2019150445 신백록

6.1절의 CNN 모형에서는 train set의 acc이 0.95에 가까운 반면 val acc은 0.71정도에 머물러있는 걸로 보아 overfitting이 발생한 걸로 생각할 수 있고, 실제 미리 분리해 둔 test set과 evaluate한 결과 0.72정도의 정확도를 보였다. 데이터는 1700여개인 데 비해 모수가 130만개를 넘어서 overfitting이 발생한 걸로 볼 수 있다.

Epoch 8/10

40/40 [==============================] - 0s 6ms/step - loss: 0.1188 - acc: 0.9592 - val\_loss: 0.9851 - val\_acc: 0.6787

Epoch 9/10

40/40 [==============================] - 0s 6ms/step - loss: 0.3281 - acc: 0.8543 - val\_loss: 0.7664 - val\_acc: 0.7104

Epoch 10/10

40/40 [==============================] - 0s 6ms/step - loss: 0.1540 - acc: 0.9512 - val\_loss: 0.9738 - val\_acc: 0.7059

9/9 [==============================] - 0s 6ms/step - loss: 0.8361 - acc: 0.7269

[0.8360813856124878, 0.7269230484962463]

InceptionV3 이전학습을 한 모형에서는 학습 결과 callback에 의해 6번째 epoch에서 학습을 멈추었고, 이를 test set과 evaluate한 결과 accuracy는 0.88정도가 나왔다. Overfitting이 있었지만 이전학습을 한 결과로 다른 모형보다 test set의 accuracy가 가장 높게 나타난 것을 알 수 있다.

Epoch 4/20

40/40 [==============================] - 1s 18ms/step - loss: 0.0078 - acc: 0.9986 - val\_loss: 0.3540 - val\_acc: 0.9321

Epoch 5/20

40/40 [==============================] - 1s 18ms/step - loss: 0.0044 - acc: 0.9993 - val\_loss: 0.3150 - val\_acc: 0.9231

Epoch 6/20

40/40 [==============================] - 1s 18ms/step - loss: 0.0140 - acc: 0.9950 - val\_loss: 0.5685 - val\_acc: 0.8597

9/9 [==============================] - 0s 8ms/step - loss: 0.7756 - acc: 0.8769

[0.7755874991416931, 0.8769230842590332]

병렬형 CNN을 적용한 모형에서는 callback에 의해 10번째 epoch에서 학습을 멈추었고, train set의 accuracy는 거의 1에 다다른 반면 validation set의 accuracy는 0.7정도 나온다. 이 모형을 test set에 적용해 본 결과, validation acc과 비슷한 0.72정도 나왔다. 일단 데이터가 모수에 비해 현저하게 적고, Input shape이 300\*300으로 이를 CNN 모형에서 flatten시킴으로써 Dense 출력층의 모수가 50여만개로 병목현상이 일어나 overfitting이 발생하였다고 볼 수 있다. 6.1절의 CNN 모형과 비슷한 정도의 성능을 보인 것을 알 수 있다.

Epoch 9/15

40/40 [==============================] - 12s 301ms/step - loss: 0.0231 - acc: 0.9961 - val\_loss: 1.0394 - val\_acc: 0.7104

Epoch 10/15

40/40 [==============================] - 12s 302ms/step - loss: 0.0402 - acc: 0.9887 - val\_loss: 0.9342 - val\_acc: 0.6923

9/9 [==============================] - 1s 82ms/step - loss: 0.9463 - acc: 0.7192

[0.9463144540786743, 0.7192307710647583]

그래서 병렬형 CNN모형의 overfitting을 해결하기 위해 input shape을 50\*50으로 줄였고, 출력층 직전에 dropout을 0.7을 주어 병목현상과 overfitting을 해결하고자 하였다.

dense\_43 (Dense) (None, 126) 20160126 flatten\_28[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_43 (Dropout) (None, 126) 0 dense\_43[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_44 (Dense) (None, 1) 127 dropout\_43[0][0]

================================================================================

학습을 하고 test set과 evaluate한 결과 약간의 성능 향상이 있었지만 inception V3 이전학습을 한 거에 비해 한참 못 미쳤다. 이 또한 데이터가 1700 여개 밖에 없어서 overfitting이 필연적으로 발생한 것으로 보인다.

9/9 [==============================] - 0s 6ms/step - loss: 1.0799 - acc: 0.7615

[1.0798544883728027, 0.7615384459495544]

마지막으로 혹시나 flatten을 하는 대신, global average pooling을 써서 출력층의 모수 숫자를 획기적으로 줄이면 어떨까 하고 시도해보았다. 전체 모수가 10만여개로 추정해야할 모수가 확실히 줄었고, 출력층의 모수 숫자도 257개로 전보다 병목현상이 덜 발생할 것으로 보았다. 이 모형을 test set과 evaluate한 결과 accuracy가 0.72로 Inception V3를 제외한 모형들과 비슷한 정도의 성능을 보이는 것을 알 수 있다. 여러 모형을 시도해보았지만 결국 이전학습을 하지 않고서는 overfitting을 해결할 수 없었다.

global\_average\_pooling2d\_6 (Glo (None, 256) 0 concatenate\_32[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_21 (Dropout) (None, 256) 0 global\_average\_pooling2d\_6[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_32 (Dense) (None, 1) 257 dropout\_21[0][0]

================================================================

Total params: 118,209

Trainable params: 118,209

Non-trainable params: 0

9/9 [==============================] - 0s 5ms/step - loss: 0.6825 - acc: 0.7231

[0.6824997663497925, 0.7230769395828247]