

CNN 알고리즘을 이용한 나이와 성별 구분 모델

신성한^o, 전홍석^{*}

^o경북대학교 컴퓨터학부,

^{*}간국대학교 컴퓨터공학과

e-mail: crossrunway01@knu.ac.kr^o, hsjeon@kku.ac.kr^{*}

Age and gender prediction model using CNN

Sung Han Shin^o, Heung Seok Jeon^{*}

^oDept. of Computer Science, Kyungpook-National University,

^{*}Dept. of Computer Engineering, Konkuk University

● 요약 ●

본 논문에서는 딥러닝 CNN 알고리즘을 이용하여 사람의 얼굴 이미지를 학습한 다음 나이와 성별을 예측하는 시스템을 제안한다. 이 시스템은 개개인마다 각기 다른 외형적 특성을 고려하여 이를 분석한 다음 이에 맞는 헤어 스타일, 옷차림을 추천할 수 있다. 해당 기술을 활용하여 메타버스 이바타 생성에 사용자의 얼굴과 같은 신체적 특성을 고려할 수 있다. 향후에는 신체 전체를 이미지화하여 보다 더 다양한 정보를 인식할 수 있도록 연구를 진행할 것이다.

키워드: 인공지능(artificial intelligence), 딥러닝(deep learning), 컴퓨터비전(computer vision), CNN(Convolution Neural Network)

I. Introduction

최근 다양한 딥러닝 알고리즘의 등장과 컴퓨터의 고성능화로 다수의 이미지 데이터를 빠른 시간 안에 정확히 분석할 수 있게 되면서 컴퓨터 비전을 활용한 다양한 사례가 등장하고 있다. 또한 메타버스의 등장으로 개인 이바타를 만들어 로블록스, 이프랜드 등의 가상 세계에서 각자의 개성을 뽐낼 수 있는 기회가 많아졌다. 실생활 중에서도 얼굴의 피부 톤을 몇 가지의 톤으로 구분하여 이에 맞는 헤어스타일, 옷의 색 조합, 화장법을 정할 수 있는 퍼스널 컬러가 각광받고 있다.

이렇듯 본인의 얼굴을 정확히 파악하는 것에 대한 관심이 높아지면서 신체 특성 분석에 인공지능을 도입하는 사례가 늘고 있다.

본 논문에서는 다양한 신체적 특성을 분석하기에 앞서 가장 기본적인 신체적 특성인 나이를 예측(Prediction)하고 성별을 분류(Classification)하는 시스템을 제안하며 이는 사람의 얼굴 사진으로 이루어진 dataset을 기반으로 학습하는 CNN 알고리즘을 이용한다.

이 시스템을 이용해 사용자가 개인 이바타를 생성하거나, 옷, 헤어스타일을 고를 때 언제 어디서나 빠르게 본인의 특성을 고려할 수 있게 된다. 향후 더 다양한 신체적 특성을 분석할 수 있는 모델을 추가하게 된다면 이 시스템은 개인 신체정보 구축에 유용하게 쓰일 수 있다.

II. Preliminaries

1. Related works

현재까지 컴퓨터 비전을 이용하여 나이와 성별 예측을 연구하며 이를 다양한 방면에 적용한 사례들이 있다.

얼굴 이미지를 통해 성별과 나이에 측정에 있어 컬러 이미지와 흑백 이미지를 비교한 사례가 있고, Google LeNet을 이용하여 나이와 성별을 예측하기 위한 그룹화 된 시스템을 제안하기도 하였다[1,3]. R-CNN에 의해 연령대 제한을 주고 GAN을 통해 얼굴의 노화 정도를 예측하는 연구사례도 있었다. 이 경우는 생성 모델을 통하여 얼굴의 노화를 예측하는 연구를 하였다[5]. CNN과 ResNet을 사용하여 실시간으로 성별과 나이를 예측하는 연구[6], 세 가지 CNN 모델을 설계한 뒤 나이와 성별 예측에 있어 정확도와 성능을 비교한 연구사례도 있다. 이 경우는 CNN 모델의 Layer의 수에 변화를 주며 어떤 차이를 보이는지 분석하였다[7]. 헤어스타일, 인종 등에 따른 다양한 요인으로 얼굴 인식에 영향을 주는 요소들을 최소화하기 위해 Multibranch CNN 알고리즘을 나이와 성별 인식에 적용한 사례도 있었다[8]. CNN을 이용한 얼굴 및 신체 인식을 직접적으로 비즈니스 모델에 적용한 경우도 있었다. 컴퓨터 비전을 활용하여 아이의 전신사진을 분석, 키를 측정하는 시스템을 애플리케이션 형태로 제안하기도 하였으며[4], CNN을 통해 실루엣 이미지를 분석하여 범죄 수사에

적용한 사례도 있었다[2].

III. The Proposed Scheme

1. 모델의 구성

본 논문에서 제안하는 시스템은 나이와 성별을 예측하는 총 두 가지 이미지 처리 모델을 포함하고 있다. 나이를 예측하는 모델은 다양한 나이대의 Label을 가지고 있으므로 예측 값이 연속적인 Regression 모델이다. 성별을 구분하는 모델은 Female, Male 두 가지를 구분하는 Classification 문제이다. 두 가지 다른 특성을 가진 예측 값을 갖기 때문에 나이와 성별을 각각 예측하는 모델을 따로 두었고 각자 별개의 Dataset으로 학습 및 검증을 한 뒤, 한 개의 Test 얼굴 사진으로 동시에 나이와 성별을 예측 할 수 있도록 했다.

2. 나이 예측 모델

2.1 Dataset

나이 예측 모델 학습에 사용된 Dataset은 Kaggle의 “Age Prediction”이다. 해당 Data는 Test와 Train 데이터로 나뉘어 있으며 각각 20살부터 50살 까지 남녀의 얼굴 사진이 274,000장 가량 저장되어 있다.

나이 예측 모델의 경우 Classification 문제가 아닌 다양한 나이대의 Label을 가진 Data를 기반으로 학습 및 검증을 하는 Regression 문제이기 때문에 예측결과가 연속성을 지닌다. 따라서 단순한 Accuracy로 정확도를 판단하기에 무리가 있어 평균 제곱근 오차 (Root Mean Squared Error)를 통해 정확도를 측정한다.

2.2 모델의 구조

나이 예측은 총 9개의 Layer로 이루어진 CNN 모델을 통해 이루어진다. (Fig. 1) 120X120 크기의 3차원 입력을 받는다. Convolution Layer에서 16개의 필터와 3X3 크기의 커널, ReLU 활성화 함수를 사용하며 Max Pooling Layer에서 2X2 크기의 풀링을 적용한다. Global Average Pooling Layer에서는 입력에 대해 평균값을 계산하여 특징 맵의 평균값을 사용하기 위해 공간 차원을 제거한다. Dense Layer에서는 64개의 뉴런을 가지며, 활성화 함수로 ReLU를 사용한다. Output Layer는 선형 활성화 함수를 사용하며, 단일 뉴런을 가진다. 앞서 설명했듯이 이 모델은 Regression 문제를 다루기 때문에 선형 활성화 함수를 사용하였다.

3. 성별 구분 모델

3.1 Dataset

나이 예측 모델 학습에 사용된 Dataset은 Kaggle의 “Gender Dataset”이다. 해당 Data는 사람의 얼굴 사진이 Test와 Train, Validation 데이터로 나뉘어 있고 각각 Female, Male로 분류되어 학습 및 검증에 활용된다. 203,000개 가량의 사진으로 구성되어

있다.

Dataset의 디렉토리에 의해 분류된 성별 Label을 기반으로 모델이 학습 및 검증 작업을 하게 된다.

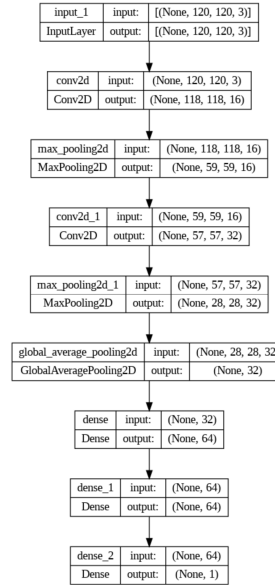


Fig. 1. Age Prediction model

3.2 VGG-16

VGG-16은 CNN 알고리즘의 일종으로 옥스퍼드 대학교의 연구팀인 Visual Geometry Group에서 개발한 것으로 16개의 합성곱 (Convolution) 레이어로 구성되어있다. 모든 합성곱 레이어에는 3x3 크기의 커널과 2 스트라이드로 이루어져있고, 스트라이드는 깊은 네트워크 구조를 가지면서도 작은 크기의 특징맵을 유지할 수 있도록 한다. ImageNet 데이터셋에 대해 224x224 크기의 입력 이미지를 처리하는 데 사용되고 대규모 Dataset에서 사전 학습된 가중치를 사용할 수 있다. 특히 이를 이용하여 다른 컴퓨터 비전 작업에 대해 전이 학습(transfer learning)을 수행할 수 있다.

VGG-16은 깊은 네트워크 구조를 가지고 있어 이미지 분류 작업에 유용하다. 따라서 female, male 두 가지로 이미지를 분류하는 성별 구분 모델의 정확도를 향상시키기 위해 기존의 CNN Layer에 VGG-16을 활용하였다.

3.3 모델의 구조

성별 구분 모델의 구조는 Fig. 2와 같이 사전에 학습된 VGG-16 모델을 기반으로 새로운 Layer를 추가하여 구성되었다. Dropout Layer를 통해 과적합을 방지하고 BatchNormalization을 통해 입력 데이터를 정규화 하였다. Max Pooling Layer에서 2X2 크기의 풀링과 2 스트라이드를 적용한 다음, Flatten Layer를 통해 3차원 데이터를 1차원으로 평탄화하였다. Dropout을 한 번 더 적용하여 일부 뉴런을 꺼줌으로써 과적합을 다시 한번 방지하였다. Output Layer에서는 다중 분류 문제의 해당 모델의 특성에 맞게 sigmoid 활성화 함수를 사용하였다.

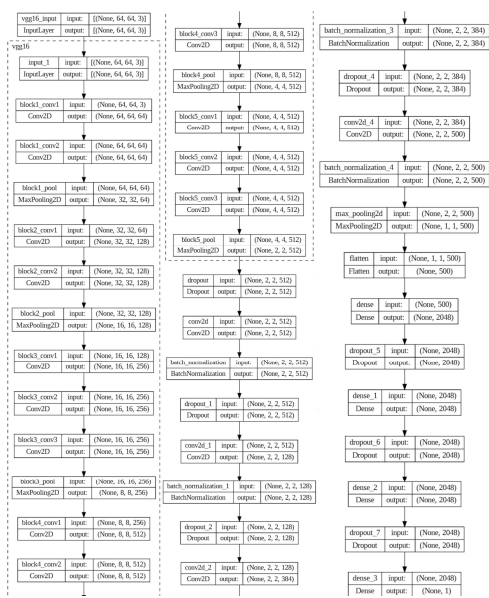


Fig. 2. Gender Detection model

3.4 실험 환경

해당 모델은 Table 1에서 보여주는 것처럼 Google Colab 환경에서 GPU 런타임으로 실험하였다.

Table 1. System Environment

Item	Value
Google Colab GPU	T4
Python	3.9.16

4. 결과

본 논문에서 제시한 모델은 나이 예측 모델과 성별 분류 모델이 각자 학습되어 하나의 이미지에 대한 분석을 하는 구조로, 정확도 역시 모델별로 측정 하였다.

나이 예측 모델의 경우 Regression 모델로 Accuracy가 아닌 평균 제곱근 오차를 구하였고 6.997의 RSME값이 도출되었다. Loss값의 경우 Fig. 4에서 보여주는 것처럼 Epoch수에 따라 점차 하향했다.

성별 분류 모델의 경우 Classification 모델로 Accuracy를 직접 도출 할 수 있었고 Fig. 3에서 보여주는 것처럼 92.66%의 정확도를 보였다. 성별 분류 모델은 VGG-16의 적용으로 보다 더 깊은 신경망을 도입하여 90%가 넘는 Accuracy를 보여주어, 실제 test 사진을 이용하여 예측 결과의 정확도가 높았다. 나이 예측 모델의 경우에는 Fig. 5. (a)의 경우와 같이 성별 분류 모델보다는 다소 아쉬운 오차를 보여주었지만 같은 나이더라도 각자 생각하는 외향의 나이가 다를 수 있는 data의 특성 상 5-6살 내외의 차이가 사용하기 불편한 수준의 오차는 아닐 것이다. 다만 dataset이 서양인 위주로 이루어져있어 동양인의 경우 Fig. 5. (b)의 경우처럼 10살 이상의 오차를 보이는 경우가 잦았다. 이는 보다 더 다양한 dataset을 확보하고 모델을 개선한다면 더 나은 결과를 얻을 것으로 기대된다.

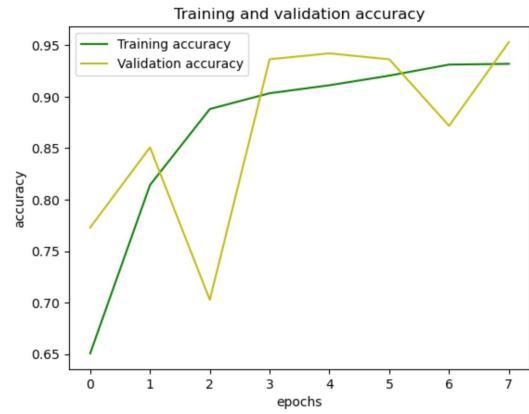


Fig. 3. Gender prediction accuracy graph

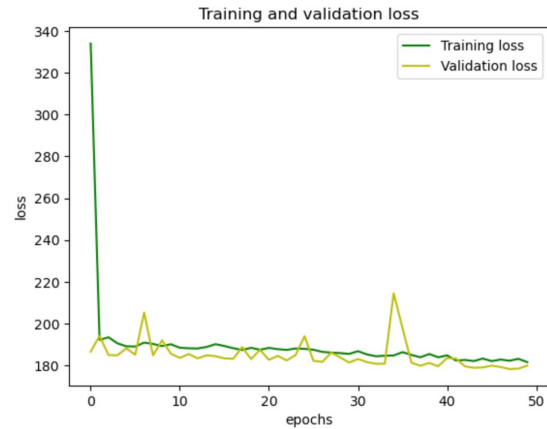


Fig. 4. Age prediction loss graph

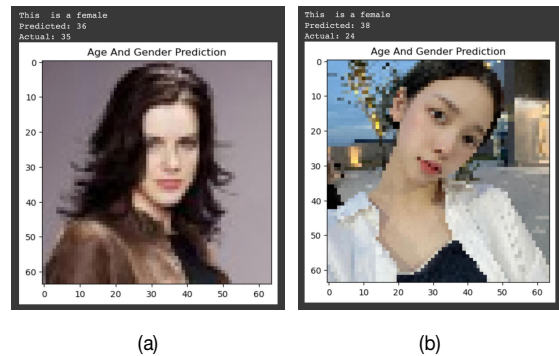


Fig. 5. Age and Gender prediction test

IV. Conclusions

본 논문에서는 CNN 알고리즘을 통해 사람의 얼굴 사진을 분석하여 나이를 성별을 예측 할 수 있는 모델을 제안하였다. 나이 예측 모델은 6.997의 RSME값을, 성별 분류 모델은 92.66%의 Accuracy를 보였다. 나이 예측모델은 이진 분류인 만큼 비교적 정확한 분류를 수행했고 성별 분류의 경우 다소 아쉬운 수치를 보여주었지만 원래 나이에 근사한 수치를 예측하였다.

최근 메타버스의 등장으로 많은 사람들이 자신을 꼭 빼닮은 개인 아바타를 생성하는 경우가 많고 퍼스널 컬러를 분석하는 등 같이 본인의 신체적 특성을 분석하는 사례가 많다. 이때 컴퓨터 비전을 활용하여 스마트폰 카메라를 이용하는 등 누구나 쉽게 어디서나 본인의 얼굴과 신체를 분석 할 수 있다면 개인 아바타 생성 혹은 화장법, 헤어스타일, 착장을 찾는 데 큰 도움이 될 것으로 기대한다.

향후 연구로는 나이 예측 모델의 성능을 강화 하고 보다 더 다양한 신체적 특성을 분석할 수 있는 추가적인 모델을 결합해 보고자 한다.

[8] Haris Setiawan and Mudrik Alaydrus and Abdi Wahab, "Multibranch Convolutional Neural Network For Gender And Age Identification Using Multiclass Classification And FaceNet Model", 2022 Seventh International Conference on Informatics and Computing (ICIC), 13 January 2023.

REFERENCES

- [1] Kyoungson Jhang and Junsoo Cho, "CNN Training for Face Photo based Gender and Age Group Prediction with Camera", 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), 21 March 2019.
- [2] Jong Won Son and Jong Hyeon Kim and Jae Hyun Kim, "Age and Gender Classification of Silhouette Image using Convolution Neural Network", Journal of The Korea Society of Computer and Information, VOL 67 NO. 01 PP. 0393 ~ 0394, November 2018.
- [3] Xuan Liu and Junbao Li and Cong Hu and Jeng-Shyang Pan, "Deep Convolutional Neural Networks-Based Age and Gender Classification with Facial Images", 2017 First International Conference on Electronics Instrumentation & Information Systems (EIIS), 22 February 2018.
- [4] Da-Yeong Yun and Mi-Kyeong Moon, "Development of Kid Height Measurement Application based on Image using Computer Vision", Journal of the KIECS. pp. 117-124, vol. 16, no. 1, Feb. 28. 2021.
- [5] Mengjiao Sheng and Zhongchen Ma and Hongjie Jia and Qirong Mao, "Face Aging with Conditional Generative Adversarial Network Guided by Ranking-CNN", 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), 25 August 2020.
- [6] Asad Mustafa and Kevin Meehan, "Gender Classification and Age Prediction using CNN and ResNet in Real-Time", 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy ((ICDABI), 20 January 2021
- [7] Mohammed Kamel Benkaddour and Sara Lahlali and Maroua Trabelsi, "Human Age And Gender Classification using Convolutional Neural Network", 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH), 19 March 2021.