-7/11 월

확랜 강의 04뒷부분과 05 절반 들음

npy파일 data를 눈으로 확인하는과정에 어려움이 이었음

npy파일의 경우 np.load(data\_path)를 통해 nii파일의 경우 nibabel.load(data\_path).get\_fdata()

주어진 LITS data를 2d로 바꿔 png로 저장

-7/12 화

확랜 강의 05마저 듣고, 관련 강의내용(conditional probability 유튜브강의)시청

PPT에 나와있는 문제들 풀이

Slice한 data를 이용해 학습 시키려했으나 data양이 많아서인지 학습속도가 매우 느림

스터디-5강까지 내용 발표(보경님),현재 task 진행상황 보고

npy파일을 png로 slice할경우 data loss가 발생하므로 안됨….

🡪해결책: array형식으로 slice마다 나누어 list형식으로 저장

-7/13 수

Png가 아닌 slice별로 list형식으로 저장 (총 22207slice)

이전에 사용한 customimagedataset과 customconvnet 코드를 가져와 흑백사진에 맞게 수정하여 학습(각 slice를 list로 저장하였기 때문에 conv2d를 통과할 때 3dimension이 입력으로 들어가게 되어 error발생🡪channel dimension추가)

학습 시켰으나 zig-zagging이 심하고 loss값이 줄어드는게 보이지 않음🡪모델수정필요

gpu out of memory🡪 batch size줄이기

Gpu memory 초기화 kill -9 PID

Accuracy가 변하지 않음-68….%

-7/14 목

network수정하여 다시 학습시켰으나 loss값이 여전히 줄지 않음

학습된 model로 train과 test loader를 넣어 predict를 출력해보니🡪 0값만 찍힘

label매칭에 문제가 있는지 확인하기위해 train\_test\_split과 dataloader를 확인했으나 이상없음

normalization이필요🡪값이 0-255가 아닌 -3000~2000사이의 픽셀값을 가지고 있기 때문에 0-255로 바꿔줘야함 그러나 min max적용시 간부위는 매우 작은 퍼센테이지를 차지하게되어 간부분을 잘볼수있게 min max를 지정값으로 두어 해당값보다 큰값들은 모두200으로 작은값은 0으로 한후 min max진행

-7/15 금

확률랜덤변수 강의 06, 07 들음

Normalization 후에 학습진행🡪loss값 zigzagging이 심하지만 줄어드는 것은 보임

매번 slice를 list에 저장하는 것이 번거로워 segmentation\_list를 얻어 label\_list를 구한뒤 txt로 저장하여 불러와서 사용

Volume\_list도 txt로 저장하고 싶었으나 용량이 너무 커서 불가

Epoch 50번 결과 Accuracy 96% 50번의 에폭 중 loss가 가자 작았던 모델을 model에 저장

-7/18 월

확률 랜덤변수강의 09들음

저장한 model을 불러와 100번 epoch반복🡪학습후 오히려 accuracy가 낮아짐

다른 network를 사용하여 학습시도

Epoch를 많이하거나, layer를 많이 쌓는다고 최적의 결과를 얻는 것이 아님 적당한 값을 찾아야함

epoch마다 학습속도가 다름…?-->60s~1200s

loss가 작은 모델을 따로 저장

-7/19 화

확률 랜덤변수강의 10들음

Model을 평가하는 지표에 대해서 공부하고 최적화한 모델을 사용해 각각에 대한 값을 구함

암을 분류하는 model이기 때문에 recall이 중요할 것으로 보임(병변이 있는 환자중 True로 예측한 비율)

각 epoch마다의 train과 validation loss의 추이와 수렴 정도를 보기위해 빈 list를 만들어 append하면서 다시 학습( epoch100)

미팅-accuracy가 너무 높게 나옴🡪하나의 volume(즉 한명의 환자)에 대한 slice가 test와 train에 나누어 들어가져 train에서 학습된 내용이 test dataset에 대해서도 유사하여 높은 accuracy

🡪split을 random하게 하는 것이 아닌 volume단위로 나눠서 test와 train dataset구성

중간평가회 대비 ppt로 정리해놓을 것!

-7/20 수

확랜 강의 11강 시청

volume으로 test, train 나눠서 학습

accuracy가 76%로 감소

valiation loss가 크게 감소하지않음🡪이유….?

보통은 model의 복잡도가 높을 때 train loss는 감소하지만 validation loss가 잘 감소하지 않음

Layer층 수를 줄여서 학습시도

Liver\_re3.ipynb에서 layer를 한층 줄이고 learning rate=0.00001로 바꿔서 학습

🡪loss값이 train과 test모두 잘 안줄어들고 zig-zagging이 심함

Liver\_re에서 layer한층 줄이고 learning rate-0.0001로 학습

-7/21 목

확랜 강의 12강 시청

ppt만들기

Pretrained model로학습

Resnet 18, Resnet 50

Layer의 의미???convolution이 포함되면 layer…?

Loss가 커도 accuracy가 높을 수 있나

-7/22 금

ppt만들기

variance와 bias에 대한 의미

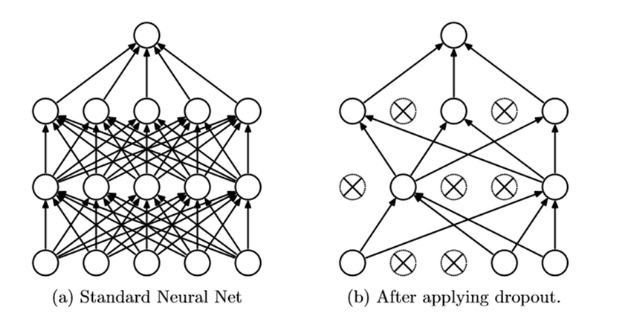
cnn과 fully connected layer의 차이

다른 모델 학습

-7/23 월

확랜 강의 13강

Dropout



효과

overfitting방지, 성능향상, 매번 다른 형태의 노드로 학습하기 때문에 앙상블 효과를 낼수있다.

모델 결합을 하게되면 학습의 성능을 개선할 수 있다

모델결합이 효과를 얻기 위해서는 서로 다른 학습 데이터를 이용해서 학습을 하거나 모델이 서로다른 구조를 가져야한다. 하지만 망이 깊은경우 복수개의 망을 학습시키는 것은 매우 힘든 작업이며 긴 연산시간을 잡아먹기 때문에 이를 위한 해결방법으로 dropout을 사용한다

여러 개의 모델을 만드는 대신에 모델 결합에 의한 투표효과와 비슷한 효과를 위해 사이클이 진행되는 동안 무작위로 일부뉴런을 생략하는데 이때 생략되는 뉴런의 조합만큼 지수함수적으로 다양한 모델을 학습하게 되어 모델 결합의 효과를 누릴 수 있다.

-7/24 화

확랜 강의 14,15강

박사님이랑 커피

PPT 금욜 4시 중간 발표

Dropout 위치에 따른 accuracy와 recall값이 달라짐을 확인

ppt만들기

-7/25 수

확랜 강의 16강

확랜강의정리

Dropout을 minmax값중 recall값이 가장 작았던 범위로 바꿔서 재학습

Vgg pretrained model을 불러와 학습

-7/26 목

확랜 강의 17강

ppt마무리

그리고 slice그림을 넣어서

네트워크가 0 (병변 없음) 이라고 예측했는데 실제로는 1 (병변 있음) 케이스인 경우

네트워크가 0 (병변 없음) 이라고 예측했고 실제로도 0인 경우

네트워크가 1 (병변 있음) 이라고 예측했는데 실제로는 0 (병변 없음) 케이스인 경우

네트워크가 1 (병변 있음)이라고 예측했고 실제로도 1 인경우