Volume 10, Issue 1, January – 2025 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14613833

Facial Expression-Based Emotion Detection for Adaptive Teaching in Educational Environments

Sathya.C[1]; Dhatchana.P[2]
Assistant Professor[1]; Student[2]
Department of Computer Science and Engineering,
Tagore Institute of Engineering and Technology,
Deviyakuruchi, Salem, Tamil Nadu, India

Abstract:- Understanding and classifying student actions within educational environments is a vital component of boosting learning results and well-being. This study presents a novel method to student activity categorisation by employing facial expression detection technologies. The technology is intended to record and evaluate pupils' facial expressions, understand their emotional states, and then classify their actions. This study investigates the application of deep learning models for face emotion identification using a dataset that includes both academic and non-academic activities. The system can recognise emotions such as happiness, sorrow, rage, and surprise. The extracted emotion traits are then used to characterise student actions, revealing whether a student is engaged, attentive, puzzled, or indifferent, among other states. This strategy has the potential to improve educational settings by offering real-time insights into student conduct and allowing for timely adjustments to improve learning experiences and outcomes. It also offers up possibilities for personalised educational support and the creation of intelligent learning systems. In this research, we will construct a system to extract face characteristics using the Grassmann method. And identify the emotions of students at certain times. Predict the active state using emotion categorisation and provide reports to the administrator. Furthermore, this technique shows potential for the creation of adaptive learning systems that react to students' emotional states, delivering extra help or challenges as needed. For example, a virtual tutor may modify the difficulty of exercises based on a student's emotional reactions, producing a dynamic and responsive learning experience.

I. INTRODUCTION

In educational settings, knowing student involvement, emotions, and activities is critical for developing effective teaching and learning tactics. Students' emotional moods are strongly related to their academic achievement and well-being. This study provides a new way to improving educational quality and providing personalised support: student activity classification based on facial expression detection. Traditional methods of monitoring student involvement and behaviour frequently rely on physical observation or self-reporting, which are subjective and restricted in scope. In contrast, our suggested approach uses computer vision and deep learning techniques to collect and analyse students' emotional responses, allowing us to categorise their behaviours more objectively and comprehensively. The heart of this technique is based on face expression detection, which has advanced significantly with the emergence of Convolutional Neural Networks (CNNs). Using these deep learning models, we can effectively identify and categorise a variety of emotions from students' facial expressions in real time, including happy, sorrow, rage, surprise, and others. These observations reveal important details regarding their emotional states. Furthermore, this technology goes beyond basic emotion identification. It correlates the recognised emotions with specific tasks, allowing us to determine if a student is actively participating in a classroom discussion, struggling with a difficult idea, or just disinterested in individual study. The real-time nature of this research allows educators and institutions to make educated decisions, change teaching tactics, and provide timely interventions to improve student learning experiences. This study also paves the way for the creation of intelligent learning systems that dynamically adapt to students' emotional states. Such systems might modify material, adjust difficulty levels, or give additional help, resulting in a more personalised and effective learning environment. However, in pursuing these breakthroughs, we must emphasise ethical issues, such as data protection, informed permission, and eliminating possible biases in emotion identification. These considerations are critical to the proper application of face expression detection in educational contexts. This research represents a significant step towards leveraging facial emotion recognition for comprehensive student activity classification. It has the potential to revolutionise education by providing real-time insights into student behaviour and fostering a more responsive and effective learning ecosystem. The existing model is shown in fig 1.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14613833

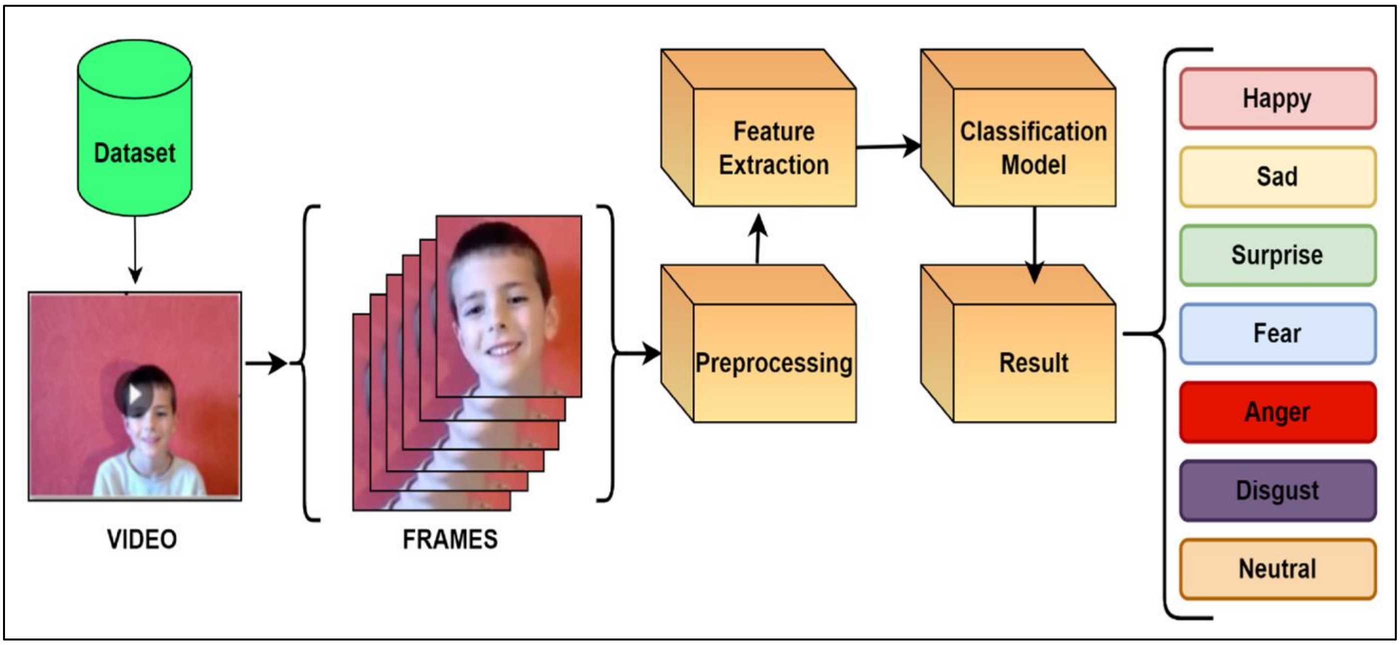

Fig 1 Video Based Facial Emotion Recognition

II. RELATED WORK

Chahak gautam, et.al,…[1] provides an innovative and effective framework for emotion recognition using feature extraction and CNN. This study demonstrates how explicit key-feature extraction from a dataset may help in successful and fast face emotion analysis. Handcrafted feature extraction is critical when you have limited data collection and want to create the model without missing any crucial or useful information. The important contribution/findings of the research are: data feature extraction acts as a useful approach for accuracy improvement, minimises the danger of overfitting, speeds up the training process, enhances data visualisation, and data processing. Emotion recognition technology is a type of facial detection and recognition that uses facial expressions as well as biophysical indications and symptoms such as pulse rate and brain activations to determine an individual's emotional state. Image-based facial expression recognition is a difficult problem, especially when it comes to determining human emotion or mood in certain scenarios, such as while enjoying or watching a series or movie, immersed in video games, shopping, or even on the battlefield. Emotions are of the biggest importance owing to an instant increase in a variety of healthcare difficulties such as depression, cancer, paralysis, and trauma. This study presents a technique to emotion identification using feature extraction and convolutional neural networks.

Mahmut dirik, et.al,…[2] suggested a study on automated facial emotion detection from facial photographs. An ANFISPSO classifier recognition model is used to create dependable decision support systems that recognise faces automatically, quickly, and robustly. The suggested technique, GPA-based normalisation, and a range of classifiers based on AU characteristics were used to compare their performance. The ANFISPSO method combines the detection and exploitation capabilities of particle swarm optimisation (PSO) and the ANFIS algorithm. The proposed ANFISPSO-based classifier had a classification accuracy of 99.6%. In conclusion, this study proposed a novel framework and highly accurate classification algorithm based on AUs for emotion recognition. The efficacy of the suggested model was assessed using a number of criteria. Compared to earlier techniques, the suggested model performed better (99.6%). The downside of this study is that it only uses static photos and does not take into account the temporal behaviour of facial expressions. Emotion identification from face photographs is a significant and active topic of research. Facial traits are commonly employed in computer vision to understand emotions, conduct cognitive science, and communicate with others. To accurately analyse facial expressions (happy, angry, sad, startled, disgusted, afraid, and neutral), a complicated system based on human-computer interaction and data is needed. It is challenging to build an effective and computationally simple technique for feature selection and emotion categorisation

Sudheer babu punuri, et.al,…[3] An efficient approach with a cutting-edge transfer learning mechanism has been proposed for face emotion identification. The system is known as EfficientNet XGBoost. The scheme's novelty is demonstrated by a specific mix of pre-trained EfficientNet architecture, fully connected layers, an XGBoost classifier, and bespoke parameter fine-tuning. The input face photos are appropriately pre-processed, and the work of feature extraction is performed using the custom model. The feature points are retrieved using different networks. To average the feature maps, global average pooling is used, and the final feature set is given into the XGBoost Classifier, which recognises class labels for different emotions. Four different datasets are utilised to validate the technique. The experimental findings for the dataset CK+ demonstrate remarkable performance with an overall accuracy rate of 100%. Furthermore, the suggested model can reliably recognise expressions while maintaining minimal latency. Datasets such as JAFFE and KDEF show an overall accuracy

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14613833

rate of 98%. Despite the uneven sample distribution in FER2013, augmentation using geometric transformation techniques resulted in a benchmark accuracy of 72.54%. To substantiate our claim, we give a comparison study of our results with previous research on current datasets. The future focus of the study would be to address the issue of boosting efficiency for unbalanced sample sets. Exploring the usage of bespoke GAN (generative adversarial networks) might be a sensible approach for recognising face emotions from the unbalanced datasets.

Ninad mehendale, et.al,…[4] devised an innovative method of face emotion recognition that takes advantage of CNN and supervised learning (possible owing to massive data). The FERC algorithm's key benefit is that it operates with multiple orientations (less than 30°) because to its unique 24-digit long EV feature matrix. The backdrop elimination provided a significant benefit in precisely assessing emotions. FERC might be the first step for numerous emotion-based applications, including deception detectors and mood-based learning for pupils. Detecting emotions from facial expressions has always been a simple task for humans, but doing it with a computer algorithm is rather difficult. With recent advances in computer vision and machine learning, it is now feasible to discern emotions in photographs. In this study, we provide a unique approach for facial emotion identification based on convolutional neural networks (FERC). The FERC is built on a two-part convolutional neural network (CNN): the first eliminates the background from the image, while the second focuses on face feature vector extraction. In the FERC model, an expressional vector (EV) is employed to identify the many forms of regular facial expressions. Supervisory data were acquired from a database of 10,000 pictures (154 individuals). It was feasible to appropriately emphasise the emotion with 96% accuracy, using an EV of length 24 values. The two-level CNN operates in sequence, and the last layer of Perceptron updates the weights and exponent values with each iteration. FERC differs from commonly used techniques using single-level CNN, resulting in improved accuracy. Furthermore, a unique backdrop removal approach done prior to EV production prevents dealing with many difficulties that may occur (for example, distance from the camera).

Aayushi chaudhari, et.al,…[5] developed a strategy for recognising emotions utilising unsupervised data that is widely available and self-supervised learning (SSL) algorithms. Using this method, we were able to save time on retraining the model or starting from scratch while also utilising currently available pretrained self-supervised learning algorithms. Using self-supervised learning as an input revealed that the created features had large dimensions and were considered high-level features, necessitating a trustworthy and in-depth fusion procedure. The results showed that we may successfully address the challenge of multimodal emotion recognition by combining self-supervised learning (SSL) with intermodality interaction techniques. Using pretrained self-supervised learning algorithms for feature extraction, we focused on enhancing the job of emotion recognition. To achieve our aim, we created a multimodal fusion methodology based on a transformer. Furthermore, we sharply defined our emotion categories by using them in two dimensions (arousal and valence). Initially, we showed that our strategy outperformed previous state-of-the-art methods by comparing our model to strong baselines from RAVDESS datasets. In the future, we intend to try recognising emotions from contextual data and categorising them into three dimensions: arousal, valence, and dominance. We also plan to put our concept to the test in the medical field to help professionals properly diagnose patients.

Liam Schoneveld et al…,[6] The study describes an enhanced deep neural network-based AVER approach. The proposed model consists of two deeper neural networks: a deep CNN model trained on knowledge distillation for FER and a tweaked and enhanced VGGish version for SER. The auditory and visual feature representations are integrated utilising a model-level fusion approach. To mimic the temporal dynamics, recurrent neural networks are employed to analyse temporally and spatially represented data. This study presents a high-performance deep neural network-based technique for AVER that incorporates a model-level fusion architecture, a modified VGGish backbone, and a visual feature extractor network. The updated face expression embedding network illustrates how to acquire robust facial expression representations by training both AffectNet and FEC concurrently. We also demonstrated that information distillation may improve facial emotion recognition even more. Our enhanced VGGish backbone feature extractor's performance suggests a viable new method for inferring emotion from audio. Furthermore, AVER has proved the usefulness of our shallow neural networks approach to multimodal fusion, surpassing state-of-the-art algorithms in predicting emotion on the RECOLA data.

Dr. P. Sumathy et al …,[7] Undoubtedly, emotion recognition will play an important role in the field of machine learning. With the recent development and widespread usage of Deep Learning and Machine Learning techniques, the prospect of creating intelligent systems that accurately grasp emotions has become more feasible. Recognition of human features and emotions is challenging due to the range of facial expressions, physical attributes, odd positions, and lighting conditions. To increase the quality of the photograph by minimising noise and illumination in the face expression images, a better image pre-processing technique utilising KNN and CA is offered. The findings of the proposed enhanced pre-processing approach with ANN classification method show that it is more accurate and has a better detection rate, sensitivity, and specificity for distinguishing human moods. When looking only at the frequency domain, a filter is one object that reduces or enhances frequencies particular to the view. The primary functions of image filters are to alter the look of images by adjusting their colours, size, shading, and other features. This filtering may be used for a variety of image processing applications, including edge enhancement, sharpening, and smoothing.

Jung Hwan Kim et al…,[8] A gadget that can reliably identify a driver's emotional expression is one way to reduce the number of fatal vehicle accidents. Ismail et al. claimed

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14613833

that furious driving increases the likelihood of an accident and endangers other people's lives. Facial expression recognition (FER) technology might protect a person from irate drivers and potentially avoid disastrous crashes. Driver moods are heavily influenced by facial emotion recognition (FER) systems. Excellent facial expression recognition in autonomous vehicles leads to reduced road rage. Even the most complex FER model performs badly in real-time testing when trained without the necessary datasets. The integrity of datasets has a greater influence on FER system performance than the accuracy of algorithms. We propose a facial image thresh (FIT) machine that uses additional capabilities that existed before to face identification and learning from the Xception method to improve the performance of FER systems for autonomous cars. In addition to the data-augmentation approach, the FIT machine needed the elimination of unnecessary facial pictures, the collection of facial images, the repair of misplaced face data, and the massive merging of original datasets.

M. A. H. Akhand et al…,[9] For a secure and safe living, smart surroundings, and a smart society, emotion recognition from facial images in unconstrained contexts (for example, public locations) where frontal view shots are not always feasible is becoming increasingly important. In order to do this, a powerful FER is required, with emotion recognition possible from a range of face perspectives, notably from varied angles. The landmark aspects of the complete image are not visible in profile views from various angles, and traditional feature extraction approaches cannot recover facial expression characteristics from side views. FER from high-resolution facial images using the DCNN model is thus regarded to be the sole way to handle such a tough assignment. The recommended FER system incorporates the TL-based technique. To make a pre-trained DCNN compatible with FER, the upper layers are replaced with dense layer(s) to improve the model's performance utilising facial expression data. The distinctive component of the proposed technique is the pipelines training process for fine-tuning: first, the thick layers are tuned, and then the remaining DCNN blocks are changed one at a time.

Avigyan Sinha et al…,[10] Facial Expression Recognition (FER) evaluates expressions in both still and moving photographs to determine the subject's emotional state. Human emotions may be communicated nonverbally through facial expressions. The algorithm classifies faces as fundamental emotions (such as anger, contempt, fear, pleasure, sadness, neutral, and surprise). In rare cases, a person's mental or physiological state of mind may also be reflected by their facial expressions (for example, exhaustion or boredom). Speech, EEG, voice quality, and text may all be used to identify emotions. Face expressions are among the most popular examples of these personality traits since they may be observed, contain a range of important features for distinguishing emotions, and are easy to discern. To create a big face collection (rather than other methods for human identification). FER can also be used with biometric identification. Technology analysis of a variety of sources, such as voice, text, device-generated health information, or blood circulation patterns inferred from images, may increase its accuracy. The project's current technology identifies emotion using traditional means such as a person's voice, facial expression, EEG, text, and so on. Nonetheless, in the history of human-computer interaction, a computer's capacity to recognise an individual's emotions is very important. We constructed a CNN model utilising data given by Jonathan Oheix to get beyond the typical technique's failure to identify emotions in HCI (computer interaction).

III. BACKGROUND OF THE WORK

There are several applications in human-computer interaction that could benefit from the capacity to perceive emotion. Face detection may be thought of as a binary categorisation of picture frames as containing or not containing a face. To learn such a classification model, we must first characterise an image in terms of characteristics that may be used to detect the presence or absence of a face on a particular picture. The present methodology typically consists of two tasks: the first is to extract ASM motion using a pyramid ASM model fitting method, and the second is to classify projected motion using Adaboost classifiers. The system then aligns three retrieved feature points, the eyes and nose section, with the mean shape of ASM and ignores the other component of the ASM against the mean face shape of ASM to estimate the geometrical dislocation information between current and mean ASM point coordinates. Then, using the Adaboost classifier, face emotions are recognised based on geometrical motion. Additionally, characteristics are extracted using Viola Jones. Viola and Jones employed wavelet-based characteristics. Wavelets are square waves with a single wavelength (one high and one low). In two dimensions, a square wave is a pair of neighbouring rectangles—one luminous and one dark.

IV. PROPOSED WORK

The suggested method for detecting student activity from facial expressions employs the Grassmann algorithm, introducing a fresh way to understanding and improving the educational experience. At its foundation, the system makes use of the Grassmann algorithm, a statistical tool designed for analysing high-dimensional data like facial expressions. The initial phase is collecting a broad dataset of students participating in various educational activities, such as in-class learning and remote education situations. This dataset collects facial photos or video frames of students participating in various activities, guaranteeing inclusion across demographics. Preprocessing procedures standardise the data by normalising pixel values and performing data augmentation to improve its diversity and quality. The Grassmann algorithm, which is well-known for its ability to reduce dimensionality and extract features, is used to recognise face emotions. This technique allows the system to extract key aspects from high-dimensional face data, such as tiny movements and expressions that communicate emotion. Grassmann's skills make it an excellent candidate for analysing and interpreting facial information associated with emotions. The Grassmann method, also known as Grassmannian subspace analysis, is a strong mathematical approach used to analyse facial characteristics in computer

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14613833

vision and facial recognition applications. One of its primary uses is dimensionality reduction, which is an important issue when dealing with high-dimensional face data. Using the Grassmann technique, it is feasible to minimise the dimensionality of facial feature vectors while keeping critical information, making it particularly useful for facial identification, emotion detection, and expression analysis. Another important use of the Grassmann method is feature extraction. It transforms high-dimensional pixel data into a more compact yet intelligible representation, highlighting significant facial traits such as texture patterns, shape variations, and landmark placements. Another aspect is subspace learning, in which the algorithm identifies subspaces within the feature space that characterise certain facial features or identities. Fig 2 shows the proposed architecture.

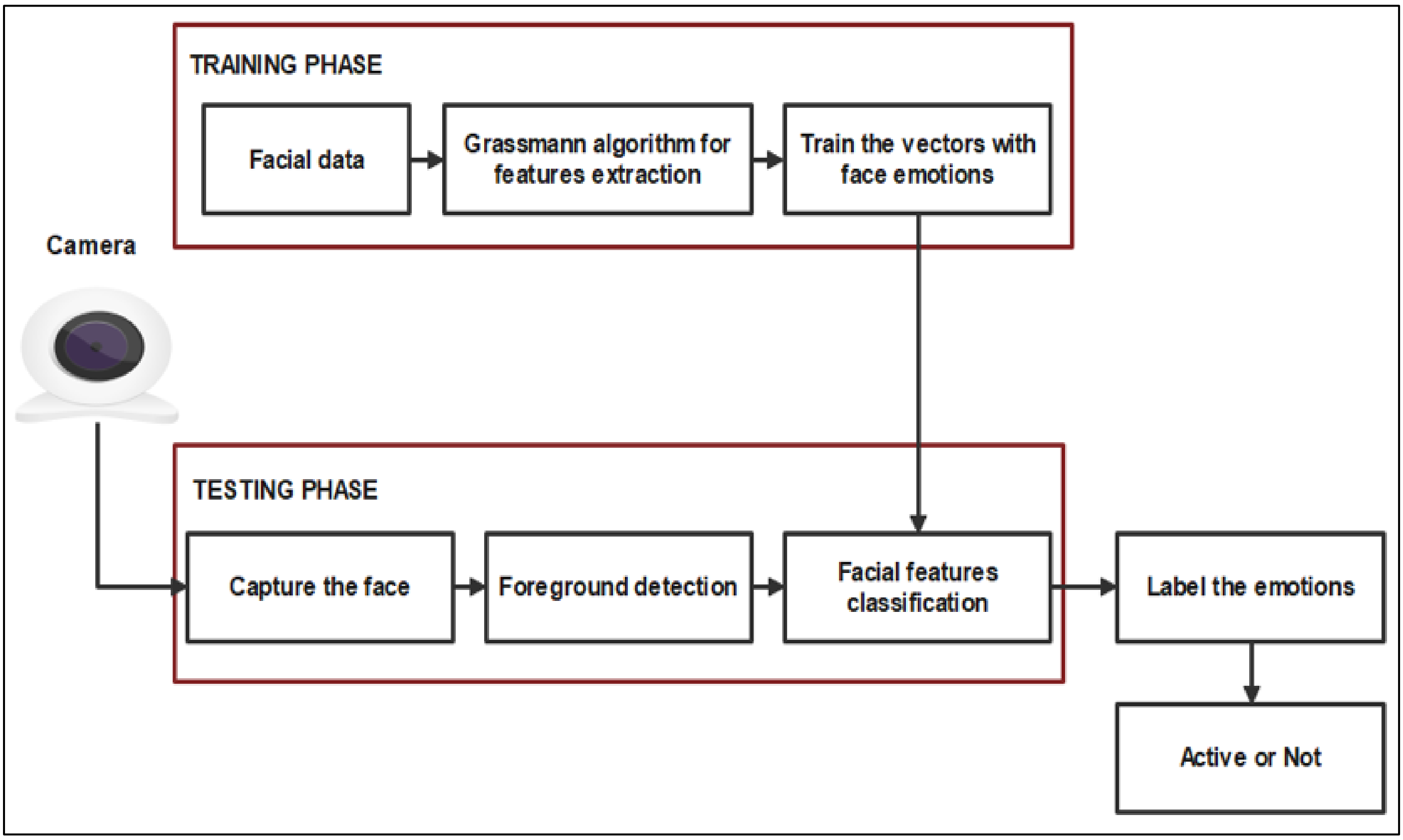

Fig 2 Proposed Architecture

➢ *Framework Construction*

- This strategy offers instructors real-time feedback, which is a big advantage. Teachers can use emotion recognition technology to assess their students' emotional responses during classes, lectures, or conversations.
- This immediate feedback allows them to make adjustment on the spot to make certain students remain engaged and understand the material.
- In this module, we can design the scaffold for students
- Student can login to the system with their details
- Admin can view the particulars about students

➢ *Features Extraction*

- Facial features extraction using Convolutional Neural Networks (CNN) is a admired approach for stress detection from facial images. CNNs are deep learning models that can robotically extract skin tone from images and learn complex patterns
- The first step in using CNN for stress detection is to collect a dataset of facial images of persons under different stress conditions.

➢ *Model Building*

- The next step is to select a CNN architecture that is suitable for facial attribute extraction.
- Popular CNN architectures for image recognition tasks embrace Sequential model. This model is pre-trained on large datasets and can be fine-tuned for stress detection.
- The CNN is then used to extract features from the facial images. This is done by fleeting the images through the CNN and extracting the activations from one of the last convolutional layers.

➢ *Activity Classification*

- Deep neural networks possess key advantages in their capability to model complex systems and utilize automatically learning skin tone through several network layers.
- As such, deep neural networks are used to carry out precision-driven tasks such as classification and classification

- Design a CNN architecture that takes in a facial image as input and predict whether the person in the image is stressed or not.
- The architecture typically consists of multiple convolutional layers followed by max-pooling layers to take out features from the images. The final layers of the network are fully connected layers that categorize the image as positive or negative.

➢ *Reports*

- In this module, provide the intelligence for all students about activity state information.
- And sentiment details stored in database for future verification.

V. EXPERIMENTAL RESULTS

The false rejection rate is the measure of the likelihood that the biometric security system will incorrectly reject an access attempt by an authorized user. A system's FRR typically is stated as the ratio of the number of false rejections divided by the number of identification attempts.

- FALSE REJECT RATE = FN / (TP+FN)
- FN =Genuine Scores Exceeding Threshold
- TP+FN = All Genuine Scores

Algorithms	FRR
Random Forest	0.42
Adaboost Classifier	0.35
CNN classifier	0.28

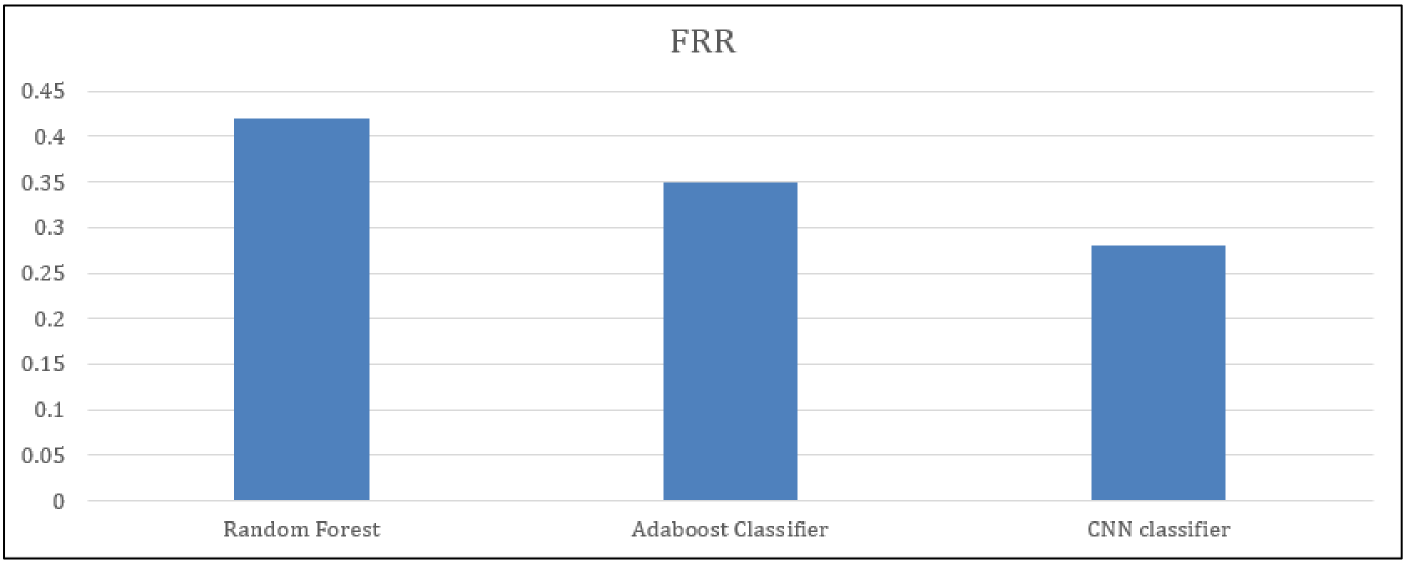

Fig 3 False Rejection Rate

The proposed system can be provided a smaller number of negative response rate than the existing algorithms such as, Random Forest, Adaboost classifier.

VI. CONCLUSION

Finally, using Convolutional Neural Networks (CNNs) to classify student behaviour based on facial expressions has significant potential for altering the educational environment. This novel technique uses deep learning to get a sophisticated knowledge of students' emotional states during various learning sessions. By analysing facial expressions and categorising emotional reactions, the technology provides real-time insights that might improve the educational experience. Using this technology, instructors may alter and personalise their teaching techniques based on real-time emotional input, which improves student engagement and understanding. The ability to detect not just fundamental emotions, but also sophisticated emotional states such as perplexity, irritation, and satisfaction broadens educators' understanding. Furthermore, the project's reach extends beyond typical classrooms to include remote and online learning contexts, offering an effective tool for measuring student engagement and well-being in a wide range of educational settings. This flexibility is particularly important in the ever-changing context of education.

REFERENCES

[1]. Gautam, Chahak, and K. R. Seeja. "Facial emotion recognition using Handcrafted features and CNN." Procedia Computer Science 218 (2023): 1295-1303.
[2]. Dirik, Mahmut. "Optimized anfis model with hybrid metaheuristic algorithms for facial emotion recognition." International Journal of Fuzzy Systems 25.2 (2023): 485-496.
[3]. Punuri, Sudheer Babu, et al. "Efficient net-XGBoost: an implementation for facial emotion recognition using transfer learning." Mathematics 11.3 (2023): 776.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14613833

[4]. Mehendale, Ninad. "Facial emotion recognition using convolutional neural networks (FERC)." SN Applied Sciences 2.3 (2020): 446.
[5]. Chaudhari, Aayushi, et al. "Facial emotion recognition with inter-modality-attention-transformer-based self-supervised learning." Electronics 12.2 (2023): 288.
[6]. Schoneveld, Liam, Alice Othmani, and Hazem Abdelkawy. "Leveraging recent advances in deep learning for audio-visual emotion recognition." Pattern Recognition Letters 146 (2021): 1-7.
[7]. Sumathy, P., and Ahilan Chandrasekaran. "An Optimized Image Pre-Processing Technique for Face Emotion Recognition System." Annals of the Romanian Society for Cell Biology 25.6 (2021): 6247-6261.
[8]. Kim, Jung Hwan, Alwin Poulose, and Dong Seog Han. "The extensive usage of the facial image threshing machine for facial emotion recognition performance." Sensors 21.6 (2021): 2026.
[9]. Akhand, M. A. H., et al. "Facial emotion recognition using transfer learning in the deep CNN." Electronics 10.9 (2021): 1036.
[10]. Sinha, Avigyan, and R. P. Aneesh. "Real time facial emotion recognition using deep learning." International Journal of Innovations and Implementations in Engineering 1 (2019).
[11]. Zhong, Yuanchang, et al. "HOG-ESRs Face Emotion Recognition Algorithm Based on HOG Feature and ESRs Method." Symmetry 13.2 (2021): 228.
[12]. Chang, Jia-Wei, et al. "Music recommender using deep embedding-based features and behavior-based reinforcement learning." Multimedia Tools and Applications 80.26 (2021): 34037-34064.
[13]. Athavle, Madhuri," Music Recommendation System Using Facial Expression Recognition Using Machine Learning, International Journal for Research in Applied Science & Engineering Technology (IJRASET), 2022
[14]. Chowdary, M. Kalpana, Tu N. Nguyen, and D. Jude Hemanth. "Deep learning-based facial emotion recognition for human–computer interaction applications." Neural Computing and Applications (2021): 1-18.
[15]. Ch, Satyanarayana. "An efficient facial emotion recognition system using novel deep learning neural network-regression activation classifier." Multimedia Tools and Applications 80.12 (2021): 17543-17568.
[16]. Mehendale, Ninad. "Facial emotion recognition using convolutional neural networks (FERC)." SN Applied Sciences 2.3 (2020): 1-8.
[17]. Ramírez, Jaime, and M. Julia Flores. "Machine learning for music genre: multifaceted review and experimentation with audioset." Journal of Intelligent Information Systems 55.3 (2020): 469-499.
[18]. Liu, Jun, Yanjun Feng, and Hongxia Wang. "Facial expression recognition using pose-guided face alignment and discriminative features based on deep learning." IEEE Access 9 (2021): 69267-69277.
[19]. Said, Yahia, and Mohammad Barr. "Human emotion recognition based on facial expressions via deep learning on high-resolution images." Multimedia Tools and Applications 80.16 (2021): 25241-25253.
[20]. Ruiz-Garcia, Ariel, et al. "Deep learning for emotion recognition in faces." International Conference on Artificial Neural Networks. Springer, Cham, 2016.