iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VI **Month of publication:** June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72535

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 13 Issue VI June 2025- Available at www.ijraset.com

Automated OMR Analyser using ML and Image Processing

Ravindra S[1], Ranganath K N[2], Nithin A[3], Manoj M H[4], M S Hemanth Kumar[5]
[1]*Associate Prof. ,* [2, 3, 4, 5]*Final Year Students Dept. of Computer, Science and Engineering, Jawaharlal Nehru New College of Engineering, Shivamogga, Visvesvaraya Technologial University – Belagavi*

Abstract: ***This work explores Optical Mark Recognition (OMR) systems built on Machine Learning and Image Processing techniques to achieve accurate and automated evaluation of answer sheets. The primary objective is to eliminate dependency on specialized scanners by utilizing devices such as mobile cameras and webcams. This system processes scanned answer sheets by isolating responses and matching them with reference answers using automated logic. Several models integrate contour detection, circle detection, and error-tolerant classification to ensure robustness even with imperfect markings. The survey shows that OMR systems deliver high accuracy in diverse lighting and marking conditions. The findings support the development of scalable and cost-effective OMR solutions that are particularly useful in academic and institutional settings.***

Keywords: Optical Mark Recognition, Image Processing, Machine Learning, Contour Detection, Education Assessment

I. INTRODUCTION

In the field of education and assessments, Optical Mark Recognition (OMR) technique is adapted for evaluating multiple-choice answer sheets. Traditional OMR systems depend heavily on specialized scanning hardware and strict formatting standards, which limit flexibility and increased costs. With advancements in Image Processing (IP) and Machine Learning (ML), more adaptive and cost-efficient OMR solutions have become feasible. Automated This automated method enhances grading efficiency, ensures reliable scoring, and reduces evaluation time. It allows answer sheets to be scanned using ordinary cameras or smartphones, making the system more accessible to institutions with limited resources. By integrating preprocessing steps like contour analysis and response identification using machine learning, the system boosts reliability even under imperfect conditions.

Current research in this area focuses on improving the tolerance of the system to varied environmental conditions and different sheet formats. With the advancements in Deep learning models and robust image technique assist in building modern OMR systems can evaluate answer sheets with minimal human intervention.

II. MACHINE LEARNING AND IMAGE PROCESSING

ML and IP are the core technologies driving the automation of OMR systems. IP techniques such as basic filtering and segmentation techniques such as color channel reduction, dynamic thresholding, and shape-based region detection help in preprocessing the scanned OMR sheet and isolating the answer bubbles accurately. These techniques ensure that the system can handle varying lighting conditions, orientations, and noise in the input image. ML algorithms are then used to classify whether a bubble is filled or not based on pixel intensity patterns and shape features. ML models can be trained to detect anomalies like multiple markings or faint fills, enhancing evaluation accuracy.

III. LITERATURE REVIEW

Research work presented by various authors related the OMR sheets analyser and evaluation described as below.

Zeki Kucukkara and Abdullah Erdal Tumer (2018) [1] proposed an OMR system that primarily utilizes image processing techniques for answer sheet evaluation.

The method begins with grayscale conversion and thresholding to prepare the scanned sheet for analysis. Contour detection and region segmentation are applied to isolate answer bubbles. The system then detects marked responses using pixel intensity-based decision logic. High accuracy is achieved without requiring expensive OMR hardware, making it cost-effective and practical for educational use.

Pooja Raundale et al. (2019) [2] developed an OMR system using OpenCV for efficient and accurate detection of marked responses on answer sheets. The methodology includes image preprocessing steps such as grayscale conversion, Gaussian blurring, and thresholding. Contour detection is used to locate bubbles, followed by sorting and mapping to specific questions. Filled options are identified based on pixel density in each bubble area. This system eliminates the need for specialized scanners and works well with standard webcams and printers.

Harendra Kumar et al. (2019) [3] proposed a machine learning-based approach for analyzing OMR sheets to enhance accuracy and reduce manual effort. Their system begins with image preprocessing, including resizing and noise removal. Features extracted are utilized to detect filled bubbles, and an ML classification model is trained to identify marked responses. The model is capable of handling partially filled and misaligned bubbles. This approach ensures flexible, cost-effective evaluation without the need for high-end scanning devices.

Qamar Hafeez et al. (2023) [4] proposes a robust OMR system designed to handle errors and noise during mark detection. The methodology includes preprocessing steps like image binarization and noise reduction to improve mark clarity. Fault-tolerant algorithm is used that accurately identify partially filled or misaligned marks. The system also integrates adaptive thresholding for dynamic mark detection under varying lighting conditions. Finally, it validates the detected marks against answer keys to ensure reliable evaluation. This paper focuses on improving OMR accuracy in imperfect scanning environments.

Effat Somaiya et al. (2023) [5] presents an OMR system using a standard webcam to capture answer sheets, making it cost-effective and accessible. The methodology involves image acquisition through the webcam, followed by preprocessing steps such as grayscale conversion and thresholding to enhance mark visibility. It uses contour detection to identify answer bubbles and classify filled versus unfilled responses. The system incorporates error handling techniques to manage variations in lighting and sheet alignment. This approach aims to provide a practical OMR solution for educators with limited resources.

L Pham Doan Tinh and Ta Quang Minh (2024) [6] introduces a fast and efficient scoring system for paper- based MCQ tests using object detection techniques. The methodology employs a Convolutional Neural Network (CNN) to quickly detect and classify filled answer bubbles on scanned sheets. Preprocessing includes image normalization and alignment correction to ensure consistent input quality. The framework also integrates a marking scheme to automatically score and generate results based on detected answers. This system focuses on improving speed and accuracy in large-scale exam scoring.

Rusul Hussein Hasan et al. (2024) [7] proposes an OMR system based on a modified Bi-directional Associative Memory (BAM) neural network for improved pattern recognition. The methodology involves preprocessing scanned answer sheets through noise filtering and normalization. The BAM network is then trained to recognize filled marks by associating input patterns with correct outputs. Fault tolerance is achieved by effectively handling partially marked or distorted bubbles. The paper demonstrates improved accuracy and robustness in recognizing marked answers compared to traditional methods.

Sujal Rooge et al. (2024) [8] presents an optical mark recognition system leveraging classical image processing methods. The methodology includes image acquisition followed by grayscale conversion and adaptive thresholding to separate marked areas from the background. Morphological operations are applied to remove noise and enhance mark contours. The system detects filled bubbles by analyzing pixel intensity and shape features within predefined answer regions. This paper emphasizes simplicity and efficiency, making it suitable for real-time OMR applications with minimal hardware requirements.

Dharmik R. C. et al. (2024) [9] propose an OMR evaluation system that combines two key components: image processing and machine learning for improved accuracy. The methodology starts with image preprocessing including noise removal and normalization, followed by segmentation to isolate answer bubbles. It then applies a dual-component classifier that uses both shape and intensity features to detect marked responses reliably. The system also includes error correction mechanisms to handle ambiguous or partially filled marks. This approach enhances detection precision and reduces false positives in OMR scoring.

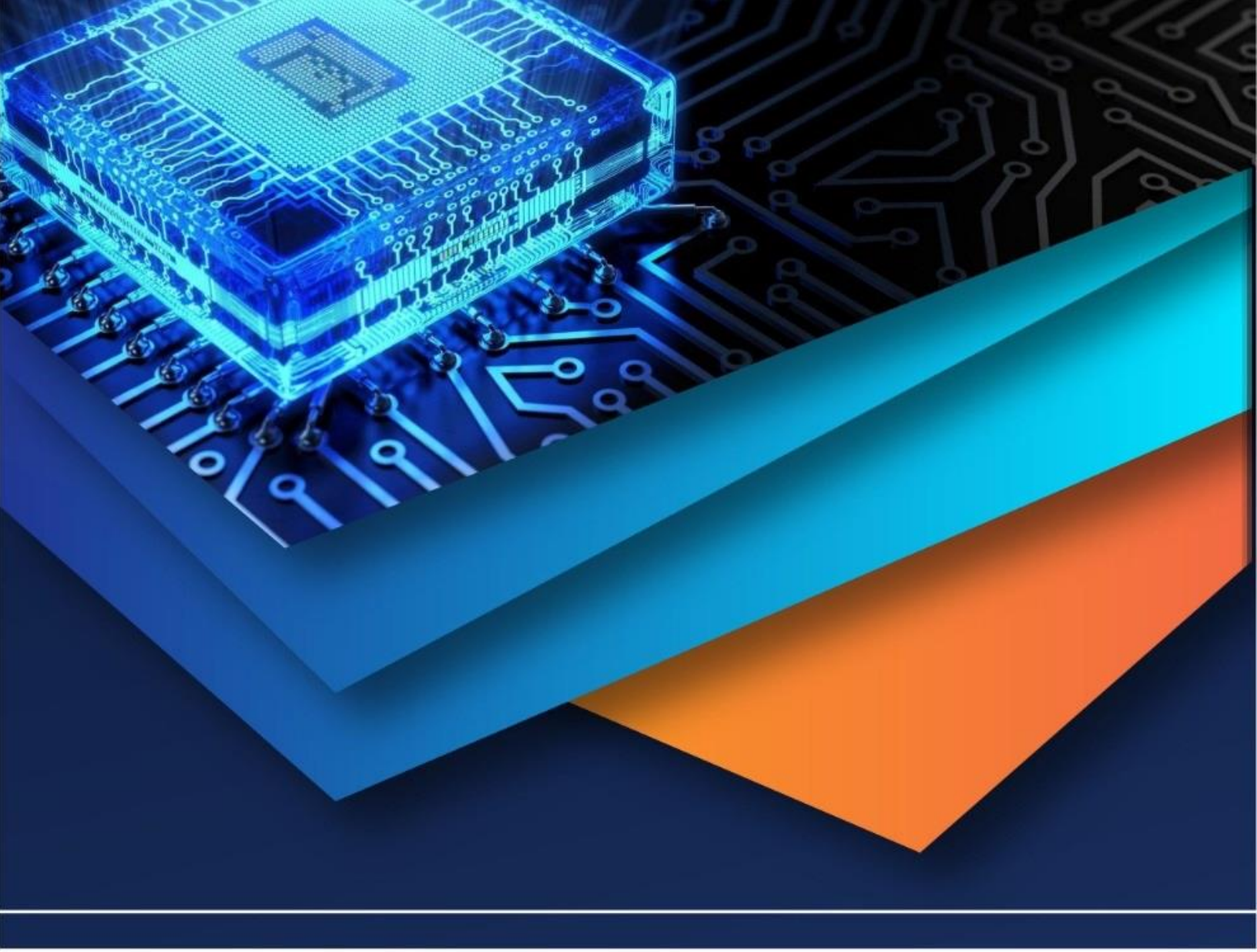
Rushikesh G. Dongare et al. (2024) [10] proposes an advanced OMR system using NL algorithms with IP techniques for improved accuracy. The methodology includes image acquisition followed by preprocessing steps such as binarization and noise reduction. Features are extracted from answer bubbles, which are then classified using supervised ML models to distinguish filled from unfilled marks. The system also incorporates data augmentation to improve model robustness against variations in answer sheet conditions. This approach aims to provide a reliable and scalable OMR solution adaptable to diverse testing environments.

IV. CONCLUSION

The paper reveals significant progress in OMR systems, with growing adoption of machine learning and deep learning techniques. While recent models incorporate CNN and real-time image analysis. Open-source tools like OpenCV, TensorFlow, have further accelerated innovation in this domain. The surveyed studies emphasize accuracy, robustness, and affordability—especially in education and assessment contexts. However, challenges such as noise, overlapping marks, and real-time processing remain areas for continued research. Notably, webcam-based and mobile-friendly solutions make OMR more accessible. The insights from related works have shaped the direction of our system, validating its practical potential. Overall, the integration of ML and Image Processing approaches is transforming OMR into a scalable, intelligent, and efficient solution.

REFERENCES

[1] Zeki Kucukkara, Abdullah Erdal Tumer, “Image Processing Oriented Optical Mark Recognition and Evaluation System”, Journal of Image Processing and Pattern Recognition, Vol. 12, 2018, pp. 68-75.
[2] Pooja Raundale, Taruna Sharma, Saurabh Jadhav, Rajan Margaye, “Optical Mark Recognition using OpenCV”, International Journal of Computer Applications, Vol. 152, 2019, pp. 45-52.
[3] Harendra Kumar, Himanshu Chauhan, Sonam Mittal, “Analysis of OMR Sheet Using Machine Learning Model”, International Journal of Advanced Research in Computer Science, Vol. 10, 2019, pp. 84-91.
[4] Qamar Hafeez, Waqar Aslam, M. Ikramullah Lali, Shafiq Ahmad, Mejdal Alqahtani, Muhammad Shafiq, “Fault Tolerant Optical Mark Recognition”, Proc. of IEEE, Vol. 45, 2023, pp. 321-330.
[5] Effat Somaiya, Alifa Sun Mim, Mohammed Abdul Kader, “Webcam-Based Robust and Affordable Optical Mark Recognition System for Teachers”, Journal of Image and Video Processing, Vol. 36, 2023, pp. 100-107
[6] Pham Doan Tinh, Ta Quang Minh, “Automated Paper-Based Multiple Choice Scoring Framework Using Fast Object Detection Algorithm”, Proc. of International Conference on Computer Vision, Vol. 40, 2024, pp. 250-257.
[7] Rusul Hussein Hasan, Inaam Salman Aboud, Rasha Majid Hassoon, Ali Saif Aldeen Aubaid Khioon, “Optical Mark Recognition using Modify Bi-Directional Associative Memory”, Journal of Computer Vision and Applications, Vol. 42, 2024, pp. 111-120.
[8] Sujal Rooge, Veeresh Kotresh Kadli, Shivanand Manjare, Varun Kumar M B, Suguna A, “OMR Detection Using Image Processing Technique”, International Journal of Advanced Computer Science, Vol. 39, 2024, pp. 180-188
[9] Dharmik R. C., Rangari S., Jain S., Nilawar A., Deshmukh G., Yeole A., “Optical Mark Recognition Evaluation System using Dual-Component Approach”, International Journal of Data Analysis, Vol. 21, 2024, pp. 203-210.
[10] Rushikesh G. Dongare, Prof. Tanuja S. Mulla, Gauri Y. Kakulte, Manjiri M. Tamkar, Shraddha S. Tambekar, “OMR System Empowered by Machine Learning and Image Processing”, International Journal of Computer Applications, Vol. 52, 2024, pp. 142-150.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY