

# 딥러닝 기반 성별 및 나이 추정 기법과 서비스 분석<sup>1</sup>

김양은<sup>0</sup>, 최유주\*

서울미디어대학원대학교 인공지능 응용소프트웨어학과

yang7319@naver.com, yjchoi@smit.ac.kr

## Analysis of deep learning-based gender and age estimation techniques and services

Yang-Eun Kim<sup>0</sup>, Yoo-joo Choi\*

AI Software Engineering, Seoul Media Institute of Technology University

\*Corresponding Author

### 요 약

딥러닝 기반 성별 및 연령 추정은 현재 다양한 응용분야에 적용하기 위하여 활발히 연구되어 왔다. 최근 몇몇 기업에서는 개인정보 보호의 문제로 성별 추정 서비스를 중단하고 있기는 하나, 사용자 및 관람객의 특성 분석등을 위하여 그 필요성이 계속하여 대두되고 있는 상황이다. 본 논문은 기존에 성별 및 나이 추정을 위한 유료서비스에 대한 특성과 사용 문제점 등을 세분화하여 조사하고 성별 및 연령 추정 기반 기존 연구들의 특성을 분석한다. 추후 이 서비스를 기반으로 전시·컨벤션 산업에서의 관심도 분석 및 재방문 의사를 파악하는데 활용하고자 한다.

### 1. 서론

나이와 성별을 추정하는 논문은 AI 분야에서 다양하게 사용하고 있는 서비스이다. 현재 딥러닝 기반 성별 분류 및 나이 추정은 가장 활발히 연구되고 있다. 하지만, 최근에는 얼굴 보안 문제로 성별을 식별하는 기능을 폐지하는 유료서비스들이 늘어나고 있다. 사생활 침해 문제가 발생하고, 나이, 성별, 지역, 상황 등에 따라 표정에 대한 해석들이 정규화되어 있지 않아서 일반화하기 어렵다는 문제가 제기되고 있다. 성별 및 나이 추정은 다양한 분야에서 사용되고 있으며 주로 영상처리 및 컴퓨터 비전 분야에서 다루고 있다.

2 절에서는 현재 시중에서 이용되고 있는 유료 서비스의 특징 등을 세분화하여 분석하고, 3 절에서는 성별 추정에 관련된 대표적인 방법들의 특성을 분석하고자 한다. 대상 연구는 RESNET 모델을 이용하여 연령 추정하는 방법, RCNN 을 사용하여 성별과 나이를 실시간으로 분석한 방법, 2022 년도 KNN 을 기반으로 새로운 rank 를 측정하여 나이를 분석한 방법 등이다.

### 2. 성별/나이추정 유료 서비스의 종류

#### 2.1 성별 추정 [1]

성별 분류는 개인의 성별을 생체 신호에 따라 결정하는데 일반적으로 얼굴 이미지를 이용하여 특징을 추출하고, 추출된 특징에 분류기를 적용하여 성별 인식을 학습한다. 즉, 성별 추론 방법은 이미지에서 얼굴의 정보를 추출하여, 얼굴의 위치, 특징, 표정 분석, 성별, 나이 등을 추정하여 도출해낸다.

#### 2.2 어플리케이션의 예

최근 수요가 높아지면서 서비스 분석에 대한 비용이 높아지고 있다. 비용에 대한 부담이 높아지면서 특정 이미지 수까지는 무료로 제공해주고 추가되는 부분을 책정하여 지원해주는 유료서비스들이 많아지고 있는 추세다. 다양한 유료서비스들이 늘어나면서 기업들의 부담을 덜어줄 수 있다.

본 논문에서는 총 8 개의 유료서비스를 조사하였으며, 각각 어떤 서비스를 진행하고 있는지에 대한 특징을 정리했으며, 서비스 가격에 대해서 정리하였다. 정리한 내용은 아래 <표 1>과 같다.

<sup>1</sup> 본 연구는 문화체육관광부 “관광서비스 혁신성장 연구개발사업”(R2022020105)의 지원에 의하여 수행되었음

<표 1> 유료서비스 종류 및 가격, 특징

서비스 종류	유료서비스 가격	유료서비스 특징
Google	매달 1,000 개 단위 무료, 그 이후는 가격대로 책정(추가비용 \$1.5)	현재는 얼굴 인식 지원하지 않음(성별 구분 불가능)
Microsoft	매달 30,000 개 무료, 그 이후로 매월 1,000 개 얼굴 당 \$0.01 개	보안 문제로 사용을 폐지한 상태(성별 구분 불가능)
Amazon	처음 백만 개의 이미지는 \$0.0012	사람을 감지하고 인지하고 식별
Clarifai	월 1,000 회 무료, 기업용은 \$30 부터 추가됨에 따라 가격 측정	얼굴 특징을 찾아 성별 분석 가능
IBM Watson Visual Recognition	서비스 사용 안함	현재는 얼굴 인식 기술의 문제점이 지적되면서 서비스를 포기한 상태
Kairos	평가 기간 동안 무료 거래 수 20,000 건 매달 \$19, \$0.02	사진이나 동영상에서 얼굴을 감지하고 감정 및 나이, 성별, 특징 감지 가능
Face me	가격 문의	연령 및 성별, 감정 파악 가능
Clova	약관에 따라 무료로 제공되고 있음	얼굴 인식 감지 및 닮은 유명한 얼굴 인식 기능

### 2.3 서비스 문제점

유료 서비스를 제공하는 과정에서 사생활 침해 문제가 발생하는 것 이외에도 나이, 성별, 지역에 따라 해석이 제각각 이므로 결과를 일반화하기 어렵다는 문제가 제기되었다. 현재 8 개 서비스 중 GOOGLE, MICROSOFT, IBM Watson Visual Recognition 은 성별 추정 서비스는 제공하고 있지 않다. Google 은 성별 구분 대신 person 이라고 개체를 통합하여 인식하고 있으며 Microsoft 와 IBM 은 서비스 자체를 중단한 상태이다.

개인정보 보호 등의 문제로 얼굴영상 기반 성별/나이 추정 서비스가 중단되고 있는 상황이지만, 방문 고객, 관람객의 특성을 설문 등과 같은 불편한 방법이 아닌 자동 인식의 필요성은 계속적으로 제기되고 있다. 이에 분석 영상의 저장 및 유출의 문제가 없는 시스템의 설계를 통한 문제점 보완이 요구되고 있다.

### 3. 성별 및 나이 추정 방법

2020 년도 CNN 모델인 RESNET 모델을 이용하여 연령 추정하는 방법, 2021 년도 RCNN 을 사용하여 성별과 나이를 실시간으로 분석한 방법, 2022 년도 KNN 을 기반으로 새로운 rank 를 측정하여 나이를 분석한 방법 등이 발표되었다.

### 3.1 ResNet 을 이용한 성별 및 나이 추정 [2]

컨볼루션 신경망(CNN)을 사용한 방법으로 얼굴 이미지에서 개인의 나이를 예측하기 위한 데이터 세트의 접근 방식으로 경험적 분석을 제시하였다.

Morph-2 데이터 세트와 CACD 데이터 세트를 사용하여 CNN 모델인 RESNET-34 모델을 사용하여 나이추정을 수행하였다. 모든 손실함수(loss function)과 뉴럴 네트워크 모델은 pytorch1.5로 구현하였다.

실험 결과는, 서수 회귀 CNN 예측 성능이 긍정적인 영향을 미쳤으며, CE-CNN, Coral-CNN 및 OR-CNN 경우, Morph-2>Afad>CACD 순으로 Morph-2가 가장 좋았다. <표 2>는 실험결과를 보여주고 있다.

<표 2> Coral-CNN 및 Niu et al의 서수 CNN에 대한 다른 테스트 데이터 세트에서 평균 불일치 수

	CORAL-CNN All predictions	OR-CNN (Niu et al., 2016) All predictions	OR-CNN (Niu et al., 2016) Only correct predictions	OR-CNN (Niu et al., 2016) Only correct predictions
<b>MORPH</b>				
Seed 0	0	2.28	1.80	2.37
Seed 1	0	2.08	1.70	2.15
Seed 2	0	0.86	0.65	0.89
<b>AFAD</b>				
Seed 0	0	1.97	1.88	1.98
Seed 1	0	1.91	1.81	1.92
Seed 2	0	1.17	1.02	1.19
<b>CACD</b>				
Seed 0	0	1.24	0.98	1.26
Seed 1	0	1.68	1.29	1.71
Seed 2	0	0.80	0.63	0.81

### 3.2 RCNN 을 이용한 실시간 성별 및 나이 추정 [3]

이 방법은 실시간 안면 분석 시스템의 시스템 수준 설계를 제시하여 객체를 감지하고 분류 및 회귀를 위한 심층 신경망 모음을 통해 사람의 연령, 성별, 얼굴 표정 및 얼굴 유사성을 인식하고 있다.

성별 및 나이 추정에 대한 데이터 세트는 500k IMDB-WIKI 와 CVPR2016 LAP 챌린지 데이터를 사용하였으며 ResNet101 및 MobileNet 과 함께 RCNN 을 사용하였다.

연구 결과는 가장 성능이 좋은 멀티태스킹 네트워크는 5.35 세 MAE 를 제공하고 가장 성능이 좋은 개별 연령 네트워크보다 약간 높았다. EfficientNet 는 더 나은 나이 MAE 및 인식 정확도를 제공하지만 계산 비용이 많이 드는 단점이 있다. [3]방법의 실험결과는 <표 3>, <표 4>와 같다.

<표 3>시스템의 여러 단계에서 정확도 및 추론 속도

Stage Network	Accuracy	FPS CPU	FPS 1050 TI	FPS 1080 TI
Age MobileNetV1	4.9MAE	31.61	148.90	147.44
Gender MobileNetV1	88.3%	31.48	150.45	149.75
Smile MobileNet1	87.2%	31.46	148.84	148.78
Multitask MobileNet1	5.67 MAE 84.2% Gender 83.6% Smile	29.80	147.06	147.20
Multitask EfficientNetB0	5.35 MAE 87.5% Gender 86.0% Smile	25.61	143.35	144.24

Paleo et al.	-	-	44.9	0.93
Red-SVM	-	-	35.9	0.96
Martin et al.	-	-	42.8	0.87
OR-CNN	56.7	0.54	38.7	0.95
CNNFOR	57.4	0.55	50.1	0.82
GP-DNNOR	57.4	0.54	46.6	0.76
SORD	59.6	0.49	-	-
DRC-ORID	-	-	44.7	0.80
Li et al	<u>60.5</u>	<u>0.47</u>	<u>54.7</u>	<u>0.66</u>
Proposed	<b>62.6</b>	<b>0.45</b>	<b>57.8</b>	<b>0.58</b>

4. 결론

본 논문에서는 딥러닝 기반 성별 분류와 나이를 추정하는 서비스에 대한 내용을 정리하였다. 성별 추정에 대한 정의를 하였고, 유료 서비스에 대해서 정리하였다. 현재 유료 서비스 중에서는 Google, Microsoft, IBM 는 성별 추정 서비스를 지원하지 않았다. 또한, 성별 및 나이 추정 논문을 분석했다. 하지만 성별만 따로 나와있는 논문은 없고 대부분은 나이를 중심으로 추정하는 논문이었다.

향후 논문을 바탕으로 성별과 나이 추정 네트워크를 구현하여 보다 높은 정확도를 향상시킬 수 있는 연구가 요구된다. 또한, 다양한 연령대와 동양인에 대한 데이터도 수집을 하여 추정 정확도를 높일 수 있는 알고리즘에 대한 연구가 요구된다. [5]

<표 4> 다양한 백본 네트워크와 멀티태스킹 네트워크 성능비교

Network	Age MAE	Gender ACC(%)	Smile ACC(%)	FPS CPU
VGG16	7.20	84.0	84.1	27.75
ResNet50	6.42	82.1	81.2	27.06
ResNet18	6.02	82.4	82.8	29.63
MobileNetV1	5.67	84.2	83.6	<b>29.80</b>
EfficientNet B0	5.35	87.5	86.0	25.61
EfficientNet B1	5.07	87.8	86.8	22.72
EfficientNet B7	<b>4.37</b>	<b>89.5</b>	<b>87.3</b>	14.63

3.3 KNN 을 기반으로 새로운 rank 를 측정하여 나이 분석 [4]

이 방법은 Moving Window Regression 이라는 새로운 서수 회귀 알고리즘을 통해 references 보다 얼마나 작고 큰지를 비교하여 상대적인 rank 를 측정하는 방법을 적용하고 있다.

ILSVRC2012 에서 각각의 p-regressor 를 훈련시키기 위해 VGG16 을 인코더로 사용하였고, 연령 추정에는 MORPH2, FG-NET 등 총 7 개의 데이터 세트를 사용하였다. UTK 데이터 세트를 사용하여 연령, 성별 및 민족에 대한 얼굴 이미지로 구성하였다. 연령대 추정에 사용되는 Adience 의 결과를 비교한 내용은 아래 <표 5>와 같다. 기존 알고리즘보다 성능이 우수하게 나왔다.

<표 5> Adience 및 HCI 데이터 세트에 대한 정확도 및 MAE

Algorithm	Adience		HCI	
	Accuracy (%)	MAE	Accuracy (%)	MAE
Frank & Hall	-	-	41.4	0.99
Cardoso et al.	-	-	41.3	0.95

참고문헌

[1] [https://link.springer.com/referenceworkentry/10.1007/978-1-4899-7488-4\\_9176](https://link.springer.com/referenceworkentry/10.1007/978-1-4899-7488-4_9176)

[2] Wenzhi Cao, Vahid Mirjalilib, Sebastian Raschka, Rank consistent ordinal regression for neural networks with application to age estimation, arXiv:1901.07884v7[cs.LG], Pattern Recognition Letters, 2020, p.7

[3] Bishwo Adhikari Xingyang Ni, Esa Rahtu, Heikki Huttunen, Towards a Real-Time Facial Analysis System, arXiv:2109.10393v1[cs.CV], 2021

[4] Nyeong-Ho Shin, Seon-Ho Lee, Chang-Su Kim, Moving Window Regression: A Novel Approach to Ordinal Regression, arXiv:2203.13122v1[cs.CV], 2022

[5] 권나현, 최하영, 박상하, 민동보, 서오석, 한국소프트웨어종합학술대회 논문집, 한국정보처리학회, 2018, p.2305