

얼굴 영상을 통한 나이 예측 시스템 보고서

2015272038 김태한

1. 주제

CNN을 사용한 3 channel 얼굴 영상 feature 추출 및 나이 예측 시스템 설계

2. 수행 과제

모바일 스마트 기기를 통한 제품 및 서비스의 확산과 행복 및 삶의 질을 목적으로 하는 소비자 집단이 증가하면서 신체와 정신의 건강한 상태를 추구하는 웰니스(Wellness) 영역이 확대되어 가고 있다. 웰니스 영역이 확대됨에 따라 사람들은 간단한 서비스를 이용해 생활 건강 관리, 피트니스·체력관리, 웰에이징(well-aging) 등이 포함된 셀프 케어로 삶의 질을 높이하고자 한다.

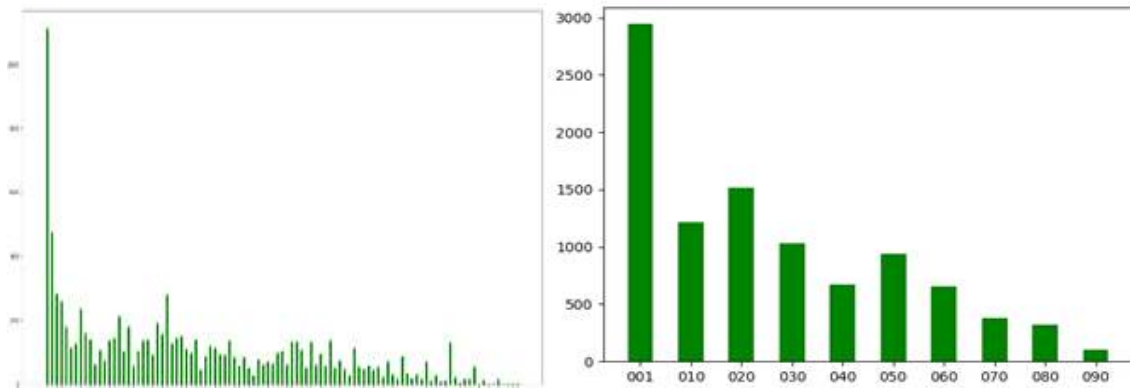
셀프케어 분야는 소비자의 현재 나이와 측정 나이의 차이에 따라 관리 방법 또한 달라지기 때문에 나이 진단의 중요성이 높아진다. 나이 진단 연구를 진행한 선행 연구 팀서는 여러 가지 얼굴 특징을 연령에 따라서 평가, 분석하였다. 그 결과를 이용하여 주름이 얼굴 나이 예측에 있어 연령과 정비례의 관계가 있음을 나타내었다. 이러한 결과 외에도 특징을 추출함에 있어서 주름 뿐만 아니라 변색, 종양을 측정하여 현재 얼굴 나이를 계산하였다. 이처럼 얼굴의 노화에는 수많은 요소들이 작용을 하고 연구팀에 따라 서로 다른 요소로 얼굴 나이를 예측하기 때문에 다양한 부분을 사용하며, 현재까지는 나이를 예측함에 있어 정확한 지표가 정해져 있지 않다. 따라서 이번 연구에서는 나이를 예측할 수 있는 방법 중 가장 기본적인 방법인 사람의 전체 얼굴 사진을 이용하여 나이를 예측하는 연구를 진행해 보았다.

3. 과제(작품)의 개발과정

과제의 개발 첫 단계에서는 우선 연구 주제와 관련하여 수업에서 배운 내용을 다시 확인하며 사용할 이론을 공부하였다. 또한 선행연구가 있는지, 비슷한 연구 방향을 띠고 있는 논문을 찾아보았고, 이를 정리하여 읽어보며 연구에 필요한 기법 및 방향을 설정하였다.

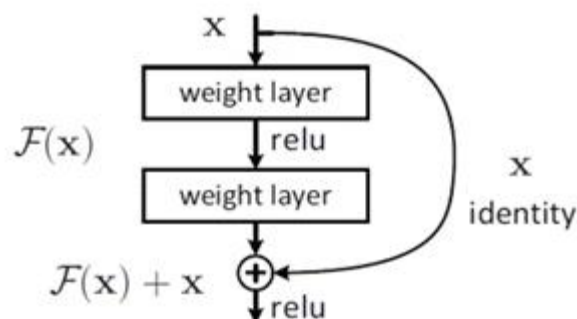
이후로 본 연구에서 사용할 3Channel 얼굴 영상 데이터를 탐색하였다. 탐색 결과, Kaggle 웹사이트(<https://www.kaggle.com/frabbisw/facial-age>)에서 총 9629장의 영상을 얻었으며, 서로 다른 영상의 크기를 224*224로 통일하였다. 전체 데이터의 80%(7703장)의 Train set과 나머지 10%(963장)의 Validation set으로 Data set을 구성하여 학습을 진행하였으며 10%(963장) Test set으로 결과를 확인하였다.

얼굴 영상과 나이라는 데이터의 특성상 나이가 많은 고령자의 데이터 수가 나이가 적은 영유아에 비해서 적은 Class 불균형이 존재한다. 이러한 Class 불균형을 해결하기 위해서 10살을 기준으로 묶어서 Class 당 부족한 데이터를 채워주었고 비교적 데이터가 많은 Class의 경우 Under Sampling을 하였다.



[그림1. Class 불균형]

다음으로 얻어진 데이터셋을 활용할 수 있는 적합한 모델을 선정하는 단계에 들어갔다. 얼굴 영상이라는 데이터의 특성상 보다 깊은 CNN모델을 요구하였고 초기에 직접 만들었던 간단한 CNN모델로는 학습이 되지 않는 것을 확인하였다. 주어진 시간 내에 완성해야하는 프로젝트의 특성 상 시간을 고려하여 2015년 ILSVRC에서 우승을 차지한 이미지 분류용 CNN모델인 ResNet모델을 선정하였다. ResNet모델은 네트워크의 깊이가 깊어짐에 따라 생기는 문제점인 Gradient vanishing을 해결하였다. Gradient Vanishing은 깊이가 깊은 모델에서 Back propagation을 진행할 때마다 weight의 미분 값이 너무 작아져 0으로 수렴하여 학습이 안되는 현상을 말한다. 이러한 문제점을 해결하기 위해서 ResNet에서는 입력 값을 출력 값에 더해 주어 미분하여도 적어도 1이상의 값이 나와 안정적인 학습이 되도록 만들었다. 따라서 보다 깊은 신경망을 만들 수 있는 장점이 있기 때문에 ResNet50모델을 선정하였다.



[그림2. Residual Block]

총 9629장의 데이터를 7703장의 Train set과 962장의 Validation set으로 나누어 8:1의 비율로 학습을 진행하였다. 모든 데이터의 값을 -1에서 1사이의 값으로 정규화를 진행했고 크기를 224*224크기로 통일시켜주었다. 10의 batch size로 30에폭 학습을 진행했다. learning rate는 0.001로 초기에 설정하여주고 20 에폭이 진행될 때 scheduler 함수를 통해 0.1이 곱해진 0.0001로 감소하게 하였다. Optimizer로 adam optimizer를 사용하였다.

다음으로 Loss Function을 만드는 단계에 들어갔다. 처음에는 Cross Entropy Error를 이용하였는데 학습이 잘 진행되지 않았고 나이의 예측 값이 하나에 수렴하지 않고 분산되거나 양분화되는 문제점이 생겼다. 따라서 Mean Squared Error와 Variance Error를 합하여 사용했다. Cross Entropy Error를 통해서 정답Class에서의 확률 오차를 줄이고 Mean Squared Error를 통해서 Prediction값의 평균값이 정답에 근접하도록 만들었다. 여기서 사용한 평균값은 밑의 수식으로 계산된다. 마지막으로 Variance Error를 통해서 Prediction의 Class가 양분화되거나 넓게 퍼지는 것을 막아 제대로 학습이 되도록 하였다. 또한 정확도 평가를 위해서 Accuracy 방식을 사용하였다.

$$Mean = [Class1 \ Class2 \ Class3 \ ... \ ClassM] \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \dots \\ P_M \end{bmatrix}$$

<수식1. Mean>

$$Variance = [(Class1 - Mean)^2 (Class2 - Mean)^2 (Class3 - Mean)^2 \dots (ClassM - Mean)^2] \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \dots \\ P_M \end{bmatrix}$$

<수식2. Variance>

$$Accuracy = \frac{TP}{}$$

<수식3. Accuracy>

4. 결론

최종적으로 ResNet50을 이용하여 제작한 model을 딥러닝을 수행해 학습시켰을 때 구성은 다음과 같다.

[optimizer : Adam(initial learning rate : 0.001), epochs : 30, batch_size : 10, Train on 7703 samples, validate on 962 samples]

Result :

Train Loss(Cross-entropy + Mean-squared + Variance) : 0.2415

Validation Loss(Cross-entropy + Mean-squared + Variance) : 1.247(at 15)

Train Accuracy : 0.9186

Validation Accuracy : 0.5901(at 20)]

| Epochs | 5 | 10 | 15 | 20 | 30 |
|------------------|--------|--------|--------|--------|--------|
| Train State | | | | | |
| Accuracy | 0.4901 | 0.5479 | 0.6131 | 0.7157 | 0.9186 |
| Loss | 1.567 | 1.319 | 1.079 | 0.7738 | 0.2415 |
| Validation State | | | | | |
| Accuracy | 0.3102 | 0.518 | 0.5783 | 0.5901 | 0.5463 |
| Loss | 3.441 | 1.397 | 1.247 | 1.267 | 2.377 |

[표1. Epoch에 따른 Train, Validation일 때의 정확도와 오차]

나는 결과를 통해서 15에폭에서 20에폭 이후에 과대적합이 일어나는 것을 확인하였고 early stop function을 통해 5에폭 동안의 Loss의 값을 확인하여 스스로 멈추도록 하였고 그 중 가장 좋은 결과를 저장하도록 하였다.

정답: 020
예측: 030

020
010

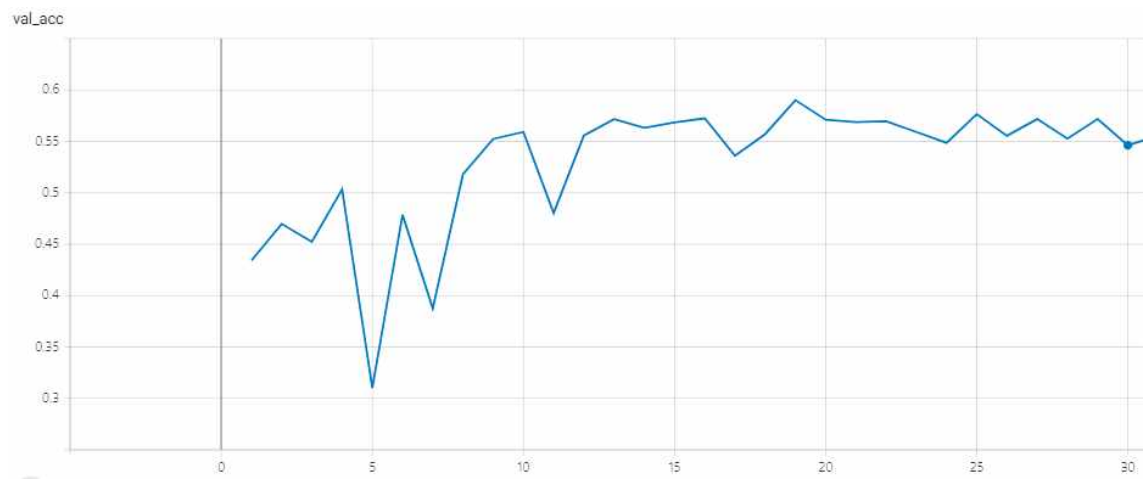
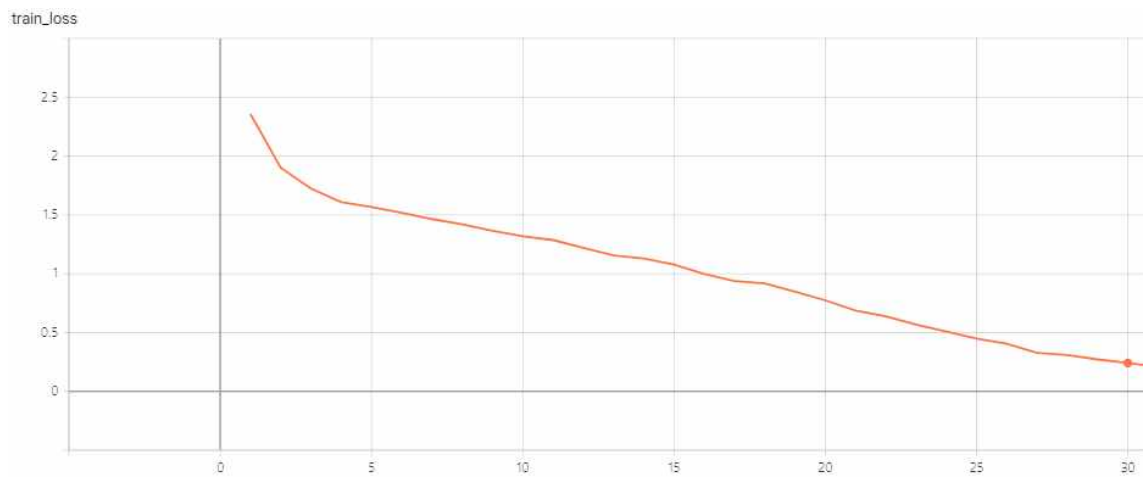
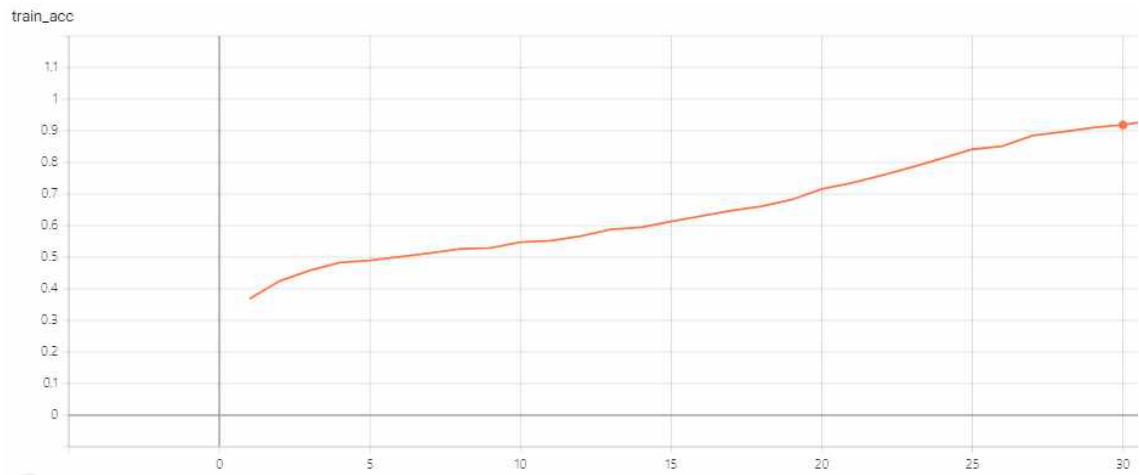
020
020

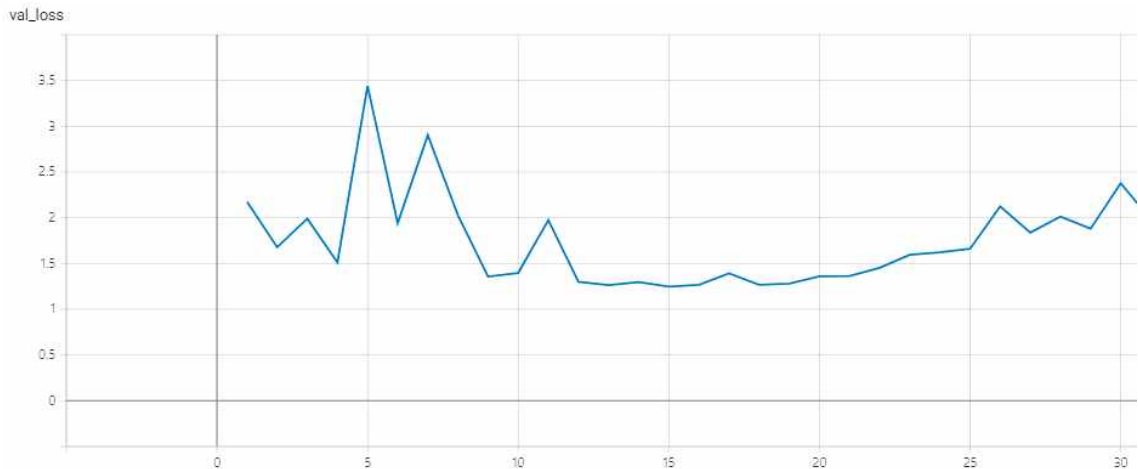
060
060

020
030



[그림3. 나와 친구들의 얼굴을 넣었을 때의 학습 결과, 김태한은 다섯 번째 그림이며, 30대로 잘못 예측하였다]





[그림4, 4, 5, 6. 위에서부터 Train Accuracy, Train Loss, Validation Accuracy, Validation Loss]

5. 프로젝트의 활용방안 및 기대효과

이 프로젝트를 진행하며 수업에서 이론으로 배웠던 CNN이 동작하여 연산을 통해 가중치가 변화하면서 학습이 이루어지는 것을 직접 확인할 수 있었으며, 웹상에서 누구나 얻을 수 있는 데이터셋을 기반으로 내가 원하고자 하는 학습과 분류의 목표를 정하고, 이에 맞게 학습을 수행하는 모델을 고르고 수정한 후 적합한 오차함수를 골라 작동하는 다음과 같은 모델을 만들 수 있었으며, 어느 정도 예측 결과가 정답에 맞춰서 학습이 이루어지는 것을 확인할 수 있었다.

나는 이렇게 얻어진 결과를 바탕으로 셀프케어 분야에 적용할 얼굴 노화 정도를 딥러닝을 이용하여 측정할 수 있다는 가능성을 볼 수 있었으며, 추후에 간단한 서비스로 제작할 수 있다면 충분한 데이터셋을 확보하여 모델을 훈련한 후, 학습된 모델을 바탕으로 입력 이미지 한 장으로 정확한 결과를 보여주는 모델도 만들 수 있을 거라고 생각이 들었다.

이 프로젝트의 활용방안으로는 폐쇄회로 상에서의 인물을 특정하게 되는 경우에 하나의 지표로 삼을 수 있을 것 같으며, 노숙자의 신원을 확보하는 등, 주민등록이 말소되는 인원의 신원을 특정하는 방법에도 어느 정도 도움이 될 수 있을 것으로 생각한다. 또한 바이오, 미용 분야에서는 정보를 입력하지 않아도 모델이 스스로 학습을 통해 얼굴과 피부를 판단하고, 결과를 바탕으로 맞춤형 서비스를 진행할 수 있는 하나의 지표로 삼을 수 있을 것으로 생각한다. 추가적으로 좀 더 심화학습을 진행할 수 있다면, 데이터셋 연구를 추가하여 좀 더 범용적으로 작동할 수 있도록 인종 및 피부색에 따른 특징들에 따른 얼굴 노화의 정도가 학습시켜 좀 더 일반화가 잘 수행된 모델을 만들고 싶다.

6. 참고논문

- 1) 이장미; 황성걸. 모바일 웰니스 제품/서비스의 현황 분석 및 기회영역 발굴. 조형미디어학, 2015, 18.1: 217-224.
- 2) KIM, Ki Bbeum, et al. Evaluation of the perception and clinical characteristics of facial skin aging according to age-group among Korean women. Asian Journal of Beauty and Cosmetology, 2019, 17.3: 387-396.
- 3) 박진영; 배정호. 얼굴 노화의 이해. Journal of Rhinology, 2012, 19.2: 87-90.
- 4) Buranasirin et al. BMC Research Notes. Development of a Global Subjective Skin Aging Assessment score from the perspective of dermatologists, 2019.
- 5) Kaiming He, Xiangyu Zhang, et al. Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016.