

Receipt Scanning with EasyOCR and ChatGPT-4o in a Mobile Finance App: an Agile Kanban Approach

M. Fiqry Septiawan^{1✉}, Siska Anraeni², Ramdaniah³

^{1, 2, 3} Informatics Engineering, Computer Science, Universitas Muslim Indonesia, Indonesia

✉ Corresponding Author : M.Fiqry Septiawan (e-mail: fikrisep130902@gmail.com)

Article Information	ABSTRACT
Article History Received : August 07, 2025 Revised : August 26, 2025 Published : October 04, 2025	Technological advancements have provided convenience for Generation Z in managing finances; however, many are still not accustomed to recording their financial activities regularly. Shopping receipts, which should serve as proof of transactions, are often ignored or poorly managed, despite their important role in tracking expenses. Therefore, this research aims to develop an Android-based financial recording application capable of handling both manual input and automated recording through receipt scanning using Optical Character Recognition (OCR) technology. The findings indicate that ChatGPT-4o significantly outperforms EasyOCR by providing more consistent accuracy and faster, stable processing, making it a more reliable solution for receipt-based financial recording. Developed using the Agile Kanban method, the application was validated through alpha testing and proven to function properly across all features. Beyond practical benefits for users, this research also contributes to the financial technology literature by demonstrating the integration of large language models (LLM) to enhance OCR performance in mobile finance applications.
Keywords: <i>Agile Kanban; Finance Tracking; Generation Z; OCR; Receipt Scanning.</i>	

INTRODUCTION

Low financial literacy remains a major challenge in Indonesia, including among Generation Z, who tend to have consumptive behavior and adopt the You Only Live Once (YOLO) principle, which is to live for the present without thinking too much about the future (Asrun & Gunawan, 2024; Lohita et al., 2022). Financial literacy is the ability to understand financial concepts and apply that knowledge effectively to manage income, expenses, and plan personal budget (Echchakoui & Heppell, 2023). Good literacy can help individuals maintain a balance between income and expenses, plan budgets wisely, and support more accurate financial decision-making (S. Bhuvanapriya et al., 2024; Ulfah, 2024).

However, financial record-keeping is often considered impractical because it requires time and attention to detail, especially when done manually through simple note-taking apps or digital spreadsheets (Felia Putri & Nurlaila, 2022). To support better financial literacy, accurate transaction recording through physical evidence such as shopping receipts plays a crucial role in monitoring an individual's financial condition (Kosadi et al., 2021). In this context, receipt scanning with OCR technology offers a more efficient alternative to manual input, providing automated extraction of transaction details that can be integrated into personal finance applications.

Receipts serve not only as proof of transactions but also as an essential means of monitoring expenses and making price comparisons (Matsui et al., 2023; Sari et al., 2023). However, many individuals neglect to store or manage receipts properly, which leads to inaccurate

expenditure records. Prior research has proposed digital receipt management using Optical Character Recognition (OCR) to address this issue (Garcia & Claour, 2021). OCR enables automatic extraction of text from receipt images (Widiantara & Romli, 2024), and tools such as EasyOCR have been widely applied for multilingual receipt and document recognition (Anthony et al., 2024; JaidedAI, 2020). Nevertheless, conventional OCR libraries still face limitations when handling poor-quality inputs, such as faded or folded receipts, resulting in reduced accuracy and unstable performance. This gap highlights the need for more robust solutions, which motivates the integration of large language models to enhance the reliability of receipt scanning in financial applications.

Previous studies have applied OCR in financial applications using tools like Google Vision API and TesseractOCR, but results have shown accuracy limitations and sensitivity to image quality, especially in faded or folded conditions (Astuti & Wicaksana, 2018; Sari et al., 2023). To address these challenges, this study compares EasyOCR with a large language model approach using ChatGPT-4o, which offers greater potential for intelligent and stable receipt-based text extraction (Bardvall & Hassle, 2024; Kim et al., 2025).

Based on these issues, enhancing financial literacy requires practical tools that make financial recording easier and more reliable. Shopping receipts, as tangible evidence of transactions, play a vital role in ensuring accurate financial tracking. By leveraging OCR-based receipt scanning, users can reduce the burden of manual entry while promoting consistent record-keeping. Therefore, this research focuses on developing an Android-based financial recording application that combines manual and automatic receipt scanning features. The study specifically aims to design and implement the application, evaluate the performance of the EasyOCR+LLaMA pipeline under different receipt conditions (normal, faded, and folded), and compare its results with ChatGPT-4o as an alternative solution for more accurate and reliable receipt-based financial recording, all within an Agile Kanban development framework that supports iterative and flexible progress.

RESEARCH METHODS

System Architecture and Agile Kanban Development

This financial recording application was developed using a client-server architecture consisting of three main components: frontend, backend, and OCR services. The frontend was built on the Android platform as a user interface that provides manual recording features, receipt scanning, transaction history management, and financial report export in PDF format. The backend is responsible for processing transaction data and communicating with the OCR services that handle text extraction from receipt images. The OCR pipeline uses EasyOCR as the primary text detection engine and LLaMA as the postprocessing stage, which corrects character recognition errors and organizes the extraction results into a structured JSON format for processing by the application.

The application development process follows the Agile Kanban approach, which enables iterative and flexible workflow management (Alaidaros et al., 2020; Hakim & Samsoni, 2024; Hartono, 2022). Each planned feature is first added to the Backlog and moved to the To-Do list when ready for development. During the Work In Progress (WIP) phase, feature development continues until it reaches the Testing & Review stage. If testing is successful, the feature proceeds to the Deployment, Acceptance, and final Launch stages. Features that fail testing are returned to the improvement phase until they meet the established standards.

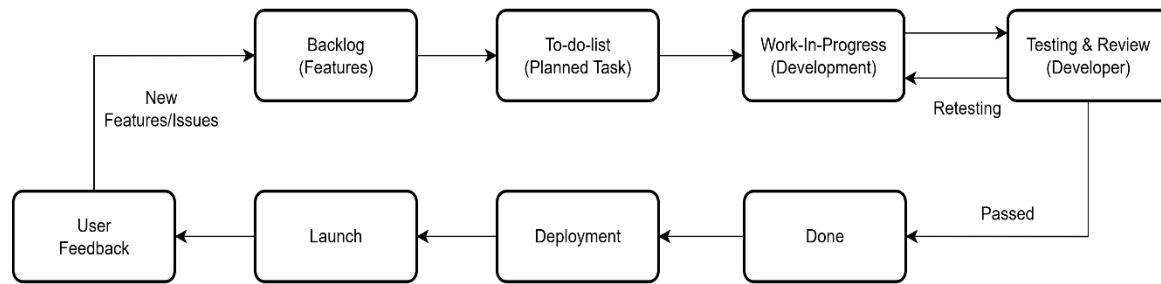


Figure 1. Workflow Agile Kanban

OCR Integration Workflow

The application supports two alternative OCR processing pipelines to extract structured information from shopping receipts. The first pipeline begins with an image preprocessing stage, including grayscale conversion, thresholding, and noise reduction to enhance input quality. After preprocessing, EasyOCR is employed as the main engine to detect and extract raw text from receipts. Since the raw OCR output often contains errors, the extracted text is further processed with the LLaMA language model using a prompt-based approach. At this postprocessing stage, LLaMA corrects spelling errors, normalizes the data format, and structures the transaction details into JSON containing the store name, date, items, and prices. This pipeline produces more accurate and consistent data, which is then subject to user validation.

As an alternative, the second pipeline leverages ChatGPT-4o as a multimodal end-to-end solution. In this pipeline, the receipt image is directly processed with a tailored prompt, enabling the model to both recognize text and generate structured JSON output in a single step. Unlike the EasyOCR+LLaMA pipeline, this approach does not require additional preprocessing or postprocessing, as ChatGPT-4o can internally handle noise and formatting inconsistencies. The extraction results from both pipelines are then subjected to validation, where automatic rules are applied and users can manually correct errors if necessary. Once validated, the information is transformed into a structured JSON format and subsequently stored securely with encryption to ensure confidentiality of sensitive financial data. Thus, both the EasyOCR+LLaMA and ChatGPT-4o pipelines converge on producing standardized digital financial records that are reliable and ready for further use.

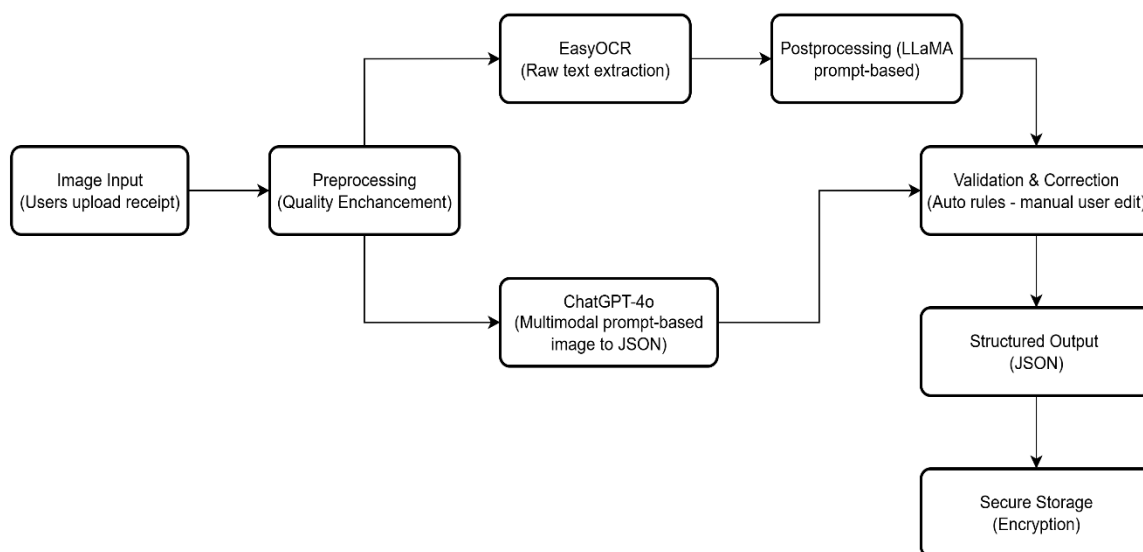


Figure 2. Workflow OCR Integration

Data and Testing Design

The application was tested using nine shopping receipt images collected from three different minimarkets (Alfamidi, Indomaret, Circle K), each representing three physical conditions: normal, faded ink, and folded. This data was used to test OCR performance, not for other purposes. Each receipt was processed using the EasyOCR+LLaMA and ChatGPT-4o pipeline, and the results were compared with manual reference text to evaluate accuracy and response speed.



Figure 3. Sample Test Receipts

Evaluation Metrics

Two main metrics are used in testing. Character accuracy is calculated using the Levenshtein Distance algorithm, which measures the similarity between OCR results and reference text. The calculation formula is:

$$\text{Character Accuracy} = \left(1 - \frac{d(a, b)}{\max(|a|, 1)}\right) \times 100\% \quad (1)$$

with a as the reference text (ground truth), b as the OCR result, and $d(a, b)$ as the Levenshtein edit distance. Response speed is measured based on the time (seconds) required for the application to generate usable JSON data from the moment it receives the receipt image.

System Functionality Test

In addition to OCR testing, black-box testing was conducted to ensure that all application features functioned as designed (Zen et al., 2024). This testing included user login, manual transaction recording, receipt scanning, transaction history management, and PDF report export. Each scenario was tested to validate that the system output matched expectations without examining the internal code.

RESULTS AND DISCUSSION

Application Overview and Key Features

The app is called FinancyQ, designed to make it easier for users to record and monitor their personal finances in a practical way, either through manual recording or automatic receipt scanning using OCR technology. The app features several key functions, including a login page for secure access, a dashboard displaying a summary of income and expenses, a manual input form, a receipt scanning feature with editable results, and the ability to export transaction reports into PDF format. The simple and intuitive user interface allows users to record and review transaction quickly.

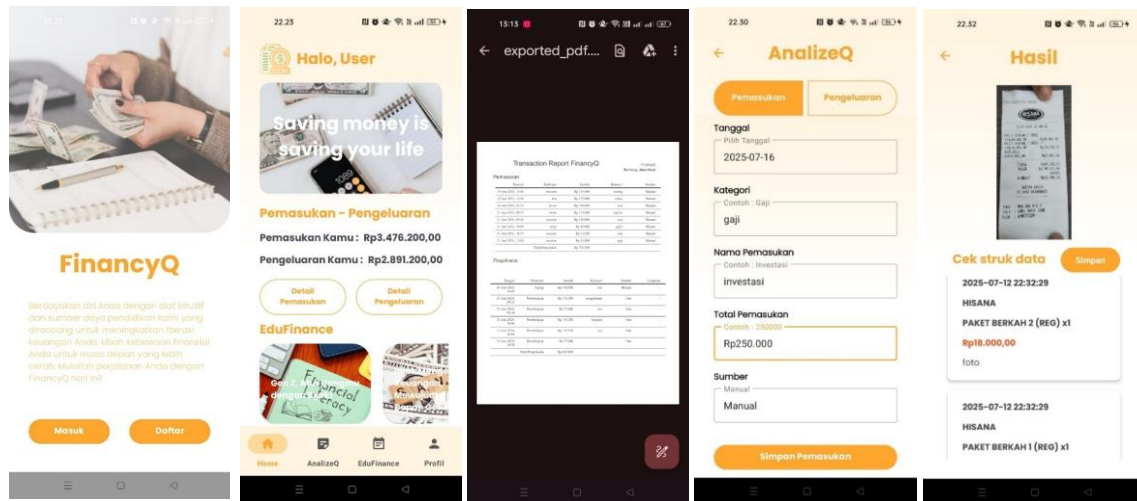


Figure 4. FinancyQ Application Interface.

OCR Accuracy Comparison

Table 1 and Figure 5 present the accuracy results of receipt scanning using two OCR pipelines. EasyOCR+LLaMA achieved 96.15% accuracy on normal Circle K receipts but dropped sharply to 63.77% on faded Indomaret receipts and also under folded conditions. By contrast, ChatGPT-4o remained highly consistent, reaching 100% in almost all cases, with only a slight decrease to 97.20% on folded Circle K receipts. These findings highlight the known limitation of traditional OCR in handling poor print quality, particularly in item and price recognition, consistent with previous studies (Bardvall & Hassle, 2024; Kim et al., 2025; Sowjanya & Vijaya Chamundeeswari, 2024). ChatGPT-4o, however, demonstrates superior robustness due to its contextual understanding and reasoning ability, enabling more reliable reconstruction of semi-structured receipts. Overall, this confirms the research objective: LLM-based models, especially ChatGPT-4o, provide higher accuracy and consistency than conventional OCR, reinforcing their potential as a more reliable solution for financial technology applications.

Tabel 1. Comparison of OCR Accuracy Across Minimarkets and Receipt Conditions.

Minimarket	Model	Normal (%)	Fading (%)	Folding (%)
Alfamidi	EasyOCR+LLaMA	80.17	66.39	89.92
	ChatGPT-4o	100	100	100
Circle K	EasyOCR+LLaMA	96.15	85.71	93.46
	ChatGPT-4o	100	100	97.20
Indomaret	EasyOCR+LLaMA	88.15	63.77	77.37
	ChatGPT-4o	98.52	100	100

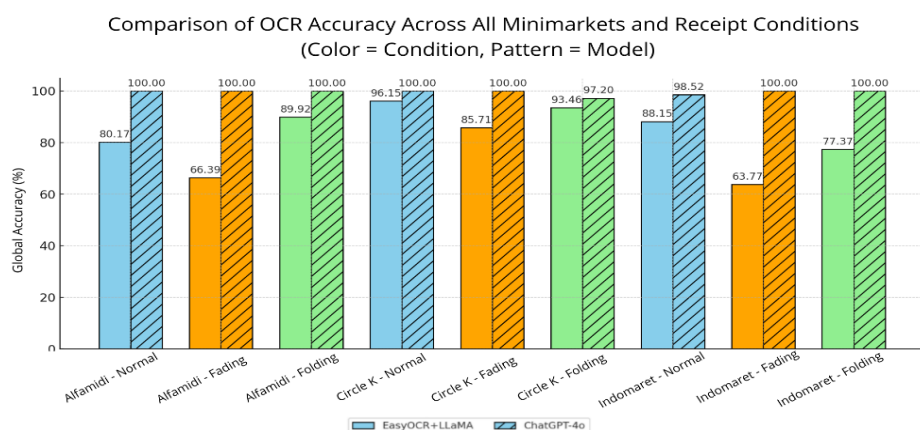


Figure 5. Bar Chart Comparing OCR Accuracy

Processing Time Comparison

In addition to accuracy, testing was also conducted to compare the processing speed between EasyOCR+LLaMA and ChatGPT-4o. The parameter measured was response time, which is the duration (in seconds) required by the application from the receipt upload process until the extracted text is ready to be displayed in the application. The test results are shown in Table 2.

Tabel 2. Comparison of OCR Response Speed Across All Minimarkets and Receipt Conditions.

Minimarket	Model	Normal (second)	Fading (second)	Folding (second)
Alfamidi	EasyOCR+LLaMA	15.60	8.69	14.60
	ChatGPT-4o	11.38	11.69	11.24
Circle K	EasyOCR+LLaMA	15.82	5.99	15.07
	ChatGPT-4o	11.11	10.01	9.85
Indomaret	EasyOCR+LLaMA	14.88	13.79	12.46
	ChatGPT-4o	10.63	9.57	12.46

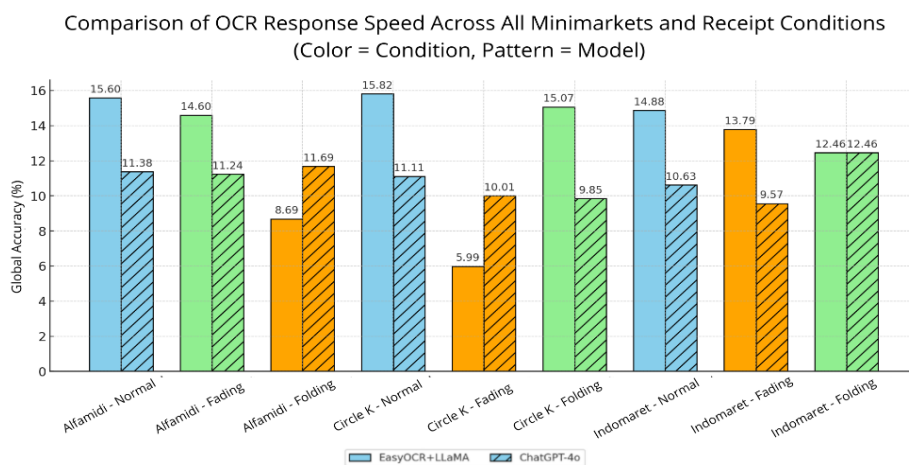


Figure 6. Bar Chart Comparing OCR Response Speed

The results show that ChatGPT-4o has faster and more stable processing times, ranging from 9.57 to 12.46 seconds under all receipt conditions. In contrast, EasyOCR+LLaMA showed variable processing times that tended to be longer, even reaching 15.82 seconds on normal receipts at Circle K. This variation was due to the need for additional preprocessing in the EasyOCR pipeline and post-extraction corrections, which were not required in ChatGPT-4o.

Kanban Contribution to Development Cycle

The development of the FinancyQ application was carried out using the Agile Kanban approach to support structured and flexible workflow management. This method helps the team monitor development progress visually, from the planning stage, through execution, validation, to feature completion. Each task is categorized into four main statuses: To Do, Work In Progress (WIP), Validate, and Complete, making the entire process easier to control and more transparent.

Kanban implementation is applied to various aspects of application development. During the UI/UX design phase, the Kanban board is used to design the initial interface, transaction input forms, profile pages, and receipt scanning features. During the backend development phase, the board facilitates scheduling modules such as transaction API development, user authentication, and database integration. Frontend development is managed in a similar manner, ensuring that each interface component is tested and refined before publication. Finally, during the OCR integration phase, Kanban helps organize the technology research process, EasyOCR testing, migration to ChatGPT-4o, and feature validation until optimal results are achieved.

Visualizations of this method's implementation can be seen in Figures 7 to 10, which display Kanban boards for each project section (UI/UX, backend, frontend, and OCR). Through this approach, the development process runs more efficiently, ensuring that the application's core features are completed gradually, validated, and ready for use according to user needs.

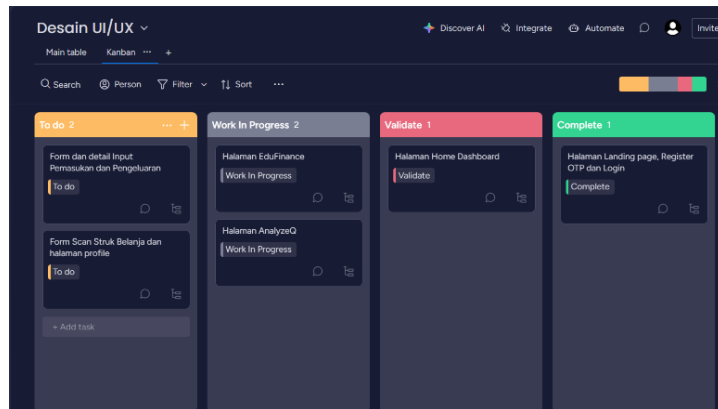


Figure 7. UI/UX Design Kanban Board

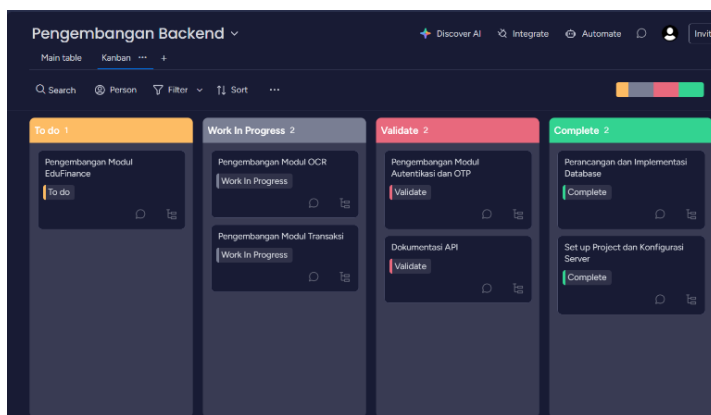


Figure 8. Backend Development Kanban Board

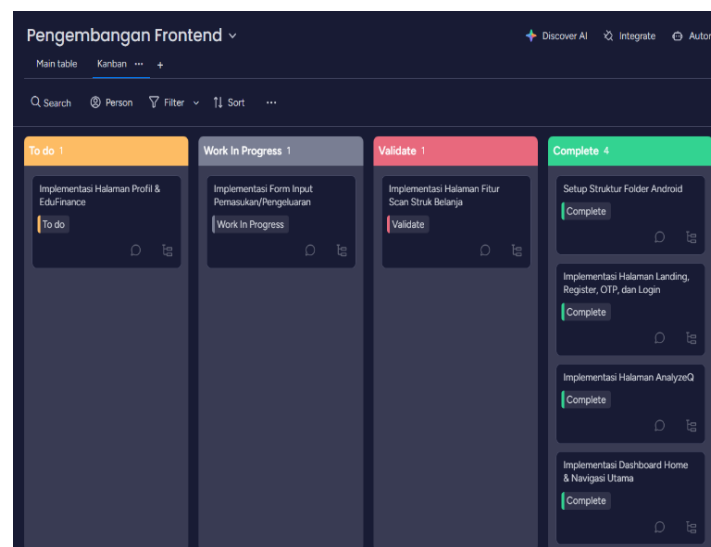


Figure 9. Frontend Development Kanban Board

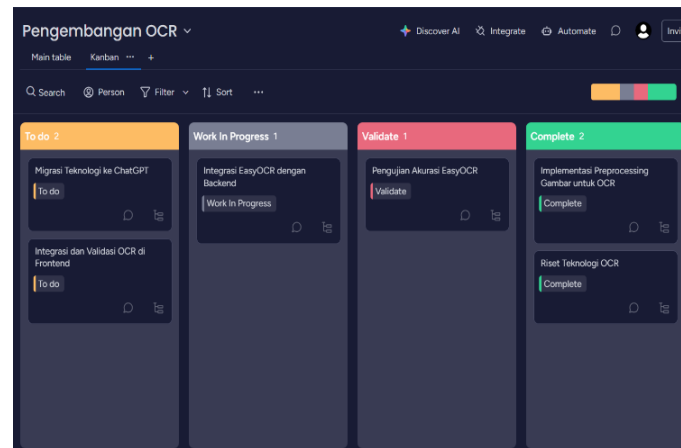


Figure 10. OCR Development Kanban Board

System Functionality Test Results

Functional testing of the system was conducted using a black-box testing approach to ensure that all application features functioned according to specifications. Testing covered key application modules such as registration and login, manual transaction input, shopping receipt scanning, transaction history management, and PDF report export. Each feature was tested with both valid and invalid input scenarios to verify the validation mechanisms and system outputs. The test results are summarized in Table 5, which shows that all features of the FinancyQ application function properly. The system can process valid inputs correctly, reject invalid data with appropriate error messages, and generate transaction reports without issues. Therefore, the application is deemed practical for use in supporting users' financial record-keeping.

Tabel 3. Black-box Testing

Features tested	Scenario Tests	Main Results	Status
Registration and Login	Valid and invalid input	Valid data is saved; incorrect input is rejected	Successful
Manual Input	All fields filled and Rp0 error	Valid data is saved; empty input/Rp0 is rejected	Successful
Receipt Scanning	Photo of receipt & non-receipt	Item data is read; error message appears if it's not a receipt	Successful
Transaction History	View, edit, delete	Data is displayed correctly; edit and delete functions work properly	Successful
Report Export	Export to PDF	PDF file is successfully generated without errors	Successful

CONCLUSION

This research successfully designed and developed FinancyQ, a mobile-based financial recording application that supports manual and automatic transaction recording through shopping receipt scanning. The system was initially built using EasyOCR combined with a post-processing model based on LLaMA. However, testing results showed that EasyOCR experienced a significant drop in accuracy on receipts with faded or folded conditions, with the lowest accuracy reaching

63.77%. As an alternative, ChatGPT-4o was tested and proven to provide significantly better and more consistent accuracy, ranging from 97.20% to 100% across various receipt conditions, with faster and more stable processing times between 9.57 and 12.46 seconds.

The Agile Kanban approach made a significant contribution to the application development process, enabling gradual feature refinement, earlier validation of OCR results, and efficient task management across UI/UX, backend, frontend, and OCR modules. As a result, the application is capable of meeting users' needs for practical and accurate financial recording, while leveraging shopping receipts as valuable transaction evidence.

Practically, this research demonstrates that integrating large language models like ChatGPT-4o into OCR-based financial tracking applications can improve text extraction performance compared to conventional OCR libraries. Further research could expand the dataset for more comprehensive testing, explore combined OCR-Large Language Model pipelines, and test system performance in real-world scenarios with more diverse receipt formats and languages.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the valuable guidance and support from our supervisors, Ms. Siska Nurnaeni and Ms. Ramdaniah, throughout the research and writing process. Their insightful advice, constructive feedback, and continuous encouragement significantly contributed to the successful completion of this work.

REFERENCES

- Alaidaros, H., Omar, M., Romli, R. B., & Romli, R. (2020). Developing an Improved Model of the Agile Kanban Method (i-KAM): A Research Design and Preliminary Results. *International Journal of Multidisciplinary Sciences and Advanced Technology*, 1(10), 32–38. <https://www.ijmsat.com/archives/ijmsat-volume-1-issue-10>
- Anthony, Herman, & Yulianto, A. (2024). Pengembangan Sistem Pengenalan Plat Nomor Indonesia Menggunakan YOLOv8 dan EasyOCR. *Jurnal Ilmiah Komputasi*, 23(4), 571–578. <https://doi.org/10.32409/jikstik.23.4.3659>
- Asrun, N. A., & Gunawan, A. (2024). Pengaruh Gaya Hidup dan Media Sosial terhadap Perilaku Konsumtif Generasi Z di Kota Medan dengan Literasi Keuangan sebagai Media Intervening. *Jurnal Manajemen Bisnis Dan Keuangan*, 5(1), 173–186. <https://doi.org/10.51805/jmbk.v5i1.205>
- Astuti, Y., & Wicaksana, K. K. (2018). Rancang Bangun Sistem Pemindaian Struk Belanja untuk Mendapatkan Rincian Belanja. *Seminar Nasional Teknologi Informasi Dan Multimedia*, 6(1), 37–42. <https://ojs.amikom.ac.id/index.php/semnasteknomedia/article/view/2111>
- Bardvall, M., & Hassle, I. (2024). *Automating Invoice Recognition: A Comparative Study of Large Language Models and OCR/ML Technologies* [KTH Royal Institute of Technology]. <https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-351177>
- Echchakoui, S., & Heppell, H. (2023). Financial Literacy: A Bibliometric Literature Review. *Advances in Management and Applied Economics*, 13(3), 129–161. <https://doi.org/10.47260/amae/1335>
- Felia Putri, D., & Nurlaila, N. (2022). Analisis Sistem Pencatatan Manual Laporan Keuangan Terhadap Kinerja Akuntan Di Perusahaan Umum Daerah Pasar Kota Medan. *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, Dan Pendidikan*, 1(6), 763–770. <https://doi.org/10.54443/sibatik.v1i6.90>
- Garcia, M. B., & Claour, J. P. (2021). Mobile Bookkeeper: Personal Financial Management Application with Receipt Scanner Using Optical Character Recognition. *2021 1st Conference*

- on Online Teaching for Mobile Education (OT4ME), 15–20. <https://doi.org/10.1109/OT4ME53559.2021.9638794>
- Hakim, A., & Samsoni. (2024). Penerapan library Tesseract Optical Character Recognition dalam fitur reimbursement menggunakan metode Kanban (Studi Kasus: PT. Kayana Nusa Teknologi). 3(3), 600–607. <https://www.journal.mediapublikasi.id/index.php/oktal/article/view/2434>
- Hartono, R. (2022). Penerapan Kanban Model Sebagai Metode Perancangan Sistem Informasi (Studi Kasus: Pemetaan Sekolah SMA/K/MA Kota Tasikmalaya). *JURNAL PETIK*, 8(1), 27–34. <https://doi.org/10.31980/jpetik.v8i1.1252>
- JaiedAI. (2020). *EasyOCR: Ready-to-use OCR with 80+ supported languages*. GitHub. <https://github.com/JaiedAI/EasyOCR>
- Kim, S., Baudru, J., Ryckbosch, W., Bersini, H., & Ginis, V. (2025). Early evidence of how LLMs outperform traditional systems on OCR/HTR tasks for historical records. <http://arxiv.org/abs/2501.11623>
- Kosadi, F., Ginting, W., & Merliana, V. (2021). Digital Receipts of Online Transactions in the Reconciliation Process and the Preparation of Financial Reports. *Journal of Indonesian Economy and Business*, 36(1), 31–50. <https://doi.org/10.22146/jieb.59884>
- Lohita, V. A. K., Suprpto, Wi., & Sahetapi, W. L. (2022). Generasi Z dalam Memanjakan Diri di Restoran All You Can Eat. *Jurnal Ilmiah Manajemen Dan Bisnis (JIMBis)*, 1(2). <https://doi.org/10.24034/jimbis.v1i2.5376>
- Matsui, A., Kobayashi, T., Moriwaki, D., & Ferrara, E. (2023). Detecting multi-timescale consumption patterns from receipt data: a non-negative tensor factorization approach. *Journal of Computational Social Science*, 6(2), 1179–1192. <https://doi.org/10.1007/s42001-020-00078-5>
- S. Bhuvanapriya, K. K. Sreedeeve, B. Vaidianathan, & S. Saranya. (2024). The Essential Role of Expense Tracking in Personal and Business Financial Management. *American Journal of Economics and Business Management*, 7(12), 1637–1655. <https://doi.org/10.31150/ajebm.v7i12.3171>
- Sari, R., Musa Adi, I., & Hidayati, A. (2023). Personal Track Your Cash: Prototipe Aplikasi Pembacaan Struk 0020 Belanja Menggunakan OCR dan Google Vision. *Prosiding Seminar Inovasi Vokasi*, 2(1), 565–573. <https://prosiding.pnj.ac.id/sniv/article/view/436>
- Sowjanya, S., & Vijaya Chamundeeswari, V. (2024). *Information Extraction Using RPA and Generative AI from Unstructured Documents: A Case of Invoices* (pp. 250–264). https://doi.org/10.1007/978-3-031-69986-3_19
- Ulfah. (2024). Pengaruh Literasi Keuangan terhadap Pengelolaan Keuangan Pribadi. *Jurnal Manajemen, Akuntansi Dan Pendidikan (JAMAPEDIK)*, 1(2), 233–240. <https://doi.org/10.59971/jamapedik.v1i2.53>
- Widiantara, I. G., & Romli, M. A. (2024). Bidirectional Long Short-Term Memory untuk Ekstraksi Informasi pada Struk Belanja. *Voteteknika (Vocational Teknik Elektronika Dan Informatika)*, 12(4), 411. <https://doi.org/10.24036/voteteknika.v12i4.130990>
- Zen, M., Irwan, Hafni, & Ananda, M. D. P. (2024). Implementasi dan Pengujian Menggunakan Metode BlackBox Testing Pada Sistem Informasi Tracer Study. *Bulletin of Computer Science Research*, 4(4), 327–340. <https://doi.org/10.47065/bulletincsr.v4i4.359>