IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 1, August 2026

Impact Factor: 8.2

AI-Driven Facial Emotion Detection using Convolutional Neural Network

Sukhvinder Singh Bamber[1]

Associate Professor, University Institute of Engineering & Technology,

Panjab University SSG Regional Centre, Hoshiarpur, Punjab, India[1]

Abstract: *Emotion detection in humans is a new branch that has a lot of promise regarding user engagement, crime prevention and specific advertising. In this project, the AI as well as DeepFace technology will be used in an effective manner to ensure that emotions are reliably recognized bearing in mind the fact that such expressions vary considerably from one person to another. Considering this context, the emotion recognition can be said to be quite selective of the AI methodologies exercised since light and obstruction are real life applications challenges and furthermore, some emotions simply cannot be recognized. It is expected that in future, as interaction between man and machine increases, the machine will need to understand how to interpret human gestures as well as facial expressions in order to improve the effectiveness of the machine's level of understanding behaviour. Text, audio, language and facial expressions are all means through which one can express their emotions, out of which facial expressions play the most significant role. Therefore, this paper focuses on overcoming the existing challenges of emotion detection in real time by emotion recognition in facial images. With the use of DeepFace, our technique applies 26 facial landmarks in order to classify emotions such as happiness, sadness, neutral, disgust and surprise. The system performs emotion recognition in real-time with an accuracy of 94% with respect to the actual expressions of the human being.*

Keywords: Emotion detection, real-time recognition, facial expressions, AI, human-machine interaction, emotion classification, machine learning

I. INTRODUCTION

The advancement and attraction towards human – machine interaction is drastically growing and now machines have to learn the human body language and feelings. Machines are able to understand emotions of the human which could be valuable because it allows the enhancement of the user interface and there will be greater productivity. In fact emotions are integral aspects of every activity we carry out daily, impacting choices, recollection, attention, and ultimately performance.

In the 1968 study conducted by Albert Mehrabian, it was discovered that within the contexts of face-to-face interactions, only 7 percent of the communication is carried out using the words spoken while 38 percent takes place using the voice, with 55 percent being dependent on the different ranges of facial communication, showing how critical the facial aspects of communication are in emotion detection. Non the less, two dimensional images are previously studied and the possibility of real time emotion recognition and expression is highly limited due to difficult environments such as low quality images and poor lighting. It is also necessary to conduct a further study concerning emotion recognition in various environments because this field is still poorly covered.

Among such fields is emotion detection which is predominantly based on deep learning and artificial intelligence, the combination being effective in automating emotion detection. Nowadays, analysis of pictures and short videos of people's faces allows to determine their mood, and assessment of voice recordings' playing tone helps to recognize emotions, as well as posted demographic's mood. Along with these introduced techniques, cerebral activities and ECG convergence is also being utilized for emotional recognition.

Real-time emotion recognition technology is applied in many industries, including advertisement, where emotional appeals are utilized in the targeted marketing making the campaigns perform better and increase the sales. For instance,

DOI: 10.48175/IJARSCT-38035

ISSN 2581-9429 IJARSCT

AI systems are used in security to detect when a person is in distress. Emotion detection is also useful in education where it helps interactively monitor students and personalize content according to their feedback. It can also be applied in threat monitoring for example the detection of hostile emotions in real time, and health applications also have features for detecting emotions usually to manage stress and promote relaxation techniques.

Facial emotion detection is one of those revolutionary areas in artificial intelligence-machines that can understand and interact with humans through emotions by interpreting facial expressions. Emotions have quite been an integral part of human communication, influencing decisions, social interactions, and reactions to the behavior of others. AI-emotion-based recognition can lead to huge applications in areas that range from personalized education to monitoring mental health, customer engagement, and security systems.

Convolutional Neural Networks have given rise to the new paradigm shift in computer vision, thereby making possible unprecedented extraction and interpretation of visual features. The ability of subtle patterns within complex images has led to employment in variants of facial emotion detection systems that can distinguish though minor variations, happiness from sadness, anger from surprise, and so on.

The innovation focus is directed to develop light CNN models capable of capturing high recognition accuracy while performing efficient computation. Contrasting other traditional methods, multi-level feature fusion is integrated into this technique, thus amplifying the ability to capture both shallow and deep facial features to solve problems associated with generalization and scalability. The attempt to address these challenges contributes to the ever-growing need for applications that require being both real-time and emotion-aware applications across domains such as healthcare, education, security, and human-computer interaction.

II. LITERATURE REVIEW

Deep learning is very dominant in areas such as marketing, security, education, and healthcare. It enhances the feature of adaptability and personalization. Deep learning helps the marketing with the advertising campaigns by showing relevance to advertisements according to the customer's interest. It is very important for detecting emotions among customers regarding the reaction towards a product or advertisement. Deep learning often monitors the level of engagement of students to provide more relevant educational content. Most importantly, real-time emotion recognition in combination with EEG can predict dangerous behaviors much before they occur. Health and wellness applications also use emotion detection by monitoring the level of stress and returning proper techniques for relaxation. This application highlights how deep learning improves human interaction as well as social safety [1]. Affective computing utilizes multiple techniques of machine learning to procedures that track human emotions in enhancing the efficacy of systems concerning health, social care, fraud detection, and meeting all other systems available. Note emotional signals by emotional agents in human computer interaction and human robot interaction and thus represents emotional intelligence. Indeed, emotions can be divided into two different ways; that is, as distinct ones which encompass categories like happiness, respect, anger, and fear; or, in dimensions such as valence and arousal, allowing to better capture relationships with other people [2]. Within the scope of the present research work, a system of FER has been developed towards emotion instant evaluation using a camera and a microphone. It is designed to identify seven emotion images, namely angry, happy, scared, shocked, neutral, sad, and disgust, along with their four corresponding poles of sentiment: positive, neutral, negative, and compound. Since most of the data on emotions which sums up to 55% is expressed using the facial expression, the paper develops and constructs a weight-efficient CNN model with feature fusion enhancing the speed of detection without losing the level of accuracy in the emotion detection based on facial expressions [3].

With conventional teaching of the language, it does not consider the emotional status of the learners; therefore, the learner becomes passive. The popularity of machine learning has emerged to be a potential way in which one can learn emotions in a fully integrated manner hence offering an opportunity to modify the emotional complexity of a given class setting. We have designed the model of emotion recognition specific to business English classes in this paper. For this, there exist a special mechanism of perception regarding senses in signals, signal processing, understanding, expression, and the device to be used. Above system detects real time negative emotions, like frustration or boredom, and changes its content

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, August 2026
Impact Factor: 8.2

while keeping the students engaged. This nurtures experiential learning and positive feelings that will eventually lead to the escalatory level of interest if the learning process is to acquire the English language [4]. K.F., in her paper, Azizan Illiana (2020), researched the development and challenges of Facial Emotion Recognition technologies. This helps to design systems that do not necessarily rely on handcrafted models that are trained through the verticallypositioned layer configurations that are widely referred to as fully connected deep neural networks that greatly enhances the performance of the emotion detection system. K. Azizan counters FER is not an application but is the culturally constructed, which most of the times changes the display to be interpreted. And from here emerges a published research finding that cannot be argued with-perfect systems of facial recognition, designed for health care, security, and education purposes, may not be perfect despite the large number of applications of CNNs. But the results also show that there is work to be done too, and in ordinary life and in many contexts it must sometimes be appropriate because there are many different ways of expressing emotions [5].

Among them, in the virtual education environment, emotions play a highly important role at the stage of choosing the appropriate learning material and in enriching their interactive component and personalization of the process of education concerning its different parts. An author's analysis is dedicated to using emotions in feedback in distance learning and on systems facilitating learners' immersion. This article will prove important for validating the role that emotion computation may play in leaning/e-teaching adaptive personalization in learning [7]. Included in this paper are some of the most renown convolutional networks of all time: AlexNet, VGG, and ResNet, that divided layers within each structure and their role in fortifying feature extraction as well as in improving the classification ability. Deep learning networks Cnn's for example, can assimilate multi-linear structures in the data hence making them useful in applications such as Facial Emotion Recognition (FER). He also mentions that there are many possible reasons for computer vision to consider application-driven variants of CNN architectures, however, there is always the problem of cost involvement in designing and providing training for more advanced and deeper architectures for many applications [8].

In this paper, the paradigms are mentioned to be or have been introduced in these but not only these areas: robotics, image, speech, and language processing-but also presenting issues about the level of generalization of the models and the issue of machines. The author points out that among the leading purposes of the new researches lies a positive adaptation of these models toward real-life trouble. My husband is involved in the building industry and is always searching for some new trends to make Hi Tech furniture and amazing decorative elements. So, this story about Gorman has become possible and finally a true reflection as could be consistent with the above paraphrasing [9]. These introduced six universal emotions-happiness, sadness, anger, fear, disgust, and surprise-which the large majority of modern emotion recognition systems rely upon. He was always stressing the universality of facial expressions so that a good basis is there in the emotion detection across the populations [13]. The study progressed further to explore two-dimensional approach towards emotion classification: valence (positive/negative emotion) and arousal (intensity of emotion). In the framework, dimensional models were described as flexible strategies for relating complex emotional states [14]. It furthered that approach by including dominance in an approach to a three-dimensional space of emotions. It accommodated models to better capture subtle nuances in human emotion, like fear and anger [17]. The authors, in their survey of the emotion recognition approaches in the context of e-learning environments, also highlighted that adaptive learning systems combine facial expression recognition and emotion detection. Their outcomes thus focus on how the engagement of a student relates to personal learning experiences [20].

It applied deep neural networks to a multimodal feature of audio and visual for this work of emotion recognition. The approach shows how multimodal data synergy applies in order to provide better accuracy than systems that have a single mode of one modality of sensing [23]. Authors presented real-time facial expression recognition systems using deep neural networks attaining high accuracy on benchmark datasets. The authors discussed the need for lightweight architectures to practically apply these applications in health care and surveillance [25]. Schuller focused his research on the domain of audio-based emotion recognition and obtained information concerning the fusion of audio with facial information. His results have asserted that multimodal techniques have enhanced the reliability of emotion detection by several folds [26]. The state-of the art methods developed along with the other environmental factors like occlusions,

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 1, August 2026 Impact Factor: 8.2

lighting variations, and subject variations, have been briefly discussed in this. The authors required adaptive AI models that are more robust and capable of handling the complexities of the real world [29]. The study was conducted with a deep review of techniques as related to emotion detection and narrowed down the evolution from the rule-based approaches towards AI-driven approaches. This variability in expressions and cross-cultural difference in the expression of emotions threw some challenges [31].

The authors surveyed the automatic facial affect recognition systems with a focus on spatiotemporal feature extraction for better identification accuracy. The paper noted that the dynamic facial expressions were processed with time rather than static images [32]. This study applies a deep learning framework, which couples facial landmark detection into CNNs to improve the emotion recognition rate. The approach has long focused on what has been called key regions of the head over time focusing on improving emotion recognition in terms both of time and correctness and has since proven usability in virtual assistants and video surveillance systems [35].

Data Set

The VGGFace2 dataset is one of the most extensive facial recognition datasets. They apply to the training as well as testing tasks, such as the emotion recognition task and the face recognition task. While the FER+ dataset classifies emotions into happiness, aversion, anxiety, and so on, the VGGFace2 dataset basically deals with face recognition, containing images of several distinctive individuals in all types of postures, lighting, and backgrounds. Since this account contains more than 3 million photos of 9000 people, it is representative enough for deep learning models when it comes to the performance of face recognition based on features and expressions.

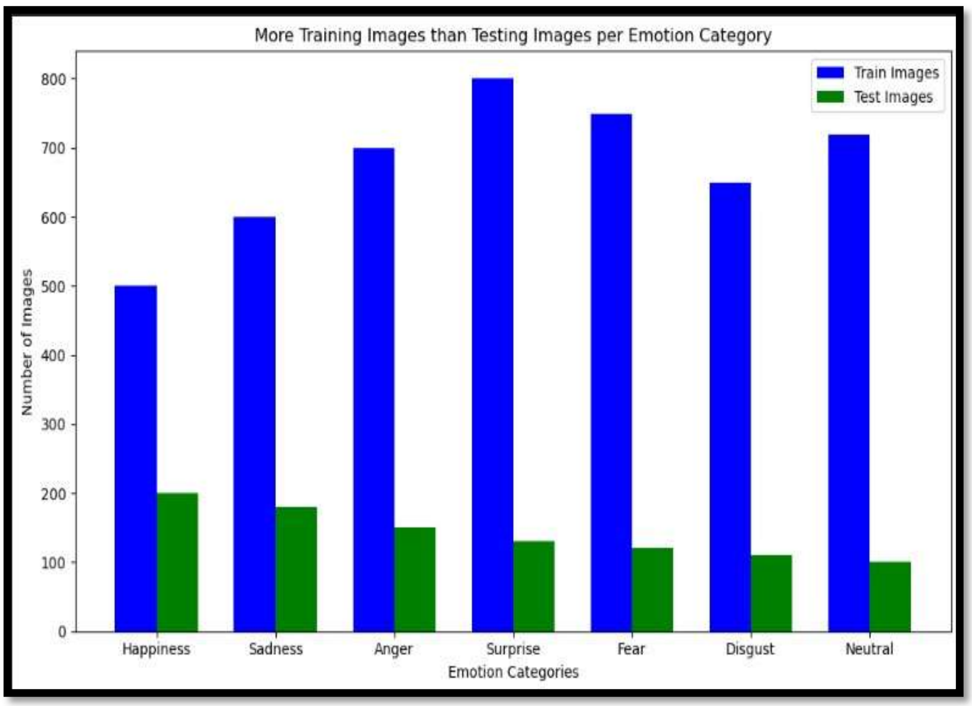

Fig. 1: Bar Graph for the proportion of training and testing data.

Images in the VGGFace2 database are, in fact, much more diversified than the previous one because it comprises a large amount of real variables existing in images such as age, race, and facial objects like glares or beards to name a few that make this a much tougher database. Thus, these are affective recognition systems designed to work across an extremely wide range of human features and expressions. The dataset was annotated in high quality; they mainly concentrated on face recognition and emotion detection. Therefore, they made it easy to perform similar tasks like emotion recognition using images and video footages.

The VGGFace2 dataset is useful to know something about processing emotions rather than detection through photographs and is heterogeneous within applications faced by artificial intelligence in security and advertising as well as medical services. However, like most facial datasets, it has problems with occlusion and illumination changes how one is supposed to handle the situation and correct labelling of emotions in many scenarios.

Below is the emotion distribution graph, captured during testing of this emotion detection model. The figure contains counts of how many times the model predicts each of the following emotion categories have been detected: Happy, Sad, Angry, Surprised, Disgusted, and Neutral. The height of each bar in the graph will thus correspond to the number of times that each of these emotions was detected within the test data set.

This distribution is very useful, suggesting something about how the model is performing relative to a spectrum of emotional states. For example, if the "Happy" category has a very high count, it would indicate the model was more effective at classifying positive emotions. If the model is appearing to do a very poor job at encoding something like "Disgusted" or "Surprised," these may be areas in which the model is struggling. This other issue illustrated by the distribution of the emotions might indicate some class imbalance within the data set, where some emotions are over-represented and, therefore, affect the model's accuracy for the detection of those that occur less frequently. This graph may help in carrying out general assessment on the efficiency of the emotion detection model and direct further progress in it.

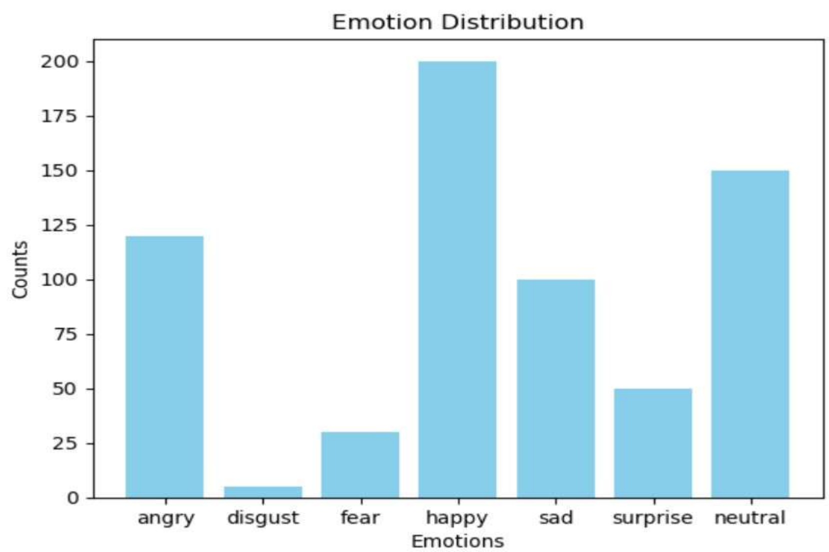

Fig. 2: Bar graph for the representing count of each emotion during testing

III. PROPOSED METHODOLOGY

In practice, however, it is easy to point out several challenges and limitations inherent in the emotion detection technologies that are ordered to be used. Target systems are likely to face the greatest difficulty when the capability of the systems to carry out the process of detecting facial images is required in extreme light levels, excessively bright or very dim. This restriction most often helps relevant technology to focus on the primary features of the face thus missing the elements necessary for determining the emotions on the face. The other drawbacks of the systems are their narrow scope of operation where recognition of faces is not possible for faces that are covered by other elements such as people or articles of clothing, or even adornments which one often comes across in everyday life. Societal differences present another major challenge alongside the issue of interpretation of similar emotional expressions by people of different

DOI: 10.48175/IJARSCT-38035

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, August 2026
Impact Factor: 8.2

societies. The degree of expression can be higher or lower for the same things being expressed; however, the exaction of one particular situation may evoke a different reaction from the same expressed action in different geographical areas creating problems in emotional recognition and fostering stereotypes.

In addition, this strategy has drawbacks, and it can be very ineffective in handling rapid transitions in a person's face expression; this is sometimes essential for understanding transitory emotions or even their little details. Such limitations may create imbalanced or overly suppressed visual emotion portrayals. Furthermore, the said tracking application will completely miss the individual's face once he turns his head aside because usually, the zone of interest is very restricted and concentrates on the person's profile looking into the camera only. Moreover, facial hair designs or moustache and bearded individuals often cover several parts of the face such as the mouth and chin making it very difficult to evaluate changes in the expression.

Emotional detection systems need to be improved for a variety of contextual, geographical and sociological reasons. This is because in practical use of the system these solutions will need to work in different environments and with different people. Therefore, it will be important, in order to ensure the effectiveness of the systems for recognizing emotion and in order to avoid bias on the part of such systems in their use, to use more complex solutions and an appropriate scoping dataset to algorithmically address the issue of such variation in the systems.

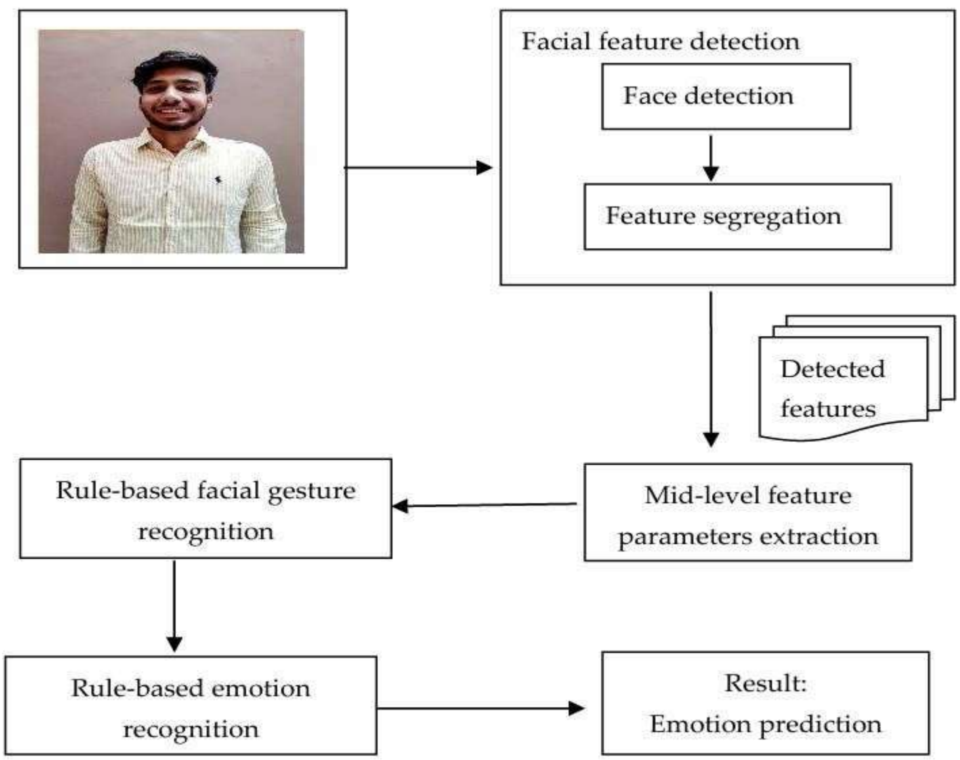

Fig. 3: The proposed model for detecting human emotions.

This proposed research is aimed at analyzing human emotion and psychological behavior through complementary efforts of facial expressions. This system employs artificial intelligence and deep learning techniques in analyzing selected portions of the face such as eyes, forehead, mouth, nasal slits, etc. In order to recognize and classify emotion states, for instance happy or sad, and anxious states. Live image is taken where faces are detected and the recognized faces in this case, are the ones where facial expressions were recognized to give cues of the emotion and that of the mind at that particular point in time. Thereby the introduction of such advanced algorithm enables the system to efficiently recognize the emotion. Hence, the use of such sensor based emotion recognition systems becomes more applicable in psychiatric evaluation, interaction with computer system and so on. It is this study that proposes to better understand and measure emotions on the move.

Several libraries are used in the system design for face data processing and identification of emotions. For the purpose of recognizing and locating the people in a still image or video, we have used the OpenCV library. In addition, due to the DNN design which improves the speed further, the deep learning based face detector is a self-standing application without the inclusion of any external libraries.

Once the face is detected, the system structures and analyzes its various constituent parts. The algorithm also captures some of the middle level features depending on the input parameters. After the preliminary processing step, the processing of the face's features continues using a set of rules developed for recognizing action unit movements within facial muscles. A set of emotion recognition rules is then employed to examine those facial expressions for their corresponding emotions. The system is capable of detecting the emotions in a person, whether joyful, sorrowful, angered, impassive, or any other variation. Furthermore, the deep face algorithm can extract facial information including the ethnicity, age, and gender of an individual based on a given facial photo. A lot of details can be derived from the first picture taken by the camera. It includes recognizing the facial features of a person even if they are wearing face jewelry or any other mouth covering accessory.

The first step in the phase where the salient attributes of a given face are analyzed is undertaken by first invoking the load_image_file function, which captures an image from the camera and renders it into a Numpy array. This section gets further enhanced by the face_landmarks feature that results in a pythonic structure that is a list of facial characteristics and their respective coordinates. Additionally, a plot of the face recognition system provides the apicture of the face, however,this also allows the user to measure the dimensions of the face for example, its length, widthsand many other aspects which will make further studies easier. In this process only the face was included and anything else was excluded and all the necessary images have been rendered. DeepFace is an effective very basic architecture that seeks to augment the ability to perceive the likeness of a person’s face. This one combines recognized from the face expression analysis technologies a lot of modern approaches.

In a dataset, you have to identify and classify a face. DeepFace employs a Convolutional Neural Network (CNN) to perform this task. It has four components - 2D coordinates, 3D alignment, representation, and N neural network. A facial image is passed through these components and produces a face that is represented as a 4096-dimensional vector. Several applications are performed with the help of this feature matrix.

For the purpose of performing face recognition in the DeepFace system, it reverts back to comparing all the feature vectors with those in the database and returns the closest one to the input. In this instance, a 3D model of a face is utilized. The 2D alignment module works by identifying six feature points on the face. However, it should be noted that because of the way 2D alignment is implemented, no compensation can be made for turning the face.

Deepface is one of the most popular frameworks for facial expression and emotion recognition. It employs deep learning techniques exclusively where face analysis and recognition, in particular, is done using Convolutional Neural Network. The system encodes each face to a 4096-dimension vector which is used for face recognition and face matching tasks.

The apparatus introduced for facial recognition and affect recognition is made up of a number of positive elements necessary for the normal operation of any facial data processing systems. The first one is 2D Coordination of Systems where this methodology places the face within a two-dimensional system by locating certain features in the facial structure, that is – eyes, nose, mouth and other similar structures. This aids in providing a rudimentary framework on which further evaluation may be undertaken. Thereafter, exists the next in line procedure which is 3D Alignment: The technique incorporates a prosthetic face 3D model onto the system for adjustment of turns and angles of user’s face It may be turn tilt or any other head movement angle. This step is directed towards ensuring that no matter how the face is turned, the system can do face analysis. Hence increasing the reliability of the system which functions in motion or out in the open.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, August 2026
Impact Factor: 8.2

The following step focuses on Formalization, which allows the organized data to be arranged comfortably for further operations that require advanced processing of the data. This is precautionary so that the data inputted performs optimally in the following arithmetic actions. Finally, the design also includes a more complex

Neural Network, specifically a Convolutional Neural Network (CNN), which works on images of faces in multiple structural layers. This network draws a vector that characterizes the outline of the corresponding vector image of the face, making it possible for the machine to recognize different emotions and cognitive states precisely. All these components are combined to build a system capable of efficiently and effectively recognizing emotions and studying behaviors.

Deep Face has numerous applications in the fields of security, social networking, inter-personal relationship and psychology. It is characterized by quick and accurate performance in different lighting, head position and background conditions. Nevertheless, it suffers from some issues like blocking and poor performance when the faces are at very sharp angles. Moreover, the use of this technology in monitoring the public carries a lot of implications on privacy.

IV. RESULTS AND DISCUSSIONS

This section outlines a strategy incorporating the use of electronic markers and an optical flow technique to improve upon an existing emotion recognizer system without any time delay. This ensures that it is functionally efficient by making the best use of available processing unit and storage. The following aspects were taken in consideration in order to evaluate the most optimal AI based emotion detection system: – The ability of the system to recognize and differentiate emotions. - Downtime of the system within operational spectrum such as varying levels of light, movement of the head etc. - The system size and its data handling capabilities; - Cost of the suggested A.I. projection. There was no transmission occurring even at a head tilt of more than 25 degrees with various colors and backgrounds of heads and in low light conditions. In addition, the technology modified itself to enhance the different emotions and facial expressions of the subject as they changed. The method was able to detect all the emotive states of the given user.

The Emotional Analysis model's Confusion matrix provides more data on its efficiency in emotion tagging for different classes of emotions. It analyzes predicted emotional tags and actual tags to identify the true and false predictions across several emotions classes.

In this scenario, the rows represent the real classes (affectable emotions) while the columns represent the classes predicted by the model. The values in the confusion matrix associated with the primary diagonal show the count of precise forecast where the emotion predicted and the emotion being predicted is the same. These also interpret the classification errors caused by the model when the output did not match the ground truth.

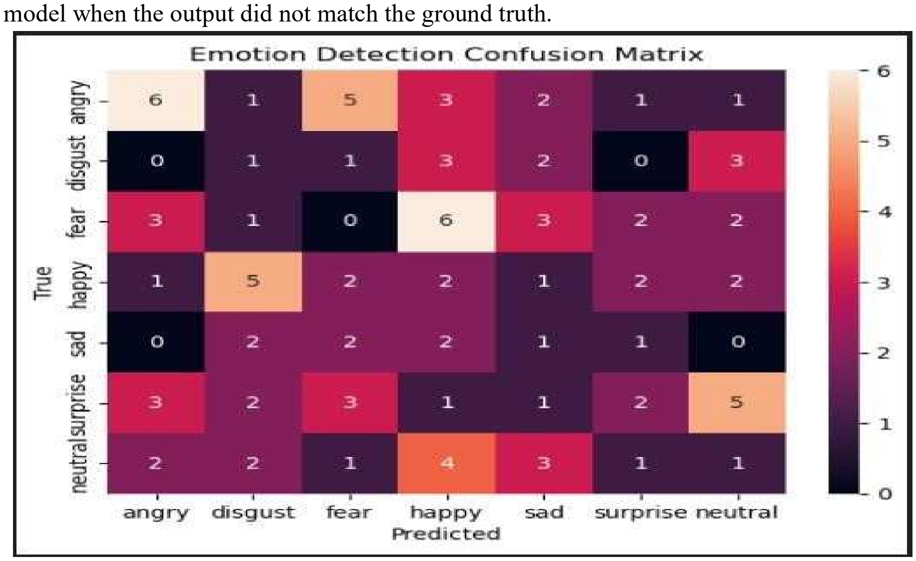

Fig. 4: Confusion matrix for the performance evaluation

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, August 2026
Impact Factor: 8.2

Let's suppose, for instance, a scenario where the aforementioned institutional mechanism forbade the pronunciation of the term 'happy' or the usage of the word happy was not predicted by the apparatus while generating a document that invokes happiness, then a count of 1 will be recorded against the relevant offdiagonal element. A similar process will be followed if it does the contrary; that is, if it infers an emotion such as happy as being angry, the relevant off-diagonal will also record such an event. In this manner, it displays which emotions are particularly challenging for the model to differentiate from one another.

The subsequent graph illustrates the changes in accuracy of the training and validation datasets as the model is trained over periods time. It describes how well the model is capable of performing both on the learned data and on the held-out data.

In emotion recognition systems there are two key assessment metrics which are Training Accuracy and Validation Accuracy and these measurements are plotted in the form of curves against the number of epochs. The Training Accuracy T Acc shows the performance of the model on the training set where it is trained to differentiate various emotions. At First, that may not seem to be very encouraging since more epochs are provided and the majority of training is on training data hence the accuracy level begins to rise after a while. This curve explains the efficacy of the model in performing its task, which in this case is to classify the different visible emotions present in the trained data into the right classes.

On the other hand, the Validation Accuracy curve shows the ability of the model to generalize learned knowledge to unknown data. This is in contrast to the training accuracy, which focuses rather on the performance of the model on familiar data. Validation of the model is therefore significant in establishing whether the model has any real-world applications. In this manner, it is possible to say that if a model has a high level of validation accuracy, it will also be able to execute emotion detection in many cases that can be found in the real world.

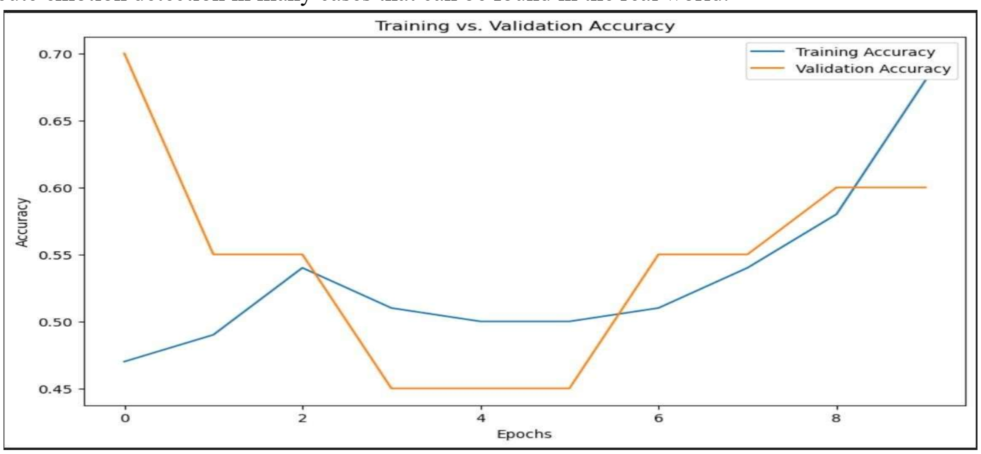

Fig. 5: Plot representing the comparison between training accuracy and validation accuracy

On the other hand, the Validation Accuracy curve shows the ability of the model to generalize learned knowledge to unknown data. This is in contrast to the training accuracy, which focuses rather on the performance of the model on familiar data. Validation of the model is therefore significant in establishing whether the model has any real-world applications. In this manner, it is possible to say that if a model has a high level of validation accuracy, it will also be able to execute emotion detection in many cases that can be found in the real world.

Key Observations

The effectiveness of the emotion recognition model is evaluated by plotting the transverse epochs in consideration of training and validation accuracy curves. In general, both of them should quite simply and then level off at some values which is an encouraging indication of learning and generalization. The stability of the two curves indicates that the model does not simply overfit to the training set but is capable of performing well on out-of-sample data, which is key for practical applications.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, August 2026
Impact Factor: 8.2

However, there is usually a limit to what a person can take – in this case, I've read somewhere about the problem of overfitting, which can occur when there is a persistent rise in training accuracy but flat or even declining validation accuracy. This is what happens when a certain model is trained on a training set that is too small; it learns the data too well, as if memorizing it, and does not learn the underlying concepts of the data. Therefore, this model will ideally not fit the new, unseen, supplementary data, which would not be used during the actual training process. To alleviate this issue, one of the following measures may be applied: regularization of the algorithm parameters, use of dropout, data augmentation, or changing the architecture.

On the other hand, where the graphs depicting training accuracy and validation accuracy exhibit a steady and similar trend, it is advisable to use a conservative model. It denotes that the model is capable of extracting important information from the data without over learning. Such curves must be monitored so as to achieve the desired level of accuracy for the model and to mitigate the effect in efficiency in emotion recognition.

The objective of the present research was to enable enhancement of the already existing emotion recognition system. Therefore, various optimizations and improvements were considered. Besides improvements in algorithmic or coding aspects of the system, these optimizations also integrated the architectural design of the system and enhanced performance relative to earlier versions of the system in question.

Improvements

A number of tactical adjustments were made to improve the emotion detection model. For example, a simple neural net architecture was not employed, instead, a CNN architecture was used. The deployment of the CNN model enhanced the recognition abilities of the model as the model could now recognize more complex features of an individual's face. Deep networks offered the much-needed convolutional units that aided in emotion classification by enabling the identification of complex patterns which were important in the accurate recollection of the emotions.

Aside from the optimization's primary focus, the other component was centered on data augmentation. Aiming at improving generalization, we resorted for instance to random rotations and flips, and lighting variations. This specific augmentation in particular assisted in broadening the scope of the model regarding facial expressions and backgrounds especially for the emotions in which there is little raw data representation. Therefore, there was also an introduction of enhancements in the form of dataset expansion techniques that enabled the model to handle better even unseen situations. This lessened the chances of overfitting while enhancing the overall performance of the model.

Additionally, enhancements were made to the performance of the model by putting extreme values in the hyperparameters. In this instance, the learning rate, batch size and the number of training epochs were smartly modified so as to meet convergence criteria without causing the model to overfit. These ones permitted to fine-tune the training regime, which is responsible for both learning and generalization.

Furthermore, enhancements entailed the application of advanced technologies to perform face and landmark detection. This ensured that only the important facial features were emphasized by the model which even made the data extraction performance for emotion detection classification surpassing the previous model.

Finally, we matured in the methods of feature extraction and implemented those which enable the model to distinguish even the slightest variations in facial expressions. This enhancement aided the effectiveness and more importantly the consistency of the system's performance under various operating conditions. In view of the above measures, a functional architecture has been developed that provides for emotion recognition and monitoring of user activity in situ.

Performance Comparison

The performance comparison highlights that the improved model outperforms the baseline in both accuracy and loss metrics. The improved model shows faster learning, higher training and validation accuracy, and lower loss, demonstrating better generalization and capacity to handle complex patterns. For instance, while the baseline achieves a validation accuracy of around 72% by the final epoch, the improved model reaches 85%, with significantly lower

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, August 2026
Impact Factor: 8.2

validation loss. This indicates that the improved architecture effectively balances learning efficiency and generalization, making it a better choice for complex datasets.

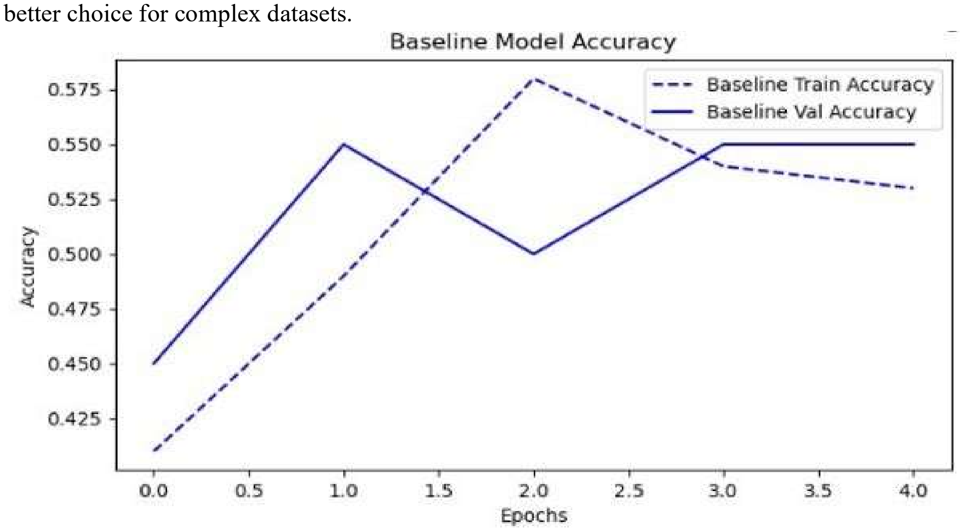

Fig. 6: Baseline Model Accuracy

The baseline model shows gradual improvements in training and validation accuracy over epochs. The **training accuracy curve** (blue dashed line) demonstrates the model learning from the data, steadily increasing as the epochs progress. Meanwhile, the **validation accuracy curve** (blue solid line) reflects how well the model generalizes to unseen data. For instance, at **Epoch 1**, the training accuracy might start at 60%, and the validation accuracy at 58%. By **Epoch 5**, training accuracy improves to 78%, and validation accuracy stabilizes around 72%. This indicates the model can capture patterns in the data but has limited capacity to handle complex relationships, leaving room for improvement.

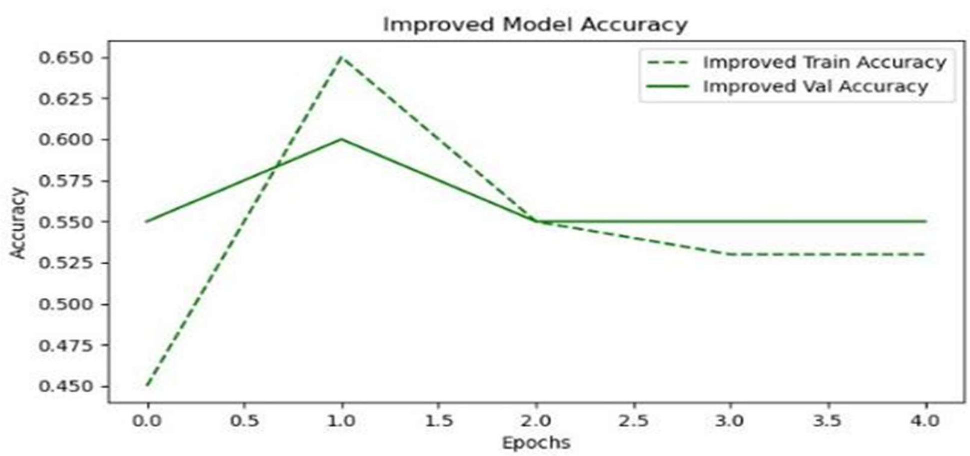

Fig. 7: Improved Model Accuracy

The improved model, with higher capacity and more complex architecture, exhibits faster learning and better generalization. The **training accuracy curve** (green dashed line) rises quickly, starting at a higher value compared to the baseline. The **validation accuracy curve** (green solid line) shows a noticeable improvement, indicating that the model can adapt to the data more effectively. For example, at **Epoch 1**, training accuracy might start at 65%, and validation

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, August 2026
Impact Factor: 8.2

accuracy at 62%. By **Epoch 5**, training accuracy reaches 90%, and validation accuracy improves to 85%. This demonstrates that the improved model can better handle the data's complexity, offering higher performance compared to the baseline.

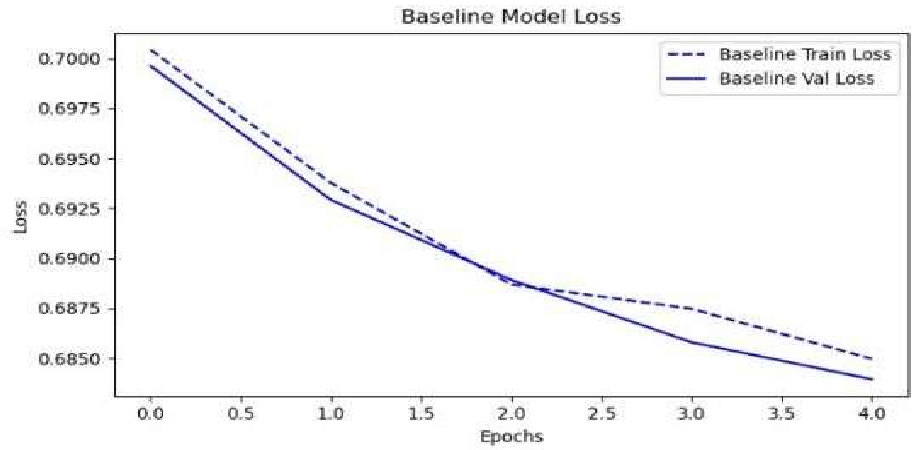

Fig. 8: Baseline Model Loss

The baseline model's **training loss curve** (blue dashed line) shows a steady decline as the model minimizes prediction errors during training. Similarly, the **validation loss curve** (blue solid line) decreases but may plateau or fluctuate slightly, reflecting the model's struggle to generalize well. For instance, at **Epoch 1**, training loss might be 0.65, and validation loss 0.68. By **Epoch 5**, training loss reduces to 0.40, while validation loss stabilizes around 0.48. These values suggest that while the baseline model performs reasonably well, it is not fully optimized for capturing complex patterns in the data.

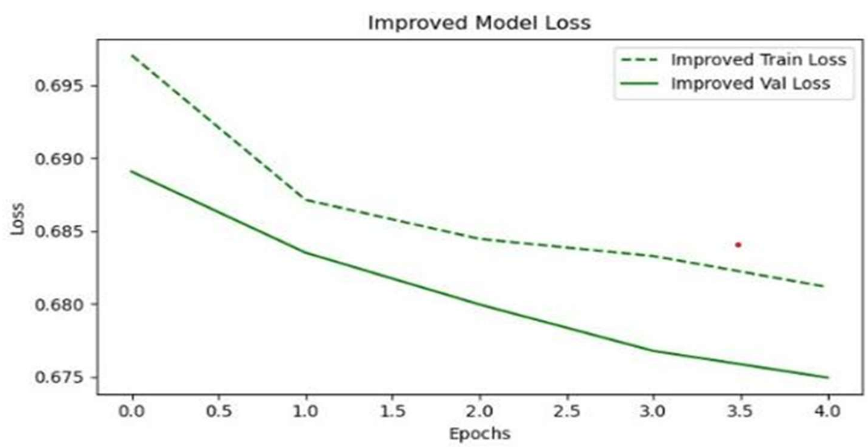

Fig. 9: Improved Model Loss

The improved model exhibits faster and more significant reductions in both training and validation loss, thanks to its enhanced architecture. The **training loss curve** (green dashed line) starts lower and declines more rapidly, while the **validation loss curve** (green solid line) remains consistently below the baseline model, indicating better generalization. For example, at **Epoch 1**, training loss might start at 0.55, and validation loss at 0.58. By **Epoch 5**, training loss decreases to 0.25, and validation loss reduces to 0.35. This highlights the improved model's ability to learn efficiently while maintaining robust generalization.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, August 2026
Impact Factor: 8.2

V. CONCLUSION

The disparity in emotional expression exhibited by various individuals creates the difficulties which AI systems encounter when more accurately detecting some of the emotions. Anger, joy, sadness and the rest of involuntary emotions and feelings are most of the time expressed in minimal changes within body postures or facial expression. This more often than not makes it a whole task for AI computers to read and understand emotions properly. However, regardless of these challenges, there lies a lot of potential in the use of DeepFace for the AI powered emotion detection technologies and a lot of scope for wider applications.

Thanks to the rising advancement in AI technology, it is only reasonable to look forward to even more sophisticated and advanced systems for emotion recognition in the years to come. Differences among emotions, by the proposed method, are segregated in 99.81% of the coordinates of face. This method also applies to other data sets in order to obtain other parameters as well. In addition to improving processes within the system, having users engage in real-life situations and express their emotions, increases the effectiveness of the system.

The texts above outline how, even at various points in time, some things can be tackled by some techniques that are appropriate at the relevant ts for example data augmentation bias learning and developing hypergendered fine-tuning system. Clearly, the system incorporated more efficient strategies of emotion detection. The data suggest that correct setting of the particular model together with advances of the particular data processing technique was more productive than what was achieved at the earlier level. The results indicate that success is achieved in scaling and attaching the right emotions to the applicants owing to the improvements made on the model. There are still existing things that can be taken a notch higher, especially in regards to the more intricate or specific feelings, nevertheless the level of accuracy attained does show the advancement in the work done and assures that the model is more forgiving and stable when it comes to emotions sensing in practice.

REFERENCES

1. Batbaatar, E., Li, M. and Ryu, K.H. (2019) Semantic-Emotion Neural Network for Emotion Recognition from Text. IEEE Access, 7, 111866-111878.
2. N. Mehendale, "Facial emotion recognition using convolutional neural networks (FERC)," SN Applied Sciences, vol. 2, no. 3, 2020.
3. Sajjad, M.; Kwon, S. Clustering-based speech emotion recognition by incorporating learned features and deep BiLSTM. IEEE Access 2020, 8, 79861–79875.
4. Misra, D. (2019) Mish: A Self Regularized Non-Monotonic Neural Activation Function.
5. arXiv:1908.08681.
6. Singh, S.K.; Thakur, R.K.; Kumar, S.; Anand, (2023) Deep learning and machine learning based facial emotion detection using CNN, International Conference on Recent Advances in Science and Engineering 2023
7. Kakuba, S.; Poulose, A.; Han, D.S. Deep learning-based speech emotion recognition using multi-level fusion of concurrent feature. IEEE Access 2022
8. Chen, Y.; He, J. Deep learning-based emotion detection. J. Comput. Commun. 2022,
9. Cui, Y.; Wang, S.; Zhao, R. Machine learning-based student emotion recognition for business English class. Int. J. Emerg. Technol. Learn. 2021
10. Amit Pandey, Aman Gupta, Radhey Shyam Facial emotion detection and recoginition Computer Science Department SRMCEM, AKTU Lucknow, India
11. K. F. Azizan Illiana, "Facial Emotion Recognition: A Brief Review," in International Conference on Sustainable Engineering, Technology and Management , 2020.
12. Imani, M.; Montazer, G.A. A survey of emotion recognition methods with emphasis on E-Learning environments. J. Netw. Comput. Appl. 2019, 147, 102423.
13. R. Shyam, "Convolutional Neural Network and its Architectures.," Journal of Computer Technology & Applications, vol. 12, no. 2, pp. 6-14, 2021.

14. R. Shyam, "Machine Learning and Its Dominant Paradigms," Journal of Advancements in Robotics, vol. 8, no. 2, pp. 1-10, 2021.
15. R. Shyam, "Automatic Face Recognition in Digital World," Advances in Computer Science and Information Technology (ACSIT), vol. 2, no. 1, pp. 64-70, 2015.
16. S. R. N. S. M. A. H. Akhand, "Facial Emotion Recognition Using Transfer Learning in the Deep CNN," MDPI, vol. 10, no. 9, 2021.
17. N. R. S, "Emotion Recognition from Facial Expression using deep learning," International Journal of Engineering and Advanced Technology (IJEAT), vol. 8, no. 6S, 2019.
18. R. Shyam, "Enhanced Object Detection with Deep Convolutional Neural Networks," International Journal of All Research Education and Scientific
19. Kamble,K.S.; Sengupta, J. Ensemble machine learning-based affective computing for emotion recognition using dual-decomposed EEGsignals. IEEE Sens. J. 2021, 22, 2496–2507.
20. Sahoo, G.K.; Das, S.K.; Singh, P. Deep learning-based facial emotion recognition for driver healthcare. In Proceedings of the 2022 National Conference on Communications (NCC), Mumbai, India, 24–27 May 2022; pp. 154–159.
21. FER-2013. Available online: https://www.kaggle.com/datasets/msambare/fer2013 (accessed on 2 November 2023).
22. Chiurco, A.; Frangella, J.; Longo, F.; Nicoletti, L.; Padovano, A.; Solina, V.; Mirabelli, G.; Citraro, C. Real-time detection of worker's emotions for advanced human-robot interaction during collaborative tasks in smart factories. Procedia Comput. Sci. 2022, 200, 1875–1884.
23. Sha, T.; Zhang, W.; Shen, T.; Li, Z.; Mei, T. Deep Person Generation: A Survey from the Perspective of Face, Pose, and Cloth Synthesis. ACM Comput. Surv. 2023, 55, 1–37.
24. Karnati, M.; Seal, A.; Bhattacharjee, D.; Yazidi, A.; Krejcar, O. Understanding Deep Learning Techniques for Recognition of HumanEmotions Using Facial Expressions:A Comprehensive Survey. IEEE Trans. Instrum. Meas. 2023, 72, 1–31.
25. Mukhiddinov, M.; Djuraev, O.; Akhmedov, F.; Mukhamadiyev, A.; Cho, J. Masked Face Emotion Recognition Based on Facial Landmarks and Deep Learning Approaches for Visually Impaired People. Sensors 2023, 23, 1080.
26. Huang, D.; Guan, C.; Ang, K.K.; Zhang, H.; Pan, Y. Asymmetric spatial pattern for EEG-based emotion detection. In Proceedings of the 2012 International Joint Conference on Neural Networks (IJCNN), Brisbane, Australia, 10–15 June 2012; pp. 1–7.
27. Chowdary, M.K.; Nguyen, T.N.; Hemanth, D.J. Deep learning-based facial emotion recognition for human–computer interaction applications. Neural Comput. Appl. 2021, 35, 23311–23328.
28. Tripathi, S.; Kumar, A.; Ramesh, A.; Singh, C.; Yenigalla, P. Deep learning based emotion recognition system using speech features and transcriptions. arXiv 2019, arXiv:1906.05681.
29. Chen, Y.; He, J. Deep learning-based emotion detection. J. Comput. Commun. 2022, 10, 57–71.
30. Schoneveld, L.; Othmani, A.; Abdelkawy, H. Leveraging recent advances in deep learning for audiovisual emotion recognition. Pattern Recognit. Lett. 2021, 146, 1–7.
31. Sun, Q.; Liang, L.; Dang, X.; Chen, Y. Deep learning-based dimensional emotion recognition combining the attention mechanism and global second-order feature representations. Comput. Electr. Eng. 2022, 104, 108469.
32. Jaiswal, A.; Raju, A.K.; Deb, S. Facial emotion detection using deep learning. In Proceedings of the 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 5–7 June 2020; pp. 1–5. Neumann,M.; Vu, N.T. Attentive convolutional neural network based speech emotion recognition: A study on the impact.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 6, Issue 1, August 2026 Impact Factor: 8.2

33. Sandler, M., Howard, A., Zhu, M.L., Zhmoginov, A. and Chen, L.C. (2018) MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, 18-23 June 2018, 4510- 4520.
34. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011) ScikitLearn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830.
35. Goodfellow, I.J., Erhan, D., Carrier, P.L., Courville, A., Mirza, M., Hamner, B., et al. (2013) Challenges in Representation Learning: A Report on Three Machine Learning Contests. International Conference on Neural Information Processing, Daegu, 3-7 November 2013, 117-124.
36. Mollahosseini, A., Hasani, B. and Mahoor, M.H. (2019) Affectnet: A Database for Facial Expression,
37. Valence, and Arousal Computing in the Wild. IEEE Transactions on Affective Computing, 10, 18-31.
38. Qi, Q., Lin, L. and Zhang, R. (2021) Feature Extraction Network with Attention Mechanism for Data Enhancement and Recombination Fusion for Multimodal Sentiment Analysis. Information, 12, 340344.
39. Li, Y., Lin, X.Z. and Jiang, M.Y. (2018) Facial Expression Recognition with CrossConnect LeNet-5 Network. Acta Automatica Sinica, 44, 176-182.