

# 인공지능 표정 인식 기술을 활용한 사용자 감정 맞춤 힐링·광고 서비스

김민식\*, 정현우\*\*, 문윤지\*\*\*, 문재현\*\*\*\*

\*국민대학교 빅데이터경영통계학과 \*\*고려대학교 응용통계학과

\*\*\*가천대학교 컴퓨터공학과 \*\*\*\*한국기술거래사회

gby1349@naver.com, officialwoo@gmail.com,

banan99@naver.com, smjhoon@gmail.com

\*\*\*\*교신저자(Corresponding Author)

## Using AI Facial Expression Recognition, Healing and Advertising Service Tailored to User's Emotion

Minsik Kim\*, Hyeon-woo Jeong\*\*, Yoonji Moon\*\*\*, Jaehyun Moon\*\*\*\*

\*Big Data Management Statistics, Kookmin University

\*\*Dept of Applied Statistics, Korea University(Sejong Campus)

\*\*\*Dept of Computer Engineering, Gachon University

\*\*\*\*Korea Technology Transfer Agents Association

### 요 약

DOOH(Digital Out of Home) advertisement market is developing steadily, and the case of use is also increasing. In advertisement market, personalized services is actively being provided with technological development. On the other hand, personalized services are difficult to be provided in DOOH and are provided by only personal information, not feelings. This study aims to construct personalized DOOH services by using AI facial expression recognition and suggesting a solution optimized for interaction between user and services by providing healing and advertisement.

### 1. 서론

옥외(OOH, out-of-home) 광고는 빠르게 디지털화되고 있다. 기술의 발전과 코로나로 인한 사회적 거리 두기는 디지털화를 가속시켰다. 특히, 2008년부터 옥외 광고에 대한 본격적인 논의가 시작된 이후 2017년부터 디지털 사이니지(Digital Signage)의 보급이 본격화되었다. [1]

이러한 변화는 광고주에게 다양한 시도를 할 수 있는 기회가 되었다. 해외에서는 미국 뉴욕의 타임 스퀘어 뿐만 아니라, 버거킹, 넷플릭스의 DOOH를 활용한 사례가 있다. 국내에서는 코엑스 K-POP 스퀘어 전광판을 활용한 디스트릭트(d'strict)의 미디어 아트 '웨이브(WAVE)'가 세계 3대 디자인 어워드 중 하나인 'iF 디자인 어워드 2021'에서 최고점으로 금상을 받았다. [2] 이처럼 세계적으로 DOOH의 다양한 시도를 통해 소비자의 이목을 끌고 있다.

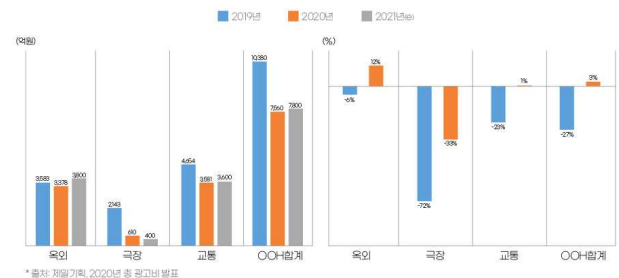
발전과 함께 개인화 서비스는 하나의 트렌드가 되었다. 광고 시장에서도 고객 맞춤형 광고들이 활발히 이루어지고 있다. 이러한 서비스들은 고객들의 정보를 바탕으로 제공된다. 이러한 광고 효과를 높이려면 고객의 기본에 따른 소비 패턴 등의 파악하는 것이 필요하다. 이에 사용자 얼굴 감지 및 분류를 통한 고객 맞춤형 광고서비스를 제안한다.

### 2. 선행 연구

디지털 옥외 광고는 기존 포스터와 같은 제한된 형식의

한계점을 극복하고 다양한 미디어 형식을 통해 효과적인 홍보가 가능해졌다. 하지만 여전히 옥외 광고의 한계점인 대중화를 극복하지는 못했다. 불특정 다수에게 노출되는 옥외 광고 특성상 많은 사람을 대상으로 하는 광고들이 나타나며, 광고 효과 또한 바로 나타나지 않는다. 이러한 특성은 코로나를 맞이하며 큰 타격을 입었다. 사회적 거리 두기로 인해 생활환경이 좁아짐에 따라 사람들에게 노출되는 기회가 줄어들었으며 이는 광고비에 영향을 주었다. 국내 광고 기업인 제일기획의 경우, (그림 1)처럼 2019년(파란색) 대비 2020년(주황색) OOH 광고비가 감소했음을 확인할 수 있다. [3]

2019년-2020년 OOH 광고비 및 전년 대비 성장률



\* 출처: 제일기획, 2020년 동 광고비 발표

(그림 1) 2019년~2020년 OOH 광고비 및 전년 대비 성장률

그리고 개인 맞춤 광고의 경우에도 한계가 존재한다. 고객의 개인정보, 소비 패턴을 통한 맞춤형 광고의 경우 소비자

의 감정은 반영되지 않았다. 기분에 따른 구매 의도에 영향을 미친다는 연구가 선행되었다는 점에서 고객의 감정 역시 광고에 매우 중요한 요소이다. [4] 고객의 감정 변화를 파악하고 이에 알맞은 광고를 추천해준다면 더 효과적일 것이다.

AI를 통한 이미지 분류 기술의 발전에 따라 사람의 표정을 통해 기분을 분류하는 연구들도 활발히 이루어졌다. [5] 이러한 연구에서는 한국인들의 얼굴이 아닌 외국인들의 얼굴 이미지들이 활용되었다. 이로 인해 해당 모형들은 한국인들에게 적용 시 정확도가 떨어진다는 문제점이 있다.

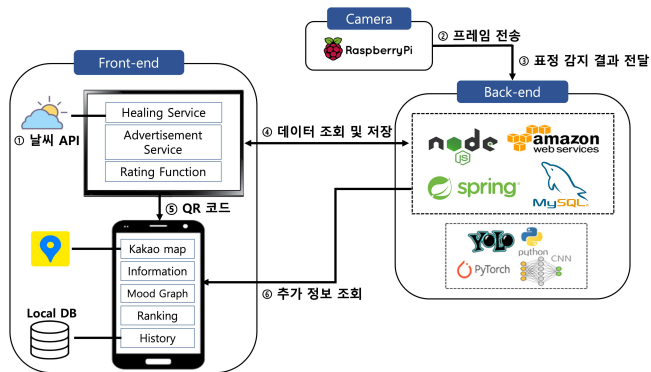
이러한 문제점들을 해결하기 위해 한국인들의 얼굴 표정을 학습시킨 모델을 탑재한 사용자 맞춤 광고서비스를 제안한다.

### 3. AI 사용자 맞춤 힐링·광고서비스

본 서비스는 옥외 광고의 개인화와 광고 측면 효과를 극대화하기 위해 패널의 카메라로 사용자의 표정을 인식해 감정 맞춤 힐링·광고서비스를 제안한다. 객체 탐지 알고리즘과 표정 인식 알고리즘을 활용하여 영상 속 사용자의 얼굴을 인식 후 표정을 감지한다. 이때 카메라에 연결된 라즈베리파이(Raspberry Pi)에서 딥러닝(Deep Learning) 서버로 영상을 전송하고 딥러닝 서버에서 처리한 분류 결과를 데이터베이스(Database, DB)에 전달한다. DB에서는 감정에 맞는 서비스 데이터를 조회하여 패널을 통해 서비스 영상을 재생한다. 이후, 해당 서비스의 구체적인 정보를 받기 원하는 사용자에게 QR 코드를 통한 모바일 웹(Mobile Web) 내에서 해당 정보를 제공한다.

#### 3.1 시스템 구성도

본 시스템은 라즈베리 파이, 딥러닝 모델 처리를 위한 딥러닝 서버, 옥외 패널과 모바일 웹을 위한 웹서버 2대로 구성되어 있다.



(그림 2) 서비스 구성도

(그림 2)는 본 시스템의 구성도이다. IoT 기기인 라즈베리파이로 실시간 캡처된 프레임을 딥러닝 서버에 전달한다. AWS의 Deep Learning AMI 인스턴스로 구동되는 딥러닝 서버는 IoT 기기인 라즈베리 파이와 직접 통신하며, 전달받

은 값을 딥러닝 모델에 전달하여 값을 도출한다.

사용자와 직접적으로 상호작용하는 웹 애플리케이션(Web Application)은 Node.js 서버를 기반으로 옥외 패널에 연결된다. Node.js 서버는 딥러닝 서버와의 통신, 날씨 API 연동, 데이터베이스 조회 및 추가, 힐링 및 광고서비스 제공 등의 기능을 구현한다. 사용자에게 세부 정보들을 제공하는 모바일 웹은 Spring Boot를 기반으로 스마트폰과 같은 휴대기기에 연결된다. Spring Boot 서버는 DB 조회 및 집계, 시각화 등을 통해 개인화된 서비스를 제공한다.

#### 3.2 딥러닝 서버

딥러닝 서버에서는 옥외 패널 상단 중앙에 배치되어있는 라즈베리 파이 카메라를 통해 실시간 프레임을 받게 된다. 서버 내 딥러닝 모델을 통해 얼굴 인식과 감정 분류 후 DB의 users 테이블에 결과를 저장하고 Node.js 서버로 데이터를 전송한다. DB에 저장된 데이터는 사용자의 일일 감정 그래프 기능과 광고서비스 만족도 랭킹 기능 등 모바일 웹에서 제공하는 서비스에 사용된다.

user_id	now_emotion	rating	today	ad_id
2	0	0	2021-9-20	16
3	1	1	2021-9-24	17
4	0	0	2021-9-28	18
5	0	NULL	NULL	18

(그림 3) 데이터베이스 테이블

(그림 3)에서 사용자를 구분하는 고유한 값인 user\_id, 사용자의 표정 결과(감정값)인 now\_emotion, 사용자에게 제공할 맞춤형 광고데이터의 고유값인 ad\_id, 해당 서비스에 대한 만족도 결과(0: 부정/불만족 1: 긍정/만족)인 rating, 옥외 패널 광고서비스를 제공한 날짜인 today로 구성된 테이블을 확인할 수 있다. 딥러닝 서버에서 패널까지 데이터가 축적된 경우, user\_id값이 5인 필드와 같이 DB에 저장되며, 옥외 패널 서비스까지 제공한 후 DB의 모습은 user\_id 값이 2, 3, 4인 필드와 같다.

#### 3.3 웹 서버

Node.js 서버는 딥러닝 서버로부터 표정 결과(now\_emotion), 사용자 id(user\_id), 광고 id(ad\_id)의 데이터를 수신하게 되면 패널의 힐링 서비스 화면(초기화면)을 광고서비스 화면으로 전환한다. 패널에서 일정 시간 동안 광고서비스가 제공된 이후, 서비스 만족도 평가와 QR 코드 화면으로 전환한다. 사용자가 만족도 평가 버튼 클릭 시 결과를 DB에 저장하고, 사용자에게 QR 코드(동적 링크)를 제공한다. 이후, 패널은 초기화면으로 전환된다.



(그림 4) 패넬를 통해 제공되는 서비스

사용자는 구체적인 광고 정보와 추가적인 서비스를 희망할 경우, 제공된 QR 코드를 통해 스마트폰의 모바일 웹에서 서비스를 이용할 수 있다. 모바일 웹에서는 광고의 구체적인 정보(주소, 설명, 연락처, 길 찾기 등)와 사용자들의 만족도에 따른 힐링 및 광고서비스의 순위 서비스를 전달한다.

#### 4. 인공지능 알고리즘 선정 및 개선

##### 4.1 사용자 얼굴 탐지 알고리즘 선정

영상 속 사용자 얼굴 검출에 사용되는 대표적인 객체 탐지(Object Detection) 알고리즘으로 CNN(Convolutional Neural Network) 와 YOLO(You Only Look Once)가 있다. CNN 계열의 알고리즘의 경우, 객체 탐지의 정확도는 높은 대신 속도가 느리다. 반면, YOLO 알고리즘은 CNN 계열 알고리즘 대비 정확도가 상대적으로 낮지만, 속도는 빠르다. 본 서비스에서는 표정 분류에 필요한 사용자 얼굴을 영상 속에서 실시간으로 감지해야 되기 때문에, 속도가 빠르며 준수한 정확도를 가진 YOLO 알고리즘을 선택하였고, 그중에서 이미 많은 검증이 이루어진 YOLOv3를 사용하였다.

##### 4.2 표정(감정) 데이터 수집

본 서비스 중 감정 분류에 사용할 CNN 모델을 구축하기 위한 학습 데이터로 AI Hub에서 제공하는 개방 데이터인 “한국인 감정 인식을 위한 복합 영상”의 데이터를 사용했다. Data-set 내에서 표정이 애매하거나 잘못 분류된 데이터를 제거하고 Train 및 Validation Data-set을 구성하였다. 그리고 Test Data는 구글 검색을 통한 데이터를 크롤링(Crawling)하여 수집하였다.

본 서비스가 옥외 서비스라는 것을 고려하여 표정은 '행복, 무표정, 슬픔' 3가지 표정으로 구성하였으며, 수집한 데이터 수는 <표 1>과 같다.

<표 1> 표정별 수집한 Data-set의 개수

	Happy (기쁨)	Blank (무표정)	Sad (슬픔)	Total
Train Data	3,200	3,200	2,840	9,240
Validation Data	800	800	710	2,310
Test Data	160	150	150	460

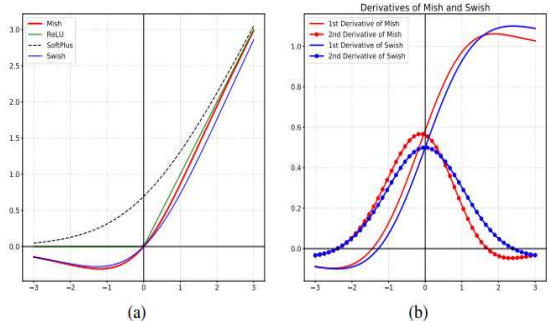
옥외 광고의 특성상 카메라 각도나, 사용자의 키에 따른 얼굴각도, 야외 밝기나 채도에 따라서 바뀔 수 있기에 데이터 증대(Data augmentation)를 적용하였다.

Train 데이터를 랜덤하게 각각 5°, 10°, 15°, 20°로 회전을 했고, 흑백 처리 및 밝기, 채도, 대비, 색조 등을 무작위로 변경하여서 과적합(Over fitting)을 해결하고자 하였다.

##### 4.3 사용자 표정(감정) 알고리즘

사람의 표정을 구분하기 위해서는 학습할 때 얼굴 영역의 데이터만 필요하다. 그래서 수집된 데이터를 YOLO 활용해 얼굴 영역을 검출하여 잘라내는 전처리 과정을 거쳤다.

감정 분류를 위한 다양한 CNN 알고리즘 중에서 대표적인 CNN 이미지 분류모델인 VGG, ResNet, MobileNet, DenseNet 계열의 모델을 학습하였다. 활성화 함수의 경우, Relu, Swish, Mish 3가지를 사용하여 성능을 비교하였다.



(그림 5) Relu(초록), Mish(빨강), Swish(파랑) 활성화 함수(a)와 Mish(빨강), Swish(파랑)의 미분 그래프(b)

$$Mish(x) = x * \tanh(\frac{1}{2} \ln(1 + e^x)) \dots (1)$$

$$Swish(x) = x * Sigmoid(x) \dots (2)$$

(1)과 (2)의 수식은 각각 Mish와 Swish 활성화 함수 공식이다. 두 함수는 Relu와 달리 모두 연속, 미분 가능하다.

그림 5. (a)를 보면 작은 음의 값을 제한적으로 허용하기에 정보 손실은 줄이고 강력한 정규화 효과가 있다. 양의 값은 제한이 없어 포화를 방지하기 때문에 깊은 신경망에서는 Relu보다 좋은 성능을 발휘한다. 특히, Mish는 Swish 비해 기울기의 Smoothing 효과와 loss는 부드럽게 형성되어 최적화와 일반화에 좋은 효과를 발휘한다. [6]

총 Epoch은 70, learning-rate는 0.0001로 하였으며, Optimizer는 Adam으로 진행해 학습을 진행하였다.

<표 2> CNN 모델과 활성화 함수에 따른 성능 비교

활성화함수 Model	Relu	Swish	Mish
MobileNet V2	70.12%	72.99%	75.22%
DenseNet	74.27%	78.76%	79.56%
Vgg19	87.84%	88.62%	89.96%
ResNet50	87.50%	89.06%	<b>90.40%</b>

<표 2>는 모델과 활성화 함수에 따른 정확도 결과를 보여 준다. ResNet(50-layers)과 Vgg(19-layers)가 다른 모델들에 비해 정확도가 높았다. 또한, 활성화 함수를 Swish와 Mish로 사용했을 때 Relu 대비 성능이 향상되었으며, 상대적으로 Mish가 Swish 비해 높았다. 본 서비스에서는 정확도가 가장 높은 Mish를 적용한 ResNet50 모델을 사용한다.

<표 3> “ResNet50 + Mish”의 Confusion Matrix(%)

	행복함	무표정	슬픔
행복함	92.99	0.64	6.37
무표정	1.38	86.90	11.72
슬픔	3.42	6.16	90.41

<표 3>은 ResNet50 모델에 활성화 함수를 Mish 적용한 모델의 감정별 혼동행렬(Confusion Matrix)을 산출하였다. 행은 실제 표정 결과를 얻은 모델의 예측 표정 결과를 나타내고, 대각선은 예측한 표정이 맞을 확률을 나타낸다. 이를 통해서 각 모델이 표정을 어떻게 예측했는지 전체적인 분포를 확인할 수 있다.

성능 결과를 비교하면, “무표정”이 “행복”과 “슬픔”에 비해서 정확도가 떨어지며, “무표정”과 “슬픔” 사이의 오분류가 상대적으로 많은 걸 확인할 수 있다. 하지만 표정들의 정확도가 85%가 넘고, 평균 정확도가 90.40%이기 때문에 좋은 모델로 판단된다.

<표 4> Test 데이터 정답과 오답 사례 사진



<표 4>는 실제 Test 중 정답과 오답 사례를 보여준다. 오답 사례 중 첫 번째 사례(행복한 표정)의 경우, 환한 미소가 아닌 옅은 미소로 인해 무표정으로 잘못 분류한 경우이며, 두 번째 사례(무표정)의 경우 입꼬리가 올라간 것을 통해 무표정을 행복으로 오분류된 것으로 확인할 수 있다. 마지막 사례(슬픈 표정)에서는 얼굴이 정면이 아니며, 벌어진 입을 통해 슬픔을 행복으로 오분류된 것으로 확인할 수 있다. 이

처럼, 오분류의 경우, 비정면 혹은 사람이 바로 판단하기도 어려움이 있을 수 있는 경우로 확인이 되었다.

5. 결론

본 서비스는 객체 탐지 알고리즘과 표정 인식 알고리즘을 활용해 사용자의 감정(표정)에 맞는 맞춤형 광고서비스를 구현하였다. 다양한 CNN 모델과 활성화 함수의 비교를 통해 정확도가 높은 알고리즘을 채택하였다.

이를 통해 버스 정류장, 백화점 등 유동인구가 많은 지역의 사용자들에게 주변 환경(날씨, 계절)을 통한 힐링 서비스와 사용자 감정에 맞는 광고서비스를 제공할 수 있다. 이는 옥외 광고에서 광고 정보가 고객에게 일방적인 전달에서 고객과의 상호작용으로 발전할 수 있다는 것을 보여준다. 기업은 개인화 맞춤 감정 마케팅, 옥외 패널과 스마트폰 연동으로 인한 접근성 향상을 통해 마케팅 효과를 극대화할 수 있고 사용자 서비스의 만족도가 향상될 것으로 기대된다.

본 서비스는 옥외 광고서비스로 국한되어 있지만, 표정 인식 알고리즘을 활용하여 다양한 플랫폼에서 사용자의 감정 인식을 활용해서 다양한 분야에서 감정 마케팅을 활용하고 개발할 수 있을 것이다.

참 고 문 헌

[1] Hwang Seol and Myoung Chun, “The Big Data Analytics Regarding the ‘Journal of OOH Advertising Research’ Articles: From 2004 to 2020”, *KOREAN ACADEMY OF OOH ADVERTISING*, 17(4), 5~28, 2020

[2] 이은주, “코엑스 광장에 몰아치던 파도, iF 디자인상 최고점”, *중앙일보*, 2021, <http://www.joongang.co.kr/article/24040835>

[3] 이혜미, “2021년 상반기 광고 시장 결산 및 하반기 전망”, *제일기획*, 8월호, No.5, 2021

[4] Yoon Yang and Hye Ree Koo, “The Effects of Mood, Message Frame and Product Involvement on Product Attitude and Purchase Intention”, *The Korean Psychological Association*, 7(1), 1~22, 2006

[5] In-kyu Choi, Ha-eun and Jisang Yoo, “Facial Expression Classification Using Deep Convolutional Neural Network”, *Journal of Electrical Engineering & Technology*, 13(1), 485~492, 2018

[6] Diganta Misra, “Mish: A Self Regularized Non-Monotonic Activation Function”, *31st British Machine Vision Conference(BMVC)*, virtual, 2020

본 논문은 과학기술정보통신부 정보통신창의인재양성사업의 지원을 통해 수행한 ICT 멘토링 프로젝트 결과물입니다.