International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 12 | Dec 2025 www.irjet.net p-ISSN: 2395-0072

OCR Result Analysis System

Harsh Singh Bhatiya [1], Mayur Pawar[2], Rohan Gaumgaonkar, Sharvan Dharme[4], Rakesh Moharle[5], Dr. Sushama Telrandhe[6]

[1,2,3,4,]U.G. Students, Department of Computer Science and Engineering, Gurunanak Institute of Engineering and Technology, Nagpur, Maharashtra, India.
[5] Assistant professor, Department of Computer Science and Engineering, Gurunanak Institute of Engineering and Technology, Nagpur, Maharashtra, India.
[6] Associate professor, Department of Computer Science and Engineering, Gurunanak Institute of Engineering and Technology, Nagpur, Maharashtra, India.

ABSTRACT- The OCR Result Analysis System automates student data entry, result generation, and report management using Optical Character Recognition (OCR). Student details such as Roll Number, Enrollment Number, Candidate Name, and academic records are extracted from scanned images or entered manually, then stored securely in a structured database. The system automatically generates Excel reports, including detailed sheets, summary reports, and category-wise analyses. A history module maintains all uploads and generated files for quick retrieval. With secure authentication and error-free processing, the system reduces manual workload, improves accuracy, and provides a reliable solution for educational result management.

1. INTRODUCTION

The OCR Result Analysis System is designed to automate the management of student records and academic reports by utilizing Optical Character Recognition (OCR) technology. It allows users to securely log in and register before accessing features that let them manually input student details or upload scanned mark sheets for OCR processing. The system then stores the extracted data in a structured database, enabling efficient and accurate result generation. Result sheets are automatically created in Excel format, simplifying analysis and report sharing for teachers and administrators In addition to result generation, the system includes a History Module to record actions such as OCR uploads, result sheet generation, and file downloads. This module ensures transparency and accountability by allowing easy retrieval of previously generated files. By automating manual tasks, the system significantly reduces workload, improves operational efficiency, and provides a reliable, secure solution for schools, colleges, and universities to manage student records and generate academic reports.

2. LITERATURE REVIEW

The study by Mehdi Rizvi, Hasnain Raza, Shahab Maqsood, and Shan Afify (2019) presents an Optical Character Recognition (OCR)-based intelligent database management system designed for examination process control. Their work applies computer vision techniques, particularly OCR and object recognition, to automate the processing of handwritten and digital text. By integrating deep learning, the system recognizes exam-related data such as signatures and written information, stores it efficiently, generates automated graphs, and reduces manual effort throughout the exam management workflow.

In another study, Eman Shaikh, Iman Mohiuddin, Ayesha Manzoor, Ghazanfar Latif, and Nazeer uddin Mohammad (2019) developed an automated grading system for handwritten answer sheets using convolutional neural networks (CNN). Trained on a dataset of 250 answer sheets, the model achieved 92% accuracy and significantly reduced manual grading time by computing scores automatically.

The work of Aarav T. Shaji, Justin Thomas T., Emils Saj, Vishnuprasad T. V., Deepa J., and Jibu Mathew (2024) focuses on document digitalization using Optical Character Recognition for student evaluation sheets. Their system converts handwritten exam digits and information into digital CSV format using CNN, helping streamline data management and reduce manual labor in the evaluation process.

Finally, Shams A. Al-Qaisy, Amal Hasan Qanaa, Rougar Khorshid, Mayar Mohamad, Salim Yahya Massad Kallul, and Ahmad Hussein Hassan (2025) emphasize enhancing security and data management in higher education through the integration of Cloud Content Management (CCM). Their study highlights challenges such as declining student engagement and resource accessibility, proposing that CCM solutions can improve collaboration, data sharing, storage, and accessibility, ultimately supporting better institutional operations.

3. METHODS AND MATERIALS

❖ MATERIALS

- **Hardware:**
- Intel i3 Processor
- 160GB Hard Disk
- 8GB RAM
- SVGA Monitor

- Standard Windows Keyboard
- Two or Three Button Mouse

- **Software:**

- 1 Operating System**:** Windows 7/8/10
- 2 Server-side Script: HTML, CSS, Bootstrap & JavaScript
- 3 Programming Language: Python
- 4 Libraries/Framework: Django
- 5 IDE / Workbench: VS Code
- 6 Technology: Python 3.6+

❖ Methods

- **Requirement Analysis:**

Requirement Analysis is the foundational phase of any software development project. It focuses on understanding what the system should do and why it should do it. This phase ensures that stakeholders, developers, and users share a common understanding.

- **System Design*:***

System design outlines the architectural blueprint of the OCR Result Analysis System. This phase involves creating models and diagrams that represent how the modules interact with each other. The system is divided into multiple layers, such as input, preprocessing, OCR extraction, post-processing, database management, and dashboard visualization. Each layer works independently but communicates seamlessly with other layers, ensuring modularity and maintainability.

- **Development:**

The application backend and frontend were coded using selected programming languages. GIS tools were integrated to enable mapping functionalities. Security features including role-based access control, digital signatures, and encryption were implemented. Blockchain integration was optionally incorporated to record transactions immutably.

- **Data Migration:**

Existing paper records were digitized through scanning and Optical Character Recognition (OCR). Legacy data underwent validation and cleansing before importation into the system to ensure data integrity.

- **Testing:**

Testing evaluates the functionality, usability, and performance of the OCR Result Analysis System. The testing process includes unit testing for individual components, integration testing for combined modules, system testing for the entire workflow, and performance testing to measure processing speed and accuracy under load. This ensures that the system performs consistently across different documents and image qualities. Accuracy testing plays a vital role in OCR systems. Engineers calculate metrics like Word Error Rate (WER), Character Error Rate (CER), and overall accuracy percentage to measure OCR performance. Usability testing ensures the dashboard is intuitive for both technical and non-technical users. Finally, security testing verifies safe file handling, controlled access, and protection of sensitive data.

- **S-Curve Analysis of Land Record Management System Implementation*:***

Testing evaluates the functionality, usability, and performance of the OCR Result Analysis System. The testing process includes unit testing for individual components, integration testing for combined modules, system testing for the entire workflow, and performance testing to measure processing speed and accuracy under load. This ensures that the system performs consistently across different documents and image qualities. Accuracy testing plays a vital role in OCR systems. Engineers calculate metrics like Word Error Rate (WER), Character Error Rate (CER), and overall accuracy percentage to measure OCR performance. Usability testing ensures the dashboard is intuitive for both technical and non-technical users. Finally, security testing verifies safe file handling, controlled access, and protection of sensitive data.

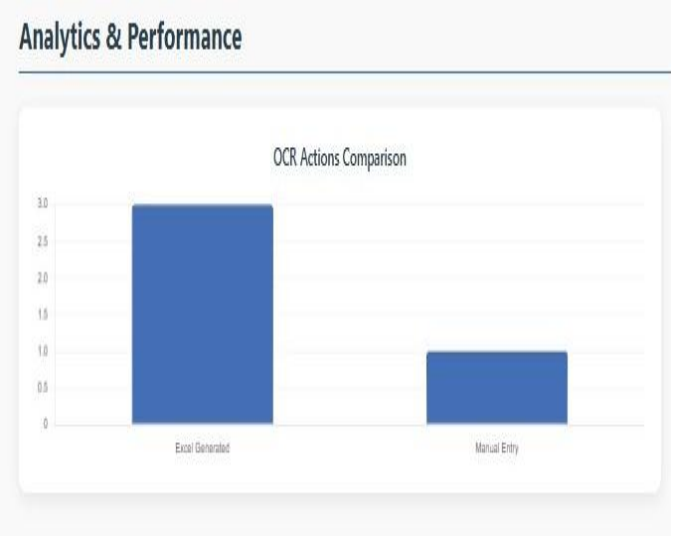

Chart -1: S-Curve Analysis of OCR Result Analysis System.

During the growth phase, accuracy increases rapidly due to model improvements, additional datasets, and optimized preprocessing. Eventually, the system enters the maturity phase, where accuracy stabilizes, and improvements become gradual. This indicates that the system has reached a performance saturation point. S-curve analysis helps researchers visualize these trends and understand system development patterns.

- **Comparative Analysis of LRMS Implementation Impact**

Comparative analysis evaluates how different OCR algorithms, models, or system enhancements affect performance. This involves comparing metrics such as

accuracy, speed, error rate, resource consumption, and language support. For example, traditional OCR models like Tesseract provide acceptable accuracy for clean text but struggle with complex or noisy images. Meanwhile, deep-learning-based models achieve significantly higher accuracy at the cost of increased computational power.

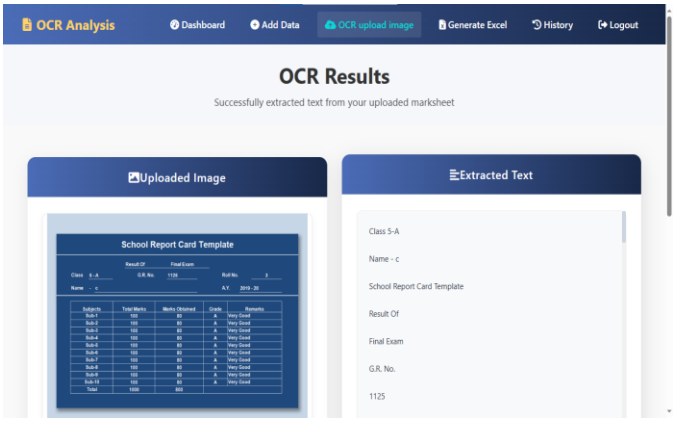

Table -1: Comparative Analysis of OCR Impact

4. DASHBOARD INTERFACE

- **Dashboard Overview and Key Metrics:**

The dashboard interface serves as the visual control center for users. It presents document upload options, extracted text, graphical accuracy reports, and performance trends. A well-designed dashboard provides users with an intuitive and user-friendly experience. Visual elements like bar charts, S-curves, accuracy graphs, and tabular data presentation help users quickly interpret results.

The dashboard also includes features such as export options, user settings, search filters, and multi-document comparison tools. Accessibility elements like adjustable font size, high-contrast themes, and keyboard navigation support ensure usability for a diverse audience. A streamlined dashboard improves overall efficiency and decision-making by presenting data in an organized format.

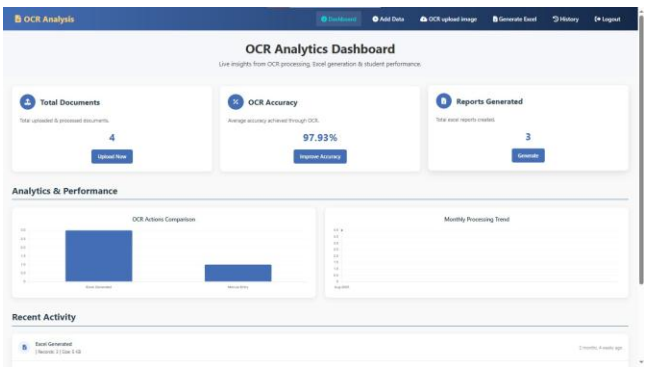

Fig-1 Dashboard Interface

- **System Features and User Accessibility:**

The OCR Result Analysis System offers multiple features, including text extraction, preprocessing, accuracy evaluation, and dashboard visualization. Additional functionalities such as batch document processing, language selection, and downloadable output further enhance the system's practicality. Multi-format support ensures compatibility with a wide range of document types, from PDFs to handwritten notes. User accessibility plays a critical role in system acceptance. Features like text-to-speech, screen reader compatibility, keyboard shortcuts, and visually accessible themes ensure inclusivity. The system follows key accessibility guidelines, such as WCAG, ensuring that differently-abled users can operate the system comfortably. Enhanced accessibility increases usability and broadens the system's user base.

5. CONCLUSION AND FUTURE SCOPE

In conclusion, the OCR Result Analysis System provides an efficient solution for automated text extraction and performance evaluation. Through its modular architecture, comprehensive dashboard, and analytical visualizations, the system successfully addresses the challenges associated with manual text processing. The integration of preprocessing techniques, OCR models, and post-processing filters ensures accuracy and reliability. Looking toward the future, the system can be expanded using advanced deep-learning architectures, real-time mobile OCR, and cloud-based distributed processing. Integration of natural language processing features, such as summarization and entity recognition, will make the system more intelligent. Additionally, automated error correction and self-learning capabilities can further improve accuracy and reduce human intervention.

6. REFERENCES

1. Mehdi Rizvi, Hasnain Raza, Shahab Tahzeeb, Shan Jaffry || Optical Character Recognition Based Intelligent Database Management System for Examination Process Control || 2019.
2. Eman Shaikh, Iman Mohiuddin, Ayisha Manzoor, Ghazanfar Latif, Nazeeruddin Mohammad || Automated Grading for Handwritten Answer Sheets using Convolutional Neural Networks || 2019.
3. Ajay T Shaju, Justin Thomas Jo, Emil Saj Abraham, K G Vishnuprasad, V Deepa, Jabin Mathew || Document Digitization Using Optical Character Recognition - A Case Study of Mark Entry Automation in Educational Institutions || 2024.
4. Shams A. Al-Qaisy, Amal Hasan Jumaa, Rozhgar Khorsheed Mahmud, Salim Yahya Kasim Khalil, Ahmad Hussein Hassan || Enhancing Secure Data Management

for Student Learning Analysis using Cloud Computing Management (CCM) in Higher Education || 2025.A secure land record management system using blockchain. (2023). IEEE Xplore Digital Library.

5. Andriy Pukas, Andrii Simak, Andriy Yushko, Sergiy Nadvynychnyy, Serhii Shandruk, Oleh Vatslavskyi || Automated Reporting Module in the Academic Staff Performance Appraisal System || 2024.User Surveys and Impact Studies. (2024). Assessment of Digital Land Record Systems in Indian States. Various State Revenue Departments.
6. Sonam Shivani Nair, Neeraj Anand Sharma || Implementation of an Automated Result Management System in a Developing Country || 2019.
7. Mahesh Jangid, Sumit Srivastava || A OCR Based Approach || 2013.
8. Cristina Alonso-Fernández, José L. Jorro-Aragoneses, Carlos M. Alaíz, Pilar Rodríguez || Learning Analytics Tools to Analyze Progress and Results With Moodle LMS Data || 2024.