

CNN을 통한 연령 및 성별 추정기

Age and Gender Estimate Application using Convolutional Neural Network

요 약

머신 러닝의 발전에 따라 머신 러닝을 이용한 각종 어플리케이션이 만들어지고 있다. 특히 Computer Vision 분야는 발전하고 있는데, Computer Vision 에서는 거의 필수적으로 Convolutional Neural Network (CNN)을 사용하고 있다. 이 Convolutional Neural Network 을 이용하여 연령 및 나이를 추정할 수 있는 어플리케이션을 만든다.

1. 서 론

이미지 속의 사람의 얼굴로부터 연령과 성별을 추정할 수 있다면, 그것은 여러 분야에서 매우 큰 도움이 될 것이다. 예를 들어, 최근 코로나19 바이러스로 인해 많은 상점에 체온을 탐지하는 기기가 설치되어 있다. 기기에는 여러 종류가 있지만, 그 중에서 열화상 카메라를 통한 체온 측정 방식은, 사람이 열화상 카메라 앞에 서서 체온을 측정을 한 뒤 체온이 일정 이하일 때 출입이 가능한 구조이다. 만약 이러한 카메라에 그 사람의 연령과 성별을 추정하는 기능을 추가한다면 어떨까? 백화점에 출입하기 위해서는 필수적으로 거쳐야하는 열화상 카메라에 그런 기능이 있다면, 체온 측정을 할 때마다 연령과 성별 정보를 데이터베이스에 저장하여 그 데이터를 영업에 활용할 수도 있을 것이다.

이처럼 인물의 연령과 성별을 추정할 수만 있다면 여러 분야에 큰 도움이 될 수 있다.

이 보고서에서는 Age and Gender Classification using Convolutional Neural Networks [1] 을 바탕으로 살짝 변형된 Network 를 구성하고, 이를 바탕으로 코드를 작성하여 사람의 얼굴 사진에 따라 연령 및 성별을 추정할 수 있는 Application 을 만들 것이다.



그림 1:

Adience dataset의 이미지. 저해상도 및 가면을 쓴 사람의 이미지도 존재한다.

이 모델의 학습을 위해서 Adience 데이터셋 (그림 1)을 사용한다. 이 데이터셋은 Age and Gender Classification using Convolutional Neural Networks [1] 에서 사용된 데이터 셋이다.

2. 네트워크 구성

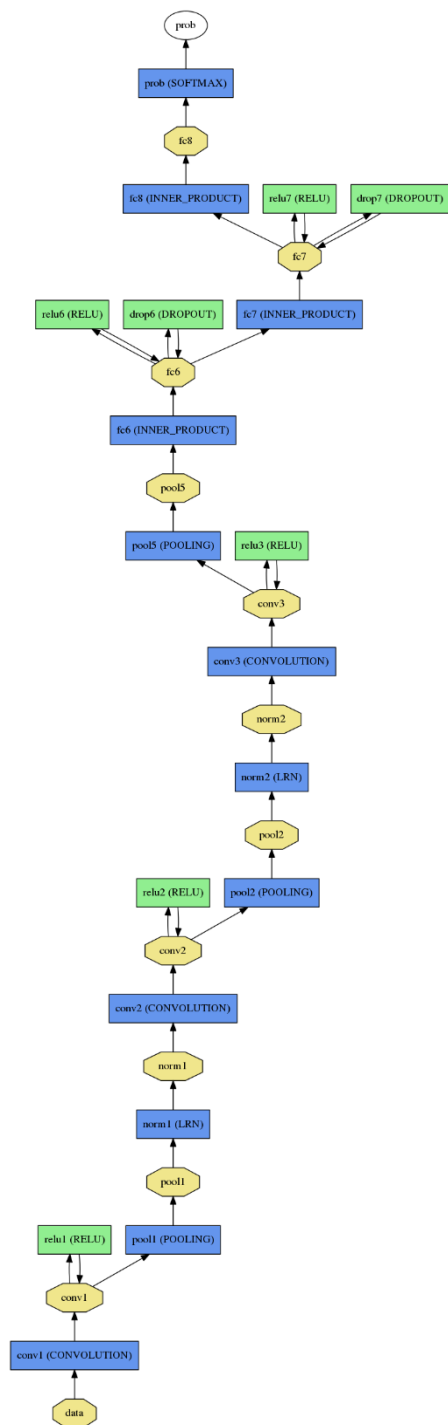


그림 2:

Age and Gender Classification using Convolutional Neural Networks[1] 에서 사용되는 network.

그림 2는 Age and Gender Classification using Convolutional Neural Networks[1] 에서 사용되는 Convolutional Neural Networks 이다.

하지만, 그림 3을 통해 알 수 있듯이 이 Network 를 사용 했을 때 제대로 학습이 되지 않았다.

```
epoch 2/10 [=====] - 21s 55ms/step - loss: 0.6909 - accuracy: 0.5368 - val_loss: 0.6914 - val_accuracy: 0.5300
epoch 3/10 [=====] - 21s 55ms/step - loss: 0.6905 - accuracy: 0.5368 - val_loss: 0.6913 - val_accuracy: 0.5300
epoch 4/10 [=====] - 21s 55ms/step - loss: 0.6905 - accuracy: 0.5368 - val_loss: 0.6917 - val_accuracy: 0.5300
epoch 5/10 [=====] - 21s 55ms/step - loss: 0.6908 - accuracy: 0.5368 - val_loss: 0.6914 - val_accuracy: 0.5300
epoch 6/10 [=====] - 21s 55ms/step - loss: 0.6907 - accuracy: 0.5368 - val_loss: 0.6913 - val_accuracy: 0.5300
```

그림 3:

그림 2의 Network 를 사용했을 때의 학습

따라서 Network 를 그림 4와 같이 변형 하였다.

각 Normalization Layer 와 Dropout Layer 를 없애주고, 마지막 Fully connected Layer 의 활성화 함수를 Softmax 에서 Sigmoid 로 변경해주었다.

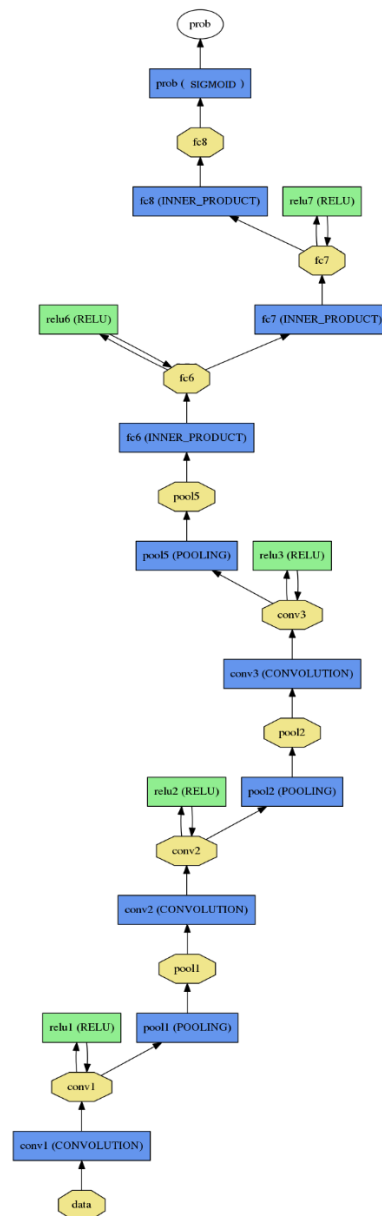


그림 4: 그림 2의 Network 를 변형시킨 Network

그림 4의 Network 를 적용했을 때 학습도 잘되고 정확도도 더 올라갔다.

이 Network 는 아래와 같이 구성되어 있다.

1. 먼저 사진이 227 X 227 사이즈로 리사이즈 된다.
2. 3 X 7 X 7 pixels 사이즈, 96개의 filter가 첫 번째 convolutional layer 에 입력된다. Rectified linear operator로는 ReLU를 사용하고 3X3 2stride 만큼 max pooling 을 한다.
3. 96 X 5 X 5 pixels 사이즈, 256 개의 filter가 두 번째 convolutional layer 에 입력된다. 그리고 똑같이 ReLU, 3X3 2 stride 만큼 max pooling을 한다.
4. 256 X 3 X 3 pixel 사이즈 384 개의filter가 세 번째 convolutional layer에 입력되고, 전과 똑같은 조건으로 max pooling 을 한다.
5. 2번의, ReLU와 512 neurons 를 포함한 Fully connected layer 를 거친다.
6. 마지막으로 SIGMOID Fully connected layer 를 통해 값을 도출한다.

3. 테스트 및 결과

먼저 데이터를 전 처리 해준 후, 훈련 데이터와 테스트 데이터로 나눈다. 이미지를 227X227으로 리사이징 후 위 네트워크와 같이 모델을 정의하여 훈련 시킨다. 이때 batch size 는 32, epochs 는 15회로 하였다. Epochs 를 10회로 할 때는 과소적합, 20회가 넘어가면 과잉적합 현상이 나타나서 15회로 하였다.

그림 4 아래와 같은 테스트 결과가 도출 되었다.

```
Gender Test loss: 0.5352568626403809
Gender Test Accuracy: 0.8124522566795349

Age Test loss: 1.7252541780471802
Age Test Accuracy: 0.46906036138534546
```

그림 5: 성별, 연령 별 테스트 결과

그리고 훈련된 모델을 통해 실제 인물의 얼굴 이미지 (그림 6) 를 입력시켜 연령과 성별을 예측해보았다.



그림 6: 테스트 이미지

결과는 그림7 과 같다.

위 인물은 남성, 1974년 생이다. 그림 6의 사진이 과거에 찍혔다는 것에 감안해도 테스트 결과의 나이와는 조금 차이가 있다.

```
[[1.]]
[[9.0900403e-07 4.4382524e-05 3.0835851e-03 1.3486947e-07 1.0000000e+00
 6.9458359e-01 9.9986815e-01 9.6667159e-01]]
성별 : 남성
나이 : 25-32
```

그림 7: 실제 이미지 테스트 결과

4. 결론

성별은 남 / 여로 나뉘는 만큼 정확도가 높았다. 하지만 연령 같은 경우에는 여러 그룹으로 나뉘고, 그 인물이 동안인지 노안인지, 화장을 했는지 와 같은 여러 요소 때문인지 정확도가 떨어지는 모습을 볼 수 있다는 점에서 한계였다.

4. 참조

[1] Gil Levi; Tal Hassner. Age and Gender Classification using Convolutional Neural Networks. 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2015