

# DGIST 하계 인턴십 활동보고서

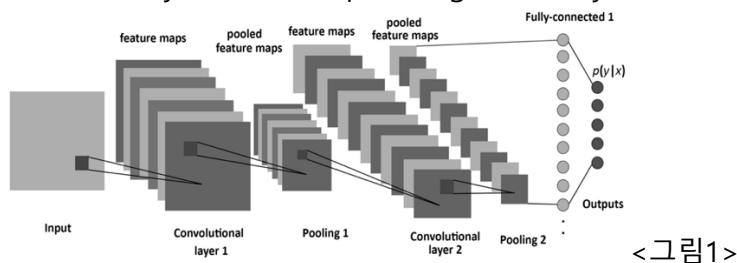
한양대학교 김재익

1. 주제 : semantic segmentation을 통한 위성사진내 도로 Detection

## 2. 배경지식

### (1) Convolution Neural Network(CNN)

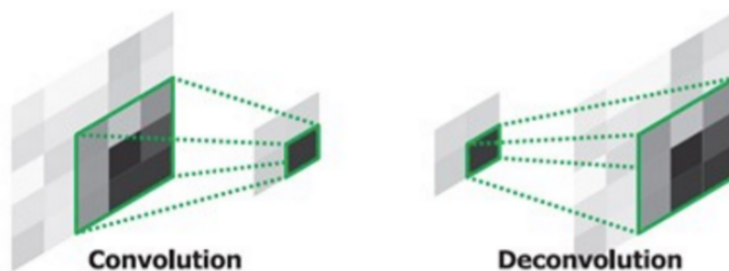
- CNN은 그림1과 같이 Convolution layer와 pooling layer의 반복과 fully-connected layer를 통해 input image를 classify하는 neural-network이다.



- Convolution layer에서는 특정 filter를 input image에 중첩시키면서 순회시켜 특정 픽셀의 특징을 뽑아내는 층으로 해석가능하다.

### (2) Deconvolution

- 위에서 언급된 Convolution개념이 특정 픽셀의 특징을 뽑아 내는것이라면, Deconvolution은 추출된 특징들을 이용해서 input의 각 픽셀값을 복원 시키는 과정이다.



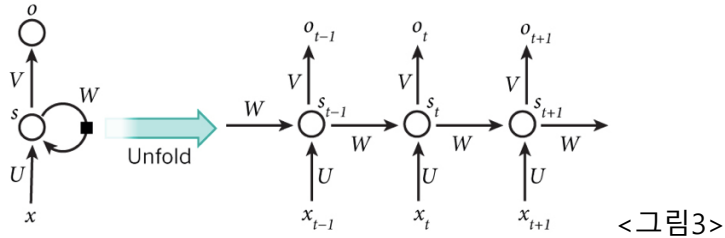
<그림2>

### (3) ResNet

- 마이크로소프트팀이 개발한 네트워크로, 스킵연결을 도입하여 깊은 층을 구성할때 생기는 신호감쇠(Vanishing Gradient)현상을 막아준다.

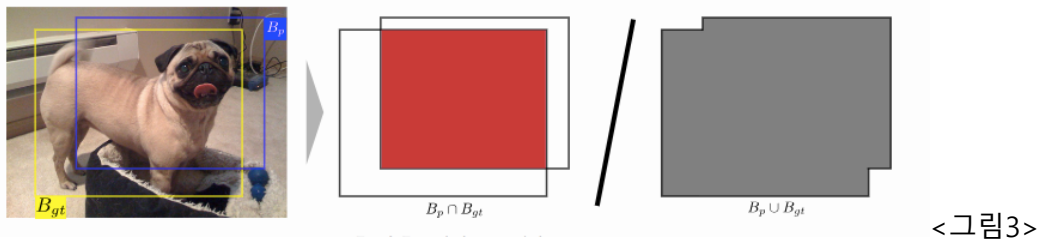
#### (4) RNN(Recursive Neural Network)

- RNN은 아래 그림3과 같은 구조로, 같은 양의 memory를 사용하는데 반해 더 깊이가 있는 신경망을 구성할 수 있다는 장점이 있다.



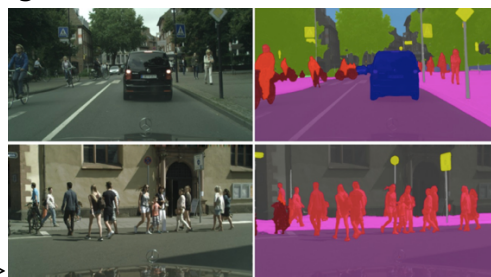
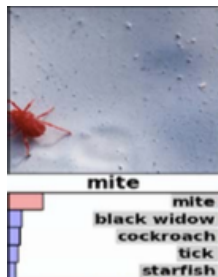
#### (5) IoU

- IoU값은 image segmentation결과의 성능을 평가하는 척도로, 실제 image와 segmentation의 교집합을 합집합으로 나눈 값이다.



### 3. 개요

- 최근 컴퓨터 비전분야에서는 image classification, detection, localization, semantic segmentation등을 수행하는 많은 deep neural network들이 개발되고 있다. 그림1과 같이 특정 image를 보고, 그 image가 어떤 class에 속하는지 classification해주는 VGG net, 그림 2와 같이 image의 각 픽셀마다 labeling을 해주어서, image의 각 부분 부분이 어떤 class에 속하는지 classification해주는 FCN, segnet, U-net등이 대표적이다.

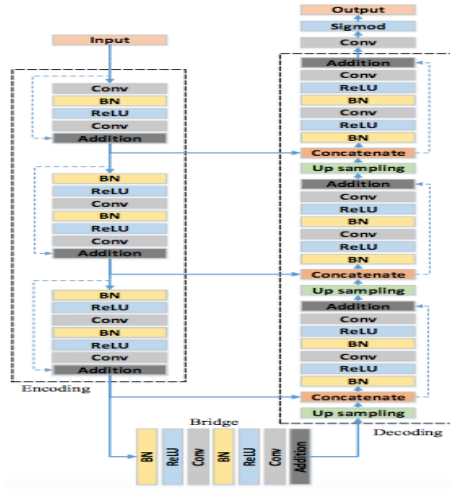


- 이 report에서는 위에서 언급된 U-Net에 마이크로소프트 팀으로 부터 개발된 ResNet을 접목시켜 깊은 neural network에서 backpropagation시에 update weight의 값이 아주 작아지는점을 보완한 ResU-Net을 구현할 것이다. 또한 ResU-Net에 RNN을 접목시킨 R2U-Net을 구현하여 기존의 neural network 모델들과 ResU-Net, R2U-Net을 이용한 결과 값을 mIoU값과 행렬출력을 통해 비교할 것이다.

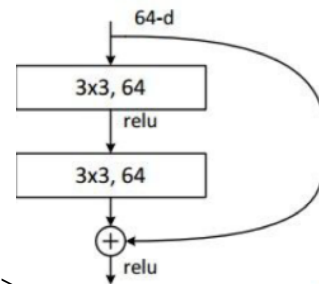
## 4. 구성 및 결과

### (1) 구성 및 코드샘플

#### ① ResU-Net



<그림5>



<그림6>

```
size_1 = x.size()
enc1 = self.encoder_1(x)
temp = self.upsample1(x)
enc1+=temp
x,indices1 = self.maxpool(enc1)
```

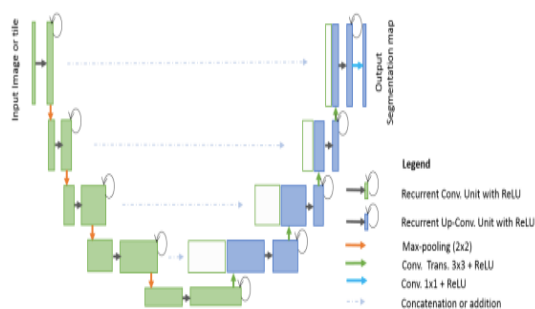
<그림7>

```
temp = torch.cat([enc4, self.up1(center)], 1)
dec1 = self.decoder_1(temp)
dec1 += self.downsample(temp)
```

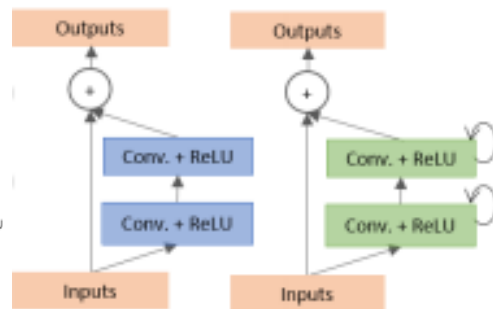
<그림8>

- ResU-Net의 코드 구성은 그림5와 같다. 그림6의 skip연결을 기존의 U-Net에 접목시켜 Vanishing Gradient Problem을 조금이나마 해결한것이 구조적인 특징이고, 각각의 encoder, decoder block의 코드구성은 그림7, 그림8과같다.

#### ② R2U-Net



<그림7>

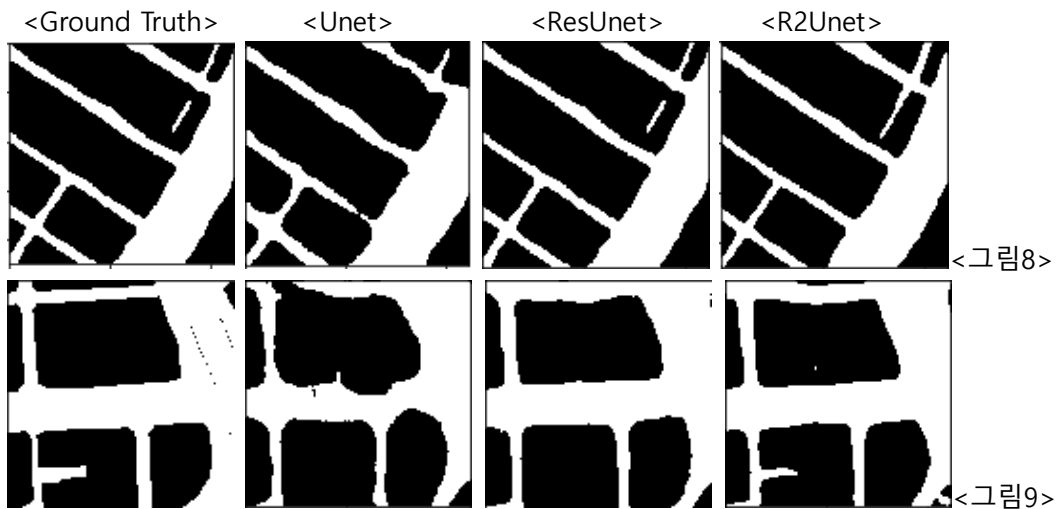


<그림8>

- R2U-Net의 코드 구성은 그림7와 같다. 그림8의 왼쪽에 있는 block처럼 Convolution layer가 반복되는것이 ResU-Net의 기본구성이었다면, R2U-Net은 RNN이라는 작은 network가 반복되어 ResU-Net과 사용되는 메모리의 양은 같지만 더욱 깊은 신경망을 구성할 수 있다는 장점을 가지고 있다.

## (2) 기존의 Neural Network와의 비교

### ① 시각적 비교



- 그림 8과 9는 label(Ground Truth)와 Neural Network들의 결과를 출력시킨 결과이다. 기존의 R2Unet, ResUnet, Unet 순으로 테두리 부분이 더 매끄러워진 모습을 보여주었고, Ground Truth에 더 가까운 모습을 보여주었다.

### ② IoU값 비교

- IoU값의 경우에는 도로의 비중이 적은 input에 대해서는 IoU값이 적게 나오는 성향을 발견할 수 있었다. 따라서 전체 image 중 도로의 비중이 1%미만인 image에 대해서는 해당 값의 계산을 해주지 않았다.
- 그결과, Unet의 경우 56.1, ResUnet은 62.45, R2Unet은 69.14의 IoU값으로, R2Unet이 가장 좋은 성능을 가짐을 확인할 수 있었다.

## 5. 결론 및 제언

- ✧ Deep Globe Challenge 수상작들의 네트워크를 공부해보고, 실제로 구현함으로써 기존에 가지고 있었던 딥러닝에 대한 개념을 정리할 수 있는 시간을 가질 수 있었다.
- ✧ 기존에 가지고 있는 지식이 그리 깊지 않았던 탓에 모델을 응용하지는 못하고 논문에서 제시한 모델을 해석하고 구현하는 정도의 연구밖에 진행하지 못했던 점이 아쉬웠다.

## 6. 출처

- Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun "Deep Residual Learning for Image Recognition"
- Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M.Taha, Vijayan K.Asari "Recurrent Residual Convolutional Neural Network based on Unet(R2U-Net) for Medical Image Segmentation"