

공학사 학위논문

얼굴 인식 로그인 시스템 연구

전남대학교 AI융합대학

IoT인공지능융합전공

이재경, 허정화

2020년 7월

얼굴 인식 로그인 시스템 연구

이 논문을 공학사 학위 논문으로 제출함

전남대학교 AI융합대학

IoT인공지능융합전공

이재경, 허정화

지도교수 김수형

이재경, 허정화의 공학사 학위논문을 인준함

전공주임 양형정 (인)

지도교수 김수형 (인)

2020년 7월

목차

(국문초록)	1
(본문)	2
1. 서론	2
2. 관련 연구	3
가. MTCNN	3
나. FaceNet	4
다. 웹 크롤링	4
3. 제안 알고리즘	5
가. 데이터(한국인 안면 이미지 AI데이터)	5
나. MTCNN을 이용한 얼굴 추출 & Facenet을 이용한 벡터화	5
다. SVM을 이용한 분류	6
라. SoftMax를 이용한 분류	6
4. 실험 및 결과	8
가. 성능평가	8
1) MTCNN 성능평가	8
2) SVM(OvO 여러 번 사용)을 이용한 분류 성능평가	10
3) SoftMax를 이용한 분류 성능평가	10
4) 최종 모델 성능평가	11
나. 결과	15
5. 결론	17
참고문헌	17

얼굴 인식 로그인 시스템 연구

이 재 경 , 허 정 화

전남대학교 AI융합대학 IoT인공지능융합전공

(지도교수 : 김수형)

(국문초록)

오늘날 4차 산업혁명이란 인공지능(AI), 사물인터넷(IoT), 로봇 기술, 드론, 자율주행차, 가상현실(VR) 등이 주도하는 차세대 산업혁명을 말한다. 이 때 핵심 개념의 인공지능(AI, Artificial Intelligence)은 인간의 인식 판단, 추론, 문제해결, 언어나 행동 지령, 학습 기능과 같이 컴퓨터 스스로 추론·학습·판단하면서 작업하는 시스템을 가리키며 이는 4차 산업혁명에서 큰 부분을 차지한다. 하지만 현재 대부분의 대학교의 포털 시스템은 인공지능과 거리가 멀다. 따라서 학교도 시대에 발맞춘 사용자의 편리함을 위한 시스템의 인공지능화가 필요하다.

본 논문에서는 사용자의 편리함을 위한 시스템의 인공지능화를 위해서 얼굴 인식을 이용한 로그인 시스템을 제안한다. 기존의 얼굴 인식 시스템은 대부분 OvO(One versus One) 매칭으로 개인이 이용하는 핸드폰, 노트북에 적합하게 사용되고 있다. 하지만 본 연구에서는 OvO 매칭이 아닌 얼굴 인식에 필요한 모델을 이용하여 여러 사람이 학습되어 있는 환경에서 한 사람을 감지할 수 있는 OvA(One versus All) 매칭 방법을 소개한다. 그리고 이러한 방법을 이용하여 사용자의 얼굴을 인식 했을 때 매칭이 잘 되는지 확인하고 이러한 방식이 얼굴 인식을 이용한 시스템과 적합한지 평가할 것이다.

(본문)

1. 서론

오늘날 4차 산업혁명이란 인공지능(AI), 사물인터넷(IoT), 로봇 기술, 드론, 자율주행차, 가상현실(VR) 등이 주도하는 차세대 산업혁명을 말한다. 이 때 핵심 개념의 인공지능(AI, Artificial Intelligence)은 인간의 인식 판단, 추론, 문제해결, 언어나 행동 지령, 학습 기능과 같이 컴퓨터 스스로 추론·학습·판단하면서 작업하는 시스템을 가리키며 이는 4차 산업혁명에서 큰 부분을 차지한다.

최근 사용자 인증 시스템에서도 인공지능을 활용한 인증방식이 시행되고 있다. 단일 요소 인증인 아이디 인증 대신 스마트폰에 내장된 하드웨어를 사용해 지문 인식, 홍채 인식, 얼굴 인식 등 인공지능을 활용한 생물학적 사용자 인증을 시행하고 있다. 특히 비접촉식 얼굴 인식 기술이 제공하는 사용자 편의성과 관리의 효율성으로 인해 그 수요가 증가하고 있으며 현재 아이디 인증을 사용하는 대학교의 포털 시스템에도 인공지능을 이용한 얼굴 인식 기술의 도입을 위해 본 연구를 진행하였다.

얼굴 인식 기술은 딥러닝 알고리즘 활성화 전 hand-crafted feature 기술이 사용되었으며 이는 얼굴의 고유한 랜드마크(landmark)를 검출하고 정규화 등의 전처리를 거친 후 특징점을 추출해 얼굴을 인식한다. 하지만 딥러닝 중심으로 기계학습이 발달함에 따라 얼굴의 고유한 특징을 더욱 효과적으로 추출할 수 있다고 알려진 Convolutional Neural Network를 기반으로 기술들이 연구되고 있으며 속도와 정확도 측면에서 이전 기술과는 비교할 수 없을 우수한 성능을 보여준다. 본 연구에서는 한국인 안면 이미지 AI데이터와 이전에 학습된 합성곱 신경망(CNN)들을 활용하여 더욱 안전성 있는 얼굴 인식 로그인 시스템을 연구한다.

구성은 다음과 같다. 2장에서는 얼굴 인식에 많이 활용되는 연구들을 살펴보고 3장에서는 사용하는 데이터와 모델 및 성능평가, 4장에서는 최종 결과, 5장에서는 고찰을 논의한다.

2. 관련 연구

본 연구에서는 얼굴 인식을 위한 방법으로 MTCNN과 FaceNet을 사용하였으며 로그인은 웹 크롤링을 사용했다.

가. MTCNN

MTCNN은 “Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks”이라는 2016년에 발행된 얼굴 인식을 위한 최첨단 딥러닝 모델 논문에 나온 것으로 P-net, R-net, O-net 으로 이루어진 3계층 CNN이다. 얼굴의 랜드마크 위치를 예측하여 찾아내며 제한적 환경에서 인식이 쉽다. 전반적인 구조는 아래와 같다.

먼저 입력 이미지를 단계별로 재조정하여 이미지 피라미드를 만들어준다. 예를 들어 300x200 크기의 이미지가 입력되면 이를 200x166, 100x66, 30x20 크기로 재조정된 이미지의 list를 만들어준다. 이렇게 해주는 이유는 작은 얼굴도 검출하기 위해서이다. 우선 이렇게 만든 이미지 피라미드를 P-net에 통과시킨다.

P-net은 12x12x3 크기의 작은 이미지를 입력받는다. 그리고 해당 영역이 얼굴인지 아닌지 각각 나타내는 face classification, 얼굴 영역을 나타내는 좌측 상단 꼭짓점의 x, y 좌표와 박스의 너비, 크기를 나타내는 4개의 bounding box regression 값, 그리고 양쪽 눈, 코, 양쪽 입꼬리의 x, y 좌표를 나타내는 10개의 landmark localization 값을 결과로 return한다. 그리고 이렇게 찾은 얼굴 영역들을 다시 원래의 이미지 크기로 되돌려준 뒤 Non-Maximum-Suppression과 bounding box regression을 적용해준다. NMS는 동일한 얼굴에 여러 번 박스가 쳐진 경우 가장 얼굴일 확률이 높은 것만 남기고 제거하는 것을 의미한다.

이렇게 P-net을 통해서 얼굴로 추정되는 박스들의 리스트를 얻고 R-net을 통해서 진짜 얼굴에 해당하는 영역들을 추려내고 bounding box regression을 더 정교하게 수행한다. 먼저 박스들을 모두 24x24 크기로 재조정된 뒤 R-net을 통과시킨다. 여기서 P-net이 전체 이미지 중 얼굴에 해당하는 부분을 추측했다면

R-net은 이를 더 정교하게 해준다. R-net에서 찾아낸 박스도 P-net과 마찬가지로 NMS와 BBR을 적용한다. 여기서 살아남은 박스들만 O-net에 전달된다.

O-net은 R-net을 통해 찾아낸 박스들을 모두 48x48 크기로 재조정된 것을 입력으로 받는다. 점점 필터의 크기를 키우면서 얼굴에 해당하는 추상적인 정보를 찾아내기 위한 용도이다. 그 뒤 4개의 Conv layer와 FC layer를 거친 뒤 세 종류의 output을 내게 되면 이것이 최종 Face Detection, Face Alignment의 결과가 된다.

이후 Multi Task Loss 기법을 사용한다. 이는 세 가지 task에 대하여 각각 loss를 구한 뒤 이를 가중치를 두어 합치는 방식을 사용한다. 본 연구에서는 이러한 MTCNN 방식을 이용하여 감지기를 만들어 얼굴 추출을 해준다.

나. FaceNet

FaceNet은 2015년 Google 연구원이 개발한 얼굴 인식 시스템으로 얼굴 그림이 주어지면 얼굴에서 고품질의 특성을 추출하여 이러한 특성을 얼굴 임베딩이라고 하는 128개의 요소 벡터로 표현한다. 이 모델은 Inception ResNet v1을 triplet loss 방식을 통해 훈련한 심층 합성곱 신경망으로, 같은 정체성(Positive)에 대한 벡터가 더 가까워지고 다른 정체성(Negative)에 대한 벡터는 더 멀어진다. 본 연구에서는 히로키 타이아의 케라스 FaceNet을 이용했다. 그의 프로젝트는 Inception ResNet v1 모델을 텐서플로에서 케라스로 변환하기 위한 스크립트를 제공한다. 본 연구에서는 이를 이용하여 FaceNet 모델을 만들고 이를 이용하여 얼굴을 128개의 요소 벡터로 표현하여 임베딩 해준다.

다. 웹 크롤링

얼굴 인식결과와 대응된 포털 아이디와 패스워드를 웹 크롤링 기능을 이용해 포털에 입력하여 로그인한다. 이 때 폼 태그를 통해 ID와 PW가 전달되므로 파라미터에 ID와 PW를 넣어 post 해주는 코드를 이용한다.

3. 제안 알고리즘

가. 데이터(한국인 안면 이미지 AI데이터)

한국인 안면 이미지 데이터는 한국정보화진흥원 (NIA) 지능정보산업 인프라 조성 사업의 세부 과제인 영상 정보 지식베이스 구축 사업의 일환으로 구축되었다. 연구 목적으로 활용되는 얼굴 데이터베이스는 대부분 해외에서 구축됨에 따라, 한국인들의 얼굴 특성에 맞는 한국형 얼굴 데이터베이스가 필요하다. 이에 본 데이터 구축은 한국인 얼굴의 통계적 특징을 파악하기 위함을 목적으로 하고 있다. 한국인 안면 이미지 데이터는 총 1000명의 데이터를 보유하고 있다. 얼굴 데이터의 실효성 제고를 위해 각도 20종, 조도 30종, 액세서리(가림) 6종, 표정 3종, 해상도 3종을 반영하여 인당 약 3만 장에 해당하는 데이터가 구축되었다.

K-FACE 데이터 셋은 'AI open innovation hub (<http://www.aihub.or.kr>)' 사이트의 안면 데이터 부분 혹은 데이터 공유 사이트 (<http://kface.kist.re.kr>) 에서 이용할 수 있다. 사이트 가입이 필요하며 이미지 신청을 하시면 내려받을 수 있다. 이미지 데이터를 내려받기 위해 데이터를 제공하는 연구기관(KIST) 사이트로 이동한다. 이 사이트에서 해상도에 따라 데이터를 요청할 수 있으며, 본 연구에서는 고화질 해상도를 요청해 사용하였고 300명의 데이터를 사용하였다.

나. MTCNN을 이용한 얼굴 추출 & Facenet을 이용한 벡터화

데이터를 학습하기 위해서는 사진에서 얼굴을 추출해야 한다. 여기에는 많은 제약이 따르는데 사용자가 액세서리를 착용했을 수도 있고, 조명이 어두울 수도 있고, 정면을 주시하지 않거나 각도가 틀어져 있을 수도 있다. 이를 위해 수작업으로 불필요한 데이터를 제거하거나 직접 사진을 편집하는 방법도 있지만 본 연구에서는 OpenCV2를 이용해 영상의 프레임을 추출하고 MTCNN을 이용해 프레임 또는 사진에서 얼굴을 추출한다.

추출 후 히로키 타이아이의 FaceNet을 이용해 얼굴을 128개 요소의 벡터로 임베딩 해주었고 본 연구에서는 이 모든 과정을 거친 데이터들을 분류모델에 학습

및 평가를 진행하였다.

아래 사진은 전체적인 모델을 그림으로 시각화하였다. [그림 3.1.]

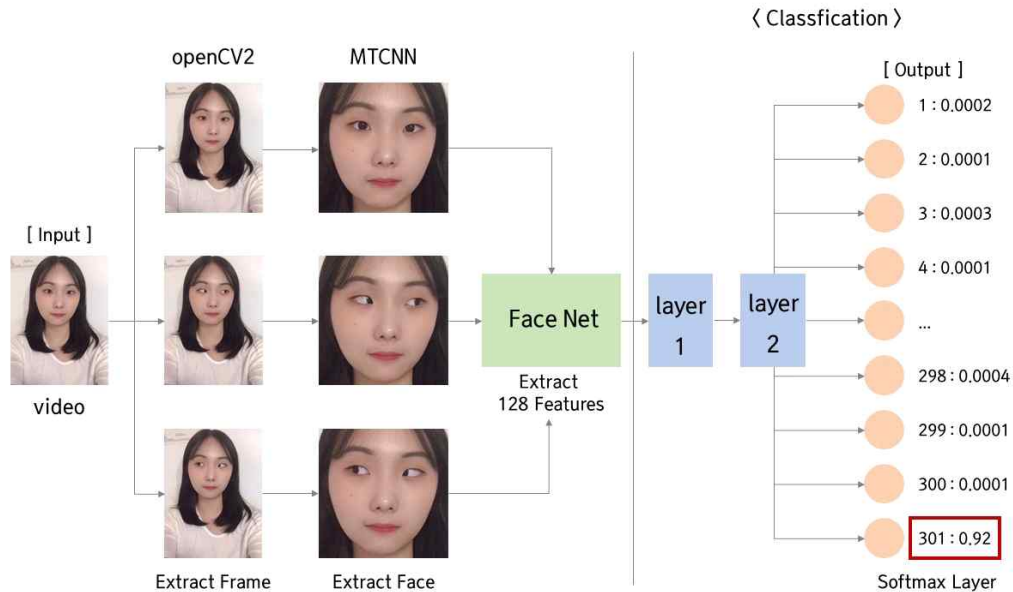


그림 3.1. 전체 모델 구조(깊은 다층퍼셉트론 다중분류기 사용)

다. SVM을 이용한 분류

SVM은 주어진 샘플 그룹에 대해 그룹 분류 규칙을 찾아내는 기법의 하나다. n 차원 데이터 공간에서 $n-1$ 로 차원이 축소된 평면인 하이퍼플레인을 찾는 데 이때 가장 가까운 data와의 거리, 즉 마진이 큰 하이퍼플레인을 찾는다. 또한, SVM은 이진 분류모델이며 이 모델로 다중분류를 할 경우, 이진 분류를 여러 번 진행해 결과를 도출한다. 본 연구에서는 128개로 벡터화된 얼굴 데이터를 누구인지 분류하도록 학습시키고 평가했다.

라. SoftMax를 이용한 분류

SoftMax는 활성화 함수로 입력되는 값을 클래스별로 확률값이 나오도록 출력시킨다. 이 확률값은 모두 더하면 1이 되며, 다중클래스 모델의 출력층에 사용하였다. 확률값이 가장 높은 클래스가 모델이 분류한 클래스이다. 따라서 본 연구에서는 다중클래스 분류를 위해 퍼셉트론 모델, 다층퍼셉트론 모델, 깊은 다층퍼셉

트론 모델을 모델링하였다. 128개로 벡터화된 얼굴 데이터를 누구인지 분류하도록 학습시키고 평가하였다. 이 때 loss는 categorical_crossentropy를 사용하고 optimizer는 SGD를 사용했다.

Layer (type)	Output Shape	Param #
Input (InputLayer)	(None, 128)	0
output (Dense)	(None, 301)	38829
Total params: 38,829		
Trainable params: 38,829		
Non-trainable params: 0		

그림 3.2. 퍼셉트론 모델

Layer (type)	Output Shape	Param #
Input (InputLayer)	(None, 128)	0
layer1 (Dense)	(None, 512)	66048
output (Dense)	(None, 301)	154413
Total params: 220,461		
Trainable params: 220,461		
Non-trainable params: 0		

그림 3.3. 다층 퍼셉트론 모델

Layer (type)	Output Shape	Param #
Input (InputLayer)	(None, 128)	0
layer1 (Dense)	(None, 512)	66048
layer2 (Dense)	(None, 512)	262656
output (Dense)	(None, 301)	154413
Total params: 483,117		
Trainable params: 483,117		
Non-trainable params: 0		

그림 3.4. 깊은 다층 퍼셉트론 모델

4. 실험 및 결과

가. 성능평가

1) MTCNN 성능평가

MTCNN의 성능을 제대로 알기 위한 평가를 하였다. 같은 사람으로 진행하였으며 위에 데이터 셋에서 설명한 것과 같이 각도, 조명, 액세서리, 표정이 다 다른 10,800장을 사용하였다. 결과적으로 각도가 많이 틀어져 있을 때, 밝기가 어두울 때 얼굴을 감지하지 못하였으며 대체로 lux가 40이하, 조명의 수평 방향이 틀어졌을 때, 사진이 옆 또는 위에서 찍힌 경우, 모자를 쓴 경우, 얼굴감지 실패 개수가 확연히 증가한다.



그림 4.1. 얼굴감지 실패 예시

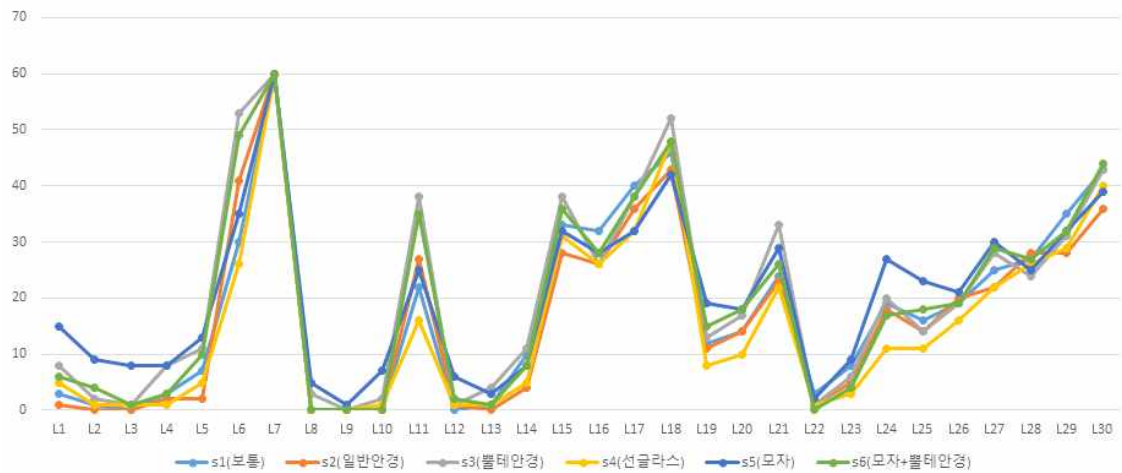


그림 4.2. 밝기에 따른 얼굴감지 실패 개수

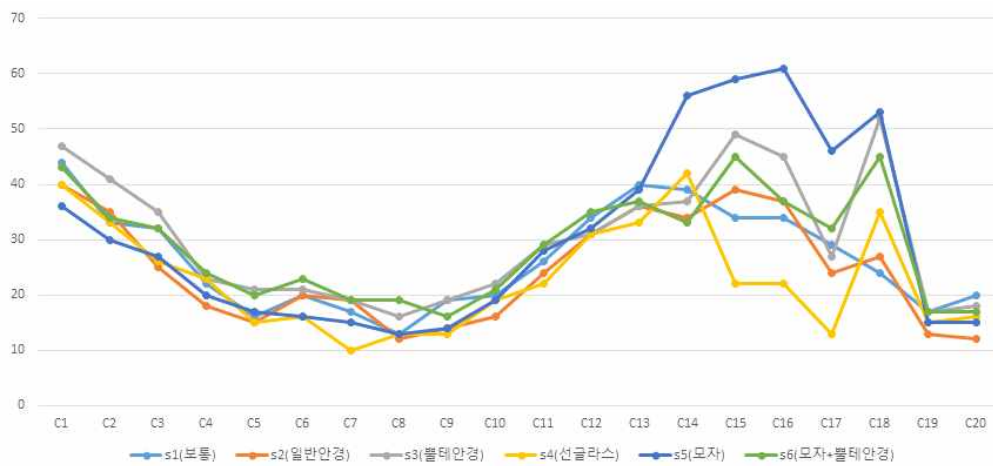


그림 4.3. 각도에 따른 얼굴감지 실패 개수

본 연구에서는 이를 극복하기 위해 사진의 밝기를 올려서 다시 한번 감지하는 형식으로 코드를 수정하였다. 정확히는 RGB값에 100을 더해주었다. 그 결과 1인 10,800장 중 3,277장에 실패하였던 것이 3,213장으로 64장 감소하였다. 나머지 실패 상황들은 애플리케이션에서 사용자에게 정면 응시 및 밝기 조정을 요청하는 형식으로 구현할 예정이다.

2) SVM(OvO 여러 번 사용)을 이용한 분류 성능평가

300명의 얼굴이 잘 나온 데이터를 사용하였으며 총 2,940개의 데이터로 학습을 시키고 1,260개의 데이터로 평가를 하였다. 학습 및 평가 정확도는 1,260개 분류를 성공적으로 하였다. 정확도가 높은 이유는 일부러 학습하기 쉬운 데이터를 사용했기 때문이다.

	precision	recall	f1-score	support
1~300	1.00	1.00	1.00	1,260
accuracy			1.00	1,260

표 4.1. SVM 성능평가

하지만 300명 중의 1명일 확률인 probability(개연성)가 굉장히 낮게 나왔으며 만약 이를 통해 애플리케이션을 만들어 사용한다면 낮은 개연성으로 인해 보안성에 문제가 있다고 판단하였다.

전체 확률을 100%라고 한다면 이 모델의 정답 평균 개연성은 3.686%가 나왔으며 이는 정답인 아닌 데이터들과 큰 차이가 없다. 따라서 SVM이 이진 분류모델이기에 이러한 한계점이 있다고 생각한다.

3) SoftMax를 이용한 분류 성능평가

300명의 얼굴이 잘 나온 데이터를 사용하였으며 총 2,940개의 데이터로 학습을 시키고 1,260개의 데이터로 평가를 하였다. 정답 데이터셋은 원-핫 인코딩(One-Hot Encoding)을 하였다. epoch를 늘릴수록 정확도가 올랐으며 이에 따라 평균 개연성도 높아졌다. 퍼셉트론 모델, 다층 퍼셉트론 모델, 깊은 다층 퍼셉트론 모델에 대한 성능평가는 epoch=300 기준으로 비교하였다.

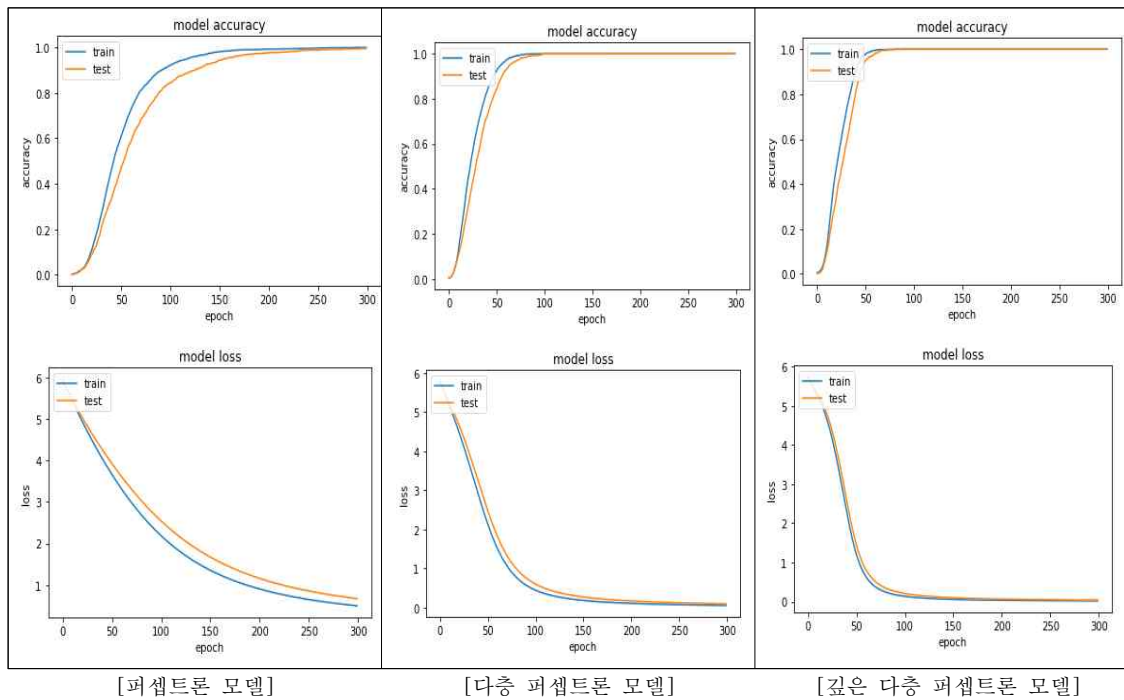


그림 4.4. 각 퍼셉트론 모델에 따른 성능평가 (epoch=300)

그래프를 보면 모델이 깊어질수록 정확도가 올라가는 속도와 loss가 줄어드는 속도가 확연히 빨라지는 것을 볼 수 있다. 정답의 평균 개연성 또한 퍼셉트론 모델은 57.899%, 다층 퍼셉트론 모델은 91.564%, 깊은 다층 퍼셉트론 모델은 96.406%로 상승하였다. 이는 정답이 아닌 데이터들과 큰 차이를 가진다고 볼 수 있다.

따라서 최종적으로 본 연구에서는 epoch=500, 깊은 다층 퍼셉트론 모델을 얼굴 인식 시스템에 사용한다. 이 모델의 평균 개연성은 97.902%이며 정확도는 1.00이다. 정확도가 높은 이유는 일부러 학습하기 쉬운 데이터를 사용했기 때문이다.

4) 최종 모델 성능평가

최종 모델에 대한 평가는 같은 사람으로 진행하였으며 위에 데이터 셋에서 설명한 것과 같이 각도, 조명, 액세서리, 표정이 다 다른 10,800장을 사용하였다. 10,800장 중 잘 분류된 사진은 총 3010장, 평균 개연성은 62.296%이며 나머지 잘 분류하지 못한 사진은 7790장, 평균 개연성은 14.209%를 가진다. 이는 데이

터에 다양한 조건이 복합적으로 적용되어 있으므로 본 모델이 얼굴이 잘 보이지 않는 사진을 인식하고 분류하기엔 한계가 있다.

	평균 개연성(%)	사진 장수(장)
분류 O	62.49649	3010
분류 X	14.20952	7790

표 4.2 최종 모델의 성능평가

따라서 누구인지 분류 후, 보다 정확한 로그인을 위한 조건으로 둘 개연성을 찾기 위해 잘 분류된 데이터에서 여러 조건을 따져 보았다.

잘 분류된 데이터(3010장) 중에서도 액세서리의 착용 여부에 따라 개연성이 굉장히 다르다. 아무것도 착용하지 않은 데이터의 평균 개연성이 79.757%로 굉장히 높았으며 가장 낮은 건 33.219%로 선글라스를 착용했을 때이다.

	액세서리	평균 개연성(%)	사진 장수(장)
1	미착용	79.757000	970
2	일반 안경	58.047645	694
3	빨데 안경	63.421099	594
4	선글라스	33.218506	318
5	모자	51.497065	251
6	모자+빨데 안경	50.839916	183

표 4.3 최종 모델의 액세서리 착용 여부 성능평가

아래 그래프를 확인해 보면 선글라스, 모자, 모자+ 안경을 착용했을 때 분류를 잘하지 못한다. [그림 4.5.]

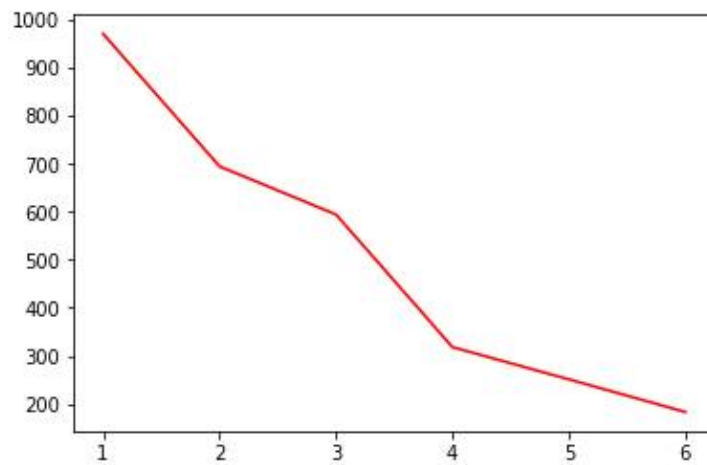


그림 4.5. 액세서리 착용에 따른 정답 분류 개수

조도의 양과 각도, 사진이 찍힌 방향에 따라서도 평균 개연성의 분포가 다양하다. [그림 4.6. 4.7.]

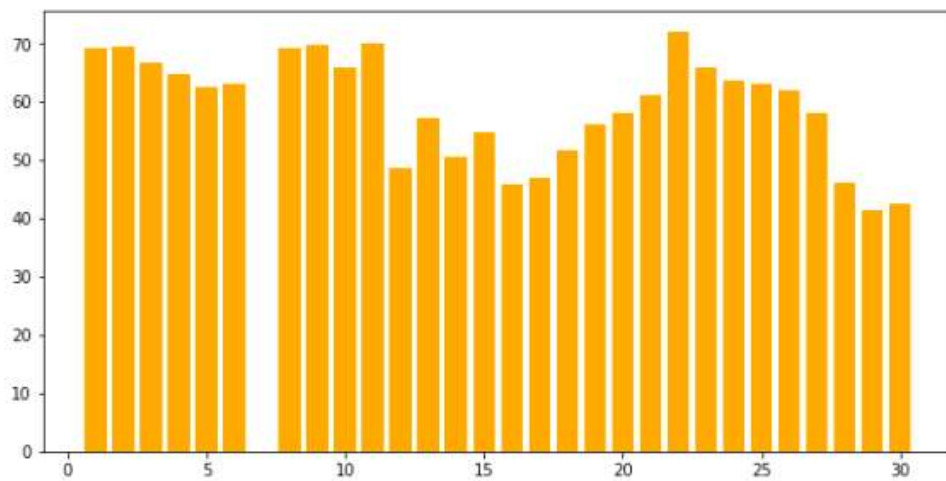


그림 4.6. 조도에 따른 평균 개연성 분포

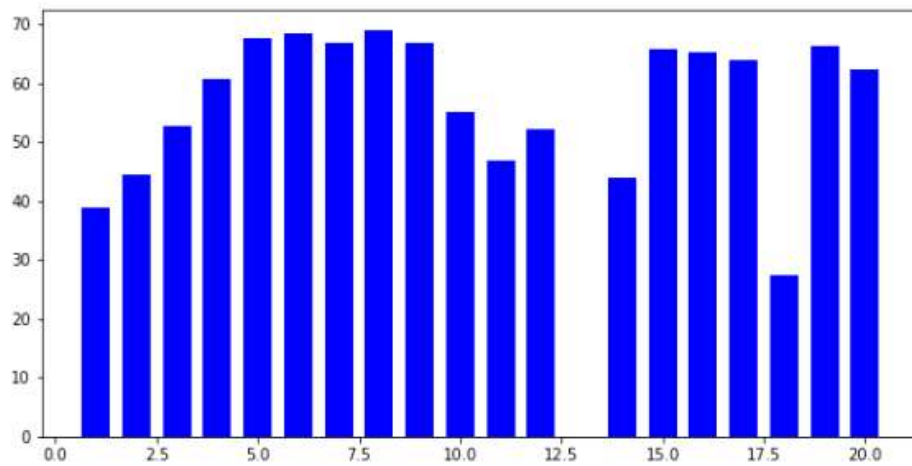


그림 4.7. 각도에 따른 평균 개연성 분포

조명이 L7의 경우에는 아예 분류하지 못하며 L22의 경우 가장 평균 개연성이 높았다. [그림 4.8, 4.9.]

각도가 C13의 경우에는 아예 분류하지 못하며 C8의 경우 가장 평균 개연성이 높았다. [그림 4.10, 4.11.]

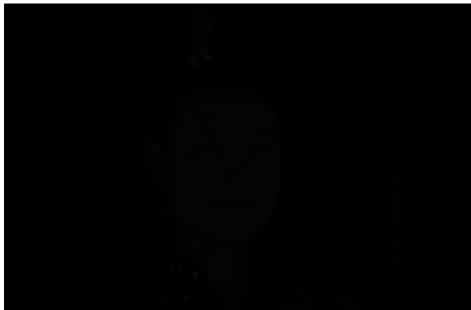


그림 4.8. L7 사진



그림 4.9. L22 사진



그림 4.10. C13 사진



그림 4.11. C8 사진

또한, 여기서 흥미로운 점은 평균 개연성이 떨어질수록 올바르게 분류된 데이터의 개수도 같이 감소하는 경향이 있다는 점이다.

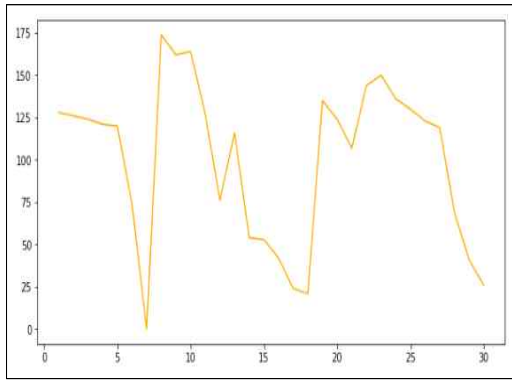


그림 4.12. 조도에 따른 분류된 데이터 개수

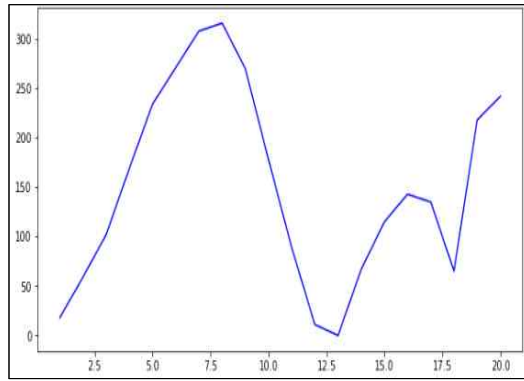


그림 4.13. 각도에 따른 분류된 데이터 개수

표정에 따라서는 찡그릴 때 평균 개연성이 37.075%로 가장 낮다. 이는 찡그릴 때 눈을 감기 때문에 누구인지 분류하기 어려움이 있어 보인다.

표정	평균 개연성(%)	사진 장수(장)
무표정	68.58038	1353
웃음	67.88565	1100
찡그림	37.07536	557

표 4.4 최종 모델의 표정에 따른 성능평가

따라서 본 연구에서는 이 모든 상황을 고려하여 평균 개연성 최댓값의 평균인 72.349%를 로그인 조건으로 둔다.

나. 결과

본 연구에서는 300명의 사진 중 실제 대학교 학생의 사진을 추가시켜 모델에 학습시켰다. [그림 3.1.] 그 후 최종 모델을 토대로 301명 중에 해당 학생을 찾아 평균 개연성의 최댓값인 72.349%가 넘는 영상의 프레임이 5번 이상이면 로그인 할 수 있도록 구현하였다.

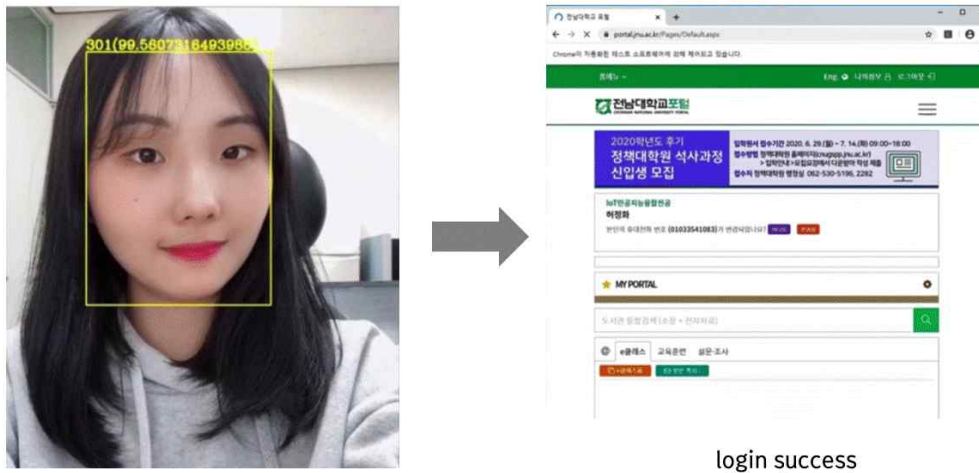


그림 4.14. 로그인 성공

[그림 4.14]는 301번 학생의 영상을 입력하였을 때 301번이라고 정확히 인식하여 로그인이 성공한 사진이다. 하지만 모델에 학습이 되지 않은 사람의 영상을 입력하면 다른 사람으로 판단하며 개연성이 매우 낮으므로 로그인에 실패하게 된다. [그림 4.15]

이로써 사람의 얼굴 영상을 이용하여 보안성 있는 로그인이 가능하다.

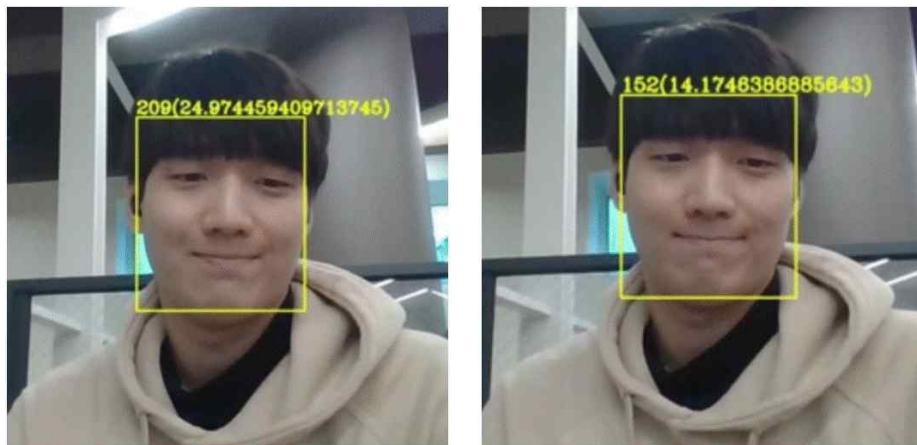


그림 4.15. 로그인 실패

5. 결론

본 연구에서는 MTCNN을 이용하여 얼굴을 추출하고, FaceNet을 이용하여 추출된 얼굴을 128개 요소의 벡터로 임베딩 해주었다. 이후 300명을 학습시킨 뒤 SVM과 SoftMax를 이용하여 얼굴을 분류해주었다. 분류기로 SVM을 이용할 경우 정확도는 높지만 낮은 개연성으로 인해 보안성 문제가 발생할 우려가 있다. 하지만 SoftMax의 경우 높은 정확도뿐만 아니라 높은 개연성을 가지기 때문에 SVM을 이용하였을 때 보다 안전하다고 판단하였다.

지금까지 대부분의 얼굴 인식 시스템은 OVA 매칭으로 개인이 사용하는 스마트폰, 노트북에 얼굴 인식을 사용했다면 본 연구로 인해 개인만이 아닌 여러 사람이 사용하는 기기 또는 웹사이트에도 얼굴 인식 시스템의 사용 가능성을 기대해 볼 수 있다.

참고문헌

[1] Zhang, Kaipeng et al. "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks." IEEE Signal Processing Letters 23.10 (2016)

[2] Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "FaceNet: A Unified Embedding for Face Recognition and Clustering." 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)

[3] "https://tykimos.github.io/2017/08/19/Numerical_Input_Multiclass_Classification_Model_Recipe/"

[4] "<https://github.com/nyoki-mtl/keras-facenet>"

[5] "https://github.com/k-face/k-face_2019"