Midterm Examination(Face Emotions Recognition)

12161235 이채영

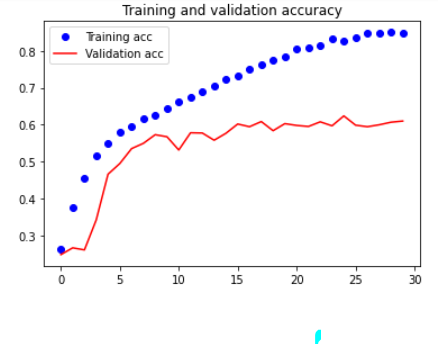
Fer2013에 주어진 유러피안 데이터 셋으로 모델을 훈련시킨 후 한국인 데이터 셋에서 모델의 정확도가 어떻게 나오는지를 확인해보았다.

**1. 주어진 유러피안 데이터로 학습시킨 모델(총 28709개의 데이터)**

-25838개의 학습 데이터와 2871개의 검증 데이터 사용

1. 초기에 매우 간단한 CNN모델을 만들었다. 그 결과 정확도는 20%에 불과하였다. 따라서 하이퍼파라미터들을 미세조정 해보았다. Layer의 개수와 배치를 바꿔보고, 손실함수와 learning rate등도 바꿔보았다. 또한 dropout layer를 추가하여 과대적합을 방지하였다. 그 결과 정확도는 무려 60%까지 향상되었다.

2. 그래프를 살펴보면 train accuracy는 계속 증가하는 반면 valid accuracy는 더 이상 증가하지 않고 있었다. 또한 train accuracy는 거의 90%에 가까워지는 반면 test data에 대한 정확도는 60%에 불과하였다. 이를 보고 train data에 너무 과대 적합 되었다고 생각했다. 따라서 더 많은 train data가 필요하다고 생각했고, 더 작은 learning rate로 과대적합을 방지해야한다고 생각하였다..





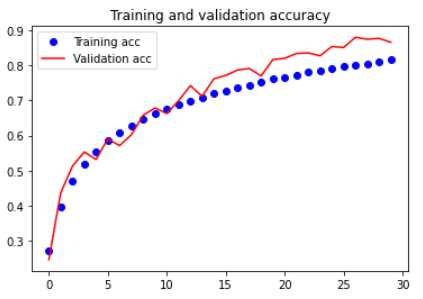
**2. 데이터 Argumentation으로 더 많은 학습 데이터로 학습시킨 모델**

**(총 12만개의 데이터)**

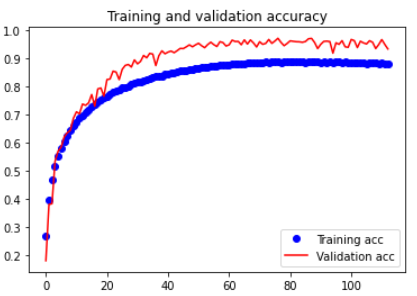
1. 우선 learning rate를 0.001에서 0.0001로 바꾸어 좀 더 미세한 weigth값 조정이 되도록 하였다. 또한 데이터셋이 부족하다 생각하여 데이터를 증가시켰다. 이미지를 좌우 반전시키는 함수를 만들어 데이터를 2배로 만들었고 ImageDataGenerator를 이용하여 데이터를 좌우로 회전, 확대, 밝기 조절 등을 하여 총 14만개의 train data로 훈련하였다.

2. 그 결과 valid data의 accuracy와 train data의 accuracy는 비슷하게 계속 증가하여 과대적합은 어느정도 방지가 되었다고 생각하였다. 하지만 test accuracy는 생각보다 너무 낮게 나왔다.

3. 학습데이터의 양이 꽤 커졌지만 정확도가 많이 올라가지 않은 원인을 생각해보았다. 처음 주어진 fer2013의 train data는 클래스별로 데이터의 불균형이 심하였다. Neutral클래스에는 약 7천개가 넘는 데이터가 있었고 Didgust클래스는 약 4백개밖에 없었다. 데이터가 한쪽으로 편향되어 있어 정확도에 좋지 않은 영향을 미쳤을 것이라 생각했다. 때문에 클래스별로 2만개씩의 데이터를 만들어 총14만개의 균형있는 데이터로 학습을 시켜보았다. 그 결과 좀더 향상된 62.58 %의 정확도를 얻을 수 있었다.



4. 하지만 위의 결과의 acc그래프와 loss그래프를 살펴보았을 때 valid데이터의 accuracy도 계속 상승하는 추세이고 loss도 계속 감소하는 추세였다. 충분한 학습이 이루어지지 않아 과소적합 될 수 있다고 생각하였다. valid accuracy의 loss가 epoch 25회동안 감소하지 않으면 학습을 조기종료를 시키는 EarlyStopping을 이용하여서 과소적합과 과대적합을 방지하였다. 미세하지만 조금 더 향상된 62.84%의 정확도를 얻을 수 있었다.





**3. 한국인 데이터로 모델 평가하기**

1. 앞서 만든 모델에 한국인 데이터 셋을 train과 test로 나누어서 test data로 정확도를 평가해보았다. 정확도는 26.22%로 매우 낮았다.



**4. 한국인 데이터로 모델 학습하기**

1. 모델을 새로 만들어 한국인 train data로 훈련시켜 정확도를 평가해본 결과 정확도는 24.59%로 매우 낮았다. 한국인 훈련 데이터가 약 600개로 매우 적었기 때문이라고 판단하였다.



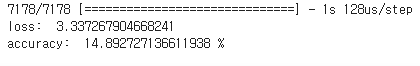
또한 유러피안 데이터로 이 모델을 평가한 결과 정확도는 23.55%로 역시 매우 낮았다.



3. 더 많은 훈련 데이터를 이용하여 훈련에 사용해보았다. 공유폴더에 있는 다른 한국인 데이터들을 이용하여 총 2824개의 데이터로 모델을 학습시켜보았다. 그 결과 정확도는 37.56%로 올랐다.



하지만 유러피안 데이터로 모델을 평가했을 때 정확도는 오히려 더 낮아졌다.



한국인의 얼굴 특징에 맞게 weigth들이 많이 조정되어 한국인과 얼굴 특징이 다른 유러피안에서 이러한 결과가 나온 것이라고 생각한다.

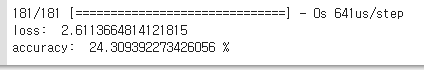
4. 한국인 데이터를 Argumentaion 한 결과 정확도는 39.34%로 늘어났다. 이렇게 많은 데이터로 학습하였는데도 정확도가 많이 향상되지 않은 이유는 잘 정제된 데이터가 아니였기 때문이라고 생각한다.



**5. 한국인 데이터와 유러피안 데이터로 모델 학습하기**

**(한국인 데이터 11만개 + 유러피안 데이터 12만개)**

1. 한국인 데이터와 유러피안 데이터 어느 것 하나만으로 모델을 학습하면 한쪽 데이터에 맞게 가중치가 조정되어 다른 한쪽에 대한 정확도가 매우 낮아지는 것을 확인하였다. 이번에는 두 데이터를 합쳐 모델을 학습시켜보았다. 그 결과 유러피안 데이터에는 62.84%의 정확도를 보였고 한국인 데이터에는 24.30%의 정확도를 보였다. 두 데이터 모두 비슷한 크기였지만 한국인 데이터에서 유독 낮은 정확도가 나왔다. 이를 보고 유러피안 데이터에 비해 한국인 데이터가 잘 정제되지 않았음을 알 수 있었다.

****

2. Convolution layer와 Pooling layer를 줄이고 Dense layer를 추가한 모델을 만들어 보았다.. 그 결과 한국인 데이터에는 37.70% 유러피안 데이터에는 59.59%의 정확도를 갖는 모델이 되었다. 두 종류의 데이터 모두에게 좀 더 적합한 모델이라고 생각한다.



6. 결론 및 고찰

1. 결과적으로 유러피안 데이터에 대해 62.84%의 정확도, 한국인 데이터에 대해 24.30%의 정확도를 갖는 모델을 만들 수 있었다. 유러피안과 한국인 데이터셋 둘 다에 정확도가 높은 모델을 만드는 것은 매우 어려웠다.

2. 학습에 사용하는 데이터의 종류에 대한 중요성을 알게되었다. 처음 유러피안 데이터로 훈련시켜 정확도가 62%였던 모델은 한국인 데이터에 대한 정확도가 26.22%에 불과했다. 한국인 데이터로 훈련시켜 정확도가 39%였던 모델은 유러피안 데이터에 대해서는 14%정확도에 불과했다. 어떤 데이터로 학습되었냐에 따라 모델의 가중치가 달라진다. 때문에 모델을 평가하는 데이터의 종류에 따라 정확도가 매우 다르게 나오는 것을 확인할 수 있었다. 유러피안과 한국인의 얼굴의 특징이 매우 다르기 때문에 Convolution과 Pooling레이어를 통해 추출되는 특징들이 다를 수밖에 없었다. 추출된 특징에 의해 가중치가 다르게 조정되었고 이 때문에 두 종류의 데이터에 대한 정확도가 다를 수밖에 없었다고 생각한다.

3. 학습 이미지의 양과 잘 정제된 데이터의 중요성을 알게되었다. 학습에 사용되는 이미지가 많을수록 정확도는 증가하였다. 하지만 데이터가 6배, 40배 늘어나도 정확도의 향상은 생각만큼 크지는 않았다. 특히 한국인 데이터에서 학습 이미지의 양을 매우 많이 증가시켰지만 정확도는 크게 증가하지 않았다. 모델 정확도가 이미지의 양을 많이 증가시켜도 높게 증가하지 않은 이유는 아마도 잘 정제되지 않은 이미지와, 잘못 레이블 된 데이터가 있었기 때문이라고 생각한다. 직접 한국인 이미지에 레이블링을 할 때에 어떤 클래스에 들어가야하는지 애매한 사진들이 있었고, 정면이 아닌 측면이 나오거나, 사진이 매우 어둡거나, 얼굴이 잘 보이지 않는 사진들이 있었다. 또한 공유 폴더의 한국인 라벨링은 모두 다른 사람들이 했기 때문에 사람마다 애매한 표정에 대해 레이블링을 다르게 하여 정확도가 생각보다 높게 나오지 않은 것으로 생각된다. 잘 정제된 데이터를 사용한다면 더 나은 정확도를 얻을 수 있을 것이라고 생각한다.

4. 또한 유러피안 데이터들은 대부분 배경이 없이 사람의 얼굴 만 잘라져 있었지만 한국인 데이터는 주변의 배경이 많았다. 이 부분 또한 한국인 데이터의 정확도가 높지 않았던 이유였다고 생각한다. 이 부분을 해결하기 위해 opencv의 Haar Cascade를 이용하여 한국인 데이터에서 얼굴만 추출해보고자 하였지만 생각보다 검출이 잘 되지않아 이 방법은 사용하지 못하였다. 만약 한국인의 얼굴부분만 크롭할 수 있다면 좀 더 정확도가 향상될 수 있을 것이라고 생각한다.