

# 인공지능 기반 수어 인식 기술

한국전자기술연구원

2021.11.26.

인공지능연구센터

박 한 무

# 목 차

---

I. 개요

II. 수어 인식을 위한 학습 데이터

III. 수어 인식 알고리즘

IV. 향후 연구 방향

V. 결론

CHAPTER

I

개요

## • 한국 수어

- **농인\*** 들이 의사소통을 하는 과정에서 **자연 발생**한 언어

\* 청각장애인을 달리 이르는 말로서 수어를 일상어로 사용하는 사람

- 한국어와는 **어휘와 문법** \* 체계가 다른 별개의 언어

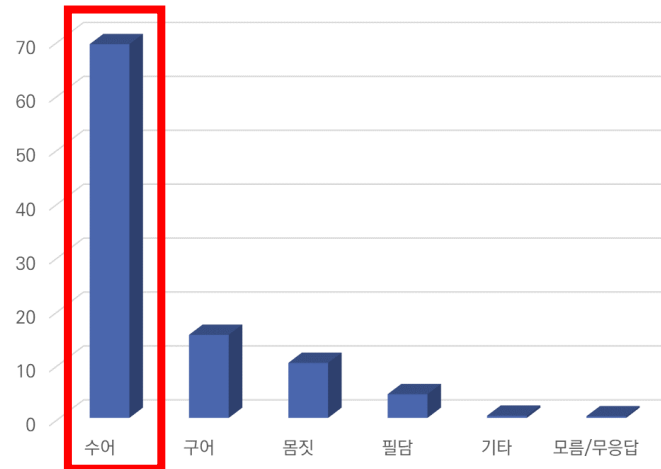
\* 한국어 어휘를 형태소 단위 치환하는 방식으로 만들어지지 않음

- 나라별로 **다른 수어**를 사용함\*

\* 국제 수어가 존재하기는 하지만 국내 농인들에게는 익숙하지 않음



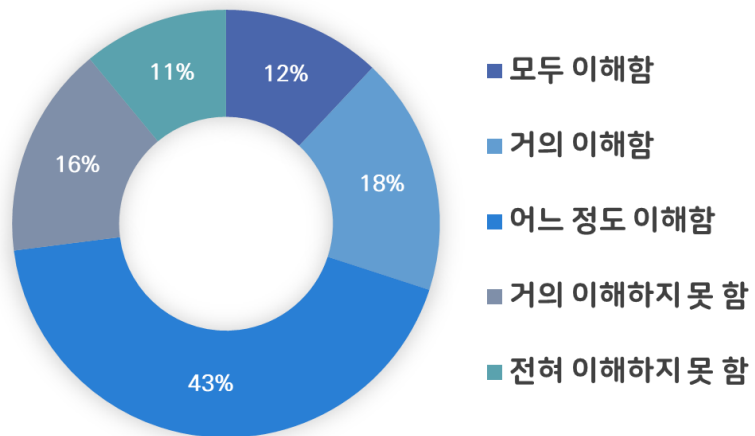
한국 수어, "안녕하세요", 국립국어원



농인의 주 의사소통 방식, 국립국어원

- 농인-청인 간 의사소통의 어려움

- 약 12% 농인만이 청인 수준으로 필담이 가능
- 약 9% 농인만이 TV 자막의 의미를 이해할 수 있음
- 농인의 문맹률은 약 30%로 추정됨

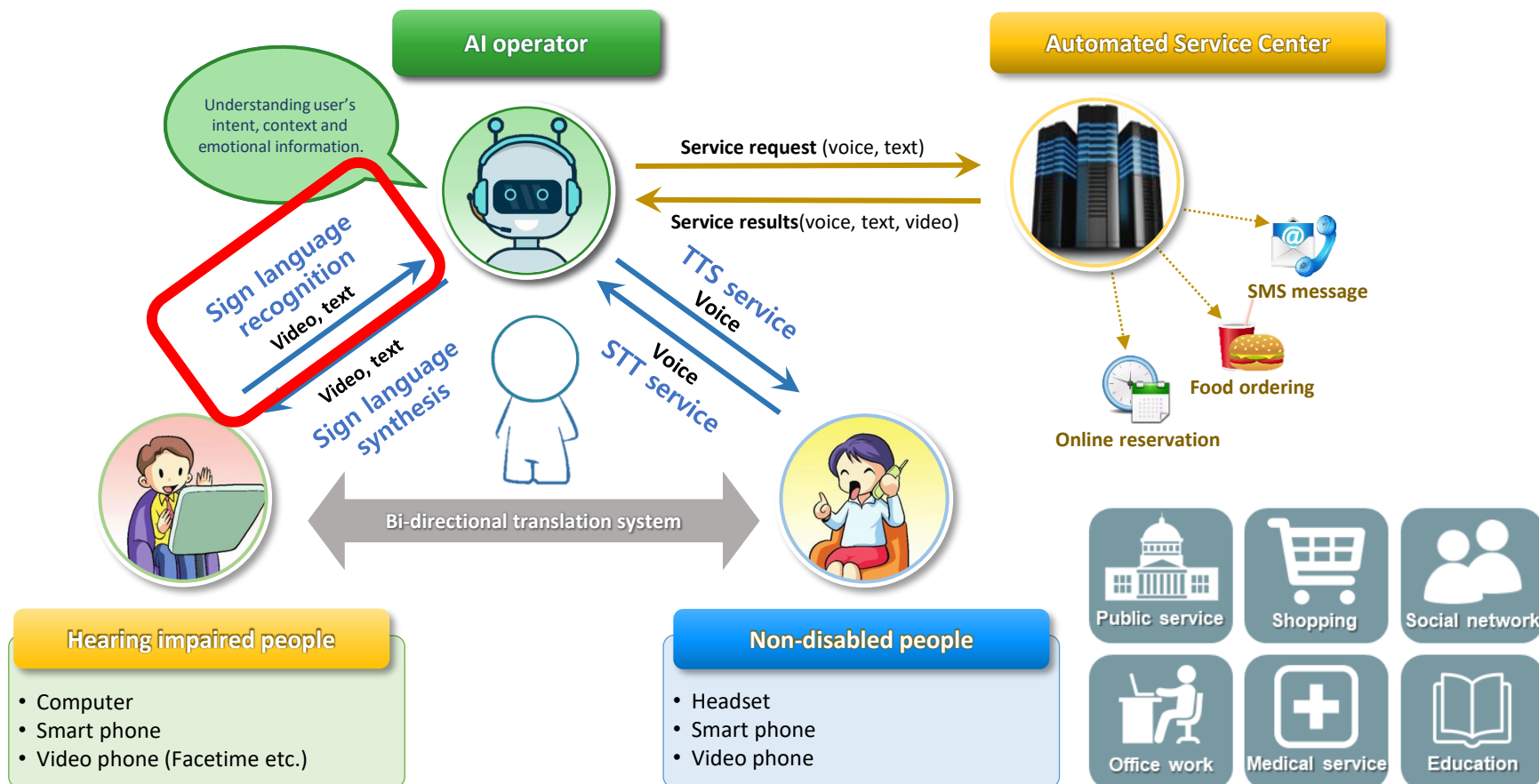


농인의 필담 이해도, 국립국어원



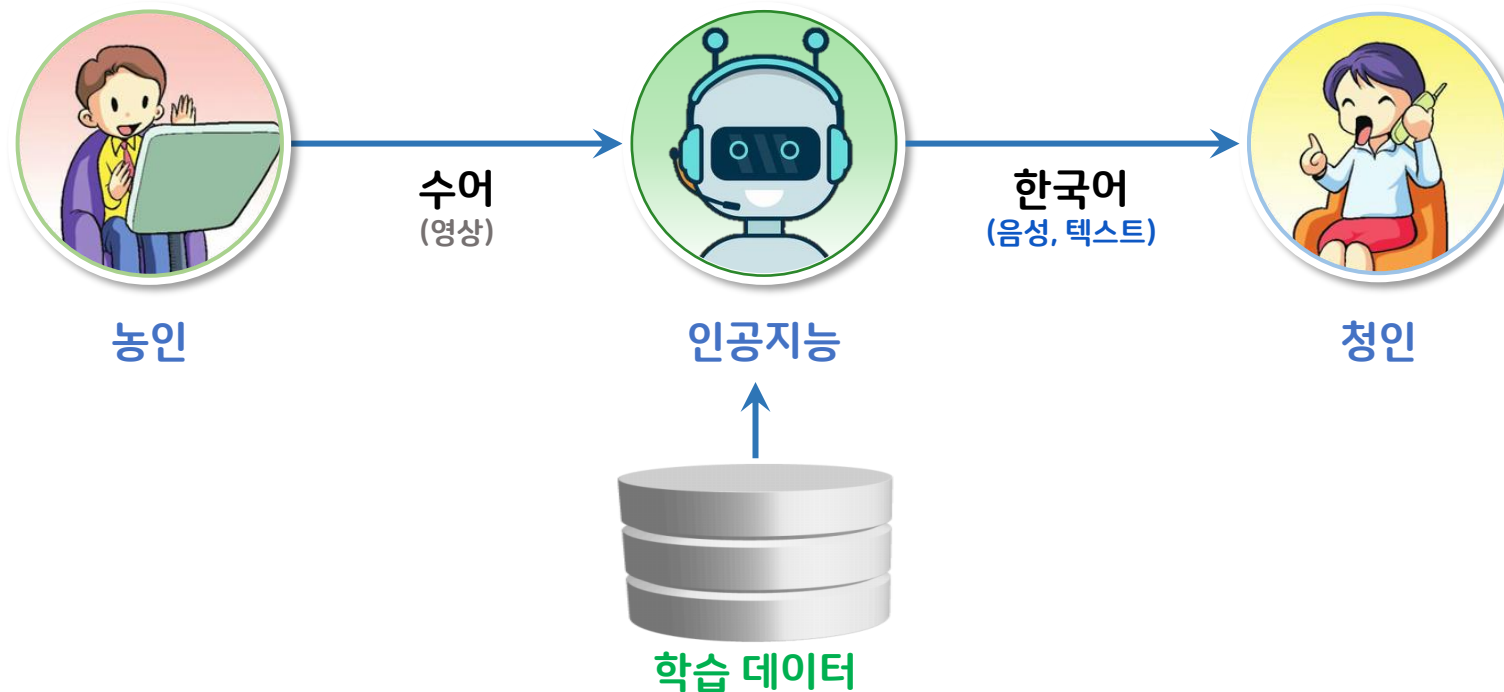
질병관리본부 정례 브리핑, 연합뉴스

• 청각장애인을 위한 양방향 통역 서비스 구조도



- 인공지능 기반 수어 인식 기술

- 인공지능이 농인의 **수어**를 이해하여 **한국어**로 번역하는 기술
- 영상을 입력으로 받아 음성 또는 텍스트로 변환
- 인공지능 모델 학습을 위한 **학습 데이터**의 구축이 필수적임



CHAPTER

## II

수어 인식을 위  
한 학습 데이터



## 1. 학습 데이터 취득

- 학습용 데이터에 반영되어야 하는 수어의 특징
  - 하나의 한국어 단어가 수어에서는 여러가지 방식으로 표현될 수 있음
  - 동일 수어 형태소임에도 농인마다 표현하는 것이 조금씩 다를 수 있음
    - 수동, 수형, 수위, 수향, 체동, 비수지 신호의 차이
  - 지역마다 다른 방식의 표현(방언)이 존재함



수어 영상 데이터, AI Hub

## 1. 학습 데이터 취득

- 학습용 수어 영상 데이터의 특징
  - 학습 데이터 취득 환경의 영향이 있음
    - 조명, 시점, 배경 등
  - 언어제공자의 외형에 따라 다른 영상이 됨
    - 옷차림, 액세서리, 생김새 등



수어 영상 데이터, AI Hub

## • 학습용 문장 셋 정제

### – 언어제공자 섭외

- 19세 이상 수어가능자
- 한국 수어를 모어로 의사소통이 가능한 자
- 언어제공자 정보 기록 (연령, 지역, 교육 기관 등)

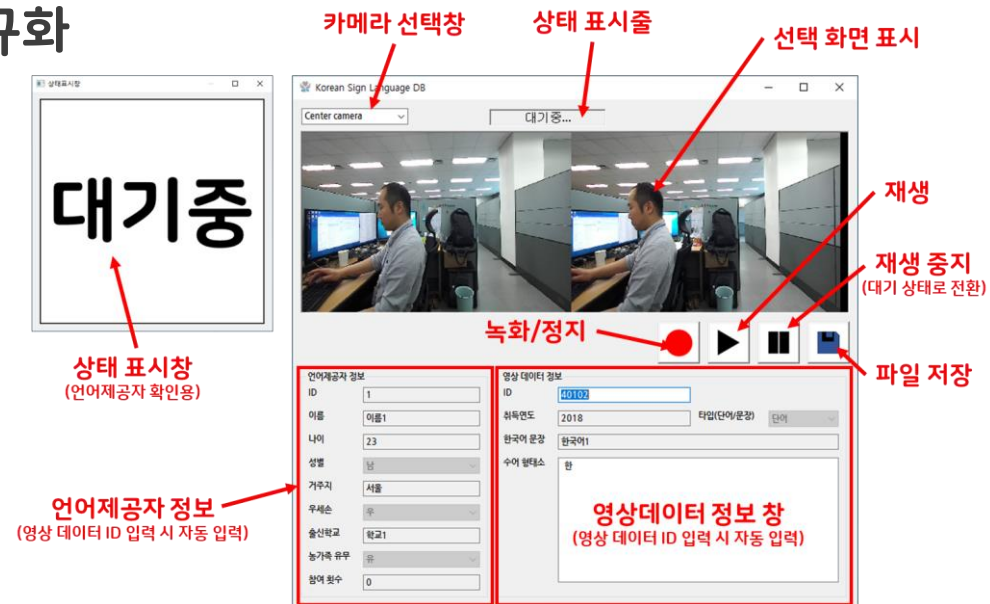
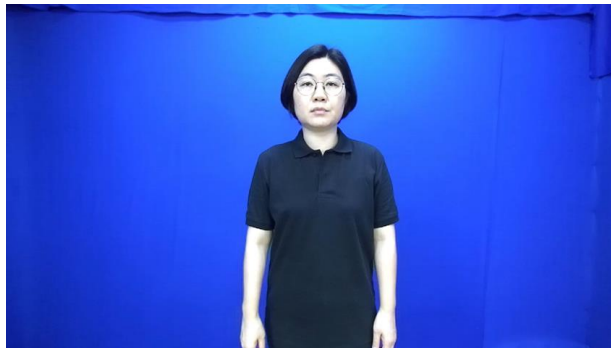
### – 한국어 문장 번안

- 다수의 수어 통역 전문가들의 자문을 통한 수지한국어 회피

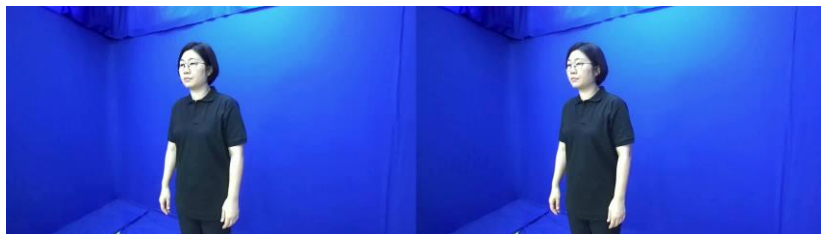
순번	구분		내용		순번	수어 표현									
55	편의시설	위치	제주 홍보관은 어디입니까?	C	1	제주	홍보	장소	어디						
56	편의시설	위치	국내선 입국장이 어디에 있습니까?	C	1	한국	안	착륙	어디						
57	편의시설	위치	국내선 출국장이 어디에 있습니까?	C	2	한국	안	이륙	어디						
58	편의시설	위치	국제선 출국장이 어디에 있습니까?	C	1	한국	안	다르다	나라	이륙	어디				
59	편의시설	위치	철체어 대여소가 어디에 있습니까?	C	1	수레	발리다	어디							
60				C	2	휠체어	발리다	어디							
61	편의시설	위치	관광안내소는 어디에 있습니까?	C	1	여행	정보	묻다	장소	어디					
62	편의시설	위치	한정식집이 어디에 있습니까?	C	1	한국	먹다	어디							
63	편의시설	위치	일식집이 어디에 있습니까?	C	1	일본	먹다	어디							
64	편의시설	위치	중식집이 어디에 있습니까?	C	1	중국	먹다	어디							
65	편의시설	위치	패스트푸드는 어디에 있습니까?	C	1	햄버거	장소	어디							
66	편의시설	위치	빵집이 어디에 있습니까?	C	1	빵	집	어디							

## • 학습 데이터 취득 시스템

- 다중 시점 스테레오 카메라 활용
  - ZED 2.0 \* 3시점(정면,좌,우)
  - 시점에 따른 영상의 다양성 반영
  - 3차원 키포인트 획득
- 조명, 배경, 취득 위치 등 정규화
- 언어제공자 복장 정규화
- 데이터 획득 SW 개발



- 학습 데이터 취득 결과



수어 문장 #1



수어 문장 #2

CHAPTER

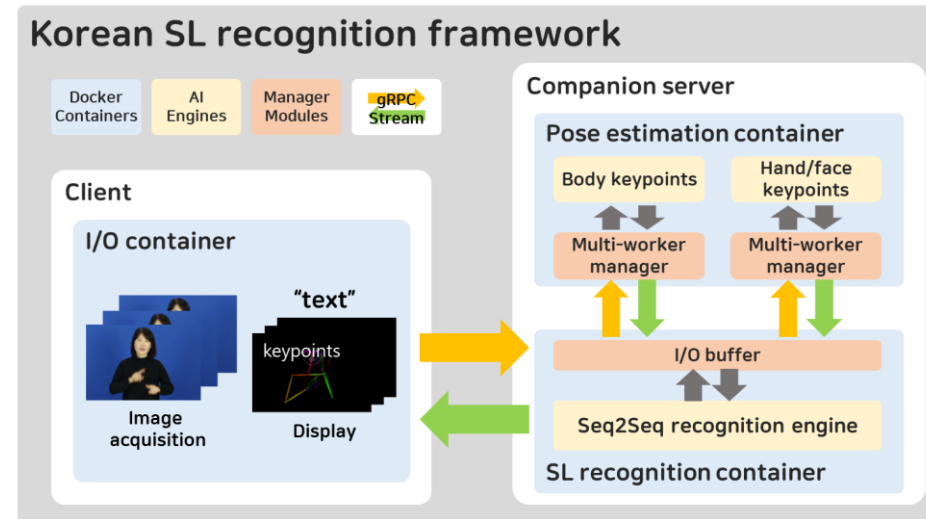
# III

## 수어 인식 알고리즘

## 1. 수어 인식 프레임워크

- 수어 인식 프레임워크 구조도

- **클라이언트**: 입력 영상 스트리밍 및 UI 제어
- 서버: 수어 인식 모듈 제어
- 연산 순서 조정을 통한 Human pose estimation 속도 최적화
  - 수어 인식을 위한 키프레임 선택 시 팔의 특징점만을 추정 (고속)
  - 키프레임 선택 후 손과 팔, 얼굴의 특징점 위치 추정 (저속)

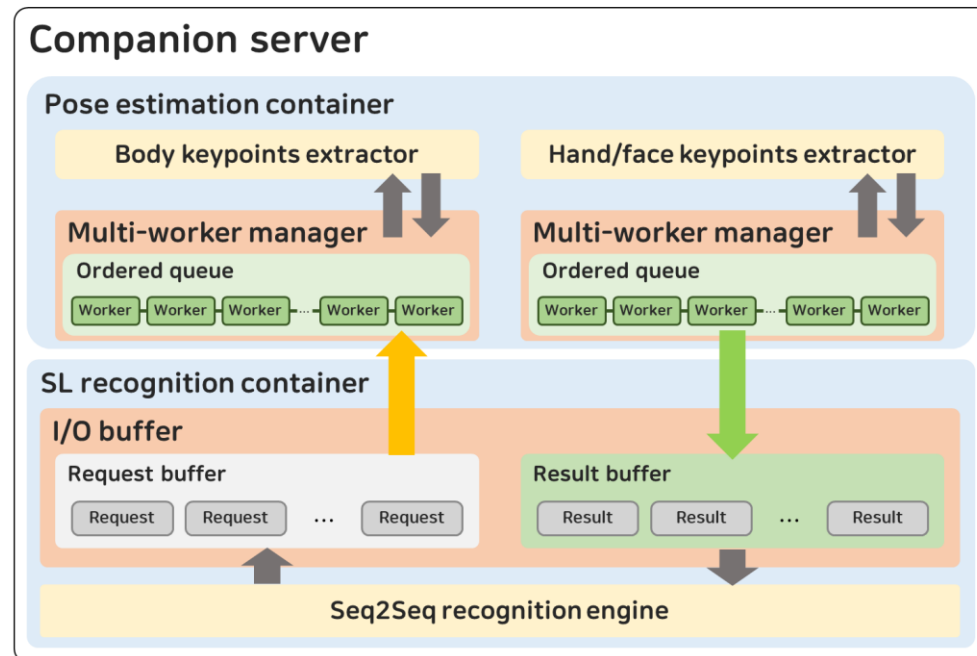


수어 인식 프레임워크 구조도

## 1. 수어 인식 프레임워크

- 수어 인식 프레임워크 구조도

- Ordered queue를 활용한 Human pose estimation 고속화
  - 다수의 이종 GPU들에 작업을 효율적으로 할당하기 위한 방식



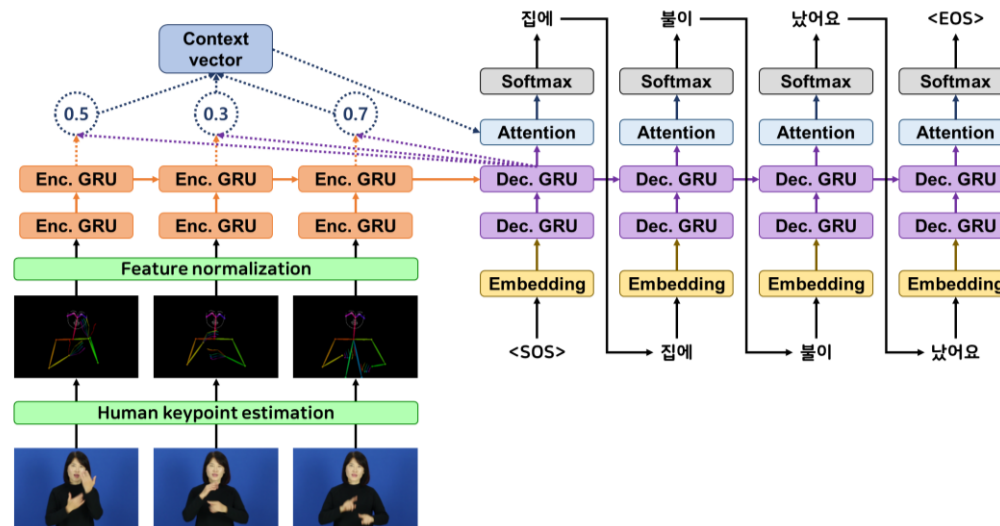
Ordered queue 기반 Multi-worker manager



## 1. 수어 인식 프레임워크

### • 수어 인식 모델

- Human keypoint 기반 수어 영상 인코딩 기술
  - 사람의 체형이나 얼굴 생김새, 주변 환경 변화에 강건
- Seq2Seq 모델 기반 수어 문장 인식 기술
  - 인코딩된 특징 벡터에 기반하여 한 단어씩 순차적으로 예측
  - Attention 매커니즘을 적용하여 번역 성능 향상



Seq2Seq 기반 수어 문장 인식 모델

## 1. 수어 인식 프레임워크

- 수어 인식 결과

### 1. 주변에 아무도 없는 경우



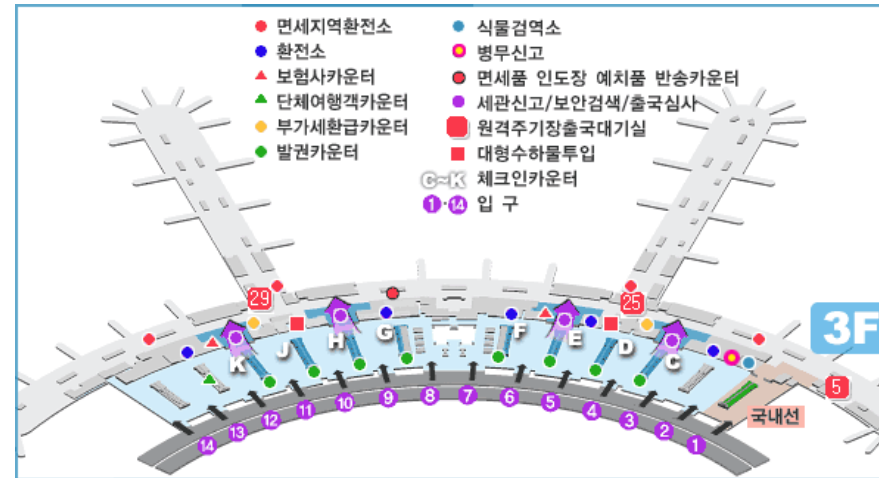
## 2. 수어 인식 기술을 활용한 배리어프리 서비스

- 청각장애인을 위한 공항 안내 서비스의 필요성

- 공항은 탑승 시설, 음식점, 편의점, 의무실, 라운지 등 다양한 시설들이 혼재되어 있는 복합 시설
- 원하는 장소로 이동하기 위해 유·무인 안내 서비스 활용이 필수적임
- 청인을 기준으로 안내 서비스가 설계되어 있어 이용이 어려움



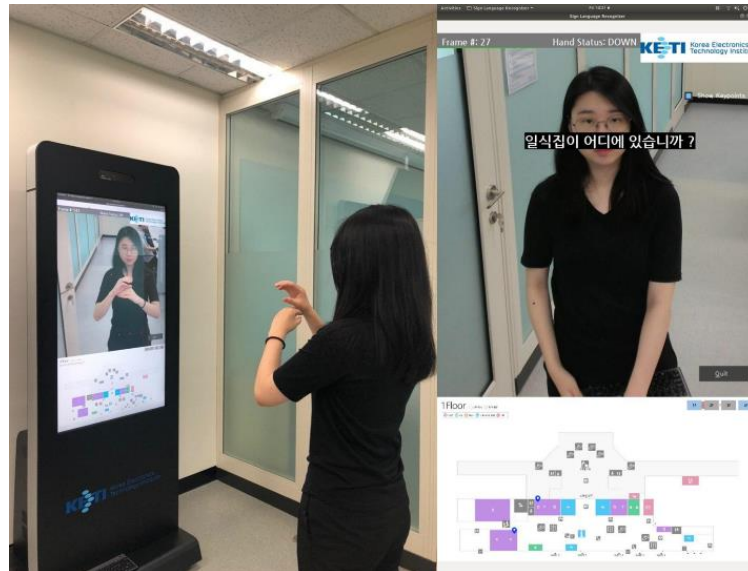
김포공항



김포공항 안내도

## 2. 수어 인식 기술을 활용한 배리어프리 서비스

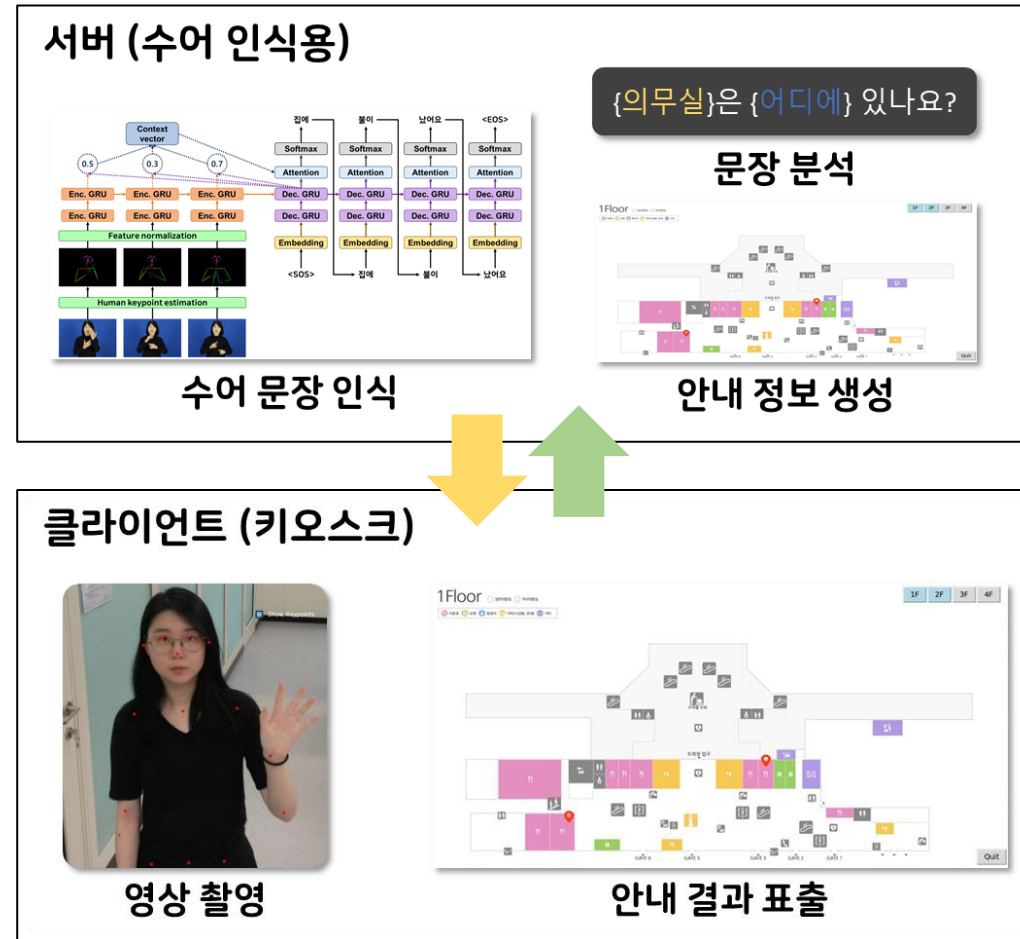
- 청각장애인을 위한 수어 인식 키오스크 개발
  - 영상 기반 마커리스 한국 수어 인식 기술을 적용
  - RGB 영상만을 활용하여 설치 장소에 제약이 적음
  - 사용자의 신체에 별도의 마커를 부착하지 않아 편의성이 좋음



청각장애인을 위한 공항 내 시설 안내 키오스크

## 2. 수어 인식 기술을 활용한 배리어프리 서비스

- 청각장애인을 위한 수어 인식 키오스크 구조도



## 2. 수어 인식 기술을 활용한 배리어프리 서비스

- 청각장애인을 위한 수어 인식 키오스크 구축 결과

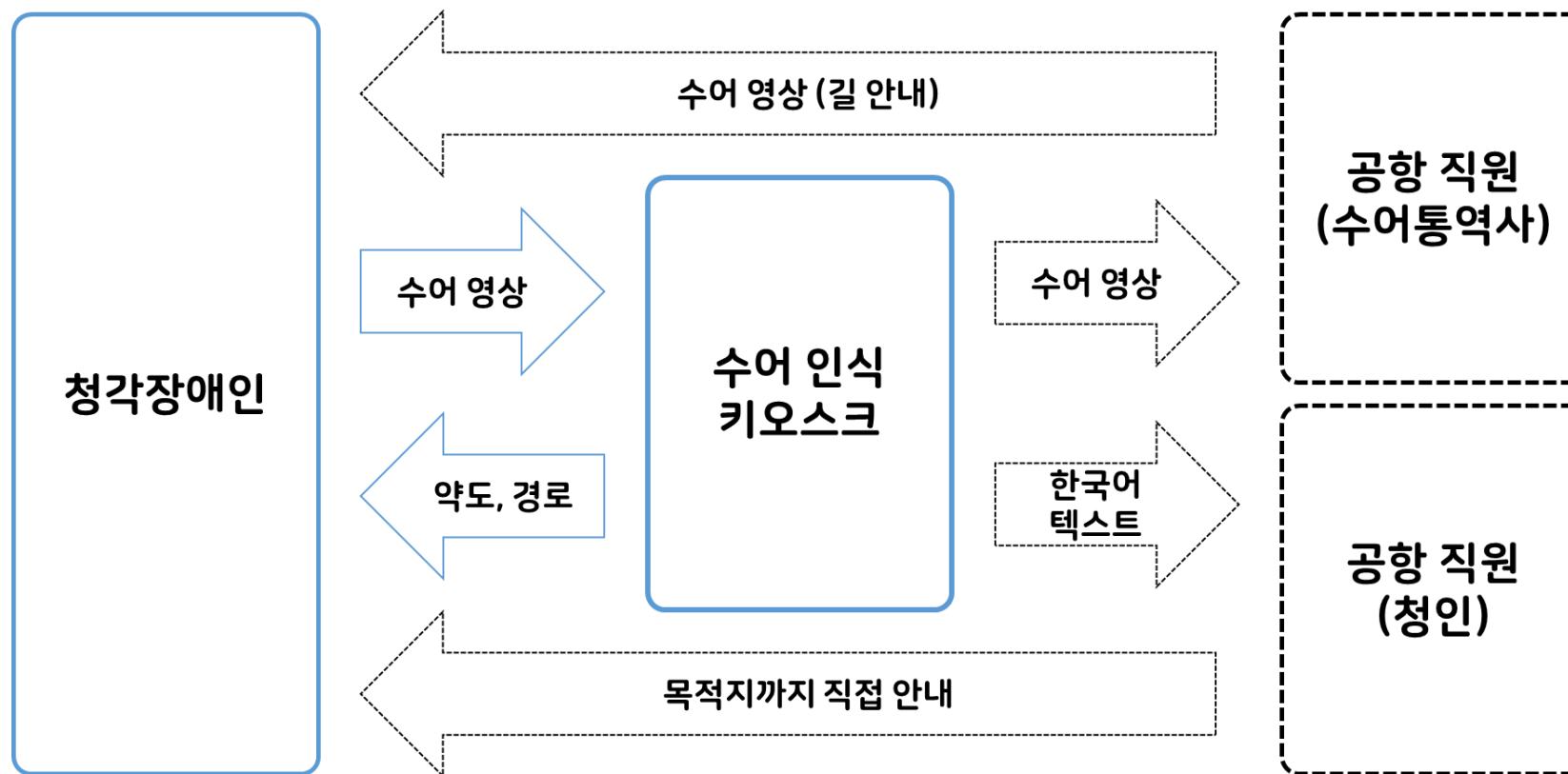
한국어 문장 표출

약도 표출



## 2. 수어 인식을 활용한 배리어프리 서비스

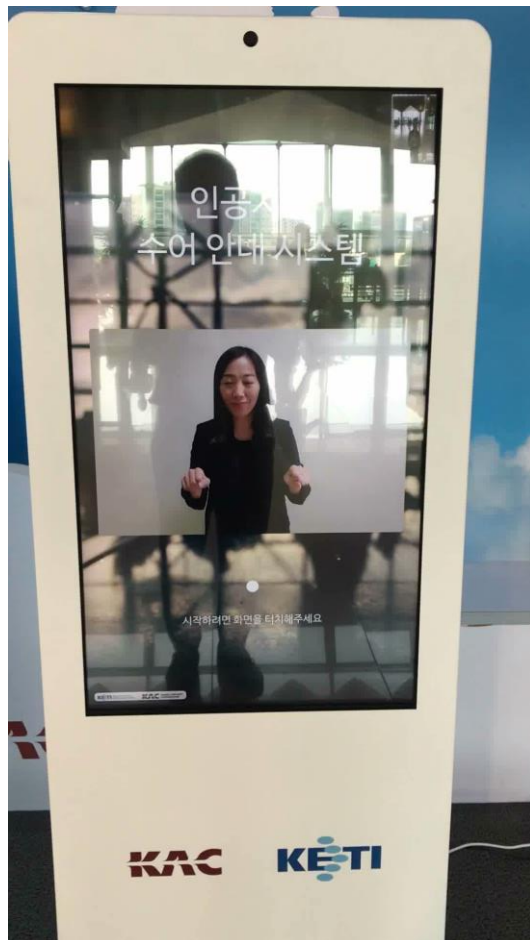
- 청각장애인을 위한 공항 안내 서비스 흐름도





## 2. 수어 인식을 활용한 배리어프리 서비스

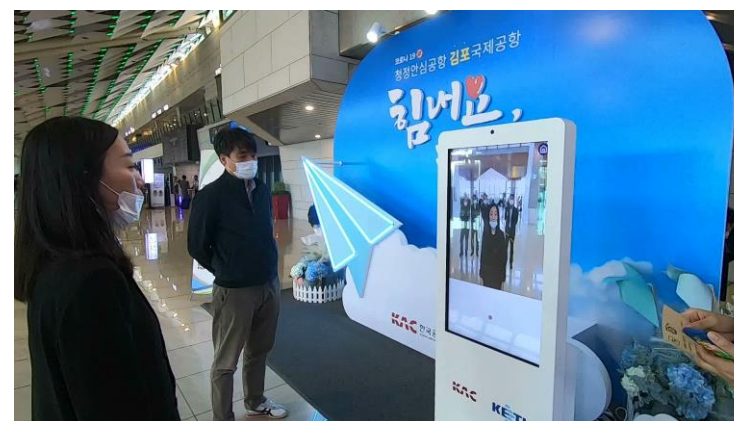
- 청각장애인을 위한 공항 안내 서비스 시연 결과



기본 안내 화면



시연 영상, "약국은 어디에 있나요?"



시연 영상, "택시승강장은 어디에 있나요?"



## 2. 수어 인식 기술을 활용한 배리어프리 서비스

- 청각장애인을 위한 공항 안내 서비스 시연 결과



세계 수어의 날(9/23) 수어 통역 컴패니언 시연

CHAPTER

IV

향후 연구 방향

## 1. 수어 인식 기술 개선 필요성

- 기존 수어 인식 기술의 한계점

- 데이터셋의 부족

- 위급 상황 105문장, 공항 안내 서비스용 145문장
    - 자연어 처리에 턱없이 부족함

- 유사 문장 인식 불가

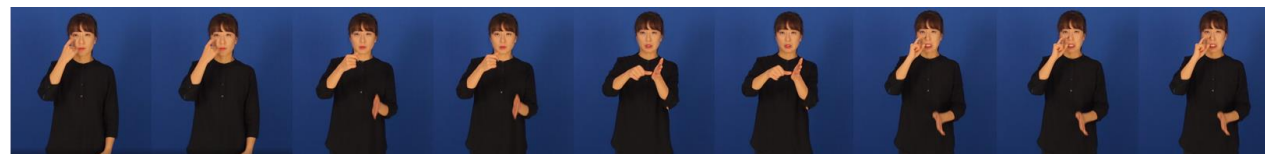
- 학습된 순서대로 발화한 수어 문장에 대해서만 인식 가능
    - 단어 하나만 교체하는 경우도 인식 불가



**수어 문장의 형태소 분석 필요!**

- 형태소 분석 기반 수어 인식 기술
  - 수어 인식 프레임워크

입력 수어 문장 영상



형태소 단위 수어 영상 분할 인식



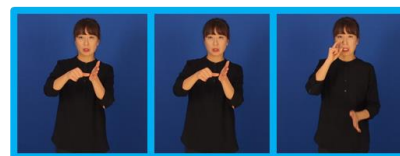
형태소 단위 수어 영상



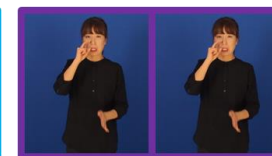
{GLOSS1}



{GLOSS2}



{GLOSS3}



{GLOSS4}

텍스트 기반 수어-한국어 번역

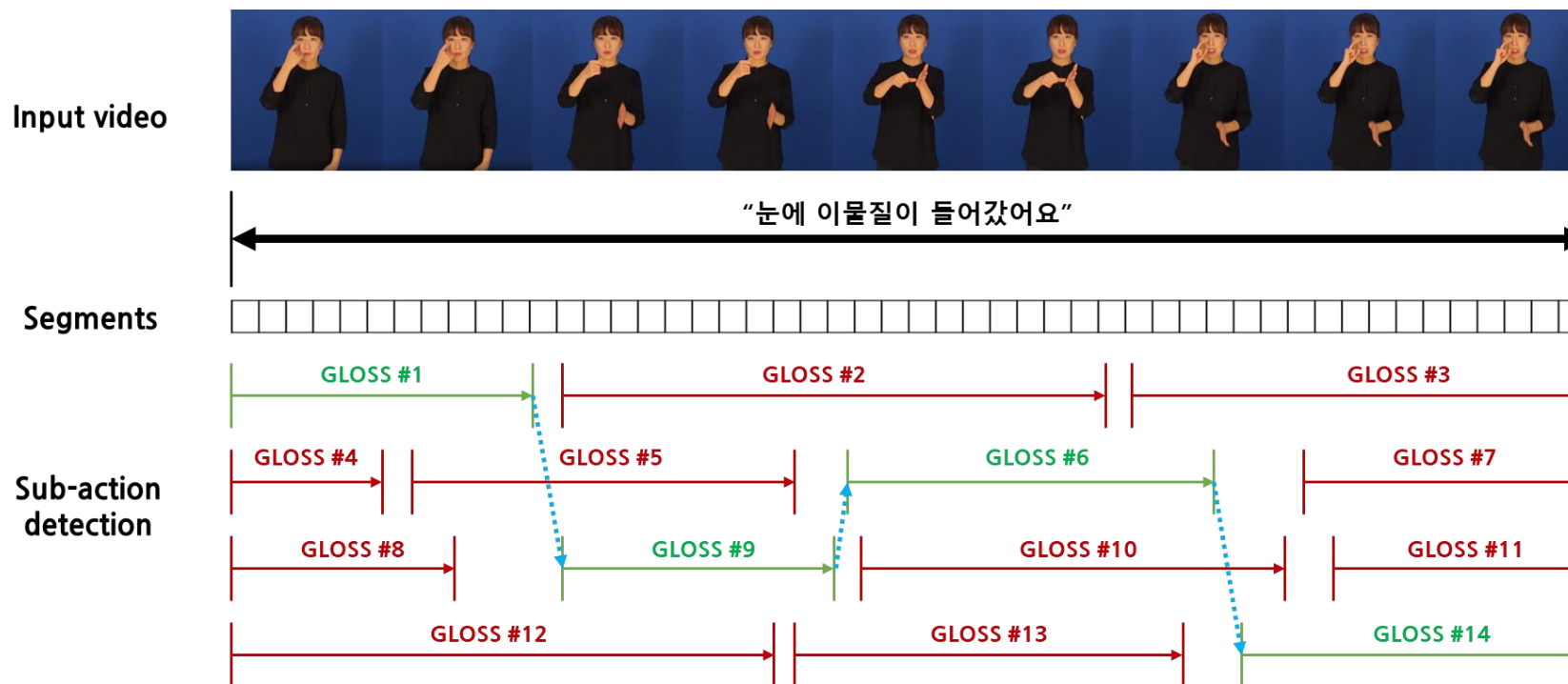


한국어 문장 생성

“눈에 이물질이 들어갔어요”

## 2. 수어 인식 기술 개선 방향

- 형태소 분석 기반 수어 인식 기술
  - 형태소 단위 수어 분할 인식 기술



## 2. 수어 인식 기술 개선 방향

- 형태소 분석 기반 수어 인식 기술
  - 텍스트 기반 수어-한국어 언어 모델 구축



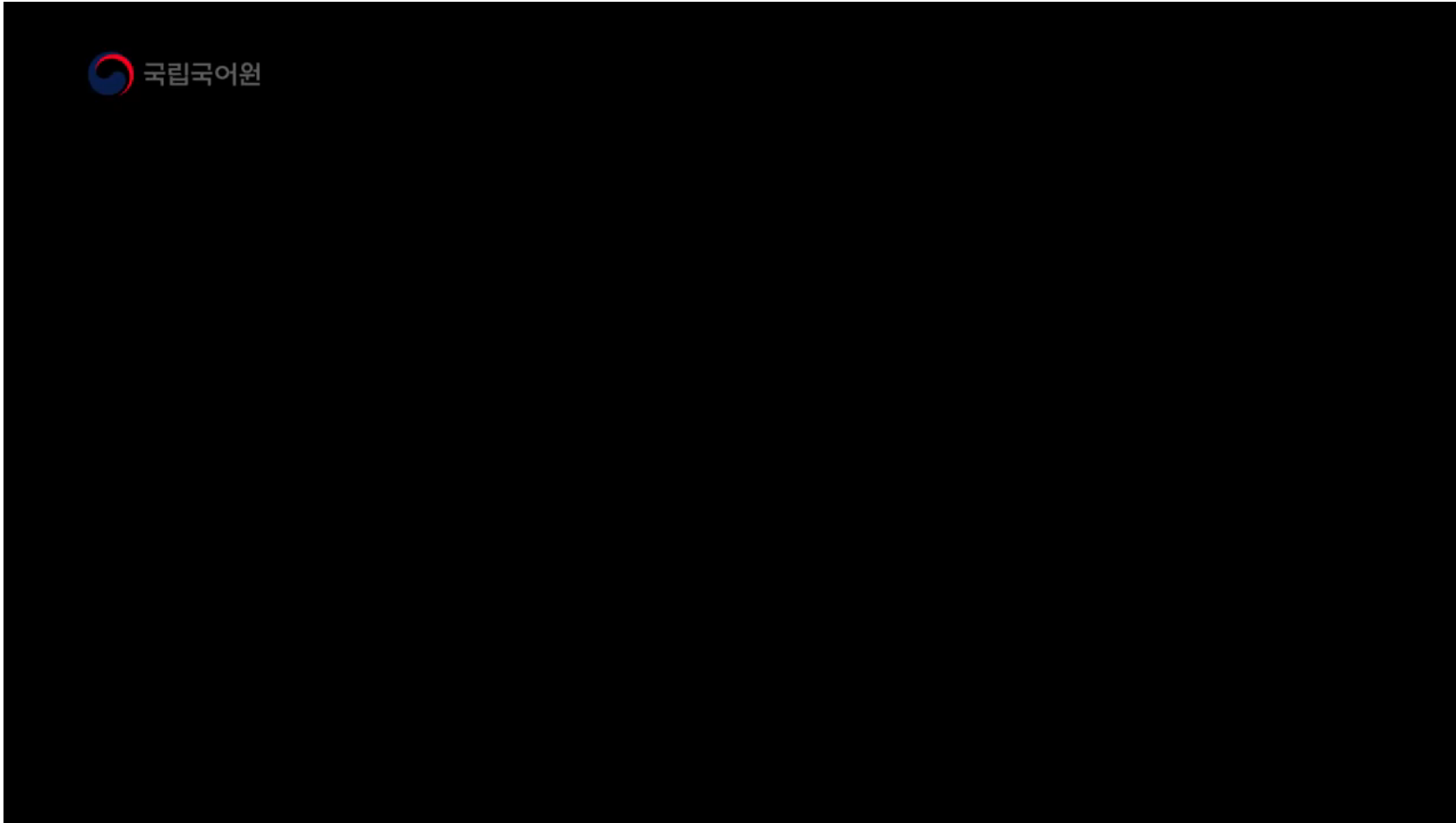
- 형태소 단위 수어 인식 기술 개발의 기대효과
  - 상대적으로 수집이 용이한 **텍스트 기반 데이터** 활용 가능
    - 데이터 구축 비용 및 수집 기간 절약
  - 어순, 어휘가 일부 다른 **유사 문장 이해** 가능
    - 인식 모델의 범용성 증대
    - 수어 인식 기술을 활용한 서비스 범위 증대
  - 수어-한국어 번역을 위한 **언어 모델** 구축
    - 수어-한국어, 한국어-수어 번역을 위한 방법론 제시

CHAPTER

V

결론





한국 수어 인식 개선 홍보 영상, 국립국어원

- [1] 국립국어원, "2017년 한국수어 사용 실태 조사," 2017.
- [2] 연합뉴스, "질병관리본부 정례 브리핑," 2020.03.21.
- [3] 한국공항공사 보도자료, "한국공항공사-전자부품연구원, 수어 통역시스템 공동개발 추진," 2019. 8. 30.
- [4] 국립국어원, "한국 수어 인식 개선 홍보 영상," 2019.