|  |
| --- |
| 국문제목  OCR 성능 향상을 위한 이미지 전처리 기법 연구  **영문제목**  A Study on Image Preprocessing Techniques for Enhancing OCR Performance  **요 약**  본 연구는 OCR(Optical Character Recognition) 성능을 향상시키기 위한 다양한 이미지 전처리 기법을 제안하고 평가한다. 전처리 과정은 이미지의 절반 크롭, 수평 크롭, 텍스트 크롭을 포함하며, 각 단계에서의 OCR 결과를 비교 분석한다. 실험 결과, 제안된 전처리 기법이 OCR의 인식 정확도를 유의미하게 향상시킴을 확인하였다. |

**1. 서 론**

OCR(Optical Character Recognition, 광학 문자 인식) 기술은 이미지 또는 문서 스캔본에서 텍스트 정보를 자동으로 추출하는 기술로, 문서 디지털화, 자동 데이터 입력, 차량 번호판 인식, 시험지 채점 등 다양한 분야에 활용되고 있다. 특히 행정, 교육, 산업 분야에서 OCR 기술의 수요는 꾸준히 증가하고 있으며, 인공지능 기술의 발전에 따라 OCR의 정확도와 실시간 처리 성능 또한 빠르게 개선되고 있다.

그러나 OCR 기술은 입력 이미지의 품질과 구성에 크게 영향을 받는다. 예를 들어 이미지에 노이즈가 많거나, 글씨가 흐릿하거나, 배경이 복잡한 경우 OCR 성능은 급격히 저하될 수 있다. 특히, 손글씨가 포함된 시험지나 설문지 등은 인식 난이도가 높고, 전처리 없이 OCR을 적용할 경우 낮은 인식률을 보일 가능성이 크다.

이러한 문제를 해결하기 위해 이미지에 대한 전처리(preprocessing) 기법이 필요하다. 전처리는 OCR의 입력 이미지에서 불필요한 정보를 제거하고, 텍스트가 포함된 영역을 명확히 하여 인식률을 높이는 역할을 한다. 전통적으로는 이진화, 노이즈 제거, 기울기 보정 등의 기법이 사용되었지만, 최근에는 딥러닝 기반의 영역 탐지 및 분할 기술을 통해 더욱 정밀한 전처리가 가능해지고 있다.

본 연구에서는 대학 시험지 스캔 이미지를 대상으로, OCR 성능을 향상시키기 위한 전처리 기법들을 제안하고 그 효과를 실험적으로 검증한다. 전처리 기법은 크게 세 가지로 구성되며, YOLOv10 기반 탐지를 활용한 절반 크롭(Half Cropped), 수평 크롭(Horizontally Cropped), 텍스트 크롭(Text Crop) 방식이 포함된다. 또한, ClovaAI에서 제공하는 OCR 학습 코드를 기반으로, MNIST 데이터셋으로 훈련한 OCR 모델을 EasyOCR 플랫폼에서 적용하였다.

실험은 다음 네 가지 조건으로 구성되며, 각각의 OCR 결과를 비교 분석하였다: (1) 원본 이미지 + 기본 OCR 모델, (2) 전처리 이미지 + 기본 OCR 모델, (3) 원본 이미지 + 훈련된 OCR 모델, (4) 전처리 이미지 + 훈련된 OCR 모델. 평가 지표로는 정확도(accuracy)와 문자 오류율(CER: Character Error Rate)을 사용하였다.

본 연구는 OCR 성능을 높이기 위한 실질적인 전처리 전략을 제시하며, 손글씨 및 복잡한 구조의 이미지에 OCR을 적용하고자 하는 다양한 분야에 기여할 수 있을 것으로 기대된다.

2. 전처리 기법

본 연구에서는 세 가지 주요 전처리 기법을 통해 OCR 인식 성능의 향상을 도모하였다. 각 전처리 기법은 YOLOv10 모델을 기반으로 한 탐지와 OpenCV를 활용한 이미지 처리 기법으로 구현되었으며, 전체적인 흐름은 다음과 같다: 원본 이미지 → 절반 크롭 → 수평 크롭 → 텍스트 크롭.

**(Figure 1)**

**1) 절반 크롭 (Half Cropped)**

원본 시험지 이미지에는 상단에 시험 정보, 중단에 문제-답안 표, 하단에 공백이 포함되어 있다. 본 연구에서는 YOLOv10을 이용하여 "question number"와 "answer"가 위치한 영역을 탐지하고, 이 부분만을 절반 크롭 방식으로 추출하였다. 이를 통해 시험지의 텍스트 정보가 있는 주요 부분만을 남기고, 배경이나 불필요한 정보는 제거되었다. 절반 크롭은 전체 이미지 크기를 줄이면서도 텍스트 영역을 명확히 하여 OCR 처리 속도와 정확도를 동시에 향상시키는 데 기여하였다.

(Figure 2)

**2) 수평 크롭 (Horizontally Cropped)**

절반 크롭된 이미지에는 문제 번호와 학생 답안이 여러 행으로 구성되어 있다. 수평 크롭은 각 행을 독립적으로 분리하여 OCR이 텍스트를 보다 효과적으로 인식할 수 있도록 하기 위한 기법이다. 이미지의 수평 라인을 기준으로 각 행의 y 좌표를 탐지한 후, 행 단위로 이미지를 자르는 방식으로 구현하였다. 수평 크롭을 통해 줄 간 간섭을 최소화하고, 행별 인식 정확도를 높일 수 있었다.

(Figure 3)

**3) 텍스트 크롭 (Text Crop)**

수평 크롭 이후의 이미지에는 각 행마다 문제 번호와 해당 답안이 병렬로 존재한다. 이들을 개별적으로 인식하기 위해 텍스트 크롭 기법을 적용하였다. YOLOv10 기반 탐지를 통해 각 숫자 블록을 탐지한 후, 해당 영역만을 잘라 OCR 입력으로 활용하였다. 특히 문제 번호와 답안이 근접하게 붙어 있는 경우, 정밀한 탐지를 통해 정확한 분할이 이루어지도록 하였으며, 이로 인해 배경 노이즈 제거와 가독성 향상 효과가 있었다.

이 세 가지 전처리 기법은 단계적으로 적용되며, 각 단계에서의 인식 결과는 실험을 통해 정량적으로 비교되었다. 다음 장에서는 이를 바탕으로 한 실험 조건 및 결과를 제시한다.

3. 실험

3.1. 실험 조건 정리

OCR 인식 성능을 분석하기 위해, 본 연구에서는 이미지 전처리 여부와 OCR 모델의 훈련 여부에 따라 총 4가지 실험 조건을 설정하였다. 아래 표는 실험 조건을 조합별로 정리한 것이다:

A screenshot of a phone

AI-generated content may be incorrect.

전처리 여부: 전처리 기법(절반 크롭, 수평 크롭, 텍스트 크롭)을 적용한 이미지인지 여부

OCR 모델: ClovaAI 코드 기반으로 MNIST 데이터셋으로 학습된 커스텀 모델과 EasyOCR의 기본 모델 비교

이 네 가지 조건에 대해 동일한 시험지 이미지를 대상으로 OCR을 수행하였으며, 각 조건별로 인식 정확도(accuracy)와 문자 오류율(CER)을 계산하였다. 다음 절에서는 이에 대한 실험 결과와 비교 분석을 다룬다.

**[학술대회 논문작성시 유의사항]**

1. 논문 페이지 수

- 참고문헌, 부록 포함하여 2~3쪽

2. 용지 및 여백처리

- 용지: A4 세로방향

- 여백: 위 쪽 30mm, 아래 쪽 20mm,

왼 쪽 10mm, 오른 쪽 10mm

**3. 논문구성**

- ① ~ ③ 항목은 1 단(column)으로 구성

① 국문제목

② 영문제목

③ 요약

- ④ ~ ⑦ 항목은 2 단으로 구성

④ 본문

- 장 및 절에 해당되는 번호는 아라비아 숫자 로

각각 1., 1.1 등과 같이 표기

- 그림의 명칭은 하단에, 표는 상단에 각각

그림 1 및 표 1로 표기

⑤ 참고문헌

- 본문 중에 참고문헌 번호를 쓰고, 그 문헌을 참고문헌란에 인용한 순서대로 기술.

- 기술 순서는 저자, 제목, 학술지명, 권, 호, 쪽수, 발행연도 순으로 작성.

⑥ 부록(해당사항이 있는 경우만 작성)

⑦ 학부생/주니어 논문은 필요시 시연동영상 URL

기재 가능.

**4. 기 타**

- 작성언어는 국문 또는 영문 중에 선택 가능

- 작성폰트는 임의사용 가능, 글자크기는 9pt 이상

- 파일은 PDF로 제출

* 사사문구는 추후 출판용 제출 시 기재

**- 심사용 논문 샘플**

[https://www.kiise.or.kr/conference02/data/sample1.pdf](http://www.kiise.or.kr/conference02/data/sample1.pdf)

**- 출판용 논문 샘플**

[https://www.kiise.or.kr/conference02/data/sample2.pdf](http://www.kiise.or.kr/conference02/data/sample2.pdf)