

# 얼굴 나이 예측을 위한 경량 CNN 모델 설계

박소정\*, 장석우\*

\*안양대학교 소프트웨어학과

e-mail: swjang7285@gmail.com

## Design of a Lightweight CNN for Facial Age Estimation

So-Jeong Park\*, Seok-Woo Jang\*

\*Department of Software, Anyang University

### 요약

얼굴 이미지를 기반으로 한 나이 예측은 조명, 표정, 촬영 각도 등 다양한 변수에 영향을 받아 정확한 예측이 어려운 문제이다. 최근 딥러닝 기술의 발전으로 합성곱 신경망(CNN)을 활용한 방법들이 높은 정확도를 보이고 있으나, 모델의 깊이에 따른 연산량 증가로 인해 실시간 시스템이나 모바일 환경에서의 적용에는 제약이 존재한다. 본 논문에서는 UTKFace와 ALL-AGE-FACE 데이터셋을 병합하고, 좌우 회전 증강기법을 적용하여 학습 데이터를 확장한 후, 128×128 크기의 흑백 이미지를 입력으로 받는 CNN 기반의 경량 모델을 설계하였다. 실험 결과, 제안 모델은 MAE 5.28, RMSE 7.65의 성능을 기록하며, VGG16 및 ResNet50 등 대형 모델과 유사한 수준의 예측 정확도를 보이면서도 약 3.7M의 파라미터 수로 경량성을 확보하였다. 이러한 결과를 통해 제안된 모델이 정확도와 연산 효율성 간의 균형을 이루었으며, 실시간 응용 환경에 적합함을 확인하였다.

## 1. 서론

얼굴 나이 추정 기술은 이미지에서 연령과 관련된 시각적 특징을 분석하여 나이를 예측하는 방식으로, 보안 시스템, 사용자 맞춤 광고, 실종자 추정, 애니메이션 이미지와 메타버스 아바타 생성 등 다양한 분야에 적용되고 있다. 그러나 얼굴 나이 예측은 표정, 조명, 촬영 각도 등 다양한 외부 변수에 영향을 받기 때문에, 안정적으로 높은 정확도를 유지하는 것이 쉽지 않다. 기존에는 사람이 정의한 규칙 기반의 특징 추출 기법이 주로 활용되었으나, 복잡한 얼굴 이미지에서는 유의미한 정보를 안정적으로 추출하기 어려웠다. 이후 딥러닝 기술의 발전으로 합성곱 신경망(CNN)이 도입되면서 이미지로부터 특징을 자동으로 학습할 수 있게 되었고, 예측 정확도는 크게 향상되었다. 다만, 깊은 구조의 CNN은 모델 크기와 연산량이 크기 때문에 실시간 시스템이나 모바일 환경에서의 활용에는 제약이 따른다. 이에 따라 최근 연구들은 모바일 환경에서의 적용을 위한 연산 효율 개선[1], 연령대 불균형 해소를 위한 데이터 증강 기법[2], 사전 학습 모델을 활용한 전이 학습 등을 통해 성능 개선을 시도하고 있다[3].

본 논문에서는 CNN 기반의 경량 모델을 설계하고, 다양한 나이의 얼굴 이미지에 대해 예측 성능과 연산 효율을 함께 평

가함으로써, 실시간 시스템에 적합한 모델의 가능성을 확인하고자 하였다.

## 2. 데이터 및 모델

본 연구의 나이 예측 모델 학습에는 Kaggle에서 제공되는 UTKFace와 ALL-AGE-FACE(AAF)[3] 데이터 세트를 병합해서 사용하였다. UTKFace는 1-116세의 약 23,000장, AAF는 2-80세의 약 13,000장으로 구성되며, AAF는 대부분 아시아권 인물로 이루어져 있다. 두 데이터는 픽셀 크기와 정렬에서 차이가 있어, 통일성을 위해 AAF 이미지를 코플 기준 200×200 크기로 리사이징하였다. 이후 연령대 불균형을 줄이기 위해 20-30대 일부 데이터를 제외하고 최종적으로 30,461장의 이미지를 사용하였다.

데이터는 학습, 검증, 테스트 세트로 6:2:2로 분할되었고, 학습 데이터에는 좌우 15도 회전 증강을 적용해 3배로 확대하였다. 이로써 모델이 다양한 조건에서 더욱 견고한 예측 성능을 확보할 수 있도록 하였다[2].

제안된 모델은 128×128 크기의 흑백 이미지를 입력으로 받아 처리하는 9개 층의 CNN 구조이다. 4개의 컨볼루션 계층에는 3×3 필터, 배치 정규화, ReLU, 맥스풀링이 적용되며, 필터 수는

계층이 깊어질수록 32에서 256까지 2배씩 증가한다. 추출된 특징은 Flatten 층을 거친 후, 256, 128개의 뉴런으로 구성된 완전 연결층(Dense layer)을 통과하고, 마지막 출력층에서는 선형 활성화 함수를 통해 연속적인 나잇값을 예측한다.

모델은 평균 제곱 오차(MSE)를 손실 함수로 사용하며, 옵티마이저는 Adam을 적용했다. 학습률은  $1e-4$ 에서 시작해, 검증 손실이 일정 에포크 이상 개선되지 않을 경우, 자동으로 절반으로 감소시키는 동적 조정 기법을 적용했고, 최소 학습률은  $1e-6$ 으로 제한해 학습 안정성과 과적합 방지를 동시에 고려하였다.

### 3. 실험 결과

제안한 CNN 기반 나이 예측 모델의 성능을 검증하기 위해 TensorFlow 기반으로 Google Colab Pro 환경에서 A100 GPU를 사용하여 학습 및 평가를 수행하였다. 비교 대상은 EfficientNetB1, MobileNetV2, InceptionV3, VGG16, ResNet50 등 사전 학습된 다섯 가지 모델이다[4]. 이들 모델은 모두 CNN 기반의 딥러닝 모델로, 구조적 특성과 특징 추출 방식에 따라 차이가 있으며, 깊이가 깊은 고성능 모델과 연산 효율을 고려한 경량 모델로 구분된다.

성능 평가는 MAE, MSE, RMSE 세 가지 회귀 지표를 기준으로 진행되었다[5]. MAE는 평균 절대 오차로 직관적인 차이를 나타내며, MSE는 큰 오차에 민감하고, RMSE는 정규 분포일 때 유용하다. 본 연구에서는 오차를 다양한 각도에서 파악하기 위해 세 지표를 함께 평가에 사용하였다[6].

결과적으로 VGG16이 가장 낮은 MAE 5.14, RMSE 7.33를 기록하였고, 제안 모델은 MAE 5.28, RMSE 7.65로 두 번째로 낮은 오차를 보였다. 다만 경량화 모델인 MobileNetV2 MAE 5.69, RMSE 8.0 및 EfficientNetB1 MAE 6.67, RMSE 9.23 보다는 높은 예측 정확도를 보였다.

모델 크기는 TensorFlow의 count\_params() 함수로 측정하였으며, 제안 모델의 파라미터 수는 약 3.7M으로, MobileNetV2의 2.27M 다음으로 작았다. 결론적으로, 제안된 모델은 상대적으로 적은 파라미터 수에서도 높은 정확도를 보여, 실시간 예측 환경에서도 효과적으로 활용될 수 있음을 확인하였다.

### 4. 결론

얼굴 나이 예측은 다양한 분야에서 활용도가 높아지고 있으나, 조명, 각도, 표정 등 다양한 변수로 인해 정확한 예측이 쉽지 않다. 특히 모바일 환경이나 실시간 시스템에서는 높은 성능과 함께 경량화된 모델이 요구된다. 기존 CNN 기반 모델은 자동 특징 학습을 통해 정확도는 높지만, 성능을 높이기 위해 네트워크가 깊어질수록 연산량이 증가해 실시간 환경에

서 활용성이 떨어질 수 있다.

이에 본 연구에서는 UTKFace와 AAF 두 데이터 세트를 병합하고, 좌우 회전 증강을 통해 훈련데이터를 확대하였으며, 입력 이미지를  $128 \times 128$  크기의 흑백 형태로 변환하여 효율을 높였다. 제안된 모델은 9층 구조로 설계되었으며, 필터 수와 네트워크 깊이를 조절하여 과도한 연산 없이도 안정적인 예측 성능을 유지할 수 있도록 구성하였다.

실험 결과, 제안 모델은 MAE 5.28, RMSE 7.65를 기록하며 VGG16, ResNet50과 유사한 성능을 보였고, 파라미터 수는 3.7M으로 MobileNetV2 다음으로 작았다. 경량 모델인 EfficientNetB1, MobileNetV2보다 높은 정확도를 보여, 경량성과 성능의 균형을 확보한 것으로 평가된다. 따라서, 제안된 CNN 모델은 비교적 적은 연산 자원으로도 높은 정확도를 유지할 수 있어 실시간 적용에 적합하다.

앞으로는 특징 추출 기반 데이터 증강을 통해 특정 연령대의 데이터를 보완하고, 모바일 환경에서의 적용 가능성과 실시간 성능을 평가할 예정이다.

#### 참고문헌

- [1] S. J. Lee and Y. H. Kim, "Face Image-based Age Prediction Using WEKA," Proceedings of the 4th Korea Artificial Intelligence Conference, pp. 66, 2023.
- [2] E. Woo, S. Y. Yeo, and Y. S. Ma, "Comparison of Data Augmentation Techniques for Deep Learning Image Classification," Proceedings of the 2024 Korea Computer Congress, pp. 792-794, 2024.
- [3] J. H. Kim and Y. T. Shin, "A Study on the Age Prediction Model Using ResNet," Proceedings of ASK 2024 Conference, Vol. 31, No. 1, pp. 803-806, 2024.
- [4] M. Hwang and T. Kim, "On-Device Inference Performance Comparison of EfficientNet and MobileNet," Proceedings of the 2024 Korea Telecommunications Conference, pp. 1153, 2024.
- [5] H. Soo Jung, "A Review and Comparison of Regression Models in Machine Learning for Drug Response Prediction Using Large-scale Public Dataset," Master's Thesis, Chungbuk National University, pp. 42, 2024.
- [6] W. Zhu, "Statistical Parameters for Assessing Environmental Model Performance Related to Sample Size: Case Study in Ocean Color Remote Sensing," Master's Thesis, Zhejiang University, pp. 42, 2024.