

전경/배경 정보의 융합을 이용한 딸기 질병의 효율적인 분류

김 동 훈*, 김 태 현*, 이 준 환^o

Efficient Classification of Strawberry Diseases Using Fusion of Foreground and Background Information

Dong-Hoon Kim*, Taehyun Kim*, Joonwhoan Lee^o

요 약

농업분야에 대한 인공지능 모델의 적용과 관련된 연구가 매우 활발히 이루어지고 있다. 본 논문에서 딥러닝 모델의 전경(foreground) 정보와 배경(background) 정보를 융합함으로써 적은 파라미터로도 딸기 질병을 효율적으로 분류하는 모델을 제안하였다. 제안하는 방법은 전경과 배경의 특징을 분리하고 두 부분의 특징사이의 관계를 포착하여 분류성능을 향상시키기 때문에 파라미터 수가 적은 단순한 모델이더라도 효율적인 분류가 가능하다. 확보한 적은 수의 딸기영상 데이터로 네트워크의 학습을 위해 데이터를 증강하였으며, ImageNet 데이터로 학습된 가중치를 전이학습 후 미세조정을 수행하였다. 제안된 방법은 7개 클래스로 분류한 딸기 질병의 분류 문제에서 95.6%의 정확도를 나타내었다.

Key Words : strawberry disease, deep learning, foreground/background features, fewer parameters, data augmentation, transfer learning

ABSTRACT

Research related to the application of artificial intelligence models to the agricultural area is very active. In this paper, we propose a model that efficiently classifies strawberry diseases with the simplified network of fewer parameters due to the fusion of foreground and background information in a deep learning model. The proposed model divides learned image features into foreground and background ones, and captures the mutual relations between two areas, so that the classification might be efficient even with fewer parameters. In the proposed method, data augmentation is applied to train a network from a small amount of gathered strawberry image data, and fine-tuning is performed after transfer learning of the weights pre-trained with ImageNet dataset. The proposed method has achieved an accuracy of 95.6% in the classification problem of strawberry diseases of seven categories.

※ 본 결과물은 농림축산식품부의 재원으로 농림식품기술기획평가원의 1세대 스마트 플랜트팜애니멀팜 산업화 기술개발사업의 지원을 받아 연구되었음(과제번호 319071-01)

• First Author : Chonbuk National University Department of Computer Science and Engineering, 학생(박사), clickmiss123@naver.com, 학생회원

^o Corresponding Author : Chonbuk National University Department of Computer Science and Engineering, 정교수, chlee@chonbuk.ac.kr, 중신회원

* Department of Agriculture Engineering, National Institute of Agricultural Sciences, 연구사, thkim8205@korea.kr
논문번호 : 202003-056-A-RU, Received March 18, 2020; Revised April 21, 2020; Accepted April 22, 2020

I. 서 론

최근 진행되고 있는 인공지능 딥러닝 모델의 발전에 힘입어 다양한 분야에 인공지능 모델의 적용을 통한 효율성 증대가 검토되고 있다. 이 중 식량 생산과 관련된 농업 분야의 경우는 질병 검출 및 분류^[1], 최적의 환경 조건을 찾아내는 표현체 분석(phenotyping)^[2], 재배 현황 분석을 위한 색인 정보^[3] 생성 등을 기계학습 기법을 이용하여 수행함으로써 생산력의 극대화와 이윤의 증대를 추구하고 있다.

농작물 피해를 최소화하기 위한 작업의 자동화를 위해서는 기계가 질병의 감염 여부를 자율적으로 판단할 수 있어야하기 때문에 분류작업의 수행이 필수적으로 포함된다. 현재 이러한 분류 작업에는 CNN(Convolutional Neural Network) 모델^[4]이 좋은 성능을 나타내어 보편적으로 사용되고 있다. 많은 CNN모델에서는 분류하기 전의 계층에서 GAP(Global Average Pooling)^[5]를 수행함으로써 특징 맵의 평균값을 구하고 해당하는 값을 통해 클래스를 분류하는 형태로 모델이 구성된다.

딸기 질병을 분류하는 문제를 사람이 직접 해결한다고 가정할 때, 단순히 질병의 병징이 나타난 위치의 정보만을 가지고 분류하기보다 해당 위치가 잎의 어느 부분인지, 주변의 열매나 줄기 상태가 어떠한지 등의 주변부의 정보까지를 종합하여 판단하는 편이 보다 높은 정확성을 확보할 수 있다. 하지만 분류의 전 단계에서 수행되는 GAP는 특징 맵에 나타나는 공간적인 정보의 평균값을 계산하기 때문에 전경 정보와 배경 정보가 희석시켜 버린다. 따라서 본 논문에서는 전경과 배경 정보를 마스크를 통해 분리 하고 이들 영역에서 분리된 GAP를 수행하고 분류단계에서의 융합을 시도하는 방법을 제안한다. 이 과정에서 증가되는 특징 맵 숫자와 부족한 학습데이터 수로 인해 학습 데이터에만 적합한 분류기가 만들어지는 과적합이 발생할 위험성이 증가하는데, 모델의 파라미터를 감소시키는 방향으로 모델을 수정하고 학습시킴으로써 과적합 방지와 동시에 효율적인 분류가 가능하도록 모델을 구성할 수 있다.

제안된 방법은 데이터 증가, 초점 손실(focal loss)^[6]의 사용, 배경/전경정보의 이용, 채널 주의 모듈의 결합을 통해 7개 카테고리로 분류한 딸기 질병의 분류 문제에서 95.6%의 정확도를 나타내었다.

본 논문은 2장에서 실험에서 사용한 데이터와 네트워크 구조를 소개하고, 3장에서는 실험 및 결과를 제시한 뒤, 4장에서 결론을 내리는 형태로 구성된다.

II. 제안된 방법

2.1 학습 데이터의 구성

병징 부분만을 포함하는 영상 패치에 병해 부류를 부가한 데이터를 사용할 경우 식물의 잎, 줄기, 열매 등과 같은 맥락정보가 누락되어 성공적인 분류기 구성이 불가능하다. 실제로 본 논문의 방법을 제안하기에 앞서 그림 1,2와 같이 병징이 나타난 부위를 중심으로 바운딩 박스를 위치시켜 crop하여 패치기반 레이블 데이터로 네트워크를 학습시켜 보았으나 좋은 결과를 얻을 수 없었다.

이런 실험결과는 낮은 해상도로 인해 분류 시 영상 전체가 가진 특징 정보를 충실히 반영하기 어렵다는 점과 그림 2와 같이 동일질병이지만 특징 분포의 차이가 큰 영상이 학습데이터에 포함되어 있기 때문으로 판단되었다. 따라서 그림 3과 같이 패치를 지정하는 바운딩 박스가 포함하는 범위를 보다 넓힘으로써 잎맥, 윤곽, 배경정보와 같은 맥락정보가 포함될 수 있도록 학습데이터를 구성 하는 것이 필요하며, 병징이 식물체의 다양한 부위에서 나타나는 질병의 경우 서로 다른 부류로 취급함으로써 분류될 종류 별로 균일한 특징 분포를 가지도록 함으로써 분류성능을 올

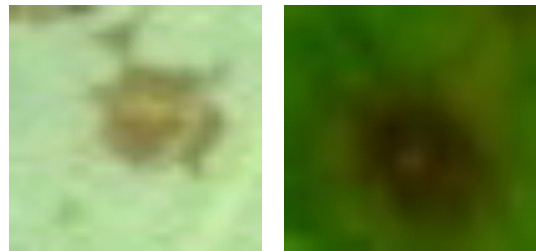


그림 1. 혼동될 수 있는 질병 (좌:세균모무늬병, 우:뱀눈무늬병)
Fig. 1. Diseases that can be confused (Left:angular leaf spot, Right:leaf spot)

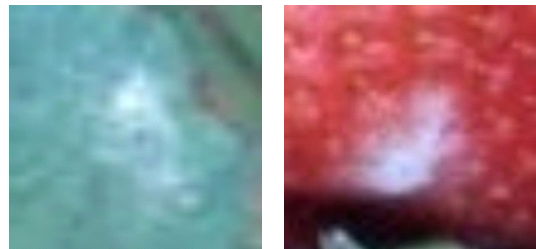


그림 2. 분포의 차이가 큰 동일 질병 (좌:흰가루병-잎, 우:흰가루병-열매)
Fig. 2. Identical diseases with large distribution differences (left:powdery mildew-leaf, right:powdery mildew-fruit)

릴 수 있을 것으로 기대되었다.

이러한 사항을 고려하여 세균무늬병(angular leaf spot), 탄저병(anthracnose fruit rot), 잿빛곰팡이병(gray mold), 잎마름병(leaf blight), 뱀눈무늬병(leaf spot), 흰가루병-열매(powdery mildew-fruit), 흰가루병-잎(powdery mildew-leaf) 까지의 7개 카테고리에 대한 분류기를 설계하였다. 즉, 흰가루 병의 경우 잎과 열매의 병징을 구분하여 부류를 나누었으며, 이에 따라 채집한 데이터에 대해 카테고리 주석을 부여하였다. 또한 실제 학습 데이터와 테스트 데이터의 구성 시 보다 많은 배경정보가 포함 될 수 있도록 Ground Truth 바운딩 박스의 크기를 1.25배 확장한 뒤 crop하고, 네트워크 입력 크기(224×224)에 맞춰 크기를 조절하였다.

한편 바운딩 박스가 포함하는 범위를 넓혀 카테고리 정보를 부여할 경우 작은 크기의 패치기반 방식에 비해 학습데이터의 수가 적어지기 때문에 데이터의 증가를 위해 영상 데이터의 90도 단위 회전과 상하, 좌우 반전을 적용하였다. 실험에서 사용한 데이터셋은 인터넷 환경으로부터의 크롤링한 데이터와 농장에 직접 찾아가서 촬영한 사진으로 구성하였다.

상대적으로 데이터를 많이 모을 수 있었던 잎에 발생한 흰가루병이나 잎마름병, 뱀눈무늬병 등에 비해 타 질병은 발병 케이스를 수집하기 어려웠기 때문에 수집할 수 있는 데이터가 한정되었다.

일반적으로 데이터 수의 불균등(imbalanced)은 학습에 좋지 않은 영향을 줄 수 있다⁷⁾. 따라서 데이터셋의 구성함에 있어서 데이터 증가를 위한 회전 및 반전은 표 1과 같이 데이터 수가 최대인 클래스와 균등한 정도가 될 수 있도록 클래스별로 증가량을 다르게 적용하였다. 데이터 증가에는 90, 180, 270도의 회전과 상하, 좌우 반전을 사용하였으므로 최대 6배로 데이터

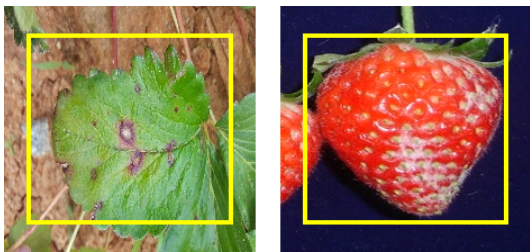


그림 3. 보다 넓은 범위를 포함하도록 구성한 데이터 (좌:흰가루병-잎, 우:흰가루병-열매, 노란 선은 기존 바운딩 박스의 의미함)
Fig. 3. Data configured to cover a wider range (left:powdery mildew-leaves, right:powdery mildew -fruit, yellow line means existing bounding box)

표 1. 각 질병별 데이터 수
Table 1. Number of data for each disease

		angular leaf spot	anthracnose fruit rot	gray mold	leaf blight	leaf spot	powdery mildew fruit	powdery mildew leaf
		train	483	228	252	591	1,140	168
	aug	1,449	1,140	1,260	1,182	1,140	1,008	1,134
test		161	76	84	197	380	56	378

를 증가시킬 수 있었다. 또한 테스트 집합으로는 데이터 증가를 적용하지 않은 원본 데이터만을 사용하였다.

가용한 데이터셋의 숫자가 충분치 않았기 때문에 모든 실험은 모델의 유효성에 대한 적절한 평가를 내릴 수 있도록 데이터셋에 4-fold cross-validation을 적용하였다.

2.2 네트워크의 구성

본 논문에서는 전경 및 배경 정보의 융합을 통한 성능 향상을 평가하기 위해 기존에 사용되고 있는 모델을 변형하여 사용하였다. 배경 정보는 그림 4와 같은 마스크를, 전경 정보는 해당 마스크의 역수를 사용하여 분리를 시도하였다. 즉 가우시안 형태의 마스크는 배경과 전경의 경계구분이 모호하며 고정 값 마스크는 경계구분이 분명하게 설정하였다.

전경/배경 정보의 분리는 네트워크의 중간 단 특징 출력값에 각각 전경/배경 마스크에 대한 곱연산을 수행하는 방식으로 수행하였다. 특징맵에 이러한 형태의 마스크에 대한 곱 연산을 수행할 경우 가우시안 마스크는 중앙으로 갈수록 전경의 비중이 높은 특징맵을, 고정 값 마스크는 전경/배경에 대한 비중이 거리에 대해 일정한 특징맵을 각각 생성할 수 있을 것이라 기대할 수 있다. 클래스의 분류는 이러한 전경/배경 정보를 가진 특징맵에 Global Average Pooling(GAP)를 수행하고, 이를 분류기에 연결하는 방식이다. 기존의 전경/배경을 분리하지 않고 GAP를 수행하는 방식은 배경에 관한 정보가 희석되고 전경의 정보가 배경 정보를 압도하게 되지만 GAP 수행 전에 각 정보를 분리함으로써 이를 방지할 수 있다. 그러나 여기서 특징맵이 너무 작아진 경우 마스크를 통한 전경과 배경정보를 정의하기 어려워 분류가 어렵기 때문에 적절한 크기의 마스크를 적용하기 위해 기존 네트워크의 계

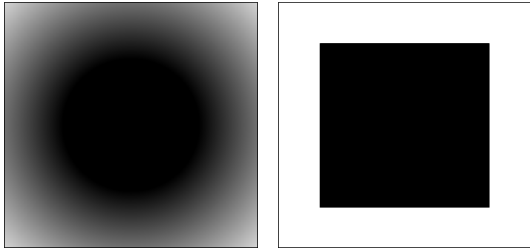


그림 4. 좌:가우시안 마스크(G-mask), 우:고정 값 마스크 (S-mask)
Fig. 4. Left:Gaussian mask(G-mask), Right:Static value mask(S-mask)

층적 합성곱 레이어의 특징맵의 크기가 어느 정도 보존되는 중간 계층 출력을 잘라서 사용하였다. 전경 정보와 배경 정보를 분리하여 특징으로 이용할 경우 특징 맵의 수가 2배가 되는데, 실험에서 사용한 데이터셋은 데이터의 수가 충분치 않아 과적합의 위험성이 증가될 수 있다.

이러한 문제의 해결을 위해서는 네트워크의 중간 계층에서 공간상으로 적당한 크기를 가지면서 특징 맵의 수가 과하지 않게 나타나는 모델이 적합하였다. MobileNetV2^[8]는 이러한 조건에 부합하며 Inception이나 ResNet, VGG 등의 모델에 비해 파라미터의 수가 현저히 적음에도 성능의 저하가 크지 않은 모델이기 때문에 본 실험에서는 해당 모델을 채택하여 문제를 해결하였다.

실험에서 사용한 가우시안 분포를 가진 마스크의 경우 중앙으로부터 최대값이 1 최소값이 0이 되도록 정규화를 수행하였고, 고정 값 마스크의 경우는 중앙으

표 2. 네트워크의 구조
Table 2. Structure of the network

input	operator	expansion factor	output channel	repeat number	stride
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	FG / BG mask	-	192	1	-
$14^2 \times 192$	avgpool 14x14	-	-	1	-
$1^2 \times 192$	conv2d 1x1	-	k	-	-

로부터 2/3 지점을 기준으로 값이 반전되도록 하여 전경과 배경의 면적이 비슷한 비율로 배분되도록 설정하였다. 모델의 구조는 stride값에 따라 그림 5와 같은 블록을 연결하는 방식으로 이루어진다. 그리고 표 2에서 repeat number가 2이상이면 stride가 2인 경우는 첫 번째 반복에서 그림 5의 우측 블록을 연결하고 다음 반복에서는 stride가 1인 블록을 연결하는 방식을 취하였다.

이상의 내용을 반영하여 실험에서 사용된 네트워크는 표 2와 같이 구성되고, 이는 그림 6와 같이 나타낼 수 있다. FG/BG mask 계층 전까지는 MobileNetV2의 구조와 동일하지만, 앞에서 소개한 목적을 위해 추가적인 병목(bottleneck)구조의 반복이 제외되어 있다.

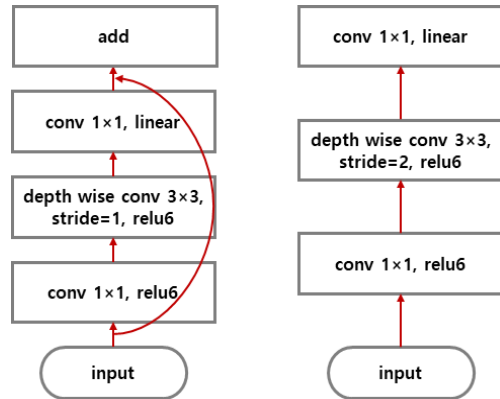


그림 5. 좌:stride가 1인 블럭, 우:stride가 2인 블럭
Fig. 5. Left:block with stride 1, Right:block with stride 2

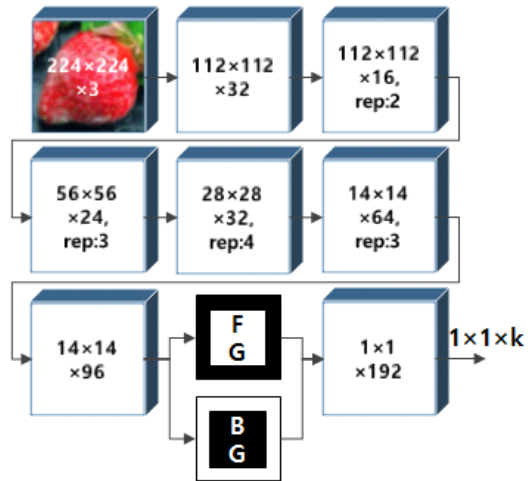


그림 6. 모델의 전체적인 구조
Fig. 6. The overall structure of the model

III. 실험 및 결과

데이터 증가 방법을 사용함으로써 학습데이터를 7,758개로 늘릴 수 있었지만, 과적합 없이 전체 네트워크를 학습시키기에는 부족한 데이터라고 할 수 있다. 이러한 상황에서는 많은 데이터를 사용하여 다양한 클래스에 대해 기학습된 가중치를 재사용하는 전이학습(transfer learning) 방법^[9]이 유용하게 사용될 수 있다. ^[10,11]의 사례를 볼 때, 식물 질병 데이터만을 따로 모아서 학습할 경우 보다 좋은 결과를 얻는데 도움이 될 수 있을 것이다. 하지만, 현재 그러한 데이터셋이 확보되지 않은 관계로 ^[12]에서와 같이 보편적인 분류 문제에서 다양한 분포를 가지는 데이터를 통해 학습된 가중치를 사용하는 방식을 채택하였다. 따라서 현재 수행하고자 하는 작업과 같은 시각적 도메인인 전체 ImageNet 데이터셋으로 학습된 가중치를 이용하여 전이학습을 수행 한 뒤, 보다 작은 학습률을 적용하여 전체 네트워크를 미세조정(fine tuning)하는 식으로 학습을 수행하였다. 최적화 도구로는 최신 기술자들이 비중 있게 반영되도록 지수 가중 이동 평균(Exponentially weighted moving average)을 사용하는 RMSProp를 사용하였고, 전이 학습에 대한 초기 학습률 파라미터는 0.001로, 미세 조정에 대한 초기 학습률 파라미터는 가중치가 거의 수렴하였다는 가정 하에 이보다 작은 0.00001로 설정하였다. 초기 실험의 경우 soft-max에 대한 cross-entropy 손실값을 통해 네트워크를 학습하였고, 그림 4의 고정 값 마스크만을

표 3. soft-max에 대한 cross-entropy 손실함수를 사용한 실험 결과
Table 3. Experimental results using soft-max with cross-entropy loss functions

	original	S-masked
average / all	0.9295	0.9402
average / class	0.8832	0.9005
angular leaf spot	0.9596	0.9627
anthracnose fruit rot	0.8191	0.8355
gray mold	0.7738	0.8333
leaf blight	0.9327	0.9480
leaf spot	0.9303	0.9382
powdery mildew fruit	0.7723	0.7902
powdery mildew leaf	0.9947	0.9960

적용해 보았다. 이상의 조건에서 전이학습과 미세조정에 각각 20,000 step의 학습을 진행하여 표 3과 같은 결과를 얻을 수 있었다. 실험의 결과에서 전경/배경정보를 융합한 경우 약 2퍼센트 정도 분류율 향상을 확인할 수 있었다. 또한, 병목계층을 조금 더 잘라내어 표 2에서 입력사이즈가 $28^2 \times 32$ 인 계층에 보다 큰 사이즈의 마스크를 적용한 실험에서는 전체 테스트 데이터에 대한 평균 분류율은 85퍼센트 정도로 낮아졌지만 전경/배경을 분리하지 않은 케이스에 비해서는 4~5퍼센트 정도 분류율이 향상 되는 것을 관찰할 수 있었다.

표 4의 실험 결과에서 열매에서 발생하는 흰가루병, 열매, 잣빛곰팡이병, 탄저병과 같은 질병들 사이에 혼동이 발생하여 분류가 잘 되지 않는 문제와 흰가루병, 열매, 잣빛곰팡이병과 같이 분류가 잘 되지 않는 특정 클래스로 인해 전체적인 성능이 하락하는 문제를 확인할 수 있었다.

해당 클래스들은 데이터 증가를 적용하여 학습 데이터의 수가 비슷하게 되도록 맞추었던 하지만 원본 데이터의 수가 적었던 클래스들임을 확인할 수 있었다. 이러한 클래스 불균형으로 인한 성능 저하 문제의 극복을 위해 틀린 예제에 대해 보다 높은 가중치를 두어 학습을 수행할 수 있도록 초점 손실을 적용하였다. cross-entropy 손실 함수가 네트워크의 출력 값 p_i 에 대해 (1)과 같은 반면, 초점 손실은 해당 클래스에 대해 가중치 α , 다른 클래스에 대해 가중치 $1-\alpha$ 를 적용함으로써 클래스 불균형 문제를 다룬다. 또한 초점 손실은 정답을 맞춘 쉬운 샘플과 어려운 샘플에 대한 손실 값의 가중치를 γ 값을 통해 스케일링되는 $(1-p_i)^\gamma$ 텀을 곱해줌으로써 조절한다. 이상을 수식으로 나타내

표 4. 고정 값 마스크를 사용한 실험 결과의 혼동행렬
Table 4. Confusion matrix of experimental results using a static value mask

true \ pred	angular leaf spot	anthracnose fruit rot	gray mold	leaf blight	leaf spot	mildew fruit	powdery mildew leaf
	ang-s	0.963	0	0.008	0.009	0.012	0
ant-r	0	0.756	0.089	0.016	0	0.059	0
gray	0	0.086	0.833	0.009	0	0.071	0
leaf-b	0.013	0.005	0.009	0.948	0.010	0.001	0.014
leaf-s	0.015	0.004	0.007	0.016	0.938	0	0.019
pow-f	0.004	0.063	0.121	0.013	0.004	0.790	0.004
pow-l	0	0	0	0.001	0.003	0	0.996

면 초점 손실을 (2)로 표현할 수 있다. α, γ 값의 조율에 앞서 아래의 실험에서는 ¹⁶⁾의 객체 검출에서 좋은 성능을 보였던 $\alpha=0.25, \gamma=2$ 를 사용하여 전경-배경 분리 방법의 유효성 테스트를 수행하였다.

$$CE(p_t) = -\log(p_t) \tag{1}$$

$$FL(p_t) = -\alpha(1-p_t)^\gamma \log(p_t) - (1-\alpha)p_t^\gamma \log(1-p_t), \quad \alpha, \gamma \geq 0 \tag{2}$$

soft-max에 대한 cross-entropy 손실함수 대신 초점 손실을 적용하고 가우시안 마스크(G-mask)를 적용하는 방식을 추가하여 앞의 실험을 반복하여 표 5와 같은 결과를 얻을 수 있었다. 해당 결과에서 초점손실을 적용함으로써 기존에 분류가 잘 되지 않던 클래스들의 분류율이 향상되어 전체 데이터에 대해서는 3퍼센트, 클래스 평균에 대해서는 5퍼센트 정도의 분류율 향상을 얻을 수 있음을 확인 가능하다. 또한 고정 값 마스크를 적용한 경우와 가우시안 마스크를 적용한 경우는 그렇지 않은 경우에 비해 0.2~0.5퍼센트 정도의 추가적인 분류율 향상이 있는 것도 확인할 수 있다. 표 3의 실험 결과와는 달리 초점 손실을 사용하는 경우 배경/전경의 특징 분리 효과는 그렇지 않았을 경우에 비해 두드러지지 않았다. 이는 초점손실에 의한 성능 향상에 의해 전경/배경 분리 효과가 상쇄되기 때문으로 판단된다.

즉, 전경/배경 정보를 융합하는 것이 유리했던 샘플들이 초점 손실을 적용하는 것이 유리한 샘플에 포함

되기 때문에 초점손실로 인한 이점을 취하지 못하면서 융합에 의한 이점을 얻을 수 있는 샘플들이 분류율 향상에 기여를 하기 때문이다. 또한 고정 값 마스크를 사용한다에 비해 가우시안 마스크를 사용한 결과의 분류율이 평균적으로 낮은 경향을 보였는데, 이러한 결과는 병징이 중앙에서 나타나는 경우보다 약간 중앙에서 벗어난 위치에서 나타나는 경우가 많기 때문에 중앙으로 갈수록 높은 가중치를 부여하는 방식이 적합하지 않았다고 해석할 수 있다.

표 5의 실험에서 사용한 초점손실의 하이퍼 파라미터 α, γ 는 객체검출에서 그 성능이 입증되었지만, 해결해야할 문제와 데이터셋이 달라졌기 때문에 다른 최적의 값이 존재할 수 있다. 따라서 보다 적합한 값을 탐색하기 위해 가장 좋은 성능을 나타낸 고정 마스크 모델을 사용하여 4-fold cross validation 중 하나의 fold에 대해 격자식으로 다양한 α, γ 값을 주는 실험을 진행하였다. $\{0.1, 0.25, 0.5, 0.75, 0.9\} \in \alpha, \{0.25, 0.5, 1.0, 2.0, 3.0\} \in \gamma$ 값을 사용하여 실험한 결과 표 6과 같은 결과를 얻을 수 있었다. 이 실험에서 클래스 별 평균 분류율과 테스트 데이터에 대한 평균 분류율 면에서 높은 값을 기록했던 $\alpha=0.5, \gamma=1.0$ 을 사용하여 전체 fold에 대해 실험한 결과는 표 7과 같다.

초점 손실의 하이퍼 파라미터를 조율한 결과 $\alpha=0.25, \gamma=2$ 를 사용한 경우에 비해 클래스 평균 분류율 측면에서 0.7퍼센트 정도의 성능 향상을 거둘 수 있었다. 마지막으로 그림 6 모델의 FG에 가중치 w 를 곱하고 BG에 $(1-w)$ 를 곱해줌으로써 전경과 배경에 대해 중요도를 조절하고자 하는 실험을 진행하였다. 초점손실의 파라미터를 $\alpha=0.5, \gamma=1.0$ 로 준 고정 마스크 모델에 4-fold cross validation 중 하나의 fold만을

표 5. 초점 손실 함수를 사용한 실험 결과($\alpha=0.25, \gamma=2$)
Table 5. Experimental results using focal loss function

	original	S-masked	G-masked
average / all	0.964	0.968	0.966
average / class	0.944	0.949	0.946
angular leaf spot	0.978	0.977	0.977
anthracnose fruit rot	0.928	0.934	0.928
gray mold	0.905	0.917	0.923
leaf blight	0.967	0.976	0.975
leaf spot	0.956	0.959	0.957
powdery mildew fruit	0.875	0.879	0.866
powdery mildew leaf	0.999	0.999	0.999

표 6. 초점손실 하이퍼 파라미터에 따른 분류 정확도
Table 6. Classification accuracy according to hyperparameter of focal loss

α	γ	average / all	average / class
0.1	0.25	0.9670	0.9553
0.25	0.25	0.9685	0.9591
0.25	0.5	0.9662	0.9596
0.25	1.0	0.9677	0.9572
0.5	0.25	0.9662	0.9544
0.5	0.5	0.9677	0.9601
0.5	1.0	0.9677	0.9604
0.5	2.0	0.9625	0.9467
0.75	1.0	0.9647	0.9528
0.9	0.25	0.9655	0.9507

표 7. 초점 손실 함수를 사용한 실험 결과의 비교
Table 7. Comparison of experimental results using the focal loss function

	S-masked ($\alpha=0.25, \gamma=2.0$)	S-masked ($\alpha=0.5, \gamma=1.0$)
average / all	0.968	0.972
average / class	0.949	0.956
angular leaf spot	0.977	0.983
anthracnose fruit rot	0.934	0.934
gray mold	0.917	0.932
leaf blight	0.976	0.984
leaf spot	0.959	0.961
powdery mildew fruit	0.879	0.902
powdery mildew leaf	0.999	0.999

표 8. 전경/배경 가중치 변화에 따른 분류율
Table 8. Classification accuracy according to weight change of foreground/background

w	average / all	average / class
0.25	0.9670	0.9550
0.5	0.9677	0.9604
0.75	0.9670	0.9519
0.9	0.9640	0.9466

사용하여 실험을 수행하였고, 표 8과 같은 결과를 얻을 수 있었다.

해당 실험에서 전경과 배경의 가중치가 1:1인 0.5의 경우에 비해 다른 비율을 적용한 경우 좋은 결과를 얻을 수 없었다. 이는 사용한 모델의 구조상 전경/배경이 동일한 특징 맵 수를 가지고 있기 때문에 곱 연산을 거치는 방식으로 가중치 부여를 할 경우 어느 한 쪽의 영향력이 거의 사라진 특징으로부터 학습이 수행되기 때문이라 해석 할 수 있다. 따라서 이러한 문제의 해결을 위해서는 전경/배경 어느 한 부분만을 사용하여 학습한 네트워크를 late fusion하는 방식 등의 다른 접근방법이 필요할 것이다.

IV. 결 론

본 논문에서는 딸기 질병 분류의 정확도 향상을 위해 전경 정보와 배경 정보를 융합하여 사용하는 방법을 제안하였다. 해당 방법의 적용을 통해

MobileNetV2구조의 학습해야 할 가중치의 수를 2.23M개에서 1/4로 줄일 수 있었고, 전경과 배경을 분리하지 않은 방법에 비해 약간의 성능 향상이 있음을 실험을 통해 확인하였다. 이 모델은 적은 파라미터를 사용했음에도 테스트 집합의 분류에서 97.2%, 클래스 평균 분류율 면에서 95.6%의 성능을 보여주었다. 제안한 방법은 맥락 정보가 중요한 분류 문제에서 유효하게 사용할 수 있음을 보여주며, 특히 성능과 자원 사용을 조율해야하는 가용 자원이 부족한 환경에서의 모델에 적용을 고려해 볼 수 있을 것이다.

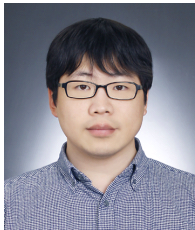
References

- [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.
- [2] F. Fiorani and U. Schurr, "Future scenarios for plant phenotyping," *Annu. Rev. Plant Biology*, vol. 64, pp. 267-291, Apr. 2013.
- [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.
- [4] C. Szegedy, et al., "Inception-v4, inception-resnet and the impact of residual connections on learning," *Thirty-first AAAI Conf. Artificial Intell.*, pp. 4278-4284, San Francisco, USA, Feb. 2017.
- [5] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv:1312.4400*, 2013.
- [6] T.-Y. Lin, et al., "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 2980-2988, Venice, Italy, Aug. 2017.
- [7] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowledge and Data Eng.*, vol. 21, no. 9, pp. 1263-1284, Sep. 2009.
- [8] M. Sandler, et al., "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE Conf. Computer Vision and Pattern*

Recognition, pp. 4510-4520, Salt Lake City, USA, Jan. 2018.

- [9] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowledge and Data Eng.*, vol. 22, no. 10, pp. 1345-1359, Oct. 2009.
- [10] H.-C. Shin, et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285-1298, May 2016.
- [10] S. Lee, et al., "Recognition and visualization of crack on concrete wall using deep learning and transfer learning," *J. Korean Soc. Agric. Eng.*, vol. 61, no. 3, pp. 55-65, May 2019.
- [12] P. Perera and V. M. Patel, "Learning deep features for one-class classification," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5450-5463, May 2019.

김 동 훈 (Dong-Hoon Kim)



2014년 2월 : 전북대학교 전자
정보공학과(컴퓨터공학) 공
학사
2017년 2월 : 전북대학교 전자
정보공학과(컴퓨터공학) 공
학석사
2018년 2월~현재 : 전북대학교

전자정보공학과(컴퓨터공학) 박사과정
<관심분야> 영상·음성분석, 감성 인식, 인공지능
[ORCID:0000-0001-6945-6952]

김 태 현 (Taehyun Kim)



2006년 8월 : 세종대학교 컴퓨
터공학과 공학사
2009년 2월 : 세종대학교 컴퓨
터공학과 공학석사
2011년 2월 : 세종대학교 컴퓨
터공학과 박사수료
2011년 8월~2019년 1월 : (주)

아이앤씨테크놀로지 책임연구원
2019년 2월~현재 : 농촌진흥청 국립농업과학원 스마
트팜개발과 농업연수사
<관심분야> HCI, 머신러닝, 감정 인식
[ORCID:0000-0002-4478-667X]

이 준 환 (Joonwhoan Lee)



1980년 2월 : 한양대학교 전자
공학과 공학사
1982년 2월 : 한국과학기술원
전자공학과 공학석사
1990년 8월 : 미국 미주리대학
전기 및 컴퓨터공학과 공학
박사

1990년 10월~현재 : 전북대학교 컴퓨터공학부 교수
<관심분야> 영상처리, 감성 분석, 인공지능
[ORCID:0000-0003-1854-9643]