**<오픈소스SW프로젝트 계획서 양식>**

**1. 과제(또는 작품) 명**

- 딥러닝을 이용한 이미지-텍스트 변환 알고리즘 및 구현

**2. 선정 동기, 목적, 필요성**

- 선정 동기 :

\* 미래 기술 사업에 있어 딥러닝과 AI 두 가지는 큰 부분을 차지[1]

\* 기술 발전 동향에 따라 최신 기술을 사용해보는 것에 초점을 둠

\* 무인 운전 자동차, CCTV에서 사용되는 딥러닝과 영상처리 기술을 사용하기로 함

- 목적 :

\* 사용자로 하여금 이미지 파일에서 원하는 글자를 텍스트 형식으로 바로 추출이 가능하게 함.

\* 현재 판매되거나 무료로 배포되고 있는 프로그램에 비해 높은 인식률로 보다 정확한 정보 추출이 가능하게 하는 것을 목표로 함

- 필요성 :

\* 시중에 사용되는 대부분 사용되는 도큐멘트 스캔 프로그램은 이미지 파일을 PDF 파일로 변환해줌.

\* 시중에 PDF파일의 내용을 Word문서나 Hwp문서로 변환하는 별도의 프로그램이 존재하나 인식률이 별로 좋지 않거나 이미지파일을 바로 변환해주지는 못함

\* 해당 기술을 사용할 경우 이미지 파일에서 바로 텍스트 추출이 가능하며, 문서 이외의 일반 사진에서도 배경이나 사물을 제외한 글자 부분만 텍스트로 변환 시켜 줌.

\* 이후 해당 기술을 플랫폼간 이동이 자유롭게 설계하여 모바일 어플리케이션이나 웹 어플리케이션으로도 활용가능 할 것

**3. 과제(작품) 관련 배경조사**

3.1 과제 관련 제품(보고서) 및 활동 현황

- 이미지에서 PDF 변환계열

\* Piksoft Inc : TurboScan



- 스마트폰 어플리케이션으로 제작된 도큐멘트 스캐너

- 스마트폰에서 촬영한 이미지 혹은 별도에 준비된 이미지 파일을 PDF파일로 변환해줌

- 촬영한 이미지의 가장자리를 자동으로 결정해주며 사용자가 원하는 부분만 파일로 변환하는 것도 가능함

\* Adobe사 : Acrobat Reader Pro



- 전 세계 많은 사람들이 사용하는 Acrobat Reader의 Pro버전

- 유료 버전 답게 변환 대상이 되는 이미지 파일의 사용자 정의 글꼴을 생성하여 해당 글꼴로 PDF파일을 생성하며 PDF파일을 자유롭게 Word나 Excel 등의 파일로 변환이 가능함

- 월정액으로 정기 구독 해야지만 해당 기능들을 사용 할수 있음

- 문서변환 계열

\* ReTIA사 : ROSE 문서인식(OCR) 서비스

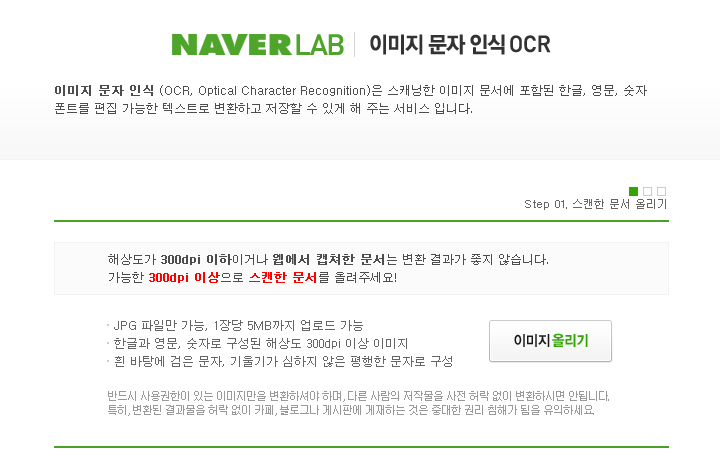


- PDF 리더를 전문으로 제작하는 회사

- 무료 시험판을 배포하고 있으나 보통 기업용 소프트웨어 제작 회사이기 때문에 일반사용자가 구매해서 사용하기에는 부담이 있음

- 별도로 온라인 서비스를 통해 PDF파일을 워드나 엑셀파일로 변환해주지만 무료 서비스는 1회 변환시 최대 3페이지이며 하루 10회 한정으로 사용이 가능함

\* 네이버렙 : 이미지 문자 인식(ORC)



- 네이버랩에서 무료로 제공하는 OCR서비스

- 무료로 제공되는 탓인지 jpg파일만 입력가능 하고 변환 포맷은 txt 포맷만 저장이 가능함 또한 해당 이미지의 해상도에 따라 결과값이 크게 변동되어 정확한 값을 얻기 힘듬

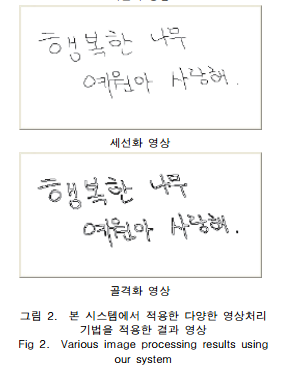
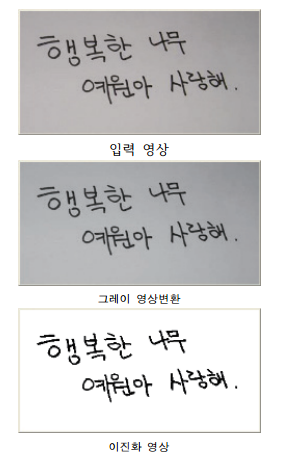
- 2014년 11월 18일부로 네이버랩 서비스가 종료됨에 따라 사용이 불가능하게 됨

3.2 과제와 기존 제품(보고서) 및 활동의 비교 분석

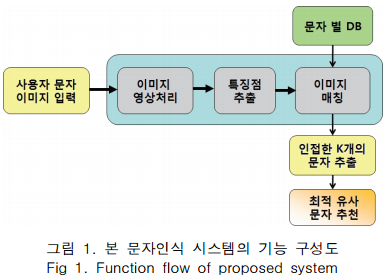
- 기존 제품들과 같이 이미지를 가공하여 텍스트를 얻는 OCR(광학적 문자인식)을 구현한다는 점에서 관련성이 있음

- Acrobat Reader가 이미지 전체를 PDF파일로 만드는 대에 비해 해당 프로그램은 이미지 내에서 글자 인 부분만 추출하여 수정가능한 텍스트 형식으로 변환하는 것에 초점을 둠

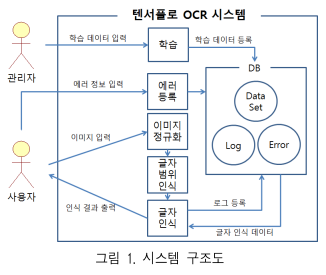
-상품 구매나 구독이 필요한 상용 소프트웨어들과 달리 누구나 무료로 사용 할 수 있도록 함.

**4. 과제(작품) 계획안**

[2]



- OpenCV의 라이브러리를 이용하여 사물 혹은 배경과 문자인 부분을 검출하기 위해 흑백 영상변환, 이진화, 세션화, 골격화 과정을 거쳐 딥러닝을 이용한 학습 시 글자인 부분을 글자로서 구분 할 수 있도록 함[2]

 - 해당 학습에는 오픈소스인 텐서플로우를 이용하여 이미지 처리에 있어 효율적인 CNN(합성곱 신경망)을 이용(CNN은 생물의 시각처리과정을 모방하여 패턴의 크기, 위치가 바뀌어도 인식 할 수 있음[3])

- 학습모델을 구성하고 반복적 학습을 통해 문자인 부분과 문자가 아닌 부분에 대해 학습 데이터를 구축함

- 딥러닝을 통해 문자인 부분을 특정 지을 수 있다면 해당 부분에 OpenCV 라이브러리와 MNIST 데이터 베이스를 이용하여 해당 부분이 어떤 글자인지 판독하도록 함

- 판독 결과를 텍스트 형식으로 배출

[4]

**5. 과제(또는 작품) 해결 방안 및 과정**

5.1 과제의 해결 방안, 과정 등 설명

- 논문과 전문서적 등을 참고하여 과제에 필요한 정보들을 수집하여 정리하며 파이썬을 기반으로 딥러닝 알고리즘 작성, MySQL과의 연동을 통해 학습 데이터를 저장하고 이를 반복하여 문자 인식 오류확률을 줄여나가도록 함

5.2 팀원의 역할

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 연번 | 소 속  학과(전공) | 학번 | 학년 | 성명 | 역 할 | 비고 |
| 1 | 컴퓨터공학과 | 2014244115 | 3 | 백승한 | 영상처리 및 딥러닝 | 조장 |
| 2 | 컴퓨터공학과 | 2014244119 | 3 | 서준덕 | DB구성 및 관리 |  |
| 3 | 컴퓨터공학과 | 2014244095 | 3 | 서덕진 | 프로그램 패키징 및 딥러닝 |  |

5.3 팀 활동 방안

- 자신이 맡은 부분과 관련된 참고문헌의 수집 및 정리, 개발 상황을 github를 통해 공유하며 주 3회 이상 오프라인 회의를 통해 현재까지의 개발 상황을 토의하고 향후 개발에 대한 아이디어를 공유, 방향성을 정해나가며 자기 평가를 실시함.

**6. 참고문헌 (필수)**

* [1] 세계적인 AI 연구진 및 전문가들이 예측한 2018년 주요 인공지능 트렌드 13, 2017.12.20, <http://blogs.nvidia.co.kr/2017/12/20/ai-headed-2018/>
* [2] 이원용 (2017). 영상처리 기법을 이용한 문자인식 시스템 개발. 융복합지식학회논문지,, 2017/07, p. 3-5
* [3] 안성만, 딥러닝 모형과 응용사례, 지능정보연구, 2016.06, p. 6-7
* [4] 윤병일, 최재성, 김병만, 이해연 (2017). 텐서플로를 이용한 OCR 시스템 개발. Proceedings of KIIT Summer Conference, 2017.06, p.3