FACE BASED EMOTION DETECTION SYSTEM USING HAAR-CASCADE AND CONVOLUTIONAL NEURAL NETWORK

Deepak Kumar Verma, Department of Computer Science and Engineering, University Institute of Engineering and Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur-208024, India. deepak300572@Gmail.com
Vibhash Yadav, Department of Information Technology, Rajkiya Engineering College Banda, India. vibhashds10@recbanda.ac.in
Alok Kumar, Department of Computer Science and Engineering,, University Institute of Engineering and Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur-208024, India. akumar.uiet@gmail.com

Abstract:

Computer vision has a tough time detecting human facial emotions. Emotion may now be discerned from a video or photograph using computer vision and machine learning. In this research, we will describe a technique for identifying facial emotion that makes use of the Haar-Cascade Classifier and convolutional neural networks. Seven distinct facial expressions were employed in the experiment based on data from FER2013, the facial expression recognition dataset. The epoch variety improves the CNN model's MSE and accuracy scores. Increased epoch value decreased MSE whereas increased accuracy increased MSE. These are the conclusions reached after doing the experiment. Consequently, CNN's proposed face emotion recognition algorithm has been shown to be effective. Researchers in this study employed Haar-Cascade, a classifier approach, and Convolution Neural Network, a deep learning method, for this study's research. Face identification uses the Haar feature-based Cascade classifier, which is an outstanding technique for object recognition. As with the CNN, the deep neural network has a vast network depth and employs the same algorithmic process as the latter. Conventional multilayer perceptron (MLP) neurons will instead be implemented in a two-dimensional form, rather than a three-dimensional shape, for the CNN model. CNN is only able to analyse two-dimensional input, such as images and audio recordings. Many layers of learning are used by CNN to interpret an input and extract a feature based on its characteristics. Using vectors of numbers, each layer is represented numerically. A convolution layer plus a pooling layer makes up this feature extraction layer. There will be a convolution layer where each neuron calculates its own weight, and the input volume will be linked to the input volume until all of the images are linked to the input volume, regardless of their size.

Keywords: Human computer Interaction, CNN, AdaBoost , Face detection, Haar feature, Emotion, Neural Network.

1. Introduction

Computer vision is a vast field within computer science that involves the analysis, comprehension, and extraction of features and information from visual data such as digital images, videos, and other forms of digital imagery[1]. There are endless application areas, fields and implementation problems sets of computer vision namely, using computer vision as a tool for image acquisition, image processing, image analytics, image recognition, object recognition, object detection, expression detection, expression identification, image segmentation, video analytics, image procession, image denoising, image reconstruction, image restoration, video reconstruction, video deconstruction, and other similar sub division in each of the mentioned domains. Face expression detection being a very cogent and effective domain of the mentioned.

[2]The domain of facial expression detection is very wide, variegated, diverse and complex due to the multiple variations of human expression and its multiple ways of analysis like facial expression analysis or emotion detection via facial expression, voice tone, voice detection or textual analysis. But in-spite of the various complexities in both acquisition and processing, the domain of computer vision has made multiple algorithms, frameworks and architectural developments to provide aid in the phenomenon of facial expression detection and emotion identification. Multiple implementations have been made to provide solutions to the various issues that occur or provide aids in solving the various road blocks in the process of facial expression detection and emotion analysis. This has lead to automatic systems and expert frameworks for computerised facial expression detection and emotion detection with endless application is multiple areas like media, entertainment, news, security, testing, education, traffic monitoring, task monitoring, content analysis of video and imagery data, healthcare,

image and video analytics, and cyber security among others. For instance, subject monitoring can be implemented in traffic or human productivity analysis to analyse the expression to determine the attentiveness in the subject of order.
But, since there are multiple issues that occur due to various complexities in the process of emotion detection and facial recognition to due human errors, image acquisition issues, image processing road blocks, feature extraction problems, segmentation complexities among other real time facial expression detection or emotion identification problems. Over the last few decades, multiple techniques have been implemented but the most cogent and practically efficient technique that became a goal post for most of the future face detection or emotion identification is one proposed by Viola and Jones. AdaBost, one of the many popular face detection techniques used is implemented very vividly by above mentioned algorithm. The technique or facial expression framework algorithm implemented by this study aims the application of a variant of Ada Bost learning algorithm but another technique is amalgamated to provide optimal performance namely Haar feature. This helps in easy, accurate and fast identification of facial features thereby contributing to the cogent identification of facial expressions, in turn leading to segmentation and accurate identification of emotions. Multiple algorithms, techniques, frameworks and architecture along with various experts systems have been discovered and implemented in the last decade as the field of computer science and computer vision has expanded exponentially to provide accurate facial expression and emotion detection for ever-changing real world human emotions and expressions. One of the few latest developments in this area being developments of human computer interaction know as Brain computer interface (BCI)- analysing and dissecting human emotions, implementing complex concepts of Convolutional Neural Network (CNN) and Deep learning (DL) [3].
Although, these newly introduced techniques and methodologies often outperform the primordial techniques but there are some methodologies that have a performance matrix higher than that of others like Deep Learning (DL). Due to the ability of Deep Learning (DL) techniques, algorithms and frameworks to handle complex and plethora of data sets each with multiple sub sets, mining abilities to perform quick analysis, providing optimal visualization results and capability of handling huge calculations. These points, come in handy when performing facial expression recognition or emotion identification, the process being multi-step. For instance, a classical methodology of facial expression recognition or emotion detection has multiple steps namely- image acquisition, image pre-processing, features extraction, feature segmentation, feature detection, emotion classification and result visualization. This is conducted with the help of neural network where, multiple neural are connected in a network kinda structure such that input data is processed via multiple layers of neural network and then computations are made to provide final output by the end of network of neurons.
Theoretically, this kind of processing as in via network of neurons known as Deep Learning (DL) comes from Machine Learning (ML), but is very different both in terms of abilities, implementations, framework and results. Deep Learning (DL) is more on the complex and dense side, having the ability to handle multiple complex and huge datasets at once. This is done by layering multiple networks of neurons or artificial neurons in numerous iterations to process the input data. These neurons try to emulate or replicate the neurons of the human brain as processing units in the expert system [4]. Below section elaborates on the use of these techniques, methodologies and framework as part of classical facial expression detection and emotion identification in step wise manner.

2. Input Image used in Methodology

The result or aim of facial expression recognition or emotion detection is to comprehend human reaction and contemplate human emotion for further analysis, comprehension, segmentation, prediction, and dissect the emotions of subject in study. Human emotions are very wide and variegates and so are their sources like video, offline images, digital images, voice recognition and textual segmentation, body moments, alternate biological and physical signals among others. But this study uses the most cogent and prominent method of facial expression recognition and emotion detection via images. This stud has divided the varied range of human emotion into three major gradations namely- Happy, Sad and Angry, represented with great precise in the figure 1 below

Figure 1: Input Images for Facial Expressions.

This research proposes a complete and cogent technique for emotion recognition by implementing it on a annotated dataset with marked areas of interest also known as regions of interests (ROI). Once local image patches and attributes are identified, weight parameters are then determined by analysis of various image classification parameters. The dataset used is digitalized film mammography database with roughly thousand image such is it smaller than a FFDM (full-field digital mammography) imagery data.

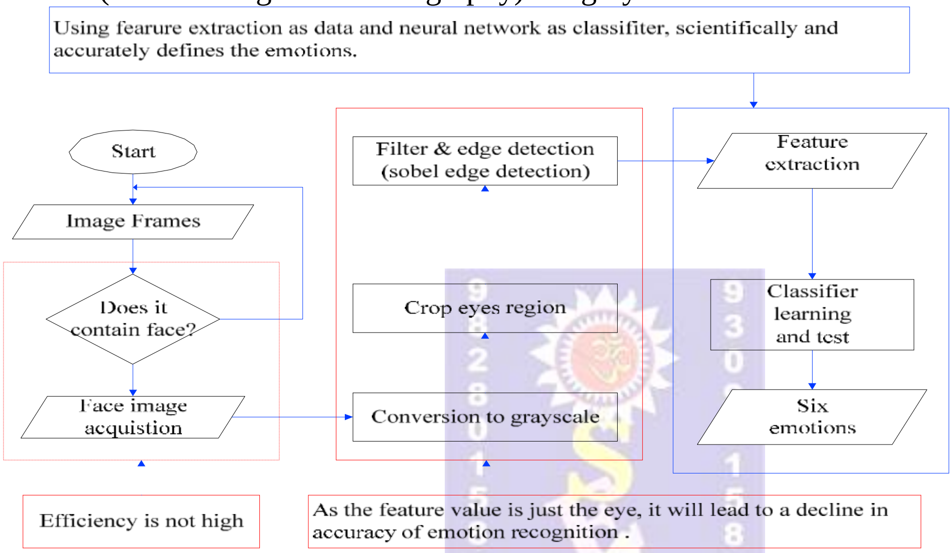


Figure 2- Facial Emotion Recognition (FCR) classical framework

The concept, application and implementation of Facial Emotion Recognition (FCR) is a milestone prominent factor in the domain of computer vision having endless applications and areas of expansion from both research and commercial implementation point of view as visualized in figure 2 above. Visual expressions lead to emotion detection and thereby providing the expert systems the ability gauge human emotion and cognition, thus becoming a very powerful tool. This research based study aims to provide a comprehensive review of multiple developments, techniques, methodologies, in the domain of FCR.

Table 1 : Captured data with their location

Data Set	Location	Description
MIT Database [163]	ftp://whitechapel.media.mit.edu/pub/images/	Faces of 16 people, 27 of each person under various illumination conditions, scale and head orientation.
FERET Database [115]	http://www.nist.gov/humanid/feret	A large collection of male and female faces. Each image contains a single person with certain expression.
UMIST Database [56]	http://images.ee.umist.ac.uk/danny/ database.html	564 images of 20 subjects. Each subject covers a range of poses from profile to frontal views.
University of Bern Database	ftp://iamftp.unibe.ch/pub/Images/FaceImages/	300 frontal face images of 30 people (10 images per person) and 150 profile face images (5 images per person).
Yale Database [7]	http://cvc.yale.edu	Face images with expressions, glasses under different illumination conditions.
AT&T (Olivetti) Database [136]	http://www.uk.research.att.com	40 subjects, 10 images per subject.
Harvard Database [57]	ftp://ftp.hrl.harvard.edu/pub/faces/	Cropped, masked face images under a wide range of lighting conditions.
M2VTS Database [116]	http://poseidon.csd.auth.gr/M2VTS/index.html	A multimodal database containing various image sequences.
Purdue AR Database [96]	http://rvl1.ecn.purdue.edu/~aleix/aleix_face _DB.html	3,276 face images with different facial expressions and occlusions under different illuminations.

The table-1 provides an elicit description various datasets that explore complexities of a facial expression detection or emotion recognition framework. These examples, aid in understanding the wide scope of computer vision and its application in facial expression recognition or emotion detection.

3. Terminology in Methodology

The classical emotion recognition using facial expressions consist of three gradations- Collating the imagery data, extracting appropriate attributes from acquired data and finally, using image classification techniques to categorize the processed images into different emotion labels. Proposed research paper elaborated and tries to innovate each of the gradation to obtain optimal performance on grounds of quality, space-time complexity and computation requirement, all while keeping in mid the real-world application of the proposed technique. The results obtained are then real time facial virtual environment (VLE) [6] for ease of display and test of applicability.

3.1 Convolution Neural Networks (CNN)

In the past few decades, there have been path breaking development in the domain of computer vision. This has led to expansion of various application and implementation areas namely- object detection, object recognition, facial expression recognition, emotion detection, video analytics, image processing, image denoising, feature extraction, feature classification among others [5]. Among these techniques, is the one of the revolutionary method known as Convolutional Neural Network (CNN), which aims to emulate the functionalities of human brain's neuron to perform multiple layers of neural networks. Such that each layer has unique framework of neurons, for processing of multiple image features and characteristics. The processing of multiple layers depend upon the number of neurons, arrangements of these respective neurons, edges to be processed among other peculiarities.

Various CNN models perform multiple tasks of image processing, in a two step format namely- training the model and performing the test on the same. Multiple image feature identification filters, feature detection, segmentation pre-processing algorithms are used. Various CNN models differ based on their variations in multiple layers and filters create numerous variations of CNN as elaborated in figure 2

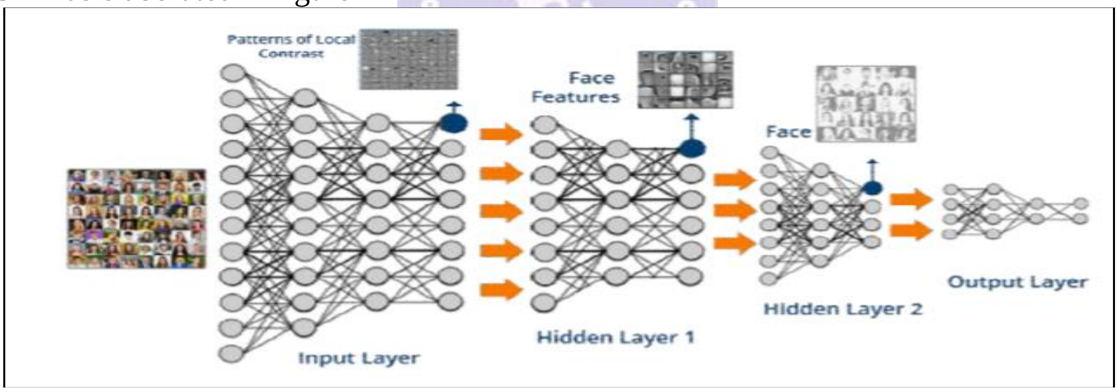


Figure 2: CNN model process based on the input image.

In the above figure 2, basic layers of CNN model are explained such that, each layer performs individualistic functions on the input images. First layers acts as an inlet for input image and deliberates the same, second layers uses functions and segmentation filters to perform such that features are all categorized in their specific gradations in the further layers, third layers does the categorization via the grid also know as image matrix with predefines custom dimensions and volume given as as (h x w x d) and the filter as (fh x fw x d) as explained in figure 3

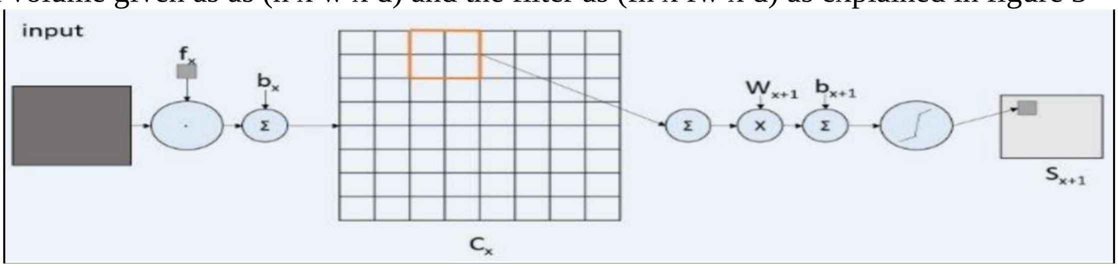


Figure 3: Process of Convolution Neural Network (CNN)

There are some critical component of Convolutional Neural Network Framework include local field of perception, weight distribution, sub-sampling of various features and filters, function redistribution based on scale of training parameters. These multiple steps provide an innate advantage to Convolution Neural Network methods like clean layer and data segmentation, classification of input images, clear tagging and categorization of images and layers provide insights to one another. The implementation of subsampling provides some added advantages like stability, size, speed, accuracy, flexible network topology, multiple variations can be created via minor alterations. Since each later has its own processing units and insights, errors and modifications can be easily passed on from one iteration to another and provide easy and quick modification- an overall agility and flexibility. Considering the input image and its gray value of 96*96 in the pre-processing gradation, creates 32 on calculation as illustrated in figure 4.

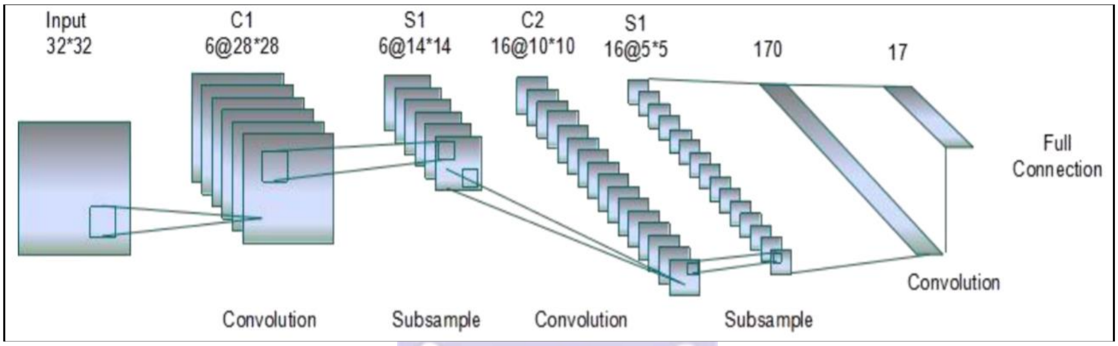


Figure 4: Stages of Convolution Neural Network (CNN) model

3.2 Haar Features

There are approximately 2 or 3 rectangles in each of Haar features. In the preliminary stage, input image are deliberated and scanned to find the Haar features. These features post identification are collated and deliberated to categorize into multiple categories using machine learning rules from AdaBoost [6]. The weights generated are then classifies into the rules of feature layers as illustrated in figure 5 below.

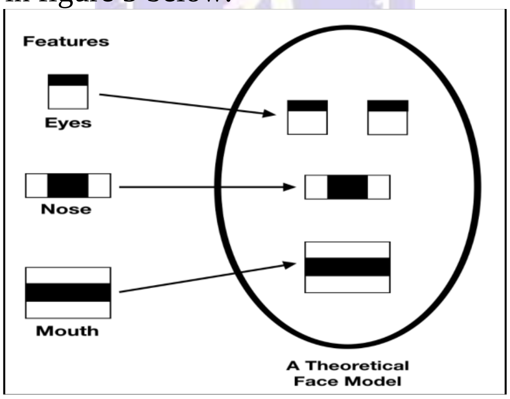


Figure 5: Haar features with their illustrations are elaborated upon above.

Every rectangle in figure 6 represents a Haar feature, when multiplied by its respective heights, provides the results summing up. The area calculated, for each rectangle is via the integral image area. A rectangular area has approximately four co-ordinates, each co-ordinate represents a different data point, such that the area calculated is the area of pixels enclosed between these data points- in whole representing the haar features. The area of the rectangle R, denoted because the rectangle integral, will be computed as follows using the locations of the integral image: L4-L3-L2+L1 (refer figure 6).

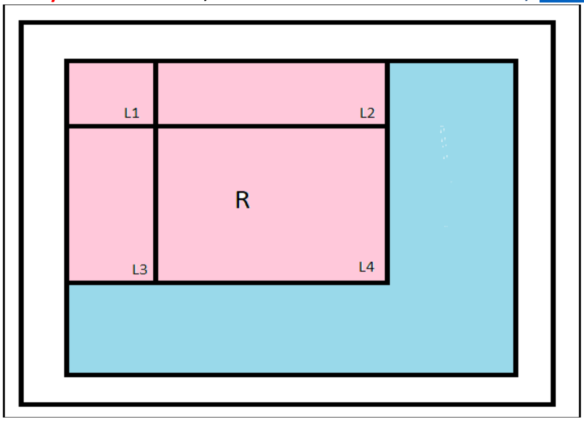


Figure 6: Calculating the area of the rectangle R, using the locations of the integral image: L4-L3-L2+L1.

One of the critical factors that determine the performance of an emotion detection or facial expression recognition algorithm is to find a taxonomic reference or in simpler words identify the predefined emotion labels to tag the identified expressions into, like tagging features in ore-determined categories. In terms of facial recognition or emotion detection, there are seven target emotion categories that are classically used to classy human emotions. This classical model, presents a cogent case in terms of scientific accuracy and model performance, making it a standard facial expression encoding system. This facial expression encoding system is known as Facial Action Coding System (FACS)- accepted widely by scientists from the domain of computer vision and from the area of psychology. These 7 categories are classified based on the visual difference visible to human eye like raising an eye brow, lines on the fore-head, pouting and pointing using lower lip, widening of the eyes, stressing on the nostrils and other supplementary information on various facials features as action points of differentiations. These facial features or facial action points Facial Action Coding System (FACS)- elaborate and elucidate the various facial moments that are visible to human eye and human cognition.

3.2.1 Haar Feature Classifier

The Haar feature rectangle also known as Haar feature classifier, is used extensively to detect Haar features and categorize them multiple classifications. Once features are classified and their respective weights are assigned, each feature is multiplied by their weights to and the results are added. The training data is manipulated and arranges according the weighted averages obtained from Haar features, which then emulated in the validation and test data set arrangements.

3.2.2 Cascade

The algorithm of facial expression detection or emotion detection proposed by Viola and Jones, aims to provide optimal performance by eliminating the facial expressions as soon as possible via quick feature classification by the cascade of processing layers. Its comparatively easier to handle multiple features and their respective details than classification of complete facial expressions in real time. In this algorithm of facial expression detection or emotion identification, each cascade has multiple parameters, these parameters eliminate various features as the cascade progresses. Via this, post multiple cascades and various iterations of elimination rounds on features. A facial expression is finally classified enough to be detected as the final cascade approaches, as elaborated in figure 7.

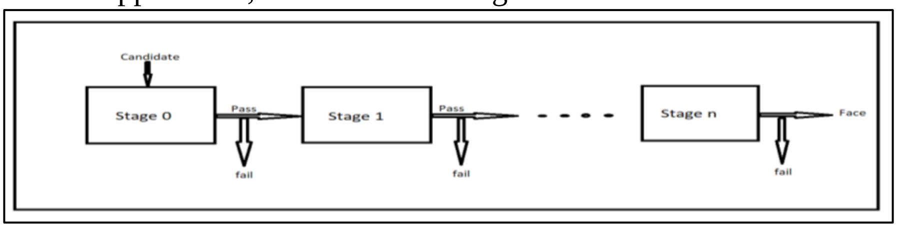


Figure 7: Cascade of stages

New points of reference are then planned into the same space and expected to have a position contingent upon which side of the opening they fall on with a class. KNN Inc. shows incredible request precision despite when simply an unpretentious proportion of planning data is open, making them particularly proper to a dynamic, astute approach to manage verbalization affirmation. Reinforce vector machines have as of late been successfully used in an arrangement of different applications, including character and content affirmation, similarly as DNA microarray data assessment.

4. Literature Survey

Image processing and computer vision has progressed exponentially in past few decades. Computer science encompasses a broad area known as computer vision, which involves the processing of visual data such as images, videos, and other forms of digital imagery to extract features and information, as well as comprehend and analyse the data[1]. There are endless application areas, fields and implementation problems sets of computer vision namely, using computer vision as a tool for image acquisition, image processing, image analytics, image recognition, object recognition, object detection, expression detection, expression identification, image segmentation, video analytics, image procession, image denoising, image reconstruction, image restoration, video reconstruction, video deconstruction, and other similar sub division in each of the mentioned domains. This section elaborates upon the latest and recent path breaking developments in the very popular branch of computer vision- Facial expression detection and human emotion identification via imagery data. Human expression gives insight to their interpersonal behaviour, mood, cognitive abilities, and aids in predicting their future actions. Applications of human emotion identification as almost endless, for instance it can be implemented and benefited from in the domain of robotics, brain computer interface, human monitoring, action prediction, administration, productivity monitoring, medicine, accident prevention, traffic surveillance, guided driving among others. The classical model of facial expression detection or emotion identification can be classified into following categories- angry, fear, happy, sad, disgust, neutral, and surprise.

The concept of Convolutional Neural Network (CNN) is widely implemented in the domain of computer vision, given its cogent performance multiple researchers have applied it in the field of facial expression detection too. The application of Convolutional Neural Network (CNN) is generally done using two steps the first step consists of removing the external background imagery data along with imge denoising and pre-processing. Post this, second gradation includes- vector extraction from the obtained image. [3]A couple of exploration works in the remedial picture blend district have started and finished throughout the last decade. Thus, various intelligent journals were circulated. We get some information about their basic responsibilities.

While there are other equally rewarding applications of facial expression identification, it was Human Computer Interface (HCI) that required level of accuracy and had the scope for path breaking results. The human facial expression for acknowledgement or apprehension was communicated to the computer interface for further interaction- forming the base for most current robotic frameworks focused on human computer interaction [4]. In a manner similar to the prior study, the current one makes use of FER. Two distinct and complex wavelets are employed as capacity extractors: transformed Gabor wavelets (GWT) and transformed dual-tree wavelets (DT-CWT) complexes (GWT). PCA or Binary Pattern Analysis may be used to take into consideration the different prospects that exist (LBP). There are a wide range of data sources that were used, including JAFFE, CK, and MUFE. K-next neighbour (KNN), neural networks (NN), and support vector machines (SVM) are some of the methodologies and frameworks that have been employed (SVM). They made ground-breaking discoveries in their profession by detecting and identifying facial expressions in real time.

[8] It is described as the reveal of a particular image inside a set of picture designs, and it states this: Artificial intelligence, machine learning, data mining, and computer vision are all part of the current drive. Face affirmation makes it possible to examine designs produced from images of the human face and analyse them... Face mining is the automated confirmation of a person's appearance based on their features, such as their eyebrows, lips, and other facial features. Human face images taken lately may be separated from their credits using evaluations based on the most current FPD and GLCM computations. FPD uses the standard of the hopping edge,

whereas GLCM uses the invariant of proclivity in its computations. Limits on capabilities have been set for precision and extraction time. MATLAB 7.0 is used to do the analysis. The recommended c technique reduced the constraints more definitively than the FPD calculation, according to the results of the experiments.
[13] With regards to gauging public activity, mental cycles, and curiosity, a single look may be a great social indicator. Recent years have seen an increase in the use of visual verification in evaluations. Some of these applications integrate human-computer interfaces with the evaluation of human understanding, the treatment of clinical patients, and other relevant fields. In this research, we look at the PCA and LDA presentation to discern seven different emotions of two people in the JAFFE (Japanese Female Face Expression) information collection, including delighted, abandoned, fair-minded, disturbed, angered, fear, and shock. Face affirmation is likely to be one of the topics we cover. Search may be used for acceptable goals, such as monitoring and human-computer cooperation, according to this assessment.
In light of the wide range of human emotions and the frequency with which they change in reaction to a person's actions, the task of identifying and recognising facial expressions and identifying human emotions takes a long time. Changes in the object's position and orientation are also constant. In order to detect and identify facial expressions and emotions, this study makes use of the concept of principle determination. DCT (digital face tracking), PCA+DWT, and LDA+DWT were used to analyse performance based on RMS error, CC, PSNR, and MSE.
[5] Human facial arrangements pass on essentially more ostensibly than articulately nuances. Acknowledgment of facial sign expects an imperative part in passing on among people and machines. It's everything except a problematic task to recognize looks by machine regardless high acknowledgment levels. Look typically finished in three-stage acknowledgment comprising of facial location, brand name extraction, and talk grouping. This paper analyzes the most current investigation on the facial development acknowledgment procedures with changed facial recognition methods, extraction and order strategies, and their results.

[14]Another study that's given cogent results include implementation of Global Wavelet Transformation to identify facial expression and comprehend human emotions. Vector dimensions are limited to the utilization of the PCA and LBP algorithms. Data set used for finding the cogency of proposed method is JAFFE data. Results from all analyses did utilizing the JAFFE information base show that in this paper, GW+LBP outperformed different strategies in a similar trial setting with a general identification pace of 90 %.

4. Related Methods and Scope

As in many facets of human life, emotions play a critical part in education. No matter where people come from or what they do for a living, feelings that are considered universal exist in all people, regardless of their background or language. It is projected that facial-expression recognition will be employed in the travel sector, security frameworks, master card checks, recognised criminal pieces of evidence, and video chat in the future as an instrument for measuring consumer loyalty. AI will be used as part of the project's investigation into how visitors' feelings and their contentment with the nature of the administration are measured while on a guided tour at a government heritage site during a guided tour. In the approach, all of the steps essential to finish the procedure are included. Emotional responses to a historic site are analysed using facial-expression recognition technology in a commercial promoting a historic location. In addition, guests were asked to anonymously rate their level of satisfaction with the guided tour. Finally, a key condition demonstrating approach is employed to demonstrate the substantial link between feelings and fulfilment. We may draw this conclusion based on our findings: facial-expression recognition data can be used to estimate client loyalty just as well as self-directed questionnaires.

4.1 Facial acquisition

To detect whether or not a face appears in a picture, this method uses a set of parameters to identify the location and size of the face. In part as a result of the range of approaches available, face localisation has evolved into a unique subfield in computer graphics. The face edge may be used to finish the facial localization process [6]. Facial localisation also often makes use of information about the texture of the face, but [7] was the first to utilise facial grayscale for this purpose. When it comes to face localisation, skin colour information is often used since it occupies a reasonably open position in the colour space [8].

4.2 Facial feature extraction

The classification algorithm and the application environment have the biggest influence on how face characteristics are extracted. For instance, the attributes of various classification techniques will vary, and the classification methods will be used in a variety of scenarios. [9] based his judgement on the 153 potential critical elements and the gap between contenders.

4.3 Emotion expression classification

Most important in classifier development are the utilisation of temporal data and which classifiers are used. Even though it does not employ temporal data, the classification strategy is nevertheless referred to as a spatial domain method. It is possible to use artificial neural networks (ANNs) as a kind of spatial approach. The whole image is used as input to a neural network or as output of image processing methods such as image Gabor filtering in image processing techniques such as picture Gabor filtering or feature illustration techniques such as PCA and ICA. [10]. The feature vector space method allows for the creation of a spatial approach utilising the general classification technique. Face expressions in a video clip were recognised using principal component analysis [11]. A support vector machine (SVM) was used as the classifier in both investigations [12] and [13].

4.4 Current emotion detection system

[14] supplied emotion detection, which specifies the necessity for a preferred orientation of the face picture in order to be recognised. The information gathered by this emotion detecting system is shown in Table 1 below.

Table 1. Current emotion detection system

Author	Method	Tested Images	Wrong Detection	Correct Detection	Accuracy(%)
[15]	A frame-work for the Recognition of Human Emotion using Soft Computing models	120	36	84	70

Emotional traits on the face may be traced back to the source of the person's eyes. In order to extract features, a filtering and edge detection approach has been presented. After that, a genetic algorithm was used to discover the ideal parameters for the final product based on the processed image (GA). For the extraction of ocular parameters using GA, a unique fitness function has been developed. The suitability of the predicted method for estimating emotions on a customized face is taken into consideration, as emotion detection techniques can serve as an expert system for continuous processing of presented data. To determine the user's mood, the training data set is fed into a back-propagation neural network.

5. Proposed Methods

Face Detection, Feature Extraction, and Face Classification are the three sequential operations required to perform the specified work. First, a camera is utilised to record the human face and then bounding box coordinates are used to monitor the exact location of the face as it is found in real time. This step uses the Adaboost algorithm and Haar cascade detection to identify the face. The Viola Jones algorithm and the haar cascade features are used in tandem for the purpose of studying human faces. The photographs show shapes, objects, and landscapes, among other things. These features are taken from the model database that has been built to do automated face recognition, and compared to the human face that has been detected. Humans also have a propensity to use CNN to identify various face expressions or emotions. Adaboost and the Haar cascade method are shown in Figure 8 as a face detection system.

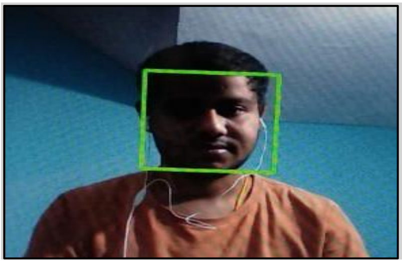
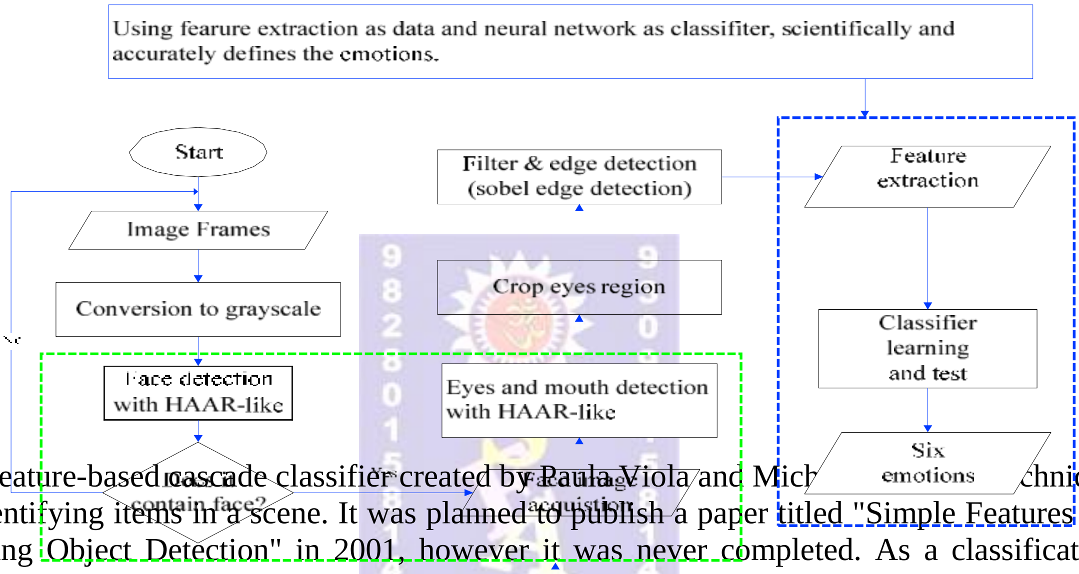


Figure 8: Face Detection

A significant amount of facial expression may be gleaned from the FER2013 dataset. In addition to the 28,709 instances in the training set, there is a 7179-example testing set. The data consists of 64 by 64-pixel grayscale photographs of the subject's face. All images have a consistent amount of space dedicated to a person's face since the face is automatically captured.



Haar feature-based cascade classifier created by Paula Viola and Mich[illegible]chnique for identifying items in a scene. It was planned to publish a paper titled "Simple Features for Boosting Object Detection" in 2001, however it was never completed. As a classification method, a Haar cascade may be described as a collection of Haar-like properties that are combined. The difference between the write's pixel value and the blank space's pixel value is the feature's definition. Its resolution is 64 x 64 pixels at the bottom. A simple face detector yields 160k possible Haar-Like possibilities. Not all of these choices, however, get the same treatment. As seen in Figure 9, the Adaboost and Haar cascade algorithms are used for face identification and feature extraction, respectively.

Figure 9: Feature Extraction

The CNN model observes the face through embedding vectors and matches the face emotion from the training model and matches it with the emotion from the training model using the VGG 16 model. The distribution platform uses the pycharm and python 3.5 software packages for face recognition and classification. Emotions are first recognised using the CNN model training and testing models, which are then used to identify the face. Anger; afraid; sad; delighted; neutral; and astonished are all examples of emotions that may be detected in real time. The VGG 16 standard was designed using the CNN paradigm for massive database classification. Fig. 10 depicts the suggested model's numerous emotion detections.

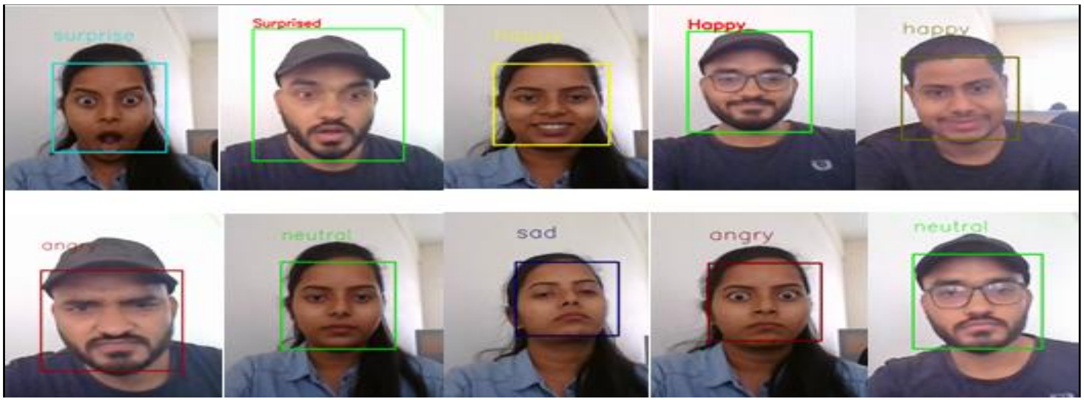
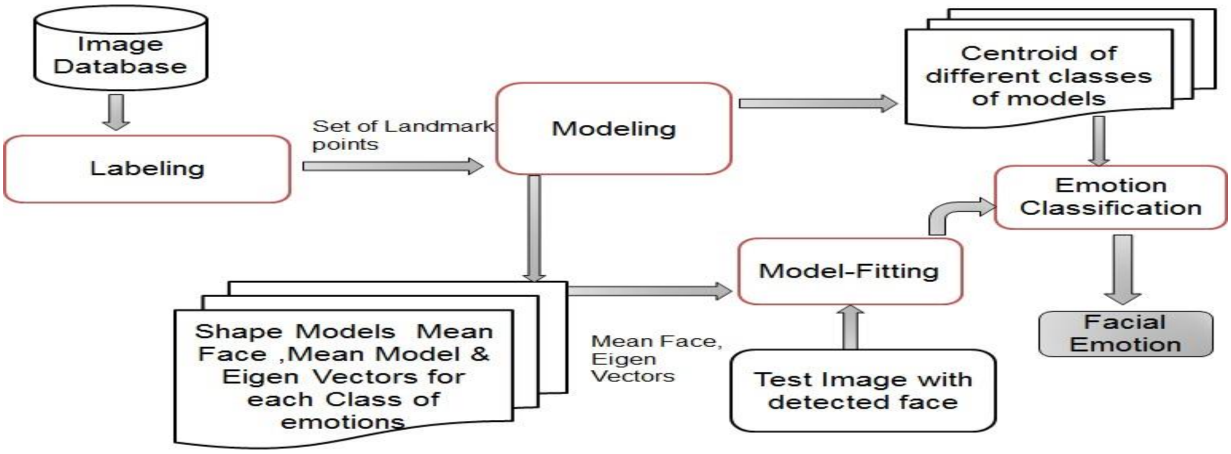


Figure 10: Emotion detection

In the accompanying image, the suggested system and its functioning are shown. It's important to note that each practical block carries out a specific function that differs from the others. Preprocessed cameras capture the image, which is then sent into the neural networks as extra data. Network architecture is used to create and train a dataset for emotion classification based on the information in the image that was gathered. Face identification and classification in real



time is shown in Figure 10 (below).

.**Figure 11:** Block diagram of Face detection and emotion classification

Figure 11 depicts the Real-time face detection and classification system. A box contains the image that was captured. In order to match the input picture, the images are trained to detect the expression of facial characteristics as Angry (fear), Sad (sadness), Happy (happiness), Neutral (neutral), and Surprise (surprise). Importing data; pre-processing; supplementing data as a feature vector; creating design model; training feature vector; and validating test model are all included in each stage of training procedure represented in this picture. Figure 12 shows a flowchart with the most important points in sequential order. Graph No. 12.

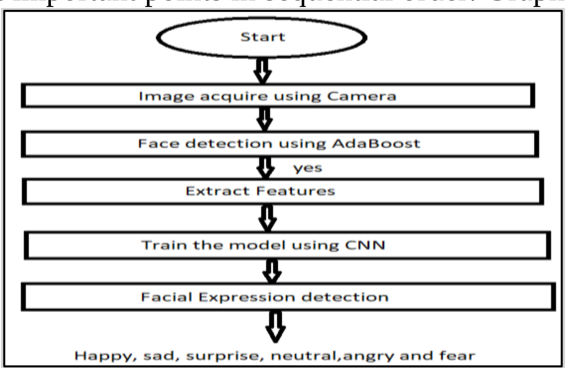


Figure 12: Face detection and emotion recognition flow chart

This application's results suggest a movement in the consistent and master field, with the estimate of customer commitment focused on the sensations of a high level system as the main point of attention. An important component of this inquiry relies heavily on the use of clear programming for face recognition and its normal assessment of sentiments in a specific historical context. It is crucial to appreciate the importance of sentiments in terms of administration quality in the movement and hospitality industries, where varied situations and programming of a more refined character have been completed. When it comes to identifying and correcting administration flaws, the newly developed application for capturing emotions in the tourist business gives professionals the data they need to better their status and project a positive image. Rather of using self-coordinated exams, artificial intelligence uses inventive programming that recognises faces and recognises emotions through facial expressions to determine pleasure. Artificial intelligence (AI) is used to monitor and measure client pleasure through sentiments in the mobility business sector, among other things, as part of this article's contribution.

7. Results and Discussion

The tests were carried out in accordance with the stated conditions and at the exact location where the patient was discovered. The emotions that have been linked to the least accuracy were researched in order to establish the psychological effect of client actions on us. We were able to better understand the participants' feelings by describing the results as dangerous lattices. The sense of hopelessness and dread was the hardest to understand. According to different sources, they expressed both confidence and surprise [15]. The most current picture game plans may be demonstrated by applying mathematical processes that need a great deal of information representation in order to reach the limit of measurable use in the restorative extraction of emotion-based pictures via their repercussions and Performance Evaluation Metrics.

We used a camera to record real-time images of a training dataset and a testing dataset for our experiments. All of our experiments used the Fer2013 dataset, which comprises 28,709 pictures. A look at Table 2 reveals our inclination to employ a range of time periods and attain a high degree of accuracy and recognition for unexplainable feelings. We may deduce from the research that a considerable decline in square error happens when the epoch of training data is greatly increased. Due to the massive quantity of testing data, we offer, every phrase must be evaluated in real time. It's clear from Table 2 that we use a wide range of training and testing datasets, as shown by the example. The results reveal that the experimenter's ability to discern emotion decreases as the epoch value increases. Similarly, model accuracy improved as the age got more prosperous. For testing and training images, the CNN technique is good, as demonstrated in Table 3.

Table 2: Mean Square Error and Model Accuracy of CNN

Epoch	N0. of Training Data	No. of Testing Data	Mean Square Error	Model Accuracy
20	28709	70	7890	72.51 %
45	28709	70	6543	77.20 %
90	28709	70	5289	81.57 %
125	28709	70	4187	85.41 %
175	28709	70	3458	87.95 %
200	28709	70	3242	88.71 %

In interpersonal relationships, subtlety and sophistication are widespread and their success is typically reliant on a variety of factors. The context, tone, and timing of the interaction, as well as the expectations of the parties involved, may all play a role. Participants who are successful

must be able to adjust to their counterpart's changing moods and behaviour during the meeting. Fortunately for humans, this ability is mostly innate, but there are variances in proficiency across people. Word selections, voice inflections, and body language may all be used by humans to quickly and instinctively deduce the emotions of others around them. The fact that humans have a similar set of fundamental emotions that are widely recognised probably has something to do with this analytical ability.
This is significant since the facial expressions used to convey these emotions are always the same. As a result, no matter what the language or cultural barrier, there will always be a fundamental set of facial expressions that people can analyse and use to communicate. Many years of research have shown that there are seven universal human face expressions that communicate the whole range of human emotions. Among the most common emotions we feel are anger, contempt, disgust, fear, joy, sadness, and surprise. In most circumstances, analysing a person's face may be a good way to determine their genuine mood and reactions in a scenario, unless they deliberately disguise their feelings.

Table 3: Facial Emotion Detection Classification

Epoch	Happy	Sad	Fear	Neutral	Angry	Surprise
20	66 %	62 %	60 %	60 %	65 %	64 %
45	71 %	68 %	65 %	70 %	70 %	68 %
90	77 %	74 %	65 %	70 %	70 %	70 %
125	82%	79 %	70 %	80 %	80 %	70 %
165	86 %	83 %	80 %	80 %	80 %	80 %
200	91 %	90 %	86 %	87 %	89 %	93 %

Results from the experiment, which measured six separate face expressions, are shown in Table 3. Each era has a distinct accuracy rate. In spite of looking at images, the facial expression model is still unable to distinguish between expressions of fear, surprise, joy, and other emotions.
Table 4 uses the whole merged image to construct the confusion matrix for Facial Emotion Detection, which is represented in Figure 1 as the confusion matrix.

Table 4. Confusion matrix of CNN

	Happy	**Sad**	**Fear**	**Neutral**	**Angry**	**Surprise**
Happy	90.33	00.00	00.00	00.00	00.00	7.74
Sad	00.00	89.33	00.00	14.67	11.33	00.00
Fear	00.00	00.00	85.33	00.00	00.00	00.00
Neutral	00.00	10.67	5.00	86.33	00.00	00.00
Angry	9.67	00.00	00.00	00.00	88.67	00.00
Surprise	00.00	00.00	9.67	00.00	00.00	92.26

Table 5 shows the accuracy result of CNN is comparatively better than other classifiers.

Table 5: Comparative result of different classifiers

Method	CNN	ANN	KNN	SVM
Accuracy	88.71 %	78.00 %	66.67 %	64.67 %

Table 6 summarises the experiment's findings on the accuracy of six distinct kinds of emotions. On the other hand, surprise and pleasure had substantially greater recognition rates than fear (see Table 6), with surprise having a much higher recognition rate. When it comes to diverse facial expressions, the characteristic value of surprise can be clearly seen, while the characteristic value of fear is less easily detectable, and so the identification rate of distinct facial expression characteristic values is lower.

Table 6. Average accuracy(%) of every emotion

	Happy	**Sad**	**Fear**	**Neutral**	**Surprise**	**Angry**
Proposed model	89.33	83.33	81.00	83.33	93.26	83.67
Current model	78.00	76.00	79.73	82.76	88.66	81

As a consequence, even after many years apart, we can still recognise the essence of our relationships, even if their facial traits have changed due to maturation, such as the growth of stubble and long hair. When it comes to applications like criminal identification and visa control, face recognition is seen as a major concern. As an example, limiting the number of faces in a data set that is well-known for its capacity to identify criminals allows us to see a particular face.

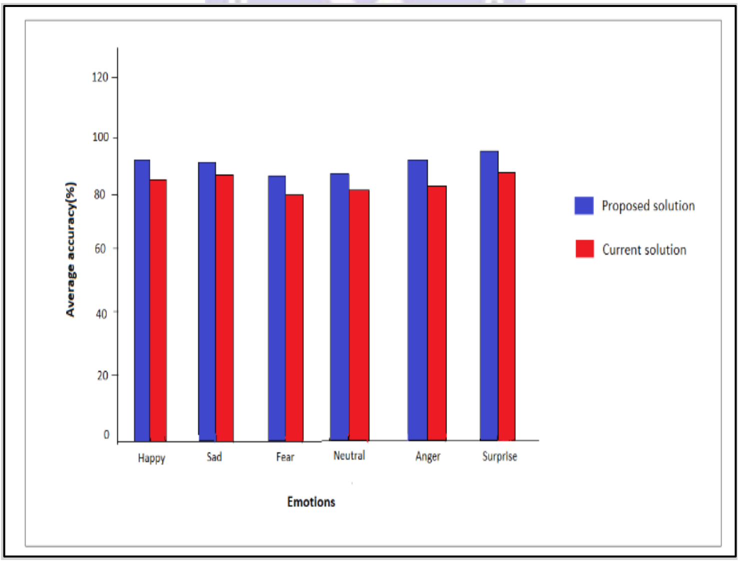


Figure 13: Average accuracy (%) of every emotion

Researchers were able to show that using facial recognition to assess client loyalty was just as beneficial as information gained from haphazard surveys (as shown in figure 13). Results show that legal and master areas are changing, with customer dependability being estimated by paying attention to how people feel about a high-level plan as the most significant modification. To begin, this research relies on the use of unambiguous programming for gaze recognition and its regular estimation of emotions in conjunction with and against a certain historical background. A knowledge of administration quality can only be gained through examining a variety of situations and programmes that have been implemented in the transportation and hospitality industries, as well as in other industries.

8. Conclusion

Using face recognition, this study proposed a model to address the issues of emotion detection based on face recognition, with attention given to the system's efficiency and accuracy at the time of the model presentation. The results obtained are quite encouraging, quick, and well-suited to real-world use. They're also quite affordable. To identify emotions in a face, this research recommended using a Convolutional Neural Network (CNN). It is possible to convey yourself via the use of six distinct facial expressions. We use real-time data analysis using the Haar-Cascade Classifier to place the spotlight on the round of the face. Haar-Cascades and neural network approaches may be used together to get the best of both worlds: high efficiency and high accuracy, respectively. Real-time data is received from a network-connected camera that is attached to a mobile device. In order to establish whether or not there are any recognisable faces in the image, the software system continuously scans it. It has a high degree of accuracy when it comes to recognising faces and the emotions expressed on those faces. Detected facial expressions may be classed as neutral if they don't exhibit any expressions, or as anger, sadness, joy, fear, or surprise if they do. As the value of each epoch rises and the model's accuracy improves, we are utilising the FER2013 database and training them to develop visualisations that will be gathered over time. However, the system accuracy typically declines if the light accessibility isn't enough to illuminate the features of the face. The distance of the faces from the camera also affects the results proportionally with the nearest faces giving better results.

9. Future Scopes

Real-time facial formation mining opens the door to a wide range of applications, such as picture designs in a given congregation of images, when done correctly within the constraints of real-world application capabilities, such as optimised computation complexities, space-time complexity adjustment, and a robust framework to handle human errors. Computer vision and picture processing are essential for success, as are information mining and AI (Artificial Intelligence). A database and human reasoning are also required.

References:

1. Yang, D., Alsadoon, A., Prasad, P. C., Singh, A. K., & Elchouemi, A. (2018). An emotion recognition model based on facial recognition in virtual learning environment. Procedia Computer Science, 125, 2-10.
2. Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). SN Applied Sciences, 2(3), 1-8.
3. Yadav, S. P. (2021). Emotion recognition model based on facial expressions. Multimedia Tools and Applications, 1-23.
4. Egger, M., Ley, M., & Hanke, S. (2019). Emotion recognition from physiological signal analysis: A review. Electronic Notes in Theoretical Computer Science, 343, 35-55.
5. Zheng, W., Tang, H., Lin, Z., & Huang, T. S. (2010, September). Emotion recognition from arbitrary view facial images. In European Conference on Computer Vision (pp. 490-503). Springer, Berlin, Heidelberg.
6. Atallah, R. R., Kamsin, A., Ismail, M. A., Abdelrahman, S. A., & Zerdoumi, S. (2018). Face recognition and age estimation implications of changes in facial features: A critical review study. IEEE Access, 6, 28290-28304.
7. Kumari, J., Rajesh, R., & Pooja, K. M. (2015). Facial expression recognition: A survey. Procedia computer science, 58, 486-491.
8. Owusu, E., Zhan, Y., & Mao, Q. R. (2014). A neural-AdaBoost based facial expression recognition system. Expert Systems with Applications, 41(7), 3383-3390.
9. Nakisa, B., Rastgoo, M. N., Tjondronegoro, D., & Chandran, V. (2018). Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors. Expert Systems with Applications, 93, 143-155.
10. Kumar, J. (2016). Affective modelling of users in HCI using EEG. Procedia Computer Science, 84, 107-114.

ISSN -2393-8048, January-June 2021, Submitted in June 2021, iajesm2014@gmail.com

11. Garg, D., & Verma, G. K. (2020). Emotion recognition in valence-arousal space from multi-channel EEG data and wavelet based deep learning framework. Procedia Computer Science, 171, 857-867.
12. Sebe, N., Lew, M. S., Sun, Y., Cohen, I., Gevers, T., & Huang, T. S. (2007). Authentic facial expression analysis. Image and Vision Computing, 25(12), 1856-1863.
13. Yan, J., Zheng, W., Cui, Z., Tang, C., Zhang, T., & Zong, Y. (2018). Multi-cue fusion for emotion recognition in the wild. Neurocomputing, 309, 27-35.
14. Shan, C., Gong, S., & McOwan, P. W. (2009). Facial expression recognition based on local binary patterns: A comprehensive study. Image and vision Computing, 27(6), 803-816.
15. Zhao, Y., Wang, X., Goubran, M., Whalen, T., & Petriu, E. M. (2013). Human emotion and cognition recognition from body language of the head using soft computing techniques. Journal of Ambient Intelligence and Humanized Computing, 4(1), 121-140.