

# Integrated OMR and OCR Solutions: An Image Processing Perspective

Nitya Nivdunge  
B.Tech Computer  
Engineering  
Mukesh Patel School of  
Technology Management &  
Engineering NMIMS  
Mumbai, India  
nityanivdunge@gmail.com

Neel Lohith  
B.Tech Computer  
Engineering  
Mukesh Patel School of  
Technology Management &  
Engineering NMIMS  
Mumbai, India  
nityanivdunge@gmail.com

Saanvi Misar  
B.Tech Computer  
Engineering  
Mukesh Patel School of  
Technology Management &  
Engineering NMIMS  
Mumbai, India  
saanvimisar@gmail.com

Dipeesha Vartak  
B.Tech Computer  
Engineering  
Mukesh Patel School of  
Technology Management &  
Engineering NMIMS  
Mumbai, India  
vartakdipeesha@gmail.com

**Abstract — Optical Mark Recognition (OMR) and Optical Character Recognition (OCR) are important technologies in the imaging industry, widely used in automated document analysis and extracting logic calculations OMR focuses on decoding marks in pre-defined locations heavily scanned, usually analytical analysis, . However, used in test grading and voting systems that use bubbles or checkboxes, OCR extracts text and interprets it from scanned documents or images, reconstructs disclosures or handwriting is processed into system-readable records This generation of automatic statistical entry, record scoring, text-heading It also serves as the basis for a checklist-based. Both OMR and OCR are built on complex algorithms for sample recognition, segmentation, and feature extraction. Using libraries like OpenCV and Tesseract for OCR, those systems use contour evaluation, thresholding, and template matching to efficiently process large amounts of data with even greater accuracy, making smart automation easier on tasks in various fields.**

**Keywords—** *Image Processing, Optical Mark Recognition (OMR), Optical Character Recognition (OCR), Document Scanning, OpenCV, NumPy, Tesseract OCR, Contour Detection, Perspective Transformation, Grid Segmentation, Pattern Recognition.*

## I. Introduction

Image processing serves as the spine for numerous pc imaginative and prescient applications, from fundamental filtering and transformation responsibilities to complicated item detection and popularity systems. The rapid evolution of this field has been powered by means of advancements in libraries like OpenCV (Open Source Computer Vision Library) and NumPy, both of which offer green, actual-time solutions for manipulating visible records. OpenCV has grow to be the cross-to preference for building real-time vision packages, attributable to its vast array of built-in capabilities. NumPy, then again, enhances OpenCV with the aid of supplying a excessive-overall performance framework

for coping with arrays and matrices, making photo manipulation quicker and more green.

Optical Mark Recognition (OMR) and Optical Character Recognition (OCR) are effective photograph processing strategies broadly used in file processing, statistics extraction, and automatic evaluation.

### A. Optical Mark Recognition (OMR) :

OMR identifies marks on documents, inclusive of stuffed bubbles on standardized bureaucracy and checks, by way of detecting contrasts in precise areas. By spotting those marked regions, OMR structures can quickly method big volumes of bureaucracy for packages like surveys, remarks bureaucracy, and multiple-desire exams, appreciably decreasing human mistakes and processing time. The approach is predicated on analyzing the alignment, shape, and function of marked regions to correctly capture responses.

### B. Optical Character Recognition (OCR) :

In evaluation, OCR is a more superior approach that reads and converts revealed or handwritten text into machine-readable digital text. OCR uses pattern popularity, photo processing, and synthetic intelligence to recognize extraordinary characters and phrases, enabling programs like report digitization, text extraction from scanned files, and automatic data entry. Tesseract, one of the most famous OCR engines, makes use of state-of-the-art gadget studying fashions to recognize characters as it should be in diverse languages and formats.

Both OMR and OCR are fundamental to automating information extraction and analysis, with OMR pleasant acceptable for structured responses like forms and OCR excelling in text reputation. These technology, specially when better by libraries like OpenCV and Tesseract, are foundational in growing green, scalable solutions throughout industries, from training to healthcare and commercial enterprise.

## II. Methodology :

### i. Image Acquisition:

The OMR procedure starts with shooting the photograph of a shape or document. This photo is commonly received thru a scanner or digital camera in a controlled putting to make certain uniform lighting and keep away from distortions. The webcam can also be used to work on multiple MCQ papers without having to save them.

**The OMR form layout normally consists of precise regions, like bubbles or checkboxes, in which respondents mark their solutions.**

Figure 1: OMR Form Template

### ii. Preprocessing the Image :

- Grayscale Conversion:** The colored image is converted to grayscale to simplify the processing given that simplest brightness levels are had to become aware of marks.
- Noise Reduction:** Noise from the scanning manner (like specks of dust) are eliminated using filters, including Gaussian or median filters, to make sure clear contours and clean edges.
- Edge Detection:** Canny edge detection is used to highlight the edges of bubbles or marks on the sheet.

The processed image is used for contour detection in the next step.

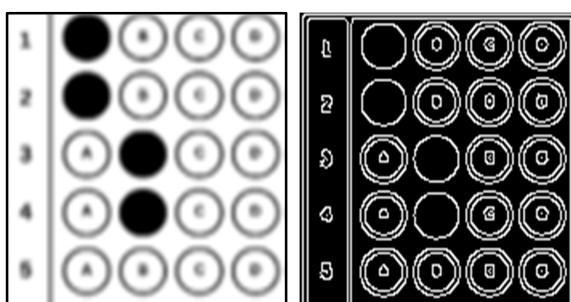


Figure 2 : Image Processing

2.1 : Noise Reduction – Gaussian Blur

2.2 : Edge Detection – Canny Edge Detection

### iii. Contour Detection :

All contours are identified from the binary (edge-detected) image, with each contour representing a potential answer bubble or bounding box on the OMR sheet.

Only the three largest rectangular contours are retained using contour area : two for questions (MCQ boxes) and one for the score area.

### iv. Region of Interest Extraction :

Corner points of the boxes are identified from the contour. Then using perspective transformation, the identified MCQ and score regions are transformed and warped into isolated images.

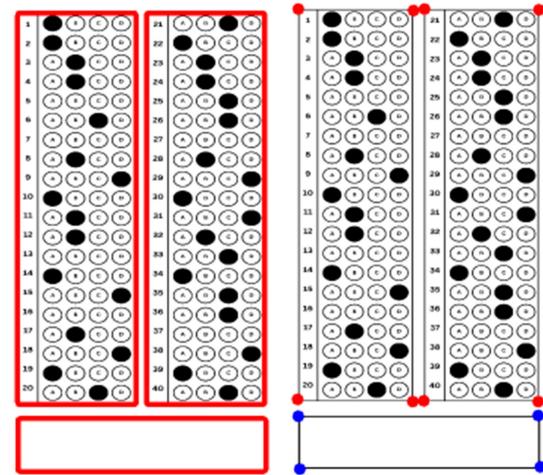


Figure 3 : Contour & Corner Points Detection

To the warped images are converted to grayscale and thresholded to create binary images where filled bubbles are highlighted.

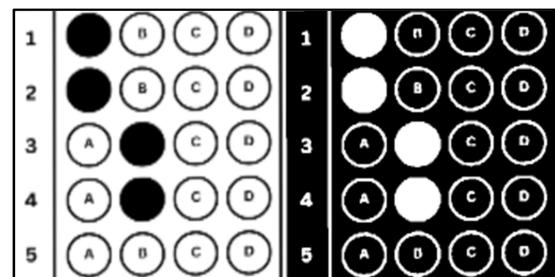


Figure 4 : Threshold Image

### v. Grid Segmentation :

Each thresholded MCQ region is divided into individual cells, with each cell representing a possible answer choice.

This is done by first equally segmenting the rows in the warped image and then segmenting each of the rows into equally spaced cells.

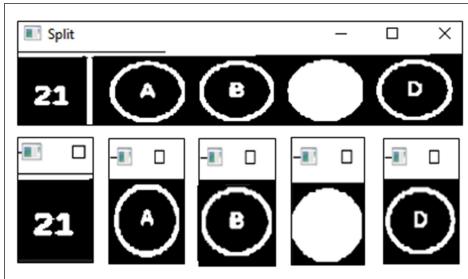


Figure 5 : Cell Formation

#### vi. Answer Extraction :

Pixel intensities are identified for each cell using cv2.countNonZero(img); higher count indicates that the cell has a filled bubble i.e selected ans. The indexes of these cells are mapped to alphabet options ( A , B , C , D ).

```
[1032. 982. 2762. 1151.]
[2677. 855. 854. 1172.]
[ 789. 825. 830. 2986.]
[ 788. 2671. 824. 1149.]
[2686. 862. 865. 1186.]
[ 801. 851. 2672. 1158.]
[ 861. 893. 864. 1284.]
[ 823. 2679. 830. 1178.]
[ 794. 830. 800. 2911.]
[2696. 862. 822. 1118.]
```

Figure 6 : Count of non-zero pixels in each cell image

Special case where no bubble is filled we check if the difference between the highest and second-highest pixel counts is below a set threshold, then answer is marked as "None" (indicating no response). The list of selected answers contains selected answers of all the questions.

#### vii. Result Analysis and Grading:

The extracted answers are compared to the provided correct answers provided initially. Right, wrong, and unanswered responses are counted, and a score is calculated based on the percentage of correct answers.

```
Collected Answers and Indices:
Q 1: ans: C, Index: 2
Q 2: ans: A, Index: 0
Q 3: ans: D, Index: 3
Q 4: ans: B, Index: 1
Q 5: ans: A, Index: 0
Q 6: ans: C, Index: 2
Q 7: ans: None, Index: -1
Q 8: ans: B, Index: 1
Q 9: ans: D, Index: 3
Q 10: ans: A, Index: 0
```

Figure 7 : Answers mapped to the selected answer cells

#### viii. Result Visualization:

The correct answers are marked on the warped MCQ images with green circle and a square around it. If selected answer is wrong then its indicated with red circle.

An inverse perspective transformation is applied to overlay the selected answers onto the original image.

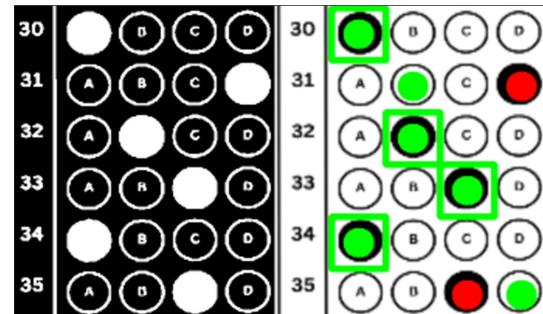


Figure 8 : Mapping of correct & wrong answers

#### ix. Final Score Display :

The final score is displayed within the score region on the original image using inverse perspective transformation.

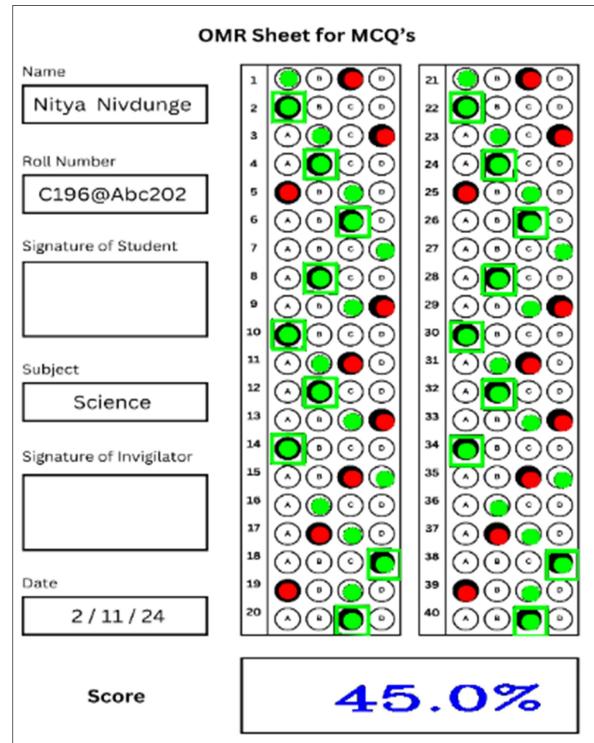


Figure 9 : Final Score Display

x. **Optical Character Recognition ( OCR ) :**

Initially the input image is cropped to the form section only. After identifying rectangular contours their areas are compared to eliminate the two largest ones which are the signature boxes.

Using Warp perspective individual images are created of the contours. These images are passed to the text extraction function that extracts text using Tesseract and returns it in string format.

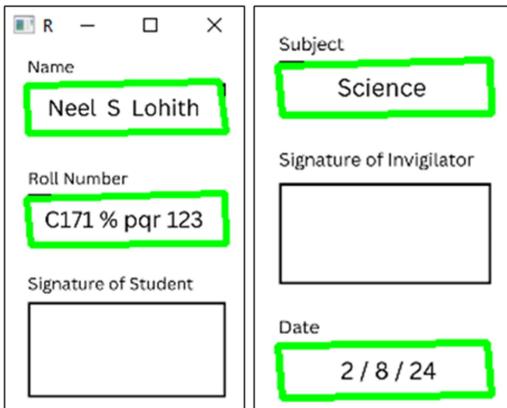


Figure 10 : Contour Identification for OCR

xi. **Database Storage :**

The text extracted from images , selected ans list , right and wrong answers , final score and grade are stored in the MySQL database.

```
Extracted text for Name: Neal S$ Lohith
Extracted text for Date: 2/8/24
Extracted text for Subject: Science
Extracted text for Rollno: C171 % par 128
```

id	Name	Rollno	Subject	Date	Scor	Grad	Selected_	Wrong		
								Ans	Right_A	Not_J
12	Saanvi Misra	C19S@XYZ	Science	28/10/24	68	C	A, A, D, B, C, C, None, B, D, A, B, B, C, A, D, No...  C, A, B, B, C, C, None, B, D, A, D, B, C, A, C, C,...	7	27	6
13	Neal S\$ Lohith	C171 % par 128	Science	2/8/24	60	C		10	24	6

Figure 11 : Extracted text is stored in Data

### III. Result Analysis :

The results of the experimental study show the efficiency of optical marker recognition (OMR) and optical character recognition (OCR) under different recognition parameters. For OMR, detection accuracy is high (100%) when 80-100% of options are filled, which is found consistently in all sample calculations. However, detection rates decrease slightly as the fill percentage decreases, with 85-90% accuracy for 50-80% filled options, even dropping to 50-65% for 0-50% ship options shows that OMR accuracy is affected by fill percentage, with higher filled channels detected better.

Table 1 : Result Analysis of OMR & OCR

Detection Criteria	Optical Mark Recognition ( OMR )		
	Example Image 1	Example Image 2	Example Image 3
100 – 80 % Option filled	100 % Detected	100 % Detected	100 % Detected
80 – 50 % Option filled	85 % Detected	90 % Detected	90 % Detected
50 – 0 % Option filled	50 % Detected	65 % Detected	60 % Detected
Detection Criteria	Optical Character Recognition ( OCR )		
	Example Image 1	Example Image 2	Example Image 3
Alphabets	90 % Detected	80 % Detected	85 % Detected
Numbers	90 % Detected	75 % Detected	80 % Detected
Special Characters / Symbols	95 % Detected	90 % Detected	85 % Detected

The number of characters detected in OCR ranges from 80% to 90% of images, indicating consistent performance with minimal variation. Numerical lines show the same coherence but with a slightly lower detection rate, especially in the example Figure 2, where the accuracy is 75%. Characters and unique symbols are identified with high accuracy, ranging from 85% to 95%, indicating the robustness of the OCR system for character recognition. Overall, the OMR and OCR systems perform well, although they exhibit some sensitivity in degrees of filling and lines, respectively. These results highlight the potential of using this technology for accurate data extraction in automated processing applications.

#### IV. Literature Review :

The methodology used in this paper integrates both Optical Mark Recognition (OMR) and Optical Character Recognition (OCR) using image processing techniques, such as contour detection and segmentation through OpenCV and Tesseract. This contrasts with Ju, Wang, and Chen's work [1], which employs Convolutional Neural Networks (CNN) on the TensorFlow platform to enhance OMR accuracy, focusing on training models to detect patterns and improve recognition in OMR applications. Similarly, Himabindu et al. [3] adopt a computer vision-based approach for OMR, applying thresholding techniques specific to detection sensitivity. However, this paper emphasizes preprocessing steps, like noise reduction and adaptive thresholding, which improve both OMR and OCR outcomes. Miguel de Elias et al. [2] provide an overview of the advancements, challenges, and limitations in OMR, emphasizing theoretical and technological progress without delving into specific software implementations. In contrast, this paper provides a practical, step-by-step methodology using distinct image processing tools, aiming for real-world adaptability in diverse layouts. Meanwhile, Kumar [4] and Patel and Prajapati [5] focus on broader OMR assessments and highlight various techniques for processing OMR sheets with 2D scanners, emphasizing traditional methods that lack the advanced adaptability found here. This paper's approach shows improved accuracy for filled and unfilled bubble detection, achieving up to 100% accuracy for highly filled options, thus aligning with and extending upon the findings of prior studies.

#### V. Advantages of Image processing techniques in this OMR & OCR application :

- i. **Enhanced Accuracy and Efficiency:** Image processing techniques, together with grayscale conversion, edge detection, and contour detection, improve the precision and velocity of Optical Mark Recognition (OMR) and Optical Character Recognition (OCR). These methods appropriately stumble on and isolate solution bubbles, enhancing the accuracy of solution extraction.
- ii. **Noise Reduction for Clearer Data:** Preprocessing techniques like Gaussian and median filtering help dispose of noise from pictures, which can also get up due to dirt, lights inconsistencies, or scanner artifacts. This makes the contours of answer bubbles and characters clearer, reducing misinterpretation and improving the reliability of the gadget.
- iii. **Edge Detection for Precise Bubble Recognition:** The use of Canny aspect detection emphasizes bubble or mark edges on OMR sheets, simplifying the mission of distinguishing stuffed bubbles from unfilled ones. This guarantees that simplest the supposed responses are identified, consequently increasing grading accuracy.
- iv. **Adaptable to Various Form Layouts:** By figuring out contours and extracting regions of interest (ROIs), the application can adapt to specific OMR layouts without requiring complete reconfiguration. This flexibility allows the equal framework to method multiple shape codecs efficiently.
- v. **Accurate Answer Detection Using Pixel Intensity:** Image processing permits the device to matter non-zero pixels in answer cells, a method that distinguishes between stuffed and unfilled bubbles correctly. This improves the detection of selected answers, even in cases of partial markings.
- vi. **Improved Data Visualization and Result Verification:** Techniques for marking accurate and incorrect answers with inexperienced or red circles assist customers quickly confirm grading accuracy. Additionally, inverse angle ameliorations project
- vii. **Text filtering for data storage:** OCR techniques, using conceptual transformation and filtering, enable the required format data (such as name and ID) to be captured efficiently and stored automatically in the database therefore reduces manual data entry, reduces human error and accelerates productivity.

Image processing techniques in OMR and OCR systems significantly enhance the speed, accuracy, and precision of data processing, ensuring answers are captured accurately from scanned forms by focusing on filled bubbles and comparing them to correct answers. Workflows are accelerated, analytical accuracy is increased, and human error is decreased because of this automation. Accurate text extraction made possible by integrated OCR simplifies data analysis and record-keeping. All things considered, this effective method maximizes extensive OMR and OCR projects, making it perfect for professional and educational contexts that demand dependable, quick data processing.

## VI. Disadvantages :

- i. **Image Quality Sensitivity:** The accuracy of OMR and OCR depends largely on the quality of scanned or captured images. Blurring, inconsistent illumination, or distortions can significantly affect the reliability of bubble detection and text extraction, leading to data processing errors.
- ii. **Form design complexity:** For effective handling, documents must be precisely and consistently aligned. System confusion may result from shape or configuration variations that depart from normal formatting, necessitating further adjustment and calibration. A key factor is picture alignment; if the form image is not in top view, the program may not be able to correctly detect the contours.
- iii. **Environment:** Imaging in OMR and OCR can be affected by environmental factors such as shadows or lighting changes, especially if a webcam or smartphone camera is used. Controlled lighting and scanning areas are usually required to achieve consistent accuracy.
- iv. **Failure to detect subtle marking errors:** OMR systems based on pixel intensity may struggle to detect subtle markings or light shadows, such as weak or incomplete filled bubbles. This can lead to misinterpretation of responses, especially if the respondent marks jumps irregularly.
- v. **Limited scalability for complex forms:** Forms with complex layouts or multiple regions of interest may require extensive preprocessing and setup, which may reduce system performance and scalability. Complex or non-standard forms may require additional software tuning to increase processing time.
- vi. **High-Performance Hardware Dependence:** Processing high-resolution images and implementing computationally intensive algorithms, such as edge detection and OCR, can be resource-demanding. Systems with lower performance may experience delays or

performance which is slower, and impacts real-time or bulk -Processing result requirements.

## VII. Improvements :

- i. **Enhanced preprocessing for consistent image quality :**  
Dynamic thresholding: Instead of using a fixed threshold, which may not be a good fit for different lighting conditions, more adaptive thresholds (e.g., Gaussian or adaptive mean thresholds) for both OMR and OCR components. Provides cycling rely on it.  
Automatic tilt and rotation correction: The inclusion of tilt correction algorithms will help improve image alignment, especially if captured in slight rotation or tilt, improving contour detection and segmentation accuracy
- ii. **Contour and cell detection optimization :**  
Contour filtering criteria: Changing the contour area criteria or adding an aspect ratio check can effectively remove unnecessary contours, reducing detection errors of post-bubbles and text areas.  
Improved cell division: The more sophisticated method of cell division based on dynamic mediators in splitCells can mean fewer changes in bubble placement, reducing the possibility of cell boundaries will come upon the affected areas
- iii. **Answer Detection Robustness :**  
Improved zero-pixel count analysis: Current countNonZero method works well but can be affected by noise. Adding a second evaluation of cell fill density or setting a dynamic threshold based on the average number of pixels can help reduce false detection when bubbles are not completely filled.  
Dealing with null responses: Use a more positive threshold that distinguishes marked from unmarked jumps. For example, if a bubble does not meet a set threshold, the response can be automatically marked as "none" to avoid misinterpretation of partial or missing responses
- iv. **OCR Text Detection Update :**  
Character separation in OCR: In the OCR.py phase, character overlap can be reduced by increasing character separation through similarity operations or bounding box transformations, resulting in omitted text among lies results. This modification will improve text recognition accuracy in situations of text density or overlapping characters high.
- v. **Code efficiency and speed :**  
Synchronization of dependent tasks: Using parallel processing for tasks such as cell splitting and pixel count analysis speeds up execution, especially useful when processing large numbers of images in time in the actual

- Task modularization: Breaking large tasks into smaller more specific tasks (e.g., separate tasks for thresholding, contour filtering, etc.) can simplify debugging and maintenance, and increase readability and scalability
- vi. **Data validation and error handling :**  
 Increased pressure in error handling: Adding inspection and error handling procedures to deal with unexpected problems (such as blank contours or misaligned cells) will increase the robustness of the system further, reducing the risk of it being on the accidental or unfair results of the act.  
 Validation of edge cases: The inclusion of validated special information, such as partially filled bubbles or open responses, will help accurately interpret ambiguous signals, increasing reliability.
- VIII. **Future Scope & Applications :**
- i. **Healthcare Sector**
    - **Automated Clinical Documentation:** Real-time conversion of handwritten notes into electronic health records (EHRs) reduces errors and enhances data accessibility.
    - **Telehealth Integration:** Virtual consultations are more efficient when patient data is captured more efficiently.
    - **Patient safety:** Prescriptions are automatically analysed to identify any mistakes and interactions.
  - ii. **Financial Services**
    - **Intelligent Document Processing:** By automating data extraction from contracts and invoices, Advanced OCR speeds up processing and lowers errors.
    - **Cheque Processing:** By precisely reading printed and handwritten information, OCR improves cheque clearing.
  - iii. **Education Sector**
    - **Automated Grading Systems:** By combining OCR with OMR, multiple tests can be graded quickly and effectively, providing valuable information about student performance.
    - **Digital Learning:** OMR makes it easier to collect data in real time from online tests, while OCR digitises textbooks.
  - iv. **Retail and E-Commerce**
    - **Intelligent Inventory Management:** OCR increases inventory accuracy by automating the reading of product labels.
    - **Analysis of Customer Feedback:** OMR examines surveys to extract insightful information about customers.
- v. **Government and Public Services**
  - **Efficient Processing:** OCR improves service delivery by streamlining the processing of applications and forms.
  - **Data analytics:** The digitisation of public documents facilitates resource allocation and well-informed decision-making.
- vi. **Legal Sector**
  - **Document Discovery:** To facilitate effective searches, OCR speeds up the digitisation and classification of legal documents.
  - **Contract Analysis:** To help with compliance monitoring, Advanced OCR extracts important clauses from contracts.

## IX.

### Conclusion :

This study effectively demonstrates the practicality and effectiveness of merging OMR and OCR technologies utilizing sophisticated image processing methods, distinguishing itself from previous research by employing a practical, software-based approach based on OpenCV and Tesseract OCR. Unlike the CNN-based methodology used in Ju, Wang, and Chen's study [1], which uses deep learning to discover patterns in OMR, this study employs contour and noise reduction techniques to improve accuracy and speed response extraction. This system displays applicability across a wide range of applications, including healthcare, education, and finance, where automated data extraction can result in considerable productivity increases. The use of noise filtering and edge detection for bubble and character recognition distinguishes this study from the more standard OMR methodologies described by Kumar [4] and Patel and Prajapati [5], which focus on simpler OMR assessments using 2D scanners.

While database storage ensures efficient data management, this study's reliance on contour detection and adaptive thresholding makes it adaptable to various layouts, overcoming the challenges of OMR technology identified by Miguel de Elias et al. [2], such as processing limitations in complex environments. However, as with many image-based systems, accuracy is still dependent on hardware quality and environmental conditions, emphasizing the need for even more robust and adaptable solutions to improve scalability and resilience in varied environments.

## X. References :

1. Y. Ju, X. Wang and X. Chen, "Research on OMR Recognition Based on Convolutional Neural Network Tensorflow Platform," 2019 11th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Qiqihar, China, 2019, pp. 688-691, doi: 10.1109/ICMTMA.2019.900157. keywords: {Q measurement;Mechatronics;Automation;OMR;Tens orFlow;CNN;Model},
2. Miguel de Elias, Erik & Tasinaffo, Paulo & Hirata Jr, Roberto. (2021). Optical Mark Recognition: Advances, Difficulties, and Limitations. SN Computer Science. 2. 10.1007/s42979-021-00760-z.
3. G. Himabindu, A. Reeta, A. S. Manikanta, and S. Manogna, "Evaluation of Optical Mark Recognition (OMR) Sheet Using Computer Vision," International Research Journal of Modernization in Engineering Technology and Science, vol. 5, no. 4, pp. 1-5, Apr. 2023.
4. S. Kumar, "A Study on Optical Mark Reader," *Galaxy International Interdisciplinary Research Journal (GIIRJ)*, vol. 3, no. 11, pp. 40-44, Nov. 2015.
5. N. V. Patel and G. I. Prajapati, "Various Techniques for Assessment of OMR Sheets through Ordinary 2D Scanner: A Survey," *International Journal of Engineering Research & Technology (IJERT)*, vol. 4, no. 9, pp. 1-6, Sep. 2015, IJERTV4IS090675.