

Multi-Tasking U-net 기반 파프리카 병해충 진단

(Multi-Tasking U-net Based Paprika Disease Diagnosis)

김서정*, 김형석**

(Seo Jeong Kim, Hyong Suk Kim)

요약

본 연구에서는 Multi-Tasking U-net를 사용하여 영역 세분화 작업(Segmentation) 과 분류 작업(Classification) 이 동시에 수행되게 함으로써 파프리카 병과 충 진단을 수행하였다. 시설 농장의 파프리카에는 병의 종류가 다양하지 않다. 이 연구에서는 비교적 발생빈도가 높은 흰가루병과 응애에 의한 피해, 정상 잎 3개의 클래스에 대해서만 진단 할 수 있도록 하였다. 이를 위한 중추 모델로는 U-net을 사용하였다. 또, 이 모델의 Encoder와 Decoder의 최종 단을 활용하여 분류 작업과 영역 세분화 작업이 각 각 수행되게하여, U-net의 Encoder가 분류작업과 영역 세분화 작업에 공유되도록 하였다. 학습 데이터로는 정상 잎 680장, 응애에 의한 피해 잎 450장, 흰가루병 370장을 사용하였다. 테스트 데이터로는 정상 잎 130장, 응애에 의한 피해 잎 100장, 흰가루병 90장을 사용하였고, 이를 통한 테스트 결과로는 89%의 인식을 얻었다.

■ 중심어 : 멀티태스킹 학습 ; 딥 러닝 ; Segmentation ; 파프리카 병해충 ; 분류

Abstract

In this study, a neural network method performing both Detection and Classification of diseases and insects in paprika is proposed with Multi-Tasking U-net. Paprika on farms does not have a wide variety of diseases in this study, only two classes such as powdery mildew and mite, which occur relatively frequently are made as the targets. Aiming to this, a U-net is used as a backbone network, and the last layers of the encoder and the decoder of the U-net are utilized for classification and segmentation, respectively. As the result, the encoder of the U-net is shared for both of detection and classification. The training data are composed of 680 normal leaves, 450 mite-damaged leaves, and 370 powdery mildews. The test data are 130 normal leaves, 100 mite-damaged leaves, and 90 powdery mildews. Its test results shows 89% of recognition accuracy.

■ keywords : Multi-Tasking Learning ; Deep Learning; Segmentation ; Diagnosis of Paprika Diseases ; Classification

1. 서론

최근 들어 국내에서도 스마트팜에 관한 연구가 활발하게 진행되고 있다. 그 중, 잡초를 인식하여 무인으로 잡초를 제거해주는 연구나 작물의 잎이나 열매의 병 유무를 진단해 주는 연구들도 있다. 이 결과를 활용하면, 제초제나 병 치료제들을 쓰지 않고도 로봇을 활용하여 자연농법을 수행할 수 있을 것이다. 따라서, 스마트팜 기술이 성숙하면, 작물들이 덜 손상

되고 화학 약품으로 오염되지 않은 건강한 자연식품의 생산이 가능할 것으로 기대되어 진다.

잡초를 제거하기 위한 연구와 병의 유무를 판단해주는 인공지능 구조로는 분류(Classification) 방식 또는 객체 검출(Object Detection) 방식을 통한 연구가 진행되고 있다 [1-2].

위에서 언급한 연구를 위해서는 인공지능의 사용이 필수적이다. 인공지능기반 영상인식의 대표적인 기술로는 분류, 객체 검출, 영역 세분화(Segmentation) 등이 있다[3-5].

* 준회원, 전북대학교 전자·정보공학부 석사과정

** 정회원, 전북대학교 지능형로봇연구소 교수

이 논문은 2020년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No.NRF-2019R1A6A1A09031717)

또한, 이 논문은 농촌진흥청“농업 과학 기술 개발 협력 연구 프로그램(No. PJ01389105)”의 지원으로 수행된 연구임

접수일자 : 2020년 01월 20일

게재확정일 : 2020년 03월 04일

수정일자 : 2020년 03월 03일

교신저자 : 김형석 e-mail : Hskim@jbn.ac.kr

기존 연구에서는 영상 촬영 시 한 가지 종류의 클래스로만 채워지도록 촬영하여, 그 영상이 어떤 클래스로 분류되는 물체를 의미하는지 파악하고자 하는 분류 작업이 주가 되었으나, 최근에는 여러 가지 종류의 클래스가 포함된 영상을 대상으로 각 객체들의 위치와 함께 클래스 분류까지 병행하는 연구로 발전하고 있다.

이 논문에서는 시설 농장의 대표적 작물의 하나인 파프리카를 대상으로 그 병해충 진단을 자동화하기 위해서, 영역 세분화 작업을 통해 병해충 위치 영역을 찾아내고, 해당 병해충 영역마다 분류를 동시에 수행 할 수 있는 Multi-Tasking U-net을 제안하였다. 이 방법에서는 영역 세분화를 위한 에리 함수와 분류 작업을 위한 에리 함수를 동시에 사용하며, U-net의 Encoder 부분이 영역 세분화 와 분류 작업을 위해 동시에 가중치 들이 학습 할 수 있도록 하였다.

본 논문에서는 파프리카 잎의 병해충을 진단하는 기술을 연구하였다. 파프리카 잎은 다른 작물에 비해 크기가 큰 편임으로 가운데 잎이 크게 나오고 배경의 잎들은 다른 작물의 잎인 듯 작고 초점이 흐리게 나오는 것들이 많았다. 또한, 잎의 뒷면을 촬영하기 위해, 잎을 손으로 잡는 경우가 많아 영상에 손가락까지 함께 찍히는 경우가 많았다. 제안한 Multi-Tasking U-net는 이런 경우에도 잘 동작하여 병해충을 잘 인식 가능하였다. 그림 1은 이런 목적으로 촬영된 파프리카 병해충 잎의 예를 보여준다. 학습이 완료된 신경회로망에 대한 테스트 결과 89%라는 높은 수준의 진단 정확도를 보여 주었다.

II. 데이터 확보

이 연구를 위해서 여러 다른 장소에 위치한 파프리카 농장을 방문하여 병해충 영상데이터를 촬영 하였다. 시설농장의 파프리카는 다양한 병 발생은 하지 않지만, 흰가루병과 응애에 의한 피해는 빈번히 일어난다. 이번 연구에서는 정상적인 잎을 포함한 흰가루병 및 응애등 총 3개의 클래스를 대상으로 분류 작업과 영역 세분화 작업을 수행하였다. 흰가루병과 응애에 대한 레이블링을 위해서, 정상적인 잎은 0, 흰가루 병은 1, 응애에 의한 피해를 받은 잎은 2로 레이블링 한 뒤,

원-핫 인코딩을 진행하였다. 영역 세분화 작업을 위한 레이블링은 Pixel-Level annotation(그림 1 오른쪽 부분의 픽셀 기반으로 영역을 레이블링 한 부분)을 만들어 사용하였고, 분류 작업을 위한 레이블링은 Image Level annotation(각 클래스에 대해 0,1,2 값을 준 레이블링 값)을 사용하였다.

이번 연구에서 영역 세분화 작업의 목표는 초점이 잘 맞춰진 영상의 중앙에 위치한 잎이었으므로, 중앙 위치의 잎 영역을 흰색인 1, 그 외 영역을 검은색으로 볼 수 있는 0의 값으로 레이블링을 하였다. 그림1은 이번 논문에서 사용한 영상들과 레이블링의 예이다.

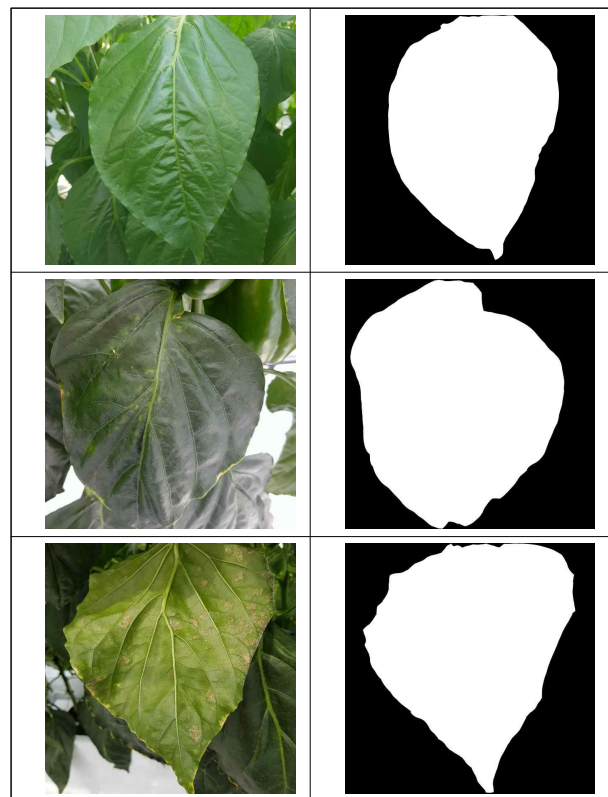


그림1. 데이터 영상의 예
(위에서부터 정상적인 잎, 응애에 의해 피해를 입은 잎, 흰가루 병이 생긴 잎)

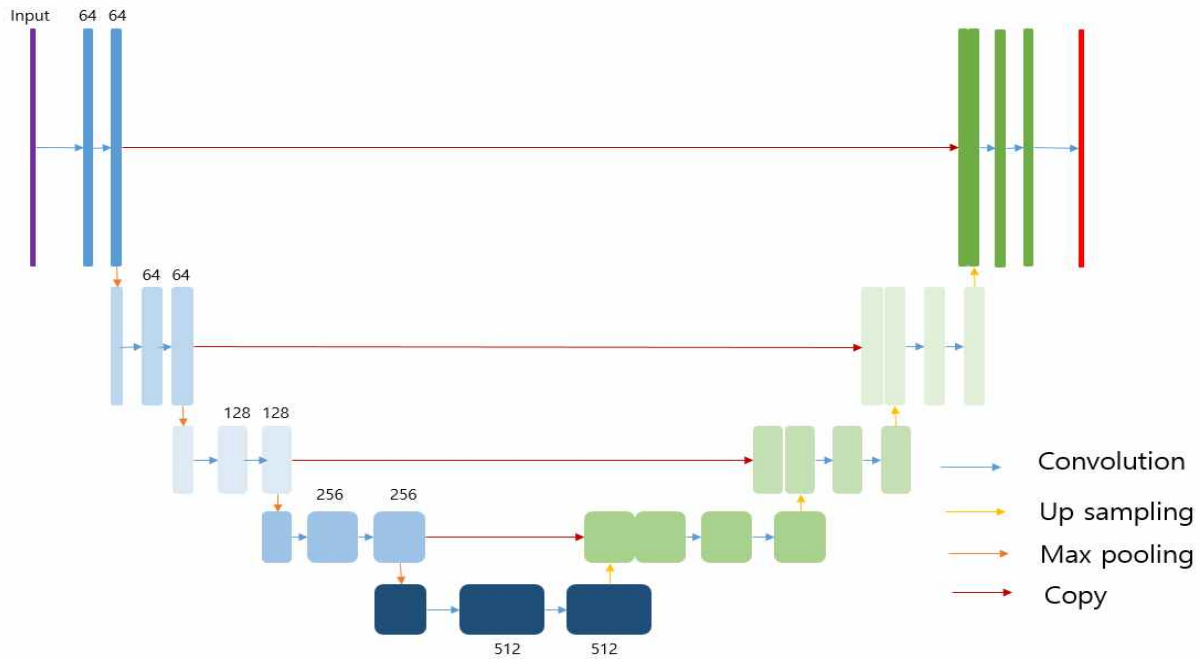


그림 2. U-net 네트워크의 구조

III. 모 델

1. U-Net

이 연구에서는 파프리카 잎의 병을 진단하기 위한 복수의 작업을 수행할 수 있도록 하기 위해서, U-net [5]를 기반으로한 Multi-Tasking U-net을 설계하였다. 모델의 구조는 그림 2와 같은 일반적으로 알려진 U-net 구조를 기본으로 확장된 구조를 하고 있다. U-net은 영역 세분화 작업을 위해 사용하는 대표적인 모델로서 Encoder 부분과 Decoder 부분 등 두 부분으로 구성되어 있다. 이 모델은 Encoder 부분에서 축소된 퓨처 맵을 Decoder 부분의 입력으로 하고, Decoder의 각 레이어 층에서는 Encoder의 동일한 레이어 층의 정보까지 결합시켜 퓨처맵 추출과 영역 세분화 작업에 좋은 결과를 보여주는 모델이다.

그림 2에서 보면 U-net은 대칭 형태의 U자 모형을 하고 있다. 일반적인 U-net은 전체적으로 컨볼루션과 디컨볼루션 레이어 층에서 각각 컨볼루션을 2번씩만 진행하며 학습을 한다. 하지만, 이번 연구에서는 Encoder 부분이 분류 작업도 진행해야 한다. 이를 위해 Encoder 부분은 [컨볼루션 - Batch normalization - Relu - 2x2 Maxpooling]를 하나의 Conv Block이라 했을 때 1, 2번 레이어 층에서는 2번의 Conv Block

3, 4, 5번 레이어 층에서는 총 3번의 Conv Block을 사용한다[3].

표 1은 제안한 모델에서 사용한 U-net의 전체적인 모습을 보여준다. 원본 영상은 R.G.B 3차원 칼라 영상이기에 처음 채널은 3차원이다. 영상들의 크기는 컴퓨터의 사양에 맞춰 최적의 상태에서 학습하기 위해 320x320 크기로 전체적으로 줄인 뒤 학습을 진행하였다. 이 때, 모든 영상은 일반화 작업(Normalize)을 통해 픽셀의 모든 값이 0 ~ 1 사이의 값으로 이뤄지도록 하였다. 첫 번째 컨볼루션을 하면서 퓨처 맵의 수는 64로 늘렸다. 이 때 Conv Block은 2번을 만을 연결하였다. 이는 VGG 네트워크의 방식을 활용하였다[3]. 이외의 파라미터는 표 1을 통해 확인할 수 있다.

퓨처 맵의 크기를 늘려주는 Decoder 부분은 사이즈를 2배씩 늘리면서 원본 크기 영상을 회복하는 걸 목적으로 한다. 파라미터 들은 표 1의 UP_1 - 5의 부분을 보면 알 수 있다. 퓨처 맵의 수는 기존 Encoder 부분에서와 똑같은 크기로 늘어나게 된다. Conv Block 역시 Encoder를 기반으로 똑같은 개수를 사용했다.

표 1. U-net Network의 전체적인 파라미터

Layer	Feature Map	size	Block 갯수
Input	3	320x320	1
Down_1	64	320x320	2
Down_2	64	160x160	2
Down_3	128	80x80	3
Down_4	256	40x40	3
Down_5	512	20x20	3
UP_1	256	40x40	3
UP_2	128	80x80	3
UP_3	64	160x160	3
UP_4	64	320x320	2
UP_5	1	320x320	2
Output	1	320x320	1

2. 제안한 Multi-tasking U-net

Multi-tasking U-net이란 두 개의 작업을 동시에 진행하는 모델이다. 먼저 위에서 언급한 일반적인 U-net을 Pixel-level로 레이블링한 목표 영상과 Cross-entropy 에러 함수를 사용하여 에러 값을 구한다.[8] 이와 동시에 U-net의 마지막 Encoder 부분 퓨처 맵에서(표 1의 Down_5) Global Average Pooling(GAP)를 통해 나온 값과 Image-level로 레이블링한 값을 통해 Cross entropy 에러 함수를 사용하여 에러를 구한다. 에러를 구하는 수식은 수식 (1)과 같다[6-7].

$$E = \lambda_1 \sum Segmentation Error + \lambda_2 \sum Segmentation Error \quad (1)$$

이번 연구에서 영역 세분화 작업에 대한 에러는 분류 작업 에러에 비해 5 - 10 배정도 높은 모습을 보였다. 이에 두 에러의 균형을 맞추기 위해 λ_1 , λ_2 를 사용하였다. 첫 번째 실험에서는 λ_1 과 λ_2 를 둘 다 1의 값을 넣어주고 사용하였고 두 번째 실험에서는 균형을 맞추기 위해 $\lambda_1 = 1$, $\lambda_2 = 5$ 의 값을 넣고 실험을 진행하였다. 전체적인 흐름은 그림3을 통해 확인할 수 있다.

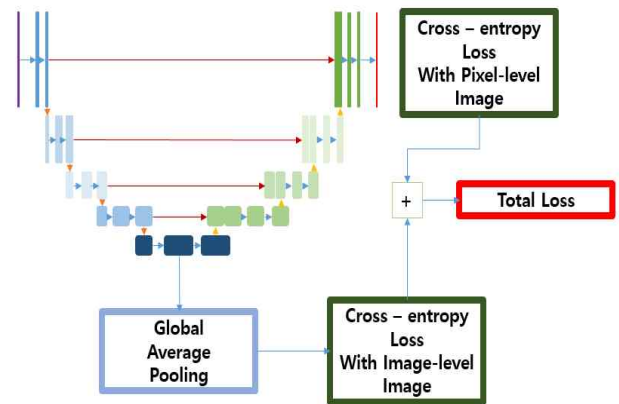


그림 3. Multi-tasking U-net의 정보 흐름 및 에러 함수

3. 모델 학습

이번 연구에서 기존 CNN모델들과 제안한 모델을 학습시키기 위한 학습 계수(Learning Rate)는 학습이 진행되면 될수록 학습 계수(Learning Rate)를 강제로 줄여주면서 학습을 진행하였다. 총 에포크의 수를 70으로 하였는데 30 에포크까지는 0.001, 50까지는 0.0005, 그 외는 0.0001의 학습 계수(Learning Rate)를 사용하여 학습이 더 잘 진행되도록 유도하였다[14]. Optimizer로는 Adam Optimizer를 사용하였고, 데이터 오그멘테이션은 플립 기술만 사용하였다. 배치 크기는 20으로 하였다. 이 때, 전이 학습 방식은 사용하지 않고 Xavier 가중치 초기화 방식 이후 학습을 진행하였다[12,13,15].

IV. 결과

파프리카 병해 영상진단을 위한 Network 선정을 위해 대표적인 Network들과 비교 연구를 수행하였다. 실험 대상 Network로는 VGG16, ResNet-50 및 Multi-Tasking U-net이었다. 표 2는 VGG16, Res-Net50 및 Multi-tasking U-net의 분류 정확도를 비교한 것이다. 이는 단순 CNN으로는 이번 문제를 해결할 수 없다는 것을 보여주며, Multi-tasking U-net이 보고 싶은 곳을 보면서 Classification을 해줄 수 있음을 보여주는 결과이다.

표 2에서 Multi-Tasking U-net 1은 영역 세분화를 위한 에러와 분류 작업을 위한 에러를 1대1로 더한 결과이고 Multi-Tasking U-net은 분류 작업에 대한 에러에 5을 곱한 뒤 더해하므로 균형을 맞춘 에러 함수를 사용한 결과이다.

표 2 CNN의 대표적인 Network VGG16, ResNet과 Multi-Tasking U-net의 성능 비교

Network	Accuracy
VGG16	68.42%
ResNet	71.68%
제안한 Multi-Tasking U-net 1	85.60%
제안한 Multi-Tasking U-net 2	90.63%

표 2와 그림 4, 5, 6 에서 기존 CNN의 모델과 제안한 모델을 비교한 이유는 기존 CNN으로는 분류작업이 되지 않는 데이터를 제안한 모델로 학습을 하면 좋은 결과를 나타낼 수 있음을 보여주기 위해서였다.

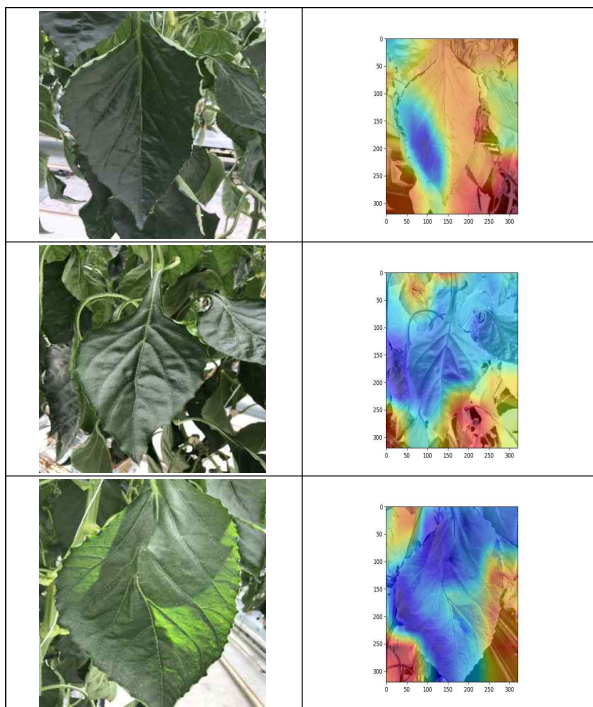


그림 4. VGG16 혹은 ResNet을 사용할 경우 네트워크 판단에 중요한 영향을 미치는 영역을 Class Activation Map(CAM)에 의해 가시화

파프리카 데이터를 CNN(VGG16, ResNet)으로 처리할 경우 정확도가 68.42%, 71.68%의 낮은 정확도를 보여주지만,

이를 Class Activation Map(CAM)을 사용하여 봤을 때, 배경을 보고 클래스를 판단하는 경우를 볼 수 있었다. 이는 보통 흰가루 병이 생기는 잎 뒤에 흰색 배경이 많고, 잎의 뒷면을 찍기 위해 이를 잡으려 하는 손가락이 나오는 경우가 많아 이것들을 보고 판단하는 경우가 많고, 그 외의 클래스도 특정 배경에 방해를 받음을 확인할 수 있었다[11].

그림 5는 Multi-Tasking U-net 1을 사용했을 때, CAM의 모습을 나타내고 있다. 전반적으로 잎의 경계를 보거나 잎 전체를 보고 병을 진단하는 모습을 보인다. 하지만 우리가 원하는 건 병의 영역만을 찾아내는 것임으로 이는 완벽한 결과라고 할 수 있다. 그림 6은 Multi-Tasking U-net 2을 사용했을 때, CAM의 모습이다. 분류작업에 대한 예리에 가중치를 두면서 잎 분류 작업에 더 학습을 하면서 잎의 중심 부분의 병들을 잘 찾아내는 것을 볼 수 있으며, 정확도 또한 오른 것을 확인할 수 있다.

그림 7은 영역 세분화 학습에 의해 모델이 잎의 영역을 예측한 결과이다. 첫 번째 영상은 원본 영상, 두 번째 영상은 정답 값(Ground truth), 마지막 영상은 모델이 예측한 영상이다. 전체적으로 잎을 잘 찾아내는 모습을 볼 수 있다[9],[10]. 가운데 잎이 아닌 뒤의 배경의 잎을 찾아내는 결과(그림 7의 3번, 5번) 또는 햇빛에 의해 잘 찾지 못하는 결과(그림 7의 6번)가 나오기도 하지만 이번 연구의 목적은 영역 세분화 작업이 아닌 찾아낸 영역을 중심으로 분류 작업을 진행하는 것이므로 충분히 만족할만한 결과라고 볼 수 있으며 IOU 값의 평균은 77%로 높은 성능을 보여주었다.

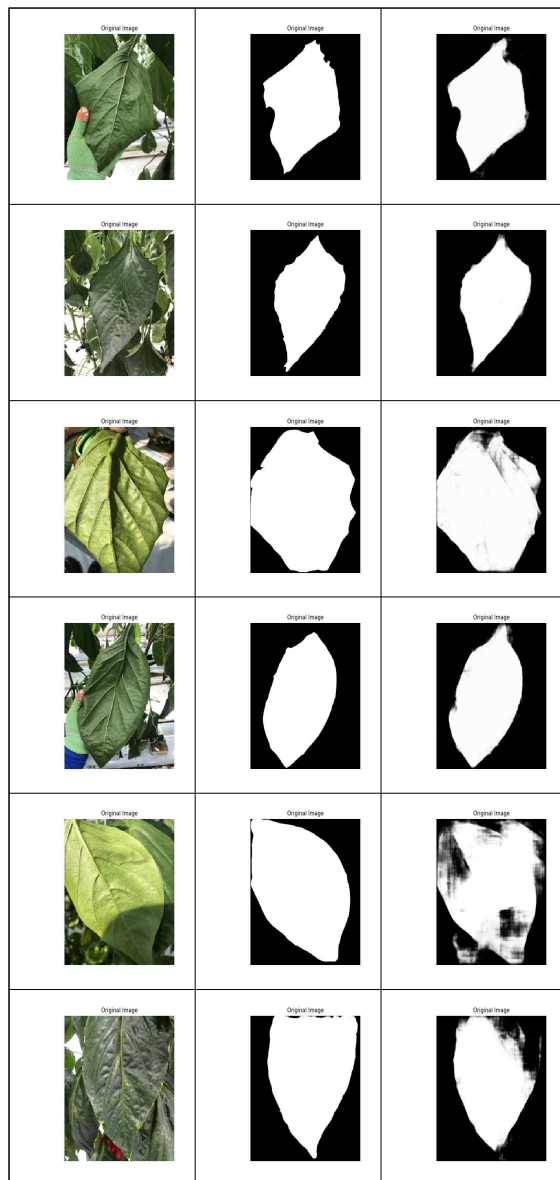


그림 7. 파프리카 잎의 Segmentation 결과

V. 결 론

본 연구에서는 Multi-Tasking U-net을 사용하여 영역 세분화 작업과 분류 작업이 동시에 수행되게 함으로써 파프리카 병해 진단을 수행하였다. 파프리카에는 병이 잘 발생하지 않으므로, 비교적 발생빈도가 높은 흰가루병 과 응애 및 정상 여부 판단 등 3개의 클래스로만 분류되게 하였다. 본 연구에서는 영상의 중심 잎의 영역을 중심으로 분류를 진행하게 함으로 배경과 불필요한 물체를 제거하게 하고 오직 잎을 보고 병의 진단을 내리게 하였고, 그 결과 기존 CNN보다 크게 향상된 결과를 냈으며, 가운데 잎의 위치를 찾아내고 그 영역의 병을 진단하게

하는 모델을 설계하였다. CAM을 통한 결과를 봤을 때, 기존 CNN보다 훨씬 더 높은 확률로 중심의 잎 부분의 병을 진단하는 모습을 보이며 테스트 결과 89.73%라는 높은 인식률을 얻었다. 이는 영역 세분화 작업과 분류작업이 동시에 학습이 될 수 있다는 것을 의미하며, 두 가지의 에러 함수를 더하여 하나의 에러 함수로 학습이 된다는 것을 보여주는 예가 될 수 있다.

또한 객체 검출 방식(Object Detection) 모델을 쓰지 않고 원하는 곳의 객체를 찾은 뒤, 그 영상의 클래스를 분류 해주는 의미 있는 연구이다.

REFERENCES

- [1] 김서정, 이재수, 김형석, “딥러닝을 이용한 양파 발의 잡초 검출 연구,” *스마트미디어저널*, 제7권, 제3호, 16-21쪽, 2018년 9월
- [2] S.Bargoti, James Underwood, “Deep fruit detection in orchards,” *2017 IEEE International Conference on Robotics and Automation(ICRA)*, Singapore, Singapore, 2017.
- [3] K.Simonyan and A.Zisserman, “Very deep convolutional networks for large-scale image recognition,” *In ICLR*, 2015.
- [4] S. Ren, K. He, R. Girshick, and J. Sun, “Faster RCNN: Towards real-time object detection with region proposal networks,” *arXiv preprint, arXiv:1506.01497*, 2015.
- [5] O. Ronneberger, P. Fischer, T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” *Proc. MICCAI*, pp. 234-241, 2015.
- [3] Hariharan, B., Arbeláez, P., Girshick, R., Malik, “Hypercolumns for object segmentation and fine-grained localization,” *Proceeding CVPR*, pp. 447 - 456, 2015.
- [4] J. Long, E. Shelhamer, T. Darrell, “Fully convolutional networks for semantic segmentation,” *Proceeding CVPR*, pp. 3431-3440, Apr. 2015.
- [5] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, “Convolutional architecture for fast feature embedding,” *Proc. of ACM MM*, pp. 675 - 678, 2014.

- [6] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," *Proc. of CVPR*, pp. 770-778, Jun. 2016.
- [7] Milletari, F., Ahmadi, S., Kroll, C., Plate, A., Rozanski, V.E., Maiostre, J., Levin, J., Dietrich, O., Ertl-Wagner, B., Bötzel, K., Navab, "N.: Hough-CNN: deep learning for segmentation of deep brain regions iMRI and ultrasound," *CoRR abs/1601.07014*, 2016.
- [8] Seyedhosseini, M., Sajjadi, M., Tasdizen, "T.: Image segmentation with cascaded hierarchical models and logistic disjunctive normal networks," *Proc. of ICCV*, pp. 2168 - 2175 , 2013.
- [9] Dong H., Yang G., Liu F., Mo Y., Guo Y. "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks In: Valdés Hernández M., González-Castro V. (eds) Medical Image Understanding and Analysis," *Communications in Computer and Information Science*, vol 723. Springer, Cham, 2017.
- [10] Drozdal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, "The importance of skip connections in biomedical image segmentation," *LABELS/DLMIA - 2016, LNCS*, vol. 10008, pp. 179 - 187, Springer, Cham, 2016.
- [11] Simard, P.Y., Steinkraus, D., Platt, J.C.: Best practices for convolutional neural networks applied to visual document analysis. In: Proceedings of Seventh International Conference on Document Analysis and Recognition, pp. 958 - 963. IEEE Computer Society (2003).
- [12] Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.* 36, pp. 61 - 78, 2017.
- [13] Milletari, F., Navab, N., Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," arXiv, pp. 1 - 11, 2016.
- [14] 서다솜, 오강한, 오일석, 유태웅, "딥러닝 설명을 위한 슈퍼픽셀 제외,포함 다중 스케일 접근법," *스마트미디어저널*, 제8권, 제2호, 39-45쪽, 2019년 6월
- [15] 박선, 김중원, "오픈 소스 기반의 딥러닝을 이용한 적조생물 이미지 분류," *스마트미디어저널*, 제7권, 제2호, 34-39쪽, 2018년 6월

 저자 소개



김서정(준회원)

2018년 전북대학교 전자공학부 학사 졸업.

<주관심분야 : 인공지능, 영상처리, 영상지능>



김형석(정회원)

1980년 한양대학교 전자공학과 학사 졸업.

1982년 전북대학교 전기공학과 석사 졸업.

1992년 University of Missouri Dept of Electrical and Computer 박사졸업.

<주관심분야 : 인공지능, 뉴럴네트워크, 영상지능>