Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/AP06.20250248
Published: August 5, 2025

Implementation Of Machine Learning Based Optical Character Recognition for Mobile Based Drug Packaging Identification

Andi Chairunnas, S.Kom., M.Pd., M.Kom.
Department of Computer Science, Faculty of Mathematics and Natural Sciences Pakuan University
Bogor, Indonesia
andichairunnas@unpak.ac.id

Hasrul SE., MM.
Department of Economic and Business
PakuanUniversity Bogor, Indonesia
hasrul@unpak.ac.id

Muslim
Department of Social Sciences and Cultural Sciences Communication Science Study Program
Pakuan University Bogor, Indonesia
muslim@unpak.ac.id

Raihan Ramadian Purwanto, S. Kom
Department of Computer Science, Faculty of Mathematics and Natural Sciences
Pakuan University
Bogor, Indonesia
raihanramadianpurwanto@gmail.com

Biyan Firmansyah S.Kom
Department of Computer Science, Faculty of Mathematics and Natural Sciences Pakuan University
Bogor, Indonesia
bianfirmansyah@unpak.ac.id

Abstract

The difficulty in accurately identifying drug information from packaging, especially for the elderly or when the packaging is damaged, poses significant health risks due to potential misuse. This research addresses this challenge by developing a comprehensive, mobile-based solution. We introduce an Android application that implements a multi- layered approach: (1) on-device Optical Character Recognition (OCR) using Google's ML Kit for fast text extraction,(2) the Jaro-Winkler algorithm to correct and enhance the accuracy of imperfect OCR results, (3) a curated database of 300 common Indonesian drugs, and (4) an innovative chatbot powered by a Large Language Model (LLM) to provide interactive and reliable drug information.

The system was evaluated through a series of comprehensive tests. The results demonstrate high functional stability and a robust Top-5 identification accuracy of 96% across 150 test cases, maintaining 98% accuracy in low-light conditions and 90% on damaged packaging. Furthermore, the application achieved a "Good" usability score of 78.67 from the System Usability Scale (SUS) and received high-to- very-high validity ratings for its content from pharmaceutical experts. This study validates the effectiveness of combining OCR with advanced text matching and a controlled LLM to create a reliable and user-friendly public health tool.

Keywords
Optical Character Recognition, Machine Learning, Jaro-Winkler Algorithm, Mobile Health, Drug Identification.

1. Introduction
The rapid advancement of digital technology has permeated various sectors, including healthcare, where mobile applications offer unprecedented support for daily activities. One critical area is access to medication information. Many individuals, particularly the elderly or those with visual impairments, face challenges in reading the small text on drug packaging. This issue is exacerbated when packaging is damaged, torn, or faded, making crucial information like the drug's name or dosage illegible. For instance, "Methylprednisolone" might be misread as "pred solone," leading to confusion and potentially hazardous medication errors. This creates an urgent need for a technological solution that can accurately interpret text from imperfect images.

Optical Character Recognition (OCR) technology presents a primary solution, enabling the digital extraction of text from images. However, OCR systems are not infallible and often produce minor errors, such as misspellings or missing characters. To address this, our research leverages the Jaro-Winkler algorithm, a string matching method highly effective for correcting errors in short strings like drug names Tresna et al. (2025).

While previous studies have explored OCR for medical reports Bhat et al. (2024) or used Jaro-Winkler for text correction Harahap et al. (2024) , a holistic solution remains a significant research gap. Many systems stop at identification and fail to provide detailed, trustworthy information interactively. This research fills this gap by developing a mobile application with a unique hybrid architecture. It combines fast, offline, on-device OCR (Google ML Kit) and Jaro-Winkler correction with an online, LLM-powered chatbot. The chatbot's knowledge is strictly limited to a curated database of 300 Indonesian drugs, preventing misinformation and ensuring reliable user support.

1.1 Objectives
The primary objective of this research is to develop and evaluate a functional mobile application that automatically identifies drug information from packaging images. This involves implementing OCR for text extraction, applying the Jaro-Winkler algorithm to enhance matching precision, and integrating an LLM-based chatbot for interactive information delivery. The system's performance will be rigorously measured through accuracy, usability, and expert validation tests to ensure it is an effective and reliable tool for users.

2. Literature Review
The development of this application is grounded in several key technological areas that require a comprehensive review. This review is structured around the core components of the system: Optical Character Recognition (OCR), advanced text-matching algorithms, and the integration of Large Language Models (LLMs) in a mobile health context.

The application of OCR in healthcare is well-documented, primarily for its ability to digitize information and automate data entry, thereby enhancing operational efficiency Andi Zulhijar et al. (2024). Research by Bhat et al. (2024) further demonstrates this by using OCR to extract data from medical lab reports for disease prediction. However, the inherent limitations of OCR, where accuracy is highly dependent on image quality, necessitate supplementary technologies Holila et al. (2024). Our research directly addresses this challenge by not relying solely on raw OCR output.

To overcome OCR inaccuracies, text correction and matching algorithms are crucial. The Jaro-Winkler algorithm is particularly well-suited for this task due to its proficiency in handling short strings and common typographical errors, a fact substantiated by multiple studies in contexts of error correction Harahap et al. (2024); Tresna et al. (2025). This makes it an ideal choice for identifying drug names, which are often short, specific, and prone to partial recognition.

Previous research on drug identification systems has explored various approaches. Some have focused on specific user groups, such as developing tools for the visually impaired using cloud-based OCR and text-to-speech APIs Mangundap et al. (2022). Others have concentrated on classification tasks, such as organizing drugs based on their therapeutic class using a combination of OCR and Convolutional Neural Networks (CNN) (Munadhif et al., 2024), or comparing the performance of algorithms like K-Nearest Neighbors (KNN) for drug type classification (Aldi et al., 2022). While these studies are valuable, they often tackle a single aspect of the problem.

A significant gap exists in providing a holistic solution that not only identifies a drug but also delivers detailed, reliable information in an interactive and user-friendly manner. The emergence of Large Language Models (LLMs) offers a paradigm shift in human-computer interaction, with their ability to understand complex instructions and learn in- context (Zhao et al., 2023). However, LLMs also present challenges like information hallucination and bias. Our work addresses this by implementing a controlled Retrieval-Augmented Generation (RAG) architecture. This approach, combined with a robust mobile development framework using Kotlin and Jetpack Compose (AryaRafa et al., 2024; Suartana, 2025), and a scalable backend like Cloud Firestore (Handoyo et al., 2022), creates an integrated system. By validating the final product through rigorous testing methods, including Black Box testing (Febriyanti et al., 2021), System Usability Scale (SUS) (Deshmukh and Chalmeta, 2024), and expert validation with Aiken's V (Utami et al., 2024), this research offers a comprehensive and trustworthy user experience that advances the state of mobile health applications.

3. Methods

This research utilized the Waterfall software development methodology due to its systematic and sequential approach, which is well-suited for projects with clearly defined requirements. The process involved sequential phases of analysis, design, development, and comprehensive testing before deployment.

The system is designed with a three-tier architecture: a client application (Android), a backend service (Google Firebase), and an external LLM service, as visualized in Figure 1. The client, built with Kotlin, handles on-device OCR and Jaro-Winkler matching. The backend uses Cloud Firestore as a database for 300 curated drugs and a Cloud Function as an intermediary to the LLM.

3.1 Evaluation Protocol

To validate the application's effectiveness, a multi-faceted evaluation protocol was implemented, comprising five distinct testing methods:

- Functional Testing: This was conducted using the Black Box method. A total of 20 test cases covering all core functionalities were executed to verify that the application operates as designed without critical errors.
- Accuracy Testing: The system's identification capability was quantitatively measured using a Top-5 Accuracy metric. This involved 150 test cases performed on 50 different drugs under three challenging conditions: ideal lighting, low light, and simulated packaging damage.
- Expert Validation: To assess content reliability, seven professional pharmacists evaluated the accuracy and clarity of the drug information. The feedback was gathered via a structured questionnaire and analyzed using the Aiken's V content validity index.
- Usability Testing: User satisfaction and ease of use were measured by administering the System Usability Scale (SUS) questionnaire to 66 general users, resulting in a quantitative usability score.
- Performance Testing: The application's efficiency was analyzed using the Android Studio Profiler. This involved monitoring and recording CPU usage, RAM consumption, and battery drain during key operational scenarios.

4. Data Collection

The data used in this study consisted of two main components: the drug information that populates the application's database and the image dataset used for testing the system's accuracy.

4.1 Drug Information Dataset

The primary dataset is a curated database containing detailed information on

- **300 types of medicine** commonly found in Indonesia, prescription drugs, over-the-counter, and limited over- the-counter drugs. The data was collected through a meticulous literature and document study, sourcing information from credible and official references, including Badan Pengawas Obat dan Makanan (BPOM), reputable digital health platforms such as Halodoc,Apotek K-24, and the "Informasi Spesialite Obat (ISO) Indonesia 2021" pharmacy reference book. To ensure data quality, each entry was cross-verified between sources, and the final, validated dataset was structured for import into the Cloud Firestore database.
- **Test Image Dataset**
  For the accuracy evaluation, a separate dataset of digital images was collected. This dataset comprised images of both primary (blister packs/strips) and secondary (boxes) drug packaging. The images were captured under three distinct, controlled conditions to simulate real-world usage scenarios: (1) ideal, well-lit conditions; (2) low-light conditions; and (3) physically damaged packaging where the text was partially obscured or torn. This dataset formed the basis for the 150 test cases used in the accuracy testing phase.

5. Results and Discussion

This section presents the comprehensive evaluation of the developed mobile application. The results are systematically organized into numerical findings, graphical representations, validation, and proposed improvements. The overall findings confirm the application's functional stability, the effectiveness of the identification method, the validity of its content from an expert perspective, and its high quality in non-functional aspects. The Black Box testing resulted in a 100% pass rate across all 20 test cases, indicating that the application is functionally stable. Compatibility testing confirmed that the application runs stably on a range of Android versions (8, 12, and 14). Furthermore, performance metrics showed efficient resource management, with a low baseline for idle states and acceptable resource peaks during intensive on-device OCR processing. A dedicated battery test showed a consumption of only 5% over 18 minutes of intensive, continuous OCR use, confirming its efficiency for typical short-term usage patterns. Each result will be detailed and discussed in the following sub-sections..

5.1 Numerical Results

The core performance of the system was quantified through six distinct testing methods. The numerical data validates the application's accuracy, usability, content reliability, functional stability, and performance.

System Accuracy The effectiveness of the identification method was the primary focus of the evaluation. As shown in Table 1, the system achieved a high overall Top-5 accuracy of 96%. The system's robustness is particularly evident in its performance under non-ideal conditions: it maintained 98% accuracy in low light and a remarkable 90% accuracy on damaged packaging. This result is significant as it demonstrates the practical value of the Jaro-Winkler algorithm in compensating for imperfect OCR outputs.

Table 1**.** System Accuracy Test Results Summary

No.	Test Condition	Total Tests	Successes(*TOP* 5)	Accuracy(%)
1.	Ideal Lighting	50	50	100%
2.	Low Light	50	49	98%
3.	Damaged Packaging	50	45	90%
Overall		**150**	**144**	**96%**

Usability The application's ease of use was measured using the System Usability Scale (SUS), a standardized method for quantifying user satisfaction. The test involved 66 respondents from diverse backgrounds. As detailed in Table 2, the application achieved an average SUS score of 78.67. According to established benchmarks, a score above 68 is considered above average. Therefore, this result places the application in the "Good" category, providing strong quantitative evidence that the user interface is intuitive and the overall user experience is positive even for users with varied technical literacy.

Table 2. System Usability Scale (SUS) Results

Respondent	SUS Score Calculation Results										Total	Total Score (x2.5)
	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10		
R1	4	4	4	2	4	3	4	4	4	3	36	90,00
R2	4	3	4	3	3	2	3	4	4	3	33	82,50
R3	3	2	3	2	3	3	3	3	3	3	28	70,00
R4	4	4	4	4	4	4	4	4	4	4	40	100,00
R5	4	4	4	3	3	4	4	4	4	4	38	95,00
R6	4	4	4	3	3	4	3	3	4	3	35	87,50
R7	3	3	3	2	4	3	3	3	3	3	30	75,00
R8	4	3	4	2	3	4	4	4	4	4	36	90,00
R9	3	3	3	4	2	4	3	3	3	2	30	75,00
R10	4	3	3	3	3	4	4	4	4	4	36	90,00
R11	3	2	4	2	3	2	4	2	4	3	29	72,50
R12	3	2	3	3	3	3	3	2	3	3	28	72,50
R13	2	3	2	2	2	2	3	2	2	4	24	60,00
R14	3	3	3	3	3	2	3	3	2	3	28	70,00

R15	4	3	4	3	4	3	4	2	4	3	34	85,00
R16	3	3	4	2	4	3	3	3	3	3	31	77,50
R17	3	3	2	3	3	2	2	3	2	2	25	62,50
R18	3	3	3	2	3	3	3	3	2	2	27	67,50
R19	3	3	4	3	4	3	3	3	3	2	31	77,50
R20	3	3	3	3	2	2	3	2	2	3	26	65,00
R21	2	3	2	3	3	3	4	3	2	3	28	70,00
R22	3	2	3	2	2	4	3	3	3	3	28	70,00
R23	3	2	3	2	3	2	4	2	3	2	26	65,00
R24	2	2	4	2	4	2	4	3	4	3	30	75,00
R25	3	2	3	2	2	3	3	3	3	3	27	67,50
R26	3	3	4	2	3	3	4	3	3	3	31	77,50
R27	2	4	4	4	4	4	4	4	4	2	36	90,00
R28	3	3	3	2	3	3	3	3	4	3	30	75,00
R29	4	4	4	4	4	4	4	4	4	4	40	100,00
R30	3	3	4	2	4	2	2	2	4	2	28	70.00
R31	4	3	3	3	3	3	3	3	3	3	31	77,50
R32	4	2	4	4	3	3	3	4	3	2	32	80,00
R33	4	4	4	4	4	4	4	4	4	4	40	100,00
R34	3	3	4	4	2	2	2	3	2	2	27	67,50
R35	3	3	3	3	3	3	3	3	3	3	30	75,00
R36	4	4	4	4	4	4	4	4	4	4	40	100,00
R37	3	3	3	3	3	3	3	3	3	3	30	75,00
R38	3	3	2	2	3	3	3	3	3	3	28	70,00
R39	4	4	4	4	4	4	4	4	4	4	40	100,00
R40	4	3	2	2	3	2	2	3	2	3	26	65,00
R41	2	3	3	3	3	3	3	3	3	4	30	75,00
R42	2	2	2	4	2	2	2	2	2	4	24	60,00
R43	2	2	2	3	2	3	2	3	2	4	25	62,50
R44	3	4	4	4	4	4	3	3	4	4	37	92,50
R45	3	4	4	4	3	4	4	3	4	4	37	92,50
R46	3	4	4	4	3	3	3	3	4	3	34	85,00
R47	2	4	4	4	4	4	3	3	4	4	36	90,00
R48	3	4	3	4	4	4	3	3	3	3	34	85,00
R49	3	4	4	4	3	3	4	4	4	4	37	92,50
R50	3	4	3	3	4	4	3	4	4	4	36	90,00
R51	3	3	3	3	4	3	4	4	4	4	35	87,50
R52	3	3	4	4	4	4	4	3	3	4	36	90,00
R53	3	4	3	4	4	4	3	3	4	3	35	87,50
R54	4	4	3	3	4	4	4	3	3	3	35	87,50
R55	4	4	3	3	4	4	3	4	3	4	36	90,00
R56	3	4	4	4	4	4	4	4	4	4	39	97,50
R57	3	4	4	4	4	4	4	4	4	4	39	97,50
R58	3	4	4	4	3	4	4	4	4	4	38	95,00
R59	4	4	4	4	3	3	3	3	4	4	36	90,00
R60	3	4	4	4	4	4	3	4	3	4	37	92,50
R61	3	4	4	4	3	4	3	4	4	4	37	92,50

R62	4	4	3	3	3	3	4	4	3	4	35	87,50
R63	3	4	4	3	4	4	3	4	3	4	36	90,00
R64	3	4	4	4	4	3	3	4	4	3	36	90,00
R65	3	3	3	3	3	3	3	3	3	3	30	75,00
R66	3	3	3	3	4	3	4	4	3	3	33	82,50
Total Score												5192.50
Average score												78,6742

$$\text{Average score} = \frac{\text{Total Score}}{\text{Total Respondents}} = \frac{5192.50}{66} = 78{,}6742$$

Content Validity To ensure the reliability of the medical information presented, a content validation test was conducted with seven pharmaceutical experts. The experts rated the accuracy, clarity, and completeness of the drug information using a questionnaire. The results were statistically analyzed using the Aiken's V content validity index, as summarized in Table 3. The analysis yielded scores ranging from 0.75 to 0.85 for all questionnaire items. These scores fall into the "High" and "Very High" validity categories, confirming that the application's content is considered accurate and trustworthy from a professional standpoint.

Table 3. Expert Content Validation (Aiken's V)

Question item	V value	Conclusion
K1	0.78	High validity
K2	0.82	Very high validity
K3	0.78	High validity
K4	0.75	High validity
K5	0.85	Very high validity
K6	0.75	High validity
K7	0.82	Very high validity
K8	0.85	Very high validity
K9	0.78	High validity
K10	0.85	Very high validity

5.2 Graphical Results

The system's architecture is graphically represented in Figure 1, illustrating the three-tier structure that separates the client application, backend services, and the external LLM engine. This design ensures modularity and scalability.

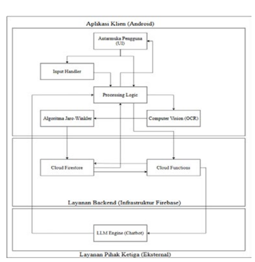

Figure 1.System architecture diagram

Figure 2 provides a visual representation of the application's user interface flow. The workflow is designed to be linear and intuitive, starting from the home screen (a), moving through the image scanning and cropping process (b), and culminating in a clear results screen (c).

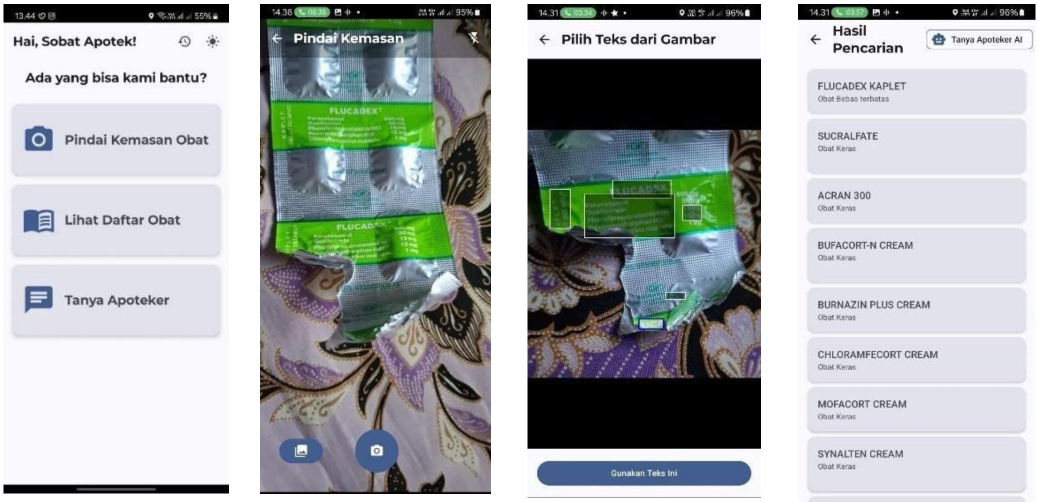

Figure 2.Application user interface flow showing (a) Home screen,

(b) Scanning process, and cropping process (c) Search results screen

To better visualize the system's robustness, Figure 3 presents the accuracy results from Table 1 in a graphical format. The bar chart clearly shows the system's high performance across all tested conditions.

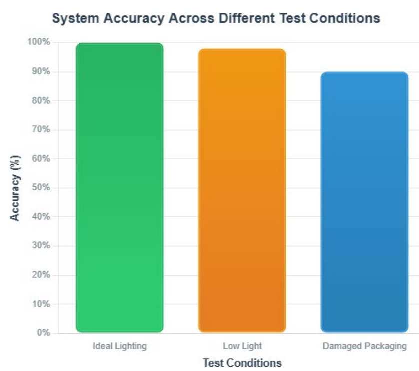

Figure 3.Bar chart of system accuracy across different test conditions

The following figures present the graphical results from the performance analysis conducted using the Android Studio Profiler. As shown in Figure 4, when the application is in an idle state on the home screen without user interaction, it demonstrates high efficiency. CPU usage remains minimal at approximately 0-1%, and memory consumption is stable at around 256 MB, indicating no unnecessary background processes are draining device resources.

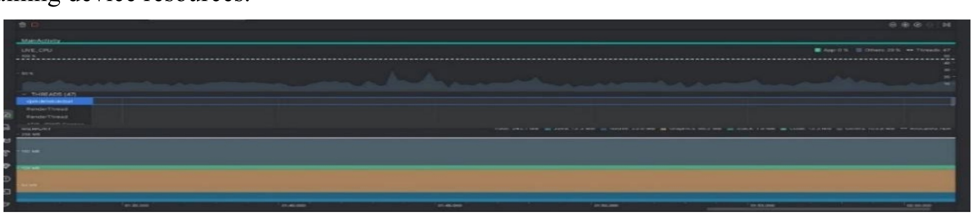

Figure 4. Performance profile of the application in an idle state.

Figure 5 captures the resource-intensive Scan OCR process. A distinct and expected peak in both CPU usage (averaging ~15% with a momentary peak up to ~30%) and memory allocation (peaking at ~585 MB) is visible. This spike corresponds to the on-device machine learning model execution. The resource usage is temporary and returns to normal levels after the process completes, confirming it is an acceptable trade-off for offline functionality.

Figure 5.Performance profile during the Scan OCR process

In contrast, Figure 6 shows the performance during a chatbot interaction. The resource footprint is minimal, with a low CPU usage of ~2-5% and stable memory at ~285 MB. This result validates the backend-heavy architecture, where the computationally expensive task of language model generation is offloaded from the user's device, keeping the client application light and responsive.

Figure 6. Performance profile during chatbot interaction

Finally, Figure 7 illustrates the performance while navigating and scrolling through lists. The low and stable resource usage, with CPU at ~1-3% and memory at ~350 MB, demonstrates efficient UI rendering and data management, contributing to a smooth and responsive user experience during standard browsing activities.

Figure 7. Performance profile during UI navigation and scrolling

5.3 Proposed Improvements

Based on the research findings and identified limitations, several improvements are proposed for future development to enhance the application's capabilities and practical value:

- Database Expansion and Enrichment: The current database of 300 drugs, while sufficient for validation, should be significantly expanded to include a more comprehensive range of medications available in the market. This would not only increase the identification success rate but also directly enhance the contextual relevance and accuracy of the chatbot's responses, a point also highlighted by expert reviewers.
- Integration of Image Classification: To move beyond text-only identification, future versions could integrate computer vision models to recognize the physical form of medications (e.g., tablets, capsules). This would add a crucial layer of verification for users and address a key limitation of the current system.
- Advanced OCR Pre-processing and Model Tuning: To improve accuracy on text with artistic fonts or in extreme lighting conditions, advanced image pre-processing techniques (e.g., binarization, noise reduction) could be implemented. Furthermore, fine-tuning the OCR model specifically on a dataset of drug packaging could yield significant performance gains.
- Development of User-Centric Practical Features: Based on expert feedback, the application's utility could be greatly enhanced by adding features that support daily medication management. This includes functionalities such as a configurable medication reminder system or a barcode scanning feature for faster and more accurate product lookup.
- Exploration of Collaborative Potentials: Future work could explore potential collaborations with external parties such as retail pharmacies. Such a partnership could enable valuable features like real-time stock availability checks or price comparisons at nearby locations, transforming the application into a more integrated healthcare tool.

5.4 Validation

The validity of this research is supported by a robust and comprehensive evaluation protocol that addresses multiple facets of the application's quality. Functional validation was unequivocally established through Black Box testing, where the application achieved a 100% pass rate across all 20 core use cases, confirming high system stability and adherence to functional requirements. Content validation was rigorously performed by seven pharmaceutical experts. The statistical analysis using Aiken's V, resulting in all items scoring between 0.75 and 0.85, provides strong quantitative proof that the medical information presented is considered to have "High" to "Very High" validity by professionals in the field. Usability validation, conducted with 66 users via the System Usability Scale (SUS), yielded a score of 78.67. This strong score provides quantitative evidence that the application is easy to use, effective, and well-received by its target audience. Finally, non-functional validation through performance and compatibility testing confirmed that the application is efficient in its resource consumption and runs reliably across a range of modern and legacy Android devices. Together, these six distinct validation methods provide a holistic and convincing proof of the system's overall effectiveness, reliability, and quality.

6. Conclusion

This research successfully developed and validated a functional, reliable, and effective mobile application for identifying medicine packaging. The system's architecture, which synergistically combines on-device OCR, the Jaro- Winkler correction algorithm, and a controlled LLM chatbot, proved highly effective in addressing the core research problem. All initial objectives were met, as evidenced by key achievements including a 96% Top-5 identification accuracy, strong performance in adverse conditions, a "Good" usability score of 78.67, and high content validity as confirmed by pharmaceutical experts. The unique research contribution lies not just in the application of individual technologies, but in their successful integration into a holistic, hybrid, and user-friendly system that is both technically robust and practically valuable. This study provides a validated technological solution that can help reduce medication errors and enhance public access to reliable health information. The findings confirm that the proposed model is a viable and powerful tool, offering a strong foundation for future advancements in mobile health technology aimed at improving patient safety.

References

Aldi, F., Nozomi, I., and Soeheri, S., Comparison of Drug Type Classification Performance Using KNN Algorithm, Sinkron: *jurnal dan penelitian teknik informatika*, vol. 6, no. 3, pp. 1028-1034, 2022.

Apotek K-24, Available: https://www.k24klik.com/, Accessed on June 15, 2025

AryaRafa, D., Wahyuni, E. D., and Arifiyanti, A. A., Rancang Bangun Aplikasi Donor Darah Darurat Donora Berbasis Android Dengan Konsep Gamifikasi Menggunakan Kotlin, *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024.

Bhat, A., Gayatri K, Kavali, M. S., and Kumar, S. N., A Web Solution for Heart Disease Prediction From Lab Reports using OCR and K-Nearest Neighbors, *International Journal of Engineering Research & Technology*, vol. 13, no. 08, 2024.

Deshmukh, A. M., and Chalmeta, R., Validation of system usability scale as a usability metric to evaluate voice user interfaces, *PeerJ Computer Science*, vol. 10, e1918, 2024.

Febriyanti, N. M. D., Sudana, A. K. O., and Piarsa, I. N., Implementasi black box testing pada sistem informasi manajemen dosen, *Jurnal Ilmiah Teknologi Dan Komputer*, vol. 2, no. 3, pp. 535-544, 2021.

Halodoc, Available: https://www.halodoc.com/obat-dan-vitamin, Accessed on June 15, 2025.

Handoyo, E. D., Santoso, S., and Surjawan, D. J., Pengembangan Aplikasi Mobile Pemesanan dan Pembayaran Makanan Berbasis Cloud Storage, *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 8, no. 1, pp. 161-174, 2022.

Harahap, R. A. A. H., Pengembangan Aplikasi Android Untuk Konversi Gambar ke Teks dengan Flutter dan OCR Menggunakan Metode Jaro-Winkler, Doctoral dissertation, Universitas Nasional, 2024.

Holila, H., Pratama, A. R., Lestari, S. A. P., and Indra, J., INTRODUCTION NATIONAL IDENTIFICATION NUMBER AND NAME ON ID CARD USING OCR (OPTICAL CHARACTER RECOGNITION) METHOD,
Jurnal Teknik Informatika (Jutif), vol. 5, no. 4, pp. 1191-1196, 2024.

Ikatan Apoteker Indonesia., *Informasi Spesialite Obat*, 23rd Edition, Pharma Tekno Solusi: Jakarta, 2021 Mangundap,
J. J., Tasripan, T., and Kusuma, H., Sistem Pengenalan Text pada Kemasan Obat untuk Membantu Penyandang Tunanetra dengan Keluaran Suara, *Jurnal Teknik ITS*, vol. 11, no. 2, pp. A128-A133, 2022.

Munadhif, I., Putri, W. P., Hasin, M. K., Rinanto, N., and Adhitya, R. Y., Implementasi OCR Dan CNN Untuk Klasifikasi Penempatan Obat Berdasarkan Kelas Terapi Di Apotek, *Jurnal Elektronika dan Otomasi Industri*, vol. 11, no. 2, pp. 421-433, 2024.

Suartana, I. M., Perbandingan Jetpack Compose dan Flutter dalam Pengembangan dan Penggunaan Resource pada Aplikasi Link D Law, *Journal of Informatics and Computer Science (JINACS)*, pp. 617-629, 2025.

Thesing, T., Feldmann, C., and Burchardt, M., Agile versus waterfall project management: decision model for selecting the appropriate approach to a project, *Procedia Computer Science*, vol. 181, pp. 746-756, 2021.

Tresna, I. M. A., Dwiyansaputra, R., and Akhyar, H., Perbaikan Kesalahan Kata Menggunakan Kombinasi Jaro- Winkler & Jaccard Similarity, *Jurnal Teknologi Informasi, Komputer, dan Aplikasinya (JTIKA)*, vol. 7, no. 1, pp. 25-36, 2025.

Utami, L., Ilahi, D. P., and Ratih, A., Analisis Indeks Aiken untuk mengetahui Validitas Isi Instrumen Scinetific Habbits of Mind, *Journal of Research and Education Chemistry*, vol. 6, no. 1, pp. 59-59, 2024.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., and Wen, J. R., A Survey of Large Language Models, *arXiv preprint arXiv*:2303.18223, 2023.

Zulhijar, A., Basri, M., and Marlina, M., Implementasi Optical Character Recognition (OCR) untuk Otomasi Penghitungan Tagihan Listrik, *Jurnal Publikasi Manajemen Informatika*, vol. 3, no. 3, pp. 162-171, 2024.

Biographies

Dr. Andi Chairunnas is a dedicated academic with a background in computer science and education, who has served as Vice Rector for Student Affairs at Pakuan University since the end of September 2023. Born in Makassar, he completed his doctoral degree at Unpak in 2019. In addition to teaching, he actively mentor's students in the field of robotics and leads student development programs and other student activities.

Dr. Hasrul is a lecturer in Pakuan University for 13 years with background in economics and business and serve as head of departement planning dan developing university. An entreprenuership is another side activity cause it impacts to students to improve their knowledge and experience.

Dr. Muslim, M.Si. is an important academic figure in the Faculty of Social and Cultural Sciences, Pakuan University, serving as Dean of FISIB covering academic, administrative, and managerial leadership functions within the faculty.

Raihan Ramadian Purwanto, S. Kom recently completed his Bachelor of Computer Science from the Faculty of Mathematics and Natural Sciences at Pakuan University, Bogor, Indonesia. His final thesis project focused on the implementation of machine learning and optical character recognition on mobile platforms for social impact, specifically in the healthcare domain. His research interests include mobile application development, machine learning, computer vision, and the practical application of AI technologies to solve real-world problems.

Biyan Firmansyah, S.Kom, currently pursuing a Master of Education degree at Pakuan University in Bogor, is known as an active figure within the Unpak student community, particularly within the Wapalapa nature lovers community. He plays a key role on the university's event committee and is involved in technology research and environmental conservation.