



Designing A Scanner Application Using Optical Character Recognition (OCR) Algorithm

Huliman^{1*}, Octara Pribadi², Akhiruddinal Hamsyi^{3*}

^{1,2,3}Informatics Engineering, STMIK TIME Medan, Indonesia

huliman.win@gmail.com^{1*}, octarapribadi@gmail.com², akhiruddinnalhamsyi@gmail.com^{3*}

Abstract

In an increasingly developed digital era, the need for efficient and accurate document management is rising. One of the solutions that can be used is a scanner application based on Optical Character Recognition (OCR) algorithms, which can convert text in images to editable digital data. This research aims to design a scanner application using OCR with an Artificial Intelligence (AI) approach to improve the accuracy of text extraction results. This application is built using React Native as the interface framework, Supabase as the backend, and AI integration through OpenAI. Testing shows that the application works well, with a success rate of 95% without any fatal errors. Meanwhile, 5% of the tests passed with minor notes related to the quality of the input image. Research prove that this application can be a practical solution in supporting document digitization, especially in campus environments and other institutions. With user-friendly features and reliable performance, this application has great potential for widespread use.

Keywords: Scanner, OCR, React Native, AI.

1. Introduction

The use of technology has brought extensive and profound impacts on human life. The rapid development of technology has made significant contributions to the efficiency of various human activities, ranging from commercial activities, education, to professional endeavors. Scanning applications are one of the manifestations of technological advancement that provide substantial benefits. Nowadays, everything is electronic-based, including documents. Many people feel reluctant to carry physical items due to the convenience offered by this technological advancement. As this has become common, scanning applications are widely used by the public.

In research conducted in 2020, by comparing the similarity levels between the original text present in the document and the extracted text, which resulted in an accuracy of an average similarity of 92.568%.[1] The subsequent research conducted in 2020 showed that the system was able to extract text based on English characters from images with complex backgrounds, achieving a word-level accuracy of 69.7% and a character-level accuracy of 81.9%.[2]

Based on the description that has been presented, a research will be conducted with the title " Designing A Scanner Application Using Optical Character Recognition (OCR) Algorithm".

2. Literatur Review

OCR is an application used to extract text from images, such as scanned documents or photos, and convert it into a machine-readable text format. The use of this character recognition system can enhance the adaptability and intelligence of computers. This system is very helpful in the ongoing digitalization initiatives of information and knowledge, for example in the creation of digital library collections and collections of ancient literature.

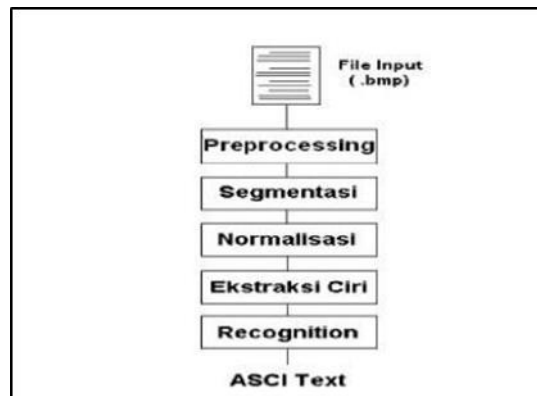


Fig. 1: OCR process in general

1. File Input

File input is in the form of a digital image file with the format *.bmp or *.jpg.

2. Preprocessing

Preprocessing is a process to remove unnecessary parts of the input image for the subsequent processes.

3. Segmentation

Segmentation is the process of separating the observation areas (regions) in each detected character.

4. Normalization

Normalization is the process of changing the dimensions of the regions of each character and the thickness of the character.

5. Feature Extraction

Feature extraction is the process of taking specific features from the observed character.

6. Recognition

Recognition is the process of comparing the features of the detected characters with the characteristic data of the characters in the database.[1]

Flowchart is a visual representation in the form of a diagram or graph that illustrates the steps and sequences of procedures in a program. Its function is to facilitate analysis, design, and coding by breaking down large problems into smaller parts to make them easier to understand and solve. This helps in evaluation and problem-solving.

1. A flowchart is usually represented with a process flow that goes from top to bottom and from left to right on the page.
2. The activities that are visualized or documented should be clearly understandable by the users.
3. There must be clarity for the beginning and the end.
4. The description of the activity stages must use words that indicate actions or deeds.
5. The steps of the activities must be in the correct order.
6. The scope of the ongoing activities must be identified and understood thoroughly through careful tracing.
7. The use of standard flowchart symbols is recommended. [3]

Digital Storage, often referred to as digital data storage, is a technology that combines computer components and storage media to store data digitally. It comes in various forms, ranging from network-attached storage, physical hard drives located inside computers, distributed cloud storage, to virtual storage.

1. Direct Attached Storage (DAS)
2. Network Attached Storage (NAS)
3. Storage Area Network (SAN)
4. Cloud
5. Cloud Hybrid [4]

Black box testing is a software testing technique used to determine the functionality of an application. The main focus of black box testing is the availability or conformity between the input for an application and the expected output for each input value.

1. Equivalence Partitioning
2. Boundary Value Analysis
3. Cause Effect Graphing
4. Decision Table Based Testing
5. Error Guessing [5]

3. Research Methods

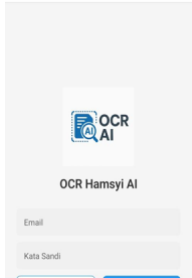
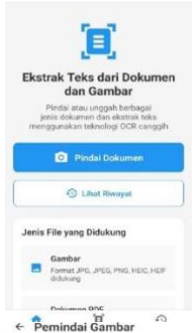


1. File Input; Researchers will collecting various data for analysis
2. Preprocessing; This process converts a color or grayscale image into a binary (black and white) image. It simplifies the image by making the text appear as solid black on a solid white background, removing any shades of gray or color. This is done by applying a threshold; any pixel value above the threshold becomes white, and any below becomes black
3. Segmentation; This process of breaking down a document image into smaller, more manageable parts, this involves identifying and separating distinct zones such as text blocks, images, tables, and columns.

4. Normalization; This process is divided to 2 process, scaling and thinning. Scaling adjusts the image's size to this optimal resolution, ensuring that the characters are clear and correctly sized for the recognition process. Thinning reduces the thickness of character strokes to a uniform width, usually a single pixel.
5. Feature Extraction; The system then analyzes the isolated character, looking for key structural features.
6. Recognitio; The system compares the image of an isolated character (called a glyph) directly to a library of pre-stored templates. If the new character's image is a very close.

4. Result and Discussion

The following will explain the appearance of OCR Hamsyi AI, which can be seen as follows.

Table 1: Results

Page Menu	Menu
	Login Section
	Main Menu
	Scan Menu
	Result Page

After the application is built, several tests are conducted with the following results.

Table 2: Result of tests

Testing Category	Number of Cases	Succeeded	Succeeded with Note	Failed
Authentication Module	5	5	0	0
OCR Module for Images	10	8	2	0
Result Storage Module	10	10	0	0
Scan History	10	10	0	0
Module Compatibility Testing	5	5	0	0
Total	40	38	2	0

Overall, the Hamsyi AI OCR application has met the established functional requirements with a success rate of 95%, with 5% noted due to incomplete characters resulting from unclear image files. There were no failed test cases, indicating that the application is ready for use. The advantages of the built application are:

1. The application successfully executes the main OCR function with a 95% success rate for printed documents and high-quality images.
2. The authentication module, results storage, and scanning history function well without significant issues.
3. The application demonstrates good compatibility across various versions of the Android operating system.
4. The application is capable of systematically recognizing and extracting image files containing more than one page.
5. The extracted text can be copied to the device's clipboard.

Meanwhile the weaknesses of the application that has been built are:

1. The application still shows limitations in recognizing and extracting handwritten text.
2. The application does not yet have the feature to read and extract information from PDF files or other document files.

5. Conclusion and Suggestions

The following outlines the conclusions from the research conducted:

1. The Hamsyi AI OCR application can be used to scan and extract data accurately. The AI support feature can interact with users regarding the contents of the documents.
2. The results of testing the Optical Character Recognition (OCR) method in black box testing show that the application's success rate reaches 95% and 5% succeeded with notes.

For further development of this application, the following suggestions are provided:

1. Further development is needed in handwriting recognition using Optical Character Recognition (OCR) algorithms.
2. This research can be continued by considering further development using more advanced AI technologies to provide a better experience in the PDF extraction process.

References

- [1] R. Rismanto, A. Prasetyo, and D. A. Irawati, 'Optimization of Image Thresholding on Optical Character Recognition in Digitalization and Document Search Systems,' *PETIR*, vol. 13, no. 1, pp. 1-11, Mar. 2020, doi: 10.33322/petir.v13i1.659.
- [2] D. Akinbade, A. O. Ogunde, M. O. Odim, and B. O. Oguntunde, "An adaptive thresholding algorithm-based optical character recognition system for information extraction in complex images," *Journal of Computer Science*, vol. 16, no. 6, pp. 784
- [3] "Utilization of Flowcharts for Business Process Description Needs," *Journal of Computer Science*, Volume 12 Number 1, March 2016.
- [4] Muttaqin et al., "Implementation of Artificial Intelligence (AI) in Life", Medan: Yayasan Kita Menulis, 2023, [Online] Available: <https://library.gunadarma.ac.id/ecollection> [Accessed: January 25, 2025].
- [5] Paradigm, V. (2022). "Use Case Diagram notations guide", Circle, 2022, [Online] Available: <https://circle.visual-paradigm.com/> [Accessed on January 25, 2025].
- [6] A. Karmakar, S. S. Roy, F. Vercauteren, and I. Verbauwhe, "Efficient finite field multiplication for isogeny based post quantum cryptography," 2017, doi: 10.1007/978-3-319-55227-9_14.
- [7] A. M. H. Pardede, M. Zarlis, and H. Mawengkang, "Optimization of Health Care Services with Limited Resources," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 9, no. 4, pp. 1444-1449, 2019, doi: 10.18517/ijaseit.9.4.8348.
- [8] A. M. H. Pardede, Y. Maulita, and R. Buatun, "Application modeling ipv6 (internet protocol version 6) on e-id card for identification number for effectiveness and efficiency of registration process identification of population," in *Journal of Physics: Conference Series*, 2018, vol. 978, no. 1, doi: 10.1088/1742-6596/978/1/012017.
- [9] S. P. Mohanty, U. Choppali, and E. Kougianos, "Everything you wanted to know about smart cities," *IEEE Consum. Electron. Mag.*, vol. 5, no. 3, pp. 60-70, 2016, doi: 10.1109/MCE.2016.2556879.
- [10] W. A. Jabbar, W. K. Saad, and M. Ismail, "MEQSA-OLSRv2: A multicriteria-based hybrid multipath protocol for energy-efficient and QoS-aware data routing in MANET-WSN convergence scenarios of IoT," *IEEE Access*, 2018, doi: 10.1109/ACCESS.2018.2882853.