

시스템 정의서(프로젝트 요약서)

(2-15) 조

작품명 (주제)	(국문) AI 기반 수어 번역 서비스		
	(영문) AI-based sign language translation service		
책 임 자	성 명	박상준	
	소 속	소프트웨어학과	
	학 번	2020039047	
개발기간	2024년 9월 2일 ~ 2025년 12월 31일		
참여학생	학번	이름	전공
	2022041028	박유경	소프트웨어학부
	2022078054	최가은	소프트웨어학부
지도교수			
작품(주제)에 대한 요약			
작품 설명	AI 기반 수어 번역 서비스는 청각 장애인과 비장애인 간의 소통을 원활하게 하기 위한 AI 기반 수어 통합 소통 플랫폼을 개발한다. 이 프로젝트는 수어 사용 자의 소통 장벽을 줄이고, 더 많은 사용자에게 수어를 학습할 수 있는 기회를 제공하는 것을 목적으로 한다.		
작품의 주요 기능	<div>1. 수어 인식</div> <div>- 실시간 수어 인식</div> <div>웹캠을 통해 실시간으로 사용자가 수어를 입력하면, AI가 이를 인식하여 텍스트로 변환한다.</div> <div>- 손 움직임 및 제스처 추적</div> <div>OpenCV와 같은 컴퓨터 비전 기술을 통해 손의 위치와 움직임을 정확하게 추적하고, 손 모양을 통해 수어 동작을 인식한다.</div> <div>2. 수어 번역</div> <div>- 미리 학습된 수어 데이터셋을 기반으로, AI 모델이 수어를 분석하고 적절한 텍스트로 매핑한다.</div> <div>3. 실시간 자막 표시</div> <div>- 번역된 텍스트는 사용자가 실시간으로 볼 수 있도록 웹 화면에 자막 형태로 표시되며, 대화가 진행됨에 따라 자막이 순차적으로 업데이트된다.</div>		

	<p>3. 양방향 실시간 통화</p> <p>- 청각 장애인은 수어를 텍스트로, 비장애인은 음성을 텍스트로 번역하여 양방향 소통이 가능하도록 한다.</p> <p>4. 생성된 자막을 음성으로 제공</p> <p>- Text-to-Speech(TTS)</p> <p>AI가 수어를 텍스트로 번역한 후, 이 텍스트를 음성 합성 기술을 사용해 자연스러운 음성으로 변환한다. 사용자는 이를 통해 청각 장애인이 수어로 한 말을 음성으로 들을 수 있다.</p> <p>5. RAG (검색 증강 생성)</p> <p>- 텍스트 요약</p> <p>벡터 데이터베이스와 AI 임베딩 모델을 사용해 생성된 텍스트들을 데이터베이스에 저장하였다가, 대화가 종료되면 대화 내용을 요약해서 보여준다.</p> <p>- 정보 검색</p> <p>대화 중에 나온 특정 단어 또는 주제에 대해 사용자가 정보를 검색한다.</p>
--	--

작품(주제)에 대한 요약 (계속)	
운영개념	<ol style="list-style-type: none"> 1. 사용자가 웹캠을 통해 수어를 입력하고, 서버에서 실시간으로 수어를 분석하여 결과를 반환하는 웹앱을 통해 동작한다. 2. AI 기반의 수어 인식 모델은 미리 학습된 데이터를 기반으로 운영되며, 실시간으로 수어 영상을 받아 텍스트로 번역한다.
기타 개발 시 고려사항	<ol style="list-style-type: none"> 1. 데이터 수집: 수어 인식의 정확도를 높이기 위해 다양한 수어 데이터셋을 확보해야 한다. 기존 공개된 수어 데이터셋을 활용하거나, 자체적으로 데이터를 수집해야 한다. 2. 다양성 확보: 국가별, 지역별로 수어가 다를 수 있으므로 여러 언어의 수어 데이터를 고려하여 다국어 지원 가능성을 염두에 두고 개발해야 한다. 3. 접근성 및 U/UX: 청각 장애인뿐만 아니라 비장애인도 쉽게 사용할 수 있는 UI/UX를 설계해야 한다. 자막 크기, 색상 조정 등 접근성 기능을 제공한다.
오픈소스활용 및 기여 방안	<p>오픈소스 활용:</p> <ul style="list-style-type: none"> - TensorFlow 및 PyTorch: AI 기반의 수어 인식 모델을 학습시키기 위한 딥러닝 프레임워크를 활용한다. - MediaPipe: 손동작 인식 및 제스처 추적을 위한 오픈소스 도구로, 정확하고 빠른 손 인식 기능을 제공한다. - OpenCV: 영상 처리와 관련된 오픈소스 라이브러리로, 손의 움직임을 분석하고 데이터를 추출하는데 활용한다.

	<p>오픈소스 기여:</p> <p>- AI 모델 개선 및 공유: 수어 인식에 대한 연구와 모델 개선 결과를 오픈소스 커뮤니티에 공유하여, 더 많은 개발자가 해당 프로젝트를 발전시킬 수 있도록 기여할 수 있다.</p>
선행기술 조사 분석	<p><특허></p> <p>양방향 수화 번역 장치 및 장치가 수행하는 양방향 수화 번역 방법: 마이크를 통해 외부에서 발생하는 음성을 인식하여 상기 음성에 대응하는 수화를 출력하는 과정과 카 메라를 통해 감지되는 수화를 인식하여 상기 수화에 대응하는 음성을 출력하는 과정을 독립적으로 수행함으로써, 실시간으로 수화-음성, 음성-수화를 자동으로 번역한다. (https://patents.google.com/patent/KR20170094668A/ko)</p>
	<p><논문></p> <p>Depth-GASLT: 깊이 정보를 활용한 글로스 프리 수어 번역(Depth-GASLT: Gloss-free Sign Language Translation Using Depth Information): 2D 수어 이미지에서 추정된 깊이 정보를 활용하여 손의 3차원 위치를 반영함으로써 수어 번역 성능을 향상시키는 Depth-GASLT라는 새로운 아키텍처를 제안한다. (https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11860644)</p> <p>객체 인식 모델 YOLOV8을 이용한 온 디바이스 시스템 기반의 수어 번역기 (Sign language translator based on on-device system using object recognition model YOLOv8): YOLOV8과 Google Coral 개발 보드를 활용 해 휴대성과 사용성을 고려한 실시간 수어 번역 및 제스처 인식 시스템을 개발하여 비수어 사용자와 수어 사용자 간의 원활한 소통을 지원한다. (https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11917832)</p>
	<p><상용 제품></p> <p>'핸드사인톡톡'</p> <p>기존의 핸드사인톡톡이라는 공공기관, 병원, 은행 등 일상생활과 관련된 안내 내용을 3D아바타가 국어문을 수어로 번역해주는 서비스가 존재한다. 하지만 해당 서비스는 정보 제공을 중점에 둔 단방향 서비스인 반면, 청각 장애인 사용자가 비장애인 간의 양방향 소통을 할 때는 적합하지 않기 때문에, 사용자 간의 실시간 소통에 초점을 둔 서비스를 개발하고자 한다.</p>
Key Words (5개) : 수어, 청각 장애인, 영상통화, AI, RAG	
지도 교수	(서명)