

Swin Transformer를 이용한 토마토 질병 진단 방법

김병준*, 정성환*, 김선형*, 김동훈*, 최강인*, 김희곤**, 정수호**, 윤가윤**, 박근호*

*한국전자기술연구원

**전라남도농업기술원

jun0420@keti.re.kr, shjeong@keti.re.kr, sh.kim@keti.re.kr, clickmiss123@keti.re.kr,
cki921@keti.re.kr, khg7136@korea.kr, aosi274@korea.kr, gayun526@korea.kr,
root@keti.re.kr

Tomato Diseases Diagnosis Method Using Swin Transformer

Byoungjun Kim*, Shunghwan Jeong*, Seonhyeong Kim*, Donghoon Kim*, Kangin Choi*,
Heegon Kim**, Sooho Jeong**, Ga-yun Yun**, Keunho Park*

*Korea Electronics Technology Institute

**Jeonnam Agricultural Research & Extension Service

요약

최근 정밀농업 기술 고도화로 시설 및 노지 생육환경의 영상분석 기반 작물 재배 연구가 진행되고 있으며, 딥러닝 모델의 발전에 힘입어 다양한 분야에 인공지능 모델의 적용 성공사례가 보고되고 있다. 본 논문에서는 토마토 작물의 질병 진단에 대해 Swin Transformer를 이용하여 정상과 9개 질병의 공간적인 특징 분석 및 오류 완화를 통한 질병 진단 정확도 개선에 관한 연구를 수행하였다.

1. 서론

질병으로부터 작물을 효과적으로 진단하는 것은 작물 생산성을 개선하고 품질을 향상시키는 중요한 수단으로 현대 농업 생산에서 중요한 문제로 간주하고 있다. 작물 질병은 여러 가지 원인에 의해 생성될 수 있고, 감염 상태 전반에 걸쳐 다양한 변화를 나타내며, 박테리아, 곰팡이, 바이러스 등을 포함하여 작물의 손상을 유발할 수 있다. 질병이 확산될 경우 생육하고 있는 작물들에 심각한 영향을 미칠 수 있으므로 작물 질병 진단 연구는 매우 중요하다.

최근 스마트팜 기술 확산으로 인해 정밀농업 기술 고도화로 시설 및 노지 생육환경 작물 재배의 연구가 진행되고 있으며, 딥러닝 모델의 발전에 힘입어 작물 계측 및 질병 진단 등 다양한 분야에 대해 인공지능 모델의 적용 사례가 보고되고 있다. [1]은 딸기 작물 질병 진단에 대하여 객체 검출 모델 중 FPN(Feature Pyramid Network) 모델을 다단 구조(Multi-scale structure)로 구성하여 오분류 감소 및 정확도 개선에 관한 방법을 제안하였다. [2]는 토마토 질병 진단에 대해 객체 검출 모델 중 Faster R-CNN을 통해 후보 질병 영역 검출 및 추출하여 CNN(Convolutional Neural Network)를 통한 오분류 감소에 관한 방법을 제안하였다. [3]은 카사바 작물 질병 진단에 대해 영상 분류 모델인 InceptionV3 모델을 이용

하여 분류 문제를 통한 작물 질병 진단 방법을 제안하였다. 하지만, 작물은 다양한 품종을 가지며 시설 및 노지 등 다양한 장소에서 생육하기 때문에 크기, 형태, 모양 등이 일정하지 않고 생육환경 내 광원, 조도 등의 문제로 인해 질병 진단은 여전히 어려운 문제로 남아있다.

본 논문에서는 Swin Transformer를 기반으로 토마토 작물에 대해 정상 및 질병을 촬영한 RGB 영상 내 전경과 배경정보의 공간적인 특징 분석 및 오류 완화를 통한 질병 진단 정확도 개선에 대한 방법을 제안한다.

2. 제안된 방법

2.1 학습 데이터의 구성

딥러닝 기반 작물 질병 진단 시 전경보다 배경이 복잡할 경우 오진단율이 증가하는 문제가 존재하고, 양질의 대량 질병 데이터 수집의 어려움을 해결하기 위해 파괴적 접근방법 기반한 정상 및 질병 시료를 획득하여 카메라로 촬영한 공개 데이터인 PlantVillage[4]의 Tomato Leaf Disease를 사용하였다. 사용한 공개 데이터는 토마토 작물의 잎에 대하여 Healthy, Bacterial Spot, Early Blight, Late Blight, Leaf Mold, Septoria Leaf Spot, Spider Mites, Target Spot, Mosaic Virus 등 정상과 9개의 질병으로 총 19,260장이며 학

습/검증/평가 데이터를 표1과 같이 구성하였다.



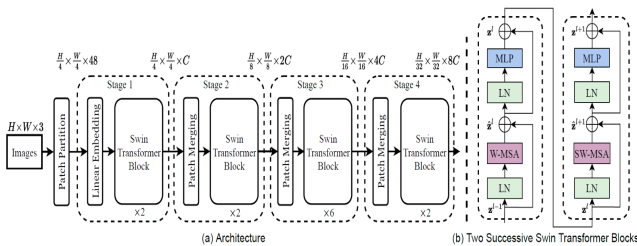
[그림 1] 토마토 질병 데이터 샘플

[표 1] 토마토 질병 진단 데이터 집합

데이터	원본	학습	검증	평가
Bacterial Spot	1,789	1,321	381	87
Early Blight	2,016	1,554	366	96
Healthy	2,022	1,561	365	96
Late Blight	1,943	1,483	368	92
Leaf Mold	1,976	1,531	351	94
Mosaic Virus	1,879	1,425	365	89
Septoria Leaf Spot	1,833	1,382	363	88
Spider Mites	1,832	1,367	374	91
Traget Spot	1,911	1,459	368	84
Yellow Leaf Curl Virus	2,059	1,595	366	98
합계	19,260	14,678	3,667	915

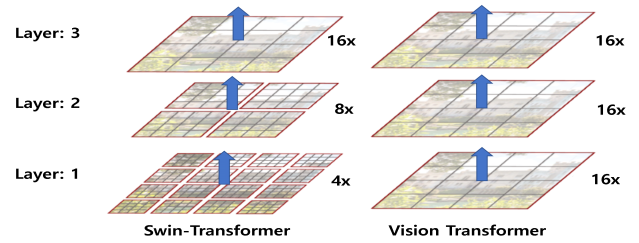
2.2 네트워크 구성

본 논문에서는 전경과 배경정보의 공간적인 특징 분석 및 오류 완화를 통한 질병 진단 정확도 개선을 위해 Swin Transformer(Shifted window Transformer)[5] 모델 중 Swin-S 기반으로 SoftMax 분류기를 통해 정상과 9개의 질병을 분류(classification)하는 구조를 사용하였다.



[그림 2] Swin Transformer 구조

Swin Transformer는 기존의 Vision Transformer의 문제점인 고정된 크기의 영상 패치(patch)를 사용하여 고해상도 및 다양한 크기의 개체에 적합하지 않은 문제를 해결한 모델로, 그림 3과 같이 작은 크기의 영상 패치 안에서만 self-attention 수행 및 레이어(layer)를 통과하면서 인접한 영상 패치들을 합쳐 계산하기 때문에 이미지 크기에 선형 비례하는 계산량과 다양한 크기 개체의 공간적인 특징 분석에 적합한 장점을 가지고 있다.



[그림 3] Swin Transformer와 Vision Transformer 구조 비교

3. 실험결과 확인 및 분석

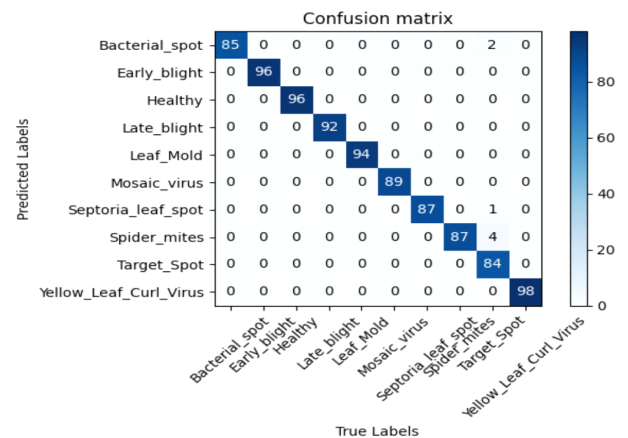
3.1 학습 및 성능평가

제안한 모델에 대해 학습, 검증, 그리고 평가를 위해 사용된 컴퓨터 환경은 Intel Xeon(r) 6230R 2.10GHz, 128 RAM, Geforce RTX 2080Ti 2대이며, Python 3.6.8, OpenCV 4.5 그리고 pytorch 1.10.2 등과 같은 개발환경에서 진행하였다.

모델 학습 시 데이터의 수가 부족하여 ImageNet1K로 사전 학습한 가중치를 이용하는 fine-tuning 기법과 매 epoch마다 데이터의 순서를 재구성(shuffle) 방법을 사용하는 학습 전략을 구성하며, 사용된 매개변수는 epoch 30, batch size 8, learning rate 0.0001, weight decay 0.00001, Adam optimizer, window size 7, patch size 4, embedding dimension 96, number of heads(3, 6, 12, 24), depths(2, 2, 6, 2) 등과 같이 구성하였다. 제안한 모델에 대하여 915개의 평가 데이터를 기반으로 성능평가 시 혼동행렬(confusion matrix)을 이용하여 정확도(accuracy) 측면으로 분석하였다.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

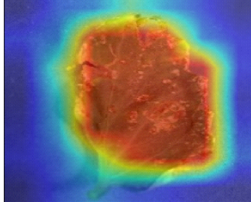
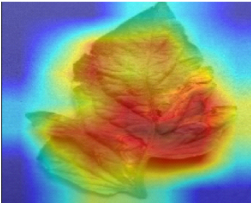
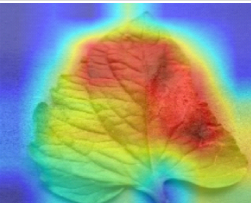
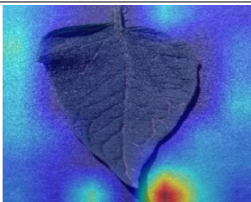
평가 데이터를 통해 제안한 모델의 질병 진단 성능 결과, 정확도 측면에서 평균 99.23%로 전경 및 배경정보의 공간적인 특징 추출과 주변 환경 오류 최소화 및 질병을 분류하는데 적합하다는 것을 알 수 있다.



[그림 4] 제안한 모델에 대한 혼동행렬 결과

3.2 Grad-CAM을 통한 질병 진단 시각화

제안한 모델의 질병 진단 관련 의사결정에 대한 정보를 분석하기 위해 Grad-CAM(Gradient-weighted Class Activation Mapping)을 이용하여 입력 영상에 대한 localization map을 생성하여 분석을 수행하였다. 분석한 결과, 질병이 존재하는 영역은 적색으로 표현되며, 질병의 형태, 모양, 그리고 크기 등의 시각화된 정보를 통하여 제안한 모델이 질병 진단에 적합하다는 것을 알 수 있다.

질병명	Grad-CAM 분석 결과	
Bacteria Spot		
Early Blight		
Late Blight		
Healthy		

[그림 5] Grad-CAM을 통한 토마토 질병 진단 시각화 결과

4. 결론

본 논문에서는 토마토 작물을 대상으로 질병 진단 정확도 개선에 관한 연구로 Swin Transformer 모델 중 Swin-S을 이용하여 정상과 9개의 질병을 분류(classification)하는 구조를 제안하였다. 915개의 평가 데이터를 이용하여 학습한 모델의 성능평가 시 평균 99.23%의 분류 성능이 도출되었고, Grad-CAM을 통한 질병 영역에 대한 localization map을 통해 질병 진단 성능을 검증하였다.

본 연구 결과를 통하여 농업 분야 인공지능 기반 영상분석 기술개발에 도움이 될 것으로 판단되며, 향후 현장 접목을 통하여 조기 질병 진단을 통한 작물 생산량 피해를 절감할 수

있을 것으로 예상된다.

ACKNOWLEDGMENT

본 결과물은 농림축산식품부, 과학기술정보통신부, 농촌진흥청의 재원으로 농림식품기술기획평가원, (재)스마트팜 연구개발사업단의 스마트팜다부처패키지혁신 기술개발사업의 지원을 받아 연구되었음(421003041HD030)

참고문헌

- [1] Kim, Byoungjun, et al. "Improved vision-based detection of strawberry diseases using a deep neural network", *Frontiers in Plant Science*, 11, 2040, 2021.
- [2] Fuentes, Alvaro F., et al. "High-performance deep neural network-based tomato plant diseases and pests diagnosis system with refinement filter bank", *Frontiers in plant science*, 9, 1162, 2018.
- [3] Ramcharan, Amanda, et al. "Deep learning for image-based cassava disease detection", *Frontiers in plant science*, 8, 1852, 2017.
- [4] Hughes, David, and Marcel Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics", *arXiv preprint arXiv:1511.08060*, 2015.
- [5] LIU, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows", *In: Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 10012-10022, 2021.