IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15321

AUTOMATED INVOICE PROCESSING USING UIPATH AND SQLITE

V. DINESHKUMAR[1], Dr. S. SHYLAJA[2]

Student , Department of Computer Applications, ,Sri Ramakrishna College of Arts & Science (Autonomous), Coimbatore – 641006, Tamil Nadu, India[1]

Assistant Professor, Department of Computer Applications,Sri Ramakrishna College of Arts & Science (Autonomous), Coimbatore – 641006, Tamil Nadu, India[2]

Abstract: The rapid growth of digital business operations has led to a significant increase in the volume of invoices that organizations must process daily. Traditional invoice handling methods rely heavily on manual data entry, which is time-consuming, prone to human errors, and inefficient when dealing with large numbers of documents. This project presents an Automated Invoice Processing System developed using UiPath Robotic Process Automation (RPA) and SQLite database technology. The system automatically processes invoice documents in PDF format by extracting relevant information using Optical Character Recognition (OCR). Key invoice details such as invoice number, invoice date, total amount, vendor type, and status are identified and validated through rule-based logic. The extracted data is then stored in a structured SQLite database for efficient management and retrieval. The system also includes mechanisms for duplicate detection, error handling, and automated file segregation. Valid invoices are stored in the processed folder, while invoices with missing or incorrect information are moved to an error folder for review. By automating repetitive tasks such as document reading, data extraction, validation, and storage, the proposed system significantly improves operational efficiency and accuracy. The implementation demonstrates how RPA combined with OCR technology can transform manual document processing into a reliable and scalable automated workflow, making it suitable for small and medium-scalebusiness-environments.

Keywords: Robotic Process Automation (RPA), UiPath, Invoice Processing, Optical Character Recognition (OCR), SQLite Database, Document Automation, Data Extraction, Workflow Automation

I. INTRODUCTION

In modern business environments, organizations receive and process a large number of invoices from vendors, suppliers, and service providers. Invoice processing plays an essential role in financial operations, as it involves verifying payment details, recording transactions, and maintaining accurate financial records. However, many organizations still rely on manual or semi-automated methods to process invoices, which require employees to read invoice documents, extract relevant information, and enter data into accounting systems. Manual invoice processing is time-consuming and susceptible to human errors such as incorrect data entry, missing fields, or duplicate records. As the number of invoices increases, managing these documents becomes increasingly difficult, leading to delays in payment processing and operational inefficiencies. Furthermore, scanned or image-based invoices require additional effort to interpret the text content manually. Recent advancements in Robotic Process Automation (RPA) and Optical Character Recognition (OCR) technologies provide an effective solution for automating repetitive administrative tasks such as invoice processing. RPA tools like UiPath allow organizations to create automated workflows that can interact with files, applications, and databases without human intervention. OCR technology enables automated systems to extract text from scanned documents and images, making it possible to process various invoice formats automatically. This project focuses on developing an Automated Invoice Processing System using UiPath and SQLite. The system reads invoice documents, extracts important data fields using OCR, validates the extracted information, stores the records in a database, and organizes processed files into appropriate folders. The goal of the project is to demonstrate how intelligent automation can improve accuracy, efficiency, and reliability in invoice processing operations.

II. OBJECTIVE

The primary objective of this project is to develop an automated system for processing invoices using Robotic Process Automation (RPA) technology. The system aims to eliminate manual intervention in invoice handling by automatically extracting important information from PDF invoice documents using Optical Character Recognition (OCR). It focuses

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15321

on identifying key invoice fields such as invoice number, invoice date, amount, and vendor details, and validating the extracted data to ensure accuracy and completeness. Another objective of the system is to store the validated invoice information in a structured SQLite database for efficient data management and retrieval. The project also aims to implement duplicate detection mechanisms to prevent repeated invoice entries and to organize processed documents by automatically moving valid invoices to the processed folder and invalid invoices to an error folder. Overall, the objective of the system is to improve efficiency, reduce human errors, and streamline the invoice processing workflow through intelligent automation

III. EXISTING SYSTEM

In many organizations, invoice processing is performed manually or through partially automated methods. In the traditional system, invoices are received in physical or digital formats such as printed documents or PDF files attached to emails. Employees manually open these documents, identify important fields such as invoice number, invoice date, vendor name, and total amount, and then enter the information into spreadsheets or accounting software.

This manual process requires considerable time and attention because every invoice must be reviewed individually. The process becomes more difficult when organizations receive invoices from multiple vendors using different document formats. As a result, employees must interpret each invoice layout manually, which increases processing time and the likelihood of errors.

One of the major issues in the existing system is the **high probability of human error**. Mistakes in entering invoice numbers, dates, or amounts may result in incorrect financial records. Additionally, duplicate invoices may be recorded because manual systems lack automated duplicate detection mechanisms. This can lead to financial discrepancies and unnecessary payments.

Another limitation of the existing system is the absence of **automated validation and document classification mechanisms**. If an invoice contains missing or invalid data, employees must manually identify and correct the issue. Furthermore, document storage is often unorganized, with invoices stored in folders without standardized naming conventions, making retrieval difficult during audits.

The existing system also lacks **OCR-based text extraction capabilities**, which means employees must manually read scanned invoices. This significantly increases processing time and reduces productivity. Due to these limitations, manual invoice processing becomes inefficient and unsustainable as invoice volumes grow.

IV. METHODOLOGY

The Automated Invoice Processing System is developed using UiPath Studio, which provides a visual workflow-based environment for implementing Robotic Process Automation. The system processes invoice documents stored in an input folder and performs a sequence of automated operations to extract, validate, and store invoice data.

The workflow begins by scanning the input folder containing invoice PDF files. For each invoice file detected, the system uses OCR technology (Tesseract OCR) to read the content of the document and convert it into machine-readable text. This enables the system to process both digital and scanned invoice documents.

Once the text is extracted, the system performs data parsing to identify key invoice fields such as invoice number, invoice date, amount, vendor type, and status. Regular expressions and string processing techniques are used to locate and extract these fields from the OCR output.

The extracted information then undergoes validation checks to ensure data completeness and correctness. If essential fields are missing or invalid, the invoice is classified as an error case and moved to the error folder. If the data is valid, the system performs a duplicate detection check by comparing the invoice number with existing records in the SQLite database.

If no duplicate is found, the validated invoice data is inserted into the SQLite database, where it is stored in structured tables. After successful insertion, the invoice document is moved to the processed folder to indicate that it has been successfully handled. Exception handling mechanisms ensure that system errors are properly managed without interrupting the workflow.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15321

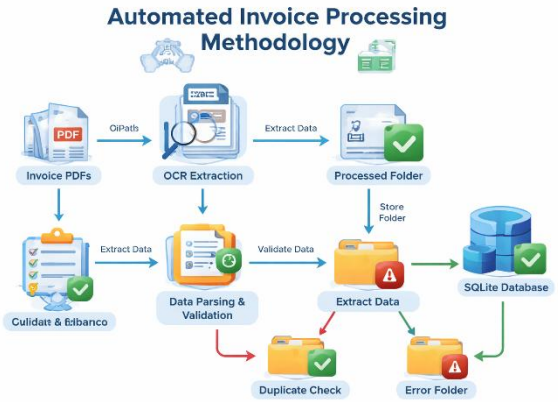

Figure representing automated invoice processing methodology.

V. RESULT AND DISCUSSION

The automated invoice processing system was tested with various invoice documents, including both digital and scanned PDF files. The system successfully extracted key invoice details such as invoice number, invoice date, vendor name, and total amount using OCR technology. The data parsing logic correctly identified and organized the extracted fields before storing them in the database.

The duplicate detection mechanism prevented the system from processing the same invoice multiple times. When a duplicate invoice was detected, the system skipped database insertion and handled the file appropriately. This feature ensured the integrity of financial records and prevented redundant data storage.

The validation module effectively identified invoices with missing or incomplete information. Such invoices were automatically moved to the error folder for manual review, ensuring that only valid and complete data was stored in the database. This significantly improved data accuracy and reduced the risk of financial inconsistencies.

Compared to manual invoice processing, the automated system demonstrated a significant improvement in efficiency and reliability. Processing time per invoice was greatly reduced, and the possibility of human error was minimized. Additionally, the structured database storage and automated file management improved document organization and retrieval.

Overall, the results confirm that the proposed system provides an efficient solution for automating invoice processing tasks in organizations that handle large volumes of documents.

VI. CONCLUSION

The Automated Invoice Processing System developed using UiPath and SQLite successfully demonstrates the effectiveness of Robotic Process Automation in handling invoice documents. The system automates data extraction, validation, and storage, thereby reducing manual effort and processing time. It improves accuracy by minimizing human errors and preventing duplicate entries. Overall, the system provides a reliable and efficient solution for automating invoice management in business environments.

REFERENCES

[1]. **UiPath Documentation**, *UiPath Studio Guide and RPA Development Concepts*, UiPath Inc., Available at: https://docs.uipath.com

[2]. **R. S. Pressman**, *Software Engineering: A Practitioner's Approach*, 8th Edition, McGraw-Hill Education, 2015.

[3]. **Tesseract OCR Documentation**, *Optical Character Recognition Engine*, Available at: https://tesseract-ocr.github.io