



국민대학교
소프트웨어융합대학
소프트웨어학부

캡스톤 디자인 I

종합설계 프로젝트


프로젝트 명	수화 人(in) 플레이어
팀 명	전지적 수화 시점 팀
문서 제목	결과보고서

Version	1.5
Date	2019-05-28

팀원	윤지은 (조장)
	정창희
	이희지
	배진영
	축이림
지도교수	임성수 교수님

CONFIDENTIALITY/SECURITY WARNING


이 문서에 포함되어 있는 정보는 국민대학교 소프트웨어융합대학 소프트웨어학부 개설 교과목 캡스톤 디자인 수강 학생 중 프로젝트 “수화 人(in) 플레이어”를 수행하는 팀 “전지적 수화 시점”의 팀원들의 자산입니다. 국민대학교 소프트웨어학부 및 팀 “전지적 수화 시점”의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

문서 정보 / 수정 내역


Filename	수행결과보고서-수화 in 플레이어.doc
원안작성자	배진영, 윤지은, 이희지, 정창희, 축이림
수정작업자	배진영, 윤지은, 이희지, 정창희, 축이림

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2019-05-20	이희지	1.0	최초 작성	최초 작성
2019-05-22	윤지은	1.1	내용 추가	수정된 연구내용 추가
2019-05-23	배진영	1.2	내용 추가	향후 추진 계획 수정
2019-05-26	정창희	1.3	내용 추가	기술 작성
2019-05-27	전체	1.4	내용 수정	전체적인 수정
2019-05-28	전체	1.5	최종 점검	최종 점검

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

목 차

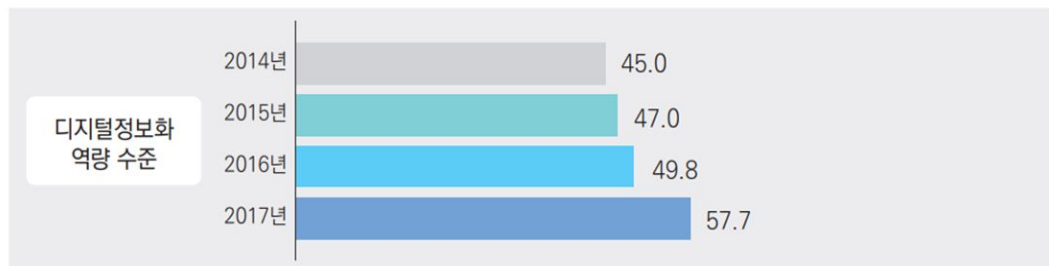
- 1 개요
 - 1.1 프로젝트 개요
 - 1.2 추진 배경 및 필요성
- 2 개발 내용 및 결과물
 - 2.1 목표
 - 2.2 연구/개발 내용 및 결과물
 - 2.2.1 연구/개발 내용
 - 2.2.1.1 단어 및 수화 DataSet 구축
 - 2.2.1.2 동영상의 음성에서 자막파일 추출 - STT
 - 2.2.1.3 수화 문법으로 재구성 - 형태소 분석
 - 2.2.1.4 수화 Data Mapping - 단어간 유사도 측정
 - 2.2.2 시스템 기능 요구사항
 - 2.2.3 시스템 비기능(품질) 요구사항
 - 2.2.4 시스템 구조 및 설계도
 - 2.2.4.1 시스템 구조
 - 2.2.4.2 시퀀스 다이어그램
 - 2.2.4.3 클래스 다이어그램
 - 2.2.5 활용/개발된 기술
 - 2.2.6 현실적 제한요소 및 해결방안
 - 2.2.6.1 하드웨어
 - 2.2.6.2 소프트웨어
 - 2.2.6.3 데이터
 - 2.2.7 결과물 목록
 - 2.3 기대효과 및 활용방안
- 3 자기평가
- 4 참고 문헌
- 5 부록
 - 5.1 사용자 매뉴얼
 - 5.2 배포 가이드
 - 5.3 테스트 케이스

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

1 개요

1.1 프로젝트 개요

1. 장애인 디지털정보화 수준 종합




'2017 디지털정보격차 실태조사'결과에 따르면

장애인 정보화 역량은 비장애인의 **절반 수준**

<그림1. 장애인의 디지털 정보화 수준>

한국정보화진흥원이 조사한 '2017 디지털정보격차 실태조사'결과에 따르면 장애인 정보화 역량은 비장애인의 절반 수준으로 나타났다. 특히 청각장애인은 어떤 장애 유형보다도 정보 습득이 어렵다. 시각장애인을 위해서는 전자점자 자료 및 점자번역기 등 솔루션이 있지만 청각장애인을 위한 수화로 번역된 영상의 제작과 배포 체계는 여전히 미흡한 상황이다. 따라서 우리는 청각장애인이 동영상을 원활하게 시청할 수 있도록 수화를 함께 제공하는 동영상 플레이어(이하 수화 in 플레이어) 제작을 목표로 한다.

수화 in 플레이어는 기존의 동영상에 ①자막, ②수화, ③번역기능을 제공한다. 결과적으로 청각장애인이 동영상 콘텐츠를 제약없이 이용함으로써 원하는 정보를 습득할 수 있도록 한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

1.2 추진 배경 및 필요성

[표 1] 연도별 장애인 현황

[단위: 천명]

	2010	2011	2012	2013	2014	2015	2016	2017	2018
청각, 언어	277	279	276	273	271	269	291	322	342

(출처 : 보건복지부 (시·도 장애인등록현황 자료))

최근에 사회적 약자에 대한 관심의 필요성이 대두되고 있다. 위 제시된 [표1]에 나온 바와 같이 2018년 기준 청각장애인의 수는 34만 명으로 지체장애 다음으로 많은 장애유형이며 그 수가 조금씩 증가하고 있지만, 사회적 여건은 열악하다. 이는 정보를 습득하는 과정에서도 문제점을 야기한다.

IT 기술 발달로 정보화 시대가 도래함에 따라 정보의 습득은 현대 사회에서 가장 중요한 부분으로 자리 잡았다. 정보는 TV, 신문, 영화, 라디오, 휴대폰 등 다양한 대중매체를 이용해 접근할 수 있으나 대부분의 경우 음성으로 제공한다. 이는 결과적으로 청인들과 정보격차를 발생시켜 청각장애인들을 정보 소외 계층으로 전락하게 만든다.

대규모의 세계적 영상 스트리밍 플랫폼인 Netflix와 Youtube는 자막을 제공하는데는 하지만 청각장애인들은 자막만으로는 영상의 내용을 정확하게 이해하기 힘들다. 이를 뒷받침 해 줄 기사의 일부 내용을 발췌하였다.


장애인정보문화누리 함효숙 활동가는 “방송을 볼 때 수화통역을 주로 보지만 중요한 내용은 자막과 같이 비교하며 보는데, 그래야 내용을 더 정확하게 알 수 있기 때문” 이라면서 “나에게 자막과 수화통역은 자유롭게 선택되어야 하는 언어” 라고 강조했다.

이날 기자회견에서 장애인정보문화누리 안세준 고문은 “나는 어려서 청각장애를 입어 수화가 모국어나 마찬가지”라면서 “자막도 방송을 보는 데 도움이 되기는 하지만 전달이 잘 안 될 때가 잦다. 그래서 방송을 볼 때 수화통역이 있는 것을 더 좋아한다”라고 밝혔다.

(“선거방송, 수화와 자막 동시에 제공하라”, 비마이너 ,
2012년 11월 23일, (<http://beminor.com/detail.php?number=4586>))

기사 내용으로 보다시피 청각장애인이 내용을 정확하게 이해하기 위해서는 자막만이 아닌 수화를 동시적으로 보여주는 것이 필요하다.

미국은 청각장애인을 위한 수화 번역 툴들이 개발된 상태이지만 우리나라는 아직 청각장애인을 위한 서비스가 부족하다. 또한 이를 활용할 수화 데이터 또한 부족한 상태이다. 현재 우리나라에서 상용화되고있는 서비스로 삼성의 ‘수화번역 도우미’가 있으나 이는 표정을 제공하지 않아 정확한 의미 전달을

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

하지 못하며, 문장을 입력하면 수화 애니메이션을 보여주는 방식으로 청각장애인을 위한 서비스는 아니다.

우리는 이러한 환경에서 청각장애인을 위한 수화 번역 영상플레이어 제작의 필요성을 느껴 본 프로젝트를 진행하고자 한다.

2 개발 내용 및 결과물

2.1 목표


본 프로젝트의 이름은 ‘수화 in 플레이어’로 사용자가 정보를 얻고자 하는 영상을 등록하면 영상의 내용을 수화로 번역하여 이를 등록된 영상과 함께 사용자에게 제공하는 것이 목표이다.

현재 시중에 나와있는 대부분의 영상매체는 자막이 함께 제공되는 영상이 적을 뿐더러 더욱이 수화 영상을 함께 제공하는 것은 뉴스나 일부 수화번역 영상 이외에 거의 찾아볼 수 없다. 또한 외국의 영상에서는 한국어 자막을 거의 찾아볼 수 없고 수화 영상이 제공된다 하더라도 한국 수화가 아닌 그 나라의 수화이기 때문에 결론적으로 청각장애인은 외국 영상을 접하는데 큰 어려움이 있다. 점차 세계가 글로벌화 되고 있고 IT 기술 발달로 영상 매체의 의존도가 높아짐에 따라 청인과 청각장애인 간의 정보 격차는 심화될 것으로 예상된다. 따라서 본 프로젝트의 기대효과는 청각장애인과 인터넷 강의나 시사/교양 프로그램 등 정보를 담고 있는 영상에서 원하는 정보를 제약없이 습득할 수 있도록 하는 것이다. 목표 달성을 위한 기술적 단계는 크게 6가지로 나뉜다.

- 1) 단어 및 수화 DataSet 구축
- 2) 동영상의 음성에서 자막 파일 생성
- 3) 영어 자막을 한국어 자막으로 번역
- 4) 한국어 문법을 수화 문법 체계 형식으로 재구성
- 5) 추출한 형태소와 수화 Data Mapping방식 구현
- 6) 서비스를 제공하기 위한 웹 페이지 구현

본 프로젝트의 주된 목적은 영상을 자막 및 수화로 번역하는 시스템 자체의 설계 및 구현이며, 모든 문장을 정확히 해석하는 것보다 전체적인 영상의 내용을 파악할 수 있을 정도의 효과를 목표로 한다.

따라서 본 프로젝트의 전제 조건은 다음과 같다. Speech To Text(이하 STT) 과정에서 발생하는 번역 시간의 지연은 고려하지 않도록 한다. STT를 이용하여 영상의 음성을 TXT파일로 변환하기 까지는 영상 길이의 ¼ 정도가 걸린다. 하지만 이 프로젝트의 핵심 기술은 자막을 수화로 ‘번역’하는 것이기 때문에 현실적인 프로젝트 진행을 위해 이 문제는 고려하지 않는다. STT나 형태소 분석기의 오인식으로 발생하는 몇몇의 번역 오류 역시 추후의 이 기술들이 개선되었을 때 자연스럽게 해결 될 문제라고 판단한다. 또한 한국어의 중의적 표현은 고려하지 않고 직역한다. 중의적 표현의 특징은

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

해학이나 풍자 등에 활용되며 의미의 다양성으로 문학 작품의 예술성을 높이고 싶을 때 사용된다. 하지만 중의적 표현을 전달하지 못하는 것이 전체 영상의 내용을 이해하는데 큰 장애가 되지 않기 때문에 중의적 표현의 처리 기능은 구현하지 않기로 한다.

2.2 연구/개발 내용 및 결과물

2.2.1 연구/개발 내용

2.2.1.1 단어 및 수화 DataSet 구축

(1) Crawling

크롤링은 웹 페이지를 그대로 가져와 해당 페이지 내의 데이터를 추출해 내는 행위를 말한다. 본 프로젝트에서는 단어마다 단어명, 품사, 의미, 수화 영상 Data를 수집하기 위해 사용한다. 국립국어원한글수어사전 홈페이지 (<http://sldict.korean.go.kr/front/main/main.do>)의 수화 영상을 크롤링하여 DB에 저장한다.

(2) SQLite / MongoDB


대용량 빅데이터를 처리하기 용이한 NoSQL인 MongoDB와 내장 DB인 SQLite를 사용한다. DB는 수어사전 홈페이지에서 제공하는 단어를 저장하는 Dictionary, 사용자가 업로드한 Video파일 및 URL, 언어 정보를 저장하는 Video로 나뉜다.

< Dictionary >

- 수어사전 홈페이지에서 제공하는 단어 약 18,000개를 저장한다.
- Collection으로는 일상생활 단어가 담겨 있는 Basic Collection, 숫자 단어가 담겨있는 Number Collection, 지화 단어가 담겨있는 Finger Collection이 있다.
- 각 Collection의 Documents에는 단어명을 나타내는 word, 의미가 담겨있는 mean, 품사정보가 담겨있는 part, 수화동영상의 url을 갖는 location, 유사도 검사에서 쓰일 참조 단어가 담겨있는 ref_word Fields가 있다.

[표 2] basic dictionary


ID	WORD	PAR T	Mean	REF_WOR D	LOCATION
751	피고인	명사	형사 소송에서, 검사에 의하여 형사 책임을 져야 할 자로 공소 제기를 받은 사람.		player/media/sign/basic/751.mp4
750	피고소인	명사	범죄 사실을 근거로 하여 고소를 당한 사람.		player/media/sign/basic/750.mp4
749	판결의경정	명사	판결의 내용에는 오류가 없으나 위산, 오기, 기타 이에 유사한 표현의		player/media/sign/basic/749.m

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전자적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

			오류나 표시의 부적당한 것이 명백한 경우에 이것을 정정하는 일.		p4
748	맞추다	용언	정도에 알맞게 하다.	정도	player/media/sign/basic/748.mp4
747	맞추다	용언	다른 사람의 의도나 의향 따위에 맞게 행동하다.	의도	player/media/sign/basic/747.mp4
746	맞추다	용언	서로 떨어져 있는 부분을 제자리에 맞게 대어 붙이다.	부분	player/media/sign/basic/746.mp4
745	맞추다	용언	다른 어떤 대상에 달게 하다.	입	player/media/sign/basic/745.mp4
744	맞추다	용언	서로 어긋남이 없이 조화를 이루다.	조화	player/media/sign/basic/744.mp4
743	맞돈	명사	물건을 사고팔 때, 그 자리에서 즉시 치르는 물건값.		player/media/sign/basic/743.mp4
742	망향	명사	고향을 그리워하며 생각함.		player/media/sign/basic/742.mp4

< Video >

- 사용자가 upload한 동영상 또는 유튜브 링크 영상의 정보를 저장하는 DB이다.
- 사용자가 업로드한 동영상에서 음성 추출, 자막 생성을 하기 위해 불러오며 재가공하게 된다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

2.2.1.2 동영상의 음성에서 자막파일 추출 - STT(Speech to Text)



<그림 2. STT 기술의 일반적인 구성>

(출처: <https://terms.naver.com/entry.nhn?docId=3386818&cid=58369&categoryId=58369>)

음성인식이란 컴퓨터가 마이크와 같은 소리 센서를 통해 얻은 음향학적 신호(acoustic speech signal)를 단어나 문장으로 변환시키는 기술을 말한다. 최근에는 많은 기업들이 머신러닝 기반 음성 처리 프로그램의 개발자 환경을 지원하는 응용프로그램 인터페이스(API)를 제공하고 있다. 대표적인 API로는 구글의 Cloud Speech-to-Text, IBM의 Watson Speech-to-Text 등이 있다.

카카오나 네이버 같은 우리나라의 소프트웨어 기업들도 개발자들은 위한 음성 API를 제공하고 있지만 자막 제작을 위한 시간 정보를 제공하지 않는다. 따라서 본 프로젝트는 time stamp를 제공하는 Google Cloud Speech-to-Text를 사용한다.

입력 동영상에서 음성인식 정확도를 위해 wav파일로 변경하고 44100Hz로 인코딩한다. 음성인식 후에는 VTT형식의 자막파일을 생성한다. VTT는 자막의 번호, 타임코드, 자막 내용으로 구성되는 구조를 가지며, 음성인식을 통해 얻어진 텍스트와 각 단어들의 time stamp를 활용하여 VTT를 작성한다.

2.2.1.3 수화 문법으로 재구성 - 형태소 분석


한국어를 수화 문법 체계로 재구성하기 위해 형태소 분석을 수행한다. 아래는 형태소 분석의 정확도를 측정한 결과이다.

[표 3] 형태소 분석 정확도

총 단어 수(개)	오류	정확도
1,484	9.14%	91.86%

총 길이 17분의 강연 영상을 테스트하였고 그 결과 [표 3]에서와 같이 비교적 높은 정확도를 보여주었다. 고유명사, 줄임말, 신조어 등에서 분석 오류가 발생했지만, 91.86%의 정확도를 보였다.

수화의 문법 구조는 전라북도교육연구정보원의 사랑의수화교실을 참고하였다. 수화에서는 사용하는 조사가 101개로 사용하는 조사가 361개인 한국어에 비해 한정적이다. 아래는 수화에서 사용하는 조사를 나타낸 표이다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

[표 4] 수화에서 사용되는 조사

구분	사용되는 조사
격조사	랑, 에서, 서, 더러, 보다, 에게, 의, 로, 이라고, 예, 처럼, 께, 으로, 한테, 라고, 게, 로서, 로써, 으로서, 으로써, 만치, 만큼, 보고, 보다, 서부터, 에서부터, 시여, 이시여, 아
서술격 조사	이니까, 이다, 야
보조사	이나, 든지, 부터, 도, 커녕, 마다, 밖에, 뿐, 만, 까지, 라도, 이라도, 라든지, 거나, 나, 나마, 이나마, 는, 다가, 대로, 든가, 들, 따라, 라야, 란, 이란, 로부터, 마는, 마저, 밖에, 부터, 뿐, 야말로, 이야말로, 에는, 에다, 에다가, 은, 이, 이라야, 이야, 조차
접속조사	과, 와, 이랑, 고, 이고, 며, 이며


수화 문장에서는 동사나 형용사가 쓰임에 따라 바뀌지 않고 그 기본형이 서술어로 쓰인다. 또한 용언의 과거형은 용언의 현재형과 ‘끝’이라는 의미의 수화를 합성하여 나타낸다. 의문문은 의문을 나타내는 종결어미 ‘느니라’라는 수화를 문장의 끝에 배치시켜 의문문을 나타낸다. 아래는 이를 반영하여 한국어를 수화의 문법 구조로 재구성한 결과이다.

[표 5] 수화 문법 구조로 재구성

구분	언어	재구성 결과
(1)	한국어	우리 생활과 민주정치 만화로 만나 봤어요.
	수화 문법	우리 생활 과 민주 정치 만화 로 만나다 보다 끝
(2)	한국어	요즘 강아지 많이 키우시죠?
	수화 문법	요즘 강아지 많이 키우다 느니라
(3)	한국어	혹시 그 강아지가 어디서 태어난 지 알고 계십니까?
	수화 문법	혹시 그 강아지 어디 에서 태어나다 지 알다 고 계시다 느니라
(4)	한국어	이번만 도와주시면. 우리 예서 서울 의대 보낼 수 있어요 어머니
	수화 문법	이번 만 도와주다 면 우리 예서 서울 의대 보내다 수 있다 어머니

2.2.1.4 수화 Data Mapping - 단어간 유사도 측정

수화 문법 체계로 재구성한 형태소들에 적절한 수화를 매핑시켜주기 위해 수화 데이터베이스에서 쿼리문을 이용하여 형태소의 의미와 맞는 수화 동영상의 경로를 가져온다. 만약 수화와 매핑을 할 때 동음이의어가 나올 경우를 대비하여 DB에 참조단어 속성을 추가적으로 저장한다. 문장에서 동음이의어가 나오면 먼저 DB에 있는 품사정보를 비교하여

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

같은 품사인 동음이의어를 선택해준다. 만약 품사가 서로 다르다면 문장 내의 동음이의어와 가장 가까운 명사를 가져온다. 그 다음 DB에 저장되어 있는 동음이의어들의 참조단어과 유사도 검사를 수행한다. 단어간 유사도를 계산하여 이 중 유사도가 가장 높은 단어를 선택한다. 아래는 단어 유사도 검사를 수행한 결과를 시각화한 표이다.

[표 6] 단어 유사도 검사 시각화


단어1		배				단어2		사과			
알고리즘 종류	ET RI	Resnik	Lin	Jiang and Conrath	Leacock and Chodorow	Wu & Palmer	Hirst and St-Onge	Banerjee and Pedersen	Patwardhan	Pekarek et al	Lin + GraS M
유사도	0.67	0.57	0.76	0.21	0.81	0.86	0.63	0.06	0.5	0.75	0.76

다음은 단어 유사도 검사를 동음이의어 처리에 적용한 결과이다. 평균 유사도란 [표 6]에서 알고리즘 별 유사도의 평균을 구한 값이다.

[표 7] 단어유사도 검사 동음이의어 처리에 적용

예문			
지나다니던 길에 애견샵이 있는데 -> 지나다니다 던 길 에 애견샵 있다			
DB 동음이의어	뜻	참조단어	평균 유사도 (길-동음이의어)
에 (1)	((체언 뒤에 붙어)) 앞말이 처소의 부사어임을 나타내는 격 조사	장소	0.04951
에 (2)	앞말이 시간의 부사어임을 나타내는 격 조사	시간	0.04105
에 (3)	앞말이 원인의 부사어임을 나타내는 격 조사	원인	0.02848
에 (4)	앞말이 조건, 환경, 상태 따위의 부사어임을 나타내는 격 조사	조건	0.01845
에 (5)	앞말이 제한된 범위의 부사어임을 나타내는 격 조사	범위	0.0

예문
그리고 그 강아지들은 도매장으로 갑니다. -> 그리고 그 강아지 들 은 도매 장 으로 가다

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전자적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28


DB 동음이의어	뜻	참조단 어	평균 유사도(도매-동음이의어)
장 (1)	농장, 옷장, 찬장, 책장 따위의, 물건을 넣어 두는 가구를 통틀어 이르는 말	옷장	0.0
장 (2)	간장, 고추장, 된장 따위를 통틀어 이르는 말	음식	0.0
장 (3)	어떤 조직체나 부서 단위의 우두머리	조직체	0.0
장 (4)	많은 사람이 모여 여러 가지 물건을 사고파는 곳. 지역에 따라 다르나 보통 한 달에 여섯 번 선다.	장사	0.06422

예문			
<p>저는 반려견 훈련사 강형욱입니다. -> 저 는 반려 견 훈련 사 강 형 욱 이다</p>			
DB 동음이의어	뜻	참조단 어	평균 유사도(훈련-동음이의어)
장 (1)	‘회사’의 뜻을 더하는 접미사	회사	0.02061
장 (2)	‘직업’의 뜻을 더하는 접미사	직업	0.03390

2.2.2 시스템 기능 요구사항

[표 8] 시스템 기능 요구사항

번호	내용	완료 여부
FR1	사용자가 영상파일을 웹 페이지에 업로드한다.	완료
FR2	사용자가 유튜브 URL을 입력하면 해당 URL로부터 영상과 자막을 다운받는다.	완료
FR3	자막과 수화영상 옵션을 선택적으로 제공받을 수 있다.	완료
FR4	업로드 된 동영상은 구글 클라우드 스토리지에 저장한다.	완료
FR5	소리추출을 위해 영상을 형식에 맞게 encoding한다.	완료
FR6	음성파일에서 구글 STT API를 사용해 음성을 텍스트로 변환한다.	완료


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

FR7	텍스트를 자막파일로 변환하여 저장한다.	완료
FR8	영어자막일 경우 구글 Translation API를 통해 한국어로 번역한다.	완료
FR9	자막 파일의 내용을 Pykomoran을 사용하여 형태소 분석을 한 뒤 수화 문법 체계와 맞게 재구성한다.	완료
FR10	단어간의 유사도검사를 통해 동음이수어를 처리한다.	완료
FR11	재구성한 자막 파일을 바탕으로 적절한 수화 동영상과 매칭 후 영상을 이어붙여 수화영상을 만든다.	완료
FR12	사용자에게 최종적으로 사용자가 등록했던 동영상과 자막, 그리고 수화 번역 영상을 함께 제공한다.	완료

2.2.3 시스템 비기능(품질) 요구사항

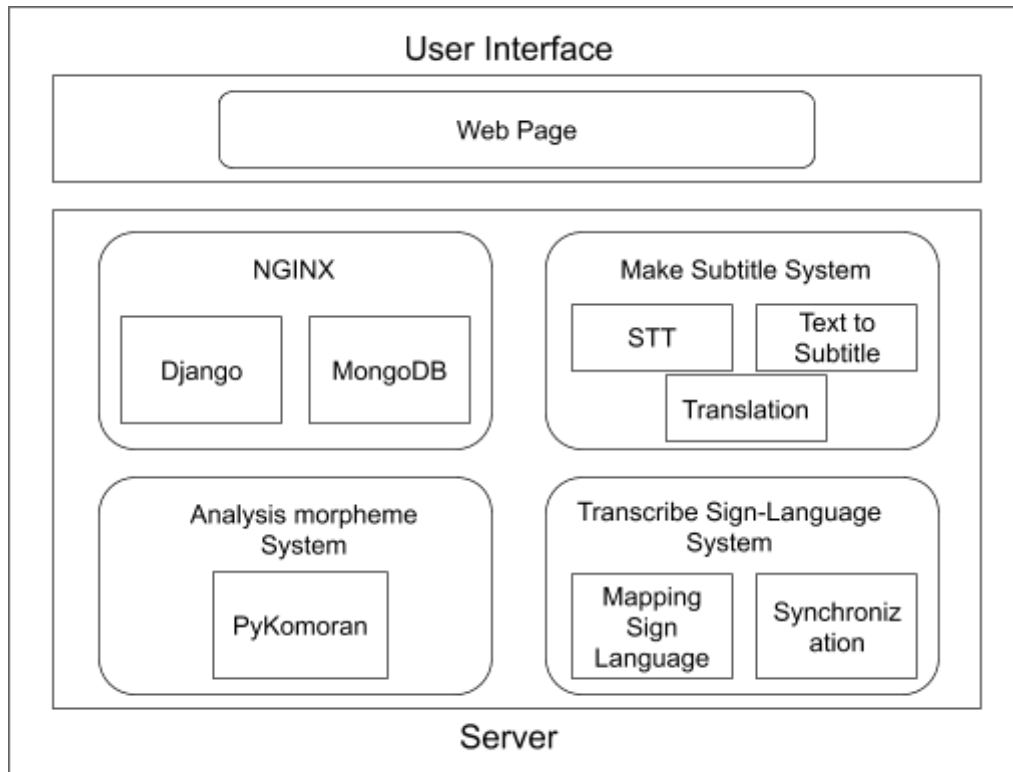
[표 9] 시스템 비기능(품질) 요구사항

번호	내용	품질 속성	달성 여부
NFR1	데이터 베이스는 적어도 15000개 이상의 수화 데이터를 담을 수 있어야 한다.	Efficiency(효율성)	달성
NFR2	시스템은 사용자에게 플레이어 매뉴얼을 제공해야 한다.	Learnability(학습성)	달성
NFR3	사용자는 몇번의 클릭으로 웹 페이지의 기능들을 파악할 수 있어야 한다.	Understandability (이해성)	달성
NFR4	시스템은 사용자가 등록한 영상의 정보를 외부에 유출하지 않도록 한다.	Legislative(합법성)	달성
NFR5	시스템은 등록된 영상을 서버 클라우드에 빠르게 저장한다.	Efficiency(효율성)	달성
NFR6	동영상 업로드 후 번역까지의 처리시간이 영상 시간의 1/20을 넘지 않도록 한다.	Efficiency(효율성)	미달성

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

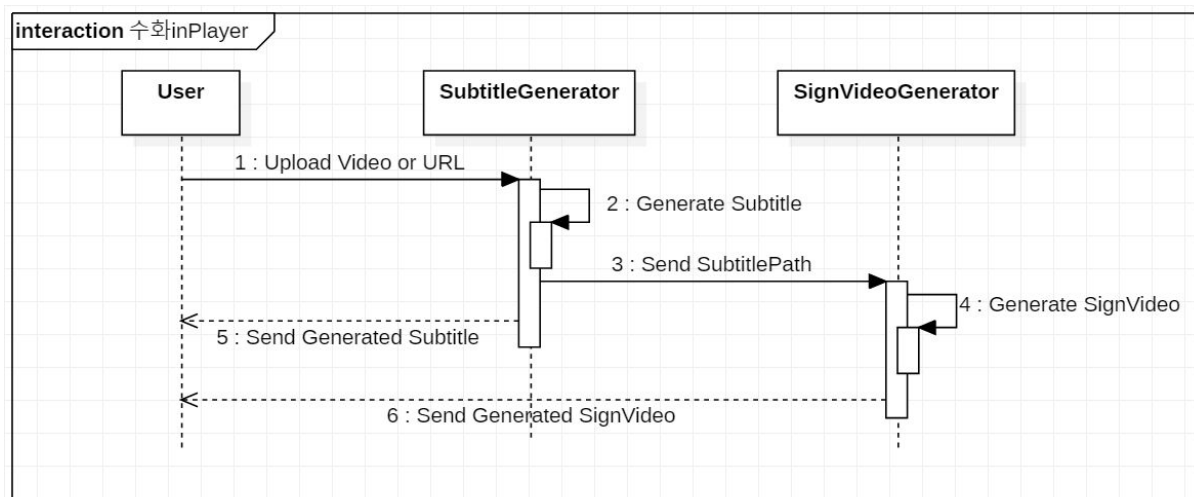
2.2.4 시스템 구조 및 설계도

2.2.4.1 시스템 구조도




<그림 3. 시스템 구조도>

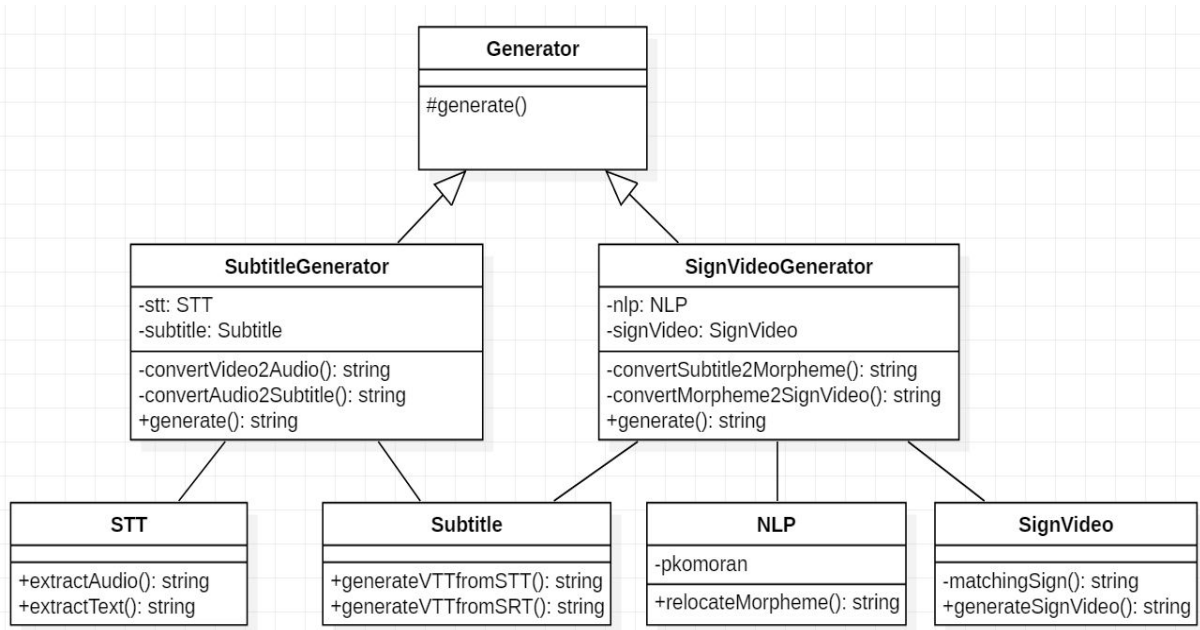
2.2.4.2 시퀀스 다이어그램



<그림 4. 시퀀스 다이어그램>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

2.2.4.3 클래스 다이어그램



<그림 5. 클래스 다이어그램>

2.2.5 활용/개발된 기술

1) Server


- AWS EC2와 Nginx를 이용하여 웹 서버 환경을 구축하였다. 개발에 사용되는 언어가 Python이기 때문에 호환성을 고려하여 Django 프레임워크를 사용하였다.
- 수화 데이터셋을 구축하기 위해서 NoSQL인 MongoDB / SQLite를 사용하며, 방대한 수화 영상을 효율적으로 저장하기 위해 Data Crawling 모듈인 BeautifulSoup를 사용한다.

2) 자막 생성

- Encoding
- Google Speech To Text
- Google Translation
- Text To Subtitle

3) 수화 생성

- PyKomoran
- ETRI OpenAPI - 어휘간 유사도 분석 API
- Make Sign Language Video

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

2.2.6 현실적 제한 요소 및 해결 방안

2.2.6.1 하드웨어

많은 양의 수화 데이터와 유저가 업로드한 동영상을 저장하고 재가공한 영상을 배포하기 위해서 서버 컴퓨터가 따로 존재해야 한다. 이를 해결하기 위해 AWS의 EC2 인스턴스의 Ubuntu 16.04 운영체제를 활용한 서버 컴퓨터를 사용한다.

2.2.6.2 소프트웨어

1) 수화번역 정확도

본 프로젝트는 영상의 음성을 수화로 번역하는데 이렇게 산출된 수화영상의 정확도를 확인하기가 어렵다. 산출된 수화 영상이 제대로 번역이 되었는지 확인하기 위해 수화 전문가나 청각장애인의 자문이 필요한데 이를 얻기가 쉽지 않다. 따라서 웹 페이지에 사용자 피드백 게시판을 만들어 만약 잘못 번역된 수화나 불편한 점이 있다면 사용자가 이를 활용해 피드백을 남기고 관리자가 승인하여 수정하도록 구현하였다.


2) 수화영상 생성시간

사용자가 동영상을 업로드하고 수화 영상이 함께 제공되기까지의 전처리 시간이 길다. 특히, STT와 수화 영상을 merge할 때 대부분의 시간이 소요된다. 이를 해결하기 위해 두 가지 솔루션을 가진다. 첫째로, 사용자가 동영상을 직접 업로드하는 방식과 함께 유튜브 링크로 업로드하는 방식을 제공한다. 만약 유튜브 영상에 자막이 있다면 서버측에서 영상과 자막 파일을 함께 저장해 STT를 사용하지 않도록 구현하였다. 둘째로, 수화 영상을 merge하는 과정에서 multiprocessing을 사용하여 시간을 축소하였다.

2.2.6.3 데이터

대한민국에서 수화에 대한 정보를 얻을 수 있는 대표적인 웹 사이트로 '국립국어원 한국수어사전'이 있다. 이 사이트는 UI 디자인이 사용하기 어렵고 수화 데이터의 업로드 시기가 오래되었다. 따라서 수화 데이터를 얻기 어렵고, 서비스를 제공하기 위한 데이터가 부족했다. 이를 해결하기 위해 부족한 데이터는 직접 촬영하여 추가한다.


2.2.7 결과물 목록

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

대분류	소분류	기능	기술문서 유/무
Web Page	웹 페이지 소개	사용자가 웹 페이지에 대한 이해도를 높일 수 있다.	무
	영상 업로드	수화로 번역하고 싶은 영상 파일을 등록한다.	무
	유튜브 URL 로드	수화로 번역하고 싶은 영상이 담긴 url을 로드한다.	무
	개선사항 피드백	사용자는 사용 후 전반적인 피드백을 남길 수 있다.	무
Server	Encoding	음성인식을 하기 위해서 mp4형식으로 바꾼 뒤 음성만 추출한다.	무
	STT	Google Speech to Text API를 사용하여 음성을 텍스트로 변환한다.	유
	Subtitle	음성인식한 텍스트파일을 사용하여 VTT 자막 파일로 변환한다.	무
	Translate	영어 영상일 경우 Google Translation API를 사용하여 한국어로 번역해준다.	무
	수화 문법으로 재구성	한국어를 수화 문법 체계로 재구성하기 위해 형태소 분석을 수행하고 재배치한다.	유
	동음이의어 처리	동음이의어가 나올 경우 품사를 비교하여 선택한다. 동일한 품사일 경우 단어간 유사도 검사를 통해 유사도가 더 높은 쪽의 단어를 선택한다.	유
	데이터 매핑	쿼리문을 작성하여 DB안에 단어의 유무를 판단하고 영상의 위치를 반환한다.	무
	수화 영상 제작	데이터 매핑에서 반환 된 수화 영상 위치를 받고 영상의 길이 정보에 맞춰 조정하여 최종 수화번역 영상을 생성한다.	무

2.3 기대효과 및 활용방안

수화 in 플레이어 프로젝트를 통해 기대하는 효과 및 활용방안은 다음과 같다. 원하는 영상을 플레이어로 재생하면 소리를 인식하여 자막과 수화를 선택적으로 서비스를 제공받을 수 있다. 즉, 청각장애인도 다양한 분야의 동영상 콘텐츠를 제한없이 이용할 수 있다. 또한 내용의

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

이해도가 향상되기 때문에 청인들과의 정보 격차 또한 줄어들 것으로 예상된다.

아울러 음성을 수화로 번역해주는 시스템은 동영상에만 국한된 것이 아닌, 강의나 회의같은 상황에서 활용할 수 있다. 실시간으로 청인의 말을 인식한뒤 수화로 번역하여 청각장애인들이 청인과 수월하게 소통할 수 있도록 해줄 것이다.

3 자기평가

청각장애인들에게 도움이 되고자 시작한 이 프로젝트는 생각보다 쉽지 않았다. 우리나라에서 제공하는 '국립국어원 한국수어사전' 사이트의 UX/UI는 낙후된 듯 보였으며 수화영상 또한 오래되어 보였다. 수화 영상은 합부로 사용할 수 없는 저작권 문제가 있었고 프로젝트를 구현하는 것도 중요하지만 데이터의 질과 양의 중요성을 크게 느낄 수 있었다. 또한 장애인들을 위한 서비스 준비가 많이 부족한 것 같아 아쉬웠다.


수화 데이터를 조정하여 최종 수화영상을 만들 때 대부분 ffmpeg를 사용하였다. ffmpeg는 속도면에서 대부분 우수하였으나 여러 연산을 처리하는 경우엔 제대로 처리가 되지 않는 부분이 있었다. 그래서 수화 영상 생성 과정 중 일부는 간편하지만 속도가 상대적으로 느린 moviePy 패키지를 사용하였다. 비디오 파일의 자세한 구조나 ffmpeg의 설정 사항을 더 학습하여 적용한다면 서비스 속도를 조금 더 향상시킬 수 있을 것으로 보인다.

우리 프로젝트는 영상에서 음성을 인식해 텍스트를 추출하고 영어를 한국어로 번역하는 API를 사용했다. API의 한계로 정확도에 문제가 있어 보여서 아쉬웠다. 하지만 우리의 목적은 모든 문장을 정확히 해석하는 것이 아니라 전체적인 영상의 내용을 파악하도록 하는 것이다. 청각장애인들을 위한 서비스 개발에 조금이라도 보탬이 되기를 바란다.

동영상으로부터 수화 영상과 자막을 제공하는 전체적인 기능이 성공적으로 구현 되었으므로 사용 가능성이 있다고 생각한다. 또한 동영상에서 음성을 추출하여 자막을 생성하는 모듈과 자막에서 다시 수화 문법으로 재구성하여 수화 번역 영상으로 생성하는 모듈이 각각 분리되어 구현되어 있으므로 각각의 결과물들도 재사용 가능성이 있어 의미있다고 생각한다.

4 참고 문헌

번호	종류	제목	출처	발행년도	저자	기타
1	논문	음성인식 기반의 사이버 수화통역사 개발	경인교육대학교	2015년	김성헌	


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서				
	프로젝트 명		수화 In 플레이어		
	팀 명		전지적 수화시점		
	Confidential Restricted		Version 1.5		2019-5-28

2	웹페이지	한국수어 사전	국립국어원	-	국립국어원	
3	논문	2017년 한국수어 사용 실태 조사 연구	국립국어원	2017년	이대섭	
4	기사	형식적인 자막방송만, 수화,해설방송은 가끔 시늉만	미디어오늘 (http://www.mediatoday.co.kr/?mod=news&act=articleView&idxno=132500)	2016년	금준경	
5	기사	“선거방송, 수화와 자막 동시에 제공하라”	BeMinor (http://beminor.com/detail.php?number=4586)	2012년	홍 권호	
6	기사	“靑 국민청원 답변 영상에 자막,수화 넣어주세요”	국민일보 (http://news.kmib.co.kr/article/view.asp?arcid=0012575346)	2018년	우승원	
7	기술개발 보고서	청각 장애인용 모바일 애플리케이션 가이드북	SlideShare (https://www.slideshare.net/teaminterface/ss-27477625)	2013	지식경제부	
8	기사	한국 수화의 구조를 밝히다 - 『한국수화 문형사전』 발간	국립국어원 (https://www.korean.go.kr/nkview/onletter/20070401/print/02.html)	2015	국립국어원	
9	웹페이지	수화교실	전북교육영상	-	전라북도교육연구정보원	

5 부록

5.1 사용자 매뉴얼

1. 수화 in 플레이어 웹 페이지(<http://13.124.80.174:8000/>)에 접속하고 시작하기 버튼을 누른다.
메인 화면에는 수화 in 플레이어 소개가 적혀있다.

 <div> 국민대학교 컴퓨터공학부 캡스톤 디자인 I </div>	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28



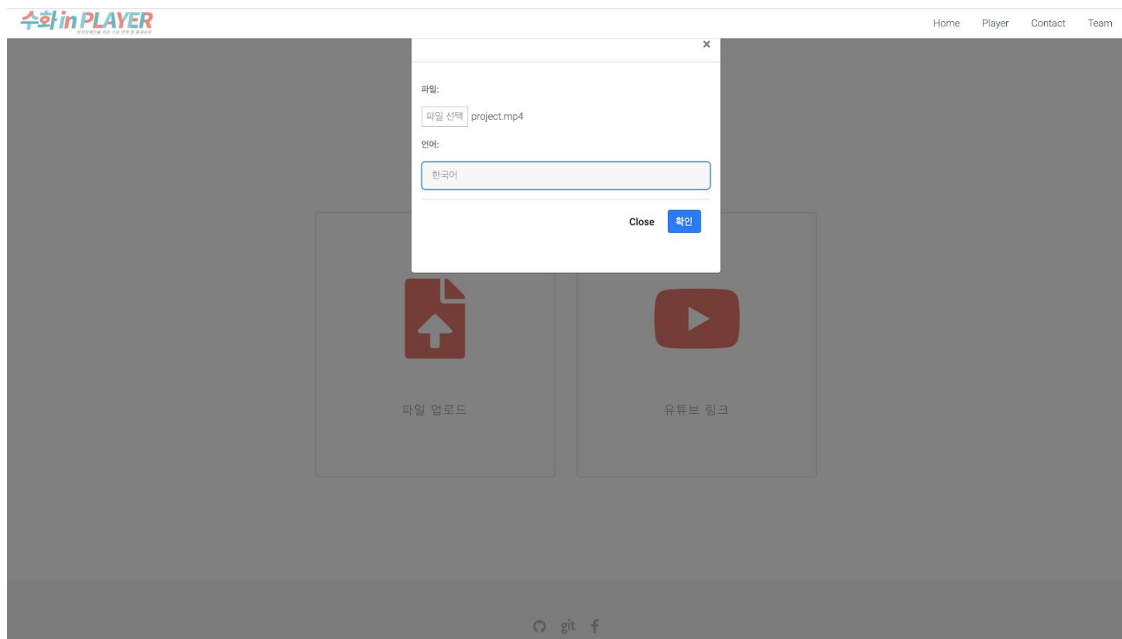
2.

a. 사용자가 영상 파일을 업로드한다.

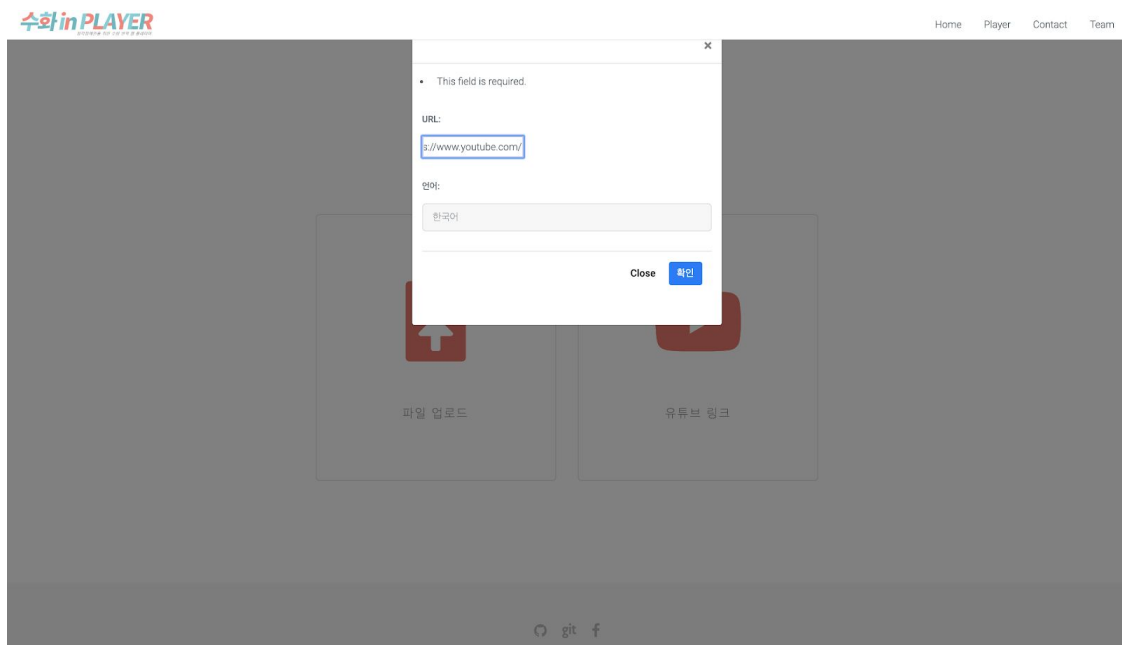
결과보고서	
프로젝트 명	수화 In 플레이어
팀 명	전지적 수화시점
Confidential Restricted	Version 1.5
	2019-5-28



국민대학교
컴퓨터공학부
캡스톤 디자인 I




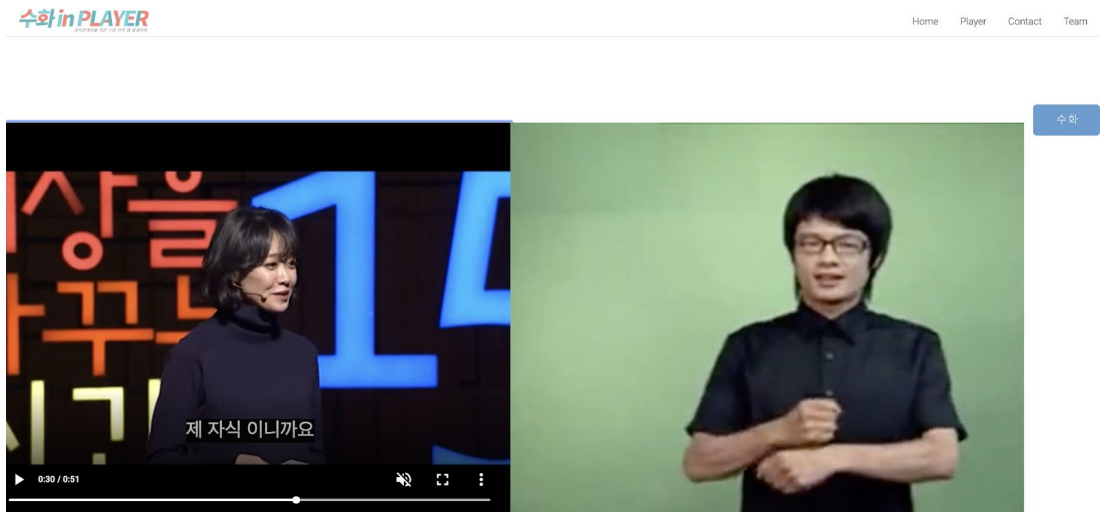
b. 사용자가 영상이 담겨있는 유튜브 URL 주소를 입력한다.



3. 영상, 자막, 수화영상을 함께 시청한다.

이때 자막과 수화 영상 서비스 중 원하는 서비스를 선택적으로 제공한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28



4. 피드백 게시판에 사용 후 후기를 작성한다. 추후 개선해야할 서비스와 기능을 위해 사용한다.

CONTACT US

오역이나 불편한 사항을 작성해주세요.
여러분의 의견이 큰 도움이 됩니다.

이름:


이메일:

휴대폰번호:

내용:


5.2 배포 가이드

수화 In 플레이어 시스템은 웹 서비스로 구현되어 있으므로, <http://13.124.80.174:8000/>로 접속하면 서비스 이용이 가능하다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

5.3 테스트 케이스

대분류	소분류	기능	테스트 방법	기대 결과	테스트 결과
Web Page	웹 페이지 소개	웹 페이지에 대한 이해도를 높인다.	웹 페이지에 들어갔을 때 잘 나오는지 확인한다.	웹 페이지 소개가 잘 되어있다.	성공
	영상 업로드	영상파일을 업로드한다.	웹 페이지에 영상 파일을 넣은 후 잘 작동하는지 확인한다.	영상파일을 넣었을 때 수화와 자막이 나온다.	성공
	URL 영상 로드	영상이 있는 URL을 입력한다.	웹 페이지에 URL을 입력후 잘 작동하는지 확인한다.	URL을 입력했을 때 수화와 자막이 나온다.	성공
	개선사항 피드백	사용 후 피드백을 남길 수 있다.	게시판에 들어가 피드백을 남겨본다.	게시판에 작성하면 DB에 저장된다.	성공
Server	Encoding	음성인식을 하기 위해서 wav형식으로 바꾼 뒤 음성만 추출한다.	동영상 파일을 넣었을 때 wav형식으로 바뀌는지 확인한다.	ex) mp4 형식이 wav형식으로 바뀌었다.	성공
	STT	Google Speech to Text API를 사용하여 음성을 텍스트로 변환한다.	음성파일을 Google stt API를 사용해서 소리를 txt파일로 기록한다. 이때 단어별 타임스탬프와 온점 기능을 추가하여 txt파일을 확인한다.	txt파일이 생성되고, 추가한 기능이 잘 작동된다.	성공
	Subtitle	음성인식한 텍스트파일을 사용하여 VTT 자막 파일로 변환한다.	음성인식하여 생성된 txt파일을 읽어서 vtt자막 파일로 생성되는지 확인한다. 그리고 생성된 vtt파일이 동영상에 자막으로 재생되는지 확인한다.	자막파일이 생성되고, 동영상에 자막이 재생된다.	성공
	Translate	영어 영상일 경우 Google Translation API를 사용하여	생성된 자막파일에서 영어문장을 Google Translation API를 사용해서 영어를 한국어로 번역이 잘 되는지 확인한다.	영어문장이 한국어로 번역되어 출력된다.	성공

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	수화 In 플레이어	
	팀 명	전지적 수화시점	
	Confidential Restricted	Version 1.5	2019-5-28

		한국어로 번역해준다.			
	수화 문법으로 재구성	한국어를 수화 문법 체계로 재구성하기 위해 형태소 분석을 수행하고 재배치한다.	재구성된 문장을 국립국어원의 수어 사전과 전북교육연구원의 수화 교실을 참고하여 알맞게 재배치 되었는지 확인한다.	한국어가 수화 문법 체계로 재구성되어 출력된다.	성공
	동음이의어 처리	단어의 동음이의어 처리하여 알맞은 단어를 선택한다.	동음이의어가 나올 경우 문장속 단어와 DB 단어 품사를 비교하여 선택한다. 동일한 품사일 경우 단어간 유사도 검사를 통해 유사도가 더 높은 쪽의 단어를 선택되는걸 확인한다.	처음 동음이의어들의 품사를 비교하고 동일한 품사일 경우 단어 유사도 검사를 통해 단어를 선택한다.	성공
	데이터 매핑	쿼리문을 작성하여 DB안에 단어의 유무를 판단하고 영상의 위치를 반환한다.	문장속 단어를 DB 단어에서 검색하여 단어의 수화 영상 위치를 가져오는지 확인한다. 한 문장을 List 단위로 만들고, List들을 합쳐 2차원 List로 출력되는지 확인한다.	단어별 수화 위치를 가져오고 문장별로 2차원 List로 출력된다.	성공
	수화 영상 제작	데이터 매핑에서 반환 된 수화 영상 위치를 받고 영상의 길이 정보에 맞춰 조정하여 최종 수화번역 영상을 생성한다.	데이터 매핑의 결과로 수화 데이터 위치 들을 받았을때, 중간산출물들이 제대로 생성되고 있는지 확인한다. 최종적으로 하나의 수화 영상이 생성되는지, 수화 영상은 제대로 재생되는지 확인한다.	수화 영상이 생성되었고, 웹에서 정상적으로 재생된다.	성공