

2020년 인공지능 학습용 데이터 교육영상 “수어 영상” 소개 및 활용

2021.03

(주)테스트웍스
팀장 이상욱





INDEX

01. 인공지능 학습용 데이터 소개
02. 저작도구 활용법
03. 데이터 학습 목적 (RCV, KETI)
04. 학습된 모델 결과 확인 방법 및 시연
05. 서비스 개발 시 학습된 모델 활용 방안(예시)

01 인공지능 학습용 데이터 소개

「인공지능 학습용 데이터 구축 사업」 소개

본 사업은 인공지능이 수어를 인식할 수 있도록
인공지능 학습용 수어 영상을 수집하고 가공하는 사업입니다.

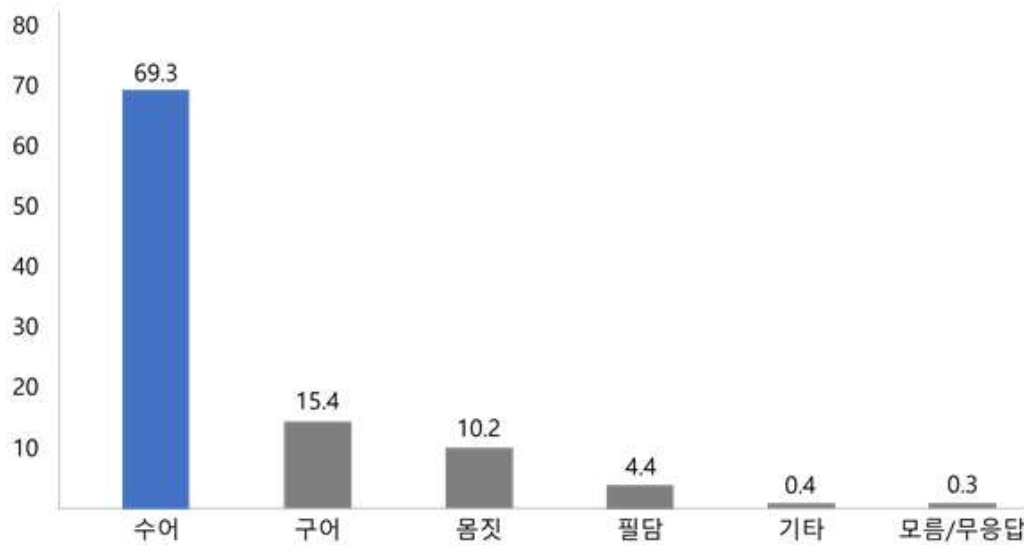
학습용 데이터의 양이 많을수록 인공지능이 다양하고 정확하게 수어를 학습할 수 있습니다.

이는 수어를 배울 때 여러 농인들과의 대화 경험이 수어실력에 영향을 미치는 것과 같습니다.

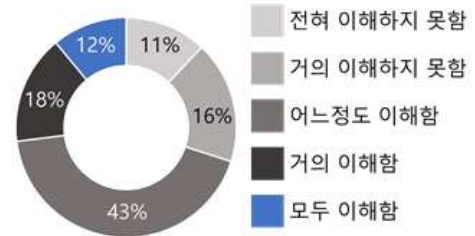
즉, 구축된 **학습용 수어 데이터**를 활용하여 **고도화된 인공지능 기술**을 통해,
각종 정보에 대한 **접근성을 확보**하며 다양한 **서비스 제공 방법**에 대한 고민이 필요합니다.

01 인공지능 학습용 데이터 소개

수어 영상 데이터 필요성



필담(문자언어) 이해도



TV 자막의 이해도



국내 청각장애인(이하 "농인") 수는 약 37만 명
 '한국 수어'는 농인 10명 중 약 7명이 사용하여 의사소통을 수행
 필담의 경우, 약 12% 농인만이 모든 의미를 완벽히 이해
 TV 자막의 경우 약 9% 농인만이 의미를 완전히 이해함

특히, **농인의 문맹률은 약 30%**에 달하는 것으로 추정되어 **한국어 문자기반의 소통이 어려움**
 → 일상생활, 경제활동, 긴급상황 등 거의 모든 분야에 걸쳐 다양한 고충을 겪음

01 인공지능 학습용 데이터 소개

수어 영상 데이터 구축 대상 - 길 찾기 / 교통 / 주소

배리어프리 키오스크



시설물 안내센터



택시 및 교통



영상 지도 서비스



수어 영상
데이터셋

- 535,000 개 수어 영상 클립
- 2개 수어 인식 모델

도메인 선정 기준

- 반영구적 데이터셋
- 재해 / 재난 긴급상황 시 위치 제공
- 시설안내/택시 등 서비스 연계
- 기구축(KETI) 긴급 및 재난 상황 관련 수어 데이터셋 연계

<길찾기/교통/주소>

실현가능성



<의료/법률/관공서>

실현가능성



<배달 서비스>

실현가능성



01 인공지능 학습용 데이터 소개

수어 통사구조(형태소와 비수지)를 고려한 수어영상 데이터 설계



특징점 추출
데이터

+

형태소 및 비수지
가공
데이터

데이터 셋 구성 요소

1. 수어 단어/문장 동영상
2. 한국어 단어/문장 병렬 코퍼스
3. 수어 형태소 시퀀스
4. 비수지 시퀀스

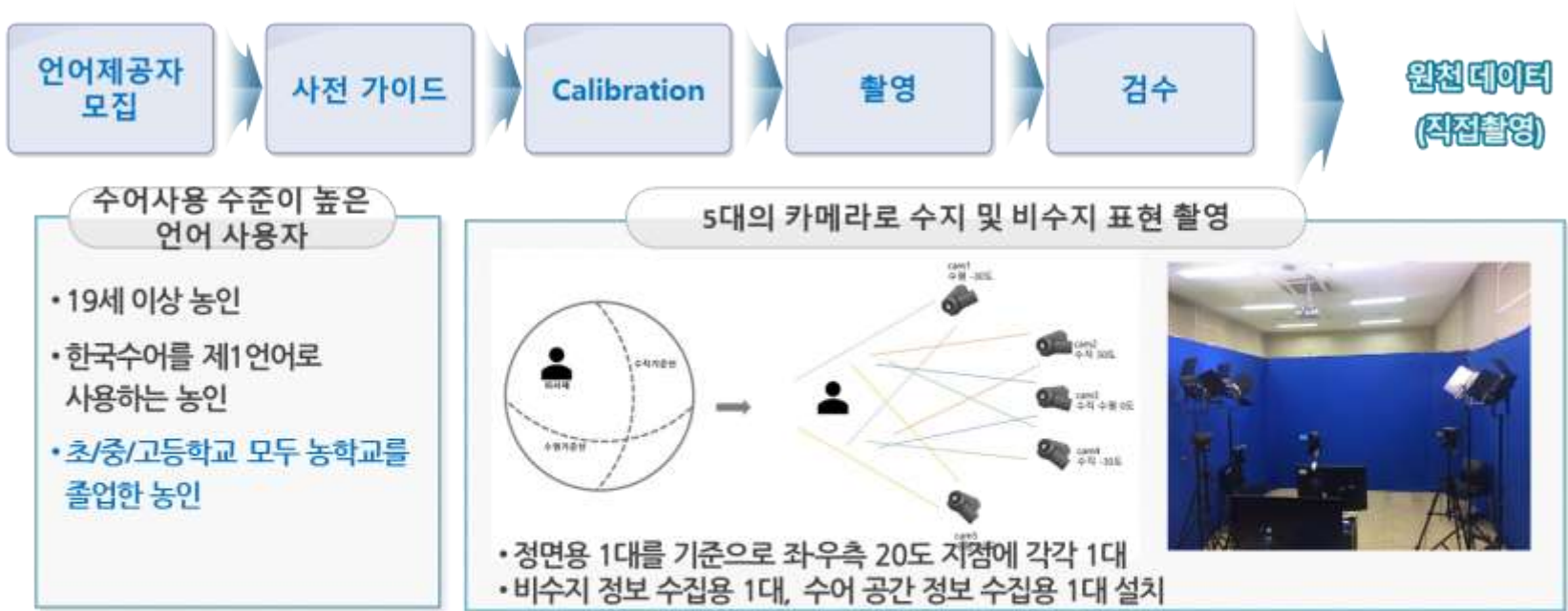


* 비수지 표현이란 수어에서 손 이외의 얼굴 표정, 몸의 움직임 등 음성 언어에서 높낮이, 강세, 장단에 해당

상황 별 도메인	주요 내용	비율
택시 탑승 상황	<ul style="list-style-type: none"> ▪ 목적지 요청 - "광화문까지 가주세요" ▪ 금액 관련 - "잔돈을 덜 거슬러 주셨어요" ▪ 시간 관련 - "3시까지 가야해서 서둘러주세요" 	<ul style="list-style-type: none"> ▪ 단어 및 문장 60%
고속터미널 이용 상황	<ul style="list-style-type: none"> ▪ 행선지 확인 - "대구 가는 버스를 여기서 기다리면 되나요?" ▪ 출발시간 확인 - "2시에 출발하는 버스는 어디서 타야 하나요?" ▪ 경유여부 확인 - "동 서울 가는 버스가 직행인가요, 경유를 하나요?" 	<ul style="list-style-type: none"> ▪ 단어 및 문장 20%
지하철 이용 상황	<ul style="list-style-type: none"> ▪ 목적지 확인 - "63빌딩으로 가려면 출구에서 어디로 가야 하나요?" ▪ 환승 확인 - "2호선으로 환승은 어디서 하나요?" 	<ul style="list-style-type: none"> ▪ 단어 및 문장 20%

01 인공지능 학습용 데이터 소개

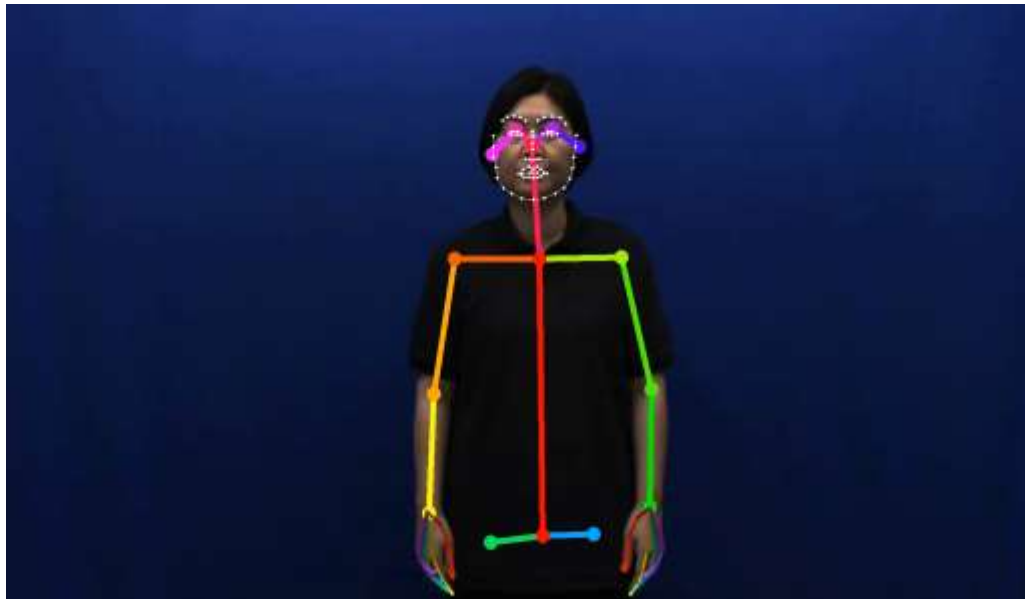
최적의 학습데이터 획득을 위한 한국 수어 사용자 모집 및 5방향 멀티 카메라 촬영을 통한 영상 획득(1/2)



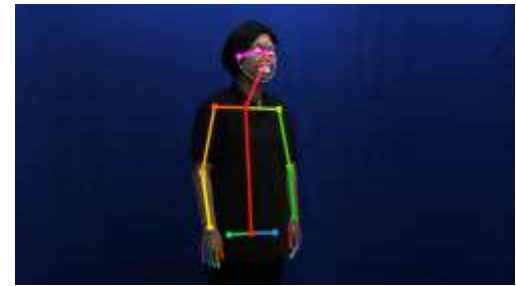
- Multi-Camera(5대) 시스템 활용, 카메라 동기화 및 Calibration 과정을 통해 다양한 각도에서 촬영된 영상 확보
- 1920 x 1080p, 60FPS 이상 해상도 촬영 및 주변 면 조명등을 추가해 충분한 셔터스피드를 확보함으로써 Motion Blur 제거
- 촬영장 내 언어제공자의 수어 속도 및 표현 확인하는 감수자 1명 및 촬영장비 및 영상자료 관리를 위한 보조인력 1명 배치, 영상 촬영 시점에서 누락 및 오류 여부를 확인하여 필요시 즉시 재촬영 진행

01 인공지능 학습용 데이터 소개

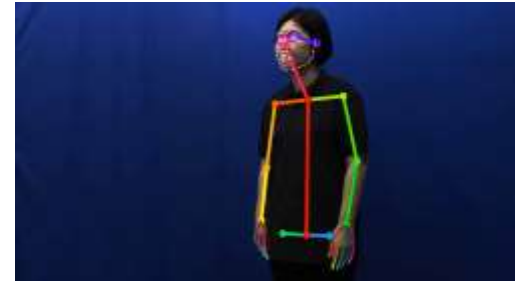
최적의 학습데이터 획득을 위한 한국 수어 사용자 모집 및 5방향 멀티 카메라 촬영을 통한 영상 획득(2/2)



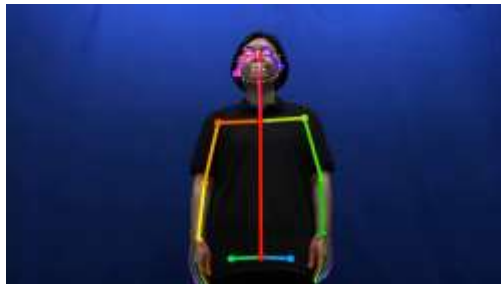
[정면]



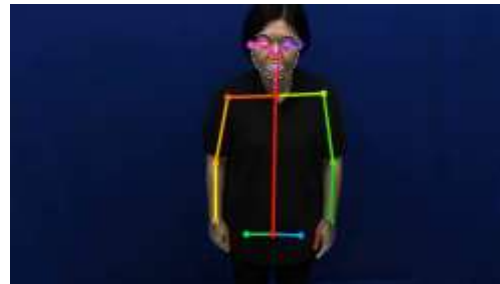
[우측면]



[좌측면]



[하단]



[상단]

- 5 채널 카메라 영상 (정면/우측면/좌측면/상단/하단)
- Pose: 25개 특징점 좌표(x,y), Confidence
 - Left/Right hand: 21개 특징점 좌표(x,y), Confidence
 - Face: 68개 특징점 좌표(x,y), Confidence

02 저작도구 활용법

저작도구 활용법 - Key point(키포인트) [1/2]



- ① 작업 표시 창 : 진행 중인 작업화면을 표시
- ② 프레임 톨 : 프레임(영상 장면) 작업에 필요한 각종 톨을 모아 놓음
- ③ 메뉴열기 : 키 포인트 가공 도구에서 제공되는 명령어들이 모여 있음
- ④ 패널 : 가공 도구에서 제공되는 기능 등을 팔레트 형식으로 모아 놓음

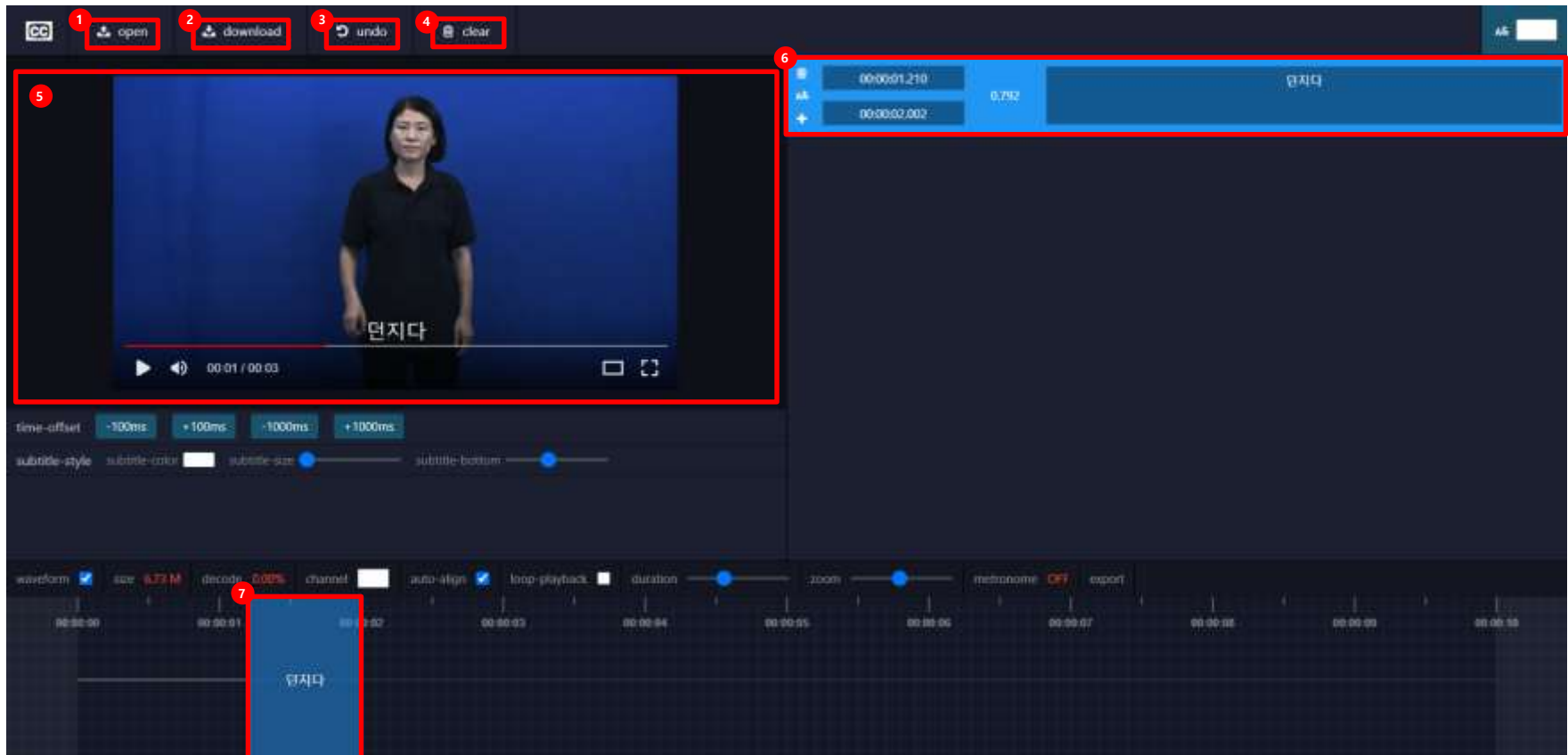
02 저작도구 활용법

저작도구 활용법 - Key point(키포인트) [2/2]



02 저작도구 활용법

저작도구 활용법 - Morpheme(형태소) [1/2]



- | | |
|---------------|--|
| ① open | : 진행 중인 동영상업로드 |
| ② download | : 작업 생성 완료 시 json 파일형식으로 결과값 저장함 |
| ③ undo | : 직전 작업을 실행 취소가능 |
| ④ clear | : 진행 중인 형태소를 모두 초기화 시킴 |
| ⑤ 수어 동영상 작업 본 | |
| ⑥ 형태소 편집 | : 형태소 설정(이름, 추가, 제거, 의문문)을 수정할 수 있는 편집 표시 창 |
| ⑦ 형태소 카드 | : 프레임 바 함께 활용하여 간단한 편집 조작 가능한 재생시간 또는 구간 조작 기능 |

02 저작도구 활용법

저작도구 활용법 - Morpheme(형태소) [2/2]

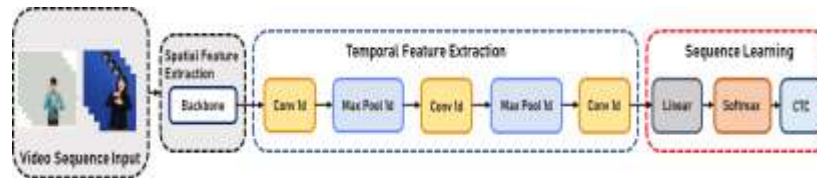


03 데이터 학습 목적(RCV)

가상 데이터를 통한
데이터 증강

데이터활용 및
인공지능 모델 개발

- 실제 도메인에서 취득한 수어 영상과 시뮬레이션을 통하여 생성한 아바타 수어 영상을 동시에 사용하여 수어 인식 모델을 학습시키고 가상 데이터의 활용 가능성을 검증
- 공개된 연속적 수어 인식 모델의 구현체가 존재하지 않았기 때문에 입력 영상으로부터 수어 문장에 대응하는 형태소 집합을 만들어 내는 연속적 수어 인식 베이스라인을 직접 구현
- 본래 형태소 별 영상에 대응하는 시간적 타이밍 정보가 존재하지 않는 상황에서 형태소의 순서를 학습시켜야 하기 때문에 Connectionist Temporal Classification(CTC) 디코더 및 손실 함수를 사용



데이터셋 증강 효과

- 각각의 보조 네트워크를 통하여 성능이 어떻게 개선되는지를 가시화하여 이를 평가하였으며, 특히 Pyramid Attention Network는 형태소의 타이밍 정확도를 높였고 Dense Gloss Supervision Network는 WER 평가 과정에서 삭제 현상을 크게 줄인다는 것을 확인.

$$WER = \frac{\#substitutions + \#deletions + \#insertions}{\#words\ in\ reference}$$

04 학습된 모델 결과 확인 방법 및 시연

데이터셋 학습 환경구축

- 가상환경 구축
- ```
$ conda env create -n environment.yml
```
- Train
- ```
$ python main.py --config-file configs/config.yaml
```
- Test
- ```
$ python main.py --config-file configs/exp_test.yaml --eval-only
```



### 데이터셋 학습 시연영상



## 05 서비스 개발 시 학습 된 모델 활용 방법

### 시범서비스 활용 방안(예시)

#### 배리어 프리(Barrier-free) 택시 서비스 개발

1. 본 과제를 통해 구축된 길 안내 관련 수어 데이터셋은 농인들이 택시에 탑승하여 목적지를 수어로 전달할 수 있는 인공지능 수어 인식 기술의 개발을 가능하게 함
2. 구체적인 주소지 뿐만 아니라 구역 내 지하철, 병원, 은행 등 주요 시설을 중심으로 경로를 설명할 수 있도록 경로 설명에 대한 인식 기술 개발 가능
3. 지화 데이터셋을 활용하여 농인들이 우편/택배 서비스를 위해 주소지를 입력하거나, 운전 시 네비게이션을 사용할 때 목적지를 편리하게 입력할 수 있는 지화 기반 초성 검색 기술을 개발 가능



< 배리어프리 택시 서비스 예시 >

#### 지하철 역사 주변의 주요 시설물 위치 안내 서비스 개발

1. 농인이 지하철 역무원에게 역사 주변의 주요 시설물(병원, 약국, 은행, 영화관 등)들의 위치를 문의하고자 할 때를 위한 인공지능 수어 인식 모델의 개발이 가능함



< 지하철 역사 내 활용 서비스 예시 >

# ..... 05 서비스 개발 시 학습 된 모델 활용 방법 .....

## 시범서비스 활용 방안(예시)



내부, 지하와 같은 공간은 GPS 정보를 제공 받기 어려움  
길안내와 지하철 역사 서비스를 사용하기에 어려움  
→ 지하철 역사 서비스 증대와 농아인을 위한 특화된 서비스 제공 필요



# 감사합니다

(주) 테스트웍스  
팀장 이상욱

E-MAIL : [sw.lee@testworks.co.kr](mailto:sw.lee@testworks.co.kr)

