

# 2024 전기 졸업과제 착수 보고서



팀주제 : 대규모 언어 모델(LLM)의 Prompt Engineering을 통한 한국어-한국수어 번역  
서비스 개발

팀명 : 수어비전

팀 번호 : 4

팀원 : 허재성, 김민혁, 한지석

## 목차

1. 주제 선정 배경 및 과제 목표
2. 요구 사항 분석 및 설계 문서
  - 2.1. 요구 사항 분석
    - 2.1.1. 사용자 요구사항
    - 2.1.2. 시스템 요구사항
  - 2.2. 개발 환경
  - 2.3. 전체 구상도
3. 진행 방안
  - 3.1. 데이터 전처리
  - 3.2. AI 모델 구성
  - 3.3. 핵심 단어 추출
  - 3.4. Android 애플리케이션 제작
4. 기대 효과
5. 현실적 제약 사항
6. 개발 일정 및 역할 분담
7. 참고 문헌

# 대규모 언어 모델(LLM)의 Prompt Engineering을 통한 한국어-한국수어 번역 서비스 개발

팀명 : 수어비전

팀 번호 : 4

팀원 : 허재성, 김민혁, 한지석

## 1. 주제 선정 배경 및 과제 목표

국립국어원(원장 송철의)의 ‘한국수어 사용 실태 조사’의 결과를 따르면, 일상적인 의사소통에서 가장 많이 사용하는 언어는 ‘수어’라고 응답한 농인은 69.3%로 조사되었다. 농인의 제1언어가 ‘수어’임을 말해주는 결과이다. 하지만 가족과의 의사소통에서는 수어 사용 비율(42.7%)이 다소 낮게 나타났는데, 이는 가족 구성원 모두가 수어에 능숙한 것은 아니기 때문인 것으로 보인다. 농인들이 수어를 일상적으로 사용하지만, 비장애인이 수어를 이해하지 못하면 소통이 어렵다는 것이 현실이다.

“부산광역시 농아인 협회”의 자문 결과에 따르면, 비장애인과 소통 어려움으로 가장 불편한 부분은 낯선 지역에서의 길 찾기와 공공기관 이용입니다. 예를 들어, 버스를 이용할 때 버스 기사와의 의사 소통이 어려워 목적지에 대한 정보를 얻는 것조차 어렵고, 관공서에서의 업무 처리가 소통 문제로 인해 지연되는 경우가 흔히 발생합니다.

AI Hub에서 제공하는 ‘수어영상’ 데이터를 기반으로 ‘Sign Language Recognition’을 위해 ‘CNN’ 기반 모델인 ‘SlowFastSign’을 학습을 한 후, 청각장애인의 수어를 카메라로 인식하고 이를 ‘Prompt Engineering’을 통해 텍스트로 번역하는 어플을 제공해 비장애인과 소통을 수월하게 하여 비장애인과 장애인의 소통문제를 완화하는 데에 기여하고자 한다.

## 2.요구사항 분석 및 설계 문서

### 2.1 요구사항 분석

#### 2.1.1 사용자 요구사항

##### 1) 수어 인식

- 사용자가 손동작을 카메라를 통해 입력하면, 시스템은 이를 실시간으로 수어로 인식해야 한다.
- 다양한 수어 동작을 인식할 수 있어야 하며, 정확도와 신속성이 보장되어야 한다.

##### 2) 텍스트 번역

- 수어로 인식된 동작은 텍스트로 번역되어야 하며, 사용자가 쉽게 이해할 수 있는 언어로 번역되어야 한다.
- 번역된 텍스트는 실시간으로 제공되어야 하며, 빠른 응답 속도가 필요하다.

#### 2.1.2 시스템 요구사항

##### 2.1 실시간 처리

- 카메라 입력을 실시간으로 처리하고, 수어 인식 및 텍스트 번역이 실시간으로 이루어져야 한다.
- 응답 시간이 최소화되어야 하며, 시스템이 지연 없이 동작해야 한다.

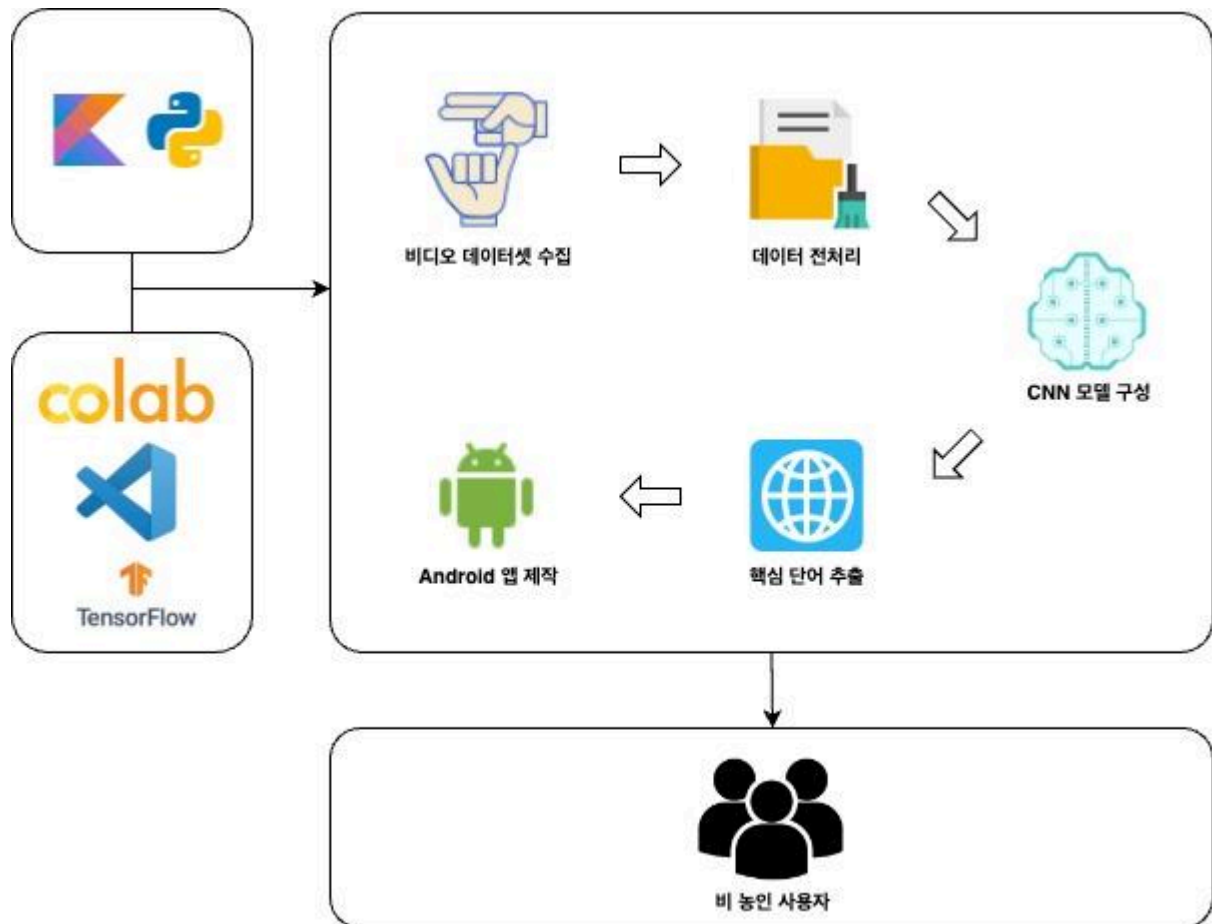
##### 2.2 정확성

- 수어 인식 및 텍스트 번역의 정확성이 보장되어야 한다. 오분류나 오역을 최소화해야 한다.

### 2.2. 개발 환경

- 개발 언어
  - Python
  - Kotlin
- 개발도구및실행환경
  - Visual Studio Code
  - Android Studio
  - TensorFlow, Colab

## 2.3 전체 구상도



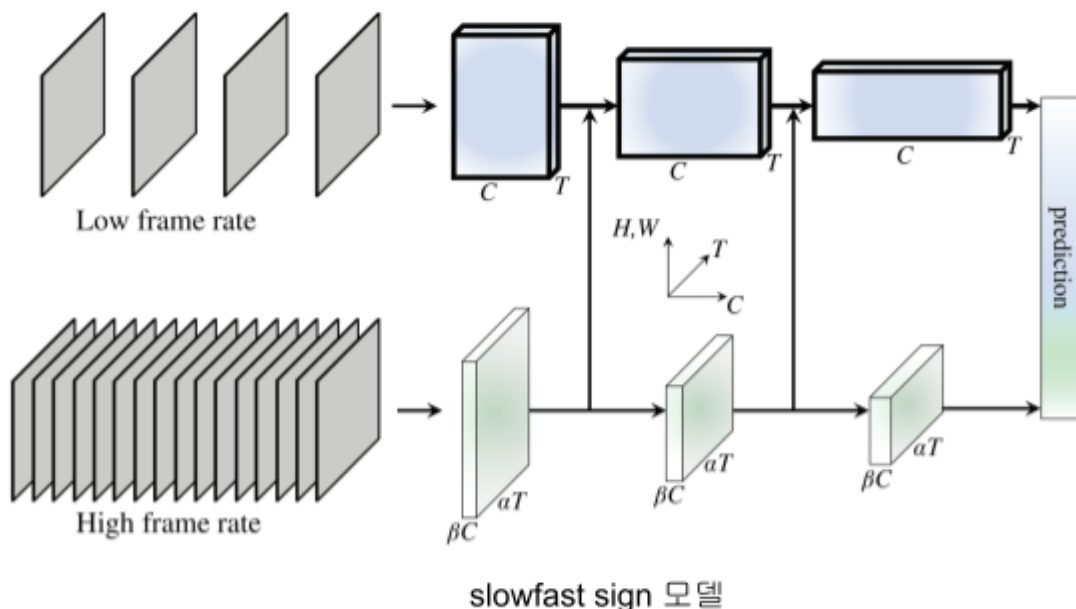
### 3.진행 방안

#### 3.1 데이터 전처리

- 이상치 및 결측치 처리 : **AI Hub**에서 제공하는 “수어 영상” 데이터의 이상치 및 결측치를 처리한다.
- 데이터 필터링 : 모델의 기능을 구체화하기 위해, 대량의 데이터 중 필요한 데이터를 추출하기 위해 데이터 필터링을 한다. 이를 위해 **OpenCV** 라이브러리에서 제공하는 다양한 함수들을 이용한다.

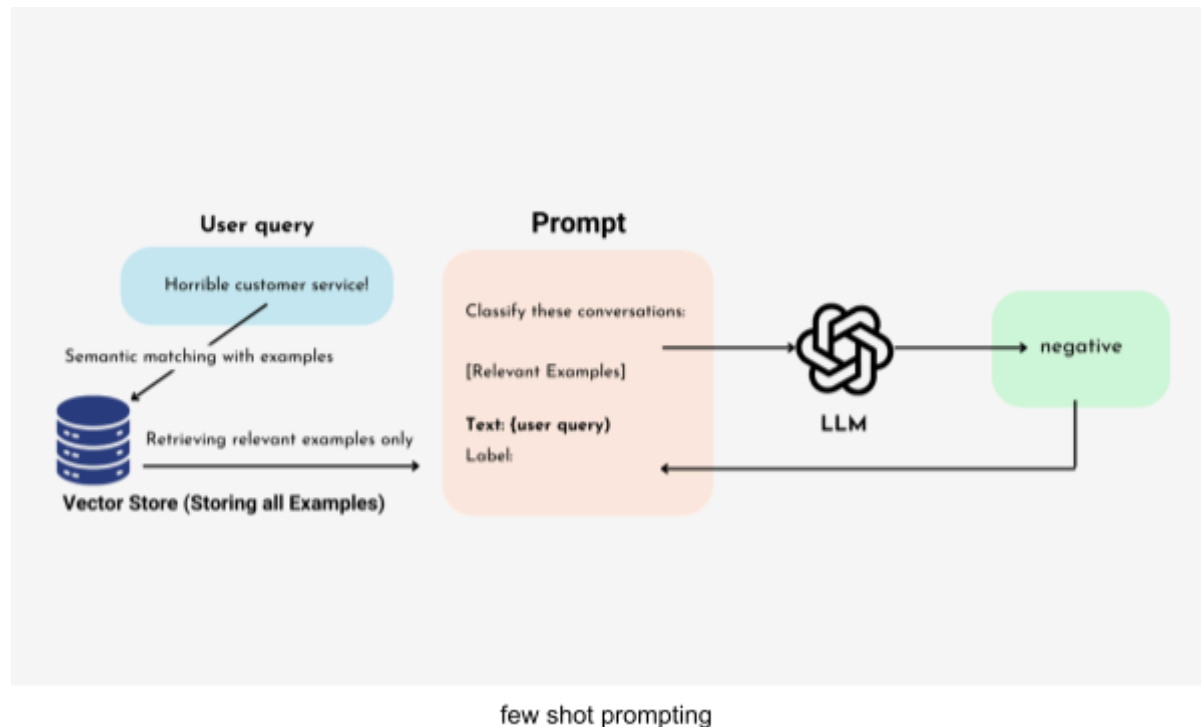
#### 3.2 AI 모델 구성

Sign Language Recognition(수어 인식)을 위해 딥러닝 기반 모델인 "SlowFastSign"을 사용한다. 이 모델은 영상 분석과 행동 인식에 쓰이는 AI 모델로, 느린 행동과 빠른 행동을 각각의 “Path”를 통해 동시에 분석하여 수어의 세밀한 움직임을 효과적으로 포착할 수 있다. 하지만 현재 존재하는 오픈소스 모델들 중에는 한글 수어를 인식하는 모델이 없기 때문에, **AI Hub**의 **KSL (Korean Sign Language)** 데이터셋을 이용하여 **SlowFastSign** 모델을 학습시킨다.



### 3.3 핵심 단어 추출

- **Prompt Engineering** : 학습된 모델에서 추출된 텍스트를 **Prompt Engineering**을 통해 사용해 핵심 단어를 추출한다. 이 때 텍스트 생성, 요약에 적합한 “GPT-3.5” 모델을 사용한다.
- **퓨샷 학습(Few-Shot Learning)** : 퓨샷 학습은 소량의 예제를 통해 모델을 특정 작업에 맞게 조정하는 방법이다. **GPT-3.5**는 퓨샷 학습을 통해 적은 양의 데이터로도 높은 성능을 발휘할 수 있다. 이를 통해 수어로부터 추출된 텍스트에서 핵심 단어를 추출하는 작업에 적용한다.



### 3.4 Android 애플리케이션 제작

메인 화면 - 녹화 화면 - 핵심 단어 추출 결과 화면으로 구성된 간단하고 깔끔한 **Android** 애플리케이션을 제작한다. **Android Studio**를 이용해서 개발하며, 핵심 라이브러리로 **Jetpack Compose**를 이용한다. 또한, 예상 사용자들의 스마트폰 버전에 대한 조사를 사전에 진행하여 다양한 기기에서 호환되도록 한다. 많은 기능과 화려한 **UI**를 제공하는 것보다 취약계층에게 접근성을 높이고 직관적인 **UI**를 제공하는 것에 초점을 맞춘다.

## 4. 기대효과

청각장애인들이 관공서에서 직접 수어로 의사소통을 할 수 있게 되어, 필요한 문서나 정보를 더 빠르고 정확하게 받을 수 있게 된다. 그리고 청각장애인들은 다른 사람의 도움이 없이도 독립적으로 필요한 업무를 처리할 수 있게 되어 서비스 접근성이 향상된다. 청각장애인들도 비장애인들과 동등하게 공공서비스를 이용할 수 있게 되어, 사회적 포용성을 증대시키고, 청각장애인들의 권리와 자존감을 높일 수 있게 된다. 수어 번역 앱을 통해 축적된 데이터를 분석하여, 청각장애인들이 자주 요청하는 문서나 서비스에 대한 인사이트를 얻을 수 있다. 이를 바탕으로 관공서 서비스의 품질을 지속적으로 개선할 수 있다.

## 5. 현실적 제약 사항

- 1) 한국 수어에 존재하는 모든 단어의 영상을 담은 수어 데이터셋이 없다. 수어는 양 손으로 표현하는 수지 신호와 표정, 시선, 입모양, 머리 및 몸의 움직임 등을 포함하는 비수지 신호로 구성되는데, 이 때 비수지 신호에서의 데이터가 매우 크기 때문이다. 또한 수어 영상에서 손의 일부분이 다른 부분을 가려서 해석이 모호해지는 자기 폐색 문제(**self-occlusion**)를 완화하기 위해 여러 방향에서 녹화한 영상이 필요하다는 점도 데이터 양 증가의 요인이다.
- 2) 설령 모든 단어의 영상을 담은 데이터 셋이 있다 하더라도, 데이터셋 내부에 존재하는 이상치, 결측치(데이터 포인트의 누락, 영상의 누락 등)를 전처리 과정에서 배제하는 일에 있어 많은 자원이 소요된다.



6. 추진 체계 및 일정

6.1 ) 개발 일정

	5월			6월					7월					8월					9월					10월		
주	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3
	데이터셋 선정 및 개발 환경 구축																									
				데이터 전처리																						
							모델 구성																			
									모델 최적화																	
														중간 보고서 작성												
																데이터 보충 가능 여부 확인										
																	오류 확인 및 최종 테스트									
																							최종 보고서 작성 및 발표준비			

## 6.2 ) 역할 분담

이름	역할
김민혁	1.데이터 전처리 2.AI 모델 정확도 개선 3.앱 개발 및 데이터베이스 설계
한지석	1. 데이터 전처리 2.sign language recognition을 이용한 학습 모델 개발 3.학습 및 검증
허재성	1. 데이터 전처리 및 수집 2.prompt engineering 학습 3.결과 분석 및 모델 최적화
공통 역할	1. 필요한 지식 학습 2. 보고서 작성

## 7. 참고 문헌

- 1) 국어생활연구원, “한국수어 사용 실태 조사”,  
[http://malsaem.kr/bbs/board.php?bo\\_table=news4&wr\\_id=147](http://malsaem.kr/bbs/board.php?bo_table=news4&wr_id=147)
- 2) Prompt Engineering Guide, “Few-Shot Prompting”,  
<https://www.promptingguide.ai/>
- 3) "mmdetection2: SlowFast Detection Configuration." GitHub, OpenMMLab,  
<https://github.com/open-mmlab/mmdetection2/blob/main/configs/detection/slowfast/README.md>. Accessed 2024-05-30.
- 4) data science dojo, “Dynamic few-shot prompting to create captivating digital content”, <https://datasciencedojo.com/blog/dynamic-few-shot-prompting/>