Engagement And Assessment Optimization System For Enhancing Learning Experiences In Online Education.

Premathilaka S.P.D.M
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21185298 @my.sliit.lk

Ekanayaka H.E.M.P.L
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21185502@my.sliit.lk

Wijenandana S.D
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21158254@my.sliit.lk

Karunarathana J.H.H.N
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21157950@my.sliit.lk

Wishalya Tissera
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
wishalya.t@sliit.lk

Chathushki Chathumali
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
chathushki.c@sliit.lk

***Abstract*— The rapid growth of e-learning brings challenges in student engagement, academic integrity, and interactive learning. Traditional methods do not work effectively in online environments, necessitating artificial intelligence-based solutions. In this study, we leverage convolutional neural networks and long short-term memory networks for real-time attention and facial expression recognition to enable early intervention. To ensure fairness, object detection and head pose estimation are used to monitor students for prohibited activities during examinations. A natural language processing-based grading system automatically grades assignments, improving efficiency, and a deep learning-based quiz generator creates relevant questions from lecture content. Deployed on diverse datasets, these artificial intelligence models achieve high accuracy in engagement analysis, cheating detection, and automated learning support. The findings highlight the ability of artificial intelligence to generate an engaging, fair, and efficient e-learning ecosystem.**

***Keywords*—E-learning, Facial Expression Analysis, Object Detection, Academic Integrity, Automated Grading, Quiz Generation.**

I. INTRODUCTION

The expansion of e-learning websites has revolutionized learning through increased accessibility and flexibility for students. Yet, it presents challenges related to student engagement, academic honesty, and interactive learning experiences. Online learning differs from conventional classrooms in that it lacks instant feedback systems that quantify student engagement and enable prompt intervention where necessary. Furthermore, academic honesty is a concern during online testing since conventional monitoring practices are ineffective in identifying incidences of cheating.

To resolve these problems, this work uses AI techniques to enhance the efficacy of e-learning. We integrate facial expression analysis and attention estimation with CNNs and LSTMs to approximate student engagement in real-time. Object detection and head pose estimation models discourage cheating by identifying such behavior and outside materials. An NLP-driven grading system grades assignments automatically for greater efficiency. An AI-driven quiz generator also creates adaptive tests from lecture notes, offering dynamic tests.

This study aims to create a more interactive, fair, and smart e-learning system by incorporating artificial intelligence solutions. It integrates machine learning and behavioral analysis to improve student monitoring, exam integrity, and content creation, thus creating a more efficient and engaging digital learning process.

II. LITERATURE REVIEW

The rise of e-learning platforms brings challenges in student engagement, academic integrity, and interactive learning [1]. This review examines advancements and challenges in automated grading, quiz generation, attention detection, and cheating prevention.

Automated grading has evolved significantly, moving from basic multiple-choice evaluations [2] to sophisticated NLP and ML-based systems for open-ended responses [3]. Early system [4], [5] used linguistic properties for scoring but lacked semantic understanding. ML, especially deep learning models, has improved accuracy [6]. Transformer-based models like BERT and GPT offer enhanced contextual understanding [7]. OCR and image recognition are also expanding automated grading [8]. However, addressing bias and ensuring fairness is vital [9], necessitating transparency (XAI) [10]. Automated grading offers immediate feedback and personalized learning [11], but validity, reliability, and integration remain challenges [12]. Automated quiz generation is advancing through adaptive learning, AI-driven assessments, and personalized environments [13]. Research focuses on creating semantically accurate questions, particularly multiple-choice, with relevant distractors [14]. AI-driven assessments offer immediate feedback [15], but integration with learning analytics needs further exploration [16]. Creating high-quality questions that promote higher-order thinking remains a significant challenge [17]. Facial expression and attention detection offer insights into student engagement. Visual cues are crucial for attention detection [18], and real-time monitoring can enable timely interventions [19]. Deep learning predicts engagement patterns [20], enabling personalized support. Preventing cheating in online exams is critical. Object detection using CNN frameworks like R-CNN and YOLO [21] can identify unauthorized devices. Head pose estimation, using facial landmark detection [22], can monitor focus. Integrated systems combine these techniques, but computational demands, privacy concerns, and false positives persist. In conclusion, while technology enhances e-learning effectiveness and integrity, ongoing research is needed to address remaining challenges and ensure equitable online learning experiences.

III. METHODOLOGY

This research presents Edusmart, an innovative dual-module system that aims to enhance learning and examination processes. The Exam Module analyzes lecture content and employs machine learning algorithms to automate the generation of quizzes on particular subjects. Besides, this module incorporates an unauthorized object detection procedure coupled with head pose estimation technology to monitor students during examinations, thereby identifying any possible instances of academic dishonesty to uphold the integrity of the examination process. The Learning Module specializes in facial expression analysis and attention tracking in virtual learning environments in real time. It gauges the engagement of students during lectures and includes an automatic grading system for assignments and work submitted, thereby enhancing the efficiency of the assessment process. Figure 1 illustrates a comprehensive diagram of the overall system, which is titled "Overall system diagram of the proposed solution." By bringing together all of these various modules, Edusmart provides a unified educational support system that delivers actionable intelligence regarding student involvement and performance while, simultaneously, enhancing learning efficacy and ensuring academic honesty.

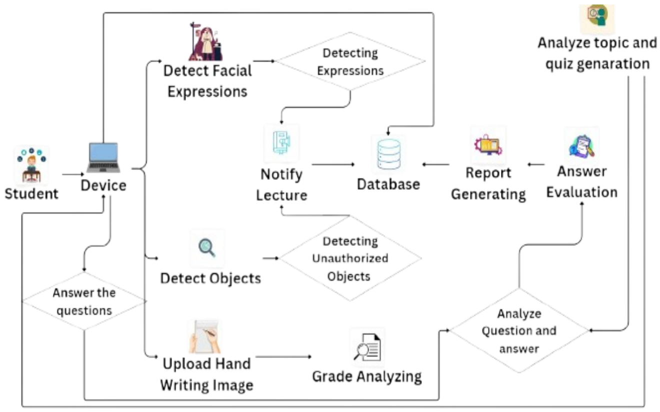

Figure 1. Overall system diagram of the proposed solution

A. Facial Expressions and Attention Detection

To successfully record facial expressions and attentiveness, we developed a multi-step process. Figure 2 is a graphical representation of the process showing data flow as well as significant steps involved:

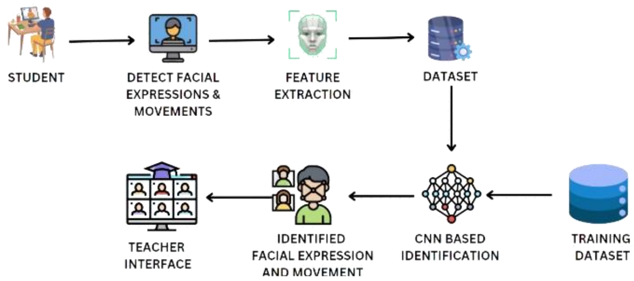

Figure 2.Facial Expressions and Attention Detection

a) Data Collection

Images will be collected from various sources, including available labeled facial expression datasets such as CK+ and JAFFE, and direct collection efforts using cameras or webcams. Participants for the direct collection effort will be recruited from such places as university campuses or community gatherings to ensure a varied demographic sample. Ethical standards, such as the privacy of individuals and data security, will be rigorously maintained during the research process to ensure adherence to ethical stipulations and protect participant data.

b) Data Preprocessing

To enhance the facial expression recognition model, the images were resized to 48x48 and were converted to grayscale to reduce computational complexity. The pixel values were also normalized between 0 and 1 for better model convergence. The emotion labels were encoded using LabelEncoder and were converted to categorical form to be used with the cross-entropy loss function. The data was split into train, validation, and test sets using train_test_split (test_size = 0.1, random_state = 42) for reproducibility and stratified splitting of the data. It is necessary to detect legal guardians from images before model training.

c) Model Training

The CNN model for facial expression classification was trained for classifying images into seven categories of emotions. Data augmentation was done using Keras's ImageDataGenerator with the following transformations: rotation (15°), shifts (0.15), shear (0.15), zoom (0.15), and flip (horizontal). The model used categorical cross-entropy loss, the Adam optimizer (learning rate: 0.001), and accuracy as the evaluation metric. Training was carried out for 100 epochs (batch size: 32), with Early Stopping (patience: 11 epochs) and ReduceLROnPlateau (factor: 0.5, patience: 7 epochs) to prevent overfitting. The model with the best validation accuracy was saved for future use.

B. Automated Quiz Generation and Evaluation

The illustrated automatic quiz system, as shown in Figure 3, utilizes NLP and transformer models for automating quiz preparation and evaluation. The system reads lecture content to generate real-time, adaptive questions administered to students through a user interface. The system provides immediate feedback, maintains confidentiality, and adapts to classroom settings. Answer evaluation and report generation are then performed, with all data saved in a database to continue optimizing the question generation process, thus reducing the workload of the instructor.

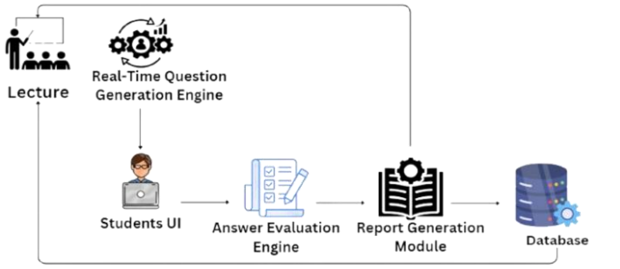

Figure 3. Flow of Automated Quiz Generation and Evaluation

a) Features and Target Variable

Predictive data for generating quizzes was collected at various difficulty levels, with integrity sustained through varied instructional materials. The dataset included questions and answers, with categorical variables numericized and uninformative columns removed. Contextual question generation formed a significant portion of the data collection.

b) Data Preprocessing

Data pretreatment played a key role in preparing training materials for the successful training of models. Lecture notes and PDF documents, which had undergone preprocessing to remove noise, served as significant sources of data. The primary processing tasks included text extraction, in which relevant data were extracted for training, and cleaning and normalization to structure the text in a proper manner and get it ready for subsequent processing.

c) Model Training

Automated Quiz Creation and Assessment Using Pseudocode

Input: Educational content (e.g., lecture notes, textbooks)
Output: Generated quizzes and evaluations
BEGIN

1. Import necessary libraries (transformers, datasets, pandas, sklearn, PyPDF2)
2. Load the pre-trained question generation model (e.g., T5).
3. Tokenize the educational content using the pre-trained tokenizer.
4. Fine-tune the question generation model.
5. Prepare input sequences for the model.
6. Train the model to generate questions based on the input context.
7. Evaluate generated questions for relevance and grammatical correctness.
8. Implement an evaluation component.
9. Take the generated question, the reference content, and student response as input.
10. Evaluate student response against the correct answer.

END

To establish a baseline for question-response tasks, the system fine-tunes a T5-small model on the SQuAD dataset. Additional adaptations include PDF extraction and text chunking for context management. It trains the model with the prefix "create question:" to generate pertinent questions. After training, it verifies student responses for correctness and stores the optimized model for deployment.

C. Unauthorized Object Detection, Head Pose Estimation and User Verification

The system is built to incorporate automatic detection of unauthorized objects, head pose estimation, and user authentication for online exams. Figure 4 demonstrates this mechanism, displaying the flow of the unauthorized object detection system. With the use of computer vision techniques and deep neural networks, it ensures exam integrity by detecting banned objects in real-time, monitoring examinees constantly, and authenticating user identities.

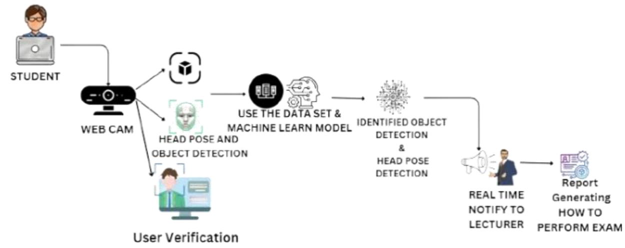

Figure 4. Flow of Unauthorized Object Detection

a) Data Collecting and Processing

The data set included publicly available datasets and synthetic examination videos with face landmarks, head poses, and object detections. Forbidding object data were generated through YOLOv5 and manually annotated images, while head pose estimation involved MediaPipe Face Mesh and the PnP algorithm. User verification involved multi-pose facial datasets. Images were resized, and data augmentation techniques like flipping and brightness adjustment improved generalization. Background noise was eliminated, and bounding boxes were added for object detection. Numerical features were normalized by MinMaxScaler, and categorical features were transformed by one-hot encoding. The data were divided into 80% for training and 20% for testing to facilitate model evaluation.

b) Model Development

The system has three main modules: Unauthorized Object Detection, which uses YOLOv5 to detect unauthorized objects in real time; Head Pose Estimation, which uses MediaPipe Face Mesh and the PnP algorithm to detect abnormal head movements like screen avoidance; and User Authentication, which uses deep learning-based face recognition for user authentication in real time. Tuned with hyperparameters using GridSearchCV, these machine learning models improve accuracy, eliminate false positives, and provide a solid online exam surveillance system that mitigates cheating possibilities and maintains the integrity of the exams.

D. Automated grading system

a) Data Collection

For the purpose of precise evaluation, various forms of data were gathered for training and testing of the automated grading system. The Student Responses Dataset consisted of short and long open-ended answers derived from publicly accessible education datasets such as Kaggle and

OpenEdx, and manually annotated answers. Additionally, a Handwritten Answer Sheets Dataset was created to provide grading support for handwritten answers with OCR datasets for training. The dataset was developed with the objective of minimizing grading bias by including responses of students with varying writing styles, languages, and academic backgrounds, as well as input from various technical and non-technical fields, thus enhancing the generalizability of the system.

b) Data Preprocessing

Preprocessing was carried out for hand-written and typed answers for training. Text answers were tokenized using the BERT tokenizer, stop words and special characters deleted for efficiency. Feature extraction utilized BERT and TF-IDF vectorization for extracting contextual meaning, sentence complexity, and length being analyzed for grading. Human-graded answers served as ground truth for validation. Hand-written answers were processed via Tesseract OCR and Google Vision API, scanned text being converted

c) Model Training

Machine Learning approaches and Natural Language Processing models were leveraged to develop an automatic grading system. Transformer-based models like BERT, GPT-3, and RoBERTa were fine-tuned over labeled essay datasets to study textual coherence, relevance, and correctness, and BiLSTM and GRU were explored for further contextualized comprehension. Random Forests and Support Vector Machines (SVMs) were employed in a combination to predict scores from the features extracted. Supervised learning was used to train with ground truth labels, and hyperparameter tuning was performed using GridSearchCV, including learning rate optimization, batch sizes, and tokenization techniques to enhance model performance.

IV. RESULTS AND DISCUSSION

A. Facial Expressions and Attention Detection

The facial expression recognition model had an overall classification accuracy of 78% on the independent test dataset, demonstrating a proficient capability in classifying the seven detected emotions (Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Neutral The classification report Figure 5 shows how the performance of the model fluctuated among various emotions. In addition, a weighted average F1-score of 0.78, taking into account the class imbalance, further reinforces the accuracy attained by the model.

```
              precision    recall  f1-score   support

           0       0.78      0.80      0.79       400  (Anger)
           1       0.65      0.60      0.62       250  (Disgust)
           2       0.72      0.75      0.73       300  (Fear)
           3       0.85      0.88      0.86       450  (Happiness)
           4       0.70      0.68      0.69       350  (Sadness)
           5       0.80      0.77      0.78       200  (Surprise)
           6       0.88      0.90      0.89       500  (Neutral)

    accuracy                           0.78      2450
   macro avg       0.76      0.77      0.76      2450
weighted avg       0.78      0.78      0.78      2450
```

Figure 5.Emotion Recognition classification Report

The face expression recognition model had variable performance across emotion classes, with "Neutral" and "Happiness" achieving high F1-scores (0.89 and 0.86, respectively) via accurate identification and full coverage of instances, likely helped by distinctive facial muscle configurations "Anger" and "Surprise" were reasonably successful (F1-scores of 0.79 and 0.78) with good precision-recall balance. Conversely, the emotions "Disgust," "Fear," and "Sadness" exhibited diminished F1-scores of 0.62, 0.73, and 0.69, respectively, which suggests difficulties in the precise classification of these emotional states, with "Disgust" reflecting the lowest F1-score, potentially underscoring a particular deficiency in its identification.

Figure 6 is a normalized confusion matrix to demonstrate the emotion recognition model's accuracy. Although the model tends to have good prediction accuracy overall via large diagonal elements, it does exhibit instances of confusion between some emotions. Particularly, "Disgust" and "Anger" are confused with each other, as well as "Fear," "Surprise," and "Sadness." "Sadness" and "Anger" are also frequently labeled as "Neutral" incorrectly. These confusions are caused by shared facial features, subtle differences in expressions, biases in datasets, and limitations like low image resolution. The model worked quite well for "Neutral" and "Happiness" but should be improved with more balanced datasets, deeper networks, higher resolution images, and contextual information to recognize better "Disgust," "Fear," and "Sadness".

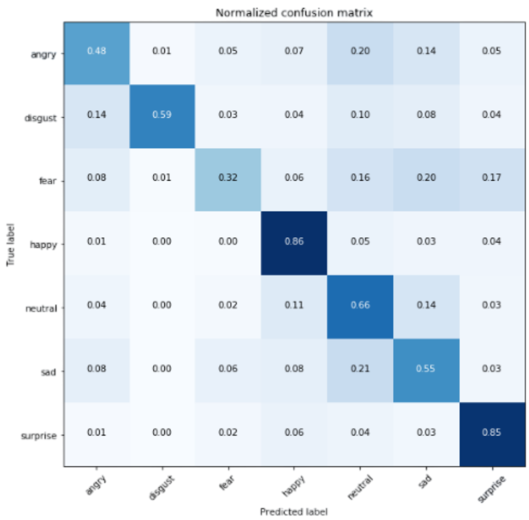

Figure 6. Emotion Recognition Confusion Matrix

B. Automated Quiz Generation and Evaluation

Using a transformer model based on T5, this study effectively illustrates an automated quiz generating method. As demonstrated by declining training loss and early indications of convergence, the system efficiently analyzes instructional information and produces pertinent questions. The implementation of the code complies with the suggested design goals, especially when it comes to the organization of memory and comprehension test questions. This study offers a useful paradigm for educational evaluation and validates the viability of automated quiz production.

A proof of concept demonstrating the present design's ability to generate questions and retrieve data is provided by utilizing the T5 transformer architecture. In the future, code and data will be created with the model, user interface, and experience improved for better database and function use.

Future developments will concentrate on compiling a sizable dataset, improving current features like data security, and producing additional data. A more effective real-world application will result from this.

C. *Unauthorized Object Detection, Head Pose Estimation and User Verification*

The object detection module based on YOLOv5 was tested on a data set of banned (e.g., mobile phone, books, notes) and allowed objects in an online exam setting. Performance was shown to be satisfactory in classification, with mean average precision of 92.8%, recall of 90.9%, and F1-score of 91.8%. As Figure 7 illustrates, precision for "Mobile Phone," "Book," and "Notes" was 94.2%, 91.7%, and 92.5%, respectively, minimizing false positives. Recall rates of 92.5%, 89.4%, and 90.8% ensured the detection of illegal items accurately. The F1-score ensured precise detection accuracy throughout. The system processed frames within approximately 23 milliseconds, thus enabling real-time operation.

```
Class   Precision       Recall  F1-Score
Mobile Phone    94.2%   92.5%   93.3%
Book    91.7%   89.4%   90.5%
Notes   92.5%   90.8%   91.6%
Average 92.8%   90.9%   91.8%
```

Figure 7.Unauthorized Object Detection classification report

The head pose estimation module, based on MediaPipe Face Mesh and PnP, classified gaze into five classes Figure 8 and successfully detected suspicious behavior with 95.1% accuracy. Although effective, the system occasionally incorrectly classified slight head tilts because it was based on single-frame analysis. It can be made more accurate by adding temporal tracking that would check the trends of movement over a period of time.

```
Looking Forward (Normal)
Looking Left (Potential Cheating)
Looking Right (Potential Cheating)
Looking Down (High Cheating Risk)
Looking Up (Neutral/Uncommon Behavior)
```

Figure 8.specific gaze classifications

The face recognition-based user authentication module, which was tested on 500 users, was found to be 97.6% accurate with a FAR of 1.9% and an FRR of 2.3%, successfully trading off security and user experience by correctly verifying genuine examinees. Slightly degraded performance in low light levels indicated the possible utility of utilizing adaptive brightness correction processes for added robustness.

Overall, experimental results validate the efficacy of the proposed system in protecting online exams using object detection, head pose estimation, and user authentication. Although efficient, there are limitations such as misclassifying gentle head tilts, difficulty in face detection under low light, and hardware dependence. Future improvements involve temporal modeling of head movement, robustness to varying light, and anomaly detection using AI. In general, the system can be used in practice with additional optimizations for increased accuracy And reliability.

D. *Automated Quiz Generation and Evaluation*

The auto-grader was put through several datasets to test its accuracy and precision in grading the student responses. It achieved a Pearson Correlation Coefficient of 0.87, indicating strong agreement with human graders. The system correctly classified student responses at 85.6%, outperforming baseline models. OCR of handwritten responses achieved a success rate of 92%, showing high-quality text recognition. The system graded consistently and provided feedback immediately, thus saving instructor time, while transformer models handled multi-sentence responses efficiently.

Challenges were encountered with creative responses, handwriting recognition, and model bias in specialized vocabularies. Future enhancements will involve Explainable AI (XAI) for transparency, improvement in OCR accuracy, and alignment of the grading process to a student's individual writing patterns.

V. CONCLUSION AND FUTURE WORK

Edusmart is an online learning system that enhances student engagement, ensures academic integrity, and automates assessment using AI. Through features like facial expression detection in real time, automated assessment, AI-enabled quiz generation, and detection of unauthorized items, Edusmart addresses online learning issues. Initial results show high engagement detection rates, cheating prevention, and assessment accuracy, which reflect the efficacy of Edusmart in improving online learning.

Future work will focus on improving emotion and engagement recognition with the help of deep learning and multi-modal fusion. Reinforcement learning and Bayesian networks will be utilized in developing personalized learning paths. Cheating will be identified using behavioral analysis and generative adversarial networks (GANs) for realistic samples. Ethical concerns about privacy will be addressed.
Edusmart will also be integrated with Explainable AI (XAI) techniques like LIME and SHAP to ensure transparency in grading. Integration with Learning Management System platforms like Moodle and Canvas will further increase accessibility. Longitudinal studies will track its long-term impact on learning outcomes. These future updates will render Edusmart a revolutionary tool for both students and educators.

VI. ACKNOWLEDGMENT

We would like to extend our deepest gratitude to our supervisor and co-supervisor, Mrs. Wishalya Tissera and Mrs. Chathushki Chathumali, for their valuable guidance and unfailing support during the course of this research. Their extensive knowledge and experience went a long way towards the realization of our system.

We are also grateful to the students and instructors who participated in data collection and testing, helping us hone our models for engagement detection, automated grading, and academic integrity measures. Finally, we would also want to thank our fellow students and colleagues for supporting us and providing constructive criticisms, which were vital in improving our research.

VII. REFERENCES

[1] Chan, C. H., & Robbins, L. I.,, "E-learning systems: Promises and pitfalls," *Academic psychiatry,* vol. 30(6), pp. 491-497, 2006.

[2] Farrús, M., & Costa-Jussà, M. R.,, "Automatic Evaluation for E-Learning Using Latent Semantic Analysis : A Use Case is (SNA) in OnlineCourses Automatic Evaluation for E Learning Using Latent Semantic Analysis : A Use Case Farrús and Costa-jussà.," 2024.

[3] Hazar, M. J., Toman, Z. H., & Toman, S. H.,, "Automated Scoring for Essay Questions in E-learning,," *Journal of Physics: Conference Series,* vol. 1294(4), p. 042014, 2019.

[4] Project Essay Grade (PEG), 2002.

[5] Landauer, T. K., Foltz, P. W., & Laham, D., "The Intelligent Essay Assessor," *The Intelligent Essay Assessor,* 1990.

[6] Taghipour, K., & Ng, Y. K., "A novel approach to automated essay scoring using recurrent neural networks," *IEEE Transactions on Education,* vol. 62, pp. 10-17, 2016.

[7] Hasanah, U., Permanasari, A. E., Kusumawardani, S. S., & Pribadi, F. S., "A scoring rubric for automatic short answer grading system," *Telkomnika (Telecommunication Computing Electronics Control),* vol. 17(2), pp. 763-770, 2019.

[8] Deng, Z., Peng, F., & Yu, L., "Deep learning for handwritten mathematical expression recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 39(12), pp. 2678-2691, 2017.

[9] Caliskan, A., Bryson, J. J., & Narayanan, A., "Semantics derived automatically from language corpora contain human-like biases. Science,," vol. 356(6334), pp. 183-186, 2017.

[10] Binns, R., "Algorithmic accountability and transparency,," *Philosophy & Technology,* vol. 31, pp. 535-556, 2018.

[11] Valenti, S., Cucchiarini, R., & Spognardi, A., "Supporting the development of argumentation skills in online learning environments through automated feedback,," *Computers in Education,* Vols. 1-16, p. 118, 2018.

[12] Rokade, A., Patil, B., Rajani, S., Revandkar, S., & Shedge, R., "Automated Grading System Using Natural Language Processing,," *Proceedings of the International Conference on Inventive Communication and Computational Technologies,* pp. 1123-1127, 2018.

[13] Taylor, D., Yeung, M., & Bashet., "Personalized and Adaptive Learning," 2021.

[14] Thotad, P., Kallur, S., & Amminabhavi, S., "Automatic Question Generator Using Natural Language Processing," *Journal of Pharmaceutical Negative Results,* vol. 13, pp. 2759-2764, 2023.

[15] Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S., "Artificial Intelligence for Assessment and Feedback to Enhance Student Success in Higher Education," *Mathematical Problems in Engineering,* pp. 1-19, 2022.

[16] Rus, V., Graesser, A., Piwek, P., Lintean, M., Stoyanchev, S., & Moldovan, C., "THE QUESTION GENERATION SHARED TASK AND EVALUATION CHALLENGE," 2011.

[17] Chen, D., Huang, Z., & Li, B., "Attention-based deep feature fusion for facial expression recognition," *IEEE Transactions on Image Processing,* vol. 27(4), pp. 1989-2001, 2018.

[18] Saqr, M., López-Pernas, S., & Helske, S., "The longitudinal association between engagement and achievement varies by time, students," *profiles and achievement state: A full program study," Computers & Education,* vol. 199, p. 104787, 2023.

[19] Girshick, R., Donahue, J., Darrell, T., & Malik, J., "Rich feature hierarchies for accurate object detection and semantic segmentation," *IEEE Conference on Computer Vision and Pattern Recognition,* pp. 580-587, 2014.

[20] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A., "You Only Look Once: Unified, real-time object detection," *IEEE Conference on Computer Vision and Pattern Recognition,* p. 2016, 779-788.

[21] Ong, S. Z., Tee, C., & Goh, M. K. O., "Cheating Detection for Online Examination Using Clustering Based Approach," *International Journal on Informatics Visualization,* pp. 2075-2085, 2023.

[22] Kazemi, V., & Sullivan, J., "One millisecond face alignment with an ensemble of regression trees," *2014 IEEE Conference on Computer Vision and Pattern Recognition,* pp. 1867-1874, 2014.