- 分別 LAB 3차 정기 미팅 -
인공지능 기법을 활용한 알츠하이머 Amyloid-PET 분류 정확도 개선

한국인공지능연구소 분별 **LAB**
김웅곤 연구원

Who am I?

現) 한국인공지능연구소 분별LAB 연구원(2018.6~)
現) 통계청 직원(2016.7~)

서강대학교 수학, 경영학 학사(2012~2017)

고교 수학강사, 대학생 전공수학, 통계 강의 경험 **多**

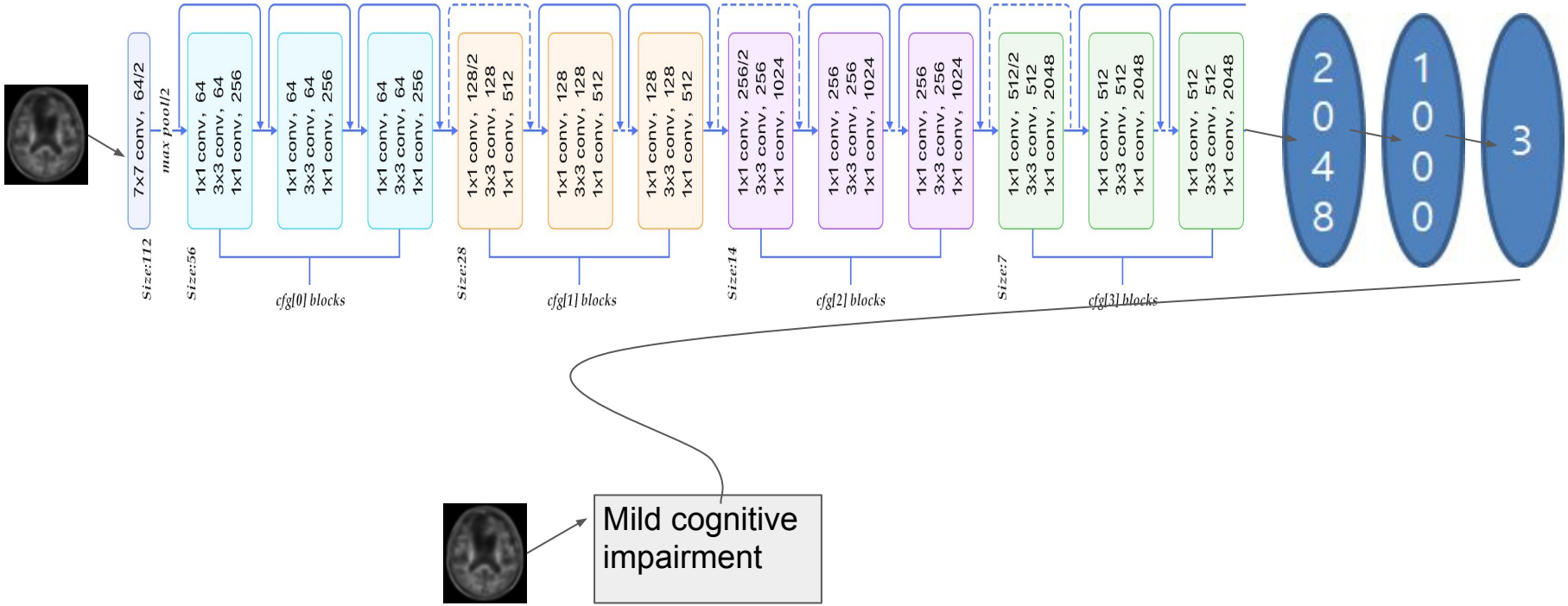
API기반 크롤러 개발 배포 **&** RL 프로젝트 수행 경험 **多**

목차

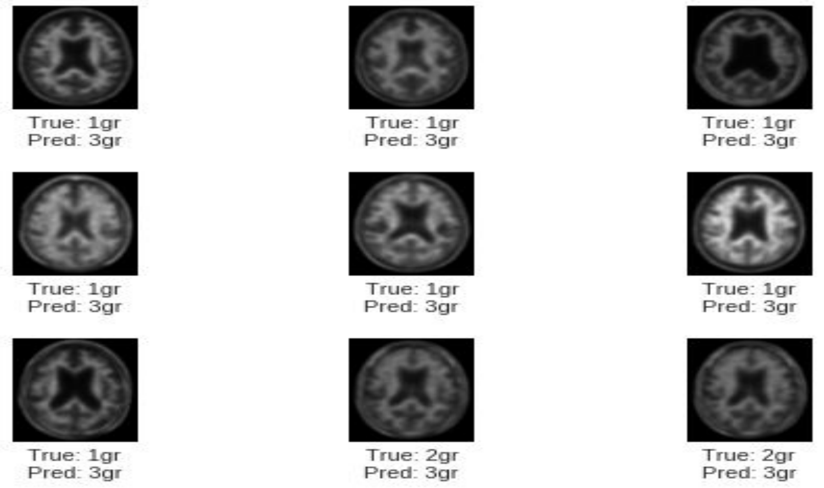
분별 LAB 중간발표

링크 :
https://docs.google.com/presentation/d/1iUJsk2KJpZ5yU6l1wnP-LfffITkVJ6kHtnkRvo32xDY/edit#slide=id.g3d222f8f52_1_14

분별 LAB 중간발표 - ResNet 50 Model



분별 LAB 중간발표 - The classification report




```
Confusion matrix:
[[66  0  7]
 [ 8  0 12]
 [ 1  2 56]]
             precision    recall  f1-score   support

          0       0.88      0.90      0.89        73
          1       0.00      0.00      0.00        20
          2       0.75      0.95      0.84        59

avg / total       0.71      0.80      0.75       152

(0) 1gr
(1) 2gr
(2) 3gr
```

F1 score 75 %, Recall 80 %
If Possible, We need an improvement!

개선 IDEA

복잡한 모델인 ResNet50 대신에

VGG16같은 쉬운 모델로 실험해 보자

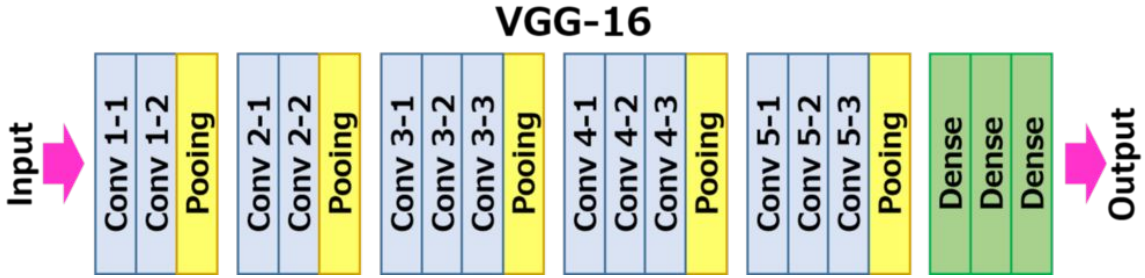
무작정 돌려보니 잘 되네????

시간 절약 + 남는 시간에 다른 곳에 투자(4 Epoch만에 테스트셋 애큐러시 89.58%)

```
Epoch 1/4
40/40 [==============================] - 10s 261ms/step - loss: 2.9676 - acc: 0.6748 - val_loss: 1.8946 - val_acc: 0.7969
Epoch 2/4
40/40 [==============================] - 9s 219ms/step - loss: 2.0105 - acc: 0.8649 - val_loss: 1.7177 - val_acc: 0.8802
Epoch 3/4
33/40 [=====================>......] - ETA: 1s - loss: 2.0094 - acc: 0.875840/40 [==============================] - 9s 219m
Epoch 4/4
40/40 [==============================] - 9s 223ms/step - loss: 2.0601 - acc: 0.8725 - val_loss: 1.7145 - val_acc: 0.8958

Accuracy: 89.58%
```

개선 IDEA - VGG16 Model

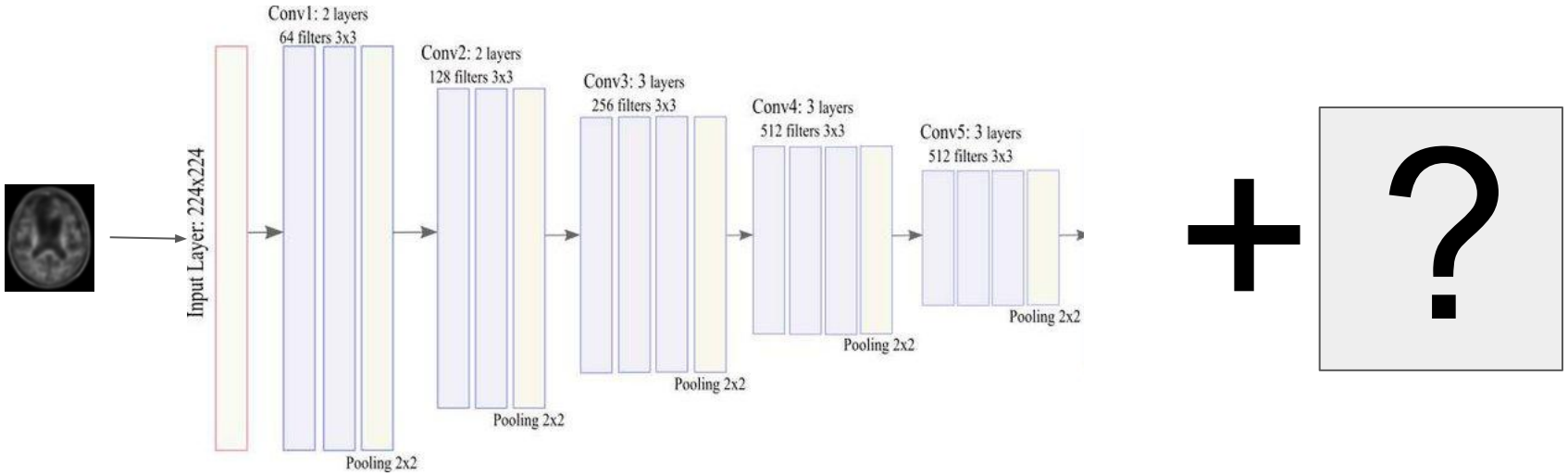
VGG16 모델로 실험!



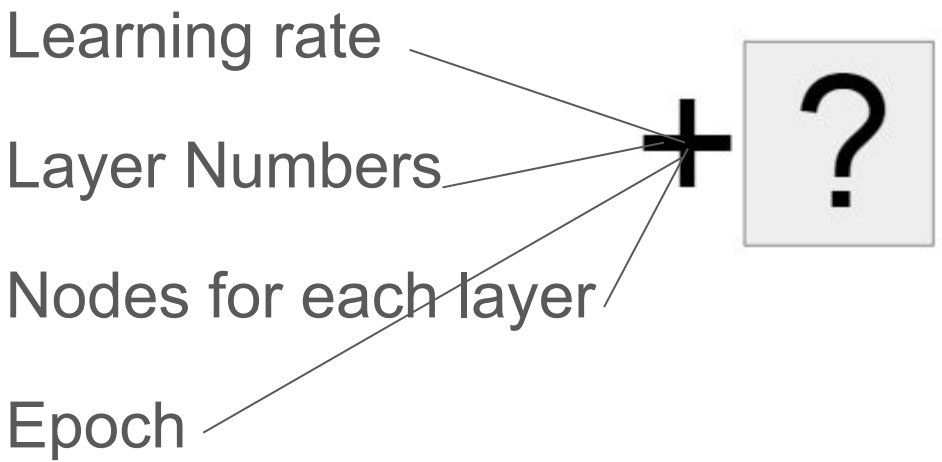
VGG16 모델의 수리적 포인트는 단순한 패치인 3 x 3 CONVOLUTION MAP을 일렬로 이어붙여서

COMPUTATIONAL COST ↓ NONLINEARITY ↑ 를 노려봄

개선 IDEA - VGG16 Model Transfer Learning



개선 IDEA - Hyperparameter Setting



어떻게
정하지??

개선 IDEA - Hyperparameter Setting

1. 아무거나 넣어봄

이것저것 넣어 보니 learning rate은 1e-5, Layer 수는 2개로.. Node 수는 1024에 dropout을 0.5만큼 넣어 주고 적절히 overffting 되지 않게 epoch을 3정도로 주니 정확도가 80%정도 나오네?? 이걸로 해야지…

개선 IDEA - Hyperparameter Setting

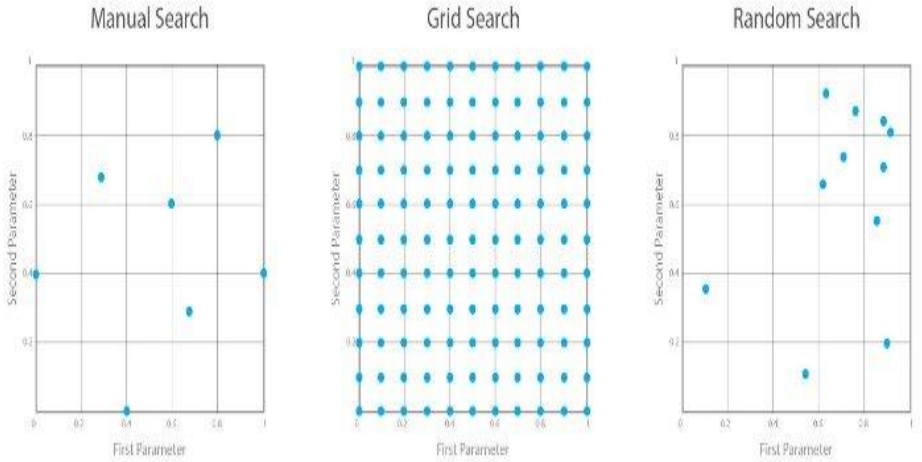
2. Grid Search

Try all Candidates

Learning rate = [1e-7 ~ 1e-2]
Num of dense layers = [1 ~ 3]
Num of nodes for each layer = [5 ~ 1024]
Num of epoch = [1 ~ 20]

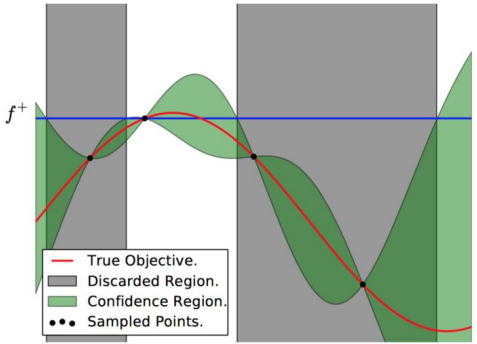
한 모델 학습시키는 데에만 10시간이 걸리는데

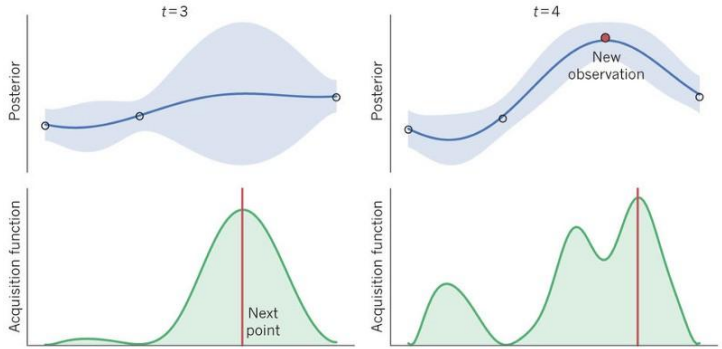
언제 다 학습시키지??...



개선 IDEA - Bayesian Optimization

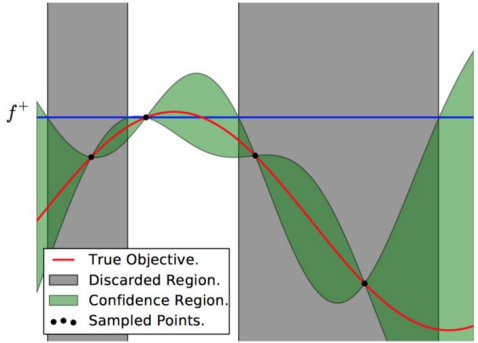
3. Bayesian Optimization

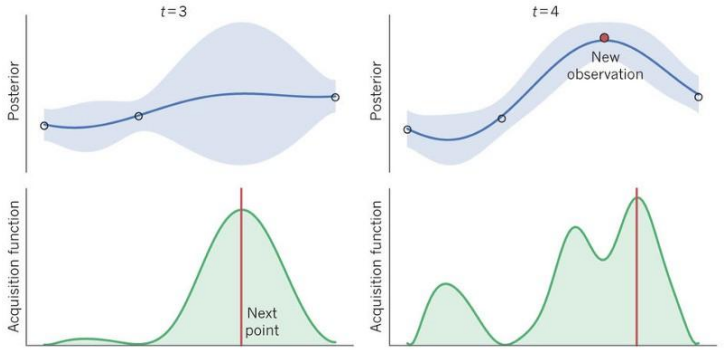




Accuracy Function = F(learning_rate, num layers, num nodes, num epochs)

개선 IDEA - Bayesian Optimization





Accuracy Function = F(learning_rate, num layers, num nodes, num epochs)

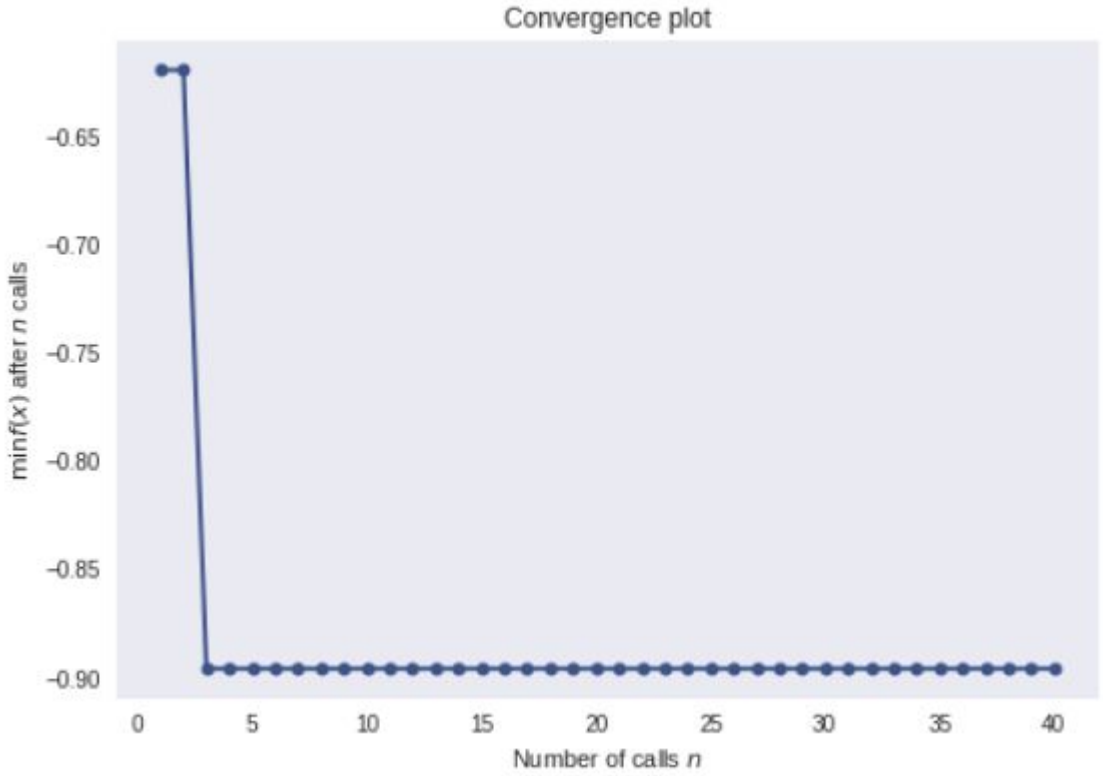
사전 결과값을 바탕으로 사후 확률인 Accuracy를 빠르게 찾아냄

개선 IDEA - Bayesian Optimization

다섯 번의 모델
학습만으로도

하이퍼파라메터들을

추정할 수 있음!



개선 IDEA - Bayesian Optimization

```
learning rate: 8.8e-04
num_dense_layers: 1
num_dense_nodes: 268
num_epoch: 4

Epoch 1/4
40/40 [==============================] - 10s 261ms/step - loss: 2.9676 - acc: 0.6748 - val_loss: 1.8946 - val_acc: 0.7969
Epoch 2/4
40/40 [==============================] - 9s 219ms/step - loss: 2.0105 - acc: 0.8649 - val_loss: 1.7177 - val_acc: 0.8802
Epoch 3/4
33/40 [=====================>......] - ETA: 1s - loss: 2.0094 - acc: 0.875840/40 [==============================] - 9s 219ms/step - loss:
Epoch 4/4
40/40 [==============================] - 9s 223ms/step - loss: 2.0601 - acc: 0.8725 - val_loss: 1.7145 - val_acc: 0.8958

Accuracy: 89.58%
```

개선 IDEA - Bayesian Optimization

Bayesian Optimization 결과

Learning rate = 8.8e-04

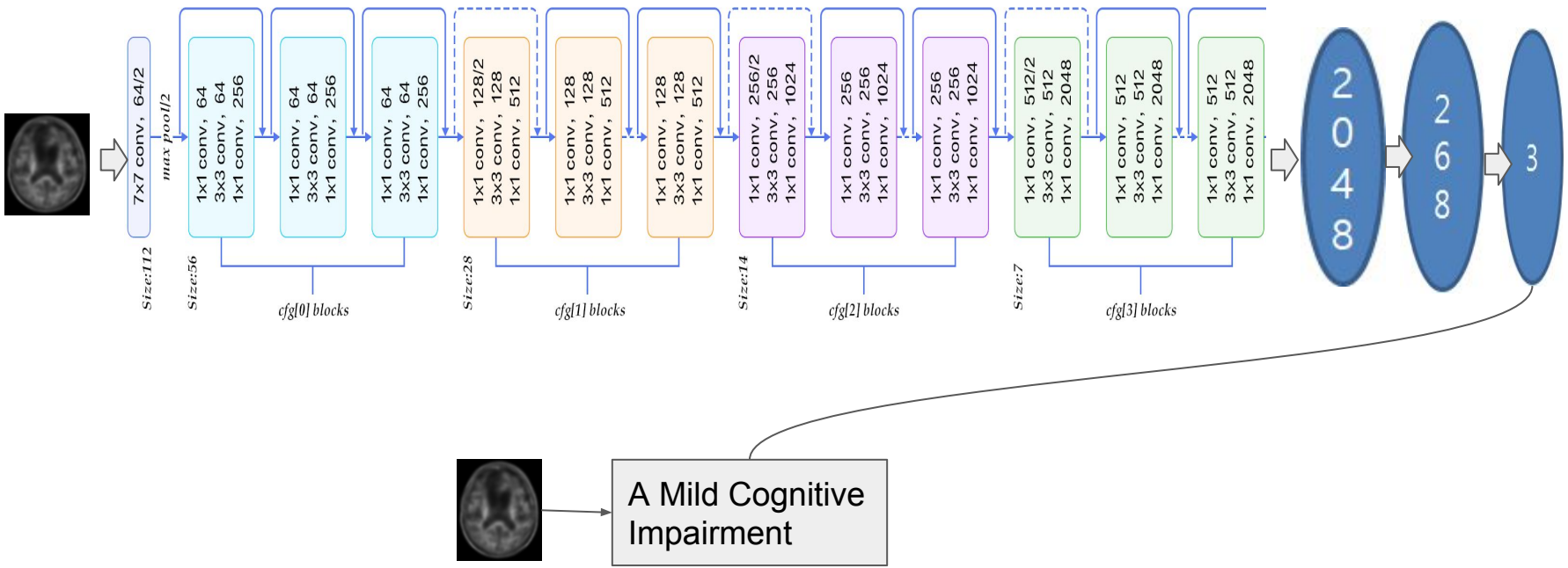
Number of dense_layer = 1

num_dense_nodes = 268

num_epoch = 4 에서

Accuracy : 89.58%

개선 결과


7x7 conv, 64/2
max pool/2
1x1 conv, 64
3x3 conv, 64
1x1 conv, 256
1x1 conv, 128/2
3x3 conv, 128
1x1 conv, 512
1x1 conv, 256/2
3x3 conv, 256
1x1 conv, 1024
1x1 conv, 512/2
3x3 conv, 512
1x1 conv, 2048
Size:112
Size:56
Size:28
Size:14
Size:7
cfg[0] blocks
cfg[1] blocks
cfg[2] blocks
cfg[3] blocks
2048
268
3
A Mild Cognitive Impairment

개선 결과

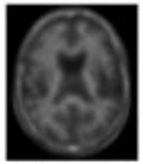
Confusion matrix:

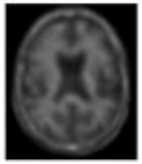
```
[[73  0  0]
 [16  0  4]
 [ 0  0 59]]
```

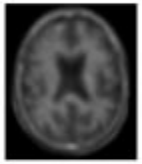
	precision	recall	f1-score	support
0	0.82	1.00	0.90	73
1	0.00	0.00	0.00	20
2	0.94	1.00	0.97	59
avg / total	0.76	0.87	0.81	152

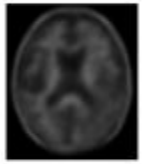

True: 2gr
Pred: 1gr

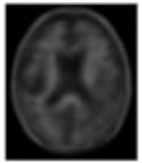

True: 2gr
Pred: 1gr

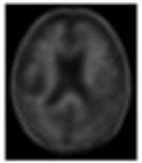

True: 2gr
Pred: 1gr

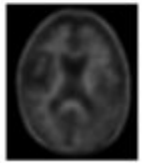

True: 2gr
Pred: 1gr


True: 2gr
Pred: 3gr


True: 2gr
Pred: 1gr


True: 2gr
Pred: 1gr


True: 2gr
Pred: 1gr


True: 2gr
Pred: 1gr

개선 결과

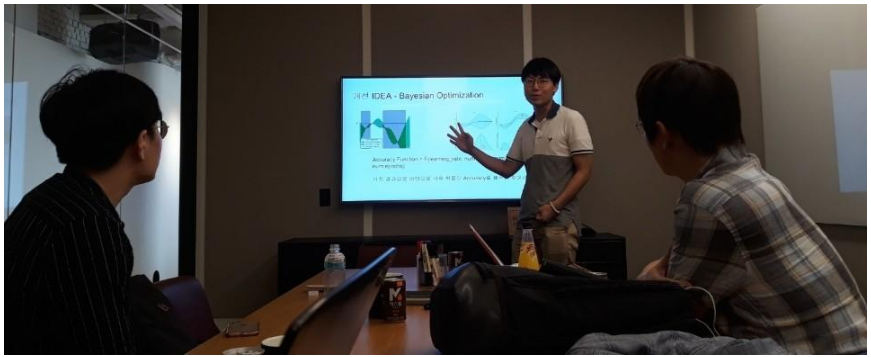
```
Confusion matrix:
[[73  0  0]
 [16  0  4]
 [ 0  0 59]]
             precision    recall  f1-score   support

          0       0.82      1.00      0.90        73
          1       0.00      0.00      0.00        20
          2       0.94      1.00      0.97        59

avg / total       0.76      0.87      0.81       152
```

0단계와 2단계는 우리가 만든 **‘알츠넷’(명칭)**가 잘 분류해 냄

그런데 1단계는 이 녀석이 정도가 Mild 하면 0단계로

정도가 약간 Severe하면 2단계로 분류해 버림. -> 사실 사람들도 비슷하지 않나?

1단계 샘플에 대한 심층 분석이 필요함



- 감 사 합 니 다 -