



스마트 팜을 위한 분할 인식 심층 합성곱 신경망 기반의 사과나무 잎사귀 질병 인식

유희진*, 손창환**

Recognizing Apple Leaf Diseases via Segmentation-Aware Deep Convolutional Neural Networks for Smart Farm

Hee-Jin Yu*, Chang-Hwan Son**

이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임
(No. 2017R1D1A3B03030853)

요 약

이 논문에서는 분할 인식 심층 합성곱 신경망 기반의 사과나무 잎사귀 질병을 인식하는 기법을 제안하고자 한다. 주된 아이디어는 잎사귀 질병이 존재하는 영역이 잎사귀 안에 포함된다는 것이다. 이 아이디어를 실현하고자 본 논문에서는 두 종류의 서브네트워크를 제시하고자 한다. 하나는 입력 영상을 배경 영역, 잎사귀 영역, 질병 영역으로 분할하기 위한 영상분할 서브네트워크이고 다른 하나는 입력 영상으로부터 질병의 종류를 예측하는 영상인식 서브네트워크이다. 각각의 서브네트워크는 먼저 인코더-디코더 타입의 아키텍처와 VGG 아키텍처로 구성된 후, 전이 학습을 통해서 훈련한다. 그리고 미리 학습된 두 종류의 서브네트워크를 연결 레이어를 통해 결합한다. 즉, 전체 네트워크를 종단 간 학습이 되기 위해, 퓨전 레이어를 사용하여 영상분할 서브네트워크에서 추정된 특징 맵을 입력 영상 위에 쌓아서 영상인식 서브네트워크에 입력한다. 실험 결과를 통해서, 예측된 영상분할 맵을 사용함으로써 기존의 VGG 네트워크에 비해 약 9%의 정인식률을 개선할 수 있었다.

Abstract

A new method of recognizing apple leaf diseases via segmentation-aware deep convolutional neural network is proposed in this paper. The main idea is that leaf diseases exist in the leaf area. To realize this idea, two subnetworks are designed. One is for the division of input image into background, leaf area, and disease area, and the other is for the prediction of leaf disease types. The two subnetworks have architecture types of the encoder-decoder network and the VGG network, respectively, and then trained separately via transfer learning. Next, the two types of subnetworks are combined through a concatenation layer. In other words, to train the entire network in an end-to-end manner, the predicted image segmentation map is stacked on the top of the input image through the concatenation layer, and then fed into the image classification subnetwork. The experimental results show that correct recognition accuracy can be increased by 9% by using the predicted image segmentation map, compared to the VGG network.

Keywords

smart farm, plant disease, deep convolutional neural networks, deep feature extraction, feature pooling

* 군산대학교 소프트웨어융합공학과

- ORCID: <https://orcid.org/0000-0001-7315-7517>

** 군산대학교 소프트웨어융합공학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0001-7077-3074>

• Received: Mar. 28, 2019, Revised: May 16, 2019, Accepted: May 19, 2019

• Corresponding Author: Chang-Hwan Son

Department of Software Convergence Engineering, Republic of Korea,

Tel.: 063-469-8915, Email: cson@kunsan.ac.kr

1. 서 론

스마트 팜이란 전통적인 농사법에 정보통신기술을 접목해서 작물의 상태와 환경을 실시간으로 체크하여 최적의 생육 환경을 조성하고 병해충 및 질병을 조기에 진단하고 대응할 수 있는 농장을 일컫는다[1]. 특히, 사물인터넷의 발달과 함께 유·무선 인터넷 보급망을 통해 다종의 스마트 센서, 즉 광센서, 가속도 센서, 전기화학 센서, 공기유량 센서, 적외선 센서를 통해 잎사귀의 각도 및 색상, 토질, 토양 산도, 스펙트럼 영상 등을 측정 및 수집한 후에, 인공지능 기법을 적용해서 작물의 생육 상태, 스트레스 지수, 병해충 유·무를 진단하고 있다. 이를 통해 농가는 노동력과 비료와 같은 자원을 최소화하여 투입하여 작물의 생산량과 품질을 최대화할 수 있다[2].

본 연구에서 다룰 주제인 작물 질병 진단 기술은 병해충으로 인한 질병 확산을 조기에 방지하고 농가의 경제적인 손실을 감소시킬 수 있기에 스마트 팜에서 반드시 필요한 기술이다. 종래의 작물 질병 진단은 전문가나 농작물에 대한 풍부한 경험이 축적된 농부가 잎사귀에 나타난 증세를 보고 판단하였다. 하지만 이 진단 방법은 시간이 많이 소요되고 비용이 요구될 뿐만 아니라 농작지가 클 경우에는 적용하기가 어렵다. 따라서 작물의 질병 진단을 자동화하는 기술이 필요하다.

최근 딥 러닝의 출현과 함께 컴퓨터 비전이 비약적인 진보를 거듭하고 있다. 특히, 2012년에 일명 알렉스 넷(Alex Net)이라 불리는 심층 합성곱 신경망(Deep Convolutional Neural Networks: DCNN)에 관한 논문이 발표된 이래로[3], 딥 러닝 기반의 영상 인식 및 검출 기술이 다양한 분야, 예를 들면 스마트폰의 홍채 인식, 무인 자동차의 보행자 검출, 의료 영상 진단 등에 적용되고 있다. 이에 발맞추어 스마트 팜 분야에서도 작물의 질병 진단을 영상 기반으로 인식하고 검출하는 시도가 진행되고 있다.

영상 기반의 작물 진단은 작물의 종류, 즉 토마토, 사과나무, 바나나 등에 상관없이 기존의 컴퓨터 비전 기술을 그대로 적용하거나 보완할 수 있다. 작물 질병 진단을 위한 종래의 컴퓨터 비전 기술은 크게 국부 특징 추출(Local Feature Extraction), 특징

풀링(Feature Pooling), 분류기 학습(Classifier Learning)으로 이루어진다. 국부 특징 추출은 영상의 국부 영역에서 특징을 묘사하기 위한 기법으로써, 예지의 방향에 따른 크기를 나타내는 히스토그램이 주로 사용된다. 대표적인 국부 특징 추출에는 SIFT[4], HOG[5], LBP[6] 등이 있다. 특징 풀링은 국부 영역에서 추출된 특징을 사용해서 부류 간 영상을 구분할 수 있는 영상 단위의 특징 추출을 말한다. 대표적인 특징 풀링은 단어 가방(Bag-of-Words)[7][8] 모델로써, 추출된 국부 특징 벡터들에서 군집화 알고리즘을 적용해서 센터 벡터들을 추출한 후, 센터 벡터와 거리가 가장 작은 센터 벡터의 인덱스에 한 표를 던져서 히스토그램을 누적하여 부류 간 영상 특징을 구별하도록 설계한다. 최근에는 피셔 벡터 인코딩(Fisher Vector Encoding: FVE)[9]이 각광을 받고 있다. 이 기법은 군집화 알고리즘 대신에 가우시안 혼합 모델을 사용해서 로그 우도(Log Likelihood)를 가우시안 혼합 모델의 파라미터로 미분한 그라디언트(Gradient) 벡터로 영상 단계의 특징을 묘사한다. 이렇게 부류 간 영상을 구분할 수 있는 특징 벡터가 주어지면, 분류기를 학습해서 원하는 분류 작업을 완성할 수 있다. 여기서 우도는 가우시안 혼합 모델의 확률을 최대로 할 수 있는 파라미터를 추정하는 확률 추정 기법을 의미한다.

하지만 종래의 수작업 기반의 특징 추출 기법은 응용 도메인이 달라지면 특징을 새롭게 모델링하거나 수정해야 되는 번거로움이 있다. 또한 응용 도메인에 대한 해박한 지식이 요구된다. 이런 문제를 해결하고자 최근에는 학습 기반으로 특징을 추출하는 접근 방식이 도입되고 있다. 즉, 딥 러닝을 통해 미리 학습된 모델을 사용해서 특정 레이어에서 추출된 특징 맵(Feature Map)을 특징 벡터로 사용한다. 이렇게 학습된 특징 벡터는 응용 도메인이 달라지더라도 유용한 특징 벡터가 될 수 있다[10]. 또는 새로운 훈련 데이터가 주어지면, 미리 학습된 모델을 미세 튜닝해서 새로운 분류 작업을 손쉽게 해결할 수 있는 전이 학습(Transfer Learning: TL)도 큰 각광을 받고 있다[11]. 왜냐하면 종래의 수작업 기반의 국부 특징 추출, 특징 풀링, 분류기 학습에 대한 깊은 이해가 없어도 전이 학습을 통해 이 일련의 과정을 단번에 해결할 수 있기 때문이다.

최근에 진행된 대부분의 작물 질병 진단에 관한 연구가 전이 학습에 기초하고 있다[12]-[14]. 그러나 전이 학습의 경우, 훈련에 사용될 데이터 셋만 다를 뿐 사용된 아키텍처는 그대로이다. 즉, 작물 질병의 종류만 다를 뿐 연구에 사용된 방법론은 거의 동일하다. 따라서 본 연구에서는 작물 질병 진단을 위한 새로운 아키텍처를 제시하고자 한다. 특히, 잎사귀 영역, 질병 영역, 배경 영역을 구분해서 작물 질병 진단의 정확도를 높일 수 있는 분할 인식 심층 합성곱 신경망을 제안하고자 한다.

1.1 제안한 접근 방법

제안한 방법의 주된 아이디어는 잎사귀 질병에 대한 정보는 잎사귀 영역에만 존재할 뿐, 배경 영역에서는 아무런 정보를 도출할 수 없다는 것이다. 만약 심층 합성곱 신경망을 학습할 때, 잎사귀 영역, 배경 영역, 그리고 질병 영역으로 구성된 분할 맵 (Segmentation Map)을 예측하여 추가적으로 활용할 수 있다면, 분할 맵에서 어떤 영역에 속한 특징 벡터가 상대적으로 더 중요한지 그리고 최종 잎사귀 질병 분류에 결정적인 역할을 해야 되는지를 알려줄 수 있다. 이를 위해, 본 연구에서는 분할 맵 예측을 위한 서브네트워크와 잎사귀 질병 분류를 위한 서브네트워크로 구성된 분할 인식 기반의 심층 합성곱 신경망을 제안하고자 한다. 두 종류의 서브네트워크는 먼저 전이 학습을 통해 미세 튜닝과정을 거친 후, 퓨전 레이어를 통해 두 종류의 서브네트워크를 결합한다. 좀 더 자세히 말하자면, 퓨전 레이어를 통해 분할 맵 서브네트워크에서 추정된 분할 맵의 결과와 입력 영상을 스택으로 쌓아 3차원 텐서를 형성하여 잎사귀 질병 분류를 위한 서브네트워크에 입력으로 준다. 그리고 종단 간 훈련 방법으로 전체 네트워크를 학습한다. 따라서 제안된 방법에서는 분할 맵 예측 서브네트워크에서 추정된 영상 분할 결과가 이후 잎사귀 질병 분류를 위한 서브네트워크에 전달됨으로써, 분할 맵에서 어떤 영역에 속한 특징벡터가 최종 잎사귀 질병 분류에 결정적인 영향을 미쳐야 되는지를 알려줄 수 있다. 즉, 분할 맵 예측 서브네트워크는 더 나은 잎사귀 질병 분류를 위한 가이드와 같은 역할을 한다. 기존

의 전이 학습 기반의 잎사귀 질병 진단은 이미 존재하는 아키텍처를 그대로 사용했지만 제안한 기법에서는 분할 인식 기반의 심층 합성곱 신경망 아키텍처를 새롭게 제안한다. 바로 이 점이 기존의 전이 학습 기법과 제안한 기법의 가장 큰 차이점이라 볼 수 있다.

1.2 기여도

첫째, 기존의 잎사귀 질병 분류 기술 중에 영상 분할을 사용하는 방식도 있다[15]. 그러나 제안한 방법처럼 종단 간 학습 방식이 아닌, 종래의 컴퓨터 비전 기술처럼 국부 특징 추출과 분류기 학습이 따로 구현된다. 또한 사용된 영상 분할 기법은 엄밀히 말해서 미리 정해진 색상만을 추출하는 군집화 기법에 속한다. 즉 제안한 방법처럼 잎사귀 영역, 배경 영역, 질병 영역을 구분하지 못한다. 또한 기존의 군집화 기반의 영상 분할 기술은 배경 영역의 색상과 잎사귀 질병의 색상과 유사할 경우 추출 성능이 떨어지는 단점이 있다. 또한 기존의 잎사귀 질병 진단 기술에서 사용된 특징은 계층적 구조에서 학습된 심층 특징이 아닌 수작업 기반의 특징이다. 이미 수많은 연구 결과에서 심층 특징 추출 기법이 수작업 기반의 특징 추출 기법보다 성능이 더 좋다는 것을 보이고 있다[11]. 그리고 기존의 연구에서는 사용된 영상은 배경색이 거의 단색이기 때문에 실제 환경을 반영하지 않고 있다. 그러나 본 연구에서 사용된 영상은 배경 영역에 잎사귀나 나뭇가지를 포함하기 때문에 좀 더 복잡하다.

둘째, 기존의 전이 학습 기법과 달리, 본 연구에서는 분할 인식 심층 합성곱 신경망이라는 새로운 아키텍처를 소개한다. 특히 잎사귀 영상에서 관심 영역인 배경 영역, 질병 영역, 잎사귀 영역을 예측하기 위한 분할 맵 예측 서브네트워크와 잎사귀 질병 분류를 위한 서브네트워크를 어떻게 각각 학습한 후, 종단 간 학습을 통해서 잎사귀 질병 분류의 성능을 제고할 수 있는지를 보이고자 한다. 또한 제안한 기법이 기존의 최첨단 기법, 즉 멀티스케일 특징 추출 및 풀링[16][17], 이선형 모델[18], 전이 학습[12]-[14]보다 더 우수한 성능을 달성할 수 있음을 확인하고자 한다.

II. 관련 연구

이 장에서는 최근 소개된 딥 러닝 기반의 영상 인식 기법을 소개하고자 한다. 이를 통해 제안한 기법과 기존의 기법과의 차이점을 이해하고자 한다.

2.1 멀티스케일 기반 심층 특징 추출 및 풀링

그림 1은 멀티스케일 기반의 심층 특징 추출, 풀링, 그리고 분류기 학습 과정을 보여주고 있다. 먼저, 스케일에 불변한 특징 추출을 위해, 영상의 크기를 가변하면서 멀티스케일 영상을 생성한다. 그리고 이 멀티스케일 영상을 사전에 학습된 VGG 모델에 입력한 후, 화살표로 표시된 레이어에서 특징 맵을 추출한다. 이 경우 512개의 28x28 특징 맵이 생성된다. 심층 특징 벡터란 특징 맵에서 동일한 위치에 있는 특징 값을 모두 모아 놓은 것을 말한다.

그림 1에서 빨간색 원으로 표시된 값을 하나의 벡터로 표현한 것이 바로 심층 특징 벡터이다. 그러나 이 심층 특징 벡터는 국부 영역의 영상 특성을 묘사할 수 있지만 부류 간 영상의 특성은 모델링하기에는 부족하다. 따라서 풀링 과정이 필요하다. 대표적인 풀링 기법에는 시각 단어 가방 모델[7][8]이나 피셔 벡터 인코딩[9] 기법이 있다. 이 방법을 사용하면 영상 단계의 특징을 모델링 할 수 있다.

그림 1에서 피셔 벡터 인코딩을 거치면 부류 간

영상을 표현할 수 있는 피셔 벡터들이 생성되고 이 벡터들이 최종 분류기의 입력으로 들어가게 된다. 본 연구에서는 서포트 벡터 머신을 분류기로 사용했고 다중 분류를 구현하기 위해 일대다 방식(One-Versus-All Manner)을 적용했다. 즉, 분류하고자 하는 부류에 속한 모든 피셔 벡터는 라벨 값을 '1'로 할당하고 나머지 모든 부류는 '-1'로 라벨 값을 할당한다.

그림 1에서 VGG 네트워크 대신에 기존의 수작업 기반의 국부 특징 추출을 사용할 수 있다. 예를 들어, VGG 네트워크 대신에 SIFT[4]를 사용해서 128차원의 특징 벡터를 추출할 수 있다. 만약 다른 조건이 모두 동일하다고 가정했을 때, 심층 특징 추출이 SIFT와 같은 수작업 기반의 특징 추출에 비해 성능이 더 우수한지 확인할 수 있다.

2.2 전이 학습

최근 연구에 따르면[11], 영상 분류를 위해 이미지 넷이라 불리는 영상 데이터베이스에 학습된 사전학습 모델(Pretrained Models)을 다른 영상 데이터베이스를 학습할 때 모델 파라미터의 초기 값으로 사용할 시 분류기의 성능을 높일 수 있다고 보고되고 있다. 이러한 학습 기법을 전이 학습이라 한다. 대표적인 사전학습 모델에는 VGG[11], ResNet[19]이 있다.

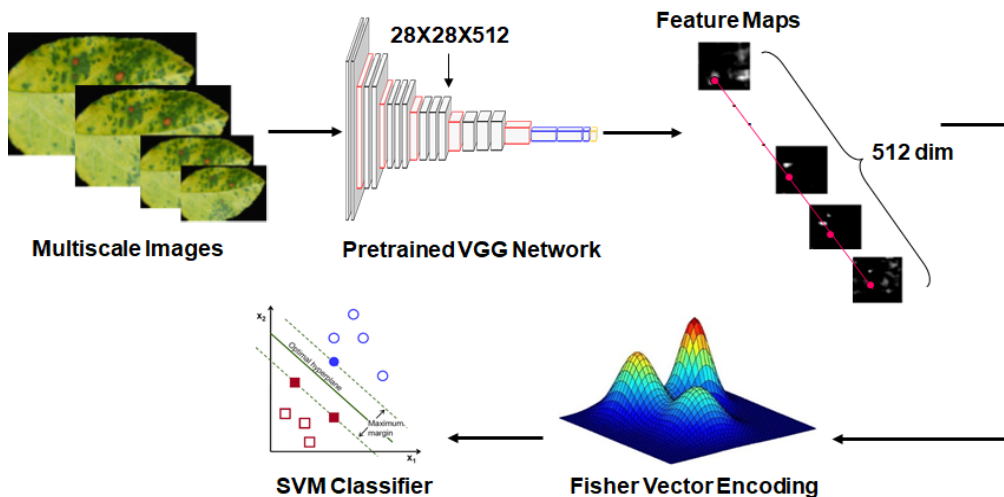


그림 1. 멀티스케일 기반 심층 특징 추출 및 풀링
Fig. 1. Multiscale-based deep feature extraction and pooling

그림 2는 1000개의 부류를 분류하기 위해 1000개의 소프트맥스 레이어(Softmax Layer)를 갖는 VGG 모델을 보여주고 있다. 본 연구의 사과나무 잎사귀 질병 진단이라는 새로운 분류 작업에 이 모델을 적용하기 위해서는 전이 학습 과정이 필요하다. 이를 위해, 먼저, 마지막 3개의 레이어를 제거한 후, 완전 연결 레이어(Fully-Connected Layer)와 소프트맥스 레이어를 추가한다. 그리고 새로운 사과나무 잎사귀 훈련 영상이 주어진 경우, 미니 배치 확률적 경사 하강법(Mini-Batch Stochastic Gradient Optimization)[20]을 사용해서 VGG 네트워크를 훈련한다. 그림 1에 제시된 멀티스케일 특징 추출 및 풀링 기법과 비교했을 때, 국부 특징 추출, 풀링, 그리고 분류기 과정이 하나로 통합된 것을 볼 수 있다.

또한 응용 도메인에 따른 특징 추출에 대한 지식이 없어도 쉽게 분류 문제를 해결할 수 있다. 하지

만 일반적으로 VGG와 같이 전이 학습은 고정된 입력 영상을 취한다. 이는 다양한 스케일에 따른 잎사귀 패턴이나 질병 무늬에 대한 정보가 상실되기 때문에 분류 성능의 저하가 발생할 수 있음을 의미한다.

2.3 이선형 모델

최근 이차 풀링(Second Order Pooling)[21]이 영상 인식에서 큰 성공을 거둔 후, 심층 합성곱 신경망 기반의 이선형 모델(Bilinear Model)이 개발되고 있다[18]. 그림 3은 국부 특징 추출과 풀링을 위한 심층 합성곱 신경망 기반의 이선형 모델을 보여주고 있다. 그림 1과 비교했을 때, 2종류의 심층 합성곱 신경망이 있는 것을 볼 수 있다.

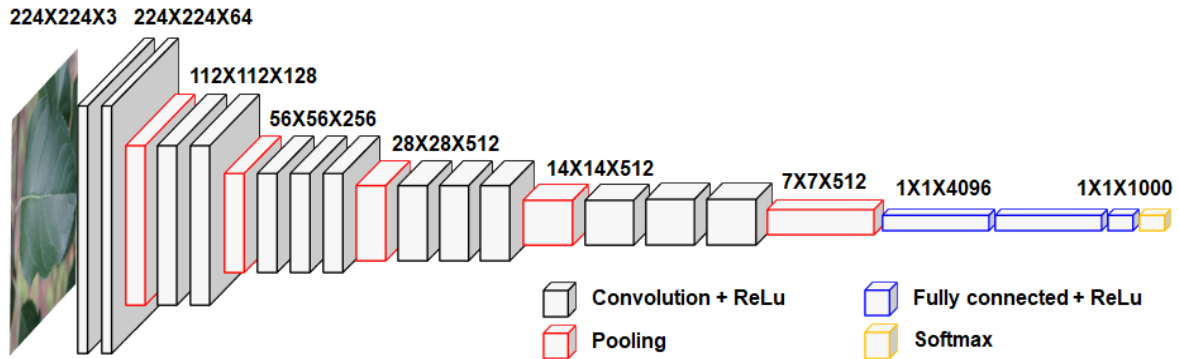


그림 2. VGG 네트워크의 아키텍처 [11]

Fig. 2. The architecture of the VGG network [11]

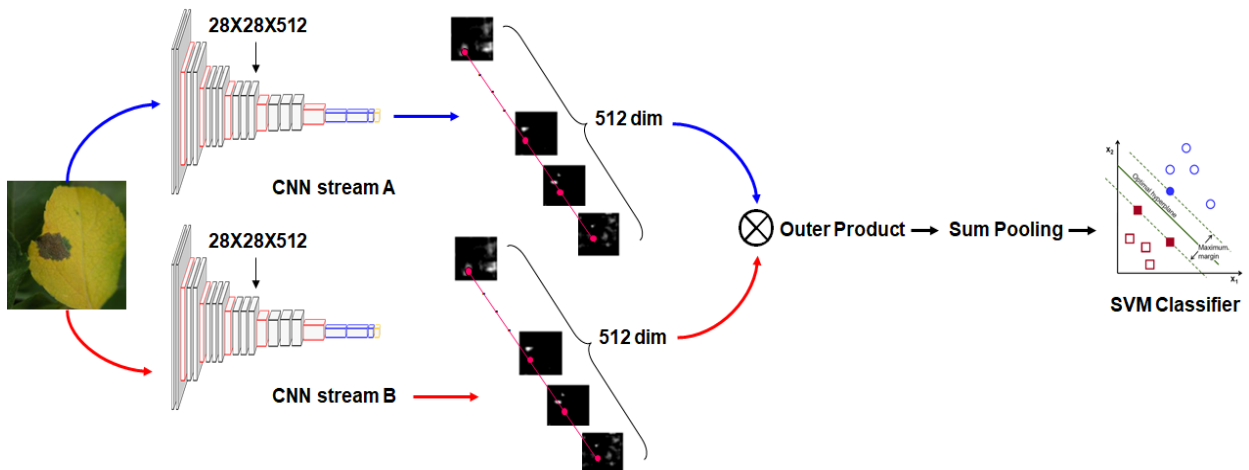


그림 3. 심층 합성곱 신경망 기반의 이선형 모델 [18]

Fig. 3. A bilinear model based on deep convolutional neural networks [18]

그림 1에서 심층 특징을 추출하는 것과 유사한 방식으로 두 종류의 심층 합성곱 신경망에서 심층 특징 벡터를 추출할 수 있다. 예를 들어, 그림 2에서 빨간색 원으로 표시된 곳에서 특징 맵의 값을 모으면 두 종류의 심층 특징 벡터를 추출할 수 있다. 그리고 두 종류의 특징 벡터를 외적(Outer Product)을 취해서 특징 요소간의 상관성을 모델링한다. 이 과정을 모든 특징 맵에 위치한 특징 벡터에 적용해서 단순히 합산한다. 이 과정을 합산 풀링(Sum Pooling)이라 부른다. 마지막으로 이 풀링된 벡터를 분류기의 입력으로 넣으면 최종 분류기를 학습할 수 있다.

III. 제안한 분할 인식 심층 합성곱 신경망

제안한 기법의 주된 아이디어는 질병이 위치하는

곳이 잎사귀 영역에 국한된다는 것이다. 반면에 배경 영역은 질병과 관련된 정보가 없다. 따라서 이러한 아이디어를 실현하기 위해서 본 연구에서는 그림 4와 같이 보이는 분할 인식 심층 합성곱 신경망 기반의 사과나무 잎사귀 질병 인식 기술을 제안한다.

그림 4에서 보이듯 제안한 네트워크에서는 두 종류의 서브네트워크가 있다. 하나는 분할 맵 예측 서브네트워크(Segmentation Map Prediction Sub-Network)이고 다른 하나는 잎사귀 질병을 예측하기 위한 VGG 서브네트워크이다. 분할 맵 예측 서브네트워크는 입력 영상을 3종류의 영역, 즉 배경 영역, 잎사귀 영역, 질병 영역으로 변환해주는 역할을 담당한다. 분할 맵 예측 서브네트워크에서 잘림 레이어(Cropping Layer)는 두 개의 특징 맵을 입력으로 받아 하나의 특징 맵의 크기를 다른 특징 맵의 크기로 변환한다.

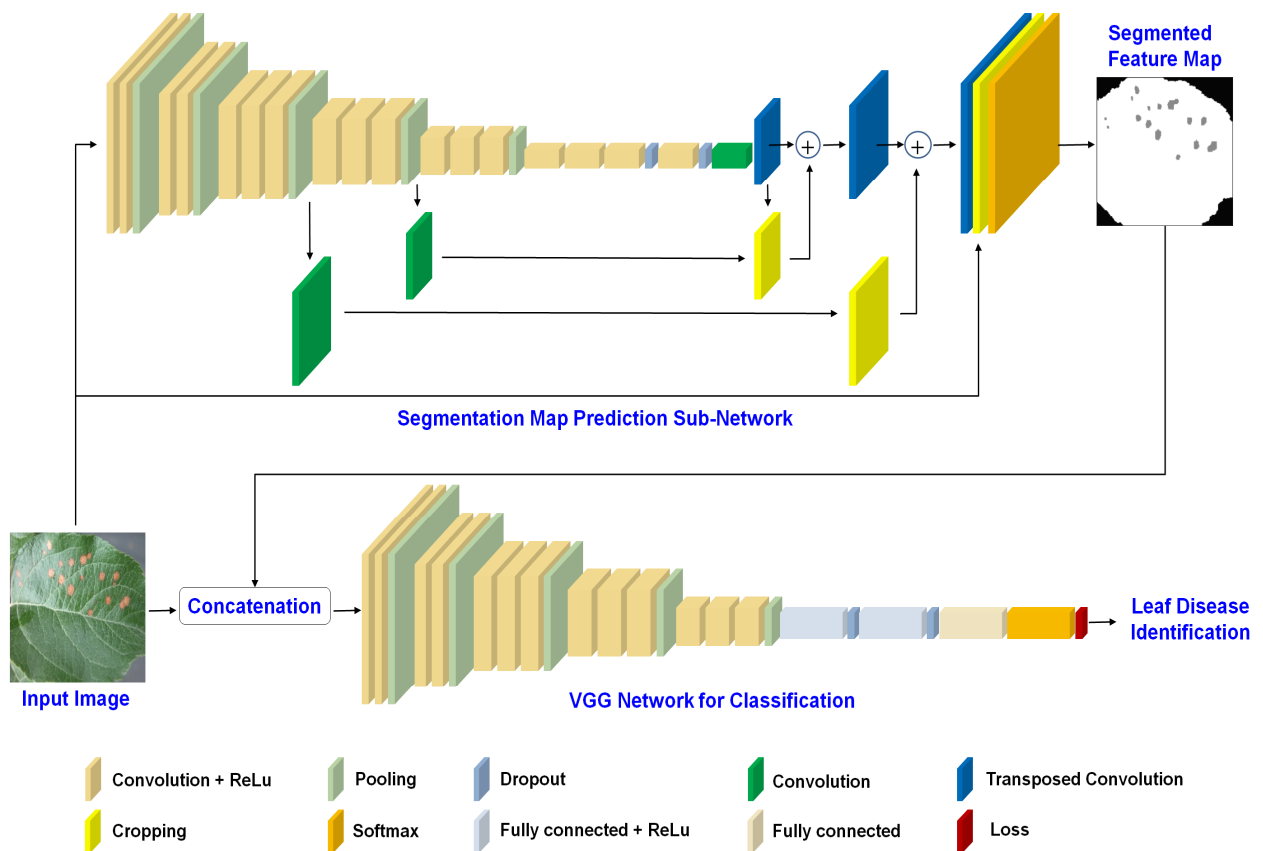


그림 4. 제안한 분할 인식 기반의 심층 합성곱 신경망
Fig. 4. Proposed segmentation-aware deep convolution neural network

그리고 전치 합성곱 레이어(Transposed Convolution Layer)는 입력 특징 맵을 크기를 확대한다. \oplus 는 두 입력 특징 맵을 요소 별로 합산한다. 연결 레이어(Concatenation Layer)는 분할 예측 서브네트워크에서 생성된 결과를 입력 영상 위에 쌓아서 3차원 텐서를 형성하고 이를 통해 분할 예측 서브네트워크와 VGG 서브네트워크를 연결한다. 그리고 종단 간 훈련을 통해 전체 네트워크를 학습하여 입력 영상으로부터 사과나무 잎사귀 질병을 인식한다.

그림 4에서 설명되지 않는 레이어, 예를 들어 합성곱 레이어(Convolution Layer), 정류선형유닛 레이어(Retified Linear Unit Layer: ReLU Layer) 등에 관한 자세한 정보는 참고문헌[11][22]을 보기 바란다.

그림 4에서 제시된 분할 맵 예측 서브네트워크는 시맨틱 분할(Semantic Segmentation)[23]에서 사용된 인코더-디코더 타입의 아키텍처이다. 하지만 시맨틱 분할의 목적은 입력 영상을 의미 있는 영역으로 나누는 것이지만 본 연구의 최종 목적은 사과나무 잎사귀의 질병 분류의 정확도를 개선하는 것이다. 다시 말하면, 본 연구에서 분할 맵 예측 서브네트워크의 역할은 입력 영상에서 배경 영역, 잎사귀 영역, 그리고 질병 영역으로 분할하여 이 정보를 VGG 서브네트워크에게 전달하는 것이다. 이를 통해, VGG 서브네트워크는 예측된 분할 맵에서 어떤 영역에 속한 특징이 최종 사과나무 잎사귀 질병 분류에 결정적인 역할을 할 것인지를 판단할 수 있다. 즉, 분할 맵 예측 서브네트워크는 더 나은 잎사귀 질병 분류를 위한 가이드와 같은 역할을 한다. 그림 4에서 분할 맵 예측 서브네트워크가 사라진다면 제안한 아키텍처는 기존의 VGG 네트워크와 동일해진다. 따라서 제안한 분할 인식 기반의 심층 합성곱 신경망의 성능을 VGG 서브네트워크와 비교해봄으로써, 분할 맵 영상을 사용하는 것이 실제로 인식을 증가에 도움이 되는지를 확인할 수 있다.

IV. 실험 결과

제안한 기법의 성능을 비교평가하기 위해, 기존의 최첨단 영상 인식 기술인 전이학습[12]-[14], 이 선형 모델[18], 멀티스케일 기반의 심층 특징 추출 및 풀링[16], SIFT 특징 추출 및 피셔 벡터 인코딩

[9] 기법을 실험하였다. 평가 척도는 정인식률(Correct Recognition Rates)로, 총 테스트 영상 개수에서 올바르게 추정된 영상 개수의 비율을 말한다. 피셔 벡터 인코딩에 사용된 가우시안의 개수는 256개이고, VGG 네트워크로써 VGG-16[11] 모델이 사용되었다. 그리고 제안한 기법은 매트랩(Matlab)으로 구현되었고 4개의 Titan-XP로 구성된 워크스테이션에서 학습되었다. 이때 훈련 영상과 테스트 영상의 비율은 1:1로 설정되었고 미니 배치 확률적 경사 하강법을 사용해서 파라미터를 갱신하였다.

4.1 사과나무 잎사귀 영상 수집

본 연구에서 사용된 모든 사과나무 잎사귀 영상은 농업진흥청 소속의 사과연구소에서 제공을 받았다. 사과나무 잎사귀 영상은 건전한 잎사귀, 갈색무늬병 잎사귀, 그리고 점무늬낙엽병 잎사귀, 총 3부류로 나누어진다. 각 부류에는 118장, 120장, 166장의 샘플 영상으로 구성되어 있다.

그림 5는 사과나무 잎사귀 영상의 예시를 보여주고 있다. 첫 번째 행은 건전한 잎사귀, 두 번째와 세 번째 행은 질병이 있는 잎사귀로써 각각 갈색무늬병 잎사귀와 점무늬낙엽병 잎사귀에 해당한다. 특히 갈색무늬병 잎사귀의 경우는 배경색의 색상과 질병의 색상이 유사한 것을 볼 수 있다. 이는 기존의 군집화 기반의 영상 분할 기법[15]은 배경 영역에 존재하는 건전한 잎사귀와 갈색무늬병의 색상을 구별하기가 어렵다는 것을 의미한다.

그리고 그림 5에서 보듯이 본 연구에서는 배경 영역이 단일 색이 아닌 나뭇가지나 잎사귀를 포함하고 있기 때문에 어느 정도 실제 환경을 반영했다고 볼 수 있다.

4.2 네트워크 학습

그림 4에 보이는 전체 네트워크를 학습하기 이전에 두 종류의 서브네트워크를 먼저 학습한다. 분할 예측 맵 서브네트워크를 학습하기 위해서는 모든 훈련 영상에 대해 영상 편집기를 사용해서 수작업으로 훈련 영상을 배경 영역, 잎사귀 영역, 질병 영역으로 구분할 필요가 있다. 이렇게 편집된 영상들

은 분할 예측 맵 서브네트워크의 로그 손실 레이어 (Log Loss Layer)에서 추정하고자 하는 목표 영상 (Target Images)으로 사용될 것이다. 모든 부류에 속

한 입력 영상에 대한 목표 영상들이 주어진 경우, 미니 확률적 경사 하강법을 사용해서 분할 예측 맵 서브네트워크를 학습한다.



그림 5. 사과 잎사귀 영상 예시: 건전 잎 영상 (첫 번째 행), 갈색무늬병 잎사귀 영상 (두 번째 행), 점무늬낙엽병 잎사귀 영상 (세 번째 행)

Fig. 5. Example of apple leaf images: normal leaf images (first row), diseased leaf images with marssonia blotch (second row), and diseased leaf images with alternaria leaf spot (third row)



그림 6. 입력 잎사귀 영상 (첫 번째 행), 목표 영상 (두 번째 행), 예측된 분할 영상 (세 번째 행)

Fig. 6. Input leaf images (first row), target images (second row), and predicted segmentation images (third row)

그림 6은 영상 편집기로 만든 목표 영상들과 분할 예측 맵 서버네트워크를 통해 추정된 결과 영상들을 비교해주고 있다. 첫 번째 행은 사과나무 잎사귀 영상이고 두 번째와 세 번째 행은 각각 목표 영상과 추정된 분할 맵이다. 비록 추정된 분할 맵에서 오류 픽셀이 존재할지라도 학습된 분할 예측 맵 서버네트워크가 우수한 성능을 보이고 있음을 알 수 있다.

실제 성능 평가에서도 약 93%의 평균 정인식률을 달성할 수 있었다. 여기서 평균 정인식률이란 각 부류 당 정인식률의 결과를 평균한 값을 말한다. 하지만 분할 예측 맵 서버네트워크의 목적은 영상 분할이 아니라, 최종 사과나무 잎사귀 질병 분류의 성능을 도와주는 것임을 유념해야 된다.

잎사귀 질병 분류를 위해 VGG 서버네트워크를 학습하기 위해서는 마지막 3개의 레이어를 제거한 후, 완전 연결 레이어(Fully Connected Layer), 소프트맥스 레이어, 로그 손실 레이어를 새롭게 추가해서 전이 학습을 실시한다. 실험 통해 VGG 서버네트워크는 약 74.7%의 정인식률을 달성할 수 있었다. 전체 네트워크는 그림 4와 같이 연결 레이어를 통해 분할 예측 맵 서버네트워크와 VGG 서버네트워크를 결합하여 앞서 영상 편집기로 생성한 목표 영상들과 부류 라벨 값을 사용해서 종단 간 학습을 한다. 이 과정을 통해, 분할 예측 맵 서버네트워크가 사과나무 잎사귀 질병 분류 목적을 위해 파라미터의 값이 변경될 것이다. 즉, 분할 맵 예측 서버네트워크는 더 나은 잎사귀 질병 분류를 위한 가이드와 같은 역할을 하게 된다.

4.3 성능 평가

표 1은 제안한 기법과 기존의 최첨단 기법의 정인식률 결과를 보여주고 있다. 먼저 제안한 기법의 정인식률의 결과가 VGG-16 전이 학습 기법[12][14]보다 약 9.6%나 증가한 것을 볼 수 있다. 이는 제안한 전체 네트워크에서 분할 예측 맵 서버네트워크를 제거할 시 9.6%의 정인식률 저하가 발생한다는 말과 동일하다. 이 결과를 통해 제안한 분할 예측 맵 서버네트워크의 사용이 최종 사과나무 잎사

귀 질병 분류의 성능을 개선하는 효과가 있음을 알 수 있다. 또한 제안한 분할 인식 심층 합성곱 신경망 기반의 사과나무 잎사귀 질병 분류 기법이 가장 우수한 성능을 달성한 것을 볼 수 있다.

그리고 멀티스케일 심층 특징 추출 및 풀링 기법이 VGG나 ResNet 전이 학습보다 성능이 우수한 것을 볼 수 있다. VGG 및 ResNet 네트워크는 고정된 영상 크기를 입력으로 받기 때문에 질병 분류에 큰 영향을 주는 잎사귀 패턴이나 질병의 텍스처 정보가 손실될 수 있다. 반면 멀티스케일 심층 특징 추출 및 풀링 기법은 다양한 스케일에서 풍부한 텍스처 정보를 활용하기 때문에 VGG나 ResNet 전이 학습보다 더 나은 성능을 얻은 것으로 분석된다. SIFT 특징 추출 및 피셔 벡터 인코딩 기법의 성능은 가장 낮은 것을 볼 수 있다. 이는 SIFT와 같이 수작업 기반의 특징 추출은 계층적 구조에서 학습된 심층 특정보다 특징 구별력이 떨어진다는 것을 말해준다.

V. 결론 및 향후 과제

본 논문에서는 사과나무 잎사귀의 질병 진단을 위해 분할 인식 기반의 심층 합성곱 기법을 새롭게 제안하였다. 잎사귀 질병이 잎사귀 영역에 존재하고 배경 영역에는 유익한 정보가 없다는 사실에 기초해서 분할 맵 예측 서버네트워크와 질병 진단을 위한 VGG 서버네트워크를 결합한 분할 인식 심층 합성곱 신경망을 제시하였다. 실험 결과에서 보듯이,

표 1. 정인식률 결과

Table 1. Results of Correct Recognition Rate

Methods	Correct recognition rates
TL using VGG network [12][14]	74.7%
TL using ResNet network [13]	76.6%
Multiscale-based deep feature extraction and pooling [16][17]	81.7%
FVE with SIFT [9]	71.7%
Bilinear model [18]	80.6%
Proposed segmentation-aware DCNN	84.34%

분할 맵 예측 서브네트워크의 사용은 VGG 서브네트워크에게 잎사귀 영역, 질병 영역, 배경 영역에서 어떤 특징 벡터가 최종 잎사귀 질병 분류에 결정적인 역할을 해야 되는지 가이드 역할을 수행함으로써, 약 9.6%의 정인식률의 개선을 도출할 수 있었다. 또한 본 연구를 통해서 두 서브네트워크를 각각 학습하고 종단 간 학습을 하는 방법론을 제시하였고 기존의 최첨단 기법들보다는 더 우수한 성능을 달성했음을 보여주었다.

References

- [1] Y.-S. Yoo, and B.-K. Kim, "Smart farm management system based on IoT," in Proc. Korean Institute of Information Technology Conference, Gumi, pp. 427-429, Jun. 2017.
- [2] <https://www.mouser.com/applications/smart-agriculture-sensors/> [Accessed: Mar. 25, 2019]
- [3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Proc. Neural Information Processing Systems, Harrah's Lake Tahoe, Nevada, pp. 1-9, Dec. 2012.
- [4] D. G. Lowe, "Distinct image features from scale-invariant key points," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, Nov. 2004.
- [5] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886-893, Jun. 2005.
- [6] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," in Proc. IEEE Transactions on Image Processing, vol. 19, no. 6, pp. 1657-1663, Jun. 2010.
- [7] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: spatial pyramid matching for recognizing natural scene categories," in Proc. IEEE International Conference on Computer Vision and Pattern Recognition, New York, NY, pp. 2169-2178, Jun. 2006.
- [8] C.-J. Kim, C.-H. Son, and D.-W. Jung, "A method of classifying species and conditions of street trees based on bag of visual words," Journal of Korean Institute of Information Technology, vol. 16, no. 3, pp. 115-123, Mar. 2018.
- [9] F. Perronnin, and C. Dance, "Fisher kernels on visual vocabularies for image categorization," in Proc. IEEE International Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, pp. 1-8, Jun. 2007.
- [10] J. Donahue et al., "DeCAF: A deep convolutional activation feature for generic visual recognition," arXiv:1310.1531 [cs.CV], Oct. 2013.
- [11] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv: 1409.1556 [cs. CV], Sept. 2014.
- [12] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Journal of Computers and Electronics in Agriculture, vol. 145, pp. 311-318, Feb. 2018.
- [13] 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," Journal of Sensors, vol. 17, no. 9, pp. 1-21, Sept. 2017.
- [14] S. P. Mohanty, D. Hughes, and M. Salathe, "Using deep learning for image-based plant disease detection," Journal of Frontiers in Plant Science, vol. 7, article 1419, Sept. 2016.
- [15] H. Ali, M. I. Lali, M. Z. Nawaz, M. Sharif, and B. A. Saleem, "Symptom based automated detection of citrus diseases using color histogram and textural descriptors," Journal of Computers and Electronics in Agriculture, vol. 138, pp. 92-104, Jun. 2017.
- [16] M. Cimpoi, S. Maji, and A. Vedaldi, "Deep filter

banks for texture recognition and segmentation," in Proc. IEEE International Conference on Computer Vision and Pattern Recognition, Boston, MA, pp. 3828 – 3836, Jun. 2015.

- [17] C.-H. Son, "Recognizing species and conditions of street trees based on multiscale and deep Feature extraction," Journal of the Institute of Electronics and Information Engineers, vol. 55, no. 4, pp. 70-78, Apr. 2018.
- [18] T.-Y. Lin, A. RoyChowdhury, and S. Maji, "Bilinear CNN models for fine-grained visual recognition," in Proc. IEEE International Conference on Computer Vision, Santiago, Chile, pp. 1449-1457, Dec. 2015.
- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, pp. 770-778, June 2016.
- [20] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv:1609.04747 [cs.LG], Sept. 2016.
- [21] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes, "Bilinear classifiers for visual recognition," in Proc. Neural Information Processing Systems, Vancouver, pp. 1-9, Dec. 2009.
- [22] Y.-S. Choi, and C.-H. Son, "Rain Removal via Deep Convolutional Neural Networks Considering Orientation and Strength of Rain Streaks," Journal of Korean Institute of Information Technology, vol. 17, no. 1, pp. 85-98, Jan. 2019.
- [23] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proc. IEEE Conference on Computer Vision and Pattern Recognition, Boston, pp. 3431-3440, Jun. 2015.

저자소개

유 희 진 (Hee-Jin Yu)



2019년 2월 : 군산대학교
소프트웨어융합공학과(공학사)
관심분야 : 컴퓨터 비전, 영상처리,
딥 러닝, 프로그래밍

손 창 환 (Chang-Hwan Son)



2002년 2월 : 경북대학교
전자전기공학부(공학사)
2004년 2월 : 경북대학교
전자공학과(공학석사)
2008년 8월 : 경북대학교
전자공학과(공학박사)
2008년 8월 ~ 2009년 12월 :
삼성전자 프린팅사업부 책임연구원
2015년 3월 ~ 2017년 2월: 캐나다 Ryerson 대학교
포닥연구원
2017년 4월 ~ 현재 : 군산대학교 소프트웨어융합공학과
조교수
관심분야 : 컴퓨터 비전, 영상처리, 기계학습, 딥 러닝