IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Emotion Driven Online Education

[1]Prof.A.H.Kagne, [2]Bhor Sankalp Pramod, [3]Bankar Asmita Narendra,[4]Shete Payal Dnyandev,[5]Sidankar Prajakta Vijay

[1]Professor at SPPU, [2]Student At SPPU, [3]Student At SPPU,[4] Student At SPPU,[5] Student At SPPU

[1]Depatment Of Computer Engineering ,

[1]Sinhgad Institutes, Pune, Maharashtra, India

Abstract: This paper suggests an interactive dashboard that uses deep learning methodologies to enhance the online classroom environment through student emotion analysis. The main challenge in digital learning is occluded or partially hidden facial data, for which CNNs and Haar cascade come into play. Haar cascade is responsible for facial feature detection and extraction by which CNNs can classify emotions.

Real-time emotion prediction is visualized on the interactive dashboard, combining inputs such as the subject, course, and instructor. Analytics on the dashboards are represented in charts and graphs and help inform teachers about student responses and how they should tweak their teaching to better engage students. It also assists in identifying students who may require extra support.

The system was validated on the CK+ dataset and was found to be highly accurate in classifying emotions such as happiness, sadness, and surprise. Other than the occasional instances in which students' faces were not visible, the system proved to be robust and consistent.

With real-time emotion-based feedback, this dashboard has transformed the dynamics of a virtual classroom into one that fosters adaptive and engaging learning, promoting the optimum achievement of educational objectives.

Keywords- Convolutional Neural Networks (CNN), emotion analysis, online learning, interactive dashboard, facial expression recognition, student engagement, deep learning, real-time analytics, virtual classroom, dataset, teaching strategies, personalized learning.

I. Introduction

This paper examines the increasing significance of emotion analysis within computer-aided learning systems, especially in environments where blended learning programs are becoming popular. Facial expression analysis, as mentioned in the paper, is hindered by occlusion when the student's facial expressions are either hidden or partially hidden; examples of these are masks, glasses, and other coverings. In this regard, the research takes advantage of the Haar cascade algorithm to detect and extract facial features. After classifying the facial regions, CNNs are then used to analyze and predict the emotions of the students. The analyzed emotions are then mapped onto an interactive dashboard that carries contextual information for reference; for instance, the subject of the lecture and the teacher.

Live visualizations in heat maps, bar charts, and scatter plots provide real-time monitoring of students' emotional conditions. Such insights allow teachers to modify their teaching methods to meet the emotional demands of the students at that moment. With Haar cascade for facial feature detection, the system captures high attention to accuracy on recognition, even if at times some regions of a student's face are blocked.

In the methodology section, the obstruction of facial information when these facial objects are in use occupies center stage together with a training procedure using the CK+ dataset for emotion recognition. One of the significant contributions of the research is a dashboard tool that implements facial emotion detection for personalized teaching interventions according to individual student needs. Other possible areas of application include health monitoring and workplace analysis.

Furthermore, the paper performs comparative methods for manipulatively oriented techniques that operate with occluded images in relation to Haar cascade. It also considers it within the frame of contextual encompassing analysis.

II. EMOTION-SCORE RECOGNITION MODEL BASED ON CNN

To train effective CNN models, we first screen CK+ datasets, then perform image pre-processing. Then, we designate labels manually to generate learning-emotion recognition datasets, and finally design convolutional networks and conduct parameter training. Finally, we will create the emotion-recognition model.

1.Data Collection: The CK+ dataset is a compilation designed for the purpose of research in the field of facial expression analysis. The images were collected under controlled conditions whereby participants were instructed to produce certain facial expressions. The dataset includes high-resolution grayscale images that have been hand labeled using the Facial Action Coding System (FACS) for accurate emotion annotation.

2.Picture Preprocessing: Facial expressions are the primary means of emotion recognition, and they might coexist with clothing or background. Nevertheless, since CNNs extract features from the entire image, they could capture environmental aspects as well, providing some degree of interference in the conscious recognition task of the data model. To rule out face-related environmental distractions, this paper employs a Haar cascade machine learning model using real-time facial detection. Haar cascade is very popular because it has proved to be an accurate system with a low false detection rate for facial recognition; the Adaboost algorithm powers it.

Due to the difference in size of extracted facial images, all detected faces are resized to a common shape before being fed to the model so that the CNN might accept them for validation. To further strengthen the robustness of the system against overfitting and, at the same time, augment the dataset, noise is added. In this case, Gaussian noise will be added to the input image, allowing the model to generalize better during training. The effect of Gaussian noise will be illustrated below on a dataset image.

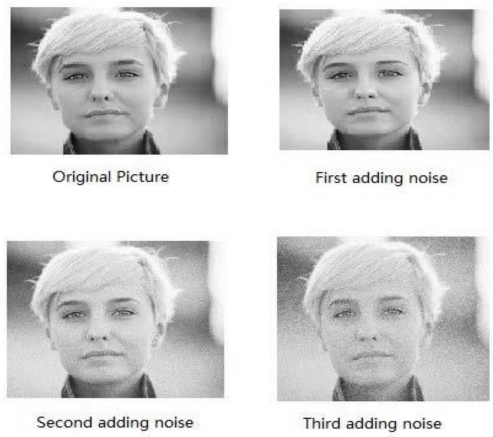

Fig.1 Images with different Guassian noise-levels

3.Caliberation of human weight scores: "The original anger, disgust, fear, happiness, sadness, surprise, and neutrality cannot be applied to this model. For example, in the context of learning, fear is nearly nonexistent, and sadness does not occur. To make the dataset compatible with the emotion-score recognition model used in this study, the calibration was manually re-adjusted based on the original annotations. The specific emotion calibration method follows the graphs below, with corresponding weights of -0.9, -0.5, 0, 0.5, and 0.9.

Designed for detecting learning-related emotions, the network consists of three main layers: an input layer, a hidden layer, and an output layer. The input layer includes 2,304 neurons, with each pixel from a 48×48 image in the dataset serving as input. The hidden layer consists of three convolutional layers, each followed by an activation function.

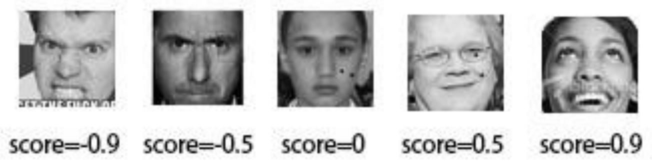

Fig. Weight-score calibration standard for learning emotion.

4.Haar cascade and CNN Model: The CNN repeatedly applies convolutions to images in the dataset by placing convolution kernels at each layer of the network structure to extract relevant features. The extracted features are then assigned corresponding weights through training parameters. Finally, the classification algorithm produces the final result. In this study, a three-layer CNN is implemented to enhance recognition accuracy. The number of convolutions is determined based on feature extraction requirements, and the activation function introduces non-linearity to the learned features, improving the network's descriptive ability.

After each convolution operation, the convolutional layer performs a pooling operation, followed by a dropout layer. The dropout layer helps reduce neuron coupling, enhances adaptability, and improves performance, mitigating common issues in traditional neural networks, such as long processing times and susceptibility to overfitting. After the dropout layer, two fully connected layers follow, playing an essential role in the network structure.

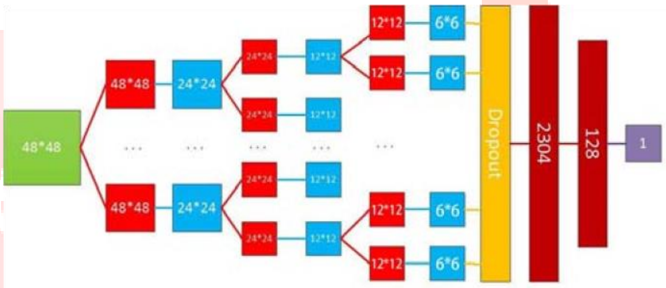

Fig. Congruent network structure for learning-emotion recognition.

The output layer consists of a neuron responsible for detecting emotional weights and utilizes the tanh function to process the output. These values are mapped to an emotional weight scale, generating predicted emotional weights.

To provide a clearer understanding of the network structure, the architecture of the convolutional network used in this study is illustrated in the figure. The number of convolutions is proportional to the amount of extracted features—the more convolutions performed, the greater the network's ability to represent images. However, excessive convolutions may capture too many specific details, increasing the risk of overfitting and reducing overall model performance.

According to the calculation, the optimal learning-emotion recognition score accuracy is ±0.14, which indicates that the identification error of the learning-emotion-recognition score is 0.14.

Layer	Number of Feature Maps	Size of Kernel	Step Length	Size of Output
Convolution 1	16	3*3	1	48*48
Pooling 1	16	2*2	2	24*24
Convolution 2	32	3*3	1	24*24
Pooling 2	32	2*2	2	12*12
Convolution 3	64	3*3	1	12*12
Pooling 3	64	2*2	2	6*6
Fully Connected 1	2,304	Fully Connected 2		128

III.REAL-TIME EMOTION RECOGNITION AND EXPERIMENTAL VERIFICATION IN ONLINE EDUCATION

A. REAL-TIME EMOTION RECOGNITION IN ONLINE EDUCATION

In online education, real-time screenshots of learners are captured, and character portraits are extracted and fed into a trained CNN model for emotion recognition to generate emotion scores. A key frame refers to a frame in which a significant character action, object movement, or change occurs within the video stream. Key frames effectively capture variations in a video . This study utilizes the open-source FFmpeg tool to extract keyframe images.

Due to differences in camera angles and distances, extracted images often have large backgrounds or incomplete character portraits. To address this, a cascaded Haar feature face-detection algorithm, based on Adaboost, is used to isolate the person in the image. Emotion-score recognition is not performed on incomplete portraits. Once a full character portrait is detected, it is resized to the specified dimensions and processed by the CNN model for emotion recognition. The resulting emotion score is then sent to the lecturer.

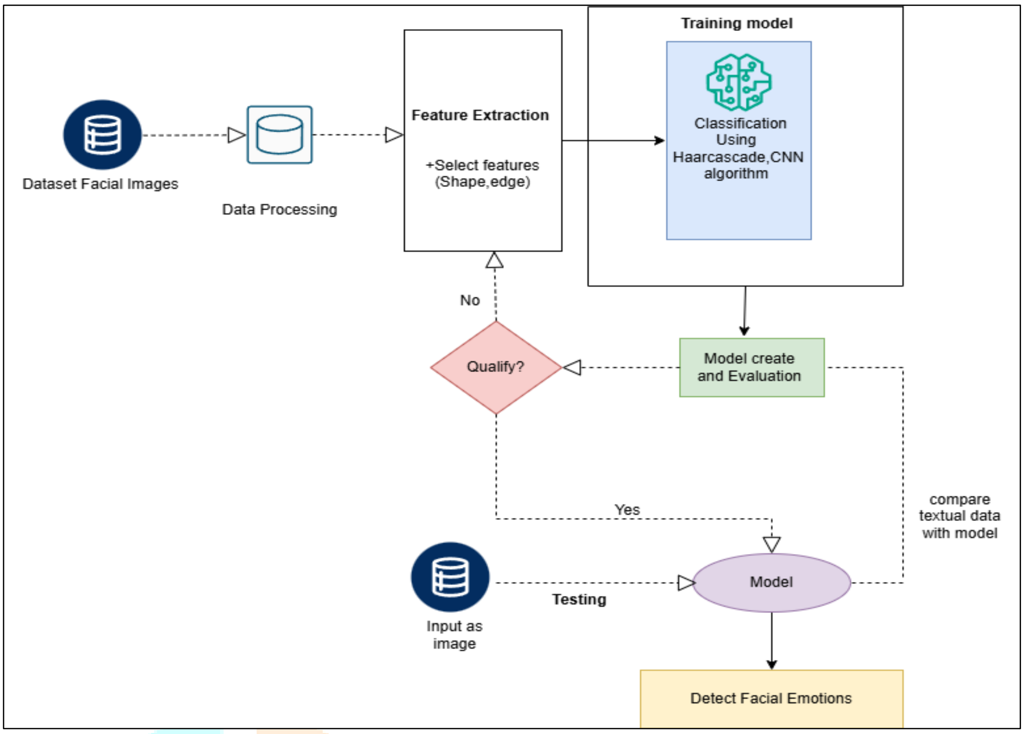

Fig.Architecture

A. Practical Applications of the Model

For practical testing, we selected the 2017 Tang Aoqing Honors Program in Science C language course at Jilin University. The class consists of 56 students, and the course has now entered the design phase. Every Tuesday and Thursday, six teaching assistants provide online guidance to students.

The 56 students were randomly divided into two groups, each consisting of 28 students and three assistant teachers. Group A used an online education platform equipped with an emotion-recognition model, allowing students to see their emotion scores on the screen. In contrast, Group B did not use an emotion-recognition model, and only the students' screens were visible

III. PROJECT SCOPE

The scope of this work is to develop a real-time emotion recognition system for virtual classrooms using a CNN model and a Haarcascade algorithm for efficient face detection. The system will detect and classify students' face expression in terms of happiness, sadness, and surprise in order to assess students' state of emotion and represent them in real-time, friendly, and graphical format for instructors to make interventions in instruction and maximize students' engagement. All preprocessing processes such as face resize and Gaussian noise for model robustness will be conducted. The work is concentrated predominantly in virtual classroom dynamics, but its application can extend to other domains such as healthcare and workplace observation. All ethical issues such as privacy, consent, and security will be addressed in a responsible use of face recognition technology. In conclusion, work will maximize students' learning through personalized interventions in terms of emotion feedback

IV. EXPERIMENTS

Loss Function: To enable the generator to produce realistic images of faces and the discriminator to discern between genuine and created images with accuracy, we intend to simultaneously optimize the generator and discriminator networks. So, we are utilizing Binary cross Entropy loss for the discriminator and L1 and L2 Loss for the generator. The discriminator loss and generator loss combined are reduced using the GAN loss.

A. GENERATOR Our Regenerative GAN trains the generator network using both L1 loss and L2 loss. It measures the typical difference, (pixel-by-pixel), between the generated and the true image. Even If there are a few minor deviations,L1loss in the context of facial reconstruction can assist in guaranteeing that the resulting image is a near match of the true image. L2 loss can ensure that the resulting image is a very good match to the original, but it might not work as well when the data contains outliers.

$$L1 = \frac{1}{bh} \sum_{i=0}^{b-1} \sum_{j=0}^{h-1} |X_i - Y_i| \tag{1}$$

$$L2 = \frac{1}{bh} \sum_{i=0}^{b-1} \sum_{j=0}^{h-1} [X_i - Y_i]^2 \tag{2}$$

where,
b-breadth
h-height
X-GANoutput
Y-original CK image

A. DISCRIMINATOR While the discriminator's job determines if an input image is real or fraudulent, the BCE loss determines the difference between the actual label and the projected output. The discriminator loss and generator loss combined are reduced using the GAN loss. This shows that the generator is trying to give samples that fool the discriminator while the discriminator is trying to classify the samples correctly.

$$BCE\ Loss = \frac{1}{b} \sum_{i=1}^{b} [log(x + log(1 - Z))] \tag{3}$$

where,
Z - regenerated image
X - original image

$$GAN\ Loss = Gen_{Loss} + Disc_{Loss} \tag{4}$$

where,

$$Gen_{Loss} - generator\ loss$$
$$Disc_{Loss} - discriminator\ loss$$

V.RESULTS:

After using three times, questionnaires for students and teaching assistants respectively were dispatched to investigate actual use consequences and in-depth interviewed several students and several teaching assistants. Overall reaction of teaching assistants in Group A was measured. In explaining students' knowledge points, whenever several students have low emotion values, assistants will adjust instruction approaches, temporarily exhibit specific examples for demonstration, or require students' answer questions through inquiry in an attempt to make contact with students. The teaching assistant clearly feels students' initiative questions and exchanges have increased too. It is no longer a traditional lecture any more. However, seeing emotion values often make them sidetracked. 53% of students in Group A stated that when have more questions and answers with teaching assistants, and have a deeper understand about knowledge points, 40% of students stated when don't dare to be interested, don't want to hear, then the teaching assistants simply posed questions, and 7% of classmates stated nothing changed.

Feedback group A The general reaction of instructors in Group B is captured below. According to their own overviews, instruction continued in a routine manner. There was no feeling of having a camera in front of them. There was no deviation in routine classes, according to students in Group B.

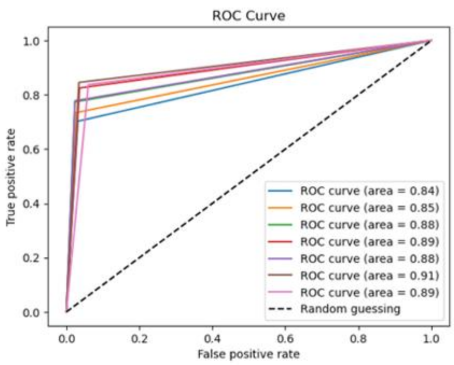

Fig.ROC Curve

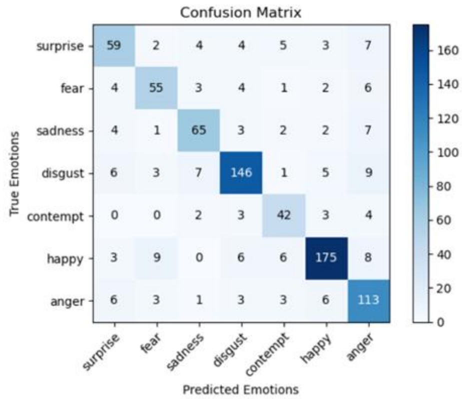

Fig.Confusion Matrix

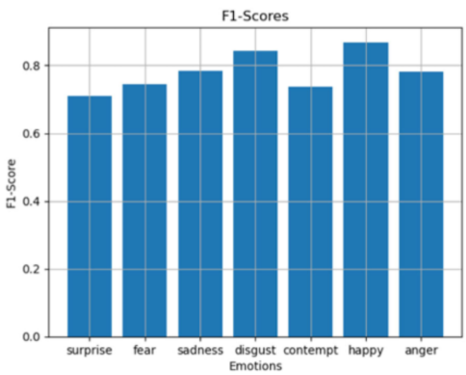

Fig.F1 score

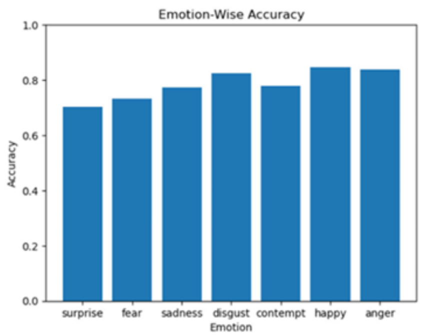

Fig.Accuracy

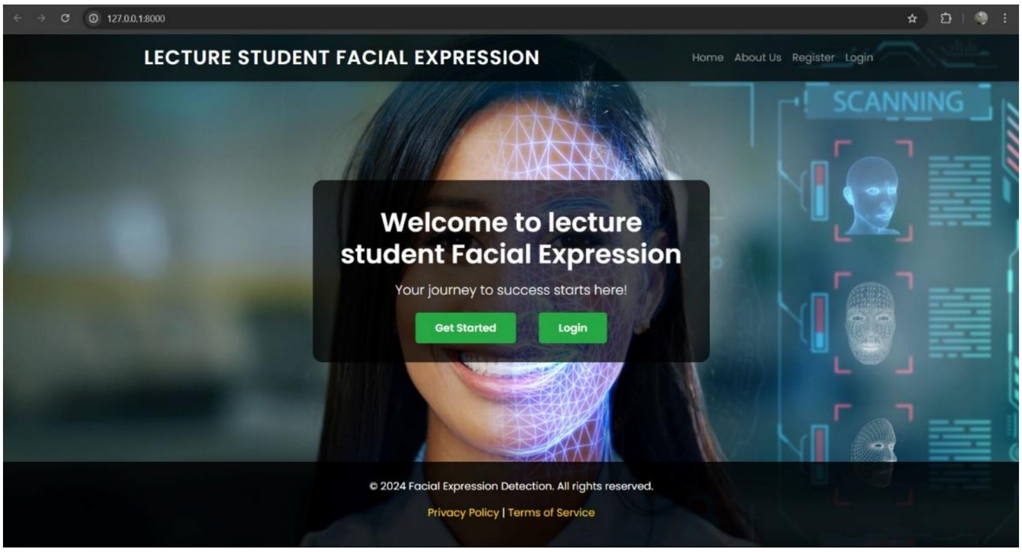

Fig.Dashboard

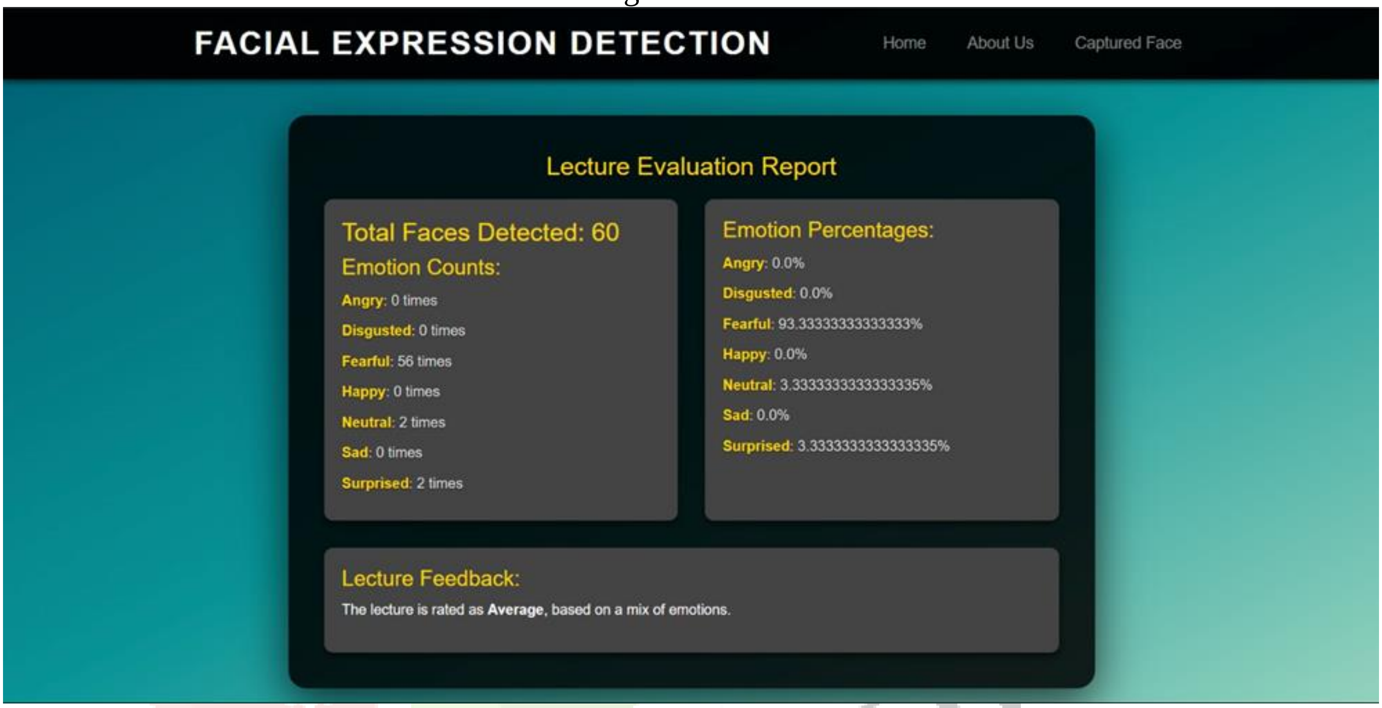

Fig.Dashboard Results

V.DISCUSSION

The results of the LSTM model demonstrate significant improvements in forecasting urban water demand. This approach offers a scalable solution that can be adapted to various urban settings. By improving water demand predictions, urban planners can optimize resource distribution and ensure sustainable water management practices. The integration of machine learning also opens avenues for future research into forecasting other urban resource needs.

VI. CONCLUSION

The research successfully developed a deep learning-based interactive dashboard to enhance the online classroom experience by analyzing students' emotions in real time. Using convolutional neural networks (CNNs) for emotion recognition and generative adversarial networks (GANs) for facial reconstruction, the system effectively detects emotions, even with occluded facial data, providing valuable insights to educators. This allows teachers to dynamically adjust their teaching strategies, improving student engagement and learning outcomes.

Validated with the CK+ dataset, the model showed high accuracy in emotion recognition and demonstrated strong generalization abilities. The dashboard's visualizations—such as heatmaps, confusion matrices, and precision-recall curves—offer educators clear, actionable data on student emotions, helping them make informed, data-driven decisions.

While the system offers significant potential to transform online learning, it also raises ethical concerns, particularly related to student privacy and data security. Ensuring the responsible use of AI in educational settings is crucial for future applications.

In conclusion, this interactive dashboard provides a powerful tool for improving the online classroom experience. However, future work must focus on refining the model's interpretability, expanding dataset diversity, and addressing ethical considerations to ensure responsible and widespread use in education.

ACKNOWLEDGEMENTS

The authors express their sincere gratitude to all individuals, institutions, and collaborators who have contributed to their respective research endeavors. The collective efforts, insights, and shared resources have been instrumental in shaping the research presented in this paper. Special thanks are extended to Prof. A.H. Kagne for her invaluable guidance and support throughout the project. Her keen interest in scrutinizing the minutiae of the project work and providing valuable suggestions have been invaluable. Additionally, Prof. S. B. Ghawate, our H.O.D., Prof S.N. Shelke provided valuable direction, enabling a thorough study of the project. The support and encouragement from the Honorable Principal, Dr. K. P. Patil, have been greatly appreciated, ensuring that all necessary resources were made available for the successful completion of this research project.

VII. REFERENCES

[1] I.Salehinand D.K.Kang,‘‘A review on drop out regularization approaches for deep neural networks within the scholarly domain,’’ *Electronics*, vol. 12, no. 14, p. 3106, Jul. 2023

[2] D. Poux, B. Allaert, N. Ihaddadene, I. M. Bilasco, C. Djeraba, and M. Bennamoun, ‘‘Dynamic facial expression recognition under partial occlusion with optical flow reconstruction,’’ *IEEE Trans. Image Process.*, vol. 31, pp. 446–457, 2022, doi: 10.1109/TIP.2021.3129120.

[3] M. Schäfer, N. Brich, J. Byska, S. M. Marques, D. Bednáa, P. Thiel, B. Kozlíková, and M. Krone, ‘‘In VA Do: Interactive visual analysis of molecular docking data,’’ *IEEE Trans. Vis. Comput. Graphics*, vol. 30, no. 4, pp. 1984–1997, Apr. 2024.

[4] K. Reese, R. Bessette, and P. Hancock, ‘‘Know Your Colors: Visual dashboards for blood metrics and healthcare analytics,’’ in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Athens, Greece, Dec. 2013, pp. 000002–000008, doi: 10.1109/ISSPIT.2013.6781845.

[5] A. Sorour and A. S. Atkins, ‘‘Big data challenge for monitoring quality in higher education institutions using business intelligence dashboards,’’ *J. Electron. Sci. Technol.*, vol. 22, no. 1, Mar. 2024, Art. no. 100233.

[6] A. Wu, Y. Wang, M. Zhou, X. He, H. Zhang, H. Qu, and D. Zhang, ‘‘Multi Vision: Designing analytical dashboards with deep learning based recommendation,’’ *IEEE Trans. Vis. Comput. Graphics*, vol. 28, no. 1, pp. 162–172, Jan. 2022, doi: 10.1109/TVCG.2021.3114826.

[7] M. N. Hasnine, H. T. Nguyen, T. T. T. Tran, H. T. T. Bui, G. Akçapınar, and H. Ueda, ‘‘A real-time learning analytics dashboard for automatic detection of online learners’ affective states,’’ *Sensors*, vol. 23, no. 9, p. 4243, Apr. 2023.

[8] J. Zhu, J. Ran, R. K.-W. Lee, K. Choo, and Z. Li, ‘‘Auto Chart: A dataset for chart-to-text generation task,’’ 2021, *arXiv:2108.06897*.