

제목 : 웹 검색엔진 및 딥러닝 기반 한글 단어 인식 OCR 시스템

1. 논문 주제 및 배경:

- 광학 문자 인식 (OCR) 기술을 활용하여 한글 단어를 인식하고 텍스트 데이터로 변환하는 시스템에 대한 연구
- 한글의 특성으로 인해 OCR 기술이 한국에서는 적극적으로 활용되지 않았으나, 본 연구에서는 한글 OCR의 정확도를 향상시키는 방법을 탐구했음

2. 주요 내용:

- OCR 정확도 향상:

한글의 언어적 특성을 고려하여 OCR 정확도를 향상시키는 방법을 연구함.

- CNN 기반 OCR 시스템:

합성곱 신경망 (CNN)을 활용하여 한글, 영어, 숫자를 훈련시키고, 복합적으로 구성된 단어에서 한글의 완성형 글자를 구분하고 인식함.

- 웹 검색 엔진 활용:

검색 엔진을 활용하여 텍스트 데이터를 검색하고 수정된 단어가 있는지 확인하여 인식 정확도를 향상시킴.

3. 적용 및 기대 효과:

- 한글 OCR 분야에서 효과적인 방법을 제시해 한국에서도 OCR을 통한 업무 효율성 증가를 기대할 수 있음.