

# 오브젝트 중심점-마스크를 사용한 instance segmentation

( An Instance Segmentation using Object Center Masks )

이종혁\*, 김형석\*\*

(Jong Hyeok Lee, Hyong Suk Kim)

## 요약

본 논문에서는 새롭게 제안하는 Multi-Path Encoder-Decoder 의 구조를 바탕으로 두개의 가지로 구성된 심층신경망을 통해서 영상 이미지에서 물체를 하나의 객체 단위로 분할 검출하는 방법을 제안하였다. 각 가지는 중심점 검출 가지(Dot branch), 객체 분할 가지(Segmentation branch)라 하고 중심점 검출 가지는 이미지로부터 각 객체의 중심점을 찾는 역할을 수행하고, 객체 분할 가지는 각 객체의 영역을 이미지로부터 분할하는 역할을 수행한다. 실험에서는 CVPPP 식물 이미지의 나뭇잎을 각각 구분하도록 학습 하였으며 중심점 검출 가지는 각 나뭇잎의 중심점들을 찾아내고, 객체 분할 가지는 원본 이미지와 찾아낸 중심점 이미지를 통하여 각 중심점에 해당하는 나뭇잎의 픽셀 분할 영역을 최종적으로 예측하게 된다. 기존의 객체 분할에서는 다양한 크기, 위치의 앵커박스를 만들어서 많은 영역( $N > 1k$ )의 물체를 확인해야하는 연산량 문제점 혹은 이미지에서 고정되지 않는 총 객체의 개수를 예측하기 어려웠던 문제가 있었다. 제안한 심층신경망에서는 중심점을 기반으로 객체를 찾아내는 효과적인 방법을 제안하였다.

■ 중심어 : 심층신경망 ; 객체분할 ; 중심점 검출 ; 물체 검출

## Abstract

In this paper, we propose a network model composed of Multi path Encoder-Decoder branches that can recognize each instance from the image. The network has two branches, Dot branch and Segmentation branch for finding the center point of each instance and for recognizing area of the instance, respectively. In the experiment, the CVPPP dataset was studied to distinguish leaves from each other, and the center point detection branch(Dot branch) found the center points of each leaf, and the object segmentation branch(Segmentation branch) finally predicted the pixel area of each leaf corresponding to each center point. In the existing segmentation methods, there were problems of finding various sizes and positions of anchor boxes ( $N > 1k$ ) for checking objects. Also, there were difficulties of estimating the number of undefined instances per image. In the proposed network, an effective method finding instances based on their center points is proposed.

■ keywords : neural network ; instance segmentation ; center detection ; object detection

## I. 서론

영상 처리에서 심층신경망이 신뢰할만한 성능을 보여줌에 따라 객체 분할 분야에서도 많은 심층신경망이 좋은 결과를 내고 있다. 영상 이미지에서 물체를 찾아내어 구분하는 방법은 일반적인 분류 보다 복잡한 문제이다. 영상 이미지에서 각 물체의 위치를 찾아내야 할 뿐만 아니라 각 물체들이 해당하는 클래스를 구분하여야 한다. 영상 이미지에서 물체를 찾아내는 방법은 영역 상자(bounding box)를 이용하여 물체의 위치를 찾아내는

물체 검출과 픽셀 단위로 클래스의 의미를 구분하는 객체 분할 방법 등이 있다. 영역 상자를 사용하는 물체 검출은 찾고자 하는 물체의 위치를 특정 영역에 한정시켜서 구분하게 된다. 객체 분할문제는 물체 검출 문제와 비슷하지만 물체 검출의 경우 각 물체의 영역에 해당하는 데이터만 예측 하는 것 과 달리 각 물체에 해당하는 픽셀들을 구분하는 문제 또한 해결해야 한다. 일반적인 객체 분할 문제는 같은 클래스에 해당하는 모든 픽셀들은 같은 값을 가지게 된다. 여기서 발생하는 문제점은 같은 클래스를 의미하지만 다른 객체인 경우(e.g. 식물의 잎을 구분하는 문제를 해결할 때 각 잎을 따로 구별해야 하는 경우, 돼지

\* 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

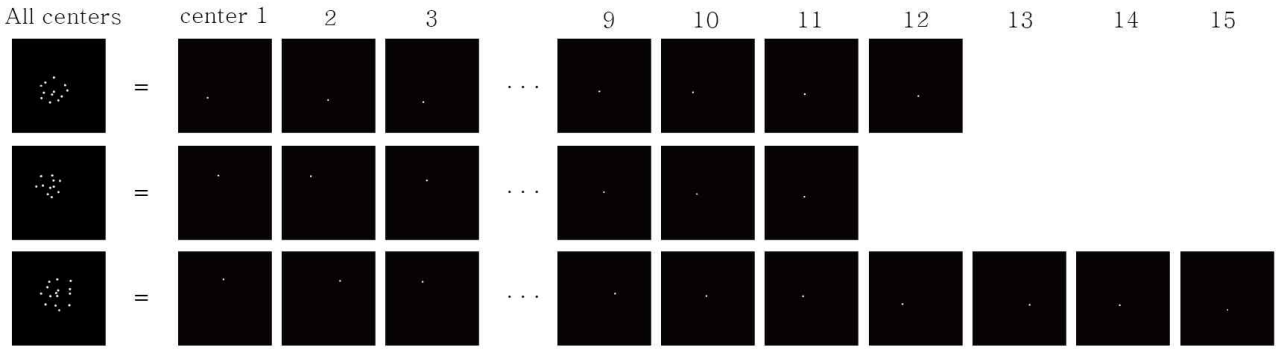


그림 1. 각 잎들의 중심점 이미지. 좌측은 모든 중심점들이 표시되어 있으며 하나의 객체영역을 분할하기 위해서 각 중심점 별로 이미지를 나눈다. 하나의 중심점 이미지는 segmentation 가지의 입력으로 들어가게 된다.

사육장에서 각 왜지들을 구분해야 하는 경우 등) 모든 물체들은 같은 클래스로 분류되기 때문에 서로 구분할 수 없다. 이를 구분하기 위해서 같은 클래스 안에서 객체를 구분하여 각각의 객체 단위의 의미로 대상을 구분할 수 있는 방법을 객체 분할(instance segmentation) 이라고 한다.

객체분할 문제를 해결하는 방법으로 여러 방식의 네트워크의 구조가 제안되어 왔지만 본 논문에서는 물체의 중심점을 기반으로 객체를 구분하는 새로운 Multi-Path Encoder-Decoder 심층신경망 구조를 제안한다[22]. 본 심층신경망은 CVPPP 2017 DATASET(A1)을 이용하여 학습을 진행하였으며, 데이터 셋 에서 각 잎의 중심점들을 기반으로 각 나뭇잎을 구분하는 것을 목적으로 설계되었다[1].

심층신경망은 크게 두 단계로 나누어서 설명할 수 있다. 첫 번째 단계는 입력된 영상 이미지를 기반으로 각 나뭇잎들의 중심점을 찾아낸다. 이 단계에서 심층신경망은 입력 영상으로부터 나뭇잎의 중심점에 해당하는 영역을 집중적으로 검출 할 수 있는 필터를 학습하게 된다. 두 번째 단계는 찾아낸 중심점 영역을 기반으로 중심점에 해당하는 하나의 나뭇잎의 영역을 분할하게 된다. 이때 첫 번째 단계 엔코더 마지막단의 특징 맵 과 두 번째 단계 엔코더 마지막단의 특징 맵이 원소별로 더해져서(element-wise add) 두 번째 단계의 디코더로 들어간다. 두 번째 단계의 디코더는 해당 중심점에 해당하는 나뭇잎의 분할된 영역을 출력하게 된다.

## II. 관련 연구

영상 이미지에서 물체를 찾는 방법은 물체 검출과 객체 분할 방법 등을 이용하는 여러 심층신경망들이 있다. 이미지에서 각 객체들의 중심점을 찾아서 물체를 검출하는 방법은 아래와 같이 연구가 진행되었다.

Corner-Net은 점을 찾는 물체 검출의 방법으로 좌 상단, 우 하단의 모서리 점 쌍을 찾아서 영상 이미지에서 물체를 찾는 다[2]. Corner-Net은 물체 검출의 문제를 해결하게 위해서 많

은 앵커박스(i.e. proposed bounding box)를 사용함으로써 물체의 모서리를 찾아서 물체를 구분하는 방법이다[3-5].

Extreme-Net은 Corner-Net보다 많은 수의 점 쌍을 그룹 지어서 물체를 찾아낸다[6]. 이렇게 찾아진 박스 영역은 단순한 사각형의 예측영역보다 좀 더 세밀한 예측영역을 구분한다. Center-Net은 Corner-Net 이 좌 상단, 우 하단의 모서리 점을 하나의 쌍으로 물체를 찾은 것과 달리 물체의 중심점에 집중하였다. 논문에서는 각 물체의 중심점을 찾은 후 물체의 영역을 찾을 수 있다고 설명한다[7].

앞서 언급한 연구는 단순히 물체의 영역을 영역박스로 표시하여 검출하는 심층신경망들이며 본 논문에서는 단순히 영역 박스가 아닌 픽셀단위로 세밀하게 측정된 물체의 영역을 검출하는 객체분할 방식의 심층신경망을 제시하고 있다.

이와 같이 객체 분할 방식의 심층신경망으로 R-CNN구조의 대표적인 객체 분할 심층신경망인 Mask R-CNN 은 물체 검출 방식의 심층신경망인 Faster R-CNN에 객체분할을 수행하는 가지를 추가하여 이미지에서 객체별 픽셀 영역을 찾아낸다 [4][12].

discriminative loss는 각 객체를 구분할 수 있는 오차로

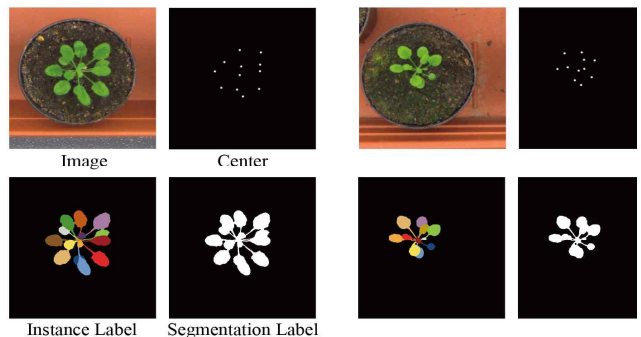


그림 2. CVPPP(A1) 데이터 셋의 구조

이미지의 각 객체는 심층신경망을 통해서 새로운 임베딩 공간에 맵핑된다. 이때 같은 객체를 의미하는 픽셀은 서로 같은 곳

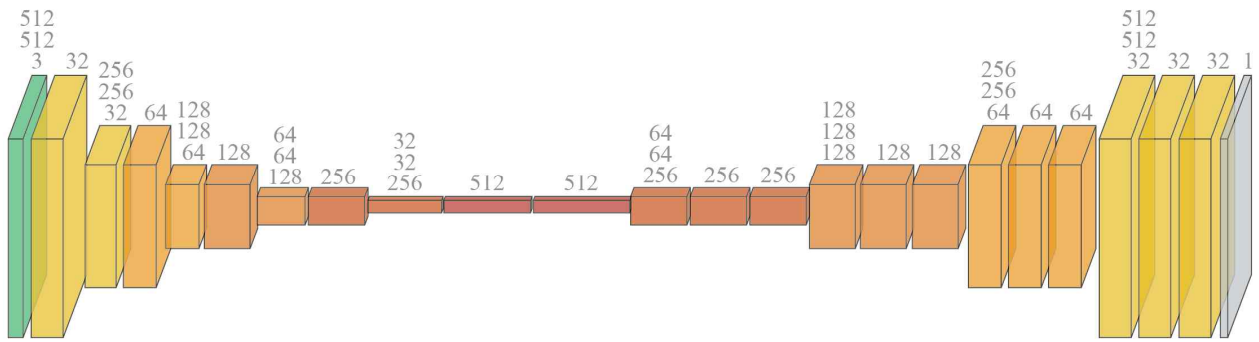


그림 3. Dot branch 의 구조

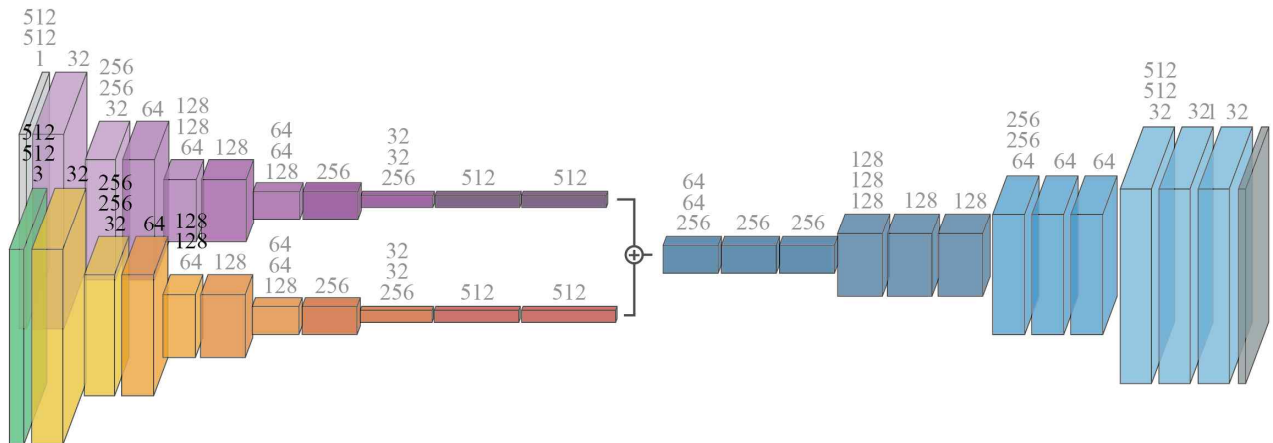


그림 4. Segmentation branch의 구조. 중간의 (+) 는 원소별 합산을 의미한다.

에 위치하게 되며 서로 다른 객체를 의미하는 픽셀들의 중심점은 서로 멀어지게 배치된다[11].

본 논문의 심층신경망의 구조는 Discriminative loss를 이용한 객체분할과 Corner-Net의 중심점을 찾는 방법에서 영감을 받아서 설계하였다. 각 객체의 픽셀들을 다른 임베딩 공간에서 하나의 중심으로 맵핑하는 동작을 본 심층신경망에서는 한 채널의 공간상에서 중심으로 맵핑된다고 표현할 수 있다. 이를 통해서 각 일들의 중심점을 찾을 수 있었고 그러한 중심점을 기반으로 Faster R-CNN과 같이 특징 맵을 공유하는 두 번째 단계를 이용하여 각 객체를 찾아낼 수 있었다[12].

### III. 제안한 방법

본 심층신경망은 CVPPP 데이터 셋의 객체 분할 문제를 해결하기 위해서 설계 되었으며 객체 분할 방식의 물체 검출은 다운-샘플링과 업-샘플링을 통한 엔코더-디코더 구조에서 좋은 성능을 보이고 있어 Multi-Path Encoder-Decoder 구조로 구성하였다[8-10][20-21]. 심층신경망은 각 객체(i.e. CVPPP 데이터 셋의 각 나뭇잎)의 중심점을 찾고, 찾은 중심점으로부터 해당하는 객체의 픽셀단위 분할 영역을 예측한다.

#### 1. CVPPP 데이터 처리

CVPPP 데이터 셋은 위에서 아래로 촬영한 식물 이미지 모음으로 촬영한 식물 데이터로 그림 2와 같은 이미지와 레이블로 이루어져 있다. 이미지 사이즈를 512 x 512 사이즈로 통일하고 이미지 내의 모든 나뭇잎의 중심점을 표시한 이미지를 통해서 각 중심점을 표시한 이미지를 생성하였다. 그림 1은 생성된 중심점 이미지를 보여주며 각 행의 가장 좌측의 이미지는 전체 잎들에 대한 중심점들의 위치를 표시한 이미지, 우측의 이미지들은 각 하나의 중심점들을 표시한 이미지를 보여준다. 표시된 중심점의 크기는 나뭇잎과 이미지의 사이즈를 고려하여 반지름 5픽셀의 원을 그려 생성하였다. 또한 2번째 단계를 정확하게 학습시키기 위해서 각 중심점에 해당하는 원을 가지는 이미지를 생성하였다(입력 이미지만다 중심점의 총 수는 나뭇잎의 개수와 같으며 그림 1과 같이 만들어진 각 중심점 이미지의 수도 다르다). 그림 2에서 각 중심에 해당하는 나뭇잎을 객체 분할을 위한 레이블을 각 분할 이미지로 나누어 구성하였다. 분할된 나뭇잎 이미지의 개수는 중심점의 개수와 동일하며 중심점으로부터 하나의 나뭇잎을 예측하는 심층신경망을 학습하는데 사용된다.

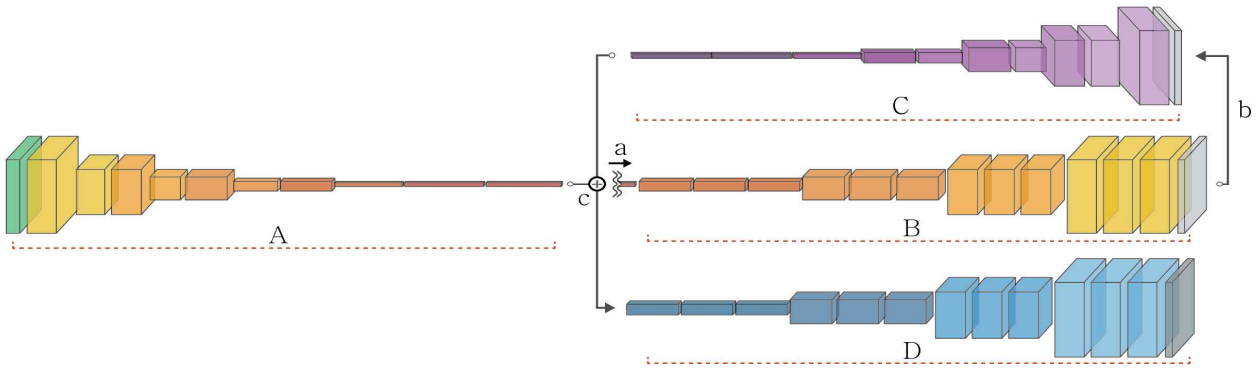


그림 5. 전체 심층신경망의 구조. a, b, c 는 데이터의 흐름을 A, B, C, D 는 각 Encoder, Decoder 영역을 의미한다.

#### Algorithm Sequence of training branch

**Input:** RGB image

**Output1:** Gray-Scale center point marked image (hypotheses)

**Output2:** segmentation of specific instance based on center point (hypotheses)

**Label1:** Gray-Scale center point marked image

**Label2:** segmentation of specific instance based on center point

initialize variables

for epoch do

  while remain center points > 0

    output1 ← Train 'Dot branch' using {input, label1}

    output2 ← Train 'Segmentation branch'

      using {output1, Dot branch middle feature map, label2}

  save model

end for

#### 알고리즘 1. 학습 순서 알고리즘

## 2. 중심점 검출 가지

중심점 검출 가지는 입력 이미지로부터 각 나뭇잎의 중심점을 찾아낸다. 그림3은 중심점 검출가지의 구조를 보여준다. 512 x 512 사이즈의 이미지를 입력으로 하는 인코더-디코더 구조를 가지고 있으면서 출력으로 그림 1의 가장 좌측의 이미지들과 같은 중심점을 나타내는 이미지를 출력한다. 중심점 검출 가지는 Center-net의 중심점 검출방법과 같이 focal loss를 이용하여 학습하였다[7][13]. 중심점 검출 가지는 학습에 따라서 입력 이미지를 1차원의 공간에 임베딩 하는 역할을 하게 된다. 그림 4에서 중심점 검출 가지의 인코더 부분의 마지막 특징 맵은 입력 이미지에 대한 특징을 담고 있고 이는 객체 분할 가지의 특징 맵과 합쳐져 다시 디코딩 된다.

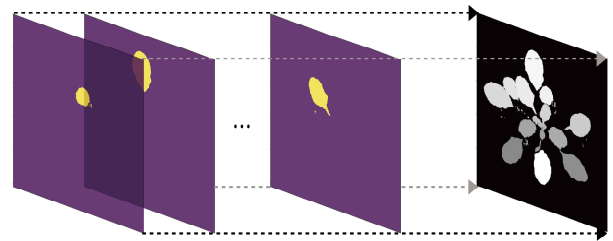


그림 6. 중복영역 제거 및 최종 결과 출력.

## 3. 객체 분할 가지

객체 분할 가지는 하나의 중심점 이미지와 원본 이미지를 통해서 중심에 해당하는 부분의 객체를 분할한다. 그림 4는 객체 분할 가지의 구조를 보여준다. 입력 이미지와 중심점 이미지(여기서 중심점 이미지는 모든 객체의 중심이 표시된 이미지가 아닌 하나의 잎에 대한 중심을 표시한 이미지를 의미한다.)는 각각 인코더 구조의 심층신경망에 들어가며 가장 마지막 단에서 원소별로 더해진다.(element-wise add). 이렇게 합쳐진 특징 맵은 하나의 디코더 구조를 통해서 원본과 같은 사이즈의 바이너리 레벨 이미지 데이터를 출력하게 된다.

## 4. 두 가지의 결합

그림 5는 중심점 검출 가지와 객체 분할 가지가 합쳐진 전체적인 심층신경망의 구조를 보여주고 있다. 중심점 검출 가지는 원본 이미지를 입력으로 가지며 데이터는 그림 5-A로 들어가서 그림 5-a 와 같이 곧바로 B-디코더를 통해서 출력 된다. 출력된 한 채널의 이미지는 그림 1 좌측의 중심점 이미지와 같은 데이터이다. 출력된 데이터는 그림1과 같이 각 중심점을 나타내는 이미지들로 분할되며 한 이미지 씩 그림 5-b 와 같이 그림 5-C의 입력으로 들어가게 된다. 그림 5-C에서 출력되는 특징

Epoch	10		80		130		250	
target image								
output image								
mIOU	0.776321		0.760910		0.794310		0.773016	

그림 7. 학습 epoch 에 따른 Test 데이터 셋의 출력 결과

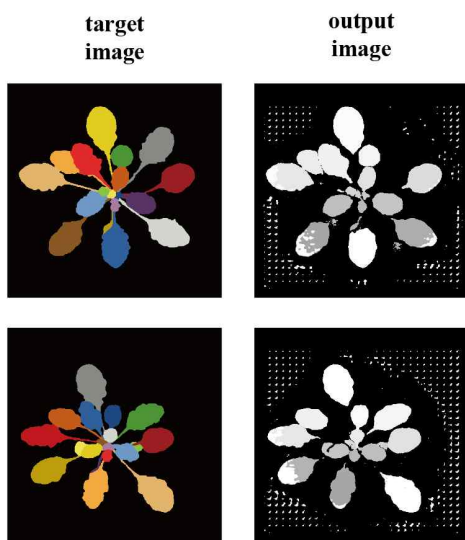


그림 8. 학습이 잘 안되었을 경우 일부 test dataset에서 발생하는 중복 노이즈 영역.

맵은 그림 5-A를 지나며 생성되었던 특징 맵과 더해진다 (element-wise add, 그림 5-c). 더해진 특징 맵은 그림 5-D를 통해서 첫 입력 이미지와 같은 너비와 높이를 가지는 한 채널의 이미지가 된다. 이렇게 출력된 이미지는 그림 5-b에서 나누어진 중심점의 개수  $N$  만큼 생기게 되며 하나의 이미지에 대해서  $\{N, H, W, 1\}$ 의 이미지를 출력한다. 출력된 이미지들은 각각 하나의 나뭇잎의 영역은 1, 이외의 영역은 0으로 나타내어지는 바이너리 스케일의 이미지이다.

마지막으로 결과를 출력하기 위해서는 출력된 결과를 이용하여 하나의 이미지로 만들어야 한다. 그림 6은 출력 이미지를 합치는 방법을 시각적으로 보여준다. 각 객체는 특정 값들을 할당 받고 하나의 이미지로 통합되어진다.

#### IV. 실험 및 결과

심층신경망의 학습은 250 epoch 동안 학습률은 0.001로 Adam optimizer를 이용하여 학습하였다[14]. 중심점 검출 가지의 오차를 이용하여 학습을 진행하고 이후 하나의 중심점에 해당하는 레이블 이미지를 통해서 객체 분할 가지를 학습하였다.

알고리즘 1은 학습 과정의 알고리즘을 보여주고 있다. 한 epoch 가 진행되는 동안 각 중심점에 대한 심층신경망의 학습이 이루어진다. 우선 RGB 채널의 이미지를 입력으로 각 나뭇잎의 중심점을 추정하는 심층신경망을 학습시킨다. 다음으로 각 중심점에 해당하는 객체 레이블을 이용하여 중심점으로 부터 하나의 객체를 학습시킨다. 이때는 바로 직전 학습했던 중심점 검출 가지의 엔코더 마지막단의 특징 맵과 원소별로 덧셈을 진행한다. 모든 중심점에 대한 학습이 진행되면 다음 이미지에 대한 학습을 동일한 방식으로 진행한다.

그림 7은 테스트 데이터 셋을 이용하여 학습된 심층신경망의 결과를 나타내고 있다. 네트워크의 평가에 각 객체의 표시된 중심점을 이용한 출력의 결과를 표시하였으며, 중심점 검출 가지의 출력은 좀 더 좋은 결과를 보여주기 위해서 출력된 이미지의 중심점들을 기준으로 5픽셀의 반지름을 가지는 원을 그려 객체 분할 가지의 입력으로 이용하였다.

객체 분할 가지의 출력을 합쳐 최종 이미지를 얻을 수 있는데 각 이미지는 서로 다른 객체를 의미하기 때문에 서로 같은 위치의 픽셀에 값을 가질 수 없다. 따라서 중복되는 영역을 제거하기 위해서 각 이미지별로 같은 위치에 값이 존재하는 이미지가 두 장 이상일 경우 해당 위치에 해당하는 모든 값을 0으로 두어 중복 영역을 제거하였다. 그림 6은 한 이미지에 해당하는 객체 분할 가지의 모든 출력을 합산하는 과정에서 그림 8과 같은 중복 영역을 제거하는 데이터 후처리 방법을 시각적으로 보여준다. 각각의 객체들은 최종 출력 이미지에서 구분이 되어야 하므로 이미지에서 나뭇잎의 최대 개수를 고려하여 0.5에서 1의 값 사이의 18단계로 구분한 값 [0.5, 0.529, 0.558, 0.588, 0.617,

Network	CVPPP TEST (A1) %
RIS + CRF[15]	66.6
IPK [16]	74.4
Brabandere et al. [17]	84.2
Ren et al. [18]	84.9
<b>Ours</b>	<b>87.1</b>
Daniel et al. [19]	90.0
Mask r-cnn [4]	90.0

표 1. 심층신경망 성능 비교 (SBD).

0.647, 0.676, 0.705, 0.735, 0.764, 0.794, 0.823, 0.852, 0.882, 0.911, 0.941, 0.970, 1] 을 각 이미지에 하나씩 할당하여 최종 이미지를 출력하였다. 따라서 최종 이미지의 각 앞은 각각 다른 강도의 값을 가지게 된다.

표 1은 객체 분할 동작을 하는 여러 심층신경망과 제한한 심층신경망의 CVPPP (A1) 데이터 셋에 대한 성능을 비교한 결과이다. Symmetric Best Dice score(SBD, %)는 모든 물체에 대해서 평균 최고 다이스 점수를 의미하며 다음의 (1), (2)의 수식에 의해서 계산되어진다.  $|*|$ 는 객체의 픽셀의 영역을 의미하며  $1 \leq i \leq M$ ,  $1 \leq j \leq N$ 의 수는 각 앞에 대한 분할 쌍을 의미한다.

$$BD(P_{g,t}, P_{predict}) = \frac{1}{M} \sum_{i=1}^M \max_{1 \leq j \leq N} \frac{2|P_{g,t(i)} \cap P_{predict(j)}|}{|P_{g,t(i)}| + |P_{predict(j)}|} \quad (1)$$

$$SBD(L_{predict}, L_{g,t}) = \min\{BD(L_{predict}, L_{g,t}), BD(L_{g,t}, L_{predict})\} \quad (2)$$

#### IV. 결론

앞선 실험으로 다른 역할을 가지는 두 개의 심층신경망 가지를 이용하여 전체 물체에서 한 객체의 중심점으로, 중심점에서 하나의 픽셀단위 객체 영역을 검출하는 심층신경망을 구성하고 심층신경망의 학습 진행정도에 따라서 심층신경망의 성능을 확인하였다.

심층신경망의 출력이 다시 심층신경망의 인풋으로 들어가는 구조로 인하여 심층신경망의 최종 출력은 심층신경망의 학습 정도에 따라서 영향을 많이 받게 된다. 그림 9는 객체의 중심점을 예측한 심층신경망의 출력과 목표이미지를 보여준다. 좌측의 이미지는 심층신경망의 실제 출력을 보여주고 우측의 이미지는

표시된 중심점 이미지를 보여준다. 가운데 이미지의 붉은 점들은 예측과정에서 잘 예측하지 못한 중심점 위치들을 나타낸다.

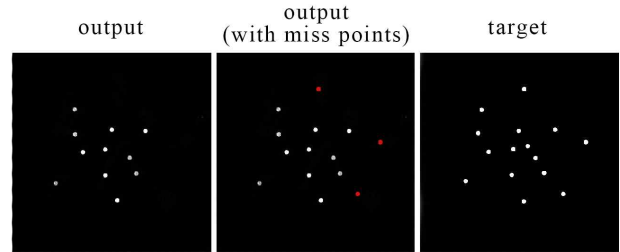


그림 9. 중심점 테스트 예측 중 소실된 중심점. 붉은 점은 예측에서 소실된 중심점들을 의미한다.

향후 연구에서는 물체의 중심점을 예측할 때 이처럼 소실되는 데이터를 방지하기 위해서 각 객체 영역으로부터 각 점을 예측하는 방법과 각 객체에 해당하는 중심점을 구별하는 방법, 중심을 찾는 심층신경망과 특징 맵을 공유하여 온전히 하나의 객체 검출에 집중하는 심층신경망 구조에 대한 연구를 하고자 한다.

#### REFERENCES

- [1] CVPPP 2017 LSC TRAINING DATASET, <https://www.plant-phenotyping.org/CVPPP2019> (accessed Dec., 2019).
- [2] Hei Law, Jia Deng, "CornerNet: Detecting Objects as Paired Keypoints," *ECCV*, 2018.
- [3] T. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, "Focal Loss for Dense Object Detection," *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2999-3007, Venice, 2017.
- [4] K. He, G. Gkioxari, P. Dollr, and R. Girshick. "Mask r-cnn," *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2980 - 2988, Oct 2017.
- [5] Joseph Redmon and Ali Farhadi. "Yolov3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.
- [6] X. Zhou, J. Zhuo, and P. Krahenbuhl. "Bottom-up object detection by grouping extreme and center points," *CoRR*, abs/1901.08043, 2019.
- [7] Kingyi Zhou, Dequan Wang, and Philipp Krähenbühl. "Objects as points," arXiv preprint arXiv:1904.07850, 2019.
- [8] Long, J., Shelhamer, E., and Darrell, T. "Fully

- convolutional networks for semantic segmentation," *CoRR*, abs/1411.4038, 2014.
- [9] Badrinarayanan. V, Kendall. A, and Cipolla. R. "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 39, pp. 2481 - 2495, 2017.
- [10] O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234 - 241, 2015.
- [11] B. D. Brabandere, D. Neven and L. V. Gool, "Semantic Instance Segmentation with a Discriminative Loss Function," *CoRR*, abs/1708.02551, 2017.
- [12] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2017.
- [13] Lin, T.Y., Goyal, P., Girshick, R., He, K., Doll'ar, P, "Focal loss for dense object detection," arXiv preprint arXiv:1708.02002, 2017.
- [14] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *CoRR*, abs/1412.6980, 2014.
- [15] Bernardino Romera-Paredes and Philip Hilaire Sean Torr, "Recurrent instance segmentation," *European Conference on Computer Vision (ECCV)*, pp. 312 - 329, 2016.
- [16] Jean-Michel Pape and Christian Klukas, "3-d histogram-based segmentation and leaf detection for rosette plants," *European Conference on Computer Vision (ECCV)*, pp. 61 - 74, 2014.
- [17] Bert De Brabandere, Davy Neven, and Luc Van Gool, "Semantic instance segmentation with a discriminative loss function," arXiv preprint arXiv:1708.02551, 2017.
- [18] Mengye Ren and Richard S Zemel, "End-to-end instance segmentation with recurrent attention," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21 - 26, Honolulu, HI, USA, 2017.
- [19] Daniel Ward, Peyman Moghadam, and Nicolas Hudson. "Deep leaf segmentation using synthetic data," *CVPPP 2018*, Newcastle, UK, Sept. 2018.
- [20] 김서정, 이재수, 김형석, "딥러닝을 이용한 양파 발의 잡초 검출 연구," *스마트미디어저널*, 제7권, 제3호, 16-21쪽, 2018년 9월
- [21] 김서정, 김형석, "Multi-Tasking U-net 기반 파 프리카 병해충 진단," *스마트미디어저널*, 제9권 제1호, 16-22쪽, 2020년 03월
- [22] 이한솔, 김영관, 홍지만, "사물인식을 위한 딥러닝 모델 선정 플랫폼," *스마트미디어저널*, 제8권, 제2호, 66-73쪽, 2019년 06월

---

 저자 소개
 

---



이종혁(준회원)

2019년 전북대학교 전자공학부 학사 졸업.

<주관심분야 : 인공지능, 영상지능, 영상처리>



김형석(정회원)

1980년 한양대학교 전자공학과 학사 졸업.  
 1982년 전북대학교 전기공학과 석사 졸업.  
 1992년 University of Missouri Dept of Electrical and Computer 박사 졸업.

<주관심분야 : 인공지능, 뉴럴네트워크, 영상지능>