

심층학습_ Alable

OCR기술 기반 시각 장애인 도움 서비스

201210302 김혜주

201810792 이민호

201810795 이정곤

지도교수님: 양희경 교수님

Contents

01. 일정 계획

02. 필요성

03. 주제

04. 기술 설명

- 알고리즘 설명
- 시연 연상

05. 사업화 가능성

06. 한계점 및 발전 가능성

필요성



필요성



여러분은 과자 뭐 있는지 보이시나요?

Do you know ~~which brands of snacks are~~ in the shelves?

필요성



점자 한계

- 시각장애인을 위한 제품 점자 : 상세한 제품 설명 부족

ex) MAX, HITE, CASS 등 다양한 맥주가 존재 → 모두 **맥주**로 점자 표기

고성표, "삼푸인지 화장품인지...시각장애인에게 '구별'을 허하라", 중앙선데이, 2019.10.26, <https://www.joongang.co.kr/article/23615660#home>



기존 시각장애인을 위한 앱의 문제점

- 튜토리얼 실행시 **대체텍스트(음성)**으로 지원 예정
- 기능을 **음성**으로 **선택**하는 기능 제공

주제

**시각장애인을 위한
제품 문자 식별 서비스**

기술 설명



Tesseract 5.0 & TTS

Tesseract 5.0 Tesseract OCR



Tesseract 란?

LSTM 기반 인식 엔진으로 구현된 OCR 오픈 소스 라이브러리



OCR 이란?

Optimal Character Recognition 의 줄임말로,
사람이 직접 쓰거나 이미지 속에 있는 **문자**를 **이미지 스캔**으로 얻은 다음,
이를 **컴퓨터가 인식**할 수 있도록 문자를 **디지털화** 하는 기술.



LSTM (Long Short-Term Memory)

RNN의 한 종류. 단일 문자가 포함된 이미지 인식에 사용되는 **CNN**은 일반적으로 일련의 문자로 이루어진 **긴 텍스트의 인식에 부적합**.

이러한 **장기 의존성(long-term dependency)** 문제 해결을 위해 탄생한 것이 **LSTM**.

Sepp Hochreiter, Jürgen Schmidhuber, "Long Short-Term Memory", Neural Computation (1997) 9 (8): 1735-1780.

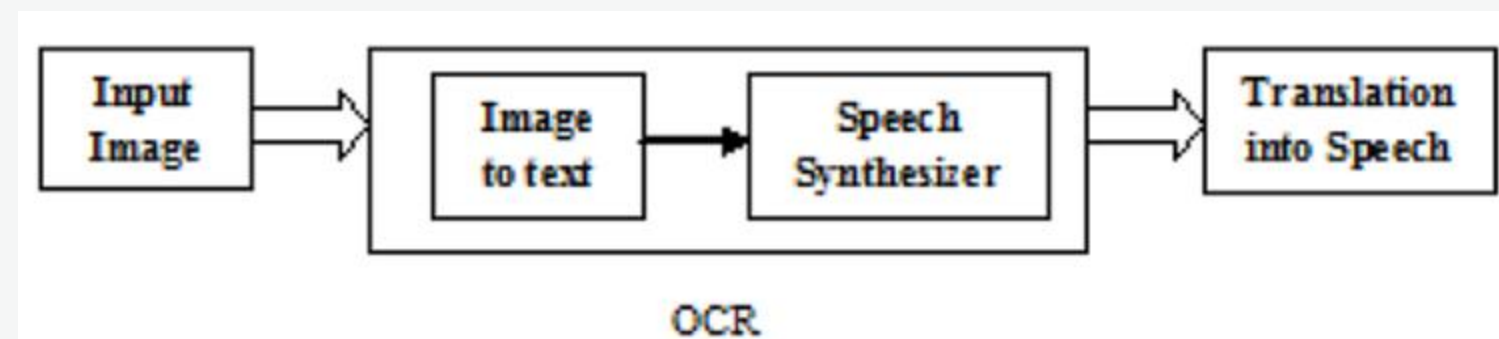


Fig.1.1 Block Diagram of OCR

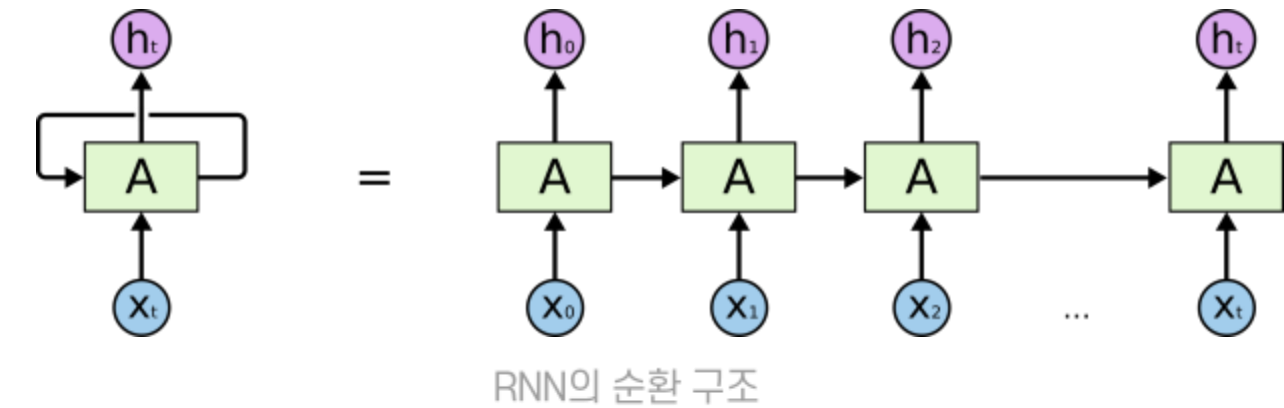
Ravina Mithe, Supriya Indalkar, Nilam Divekar, "Optical Character Recognition", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-2, Issue-1, March 2013

RNN (Recurrent Neural Network)



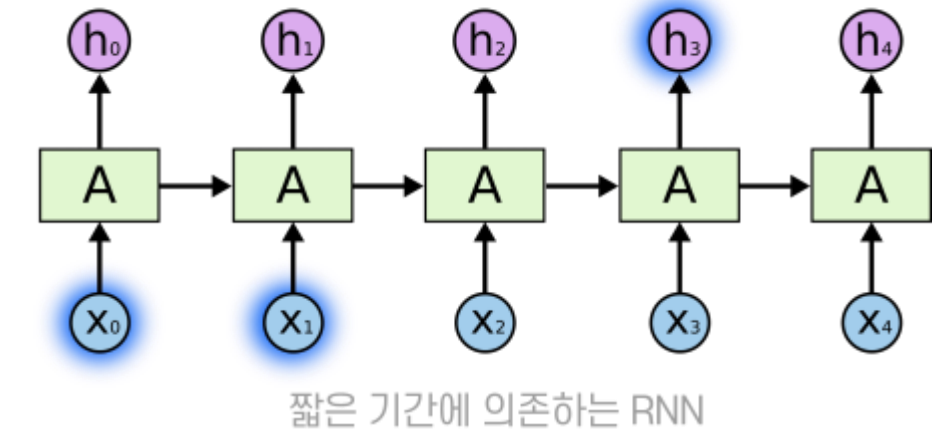
RNN이란?

Hidden state가 방향을 가진 Edge로 연결되어 순환구조를 이루는 인공신경망의 한 종류. 문장과 같은 'Sequence Data'를 분석하는데 유용하며, Input과 Output의 길이에 제약받지 않고 유연한 구조를 가질 수 있다는 점이 장점.



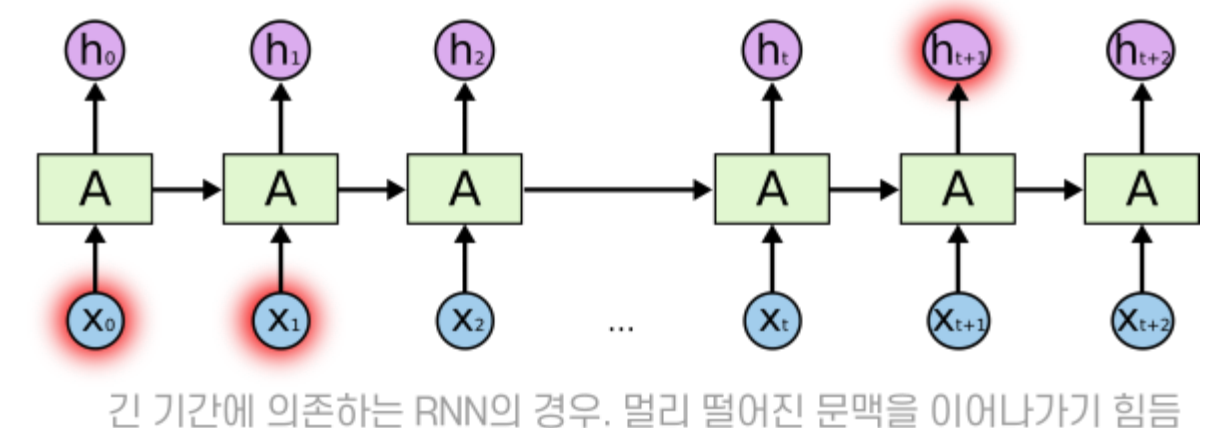
짧은 시간에 의존하는 RNN

필요한 정보를 얻기 위한 시간 격차가 크지 않다면 RNN도 지난 정보를 바탕으로 학습 가능. 하지만 더 많은 문맥을 필요로 하는 경우, 필요한 정보를 얻기 위한 시간 격차는 늘어나며 이 격차가 클수록 RNN은 학습하는 정보를 계속 이어나가기 힘들.



LSTM (Long Short-Term Memory)

RNN의 한 종류. 단일 문자가 포함된 이미지 인식에 사용되는 CNN은 일반적으로 일련의 문자로 이루어진 긴 텍스트의 인식에 부적합. 이러한 CNN과 RNN의 긴 기간 의존성(long-term dependency) 문제 해결을 위해 탄생한 것이 LSTM.

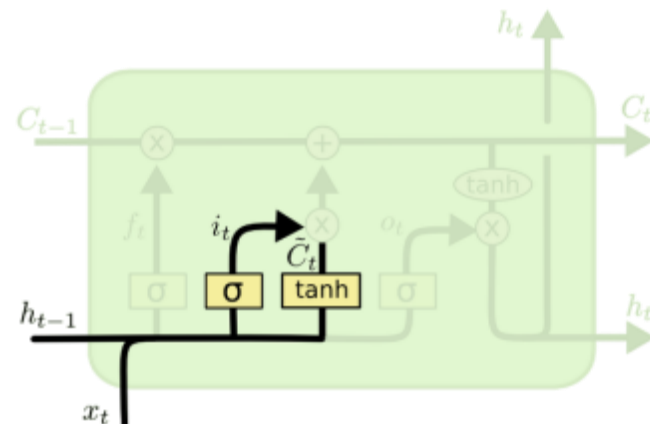
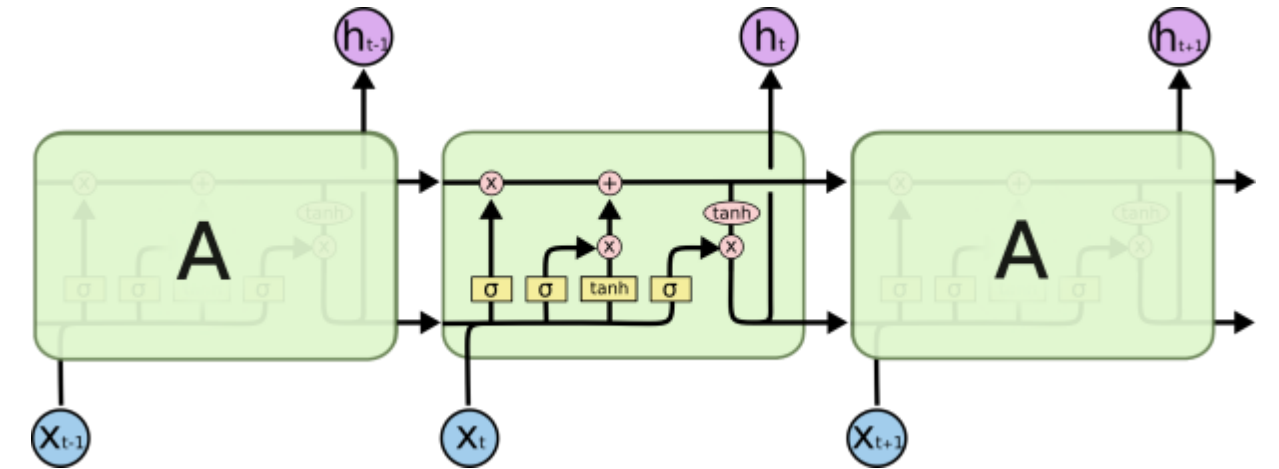


LSTM (Long Short-Term Memory)

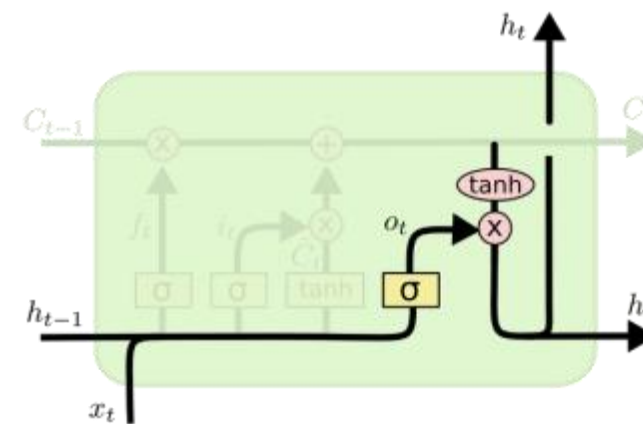


긴 의존 기간을 필요로 하는 학습을 수행할 능력을 가진 LSTM

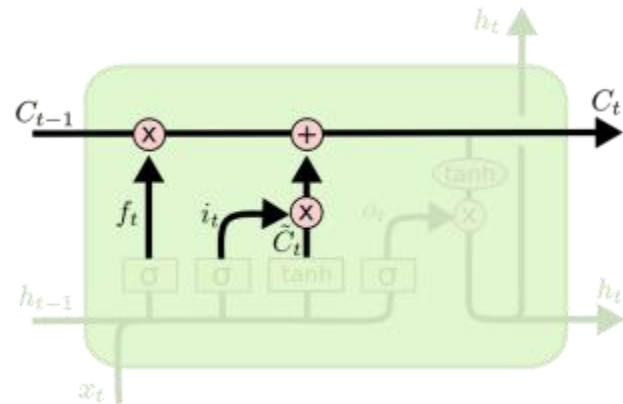
LSTM도 RNN과 같이 체인 구조를 가지고 있지만, 각 순환 모듈은 다른 구조를 가짐. 기본 LSTM 모델의 경우, 단순한 neural network layer 한 층 대신에 4개의 layer가 서로 정보를 주고 받는 구조.



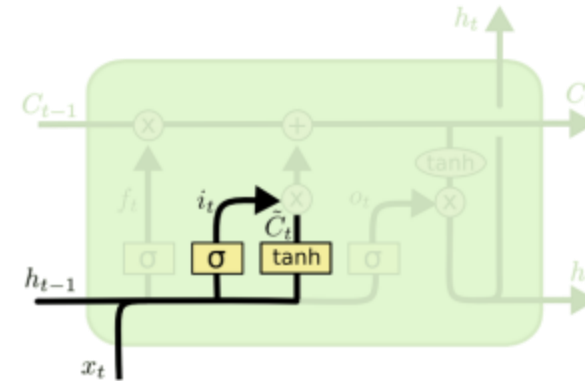
1. 어떤 정보를 버릴 것인지를 정하는 forget gate layer



2. 새로운 정보 중 어떤 것을 cell state에 저장할 것인지를 정하는 input gate layer



3. 과거 state를 업데이트해서 새로운 cell state를 만드는 cell state update layer



4. 무엇을 output으로 내보낼 지 정하는 output gate layer

TTS(Text to Speech)



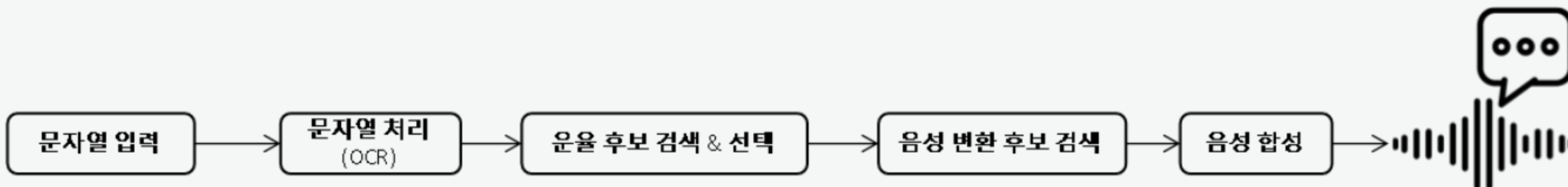
TTS 란?

음성 합성(音聲合成, speech synthesis)은 말소리의 음파를 기계가 자동으로 만들어 내는 기술



gTTS (Google-Text-to-Speech)

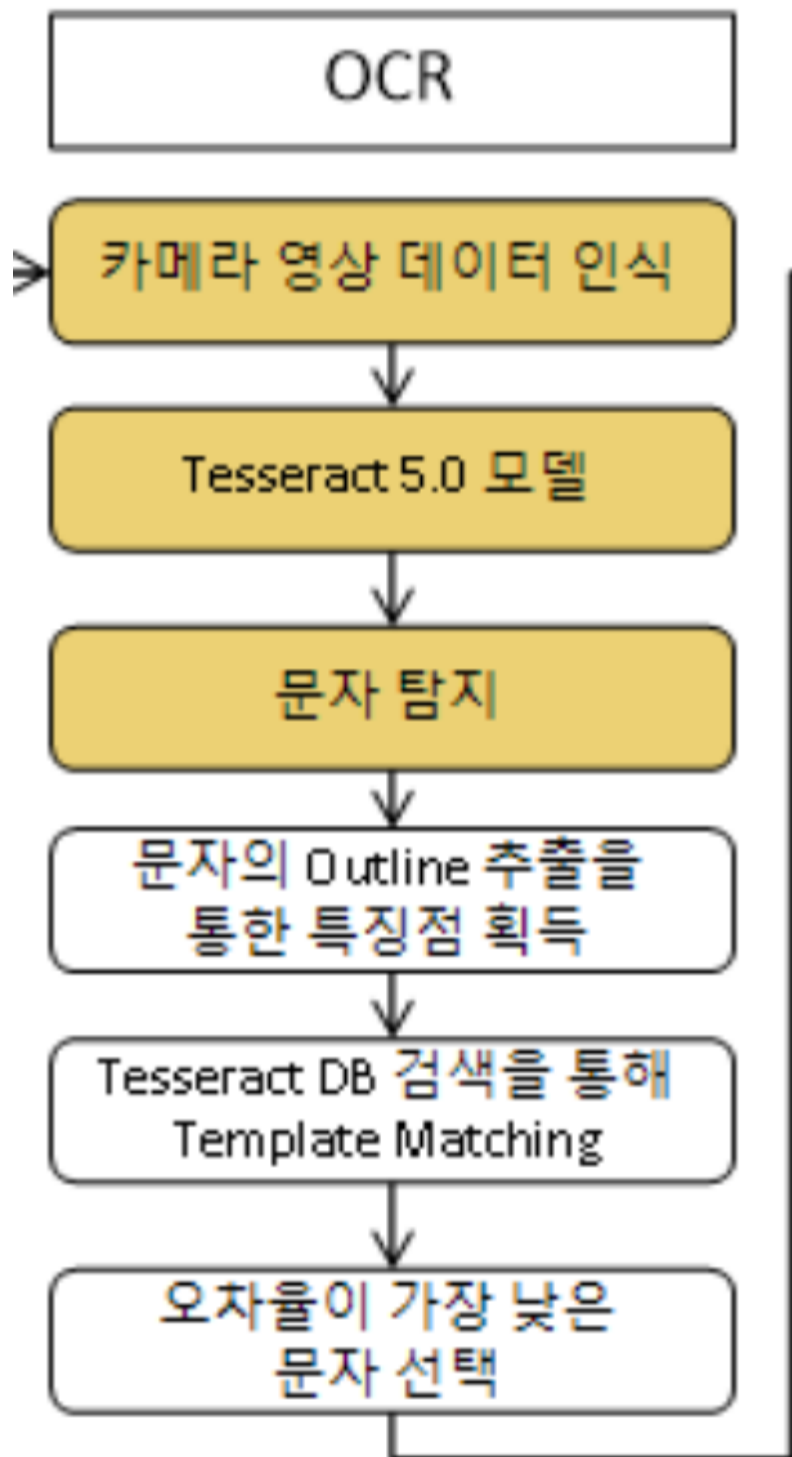
구글 텍스트 음성 변환(Google Text-to-Speech)은 안드로이드 운영 체제용으로 구글이 개발한 스크린 리더 애플리케이션이다. 애플리케이션에 화면 상의 텍스트를 읽는 기능을 제공하며 수많은 언어를 지원한다.



서비스 구현



서비스 구현



1. Preprocessing

들어온 이미지를 처리하기 위해 히스토그램 평활화, 이진화 작업을 통하여 보다 효율적인 출력물을 얻기 위한 전처리 작업.



Lenna



Histogram Equalization

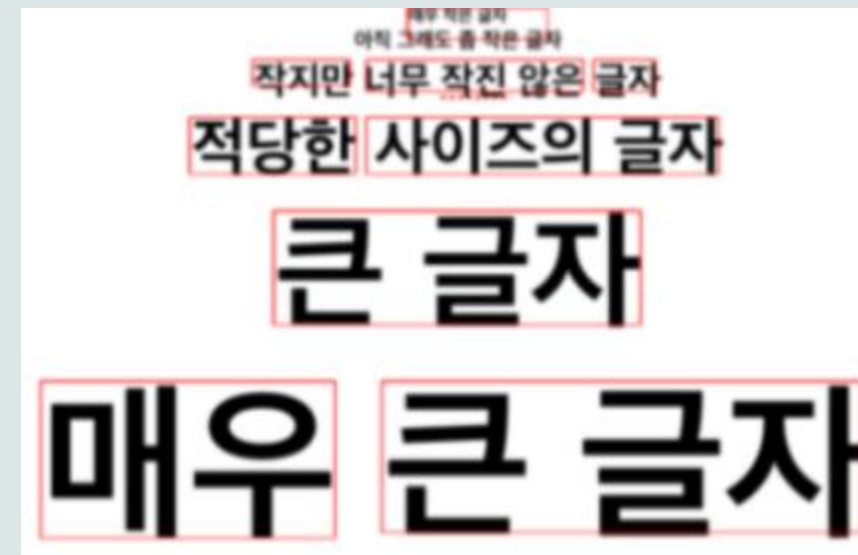


Binarization



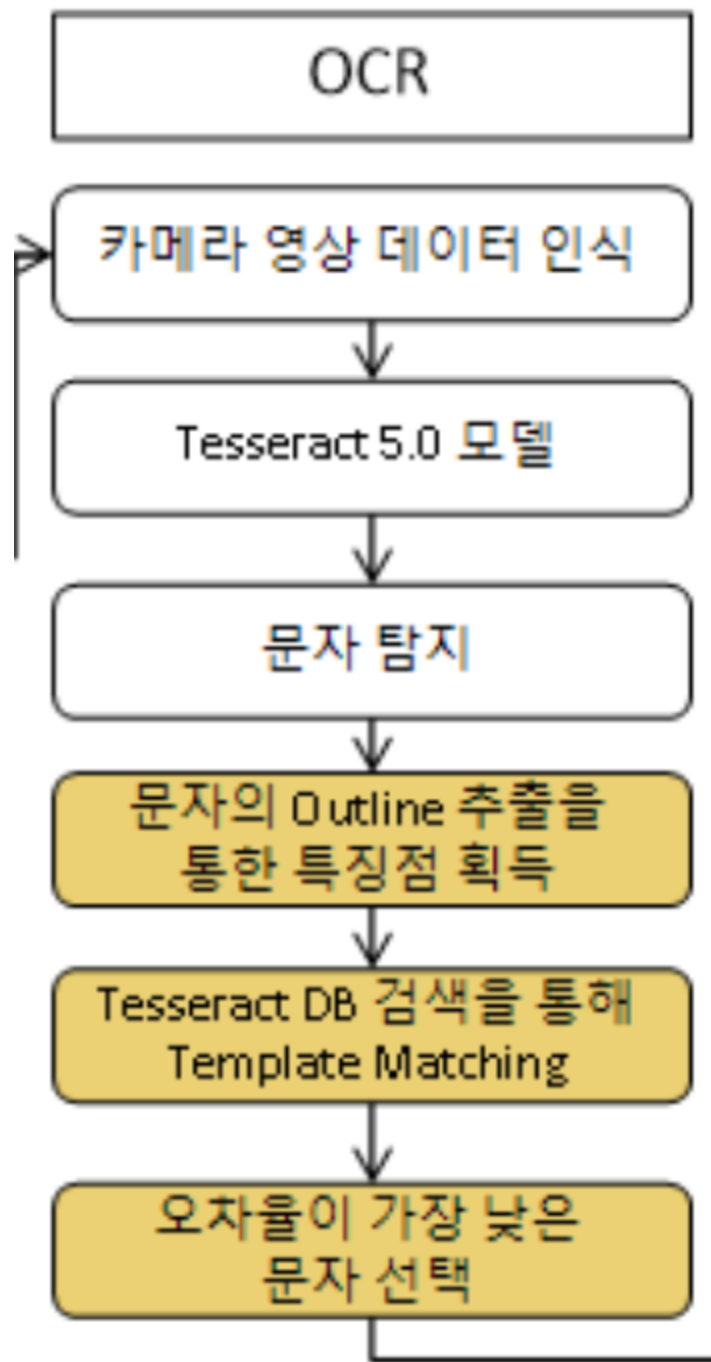
2. Segmentation

전처리 작업을 거친 뒤 이미지에서 글자 단위로 혼합된 Text를 분류하는 Segmantation 작업이 이루어짐.



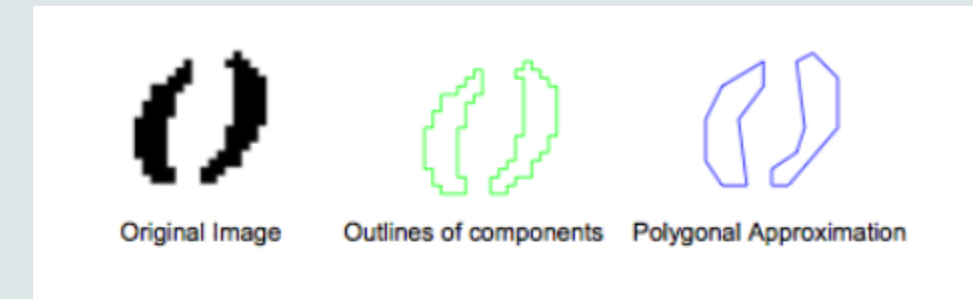
Segmentation-based Model

서비스 구현



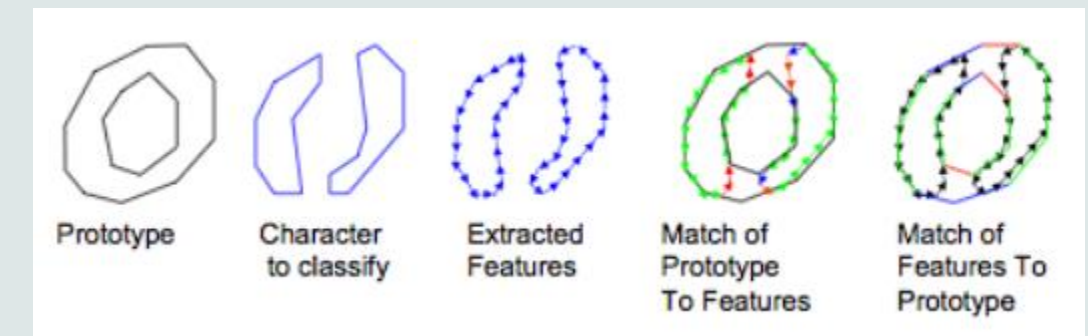
3. Outline 추출

원래의 이미지에 Outline을 생성.



4. 특징점 추출 및 Template Matching

문자의 특징점이 추출되면 Tesseract 데이터베이스 검색을 통하여 특징점이 비슷한 문자들과의 Template Matching 후 특징점과 원본 이미지의 오차율이 가장 낮은 문자를 선택.



시연 영상



Start

Snap Shot

사업화 가능성

1. 정부 및 지자체 공익사업 공모
2. 민간 재단 공익사업 공모
3. 기업의 공익사업

사업화 가능성

1. 정부 및 지자체 사업 공모

1. 정부 및 지자체 사업

시각장애인 원격생활안전 서비스

엔젤아이즈

‘엔젤아이즈’는 서울특별시가 시각장애인을 위해 만든 원격생활 안전 앱이다. 웨어러블 카메라와 앱을 통해 실시간 영상이 보호자나 서울시 도우미 등에게 전송돼 시각장애인을 도와준다. 보호자가 아니더라도 시각장애인을 돕고 싶은 사람은 누구나 ‘엔젤아이즈 도우미’ 앱에 자원봉사자로 등록할 수 있다. 도움 가능 시간을 설정해 '시각장애인의 눈'이 되어 줄 수 있다. 서울시 도우미 안내 시간은 오전 9시부터 오후 6시까지다.



(애플스토어, 구글플레이)

김병훈, " 장애인과 함께하는 '반려 앱' 31개, BLOTTER, 2018.04.19, <https://www.bloter.net/newsView/blt201804190001>

[사례1]

서울특별시

시각장애인을 위한 복지

시각장애인을 위한 앱

엔젤 아이즈 개발

사업화 가능성

1. 정부 및 지자체 사업 공모

나라장터
Korea ON-Line E-Procurement System

통합검색

검색

· 사이트맵 · 원격지원

나라장터
e-고객센터

나라장터
관련사이트

입찰정보

수요기관

조달업체

로그인

통합검색

입찰공고

개찰결과

최종낙찰자

발주계획

사전규격

계약진행현황

전자서고

☐ 전체

☐ 물품

☐ 공사

☐ 용역

☐ 리스

☐ 외자

☐ 비촉

☐ 기타

☒ 민간

☐ 연기

상세검색

☐ 입찰마감건 제외

입찰공고 검색 결과 (177건)

정확도순 | 공고일순 | 개찰일순

민간

[20200510293-00]

시각장애인의 책가뭉 해결을 위한 e북 제작 웹사이트 개발 구축

▼ 펼치기

일반

입찰마감 2020/05/15 10:00

|

공고 2020/05/11 10:46

|

개찰 2020/05/15 11:00

|

수요기관 아이티로시각장애...

민간

[20200440254-00]

시각장애인의 책가뭉 해결을 위한 e북 제작 웹사이트 개발 구축

▼ 펼치기

일반

입찰마감 2020/05/08 17:00

|

공고 2020/04/28 12:04

|

개찰 2020/05/08 18:00

|

수요기관 아이티로시각장애...

민간

[20190512486-00]

시각장애인 책가뭉해결을 위한 e북 클라우드 구축운영

▼ 펼치기

일반

입찰마감 2019/05/15 17:00

|

공고 2019/05/10 10:46

|

개찰 2019/05/16 18:00

|

수요기관 아이티로시각장애...

민간

[20190447091-00]

시각장애인 책가뭉해결을 위한 e북 클라우드 구축운영

▼ 펼치기

일반

입찰마감 2019/05/09 17:00

|

공고 2019/04/29 12:04

|

개찰 2019/05/09 18:00

|

수요기관 아이티로시각장애...

민간

[20180533118-00]

시각장애인의 책 가뭉 해결을 위한 e북 제작 웹사이트 ...

▼ 펼치기

긴급

입찰마감 2018/05/29 17:00

|

공고 2018/05/25 15:10

|

개찰 2018/05/29 18:00

|

수요기관 아이티로시각장애...

민간

[20180516540-01]

시각장애인의 책 가뭉 해결을 위한 e북 제작 웹사이트 ...

▼ 펼치기

변경

입찰마감 2018/05/23 17:00

|

공고 2018/05/14 16:48

|

개찰 2018/05/23 18:00

|

수요기관 아이티로시각장애...

민간

[20170715097-00]

목포시 시각장애인 생활이동지원센터 차량 구입

▼ 펼치기

일반

입찰마감 2017/07/21 17:00

|

공고 2017/07/13 16:33

|

개찰 2017/07/21 18:00

|

수요기관 목포시 시각장애...

민간

[20170704272-00]

목포시 시각장애인 생활이동지원센터 차량 구입

▼ 펼치기

일반

입찰마감 2017/07/13 15:00

|

공고 2017/07/05 14:11

|

개찰 2017/07/13 16:00

|

수요기관 목포시 시각장애...

민간

[20150722751-00]

점자프린터 구입을 통한 시각장애인 정보접근성 강화 사업

▼ 펼치기

일반

입찰마감 2015/07/28 10:00

|

공고 2015/07/24 12:31

|

개찰 2015/07/28 11:00

|

수요기관 (사)한국장애인소...

나라장터
국가종합전자조달

입찰 방식으로 사업 진행

사업화 가능성

1. 정부 및 지자체 사업 공모



정부 및 지자체 공모
(공모시기 1월~3월 | 사업추진 3월~12월)
자세한 정보 <https://www.snpo.kr/> 사이트에서 확인



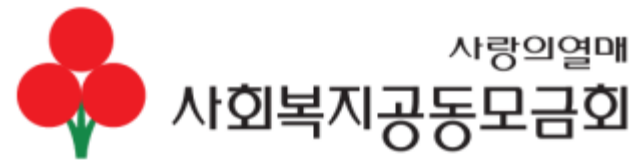
비영리법인/민간 단체

사업화 가능성

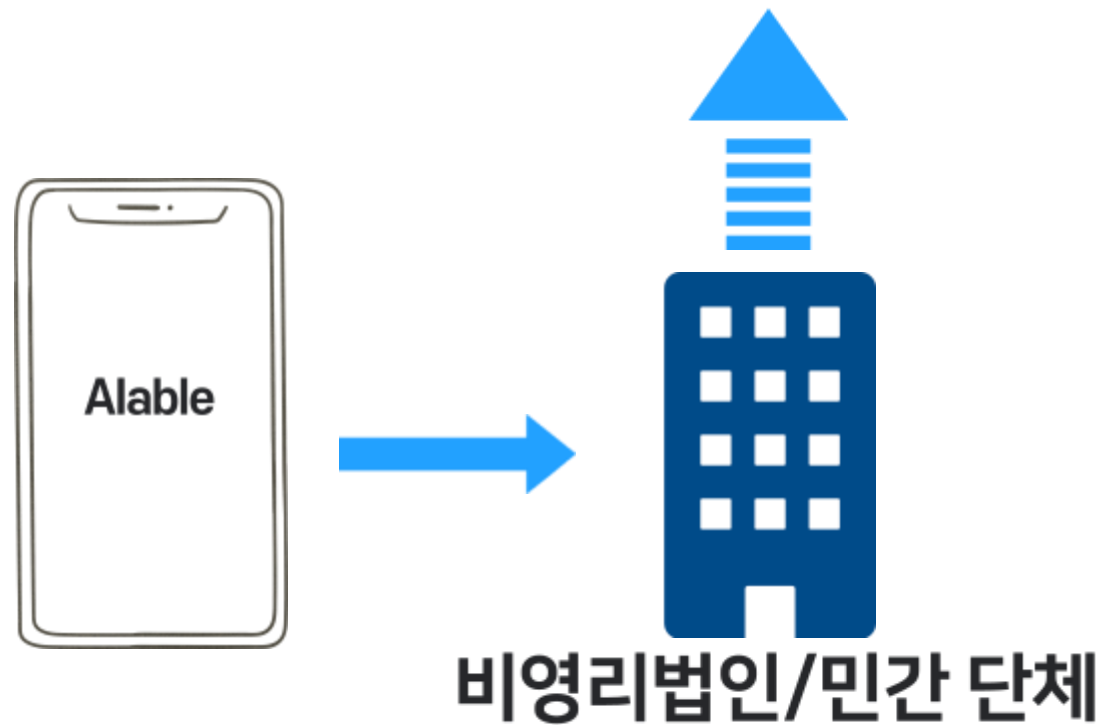
2. 민간재단 공익사업 공모



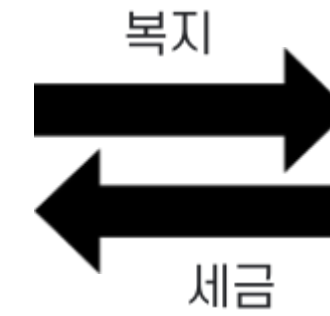
민간재단 공모
(공모시기 기관별 사이 | 사업추진 3월~12월, 차년도 1~12월)
자세한 정보 <https://www.snpo.kr/> 사이트에서 확인



아산사회복지재단
Asan Foundation



사업화 가능성



정부 및 지자체 사업의 한계

사업 공고가 안 날 수도 있음



기업 **공익** 사업

사업화 가능성

3. 기업의 공익사업

3. 기업의 공익 사업



사회적 공헌을 통한
기업의 이미지 개선



시각 장애인

한계점

- 어플리케이션으로 미구현 된 상태
- 제품 글자 인식률 저하
->학습 데이터(제품 글자)의 부족

발전 가능성

- 제품 사진 dataset을 충분히 축적 후 학습을 통한 제품 글자 인식률 향상
- 어플리케이션 구현 시, 기존 어플리케이션과 다르게 음성 인식 리모트 컨트롤 기능 추가
 - > 직접 사용하는 주체인 시각장애인을 위한 기능 구현

감사합니다