ML Driven Emotion Identification for Feedback Analysis in E-Learning Platforms

U.Venkataramana Reddy[1],Y.Sanjeeva Reddy[2],Y.Mani Baskar Reddy[3]

[1]U. Venkataramana Reddy,
venkataramanareddy.cse@gmail.com,
Dr. M.G.R Educational and Research and institute

[2]Y. Sanjeeva Reddy,
reddysanjeev304@gmail.com,
Dr. M.G.R Educational and Research and institute

[3]Y. Mani Baskar Reddy,
manibhaskarreddybasannagari@gmail.com,
Dr. M.G.R Educational and Research and institute

Dr. G. Soniya Priyatharshini
soniya.cse@drmgrdu.ac.in
Dr. M.G.R Educational and Research and institute

Mr. Saravanan Elumalai,
Saravanan.e@drmgrdu.ac.in
Dr. M.G.R Educational and Research and institute

Abstract: ***Feedback analysis for E-Learning platforms using deep learning methods for facial expression analysis is proposed in this work. They suggests a nonintrusive technique to evaluate the psychological impacts of elearning platforms by considering feedback data of users in the form of videos and analyzes emotions from video and decides the user feedback level. This process automates the understanding of feedback in a short amount of time for anyone, without requiring specific knowledge of the system.Convolutional Neural Networks (CNNs), a category of deep learning, has become increasingly popular for its ability to precisely analyze and comprehend emotions from facial photos. These models may efficiently learn and extract essential features by being trained on large datasets of tagged facial expression photos. This application takes video feedback as input for various E-learning platforms and analyzes emotions and displays the predicted emotions on video. A publicly accessible dataset is used to test the proposed technique, and the results show that it achieves 90.3% accuracy. This indicates that the approach is highly effective in accurately recognizing facial emotions.***

Keywords: Convolutional Neural Network (CNN) face images, Artificial Intelligence (AI), E learning, Machine learning, Deep learning.

1. Introduction

In several areas, including the analysis of the facial expressions to identify emotions and give video feedback and understand if user is happy or not. The ability of computers to recognize emotions is crucial for humancomputer interaction and can significantly enhance the ELearning environments. The prevalence of online learning platforms has risen owing to their adaptability and accessibility for students. Nevertheless, the efficacy of individualized and

captivating learning experiences is impacted by their frequent inability to recognize and address learners' emotions. To solve this, we are developing emotion identification systems through deep learning approaches. These systems effectively identify and classify emotions in real-time by analyzing facial expressions using specialized neural networks, such as CNNs and are utilized for analyzing users feedback [10] given to the respective online learning platforms.
This engenders captivating prospects for the customization and adaptability of learning by integrating emotion recognition technology into E-Learning platforms. This research endeavors to scrutinize the potentialities of utilizing deep learning algorithms and assessing facial cues in feedback analysis within digital learning ecosystems for emotion identification. We will probe into the advantages of integrating emotion recognition in elearning paradigms and delve into its pivotal role in augmenting human-computer interaction.
Furthermore, we will go over the strategies and tactics employed by deep learning systems, with an emphasis on how CNNs interpret facial expressions[13]. We may obtain immediate knowledge into learners' emotions by incorporating deep learning-based[16] emotion identification into feedback system for E-Learning platforms.
This capability enables the provision of tailored content and personalized support. Deep learning architectures can also engender educational artifacts that contribute to the optimization of the pedagogical continuum. While there exists a plethora of advantages to the utilization of emotion detection technology, it is imperative to judiciously navigate its implementation, ensuring due regard for user autonomy and the safeguarding of privacy. Moreover, this investigation will undertake an in-depth examination of the ethical quandaries and obstacles intrinsic to the incorporation of emotion recognition mechanisms within ELearning systems.

I. RELATED WORK

An approach for real-time identification of emotions using virtual indicators, facial features, and EEG inputs is mentioned in the article [1]. The study concentrated on children with autism and those with physical disabilities. Six facial expressions and EEG signals were classified by the program using CNN and LSTM classifiers. Fifty-five undergraduate students participated in the study to recognize face emotions, and 19 of them collected EEG signalsAn optical flow algorithm was employed to track virtual markers affixed to the subject's facial features. Concurrently, fourteen signals derived from the EEG signal reader were harnessed for emotional classification. Additionally, the spatial displacement between the individual's facial landmarks and each marker point served as a discriminative feature for expression classification. The significance of recognition of faces in several programs, including management of databases, authentication of identities, and security, is covered in this paper [2]. The paper describes a neural network (NN) framework and haar cascading detection-based neural network training algorithm for precise face recognition as well as localization. Three goals of the proposed work are to identify, recognize, and categorize faces using OpenCV, code written in Python, and a dataset. An experiment was carried out to define the emotional cues of several pupils, and the outcomes show how effective the face analysis method is. Finally, a measurement is made of the computerized face detection and identification accuracy.
They suggested [3] a LeNet structure for face emotion identification that is built on RNNs. They combined three datasets (JAFFE, KDEF, and our unique dataset) first. Subsequently, they programmed the LeNet architecture to categorise emotions. In this report, they classified seven distinct sentiments using facial lexes with a reliability of 86.43% and a validation accuracy of 81.81%.
They talk about [4] how social contact between humans and computers can be enhanced by computer-animated entities and robots. Due to the prevailing uncertainty concerning scalability, real-time, face-to-face interactions necessitate reliance on sensory-rich perceptual primitives rather than costly symbolic inference techniques. The methodology depicted in the study identifies frontal faces and encodes[5] them in real-time across seven dimensions.Its performance is improved by using SVM classifiers and boosting approaches, and it has been evaluated on a set of images of posed expressions. The outputs of the system vary smoothly over time, offering a potentially useful approximation to codify the motion of the facial emotional cues fully automatically and inconspicuously.
This research [7,8] uses an autonomous facial feature trackers for face localization and extracting features to confront the difficult problem of real-time emotion identification through expressions of the human face in the live video [9]. To infer emotions, a Support Vector Machine classifier is fed with the extracted facial traits [12]. The outcomes of tests

assessing the method's accuracy[15] for a range of circumstances, such as person-dependent and person-independent acknowledgment, are presented in the study. The findings underscore the efficacy of the proposed methodology in facilitating fully automated and imperceptible real-time facial emotion identification in video streams. The importance of searching for intelligent and emotional man-machine[17] interactions is covered in the paper's conclusion, along with some potential directions for subsequent research.

II. PROPOSED MODEL

Convolutional Neural Networks (CNNs) are a class of deep neural networks that have proven to be highly effective in various computer vision tasks, including face emotion detection.

A. Input Layer:

The input to the CNN is an image containing a face. The image is usually characterized as a 3D tensor with dimensions (height, width, channels), where channels typically represent the RGB- Red, Green, Blue color channels.

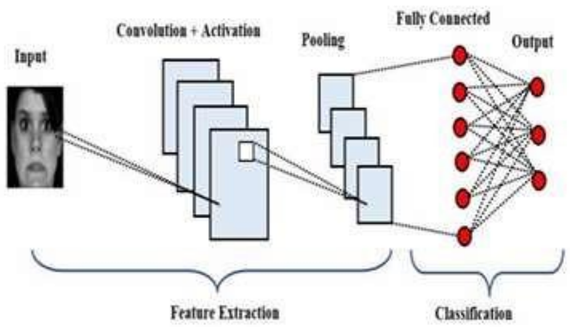

Fig. 1. Face emotion recognition model using Deep convolutional Neural Network Architecture

B. Convolutional Layers:

CNNs use convolutional layers to distinguish repetitions as well as features in the input image. Convolution involves sliding a minor filter (also known as kernel) over the input image, followed by cyphering the dot product at each step. This helps the network identify local patterns such as textures, edges, and simple shapes. The result of the convolutional layer is referred to as a feature map, elucidating the presence of specific features within the input. The formula for computing the output size of a
convolutional layer is given by:

$$\boldsymbol{Output\ Size} = \boldsymbol{Input\ Size} - \boldsymbol{Filter\ Size} + \mathbf{2} \times \boldsymbol{Padding}\ /\boldsymbol{Stride} + \mathbf{1}$$

C. Activation Function

After each convolution operation, an activation function (e.g., ReLU - Rectified Linear Unit) is applied element-wise to introduce non-linearity into the model.

D. Pooling Layers:

Pooling layers (often max pooling) are utilised to reduce the spatial dimensions of the feature maps while retaining the critical information. Pooling helps in making the detection of features invariant to scale and orientation changes.

E. Flattening:

Following numerous convolutional layers combined with pooling operations, the feature maps are transformed into a onedimensional vector by flattening them. This step prepares the data for the fully connected layers.

F. Fully Connected Layers:

The flattened vector undergoes processing in one or more fully connected layers, which constitute conventional neural network layers in the architecture. These layers learn complex relationships between the extracted features.

G. Output Layer:

The final fully connected layer yields the network's output, commonly presented as a probability distribution across distinct classes, such as emotions in this context. The Softmax activation function is often used here. The formula for the Softmax function for a class i is given by:

$$Softmax(z) = \frac{e}{\sum e}$$

where C is the number of classes, and z_i is the unnormalized logit for class i.

1) Proposed Model of Convolutional Neural Network

In the realm of image analysis, Convolutional Neural Networks (CNNs) stand out as the predominantly employed method. CNNs are distinct from multi-layer perceptrons (MLPs) in that they feature convolutional layers, which are hidden layers. The suggested approach is predicated on a CNN framework with two levels. Background elimination is the initial level that is advised; it is utilized to extract emotions from an image, as Fig. 1 illustrates. In this case, the main expressional vectors (EV) is extracted using the standard CNN network module. The expressional matrix (EV) is generated through the acknowledgment of pertinent face points of interest. EV and expression alterations are directly correlated. Utilizing a basic perceptron modeling unit on a face image devoid of the background leads to the extraction of the EV.

The final stage in the suggested training model also includes a non-convolutional perception layer. After receiving the information being provided (or image), every layer of convolution processes it and conveys the outcomes to the subsequent layer. As seen in Fig. 2, the above change is a convolution procedure. Pattern recognition is a feature shared by all the employed convolutional layers. Each convolutional layer contained four filters. In the majority of instances, each input image forwarded to the initial CNN segment, responsible for removing the background, comprises forms, textures, objects, and edges alongside the facial features. At the initial stages of the layer of convolution 1, the circle detector, edge detector, and sector detector filters are utilized. The second segment of the CNN filtering process captures facial features, encompassing the nose, cheeks, lips, eyes, and ears, subsequent to the identification of the face.

2) Data

The face emotion images dataset is available in Kaggle repository. This face emotion dataset utilized within our model includes seven categories of emotions. Each image has a variable amount of pixel size. As illustrated in the dataset split table (Table 1) below, each class is represented by an equal number of training samples. This certifies that there is no bias in the distribution, thereby considering the sample size for each of the train-testvalidation sets.

TABLE I. FACE EMOTION DATASET TABLE

Num. of Classes	7
Num. of training images	28709
Num. of Validation Images	1435
Total Images	30144

III. RESULTS AND DISCUSSIONS

TensorFlow and Keras with OpenCV library is utilised for training CNN model by using Adam optimizer and Relu optimizer. Dataset is trained with 100 epochs. Data is loaded to

VOLUME: 09 ISSUE: 04 | APRIL - 2025 SJIF RATING: 8.586 ISSN: 2582-3930

jupyter notebook and images are preprocessed with image data generator library and dataset is split into training as well as testing sets with 80/20 ratio and CNN model is initialized with given parameters and trained with 100 epochs which shows accuracy.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_3 (Conv2D)            (None, 222, 222, 32)      896
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 111, 111, 32)      0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 109, 109, 64)      18496
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 54, 54, 64)        0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 52, 52, 128)       73856
_________________________________________________________________
max_pooling2d_5 (MaxPooling2 (None, 26, 26, 128)       0
_________________________________________________________________
flatten_1 (Flatten)          (None, 86528)             0
_________________________________________________________________
dense_2 (Dense)              (None, 128)               11075712
_________________________________________________________________
dense_3 (Dense)              (None, 7)                 903
=================================================================
Total params: 11,169,863
Trainable params: 11,169,863
Non-trainable params: 0
_________________________________________________________________
```

Fig. 2. Model summary for Facial emotion detection

```
21/21 [==============================] - 21s 1s/step - loss: 0.4845 - accuracy: 0.8376 - val_loss: 3.4826 - val_accuracy: 0.2
922
Epoch 39/50
21/21 [==============================] - 21s 996ms/step - loss: 0.4442 - accuracy: 0.8525 - val_loss: 4.1884 - val_accuracy:
0.3247
Epoch 40/50
21/21 [==============================] - 20s 970ms/step - loss: 0.4247 - accuracy: 0.8480 - val_loss: 4.1700 - val_accuracy:
0.2987
Epoch 41/50
21/21 [==============================] - 21s 1s/step - loss: 0.4168 - accuracy: 0.8599 - val_loss: 4.0242 - val_accuracy: 0.2
987
Epoch 42/50
21/21 [==============================] - 21s 1s/step - loss: 0.4030 - accuracy: 0.8659 - val_loss: 4.3613 - val_accuracy: 0.2
727
Epoch 43/50
21/21 [==============================] - 21s 993ms/step - loss: 0.3802 - accuracy: 0.8495 - val_loss: 3.8534 - val_accuracy:
0.3312
Epoch 44/50
13/21 [=================>............] - ETA: 7s - loss: 0.3591 - accuracy: 0.9012
```

Fig. 3. Depicts the accuracy & loss of the value

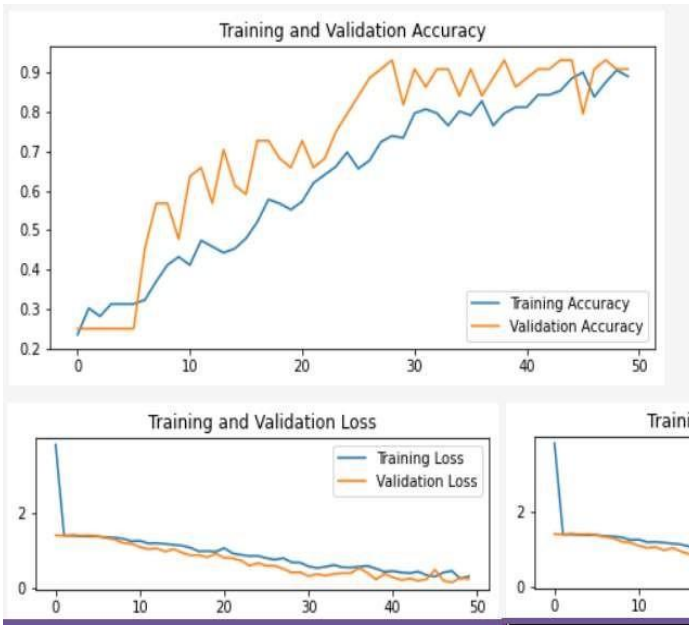

Fig. 4. Training loss and accuracy

The above figure 4 displays the accuracy of the model and loss and we can observe that loss decreases over time and accuracy increases in both training and validation.

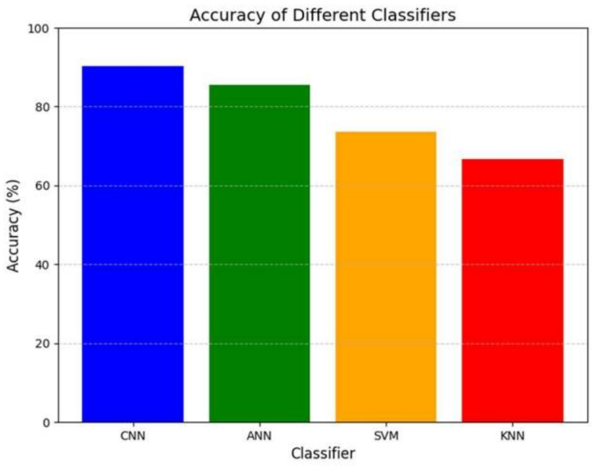

Fig. 5. Accuracy Comparison chart

The comparison chart, Fig 5, assesses the performance of CNN, ANN, SVM, and KNN classifiers for video-based emotion recognition.

The provided comparison evaluates the performance of different classifiers for emotion detection in videos, a crucial task in affective computing and human-computer interaction research. Four classifiers, namely Convolutional Neural Network (CNN), Artificial Neural Network (ANN), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), were considered for their ability to accurately recognize emotions depicted in video data.

Among the classifiers, the Convolutional Neural Network (CNN) exhibited the highest accuracy of 90.3%, showcasing its effectiveness in learning intricate spatial features from video frames and capturing complex patterns related to emotional expressions. This underscores the significance of deep learning approaches in video-based emotion recognition tasks, aligning with contemporary research trends leveraging deep neural networks for multimodal emotion analysis.

The Artificial Neural Network (ANN) demonstrated a commendable accuracy of 85.50%, indicating its capability to model nonlinear relationships and extract relevant features for emotion classification. Although slightly lower than the CNN, ANN's performance highlights the efficacy of traditional neural network architectures in processing sequential data, albeit with potentially lesser model complexity.

Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) classifiers achieved accuracies of 73.50% and 66.75%, respectively. While SVM leverages the notion of maximizing margin between classes for classification, and KNN relies on proximity-based voting, their accuracies indicate relatively lower performance compared to deep learning models.

Overall, the comparison underscores the dominance of deep learning techniques, particularly Convolutional Neural Networks, in achieving superior accuracy for emotion recognition in videos.

IV. CONCLUSION

This paper introduces a feedback analysis system predicated on the analysis of human facial expressions within user-uploaded videos pertaining to an E-learning platform. To facilitate this analysis, we have devised a web application leveraging the Flask framework. Designed to accommodate numerous E-learning users seeking to furnish feedback on diverse E-learning platforms, this web application allows users to upload videos and access real-time feedback, enabled through the detection of emotions within the uploaded content. To extract human facial expressions from feedback videos and identify emotions, we propose employing a dual-layer approach of convolutional neural networks. This algorithm uses 2800 photos from the Face Emotion Image dataset to classify seven different human facial emotions. Human emotions include happiness, sadness, disgust, anger, fear, surprise, and neutrality. The best fit and data generalizability of the model is demonstrated by its comparable training and validation accuracy. We demonstrate feedback analysis form video uploaded by users in E learning platform.

REFERENCES

[1] Ayvaz, Uğur; Gürüler, Hüseyin; Devrim, Mehmet Osman "USE OF FACIAL EMOTION RECOGNITION IN E-LEARNINGSYSTEMS"
[CrossRef], Sep 2017.

[2] D'Errico Francesca, Paciello Marinella, De Carolis Bernardina,Vattanid Alessandro Palestra Giuseppe, Anzivino Giuseppe CognitiveEmotions in E-Learning Processes and Their Potential Relationshipwith Students' Educational Adjustment, 2018

[3] Hassouneh Aya, A.M. Mutawa, and M. Murugappan "Developmentof a Real-Time Emotion-Recognition System Using FacialExpressions & EEG based on machine-learning and deep neuralnetwork" , 2020

[4] Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi "A realtime face emotion classification and recognition using deeplearning-model", 2020

[5] Auto Encoders and Decoders Techniques of Convolutional Neural Network Approach for Image Denoising In Deep Learning,P Chilukuri, JRA Kumar, R Anusuya, MR Prabhu,Journal of Pharmaceutical Negative Results 13 (4), 1036-1040, 2022

[6] Mehmet Akif Ozdemir, Berkay Elagoz, Aysegul Alaybeyoglu, RezaSadighzadeh and Aydin Akan "Real Time Emotion Recognition fromFacial Expressions Using CNN Architecture" , Oct 2019.

[7] Isha Talegaonkar, Kalyani Joshi, Shreya Valunj, Rucha Kohok andAnagha Kulkarni "Real Time Facial Expression Recognition usingDeep Learning", 2019

[8] Marian Stewart Bartlett, Gwen Littlewort, Ian Fasel and Javier R.Movellan "Real Time Face Detection and Facial ExpressionRecognition: Development and Applications to Human ComputerInteraction.", 2003

[9] Aneta Kartali, Miloš Roglić, Marko Barjaktarović, MilicaĐurić-Jovičić and Milica M. Janković "Real-time Algorithms forFacial Emotion Recognition: A Comparison of DifferentApproaches"., 2018

[10] Ati Jain, Hare Ram Sah "Student's Feedback by emotion and speechrecognition through Deep Learning", 2021

[11] Wisal Hashim Abdulsalam, Rafah Shihab Alhamdani, andMohammed Najm Abdullah "Facial Emotion Recognition fromVideos Using Deep Convolutional Neural Networks" , 2019

[12] PHILIPP MICHEL AND RANA EL KALIOUBY "REAL TIME FACIAL EXPRESSIONRECOGNITION IN VIDEO USING SUPPORT VECTOR MACHINES" , 2003

[13] David Dukić and Ana Sovic Krzic "Real-Time Facial
ExpressionRecognition Using Deep Learning with Application in the ActiveClassroom Environment", 2022

[14] A.V.Sriharsha et al. "Electroencephalography image classification using convolutional neural networks." In The International Conference on Innovations in Computing Research, pp. 42-52. Cham: Springer
International Publishing, 2022.

[15] T. S, R. B. N and M. K. R, "Machine Learning Algorithm to Detect EEG based Emotion states using Virtual-Video stimuli," 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), Bangalore, India, 2023

[16] K. Chandni, A. P. Singh, A. K. Rai and A. S. Chauhan, "Image Recognition Using Machine Learning Techniques," 2023 International Conference on Sustainable Emerging Innovations in Engineering and Technology (ICSEIET), Ghaziabad, India, 2023

[17] C. -S. Jiang and Z. -T. Liu, "Emotion Recognition via Environmental Context and Human Body," 2023 IEEE 6th International Conference on Industrial Cyber-Physical Systems (ICPS), Wuhan, China, 2023.