

인공지능을 활용한 수어 번역 서비스

수어랑 말해랑

TEAM_SALT



팀장 : 김군순

팀원 : 김다희, 김훈종, 이대섭, 백지수

목차

Contents

- 01 제안 배경
- 02 설계 과정
- 03 구현 과정
- 04 발전 방향
- 05 참고문헌 및 팀원소개

01

제안 배경

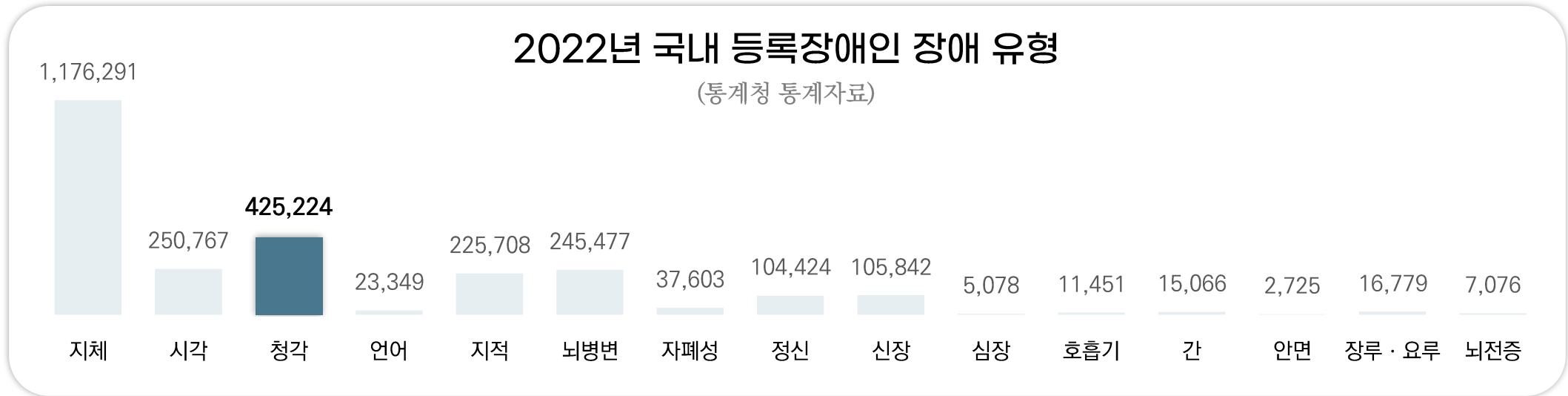
필요성 / STP분석 / 유사 서비스

01 제안 배경 필요성

제안의 배경

> 청각장애인의 수 증가

- 2022년 기준 국내 청각장애인의 수는 약 42만명으로 전체 등록 장애인 중 두 번째로 많은 수로 집계됨



- 청각장애인 중 수화언어(이하 수어)를 사용하는 이들을 농인이라고 칭함. 여기서 수어는 특정 언어에 기반을 두는 것이 아닌 고유한 언어임
- 한국농아인협회에서는 국내 농인 중 문맹률을 10~20%로 추정하고 있고, 통계청 조사 결과, 약 40%의 농인은 자막서비스를 이용하지 않고 자막 서비스를 이용할 경우 내용을 모두 이해하는 경우는 12.3%임

01 제안 배경 필요성

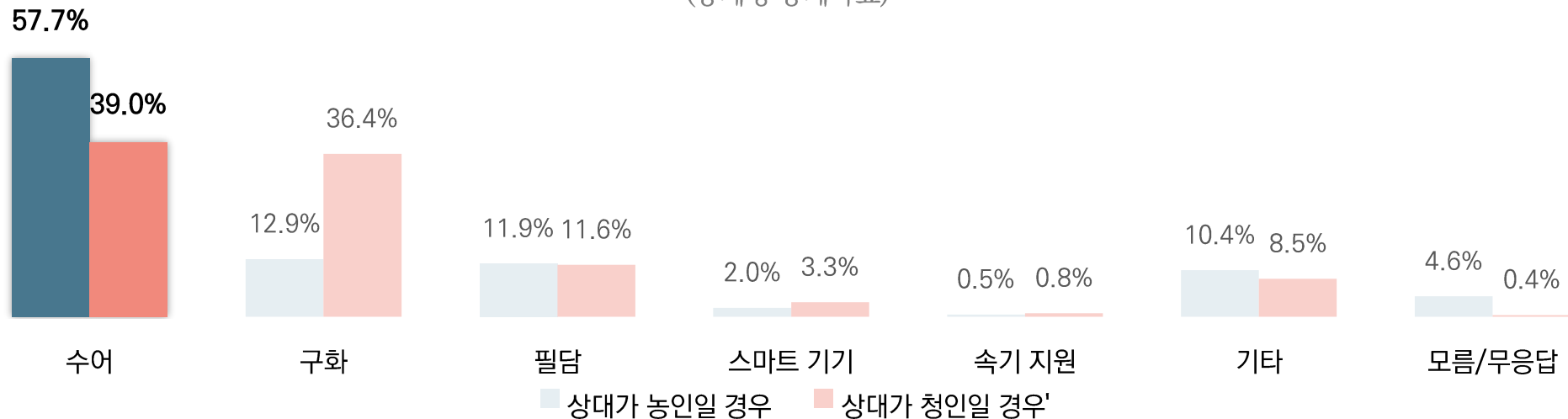
제안의 배경

> 수어의 사용정도

- 수어는 다른 언어에 기반을 두지 않는 고유 언어이기 때문에 농인-청인 간 의사소통 및 정보전달에 있어 한국어 기반인 필담, 자막은 의사소통의 대안이 되기에 어려움이 있음
- 통계청 조사결과 '원하는 의사소통 방법' 문항에서 '상대가 농인인 경우', '상대가 청인인 경우' 모두 '수어'를 사용한 의사소통을 가장 선호함

원하는 의사소통 방법

(통계청 통계자료)



01 제안 배경 필요성

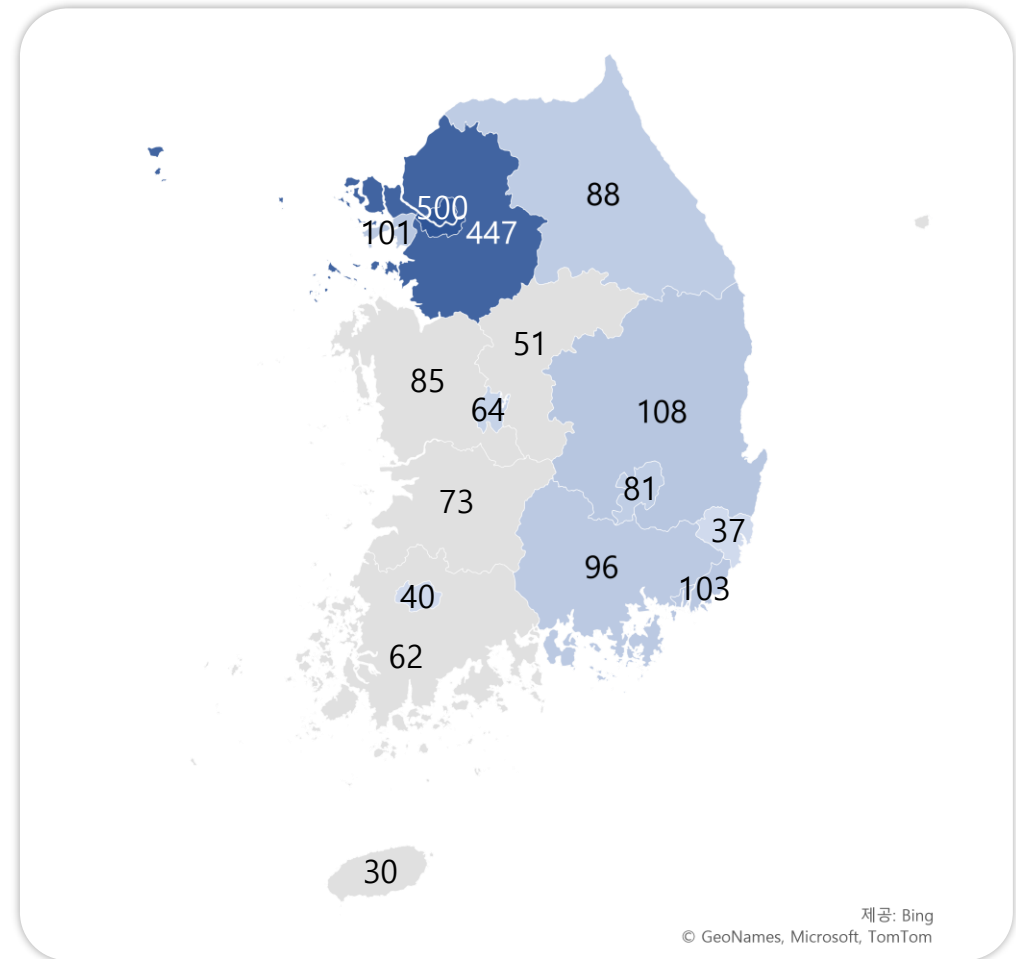
제안의 배경 및 필요성

> 수어통역사 현황

- 한국수어통역사협회에 따르면 수어통역사로 등록되어 있는 인원은 2022.12.31 기준으로 **1,973명**으로 청각 장애인 수에 비해 인력이 많이 부족함
- 특히, **지방은 인력이 현저히 떨어져** 지방의 농인들은 통역서비스를 누리기 어려움

> 필요성

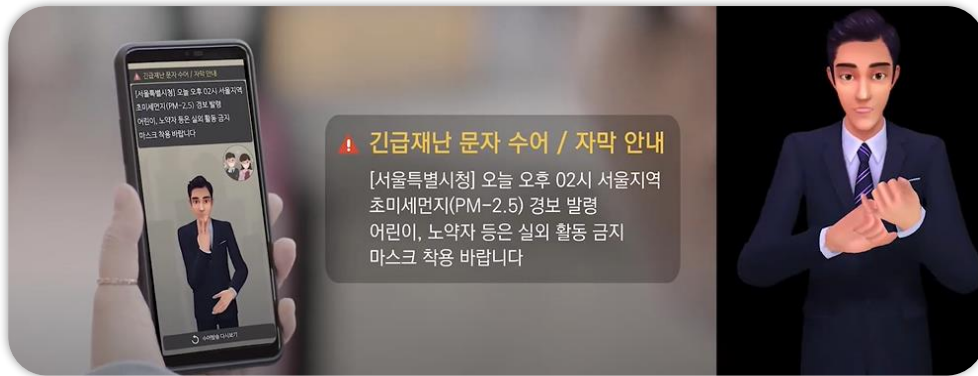
- 사회 전반에 걸친 **청인 중심의 정보전달 체계**로 인하여 일상생활, 경제활동, 긴급상황 등의 분야에서 많은 농인들이 고충을 겪고 있음
- 농인들 또한 사회의 구성원으로서 청인들과 같은 분야에서 **동등한 서비스**를 제공받는 것은 기본권임
- 차별금지법을 통해 이러한 기본권을 보장하려 하지만 실질적으로 부족한 상황이기 때문에 **수어 번역 서비스**를 제공하여 기본권을 보장받을 수 있도록 도와려고 함



01 제안 배경 유사서비스

유사서비스

[수어통]



응급안내, 재난문자, 시설안내 등 다중이용시설 응급 안내방송을 자막이나 수어(수어 애니메이션)로 전달

[충남대학병원 아바타 수어 키오스크]



아바타를 통한 이용안내와 같은 정보전달을 함

[누리뷰 키오스크(대전시 통합민원안내시스템)]



아바타를 통해 음성 및 수어로 간단한 소통이 가능함

01 제안 배경 유사서비스

차별성

아쉬운 점

키오스크의 경우 특정,
시설에서만 사용가능

농인만 이용하는 서비스

사용자 간의 소통 불가



개선 목표

웹에 접근만 하면
사용 가능

청인도 이용하는 서비스

커뮤니티를 통해
사용자 간, 운영자와의 소통 가능



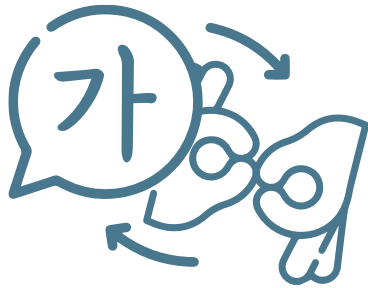
02

설계 과정

주요 서비스 / 서비스 흐름도 / ERD

02 설계 과정 주요 서비스

주요 서비스



수어 번역

메인 기능

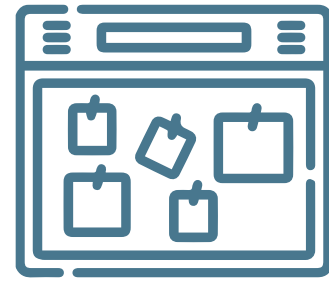
- 농인의 수어 동작을 캠으로 인식하고 분류
- 분류된 단어들을 조합해 문장으로 출력



수어 사전

서브 기능

- 한국 수어사전에 등록된 단어들을 조회



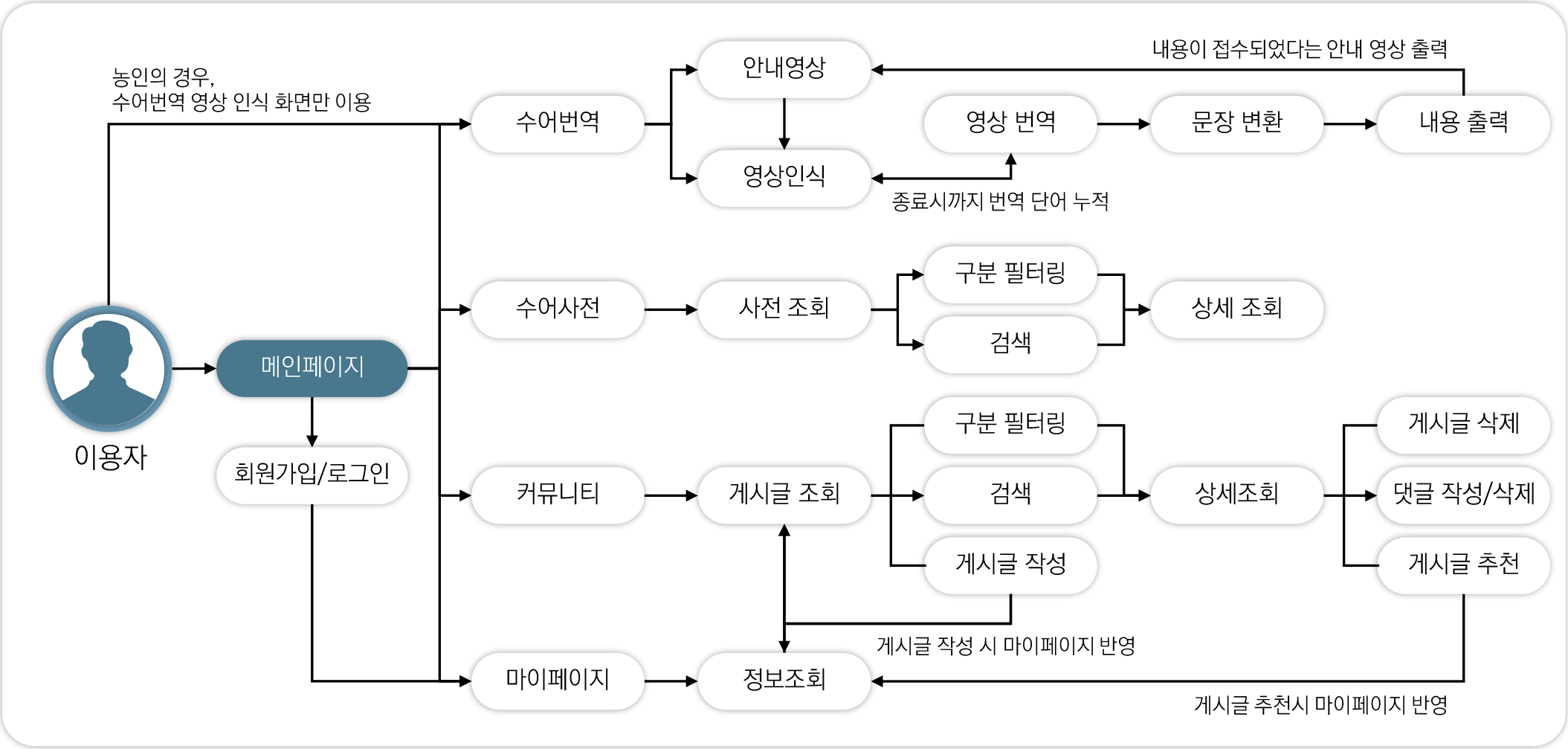
커뮤니티

서브 기능

- 서비스의 지속적인 발전을 위한 기능
- 이용시설, 이용자 등이 서비스가 더 발전할 수 있도록 건의/이용자간 의견 교환

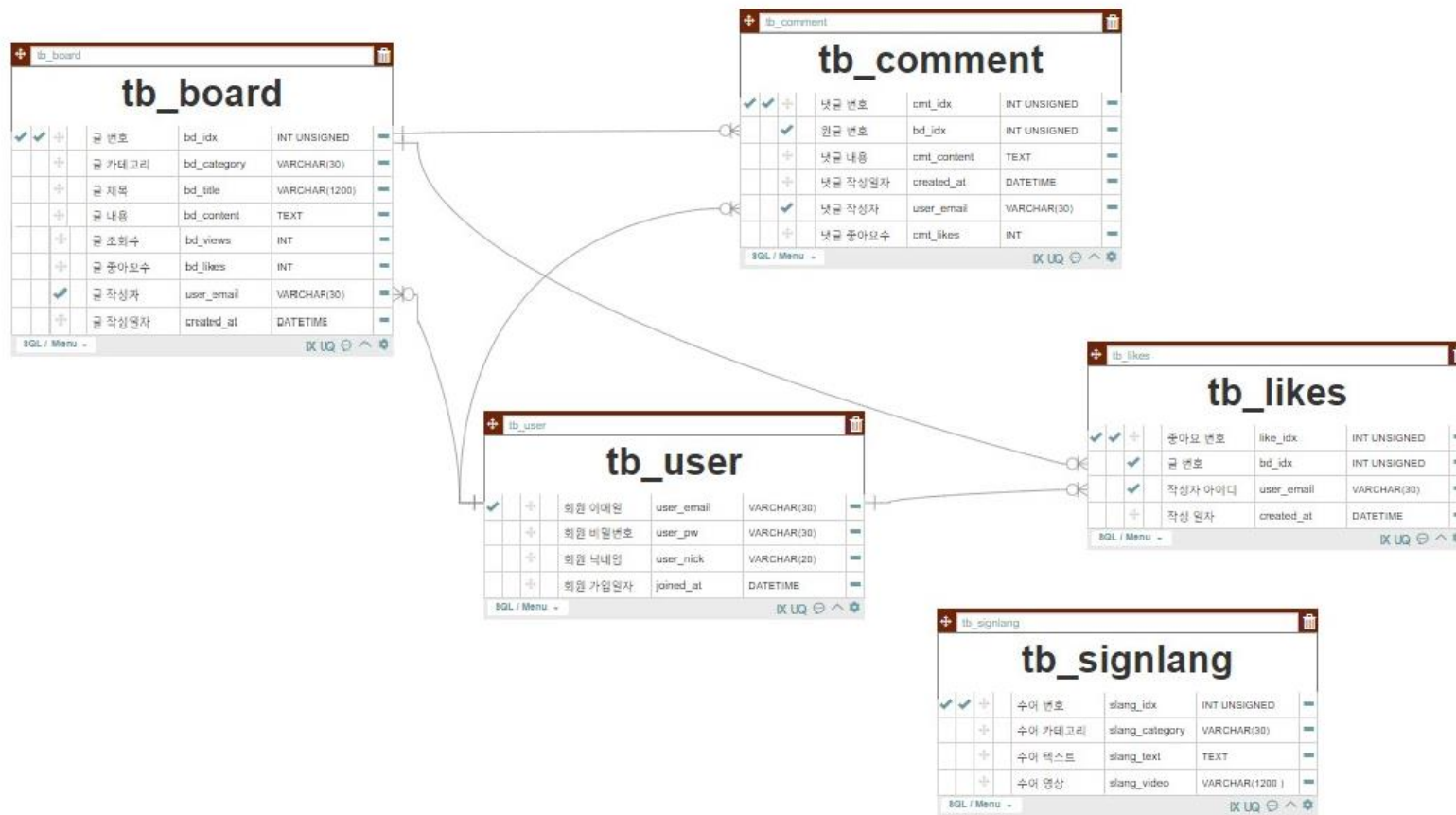
02 설계 과정 서비스 흐름도

서비스 흐름도



02 설계 과정 ERD

ER DIAGRAM



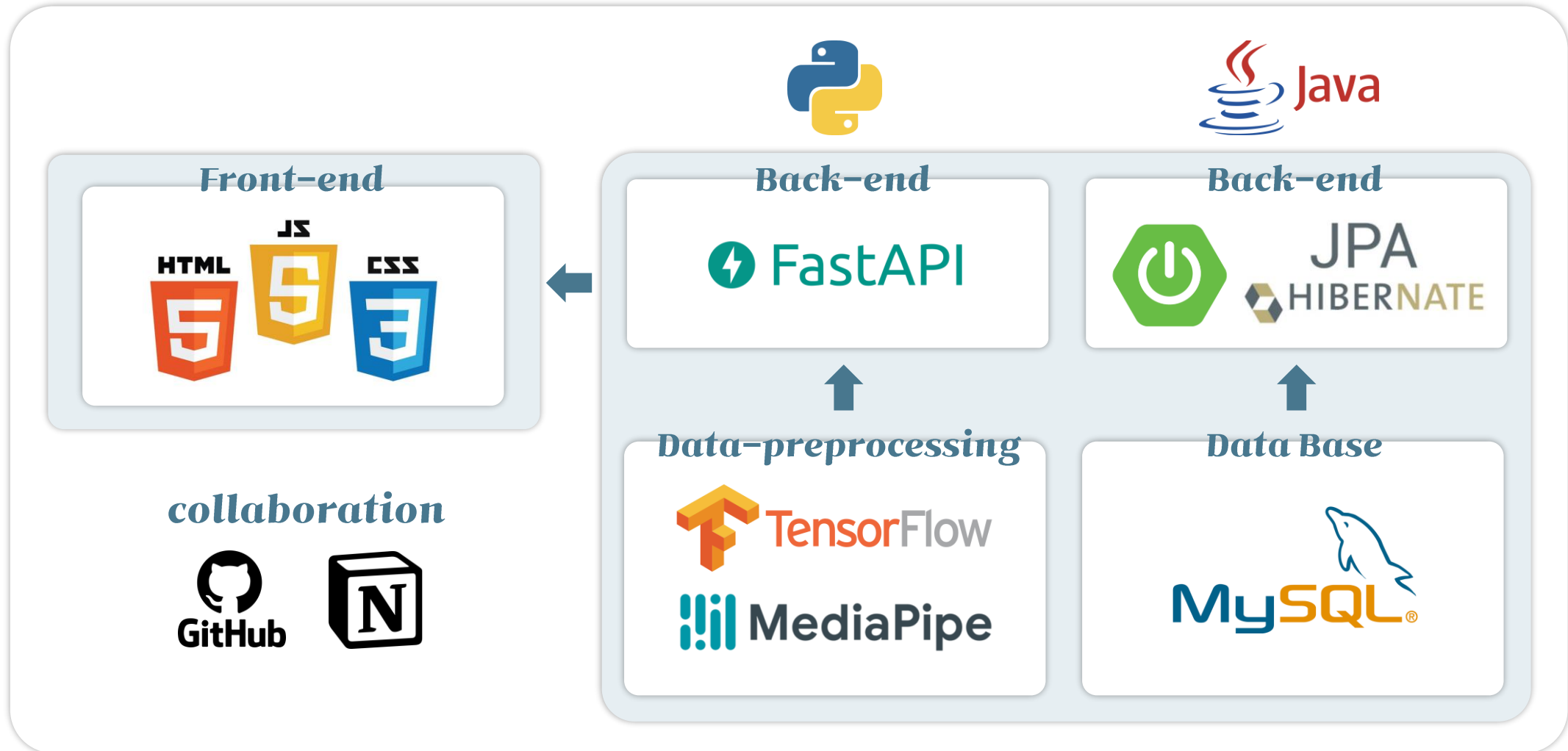
03

구현 과정

사용 언어 및 도구 / 아키텍처 / 데이터 수집 /
모델링 / 주요 기능 / 시연 영상 / 트러블 슈팅

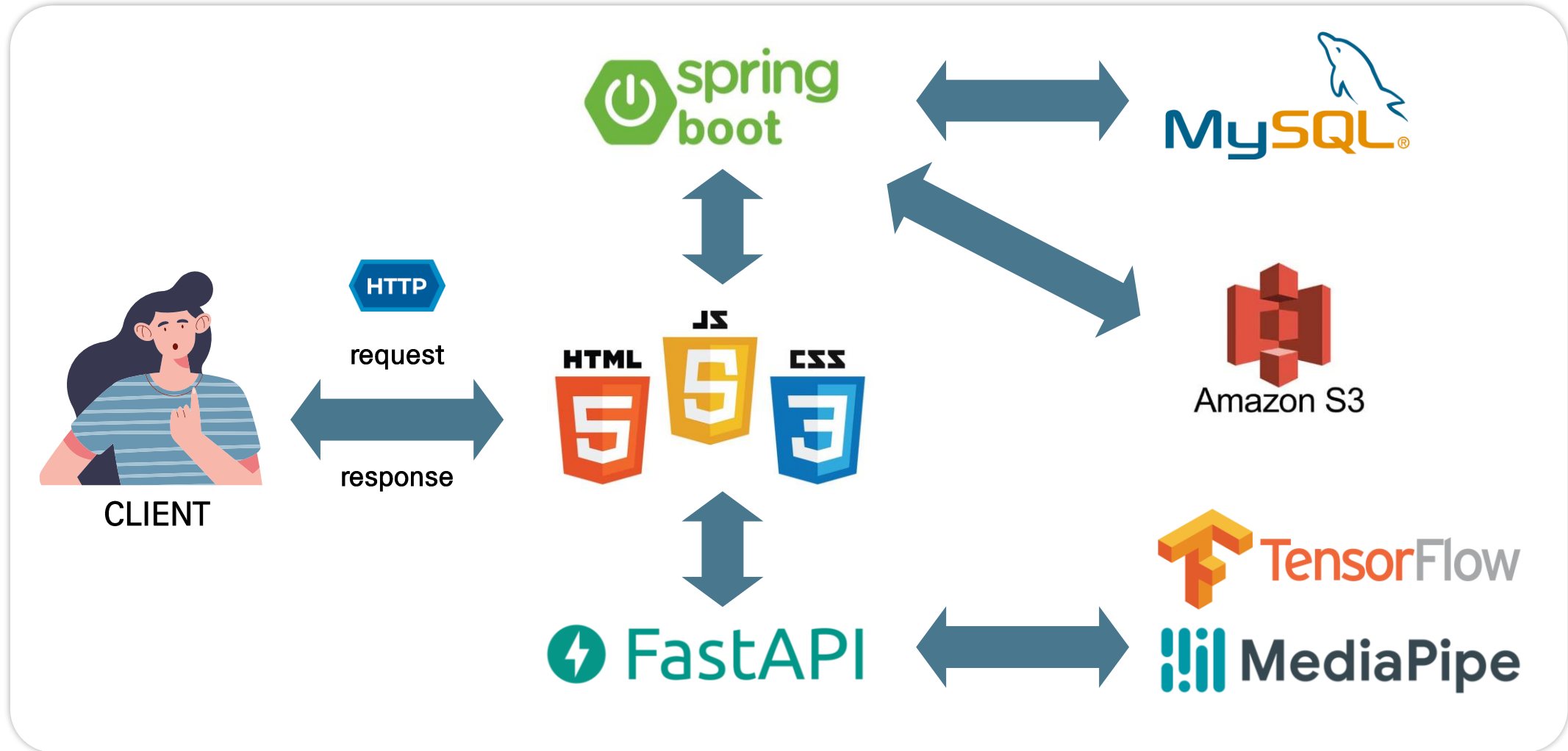
03 구현 과정 사용언어 및 도구

사용 언어 및 도구



03 구현 과정 아키텍처

아키텍처



03 구현 과정 데이터 수집

데이터 수집

> AI-Hub 영상 데이터

AI Hub

AI 데이터찾기

AI 허브소개

참여하기


커뮤니티

AI 개발지원

고객지원

마이페이지

로그아웃



#인공지능 #데이터셋 #수어영상 #기계학습 #키포인트검출 #수어

수어 영상

용량 : 2.63 TB

영상 데이터
536,000건 데이터

분야 한국어 유형 텍스트, 비디오

구축년도 : 2020 갱신년월 : 2021-06 조회수 : 18,412 다운로드 : 3,763 용량 : 2.63 TB

다운로드

↓ 샘플 데이터 ?

관심데이터 취소 83

소개

파일 목록 (API 다운로드)

※ 내국인만 데이터 신청이 가능합니다.

데이터 토론

문의하기

목록

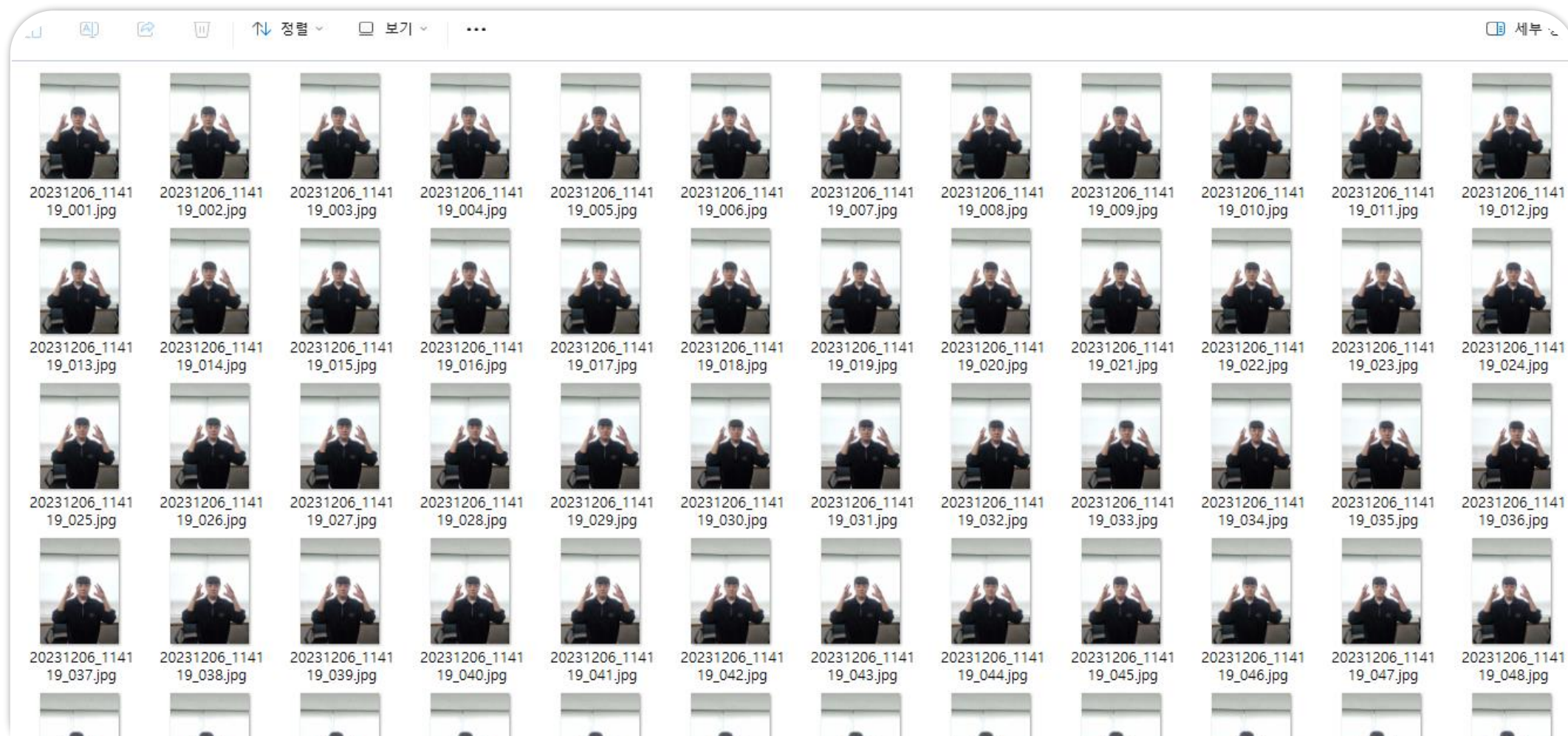
데이터 개요

메타데이터 구조표

03 구현 과정 데이터 수집

데이터 수집

> 직접 촬영 이미지 약 5,000건 이미지 데이터

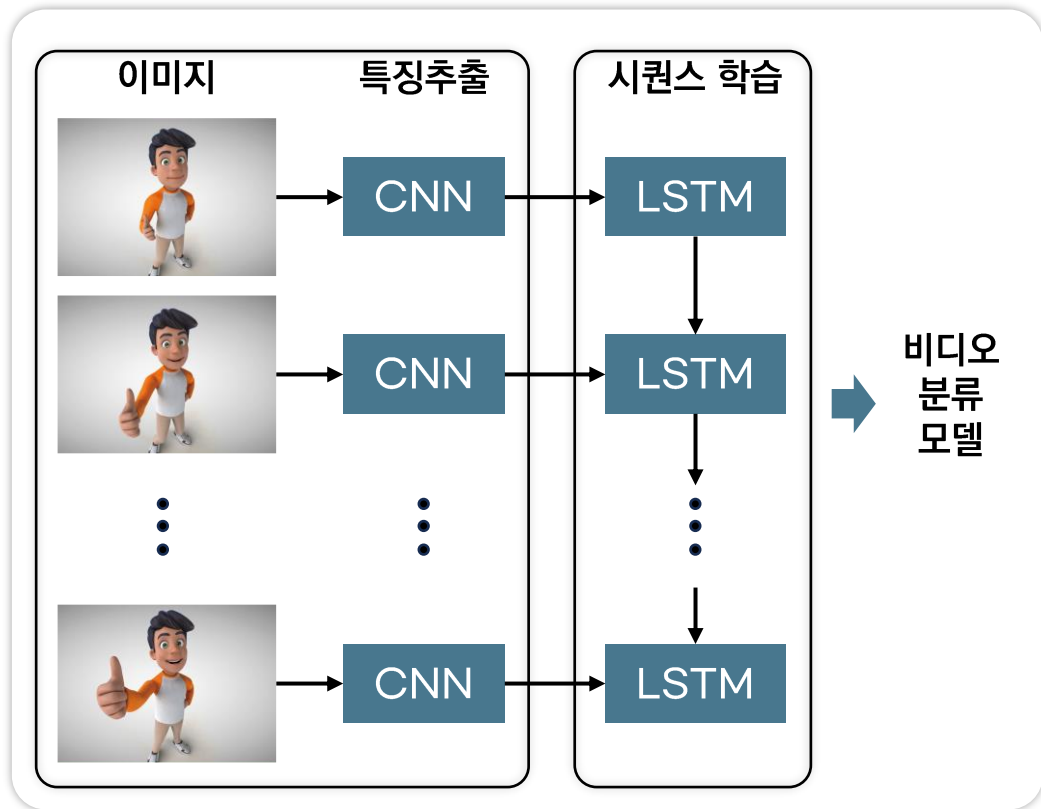


03 구현 과정 모델링

모델 선정

> CNN-LSTM 모델

- 이미지 특징 추출에 적합한 CNN과 시계열 데이터 학습에 적합한 LSTM 하이브리드 모델



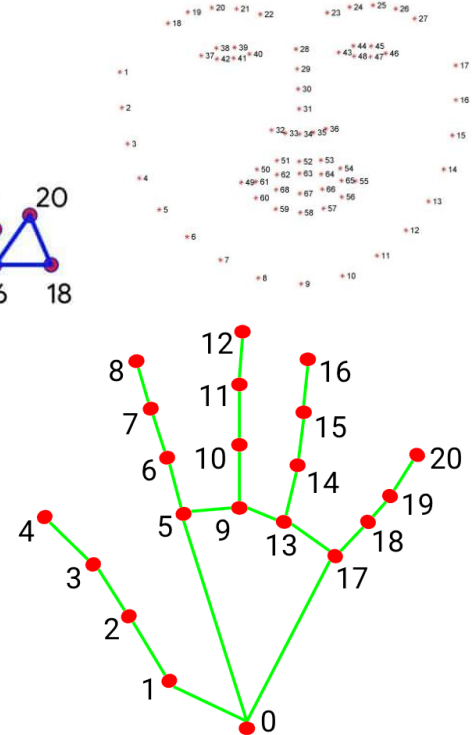
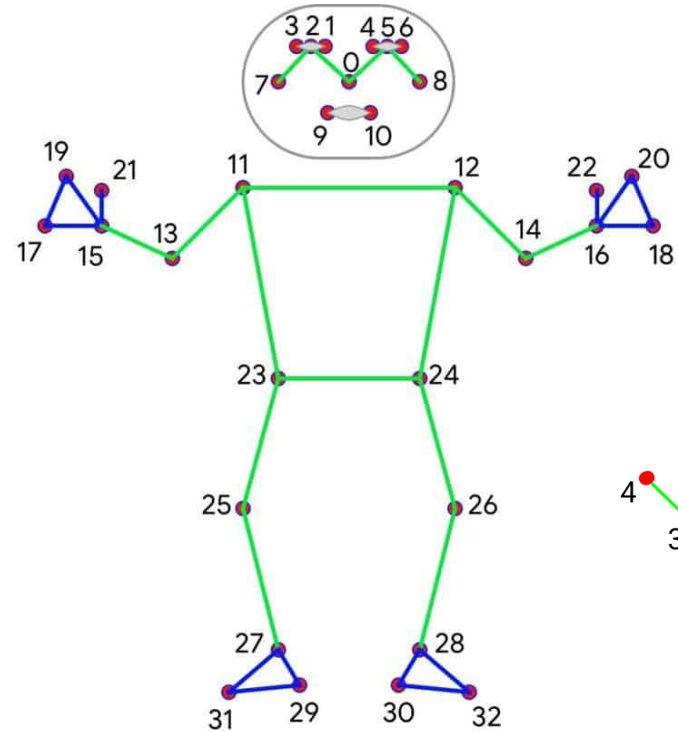
- VGG(Visual Geometry Group)
깊이가 깊어질수록 더 세밀한 특징을 학습하나 계산 비용 증가
- ResNet (Residual Networks)
잔차 연결을 통해 깊은 네트워크에서도 효율적으로 학습
- DenseNet (Densely Connected Convolutional Networks)
특성 재사용을 촉진하는 구조로 특성의 다양성을 유지하면서 효율적으로 학습
- EfficientNet
균형 잡힌 방식으로 네트워크의 깊이, 너비, 이미지 해상도를 동시에 확장

03 구현 과정 모델링

도구 선정

> MediaPipe

- 사람 동작 인식 도구
- 수어 동작 인식에 유용한 pose, hand, face detection을 사용



03 구현 과정 모델링

모델링

> CNN-LSTM 모델과 MediaPipe 데이터 결합하여 학습

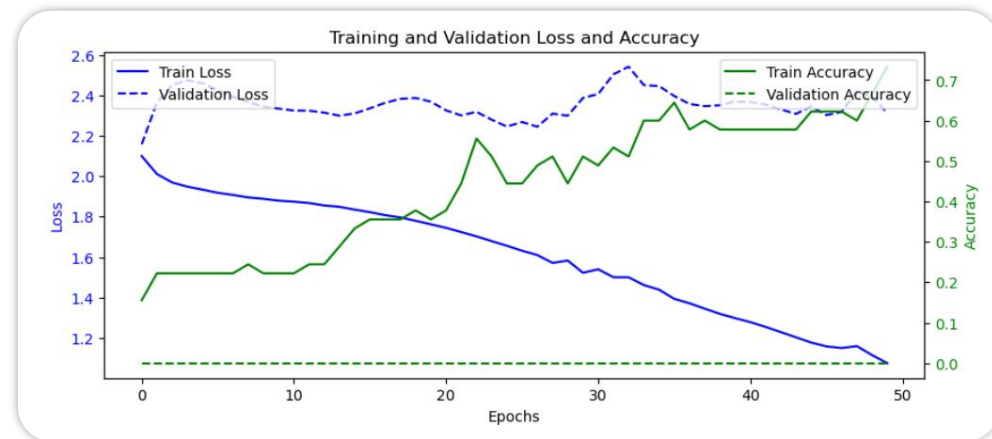
- 분류 테스트에 사용할 5개 클래스, 각 60개 영상 준비
- 손실, 정확도 그래프를 분석했을 때, 과적합 양상이 나타남을 확인



데이터 증강 (-5~5도 회전 / 밝기 조절 / 노이즈)

조기 종료, 드롭아웃 설정

CNN모델 변경 / 특징점 추출 값 변경

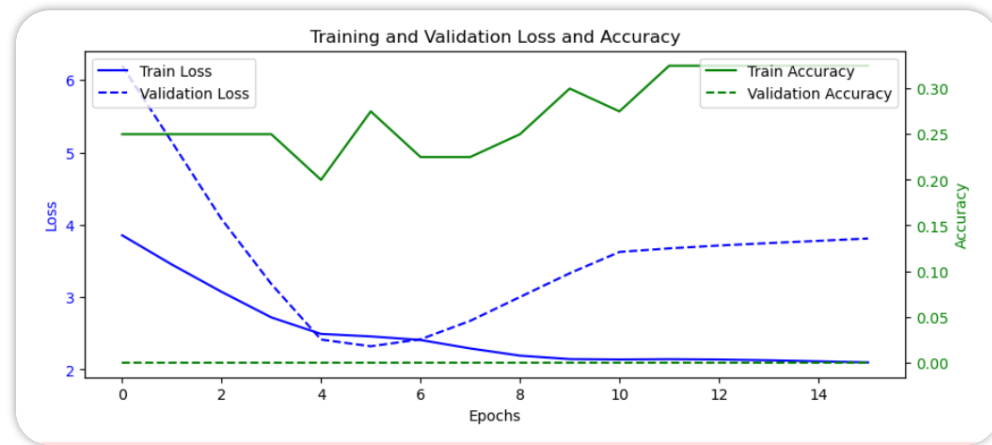


- 조금 더 일반화되는 경향과 과적합 되는 속도가 느림
- 그러나 손실은 더 높은 값을 유지, 정확도는 매우 낮아 여전히 학습이 잘되지 않음



데이터의 양 부족 또는 데이터 전처리에 문제가 있다고 파악

영상에서 데이터를 추출하는데 이미 하루 이상 소요되기 때문에 다양한 시도는 물리적으로 불가능

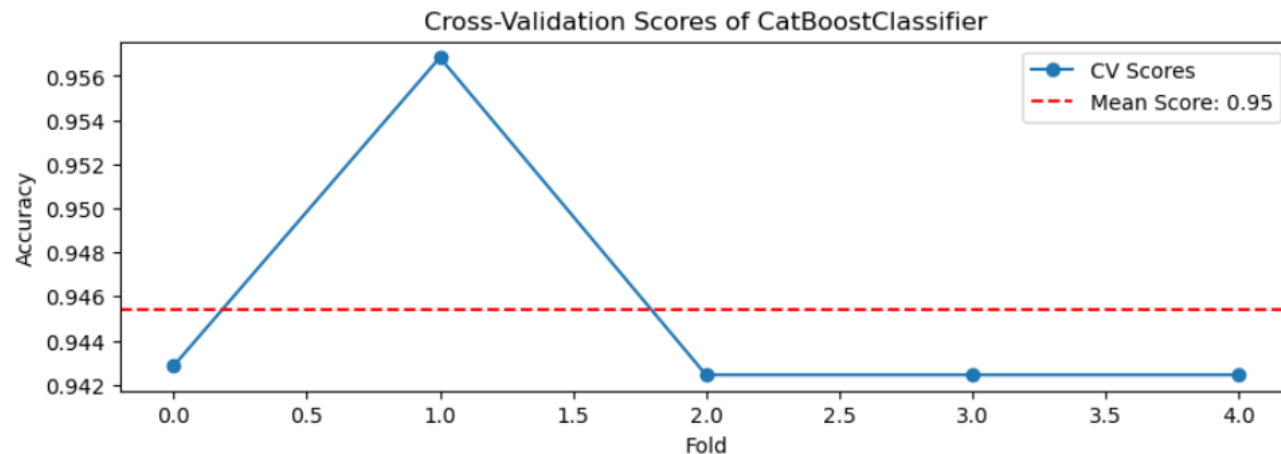


03 구현 과정 모델링

모델링

> MediaPipe 데이터만 학습

- 수어 동작에는 손가락, 팔 각도, 얼굴 표정 등 다양한 요소가 포함되지만 핵심 동작이 있음에 초점을 맞춰 이미지 데이터 제작
- 훈련 데이터에 10개 클래스, 각 200장씩 준비 후 데이터 증강
- Meidapipe로 Pose중 상반신 랜드마크, Hand 랜드마크 추출 후손가락의 각도 및 팔의 각도 계산하여 학습
- 검증결과가 좋았으나 실제 웹 캠으로 테스트 시 손가락의 모양이 비슷한 경우, 정확도가 떨어짐
- 보완을 위해 상반신, 각 손가락 Marker의 좌표 x, y, z,값을 추출해 학습 데이터에 포함



03 구현 과정 주요 기능

개요

> 주요기능

- 수어 번역 / 수어 사전(R) / 커뮤니티 CRUD / 회원가입 CRUD
- 기본 커뮤니티 웹 사이트 기반의 수어를 번역해주는 서비스가 추가

수어랑
말해랑

수어번역

수어사전

커뮤니티

공지사항

+
회원가입

로그인



한국수어(手語)

한국수화언어의 줄임말로
한국수어는 한국어와는 문법체계가 다른
대한민국의 농인의 **고유한 언어**입니다.



수어랑말해랑의 목적



수어랑말해랑은 농인들 또한 사회서비스를
누릴 수 있도록 수어를 번역하는 서비스를
제공하려 합니다.

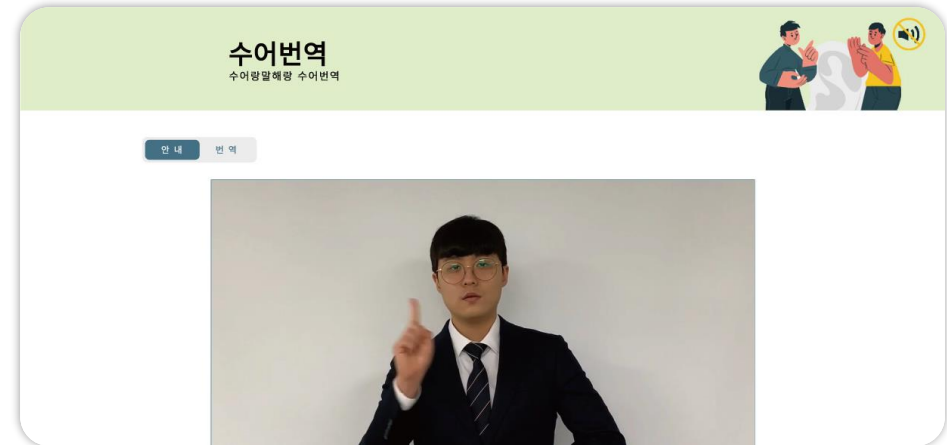
특히, 사회서비스를 누리는데 필수적인
교통분야에 접근성을 높이며 청인들이
농인들에게 잘 대처할 수 있도록 수어내용에

03 구현 과정 주요 기능

메인 기능

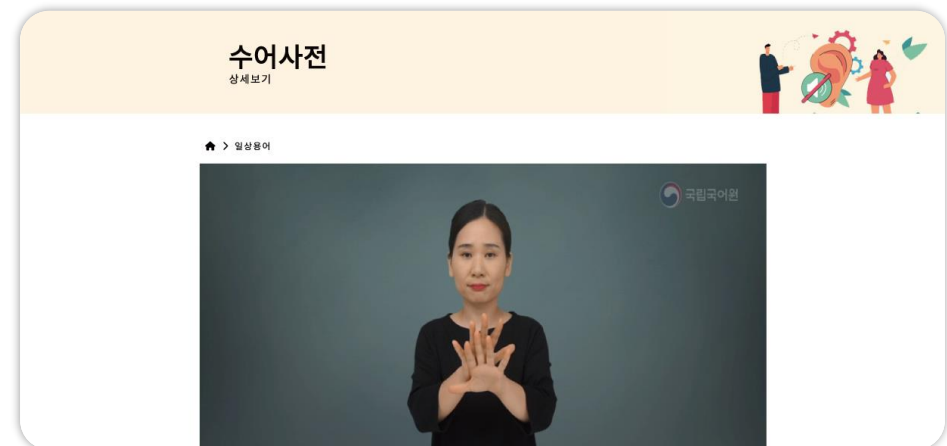
> 수어 번역

- 농인의 수어를 영상으로 인식해 번역하여 전달하는 기능
 1. ‘안내’ 탭에선 항상 ‘무엇을 도와드릴까요?’라는 영상이 반복된다.
 2. ‘번역’ 탭에선 웹 캠을 트는 ‘시작’ 버튼과 캠을 종료하는 ‘종료’버튼이 있다.
 3. 번역을 마친 후 ‘종료’ 버튼을 클릭하면 ‘접수되었습니다. 잠시만 기다려주세요.’를 수어 영상으로 출력한다.



> 수어 사전

- 한국어에 대응하는 수어를 검색하고 조회하는 기능
 1. 리스트에는 썸네일과 수어에 대응하는 한국어 뜻이 적혀 있다.
 2. 일상 단어 / 전문 단어 탭으로 구분하여 필터링 된다.
 3. 원하는 단어를 검색할 수 있다.
 4. 영상으로 상세조회 할 수 있다.



03 구현 과정 주요 기능

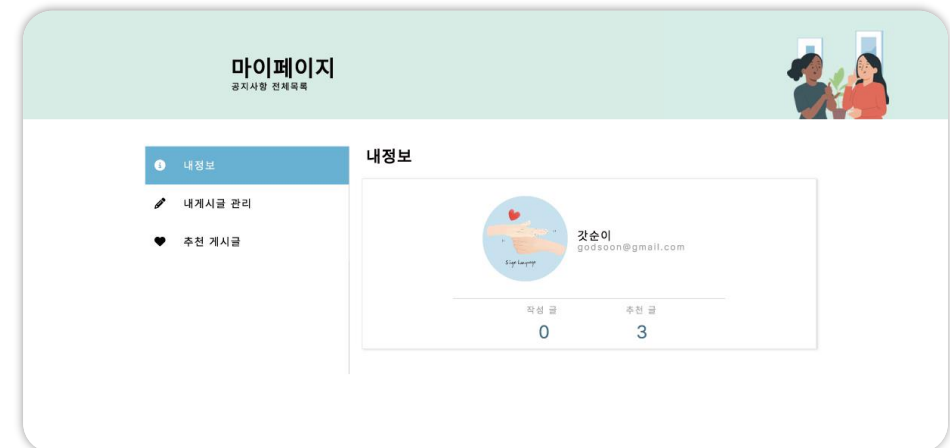
서브 기능

> 커뮤니티

- 이용자들간의 교류 및 서비스에 건의를 하는 기능
- 서비스의 지속 발전에 필요한 기능
 1. 게시글 조회 / 작성 / 수정 / 삭제를 할 수 있다.
 2. 댓글 조회 / 작성 / 수정 / 삭제를 할 수 있다.
 3. 회원 1명당 하나의 게시글에 공감 표현을 남길 수 있다.
 4. 회원 닉네임 또는 게시글 제목으로 검색할 수 있다.

> 회원관리

- 회원가입 후 게시글 관련 활동을 하는 기능
 1. 회원가입 / 탈퇴 / 정보수정을 할 수 있다.
 2. 내가 작성한 글 / 추천 글을 관리 및 조회 할 수 있다.



03 구현 과정 시연 영상



03 구현 과정

트러블 슈팅

모델링

> 물리적, 시간적 제약

- 본래의 기획의도와 목표에 맞는 성과를 구현하기 위해서는 현 원의 수강생이 감당할 수 없는 장비, 인력, 시간이 필요

CNN-LSTM

CNN-LSTM
+
Mediapipe-LSTM

CNN
+
Mediapipe

Mediapipe-
LSTM

Mediapipe

- 시도해 본 여러 학습 데이터 전처리 테스트 결과, 현실적으로 가능한 데이터 전처리 전략 선택

1. 학습시킬 테스트 수어 동작 선정 및 핵심 동작 선정
2. 직접 원시 데이터 생성 및 데이터 증강
3. Mediapipe로 pose, hand 랜드마크 추출
4. 추출한 marker의 x, y, z값 및 손가락, 팔 각도 저장
5. 평균값의 4배 이상 벗어나는 이상치 제거
6. knn, decision tree, catboost, lgbm 모델 테스트
7. train /test 값을 바꾸며 자체검증 후, 가장 정확도가 높은 catboost모델 선정

Out[14]: 0.9970149253731343



04

발전 방향

발전 방향 및 기대 효과

04 발전 방향

발전 방향 및 기대효과

발전 방향 및 기대효과

> 실시간 농인-청인의 의사소통 지원

- 현재의 모델은 프로토타입 수준이나, 데이터셋의 크기와 시간에 따라 정확도는 증가할 것임
- 분류 동작이 늘어나고, 다른 문장 생성 모델과 같이 성장하면서 동작 분류 모델과 문장 생성 모델의 상호보완으로 단일 모델 사용보다 오차발생을 줄일 수 있음
- 모든 수어 단어 데이터를 확보하여 청인이 전달한 텍스트에 알맞은 수어 영상을 조합하여 출력함으로써 실시간 의사소통을 지원할 수 있음

> 사회 서비스의 확대

- 사회 서비스를 누리는데 필수적인 수어만 분류하더라도 농인들의 고통이 많이 감소할 것으로 보임
- 청인과 농인간의 소통 간극을 줄이고, 농인들의 다양한 사회 참여 기회 증가
- 교육 및 학습, 의료서비스 등 다양한 분야에서 서비스 개선 및 편의 제공을 위해 사용

05

참고문헌 및 팀원 소개

참고문헌 / 팀원소개

참고문헌

> 모델링

- 조수민·조성연·신소연·이지향, 「청각 장애인을 위한 수어 영상-자연어 번역 서비스 및 모바일 어플리케이션 구현」, ACK 2021 학술발표대회 논문집, 28권 2호, p1012-1015
- 김미서·유문희·장지현·최정윤·나고은·김미숙·나정조, 「딥러닝 기반 수어 교육 온라인 플랫폼 구현」, 디지털콘텐츠학회논문지, Vol.23, No. 11, 2022, p2147-2157
- 김신영·엄서정·유선영·김수정·이경미, 「미디어파이프와 장단기기억을 이용한 수화동작인식 앱 개발」, 디지털콘텐츠학회논문지, Vol.24, No. 1, 2023, p111-119



팀장 **김군순** 경영학부

PM / Data Modeling / DB
Back-end / Front-end

- 프로젝트 기획 및 설계 및 일정관리
- 로고 디자인 및 화면 디자인 초안 설계

- 모델링 전략 설계
- Mediapipe, Tensorflow를 활용한 영상 / 이미지 데이터 가공
- 각 선정한 모델에 맞는 데이터 전처리
- 수어 번역 인공지능 학습/검증 및 활용

- 데이터베이스 설계
- JPA를 활용한 테이블 생성

- 스프링 시큐리티 활용, 보안/인가 관리
- FastAPI를 통한 파이썬 웹 개발
- 스프링부트, FastAPI 연결
- CORS설정

- 수어 번역 시작/중지에 따른 서버 요청
- openCV 실시간 스트리밍 송출
- 수어 영상 서버 전송 및 번역

- 모델링 / 백 / 프론트 에러 디버깅



팀원 **김다희** 사회복지

Back-end / Front-end
Data Modeling

- 프로젝트 기획 및 설계

- 게시글 CRUD 기능 구현
- 댓글 CRUD 기능 구현
- 회원 CRUD 기능 구현
- 수어사전 검색 기능
- 게시글 검색 기능
- 백엔드 에러 디버깅

- 수어 사전 카테고리 분류 비동기 처리
- 게시글 페이징
- 회원가입 이메일/닉네임 중복확인
- 프론트엔드 에러 디버깅

- 데이터 증강
- 수어 번역 단어 문장 변환



팀원 **김훈종** 전기공학

DB / Back-end
Data Modeling / Front-end

- 프로젝트 기획 및 설계

- 데이터베이스 설계
- JPA를 활용한 테이블 생성
- 크롤링 데이터 Amazone S3저장
- MySQL DB 에러 디버깅

- 게시글 CRUD 기능 구현
- 회원 CRUD 기능 구현
- 백엔드 에러 디버깅

- 데이터 증강
- 수어 번역 단어 문장 변환

- Ajax를 활용한 게시글 추천 비동기 처리
- 프론트엔드 에러 디버깅



팀원 **이대섭**

Front-end

- 프로젝트 기획 및 설계

- 메인 페이지 구현
- 수어 번역 페이지 구현
- 수어 사전 페이지 구현
- 커뮤니티 페이지 구현
- 공지사항 페이지 구현
- 마이 페이지 구현

- 수어 번역 페이지 구분 탭 비동기 처리
- 프론트엔드 에러 디버깅



팀원 **백지수** 식품영양

Front-end

- 프로젝트 기획 및 설계

- 커뮤니티 페이지 구현
- 로그인 / 회원가입 페이지 구현
- 마이 페이지 구현