iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.82235

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Smart Retail Analytics: Automated Invoice Intelligence System

Prajwal Gaikwad[1], Aryan Hase[2], Omkar Chaskar[3], Suraj Gore[4], Kapil Tamhane[5]
[1, 2, 3, 4]*Dept. of Computer Engineering Pravara Rural Engineering College Loni, India*
[5]*Assistant Professor, Dept. of Computer Engineering Pravara Rural Engineering College Loni, India*

Abstract: ***The growing volume of invoice documents in retail and financial sectors, often available in unstructured and semi-structured formats, has made traditional processing methods inefficient and error-prone, while manual data entry further increases delays and inaccuracies in financial operations. To ad-dress these challenges, this paper presents an automated invoice intelligence system that integrates Optical Character Recognition (OCR), Natural Language Processing (NLP), and Robotic Process Automation (RPA) for end-to-end invoice processing. The pro-posed system extracts key information such as invoice numbers, vendor details, GST information, dates, total amounts, and line items from scanned images and PDF documents, converting them into a structured format suitable for storage and analysis. A key contribution of this work is the integration of data extraction with analytical visualization, enabling organizations to derive meaningful insights from financial data. Additionally, the system incorporates workflow monitoring to evaluate performance and identify inefficiencies. Experimental results demonstrate that the proposed approach improves extraction accuracy, reduces man-ual effort, and minimizes processing time, thereby providing a scalable and efficient solution for intelligent invoice management.***
Keywords: Optical Character Recognition (OCR), Natural Language Processing (NLP), Robotic Process Automation (RPA), Invoice Processing, Document Automation, Data Extraction, Business Intelligence, Retail Analytics, Data Visualization.

I. INTRODUCTION

In recent years, the rapid expansion of digital transactions in the retail and financial sectors has led to a significant increase in the volume of invoice data generated on a daily basis. Despite advancements in digital technologies, a considerable portion of invoices still exist in unstructured or semi-structured formats, such as scanned documents, images, and PDF files. Traditional invoice processing methods rely heavily on manual data entry, which is time-consuming, prone to errors, and inefficient when handling large-scale data. These limitations affect data accuracy, processing speed, and overall operational efficiency.

To overcome these challenges, intelligent automation tech-niques have emerged as effective solutions for modern in-voice management. Technologies such as Optical Character Recognition (OCR), Natural Language Processing (NLP), and Robotic Process Automation (RPA) enable automated extrac-tion, interpretation, and processing of invoice data. OCR con-verts textual information from images into machine-readable form, while NLP facilitates the identification and classifica-tion of relevant invoice fields. In addition, RPA automates repetitive tasks involved in data handling and workflow ex-ecution, thereby reducing manual intervention and improving efficiency. The integration of these technologies enables the devel-opment of smart invoice processing systems that go beyond simple data extraction by incorporating analytical capabilities. By transforming raw invoice data into structured formats, orga-nizations can achieve efficient storage, retrieval, and analysis. Furthermore, data visualization tools allow stakeholders to gain meaningful insights into financial operations, supporting better decision-making and helping identify trends, anomalies, and inefficiencies. In this paper, a Smart Retail Analytics system is proposed to automate invoice intelligence using a combination of OCR, NLP, and RPA techniques. The system is designed to accu-rately extract key information from diverse invoice formats and convert it into structured data for further analysis. Additionally, workflow monitoring mechanisms are incorporated to evaluate system performance and enhance operational efficiency. The proposed approach aims to reduce manual effort, improve accuracy, and provide a scalable solution for intelligent invoice management in retail environments.

II. LITERATURE REVIEW

Several studies have explored the automation of invoice processing using Optical Character Recognition (OCR) and rule-based techniques. Early approaches primarily relied on OCR systems to extract textual information from scanned financial documents, which significantly reduced manual ef-fort. However, these methods often depended on predefined templates, limiting their ability to handle invoices with varying layouts and formats.

Rule-based systems demonstrated good performance on structured documents but showed reduced accuracy when applied to semi-structured or unstructured data. To address these limitations, researchers introduced Nat-ural Language Processing (NLP) techniques for more flexi-ble information extraction. NLP-based approaches enable the identification of key invoice fields such as invoice numbers, dates, and vendor details, even in unstructured documents. In particular, Named Entity Recognition (NER) models have been widely used to improve the accuracy of extracting financial entities. While these methods provide better adaptability com-pared to traditional approaches, they are generally limited to data extraction and do not incorporate end-to-end workflow automation.

Machine Learning (ML) and deep learning techniques have further enhanced the capabilities of invoice process-ing systems. Supervised learning models have been applied for invoice classification and data extraction, offering im-proved adaptability across different formats. Additionally, deep learning-based text recognition methods have demonstrated strong performance, especially when dealing with low-quality or noisy scanned documents. Despite these advantages, such approaches often require large training datasets and signifi-cant computational resources, which may limit their practical deployment.

Robotic Process Automation (RPA) has been widely adopted to automate repetitive tasks involved in invoice pro-cessing, such as data entry and validation. By integrating RPA with OCR, partially automated systems have been developed to reduce processing time and manual intervention. However, many of these systems remain rule-based and lack the ability to adapt dynamically to diverse invoice formats or perform intelligent decision-making.

In addition to extraction and automation, Business Intelli-gence (BI) and process mining techniques have been used to enhance data analysis and workflow monitoring. BI tools en-able the visualization of financial data, allowing organizations to derive actionable insights, while process mining techniques help identify inefficiencies in operational workflows. Neverthe-less, these techniques are often applied independently and are rarely integrated with automated invoice processing systems. Overall, existing research indicates that while OCR, NLP, ML, and RPA have been effectively utilized in isolation, there is still a lack of a unified system that combines these technologies into a comprehensive framework. Most existing solutions focus on specific components such as data extraction or task automation, rather than providing a complete end-to-end solution. This highlights the need for an integrated system capable of handling extraction, automation, monitoring, and analysis within a single platform.

III. PROBLEM STATEMENT

Invoice processing remains a critical component of financial operations; however, it continues to face several practical challenges in real-world applications. Organizations often deal with invoices in multiple formats, including scanned images, PDF documents, and semi-structured layouts, which makes consistent data extraction complex and time-consuming. De-spite the adoption of existing digital solutions, achieving high accuracy and efficiency in invoice processing remains a significant challenge.

A major limitation is the continued reliance on manual processing, where human intervention is required for data entry and verification. This not only increases processing time but also introduces errors that affect data reliability. In addi-tion, the variability in invoice formats further complicates the extraction process, as traditional systems struggle to generalize across diverse document structures. Another important issue is the lack of analytical capabilities in existing systems, which primarily focus on data storage rather than generating mean-ingful insights such as spending trends or vendor performance analysis. Furthermore, many processes involved in invoice handling, including data extraction and validation, are not fully automated, resulting in reduced operational efficiency and slower workflows.

These challenges have a direct impact on business perfor-mance. In the retail sector, it is estimated that organizations can lose up to 15% of their annual revenue due to inefficient inventory management, including overstocking and stockouts caused by the lack of real-time predictive insights. This highlights the need for an intelligent system that not only automates invoice processing but also supports data-driven decision-making.

To address these issues, this work aims to develop a com-prehensive and intelligent invoice management system with the following objectives. First, to automate invoice processing using AI-enabled OCR techniques in order to eliminate manual data entry and improve accuracy. Second, to incorporate predictive analytics for inventory management, enabling the forecasting of stockouts in advance. Third, to provide real-time financial insights, including ledger tracking and profitability analysis, to support better financial management. Finally, to integrate intelligent decision-support capabilities that generate actionable business recommendations based on live financial data.

The proposed approach is designed to improve efficiency, reduce manual effort, and provide a scalable solution for modern retail and financial environments.

IV. PROPOSED SYSTEM AND ARCHITECTURE

The proposed system presents an intelligent and automated solution for invoice data extraction and analysis by integrating Optical Character Recognition (OCR), Natural Language Pro-cessing (NLP), Robotic Process Automation (RPA), and Busi-ness Intelligence techniques. It is designed to process invoices in various formats, including PDF files, scanned images, and digital documents, and transform them into structured data suitable for further processing and analysis.

A key contribution of the proposed approach is the inte-gration of real-time analytics with invoice processing. Unlike conventional systems that focus primarily on data extraction, the proposed system enables users to derive meaningful in-sights from financial data. In addition, it incorporates process monitoring capabilities to evaluate system performance and identify operational inefficiencies. The combination of data extraction, automation, analytics, and monitoring within a unified framework enhances the overall effectiveness of the system.

The system is composed of several interconnected compo-nents that work collaboratively to achieve end-to-end automa-tion. Initially, the input module accepts invoice documents in multiple formats. These documents are then processed by the OCR engine, which converts unstructured content into machine-readable text. The extracted text is further analyzed by the NLP processing unit to identify and extract key invoice attributes such as invoice number, vendor information, date, and total amount. The processed data is subsequently stored in a structured database, enabling efficient storage and retrieval. To enhance operational efficiency, an RPA module is in-corporated to automate repetitive tasks such as data valida-tion and data entry, thereby minimizing manual intervention. Furthermore, an analytics dashboard provides visualization and reporting capabilities, allowing users to analyze invoice data, identify patterns, and gain actionable insights. A process monitoring component is also integrated to track workflow execution and detect performance bottlenecks.

The overall workflow begins with the submission of invoice documents through the input module. The OCR engine ex-tracts textual content, which is then processed by the NLP unit to identify relevant information. The extracted data is stored in the database, after which RPA mechanisms automate subsequent processing steps. Finally, the analytics dashboard and monitoring module provide insights and performance evaluation. This structured workflow ensures efficient and accurate processing of invoice data.

The system is implemented using modern technologies that support scalability, automation, and efficient data handling. As a result, it significantly improves processing accuracy, reduces execution time, and provides a reliable solution for intelligent invoice management.

A. System Architecture

The system architecture of the proposed Intelligent Invoice Data Extraction and Analytics System is designed in a modular and layered manner to ensure efficient and seamless processing of invoice data. The architecture facilitates the transformation of unstructured invoice documents into structured information and meaningful insights through a sequence of interconnected layers.

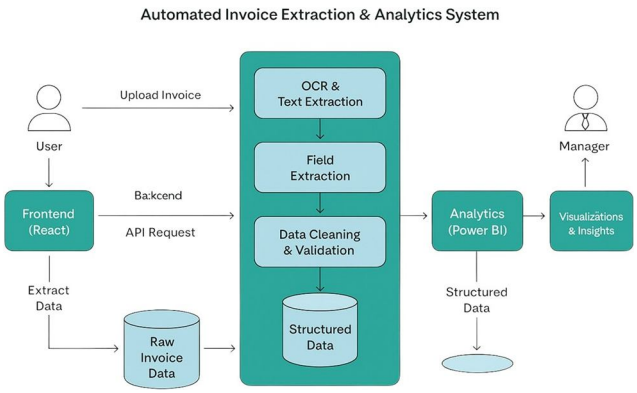


Fig. 1. System Architecture

As illustrated in Fig. 1, the process begins with the input layer, which handles invoice documents in multiple formats such as PDF files, scanned images, and digital records. These documents are then forwarded to the OCR layer, where Optical Character Recognition techniques are applied to extract textual content and convert it into machine-readable form.

The extracted text is subsequently processed by the NLP layer, which identifies and classifies key invoice attributes, including invoice number, vendor details, date, tax informa-tion, and total amount. This layer ensures accurate informa-tion extraction despite variations in document structure and formatting.

The structured data is then stored in a centralized database, which supports efficient storage, retrieval, and management of invoice information. To support data-driven decision-making, an analytics dashboard is integrated into the system, providing visualization and reporting features that help users analyze financial data and identify trends.

In addition, the architecture incorporates an RPA-driven workflow to automate repetitive tasks such as data validation and process execution, thereby reducing manual effort and improving operational efficiency. A process monitoring mech-anism is also included to continuously track system perfor-mance, detect bottlenecks, and optimize workflow execution. Overall, the proposed architecture ensures a seamless flow of data across all components, resulting in a scalable, flexible, and efficient system for intelligent invoice processing and analytics.

V. METHODOLOGY

The proposed system follows a structured and sequential methodology to automate invoice processing and enable intelligent data analysis. The workflow integrates data acquisition, preprocessing, extraction, validation, storage, automation, and analytics into a unified pipeline to ensure efficient and accurate processing of invoice data. The overall process flow, illustrated in Fig. 2, begins with data acquisition and concludes with data visualization and decision support.

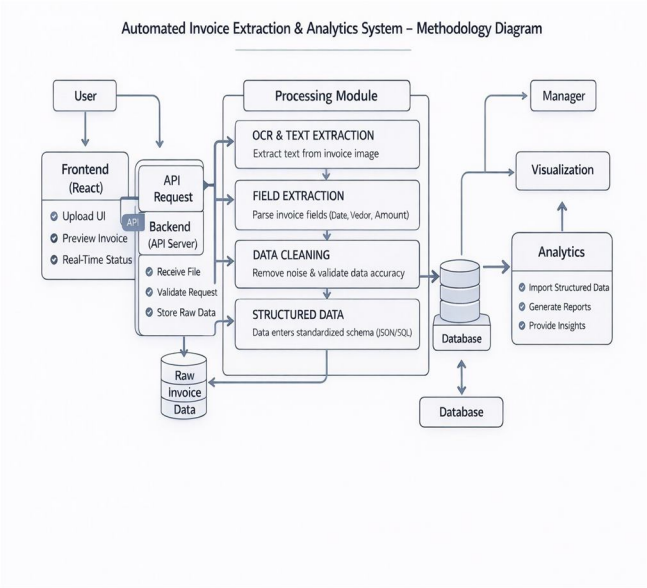


Fig. 2. Methodology

A. Data Acquisition

In this stage, invoice data is collected from multiple sources in various formats, including PDF files, scanned documents, and digital records. These invoices may vary significantly in terms of layout, structure, and quality depending on the source. Therefore, the system is designed to handle both structured and unstructured data without requiring manual preprocessing. This stage establishes a consistent input base for subsequent processing steps, directly influencing the overall accuracy of the system.

B. Data Preprocessing

The collected documents undergo preprocessing to enhance their quality prior to text extraction. Since input files may contain noise, distortions, low resolution, or uneven lighting, several image processing techniques such as noise removal, resizing, deskewing, and grayscale conversion are applied. These operations improve text clarity and standardize docu-ment formats, thereby increasing the accuracy of the OCR process.

C. Text Extraction

Following preprocessing, the OCR module extracts textual information from the documents and converts it into machine-readable format. This process involves recognizing characters, words, and structural patterns within the invoice, resulting in a digital representation of the original content. The extracted text serves as the foundation for further processing and analysis.

D. Information Extraction

In this stage, Natural Language Processing techniques are applied to identify and extract key invoice attributes, including invoice number, vendor details, date, GST information, and total amount. The system analyzes textual patterns, keyword positions, and contextual relationships to accurately capture relevant information. The extracted data is then organized into a structured format for downstream processing.

E. Data Validation

The extracted information is validated using predefined rules and logical checks to ensure accuracy, consistency, and com-pleteness. This includes verifying field formats, identifying missing or inconsistent values, and removing duplicate entries. The validation process enhances data integrity and ensures that only reliable data is forwarded for storage and analysis.

F. Data Storage

Validated data is stored in a structured database designed for efficient organization, retrieval, and management of invoice records. This structured storage approach supports scalability and ensures seamless integration with analytical and reporting modules.

G. Automation

Robotic Process Automation is utilized to automate repet-itive tasks such as data entry, workflow execution, and sys-tem integration. This reduces manual intervention, minimizes errors, and improves overall process efficiency by ensuring consistent execution of routine operations.

H. Data Analysis

The stored data is analyzed to extract meaningful insights related to financial transactions, spending patterns, and poten-tial anomalies. Analytical techniques are applied to identify trends and irregularities, enabling improved financial monitor-ing and decision-making.

I. Data Visualization

The final stage presents the analyzed data through inter-active dashboards and reports. These visualizations provide clear representations of financial metrics, trends, and key performance indicators, allowing users to interpret insights effectively and make informed business decisions.

VI. ALGORITHMS AND MODELS USED

To ensure accurate extraction, processing, and structuring of invoice data, the proposed system incorporates a combination of algorithms and computational models. These techniques enable efficient handling of unstructured data, improved in-terpretation of invoice content, and enhanced overall system performance and reliability.

A. Optical Character Recognition (OCR) Algorithm

The OCR algorithm is employed to extract textual informa-tion from invoice images and scanned documents. The OCR pipeline consists of three primary stages: preprocessing, text detection, and text recognition. In the preprocessing stage, image enhancement techniques such as grayscale conversion, noise reduction, and thresholding are applied to improve image quality. The processed image is then analyzed to detect text regions, followed by character recognition to convert the extracted content into machine-readable form. Post-processing techniques are further applied to organize the extracted text into structured financial data using rule-based methods.

B. Image Preprocessing Techniques

Image preprocessing plays a crucial role in improving the accuracy of text extraction. Various techniques, including noise reduction, binarization, skew correction, and edge detection, are applied to standardize the input documents. These op-erations enhance text clarity and reduce distortions, thereby improving the performance and reliability of the OCR process.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

C. Named Entity Recognition (NER) Model

Named Entity Recognition models are utilized to identify and categorize key information from the extracted text, such as invoice number, vendor name, date, GST details, and total amount. These models leverage contextual relationships and linguistic patterns within the text to accurately label relevant entities. The structured output generated by the NER model facilitates efficient downstream processing and analysis.

D. Pattern Matching Techniques

Pattern matching methods, including regular expressions and rule-based approaches, are used to identify specific data formats such as dates, invoice numbers, and monetary values. These techniques rely on predefined patterns and logical rules to extract structured information from unstructured text. The use of pattern matching enhances the consistency and accuracy of the extracted data.

E. Classification Models

Machine learning-based classification models are employed to categorize invoice components and improve data extraction accuracy. These models are trained on labeled datasets to rec-ognize patterns and variations in invoice formats. By learning from diverse data samples, the classification models enhance the system's ability to generalize across different invoice structures, resulting in improved performance and robustness.

VII. AI & ML TECHNIQUES USED

Intelligence (AI) and Machine Learning (ML) techniques play a significant role in enhancing the efficiency and accuracy of invoice processing systems. The proposed system incorporates multiple AI and ML approaches to enable intelligent interpretation of unstructured invoice data and to improve adaptability across diverse document formats. These techniques contribute to robust data extraction, classification, and analysis, thereby improving overall system performance.

A. Natural Language Processing (NLP)

Natural Language Processing techniques are employed to analyze and interpret textual data extracted from invoices. NLP enables the identification of key entities such as invoice num-bers, vendor details, dates, and financial values by leveraging contextual understanding and linguistic patterns. By analyzing sentence structure and semantic relationships, NLP facilitates the transformation of unstructured text into structured and meaningful information, thereby improving extraction accu-racy and supporting efficient downstream processing.

B. Deep Learning Approaches

Deep learning models are utilized to enhance both text recognition and information extraction, particularly in scenar-ios involving complex invoice layouts and varying document formats. These models, trained on large datasets, are capable of learning hierarchical patterns and structural features within documents. As a result, they provide improved robustness and accuracy in extracting relevant information, even from low-quality or noisy inputs.

C. Supervised Learning

Supervised learning techniques are applied to classify and extract invoice data using labeled training datasets. These models learn the relationship between input features and corresponding outputs, enabling accurate classification of in-voice components across different formats. Over time, the models improve their performance by adapting to variations in invoice structures, thereby ensuring consistent and reliable data processing.

D. Feature Extraction

Feature extraction techniques are used to identify and rep-resent important characteristics of invoice data, such as key-words, patterns, and numerical values. These features provide meaningful input to machine learning models, enhancing their ability to classify and process invoice information accurately. Effective feature representation contributes to improved sys-tem efficiency, reliability, and overall performance.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

VIII. RESULTS AND DISCUSSION

An experimental evaluation of the proposed Smart Re-tail Analytics–Automated Invoice Intelligence System was conducted to assess its effectiveness in automating invoice processing and extracting data from diverse invoice formats. The evaluation included multiple input types, such as scanned invoices, PDF documents, and photographic images, to exam-ine system performance under varying structural and quality conditions.

The integration of Optical Character Recognition (OCR) and Natural Language Processing (NLP) significantly en-hanced the system's ability to convert unstructured invoice data into structured information [2], [3]. The OCR component demonstrated reliable performance when processing high-quality and moderately degraded documents. Image prepro-cessing techniques, including noise removal, skew correction, and binarization, contributed to improved text extraction ac-curacy [2]. However, performance degradation was observed in cases involving low-resolution images, distorted documents, or invoices containing handwritten or overlapping text.

The NLP module effectively extracted and classified key invoice attributes such as invoice number, vendor details, date, GST information, and total amount. By leveraging contextual understanding and language modeling techniques, the system maintained consistent performance across different invoice layouts [3], [6]. Slight variations in accuracy were noted when handling highly irregular or non-standard document formats.

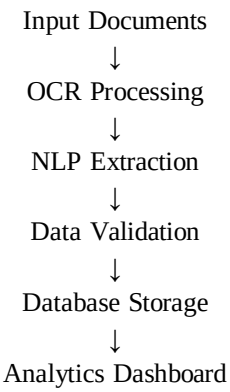
From an operational perspective, the integration of Robotic Process Automation (RPA) enabled seamless end-to-end au-tomation of invoice processing tasks. This significantly re-duced processing time compared to manual approaches while minimizing human intervention [4], [5]. As a result, the system achieved higher throughput and consistent performance when handling large volumes of invoices. The implementation of automated validation mechanisms ensured that extracted data was accurate and consistent before being stored in the database. This reduced errors commonly associated with manual data entry and improved the overall reliability of financial records. Furthermore, the system supported the development of a centralized and structured data repository, enabling efficient storage and retrieval of invoice information. The integration of data visualization tools facilitated the creation of interactive dashboards, allowing users to analyze financial metrics such as expenditure trends, vendor distribution, and temporal patterns, thereby supporting informed decision-making [6].

In addition, process monitoring techniques were applied to evaluate workflow performance and identify inefficiencies. The analysis revealed bottlenecks, delays, and redundant operations within the processing pipeline, highlighting opportunities for system optimization.

Despite its effectiveness, certain limitations were identified. The performance of the OCR component is highly dependent on input quality, particularly for low-resolution or noisy doc-uments [2]. Moreover, invoices with complex layouts, nested tables, or non-standard formatting present challenges for accu-rate information extraction. These limitations indicate potential areas for improvement through the adoption of advanced deep learning-based document understanding models [1].

Overall, the experimental results demonstrate that the inte-gration of OCR, NLP, and RPA provides an efficient, scalable, and reliable solution for automated invoice processing across diverse real-world scenarios.

. *Performance Flow Diagram*



The experimental evaluation conducted supports the conclu-sion that the combined use of OCR, natural language process-ing (NLP), and robotic process automation (RPA) technologies generates an accurate, time-efficient, and strongly automated solution for processing multiple types of invoices. As such, it is an effective, scalable, and reliable means of processing different types of invoices.

IX. CONCLUSION

This paper presented a Smart Retail Analytics–Automated Invoice Intelligence System designed to streamline and auto-mate the extraction, processing, and analysis of invoice data. The proposed system integrates Optical Character Recognition (OCR), Natural Language Processing (NLP), and Robotic Pro-cess Automation (RPA) to efficiently handle both structured and unstructured invoice formats.

The system is capable of extracting key invoice attributes, including invoice number, vendor details, date, GST informa-tion, and total amount, and converting them into structured data for centralized storage. This automation significantly reduces manual effort, minimizes human error, and improves overall processing efficiency. In addition, the integration of data visualization tools enables the generation of interactive dashboards, providing valuable insights into financial trends, vendor distribution, and expenditure patterns.

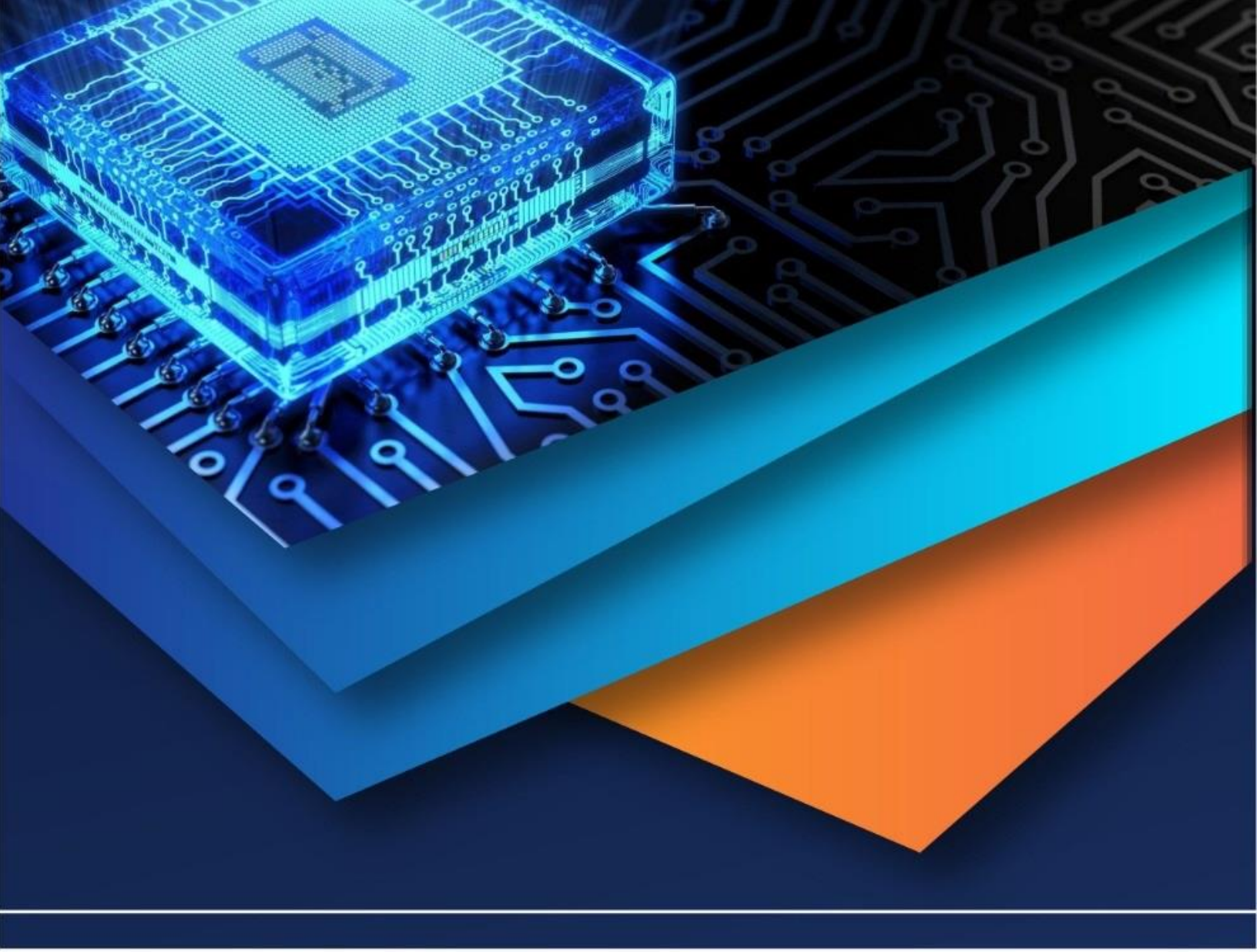
Furthermore, the incorporation of process monitoring tech-niques facilitates the identification of workflow inefficiencies and supports process optimization. The experimental results demonstrate that the proposed approach achieves reliable per-formance and can effectively process invoices across varying formats and conditions.

Despite its effectiveness, certain limitations were observed. The performance of the OCR component is influenced by input quality, particularly in cases involving low-resolution or noisy documents. Additionally, invoices with complex layouts and non-standard structures present challenges for accurate information extraction. These limitations indicate potential areas for improvement through the integration of advanced deep learning-based document understanding techniques.

In conclusion, the proposed system provides a scalable and efficient framework for intelligent invoice processing and financial analytics. Future work may focus on enhancing model robustness, improving handling of complex document structures, and incorporating more advanced AI-driven tech-niques to further optimize system performance.

REFERENCES

[1] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
[2] R. Smith, “An overview of the Tesseract OCR engine,” in Proc. Int. Conf. Document Analysis and Recognition (ICDAR), 2018, pp. 629–633.
[3] J. Devlin et al., “BERT: Pre-training of deep bidirectional transform-ers for language understanding,” in Proc. NAACL-HLT, 2019, pp. 4171–4186.
[4] M. D. Atkinson and J. E. Dietz, “Robotic process automation: A review of the technology and its applications,” IEEE Access, vol. 8,
[5] pp. 123456–123467, 2020.
[6] G. Singh and P. Sharma, “Automated invoice processing using OCR and machine learning,” in Proc. IEEE Int. Conf. Computational Intelligence, 2021, pp. 210–215.
[7] R. Shafique et al., “A survey on intelligent document processing using AI techniques,” IEEE Access, vol. 11, pp. 45678–45695, 2023.
[8] Z. Li, W. Tian, C. Li, Y. Li, and H. Shi, “A structured recognition method for invoices based on the StrucTexT model,” Applied Sciences, vol. 13, no. 12, 2023.
[9] J. Redmon and A. Farhadi, “YOLOv3: An incremental improvement,” arXiv preprint arXiv:1804.02767, 2018.
[10] S. Deshpande and M. Rajalbandi, “Intelligent document processing: AI-powered RPA for multilingual OCR of receipts,” Int. J. Sci. Res. Archive, vol. 14, no. 1, pp. 1164–1166, 2025.
[11] H. Martinez, “Automating invoice processing through intelligent OCR and entity extraction,” Int. J. Artificial Intelligence and Machine Learn-ing, vol. 4, no. 2, pp. 86–92, 2025.
[12] P. Malla, “A scalable enterprise framework for AI-driven invoice pro-cessing using document intelligence,” Int. J. AI, Big Data and Manage-ment Studies, vol. 6, no. 4, 2025.
[13] L. Gawade et al., “Invoice data extraction using OCR and large lan-guage models,” Int. J. Research in Applied Science and Engineering Technology, 2025.
[14] A. Singh and P. Singh, “Improving financial invoice workflows with RPA and OCR using multimodal techniques,” Journal of Electrical Systems, vol. 20, no. 11, 2024.
[15] M. H. Emel et al., “Efficient and accurate date extraction from invoices using OCR and object detection,” Advances in Artificial Intelligence Research, vol. 4, no. 1, pp. 10–17, 2024.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY