

# CNN을 이용한 심리 상담 서비스에 관한 연구

\*김정욱, 강병헌, 김민규, 유승한  
인천대학교 컴퓨터공학과

[neity16@daum.net](mailto:neity16@daum.net), [tsl\\_kiss@naver.com](mailto:tsl_kiss@naver.com), [kimm3161@naver.com](mailto:kimm3161@naver.com), [gankyoun@naver.com](mailto:gankyoun@naver.com)

## Psychological Counseling Service using CNN (Convolutional Neural Network)

\*Jungwook Kim. Byunghun Kang. Mingyu Kim. Seunghan Yoo  
Department of Computer Engineering, Incheon National University

### 요 약

CNN(Convolution Neural Network)은 합성곱(Convolution)을 이용해서 시각적 이미지를 분석하는데 사용되는 인공지능 기술이다. 본 논문에서는 CNN을 이용한 실시간 심리 상담 서비스에 대해 논한다. 상담 서비스에 심리학과 CNN을 접목시킴으로써 내담자의 사진을 심리학적 비언어 행동을 기반으로 분석하여 내담자의 예상 심리를 파악하고, 유의미한 상담 자료를 생성해 상담의 질을 향상시킬 수 있도록 한다.

### 1. 서 론

#### 1.1 코로나19와 비대면

최근 코로나19로 인해 많은 일상이 바뀌었다. 그리고 그 변화된 방식의 중심에는 만나지 않는 방식의 '비대면'이 있다. 학생들의 등교가 힘들어지자 비대면 수업이 이제는 당연해졌고, 심지어 채용을 하는 기업에서도 온라인 면접이나 모든 절차 혹은 일부 절차를 비대면 방식으로 진행하고 있다. 이렇듯 많은 일상이 바뀐 지금의 시대를 '포스트 코로나'시대라고 부른다. 이제는 코로나19가 있기 전의 상황으로는 돌아가기 힘들고, 앞으로 '비대면'은 더욱 많은 방식으로 우리의 생활에 반영 될 것이다. 게다가 이렇게 생활 방식이 변화함에 따라 하루가 다르게 기술도 함께 변하고 있는 모습을 볼 수 있다. 그리고 본 논문에서 다루는 상담 시스템은 결과적으로 사용자에게 '카메라'와 제공되는 'S/W'만 있다면 비대면 방식으로 진행이 가능 한 시스템이기 때문에 이러한 시대적 상황에도 충분한 사용자의 수요를 확보하고 필요성을 가지고 있다고 말할 수 있다.

#### 1.2 다양한 상담 유형

상담에는 다양한 유형이 있다. 작은 범위는 가정에서 이루어지는 '부모-자식' 간의 상담이나, 학교에서 이루어지는 '교사-학생' 간의 상담, 그리고 동급생들 사이에서

발생하는 '친구-친구' 간의 상담 등이 있다. 그리고 큰 범위로서는 여러 내담자와 한명의 상담자로 이루어진 '집단상담'의 형태도 있다. 여기서 중요한 점은 이러한 다양한 상담에서 '상담자'는 전문가가 아닌 경우가 더 많다는 것이다. 그래서 '비전문가'들이 상담을 진행하는 상황에 있어서 '전문가'의 역할을 해주는 S/W를 만들기 위해 연구하였고, '내담자'의 상담 화면에서 비언어적인 행동들을 심리학에 근거하여 분석한 뒤 현재 상황에 맞추어 내담자가 느낄 수 있는 심리를 예측해보려고 한다. 이어서 해당 연구의 필요성과 시스템 구조 및 구체적인 개발 방법에 대해 논한다.

### 2. 본 론

#### 2.1 인공지능 및 심리학과 행동분석의 융합

[1][2] 행동분석은 사람이나 동물을 대상으로 다양한 상황에서 의미하는 바를 분석하는 것으로 인공지능을 적용한 행동 분석의 사례는 많다. 하지만 본 연구의 차이점은 과학적인 분석보다는 심리학을 융합한 데에 있다. [3] 심리학을 기반으로 비언어적 행동을 분석하는데 단순히 어떤 행동이 어떤 의미를 가진 것에 그치는 것이 아니라, '상담'이라는 상황에 맞추어 의미를 다시 한 번 분석한다. 예를 들어 상담자가 '어머니'라는 키워드를 언급했을 때 내담자가 턱을 만졌고 심리학적 분석으로 '불안'

과 ‘초조’ 라는 키워드가 나왔다고 가정해보자. 그렇다면 우리는 해당 키워드만 보여주는 것이 아니라, “현재 대화 주제에 대해 불편해하고 있을 수 있음” 이라고 띄워준다. 즉, 특정 감정에 대한 키워드를 중심으로 ‘상담’이라는 상황에 맞추어 의미를 녹여내서 상담자가 빠르게 상황을 인지하고 판단할 수 있도록 한다. 따라서 심리학을 인공지능과 행동분석과 융합하였기 때문에 보다 더 유의미한 자료를 만들 수 있는 모습을 볼 수 있다.

## 2.2 시스템 구조

오픈소스 영상처리 라이브러리인 OpenCV를 이용하여 ‘내담자’의 행동을 1차적으로 분류하여 인식 및 캡처한다. 그리고 해당 이미지를 Keras라는 인공지능 라이브러리를 통해 구현된 CNN모델에 입력되어 2차 세부 분류가 이루어진다. 그리고 분류된 이미지와 결과는 AWS의 스토리지인 S3에 저장된다. 그리고 마지막으로 SpringBoot로 제작된 Web Server를 통해 S3에 저장된 이미지와 분류된 레이블 값을 읽어서 웹을 통한 실시간 상담 서비스의 상담 화면에 띄워주게 된다. 그렇게 되면 상담자는 해당 내용을 보고 판단하여 상담 기록에 기입하게 된다.

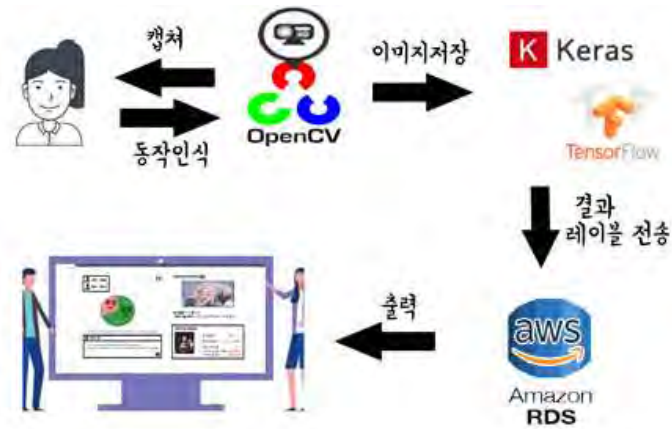


그림 1. 전체 시스템 구조

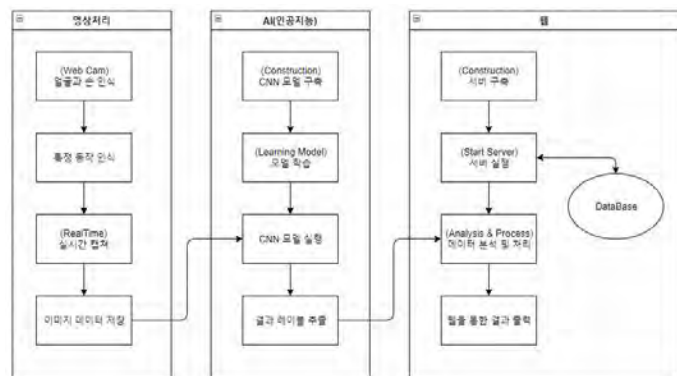


그림 2. 아키텍처 구성도

## 2.3 개발 방법

### 1) 영상처리

해당 프로젝트에서 영상처리의 역할은 웹캠을 통해 실시간으로 촬영되는 내담자의 이미지에서 사전에 정의한 비언어적 행동을 내담자가 취했을 때 해당 영상처리 시스템에서 자동으로 탐지하고 캡처한 뒤 CNN 모델을 통해 분류하도록 Flask 서버로 전송하는 것이다.

사전에 정의한 심리학적 의미가 담긴 비언어적 행동을 실시간 영상에서 자동으로 탐지하도록 구현하기 위해서 [4] TF-openpose라는 라이브러리를 활용했으며, 해당 라이브러리는 영상 속의 사람을 탐지해 얼굴의 눈, 코, 입, 귀 그리고 신체의 각 joint(관절)의 좌표값들을 Centers[] 라는 배열에 담아 Main 함수로 반환해주고 영상에 프레임 단위로 탐지된 부위들을 점으로 연결해 뼈대 형태로 나타내준다.

이렇게 TF-openpose를 활용해 프레임 단위로 전달받은 각 신체부위의 좌표들의 거리를 이용하여 내담자가 비언어적 행동(해당 프로젝트에서는 턱, 코, 이마를 만지는 각각의 행위를 비언어적 행동으로 정의)을 취할 때 영상처리 시스템이 자동으로 탐지하도록 알고리즘을 구현했으며, CNN 모델의 분류 정확도를 높이기 위하여 상담자가 내담자의 화면에서 불필요한 배경 영역을 제외하고 상담반에 해당하는 영역만 이미지로 전송하게끔 ROI(관심 영역)를 직접 지정하도록 구현했다.

이미지를 전송하는 부분엔 Rest API가 활용되었으며 영상처리 시스템이 내담자의 비언어적 행위를 탐지할 때마다 실시간으로 해당 프레임이 행동 분류를 위한 Flask 서버로 전송되게 된다.

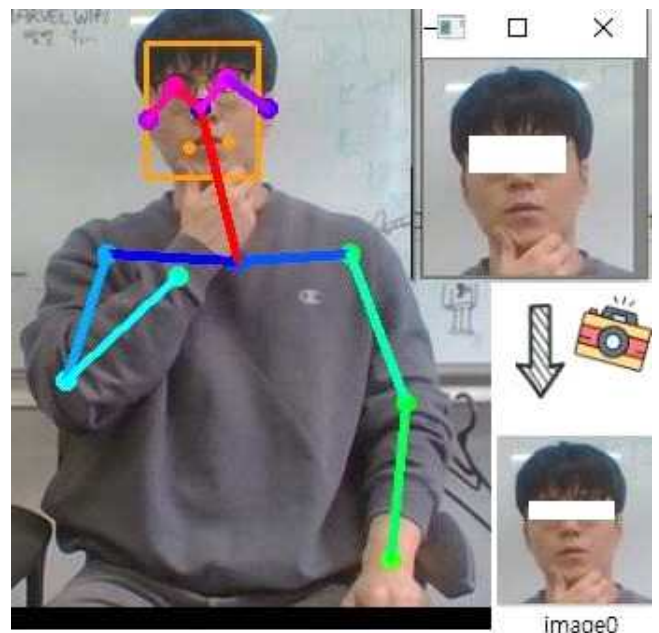


그림 3 내담자의 화면 및 ROI 영역

## 2) 인공지능

최종 목표는 우리가 원하는 동작을 분류하는 CNN모델을 구축하는 것이다.[5] python언어를 기반으로 인공지능 라이브러리인 Keras를 사용하였다.[6] 그리고 모델의 정확도 향상을 위해 사전에 훈련된 신경망 모델인 VGG16을 선택했다. 총 1400만장, 1000가지 종류를 가진 Image로 학습된 VGG16은 [합성곱 기반] 과 [조밀연결 계층] 두 가지로 구성되어 있다. 우리가 원하는 동작을 분류하기 위해서 일반적인 특징을 가지는 [합성곱 기반]에 새로운 분류기를 추가해서 CNN 모델을 구성했다. 해당 모델은 over-fitting과 gradient vanishing을 개선하기 위해 Relu 활성화 함수를 사용했다. 그리고 적합한 모델 구축을 위해서 ‘미세조정’ 기법을 사용해서 VGG16의 일부 Layer의 가중치 동결을 해제하고 우리의 데이터로 추가 학습을 시켰다. 그리고 HyperParameter의 최적화를 위해서는 Kas에서 제공하는 Tuner를 사용했다.

학습 데이터 형태가 ‘이미지’이기 때문에, 데이터를 모으기 위해 Google’s Crawler를 사용하여 국내 포털 및 해외 포털에서 이미지를 준비했다. 이렇게 준비된 이미지는 여러 단계에 걸쳐서 필터링 된다. 1차로 수작업을 통해서 관련 없는 이미지를 거른다. 그리고 2차로 Visipics라는 S/W를 통해 중복 데이터를 거른다. 그리고 3차로 이미지의 사람 범위를 모두 같은 영역으로 맞추기 위해서 이미지 수정 과정을 거친다.

이렇게 모델을 구축하고 학습시킬 데이터를 선별하여 구축한 모델의 최종 정확도는 80중, 후반의 Accuracy를 유지하고 있다.

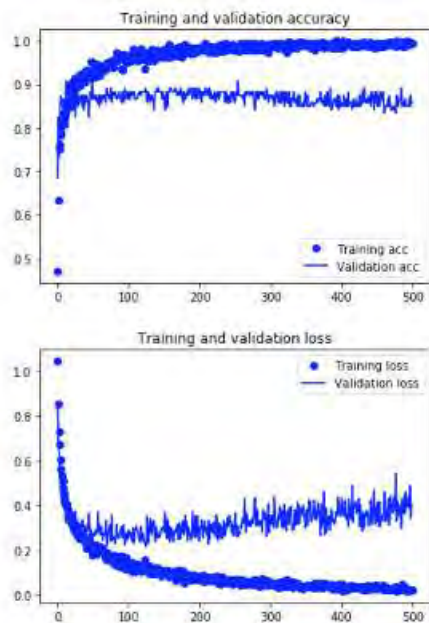


그림 4. 최종 CNN Accuracy

## 3) 웹

웹은 Java 기반의 SpringBoot를 사용하여 실시간 상담 서비스를 구축하기 위한 기본적인 ‘상담자’, ‘내담자’, ‘게시판’ 3가지 객체에 대한 DB를 구성했다. 그리고 실시간으로 인공지능에서 분류된 이미지와 레이블 같은 자원들을 저장하고 사용하기 위해 AWS의 S3 Storage를 사용했다. 따라서 웹에서 핵심은 실시간 Resource를 S3객체를 통해 저장하고, 실시간 상담 화면에서 가져오는 부분이라고 할 수 있다.



그림 5. 분류된 이미지

## 4) 시스템 연동

앞에서 구축된 [영상처리] [인공지능] [웹]간 서로 데이터를 주고받기 위한 기술 스택은 다음과 같다. 우선 CNN 모델을 ‘Load’ 하고 ‘Use’할 수 있는 Flask 서버를 구축했다. Flask 서버에서는 들어온 이미지를 CNN 모델의 입력 형식인 150x150에 맞추어 이미지를 resizing하고 모델에 입력하고 해당 동작에 따라 레이블을 분류한다. 그리고 영상처리에서 캡처된 이미지나 인공지능에서 분류된 데이터를 넘길 때에는 SpringBoot 서버에 REST API를 통해 json 방식의 데이터로 구성하여 통신했다.

## 3. 결론 및 향후 연구방향

본 논문에서는 CNN을 이용하고, 심리학을 기반으로 한 심리 상담 서비스에 대한 연구의 필요성과 시스템 구조 및 구체적인 개발 방법에 대하여 논하였다. 내담자의 행동을 분석하고 심리학적 키워드를 추출한 후 상담이라는 상황에 맞는 텍스트로 변환하여 상담자에게 제공해준다. 이렇게 유의미한 자료를 제공함으로써 상담의 질적 향상을 이룰 수 있다는 사실을 알았다. 하지만 더 유용성과 효율성을 가지기 위해서 S/W가 분류하고 보여줄 수 있는 범위는 넓어져야 한다. 따라서 여기서 우리가 단순 영상처리를 통한 행동분석이 아닌 인공지능을 접목시킨

이유가 등장한다. 바로 시스템의 확장 때문이다.

현재 인공지능을 접목시켜 사람의 감정을 읽는 Microsoft의 표정 분석 시스템이 오픈소스로 존재한다. 그리고 IBM사의 Personality Insights는 글을 통해서 사람의 개성, 필요, 가치관 등을 분석하는 시스템으로 글을 통해서 사람을 판단한다고 할 수 있다.[7] 이렇게 인공지능을 통한 다양한 측면의 분석을 추가 적용한다면 앞서 행동뿐만 아니라 더욱 넓은 범위에서 더 다양한 자료를 상담자에게 제공할 수 있다. 따라서 우리는 점차 다양한 인공지능을 접목시켜 보다 완성도 있고 유용한 상담 시스템을 개발할 계획이다.

‘본 논문은 과학기술정보통신부 정보통신창의인재양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트의 결과물입니다.’

#### 4. 참고 문헌

- [1]. 이은지. "인간과 인공지능의 상호작용." 국내박사학위논문 고려대학교 대학원, 2020. 서울
- [2]. 박연수. "상담지향 챗봇 서비스의 담화 연구." 국내석사학위논문 이화여자대학교 대학원, 2019. 서울
- [3]. 오시현. "Social-Emotional Intervening Pathways between Attachment Theory and Mental Health Symptoms: Finding Relevant Treatment Targets in Youth." 국내석사학위논문 덕성여자대학교 대학원, 2019. 서울
- [4]. Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh. Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 7291-7299
- [5]. 팜동. "CNN 기반 얼굴감정인식과 새로운 손실함수에 대한 연구." 국내석사학위논문 숭실대학교 대학원, 2019. 서울
- [6]. Phung Van Hiep. "CNN-based Model Design for High Accuracy Image Classification on Cloud and Crop Pest." 국내박사학위논문 Hanbat National University, 2020. 대전
- [7]. 장진후. "딥러닝 기법을 이용한 자기소개서 성향 분석." 국내석사학위논문 단국대학교 정보·지식재산대학원, 2019. 경기도