


# 「2025 창의적 종합설계 경진대회」 과제보고서

대 학 명	대구대학교	팀 명	국문	굿바이				
			영문	GoodBye				
작 품 명	국문	싸인톡: 인공지능 기반 수어인식 프로그램						
	영문	SignTalk						
제작기간	2025년 07월 22일 ~ 2025년 10월 27일							
참가분야	<input type="checkbox"/> 일 반 <input type="checkbox"/> 다학제융합형 <input type="checkbox"/> 산학협력형* (산업체과제수행, 산업체인사멘토링, 작품제작지원, 기타) <input checked="" type="checkbox"/> 사회기여형* (사회적약자 기여형, 지역사회 기여형 <b>중 택1</b> ) *산학협력형 및 사회기여형의 경우, 확인서(붙임 참조) 필수제출							
지식재산권	<input type="checkbox"/> 해당 (특허등록(등록번호 기입), 특허출원(출원번호 기입), 실용신안(등록번호 기입), 기타(지재권유형기입) <b>중 택1</b> ) <input checked="" type="checkbox"/> 해당없음							
포상실적	대회명	주최/주관	상격	수상일시				
		/						
		/						
작품 내용요약	<p>SignTalk은 AI 기반 한국 수어 지문자 실시간 인식 및 학습 플랫폼입니다. Flutter 크로스플랫폼 앱으로 Android, iOS, Web에서 동일하게 작동하며, MediaPipe 손 랜드마크 추출과 딥러닝 모델을 결합하여 실시간 수어 인식을 제공합니다.</p> <p>5단계 커리큘럼으로 체계적 학습을 지원하며, 각 단계마다 즉각적인 피드백과 손모양 분석을 제공합니다. 정적 자모는 MLP 모델로, 쌍자음과 복합모음은 Bidirectional LSTM 모델로 인식하여 인식 정확도를 높였습니다. 4단계 퀴즈 시스템(날말/초급/중급/고급)으로 학습 효과를 검증하고, JWT 기반 사용자 인증과 SQLite 데이터베이스로 개인별 학습 진도를 관리합니다.</p> <p>주요 특징은 다음과 같습니다. 첫째, 스마트폰만 있으면 24시간 언제 어디서나 무료로 학습할 수 있어 수어 학습의 진입 장벽을 낮췄습니다. 둘째, TFLite 경량화로 모바일 및 임베디드 환경에서도 실시간 추론이 가능합니다. 셋째, 데이터 증강 기법으로 다양한 손 크기와 표현 방식을 학습하여 실사용 환경에서 강건한 성능을 보장합니다. 넷째, 스킵 기능과 복습 모드로 학습자의 수준과 속도에 맞춘 개인 맞춤형 학습을 지원합니다. 이를 통해 농인과 비장애인 간 소통 장벽을 낮춰 포용적 사회 구현에 기여합니다.</p>							
작품사진								
								

## 1. 과제의 목적, 개발동기 및 필요성

SignTalk은 AI 기반 한국 수어 지문자 실시간 인식 및 학습 플랫폼입니다. 본 과제의 목적은 수어 학습의 시간·장소·비용 제약을 제거하고, 스마트폰만으로 40개 한국 수어 지문자를 체계적으로 학습할 수 있는 환경을 제공하는 것입니다. 사용자가 카메라 앞에서 손모양을 보이면 AI가 실시간으로 인식하여 정확도를 분석하고 즉각적인 교정 피드백을 제공합니다. 정지된 손모양뿐만 아니라 연속 동작이 필요한 쌍자음과 복합모음까지 인식하여 한국 수어 지문자 전체를 학습할 수 있습니다.

수어를 배우고 싶어도 시작조차 어려운 현실이 개발의 출발점이었습니다. 수어 학원은 전국 50개소 미만으로 접근성이 극히 낮고, 학원까지 이동하는 시간과 비용도 부담입니다. 수강료는 월 10만원 이상으로 경제적 부담이 크며, 주 1~2회 정해진 시간에만 학습 가능해 직장인이나 학생에게는 현실적으로 어렵습니다. 독학을 시도해도 올바른 손모양을 확인하는 과정이 복잡하고 즉각적이지 않아 학습 효과가 떨어지고 지속성을 유지하기 어렵습니다. 이러한 높은 진입 장벽이 수어 입문자들의 첫걸음을 가로막고 있습니다. SignTalk은 이 문제를 해결하기 위해 탄생했습니다.

국내 주요 서비스와 비교 시 명확한 필요성이 보입니다. 한국수어사전 앱(국립국어원)은 8,000개 단어의 동영상 사전을 제공하지만 일방향 정보 제공에 그치며 사용자 인식 기능이 전혀 없습니다. 손말이음(한국농아인협회)은 실시간 통역 서비스이나 평균 10분 대기 시간과 평일 09:00~18:00 운영 시간 제한이 있습니다. 온라인 강의는 월 비용과 실습 피드백 부재가 한계입니다.

독학을 시도해도 올바른 손모양을 확인하는 과정이 복잡하고 즉각적이지 않아 학습 효과가 떨어지고 지속성을 유지하기 어렵습니다.

SignTalk은 기존 서비스들과 달리 양방향 학습(인식+피드백), 5단계 체계적 커리큘럼, 24시간 무료 이용, 개인별 진도 추적을 제공하여 수어 입문의 문턱을 낮추고자 했습니다.

## 2. 과제의 해결방안 및 수행과정

한국 수어 지문자 40개는 정적 자모음 31개와 동적 자모음 9개로 분류됩니다. 정적 자모음(ㄱ, ㄴ, ㄷ 등)은 정지된 손모양으로 단일 프레임 인식이 가능하지만, 쌍자음 5개(ㄲ, ㄳ, ㄴㄹ, ㄷㄹ, ㄹㄹ)와 복합모음 4개(ㅏ, ㅑ, ㅓ, ㅕ)는 연속 동작으로 시간적 순서 정보가 필요합니다. 이를 해결하기 위해 하이브리드 AI 모델 전략을 수립했습니다. 정적 자모음은 Dense Neural Network로, 동적 자모음은 Bidirectional LSTM으로 인식하는 이중 모델 구조를 설계했습니다.

개발 과정은 다음과 같습니다. 먼저 MediaPipe Hands로 손 랜드마크를 추출하는 전처리 파이프라인을 구축했습니다. 정적 모델은 21개 랜드마크의 좌표를, 시퀀스 모델은 손목과 검지 끝의 움직임 정보를 특징으로 사용했습니다. 데이터 부족 문제는 노이즈 추가, 스케일 변경, 속도 조절 등의 증강 기법으로 해결하여 원본 대비 5배 확장했습니다. 특히 혼동되기 쉬운 10개 자모에 집중 증강을 적용했습니다.

학습 과정에서 과적합 문제가 발생하여 Early Stopping과 Dropout을 적용했고, Z-score 정규화로 다양한 환경에서도 일관된 인식 성능을 확보했습니다. 초기에는 매 프레임마다 AI 인식을 수행했으나 CPU 사용률이 80%를 초과하여, 인식 주기를 0.15초로 조정하고 카메라 버퍼를 최소화하여 CPU 사용률을 30% 이하로 개선했습니다. 동적 자모 인식을 위해 사용자별 시퀀스 버퍼를 구현하여 프레임들을 수집하고, 충분한 데이터가 모이면 LSTM 모델로 예측하는 방식을 채택했습니다.

## 3. 과제의 내용

본 과제는 딥러닝 기반 컴퓨터 비전 기술을 활용한 한국 수어 지문자 실시간 인식 및 학습 플랫폼입니다.

니다. 이론적 근거는 Dense Neural Network의 공간적 패턴 인식 능력과 Bidirectional LSTM의 시계열 데이터 처리 능력을 결합한 하이브리드 아키텍처에 있습니다. MediaPipe Hands 모델을 통해 손 랜드마크 21개 지점의 2차원 좌표를 실시간으로 추출하고, Z-score 정규화를 적용하여 조명, 카메라 각도, 손 크기 등 다양한 환경 조건에서도 일관된 인식 성능을 보장합니다. 특히 Bidirectional LSTM은 과거와 미래 정보를 동시에 활용하여 시간적 순서가 중요한 쌍자음과 복합모음 인식에 핵심적인 역할을 합니다.

시스템은 3계층 아키텍처로 설계되었습니다. 프레젠테이션 계층은 Flutter 3.9.2 기반 크로스플랫폼 앱으로 iOS, Android, Web을 지원합니다. 반응형 UI 설계로 다양한 화면 크기에 대응하며, Provider 패턴으로 상태 관리를 구현했습니다. 비즈니스 로직 계층은 Flask 3.1.2 서버로 JWT 기반 사용자 인증, 학습 진도 관리, AI 추론, 퀴즈 시스템을 처리하는 15개 RESTful API 엔드포인트를 제공합니다. 주요 API로는 회원가입/로그인, 손모양 분석, 학습 진도 조회/업데이트, 퀴즈 결과 저장, 통계 조회 등이 있습니다. 데이터 계층은 사용자 정보와 학습 기록을 저장하는 SQLite 데이터베이스와 TensorFlow 2.19.1 기반 AI 모델로 구성됩니다.

정적 모델은 5층 Dense 신경망 구조로 설계되었습니다. MediaPipe가 추출한 21개 랜드마크의 x, y 좌표(총 42차원)를 입력받아 31개 기본 자모음(ㄱ~ㅎ, ㅏ~ㅣ 등)을 분류합니다. Dropout 레이어를 각 Dense 레이어 사이에 배치하여 과적합을 방지하고, Adam 옵티마이저와 Categorical Crossentropy 손실 함수를 사용했습니다.

시퀀스 모델은 3층 Bidirectional LSTM과 2층 Dense 신경망으로 구성됩니다. 손목(랜드마크 0번)과 검지 끝(랜드마크 8번) 2개 지점의 x, y 좌표, 이전 프레임 대비 변화량(dx, dy), 이동 속도(spд\_sum) 총 5개 특징을 각 지점마다 추출하여 10차원 특징 벡터를 생성합니다. 최대 30프레임(약 1.5초)의 시퀀스 데이터를 수집하여 9개 쌍자음/복합모음(ㅁ, ㅂ, ㅅ, ㅈ, ㅊ, ㅋ, ㆁ, ㆅ, ㆆ)을 분류합니다. 양방향 LSTM은 순방향과 역방향 정보를 모두 학습하여 동작의 시작과 끝을 정확히 파악합니다.

또한, 확장 가능성을 염두해 두 모델 모두 TFLite 포맷으로 경량화하여 모바일 및 라즈베리파이 같은 임베디드 환경에서도 실시간 추론이 가능합니다. 모델 크기는 정적 모델 약 2MB, 시퀀스 모델 약 3MB로 앱 용량 부담이 적습니다.

정적 모델은 3,298개 원본 샘플을 수집하여 데이터 증강 후 약 16,000개로 확장했습니다. 학습 데이터 2,638개, 검증 데이터 660개로 8:2 분할하여 학습했으며, 검증 데이터 기준 99.6% 정확도를 달성했습니다. 시퀀스 모델은 451개 원본 샘플을 수집하여 증강 후 약 2,255개로 확장했습니다. 학습 데이터 361개, 검증 데이터 90개로 분할하여 학습했으며, 검증 데이터 기준 99.8% 정확도를 기록했습니다. 실제 사용 환경에서는 다양한 사용자와 환경 변수로 인해 약 90%의 정확도를 보입니다.

시스템 성능은 30fps 카메라 스트리밍과 0.15초 인식 주기로 즉각적인 피드백을 제공하며, CPU 사용률 30% 이하, 메모리 200MB로 배터리 효율적입니다. 정적 자모음은 단일 프레임으로 즉시 인식하고, 동적 자모음은 최소 5프레임 수집 후 예측하여 사용자에게 진행률을 표시합니다.

#### 4. 기대효과 및 활용 방안

국내 청각 장애인은 약 40만 명이며, 이들과 소통하고자 하는 잠재 학습자 규모가 상당합니다. 현재 수어 교육은 학원 수가 제한적이고 월 10만원 이상의 수강료 부담으로 접근성이 낮은 상황입니다. 이에 누구나 무료로 이용할 수 있는 공공 학습 플랫폼을 제공함으로써, 장애인 인식 개선과 포용적 사회 구현에 기여하고자 합니다. 본 서비스는 완전 무료로 제공되어 모든 이용자가 비용 걱정 없이 수어를 배울 수 있도록 하며, 지역 사회의 수어 교육 인프라 구축과 사회적 가치를 실현하는 데 중점을 둡니다. 특히 스마트폰 사용이 어려운 고령층과 아동을 위해 TFLite 경량화 모델을 탑재한 저가형 학습 전용 태블릿 기기 등을 활용해 교육 접근성을 더욱 확대할 계획입니다. 현재는 수어 입문자를 위한 지문자 수어부터 시작해 점차 단어 및 문장 수어 교육으로 확대하여, 입문자에서 고급자까지 모든 수준의 학습자가 체계적으로 수어를 마스터할 수 있는 완전한 학습 생태계를 구축할 예정입니다. 또한 저사양 기기에서도 원활히 작동 가능한 TFLite 경량화 덕분에 도서관, 주민센터, 문화센터 등 공

공장소에 태블릿 키오스크를 설치하여 공공 인프라에서의 활용도 가능합니다. 교육기관에서는 초·중·고 장애인식 개선 교육 및 수어 체험 프로그램으로, 공공기관에서는 민원 창구 직원 대상 기초 수어 교육으로, 복지시설에서는 청각 장애인 복지관과 노인복지관의 무료 학습 도구로 활용될 수 있습니다. 향후 수어를 잘 모르는 사람이 말하는 것을 애니메이션으로 수어 표현하는 기능, 수어하는 사람의 동작을 실시간으로 글로 변환하는 양방향 소통 기능으로 확장할 계획입니다. 이를 통해 농인-비장애인 간 소통 장벽을 낮춰 포용적 사회 구현에 기여할 것으로 기대됩니다.

### 5. 비용분석 및 팀원 간 역할 분담

비용 분석			팀원 간 역할 분담		
항목	세부항목	소요비용 (천원)	성명	역할	참여도 (%)
재료비			권채은	초기 전체 개발 및 백엔드, 모델 학습, 형상관리, 문서작업	50
시제품가공비			김하령	프론트엔드, 데이터 수집	25
...			권소진	데이터 수집, 데이터 관련 처리, 발표	25
기타 경비			...		
합 계			합 계		100%

### 6. 참고문헌