

## 딥러닝 기반 수어 교육 온라인 플랫폼 구현

김 미 서<sup>1</sup> · 유 문 희<sup>1</sup> · 장 지 현<sup>1</sup> · 최 정 윤<sup>1</sup> · 나 고 은<sup>2</sup> · 김 미 숙<sup>3</sup> · 나 정 조<sup>4\*</sup><sup>1</sup>덕성여자대학교 IT미디어공학과 학사과정 <sup>2</sup>뉴욕시립대학교 헌터칼리지 상담학과 교수<sup>3</sup>강릉원주대학교 컴퓨터공학과 강사 <sup>4\*</sup>덕성여자대학교 IT미디어공학과 교수

## Implementation of deep learning-based sign language acquisition online platform

Mi-Seo Kim<sup>1</sup> · Moon-Hee Yoo<sup>1</sup> · Ji-Hyeon Jang<sup>1</sup> · Jeong-Yun Choi<sup>1</sup> · Go-Eun Na<sup>2</sup> · Mi-Suk Kim<sup>3</sup> · Jung-Jo Na<sup>4\*</sup><sup>1</sup>Bachelor's course, Department of IT Media Engineering, Duksung Women's University, Seoul 01369, Korea<sup>2</sup>Professor, Department of Counseling, Hunter College, CUNY, New York 10065, USA<sup>3</sup>Instructor, Department of Computer Engineering, Gangneung-Wonju National University, Gangwon 26403, Korea<sup>4\*</sup>Professor, Department of IT Media Engineering, Duksung Women's University, Seoul 01369, Korea

### [요 약]

청각 장애인들은 수어를 주된 언어로 사용하지만, 국내 수어 교육 관련 콘텐츠와 수어 습득 환경은 접근성이 낮고 열악하다. 2022년 수어 교육 관련 서비스 플랫폼은 총 15개로, 수어 학습 형태는 주로 강의를 보고 동작을 따라하도록 유도하고 있다. 하지만 학습자에게 동작에 대한 실시간 피드백은 제공하지 못하고 있다. 따라서 본 논문에서는 딥러닝을 이용한 수어 동작 판별 모델을 통해 학습자가 올바른 동작을 취하고 이에 대한 결과를 피드백으로 제공하는 실시간 플랫폼을 제안한다. 수어 동작 판별 모델은 CNN과 LSTM을 결합한 모델로, OpenPose와 MediaPipe를 사용하였으며 팔과 손, 손가락을 추적한 데이터로 학습시켜 구현하였다. 본 수어 교육 플랫폼은 딥러닝을 통해 학습자가 스스로 올바른 수형인지를 판단할 수 있으며, 양방향 학습이 가능하여 수어를 익히는 데에 도움이 될 수 있다.

### [Abstract]

Sign language is the main communication method for people with hearing impairments. However, the educational system and learning content for sign language acquisition have been characterized by a lack of accessibility and low quality in Korea. In 2022, 15 educational platforms regarding sign language acquisition exist. The formats mainly consist of watching videos and following pre-recorded hand movements, but real-time feedback is not provided about learned content. Thus, this study proposes a platform to teach sign language that provides real-time feedback on the proper hand movements of sign language. Based on CNN and LSTM, this model of distinguishing sign language movements uses a data-tracking system that tracks learners' arms, hands, and fingers using OpenPose and MediaPipe. Through deep learning, the platform can help learners identify accurate hand movements and learn sign language interactively.

**색인어** : 딥러닝, 동작 인식, 교육, 수어, 청각 장애인**Keyword** : Deep Learning, Motion Recognition, Education, Sign Language, People with hearing impairment<http://dx.doi.org/10.9728/dcs.2022.23.11.2147>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 31 August 2022; Revised 07 October 2022

Accepted 31 October 2022

**\*Corresponding Author; Jung-Jo Na****Tel:** **E-mail:** [jungjona@duksung.ac.kr](mailto:jungjona@duksung.ac.kr)

## I. 서 론

### 1-1 연구 배경 및 목적

지난 2020년 3월 9일 한국농아인협회, 한국농아대학생연합회 등 청각 장애 관련 단체들이 그들의 학습권 보장을 촉구하기 위한 ‘코로나19에 따른 청각장애 학생들의 학습권 보장 촉구 기자회견’을 열었다. 코로나바이러스 감염증(COVID-19) 사태가 장기화됨에 따라 우리 사회는 많은 변화를 겪고 있다. 그중 가장 큰 변화는 사회적 거리두기의 지속화에 따른 비대면 환경이 일상이 된 것이다. 하지만 이러한 변화는 청각장애인들에게 소외감과 학습권 침해라는 변화를 초래하였으며, 그들의 학습 환경이 얼마나 열악한지 드러나게 되었다.

청각 장애인들의 수어 학습 환경은 대부분 오프라인을 통해서 한정적으로 이루어지고 있었다. 국내 장애인 등록 현황 [1]에 따르면, 2021년을 기준으로 등록된 청각 장애인의 수는 411,749명으로 지체 장애인 다음으로 두 번째로 높은 비율을 차지하고 있음에도 불구하고, 청각장애인 고용차별 및 고용개선방안 실태조사[2]에 따르면 그들의 수어 습득 경로로 과반이 수어를 학교 선배나 친구로부터 배웠다고 응답하였다. 한국농아인협회를 통해서 수어 교실을 수강할 수도 있지만 자격증 취득을 위한 수업이 과반이거나 참여 인원이 10명 내외의 소수로 진행되고, 수강료를 내고 별도의 수강 신청을 해야 하는 등 복잡한 과정이 필요로 된다. 청각 장애인의 92.8%가 후천적으로 발생한다는 점[3]을 고려하였을 때, 청각 장애인들의 적극적이고 자발적인 노력이 없이는 수어 학습이 어려운 실정이다.



이처럼 열악한 그들의 수어 학습 환경에 비대면 상황까지 더해지면서 그들의 학습권은 보장받지 못하고 있다. 이에 본 논문에서는 청각장애인들이 시공간이나 비용 등 학습에 제약받지 않고 제1 언어인 수어를 배울 기회와 환경을 온라인으로 제공함으로써 보다 나은 학습 환경에 기여하고자 서비스 플랫폼을 기획하게 되었다.

### 1-2 연구 필요성

최근 동작 인식과 딥러닝에 기반한 한국수어 서비스로는 주로 실시간 수어 통역 서비스에 대한 연구가 활발하게 이루어지고 있으며 실제 공공장소 내 실시간 수어 통역 서비스가 제공되고 있다. 교육을 목적으로 하는 연구로는 립모션 센서를 이용하거나 지화(지문자, 지숫자)를 인식하는 등의 연구가 이루어지고 있으나 고가의 장비가 필요하거나 수어의 연속적인 동작을 인식하지 못한다는 한계점이 있다. 수어를 학습할 수 있는 온라인 환경으로는 국립국어원 한국수어사전 웹사이트와 전라북도 교육청 교육 연구정보원에서 지원하는 전북교육 영상 웹사이트 내 ‘수화 교실’과 애플리케이션 「사랑의 수화 교실」이 있다.

표 1. 국내 수어 통역 서비스 사례

Table 1. Domestic sign language interpretation service cases

| Title   | Image   |
|---|---|
| Gimpo Airport AI Sign Language Information System |  |
| Seoul Subway Video Phone                          |  |

각 환경에서는 수어 단어 및 일상생활 수어 동작에 대한 영상을 제공하지만, 이는 사용자가 강의 영상을 보는 것에서 그쳐 수어 동작을 따라 해보는 것을 유도하거나 동작에 대한 피드백이 불가하다. 이러한 한계를 극복하고자 웹캠을 이용해 수어 동작을 촬영하고 동작에 대한 결과를 제공함으로써 수어 학습자의 교육 접근성을 높이고 양방향 소통이 이루어지는 플랫폼을 구현하고자 한다.

## II. 수어 교육 서비스

### 2-1 수어 정의 및 분류

청각 장애인은 선천적이거나 후천적인 요인으로 청각에 이상이 생겨 소리를 듣지 못하는 사람을 뜻한다 [4]. ‘수화’(手話, Sign language)는 이런 청각 장애가 있는 사람들이 손과 손가락의 모양, 손바닥의 방향, 손의 위치, 손의 움직임을 달리하여 의미를 전달하는 언어이다 [5]. 최근에는 이를 강조하기 위해 ‘수화’ 뒤에 ‘언어’를 붙여 ‘수화언어’라고 하거나 줄여서 ‘수어’라고 한다. ‘한국수어’는 한국어와는 문법 체계가 다른, 대한민국 농인의 고유한 언어이다 [6].

### 2-2 수어 교육에 대한 환경 분석

청각 장애인들은 음성 언어 대신 수어를 주로 제1 언어로 사용한다. 우리나라의 수어는 2016년 ‘한국수화언어법’이 제정되면서 ‘한국수화언어’는 국어와 동등한 자격을 가진 농인의 고유한 언어가 되었다 [7]. 하지만 국내 청각 장애인 수에 비해 수어를 배울 수 있는 환경은 제한적이다.

교육부 내부 자료에 따르면 농학교 교사 98.5%가 “수어를 할 수 있다”고 응답했지만, 응답자 중 수화 통역사 자격증을

가지고 있는 사람은 6.1%에 불과했다. 실제 수어를 구사하지 못하는 교사들은 입말로 수업을 진행하기 때문에 학생들은 수업을 따라가기에 어려움을 겪는다 [8].

2022년 10월 현재 수어 사전과 통역을 모두 포함하여 한국 수어를 학습할 수 있는 서비스 플랫폼은 총 15개가 존재한다.

수어 교육 애플리케이션은 안드로이드 운영체제를 위한 Google 플레이 스토어에 등록된 앱의 수는 8개, iOS 운영체제를 위한 Apple 앱스토어에 등록된 앱의 수는 3개이다. 수어 교육 웹사이트는 4개이며, 웹은 모두 지방자치단체에서 운영하고 있다. 수어 교육 플랫폼 개발을 위해서 제공되고 있는 서비스 플랫폼을 <표 2>와 같이 수어 동작에 대한 강의 영상을 제공하는지, 수어 동작을 잘 습득하였는지 확인 또는 테스트를 제공하는지, 사용자의 수어 동작을 실시간으로 인식하는지 3가지 조건으로 비교하였다. 이때, Google 플레이 스토어의 서비스 플랫폼은 스토어 내 상위 5개를 선정하여 비교를 진행하였다.

**표 2.** 국내 수어 교육 서비스 플랫폼의 비교분석

**Table 2.** Comparative analysis of sign language education platform in Korea

| No. | Platform | Name  | Condition 1: applying class video | Condition 2: tracking and applying questions | Condition 3: real-time recognition | Purpose of provision  |
|-----|----------|---|-----------------------------------|--|------------------------------------|-----------------------|
| 1   | Web      | Korean Sign Language Dictionary                         | O                                 | X  | X                                  | Dictionary, Education |
| 2   |          | Sign Language Standard Dictionary                       | X                                 | X  | X                                  | Dictionary            |
| 3   |          | Sign Language Class                                     | O                                 | X  | X                                  | Dictionary, Education |
| 4   |          | Gangwon-do Sign Language Cultural Center                | O                                 | X  | X                                  | Education             |
| 5   | Android  | Love Sign Language Class                                | O                                 | X  | X                                  | Dictionary, Education |
| 6   |          | Open Sign Language                                      | O                                 | X  | X                                  | Dictionary, Education |
| 7   |          | Sign Language translator                                | O                                 | X  | X                                  | Translation           |
| 8   |          | Sign Language Game                                      | X                                 | O  | X                                  | Education             |
| 9   |          | Sign Language Translator (AI Sign Language Recognition) | X                                 | X  | O                                  | Translation           |
| 10  | iOS      | Sign Language Dictionary                                | O                                 | X  | X                                  | Dictionary            |
| 11  |          | Su-Hangul   | X                                 | O  | X                                  | Education             |
| 12  |          | ByHands   | O                                 | X  | X                                  | Education             |

비교 결과 수어 사전의 제공과 이를 통한 학습을 목적으로 한 서비스가 가장 많았으며 실시간으로 사용자의 동작을 인식하는 서비스는 1개뿐이었고, 3가지 조건을 모두 충족하는 서비스는 존재하지 않았다. 따라서 본 논문에서는 <표 2>에서 제시한 3가지 조건을 만족하는 서비스 플랫폼을 제안한다.

‘한국수화언어법’에 명시한 것처럼 국가는 청각 장애인의 교육받을 권리를 보장해 주어야 하며[7], 국가와 지방자치단체 또한 교육 서비스 제공에 대한 노력을 꾸준히 하여야 한다. 이에 본 수어 교육 서비스 플랫폼은 열악한 수어 교육 환경 속에서 청각 장애인을 위한 수어 교육 환경 조성을 위하여 설계되었다.

## 2-3 선행연구

본 논문에서 다루고자 하는 영상을 이용한 수어 동작 인식과 관련된 논문 중 2021년도 이후에 발간된 국내 등재 학술지 논문 3편을 사용된 딥러닝 모델을 중점으로 비교해보았다.

‘딥러닝 기반 OpenPose를 이용한 수화 동작 인식에 관한 연구’에서는 수화 영상을 이미지로 추출한 데이터 셋을 사용하여 CNN 모델을 학습시켜 수화 동작을 인식하였다. 이때 OpenPose를 이용하여 얼굴, 손, 손의 위치 등 비수지 요소를 포함한 수화 영상을 포함 시켜 모델의 인식률을 개선했다[9]. ‘청각장애인의 수어 교육을 위한 MediaPipe 활용 수어 학습 보조 시스템 개발’에서는 OpenCV 라이브러리와 MediaPipe를 이용해 손과 손가락을 추적하여 수어 동작과 Landmark를 추출하였다. 이후 수어 동작 이미지는 CNN 분류기에, Landmark는 SVM 분류기를 통해 분류하여 두 결과가 일치하는지 비교 후, 올바른 수형인지 사용자에게 알려주는 시스템을 개발하였다[10]. ‘CNN기반의 온라인 수어통역 상담 시스템에 관한 연구’에서는 사용자의 수어 동작을 인식한 뒤 이를 텍스트로 변환하여 사용자에게 제공하는 시스템을 개발하였다. OpenCV의 Teachable Machine 기술을 이용하여 사용자의 수어 동작을 인식한 후 라벨링 과정을 거친 뒤 CNN 모델로 학습을 진행하였다[11].

**표 3.** 수어 동작 인식 선행연구

**Table 3.** A prior research on sign language gesture recognition

| Title   | Model | Pose estimation   | Research purpose                                |
|---|-------|-------------------|---|
| A Study on Korea Sign Language Motion Recognition Using OpenPose Based on Deep Learning                               | CNN   | OpenPose          | Sign language gesture recognition               |
| Development of a Sign Language Learning Assistance System using Mediapipe for Sign Language Education of Deaf-Mutlity | CNN   | MediaPipe, OpenCV | Converting text after recognizing sign language |
| CNN-based Online Sign Language Translation Counseling System  | CNN   | OpenCV            | Translation after recognition of sign language  |

## 2-4 서비스 플랫폼 기술 동향

OpenPose와 MediaPipe는 사람의 동작을 예측하는 기술이다. CNN과 LSTM은 딥러닝 알고리즘의 한 종류로 본 서비스 플랫폼에서는 수어 동작을 인식하고 분류하는 데에 사용된다.

### 1) OpenPose

OpenPose는 Carnegie Mellon University에서 개발한 오픈소스 라이브러리이다. 사진 또는 영상에서 실시간으로 인체, 손, 얼굴 및 발의 키포인트(총 135개)를 추론할 수 있고, 이를 입력된 사진 또는 영상 위에 시각화하여 보여준다.

### 2) MediaPipe

MediaPipe는 구글에서 개발한 오픈소스 프레임워크이다. 사진 또는 영상에서 실시간으로 사람의 동작을 감지하여 몸 관절에서 33개, 손 관절에서 21개의 3D 랜드마크를 추론할 수 있다.

### 3) CNN

CNN(convolutional neural network)은 인간의 시신경 구조를 모방한 기술로 이미지 분류에 주로 사용되는 신경망이다. CNN은 입력 데이터의 특징을 추출하는 Feature Extraction을 수행하고 Convolution Layer와 Pooling Layer이 여러 번 쌓여 특징을 추출하게 된다.

### 4) LSTM

LSTM(long short term memory)은 RNN(recurrent neural network)의 한 종류로 최근에 많이 활용되고 있는 신경망이다. LSTM은 기존의 RNN 모델의 단점인 장기 의존성 문제를 해결한다. 시계열 데이터 중에서도 긴 기간을 가지고 있는 데이터를 활용할 때 적합한 학습 수행 능력을 갖추고 있다. Carnegie Mellon University에서 개발한 오픈소스 라이브러리이다.



그림 1. 인간 동작 예측 기술 (a) 오픈포즈 (b) 미디어파이프

Fig. 1. human pose estimation (a) OpenPose (b) MediaPipe

## III. 수어 교육 서비스 플랫폼 개발

### 3-1 서비스 플랫폼 주요 기능 소개

본 서비스 플랫폼은 수어를 학습하고자 하는 누구에게나 수어를 자율적으로 학습할 기회를 제공하고 수어 교육의 접근성을 높이고자 개발되었다. 서비스의 주요 기능으로는 수강하기, 테스트하기, 마이페이지 내 수강 현황, 성적표, 오답 노트로 구성이 되어있다.

#### 1) 수강하기

수강하기에서는 사용자가 원하는 단계와 커리큘럼을 선택하여 강의 영상을 시청한다. 시청 후 사용자는 해당 단어에 대한 수어 동작을 웹캠으로 촬영하고 동작에 대한 결과를 제공한다. 결과는 사용자가 행한 동작에 대해 가장 유사하다고 판단된 상위 3개의 단어를 표기한다. 이를 통해 사용자는 자신의 동작이 어떤 단어와 유사한지 알 수 있다.

수강하기의 단계는 언어의 학습과 응용 순서를 고려하여 초급, 중급, 고급으로 나누어지며 각 단계 내에서 커리큘럼으로 세분화된다. 초급 단계는 지문자로 고유명사를 표현할 수 있도록 [ㄱ], [ㄴ], [ㄷ] 자음 3개와 [ㅏ], [ㅑ], [ㅓ] 모음 3개로 구성되어 있다. 중급 단계는 일상생활에서 많이 쓰이는 표현으로 간단한 의사소통을 표현할 수 있도록 ‘안녕하세요’, ‘괜찮습니다’, ‘미안합니다’, ‘감사합니다’, ‘좋아하다’, ‘싫어하다’ 6개로 구성되어 있다. 고급 단계는 문장을 구성하는 성분인 단어들로 국립국어연구원의 한국어 학습용 어휘선정 결과 보고서를 참고[12]하여 선정하였으며 ‘마음’, ‘사람’, ‘생각’, ‘친구’, ‘학교’로 구성되어 있다.

#### 2) 테스트하기

테스트하기는 사용자가 원하는 단계를 선택하여 진행할 수 있다. 테스트 문제는 각 단계의 단어들이 무작위로 제시된다. 테스트는 각 단계에서 5개의 단어로 평가된다. 사용자는 제시되는 각 단어를 10초 이내로 수어로 나타낸 후, 다음 버튼을 클릭하고 5개 단어 모두를 수행한 후 제출하기 버튼을 누르게 된다. 이후 테스트의 결과는 곧장 실행화면에서 확인할 수 있다. 테스트 결과는 서버에 자동 저장이 되어 추후 성적표와 오답 노트에서 다시 확인할 수 있다.

#### 3) 마이페이지

마이페이지에서는 수강 현황, 성적표, 오답 노트를 확인할 수 있다. 수강 현황은 마이페이지의 첫 화면에서 그래프로 제공된다. 성적표는 테스트에 대한 결과를 확인할 수 있고 오답 노트에서는 테스트에서 틀린 단어들을 확인하고 복습할 수 있다.



### 3-2 플랫폼 디자인 개발

#### 1) 기능 흐름도

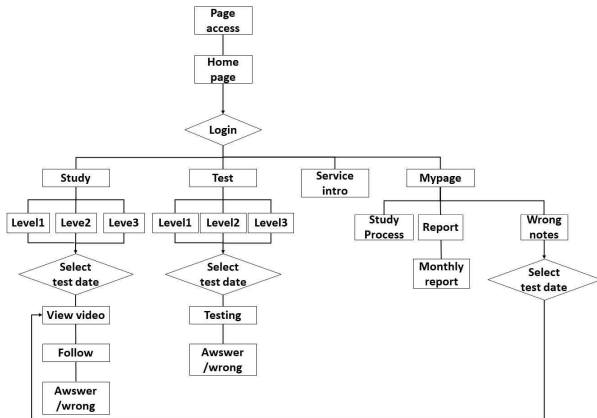


그림 2. 기능 흐름도  
Fig. 2. work flow

<그림 2>은 사용자의 상황에 따른 이용 방식을 나타낸 기능 흐름도이다. 페이지 접속을 한 후 로그인 한 사용자만 이용이 가능하도록 서비스를 제공한다. 로그인 후 플랫폼을 처음 이용하는 신규 사용자를 위하여 ‘서비스 소개’에서 서비스에 대한 소개와 이용 방법을 제공한다. 서비스 이용 방법을 익힌 후에는 수강하기에서 사용자가 원하는 단계를 선택하여 학습하고자 하는 단어의 영상을 시청할 수 있다. 영상 시청이 끝난 후에는 ‘따라하기’를 통하여 동작에 대한 결과를 확인할 수 있다. 더불어 테스트하기에서 자신의 학습 이해도를 확인할 수 있다. 테스트하기를 통하여 나온 결과는 마이페이지의 성적표와 오답 노트를 통해서 다시 확인할 수 있다.

#### 2) 사용자 웹 인터페이스

<그림 3>, <그림 4>, <그림 5>, <그림 6>은 전체 플랫폼에 대한 사용자 인터페이스로 사용자들이 보게 될 시각적 화면이다. <그림 3>과 같이 홈, 수강하기, 테스트하기, 마이페이지의 UI를 디자인하였다. <그림 3>은 홈 화면으로 각 주요 기능으로 이동할 수 있는 메뉴와 간단한 서비스 소개 글이 적혀있다. <그림 4>는 수강하기 화면으로 강의를 수강할 수 있는 메인 재생 화면과 해당 커리큘럼 내의 단어를 선택할 수 있는 사이드바와 강의 이동 버튼, 따라하기 버튼으로 이루어져 있다. <그림 5>는 테스트하기 화면으로 테스트를 진행할 수 있는 웹캠 화면과 사용자가 임의로 촬영을 종료할 수 있는 촬영 종료 버튼, 테스트 자체를 종료할 수 있는 테스트 종료 버튼으로 이루어져 있다. <그림 6>은 마이페이지 화면으로 강의 수강 현황을 알 수 있는 수업 진행 상황을 나타낸 그래프와 성적표, 오답 노트 페이지로 이동할 수 있는 버튼으로 이루어져 있다.



그림 3. 사용자 인터페이스 1  
Fig. 3. User Interface 1

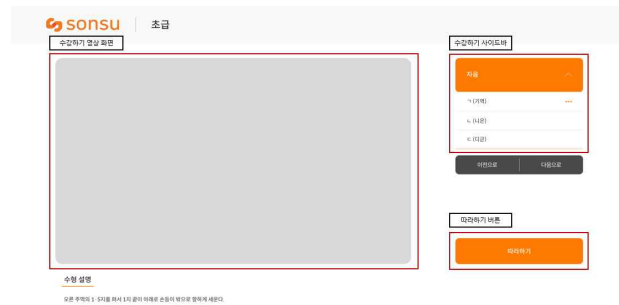


그림 4. 사용자 인터페이스 2  
Fig. 4. User Interface 2



그림 5. 사용자 인터페이스 3  
Fig. 5. User Interface 3



그림 6. 사용자 인터페이스 4  
Fig. 6. User Interface 4

### 3-3 딥러닝 기반 플랫폼 적용 기술

#### 1) 시스템 아키텍처

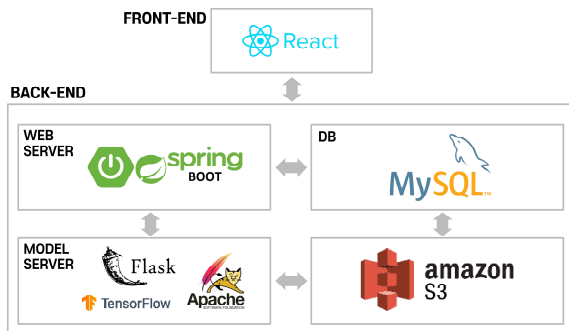


그림 7. 시스템 아키텍처  
Fig. 7. System Architecture

<그림 7>은 해당 웹 서비스를 제작하기 위해 사용한 기술들을 표현한 시스템 아키텍처이다.

사용자는 pc를 이용하여 서비스에 접속 후 웹서버에 요청을 보낸다. 웹서버는 요청에 따라 단계 및 커리큘럼 정보, 단어 정보 등이 저장되어있는 데이터베이스 또는 수어 영상 및 이미지가 저장되어있는 클라우드 스토리지에서 필요한 데이터를 가져오거나 데이터의 삽입 요청을 처리하고 그 결과를 클라이언트에게 보낸다. 클라이언트가 요청 결과를 화면에 출력하면 사용자가 보낸 요청의 결과를 시각적으로 확인할 수 있다.

서비스의 핵심 기능인 딥러닝을 기반으로 하는 사용자의 수어 동작 판별 기능은 python을 기반으로 하는 모델 서버를 따로 구축하여 Java 기반의 서버인 웹서버와 연동할 수 있도록 구성하였다. 사용자가 수어 동작을 수행하고 수행한 단어의 정보와 사용자의 영상을 모델 서버에게 보낸다. 모델 서버에서 딥러닝 수어 동작 판별 모델을 통해 사용자의 동작을 판별하여 결과값을 데이터베이스에 저장하고, 클라이언트의 요청사항에 따라 결과값을 전달한다.

#### 2) 주요 기술 진행 흐름

##### (1) 수강하기

사용자는 수강하기에서 강의 영상을 시청하고, 배운 수어 동작을 따라 할 수 있다. ‘따라하기’ 페이지에서 ‘촬영 시작하기’ 버튼을 눌러 웹캠 촬영을 시작해 수어 동작을 수행한다. 사용자가 ‘촬영 끝내기’ 버튼을 누르면 클라이언트는 모델 서버에게 사용자의 영상에 대한 결과를 요청하게 된다. 모델 서버는 먼저 전달받은 사용자의 영상에서 5프레임당 1장을 추출한다. 추출한 이미지 데이터들을 수어 동작 판별 모델에 입력하여 동작에 대한 결과를 얻어 클라이언트에 전달하게 된다. 클라이언트에 결과가 전달되었다면 웹 페이지에서 ‘결과보기’ 버튼이 활성화된다. 사용자가 ‘결과보기’ 버튼을 클릭하게 되면 수강 현황을 체크하기 위해 웹 서버에 요청을 보내게 되고, 웹 서버는 결과에 따른 화면을 제공한다. 결과는 사용자

가 진행한 동작이 정답과 얼마나 유사한지와, 정답 외에 어떤 단어와 유사했는지를 퍼센트로 제공해 준다.

##### (2) 테스트하기

사용자가 단계를 선택하여 ‘시작하기’ 버튼을 누르면 웹 서버에서 테스트 단어를 랜덤으로 제공해준다. 테스트가 시작되면 사용자는 제시되는 단어를 보고 수어 동작을 수행한다. 그 후 수강하기와 동일하게 모델 서버가 사용자의 동작에 대한 결과를 추출한다. 테스트하기에서는 결과를 클라이언트에 전달하지 않고 웹 서버에 전달하여 데이터베이스에 저장하게 되는데, 이는 테스트를 마친 후에 전체 결과를 제공하기 때문이다. 사용자가 테스트를 마친 후 ‘채점하기’ 버튼을 누르면 문제 번호와 함께 정답, 오답 여부를 확인할 수 있다. 또한 저장된 결과를 이용하여 사용자에게 성적표와 오답 노트 기능을 제공할 수 있다.

#### 3) 개발 환경

본 연구의 개발환경으로 컴퓨터 CPU는 Intel Core i7 processor, 메모리 16GB를 이용하였으며, 웹캠을 이용하였다. 클라이언트와 서버 환경으로는 React (18.1.0), Spring boot (2.7.1), Flask (2.1.2)이고 데이터베이스로 MySQL(8.0.27)을 이용하여 구현했다.

표 4. 개발환경

Table 4. Development Environment

| Variables |             | Specification                                     |
|-----------|-------------|---|
| H/W       | CPU         | Intel Core i7 processor                           |
|           | Memory      | 16 GB   |
|           | Device      | Wecam   |
| S/W       | Environment | Spring boot(2.7.1), Flask (2.1.2), React (18.1.0) |
|           | Data Base   | MySQL (8.0.27)                                    |

#### 4) 딥러닝 수어 동작 판별 알고리즘

##### (1) 데이터 셋

본 서비스 플랫폼은 교육을 목적으로 하므로 표준화된 수어 동작의 학습을 위해서 국립국어원의 표준 한국 수어 영상을 기준으로 데이터 셋이 구축되었다.

초기에는 데이터 구성으로 일상생활 속 많이 쓰이는 문장 6개, 국립국어원에서 제공되는 한국어 학습용 어휘 목록 내 자주 사용되는 명사 어휘 5개를 채택하여 총 11개의 Label로 구성하였다. 각 Label에 대한 원본 데이터로 국립국어원 한국수어사전에서 제공되는 영상과 AI HUB에서 제공되는 한국수어 영상 데이터셋을 사용하여 총 11개의 영상을 수집하였다. 추가로 수어는 팔과 손을 이용해서 표현하는 언어이기 때문에 정확한 동작의 인식과 딥러닝 모델의 학습률 향상을 위해서 동작 인식 기술인 MediaPipe를 적용하여 수어 동작 자의 팔

과 손가락이 인식된 영상을 추가하였다. 따라서 각 Label에 대한 데이터로 22개의 영상을 수집하였다. 수집된 영상들은 불필요한 동작을 학습에서 제외하고 데이터의 균일화를 위해서 수어 동작의 시작과 끝을 잘라 핵심 동작만을 추출한 뒤, 동작을 반복하거나 배속을 조정하는 과정을 통해 모두 3초의 영상으로 통일시켰다. 이후 영상 데이터를 핵심 동작으로 구성된 이미지 시퀀스 데이터로 변환하기 위해서 30fps 영상 기준 5프레임당 1장, 60fps 영상 기준 10프레임당 1장을 추출하여, 한 영상당 16장으로 이루어진 시퀀스 데이터로 구성하였다. 따라서 한 Label 당 총 352장의 이미지 데이터를 확보하였다. 하지만 해당 데이터로 학습시킨 모델의 accuracy 값이 0.6629로 측정되었다. 학습 데이터가 부족하다고 판단되어 정확도의 개선을 위해 데이터셋의 양을 늘리기로 하였다.

수정된 데이터셋의 경우에는 데이터 구성에 언어의 기초가 되는 자음과 모음 각각 3개를 추가하여 총 17개의 Label로 구성하였다. 자모음의 경우에는 AI HUB에서 데이터셋을 제공하지 않아 직접 촬영을 진행하여 총 11개의 영상으로 통일시켰다. 부족한 데이터셋을 늘리기 위해 동작 인식 기술인 OpenPose를 추가하여 각 Label에 대한 데이터로 총 33개의 영상을 수집하였다.

생성된 이미지 시퀀스 데이터는 CNN의 데이터 셋으로 사용되는데, 이때 부족한 학습 데이터를 더 늘려 모델의 성능 향상과 과적합 발생을 낮추기 위해서 Data Augmentation 기법을 적용하였다. Image Augmentation Library인 Albumentations를 이용하여 수어 동작의 의미를 해치지 않는 선에서 각 이미지 시퀀스별로 Augmentation을 해주었다. 사용된 조건으로는 회전  $\pm 5^\circ$ , 이동  $\pm 5$  px, 크기 조정  $\pm 0.3$ 배, 밝기  $\pm 0.2$ , 회전 +5 이동 +5로, 총 9개의 조건을 2번의 조정을 통해 10배의 Augmentation을 수행하였다. 최종적으로 한 수어 영상 Label에 대해서 528장의 원본 이미지와 5,280장의 변형된 이미지로 총 5,808장의 이미지를 구축하였다.

전 과정들을 통해 생성된 이미지 데이터들은 CNN 모델의 학습 데이터 양을 보장하기 위해서 Train 80%, Validation 20% 비율로 분류하여 데이터 셋으로 구성하였다.

## (2) 모델

수어는 시간에 따른 동작의 변화가 중요하므로 이를 반영할 수 있도록 CNN과 LSTM을 결합하여 모델을 구성했다. 수집한 데이터의 특징을 학습시키기 위해 ImageNet 데이터베이스의 영상에 대해 훈련된 컨볼루션 신경망인 VGG16 모델을 이용해 특징을 추출하고, 이미지 별 추출된 특징을 시계열 데이터 학습에 적합한 LSTM을 통해 특징을 학습시켰다.

옵티마이저는 SGD, 손실함수로는 categorical\_crossentropy를 사용하였으며 마지막으로 훈련된 데이터를 클래스로 분류하기 위하여 밀집(Dense) 레이어를 적용하여 클래스를 분류했다. 클래스 분류를 위해 마지막 밀집 레이어에는 softmax를 적용하였다.

## ① 실험 환경

구현 환경으로 컴퓨터 CPU는 Intel Core i7-10700, GPU는 NVIDIA GeForce RTX 3060 Ti를 이용하였고 RAM은 32GB에서 딥러닝을 학습시켰다. Hyper-parameter의 경우 모델 batch size는 128로, epoch는 200으로 설정하여 학습시켰다.

## ② 모델 결과

<그림 8>, <그림 9>은 딥러닝 수어 동작 판별 모델의 학습 곡선 및 학습 정확도를 보여준다. 학습시킨 모델의 loss 값은 0.064이고, val\_loss 값은 0.046로 측정되었다. accuracy 값은 0.979이고, val\_accuracy 값은 0.984로 측정되었다.

## ③ 테스트 결과

<표 5>는 17개의 Label 중 ‘ㄷ’, ‘안녕하세요’, ‘마음’에 대한 테스트 결과이다. 데이터 아래에 표기된 배열의 Label은 순서대로 ‘ㄱ’, ‘ㄴ’, ‘ㄷ’, ‘ㄹ’, ‘ㅁ’, ‘ㅂ’, ‘안녕하세요’, ‘괜찮습니다’, ‘미안합니다’, ‘감사합니다’, ‘좋아하다’, ‘싫어하다’, ‘마음’, ‘사람’, ‘생각’, ‘친구’, ‘학교’이다. 배열의 값은 분류 모델의 값을 나타낸다. 총 평균 분류 모델의 값은 약 32%로 train accuracy에 비해서 낮은 확률로 예측한 것을 알 수 있다.

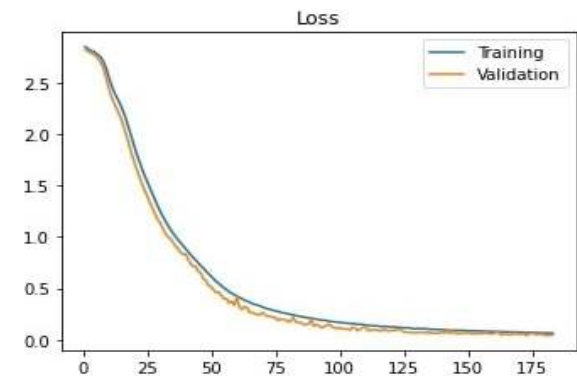


그림 8. 수어 동작 판별 모델 학습 손실 곡선

Fig. 8. sign language distinction model training loss curve

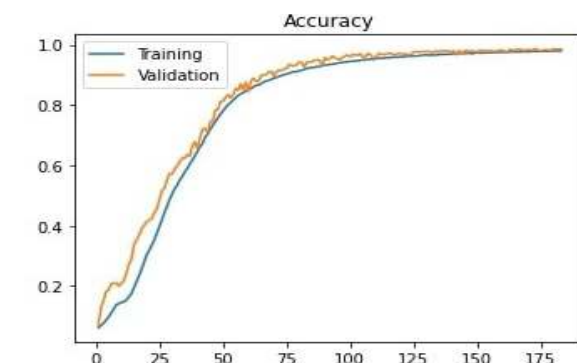


그림 9. 수어 동작 판별 모델 학습 정확도 곡선

Fig. 9. sign language distinction model training accuracy curve

표 5. 딥러닝 모델 테스트

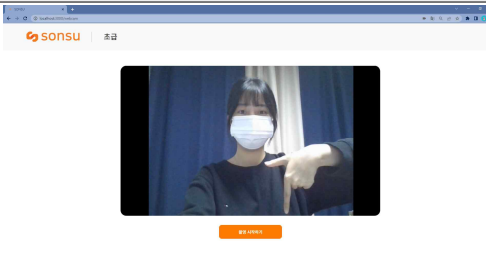
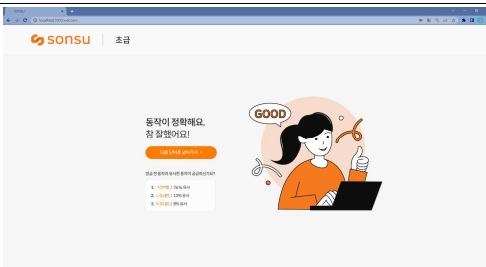
Table 5. Deep Learning Model Test

\*It is the table within Korean because it is a sign language for Korean.

| Label   | Image  |
|---------|--|
| 'ㄷ'     |   |
| '안녕하세요' |   |
| '마음'    |  |

표 6. 서비스 실증 테스트

Table 6. Service Demonstration Test

| Title   | Image   |
|---|---|
| Imitating and filming the user's sign language gestures |  |
| Results according to the user's sign language actions   |  |

### 3-4 실증 테스트

본 논문에서 제시한 사용자 웹 인터페이스와 기술들을 적용하여 딥러닝 기반 수어 교육 플랫폼의 구현을 완료하였다. <표 6>를 통해서 딥러닝 모델이 웹 서비스에서 잘 구현되고 있는 것을 확인해볼 수 있다.

## IV. 결론

본 논문에서는 딥러닝에 기반한 수어 교육을 목적으로 단편적인 강의 수강에서 그치는 것이 아니라 직접 수어 동작을 수행해보도록 유도하고, 동작에 대한 결과를 제공함으로써 양방향적인 소통이 이루어지는 교육 서비스 플랫폼을 구현하였다.

서비스의 주요 기능으로는 수강하기, 테스트하기, 마이페이지 내 오답 노트와 성적표가 존재한다. 수강하기에서 사용자는 자신이 학습을 원하는 단계를 선택하여 강의를 수강할 수 있으며, 이때 단계는 언어의 학습과 응용 순서를 고려하여 지문자와 고유 명사를 표현할 수 있는 초급, 일상생활 속 간단한 의사소통을 표현할 수 있는 중급, 문장의 구성 성분인 단어를 표현할 수 있는 고급으로 나누어져 있다. 사용자는 수강하기를 통해 원하는 강의를 수강한 뒤 따라 하는 모습을 웹캠으로 촬영하여 동작에 대한 결과를 피드백 받을 수 있다. 테스트하기 또한 초급, 중급, 고급 단계로 나누어져 진행되며, 해당 단계 내에서 무작위로 단어가 선정되어 평가를 진행한다. 사용자는 테스트하기를 통해서 본인이 해당 수어를 얼마나 숙지하고, 수행할 수 있는지 확인할 수 있다. 테스트하기의 결과는 성적표를 통해서 모두 조회할 수 있으며, 그중 틀린 수어 단어에 대해서는 오답 노트를 통해서도 확인이 가능해 사용자의 복습 과정 또한 돕는다.

적용된 주요 기술로는 동작 인식 알고리즘인 OpenPose와 MediaPipe가 사용되었다. 본 동작 인식 알고리즘은 사진 또는 영상에서 실시간으로 사람의 동작을 감지하여 신체 부위 및 관절을 인식해 시각적으로 보여주는데, 이를 수어 동작 영상에 적용하여 수어 동작에 사용되는 팔과 손목, 손, 손가락 관절 인식에 정확도를 더했다. 적용된 영상은 수어 동작 판별 모델을 학습시키기 위한 데이터 셋으로 사용되었다.

수어 동작 판별 모델에서 수어는 시간에 따른 동작의 변화가 의미를 가지기 때문에 이를 반영할 수 있도록 시퀀스로 이루어진 이미지 데이터 셋과 CNN과 LSTM을 결합하여 구성하였다. 데이터 셋으로는 동작 인식 알고리즘이 적용된 영상을 이미지화한 시퀀스 데이터와 모델의 성능 향상 및 과적합 발생을 낮추기 위해 Augmentation을 거친 이미지 시퀀스 데이터가 사용되었으며, 최종적으로 하나의 수어 단어 Label 당 8,976장으로 구성된 총 17개의 Label을 가지는 데이터 셋을 구축하였다. 학습 모델은 ImageNet 데이터베이스 영상에 대해 훈련된 컨볼루션 신경망인 VGG16 모델을 통해 특징



을 추출하고, 이를 시계열 데이터 학습에 적합한 LSTM을 통해 특징을 학습시켰다. 이때 옵티마이저는 SGD, 손실함수로 categorical\_crossentropy를 사용하였으며, 클래스 분류를 위해 밀집 레이어에는 softmax를 적용하였다. 구성된 모델은 0.979의 정확도를 가진다. 하지만 보여진 정확도에 비하여 실증 테스트상에서의 인식률이 높지 않은 차이를 보였다. 이를 극복하기 위해 추후 더 다양한 데이터 셋 구축 및 증폭 작업을 통하여 동작 인식을 개선을 위한 연구가 필요해 보인다.

사용자 인터페이스는 청인과 청각 장애인은 시각적인 부분에서는 큰 차이가 없지만 문자를 인식하는 능력에서 차이가 존재하기 때문에 이를 고려하여 웹 UI를 디자인하였다. 페이지 내에서 사용되는 문장은 최대한 간단하고 명료한 문장들로 구성하였으며, 도표를 활용하여 시각적으로 눈에 들어올 수 있도록 디자인하였다. 또한 밝은 주황색 색상을 주요 색상으로 사용하고 일러스트를 사용하여 밝고 캐주얼한 UI를 연출하여 접근에 부담을 느끼지 않도록 디자인하였다.

서비스 사용자들에게는 온라인 환경으로 웹 페이지와 웹캠을 통해 학습 환경이 제공된다. 따라서 기존의 농학교, 수어 교실을 통해 수어를 학습하던 환경에 비해 누구나 쉽게 접근할 수 있으며 립모션과 같은 비싼 동작 인식 장비 없이 학습에 참여할 수 있어 시간공간과 비용의 제약을 완화할 수 있다. 또한 기존의 온라인 수어 학습 환경에서는 단순히 강의 영상만을 제공하여 사용자가 시청을 통해서만 학습하는 단방향적인 소통이 이루어지는 방식이었지만, 본 서비스 플랫폼은 강의 영상 시청과 동작 수행을 유도하고 수어 동작에 대한 피드백을 제공한다. 따라서 사용자들이 단순히 강의 영상의 시청에서 그치는 것이 아니라 직접 동작을 수행해보고 결과를 제공받아 확인하는 주도적이고 양방향적인 소통이 이루어져 사용자의 학습률을 높일 수 있다. 본 연구는 모델의 정확도에 비해서 실증 테스트의 정확도가 떨어지는 한계점이 있지만 수어 동작 강의와 수어 동작 인식을 통한 문제 풀이 및 테스트 시행이 모두 가능한 플랫폼이 존재하지 않았다는 점에서 새로운 수어 교육 플랫폼 제안에 의의가 있다.

본 서비스 플랫폼은 수어 학습을 희망하는 청각 장애인들을 위해 기획되었다. 그러나 최근 청인(청각장애가 없는 비장애인)들의 수어에 대한 관심이 높아짐에 따라 수어 학습을 희망하거나 장애를 지닌 분들과의 소통을 원하는 청인들을 위하여도 본 교육 플랫폼의 활용도는 높다고 할 수 있다. 이에 본 서비스 플랫폼을 통해 수어에 대한 인식이 개선되고 더 많은 청각 장애인과 청인의 소통이 이루어질 기회의 제공을 기대해볼 수 있다. 또한 수어를 위한 학습 교재나 수어를 구사할 줄 아는 교사가 부족한 농학교에서의 학습 보조 도구로써 활용도 기대해 볼 수 있다. 현재는 단어와 짧은 인사 문장 위주의 동작만 학습이 진행되었지만, 추후 한국 수어의 문장 성분이나 문법 등 문장 위주의 모델 학습을 추가로 진행한다면 실생활에서 사용되는 구문과 더 긴 문장으로서의 의사 표현이 가능한 서비스로 확장될 수 있을 것이다.

본 서비스 플랫폼을 통해서 수어가 국어와 동등한 자격으

로서 자리매김하고 더 많은 청각 장애인들이 그들의 제1 언어 학습권을 보장받는 기회가 될 것이다.

## 참고문헌

- [1] KOREAN Statistical Information Service (KOSIS). Number of The Registered Disabled-by Year, Types of Disability and Gender (The Whole Country). [Internet]. Available: [https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT\\_11761\\_N001&conn\\_path=I2](https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT_11761_N001&conn_path=I2)
- [2] National Human Rights Commission of Korea. A Survey on Employment Discrimination and Employment Improvement for the Hearing Impaired. [Internet]. Available: <https://www.humanrights.go.kr/site/program/board/basicboard/view?menuid=001003001004001&pagesize=10&boardtypeid=16&boardid=7604958>
- [3] KOREAN Statistical Information Service (KOSIS). Causes of Disability by Disability Type of Persons with Home Disabilities. [Internet]. Available: [https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT\\_11732S0108](https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT_11732S0108)
- [4] National Institute of Korean Language Republic of Korea. Standard Korean Language Dictionary. [Internet]. Available: [https://stdict.korean.go.kr/search/searchView.do?word\\_no=513979&searchKeywordTo=3](https://stdict.korean.go.kr/search/searchView.do?word_no=513979&searchKeywordTo=3)
- [5] National Institute of Korean Language Republic of Korea. Standard Korean Language Dictionary. [Internet]. Available: [https://stdict.korean.go.kr/search/searchView.do?word\\_no=519022&searchKeywordTo=3](https://stdict.korean.go.kr/search/searchView.do?word_no=519022&searchKeywordTo=3)
- [6] National Institute of Korean Language Republic of Korea. Introduction of Sign Language. [Internet]. Available: [https://www.korean.go.kr/front/page/pageView.do?page\\_id=P000300&mn\\_id=202](https://www.korean.go.kr/front/page/pageView.do?page_id=P000300&mn_id=202)
- [7] Korean Law Information Center. Korean Sign Language Act. [Internet]. Available: <https://www.law.go.kr/%EB%B2%95%EB%A0%B9%ED%95%9C%EA%B5%AD%EC%88%98%ED%99%94%EC%96%B8%EC%96%B4%EB%B2%95>
- [8] HanKookIlbo. "a school for the deaf class turned a deaf ear to deaf appeals." [Internet]. Available: <https://www.hankookilbo.com/News/Read/201910011502091887>
- [9] In-Hye Kim and Il-Hong Jung, "A Study on Korea Sign Language Motion Recognition Using OpenPose Based on Deep Learning", The Journal of Digital Contents Society,

Vol. 22, No. 4, pp. 681-687, April. 2021.

<http://dx.doi.org/10.9728/dcs.2021.22.4.681>

- [10] Jin-Young Kim and Hyun Sim, "Development of a Sign Language Learning Assistance System using Mediapipe for Sign Language Education of Deaf-Mutlity", The Journal of the Korea Institute of Electronic Communication Sciences, Vol. 16, No. 6, pp. 1355 - 1362, Dec. 2021.

<http://dx.doi.org/10.13067/JKIECS.2021.16.6.1355>

- [11] Won-Cheol Park and Koo-Rack Park, "CNN-based Online Sign Language Translation Counseling System", Journal of Convergence for Information Technology, Vol. 11, No. 5, pp.17 - 22, May. 2021.

<https://doi.org/10.22156/CS4SMB.2021.11.05.017>

- [12] National Institute of Korean Language Republic of Korea. Korean vocabulary selection for study reports. [Internet]. Available:

[https://www.korean.go.kr/front/reportData/reportDataView.do?mn\\_id=207&report\\_seq=581](https://www.korean.go.kr/front/reportData/reportDataView.do?mn_id=207&report_seq=581)



**김미서 (Mi-Seo Kim)**

2019년~현 재 : 덕성여자대학교 IT미디어공학과 (공학사)

※ 관심분야 : 인공지능, 인터랙션 디자인, 컴퓨터 비전



**유문희 (Moon-Hee Yoo)**

2018년~현 재 : 덕성여자대학교 IT미디어공학과 (공학사)

※ 관심분야 : 인공지능, 인터랙션 디자인, 컴퓨터 비전



**장지현 (Ji-Hyeon Jang)**

2019년~현 재 : 덕성여자대학교 IT미디어공학과 (공학사)

※ 관심분야 : 인공지능, 인터랙션 디자인, 컴퓨터 비전



**최정윤 (Yun-Jeong Choi)**

2019년~현 재 : 덕성여자대학교 IT미디어공학과 (공학사)

※관심분야 : 인공지능, 인터랙션 디자인, 컴퓨터 비전



**나고은 (Go-Eun Na)**

2004년 : Iowa State University 심리학 (학사)

2008년 : The University of Iowa 학교상담학과 (석사)

2012년 : University of Maryland 상담교육학과 (박사)

2021년~현 재: 한국상담학회 국제지역분과 교육이사

2018년~현 재: Journal of Asian Pacific Counseling 부편집위원장

2017년~현 재: Hunter College, CUNY 상담학과 교수

※관심분야 : 학교상담, 다문화상담, VR 상담



**김미숙 (Mi-Suk Kim)**

2000년 : 강릉원주대학교 컴퓨터과학과 (이학석사)

2018년 : 강릉원주대학교 컴퓨터공학과 (공학박사)

2006년~2017년: (유)코디아아이 대표이사

2020년~2021년: 하이테크 이사 T/F팀

2000년~현 재: 강릉원주대 컴퓨터공학과 강사

※관심분야 : 인공지능, 인공생명, 빅데이터



**나정조 (Jung-Jo Na)**

2001년 : 덕성여자대학교 동양화, 시각디자인과 (예술학사)

2006년 : Pratt Institute, Digital Arts (예술학 석사)

2020년 : 가천대학교, 예술치료학 (미술학 석사)

2011년 : 숭실대학교, 미디어아트 (공학박사)

2021년~현 재: 한국영상학회 이사

2021년~현 재: 디지털미디어연구소 부소장

2020년~현 재: 덕성여자대학교 IT미디어공학전공 교수

※관심분야 : 예술치료, UI/UX, 미디어아트, 모션그래픽스