

Sign_Language.txt

건 국 대 학 교



팀명 : Sullivan

팀원 : 이찬민, 박승민, 문지영

소속학과 : 컴퓨터공학부

지도교수 : 임창훈 교수님



팀명 : Sullivan

- 팀 소개 및 팀원 소개

청각장애인을 위한 수어 번역 시스템을 개발하기 위해 팀을 결성했으며, 팀장인 이찬민은 수어 번역을 위한 딥러닝 서버 모델, 박승민은 웹 프론트, 문지영은 수어 시나리오 및 채팅방 DB를 담당했습니다.

팀명은 청각장애인이었던 헬렌 켈러가 장애를 극복하고 사회운동가로 성장할 수 있게 가르쳤던 앤 설리번의 이름을 따서 지었습니다.

- 팀원 소개 및 정보

이름	소속(학과)	역할	이메일	연락처
이찬민	컴퓨터공학부	백엔드, 웹 프론트, 딥러닝	chemwang@konkuk.ac.kr	010-9627-0516
박승민	소프트웨어학과	웹 프론트	smlion1997@gmail.com	010-8785-5741
문지영	컴퓨터공학부	시나리오작성, 채팅방 DB	gracemjy@konkuk.ac.kr	010-7151-7279

- 작품 개요

· 개발배경

과거 청각장애인들의 의사소통인 수화(이하 수어)를 비장애인이 이해하는데 사전지식이 있는 경우를 제외하면 사실상 불가능해서, 청각장애인들이 실생활에 어려움이 많다는 기사를 접한 적이 있습니다.

이에 청각장애인들이 특정한 장소 및 상황에서 비장애인과 소통을 필요로 하는 곳을 선정해서 실시간으로 수어를 번역해서 텍스트로 보여주는 시스템을 구축하면 어떨까 하는 생각으로 프로젝트를 시작하게 되었습니다.

· 목표 시스템

1. 청각장애인이 수어를 카메라 앞에서 시연하면, 텍스트 형식으로 화면에 출력이 된다.
2. 비장애인은 키보드를 통한 채팅으로 청각장애인에게 말한다.
3. 기본적인 장소는 병원으로 상정한다.

- 작품 구성 및 상세내용

· 작품의 구성

청각장애인이 수어를 시연하면 웹서버 내의 딥러닝 모델이 이를 분석하여, 90% 이상의 정확도를 일정 시간 이상 확인할 경우 단어로 인식해서 그대로 영상에 출력되게 한다.

상호간의 채팅내역은 DB로 전송되어 저장한다.

· 상세 내용

영상처리 및 딥러닝을 위한 라이브러리로는 Keras가 사용되었으며, 데이터 생성을 위해 Google MediaPipe Holistic을 사용해서 제작되었으며, OpenCV 창에서 확인할 수 있다.

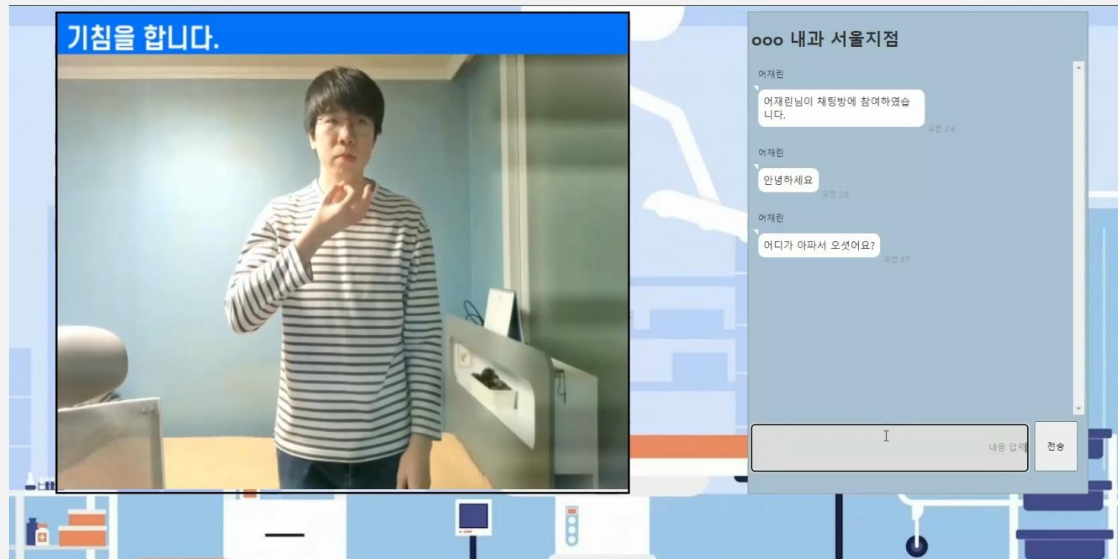
단어 별 30 번의 반복촬영을 하고 전체 데이터를 1000번 반복하며 학습시켜서 최종적으로 모델을 만들었다.

DB는 NoSQL인 MongoDB를 사용하였다.

- 개발 세부 내용(설계서)

· 서비스 시나리오

1. 채팅방 입장
2. 카메라를 통해 수어 영상을 받아와서 예측 서버로 보냄
3. 학습된 모델을 이용하여 수어 예측을 함
4. 그 결과를 텍스트로 영상 위에 보여줌
5. 상대방은 키보드를 통해 채팅이 가능

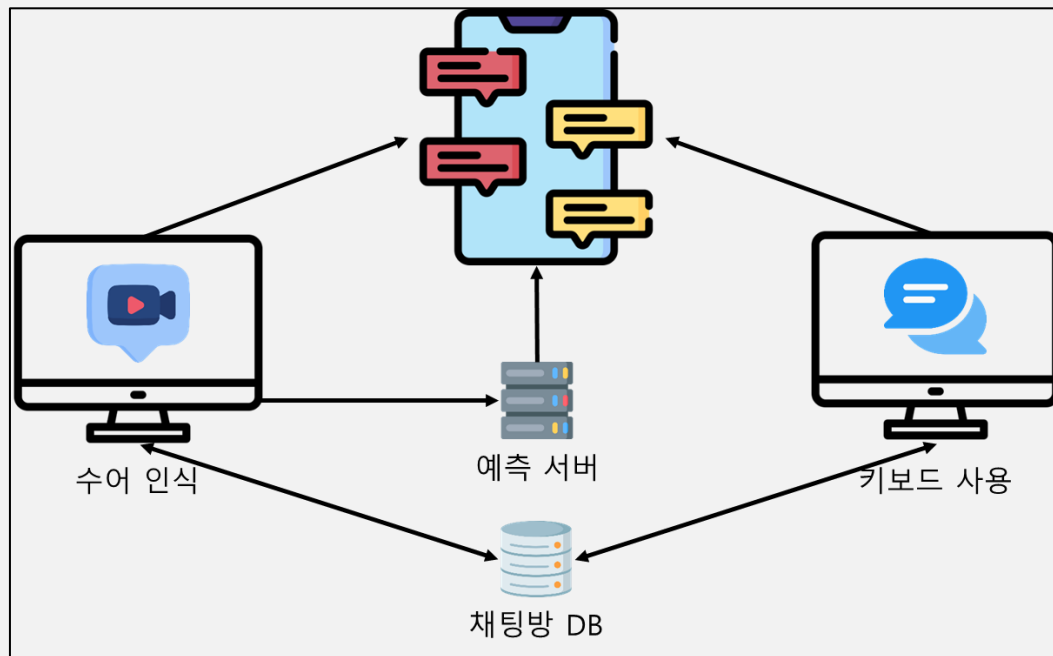


- 개발 세부 내용(설계서)

- 시스템 설계

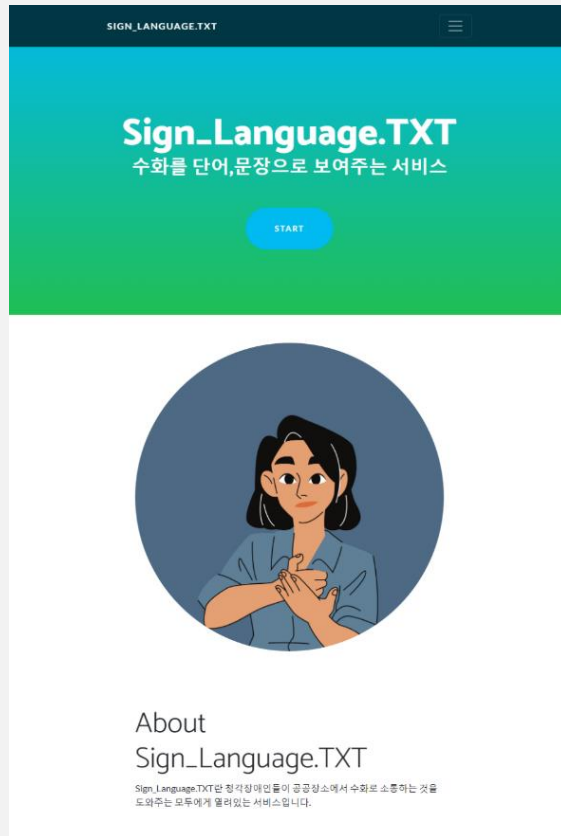
카메라를 통해 청각장애인의 수어를 인식하고 예측 서버로 전달되며, 예측 결과는 웹서버를 통해 웹 영상 위에 텍스트로 출력이 된다.

각 채팅방은 DB에 저장되고 채팅 내역 또한 DB에 저장된다.

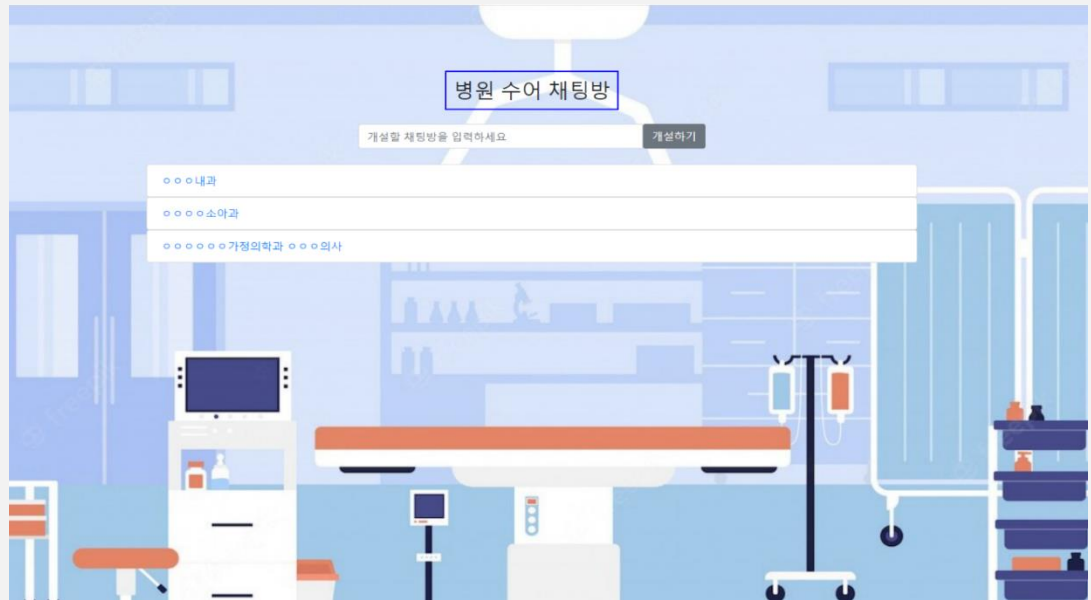


- 개발 세부 내용(설계서)

· UI 설계



메인 페이지



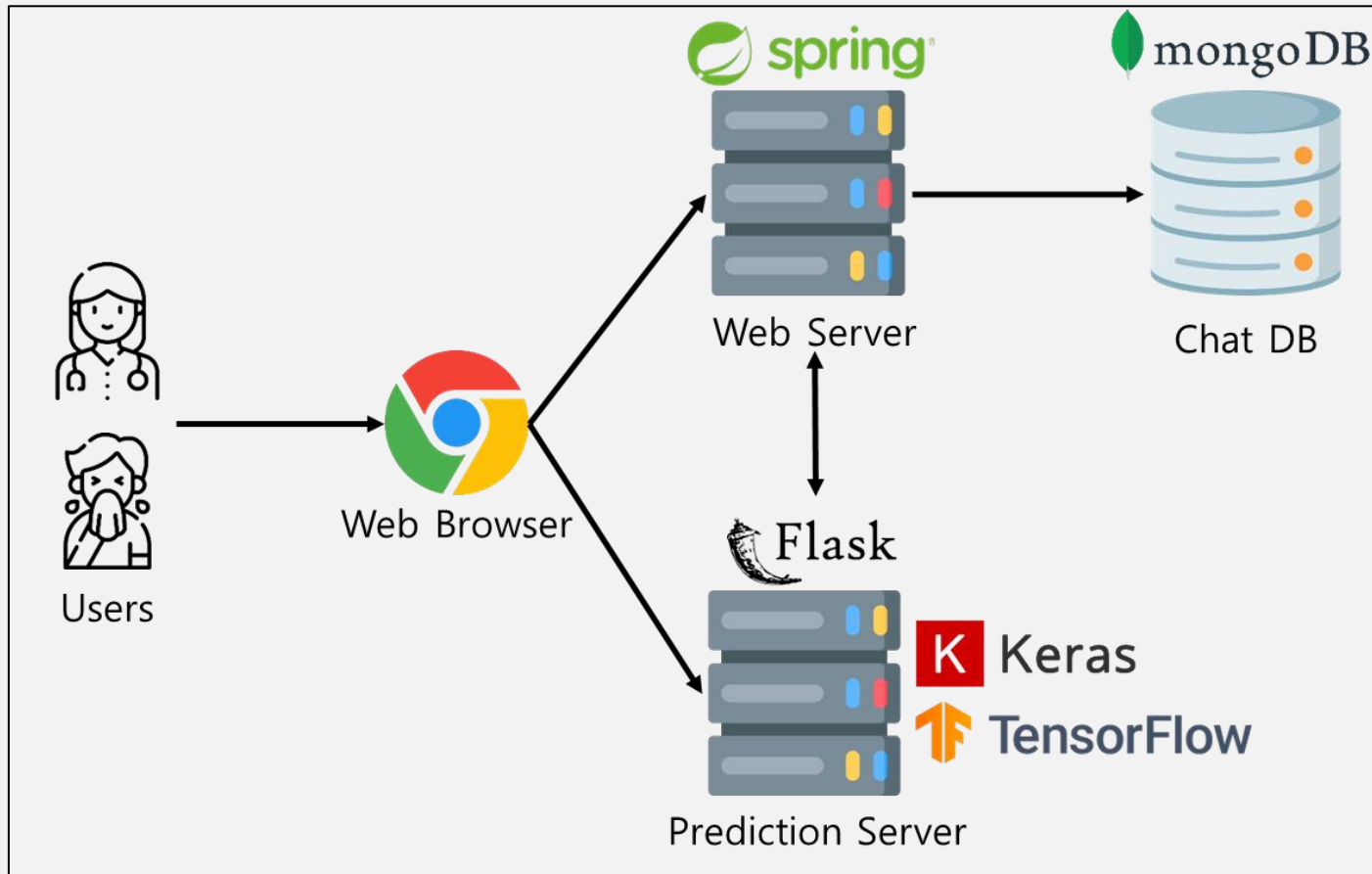
채팅방 개설 페이지

- 개발 세부 내용(설계서)

- UI 설계
- 메인 페이지
 - 채팅방 개설 페이지로 이동할 수 있다.
- 채팅방 개설 페이지
 - 채팅방을 개설하거나 기존 채팅방을 선택하여 입장할 수 있다.
- 채팅방 페이지
 - 카메라와 연결된 OpenCV 창 또는 키보드를 통해 채팅할 수 있다.

- 개발 세부 내용(설계서)

• Interface 설계



- 개발 세부 내용(설계서)

- Interface 설계
 - Spring Boot를 이용해 웹 서버를 구축했고 유저의 request를 처리
 - 웹서버의 경우 채팅방 생성, 채팅, DB 관리 등의 역할을 수행
 - 딥러닝 서버는 Flask framework를 사용
 - 웹 서버와 딥러닝 서버 간의 통신은 예측 서버에서 예측결과를 POST 하는 방식으로 커뮤니케이션 가능
 - DB는 MongoDB를 사용했고 Spring에 연결하여 데이터를 관리

- 개발 세부 내용(설계서)

- 기타 모듈 설명 등
- MediaPipe Holistic
 - 포즈, 얼굴 랜드마크, 손 추적에 대한 실시간 인식은 제스처 및 수화 인식, 스포츠 분석과 같은 작업을 위한 빠르고 정확한 솔루션을 제공
 - MediaPipe Pose, MediaPipe Face Mesh, MediaPipe Hands 모델을 각각 활용하여 총 543개의 landmark를 생성
 - 각 손 당 33개의 pose landmark, 486개의 face landmark, 21개의 hand landmark로 구성됨



Fig 1. Example of MediaPipe Holistic.

- 구현결과 요약

기본적으로 컴퓨터 1 대와 카메라 1 대를 필요로 한다. 청각장애인이 수어를 수행하면 카메라로부터 영상을 Flask framework 기반의 예측서버에 전달되어 영상처리 및 예측을 한다. 예측된 결과는 영상위에 텍스트로 출력이 된다.

Spring Boot framework 기반의 웹 서버에서 영상과 채팅을 통해 청각장애인과 상대방이 소통을 할 수 있다. 결과적으로, 청각장애인과의 소통을 돕고 의사소통에 있어서 불편함을 해소할 수 있다.

- 프로젝트 성과에 대한 기대효과

Sign_Language.txt의 목표는 수어를 텍스트로 변환하는 것이다. 이를 통해 청각장애인이 수어를 모르는 다른 사람과의 소통을 돕고자 한다. 특히, 공공 장소로 한정하여 소통이 원활하지 않아 생기는 불편을 줄이고자 한다.

첫 번째 시나리오인 병원에서 환자인 청각장애인과 의사의 소통의 정확도를 높여 좋은 의료서비스를 받을 수 있을 것이라 기대한다.

- 활용가능성 및 확장 가능성

병원에 이어서 공항, 도서관 등 공공장소를 추가할 수 있다. 다양한 장소와 상황에서 청각장애인과 의사소통으로 일상생활에서 말이 통하지 않는 답답함을 해소할 수 있다.

병원을 기준으로 한 모델은 현재 12개의 단어를 학습했다. 이후, 더 많은 단어를 학습하여 수어 예측할 수 있는 단어의 개수를 늘려 편의성을 증가할 수 있다.

- 최종 결과물에 대한 동영상 / 사진

1. 동영상 링크

- <https://youtu.be/eXK0KxhbSDg>

2. 가장 '하이라이트' 장면

- "00:30~01:28"

3. 관련 이미지

