

Jetson Nano 기반 실시간 수어-문자 번역 웨어러블 기기

I. 작품 개요

● 프로젝트 선정배경

- 청각장애인의 의사소통 삶의 질이 가장 낮은 것으로 조사됨
- 수어 사용자 90.8%가 사용하기 가장 적절한 언어는 한국 수어라고 응답함
- 공공 서비스 환경에서의 농인 정보 소외 문제

● 프로젝트 목표

- 수어를 자연스러운 한국어로 실시간 번역하는 플랫폼
- RAG 기반 GPT 연동을 통해 문맥보정
- Jeston Nano 기반 휴대용 장치 개발

II. 작품 내용

● 소프트웨어

① 모델 채택 과정

❖데이터 구성

AI-HUB에서 교통 관련 수어 단어 10개 선정 후 750개의 영상 수집하였음. 재생 속도 변형(0.85~1.2배)을 통해 7배 규모로 데이터 증강하였고 총 5,250개의 데이터 확보

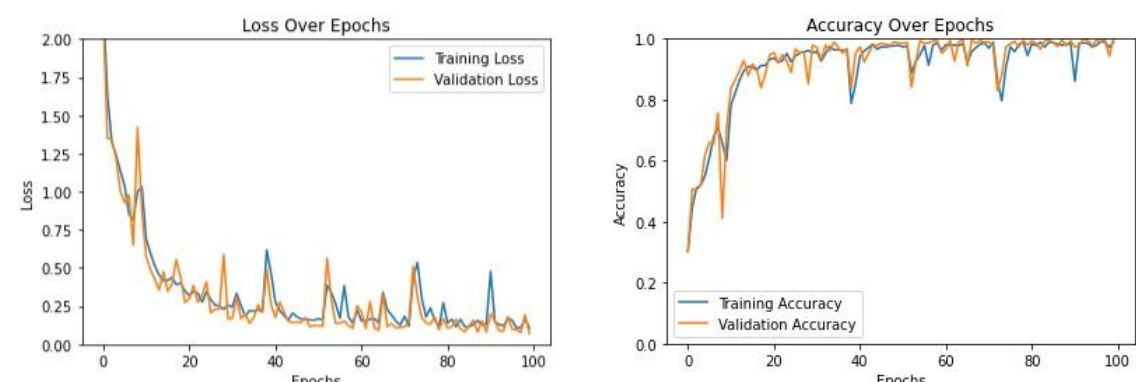
❖모델 호환성 확보

초기 모델은 colab 환경에서 LSTM 2계층 모델(256,128, Dropout 0.2, Dense 256) 기반으로 학습. 실행 과정 중 Jetson과 Tensorflow 버전 호환성 문제가 발생하였음. Jetson과 동일한 버전의 로컬 가상환경에서 재학습으로 호환성 문제를 해결

❖모델 성능 개선

실시간 추론 성능이 낮게 측정되었으며, 이는 Jetson의 하드웨어 성능 제한으로 인한 낮은 FPS와 라이브러리 버전 차이에서 기인한 것으로 분석됨. 이를 개선하기 위해 3가지 실험 모델을 구성하고, 각각의 정확도를 비교

정확도 측정은 Jetson Nano에 연결된 카메라를 통해 10개의 단어를 각 10회씩 실시간 번역하여 총 100회 테스트하였음. 3가지의 실험 모델 중 정확도가 84%로 가장 높았던 LSTM 3계층 모델(256,128,64, Dropout 0.2, Dense 256)을 최종 모델로 선정함



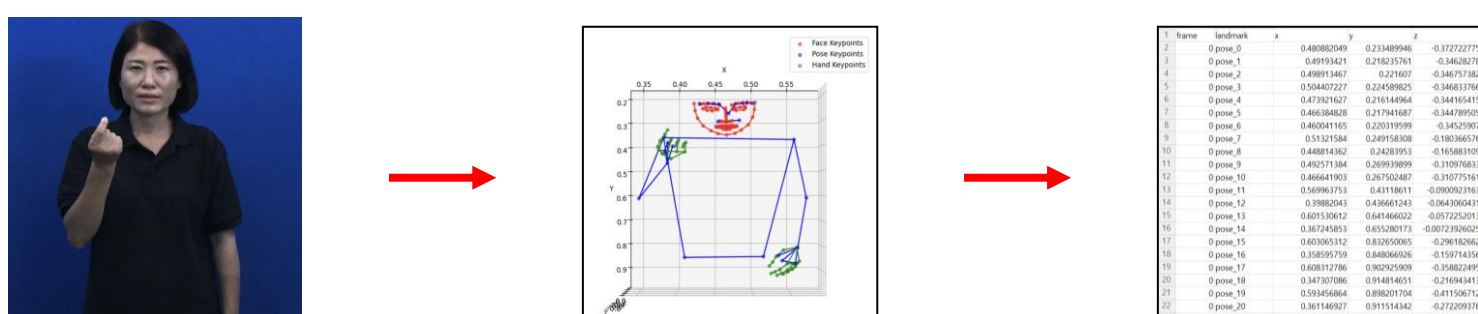
최종 모델 Loss 및 Accuracy plot

구분	모델구성	변경사항	정확도	결론
A. 경량화 모델	LSTM => GRU	홀수 프레임만 학습	53%	FPS 개선 효과 없음
B. 데이터 추가 모델	기존 LSTM	직접 촬영 데이터 추가	47%	인식률 저하
C. 계층 수 추가 모델	LSTM 3계층	계층 수 확장	84%	정확도 향상



3가지 실험모델 구성 및 세부정확도

② 실시간 키포인트 추출과정



❖ 키포인트 추출

MediaPipe 라이브러리를 사용하여 전신 25개 + 왼손 21개 + 오른손 21개 + 얼굴 68개 = 총 135개로 구성
각 키포인트는 (x, y, z) 좌표로 구성 → 프레임당 405개 데이터 생성

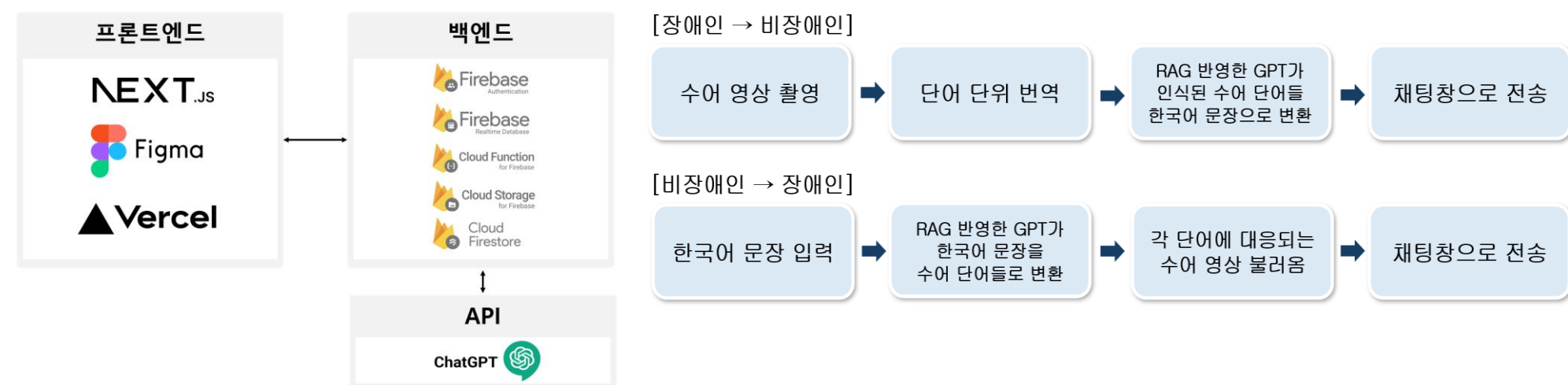
❖버퍼 저장 및 정규화

5초간 프레임 누적 후 추론, 추론 시 최근 30프레임만 사용 (부족 시 zero-padding) → 입력 형식: (30, 405)

❖TFLite 기반 추론

정확도 확률이 0.7 이상일 경우 해당 단어 인식, 그 외는 “waiting” 상태로 처리 → 안정성 확보

③ 애플리케이션 개요



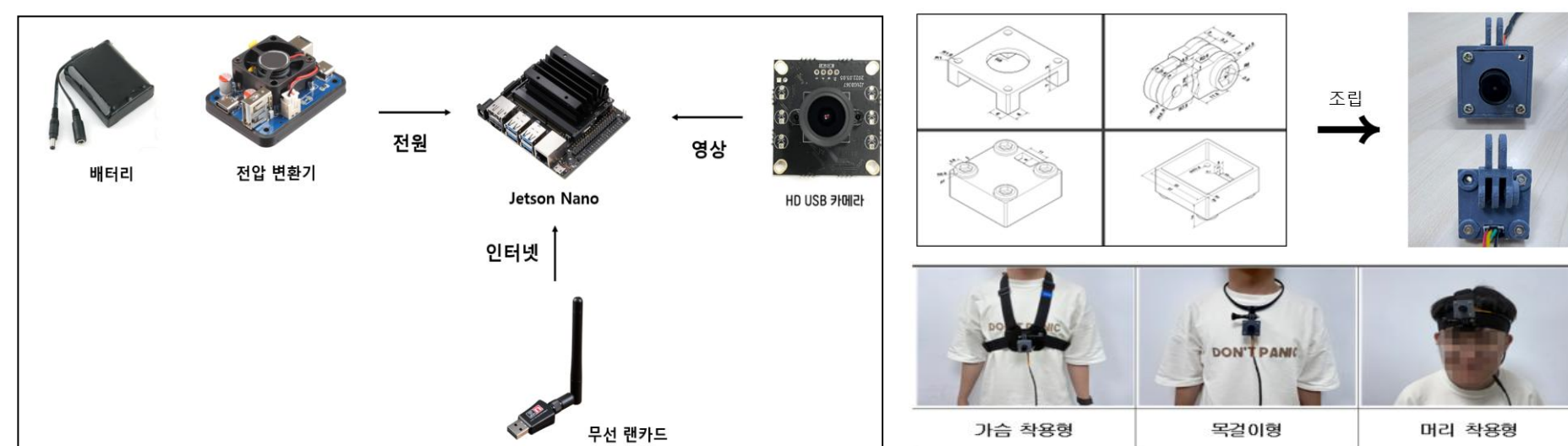
애플리케이션 개발도구

애플리케이션 흐름도

❖ RAG(Retrieval-Augmented Generation)기법 적용

국립국어원의 ‘한국어-한국수어 병렬 말뭉치’ 중 교통 상황 관련 약 2,100개 문장을 선별하여 Firestore에 저장하였음. 이후 각 문장을 임베딩한 뒤, 입력 문장과 유사한 문장을 검색하고, 이를 기반으로 GPT API를 활용해 자연스러운 한국어 문장으로 변환하였음. 최종적으로 생성된 문장은 사용자에게 전송되었으며, 수어 영상 제작 시 필요한 수어 단어 리스트 추출에도 활용되었음

● 하드웨어



하드웨어 구성요소

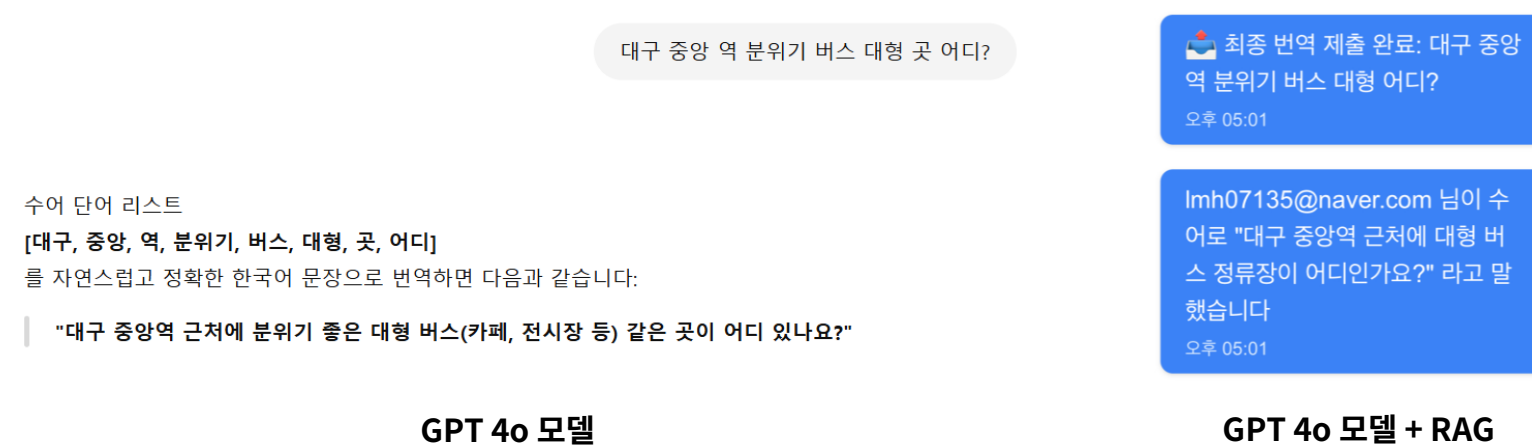
카메라 모듈 제작 및 착용사진

❖ 카메라 모듈 케이스 제작 및 착용

착용 방식은 액션캠 형태의 거치형 장비를 기반으로 하였으며, 해당 장비에 카메라 모듈을 결합할 수 있도록 전용 케이스를 제작하였음. AutoCAD를 활용하여 설계 도면을 작성한 후, 3D 프린터를 이용해 케이스를 출력·제작함. 착용 방식은 가슴 착용형, 목걸이형, 머리 착용형의 총 3가지로 구성되어 있으며, 사용자 기호에 따라 선택적으로 착용할 수 있도록 하였음

III. 작품 결과

● RAG 성능 비교



GPT 4o 모델

GPT 4o 모델 + RAG

● 최종 성능 평가

통합 안정성	수어 입력부터 채팅방 전송까지 전체 프로세스가 중단 없이 정상 작동하는 비율
통합 정확도	사용자의 수어 표현이 의도한대로 번역되어 채팅 메시지로 전달된 비율

성능은 최종 시스템의 전체 동작 흐름(수어 입력 → 단어 추출 → 문장 변환 → 채팅방 전송)의 정상 수행 여부에 대하여 평가하였으며, 시스템 성능은 통합 안정성, 통합 정확도 항목을 기준으로 평가함
총 50회의 테스트 중 모든 테스트가 중단 없이 완료되어 **통합 안정성은 100%**로 평가됨. 총 50회의 테스트 중 38회에서 수어 단어를 정확히 인식 후 변환하여 **통합 정확도는 76%**로 평가됨

IV. 향후계획 및 기대효과

● 기대효과

- 수어 비사용자와의 실시간 소통 가능
- 공공서비스 환경에서의 정보 격차 해소
- 청각장애인 고용/상호작용 환경 지원

● 향후계획

- 수어 구간 자동 분리 기능 도입
- 수어 단어 범위 확대 (교통 → 일상전반)