, Fitrianto R, Wardani DT. "Implementation of Convolutional Neural Network Method for Classification of Diseases in Tomato Leaves.", In 2019 Fourth International Conference on Informatics and Computing (ICIC), 0, pp. 1-5, Oct 2019.
- [5] Hughes D, Salathé M. "An open access repository of images on plant health to enable the development of mobile disease diagnostics.", arXiv preprint arXiv:1511.08060, Nov 2015.