

ResNet을 이용한 나이 예측 모델 연구

김지훈¹, 신용태²

¹승실대학교 컴퓨터학과 석사과정

²승실대학교 컴퓨터학과 교수

jihunthank@soongsil.ac.kr, shin@ssu.ac.kr

A Study on the Age Prediction Model Using ResNet

Ji-Hun Kim¹, Young-Tae Shin²

^{1,2}Dept. of Computer Science and Engineering, Soong-Sil University

요 약

본 연구는 디지털 기술과 인공지능의 발전을 배경으로, ResNet 모델을 활용하여 얼굴 인식 및 나이 예측 시스템을 개발하고 평가한다. ResNet의 잔차 학습과 스킵 연결 기능은 깊은 신경망에서 발생할 수 있는 기울기 소실 문제를 해결하여 모델의 학습 효율을 높이는 데 중요한 역할을 한다. 또한 All-Age-Faces Dataset을 이용하여 나이 예측에서 아시아 인종에 대한 편향 없이 고르게 좋은 성능을 보여주는 것을 목표로 한다.

1. 서론

최근 디지털 기술의 발전은 일상생활의 많은 측면에서 혁신을 가져왔으며, 특히 인공지능 기술의 발전은 이미지 인식과 처리 분야에서 중대한 진보를 이루었다. 얼굴 인식 기술은 이러한 발전의 중심 중 하나에 있으며 보안, 개인화된 서비스 제공, 인터페이스의 접근성 향상 등 다양한 분야에서 활용되고 있다. 본 연구는 얼굴 인식 기능 중에서도 Resnet 모델을 이용하여 얼굴 인식을 통한 나이 예측 모델을 개발하는 것에 초점을 맞추고 있다.

얼굴 인식 기술의 주목할 만한 응용 중 하나는 나이 예측이다. 최근 사회에 고연령 계층의 디지털 격차에 대한 문제점이 대두되고 있다[1]. 이때 앞서 언급한 나이 예측 기술을 이용하여 사용자의 나이를 판별, 연령대별 맞춤 인터페이스를 제공한다면 나이와 상관없이 사용자에게 더 나은 경험을 제공하게 될 것이다.

또한, 보안 분야에서도 얼굴 인식 기술은 큰 잠재력을 지닌다[2]. 그 예시로 만약 범죄자의 몽타주를 나이 예측 기술을 통해서보다 정확히 파악한다면 이를 통해 더 빠르고 정확하게 범죄자를 추적, 최종적으로 공공의 안전을 강화하는 데 큰 역할을 할 것이다. 하지만 얼굴 인식 기술의 오류는 보안적으로

심각한 부작용을 초래할 수도 있다. 예를 들어, 2019년 4월 미국 디트로이트에서는 얼굴 인식 시스템의 오류로 인해 한 40대 아프리카계 흑인 남성이 절도 범으로 오해를 받아 체포된 사건이 있었으며, 2016년에는 아시아계 뉴질랜드인이 얼굴 인식 기능 오류로 인해 여권 사진 갱신을 거부당한 일도 있었다. 이러한 사건들은 얼굴 인식 기술에서 인종 편향을 해소하고, 정확하게 인식하는 것도 얼마나 중요한지를 강조한다.

본 연구는 ResNet을 사용하여 얼굴 인식과 나이 예측 모델을 만들어 효율성을 평가하고, 향후 연구에서 이 모델을 기반으로 연령대별 맞춤 서비스를 제공할 수 있도록 하는 것이 목적이다.

2. 선행연구

본 장에서는 기존 나이 예측 모델 연구의 방법에 대해 구체적으로 이야기하고 그 한계에 대해 알아보려 한다.

나이 예측은 컴퓨터 비전과 기계 학습 분야에서 활발히 연구되어 온 주제 중 하나로, 개인의 얼굴 이미지만을 사용하여 그 나이를 추정하는 기술이다. 초기 모델들은 주로 선형회귀나 로지스틱 회귀와 같은 전통적인 통계적 방법을 사용하여 나이를 예측하였으나, 이러한 모델은 제한된 특성 추출능력과 낮

은 정확도의 문제를 내포하고 있었다[3].

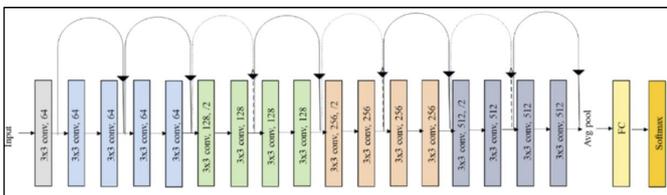
심층 학습의 등장과 함께 나이 예측 분야에 혁신이 일어났다. 특히 깊은 신경망 아키텍처는 복잡한 얼굴 특성을 더 정확하게 학습할 수 있는 능력을 제공하였으며, 이는 예측 정확도를 크게 향상하게 시켰다. 예를 들어, “DeepEXpectation(이하 DEX)” 모델은 VGG-16 네트워크를 기반으로 하여 나이를 예측하고 매우 높은 정확도를 달성하였다. 그러나 이 모델은 대량의 데이터에 대한 처리가 필요하며, 학습 시간이 매우 길다는 한계점 역시 드러내었다[4].

또한, 대부분의 연구는 서구 중심의 데이터 세트에 기반하여 개발되었으며, 이로 인해 다양한 인종의 얼굴 특성을 포괄적으로 반영하지 못하는 문제점을 지니고 있다. 이는 모델의 일반화 능력을 제한하며, 특히 아시아와 같은 비서구 지역의 인구에 대해 낮은 예측 정확도를 보이는 결과를 초래하였다[5].

본 연구는 이러한 기존 연구들의 한계를 극복하고자 한다. 선행된 연구에서 소외된 아시아 인종을 포함하는 데이터 세트를 사용하여 모델의 일반화 능력을 강화하고, 최신 심층 학습 기술을 활용하여 학습 효율을 개선할 계획이다. 이를 통해 더욱 정확히 신뢰할 수 있는 나이 예측 모델을 개발하고자 한다.

3. ResNet[6]

ResNet (Residual Network)은 딥러닝 분야에서 혁신적인 발전을 가져온 CNN 아키텍처의 일종으로, ResNet은 특히 깊은 신경망을 구축하는 데 있어서 발생할 수 있는 기울기 소실(Gradient Vanishing) 문제를 해결하는 데 중점을 두고 설계되었다. 많은 경우 보통 ResNet-50 이상의 레이어가 사용되지만 본 연구에서는 데이터의 양이 비교적 부족하여 과도하게 많은 레이어는 overfitting 현상을 일으킬 수 있기에 ResNet-18 모델로 변형하여 이용하였다.



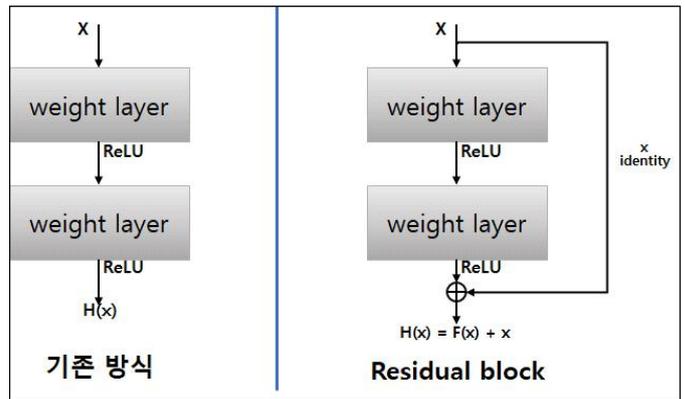
(그림 1) ResNet-18 모델의 레이어 구성도

이 모델의 핵심 특징은 ‘잔차 학습(Residual Learning)’과 ‘스킵 연결(Skip Connection)’이다. 먼저, 잔차 학습이란 네트워크가 출력을 직접 학습하는 대신, 출력과 입력의 차이를 학습함으로써 학습

과정을 용이하게 하는 것을 의미한다.

두 번째로 스킵 연결은 입력을 몇 개의 레이어를 건너뛰고 직접 출력에 더하는 방식으로, 네트워크가 깊어져도 기울기가 쉽게 소실되지 않도록 해준다.

이러한 구조 덕분에 ResNet은 신경망이 깊어져도 효율적인 학습이 가능하며, 이는 깊은 네트워크에서도 높은 성능을 달성할 수 있게 해주기에 본 연구에서 알고리즘 모델로 선택하였다.

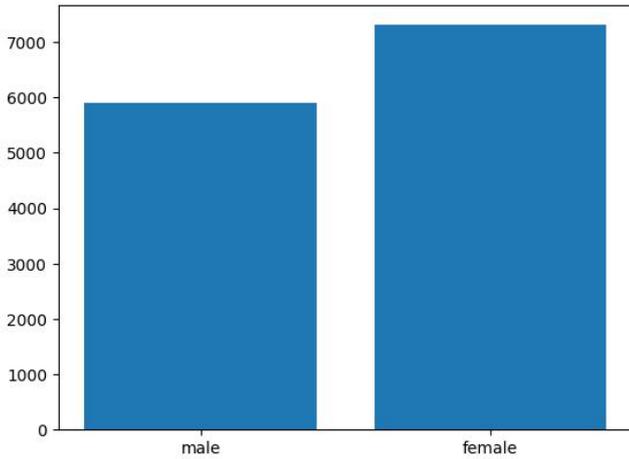


(그림 2) 기존 방식의 학습 방식과 ResNet의 잔차 학습과 스킵 연결

4. 학습 데이터

4.1 All Age Face Dataset(AAF)[7]

All-Age-Faces(이하 AAF)는 2세부터 80세까지의 다양한 연령대를 포함하는 총 13,322장의 얼굴 이미지를 담고 있는 데이터베이스이다. 이 데이터 세트는 특히 아시아 인종의 여성 얼굴 이미지 7,381장과 남성 얼굴 이미지 5,941장을 다양하게 포함하고 있다. 이 컬렉션은 나이 예측 및 성별 분류와 같은 작업을 위한 기계 학습 알고리즘의 개발 및 평가를 위해 설계되었다. 본 데이터 세트의 주요 특징으로는 각 이미지가 개인별로 고유하며, 개인의 일련 번호와 나이를 나타내는 특정 형식으로 명명되어 모델에 학습시키기 특화되었다는 것이다. 또한, 성별 레이블이 제공되어 여성은 0, 남성은 1로 표시된다, 본 데이터 세트는 아시아 인종을 위주로 구성되어 있어서보다 아시아인에 적합한 학습을 도와줄 수 있기에 사용되었다. 다음 (그림 3)은 AAF의 남녀 비율을 그래프로 보여준다.

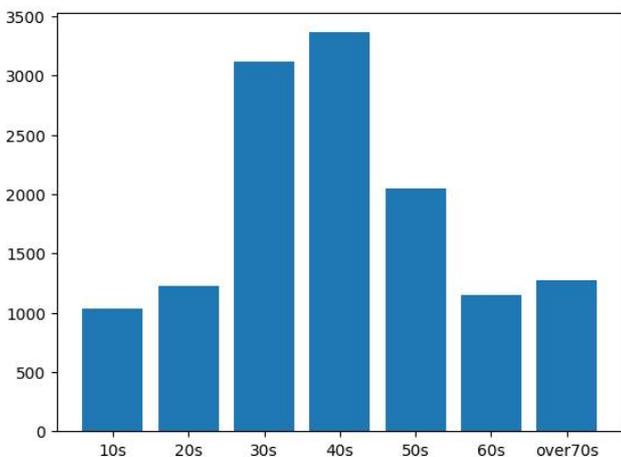


(그림 3) AAF의 남녀 성비 그래프

4.2 데이터 전처리

본 연구에서는 앞서 설명한 AAF 데이터 세트를 학습시키기 전에 데이터를 더욱 정교하게 처리하기 위해 표준화와 데이터 증강과 같은 보편적인 데이터 전처리 기법을 적용하였다. 표준화는 데이터의 스케일을 일정하게 조정하며 모델이 각 특성의 중요도를 올바르게 인식하고 효율적으로 학습할 수 있도록 돕는 역할을 한다. 또한, 데이터 증강은 이미지를 다양하게 변형하여 학습 데이터의 다양성을 향상시키고, 모델이 다양한 환경에서의 패턴을 학습할 수 있도록 한다.

그뿐만 아니라, 본 연구에서는 예측의 정확도를 높이기 위해 데이터를 미성년자부터 70대까지의 나이 그룹으로 범주화하였다. 이는 나이에 따른 다양한 특성을 고려하여 모델이 더 다양한 연령대의 데이터에 대해 학습하고 예측할 수 있도록 돕는 역할을 한다. 범주화된 데이터는 (그림 4)와 같다.



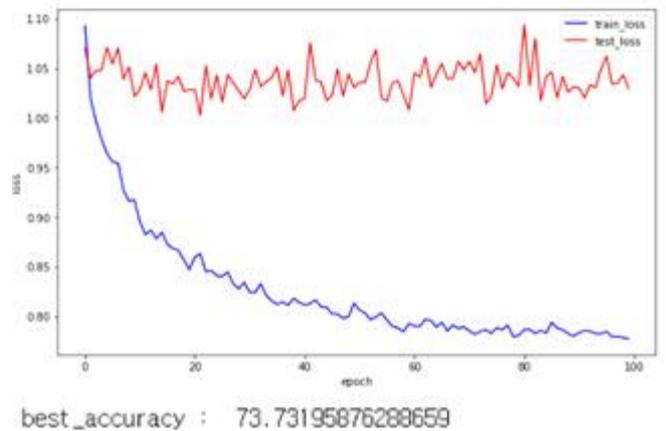
(그림 4) AAF의 연령대별 범주화 비율

5. 실험 및 결과

학습 모형은 다음과 같다. 본 연구에서는 ResNet18 모델을 사용하여 얼굴 이미지를 통한 나이 예측을 수행한다. 모델의 마지막 연결층(fc)은 512개의 입력 특성을 받아 512개의 뉴런으로 구성된 Dense Layer로 시작하여, ReLU 활성화 함수와 30%의 Dropout을 거쳐 7개의 출력을 가진 Softmax Layer로 끝난다. 이 구조는 최종적으로 클래스 확률을 출력하여, 각 이미지의 나이 범주를 예상한다.

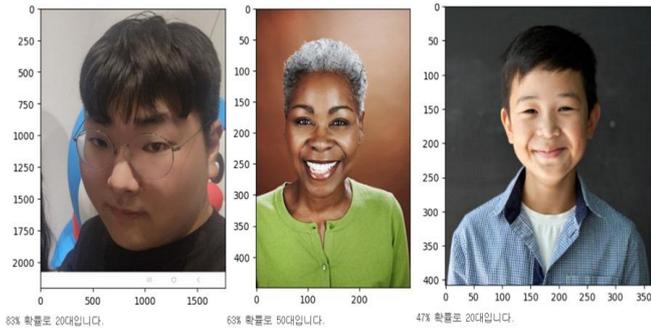
데이터 세트는 전처리 과정을 거쳐, train_set와 test_set로 나뉘어 진다. train_set와 test_set의 비율은 8:2로 설정하였다. train_set는 데이터 증강을 통해 두 배로 확장한다. 이렇게 확장된 이미지는 클래스 불균형을 해소하기 위해 가중치가 적용된 무작위 sampling으로 배치가 구성된다. test_set 역시 동일한 가중치 sampling 방식을 사용하여 과적합을 방지하고 모델의 일반화 성능을 평가한다.

학습 과정은 Cross Entropy Loss 함수와 Adam 최적화 알고리즘을 사용하여 진행된다. 학습률과 가중치 감소는 각각 0.0001로 설정된다. 모델은 각 epoch마다 성능을 평가하여 가장 좋은 검증 결과를 보인 시점의 모델을 저장한다. 이 과정에서 학습률 스케줄링은 사용되지 않는다. 학습된 모델은 최종적으로 test_set에 대한 정확도와 손실을 출력하며, 이를 통해 모델의 성능을 평가한다.



(그림 5) 매 epoch마다 기록된 loss값과 가장 높은 정확도를 보인 버전의 정확도(약 73.3%)

train_loss는 학습이 잘되어 점차 감소하는 모습을 보이지만 test_loss는 overfitting 되는 경향을 보여주었다. 모델의 정확도는 약 73% 정확도를 보여주었다.



(그림 6) 실제 사진을 이용한 정확도 측정 테스트(순서대로 20대, 50대, 10대)

다음은 실제 사진을 이용해서 모델의 정확도를 테스트 하였다. 특히 인종 편향에 대한 테스트를 위해 기존의 연구에서 많이 이용된 백인이 아닌 흑인과 아시아 인종에 사진을 대입하여 정확도를 측정하였다. 그 결과 먼저 아시아 인종 20대와 흑인 50대의 경우 각각 83%, 63% 확률로 정확하게 측정하였다. 그러나 10대의 경우 47% 확률로 20대를 예측하여 잘못된 예측을 보여주었다.

6. 결론

본 연구에서는 ResNet 모델을 사용하여 얼굴 인식을 통한 나이 예측 시스템을 개발하였다. 실험 결과, train_loss는 만족스러운 수준으로 낮게 유지되었으나, 검증 손실에서는 overfitting 되는 경향이 나타났다. 또한 20대에서 50대는 비교적 정확한 예측을 하였으나 미성년자와 60대 이상의 나이대에서는 비교적 부정확한 예측을 보여주었다. 이는 학습 데이터의 부족이 원인으로 추정된다. 모델의 테스트 데이터 세트에 대한 정확도는 약 73% 이상 달성하였으며 동양인의 얼굴 인식에 있어서도 효과적인 모습을 보여주었기에 ResNet 모델이 나이 예측과 얼굴 인식 분야에서 유용하게 사용될 수 있음을 시사한다. 하지만 test_loss에서 overfitting되는 경향을 보여주었기에 향후 연구에서는 더 풍부한 학습 데이터를 이용하고 모델 구조의 최적화를 하여보다 정확한 나이 예측 모델을 만들어야 할 필요가 있을 것이다.

이 연구는 나이 예측 및 얼굴 인식 분야에서의 AI 기술 적용 가능성을 확인하였으며, 추가 연구를 통해 이 기술의 실용성과 정확성을 더욱 높일 수 있을 것으로 기대된다.

사사문구

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과로 수행되었음"
(IITP-2024-2020-0-01602)

참고문헌

- [1] REPETTI, Marion; FELLAY-FAVRE, Elisa. Ageism and the digital divide in Switzerland during COVID-19: Lessons for the post-pandemic world. *Journal of Aging Studies*, 2024, 69: 101227.
- [2] GOKULAKRISHNAN, S., et al. An optimized facial recognition model for identifying criminal activities using deep learning strategy. *International Journal of Information Technology*, 2023, 15.7: 3907-3921.
- [3] FU, Yun; HUANG, Thomas S. Human age estimation with regression on discriminative aging manifold. *IEEE Transactions on Multimedia*, 2008, 10.4: 578-584.
- [4] ROTHE, Rasmus; TIMOFTE, Radu; VAN GOOL, Luc. Dex: Deep expectation of apparent age from a single image. In: *Proceedings of the IEEE international conference on computer vision workshops*. 2015. p. 10-15.
- [5] PANIĆ, Nenad; MARJANOVIĆ, Marina; BEZDAN, Timea. *Ethnic Representation Matters: Investigating Bias in Facial Age Prediction Models*. 2024.
- [6] HE, Kaiming, et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. p. 770-778.
- [7] CHENG, Jingchun, et al. Exploiting effective facial patches for robust gender recognition. *Tsinghua Science and Technology*, 2019, 24.3: 333-345.