# 환경 정보와 인간의 얼굴 정보에 기반한 이미지 추천 시스템

원혜민*, 황지혜**, 김지수**, 곽노준**
아주대학교 전자공학과*, 서울대학교 컴퓨터 지능 및 패턴인식 연구실**
dnjsgpals@ajou.ac.kr, {hjh881120, kimjiss0305, nojunk}@snu.ac.kr

## 요 약

인간이 느끼는 대부분의 감정은 표정을 통해 표현된다. 또한 주변 환경, 연령, 성별 등에 의해 다양하게 표현 될 수 있다. 따라서 사용자의 현재 상황에 따라 적절한 정보를 제공하기 위해서는 다양한 요인들을 고려해야한다. 본 논문에서는 사용자의 주변 환경, 표정, 연령, 성별 정보를 기반으로 사용자의 기분 상태에 따른 적절한 이미지 정보를 제공하는 시스템을 개발했다. 환경 정보는 기상청 웹사이트, OpenWeatherMap 웹 사이트, 스마트폰을 통해 수집했다. 표정, 연령 및 성별은 Caffe 심화 학습 알고리즘을 사용하여 분류했다. 선호하는 이미지는 사람들이 느끼는 감정에 따라 다를 수 있으므로 사용자 설문 조사를 통해 프로그램 사용자의 만족도를 높였다. 이 설문 조사는 다양한 연령대와 성별로 실시되었으며, 이를 기반으로 프로그램은 사용자 만족도와 안정성을 향상시키기 위해 시스템이 수정되었다. 실험에서 표정 인식은 89 %의 정확도를 보였고 연령 인식은 84 %의 정확도를 보였으며 성별 인식은 97 %의 정확도를 보였다. 또한 사용자가 이미지 추천 시스템을 사용했을 때의 사용자 점수는 10 점 만점에 8.5 점이 획득되었다.

## 1. 서론

요즘은 각 개인에게 맞게 맞춤 서비스를 제공하는 것이 보편적이 되었다. 이를 위해 일상 생활에서 제공되는 제품이나 서비스에서 인간의 감정을 자동으로 인식하고, 감정 정보의 데이터를 디지털화하며 감정 정보를 사용자의 상황에 따라 처리하여 개인화된 감정 서비스를 제공하는 감성 공학 기술이 강조되고 있다.

감성공학은 인간과 기계의 효율적인 상호 작용을 통해 인간의 감성을 정량적으로 측정하고 이를 과학적으로 분석 및 평가하여 편리하고 편안하며 안전하며 인간의 삶을 편안하게 돕는 기술이다. 본 논문에서는 감성 공학의 여러 기술 중 생체 인식 기술을 사용하는 인간 감성 기술과 심층 학습과 같은 알고리즘을 사용하여 인지된 감정을 추측, 예측 및 표현하는 기술을 사용했다.

우리의 시스템은 센서 처리 기술, 웹 데이터 획득 기술과 신경망을 이용한 분류 기술과 얼굴 인식을 위한 생체 인식 기술을 기반으로 한다. 기존의 논문은 딥러닝 알고리즘을 통해 인간의 표정을 인식하여 감정만을 인식한다. 그러나 사람의 감정만 인식하고 피드백을 제공하는 것은 개인에게 적절한 정보를 제공하기에 충분하지 않다. 감정 인식을 통한 피드백은 다양한 요인을 고려해야 한다. 본 논문에서는 표정인식을 통한 감정인식뿐만 아니라 그 감정에 영향을 미칠 수 있는 다양한 정보들을 이용하여 적절한 이미지 정보를 제공해주는 시스템을 구축했다. 이를 위해 본 논문에서는 사용자의 얼굴에서 나이와 성별 정보를 실시간으로 추정하고 20 개의 환경 정보를 수집했다. 표정, 연령, 성별 인식에는 합성곱 신경망(Convolutional Neural Network: CNN) 알고리즘이 사용되었으며, 모델 아키텍처로는 표정 인식에는 GoogLeNet, 연령 및 성별 인식에는 VGG16 이 사용되었다. 또한 스마트폰의 센서, 기상청의 RSS (Really Simple Syndication), OpenWeatherMap 을 통해 20 개의 환경 정보를 수집했다.

## 2. 관련 연구

### 2.1 표정인식

표정은 인간 감정 인식의 가장 중요한 특징 중 하나이다. 표정인식은 보편적으로 분노, 슬픔, 놀람, 행복, 혐오 및 두려움과 같은 6 가지 기본 표현 중 하나를 출력으로 제공한다. 표정 인식 시스템은 정적 이미지와 동적 이미지 시퀀스로 작업하는 두 가지 범주로 나눌 수 있다. 정적 기반 방법은 현재 입력 이미지에 대한 정보 만 포함하는 특징 벡터를 사용하기 때문에 시간 정보를 사용하지 않는다. 반면 한편, 동적 기반 방법은 이미지의 시간 정보를 사용하여 하나 이상의 프레임에서 캡처된 표정을 인식한다. 얼굴 표정 인식은 얼굴 찾기, 얼굴 데이터 추출, 얼굴 표정 인식 순으로 수행된다[1].

Ali 등은 부스트 신경망(boosted neural network)을 이용한 다민족 표정 인식을 수행했다[2]. 이 논문에서는 일본인, 타이완인, 백인 및 모로코인을 포함하는 JAFFE, TFEID 및 RaFD 데이터베이스를 사용하고, 두 개의 실험에서 다섯 가지 표정(분노, 행복, 슬픔, 놀라움 및 공포)을 인식했다. 첫 번째 실험에서는 모든 데이터를 합쳐서 실험하여 93.75%의 정확도를 얻었다. 두 번째 실험에서는 학습 데이터로 TFEID 와 RaFD 데이터베이스를 사용하고 JAFFE 데이터베이스를 테스트 데이터로 사용하여 48.67 %의 정확도를 얻었다.

### 2.2 나이 및 성별 인식

나이 및 성별 인식에는 정확한 성별과 연령 정보가 포함된 데이터베이스가 필요하다. 일반적으로 FG-NET [3], MORPH [4], UIUC-IFP-Y [5], FERET [6] 등이 연령 및 성별 추정 데이터베이스로 많이 사용된다.

Eidinger 등은 LBP 와 FPLBP 알고리즘을 사용하여 나이와 성별을 인식했다[7]. 데이터베이스로는 FG-NET 과 MORPH 데이터셋이 사용되었고, 연령 인식의 정확도는 45.1±2.6%, 성별인식은 77.8%의 정확도를 얻었다.

Chen 등은 연령 인식에 최적화된 ranking-CNN 알고리즘을 개발했다[8]. 그들은 실험을 위해 MORPH 데이터셋을 사용했고 인식 허용 범위를 나타내는 변수를 사용하여 실험을 수행했다. 결과적으로 변수 범위가 6 인 경우 정확도는 89.90%, 7 인 경우 92.93 %의 정확도를 얻었다.

### 2.3 추천 시스템

추천 시스템은 사용자가 관심을 가질 만한 제품이나 서비스를 자동으로 찾아줄 수 있는 개인화된 방법이다[9]. 추천 시스템에 입력되는 내용은 사용자에 대한 기본 정보, 아이템에 대한 특징, 사용자의 과거 사용 내역과의 상호 작용 및 시간 및 공간 데이터와 같은 기타 추가 정보를 포함한다. 이러한 내용들을 통해 추천 시스템은 입력 데이터 유형에 기반한 하이브리드 추천 시스템 등을 제공한다.

Lei 등은 사용자가 선호하는 이미지를 추천하는 시스템을 개발했다[10]. 이 논문에서는 Flickr 에서 API 로 크롤링된 데이터베이스를 사용했고, 듀얼 넷 딥 네트워크 모델(dual-net deep network model)을 통해 이미지를 분류하여 긍정적, 부정적 이미지를 나누어 사용자에게 보여주는 시스템을 구축했다.

## 3. 알고리즘

본 논문에서는 합성곱 신경망(Convolutional Neural Network : CNN)을 사용하여 표정, 연령, 나이 인식을 수행했다. 모델 아키텍쳐로는 GoogleNet 과 VGG16 을 사용했다. CNN 은 분류분야에서 보편적으로 많이 사용되는 것이고, 모델 아키텍쳐는 실험을 통해 대상을 잘 분류하는 모델을 선택하여 사용했다.

### 2.1 Convolution Neural Network (CNN)

합성곱 신경망(Convolutional Neural Network : CNN)은 여러 개의 합성곱 계층(convolutional layer)과 통합 계층(pooling layer), 완전 연결 계층(fully connected layer)들로 구성된 네트워크 이다. CNN 은 크게 세 단계를 거쳐 수행된다. 먼저 고정된 input 이미지 사이즈에 대해 합성곱 필터와 activation function(e.g. ReLU)를 통해 새로운 특징 맵을 형성하고, 이러한 레이어를 반복적으로 쌓는다. 그 후 분류를 위해, 완전 연결 계층을 연결하여 각 label 에 대한 점수를 출력으로 가진다. 마지막으로 역전파(Back propagation)를 통해 각 레이어의 파라미터들을 학습한다.

### 2.2 모델 아키텍쳐

GoogLeNet 은 “inception”이라는 개념을 지닌 깊은 합성곱 신경망(deep-convolutional neural network) 아키텍처로 22 개 계층 구조를 가지고 있다[11]. 이 아키텍처는 계산을 일정하게 유지하면서 네트워크의 깊이와 폭을 늘려 컴퓨터의 리소스를 효율적으로 활용하도록 설계되었다. 이러한 아키텍처 결정은 헤비안 규칙(hebbian rule)과 다중 스케일 프로세싱에 기반하여 최적화된다.

VGG16 은 16 개 가중치들의 층으로 이루어진 매우 작은 합성곱 필터를 사용한다[12]. 이 모델 아키텍쳐는 필터 크기 3x3, 스트라이드 1, 제로 패딩 1 및 2x2 필터 크기 (패딩 없음)를 가진 풀 레이어로 구성된다. 이미지 분류 성능은 GoogLeNet 보다 약간 낮지만 빠르고 간단하며 변형에 용이하다.

## 4. 실험 및 결과

사람의 얼굴에서는 얼굴 표정, 연령, 성별 등의 많은 정보를 얻을 수 있다. 본 논문에서는 합성곱 신경망(Convolutional Neural Network : CNN)을 통해 Cohn-Kanade(CK+) 데이터셋과 MORPH 데이터셋의 얼굴 데이터를 학습하고 테스트함으로써 표정, 연령, 성별을 인식했다.

### 4.1 Dataset

본 논문에서는 총 두 개의 데이터셋이 사용되었다.

표정 인식을 위해 사용된 Cohn-Kanade(CK+) 데이터셋은 성인의 8 가지 표정에 대한 데이터를 가지고 있다[13]. 이 데이터셋에는 중립, 슬픔, 놀람, 행복, 공포, 분노, 경멸 및 혐오감 등 8 가지 표

정 데이터가 있다. 이것은 640x480 크기의 회색 채널 이미지로 123 명의 피실험자들에 대해 8 가지 표정의 비디오 시퀀스로 이루어져 있다.

표 1. CK+ dataset

| | Train | Validation |
|---|---|---|
| Subjects | 100 | 23 |
| frame (image) | 1330 | 305 |

연령과 성별 인식에는 MORPH 데이타셋(학술적 버전)이 사용되었다[4]. 이 데이터셋에는 대상의 연령과 성별, 인종 정보 등의 개인정보가 들어가 있고 총 55,134 개의 데이터로 이루어져 있다.

표 2. MORPH dataset

| | 남자 | 여자 | 총 |
|---|---|---|---|
| 20 세 미만 | 6,638 | 831 | 7,469 |
| 20 ~ 29 | 14,016 | 2,309 | 16,325 |
| 30 ~ 39 | 26,447 | 2,910 | 25,357 |
| 40 ~ 59 | 10,062 | 1,988 | 12,050 |
| 50 세 이상 | 3,482 | 451 | 3,933 |
| 총 | 46,645 | 8,489 | 55,134 |

### 4.2 표정인식

본 논문에서는 각 비디오 시퀀스에서 일부 프레임들을 추출해 낸 후, 각 프레임에 대해 haar-like feature 기반의 얼굴 찾기 알고리즘을 적용하여 얼굴영역만 잘라 낸 것을, 학습 및 Validation image data 로 사용하였다. CNN 알고리즘을 사용하였고 GoogleNet 을 모델로 사용해 실험을 진행했다. 그 결과 8 가지의 표정 중 중립에 대한 데이터들을 제외하고 Validation set 에 대해 약 89%의 성능을 보였다.


그림 1. 표정인식

### 4.3 연령 및 성별 인식

본 논문에서는 연령 및 성별인식을 위해 CNN, 모델 아키텍쳐 VGG16 을 사용했다. 두 데이터의 특성 차이로 인해 loss layer 는 서로 다른 것들을 사용했다. 성별 인식에는 multiclass classification 에서 많이 사용되는 softmax loss 를 사용했다. 그러나 연령인식에는 연령을 세분화하여 분류하기 위해 regression loss 중 하나인 Euclidean loss 사용했다.

연령인식은 데이터에서 연령의 범위를 16 세에서 50 세까지 총 45 개의 class 로 분류하여 실험을 진행했다(16 세 이하는 16 세 class 로, 50 세 이상은 50 세 클래스로 분류). 학습과 테스트를 3:1 비율로 분류하였고 정확도 측정에서 +-5 세까지의 오차율 범위 지정하여 실험을 진행했다. 그 결과 약 84%의 인식률을 보였다.

표 3. 연령별 accuracy

| | 데이터 개수 | Accuracy |
|---|---|---|
| 10 대 | 1,493 | 0.96 |
| 20 대 | 3,266 | 0.88. |
| 30 대 | 3,071 | 0.83. |
| 40 대 | 2,409 | 0.80 |
| 50 대 | 720 | 0.70 |
| 60 대 | 67 | 0.55 |

성별인식은 총 55,347 장의 데이터 중에서 학습 데이터는 50,170 장, 테스트 데이터는 5,177 장으로 분류하여 사용했다. 그 결과 테스트 데이터에 대해 여자 약 99%, 남자 약 96%로 총 97%의 인식률을 보였다.

표 4. 성별 별 accuracy

| | 데이터 개수 | Accuracy |
|---|---|---|
| 여자 | 3231 | 0.99 |
| 남자 | 15192 | 0.96 |

## 5. 환경 정보와 인간의 얼굴 정보에 기반한 이미지 추천 시스템

이미지 추천 시스템은 기존 보유 공간(광고용 대형월, 체험/전시관 인터랙티브 비디오월, 버스정류장 디지털 광고보드 등)에 적용함으로써 현재 필요한 정보를 바로 제공하는 것이 가능하게 한다.

인간의 감정은 날씨와 주변 환경에 따라 변화할 수 있고 보통 얼굴표정으로 많이 표현된다. 사용자에게 알맞은 이미지를 추천하기 위해서는 현재의 외부 및 주변 환경, 사용자가 느끼고 있는 감정, 사용자의 성별, 사용자의 연령 등 다양한 요소들을 고려해야한다. 본 논문에서는 15 개의 환경정보와 인간의 얼굴에서 얻을 수 있는 표정, 연령, 성별 정보를 이용해 이미지를 추천하는 시스템을 구현했다.

실험은 Windows 8.1, Intel 코어 i7, RAM 24GB 및 Titan X (파스칼)가 설치된 PC 에서 수행되었으며 카메라는 Logitech C920 이 사용되었다. 실험은 조도가 평균 150lux 이상인 곳에서 수행되었고 카메라는 정면에 설치되었다.

### 5.1 환경정보 수집

본 논문에서는 외부 환경 정보 (날씨)와 주변 환경 정보를 웹 및 센서를 통해 실시간으로 수집했다.

외부 환경은 기상청 및 Open-WeatherMap 웹사이트를 통해 수집되었다. 기상청 웹에서는 RSS 를 통해 기상 데이터가 업데이트 될 때마다 새로운 날씨 정보 12 개를 받았다. Open-WeatherMap 에서는 공개 API 를 통해 3 가지 기상 정보를 실시간으로 받았다.

표 5. 수집되는 환경정보

| 수집 데이터 | 출처 |
|---|---|
| 온도 | 기상청 |
| 일중 최고온도 | |
| 일중 최저온도 | |
| 하늘상태 | |
| 강수 상태 | |
| 날씨 | |
| 강수확률 | |
| 12 시간 예상 강수량 | |
| 12 시간 예상 적설량 | |
| 풍속 | |
| 풍향 | |
| 습도 | |
| 기압 | Open-WeatherMap |
| 일출 시간 | |
| 일몰 시간 | |

주변 환경 정보는 Android 기반 스마트 폰의 센서를 사용하여 수집했다. 이를 통해 사용자 주변에 대한 조도, 대기압, 소음, 방향(회전각도), 가속도의 5 가지 정보를 받아 사용자가 특정 공간 환경하에서 감정, 행동 변화 등을 유발할 수 있는 정보를 수집했다.


그림 2. 안드로이드 앱 데이터 출력 결과

### 5.2 실험결과

본 논문에서는 사용자의 주변 환경 정보, 감정 상태, 나이 및 성별 정보를 결합하여 적절한 이미지 정보를 제공하는 시스템을 개발했다. 사용자의 현재 주변 정보를 수집하여 표 정, 성별, 연령 정보를 인식하고 그 정보들을 통해 사용자에게 필요한 정보를 제공한다. 이를 위해 if-then 규칙을 적용하여 결과를 추론하여 나오는 이미지가 규칙에 맞게 출력되는지 측정했다. 또한 사용자 평가를 통해 편리성, 적합성, 정확성을 평가했다.

이 시스템을 테스트하기 위해 20 대와 40 대의 남녀 (남자 9 명, 여자 3 명)가 총 12 명이 실험에 참여했다. 테스트 후 총 5 가지 항목의 설문조사를 수행하여 프로그램 만족도를 조사했다. 사용자의 사용성과 편리성은 사용자의 주관적인 평가이지만 실험 후 가능한 한 많은 사람들의 만족을 이끌어 내기 위해 동일한 실험자에게 총 두 번의 실험이 수행하면서 프로그램을 개선하였다. 첫 번째 실험 후 설문조사 결과에 따라 이미지와 프로그램 UI 를 개선했고 그 결과 두 번째 실험에서는 만족스러운 결과를 얻었다. 본 논문에서는 다음과 같은 질문을 통해 프로그램 사용의 편리성, 제안된 이미지의 적합성 및 프로그램의 정확성을 살펴보았다.

- 사용자 정보: 연령, 성별
- 사용한 프로그램은 편리했습니까? (1~10)
- 프로그램에서 제시한 영상은 본인의 상황에 부합하게 적절했습니까? (1~10)
- 본인의 연령, 성별, 표정에 관하여 프로그램이 정확하다고 판단하십니까? (1~10)
- 프로그램에 관한 의견을 적어주세요.

모든 설문조사는 익명으로 진행되었으나 연령과 성별에 따른 반응을 보기 위해 두 항목에 대한 개인 정보를 수집했다. 다음으로, 프로그램의 인터페이스가 사용하기 쉽고, 프로그램에서 보여지는 결과 이미지가 적절한지, 사용자의 표정과 연령 인식이 잘 수행되는지 여부를 파악하기 위해 세 가지 질문을 넣었다. 마지막으로 사용자의 프로그램에 대한 의견을 통해 프로그램의 개선점이나 단점에 대한 피드백을 받았다

표 6. 설문조사 결과

| | 질문 1 | 질문 2 | 질문 3 | 평균 |
|---|---|---|---|---|
| 총 | 8.4 | 8.8 | 8.2 | 8.5 |
| 여자 | 7.7 | 9 | 8 | 8.2 |
| 남자 | 8.7 | 8.7 | 8.2 | 8.5 |
| 20 대 | 8.5 | 9 | 8.3 | 8.6 |
| 30 대 이상 | 8.3 | 8.5 | 8 | 8.3 |

총 12 명의 사용자가 설문조사에 응하였고 3 가지 질문에 대한 평균 점수는 10 점 만점에 8.5 점이었다. 여성의 경우, 프로그램의 편리성 부분에서

낮은 점수를 보였는데 이는 UI 문제로 인한 것으로 앞으로 더 개선할 계획이다. 또한 사용자의 의견에 얼굴 표정 인식률에 대한 많은 피드백이 있었다. 이는 사용된 두개의 데이터베이스에 백인과 아프리카계 사람들의 데이터가 아시아인 데이터 보다 압도적으로 많고, 인종간 표정을 짓는데 사용되는 얼굴 근육에 차이가 있어 발생한 문제이다. 이 설문 조사는 다양한 연령대와 성별을 대상으로 실시되었으며, 이를 토대로 우리는 시스템의 성능을 향상시켰다.

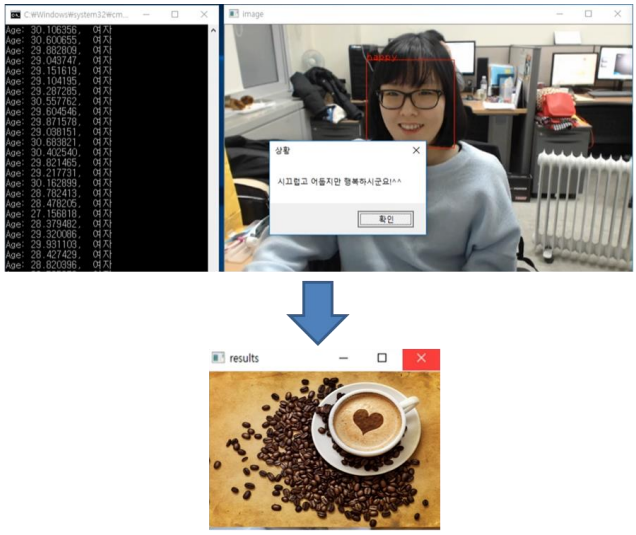


그림 3. 이미지 추천 시스템 예: 외부 날씨 좋음, 주변 소음 심함, 어두움, happy, 30 대, 여자

## 6. 결론

본 논문에서는 인간의 감정을 실시간으로 인식하여 성별, 연령 및 주변 환경에 따라 적절한 이미지 정보를 제공하는 시스템을 구축했다. 이를 위해 카메라로 8 가지 표정, 성별, 나이를 실시간으로 인식하고, 기상청 RSS 데이터와 OpenWeatherMap 을 통해 15 개의 환경 정보를 수집했다. 데이터베이스로 CK+ 데이터셋 및 MORPH 데이터셋이 사용되었고, CNN 알고리즘과 GoogLeNet 및 VGG16 모델 아키텍처가 사용됐다. 실험결과, 표정 인식 89%, 연령 인식 84%, 성별 인식 97%의 정확도를 얻었다. 또한 실시간으로 동작하도록 설계하여 사용자 평가를 받은 결과 10 점 만점에 8.5 점을 받았다. 실시간 테스트 결과, 인종간의 사용되는 안면 근육의 차이 때문에 시스템의 표정인식 부분의 정확성에 중요한 영향을 미쳤다. 이는 실험에 사용된 학습 데이터에 아시아인의 표정 데이터가 다른 인종에 비해 현저히 부족하여 발생한 문제이므로 앞으로 더 많은 데이터를 수집, 추가하여 개선해 나갈 것이다. 또한 향후 다양한 상황에서 더 많은 이미지를 추천, 제공하기 위해 시스템을 개선해 나갈 것이다.

## 감사의 글

본 연구는 한국연구재단의 차세대정보컴퓨팅기술개발사업에 의해 진행되었음. (2017M3C4A7077582).

## 참고문헌

[1] Anil K Jain and Stan Z Li, Springer : Handbook of face recognition, 2011.

[2] G. Ali, M. A. Iqbal, and T. S. Choi, “Boosted NNE collections for multicultural facial expression recognition,” Pattern Recognition, 55, pp. 14–27, 2016.

[3] G. Panis, A. Lanitis, “An overview of research activities in facial age estimation using the fg-net aging database,” In ECCV, pp. 737–750, 2014.

[4] K. Ricanek, T. Tesafaye, 2006. “Morph: A longitudinal image database of normal adult age-progression,” *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 341-345, 2006.

[5] Y. Fu, T.S. Huang, "Human Age Estimation with Regression on Discriminative Aging Manifold", *IEEE Trans. Multimedia*, no. 4, pp. 578-584, June 2008.

[6] P. J. Phillips, H. Wechsler, J. Huang, and P. J Rauss, “The FERET database and evaluation procedure for face-recognition algorithms,” Image and vision computing, vol. 16, no. 5, pp. 295–306, 1998.

[7] E. Eidinger, R. Enbar and T. Hassner, “Age and gender estimation of unfiltered faces,” IEEE Transactions on Information Forensics and Security, vol. 9, no. 12, pp. 2170–2179, 2014.

[8] S. Chen, C. Zhang, M. Dong, J. Le and M. Rao, “Using ranking-cnn for age estimation,” In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[9] G. Adomavicius, A. Tuzhilin, 2005. “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions,” IEEE transactions on knowledge and data engineering, vol. 17, no. 6, pp. 734–749, 2005.

[10] C. Lei, D. Liu, W. Li, Z. J. Zha, and H. Li, “Comparative deep learning of hybrid representations for image recommendations,” In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2545–2553, 2016.

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, “Going deeper with convolutions,”, In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[12] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition", *Proc. Int. Conf. Learn. Representations*, 2015.

[13] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, “The extended cohn-kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," Computer Vision and Pattern Recognition Workshop on Human-Communicative Behavior, 2010.