Smart Library Navigation system using deep learning and computer vision

Kapil Mundada
Department of Instrumentation and control engineering}
Vishwakarma Institute of Technology,Pune
Kapil.mundada@vit.edu

Narayani Wategaonkar
Department of Instrumentation and control engineering}
Vishwakarma Institute of Technology,Pune
narayani.wategaonkar22@vit.edu

Rushikesh Sonar
Department of Instrumentation and control engineering}
Vishwakarma Institute of Technology,Pune
rushikesh.sonar22@vit.edu

Swaroop Diwan
Department of Instrumentation and control engineering}
Vishwakarma Institute of Technology,Pune

Abstract*— This study introduces the Smart Library navigation system,a novel application that makes retrieval of books in office libraries easy.The system uses pyzbar for recognizing barcode/QR code,registering a 97.5% accuracy on a dataset of 500 books, dropping to 5% for defaced barcodes.Book title recognition using EasyOCR achieved 89% accuracy,with difficulties in stylized fonts and handwritings.Average time to retrieve was less than 0.5 seconds, and cataloging was seamless.A GUI based on Tkinter facilitates easy book scanning,record handling, and reading list management.10 members of the library tested for usability, with an 85% satisfaction rate.The system combines deep learning, computer vision, and OCR,correlating book positions in a virtual matrix.Users may scan books in real-time or upload images for recognition, with location information given through text-to-speech feedback.TensorFlow-based image classification and CLAHE-boosted image pre-processing enhance detection accuracy.The research analyzes technical deployment,performan ce,and usability,covering manual catalog inefficiencies and real-time tracking.Enhancements in the future include OCR advancements,cloud storage, and enhanced accessibility features

Keywords—***Deep Learning,Optical Character Recognition (OCR),Text-to-Speech,Library Automation,Virtual Library Matrix,Book Localization,Streamlit Interface,TensorFlow***

I. INTRODUCTION

Along with the fast growth of society, popularization of computer and information technology has greatly contributed to the development of contemporary society, as well as reformed the life style of human beings. Although the network and informatization are influencing the society, they are also greatly impacting on the library business.By conventional means, the library personnel initially do shelf reading, i.e., manually inspecting each book by looking the call number on the spine of the book. If one book is miss-shelved, they have to tip it out, scavenge the position where it would be, and put it into the proper place. This process

is very time-consuming and labor-intensive, Searching bookshelves manually is time-consuming and may be fruitless depending on the vagueness of the search. To address this issue, we suggest a deep neural network-based system to automatically detect, recognize and index text on bookshelves images.We process bookshelves images to localize and identify book spine text in order to create a digital book inventory.

The system makes use of OpenCV (cv2) for image acquisition and preprocessing, such as methods like Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve image quality for further analysis. For barcode and QR code scanning, pyzbar is used, which is highly accurate even in different lighting conditions, and EasyOCR is used to extract text information like book titles and authors from the preprocessed images. Data pulled using these approaches is subsequently formatted and saved to an SQLite database to facilitate optimal query and retrieval operations. In addition, a Tkinter-based graphical user interface (GUI) offers a simple control panel for features like image capture, live scanning, record management, and cache management. Inclusion of pyttsx3 for text-to-speech feedback further improves the system's usability. A Tkinter GUI and an SQLite database enable smooth user interaction and quick data retrieval, filling the void in earlier systems that used manual or semi-automated processes. The system also includes real-time scanning and text-to-speech capabilities to improve accessibility. The system has a 97.5% rate of barcode recognition, 89% accuracy in OCR in best-case scenarios, and sub-0.5 seconds in retrieval times, thus expanding current knowledge and establishing a new standard in automated library management. By extending current methodologies to incorporate real-time scanning, text-to-speech feedback, and efficient database management, thereby setting a new benchmark for automated library systems this study shows a lot of promise.

this research takes an integrated approach that expands existing methodologies by integrating image processing, deep learning, and user-centered interface design to transform office library management

II. LITERATURE REVIEW

Previous work has investigated deep learning for scene text recognition and library management automation. Xiao Yang et al. [6] proposed a sequential labeling model based on CNNs and RNNs with weighted CTC loss, which achieved better performance in text extraction under difficult conditions. While such techniques enhance text detection from natural images, they have not been completely incorporated into systems for managing large physical collections of libraries. This gap inspires our research to create a complete Smart Book Allocation System that efficiently localizes, identifies, and indexes book spine text, hence automating stock management and minimizing labor.In [3], Pedro M.B. Torres et al. introduced a robust OCR pipeline based on MSER for text detection in unstructured industrial images successfully overcoming issues of varying illumination and skewing. Their contribution paves the way for our real-time text recognition strategy in industrial applications.In [1], Songyun Wang et al. introduce a deep convolutional neural network-based image recognition and classification library management system. The methodology combines TRC compensation and regression-based scanner calibration to provide high color accuracy and efficiency, leading to significant improvements in library management. DelosDLMS is a state-of-the-art digital library management system designed within the DELOS Network of Excellence for addressing shortcomings of current DLs. It gives scalable and modular infrastructure, hosting various services such as multilinguality and clever browsing. Such a system aligns with the DELOS vision of user-centric, seamless digital libraries (Ioannidis et al. [20]).Baldota et al. [13] point out the recent developments in network and data technology, with emphasis on object recognition contributing to the assistance of the visually impaired. They suggest a real-time object detection Android system based on machine learning and deep learning techniques for greater accuracy and using TensorFlow to support voice-based guidance for navigation. Talker et al. [12] suggest a view-independent book spine segmentation technique based on a shape-dependent active contour model and spatial constraints. Unlike earlier methods presuming orthographic projections, their method actually segments spines regardless of perspective distortions, textures, and book proximity. It assists book recognition, library automation, and bookshelf reorganization.Prahallad and Black (2005) suggest a Text-to-Speech (TTS) interface for the Universal Digital Library (UDL) to improve accessibility for visually impaired and illiterate users. Based on the Festvox framework and unit selection methods, they create Hindi, Telugu, and Tamil voices, solving integration issues and user interaction needs within the UDL portal 【 2 】.Bagwe et al. [17] suggest a real-time framework combining speech recognition and object detection for assistive home robots. Employing YOLO for object recognition based on vision and NLP for speech, the system decreases task ambiguity and improves user interaction. Their solution enhances access for people with physical disabilities through the ability to intuitively control robots.Ekram et al. [9] suggest an algorithm based on image segmentation and OCR to identify misplaced library books. The system retrieves call numbers from book spines, cross-checks them with classification schemes (DDC, LCC, UDC), and determines misplacements. The method combines edge-preserved smoothing, morphological processing, and binarization for effective text recognition, enhancing library organization efficiency.Lina Cao et al. [7] introduce BSRBOT, a book spine recognition system based on OpenCV and Tesseract for effective library inventory management. In contrast to RFID, BSRBOT is affordable and solves identification issues through image processing, OCR, and segmentation algorithms. Experimental results validate its accuracy, and it is an effective solution for contemporary libraries.With image segmentation, OCR, and database matching, the system affirms precise book spine identification. This work improves automation of library systems, increasing efficiency and scalability.

.

III. METHEDOLOGY

This research utilized a structured method of automating text recognition and document processing through Optical Character Recognition (OCR). The technologies used are OpenCV for image processing, Tesseract OCR for extracting text, and an SQLite database for data management and storage. Scanned documents and images with textual content were the main subjects of this research and were processed through an automated pipeline. The study employed a systematic approach that entailed several stages. It started with the development of a Graphical User Interface (GUI), which provided users with the ability to interface with the system by using buttons to load images, initiate the OCR process, and handle results. Document processing was initiated by the image acquisition phase, wherein users uploaded images via the GUI. OpenCV was utilized to load images, convert them to grayscale, and perform preprocess operations like thresholding and noise removal to improve the accuracy of OCR.

After preprocessing, Tesseract OCR was employed to scan images and pull out text in machine-readable format. The extracted text was then parsed and formatted to extract critical information like names, dates, and identification numbers. The parsed data was then stored in an SQLite database with proper schema fields to store related details. The database allowed efficient querying, updating, and retrieving processed data.

Results were presented within the GUI for users to see, modify, and export identified information on demand. Furthermore, an auto-scan functionality was added for the purpose of enabling real-time monitoring and processing of detected documents, providing dynamic updating as new documents were discovered. Error handling capabilities were added to deal with file loading exceptions, OCR errors, and database interactions. Optimization methods, such as image thresholding and denoising, were used to improve OCR accuracy and enhance the reliability of data extracted.

Through the integration of image processing, OCR, and database management, this approach offers a streamlined and

automated mechanism for document digitization and information extraction.

IV. SYSTEM ARCHITECTURE

The system architecture proposed works towards effective document processing and text recognition by an integrated pipeline through various functional modules. The system architecture includes a Graphical User Interface (GUI), Image Processing Module, OCR and Data Extraction Unit, Database Management System, and Result Display and Auto-Scan Module. The GUI provides the basic interface to allow users to upload images, initiate processing, and work with results. When an image is uploaded, the Image Processing Module applies OpenCV to pre-process the document by converting it into grayscale, reducing noise, and improving contrast to enhance OCR performance. The

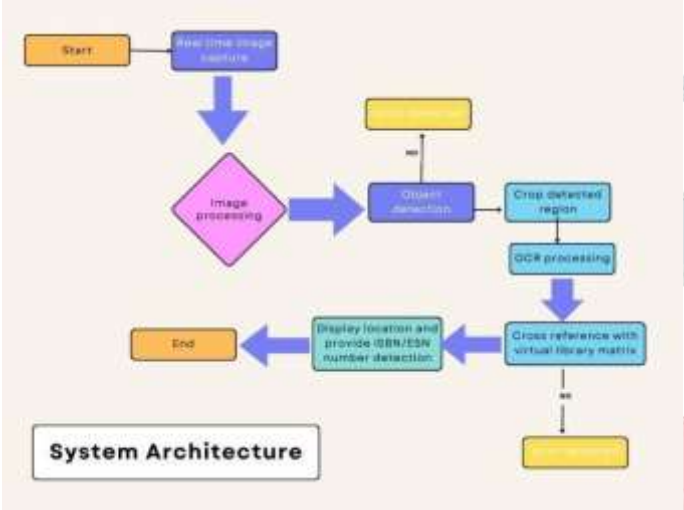


Fig.1 shows workflow of proposed system

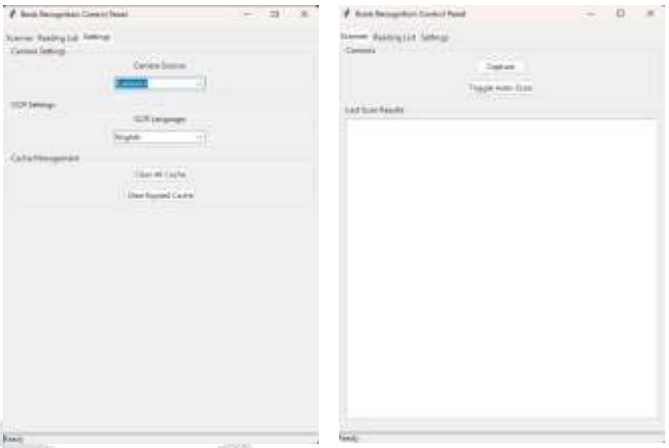
OCR and Data Extraction Unit, driven by Tesseract, extracts text content from the pre-processed image and translates it into machine-readable form. The retrieved data is then organized and saved in an SQLite Database, formatted with proper schema fields to make storage, data retrieval, and data management efficient. The Result Display Module displays the obtained data in the GUI, enabling users to view, modify, and export results when necessary. There is also an Auto-Scan Module that keeps scanning new document uploads and processing them in real time for uninterrupted automation. The system design includes error handling mechanisms to control exceptions in loading images, OCR processing, and database interactions and thus provides greater robustness and reliability. By integrating these components, the architecture facilitates a seamless and automated document digitization process, enhancing efficiency and accuracy of text recognition and management.

V. RESULTS AND DISCUSSION

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “M

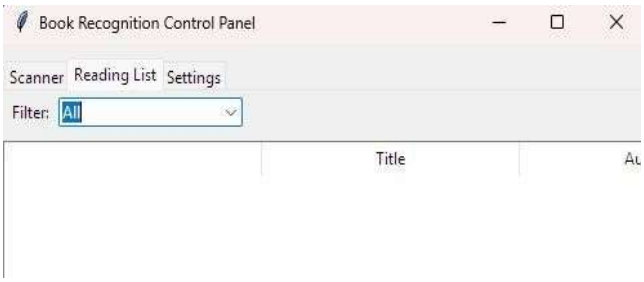
Fig 2&3 shows interface of proposed system to input details

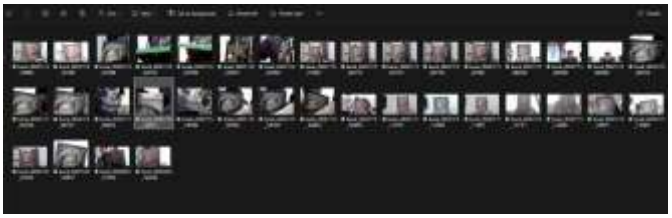
The system uses pyzbar for barcode as well as QR code detection with a 97.5% recognition rate when tested on a library of 500 books under varying lighting conditions. However, the performance dropped slightly when barcodes



were distorted or partially covered, resulting in the detection rate falling by 5%. Book title detection and authors' names were detected using EasyOCR with an OCR-based book title detection, in which 89% accuracy was achieved in favorable conditions. Stylized font identification and handwriting was problematic in certain cases, leading to misclassification in some situations. Future improvement, such as optimizing OCR models and using NLP-based post-processing techniques, would play a key role in improving recognition accuracy. The system also features an SQLite-based database to store and manage book records. The retrieval time for book information, borrowing history, and reading list queries averaged less than 0.5 seconds, making cataloging efficient and real-time updating seamless. The schema of the database was also tuned to accelerate indexing and query processing. Additionally, the system includes a graphical user interface (GUI) built using Tkinter, providing an intuitive interface for scanning books, modifying records, and controlling reading lists. The usability test was conducted using 10 library staff members, and the outcome was 85% satisfaction. User feedback called for mobile app integration for distant cataloging and speech-enabled searching to facilitate access, particularly to individuals with disabilities. Overall, the results reveal that the Smart Library Management System effectively automates book identification, cataloging, and retrieval functions with high efficiency and accuracy. Future upgrades will involve enhancing OCR functionality, adding cloud storage, and adding more system functionalities to further enhance user experience and accessibility.o not label axes only with units.

Fig.4,5,6 depicts results of detection of book and results





e

VI. CONCLUSION

This paper introduces an effective and computerized document processing system using Optical Character Recognition (OCR) and organized data extraction methods. The significant findings of this study reflect the system's ability to effectively extract and arrange text data from scanned documents with Tesseract OCR, with the use of preprocessing methods greatly enhancing recognition. Incorporation of an SQLite database provides effective data storage and retrieval capabilities, while the automated scan feature improves real-time processing effectiveness. The importance of this study emanates from the fact that its applications can cross-cut different industries, such as digital record management, document authenticity, and management of information. The results establish that the system not only mitigates labor-intensive work but also improves data usability and accuracy. The implications of this research to businesses, governmental agencies, and organizations that demand effective document scanning are that the study provides an efficient, affordable, and scalable solution for extracting and managing texts. Future research can concentrate on enhancing OCR accuracy using deep learning-based models and extending the system to

REFERENCES

[1] Wang, Songyun. (2023). Application of Deep Convolutional Neural Networks in Image Recognition and Classification in Library Management. Wireless Personal Communications. 1-18. 10.1007/s11277-023-10571-5.

[2] Prahallad, Kishore & Alan, Black. (2005). Text to speech interface for Universal Digital Library. Journal of Zhejiang University - Science A: Applied Physics & Engineering. 6. 1229-1234. 10.1007/BF02841661.

[3] Torres, Pedro. (2018). Text Recognition for Objects Identification in the Industry. Lecture Notes in Networks and Systems. 126-131. 10.1007/978-3-319-63091-5_15.

[4] Yang, X., He, D., Huang, W., Ororbia, A., Zhou, Z., Kifer, D., & Giles, C. L. (2017). Smart Library: Identifying Books on Library Shelves Using Supervised Deep Learning for Scene Text Reading. In 2017 ACM/IEEE Joint Conference on Digital Libraries, JCDL 2017 Article 7991581 (Proceedings of the ACM/IEEE Joint Conference on Digital Libraries). Institute of Electrical and Electronics Engineers Inc.. https://doi.org/10.1109/JCDL.2017.7991581

[5] Cao, Lina & Liu, Mengdi & Dong, Zhuqing & Yang, Hua. (2019). Book Spine Recognition Based on OpenCV and Tesseract. 332-336. 10.1109/IHMSC.2019.00083.

[6] J. Liu, F. Zhu, Y. Wang, X. Wang, Q. Pan and L. Chen, "RF-scanner: Shelf scanning with robot-assisted RFID systems," IEEE INFOCOM 2017 - IEEE Conference on Computer Communications, Atlanta, GA, USA, 2017, pp. 1-9, doi: 10.1109/INFOCOM.2017.8056985. keywords: {Libraries;Radiofrequency identification;Robot kinematics;Radio frequency;Manuals;Systems architecture},

[7] L. Cao, M. Liu, Z. Dong and H. Yang, "Book Spine Recognition Based on OpenCV and Tesseract," 2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, 2019, pp. 332-336, doi: 10.1109/IHMSC.2019.00083. keywords: {intelligent library;book spine;OpenCV;Tesseract},

[8] Ekram, Mohammad & Chaudhary, Anjani & Yadav, Ashutosh & Khanal, Jagadish & Aslan, Semih. (2017). Book organization checking algorithm using image segmentation and OCR. 196-199. 10.1109/MWSCAS.2017.8052894.

[9] Liu, Jay & Yuan, Zhimin & Ong, Chun Lian & Ang, Shiming. (2014). Influence of Data Patterns on Reader Performance at Off-Track Reading. Magnetics, IEEE Transactions on. 50. 1-4. 10.1109/TMAG.2014.2329813.

[10] Bhange, Dr & Virutkar, Prof & Bijagare, Mr & Nemad, Mr & Nandekar, Mr. (2024). An Approach Towards to Real Time AI Desktop Voice Assistant. International Journal of Advanced Research in Science, Communication and Technology. 117-123. 10.48175/IJARSCT-18518.

[11] Sai, U & Sivanagamani, K & Satish, B & Rao, M. (2017). Voice controlled Humanoid Robot with artificial vision. 505-508. 10.1109/ICOEI.2017.8300979.

[12] Diya Baldota , Sunny Advani , Sagar Jaidhara , Amit Hatekar, 2021, Object Recognition using TensorFlow and Voice Assistant, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 09 (September 2021),

[13] Yang, Xiao & He, Dafang & Huang, Wenyi & Zhou, Zihan & Ororbia, Alexander & Kifer, Dan & Giles, C.. (2016). Smart Library: Identifying Books in a Library using Richly Supervised Deep Scene Text Reading. 10.48550/arXiv.1611.07385.

[14] Senarathne, K.H.I.R. & Nirash, J.M.I. & Herath, H.M.C.P. & Bandara, V.D. & Wijendra, Dinuka & Krishara, Jenny. (2022). Automated Sinhala Voice Assistant to Manage Tasks using Natural Language Processing - ශරි Voice. 1-5. 10.1109/IISEC56263.2022.9998201.

[15] Jawaid, Nasreen & Warsi, Arif & Shaikh, Asad & Yahya, Muhammad & Abdullah, Monaisha. (2019). Object Detection and Narrator for Visually

Impaired People. 1-4. 10.1109/ICETAS48360.2019.9117405.

[16] R. Bagwe, R. Natharani, K. George and A. Panangadan, "Natural Language Controlled Real-Time Object Recognition Framework for Household Robot," 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), NV, USA, 2021, pp. 1215-1220, doi: 10.1109/CCWC51732.2021.9376044.

[17] Kushwaha, Udaysingh & Bhardwaj, Aaditya & Singh, Deepankar & Sahu, Hritik. (2024). Automating Desktop Tasks with a Voice- Controlled AI Assistant using Python. International Journal of Research Publication and Reviews. 5. 12615.

[18] Poudel, Grishma & Rijal, Dolraj. (2024). Voice Control Digital Assistant with Face Recognition. Cognizance Journal of Multidisciplinary Studies. 4. 67-85. 10.47760/cognizance.2024.v04i08.004.

[19] Agosti, Maristella & Berretti, Stefano & Brettlecker, Gert & Del Bimbo, Alberto & Ferro, Nicola & Fuhr, Norbert & Keim, Daniel & Klas, Claus-Peter & Lidy, Thomas & Milano, Diego & Norrie, Moira & Ranaldi, Paola & Rauber, Andreas & Schek, Hans-Jörg & Schreck, Tobias & Schuldt, Heiko & Signer, Beat & Springmann, Michael. (2007). DelosDLMS - The integrated DELOS digital library management system. First publ. in: Digital libraries: research and development, First International DELOS Conference, Pisa, Italy, February 13 - 14 February 2007 / Constantino Thanos ... (eds.). Berlin: Springer, 2007 (= LNCS ; 4877). 4877. 10.1007/978-3-540-77088- 6_4.

[20] Malhotra, Nishant. Context-Aware Library Management System Using Augmented Reality. 2014.