

책장 이미지 기반 책 제목 인식 및 재배치 가이드 솔루션 제안

정윤우 · 안형욱 · 이국정 · 이예은 · 이지유*

한국폴리텍대학 성남캠퍼스

An Image-Based Solution for Book Title Recognition and Automatic Bookshelf Rearrangement

Yoon-wu Jung · Hyung-wook Ahn · Guk-jeong Lee · Ye-eun Lee · Ji-yoo Lee*

Korea Polytechnics Sungnam Campus

E-mail : ch830ok@gmail.com

요 약

본 논문은 이미지 기반 객체 인식과 OCR 기술을 활용하여 책장 이미지를 분석하고, 인식된 도서 정보를 바탕으로 도서의 재배치 가이드를 제공하는 솔루션을 제안한다. 기존의 OCR 기반 접근 방식이 책 제목이나 기본적인 메타데이터 추출에 그쳤다면, 본 논문에서는 객체 인식과 OCR로 추출한 도서 정보와 배열 정보를 연계하여 책 배열을 최적화하고, 사용자에게 2차원 재배치 가이드를 제공하는 데까지 확장하였다. 나아가 최종적으로는 도서가 재배치된 형태를 시각적으로 복원한 이미지를 생성하여 사용자에게 제공함으로써, 공간 활용과 정리 작업의 효율성을 동시에 향상시키는 것을 목표로 한다.

ABSTRACT

This paper proposes a solution that analyzes bookshelf images using image-based object detection and OCR technologies, and provides a visual guide for rearranging books based on the recognized information. While conventional OCR-based approaches have primarily focused on extracting book titles or basic meta data, this study extends the scope by integrating object detection results with OCR-extracted book and layout information to optimize the arrangement of books. Furthermore, the proposed solution generates a visual reconstruction of the rearranged bookshelf, ultimately aiming to improve both space utilization and organization efficiency.

키워드

Bookshelf Recognition, Object Detection, OCR, Automatic Arrangement , Space Optimization

I. 서론

현대인의 서재는 여전히 중요한 지식 자원으로 남아있으나, 도서의 종류와 수량, 배열 상태를 한눈에 파악하기 어려운 문제가 지속되고 있다[1]. 이러한 문제는 개인 서재뿐만 아니라 공공 도서관에서도 사용자의 도서 검색

및 관리자의 도서 정리의 불편함으로 이어진다[2].

이와 같은 배경에서, 본 연구는 이미지 기반 객체 인식과 OCR, 2D Packing 기술을 결합하여, 사용자가 촬영한 책장 사진에서 도서 정보를 추출하고, 이를 바탕으로 도서 정보 파악 및 재배치 정돈 가이드를 제공하는 솔루션을 제안한다.

* iamjiyoo@kopo.ac.kr

이를 통해 사용자는 배열 상태를 직관적으로 인식하고, 도서 정리의 반복적인 부담을 줄이고, 체계적인 관리가 가능해질 것으로 기대된다.

II. 관련 연구

기존 도서 관리 방식은 RFID나 바코드 기반으로 각 도서에 태그를 부착하고 전용 리더기를 사용해야 하여 초기 구축 비용이 높고 유지 관리가 번거롭다[3][4]. 이러한 문제점을 극복하기 위해 최근에는 YOLO, Faster R-CNN 등 딥러닝 기반 객체 인식 기법이 활용되어 책장 전체 이미지를 분석, 도서의 위치와 크기를 자동으로 인식하는 연구가 진행되고 있다[5][6].

또한, 도서 제목 자동 추출을 위해 책등 영역을 ROI로 설정하고 EasyOCR, Tesseract 등을 이용한 OCR 기술이 적용되었었다[7][8]. 특히, 다양한 조명, 배경 조건에서는 EasyOCR이 Tesseract보다 더 높은 텍스트 인식 정확도를 보인다[9]. 아울러, Transformer 기반의 TrOCR[10] 및 SRN[11]과 같은 모델을 통해 한글·영문 혼용 환경에서도 높은 인식률을 확보하려는 시도가 이루어지고 있다.

반면, 기존의 도서 인식 및 OCR 연구는 대부분 도서 객체 검출과 책등 텍스트 추출에 초점을 맞추었으나, 이러한 연구들은 책장의 책을 재배치하는, 공간 배치 최적화 문제와는 직접적으로 연결되지 않았다.

본 연구는 도서 인식 및 제목 추출과 재배치 최적화 연구의 결합을 통해, 분리되어 있던 두 축의 기술을 통합하는 새로운 응용 가능성을 제시한다.

III. 제안 솔루션 구성

본 논문은 이미지 기반 객체 인식과 OCR 기술을 활용하여 책장 이미지를 분석하고, 인식된 도서 정보를 바탕으로 도서의 재배치 가이드를 제공하는 확장형 솔루션을 제안한다.

이러한 구조는 OCR결과와 공간배치 방식에서 객체인식정보를 활용하지만 서로 독립적으

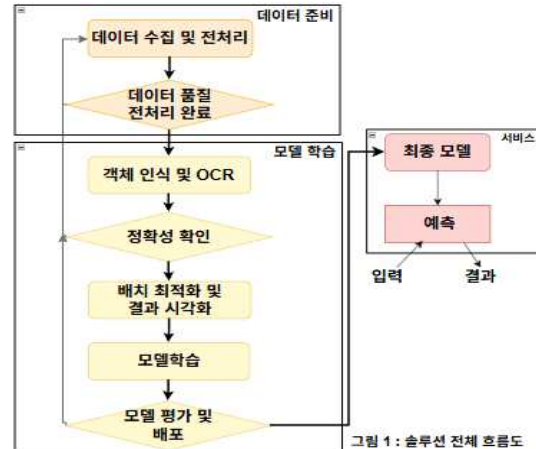


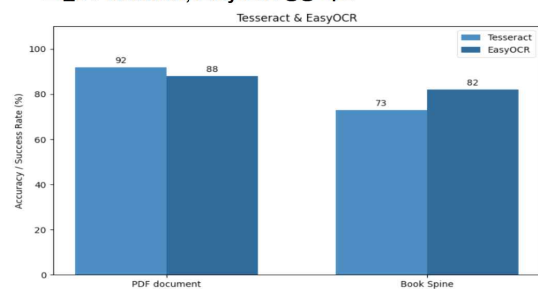
그림 1: 솔루션 전체 흐름도

로 처리하던 기존 접근에서, 각 단계에서 얻은 데이터를 통합적으로 활용하여 사용자의 공간을 재구성하는 정보흐름을 구현한다. 전체 솔루션의 흐름은 그림1과 같으며, 책장 이미지 입력부터 도서 객체 탐지, 제목 추출, 정렬 기준에 따른 재배치, 결과 시각화까지의 일련의 절차로 구성된다.

3-1 OCR 성능비교

Tesseract와 EasyOCR의 성능을 비교하기 위해, 정적인 문서인 PDF 문서 인식 정확도와 책등 영역에서의 텍스트 인식 성공률을 각각 평가하였다. 그 결과는 아래 그림2와 같다. 이러한 결과를 고려하여, 본 연구에서는 EasyOCR을 채택하여 솔루션을 구성할 예정이다.

그림2 : Tesseract, EasyOCR 성능 비교



3-2 유전알고리즘 기반 도서 재배치 가이드

OCR과 객체 인식 과정을 통해 확보된 도서의 위치 및 크기 정보를 기반으로, 책장 공간을 최적으로 배치하는 가이드를 생성한다. 본 연구에서는 이 문제를 2D Packing 문제로

정의하고 유전알고리즘 기반 최적화 기법을 통해 효율적인 재배치 구성을 탐색하고자 한다. 재배치 과정은 객체 인식 단계에서 각 도서의 경계 박스 좌표와 크기 정보를 추출하고, OCR을 통해 인식한 메타데이터를 통합하여 활용한다.

재배치 최적화의 적합도(Fitness)는 공간 활용률(Utilization)과 안정성(Alignment)의 가중합으로 정의하며 적합도 함수는 아래 수식1과 같이 나타낼 수 있다.

수식1

$$Fitness = \alpha \times Utilization + \beta \times Alignment$$

$$Utilization = \frac{\text{도서 배치 면적의 합}}{\text{책장 전체 면적}}$$

$$Alignment = 1 - \frac{\sum | \text{개별책의 높이}_i - \text{높이평균} |}{N}$$

최적 적합도 조합을 탐색하기 위해 유전알고리즘을 적용하며, 실제 데이터셋을 통한 구현과 평가에 관해서는 향후 연구를 통해 구체화하고자 한다.

IV. 결론 및 향후 연구

본 논문에서는 객체 인식, OCR, 2DPacking 기술을 통합하여, 도서의 정보추출을 넘어 효율적인 도서 재배치 가이드를 제공하는 솔루션을 제안하였다. 기존의 OCR 기반 정보추출 방식을 공간 최적화와 결합하여, 도서의 위치 및 크기를 고려한 재배치 과정을 최적화하는 방향으로 확장하였다. 이 솔루션은 개인 사용자뿐만 아니라 도서관, 아카이브 등 대규모 도서 보관 공간에서도 자동화된 재배치 제안 기능을 통해 관리 효율성을 증대시킬 수 있을 것으로 기대한다.

향후 연구에서는 실제 데이터셋을 기반으로 최적화 성능을 평가하고, 다양한 책장 구조와 사용자 맞춤형 기준에 맞춘 재배치 알고리즘을 추가적으로 개발할 계획이다. 또한, 장기적인 도서 관리 시스템과의 연계를 통해 지속적인 관리와 최적화를 지원하는 방안을 모색할 예정이다.

References

- [1] Liu et al., Smart bookcase based on image recognition E3S Web of Conferences 233:04038 (2021).
- [2] 정연찬, 홍성균, 김신우. (2013). 도서관에서 길을 잃다: 도서관 책찾기의 문제점 탐색 및 대안의 제안. 한국 HCI 학회 학술대회.
- [3] Shuo Zhou, T.et al."Library On-Shelf Book Segmentation and Recognition Based on Deep Learning"
- [4] Anderson, P., & Carlson, A. (2012). Chapter 2: RFID Costs, Benefits, and ROI. Library Technology Reports, 46(4), 18-24
- [5] Redmon, J. et al. "You Only Look Once: Unified, Real-Time Object Detection", CVPR, 2016.
- [6] Smith, R."An Overview of the Tesseract OCR Engine", ICDAR, 2007.
- [7] Shuo Zhou, T.et al."Library On-Shelf Book Segmentation and Recognition Based on Deep Learning"
- [8] M. P. Nevetha, et al. "Automatic Book Spine Extraction and Recognition for Library Inventory Management", ACM Multimedia, 2015.
- [9] Vedhaviyassh, D.R., et al. "Comparative Analysis of EasyOCR and TesseractOCR for Automatic License Plate Recognition using Deep Learning Algorithm." IEEE ICECA 2022.
- [10] Minghao Li, Yiheng Xu, Lei Cui, et al., "TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models," arXiv preprint arXiv:2109.10282, 2021.
- [11] Bowen Cheng, Bin Xu, Zhongyu Wei, et al., "Semantic Reasoning Network for Scene Text Recognition," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 12313-12322, 2020.