**컴퓨터비전 텀프로젝트 [U-Net 기반 인도 보행 영상에서의 점자블록 Segmentation]**

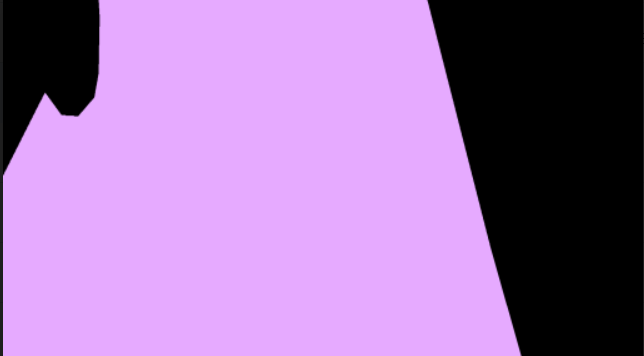
12181737 김기원 12201942 이해수

1. **개요**

시각 장애인의 보행권을 보장하기 위해 필수적인 인도의 점자블록들이 끊겨있는 경우나, 낙엽이나 쓰레기로 인해 제대로 된 기능을 하지 못하는 경우들이 존재한다. 따라서 인도를 보행하며 얻어진 데이터에서 점자블록에 집중하여, 이를 직접 segmentation 하는 과정을 통해 이상 점자블록을 판별하고 보행 가능 구역을 파악할 수 있는 모델을 얻어내는 것을 프로젝트의 목표로 한다.

1. **데이터셋 구축**

학습을 위한 데이터셋은 ai hub에서 제공되는 인도보행 영상 중 인도 노면 상태 등의 정보를 폴리곤 형태로 어노테이션하고 마스킹한 데이터인 surface masking data를 이용하였다. ai hub에서 받아온 dataset을 살펴보면 우선 jpg로 저장된 input image가 존재하며, 해당 image들에 맞게 labeling 된 png 형식의 mask image가 존재한다.

학습에 앞서 속도 향상을 위해서 데이터 전처리로 resize를 먼저 진행하였다. 기존 이미지가 1920x1080 크기로 아주 큰 고해상도의 image였으므로 512x512로 resize하여 저장하기 위하여 opencv 함수를 이용하였다.

텍스트이(가) 표시된 사진

자동 생성된 설명

//jpg로 저장된 input image resize를 진행하는 코드

텍스트이(가) 표시된 사진

자동 생성된 설명

//png로 저장된 label image resize를 진행하는 코드

dataset에는 점자블록 외에도 다양한 이동 장애물이나 노면에 대한 정보들이 저장되어 있어 이 중 점자 블록에 대한 data만 추출하여 저장하는 처리를 진행하고자 하였다.

텍스트이(가) 표시된 사진

자동 생성된 설명

//전체 dataset 내에서 ‘점자블록이 포함된’ image들을 masking하는 코드

텍스트이(가) 표시된 사진

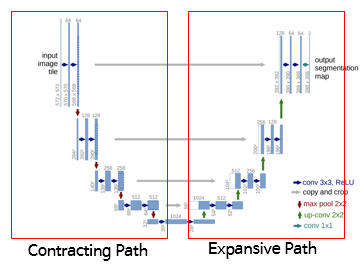
자동 생성된 설명

//masking 처리된 점자블록 data들을 하나의 폴더에 저장하는 코드

1. **모델 분석**

학습 모델로는 semantic segmentation에서 널리 사용되는 모델인 U-Net을 선정하였습니다. 코랩에서 tensorflow와 keras를 이용하여 학습하였다.

U-Net은 바이오메디컬 분야에서 image segmentation를 위해 만들어진 모델입니다. fcn (fully convolutional network) 를 기반으로 확장한 end-to-end 모델로, 구조가 U자 형태기 때문에 U-Net으로 불린다.



구조 왼쪽의 Contracting path는 convolution + ReLU + Pooling을 통해 input 이미지의 특징을 추출하는 작업을 거친다. 구조 오른쪽의 Expanding path는 contracting path의 feature map을 기반으로 upsamping을 거친 후 이를 contracting path의 feature map의 context와 결합하여 localization을 진행한다.

Contracting path에서는 3x3 컨볼루션을 2번, max pooling을 진행한다. downsampling 시 feature 크기는 절반이 되고 channel의 수는 2배가 된다. 이때 컨볼루션에 ReLU 함수를 이용한다.

Expansive path에서는 2x2 up 컨볼루션으로 feature map 크기를 2배로 늘리고, up-sampling을 통해 채널수를 다시 절반으로 줄인다. 각 expanding step에서는 skip connection의 방식으로 concatenation을 진행한다. unet은 convolution 과정에서 padding을 하지 않으므로 최종 output segmentation map 크기가 작아지는 것이 특징이다.

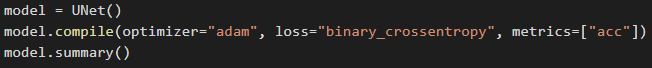
model에서 Contracting path, 즉 down\_sampling을 4번 진행하고, Expansive step으로 up\_sampling을 4번 진행한다. fcn과 다르게 unet에서는 channel 방향으로 concatenation하여 channel을 늘린다. end-to-end 구조를 이용하는 unet은 더욱 빠르고 우수한 성능을 보이므로 해당 모델을 택하게 되었다.

**4. 결과**

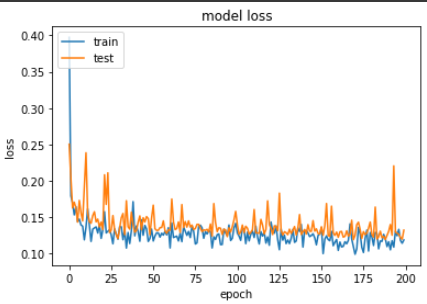
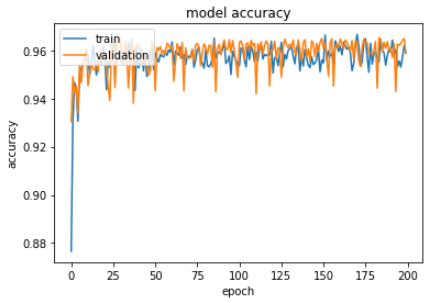
1) BaseLine

텍스트이(가) 표시된 사진

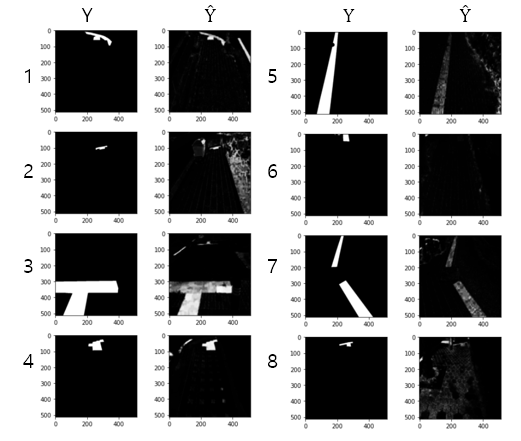
자동 생성된 설명



첫번째 모델은 Base 라인으로 hyper parameter를 다음과 같이 설정하였다.



accuracy, loss커브를 보면 전반적으로 학습이 잘 이루어 지고 있는 모습을 확인 할 수 있다. 학습 후 test set으로 evaluate 해본 결과 0.958의 결과가 나왔다.

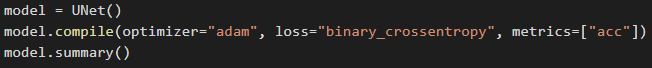


test set의 sample을 확인 해보면 8개의 샘플 중 5개가 segmentaton에 성공한 것을 확인 할 수 있다.

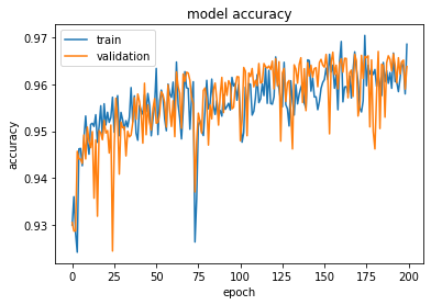
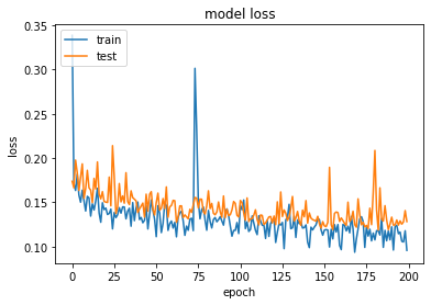
2) BaseLine + Suffle

텍스트이(가) 표시된 사진

자동 생성된 설명

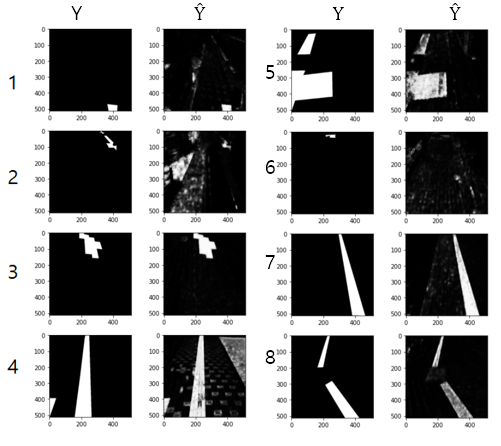






다음은 baseline에 데이터 셔플을 추가한 모델이다. 원래는 mini batch를 활용한 학습에서는 매 에폭마다 data를 셔플하는것이 원칙이지만 이번 baseline을 학습시키고 분석하는 도중 셔플이 진행이 안 되었기에 추가하여 테스트 해보았다. 다른 파라미터들은 baseline에서와 동일하다.

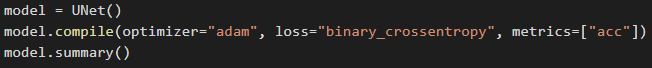


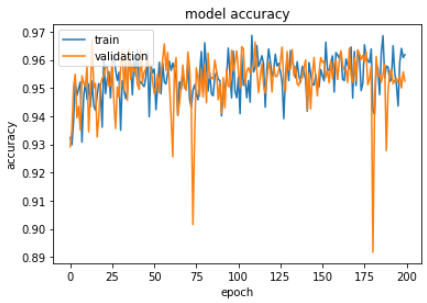
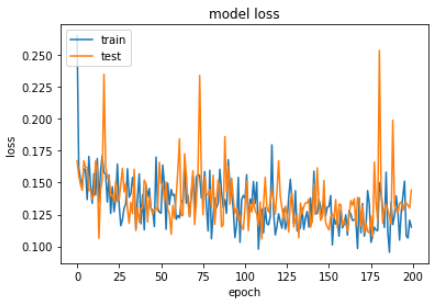
sample image를 확인해 본 결과 8개의 샘플 중 7개를 segmentation 하는 데 성공하였다.

3) Large Batch + Shuffle

텍스트, 시계이(가) 표시된 사진

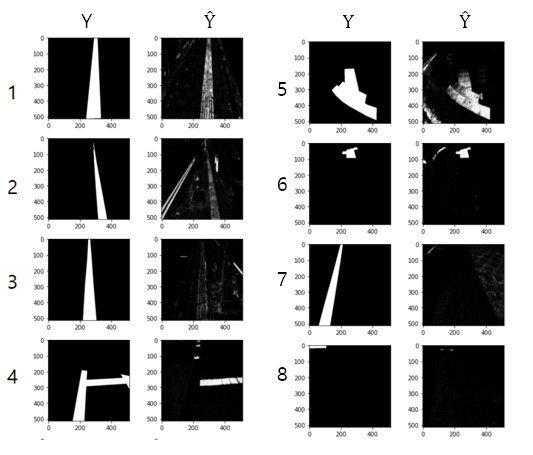
자동 생성된 설명





마지막은 셔플을 추가한 베이스 라인에서 배치사이즈를 2배 키운 모델이다. 배치 사이즈가 커진 만큼 그래프에서 튀는 부분이 조금 더 많아진 것을 확인할 수 있었다. emulation 결과로 0.953의 결과가 나왔다.



sample 이미지를 확인해 본 결과8개의 샘플 중 4개를 segmentation 하는 데 성공하였다.

1. **고찰 및 활용 방안**

테이블이(가) 표시된 사진

자동 생성된 설명

위 세 모델의 결과를 정리한 표를 보면 Baseline에 Shuffle을 추가한 모델, Baseline 모델, Batch size를 늘린 모델 순서로 Test Accuracy가 측정되었다. 당연히 기존 모델보다 셔플을 추가한 모델이 더 좋은 결과가 나왔고 의외로 배치사이즈를 늘린 모델이 가장 안좋은 결과가 나왔다.

bag of tricks 논문에서 소개된 Linear scale ReLU 기법에 따라 Batch Size를 늘렸을 때 Learning rate도 늘린다면 더 좋은 결과가 나올것으로 예상된다.

원래 Image Augmentation을 적용하여 결과를 비교해보려 했지만 Custom Data set에대한 이해도가 부족해 구현하지는 못했습니다만 Augmentation을 구현하여 적용한다면 가장 좋은 결과가 나올 것으로 생각한다.

해당 프로젝트를 통해 점자블록 인식 정확도를 높여 시각장애인을 위한 어플이나 기타 기기에 활용할 수 있을 것이다. 나아가 파손 정도를 확인하여 이를 개선할 수 있는 서비스를 개발하는 데에도 도움이 될 수 있으리라 기대할 수 있을 것이다.

1. **참고문헌 & 코드**

[1] UNet source code   
https://github.com/nikhilroxtomar/UNet-Segmentation-in-Keras-TensorFlow/blob/master/unet-segmentation.ipynb

[2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, “Convolutional Networks for Biomedical Image Segmentation“