International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 03 | Mar 2026 www.irjet.net p-ISSN: 2395-0072

OCR CALCULATOR FOR EFFICIENT EVALUATION OF STUDENT LAB PERFORMANCE

Dr. Ch. V. Phani Krishna[1], D. Sreeja Reddy[2], Ch.Alekhya[3], B.Dinesh Reddy[4], D.Lakshmikanth[5]

[1] Professor & HOD, Department of Computer Science and Engineering
[2,3,4,5] B.Tech Students, Department of Computer Science and Engineering
Teegala Krishna Reddy Engineering College , Telangana, India

Abstract - *In educational institutions, evaluating student performance in laboratory courses requires faculty members to manually record and calculate marks across multiple sessions. This manual process is time-consuming, error-prone, and inefficient. To address these challenges, this research proposes an OCR-based web application that automates the extraction and calculation of student laboratory marks. The system uses Optical Character Recognition (OCR) and image processing techniques to capture and analyze images of mark sheets. Techniques such as image preprocessing, cell segmentation, and digit recognition are employed to accurately extract numerical values from structured tables. The extracted marks are automatically processed to compute total scores and average performance of students. The proposed system integrates computer vision and machine learning techniques, including Convolutional Neural Networks (CNNs), to improve recognition accuracy. The web application provides a user-friendly interface that allows faculty members to capture images of lab evaluation sheets and instantly obtain calculated results. This automated solution significantly reduces manual workload, minimizes human errors, and improves the efficiency of academic evaluation processes. The system also ensures standardized and reliable assessment of laboratory performance. The proposed OCR calculator can be effectively adopted by educational institutions to enhance digital transformation in academic evaluation systems.*

***Key Words*: Optical Character Recognition, OCR, Digit Recognition, Image Processing, CNN, Web Application.**

1.INTRODUCTION

Optical Character Recognition (OCR) is a powerful technology that converts printed or handwritten text from images or scanned documents into machine-readable digital text. It has become an essential tool for digitizing physical documents and automating data extraction processes. OCR systems use a combination of image processing, pattern recognition, and machine learning techniques to analyze images and recognize characters with high accuracy.

The OCR process typically includes several stages such as image acquisition, preprocessing, segmentation, feature extraction, and character recognition. During preprocessing, techniques such as noise removal, thresholding, and image enhancement are applied to improve image quality. After preprocessing, segmentation techniques isolate individual characters or digits from the image. Machine learning models, particularly deep learning algorithms like Convolutional Neural Networks (CNNs), are then used to recognize the extracted characters.

OCR technology has been widely adopted in many sectors including banking, healthcare, document management, and education. In educational institutions, OCR can be used to digitize academic records, automate grading processes, and simplify administrative tasks. Despite these advancements, many laboratory evaluation systems still rely on manual data entry and calculation of marks, which can be time-consuming and prone to human error.

Faculty members often need to manually record marks for each student across multiple laboratory sessions and later calculate the total and average scores. This process becomes increasingly difficult when dealing with large numbers of students. Therefore, there is a strong need for an automated system that can simplify and speed up this evaluation process.

This project proposes an OCR-based calculator web application designed to automate the evaluation of student laboratory performance. The system captures images of mark sheets, extracts numerical values using OCR techniques, and automatically calculates the total and average marks. By integrating image processing, machine learning, and web application technology, the proposed system improves efficiency, reduces manual workload, and ensures accurate student performance evaluation.

1.1 Overview

Optical Character Recognition (OCR) plays a significant role in converting physical documents into digital formats. It enables the automatic extraction of textual information from images and scanned documents, allowing the data to be stored, edited, and processed electronically. Modern OCR systems rely on advanced algorithms, deep learning models, and pattern recognition techniques to achieve high accuracy in character recognition. In the context of educational institutions, OCR can simplify many administrative tasks, including grading, document digitization, and record management. By using OCR technology, information from printed or handwritten documents can be quickly converted

into digital data that can be easily processed by computer systems. The proposed OCR calculator system uses image processing techniques to capture images of laboratory mark sheets and extract the marks assigned to students. The system then performs automatic calculations to determine the total marks and average scores. This automation significantly reduces manual effort and improves the speed and accuracy of the evaluation process.

1.2 Motivation

In educational environments, evaluating student laboratory performance often requires manual entry and calculation of marks. Faculty members must record marks for multiple students across different lab sessions and calculate the final consolidated results. This process is repetitive, time-consuming, and prone to errors. Additionally, manual evaluation processes increase the workload for faculty members and reduce the time available for teaching and research activities. Automating this process can significantly improve efficiency and accuracy.

OCR technology offers a practical solution for this challenge by enabling automatic extraction of numerical data from mark sheets. By integrating OCR with web application technology, faculty members can easily capture images of mark sheets and automatically calculate student performance metrics. This motivates the development of an OCR-based web application that simplifies the evaluation process.

1.3 Problem Definition

In laboratory-based courses, faculty members must evaluate student performance based on marks obtained in multiple lab sessions. Traditionally, this process involves manually recording marks from lab evaluation sheets and calculating consolidated scores for each student. This manual approach is inefficient and susceptible to human errors, especially when dealing with a large number of students. Furthermore, the time spent on manual calculations reduces the time available for academic activities such as teaching, mentoring, and research. Errors in manual data entry can also affect the accuracy and fairness of student assessments.

To address these issues, there is a need for an automated system that can capture marks directly from evaluation sheets and perform calculations automatically. The proposed OCR calculator system aims to solve this problem by using image processing and OCR techniques to extract numerical values and compute performance metrics instantly.

1.4 Objectives

The main objectives of the proposed system are:

1. To develop a web application that automates the evaluation and consolidation of student laboratory marks using OCR technology.
2. To improve efficiency by reducing the manual workload involved in recording and calculating marks.
3. To enhance accuracy by minimizing human errors in data entry and calculation.
4. To implement robust image processing techniques for preprocessing and segmentation of mark sheet images.
5. To provide instant calculations of total marks and average scores based on extracted data.
6. To design a user-friendly interface that allows faculty members to easily capture and process images of mark sheets.

2. PROPOSED SYSTEM

The proposed system introduces an automated OCR-based web application designed to simplify and improve the evaluation process of student laboratory performance. Traditional evaluation methods require faculty members to manually record marks from laboratory sheets and calculate totals and averages. This process is time-consuming and susceptible to errors. The proposed system addresses these limitations by using Optical Character Recognition (OCR) and image processing techniques to automatically extract and process numerical data from mark sheets.

The system allows faculty members to capture images of laboratory evaluation sheets using a web application. The captured images are processed through several stages including image preprocessing, cell segmentation, digit recognition, and data consolidation. These processes enable the system to accurately identify marks from structured tables and automatically compute performance metrics such as total marks and average scores.

By integrating computer vision algorithms and machine learning techniques such as Convolutional Neural Networks (CNNs), the proposed system ensures reliable recognition of numeric characters even under varying lighting conditions and image quality. The automated system significantly reduces manual effort, improves accuracy in evaluation, and provides instant feedback to educators.

2.1 Web Application Development

The proposed system is implemented as a web application designed to provide a user-friendly interface for faculty members. The web app acts as the central platform where users can capture images of laboratory mark sheets and process them automatically.

The application allows users to upload or capture images using the device camera. Once the image is captured, the system processes the image and extracts relevant data. The interface is designed to be simple and intuitive so that

faculty members can easily operate the system without requiring technical expertise.

The web application also displays the calculated results, including the extracted marks, total scores, and average performance of students. This helps educators quickly analyze student performance and maintain digital records.

2.2 Image Processing

Image processing plays a crucial role in improving the quality of captured images before performing OCR. The captured images may contain noise, shadows, or distortions that can affect recognition accuracy. Therefore, preprocessing techniques are applied to enhance image quality.

These preprocessing techniques include noise reduction, grayscale conversion, thresholding, and image normalization. Edge detection algorithms are also applied to identify the structure of tables and grids present in the mark sheet.

By enhancing image clarity and structure, the preprocessing stage ensures that the OCR system receives clean and well-structured input data, thereby improving the accuracy of digit recognition.

2.3 Cell Segmentation

Cell segmentation is the process of identifying and isolating individual cells within the table structure of the mark sheet. Since marks are usually written in tabular format, detecting the boundaries of each cell is essential for extracting the correct values. The system uses techniques such as edge detection and Hough Transform to detect lines and grid structures within the image. Once the grid structure is identified, the image is divided into individual cells corresponding to different marks. Each segmented cell is then processed individually to extract the digits present within it. This segmentation step ensures that the OCR system focuses only on relevant regions containing marks, thereby improving recognition accuracy.

2.4 Optical Character Recognition

Optical Character Recognition (OCR) is the core component of the proposed system. The OCR module is responsible for extracting textual or numerical information from the segmented cells.

The OCR engine analyzes the processed image and identifies characters using pattern recognition techniques. Machine learning models, particularly Convolutional Neural Networks (CNNs), are used to recognize digits accurately. The OCR module converts the detected characters into machine-readable text.

This recognized text represents the marks assigned to students in laboratory sessions. The OCR module ensures accurate extraction of numerical data even when digits are handwritten or slightly distorted.

2.5 Data Consolidation and Calculation

After the OCR process extracts the numerical values from the mark sheet, the data consolidation module organizes the recognized numbers into a structured format. The extracted marks are mapped to the corresponding students and lab sessions. The system then performs automatic calculations such as computing total marks and average scores for each student. These results are displayed within the web application for quick analysis. This automated calculation process eliminates manual effort and reduces the possibility of human errors in mark consolidation. It also allows educators to instantly evaluate student performance and maintain digital records efficiently.

3. IMPLEMENTATION DETAILS

The implementation of the proposed OCR-based calculator system involves integrating image processing techniques, machine learning algorithms, and web application development. The system is designed to automatically capture images of laboratory evaluation sheets, extract numerical marks using OCR, and calculate performance metrics such as total and average marks. The implementation focuses on ensuring accurate digit recognition, efficient data processing, and a user-friendly interface for faculty members.

The system follows a structured workflow where each stage performs a specific task such as image capture, preprocessing, segmentation, recognition, and calculation. These components work together to automate the evaluation process and provide instant results.

3.1 System Architecture

The system architecture illustrates the overall structure and workflow of the proposed OCR-based evaluation system. It shows how different components interact with each other to process input images and generate calculated results. The architecture consists of several modules including image capture, image preprocessing, cell segmentation, OCR processing, digit recognition, and data consolidation. Initially, the user captures an image of the laboratory marks sheet using the web application. The captured image is then sent to the preprocessing module, where noise reduction and image enhancement techniques are applied.

After preprocessing, the system performs cell segmentation to detect table structures and isolate individual cells containing marks. Each segmented cell is then processed by the OCR module to recognize digits using machine learning algorithms such as Convolutional Neural Networks (CNNs).

The recognized numerical values are then sent to the data consolidation module, where calculations such as total marks and average scores are performed. Finally, the calculated results are displayed to the user through the web application interface.

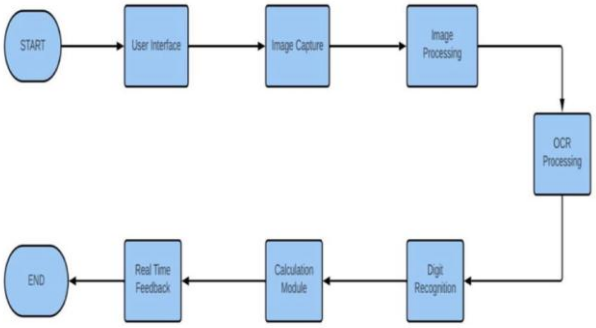


Figure 3.1 Block Diagram of Proposed Workflow

3.2 Image Acquisition

The first stage of the system involves capturing images of laboratory evaluation sheets using the web device camera. The application allows users to either capture images in real time or upload previously captured images from the device storage.

The image acquisition module ensures that the captured images have sufficient clarity and brightness for further processing. Proper alignment and focus are important to ensure accurate recognition of digits during later stages.

3.3 Image Preprocessing

Image preprocessing is performed to improve the quality of the captured images and remove unnecessary noise. Since raw images may contain distortions, shadows, or uneven lighting, preprocessing techniques are applied to enhance the image for accurate OCR processing. The preprocessing stage includes operations such as grayscale conversion, noise reduction, thresholding, and image normalization. These operations help highlight textual and numerical regions within the image while removing unwanted background elements. By improving image quality, the preprocessing stage increases the accuracy of segmentation and character recognition.

3.4 Table Detection and Cell Segmentation

In laboratory evaluation sheets, marks are usually recorded in tabular form. Therefore, detecting the table structure and isolating individual cells is an important step in the implementation. The system uses techniques such as edge detection and the Hough Transform to identify horizontal and vertical lines in the image. These detected lines are used to locate the boundaries of the table and divide it into individual cells. Each segmented cell corresponds to a specific mark entry. These cells are extracted and forwarded to the OCR module for digit recognition.

3.5 OCR and Digit Recognition

The OCR module is responsible for recognizing characters and digits from the segmented cells. The system analyzes each cell image and identifies numerical values using pattern recognition and machine learning techniques. A Convolutional Neural Network (CNN) model is used to classify the digits extracted from the images. The CNN model is trained using datasets such as MNIST, which contain large collections of handwritten digits.The OCR module converts the detected digits into machine-readable text, which represents the marks assigned to students in different laboratory sessions.

3.6 Data Processing and Result Generation

After digit recognition, the extracted numerical values are processed by the data consolidation module. This module organizes the recognized marks into a structured format corresponding to individual students and lab sessions. The system automatically performs calculations such as total marks and average scores for each student. These calculated results are then displayed on the web application interface, allowing faculty members to quickly analyze student performance. The automated calculation process significantly reduces manual workload and ensures accurate evaluation of laboratory performance.

4. RESULTS AND PERFORMANCE ANALYSIS

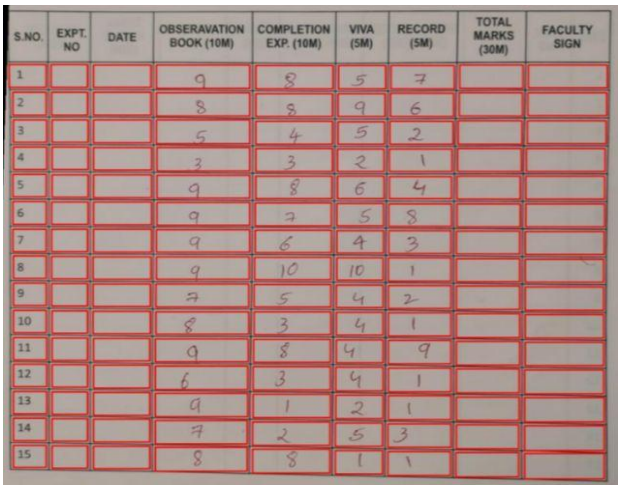
The proposed OCR-based calculator system was implemented and tested to evaluate its effectiveness in extracting and processing laboratory marks from evaluation sheets. The system was evaluated based on its ability to accurately recognize digits, process tabular data, and compute total and average marks automatically. The results demonstrate that the proposed system significantly reduces manual effort and improves the accuracy of student performance evaluation.

The testing process involved capturing images of laboratory mark sheets using a web application and processing them through the developed application. The system successfully performed image preprocessing, table detection, cell segmentation, digit recognition, and result calculation. The extracted marks were then used to compute the total and average scores for students.

4.1 OCR Recognition Results

The OCR module was tested using different mark sheet images containing numerical values written in tabular format. The system successfully detected table structures and segmented individual cells containing marks. Each cell was processed through the OCR engine to recognize numerical digits.

The experimental results showed that the OCR system was able to accurately recognize digits from well-structured tables. Image preprocessing techniques such as grayscale conversion, thresholding, and noise removal significantly improved recognition accuracy. The use of Convolutional Neural Networks (CNNs) further enhanced the performance of digit recognition.



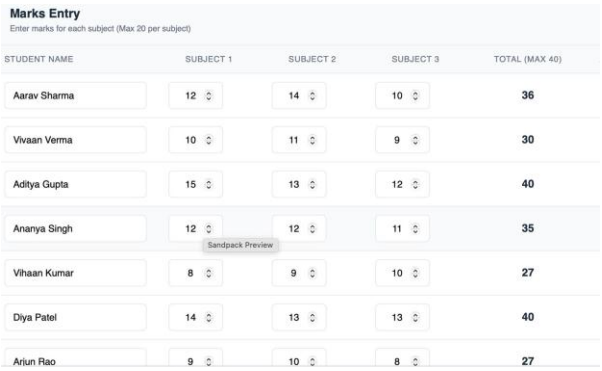
S.NO.	EXPT. NO	DATE	OBSERAVATION BOOK (10M)	COMPLETION EXP. (10M)	VIVA (5M)	RECORD (5M)	TOTAL MARKS (30M)	FACULTY SIGN
1			9	8	5	7		
2			8	8	9	6		
3			5	4	5	2		
4			3	3	2	1		
5			9	8	6	4		
6			9	7	5	8		
7			9	6	4	3		
8			9	10	10	1		
9			7	5	4	2		
10			8	3	4	1		
11			9	8	4	9		
12			6	3	4	1		
13			9	1	2	1		
14			7	2	5	3		
15			8	8	1	1		

Fig - 2: OCR Digit Recognition Input

The above figure illustrates the input of the OCR system where digits are extracted from the segmented cells of the mark sheet.

4.2 Automated Calculation of Marks

After the OCR module extracted numerical values, the system automatically calculated the total marks and average scores for each student. The automated calculation module processed the recognized digits and generated consolidated results. sThis feature eliminates the need for manual calculation and ensures accurate evaluation of student laboratory performance. The system displays the results instantly on the web application interface, enabling faculty members to quickly analyze the performance of students.



Marks Entry

Enter marks for each subject (Max 20 per subject)

STUDENT NAME	SUBJECT 1	SUBJECT 2	SUBJECT 3	TOTAL (MAX 40)
Aarav Sharma	12	14	10	36
Vivaan Verma	10	11	9	30
Aditya Gupta	15	13	12	40
Ananya Singh	12	12	11	35
Vihaan Kumar	8	9	10	27
Diya Patel	14	13	13	40
Arjun Rao	9	10	8	27

Fig- 3: Automated Marks Calculation Interface

The figure above shows the final output generated by the system, where extracted marks are processed to calculate total and average scores.

4.3 Performance Evaluation

The performance of the proposed system was evaluated based on recognition accuracy, processing speed, and usability. The results indicate that the system achieves high accuracy in recognizing digits from structured mark sheets. The preprocessing and segmentation techniques helped improve OCR accuracy by providing clean input images. Additionally, the web application interface allowed faculty members to easily capture images and obtain results without requiring complex technical knowledge. The overall system performance demonstrates that the proposed solution is efficient, reliable, and suitable for practical use in educational institutions. The implementation of the OCR-based evaluation system significantly reduces manual workload, minimizes calculation errors, and improves the overall efficiency of laboratory performance assessment.

5. CONCLUSION

In this research, an OCR-based calculator system was developed to automate the evaluation of student laboratory performance. The proposed system utilizes Optical Character Recognition (OCR), image processing techniques, and machine learning algorithms to extract numerical values from laboratory mark sheets and perform automatic calculations. By capturing images of evaluation sheets through a web application, the system processes the images using preprocessing, cell segmentation, and digit recognition techniques to accurately identify the marks assigned to students. The implementation of this system significantly reduces the manual effort required for recording and calculating student marks. It also minimizes human errors that may occur during manual data entry and computation. The use of deep learning techniques, particularly Convolutional Neural Networks (CNNs), improves the accuracy of digit recognition and ensures reliable extraction of numerical data from structured documents. The experimental results demonstrate that the proposed system can effectively detect table structures, extract marks from individual cells, and compute total and average scores automatically. The web application provides a user-friendly interface that allows faculty members to easily capture images and obtain instant results.

Overall, the OCR calculator system enhances the efficiency, accuracy, and reliability of student laboratory performance evaluation. The proposed solution can be effectively adopted by educational institutions to simplify the assessment process and support digital transformation in academic evaluation systems.

6. FUTURE WORK

Although the proposed OCR-based calculator system successfully automates the evaluation of student laboratory performance, several improvements can be implemented in future work to enhance its functionality and performance. One possible enhancement is improving the accuracy of character recognition for handwritten digits. While the current system performs well with clearly written digits, more advanced deep learning models and larger training datasets can be used to improve recognition accuracy for different handwriting styles. Another future improvement is the integration of cloud-based storage systems. By storing extracted data and calculated results in a cloud database, institutions can maintain centralized academic records and access them from multiple devices. This would also allow faculty members to track student performance over time. The system can also be extended to support recognition of complete text fields such as student names, roll numbers, and subject codes in addition to numerical marks. This would allow the system to automatically generate complete digital records of laboratory evaluations. Furthermore, the application can be integrated with existing Learning Management Systems (LMS) or academic management software used by educational institutions. This integration would enable automatic updating of student marks in institutional databases. In the future, the system can also incorporate advanced features such as real-time analytics, performance dashboards, and automated report generation. These features would help educators analyze student performance more effectively and improve decision-making in academic evaluation. Overall, future enhancements will focus on improving recognition accuracy, expanding system capabilities, and integrating the solution with modern digital education platforms.

REFERENCES

[1] Abhash Sinha, Martin Jenckel, Syed Saqib Bukhari and Andreas Dengel, "Unsupervised OCR Model Evaluation Using GAN," 2019 International Conference on Document Analysis and Recognition (ICDAR), pp. 1256–1259, IEEE, 2019.

[2] Venkata Naga Sai Rakesh Kamisetty, P. Jeyanthi, Bodapati Sohan Chidvilas, V. Maria Anu, S. Revathy, and L. Mary Gladence, "Digitization of Data from Invoice using OCR," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), pp. 1842–1847, IEEE, 2022.

[3] Áron Zoltán Kaló and Miklós László Sipos, "Key-Value Pair Searching System via Tesseract OCR and Post Processing," 2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI), pp. 978–1–7281–8053–3, IEEE, 2021.

[4] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner, "Gradient-Based Learning Applied to Document Recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[5] Ray Smith, "An Overview of the Tesseract OCR Engine," Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR), vol. 2, pp. 629–633, IEEE, 2007.

[6] A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, "Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks," Proceedings of the 23rd International Conference on Machine Learning, pp. 369–376, 2006.

[7] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.

[8] D. Doermann, "The Indexing and Retrieval of Document Images: A Survey," Computer Vision and Image Understanding, vol. 70, no. 3, pp. 287–298, 1998.