Webcam Based Robust and Affordable Optical Mark Recognition System for Teachers

Effat Somaiya[1], Alifa Sun Mim[2], Mohammed Abdul Kader[3]
[1,2,3]Department of Electrical and Electronic Engineering, International Islamic University Chittagong, Bangladesh

Article Info

Article history:

Received Aug 30, 2024
Revised Oct 18, 2024
Accepted Nov 5, 2024

Keywords:

OMR
Image Processing
Grading Automation
OpenCV
Webcam

ABSTRACT

The growing need for efficient automated grading solutions has driven advancements in optical mark recognition (OMR) systems for multiple-choice assessments. This paper introduces a novel webcam-based OMR system that employs advanced image processing and computer vision techniques to eliminate the dependency on specialized hardware. The proposed system enhances image quality, extracts relevant data, and accurately processes marked responses through a robust pipeline of preprocessing, segmentation, and recognition algorithms. Addressing challenges such as inconsistent handwriting styles and varying lighting conditions, the system demonstrates high accuracy and reliability, achieving an impressive accuracy rate of 100%. Experimental validation highlights significant improvements in grading efficiency, reduced human error, and enhanced consistency when compared to manual grading methods. The scalability of the system makes it applicable to remote learning environments, online exams, and large-scale assessment scenarios. Future research directions include integrating machine learning techniques to extend the system's capabilities to subjective assessments and potential collaborations with educational institutions and online platforms. This research contributes to the field by providing an accessible and scalable automated grading solution that optimizes assessment workflows and improves the educational experience.

Corresponding Author:

Mohammed Abdul Kader
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chattogram-4318, Bangladesh
Email: kader05cuet@gmail.com

1. INTRODUCTION

Computer-based image processing and recognition technologies have become essential across various sectors, including automation, security, and biomedical applications [1]. Despite their widespread use, the application of these technologies in educational settings has been relatively limited. One significant opportunity for advancement is in the automated grading of multiple-choice exams, a task traditionally performed by optical mark recognition (OMR) systems. OMR systems are designed to process data from specially prepared forms, efficiently handling large volumes of documents such as exams and surveys [2, 3]. However, these systems often come with high costs due to the expensive optical scanners and maintenance required. This financial burden poses a substantial challenge, particularly for schools and universities in developing countries, where budget constraints make it difficult to procure and maintain such equipment [4]. Consequently, the high cost of traditional OMR systems limits their accessibility and prevents the broader implementation of automated grading solutions, hindering potential improvements in efficiency and accuracy in educational assessments.

In recent years, research has focused on making optical mark recognition (OMR) systems more affordable and accessible by using readily available hardware, such as webcams. For instance, Eyegrade offers a low-cost solution using a standard webcam for both mark detection and optical character recognition [4]. Although OMR is widely used in research facilities, universities, schools, and other organisations to process hand-filled documents such as timekeeping records, elections, surveys, questionnaires, exams, and replay cards, the technology often remains costly. Schools, in particular, face challenges in acquiring and using this equipment due to budget constraints. To address this, a mobile application was developed to assess students' response sheets using a phone or tablet, eliminating the need for optical devices [4]. This software exhibited high accuracy, recognizing forms with 90% to 99.7% accuracy within 20 seconds. Research into webcam-based grading has explored improving image processing algorithms, machine learning for accurate grading, user experience, scalability, system robustness, and data security. For example, Eyegrade's solution eliminates the need for response bubbles by detecting marks and recognizing handwritten student IDs [5]. Another study developed a real-time OMR system using a webcam to read handwritten data from plain paper with high accuracy [6]. Classical OMR techniques include template matching, pixel-based analysis, and thresholding methods. Mitsuji et al. introduced a template-matching processor using pixel sums and squared deviations [7]. Zou et al. enhanced this with a new thresholding method using arctangent Hausdorff distance for improved pattern matching [8]. Images in optical form were recognised by identifying key points and establishing the image's alignment with these points [9]. Simplified methods have further advanced OMR technology. A study explored template matching with an affordable webcam, accurately grading answers by comparing cropped regions with template images [10]. Another study from the same year used a webcam and a small OMR form [6]. In 2019, researchers developed an OMR system using OpenCV to automate MCQ grading by capturing images, detecting bubble contours, and matching them with correct answers [11].

Building on previous advancements, this research develops an optical mark recognition (OMR) system using a standard webcam and the OpenCV library for image processing. The proposed system addresses traditional OMR's challenges by offering a fully automated and cost-effective alternative, eliminating the need for expensive hardware while maintaining high accuracy and adaptability across various environments [6, 12]. By integrating advanced image processing techniques such as adaptive thresholding [13, 14], the system can dynamically adjust to varying conditions. This approach enhances the efficiency and accessibility of grading processes in educational institutions, making it a valuable tool for those with limited financial resources [15–17].

The novelty of this research lies in its ability to replace traditional OMR scanners with a more accessible solution that requires only a standard webcam and open-source software. This research provides a smartphone with a software camera-based grading approach as an efficient and accessible alternative to standard grading systems. Unlike previous systems that often required manual adjustments or specialized equipment, this system offers a fully automated process that can be easily implemented in resource-constrained environments. This makes it a valuable contribution to the fields of image processing and educational technology, addressing a critical need for cost-effective and scalable solutions in automated exam grading.

2. METHODOLOGY

Webcam-based Optical Mark Recognition (OMR) automates data extraction from forms, such as surveys, tests, and questionnaires, using image processing techniques in Python and OpenCV. This approach reduces the need for manual data entry and minimises human error. Python's extensive libraries enable efficient integration of OMR algorithms, while OpenCV provides essential tools for image analysis, including contour detection and region identification. The methodology consists of image acquisition via webcam, preprocessing to improve image quality, feature extraction to identify marked regions, and result interpretation. OpenCV's image processing capabilities are critical for accurately extracting and analysing features. The presented diagram depicts the sequential steps involved in the proposed work's methodology, as shown in Figure 1.

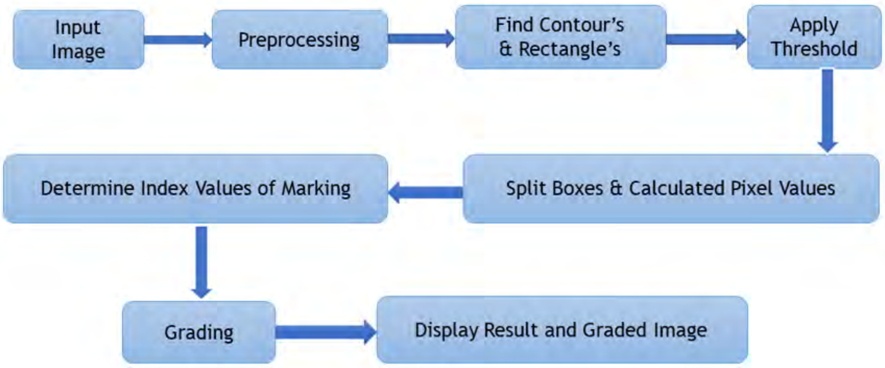

Figure 1. Block Diagram of Proposed Methodology

2.1. Image Capture and Preprocessing

The methodology begins by capturing an image of the Optical Mark Recognition (OMR) answer sheet using a standard webcam. The preprocessing phase follows, aimed at enhancing the image quality for subsequent analysis. The initial step involves resizing the captured image to standardize the dimensions across all samples, ensuring uniform processing. During this step, duplicate copies of the resized image are created to preserve intermediate results, facilitating comparison and verification of processing stages, as illustrated in Figure 2.

Subsequently, the resized image undergoes conversion to grayscale, which simplifies the image by reducing it to a single intensity channel. The grayscale value is computed using the following formula:

$$\text{Grayscale Value} = (R \times 0.2999) + (G \times 0.587) + (B \times 0.114) \tag{1}$$

Where:

- **R** is the (Red) value of each pixel which contributes about 29.99%,
- **G** is the (Green) value of each pixel which contributes about 58.70%,
- **B** is the (Blue) value of each pixel which contributes about 11.40%.

This conversion step reduces color complexity and prepares the image for more efficient further processings as illustrated in Figure 3.

To mitigate the impact of noise on image quality, a Gaussian blur is applied to the grayscale image. This technique smooths the image and is mathematically expressed as follows:

$$G_{xy} = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{2}$$

Where:

- $\mathbf{G}_{xy}$ is the Gaussian value at point (x,y),
- **x** and **y** are spatial coordinates,
- σ controls the spread (standard deviation) of the function,
- **2**π normalizes the function so its total area is 1,
- The exponential function **exp** gives G_{xy} its Gaussian (bell-shaped) profile. The negative exponent $(-(x^2 + y)^2 / (2\sigma^2))$ causes G_{xy} to decrease as (x,y) moves away from the origin.

Gaussian blurring facilitates more accurate edge detection by reducing high-frequency components, such as noise, that could interfere with subsequent processing steps. This step helps improve the accuracy of edge identification, as shown in Figure 4.

Finally, the Canny edge detection algorithm is applied to the blurred grayscale image. This algorithm identifies significant edges and boundaries, isolating regions of interest for further analysis, including contour detection and region identification, as shown in Figure 5. The edge-detected image provides the necessary input for extracting contours and regions of interest in the next steps.

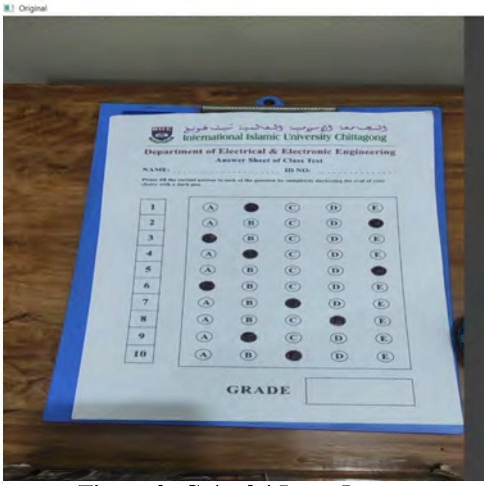

Figure 2. Colorful Input Image

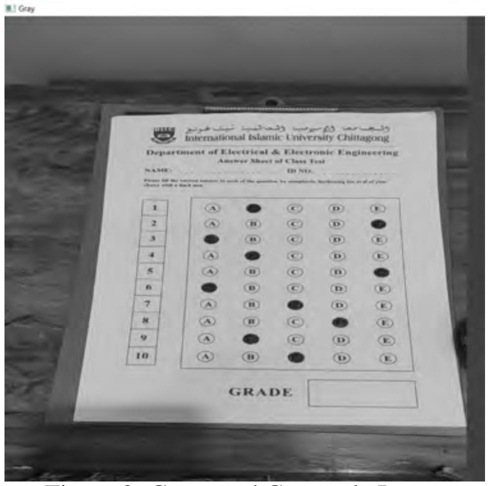

Figure 3. Converted Grayscale Image

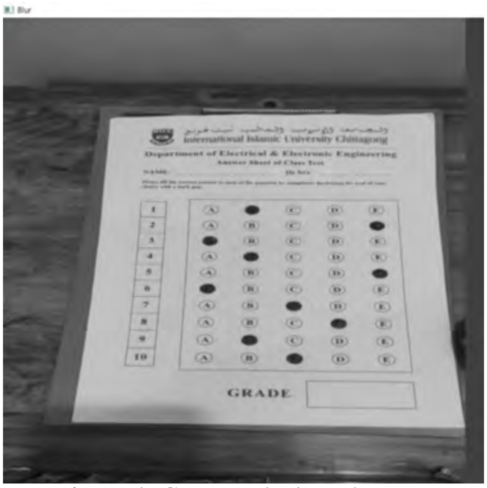

Figure 4. Converted Blurred Image

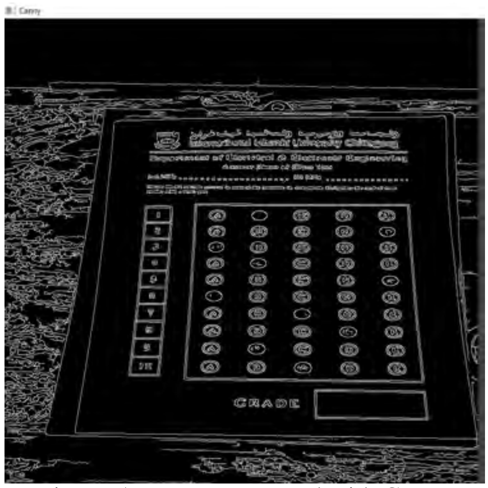

Figure 5. Image Processed with Canny Algorithm

2.2. Contour Detection and Region Identification

Following the preprocessing steps, the next phase involves detecting contours to identify key regions of interest on the OMR answer sheet. OpenCV's contour detection method is utilized to locate rectangular shapes that correspond to the answer sheet. This is achieved through contour approximation, which helps filter out non-relevant contours based on criteria such as aspect ratio and contour area.

The largest contour, which corresponds to the answer sheet, is then selected. To facilitate alignment and further processing, the corner points of this contour are extracted and displayed in Figure 6. In addition, the specific region designated for displaying grades is depicted in Figure 7. These extracted contours are crucial for ensuring precise alignment and minimizing distortion in the subsequent steps, which are necessary for accurate grading.

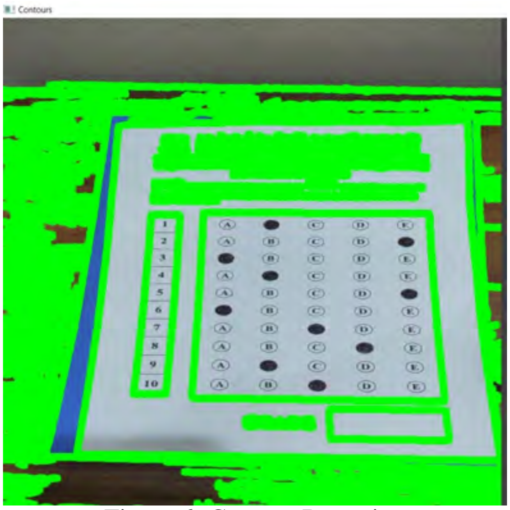

Figure 6. Contour Detection

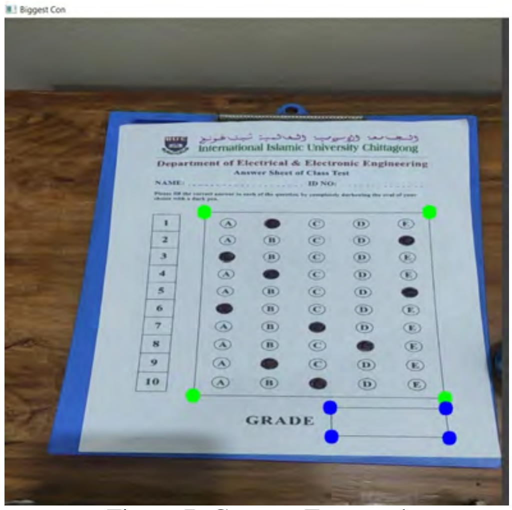

Figure 7. Contour Extracted

2.3. Perspective Transformation

To correct any distortions and ensure proper alignment of the answer sheet, a perspective transformation is applied. This process involves calculating a transformation matrix using the extracted corner points of the sheet. The matrix maps the quadrilateral region to a rectangular shape, effectively rectifying the image. The transformed image is then displayed with the grading section properly aligned, as shown in Figure 8, which illustrates the corrected perspective of the answer sheet and grade display area.

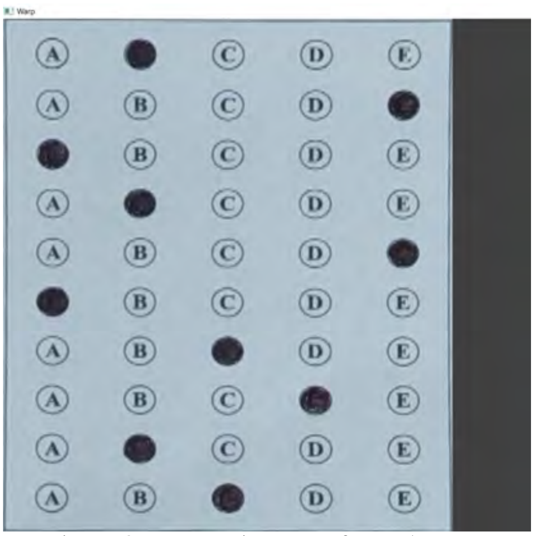

Figure 8. Perspective Transformed Image

2.4. Thresholding

Otsu's method is applied to threshold the grayscale image dynamically, converting it into a binary image. The following is an expression for the thresholding formula used in image processing:

$$I_{\text{out}}(x, y) = \begin{cases} 255, & \text{if } I_{\text{in}}(x, y) > T \\ 0, & \text{if } I_{\text{in}}(x, y) \leq T \end{cases} \quad (3)$$

Where:

- $I_{\text{in}}(x, y)$ is the input image intensity at pixel coordinates (x, y),
- $I_{\text{out}}(x, y)$ is the output binary image intensity at pixel coordinates (x, y),

- T is the threshold value.

This thresholding method adapts to varying lighting conditions, ensuring accurate detection of marked responses. The result of this process is displayed in Figure 9.

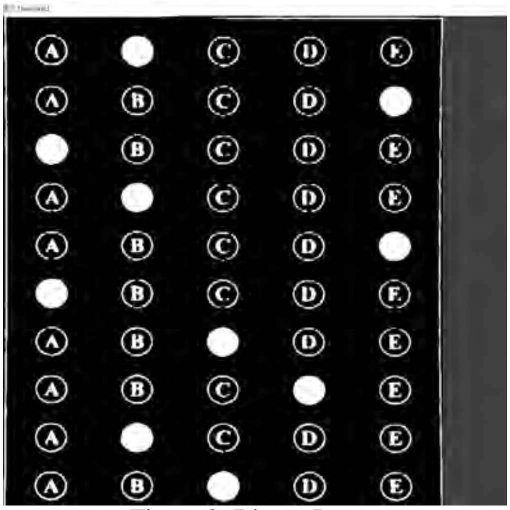

Figure 9. Binary Image

2.5. Answer Box Segmentation and Pixel Analysis

The binary image is segmented into individual answer boxes, and the number of non-zero pixels in each box is counted to detect the marked answers. This allows for precise identification of selected answers based on the density of markings. Figure 10 illustrates the pixel data stored in an array, representing the density of markings in each answer box. The system leverages this data to differentiate between marked and unmarked answers, ensuring accurate grading even in cases of partial marks or incomplete selections.

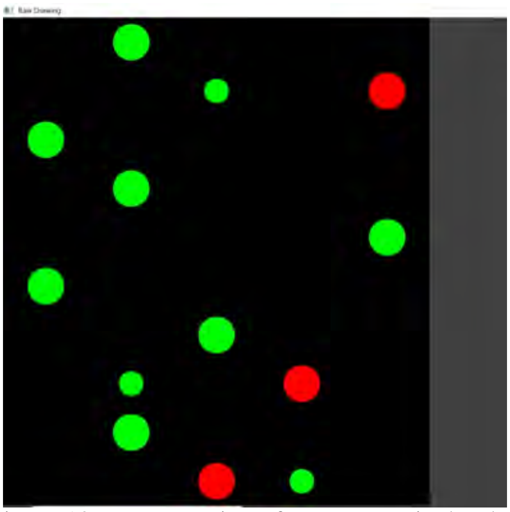

Figure 10. Raw Drawing of Non-Zero Pixel Value

2.6. Mark Detection and Grading

Marked answers are identified by selecting the box with the highest pixel count for each question. These marked responses are compared to the correct answers, assigning a score of 1 for correct responses and 0 for incorrect ones. The final grade is calculated as the percentage of correct answers out of a total and it is presented in Figure 11. The threshold for detection was set at 145, optimizing the system's performance for different lighting conditions.

Figure 11. Graded Image

2.7. Result Visualisation and Output

The system visualises the grading results by overlaying circles on the transformed answer sheet to mark correct and incorrect answers. It compares the user's responses with the correct answers and marks the results as green for correct and red for incorrect, while also showing the correct answer for any wrong response. The final output includes the evaluated answer sheet, displaying the marked answers, correct answers, and the computed grade. The system also generates a file with the results and updates the grade display. Intermediate images, such as the original, grayscale, contour, and transformed images, are stacked to showcase the processing steps visually.

3. RESULTS AND DISCUSSION

The proposed system, utilizing a webcam and advanced image processing techniques with Python and OpenCV, provides an efficient solution for automating the grading of multiple-choice assessments. It is characterized by high accuracy, fast processing, and robustness to varying lighting conditions and sheet orientations. These attributes ensure reliable and consistent grading performance in diverse examination environments. Additionally, the system minimizes the potential for human error and significantly reduces grading time, enabling rapid feedback for students. The system's flexibility allows it to accommodate different answer sheet formats, making it adaptable for use in a variety of educational contexts. Furthermore, its ease of integration with other educational platforms enhances its applicability in automated assessment systems.

3.1. System Performance

The system demonstrates excellent performance, achieving 100 percent accuracy in recognizing all marked answers and processing each form in under 2 seconds. During the prepossessing phase, the images are resized, converted to grayscale, blurred using Gaussian filtering, and processed with Canny edge detection to enhance contour identification. Thresholding techniques are then applied to segment the answer bubbles, with marked answers identified by analyzing pixel intensities. Grading is done by comparing the selected answers to the answer key, and the final score is displayed on the processed answer sheet. Additionally, the "Stacked Images" display in Figure 12 illustrates the various stages of the optical mark recognition process, providing a comprehensive view of the system's workflow.

Figure 12. Layout of Multiple Images

The final output, as shown in Figure 13, visually highlights correct and incorrect responses, providing users with a clear assessment of their performance.

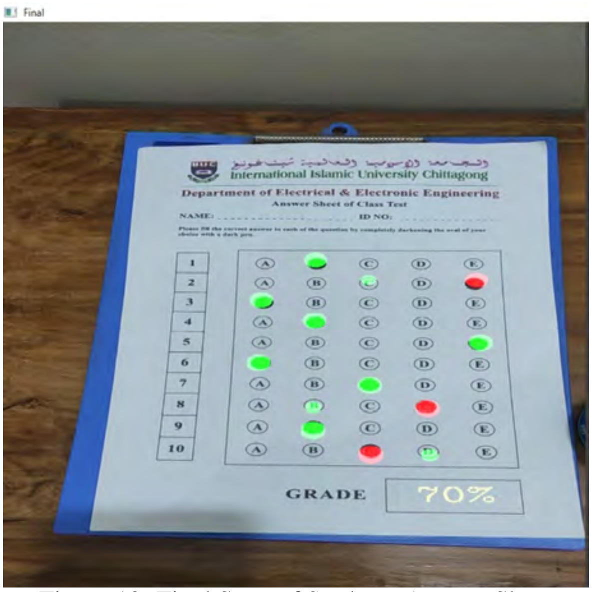

Figure 13. Final State of Students Answer Sheet

3.2. Orientation Handling

The system effectively handles different orientations of answer sheets, as demonstrated in Figures 14 through 17. This flexibility ensures accurate grading regardless of how the sheets are positioned, benefiting diverse testing environments.

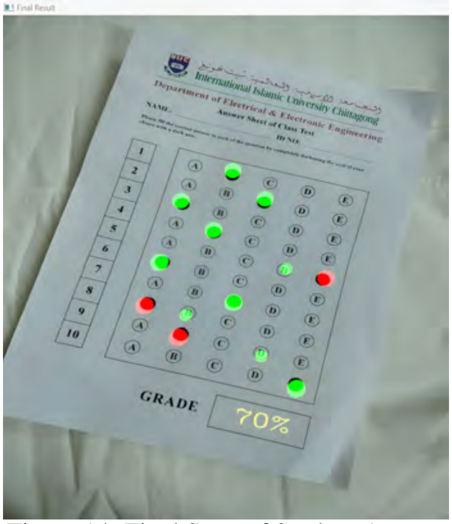

Figure 14. Final State of Student Answer Sheet (approximate +30 degree)

Figure 15. Final State of Student Answer Sheet (approximate -30 degree)

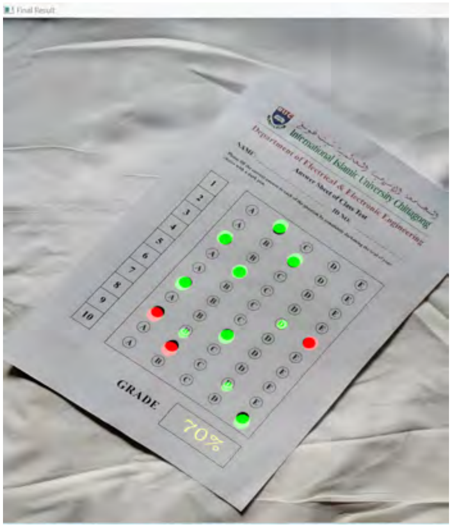

Figure 16. Final State of Student Answer Sheet (approximate +45 degree)

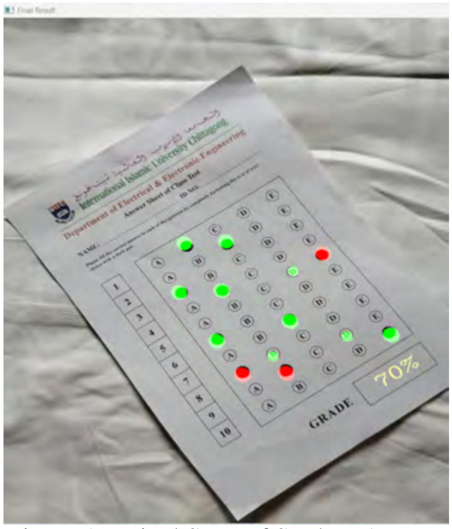

Figure 17. Final State of Student Answer Sheet (approximate -45 degree)

When the grade position on answer sheets changes, the system detects and adjusts accordingly as shown in Figure 18.

Figure 18. Position of Grade Changed in the Layout

3.3. Adaptability to lighting conditions

The system dynamically adjusts thresholds for individual image blocks, enabling it to handle varying lighting conditions. It performs consistently well in low-light environments, as presented in Figure 19 and 20, ensuring reliable results across different scenarios. The combination of adaptability, efficiency and diverse environments makes the system a robust solution for educational institutions.

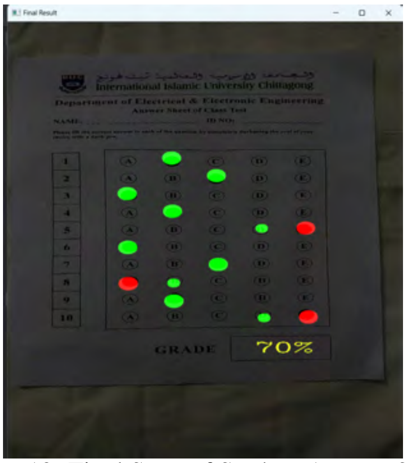

Figure 19. Final State of Student Answer Sheet (Under Low Light)

Figure 20. Final State of Student Answer Sheet (in different orientation under low light)

3.4. Performance of densely packed columns and rows

The system efficiently handles densely packed rows and columns of answer sheets, as shown in Figure 21. High performance and dependability are maintained by the system even in settings with densely packed rows and columns.

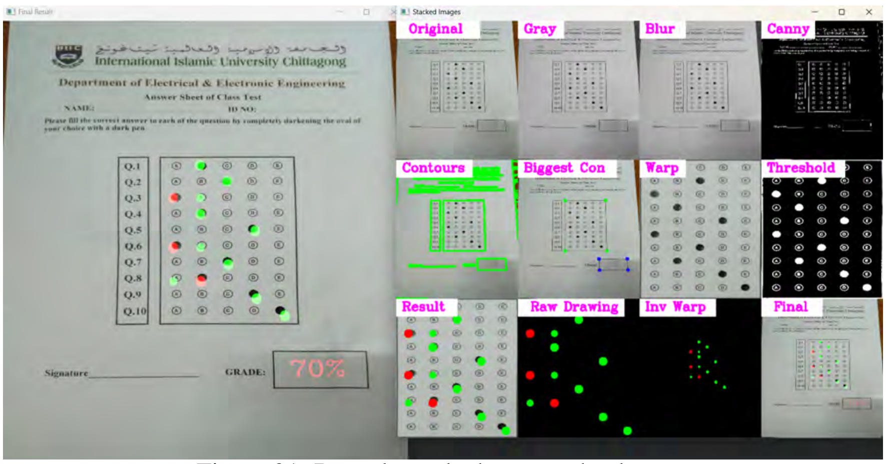

Figure 21. Densely packed rows and columns.

3.5. Challenges and Improvements

While the system excels in most cases, challenges arise with blank answer sheets and those filled in parallel formats, which can complicate response identification as illustrated in Figure 22 and 23. Rotated answer sheets (90° and 180°) also pose alignment issues, leading to potential errors and it is presented in Figure 24 to Figure 26. These cases highlight areas for further refinement.

Figure 22. Blank Paper

Figure 23. Parallel Format

Figure 24. Descending Format (Rotated 180 degree)

Figure 25. 90 degree Rotated (Left Side)

Figure 26. 90 degree Rotated (Right Side)

3.6. Comparison with Previous Work

Table 1. The following table summarizes the system's performance compared to previous research

Previous Work	Process Time	Accuracy	Device Used	Low Light Handling	Different Angle Test	Grade Position Changed
Huseyin Atasoy [6](2015)	20s	98.86%	Computer	Yes	N/A	N/A
Kamuju AbhiSubrahmanyam [18](2022)	1s	99.76%	Computer	N/A	N/A	N/A
Nalan Karunanayake [10] (2015)	<2s	97.6%	Webcam	N/A	N/A	N/A
Azman Talib [19] (2015)	5s	>96%	1.3 MP Webcam	N/A	N/A	N/A
G. Himabindu [20] (2023)	60s max	99.89%	Computer	N/A	N/A	N/A
Our Work	<2s	**100%**	**Mobile and Computer**	**Yes**	**Yes**	**Yes**

Our system outperforms previous models in accuracy, speed, and versatility, making it a robust solution for various educational and testing environments.

3.7. Future Work

Future enhancements will focus on refining the system's ability to handle blank and rotated answer sheets. Additionally, the integration of machine learning techniques will expand its capabilities to include subjective assessments, making the system even more comprehensive and adaptable.

4. CONCLUSION

The optical mark recognition (OMR) system developed using image processing techniques has shown high accuracy and efficiency in automating the grading of multiple-choice exams. The system has proven adaptable to various answer sheet designs, demonstrating its robustness. However, challenges such as overlapping or incomplete marks and varying lighting conditions highlight areas for further improvement. Error analysis has identified potential sources of inaccuracies, including image noise and thresholding issues. Addressing these will improve grading accuracy. The system processes answer sheets quickly, making it suitable for large-scale assessments. Its advantages over manual grading include significant time savings and reduced human error, making it a reliable tool for educational and testing institutions. Future work will focus on overcoming current limitations, integrating machine learning to enhance recognition accuracy, and expanding capabilities to handle subjective assessments. Optimizing computational efficiency will further improve processing speed and scalability. In conclusion, this OMR system provides a valuable solution for automated grading, with ongoing research poised to enhance its performance and impact across various assessment domains.

REFERENCES

[1] G. Samtaş and M. Gülesin, "Sayısal görüntü işleme ve farklı alanlardaki uygulamaları," *Ejovoc (Electronic Journal of Vocational Colleges)*, vol. 2, no. 1, pp. 85–97, 2011.

[2] H. M. Parul and M. Kaur, "A novel optical mark recognition technique based on biogeography based optimization," *International Journal of Information Technology and Knowledge Management*, vol. 5, no. 2, pp. 331–333, 2012.

[3] Z. Küçükkara and A. E. Tümer, "An image processing oriented optical mark recognition and evaluation system," *International Journal of Applied Mathematics Electronics and Computers*, vol. 6, no. 4, pp. 59–64, 2018.

[4] A. S. Yüksel, İ. A. Çankaya, M. A. Yalçınkaya, and N. Ateş, "Mobile based optical form evaluation system," *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 22, no. 2, pp. 94–99, 2016.

[5] J. A. Fisteus, A. Pardo, and N. F. García, "Grading multiple choice exams with low-cost and portable computer-vision techniques," *Journal of Science Education and Technology*, vol. 22, pp. 560–571, 2013.

[6] H. Atasoy, E. Yildirim, Y. Kutlu, and K. Tohma, "Webcam based real-time robust optical mark recognition," in *Neural Information Processing: 22nd International Conference, ICONIP 2015, Istanbul, Turkey, November 9-12, 2015, Proceedings, Part II 22*. Springer, 2015, pp. 449–456.

[7] L. Di Stefano and S. Mattoccia, "Fast template matching using bounded partial correlation," *Machine Vision and Applications*, vol. 13, pp. 213–221, 2003.

[8] Y. Zou, F. Dong, B. Lei, L. Fang, and S. Sun, "Image thresholding based on template matching with arctangent hausdorff distance measure," *Optics and Lasers in Engineering*, vol. 51, no. 5, pp. 600–609, 2013.

[9] I. A. Belag, Y. Gultepe, and T. M. Elmalti, "An image processing based optical mark recognition with the help of scanner," *Int. J. Eng. Innov. Res*, vol. 7, no. 2, pp. 2277–5668, 2018.

[10] N. Karunanayake, "Omr sheet evaluation by web camera using template matching approach," *International Journal for Research in Emerging Science and Technology*, vol. 2, no. 8, pp. 40–44, 2015.

[11] P. Raundale, T. Sharma, S. Jadhav, and R. Margaye, "Optical mark recognition using open cv," *International Journal of Computer Applications*, vol. 178, pp. 9–12, 2019.

[12] C. Herrera, S. Del Campo, A. Dueñas, J. León, G. Kemper, and C. del Carpio, "Implementation of a digital image processing algorithm on a colibri imx6 embedded industrial system for optical mark recognition," in *Brazilian Technology Symposium*. Springer, 2017, pp. 9–22.

[13] D. Sas and K. Saeed, "Comprehensive performance evaluation of various feature extraction methods for ocr purposes," in *Computer Information Systems and Industrial Management: 14th IFIP TC 8 International Conference, CISIM 2015, Warsaw, Poland, September 24-26, 2015, Proceedings 14*. Springer, 2015, pp. 411–422.

[14] T. F. Carolan, P. Bilazarian, and L. Martin, "Automated data collection and assessment for debriefing distributed simulation-based exercises (ddsbe)," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 48, no. 22. SAGE Publications Sage CA: Los Angeles, CA, 2004, pp. 2567–2570.

[15] I. A. Belag, Y. Gultepe, and T. M. Elmalti, "An image processing based optical mark recognition with the help of scanner," *Int. J. Eng. Innov. Res*, vol. 7, no. 2, pp. 2277–5668, 2018.

[16] T. V. Ha and N. T. Thu, "An application of image processing in optical mark recognition: An application of image processing in optical mark recognition," *Vietnam Journal of Agricultural Sciences*, vol. 3, no. 4, pp. 864–871, 2020.

[17] S. Feroz, U. Poornima, G. L. Shruthi, and R. Harshitha, "A survey on image processing techniques," *Journal of Emerging Technologies and Innovative Research*, vol. 3, no. 2, pp. 90–95, 2016.

[18] K. AbhiSubrahmanyam, M. S. Vineeth, Y. Mani, S. Suhith, K. PrabhathBandi, and Student, "An optical mark recognition and evaluation system based on image processing," 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:251017678

[19] A. Talib, N. Ahmad, and W. Tahar, "Omr form inspection by web camera using shape-based matching approach," *International Journal of Research in Engineering and Science*, vol. 3, no. 4, pp. 29–35, 2015.

[20] G. Himabindu, A. Reeta, A. Manikanta, and S. Manogna, "Evaluation of optical mark recognition (omr) sheet using computer vision," *International Research Journal of Modernization in Engineering Technology and Science*, no. 04, pp. 5–9, 2023.

BIOGRAPHIES OF AUTHORS

Effat Somaiya holds a degree in Electrical and Electronics Engineering (EEE) from the International Islamic University Chittagong (IIUC). Her research interests include machine learning and digital signal processing, with an emphasis on practical applications. Her final year thesis, "Webcam-Based Robust and Affordable Optical Mark Recognition System for Teachers," highlights her commitment to advancing accessible educational technology. Effat can be contacted via email at effatsomaiya28@gmail.com.

Alifa Sun Mim received her B.Sc. degree in Electrical and Electronic Engineering from International Islamic University in Chittagong, Bangladesh. She has published her research on IoT-based systems at an international conference in Algeria, with details available in the Microsoft's Conference Management Toolkit (CMT). Her research interests focus on the Internet of Things (IoT), embedded systems, machine learning, and renewable energy. She can be contacted via email at: alifamim219@gmail.com.

Mohammed Abdul Kader is currently working as a researcher on a research project under the Advanced Nuclear Technology Research Group at the Universitat Politècnica de Catalunya (UPC) in Barcelona, Spain. He is also a part-time PhD candidate in the Department of Electrical and Electronic Engineering at Chittagong University of Engineering and Technology, Bangladesh. Additionally, he holds the position of Assistant Professor in the Electrical and Electronic Engineering Department at the International Islamic University in Chittagong, Bangladesh. He received his undergraduate and graduate degrees in Electrical and Electronic Engineering from Chittagong University of Engineering and Technology. He is the author or co-author of more than 35 research papers listed in the Scopus database. His research interests include robotics, embedded systems, machine learning, and digital signal processing. He can be contacted via email at kader05cuet@gmail.com.