

EAST 모델과 OCR을 이용한 실시간 문자 탐지 시스템

최예준^o, 문미경^{*}

^o동서대학교 소프트웨어학과,

^{*}동서대학교 소프트웨어학과

e-mail: dichieasy1212@naver.com^o, mkmoon@dongseo.ac.kr^{*}

Real-time Character Detection System Using EAST Model and OCR

Ye-Jun Choi^o, Mikyeong Moon^{*}

^oDept. of Software, Dongseo University,

^{*}Dept. of Software, Dongseo University

● 요약 ●

웹페이지나 디지털 문서에는 특정 단어나 특정 문구를 검색하는 기능이 있다. 인쇄된 도서나 참고서 등과 같은 인쇄물에는 실시간으로 특정 단어나 특정 문구를 찾는 기능이 없어 어려움을 겪는 경우가 많다. 본 논문에서는 텍스트를 감지(Detection)하는 EAST 모델과 텍스트를 인식(Recognition)하는 EasyOCR을 활용한 실시간 문자 탐지 시스템의 개발내용에 대해 기술한다. 이 시스템을 통해 사용자는 인쇄물에서 실시간으로 원하는 단어나 문구를 찾아 필요한 정보를 빠르게 읽는 것에 효과적일 것을 기대한다.

키워드: 텍스트(text), 감지(Detection), 인식(Recognition), OCR, EAST

I. Introduction

온라인에서는 웹페이지나 디지털 문서 등에서 사용자가 찾고 싶은 특정 단어나 특정 문구를 검색하면 찾을 수 있는 기능이 있다. 그러나, 오프라인에서는 인쇄된 도서, 참고서 등과 같은 인쇄물에서 실시간으로 특정 단어나 특정 문구를 찾는 기능이 없어 불편함이 있다. 특정 단어나 특정 문구를 검색하기 위해 OCR 기술을 사용하고자 한다. OCR은 문서 전체를 읽어 문자를 인식하는 기술이며 문서 처리 분야, 법률 서비스, 교육 분야 등 다양한 분야에서 활용되고 있다. 본 논문에서는 도서나 참고서 등과 같은 인쇄물에서 카메라를 통해 실시간으로 특정 단어나 특정 문구를 쉽고 빠르게 찾을 수 있는 시스템의 구현 내용을 기술한다.

II. Preliminaries

OCR은 문서 전체를 읽어 문자를 인식하는 기술이다. 그러나 문서 전체를 읽어 문자를 인식하기 때문에 불필요한 영역까지 인식되는 단점이 존재한다. 이를 보완하기 위해 맞는 템플릿을 사용하거나 AI 학습데이터를 기반으로 이미지의 문자를 인식하는 규칙을 스스로 만들어 학습하여 필요한 영역만 인식해야 한다[1, 2]. 문서 전체가 아닌 필요한 영역만을 인식하기 위해 연구[2]는 책등 이미지를 찍어 YOLO를 이용하여 책제목 영역을 추출하여 OCR을 적용한 연구이다.

연구[3]은 YOLO와 EAST 모델이 깨끗한 문서의 문자 영역 탐지 성능이 비슷하다는 것을 보였다. 본 논문에서는 EAST 모델과 OCR을 같이 이용하여 카메라를 통해 실시간으로 오프라인 인쇄물의 문자를 인식하는 시스템을 구현한다.

III. The Proposed Scheme

본 연구에서는 인쇄물의 텍스트를 감지할 수 있는 EAST 모델과 인쇄물의 텍스트를 인식할 수 있는 EasyOCR을 활용하여 Python 언어로 개발한다. 텍스트 감지 모델은 일반적으로 파이프라인 안에 다수의 단계와 구성요소로 되어 있다. 이는 철저한 조정이 필요하며 성능이 최적화되지 않고 전체 파이프라인의 처리 시간 증가 등 악영향을 미치게 할 수 있다. EAST 모델은 불필요한 중간 구성요소와 단계를 줄여 연산 시간과 빠른 속도, 성능 등을 보여주어 문자 탐지 분야에 주목을 받고 있다[4]. EasyOCR은 Naver Clova AI의 CRAFT, VGG, ResNet을 기반으로 개발되어 80개 이상의 언어를 지원하고 있으며 손쉽게 수행할 수 있는 Python 패키지이다[5].

3-1. 텍스트 영역 추출

Python 라이브러리 EasyOCR과 OpenCV를 설치해 개발 환경을 구축한다. EAST모델은 사전 훈련된 EAST 모델의 pb파일을 불러와 사용한다. 사용자가 검색하고 싶은 단어를 입력하면 카메라를 이용해 실시간으로 화면의 프레임을 가져와 사용한다. 프레임은 EAST 모델의 인식률을 높이기 위해 크기를 320x320으로 줄이고 filter2D를 사용하여 선명하게 한다. 그리고 가우시안블러(GaussianBlur)를 사용하여 노이즈를 제거한 뒤, normalize를 사용하여 프레임의 대비를 높이는 순서로 전처리하여 진행한다. 전처리가 된 프레임은 EAST 모델을 통해 텍스트의 단어마다 x, y의 좌표를 받아 영역을 생성한다.

3-2. 텍스트 영역 인식

생성된 영역들은 EasyOCR의 인식률을 높이기 위해 색상을 회색으로 변경하고 노이즈를 제거하는 전처리 과정을 거친다. 사전에 입력받은 사용자 입력 단어와 EasyOCR을 통한 텍스트 영역의 단어들을 비교한다.

3-3. 개발 결과

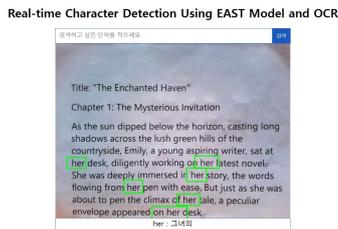


Fig. 1. result

EasyOCR을 통해 인식한 텍스트 영역의 단어 중 사용자가 원하는 단어가 있으면, 프레임에 해당 텍스트 영역을 표시하여 송출하게 한다. Fig 1.은 본 연구의 결과화면으로 사용자가 'her' 단어를 입력하여 5개의 해당 텍스트 영역을 추출하여 시각적으로 표시하여 보여주고 있는 화면이다.

IV. Conclusions

본 논문에서는 도서나 참고서 등과 같은 인쇄물에서 카메라를 통해 실시간으로 특정 단어나 특정 문구를 쉽고 빠르게 찾을 수 있는 시스템을 구현했다. 향후 연구에서는 손글씨 등과 같은 다양한 텍스트를 신속하고 정확하게 감지하고 인식할 수 있도록 시스템을 보완 확장할 계획이다.

ACKNOWLEDGEMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01817)

REFERENCES

- [1] W.J. Kim, "A Study on the Prediction for the OCR Technology Development Trajectory based on the Patent and Article Information", Ph.D. thesis of Korea National University of Technology and Education, Graduate School of General Studies.
- [2] S.Y. Kim et.al., "Multi-book Label Detection Model using Object Detection and OCR", Journal of Korean Institute of Information Technology. Vol. 21, No. 2, pp.1-8, 2023.
- [3] C.Y. Park et.al., "YOLO, EAST : Comparison of Scene Text Detection Performance, Using a Neural Network Model", KIPS Trans. Softw. and Data Eng, Vol.11, No.3 pp.115~124, 2022.
- [4] Xinyu Zhou et.al., "EAST: An Efficient and Accurate Scene Text Detector", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.5551-5560, 2017.
- [5] C.Y. Kim et.al., "A Study on the Survey and Demonstration Test of OCR based on Open Source for AI OCR", Proceedings of Symposium of the Korean Institute of communications and Information Sciences, pp.851-852, 2023.