

Developing an Optical Character Recognition Application for Mobile-Based Recognition and Translation of Arabic-Malay Characters

Fitri Catur Utami^{*,a,1}, Fajri Profesio Putra^{a,2}

^aPoliteknik Negeri Bengkalis, Bengkalis, Riau

*Corresponding Author: fitricutami@gmail.com

Abstract

This research develops an Optical Character Recognition (OCR) application to recognize and translate Arabic-Malay script, widely used in Bengkalis and surrounding regions. The application is designed to facilitate learning and interaction with the script through features for text recognition and translation. To enhance OCR accuracy, advanced image preprocessing techniques and a comprehensive training dataset were employed. Testing under various lighting conditions showed an average recognition accuracy of 25%. Performance was optimal under well-lit conditions, achieving 100% accuracy at 1639 lux, but accuracy significantly declined in low-light environments, particularly at 3.5 lux, where it dropped to 0%. While this result demonstrated the system's potential, a comparison with similar OCR systems indicates significant room for improvement, particularly in recognizing complex or handwritten Arabic-Malay characters. The results highlight challenges related to lighting conditions, including variations in light sources and angles, which greatly affect recognition accuracy. Indeed, future work should focus on expanding the training dataset with diverse samples, refining preprocessing techniques such as automatic brightness adjustment, and implementing advanced machine learning models like Convolutional Neural Networks (CNNs). This application holds significant potential for practical uses in education, translation services, and local communication, contributing to preserving and promoting the Arabic-Malay script as part of cultural heritage in the modern era.

Keywords: OCR, Arabic-Malay Script, Learning, Light Conditions, Light Sources, Dataset Expansion

I. INTRODUCTION

The Arabic-Malay script, also known as Jawi, is derived from the Arabic script and has been adapted to accommodate the phonetic needs of the Malay language. This adaptation involves the integration of the Arabic alphabet with modifications, including the addition of five unique characters specific to the Malay community [1], [2], [3]. The script, which comprises 29 letters from "alif" to "ya," began to flourish with the arrival of Islamic culture in the Malay Archipelago, particularly during the reign of the Samudra Pasai Kingdom and the Malacca Sultanate. Historically, Jawi served as a primary medium for communication, education, and governance in Malay-speaking regions. In Riau Province, especially in Bengkalis, the Arabic-Malay script was once integral to the school curriculum from elementary to secondary levels. However, the introduction of the 2013 national curriculum led to its removal from most schools, leaving only a few institutions to continue its teaching. This shift has created significant barriers for younger generations, particularly elementary school students, who now lack exposure to the script in formal education [4], [5].

Recognizing the importance of Jawi as a cultural heritage, the Riau Provincial Government introduced Regulation No. 46 of 2018, mandating its use in various public settings, such as road signs, government offices, and official documents [6], [7]. Despite these efforts, the familiarity and proficiency of the younger generation with Jawi remain limited, posing a substantial challenge to its preservation. To address this issue, advancements in technology offer practical solutions. Android-based Optical Character Recognition (OCR) technology provides a means of recognizing and digitizing Jawi text from images, enabling users to interact with the script in real time. This innovation not only simplifies access to Arabic-Malay script but also serves as a more efficient alternative to traditional dictionaries and digital translators, which are often time-consuming and limited in their capabilities [8]. By leveraging technology, the Arabic-Malay script can be reintroduced to a wider audience, ensuring its relevance and survival in modern contexts.

The Arabic-Malay recognition application developed in this research is equipped with a learning module that includes several key features. The learning module begins with a general introduction to the Arabic-Malay script, followed by an introduction to the letters, each accompanied by example words. Additionally, a writing feature allows users to follow displayed words on the screen through an image. Users can write the words by tracing them with their fingers directly on the device screen. The application is also integrated with educational

videos through YouTube, offering further guidance on learning the Arabic-Malay script, thereby simplifying the learning process.

Several studies have focused on similar OCR-based systems, such as the research conducted by [9], [10], [11], which addresses the challenges of processing and scanning physical documents, particularly in situations where documents are lost or damaged. This study shows the potential of OCR technology for improving document processing efficiency. Similarly, [12], [13], [14] explored the use of Augmented Reality (AR) for introducing the Arabic-Malay script, providing an innovative approach to learning, although it does not utilize OCR technology. Furthermore, [15], [16], [17] developed an image-to-text translator using OCR to address language barriers for tourists in Indonesia. However, none of these studies specifically addresses the integration of OCR technology for the Arabic-Malay script. This gap in research presents an opportunity for the current study, which combines OCR-based text recognition with interactive learning features and multimedia support to facilitate the learning of Arabic-Malay script and its preservation. Therefore, this research aims to develop an Android-based OCR application that can function as both an introduction to Arabic-Malay script and a flexible learning tool accessible anytime. This application is expected to contribute to the preservation of the Arabic-Malay script, particularly in Bengkalis, by utilizing modern and user-friendly technology. Through this application, it is expected that younger generations can easily recognize and learn the Arabic-Malay script without relying on formal educational media.

II. METHOD

This research focused on developing an Optical Character Recognition (OCR) system for recognizing Arabic-Malay script. The data used in this study consisted of various images of Arabic-Malay script, which were collected from Malay cultural institutions and documents. The dataset included a variety of text styles and character shapes to ensure the system could recognize different forms of Arabic-Malay script. The images were preprocessed to remove noise, enhance contrast, and improve the clarity of the characters, making them more suitable for accurate OCR processing. Preprocessing techniques such as binarization (conversion to black-and-white) and segmentation were applied to isolate individual characters for recognition.

The tools used in this research included both hardware and software components. The hardware included a laptop with an Intel Core i7 processor, 8GB RAM, and a 273GB SSD, along with an external camera for capturing images of Arabic-Malay text. The software used for development included Windows 10 Pro as the operating system, Dart for programming the Android application, and Android Studio for building and testing the app. Flutter was employed as the cross-platform framework for mobile development, while TensorFlow Lite was used for running the machine learning models in the application. Visual Studio Code served as the code editor for the project.

The development of the OCR system followed the Prototype Method, which was iterative and allowed for continuous user feedback. The process began with the gathering of system requirements through interviews with experts in Malay culture and the analysis of Arabic-Malay documents. This information helped to define the key features of the application. A quick design prototype was then created, focusing on the primary functions of image input, text recognition, and translation. The initial prototype was tested with users who provided feedback on its usability and accuracy. Based on their feedback, refinements were made to the OCR model and the user interface. The final prototype was developed to handle a wide range of text styles and varying lighting conditions, improving its overall accuracy and usability.

The OCR system was developed using Teachable Machine, which incorporated Convolutional Neural Networks (CNNs) for pattern recognition and produced models in TensorFlow Lite (TFLite) format for deployment. The process began with dataset preparation and preprocessing, including noise reduction, binarization, and character segmentation, ensuring that the images were optimized for training. Feature extraction and model training were conducted by using the Teachable Machine platform to leverage CNNs to identify unique patterns in Arabic-Malay characters. Once trained, the model was exported in TFLite format and integrated into the Android application, enabling efficient and real-time text recognition on mobile devices.

The system architecture included several key components, from image input to text recognition and translation. Users could capture or upload images containing Arabic-Malay script, which the system processed in real-time [18], [19]. The system is optimized to handle various types of text inputs and ensure fast, accurate recognition. The mobile application also offered features such as text recognition, where users can capture or upload images of Arabic-Malay script, and text translation, which instantly converts the recognized text into Indonesian. Additionally, the app provided educational tools that allowed users to practice writing Arabic-Malay characters and receive feedback on their progress, making it a valuable resource for both learners and general users.

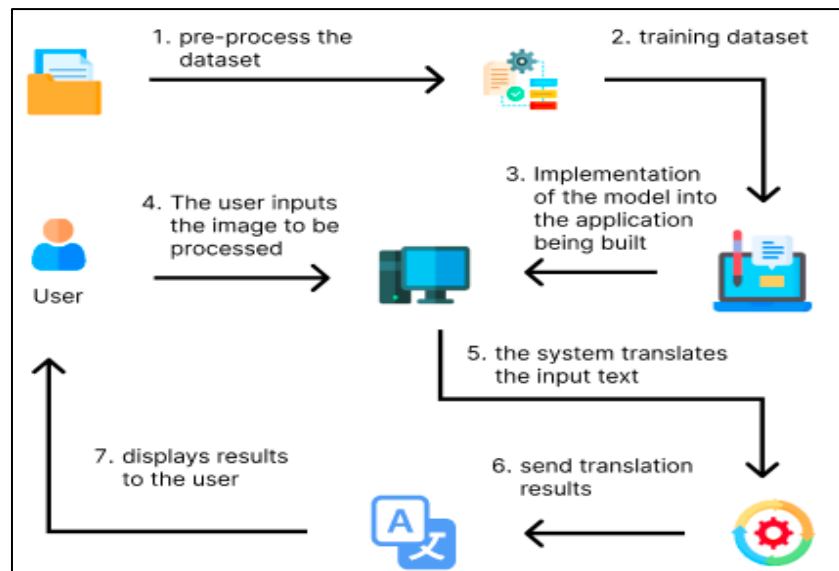


Figure 1 System Architecture

Figure 1 illustrates the workflow of an image-to-text translation system. It begins with the pre-processing of the dataset, where the data are organized and cleaned for optimal use in model training. Once pre-processed, the dataset is used to train the model, enabling it to recognize and interpret features from images, such as Arabic-Malay text. After the model is trained, it is implemented into the application, allowing the application to leverage the model for real-time use. The user then inputs an image containing text that they wish to be translated. The system proceeds to translate the input text using the trained model and then sends the translation results to the processing stage for further refinement. Finally, the application displays the results to the user, completing the image-to-text translation process.

III. RESULTS AND DISCUSSION

A. *Developing an Optical Character Recognition Application*

The evaluation of the Optical Character Recognition (OCR) application for Arabic-Malay script recognition was conducted through systematic and controlled testing under diverse lighting environments. The main objective of these evaluations was to investigate how well the OCR system could maintain its accuracy when exposed to variable lighting conditions, which are commonly encountered in real-world scenarios. The tests were structured to simulate both optimal and suboptimal lighting situations, ranging from bright natural daylight to dim artificial illumination. This approach ensured a comprehensive assessment of the application's capacity to process and recognize characters accurately, despite potential visual impairments caused by inconsistent lighting.

During the testing phase, the system's performance was measured by examining its ability to correctly identify and interpret individual Arabic-Malay characters. This included analyzing its recognition rate, the frequency of errors, and its responsiveness to minor distortions in the script due to lighting glare or shadow interference. Particular attention was given to diacritical marks and character shapes, which are crucial in Arabic-Malay orthography and may be easily misread under poor lighting. The analysis revealed patterns in how lighting affected recognition accuracy, with notable declines in precision when the script was captured under low or uneven light sources. These results underscore the importance of optimal lighting for reliable OCR function.

The outcomes of these tests highlighted both the strengths and limitations of the OCR application. On one hand, the system demonstrated strong accuracy in well-lit conditions, reliably interpreting standard script with minimal errors. On the other hand, its performance decreased in dim or overly bright settings, indicating a need for improved adaptability to environmental changes. This limitation suggests that while the OCR application is highly effective under controlled conditions, further enhancements in image preprocessing and adaptive thresholding may be necessary to improve its robustness. Overall, the evaluation emphasized the significance of environmental variables in OCR technology and provided essential feedback for future development focused on increasing its versatility in real-world use cases.



Figure 2 Translate Text Page

Figure 2 illustrates a user-friendly interface designed specifically for text translation within the application. The page is titled “*Terjemah Teks*” (Translate Text), clearly indicating its primary function. At the top portion of the screen, users are presented with two intuitive options: a camera icon that allows them to directly capture an image of the text they wish to translate, and an image icon that enables the selection of existing images from the device’s gallery. These features offer flexibility and convenience, catering to both real-time translation needs and pre-existing content. This dual-option setup is particularly useful for users who may encounter text in different environments, whether needing to quickly translate signage, documents, or other visual information on the go.

In the middle section of the interface, a central white box displays an image of Arabic text, demonstrating the app’s ability to process non-Latin scripts. Just below this image, a text field reveals the translated result, which in this case is the word “*garam*” (salt). This simple yet clear example highlights the effectiveness of the app’s OCR (Optical Character Recognition) and translation capabilities. The streamlined layout ensures that users can easily view both the source text and the resulting translation without confusion. By minimizing the number of steps required and presenting the information in a clean, organized format, the interface enhances the overall user experience, promoting ease of use, accessibility, and efficiency in multilingual communication.

B. Performance Under Lighting Conditions

Table 1 presents a detailed summary of the lighting test results aimed at evaluating the performance of character recognition under a specific luminance level, namely 1639 lux. This table systematically documents the outcomes of four separate image inputs, each containing a different word with varying lengths, to assess how consistent lighting conditions influence the accuracy of Optical Character Recognition (OCR). The parameters recorded for each test case include the luminance value, the image displayed, the word contained in the image, the total number of characters in each word, and the number of characters correctly recognized by the system. The words tested, “Duda,” “Gagah,” “Lapang,” and “Pekanbaru,” represent a range of character lengths, from short (4 characters) to longer words (9 characters), thereby providing a broader perspective on system accuracy across word complexities. The results indicate that all characters in each word were correctly identified, suggesting that the application performs with perfect accuracy under the highlighting level of 1639 lux. These findings imply that sufficient and consistent lighting significantly enhances OCR reliability, and they serve as a baseline for comparing performance under lower lighting conditions in subsequent tests.

Table 1 Lighting Test Table

No	Luminance (Lux)	Images	Word	Number of Characters	Correct Character
----	-----------------	--------	------	----------------------	-------------------

1	1639		Duda	4	4
---	------	--	------	---	---

No	Luminance (Lux)	Images	Word	Number of Characters	Correct Character
2	1639	گاجه	Gagah	5	5
3	1639	لاڦڻ	Lapang	6	6
4	1639	پڪانبھارو	Pekanbaru	9	9

Table 1 presents the results of a lighting test aimed at evaluating the effectiveness of an Optical Character Recognition (OCR) system under a luminance level of 1639 lux. This luminance level represents a well-lit environment, suitable for accurate visual processing by both human observers and automated recognition systems. The table includes four sample words: “Duda,” “Gagah,” “Lapang,” and “Pekanbaru,” which vary in length from 4 to 9 characters. Each word was extracted from an image and processed by the OCR system to assess how accurately the characters could be recognized. The data shows that, in all four cases, the number of correctly recognized characters matches exactly the total number of characters in each word, indicating a 100% character recognition rate across all tested samples. This consistency suggests that the OCR system functions with high precision when exposed to optimal lighting conditions.

From a technical and analytical perspective, these results validate the hypothesis that higher luminance levels contribute positively to OCR performance. The absence of recognition errors across varying word lengths also implies that the system is not significantly affected by the complexity of input (in terms of character count) under such lighting conditions. Moreover, the uniform success rate across different words demonstrates the robustness and reliability of the application’s text recognition algorithm in well-illuminated settings. These findings are critical for the development of translation or scanning applications where lighting can be a determining factor in user satisfaction and output quality. The high accuracy achieved in this controlled environment serves as a strong baseline for future comparative analysis involving lower luminance levels, where recognition accuracy may decline and reveal the lighting thresholds of the OCR system’s operational effectiveness.

Furthermore, Table 2 presents a comprehensive analysis of Optical Character Recognition (OCR) accuracy across varying lighting levels, measured in lux, to assess the system’s performance under different illumination conditions. The table includes six key metrics: the number of words tested, the number of words correctly recognized, accuracy, precision, recall, and F1 score. Each lighting level, ranging from well-lit conditions (1639 lux) to extremely low light (3.5 lux), was tested using four word samples to evaluate the OCR’s effectiveness. The results demonstrate that the OCR system achieves perfect accuracy (100%) in high and moderate lighting conditions (1639 lux and 485 lux), while its performance significantly declines under poor lighting. At 180 lux, accuracy drops to 75%, and at 10 lux, it falls further to 50%, with corresponding decreases in recall and F1 score. At the lowest illumination level of 3.5 lux, the system fails to recognize any words, resulting in 0% across all performance metrics. This data clearly illustrates the critical role lighting plays in ensuring optimal OCR functionality.

Table 2 Illustrates the Analysis of OCR Accuracy

Lighting Level (Lux)	Number of Words Tested	Number of Words Correctly Recognized	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1639	4	4	100	100	100	100
485	4	4	100	100	100	100
180	4	3	75	100	75	85,71
10	4	2	50	100	50	66,67
3,5	4	0	0	0	0	0

The OCR system was tested under five lighting levels: 1639 lux (bright daylight), 485 lux (moderate light), 180 lux (dim light), 10 lux (low light), and 3.5 lux (very low light). The results indicate that the system performed optimally in bright lighting, achieving 100% accuracy at 1639 lux and 485 lux. However, performance declined significantly as lighting worsened, with accuracy dropping to 50% at 10 lux and failing at 3.5 lux. This demonstrates the critical role of adequate lighting in the OCR system's ability to recognize characters accurately. The results indicate that the system performed optimally in bright lighting, achieving 100% accuracy at 1639 lux and 485 lux. However, performance declined significantly in low-light environments, with accuracy dropping to 50% at 10 lux and failing at 3.5 lux. This suggests that both the type of light (e.g., LED, fluorescent, or natural light) and the angle of illumination significantly impact character recognition accuracy. Future implementations should incorporate preprocessing techniques like automatic brightness adjustment or histogram equalization to enhance the system's robustness under challenging lighting conditions [20], [21].

C. Character Recognition Reliability

Table 3 presents a series of examples from the reliability testing conducted to evaluate the performance of character recognition in Arabic-Malay text. This table is instrumental in assessing how accurately the system can interpret and process Arabic-script words commonly used in the Malay language, which often include unique diacritical marks and script features that pose challenges to OCR technology. Each row in the table lists an Arabic-Malay word alongside the total number of characters it contains and the number of those characters that were successfully recognized by the system. These data points serve to illustrate the consistency and robustness of the OCR algorithm when applied to texts that are structurally and linguistically distinct from Latin-based scripts.

The table not only provides a quantitative perspective on the character recognition process but also contributes to understanding the practical applicability of the OCR system in real-world contexts where Arabic-Malay writing is commonly used, such as religious, educational, or cultural materials. By comparing the total characters with the correctly identified ones for each word, researchers can identify patterns of strengths or weaknesses in the system's ability to detect particular script features. High recognition rates suggest system reliability and suitability for deployment in Malay-speaking regions, while any inconsistencies or errors could inform future improvements in image preprocessing, language modeling, or font training. Overall, Table 3 functions as an essential component in validating the OCR tool's reliability for multilingual and non-Latin script environments.

Table 3 Example Reliability Testing

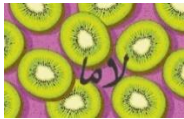
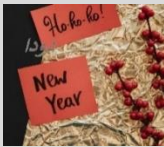

No	Image	Word	Total Character	Correct Character
1		Lama	4	1
2		Duda	4	0
3		Jalan	5	2

Table 4 presents a detailed analysis of Optical Character Recognition (OCR) accuracy derived from reliability testing focused specifically on Arabic-Malay text. This table plays a critical role in evaluating the system's consistency and effectiveness in handling non-Latin scripts, particularly those incorporating Arabic

characters used in Malay writing. For each test case, the table outlines the total number of characters present in the original text and compares it to the number of characters accurately recognized by the OCR system. This comparison allows for the calculation of an accuracy percentage, offering a quantifiable measure of the system's recognition reliability for each instance. The results in Table 4 provide valuable insights into how well the OCR system can process varying levels of textual complexity and character count in Arabic-Malay scripts. High accuracy percentages indicate strong performance and suitability for practical applications such as document digitization, translation, or educational tools. Conversely, any notable drop in recognition accuracy highlights specific challenges that may require algorithm refinement or additional training data. Overall, this table serves as an important benchmark in validating the OCR system's capability for Arabic-Malay text, informing future improvements and ensuring its applicability across multilingual contexts where script diversity is a significant factor.

Table 4 Analysis of OCR Accuracy Based on Reliability Testing

No	Image Text	Total Characters	Characters Recognized Correctly	Accuracy (%)
1	Lama	4	1	25
2	Duda	4	0	0
3	Jalan	5	2	40
4	Jalan Gatot Subroto	17	3	17.65
5	Bahasa	6	2	33.33
6	Negri Deli	9	2	22.22
7	Darma	5	1	20
8	Nama	4	2	50
9	Arab Melayu	10	3	30
10	Jawi	4	1	25

The reliability testing results indicate that the OCR system's performance was inconsistent, achieving an overall accuracy of 25% across ten tested images. This percentage represents the ratio of correctly recognized characters to the total number of characters across all test samples. Specifically, out of 40 total characters in the test images, only 10 characters were accurately recognized [22]. Variations in text length, character size, and input quality significantly influenced the system's ability to recognize characters correctly. Longer and more complex texts, such as "Jalan Gatot Subroto," exhibited lower recognition rates, while simpler texts, like "Nama," showed relatively higher accuracy. These findings suggest that the OCR system struggled to generalize across diverse inputs and required improvements in training data diversity, preprocessing techniques, and recognition algorithms to enhance its reliability and applicability in real-world scenarios.

The reliability testing results indicate that the OCR system's performance was inconsistent, achieving an overall accuracy of 25% across ten tested images. This percentage represents the ratio of correctly recognized characters to the total number of characters across all test samples. Specifically, out of 40 total characters in the test images, only 10 characters were accurately recognized. Variations in text length, character size, and input quality significantly influenced the system's ability to recognize characters correctly. The 25% overall recognition accuracy indicates substantial limitations, especially for complex or handwritten Arabic-Malay characters. Comparative analysis with other OCR systems revealed that this result underscores the necessity for more robust training datasets and advanced preprocessing techniques. For instance, augmenting the dataset with synthetic images or diverse text styles could improve the system's ability to generalize across various input types.

D. Application Interface and Functionality

The OCR application enables users to upload or capture images of Arabic-Malay text for recognition and translation. The system processes the text and provides a digital output that can be translated into Indonesian. The application also includes an intuitive interface that facilitates user interaction, allowing individuals to learn and practice writing Arabic-Malay characters through integrated educational modules. The results demonstrate that the OCR system shows strong potential for recognizing Arabic-Malay script in controlled environments with optimal lighting. The system's 100% accuracy under bright conditions (1639 lux and 485 lux) underscores the feasibility of OCR technology for recognizing Arabic-Malay text. However, the significant decline in performance in low-light conditions, particularly at 3.5 lux, emphasizes the importance of improving the system's

preprocessing capabilities. Techniques such as automatic brightness adjustment and noise filtering could enhance the OCR's robustness under challenging lighting conditions.

The 25% overall accuracy for character recognition highlights limitations in the OCR system's ability to handle diverse and complex text styles. The current model performs well with simple characters and high-quality images, but struggles with more intricate and handwritten characters. Expanding the training dataset to include a wider range of Arabic-Malay text samples, along with incorporating advanced machine learning algorithms like deeper Convolutional Neural Networks (CNNs), could significantly improve the system's recognition capabilities. The user-friendly design of the OCR application facilitates its adoption by a broad audience, from beginners learning Arabic-Malay script to more experienced users. By integrating real-time text recognition and translation features, the application serves as a valuable educational tool [23], [24]. However, further enhancements to the system's recognition accuracy and its ability to handle complex images are necessary to make the application more effective and reliable. Indeed, the OCR system demonstrates potential for supporting Arabic-Malay script recognition and education. Nevertheless, future efforts should focus on improving performance in low-light environments, expanding training datasets, and enhancing model robustness to ensure reliable recognition across a variety of conditions [25], [26]. These refinements will enable the application to serve as a comprehensive tool for preserving and learning Arabic-Malay script.

IV. CONCLUSION

This study developed an Optical Character Recognition (OCR) application for recognizing and translating Arabic-Malay script, widely used in Bengkalis and its surrounding regions. The application has successfully facilitated learning and interaction with Arabic-Malay text through integrated recognition and translation features. The system demonstrated optimal performance in well-lit conditions, achieving 100% accuracy under 1639 lux and 485 lux lighting levels. However, performance declined significantly in low-light environments, particularly at 3.5 lux, highlighting the importance of lighting conditions for reliable OCR functionality. The overall recognition accuracy of the OCR system was 25%, with challenges observed in handling complex and handwritten Arabic-Malay characters. These findings indicate the need to enhance the system's training dataset to include more diverse and representative samples of Arabic-Malay text. Additionally, integrating advanced machine learning models, such as deeper Convolutional Neural Networks (CNNs), could improve the system's ability to recognize complex and low-quality text images.

Future work should focus on expanding the training dataset to include diverse Arabic-Malay text samples, such as handwritten characters and varying font styles. Additionally, integrating preprocessing techniques like adaptive thresholding or brightness adjustment could enhance image quality for low-light recognition. Another promising approach is utilizing deeper Convolutional Neural Networks (CNNs) for more robust and accurate text recognition. One potential enhancement is the integration of a light sensor to detect and adjust for low-light conditions dynamically. This would allow the application to recommend additional lighting or optimize the preprocessing stage based on detected light levels. In short, the OCR system has potential applications beyond learning and interaction. For example, it could be employed in educational tools for schools that teach Arabic-Malay script, translation services for documents, or local communication in communities that frequently use the script. These applications could help preserve the cultural heritage of Arabic-Malay script while providing practical solutions for its use in modern contexts.

REFERENCES

- [1] H. K. Sari, M. Audina, S. M. Sunarso, And H. Faizah, "Analisis Penulisan Arab Melayu Pada Tugu Nama Balai Bahasa Provinsi Riau," *J. Ris. Pendidik. Dan Bhs.*, Vol. 2, No. 1, Pp. 39–45, 2022, <http://doi:10.58192/Insdun.V2i1.407>.
- [2] M. S. Asih And H. Medan, "Pengenalan Huruf Pada Citra Digital Menggunakan Algoritma Template Matching," No. 70, 2017.
- [3] I. G. M. D. I Kadek Dharma Laksana, I Made Gede Sunarya, "Aplikasi Ocr (Optical Character Recognition) Tulisan Tangan Aksara Bali Dengan Metode Contour Analysis," *Kumpul. Artik. Mhs. Pendidik. Tek. Inform.*, Vol. 4, Pp. 49–58, 2015.
- [4] A. N. Hasan Nindya Murwato, Suci Aulia, "Perancangan Translator Image To Text Dengan Menggunakan Metode Optical Character Recognition Berbasis Matlab," Vol. 6, No. 1, Pp. 629–636, 2020.
- [5] A. Setiawan, H. Sujaini, And A. B. Pn, "Implementasi Optical Character Recognition (Ocr) Pada Mesin Penerjemah Bahasa Indonesia Ke Bahasa Inggris," *J. Sist. Dan Teknol. Inf.*, Vol. 5, No. 2, Pp. 135–141, 2017.

- [6] F. O. Rahmalisty, S. Aulia, And S. Hadiyoso, "Perancangan Language Translator Image To Text Menggunakan Metode Optical Character Recognition Berbasis Pengolahan Citra Language Translator Image To Text Design Using Optical Character Recognition Method Based On Image Processing," Vol. 9, No. 1, Pp. 187–195, 2023, [Online]. Available: <https://Openlibrarypublications.Telkomuniversity.Ac.Id/Index.Php/Appliedscience/Article/View/19395>
- [7] Nehru, Y. R. Hais, And A. C. Devina I, "Implementasi Canny Filter Optical Character Recognition(Ocr) Untuk Identifikasi Tanda Nomor Kendaraan Bermotor (Tnkb)," *J. Eng.*, Vol. 6, Pp. 1–9, 2024, [Online]. Available: <https://Online-Journal.Unja.Ac.Id/Jurnalengineering/Article/View/31406/17489>
- [8] T. E. Prasetiyo, "Rancang Bangun Aplikasi Penerjemah Bahasa Indonesia-Jawa Menggunakan Optical Character Recognition," *Comput. Its Appl. J.*, Vol. 4, No. 1, Pp. 37–40, 2021, [Online]. Available: <https://Www.Nesabamedia.Com/Pengertian->
- [9] P. Haryani, "Augmented Reality (Ar) Sebagai Teknologi Interaktif Dalam Pengenalan Benda Cagar Budaya Kepada Masyarakat," *J. Simetris*, Vol. 8, No. 2, P. 807, 2017.
- [10] Nurhadi And Mulyadi, "Rancang Bangun Aplikasi Augmented Reality Berbasis Face Tracking Untuk Mendeteksi Wajah Peserta Wisuda," Vol. 13, No. 1, Pp. 1189–1199, 2018.
- [11] E. Roza, "Aksara Arab-Melayu Di Nusantara Dan Sumbangsihnya Dalam Pengembangan Khazanah Intelektual," *J. Perad. Islam*, Vol. 13, No. 1, P. 177, 2017, <http://doi: 10.21111/Tsaqafah.V13i1.982>.
- [12] F. Qomariyah And L. Nisa, "Aplikasi Pengenalan Tulisan Tangan Huruf Arab Menggunakan Kombinasi Fitur Dan Tts Untuk Anak Usia Dini," *J. Apl. Teknol. Inf. Dan Manaj.*, Vol. 2, No. 2, Pp. 41–55, 2021, <http://doi: 10.31102/Jatim.V2i2.1240>.
- [13] W. Hamidah, Irawan, Hasbullah, And Kaswar, "Deteksi Nominal Uang Kertas Menggunakan Ocr (Optical Character Recognition)," *J. Ilmu Komput. Dan Teknol. Inf.*, Vol. 7, No. 2, Pp. 72–76, 2022, <http://doi: 10.36805/Technoexplore.V7i2.2123>.
- [14] S. M. Novriyanto, "Perangkat Ajar Tulisan Arab Melayu Berbasis Multimedia (Studi Kasus: Smpn 023 Pekanbaru)," Pp. 217–224, 2011.
- [15] A. Sri Wulandari And S. Mufti Prasetiyo, "Perancangan Aplikasi Pembelajaran Pengenalan Angka 3 Bahasa (Indonesia, Arab & Inggris) Berbasis Android Menggunakan Metode Prototype," *J. Ilmu Komput. Dan Sci.*, Vol. 1, No. 08, Pp. 1131–1145, 2022.
- [16] F. H. Sadrakh Ginting, "Penggunaan Mobile Vision Api Untuk Pengenalan Teks Dalam Sebuah Image Berbasis Android," Pp. 1143–1152.
- [17] I. N. T. Lestari And D. I. Mulyana, "Implementation Of Ocr (Optical Character Recognition) Using Tesseract In Detecting Character In Quotes Text Images," *J. Appl. Eng. Technol. Sci.*, Vol. 4, No. 1, Pp. 58–63, 2022, <http://doi: 10.37385/Jaets.V4i1.905>.
- [18] Kusanantoro, T. Rohana, And D. Sulistya Kusumaningrum, "Implementasi Metode Tesseract Ocr(Optical Character Recognition) Untuk Deteksi Plat Nomor Kendaraan Pada Sistem Parkir," *Sci. Student J. Information, Technol. Sci.*, Vol. 3, No. 1, Pp. 59–67, 2022.
- [19] Ikrimach And M. Retnowo, "Implementation Of Multi-Criteria Decision Making (Mcdm) For New Student Majors Recommendations," *Int. J. Eng. Technol. Nat. Sci.*, Vol. 4, No. 1, Pp. 52–58, 2022, <http://doi: 10.46923/Ijets.V4i1.144>.
- [20] W. Pujianti, M. Nasir, And S. Mawarni, "Perancangan Aplikasi Pembelajaran Arab Melayu Berbasis Augmented Reality Untuk Siswa Tingkat Sekolah Dasar," *Semin. Nas. Ind. Dan Teknol.*, Pp. 37–44, 2019, [Online]. Available: <https://Eprosiding.Snit-Polbeng.Org/Index.Php/Snit/Article/View/67>
- [21] M. Rizal Toha And A. Triayudi, "Penerapan Membaca Tulisan Di Dalam Gambar Menggunakan Metode Ocr Berbasis Website (Studi Kasus: E-Ktp)," *Jst (Jurnal Sains Dan Teknol.*, Vol. 11, No. 1, Pp. 175–183, 2022, <http://doi: 10.23887/Jstundiksha.V11i1.42279>.
- [22] M. R. Fauzi, N. Agus, And A. Ajulian, "Mengubah Tulisan Tangan Menjadi Text Digital Ocr (Optical Character Recognition) Dengan Menggunakan Metode Segmentasi Dan Korelasi," Pp. 1–5, 2014.
- [23] Baihaqi, F. Arnia, And R. Muharar, "Dataset Kata Jawi Untuk Sistem Pengenalan Tulisan Tangan Jawi Kuno," *Serambi Eng.*, Vol. Vii, No. 3, Pp. 3546–3554, 2022.
- [24] E. S. Susanto, F. Hamdani, F. Nuryansah, And N. Oper, "Pengembangan Aplikasi Smart-Book Sebagai Media Pembelajaran Bahasa Inggris Anak Berbasis Ar (Augmented Reality)," *J. Mnemon.*, Vol. 5, No. 1, Pp. 64–71, 2022, <http://doi: 10.36040/Mnemonic.V5i1.4438>.
- [25] A. Firdaus *Et Al.*, "Implementasi Optical Character Recognition (Ocr) Pada Masa Pandemi Covid-19," *J. Jupiter*, Vol. 13, No. 2, Pp. 188–194, 2021.
- [26] K. Banu, D. Andreas, W. Anggoro, And A. Setiawan, "Ocr: Masa Depan Pengenalan Karakter Optik Dan Dampaknya Pada Kehidupan Modern," *J. Teknol. Inf.*, Vol. 9, No. 2, Pp. 147–156, 2023, <http://doi: 10.52643/Jti.V9i2.3798>.