Journal of Intelligent Systems and Internet of Things
Vol. 17, No. 02, PP. 488-505, 2025
DOI: https://doi.org/10.54216/JISIoT.170231

An Intelligent Student Performance Monitoring System Using Interactive GUI and Multi-Criteria Evaluation

M. E. ElAlmi[1,*], A. F. Elgamal[2], Samar O. AbouElwafa[3]

[1]Prof. of Computer and Information System, Computer Teacher Department, Faculty of Specific Education, Mansoura University, Mansoura, Egypt

[2]Prof. of Computer Science, Computer Teacher Department, Faculty of Specific Education, Mansoura University, Mansoura, Egypt

[3]Demonstrator of Computer Teacher Department, Faculty of Specific Education, Mansoura University, Mansoura, Egypt

Emails: moh_elalmi@mans.edu.eg; amany_elgamal@mans.edu.eg; samaroday@mans.edu.eg

Abstract

Student performance during the lecture needs to be closely watched to ensure effective learning takes place. This helps the lecturer monitor the performance of the students in real time. By observing the performance of the students, the lecturer can detect the ones who find performance difficult and assist them accordingly. Besides this, the lecturer can also modify the method of teaching whenever needed. By understanding that their performance can be checked through the system, the student remains motivated to perform even in class. The study will help to develop a system that can be used to monitor the performance of the student during the real time lecture using sound and image processing. The method of developing the system involves the use of two methods: image processing and sound processing. The image processing technique can be used to detect the image of the student, while the sound processing technique will be used to detect the sound of the student during the performance. In the proposed system, Gray Level Co-occurrence Matrix technique has been used along with the Viola-Jones method to detect images along with the weighted Euclidean distance method used in image processing. Additionally, the Mel Frequency Cepstral Coefficients method has been used to detect the relevant sound along with the classification method involving the K-Nearest neighborhood method. The experiment has shown the efficiency of the system developed because the accuracy of image and sound identification of the student was at an average of 89% and 90% respectively. All of this helped to ascertain the efficiency of the system in the development of the research study.

Received: January 30, 2025 Revised: March 28, 2025 Accepted: June 30, 2025

Keywords: Student performance; Monitoring performance; Performance measurement; Viola-Jones; MFCC; K-NN; Weighted Euclidean Distance; Real time learning; Image processing; Sound processing

1. Introduction

The sudden global transition being witnessed in the shift towards online/offline learning due to recent technological advancements across the globe has greatly impacted the learning environment. Although this has opened avenues towards unparalleled flexibility and accessibility, it also presents a crucial challenge regarding the genuine measurement of student performance. The point of conflict remains in the traditional method of invigilation’s ineffectiveness at being scalable in the technological dimension. This creates a dilemma across the education sector regarding the enormous requirement of scalable and efficient methods being employed to authenticate student identity and track live lectures to avoid instances of malpractice concerning remote qualifications.

The traditional method of proctoring involves only video streaming as part of the solution and can be prone to advanced levels of imitation tactics, video playback recording, and group cheating. They can also result in the production of fake results or the failure to detect cheating practices through audio-based activities and can lead to a deterioration of security along with a distrustful environment of the student community. The research problem can be accurately articulated as follows: developing an intelligent and combined monitoring system that involves multi-modal verification in terms of constant image identification, sound detection, and real-time sound processing.

However, recent research has been tackling these problems using advanced AI models. In one study, Aly, Mohammed (2025), a novel online learning platform was presented that utilizes facial expression detection algorithms to monitor the performance of pupils in the classroom. By intermittent image capturing and image data extraction algorithms, the platform utilizes ResNet50, CBAM, and TCNs to boost facial expression detection algorithms. In which the platform achieved a level of accuracy of 91.86% accuracy of RAF-DB Dataset, 91.71% accuracy of the FER2013 Dataset, 95.85% of the CK+, and 97.08% of the KDEF Expression Dataset. In evaluating the platform's performance through state-of-the-art models using RAF-DB Dataset, CK+, CK+, and KDEF Expression Datasets revealed the importance of the findings to the educational sector [1]. In parallel efforts, Adeyemi J. O., Ogunlere S. O., & Akwaronwu (2025) conceived a system to automatically detect cheating and targeted pupils to rectify the limitations of cheating detection practices in the classroom. The system can detect cheating practices in the classroom in real time using ML algorithms and manual practices. According to the PICOS framework, the research is aimed at students who struggle with exam cheating, focuses on developing a detection system, comparing traditional monitoring techniques to the new system, improving accuracy and fairness in identifying cheating, and collects evidence using systematic review and meta-analysis methods [2]. Alnasyan, B., Basheri, M., Alassafi, M. et al (2025) developed KANFormer model, a novel deep learning (DL) architecture that integrates Kolmogorov–Arnold networks (KAN) with multi-head self-attention (MHSA) mechanisms to captures complex patterns within large-scale student datasets, enabling more accurate predictions of student performance. KANFormer processes student demographic, academic, and engagement features from the Open University learning analytics dataset (OULAD) while applying a feature elimination process to enhance dataset quality, reduce redundancy, and mitigate multicollinearity, ensuring optimal predictive performance. The experiments evaluate the model performance on binary and multi-class classification problems regarding the identification of successful and unsuccessful students, as well as the performance of each student. Ablation experiments are used to determine the effect of MHSA and KAN on model performance [3].

According to all these contemporary studies, this research proposes a novel, end-to-end remote student performance tracking and monitoring system. First, we implement an image verification system with anti-spoofing capabilities to ensure the student’s persistent presence. Second, we integrate a real-time sound verification and ambient sound classification module to detect unauthorized sound. Finally, we design the system with a focus on computational efficiency and user privacy. By fusing these verified sound and image modalities into a single, coherent platform, this research aims to set a new benchmark for secure and fair remote assessment.

The remaining sections of the paper are organized as follows: Section 1 presents a literature review. Section 2 describes the design of the proposed system and the approach used. Section 3 describes the implementation phase of the proposed system. Section 4 describes the experimental work. Section 5 interprets the results of the study while section 6 presents the conclusion of the study.

2. Related Work

Song, Xiaohuan had also identified the vital data requirements involved in the performance of the students as well as the courses in higher education. The research used the method of qualitative thematic analysis of the results of the interviews of the relevant experts to identify the important data factors involved in the development of Learning Analytic (LA) tools. The vital findings from the research mentioned the usage of RDF and graph databases in aggregating the heterogeneous learning data of the e-learning environment [4].

The SDLC approach had been applied in the student performance monitoring system developed by Jenefie A. Conde and Jerry I. Teleron. The study used the website-based platform and the data analytics approach. The system had been effective in classifying the student performance levels and offered the analytics of the student performance levels. The study found improved decision-making for the teachers and the efficiency of the academic management [5].

ADEWUMI Moradeke Grace suggested a student attendance system through UI/UX Design that would promote engagement and efficiency in the tracking of student attendances. The project applied facial recognition software and gamification components along with real-time analytics achieved through the usage of interfaces from the software ‘Figma’. The project was able to achieve accessibility and user satisfaction of the students and the administration [6].

Gomathy, C. K., A. Thamil Neethi, and S. Sai Krishna had proposed the concept of SmartGrade. It was an automated system that used the concept of machine learning to predict and evaluate the performance of the students. The concept used the Flask web framework and the SQLite database to efficiently handle academic data and the Random Forest algorithms. The system had role-based dashboards and notification services using WhatsApp,email/SMS. The results revealed a high level of accuracy of 95% and improved student, teacher, and parent communication [7].

The study of Matus, Marko, Josipa Bađari, and Igor Balaban used Moodle’s LMS system’s log files for the detection of factors predicting performance in the blended learning environment. The study showed that the engagement of users in self-assessment tasks had a slight effect directly on performance, but the results of the mid-term exams had a strong direct effect on performance. The study presents the effectiveness of the results of LMS analytics in the early detection of at-risk students [8].

A system to recognize the actions of the student in the images taken through the webcam has been suggested by Arnab Dey et al. This work differs from the already existing student action recognition through video because it only makes use of images. Recognizing student actions from images intensifies the complexity of the problem. To meet this challenge, a novel deep learning model named AdaptSepCX Attention, had designed for student action recognition in online learning environments. The proposed model outperforms other models such as NASNet Mobile, ConvXNet, and MobileNetV2 in student action recognition. Student’s action recognition has broader implications beyond the online classroom. This research enhances the understanding of student involvement in online learning and offers effective solutions for recognizing actions from images [9].

KOMMINENI, KIRAN KUMAR had proposed a method to develop a comprehensive embedded class attendance system using facial recognition by showing whether the face of the person is the student for that specified class or not. The system is based on the machine learning algorithm to be implemented on python language and using computer/laptop camera for the input image of the students or a normal outer camera. The system is programmed to handle face recognition by implementing the Local Binary Patterns algorithm LBPs. The average accuracy of recognition is 100 % for good quality images, 94.12 % for low-quality images and 95.76 % for Yale face database when two images per person are trained [10].

Maryam A. Sulaiman had introduced an online proctoring system that employs deep learning to supervise physical locations without requiring a human proctor. The system utilizes face detection and recognition algorithms to implement biometric procedures, including facial recognition. The system presents an approach for face recognition training incrementally. This eliminates the requirement for an additional step, resulting in reduced computation cost and time. To achieve high accuracy, the suggested model evaluated three distinct face detectors: Yoloface, MTCNN, and HOG. The assessment of the suggested model identified that the HOG approach has surpassed the others [11].

N. M. Alruwais and M. Zakariah had introduced a revolutionary deep learning-based approach for detecting student activities in e-class. The suggested model is based on CNN. It performs better than current methods for identifying anomalous behavior that tracks student activity in the e-class using computer vision and hardcoded algorithms. Using the UPNA Head Pose Database and a CNN model. The accuracy equals 99% for the work on ‘‘student recognition and activity monitoring in E-classes using deep learning.” To increase the model’s accuracy, the study used many approaches of deep learning, including data augmentation, transfer learning, and fine-tuning. The study’s results show how deep learning approaches may enhance student engagement and learning outcomes in online learning environment [12].

Seung Ho Seo, Min Young Kim and Yong Kim had proposed a facial expression recognition-based online video lecture system using an AI-enhanced facial recognition engine. The system aims to deliver learners' real-time participation data to instructors and help instructors concentrate on their teaching activities during video lectures. The system aims also to assist instructors in managing learner information and to enhance the efficiency and effectiveness of classes. The system uses SIFT or HOG to extract face features, and the extracted features are analyzed for emotions using CNN or SVM, and then the analysis is used to assess interest levels in the class. A pipeline using deep learning models for facial expression analysis was designed in the system, and TensorFlow was used as the AI engine for analyzing the attitudes of learners. It is expected to elevate the satisfaction levels of instructors and learners, enhancing participation and learning outcomes. All of them contribute addressing the reducing the learning disparities precipitated by COVID-19 [13].

Vijaya U. Pinjarkar, et al had discussed a student engagement monitoring system in online teaching mode. The system aims to have a score of attentiveness and an average score per student. According to the results system will give feedback, recommend activities and observe the student's behavior also monitor the clicks of mouse and keyboard. If the student is not attending, the system will notify the instructor and track website snapshots, the browser tabs and other activities. The system uses React.js to create face recognition model, Computerized Face Recognition technology to verify students from their facial biometric pattern and data, DAiSEE dataset to identify the students' boredom and Head Pose Estimation, and a computer vision technique which is Head Pose Estimation for predicting and tracking a human being using DL-based method for face detection, including Faster R-CNN or SSD. The information collected from the system is used to improve classroom, lecture training, student performance, and what questions should be asked during and after each lecture [14].

Yingying Lou and Fan Li had proposed an evaluation model of learning status of online education students based on a double improved neural network. In the suggested model, R-Net and O–Net in MTCNN can identify and extract students’ faces, and ERC is applied to the neural network structure to judge the learning status of the students with eye movement and facial expression information and obtain the result through weighted calculation. Finally, the results are evaluated, and finally the method is tested for its accuracy in practical applications. The results show that in the Fer2013 facial expression data set, the loss value reaches the lowest value and tends to be stable when the method iterates about 80 times with low recognition accuracy rate. The scoring results obtained by the proposed method can prove that the proposed algorithm has high accuracy and can apply in real environment [15].

The following section illustrates the proposed system design and methodology.

3. System Design and Methodology

The proposed system aims to monitor the performance of students enrolled in a designated course. It achieves this through several processes: a secure identity verification process, generation of adaptive exams process, finally analytical report process.

A. System Architecture

The system architecture layers are illustrated in Fig. 1, which consist of five layers: Graphical User Interface (GUI), Validation, Evaluation, Analytics Performance, and Data Layer.

1) GUI layer: Responsible for all user interaction including data input, test interface, also evaluation and results display in real time.
2) Validation Layer: Within this layer, the student's account is logged in and verified to ensure that he is enrolled in the designated course. Subsequently, the student's image is captured to confirm his identity then, the student's voice is recorded, and the system verifies that both the voice and the image belong to the same student. Once the student's identity has been properly authenticated, the student can answer the examination questions. After the examination has been completed, the student receives a performance analysis report of his results.
3) Evaluation layer: Within this layer, student assessment is conducted according to a sequence where questions are randomly generated for each student, answers are automatically checked, student scores are calculated, and then the results are displayed.
4) Performance analysis layer: A comparative analysis is conducted between the students' current test performance and their historical results past performance records from the database. This evaluation determines their performance trend and development level (e.g., Excellent, Very Good, and Good). The final outcome is a comprehensive report illustrating the student's academic progress report.
5) Data layer: Supports saving and loading of system data in database for long-term storage of performance analysis. Figure 1 illustrates the layers of the proposed student performance monitoring system. Data flow is illustrated in Figure 2.

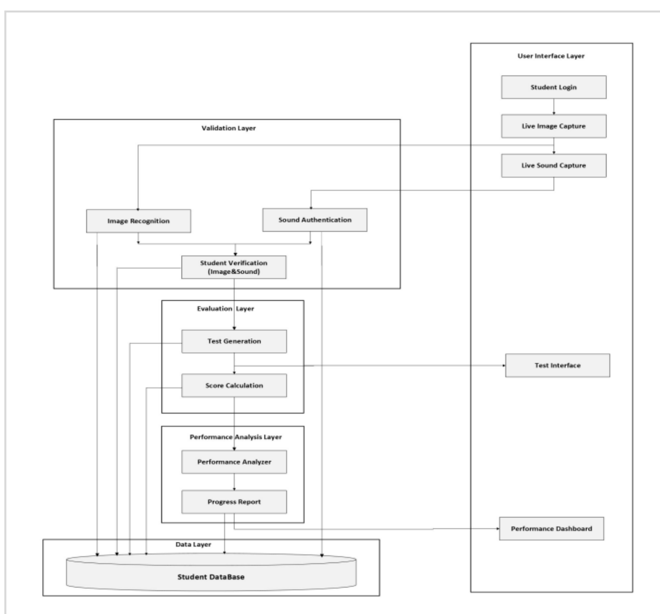

Figure 1. Layers of Student Performance Monitoring System

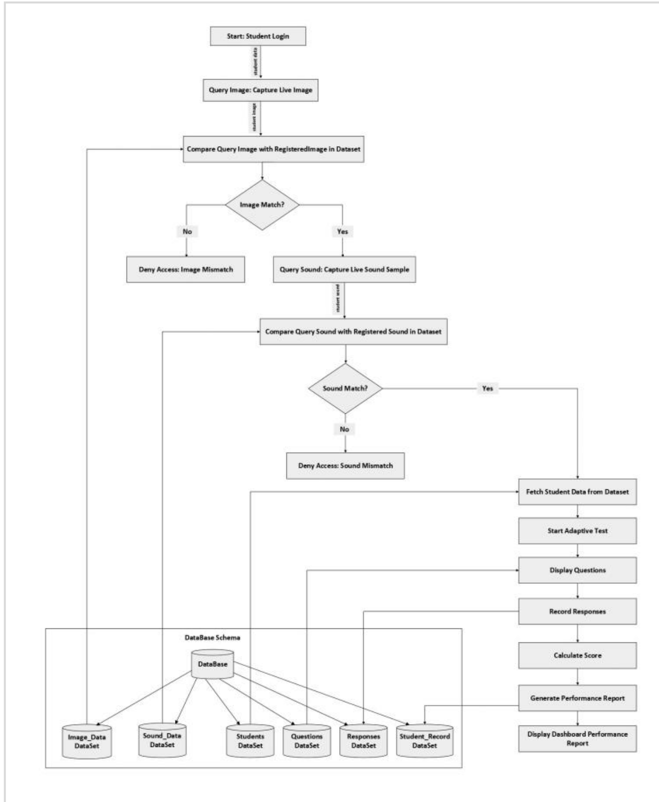

Figure 2. Data Flow of Student Performance Monitoring System

B. System Techniques

The proposed system is based on a set of image and sound processing techniques, as follows:

1) Gray Level Co-occurrence Matrix: GLCM is used to extract student face texture features. These features are calculated using the following equations [16,17]:

$$\text{Energy} = \sum_{i,j} p(i,j)^2 \quad (1)$$

Where (i, j) are a single pixel and p(i,j) is the probability that two pixels with a specified separation have grey levels i and j.

$$\text{Contrast} = \sum_{i,j} |i-j|^2 p(i,j) \quad (2)$$

$$\text{Correlation} = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j} \quad (3)$$

Where (μ_i, μ_j) represent the horizontal and the vertical means in the matrix and (σ_i, σ_j) represent standard deviation.

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (4)$$

2) Viola-Jones: The Viola-Jones face detection algorithm is a foundational computer vision technique used to detect faces in images. It was one of the first object detection frameworks to provide real-time performance. The computer vision system toolbox in MATLAB 2024 contains Vision. Cascade Object Detector System function which detects facial based on Viola-Jones face detection algorithm [18-20].

3) Weighted Euclidean Distance: There are many methods used for pattern matching. The Weighted Euclidean Distance measure is the technique used in the proposed system. It is one of the fundamental and widely used techniques in matching. The equation of WED measure can be written as follows [21,22]:

$$d(v, v^k) = \sqrt{\sum_{i=1}^{n} p_i (v_i - {v^k}_i)^2} \qquad (5)$$

Where (v_i) is used to balance the variations in the dynamic range, (p_i) is the weight added to the component and (k)is the matched image index.

$$p_i = \frac{N}{\sum_{K=1}^{N} ({v_i}^k - \bar{v_i})^2} \qquad (6)$$

Where (N) is the number of images in databases.

$$\bar{v_i} = \frac{\sum_{K-1}^{N} {v_i}^k}{N} \qquad (7)$$

4) MFCC: Mel-frequency cepstrum coefficient is one of the best distinctive features of emotion recognition problems. MFCC is based on the characteristics of the human ear's hearing & perception, which uses a nonlinear frequency unit to simulate the human auditory system [23-25]. MFCC consists of some steps. Each step has its function and mathematical approaches as following [26-28]:

5) Fast Fourier Transform (FFT): is applied to convert each frame of N samples from time domain into frequency domain. The result after this step is often referred to as spectrum or periodogram. The FFT is defined on the set of N samples {xn}, as follows:

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{-2\pi ikn}{N}} \quad , k = 0, \ldots, N-1 \qquad (8)$$

Where (X_k) is complex numbers and the resulting sequence {XK} is interpreted as follows: Positive frequencies 0 < f < Fs / 2 correspond to 0 < n < N / 2 - 1 values, while negative frequencies – Fs /2 < f< 0 correspond to N/2+1 < n < N-1.

6) The magnitude frequency response of each frame: is multiplied by 20 numbers of triangular band pass filters to obtain the corresponding log energy of each triangular filter. The mid-point of each filter represents the Mel frequency. The Mel frequency (Mel(f)) is computed from the linear frequency as:

$$Mel(f) = 2595 * \log_{10}[1 + \frac{f}{700}] \qquad (9)$$

Where (f) is the actual frequency in Hz.

7) Discrete Cosine Transform (DCT): is applied on the 20-log energy (Sk) obtained from the triangular filters to have L mel-scale cepstral coefficients. It is obtained by the following equation:

$$c_n = \sum_{k=1}^{K} (\log S_k) \cos\left[n\left(k - \frac{1}{2}\right)\frac{\pi}{k}\right] \quad n = 1, , , L \qquad (10)$$

Where (K) is the number of the subbands and (L) is the desired length of the cepstrum

The set of coefficients is called acoustic vector. Therefore, each input utterance is transformed into a sequence of acoustic vector. The statistical features mean (Mm), standard division (Msd), median (Mmd), skewness (Msk) and kurtosis (Mku) are calculated from acoustic vector. Consequently, five proposed features are extracted from signal MFCC. These features are:

F = {Mm, Msd, Mmd, Msk, Mku} (11)

8) K-NN: is one of the most popular and simple supervised learning algorithms. It is used in various pattern recognition and classification problems [29-31]. KNN is classified based on the majority of the K-nearest neighbour category. The distance between the query sample and all training samples is computed. Then, it is sorted, and the closest neighbours are determined based on the minimum distance Kth. The category Y of the nearest neighbours is collected. The majority category of the nearest neighbours is used to predict the value of the query sample. The Euclidean is the most widely used distance function with K-NN, as follows [32-34]:

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (12)$$

Where (x) features vector of dimension n for the query sample and (y) features vector of dimension n for training samples.

The following section illustrates the implementation phase of the proposed system.

4. Implementation Phase

MATLAB R2024a was utilized to implement the proposed system with Guide. Several important toolboxes are installed in MATLAB to support computer vision. The Computer Vision Toolbox provides algorithms and functions for object detection, feature extraction, and image analysis, which are essential for processing visual data. The image acquisition toolbox allows cameras and other imaging devices directly to MATLAB for real-time image and video capture. Additionally, the image acquisition toolbox support package for OS generic video interface makes it possible to acquire video from standard operating system devices, ensuring compatibility with a wide range of hardware. These tools together give a strong environment for implementing and testing the proposed system. Additionally, conditional logic was utilized to classify performance levels based on total scores. Pseudocode of matching process for student image and sound is illustrated in Alogrithm 1 and Alogrithm 2.

Algorithm 1: Pseudocode for Student Image Matching

```
Begin
    Set feature_vector = CREATE_ARRAY(1, 4)
    Set feature_vector[1] = max_cont
    Set feature_vector[2] = max_en
    Set feature_vector[3] = max_corr
    Set feature_vector[4] = max_hom
    Load_Database('image_db.mat')
    Set similarity_scores = CREATE_ARRAY(50, 1)
    Set total_sum = 0
    Set weights = [1.0, 0.7, 0.5, 0.3]
    For i = 1 To number_of_image do
        Set contrast_diff = (GLDB[i].cont - feature_vector[1])^2
        Set energy_diff = (GLDB[i].en - feature_vector[2])^2
        Set correlation_diff = (GLDB[i].corr - feature_vector[3])^2
        Set homogeneity_diff = (GLDB[i].hom - feature_vector[4])^2
```

```
            Set similarity_scores[i] = SQRT(weights[1]*contrast_diff + weights[2]*energy_diff +
            weights[3]*correlation_diff + weights[4]*homogeneity_diff)
            Set total_sum = total_sum + similarity_scores[i]
        End
    Set [sorted_scores, sorted_indices] = SORT(similarity_scores)
    Set verification_score = 1 - (sorted_scores[1] * 10 / total_sum)
        If verification_score >= 0.90 then
            Set_Text(handles.r1, 'Verified')
        Else
            Set_Text(handles.r1, 'Not verified')
        End
    End
```

Algorithm 2: Pseudocode for Student Sound Matching

```
Begin
    Set k = 4
    Set query_features= [CONVERT_TO_DOUBLE(mfccmean),CONVERT_TO_DOUBLE(mfccstd),
    CONVERT_TO_DOUBLE(mfccsmed),CONVERT_TO_DOUBLE(mfccssk),
    CONVERT_TO_DOUBLE(mfccsku)]
    Set database_features = LOAD_FEATURE_DATABASE('sound_db.xlsx')
    Set x1 = database_features.mean , Set x2 = database_features.std , Set x3 = database_features.median ,
    Set x4 = database_features.skewness , Set x5 = database_features.kurtosis , Set IDs = database_features.ID
    Set names = database_features.name, Setlevels = database_features.level
    Set departments= database_features.department
    Set classes = database_features.class
    For i = 1 To LENGTH(database_features) do
        Set squared_differences = [(query_features[1] - x1[i])^2, (query_features[2] - x2[i])^2,
```

```
        (query_features[3] - x3[i])^2, (query_features[4] - x4[i])^2,(query_features[5] - x5[i])^2]
        Set euclidean_distance[i] = SQUARE_ROOT(SUM(squared_differences))
        End
    Set total_distance = SUM(euclidean_distance)
For i = 1 To LENGTH(euclidean_distance) do
        Set similarity_scores[i] = 1 - (euclidean_distance[i] * 10 / total_distance)
        End
Set combined_data = COMBINE_COLUMNS(euclidean_distance, IDs, classes, names, similarity_scores)
Set sorted_results = SORT_ROWS(combined_data, BY_COLUMN: 1)
Set top_k_classes = EXTRACT_COLUMN(sorted_results[1:k], COLUMN: 3)
Set predicted_class = CALCULATE_MODE(top_k_classes)
    Switch (predicted_class)
            Case 1: Set name = student_name, Set department = student_department
             Set level = student_level, Set ID = student_id
            Case 2: Set name = student_name, Set department = student_department
             Set level = student_level, Set ID = student_id
            // ... Additional cases 3 through 99
            Case 100: Set name = student_name, Set department = student_department,
             Set level = student_level, Set ID = student_id
    End
Set display_data = EXTRACT_COLUMNS (sorted_results, COLUMNS: [2, 3, 5])
Set _Table_Data('table2', display_data)
Set _Table_Properties('table2', COLUMN_WIDTH: 100, HEADERS: ['Index', 'Class', 'Similarity'])
SHOW_PANEL('panel1'), Set_Text (handle2, name), Set_Text(handle3, department)
Set_Text(handle4, level), Set _Text(handle5, ID), Set sc2= name
End
```

The following section illustrates the experimental work of the proposed system.

5. Experimental Work

The proposed system was applied to a sample of fourth-year students at the Faculty of Specific Education, Mansoura University, in Software Engineering course. The sample of proposed system screens is illustrated in figures from Figure 3 to Figure 8.

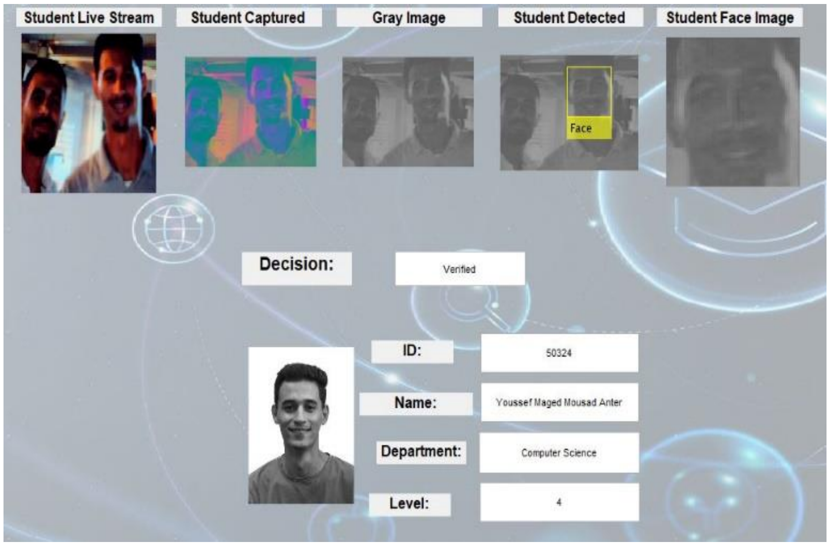

Figure 3. Sample of Student Image Detection

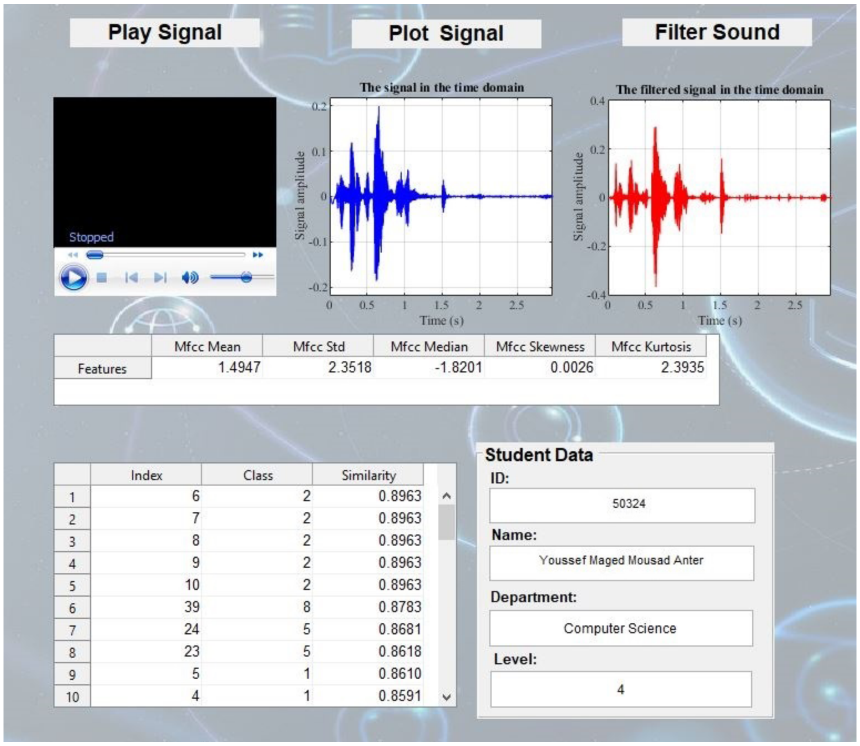

Figure 4. Sample of Student Sound Detection

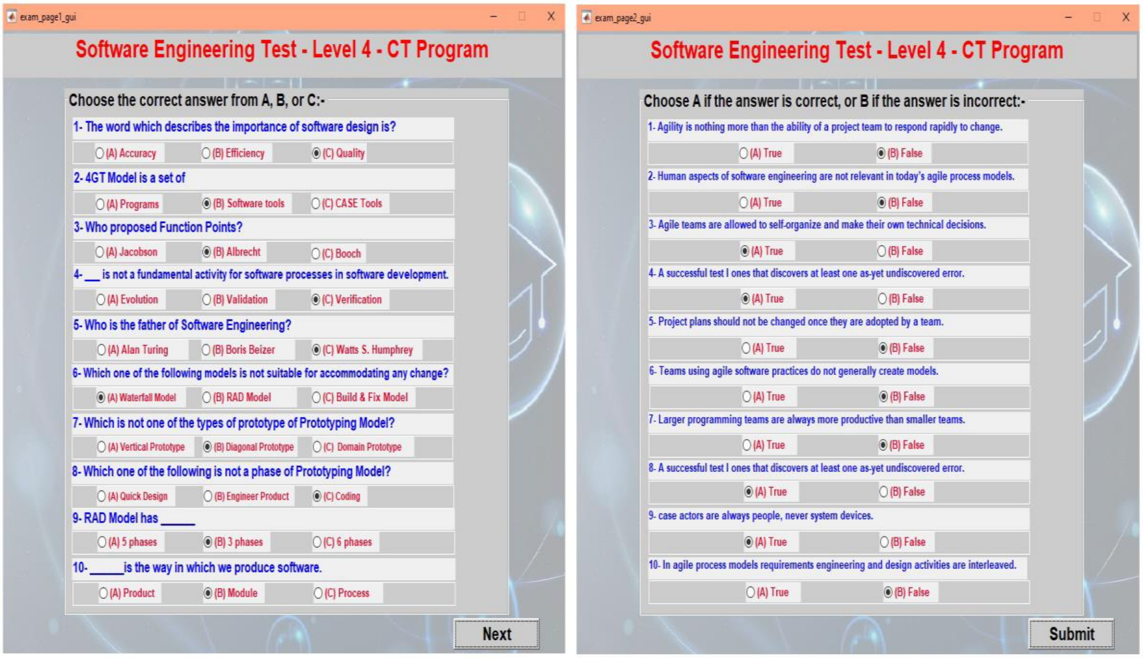

Figure 5. Sample of Student Test

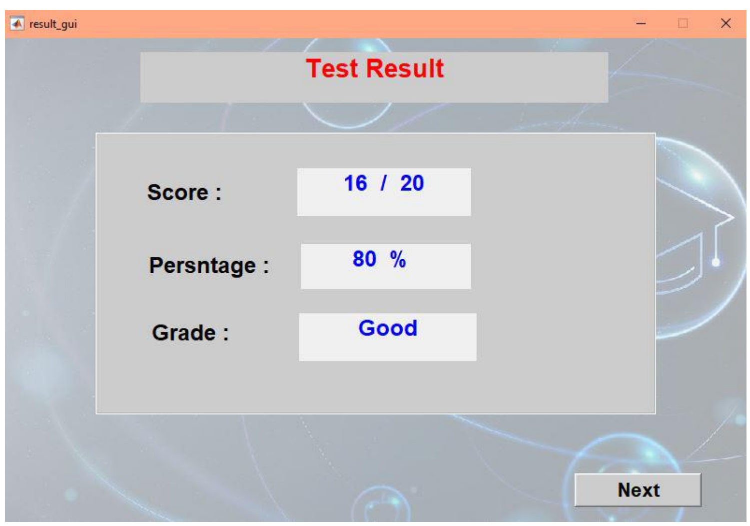

Figure 6. Sample of Test Result

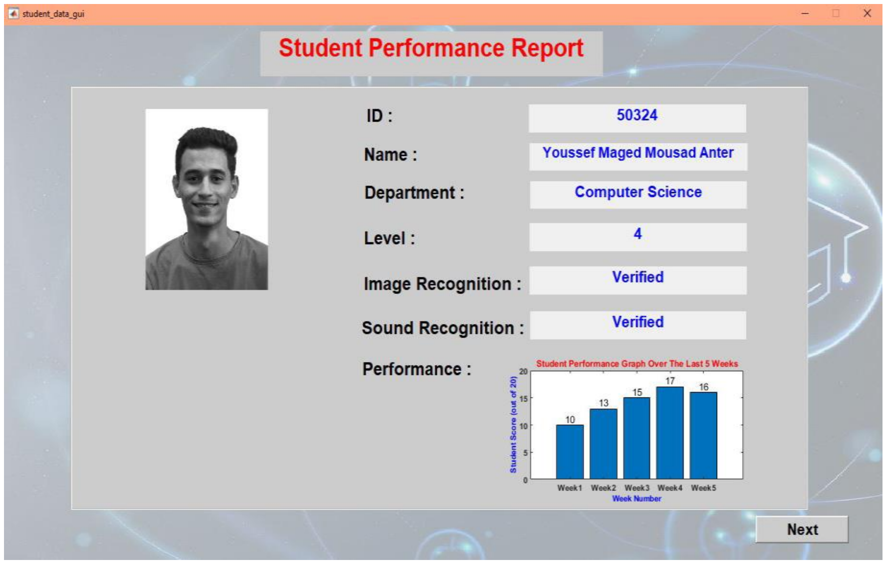

Figure 7. Sample of Student Performance Report

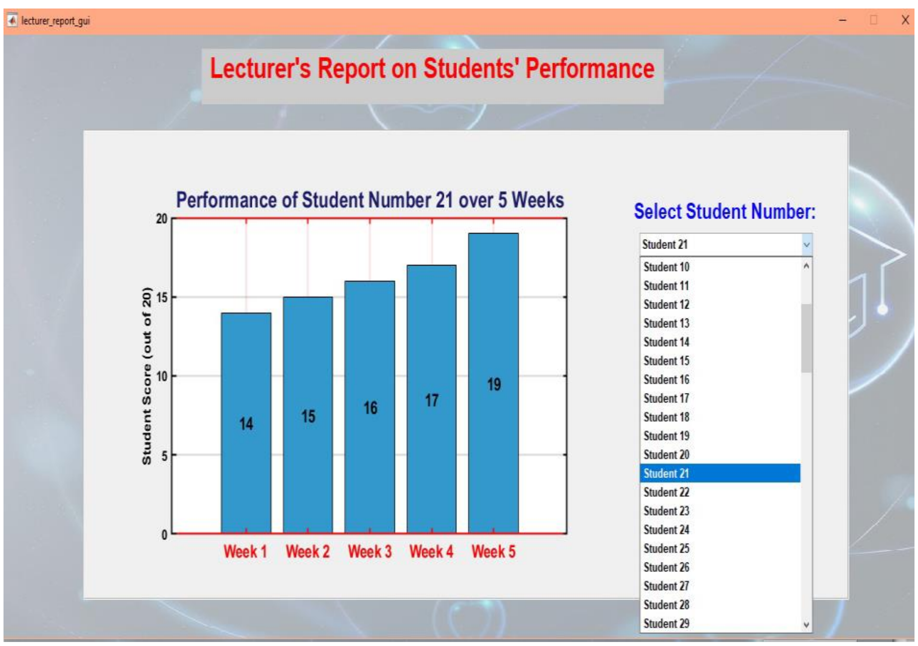

Figure 8. Sample of Lecturer Report on Student Performance

The following section illustrates the results.

6. Results

The results are explained as follows:

A. Evaluation of the Image Recognition Part

Three metrics are used for evaluation purposes. Precision (P) calculates the ration of relevant images retrieved to the total number of images retrieved (relevant and non-relevant images). Recall (R) calculates the ratio of no. of relevant images retrieved to the total number of relevant images in Database. F-Measure (F-Score) calculates by computing the Mean of precision and recall. These metrics equations are as follows [35]:

$$P = \frac{TP}{TP + FP} = \frac{\text{True Positive}}{\text{True Positive + False Positive}} \quad (13)$$

$$R = \frac{(TP)}{(TP) + (FN)} = \frac{\text{True Positive}}{\text{True Positive + False Positive}} \quad (14)$$

$$F - \text{Score} = 2 \times \frac{2(P\times R)}{(P+R)} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (15)$$

Fig. 9, Fig. 10 and Fig. 11 show precision, recall and f-measure curves of student images retrieval.

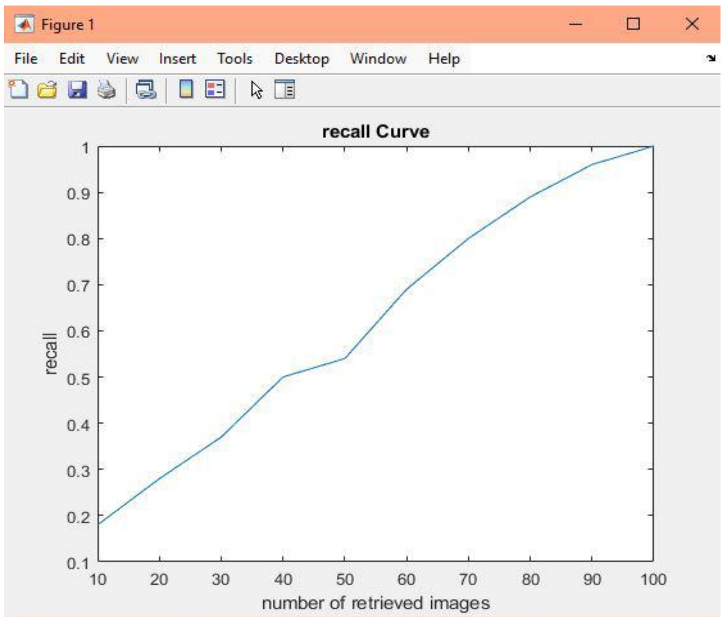

Figure 9. Recall curve

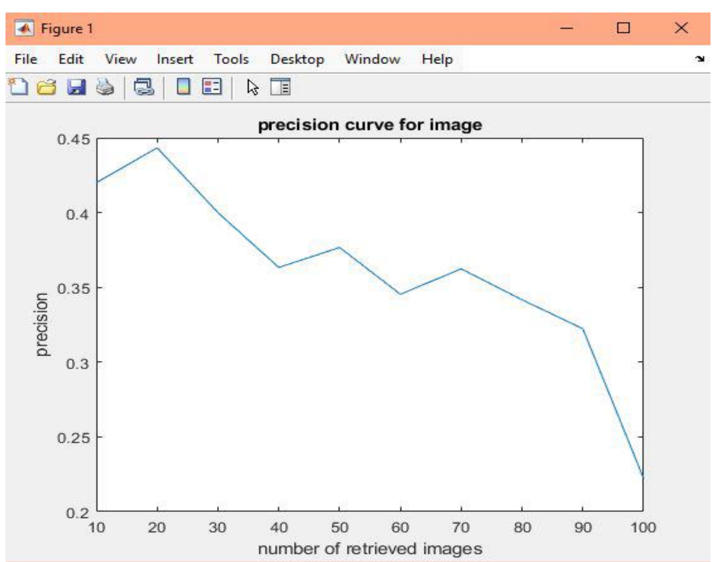

Figure 10. Precision curve

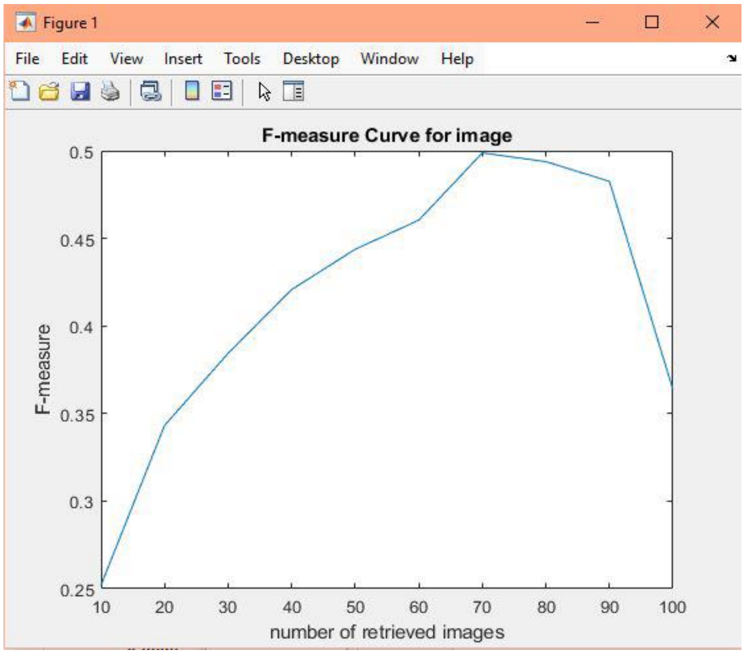

Figure 11. F_measure curve

Table 1 shows the accuracy and the average accuracy of the students' image recognition, where number of trials is equal to 10:

Table 1: The accuracy and the average accuracy of the students' image recognition

Image number	**Number of good trials**	**Accuracy (%)**
1	9	90
2	8	80
3	10	100
4	10	100
5	8	80
6	9	90
7	7	70
8	10	100
9	8	80
10	10	100
11	9	90
12	8	80
. . .		
98	10	100
99	9	90
100	9	90
Average accuracy		**89**

The accuracy formula is represented by the following equation:

$$\text{Accuracy} = \frac{\text{Number of good trials}}{\text{Total Number of trials}} \times 100 \qquad (16)$$

B. Evaluation of the Sound Recognition Part

Table 2 shows the accuracy and the average accuracy of the students' sound recognition, where number of trials is equal to 10:

Table 2: the accuracy and the average accuracy of the students' sound recognition

Sound number	Number of good trials	Accuracy (%)
1	9	90
2	10	100
3	8	80
4	10	100
5	9	90
6	10	100
7	10	100
8	9	90
9	8	80
10	10	100
11	10	100
12	10	100
. . .		
98	10	100
99	10	100
100	9	90
Average accuracy		**90**

C. Evaluation of the Proposed System (Image & Sound)

Table 3 shows the ability of the proposed system to identify students from their images and sounds, thus making this system efficient and reliable.

Table 3: The ability of the proposed system to identify students from their images and sounds

Student Number	Image Detected	Sound Detected	Final Decision
1	True	True	True
2	True	False	False
3	True	True	True
4	True	True	True
5	True	True	True
6	True	True	True
7	True	True	True
8	True	True	True
9	False	True	False
10	True	True	True
11	True	True	True

Student Number	Image Detected	Sound Detected	Final Decision
1	True	True	True
12	True	True	True
. . .			
98	True	True	True
99	True	False	False
100	True	True	True

The following section illustrates the discussion of this study.

7. Discussion

Monitoring students' performance during real time lectures plays a vital role in enhancing the quality of education. In this study, a model that combines image processing and sound processing was introduced. This model enables a trustworthy assessment of student engagement and participation. Using the GLCM method and Viola-Jones detector weighted Euclidean distance method for image processing and the MFCC method combined with the KNN technique for sound processing, the model combined the visual and auditory components of the method and efficiently monitored performance accurately. Indicatemental results proved the efficiency and effectiveness of this study's method and showed that the suggested approach can be applied in real-time learning settings when the average accuracy of student image and sound recognition reaches 89% and 90%, respectively.

References

[1] M. Aly, "Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model," *Multimedia Tools and Applications*, vol. 84, no. 13, pp. 12575–12614, 2025.

[2] J. O. Adeyemi, S. O. Ogunlere, and B. G. Akwaronwu, "Real-time detection of examination malpractices using convolutional neural networks and video surveillance: A systematic review with meta-analysis," *British Journal of Computer, Networking and Information Technology*, vol. 8, no. 2, pp. 15–50, 2025.

[3] Alnasyan, M. Basheri, M. Alassafi, and K. Alnasyan, "Kanformer: An attention-enhanced deep learning model for predicting student performance in virtual learning environments," *Social Network Analysis and Mining*, vol. 15, no. 25, pp. 1–21, 2025.

[4] X. Song, "Emotional recognition and feedback of students in English e-learning based on computer vision and face recognition algorithms," *Entertainment Computing*, vol. 103, no. 13, pp. 4860–4874, Jul. 2025.

[5] J. A. Conde and J. I. Teleron, "Student performance indicator monitoring systems through system development life cycle (SDLC) method," *International Journal of Advanced Research in Arts, Science, Engineering & Management*, vol. 12, no. 1, pp. 274–285, 2025.

[6] Adewumi and M. Grace, "Development of a student monitoring attendance for improved learning using UI/UX design," *Journal of Liaoning Technical University (Natural Science Edition)*, vol. 19, no. 3, pp. 148–163, 2025.

[7] K. Gomathy, A. T. Neethi, and S. S. Krishna, "Automating student performance analysis using machine learning," *Singaporean Journal of Scientific Research*, vol. 17, no. 1, pp. 62–70, 2025.

[8] M. Matus, J. Bađari, and I. Balaban, "Using LMS records to track student performance: A case of a blended course," in *Proc. 17th Int. Conf. Computer Supported Education (CSEDU 2025)*, vol. 1, 2025, pp. 283–290.

[9] Dey, A. Anand, S. Samanta, B. K. Sah, and S. Biswas, "Attention-based AdaptSpecX network for effective student action recognition in online learning," *Procedia Computer Science*, vol. 233, pp. 164–174, 2024.

[10] K. Kommineni and K. Kumar, "Machine learning approach for face recognition-based attendance system," *Materials Science and Technology*, vol. 23, no. 2, pp. 65–69, Feb. 2024.

[11] M. A. Sulaiman, "Development of an electronic examination platform using face recognition methods," *Science Journal of University of Zakho*, vol. 12, no. 3, pp. 308–315, 2024.

[12] N. M. Alruwais and M. Zakariah, "Student recognition and activity monitoring in e-classes using deep learning in higher education," *IEEE Access*, vol. 12, pp. 66110–66128, 2024.

[13] S. H. Seo, M. Y. Kim, and Y. Kim, "Design and implementation of an online video lecture system based on facial expression recognition," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 14, no. 3, pp. 866–872, 2024.

[14] V. U. Pinjarkar, U. S. Pinjarkar, H. N. Bhor, Y. V. Mahajan, V. R. Patil, S. D. Rajput, *et al*., "Student engagement monitoring in online learning environment," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 1, pp. 292–298, 2024.

[15] Y. Lou and F. Li, "Design of an online education student learning status evaluation model based on dual-improved neural networks," *Intelligent Systems with Applications*, vol. 22, pp. 1–13, 2024.

[16] P. R. Nababan, T. Haryanto, and S. H. Wijaya, "Classification of coral images using support vector machine with gray-level co-occurrence matrix feature extraction," *JOIV: International Journal on Informatics Visualization*, vol. 9, no. 3, pp. 936–946, 2025.

[17] H. H. Fauziyah, D. R. Ningtias, B. Wahyudi, and J. N. D. Simanjuntak, "Identification of lung cancer using gray-level co-occurrence matrix and artificial neural network with backpropagation algorithm," *Journal of Soft Computing Exploration*, vol. 6, no. 1, pp. 51–61, 2025.

[18] T. V. Reddy, S. Kumar, V. S. Reddy, T. L. Kayathri, R. Suresh, and D. Srikar, "Harnessing Viola–Jones for effective real-time crowd monitoring based on image processing techniques," in *Proc. Int. Conf. Futuristic Technology (INCOFT 2025)*, vol. 3, pp. 368–376, 2025.

[19] C.-H. Choi, J. Han, H. W. Oh, J. Cha, and J. Shin, "EOS: Edge-based operation skip scheme for real-time object detection using Viola–Jones classifier," *Electronics*, vol. 14, no. 2, pp. 1–22, 2025.

[20] E. E. ElAlfi, A. E. Shalaby, M. E. E. ElAlmi, and S. A. Shaban, "Novel advisory system for the psychological guidance of university students," *International Journal of Computer Science Trends and Technology*, vol. 6, no. 3, pp. 219–237, 2018.

[21] N. G. Ramadhan, W. Maharani, A. A. Gozali, and Adiwijaya, "Enhancing SMOTE using Euclidean weighting for imbalanced classification datasets," *Journal of Applied Data Sciences*, vol. 6, no. 3, pp. 2207–2220, 2025.

[22] M. Dahria, S. Defit, and Yuhandri, "Development of Euclidean distance algorithm for ANFIS optimization in IoT-based pond water quality prediction," *Journal of Robotics and Control*, vol. 6, no. 4, pp. 1777–1789, 2025.

[23] P. Sari, A. Fadlil, and T. Sutikno, "Islamic sound recognition using MFCC and SVM: Case study on takbir and sholawat," *Journal of Computer Networks, Architecture and High Performance Computing*, vol. 7, no. 3, pp. 825–839, 2025.

[24] Haryaveda, Almeranda, N. Zahron, and R. Kurniawan, "Bird sound quality analysis for chirping masters using MFCC and SVM classification algorithm," *Journal of Computer Science, Information Technology and Telecommunication Engineering*, vol. 6, no. 2, pp. 978–987, 2025.

[25] S. Sultoni, B. Darmawan, and Supriono, "Speaker recognition system using MFCC and HMM methods," *International Journal of Informatics and Computation*, vol. 7, no. 1, pp. 206–218, 2025.

[26] S. Joysingh, S. Johanan, P. Vijayalakshmi, and T. Nagarajan, "Significance of chirp MFCC as a feature in speech and audio applications," *Computer Speech & Language*, vol. 89, pp. 1–11, 2025.

[27] Y. Yan, L. van Bemmel, F. M. E. Franssen, S. O. Simons, and V. Urovi, "Developing a multi-feature fusion model for exacerbation classification in asthma and COPD," *Computer Methods and Programs in Biomedicine*, vol. 268, pp. 1–12, 2025.

[28] Subhash, Darshana, B. Premjith, and V. Ravi, "A robust accent classification system based on variational mode decomposition," *Engineering Applications of Artificial Intelligence*, vol. 139, pp. 1–20, 2025.

[29] Kasthuri, A. Suruliandi, E. Poongothai, and S. P. Raja, "Evaluating feature extraction and classification techniques: A comparative approach to face annotation," *Engineering Applications of Artificial Intelligence*, vol. 162, pp. 1–23, 2025.

[30] Guo, Y. Zhang, S. Zhang, and W. Xiao, "Classification model for blast furnace status based on multi-source information," *Engineering Applications of Artificial Intelligence*, vol. 141, pp. 1–10, 2025.

[31] M. I. H. Siddiqui, A. H. Sakib, S. Akter, J. Debnath, and M. R. Mahmud, “Comparative analysis of traditional machine learning vs deep learning for sleep stage classification,” *International Journal of Science and Research Archive*, vol. 15, pp. 1778–1789, 2025.

[32] Kannan, M. J. Saikia, S. Kumar, and S. Datta, “Detection of valvular heart diseases from PCG signals using machine and deep learning models: A review,” *IEEE Access*, vol. 13, pp. 110344–110359, 2025.

[33] H. A. A. Al-Khamees, N. S. Sani, A. S. Gifal, L. X. W. Liu, and M. I. Esa, “A dynamic model using k-NN algorithm for predicting diabetes and breast cancer,” *Computers in Biology and Medicine*, vol. 192, pp. 1–14, 2025.

[34] S. Andriyanto, S. Muharni, and Sulistiyanto, “Analisis data sintetis polycystic ovary syndrome menggunakan algoritma Naïve Bayes dan k-NN,” *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 1, pp. 144–149, 2025.

[35] S. Beddar-Wiesing, A. Moallemy-Oureh, M. Kempkes, and J. M. Thomas, “Absolute evaluation measures for machine learning: A survey,” *arXiv preprint*, arXiv: 2507.03392, 2025.