

수어비전

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

Contents

01	_____	배경
02	_____	전체 구조
03	_____	모델
04	_____	Prompt Engineering
05	_____	어플리케이션
06	_____	결론

01 배경

연구 배경

2016년 한국수화언어법 시행 이후, 농인의 한국수어 사용 실태를 조사함으로써 수어 환경 조성을 위한 기초 데이터를 제공했다.

문제 제기

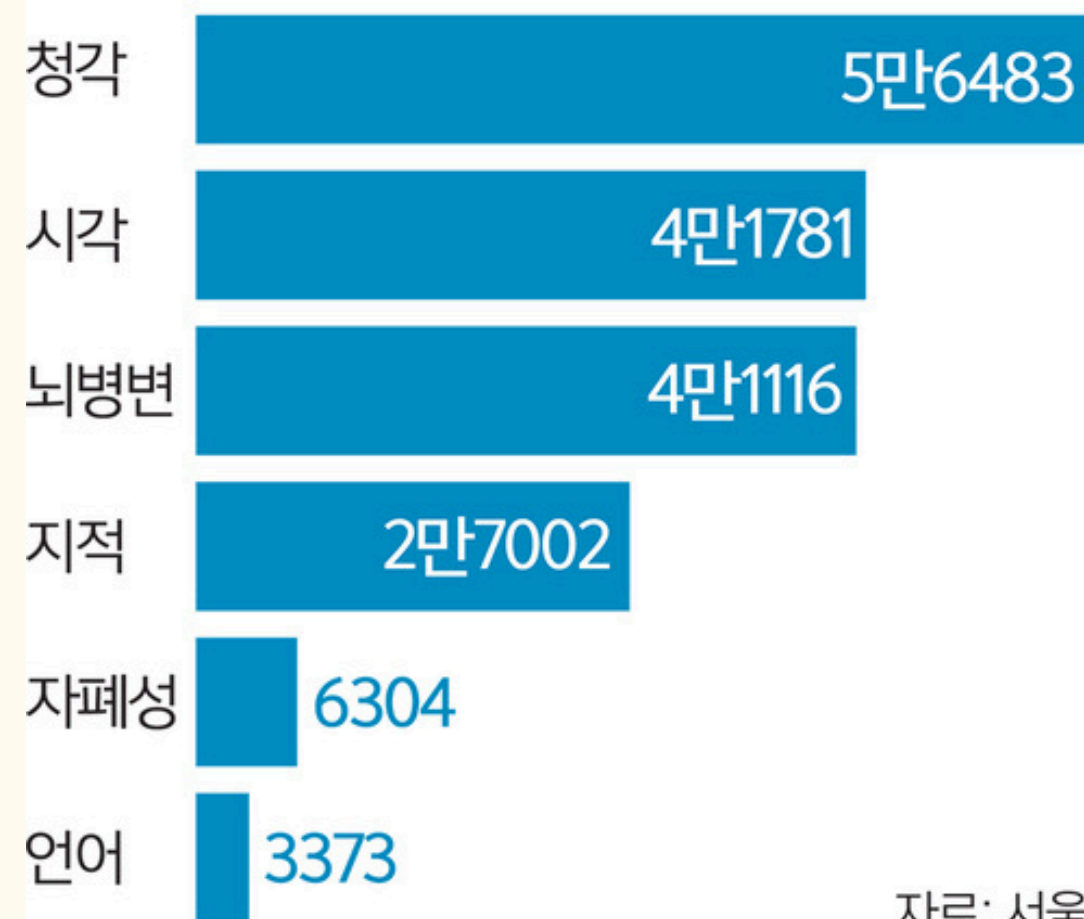
청각장애인, 특히 고령층이 한글 해독 능력이 부족해 수어에 의존하고, 비장애인과 의사소통에서 필담, 몸짓 등 불편함을 겪으며, 수어 통역 서비스도 제한적이다.

연구 목표

AI Hub의 수어 영상 데이터를 활용하여 CNN 기반 SlowFastSign 모델을 통해 실시간 수어 인식 애플리케이션을 개발하여 청각장애인과 비장애인 간 원활한 소통을 지원한다.

의사소통에 어려움을 겪는 유형의 장애인 (단위: 명)

*2020년 7월 서울시 등록 기준

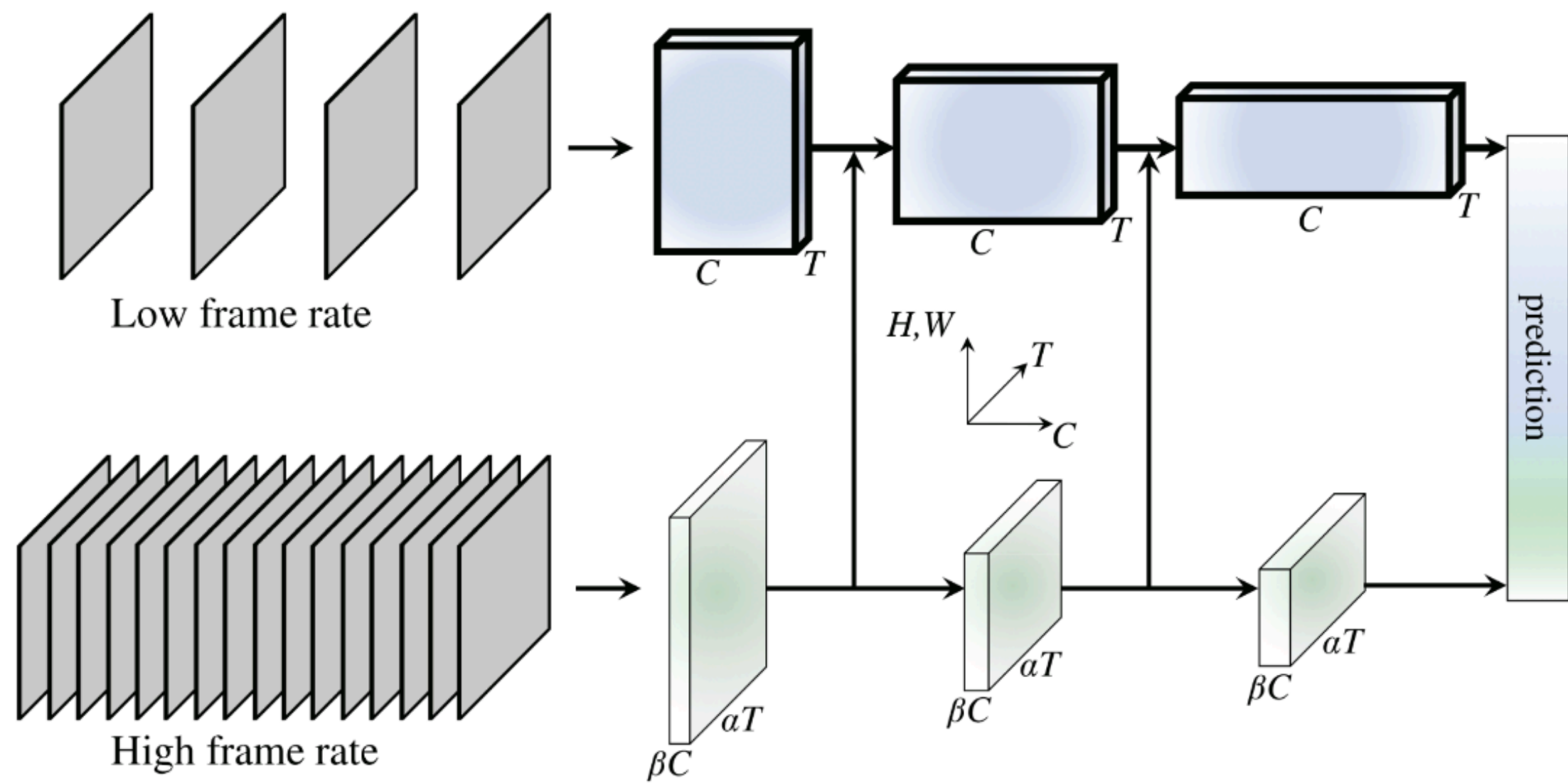


자료: 서울시

02 전체 구조



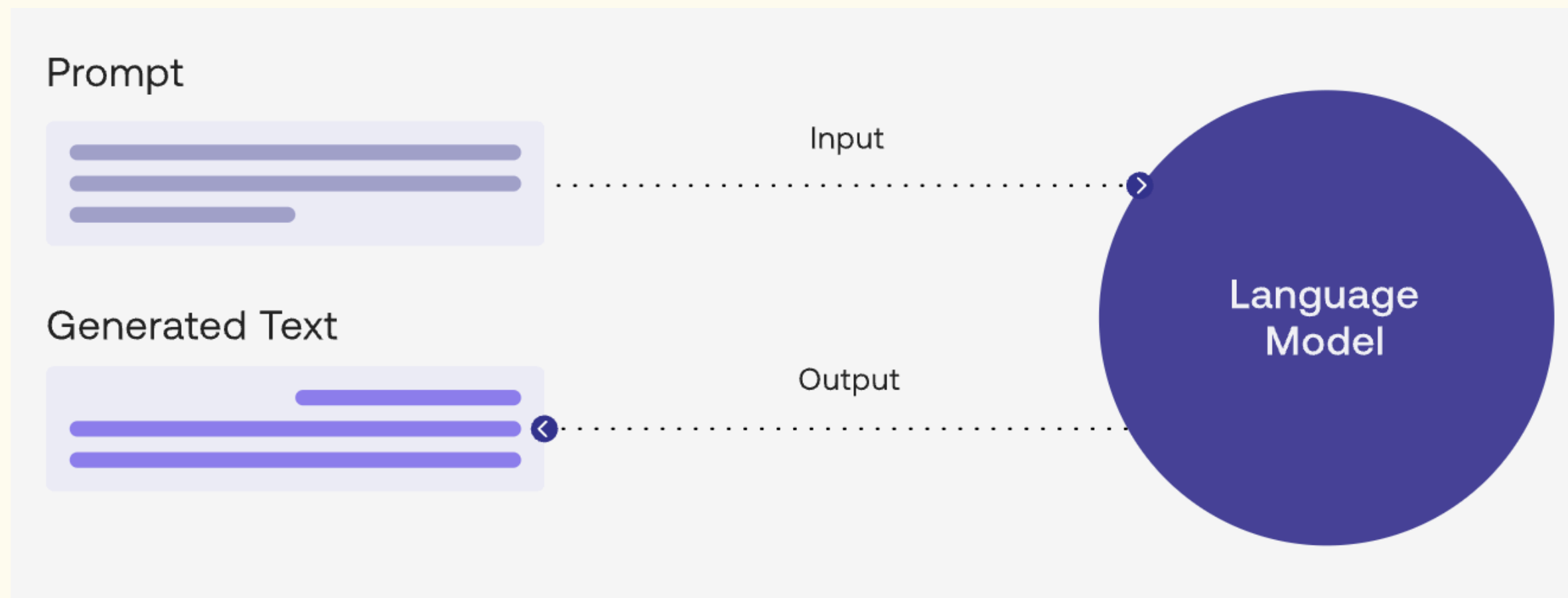
03 모델



slowfast sign 모델

04 Prompt Engineering

- Prompt Engineering은 특정한 목표를 달성하기 위해 대형 언어 모델(LLM)에 입력하는 프롬프트(Prompt)를 전략적으로 설계하는 기술이다.
- 효과적으로 프롬프트 엔지니어링을 활용하면, 대형 언어 모델(LLM)이 더 정확한 결과를 출력하도록 유도할 수 있다.
- 모델을 효과적으로 사용하기 위해서는 일반적으로 지침과 몇 가지 예시를 프롬프트에 포함시키며, 이렇게 작업을 수행하는 방법을 예시로 보여주는 방식을 “퓨샷 학습(few-shot learning)”이라고 한다.



04 Prompt Engineering

- 본 프로젝트에서는 프롬프트 엔지니어링을 효과적으로 수행하기 위한 LLM으로써 ChatGPT로 널리 알려진 openai 사의 API를 이용하였다.
- 해당 API를 통해 제공되는 모델을 더욱 효과적으로 사용하기 위해 Fine-Tuning을 진행하였다.



Fine-tuning

Fine-tuning lets you get more out of the models available through the API by providing:

- Higher quality results than prompting
- Ability to train on more examples than can fit in a prompt
- Token savings due to shorter prompts
- Lower latency requests

04 Prompt Engineering

- AI모델을 통해 생성될 수 있는 예상 글로스 집합과, 각 글로스 집합이 바람직하게 해석된 결과 2000여개를 openai API에 학습시켰다.

```
1 {"prompt": "유턴 맞다", "completion": "이 정차지에서 돌아가야 합니다."}
2 {"prompt": "공항 버스 보다 전 곳 차내리다", "completion": "공항 전에 내려주세요."}
3 {"prompt": "버스 곳 내리다 맞다", "completion": "이 정류장에서 내려주세요."}
4 {"prompt": "아이 실종 잃어버리다", "completion": "아이가 실종되었습니다."}
5 {"prompt": "여기 내리다 맞다", "completion": "여기서 내려주세요."}
6 {"prompt": "우산 실종 잃어버리다", "completion": "우산을 잃어버렸어요."}
7 {"prompt": "차밀리다 시간 괜찮다", "completion": "출발 시간에 여유가 있습니다."}
8 {"prompt": "서울대학교 길 찾다 방법", "completion": "서울대학교로 가는 방법을 알려주세요."}
9 {"prompt": "언덕 저기 곳 다음 도착 차내리다", "completion": "저기 언덕 아래 정차해주세요."}
10 {"prompt": "서울농아인협회 내리다", "completion": "서울농아인협회에서 내려주세요."}
11 {"prompt": "카드 고장 잘 안되다", "completion": "카드가 작동하지 않아요."}
12 {"prompt": "핸드폰 없다 실종", "completion": "핸드폰을 잃어버렸어요."}
13 {"prompt": "군청 곳 다음 내리다", "completion": "군청 다음에 내려주세요."}
14 {"prompt": "버스 곳 다음 내리다", "completion": "다음 정류장에서 내려주세요."}
15 {"prompt": "핸드폰 번호 필요", "completion": "전화번호를 알려주세요."}
16 {"prompt": "영등포 곳 찾다 곳 방법", "completion": "영등포로 가는 방법을 알려주세요."}
17 {"prompt": "아이 잃어버리다 실종", "completion": "아이를 잃어버렸어요."}
18 {"prompt": "화장실 무엇", "completion": "화장실이 어디예요?"}
19 {"prompt": "지갑 실종", "completion": "지갑을 잃어버렸어요."}
20 {"prompt": "지름길 가다 내리다", "completion": "지름길로 가서 내려주세요."}
```


04 Prompt Engineering

- 파인튜닝 전 후로 해석 결과의 정확도가 많이 향상되었다.

단어가 포함된 텍스트 파일의 이름을 입력하세요: /home/viplab/Desktop/sign_language_project/app/AihubModel/output.txt

생성된 택시 상황 문장: 지하철이 막혀서 다음 정류장까지 가야 해요. 혹시 어떻게 가야 하는지 알려주실 수 있나요?



글로스 : 여 기 내 리 다 맞 다

번역 : 여 기 내 리 면 될 까 요 ?

글로스 : 공 항 버 스 보 다 전 곳 차 내 리 다

번역 : 공 항 버 스 보 다 앞 에서 내 려 주 세 요 .

글로스 : 버 스 곳 내 리 다 맞 다

번역 : 버 스 정 류 장 에서 내 리 면 돼 요 .

05 결론

- 수어 : 수어는 손과 신체적 신호를 이용하는 시각적 언어이며, 비수지신호가 중요한 역할을 한다. 한국 수어 데이터셋은 다양한 각도의 영상을 제공하여 의사소통 문제 해결을 돕는다.
- 모델 학습: SlowFast Sign 모델을 파인튜닝하여 성능을 향상시켰으며, 환경설정 및 데이터 전처리에서 많은 시행착오를 겪었다. 그럼에도 불구하고 적절한 하이퍼파라미터를 조정하여 모델 학습을 성공적으로 완료할 수 있었다.
- 자원 한계 : GPU 성능을 높이고 메모리 용량을 확장하여 배치 크기 및 num_workers 값을 조정, 학습 속도를 개선하고 모델의 일반화 성능을 향상시킬 필요가 있다.

수어비전

감사합니다.
