A Comprehensive MATLAB based Optical Mark Reader and Database Generation System for the Windows Operating System

Samrina Zamir[1,*], Saba Javed[2], Saqib Jamshed Rind[3],

[1]College of Computing and Information Sciences, KIET, Karachi, Pakistan
[2]Department of Electronic Engineering, NEDUET, Karachi, Pakistan
[3]Department of Automotive Engineering, NEDUET, Karachi, Pakistan
[*]Corresponding author: zsamrina@gmail.com

Abstract

The importance of software systems in business cannot be overstated. Often, there are numerous processes involved in developing software solutions, including problem definition, solution design, building, testing, and maintenance. The objective of this study is to propose an optical mark reader (OMR) that is practical, affordable, and compatible with Windows. The OMR can be used to process data in polling places, schools, elections, and other comparable places. It was created especially for grading multiple-choice tests. It may be easily integrated with a multipurpose printer and comes with an updated tool for image scanning and reading. The article also provides a database that generates thorough reports and was built in a soft environment. The paper uses an aptitude test grading system for a university to show how well the system works, and the database successfully creates a detailed report of the student's data and final results. The system was created utilizing state-of-the-art software, including MATLAB and SQL server.

Keywords—Optical mark reader, Database generation, Image Processing, Grading System, MATLAB

1 Introduction

An optical mark reader (OMR) is a device that can read markings made on paper using some ink or pencil. The markings are typically in the form of bubbles, squares, or other shapes that can be filled in by the user[1]. The OMR uses a light source and sensors to detect the presence/absence of marks in the specific locations on the paper. Next, it interprets the marking as data, such as multiple-choice answers, and records them in a database or computer System. OMR has a variety of applications other than multiple-choice paper grading, such as processing surveys and tallying votes in Elections[2]. The system can process large sets of data quickly and accurately[3].

The existing and well-established OMR systems mostly work with a dedicated scanner that gleams a beam of light onto the sheet that creates a contrasting reflectivity at decided locations on the page[4]. This data is then used to distinguish the spotted areas because they reflect less light than the blank areas of the sheet. The paper aims to present the design and development of a user-friendly OMR with a database generation system. The proposed system provides an easy-to-use GUI for the user and a complete database generation system that is capable of generating several compiled reports as per the requirement. The complete block diagram of the proposed OMR system is given in Fig. 1.

The rest of the paper is organized as follows, Section II describes the sheet designing software, its mechanism, and its specifications. Sections III and IV discussed the image processing and representation techniques respectively. Section V discusses the key features of GUI and its development followed by Section VI which explains the database generation.

2 Sheet Designing

There have been numerous paper size standards across different countries and time periods. However, a widely accepted international ISO standard currently exists,

ISSN: 2523-0379 (Online), ISSN: 1605-8607 (Print)
DOI: https://doi.org/10.52584/QRJ.2102.05

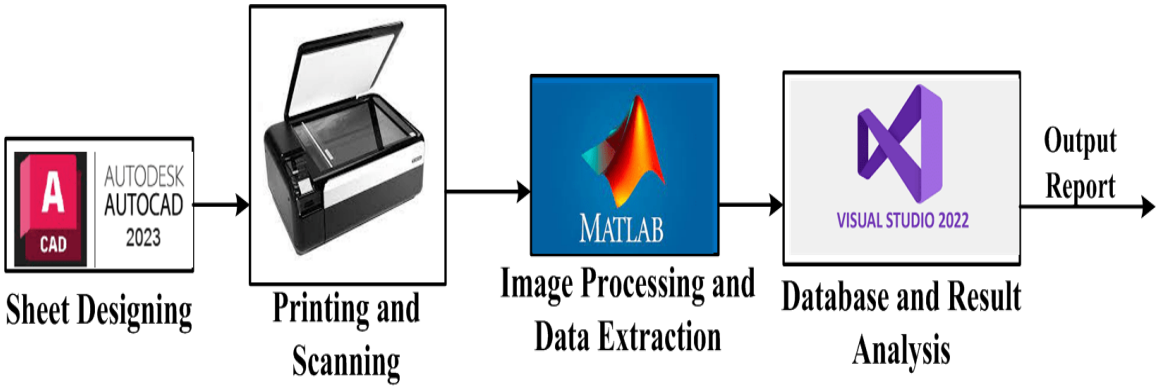

Fig. 1: A complete block diagram of the Proposed OMR system

which includes A4, B3, C4, and others. A4 is the most commonly used paper size among them. It is to be noted that the key feature of an OMR system is its scaling, where dividing a sheet with an aspect ratio of $\sqrt{2}$ into two equal halves parallel to its shortest sides will result in the halves having the same aspect ratio of $\sqrt{2}$. Various software programs are available for designing such sheets, and the choice of software ultimately depends on the user's preference for convenience, ease of installation, and familiarity with use and handling. Following this fact, in this paper, the author has used AutoCAD for sheet design, a software application for architects, engineers, graphic designers, and other professionals to create two-dimensional (2D) and three-dimensional (3D) models of buildings, products, and other designs. This tool is well-thought-out to be the best and most widely used software in the design, engineering, construction manufacturing, media, and entertainment disciplines [5].
Fig. 2. shows the AutoCAD designed sheet. The following instructions were followed while designing the sheet.

1) Specify the size of the sheet (Letter/ A4)
2) Specify the number of regions
3) Specify the number of rows and columns for each region
4) Specify the pixel standard for each region
5) Specify the pixel spacing for both horizontal and vertical arrangements
6) Proper and standard Margin across the sheet.

3 Image Processing Techniques

Applications for optical mark readers (OMR) software depend substantially on image processing[6]. An OMR system can precisely capture and process data from scanned papers that contain multiple-choice answers with the aid of image processing algorithms. Image processing, in particular, assists in improving the quality of scanned images, extracting pertinent information, and removing noise and undesirable markings [7]. Image enhancement, segmentation, feature extraction, and pattern recognition are just a few of the capabilities and methods available in the versatile image processing software MATLAB [8]. An OMR system can improve accuracy, processing speed, and overall performance by utilizing the capabilities of MATLAB [9]. Therefore, it is essential to incorporate image processing methods into an OMR software application to guarantee accurate and trustworthy data processing. For image processing, picture reading, image editing, extraction, filtration, and data comparison in this paper, as well as for the development of output, the author used MATLAB R2022b.

4 Image Representation

MATLAB offers a variety of image representation formats, such as grayscale, indexed, and true-color RGB formats [10]. Each member in a matrix of M×N elements that represents a grayscale image represents the pixel's intensity level. The image's height and width are represented by the M and N dimensions, respectively. Indexed images, on the other hand, use two matrices - a colormap of size K×3, where K is the number of colors, and an indexed matrix of size M×N. The colormap contains the values for each color, and the indexed matrix identifies the pixels according to their corresponding values in the colormap. True-color RGB images are represented as a 3D double matrix, where each pixel has three values for red, green, and blue, corresponding to the color channels, along with the three dimensions. It is worth noting that grayscale format is commonly used for image processing tasks[11].

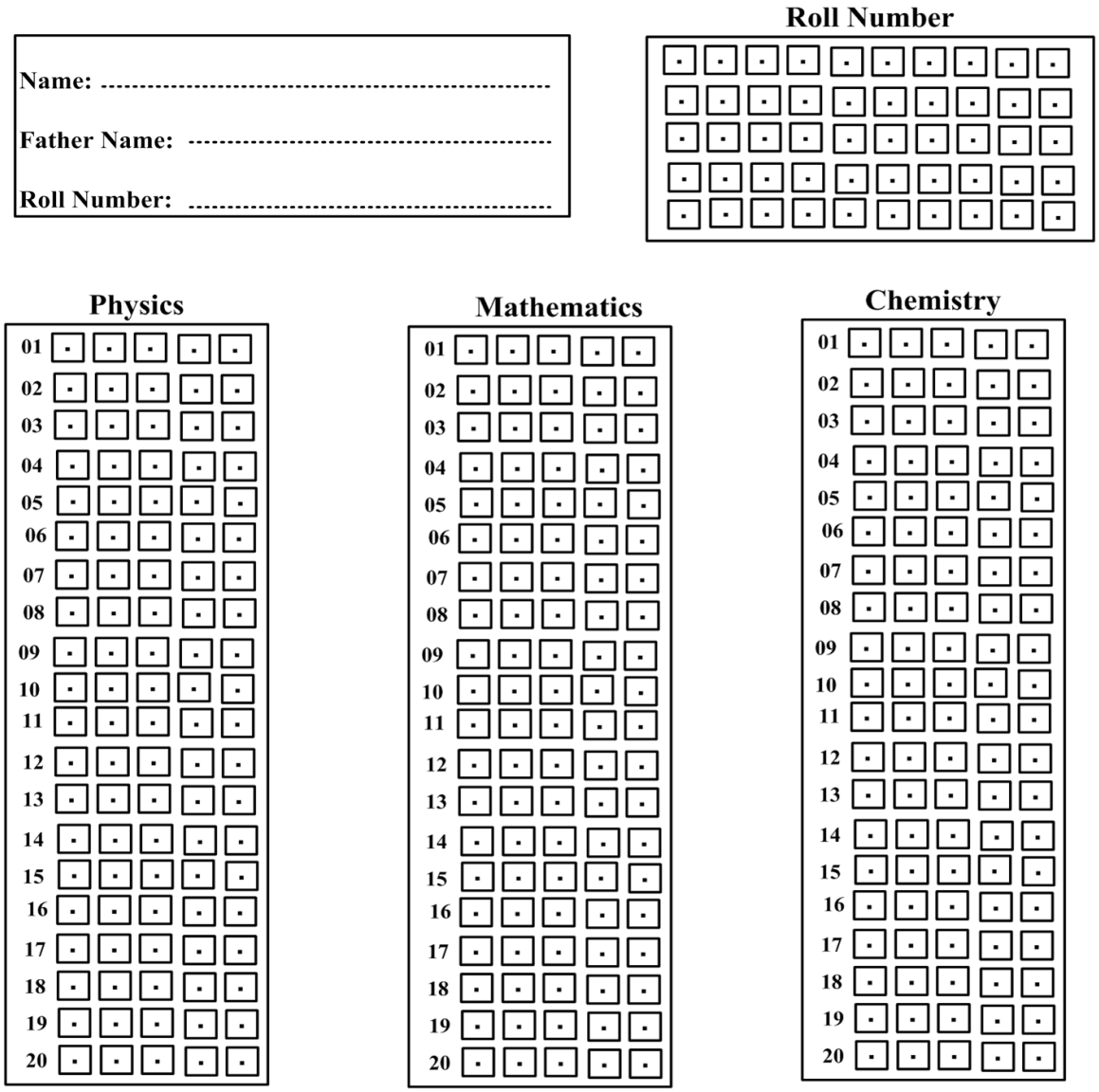

Fig. 2: Scanned Image of an AutoCAD Design Sheet

5 Development of GUI for Paper Scanner

The author has utilized MATLAB Lab Windows to create a graphical user interface (GUI). This software environment is known for its ease of use and similarity to other GUI development tools. Lab Windows operates on an event-driven model and utilizes the same libraries as LabVIEW, a popular data acquisition program. The primary focus of Lab Windows is to facilitate the development of data acquisition programs. Algorithm 1 and Fig. 3 show the steps used in the development of the Graphical User Interface and the GUI of the complete system, respectively. As seen in Fig. 4. The start and end range specifies the number of papers to be scanned. An indicator labeled as "All Selected" is also given to ensure that all the specified papers are selected, in case of any missing paper or error the "Select LED (Error) will glow. The process will start as soon as the "Strat Process" button is pressed. The relevant two indicators "Progress LED" and "Complete LED" blinks throughout the process and at the completion of the process respectively. The log at the right displays the papers that have been scanned so far. All the results from the scanned papers are exported to a text file. This is later on processed to generate the output.

Algorithm 1: Pseudo Code for Designing the GUI of Paper Scanner

1) Start of OMR Algorithm.
2) First the image is input through the scanner and the GGB image is converted to grayscale.

3) Store in Arrays. The label characters are under corresponding sections on the OMR sheet.
4) Label all circle bubbles under under corresponding sections on the OMR sheet.
5) Detect approximated pixel positions for cubes.
6) Estimate the distances between all the selected center/middle pixels of all circle bubbles.
7) Apply nested loop structure for the column and row-wise vertical and horizontal style of shading.
8) Declare the threshold value.
9) Extract the data
10) Import in text files
11) Generate the output.
12) End of OMR Algorithm.

6 Database and Result Generation

The proposed OMR utilizes SQL server management Studio (SSMS 19.0. 1) for database generation and evaluation. SSMS is software that operates databases, providing storage, access, security, backup, and other facilities. The programming software used for GUI development of result and database generation systems is Microsoft Visual Studio 2022. Fig. 4 shows the flow-chart which describes the sequential process of database and result generation system. As can be seen in Fig.4 the database generation system is responsible for saving all the primary information such as venue, city, and date related to the test, student information that includes the student's bio-data, and relevant details. Finally, the questionnaire (questions and answers both) is saved in the database. This stored information is then exported to a text file and is used later for final report generation.

Fig. 5 shows the main mania and the subsequent tabs. For instance, the student information tab contains the entire information (both personal and academic) of students along with the reports in two formats either Individual student reports or complete reports containing information of all students. The test information tab contains the primary information related to the test. Whereas, the process information menu is used for processing the saved result/information and eventually generating the individual or complete report as per requirement. Each form design has a toolbar at the top of the form with different icons referring to different control actions as listed in Fig. 6.

Please note that the icon representation and action are the same for all form designs. Out of 240 sheets, 239 were accurately examined using OMR software, whereas 1 sheet had an error as a result of a software glitch (see Fig. 7). The figure's vertical axis shows a success rate of 100% for 239 copies, indicating an exceptionally high success rate for the majority of copies and demonstrating a high level of effectiveness or reliability in the process being measured. The horizontal axis represents the number of copies in the range of 1 to 240. OMR software has consistently detected and interpreted marks on scanned papers with a success rate of 1, which is 239 out of 240 copies. This indicates a high level of accuracy.

Out of 240 sheets, 239 were accurately examined using OMR software, whereas 1 sheet had an error as a result of a software glitch (see Fig. 7). The figure's vertical axis shows a success rate of 100% for 239 copies, indicating an exceptionally high success rate for the majority of copies and demonstrating a high level of effectiveness or reliability in the process being measured. The horizontal axis represents the number of copies in the range of 1 to 240. OMR software has consistently detected and interpreted marks on scanned papers with a success rate of 1, which is 239 out of 240 copies. This indicates a high level of accuracy.

Fig. 8 shows a sample report generated by the proposed OMR system for test data. The following distinguished features can be observed.

1) Auto Resolved resolution settings.
2) Built-in printer and scanner function for all sorts of models. Therefore, it is easy to use with any printing or scanning machine.
3) Programming is done in C using Microsoft Visual Studio 2022 the latest version of Microsoft has enormous features and uses.
4) More user-friendly software.
5) Any locally installed printer and scanner could be used for the printing and scanning of sheets.
6) Compatibility with all the latest versions of Windows due to the use of running engineering tools such as MATLAB and Microsoft Visual Studio 2022
7) The OMR-based system has an Editable Sheet.

7 Conclusion and Future Recommendations

This paper introduces a fast and user-friendly optical mark reader (OMR) system for evaluation purposes. The proposed system employs advanced tools, such as AutoCAD, MATLAB, Visual Studio, and SQL database, for designing, image processing, and result compilation/generation. The proposed algorithms enable the system to generate accurate results for a large number of forms after reading and processing. The system's user interface is designed to be simple and

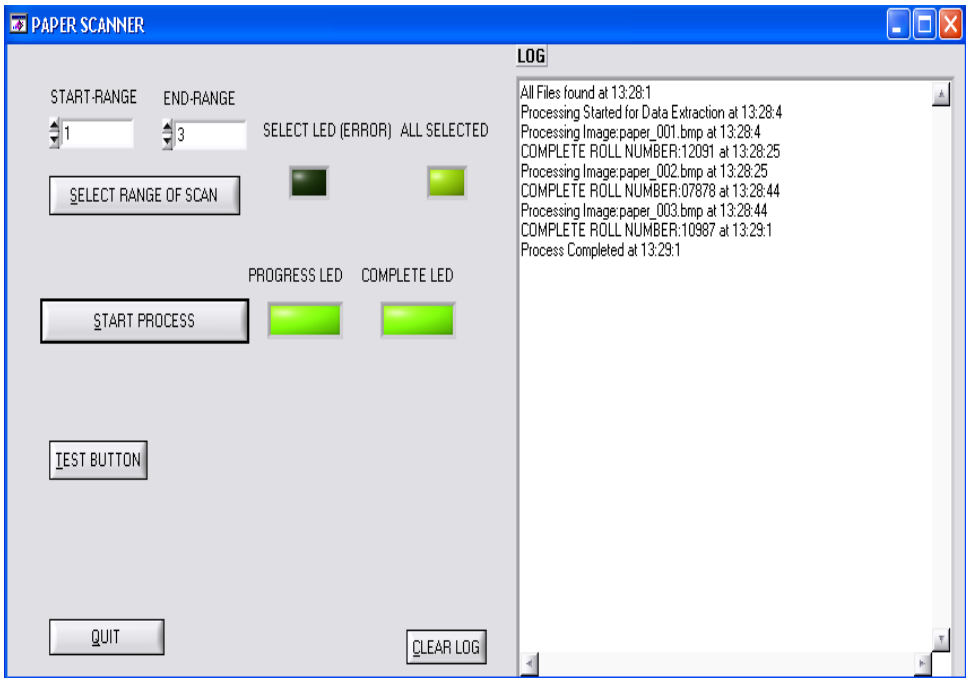

Fig. 3: Visual shot of the GUI Used

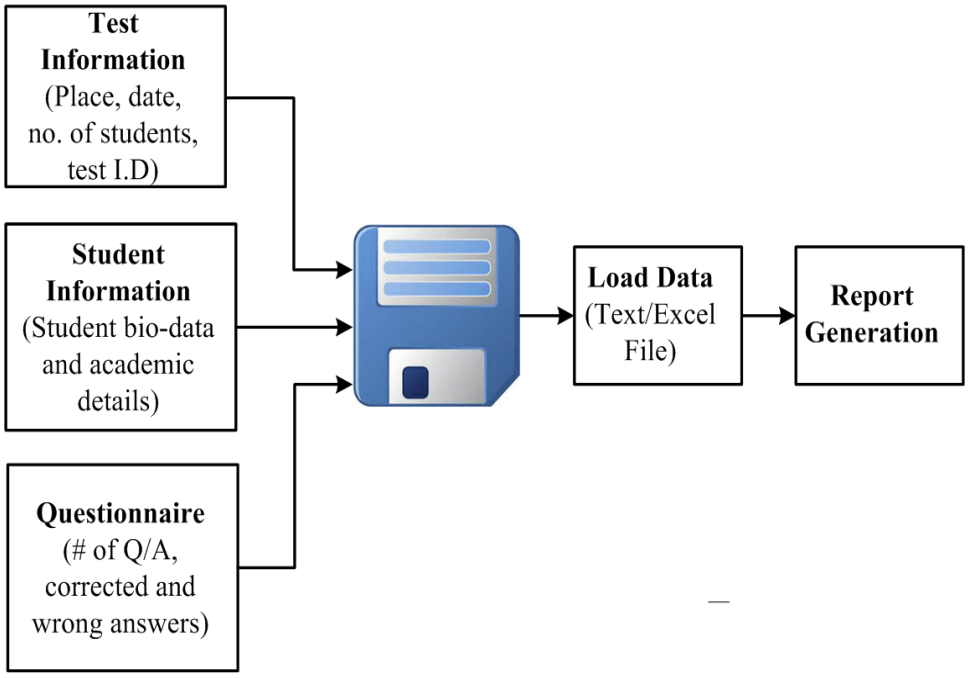

Fig. 4: Complete Flowchart of Database and Result Generation

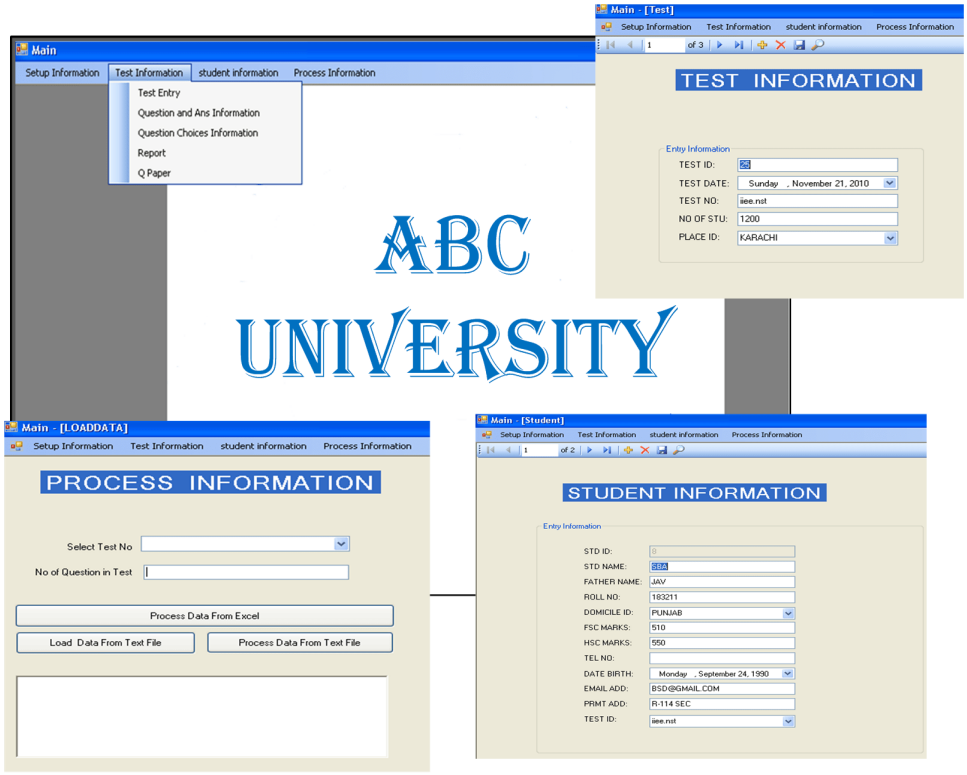

Fig. 5: A visual representation of the main form and subsequent tabs

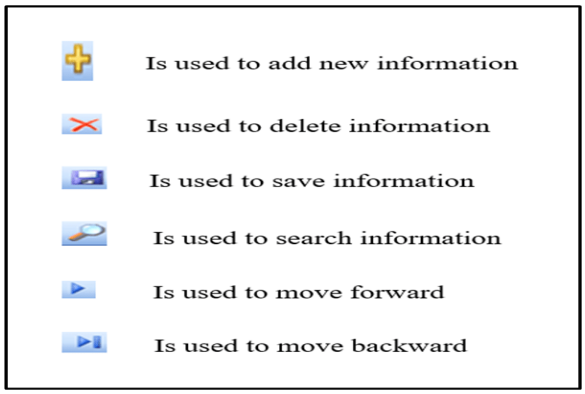

Fig. 6: Tools in the form and corresponding actions they perform

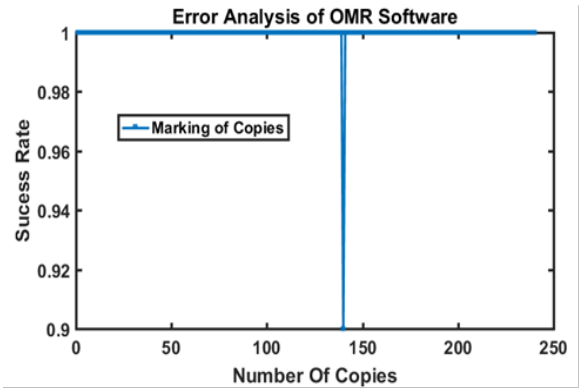

Fig. 7: Error Analysis of OMR Software

intuitive, making it suitable for use in various educational institutions, particularly universities. While the proposed system is tested in educational settings, it has the potential to be implemented in other real-time scenarios, such as attendance systems, voting systems, university admission form evaluation, and consumer and community surveys.

References

[1] Z. Küçükkara and A. E. Tümer, "An image processing oriented optical Mark recognition and evaluation system," Int. J. Appl. Math. Electron. Comput., vol. 6, no. 4, pp. 59–64, 2018.

[2] G. Nagy and D. Lopresti, "The role of document image analysis in trustworthy elections," in Advances in Digital Document Processing and Retrieval, WORLD SCIENTIFIC, 2014, pp. 51–81.

[3] R. Avyodri, S. Lukas, and H. Tjahyadi, "Optical character recognition (OCR) for text recognition and its post-processing method: A literature review," in 2022 1st International Conference on Technology Innovation and Its Applications (ICTIIA), 2022.

[4] N. V. Patel, G. I. Prajapati, and SVMIT,Bharuch., "Various techniques for assessment of OMR sheets through ordinary 2D scanner: A survey," Int. J. Eng. Res. Technol. (Ahmedabad), vol. V4, no. 09, 2015.

[5] M. P. Calado, A. A. Ramos, and P. Jonas, "An application

Fig. 8: A visual representation of compiled Reports generated by the software

to generate, correct and grade multiple-choice tests," in 2019 6th International Conference on Systems and Informatics (ICSAI), 2019.

[6] S. B. Gaikwad, 'Image Processing Based OMR Sheet Scanning,' International Journal of Advanced Research in Electronics and Communication Engi-neering (IJARECE), vol. 4, no. 3, pp. 519-522, 2015

[7] X. Wang, T. Chen, D. Li, and S. Yu, "Processing methods for digital image data based on the geographic information system," Complexity, vol. 2021, pp. 1–12, 2021.

[8] R. Ahad, R. Toufiq, and S. U. Zaman, "Information retrieval from damage OMR sheet by using image processing," in 2019 3rd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), 2019.

[9] L. Bansal, A. Sood, and H. Wani, "OMR Sheet Recognition," Int. J. Inf. & Comput. Technol., vol. 8, no. 1, pp. 11-32, 2018.

[10] Y. S. S. S. Reddy, A. S. Srinivas, and G. M. Krishna, 'OMR Evaluation using Image Processing,' International Journal of Innovations & Advancement in Computer Science, vol. 7, no. 4, 2018.

[11] N. Subramanian, O. Elharrouss, S. Al-Maadeed, and A. Bouridane, "Image steganography: A review of the recent advances," IEEE Access, vol. 9, pp. 23409–23423, 2021.