

11조

Sign_Language.txt

시스템 설계서



과목명	졸업프로젝트1
담당교수	김두현 교수님
지도교수	임창훈 교수님
학과	컴퓨터공학과
팀원	201815004 이찬민
	201611205 박승민
	201814119 문지영

목차

1. Architecture

- 1.1 System Architecture
- 1.2 Software Architecture

2. Component Diagram

- 2.1 Component Definition
- 2.2 Component Interface
- 2.3 Component Algorithm

3. Sequence Diagram

- 3.1 Use-Case-Sequence Diagram

4. User Interface Configuration

- 4.1 메인
- 4.2 통역 화면

5. ER Diagram

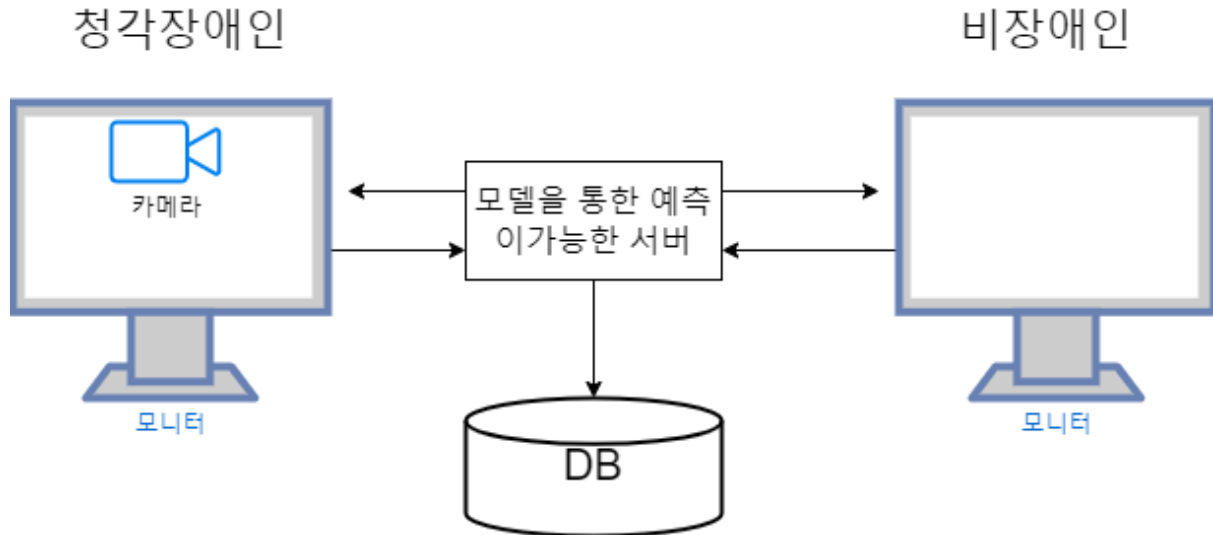
- 5.1 서비스, 수어 및 채팅 관계도

6. Programming Strategy

- 6.1 Programming Language
- 6.2 Programming Tool

1. Architecture

1.1 System Architecture



- 청각장애인 시작

내장 카메라를 통해 청각장애인의 수어를 인식합니다. 수어는 모델을 통한 예측 가능한 서버에서 수어 예측 모델로 수어 동작에 해당하는 텍스트 형태로 변환됩니다. 이때, 예측 결과 및 정확도 DB 관리를 통해 DB에 저장합니다. 그 다음, 번역된 텍스트는 대화 내역 DB 관리를 통해 DB에 저장됩니다.

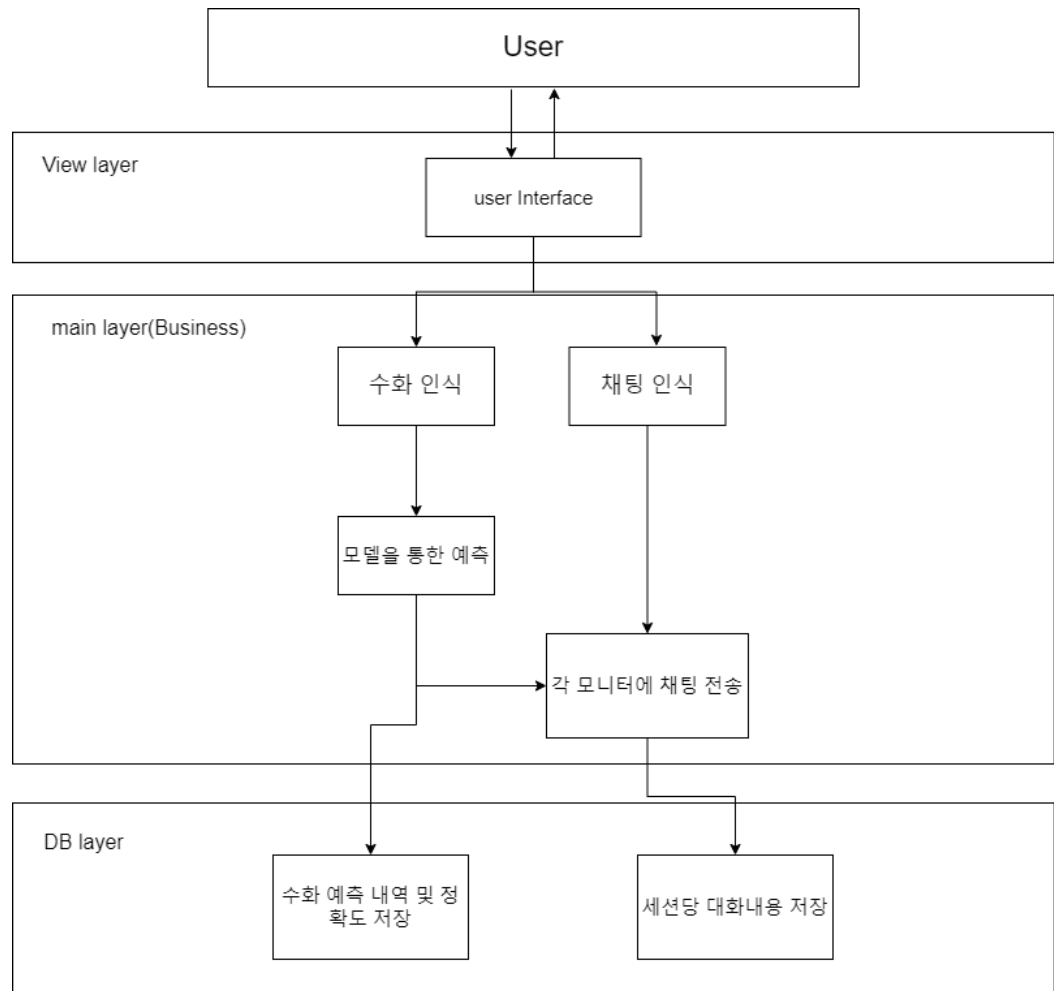
- 비장애인 시작

키보드를 통해 비장애인의 텍스트를 입력 받습니다. 그 후, 대화 내역 DB 관리를 통해 DB에 저장됩니다.

- 대화 창(모니터)

대화 내역 DB 관리를 통해 청각장애인과 비장애인의 대화 내역을 대화 창에 업데이트하여 모니터를 통해 소통이 가능합니다.

1.2 Software Architecture



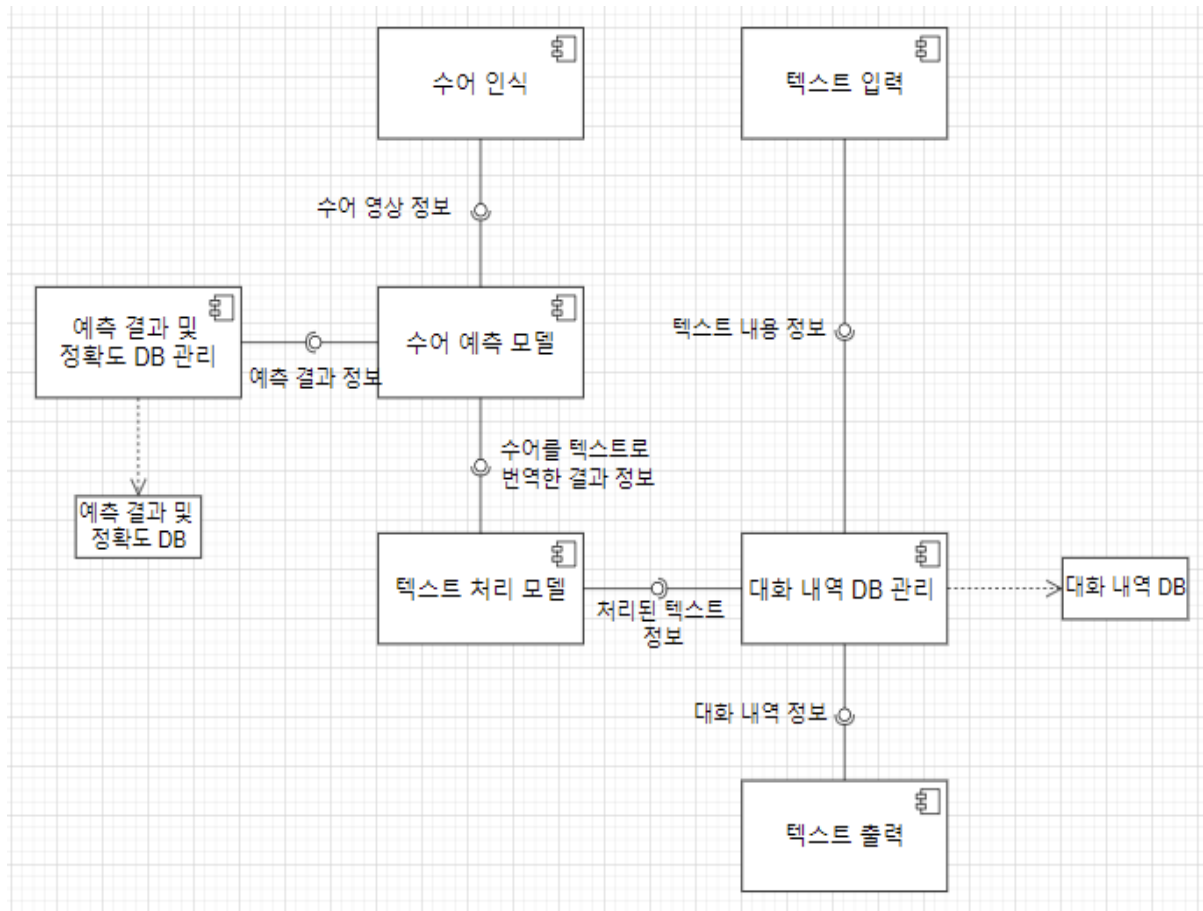
View Layer, Main Layer, DB Layer 로 총 3 개의 layer 로 이루어져 있습니다.

사용자는 View Layer 의 user interface 을 통해 프로그램을 사용할 수 있습니다.

Main Layer 에서 수화 인식, 모델을 통한 예측 순으로 이어지고, 각 모니터에 채팅 전송 및 DB Layer 의 수화 예측 내역 및 정확도 저장으로 데이터를 전송합니다.

채팅 인식의 경우, 각 모니터에 채팅 전송으로 데이터를 전송합니다. 마지막으로 각 모니터에 채팅을 업데이트하기 위한 세션당 대화 내용 저장이 필요합니다.

2. Component Diagram



컴포넌트 ID	컴포넌트명	개요	관련 유스케이스
SL_RCG	수어 인식	청각장애인이 시연하는 수화 영상을 입력 받음	수어 입력
SL_MDL	수어 예측 모델	Google MediaPipe를 사용하여 입력 받은 영상을 모델로 분석함	수어를 텍스트로 번역
TXT_MDL	텍스트 처리 모델	분석으로 나온 텍스트를 문법적으로 이해할 수 있을 정도로 검수	수어를 텍스트로 번역, 번역한 텍스트 출력
TXT_IN	텍스트 입력	비장애인의(의사)의 텍스트를 입력 받음	텍스트로 입력
TXT_OUT	텍스트 출력	분석된 수어의 텍스트와 비장애인이 입력한 텍스트를 출력	텍스트 화면 출력

PRD_DB_MNG	예측 결과 및 정확도 DB 관리	수화영상을 분석할 때 나온 텍스트, 정확도를 저장 및 관리	
CNV_DB_MNG	대화 내역 DB 관리	청각장애인과 비장애인간의 대화 내역을 저장	텍스트로 입력, 텍스트 화면 출력

2.1 Component Definition

컴포넌트 ID	SL_RCG	컴포넌트명	수어 인식
컴포넌트 개요	청각장애인이 시연하는 수화 영상을 입력 받음		
내부 클래스			
클래스명	비고		
RawSLData	카메라로부터 수어 영상을 받음. 받은 수어 영상에서 영상 처리할 데이터를 얻음.		
인터페이스 클래스			
ID	인터페이스명	오퍼레이션명	
SL_DAT	수어 영상 정보	SLData	

컴포넌트 ID	SL_MDL	컴포넌트명	수어 예측 모델
컴포넌트 개요	Google MediaPipe를 사용하여 입력 받은 영상을 모델로 분석함		
내부 클래스			
클래스명	비고		
SLData	수어 영상 데이터를 예측하기 전에 전처리 과정		
PredictionModel	수어 예측 모델을 통해 수어 동작에 해당하는 텍스트가 결과값 나옴		
SLResult	DB에 저장할 예측 결과 및 정확도 데이터를 처리		
인터페이스 클래스			
ID	인터페이스명	오퍼레이션명	
PRED_DAT	예측 결과 정보	PredictionData	
TRS_DAT	수어를 텍스트로 번역한 결과 정보	TranslationData	

컴포넌트 ID	TXT_MDL	컴포넌트명	텍스트 처리 모델
컴포넌트 개요	분석으로 나온 텍스트를 문법적으로 이해할 수 있을 정도로 검수		
내부 클래스			
클래스명	비고		
TextProcessingModel	예측 결과 정보를 읽기 쉽게 텍스트를 처리 텍스트 처리된 결과를 대화 내역을 전송		
인터페이스 클래스			
ID	인터페이스명	오퍼레이션명	
CHK_TXT_DAT	처리된 텍스트 정보	CheckedText	

컴포넌트 ID	TXT_IN	컴포넌트명	텍스트 입력
컴포넌트 개요	비장애인(의사)의 텍스트를 입력 받음		
내부 클래스			
클래스명	비고		
RawText	텍스트를 입력 받음.		
인터페이스 클래스			
ID	인터페이스명	오퍼레이션명	
RAW_TXT	텍스트 내용 정보	RawText	

컴포넌트 ID	TXT_OUT	컴포넌트명	텍스트 출력
컴포넌트 개요	분석된 수어의 텍스트와 비장애인이 입력한 텍스트를 출력		
내부 클래스			
클래스명	비고		
PrintConversation	대화창에 번역된 수어 텍스트와 입력된 텍스트를 출력		

컴포넌트 ID	PRD_DB_MNG	컴포넌트명	예측 결과 및 정확도 DB 관리
컴포넌트 개요	수화영상을 분석할 때 나온 텍스트, 정확도를 저장 및 관리		
내부 클래스			
클래스명	비고		
PredictionDBManager	예측 결과 및 정확도 DB를 관리		
인터페이스 클래스			
ID	인터페이스명	오퍼레이션명	
STR_PRD_DB	예측 결과 및 정확	PredDB	

	도 DB에 저장	
--	----------	--

컴포넌트 ID	CNV_DB_MNG	컴포넌트명	대화 내역 DB 관리
컴포넌트 개요	청각장애인과 비장애인간의 대화 내역을 저장		
내부 클래스			
클래스명	비고		
ConvHistoryDBManager	대화 내역 DB 관리		
인터페이스 클래스			
ID	인터페이스명	오퍼레이션명	
CNV_DAT	대화 내역 정보	ConvData	
STR_CNV_DB	대화내역 DB에 저장	ConvDB	

2.2 Component Interface

인터페이스 ID	SL_DAT	인터페이스명	수어 영상 정보
오퍼레이션명	SLData		
오퍼레이션 개요	수어 인식 컴포넌트에서 실시간으로 받은 수어 영상 데이터를 저장하고 수어 예측을 위한 SLData를 호출		
파라미터	입력 받은 영상 객체 : Object		
반환값	없음		

인터페이스 ID	PRED_DAT	인터페이스명	예측 결과 정보
오퍼레이션명	PredictionData		
오퍼레이션 개요	수어 예측 모델 컴포넌트에서 입력 받은 영상을 수어 예측 모델을 통해 나오는 예측 결과 및 정확도 값을 예측 결과 및 정확도 DB 관리 컴포넌트로 보냄		
파라미터	입력 받은 영상 객체 : Object		
반환값	없음		

인터페이스 ID	TRS_DAT	인터페이스명	수어를 텍스트로 번역한 결과 정보
오퍼레이션명	TranslationData		

오퍼레이션 개요	수어 예측 모델 컴포넌트를 통해 수어 영상을 텍스트로 변환된 이후 문법적인 확인을 위해 CHK_TXT_DAT 호출
파라미터	입력 받은 영상 객체 : Object
반환값	없음

인터페이스 ID	CHK_TXT_DAT	인터페이스명	처리된 텍스트 정보
오퍼레이션명	CheckedText		
오퍼레이션 개요	텍스트 처리 모델 컴포넌트에서 번역된 텍스트를 읽기 좋게(주어-목적어-동사 순 또는 단어의 나열) 처리한 값을 대화 내역 DB 관리 컴포넌트로 보냄		
파라미터	문법 확인된 텍스트 : String		
반환값	없음		

인터페이스 ID	RAW_TXT	인터페이스명	텍스트 내용 정보
오퍼레이션명	RawText		
오퍼레이션 개요	텍스트 입력 컴포넌트에서 비장애인의 텍스트를 입력 받아 대화 내역 DB 관리 컴포넌트로 보냄		
파라미터	입력된 텍스트 : String		
반환값	없음		

인터페이스 ID	CNV_DAT	인터페이스명	대화 내역 정보
오퍼레이션명	ConvData		
오퍼레이션 개요	대화 내역 DB 관리 컴포넌트에서 대화 내역 중 대화창에 표시하기 위한 정보들을 텍스트 출력 컴포넌트로 보냄		
파라미터	없음		
반환값	텍스트 : String		

인터페이스 ID	STR_PRD_DB	인터페이스명	예측 결과 및 정확도 DB에 저장
오퍼레이션명	PRDDB()		
오퍼레이션 개요	예측 결과 및 정확도를 예측 결과 및 정확도 DB에 저장		
파라미터	예측된 수어 텍스트 : String		
반환값	Boolean		

인터페이스 ID	STR_CNV_DB	인터페이스명	대화 내역 DB에 저장
오퍼레이션명	CNVDB()		
오퍼레이션 개요	대화내역을 대화내역DB에 저장		
파라미터	입력된 텍스트 : String		
반환값	Boolean		

2.3 Component Algorithm

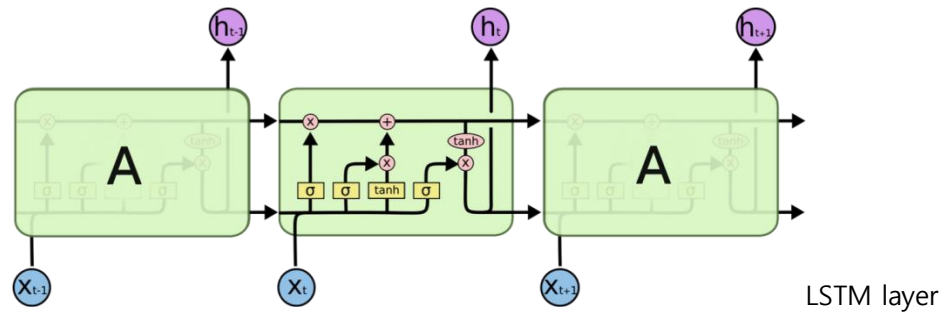
2.3.1 수어 예측 모델 생성

2.3.1.1 데이터셋 생성 알고리즘

- ▶ OpenCV를 이용하여 웹캠을 통한 video를 시작한다.
- ▶ 비디오창이 뜨면 몸에 Mediapipe Holistic을 이용하여 사람의 몸, 얼굴, 손에 point들을 찍어준고 이어준다.
- ▶ 지정된 시간동안 수어를 수행하고 그 영상으로부터 포즈, 얼굴, 왼손, 오른손을 분리하고 각 point들로부터 값을 추출하여 np array형태로 저장한다.
- ▶ 위의 과정을 30번 반복하여 한 단어당30개의 np array를 만들어낸다.

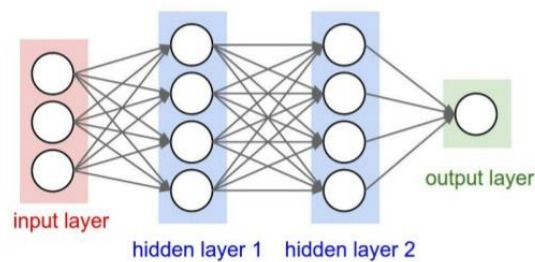
2.3.1.2 모델 생성 알고리즘

- ▶ 한 단어당 만들어진 30개의 np array를 하나의 window로 만들고 라벨을 붙여준다.
- ▶ Keras의 sequential 모델을 통해 layer를 구축
- ▶ RNN의 히든state에 cell-state를 추가한 구조인 LSTM layer 와 Dense layer를 이용하여 모델을 만든다.



Fully Connected(Dense) Neural Network

- Typical 3-layer **fully connected** neural network



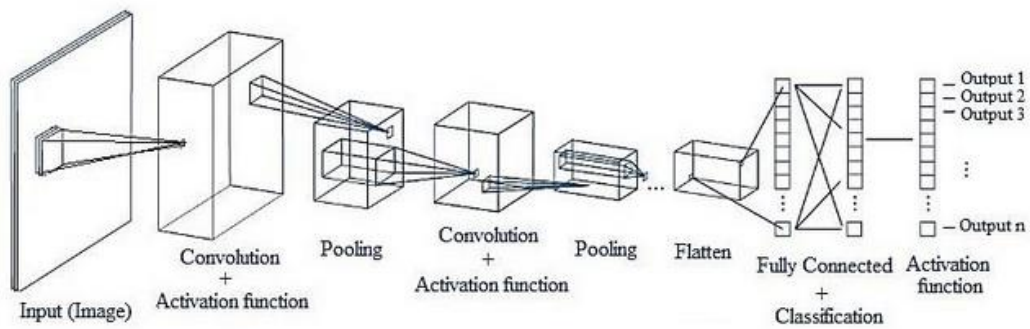
Dense layer

- ▶ optimizer로는 학습률, 방향 두가지를 모두 고려한 Adam을 사용하고 출력층에서 softmax 함수를 사용하는 다중 클래스 분류일 경우 사용하는 범주형 교차 엔트로피 손실 함수를 사용하여 모델을 컴파일한다.
- ▶ Epoch와 batch를 적절하게 설정하여 최대한 accuracy가 잘나오도록 한다.
- ▶ 모델을 통해 학습시키고 weight은 저장한다.

2.3.1.3 인식 실패시

- ▶ Layer의 교체 또는 정확한 데이터셋 사용 등 인식률을 높이는 과정을 더 추가한다.

2.3.1.4 인식률이 계속 낮을 경우



- ▶ 위의 RNN이 아닌 Convolutional Neural Network(CNN)을 이용 시도 한다.
- ▶ layer로는 Convolutional layer, Pooling layer 등을 사용한다.
- ▶ 완성후 RNN과 CNN중 더 나은 인식률을 보여주는 것으로 선택한다.

2.3.2 수어 예측 판정 알고리즘

2.3.2.1 알고리즘

- ▶ 카메라나 웹캠을 이용하여 수어를 인식하고 사람부분에 mediapipe holistic 을 이용하여 landmark(point)들을 찍는다.
- ▶ 찾아진 landmark로부터 keypoint를 추출하고 30프레임이 된 경우 예측을 수행한다.
- ▶ 수어들 중에 정확도 값이 85%를 넘어가는 것이 있다면 그 단어로 예측을 판정한다.

2.3.3 수어 인식 및 예측 결과

2.3.3.1 알고리즘

- ▶ 청각 장애인의 수어를 카메라나 웹캠을 통해 인식한다.
- ▶ 받아진 실시간 영상을 모델을 통해 예측한다.
- ▶ 예측 판정 알고리즘을 통해 판정을 완료한다.
- ▶ 결과를 채팅창에 보여준다.

2.3.3.2 예측이 틀린 경우

- ▶ 청각장애인 측에서 확인한다.
- ▶ 틀린 예측결과를 삭제한다.
- ▶ 다시 수어 인식을 통해 재시도를 하도록 한다.

2.3.4 DB저장 관련 서버

2.3.4.1 알고리즘

- ▶ 수어를 예측할 때 마다 예측 결과와 정확도를 DB에 저장한다.
- ▶ 전체 채팅 내역을 DB에 저장한다.

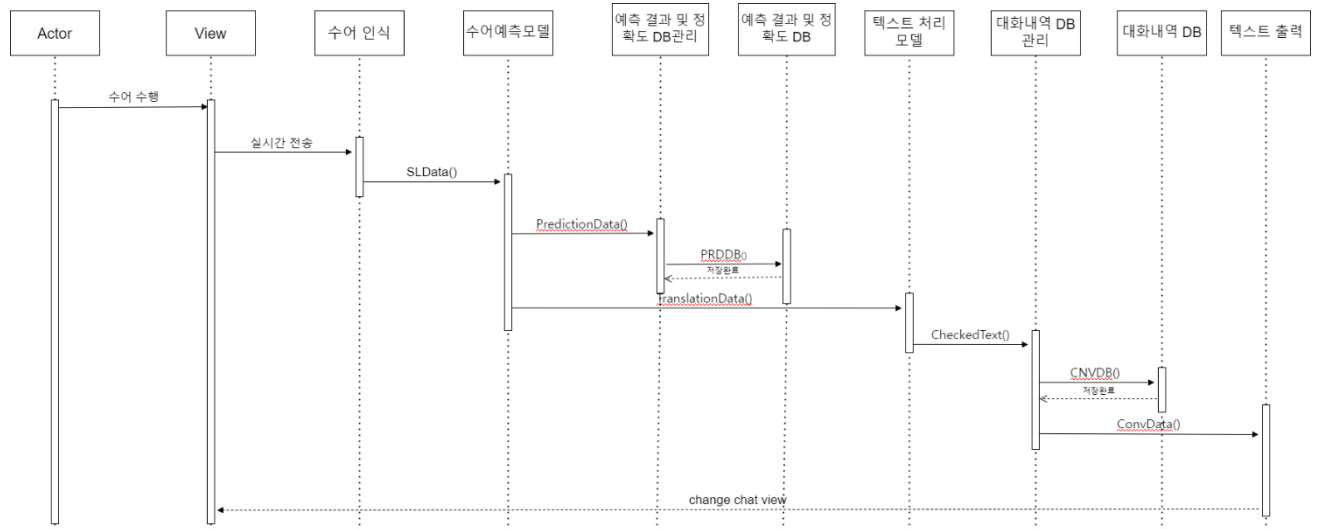
2.3.4.2 DB오류로 인한 저장 불가 해결

- ▶ 저장소를 2개를 운용하여 한쪽에 문제가 생기더라도 계속 저장할 수 있도록 한다.

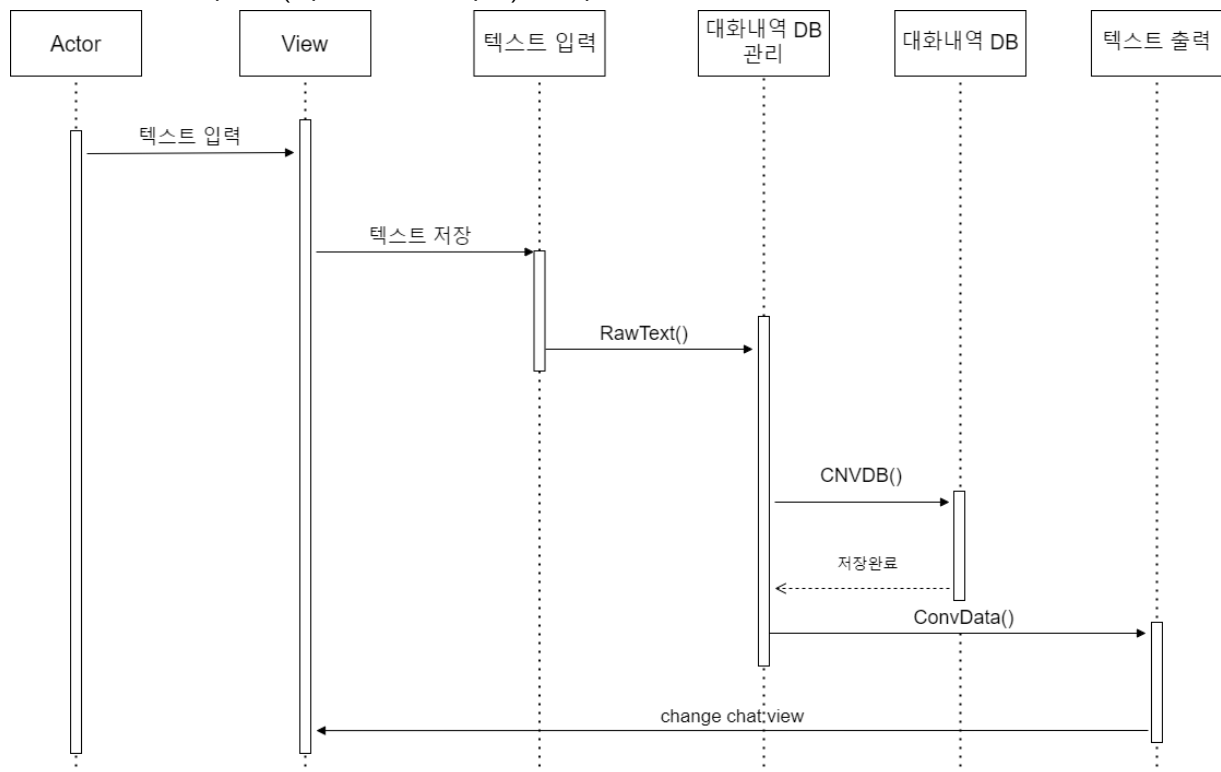
3. Sequence Diagram

3.1 Use-Case-Sequence Diagram

3.1.1 수어 인식 및 예측 후 출력



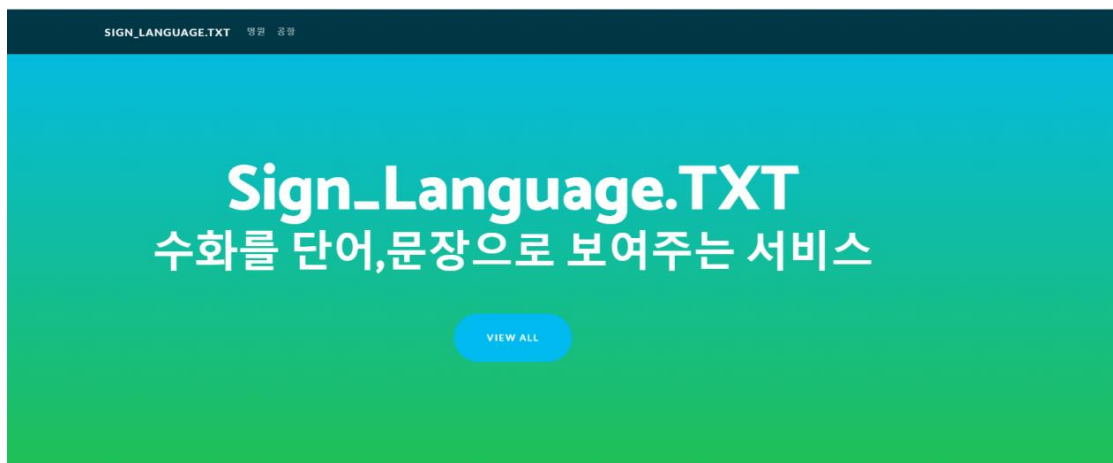
3.1.2 텍스트(키보드로 입력된) 출력



4. User Interface Configuration



4.1. 메인

화면 ID	TS_WM_00	화면명	메인 페이지
관련 유스케이스 ID	Use-case 다이어그램에 나와 있지 않은 메인 페이지		
관련 시퀀스도 ID	Sequence 다이어그램의 actor와 view부분		
화면유형	서비스 요청	메뉴경로	초기화면
화면개요	서비스를 사용하기 위해 웹에 접속했을 때 보여질 화면 Navbar옆의 사용처를 클릭하거나 View버튼을 클릭해 아래화면으로 이동 후 스크롤하여 사용할 서비스를 START버튼을 눌러 시작하도록 함		



About

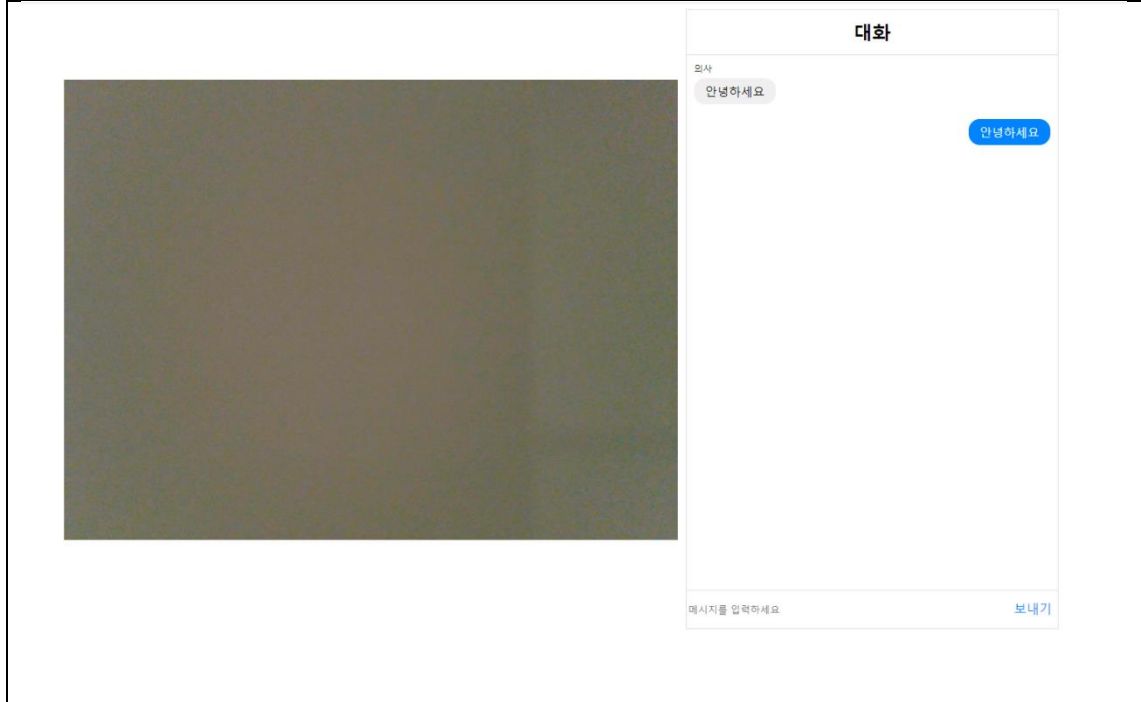


<div> <div> SIGN_LANGUAGE.TXT 변환 공통 </div> <div> <div> About Sign-Language.TXT </div> <div> Sign-Language.TXT란 청각장애인들이 공공장소에서 수화로 소통하는 것을 도와주는 모두에게 열려있는 서비스입니다. </div> </div> <div>  </div> <div>  </div> <div> <div> 병원에서 사용해 보세요. </div> <div> 수화를 사용하여 의사와 의사소통해보세요. </div> <div>START</div> </div> </div>				
입출력 항목				
항목명	컨트롤명	타입 및 길이	속성	Validation Check
수화 이미지	Sign_language(1)_img	ImageView	O / R	
의사 이미지	doctor_img	ImageView	O / R	
View 버튼	view_service_btn	Button	I	필수 항목
START 버튼	service_start_btn	Button	I	필수 항목
*속성 I:Input, O:output, R:ReadOnly, E:Editable, H:Hidden				
처리내용				
<ul style="list-style-type: none"> View 버튼: 사용자가 버튼을 클릭하여 현재 제공하는 서비스가 무엇인지 사용상황에 맞춰 사용할 서비스를 고를 수 있도록 아래쪽 view로 스크롤하여 내려준다. START 버튼: 장소에 따라 여러가지로 나뉜 서비스로 연결을 해준다. 				
기술적 고려사항				
<ul style="list-style-type: none"> view버튼 클릭 이후 스크롤하여 원하는 서비스를 선택하여 시작 가능하게 하는 경우 이외에도 네비게이션 바 옆에 사용 가능한 서비스를 나열해 놓음으로써 더 빠른 사용이 가능하도록 한다. 				

4.2. 통역 화면

화면 ID	MS_FS_00	화면명	수어 번역 및 채팅
관련 유스케이스 ID	텍스트로 입력, 자주 사용하는 표현은 버튼으로 선택, 수어를 텍스트로 번역, 번역한 텍스트 출력		
관련 시퀀스도 ID	3.1.1 수어 인식 및 예측 후 출력 3.1.2 텍스트(키보드로 입력된) 출력		

화면유형	입력, 서비스 요청	메뉴경로	메인 화면의 Start버튼 클릭 또는 navbar옆의 장소이름 선택
화면개요	비장애인 측은 채팅을 사용하여 청각장애인 측에게 보여주고 청각 장애인 측은 웹캠 또는 카메라를 통해 보여지는 화면에 수어를 수 행하도록 하는 화면		

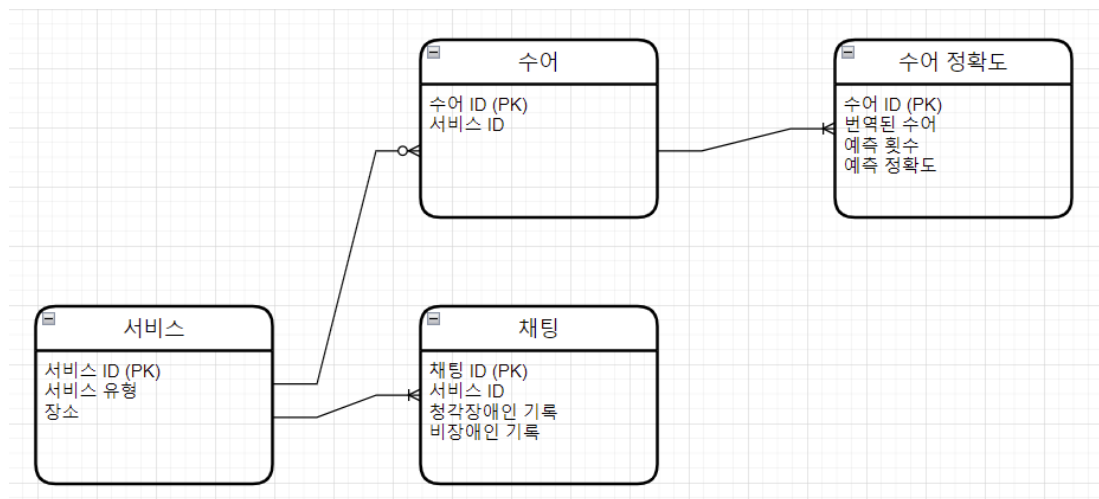


입출력 항목				
항목명	컨트롤명	타입 및 길이	속성	Validation Check
카메라(웹캠) 화면	camera_view	Video	I / R	필수 항목
채팅 화면	chat_view	ChatView	O / R	필수 항목
메시지 입력 창	chat_input_text	String	I / E	필수 항목
메시지 보내기	chat_input_btn	Button	I	필수 항목
*속성 I:Input, O:output, R:ReadOnly, E:Editable, H:Hidden				
처리내용				
<ul style="list-style-type: none"> - 카메라(웹캠) 화면: 기본적으로 그냥 켜져 있는 상황에서 청각장애인 측에서 대답해야 할 때 영상을 인식한다. - 채팅 화면: 비장애인 측에서 친 메시지의 내용과 청각장애인 측의 수어의 내용을 모두 화면에 출력하도록 한다. - 메시지 입력 창: 비장애인 측에서 키보드를 이용하여 메시지를 입력한다. - 메시지 보내기: 				

입력된 메시지를 채팅내역으로 보내준다.
기술적 고려사항
<ul style="list-style-type: none"> - 수어 예측이 틀렸을 경우 다시 인식될 수 있도록 해야 한다. - 채팅내역과 수어 예측 내용들이 데이터베이스에 저장되어야 한다. - 새로고침시 채팅내역을 리셋한다. - 수어 예측이 틀렸을 경우 다시 시도할 수 있도록 한다. - 사용자간의 메시지의 내용이 loss 되지 않도록 socket 통신을 구현하도록 한다.

5. ER Diagram

5.1. 서비스, 수어 및 채팅 관계도



6. Programming Strategy

6.1. Programming Language

Webpage : Javascript(+ HTML, CSS), **Server** : Java

Deep learning: Python

6.2. Programming Tool

VS Code, Anaconda