MACHINE LEARNING TECHNIQUES FOR DETECTING AND RECOGNISING EMOTIONS BY FACIAL EXPRESSIONS

Dr.Sarojini Yarramsetti[1], Dr. Shiney Chib[2], Vaishali Singh[3], Galiveeti Poornima[4], Mrs. Rupali Suraskar[5], Dillip Narayan Sahu [6*]

[1]Associate Professor, Department of CSBS, Nehru Institute of Engineering and Technology, Coimbatore, India.

[2]Director, Department of Management, Datta Meghe Institute of Management Studies, Nagpur, India.

[3]Assistant Professor, Department of Computer Science & Engineering (CSE), Ambalika Institute of Management and Technology (AIMT), Lucknow, UP, India.

[4]Assistant Professor, School of CSE & IS, Presidency University, Bangalore, Karnataka, India.

[5]Teaching Associate, Department of Electronic, Larsen and toubro Edutech, Nagpur, Maharashtra, India.

[6*]Assistant Professor, Department of MCA, Gangadhar Meher University, Odisha, India.

*Email: dillip1seminar@gmail.com

Abstract

Non-verbal communication, such as facial expression, also includes the use of body language and vocal inflection to communicate emotion. Facial expressions may be used to many different purposes. Computer science, Biotechnology, Psychology, Chemistry, and Pharmacy all get additional interest as a result of face expression recognition. Expressions used in HCI studies to better understand human-computer interactions. Facial expression identification paves the way for precise extraction of emotional characteristics. Static picture facial expression identification methods neglect the dynamic and static properties of facial organs and muscle movements, as well as the geometry and visual elements of facial emotions. By performing patch matching procedures and choosing critical patches, we are able to overcome this constraint utilising extracted 3D Gabor features based on patches. Positive findings from the testing phase include an increased Correct Recognition Rate (CRR), better performance when taking facial characteristics and bodily movements into account, fewer incorrect face registrations, and shorter processing time. The suggested method provided the greatest CRR on the JAFFE and Cohn-Kanade AU-Coded Facial Expression datasets, establishing a clear advantage above state-of-the-art methods.

Keywords: human-computer interaction, emotion detection, and automated facial expression recognition systems.

1. Introduction

Because human emotion is employed in a wide variety of real-time applications, the study of methods to automatically recognise facial emotions is an important area of research.In general, human express his/her emotions in interpersonal communications through various forms, which includes facial expression, voice tone, body gesture and posture. There are a number of indicators that may be used to quantify how a person is feeling. Human-Computer Interaction (HCI) as an emotional computing system and other uses have attracted the attention of many academic and industrial research organisations. An individual's emotional condition may be approximately determined by observing facial expressions. Since this is the case, facial expressions are often employed as inputs in automated Facial Expression Recognition (FER) systems, which are then put to use in a wide range of real-time applications, including video games, police interrogations, psychiatry, animations, etc. Actually, in 1872, Darwin put up a notion regarding the similarities between animal and human facial expressions. Massive research on human facial expressions conducted by Ekman and Friesen (1978) provide credence to this hypothesis. They identified six typical facial expressions that cover the whole range of human emotion: happiness, sadness, surprise, anger, fear, and disgust [1]. They have also researched facial expressions in many civilizations, including some considered to be very remote or primitive, and have found evidence to support the notion.

Human emotions as determined by facial expression recognition have been proven to be consistent across cultures despite cultural norm variances [2]. Popular among FER researchers is the framework developed by Ekman and Friesen, the Facial Action Coding System (FACS). Facial Action Units(AUs) are the building blocks of this paradigm for analysing facial expressions. FACS analyses still photographs,to manually characterise a person's facial expression. Features are extracted based on the degree of facial distortion. Ko and Sim's (2010) Dynamic Bayesian Network (DBN), Wang and Li's (2015) Local Binary Pattern (LBP), Dalal and Triggs's (2005) Histogram of Oriented Gradients (HOG), and Lowe's (1999), Scale Invariant Feature Transform (SIFT) are all useful tools for identifying and quantifying facial emotions[3].A large majority of real-time face emotion identification algorithms rely on appearance characteristics or geometry data. The geometry of a person's face is determined by the arrangement of its many parts, such as the forehead, eyes, eyebrows, lips, nose, chin, etc. Facial texture, including wrinkles, furrows, and other creases, is used to extract appearance-based elements. According to psychologists, various facial features each communicate unique emotions. For instance, the eyes and eyebrows in the top region of the face provide more data for identifying anger and fear. Conversely, the mouth is the primary tool for expressing negative emotions like disgust and positive ones like delight, while the upper and lower portions of the face are responsible for expressing surprise.

Despite being a simple and obvious procedure for humans, expression recognition is a challenging issue for human-computer interaction (HCI) based applications including mobile computing, gaming, health monitoring, and robotics[4]. The authors Li and Ji, (2005) have proposed an approach based on probabilistic framework to dynamically model and recognize the user’s emotional state and responds to assist them on time. The authors Picard et al., (2001) have shown the significance of affective psychological states on human emotions. Using fuzzy logic and regression trees, the authors Rani et al., (2007) offer a method that combines various psychological indices to recognise a particular emotion (anxiety) in real time, and then compare and contrast the two methods' efficacies [5].From the above, it is clear that emotion identification in human faces has emerged as a promising area of study for cutting-edge and real-time applications in fields including machine learning, computer vision, human cognition, and pattern recognition.Seven specific feelings extensively used for interpretation of expression is depicted in Figure.1., which Darwin (Darwin, 1872)described. Consequently, the appropriate teaching and evaluation resources on facial expressions can be made available in most existing systems where attempts have taken to understand such simple words, these specific feelings based on classifying them[6].

Researchers also putmuch effort into calculating facial anatomy, defining the facialappearance and its activity, and even classifying facial looks. According to Tian et al., the general structure for facial glance perception is well established to be identical to that and provided in Figure.2. The method of general emotional recognition for the face consisted of facial development, facial identification, definition, and facial expression interpretation[7]. The first step is face acquisition, in particular, to focus on different attempts to get face area from pictures or videos and include all facets in the model of a reference with face markings, extract characteristics for a defined body language and categorize the terms. The second step is the use of facial appearance extraction and representation techniques. Affect recognition is essential to understand one's facial structure[8]. It can be seen as generating excellent characteristics to define the presence, form,and motion of facial expressions well. More specifically, facial look applications designed to describe the structure of the face moreaccurately. Generally, they are fundamental properties, e.g., texture,form, and colour.

Figure.1: Seven specific feelings extensively used for interpretation of expression

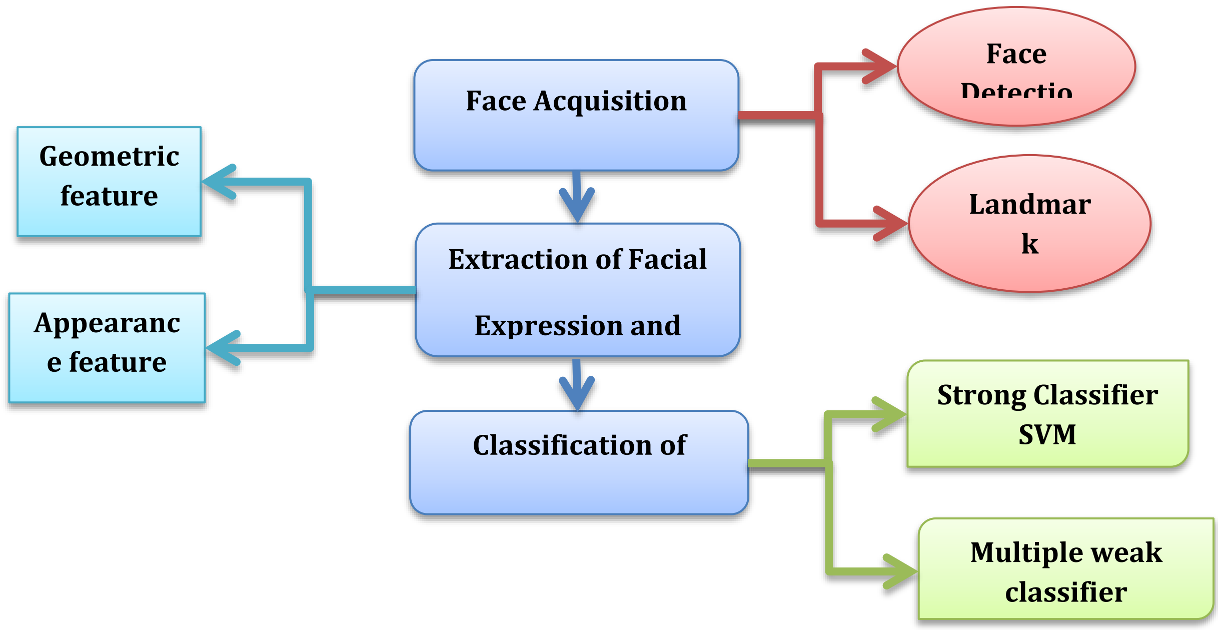

Figure.1: Flowchart of Facial Expression Recognition.

2. FACIAL EXPRESSION OF EMOTION

Facial Expressions (FEs), together with other audible, linguistic, and physiological cues, allow people to communicate with one another. A facial elastogram is a correlation between two or more facial muscle changes that occur subcutaneously. They play a crucial role in evaluating our mental states and emotions, including where our attention is now focused, how we feel, and whether or not we understand or disagree with a given signal[9]. When compared to the face and other mental, social, and physiological indicators, they are one of the most effective means of transmitting and describing the shown emotion.

Boulogne, in 1862, conducted experimental manipulations of FE activations by administering galvanic electrical stimulations directly to the face muscles; this marked the beginning of scientific inquiry into the FE analysis of human emotions. Boulogne's illustration of his theory that various facial muscles are responsible for each emotion supported his position. Then, in 1872, Darwin and Prodger suggested the idea of universal FEs in man and animals, speculating that there could be a few number of basic emotions that are exhibited consistently throughout cultures and geographies [10]. Research on FEs conducted in depth by Ekman and Friesen in 1978 lends credence to the idea that the facial expressions are universal in nature. The seven universal FEs are (1) anger, (2) disgust, (3) fear, (4) happiness, (5) sorrow, (6) surprise, and (7) neutral. The most common framework for studies of emotion recognition is psychological ideas on the universality and interpretation of facial expressions in terms of fundamental emotion categories. It has the benefit that, notwithstanding the inequalities imposed by societal norms, FEs referring to fundamental emotions are easily recognised and characterised by people.

Before Suwa's(1978) presentation of an early work on automated FE analysis from photographs, FE analysis was largely the territory of psychologists doing research. When it comes to computer vision research in the 1990s, automated Facial Expression Recognition (FER) becomes a hot and difficult area. In order to classify abstract emotional states, it has been posed as a pattern recognition and learning issue including the extraction and encoding of face movements and feature deformations from pictures or sequences [11]. Face detection, feature extraction and representation, and recognition are the three cornerstones of the framework for automated FER. Mase (1991) was the first to employ optical flow for automated FER, and subsequent work by Essa and Pentland (1994) and Iwano et al. (1996) identified the direction of movements induced by Fes and grouped them into fundamental emotional categories by optical flow calculation. Picard (1997) developed a number of models for emoji recognition and introduced many novel uses for affective wearable computers. Improvements in automated FER are made possible by advances in related fields including psychology, human movement analysis, face detection and tracking, and face recognition.Nevertheless, computers still have a long way to go before they can match human emotion recognition efficiency, since they are unable to do so, for either the user or themselves. Assisting the computer with a sense of human feeling may lead to a new degree of affection in Human Computer Interaction (HCI).

It's not easy task to automate face expression analysis. Several options must be given, and its performance is affected by a number of variables. Such differences are shown in Figure.3. Firstly, Labels or face visual features can be used for describing FEs.Second, we may classify emotions as either "basic" or "non-basic," depending on their complexity.The third aspect is how posed or natural FE seems. Fourth, how fast and subtle the FE is. Fifth, Images or videos as the input source. Sixth, identifying emotions is impacted by both internal and external variables.

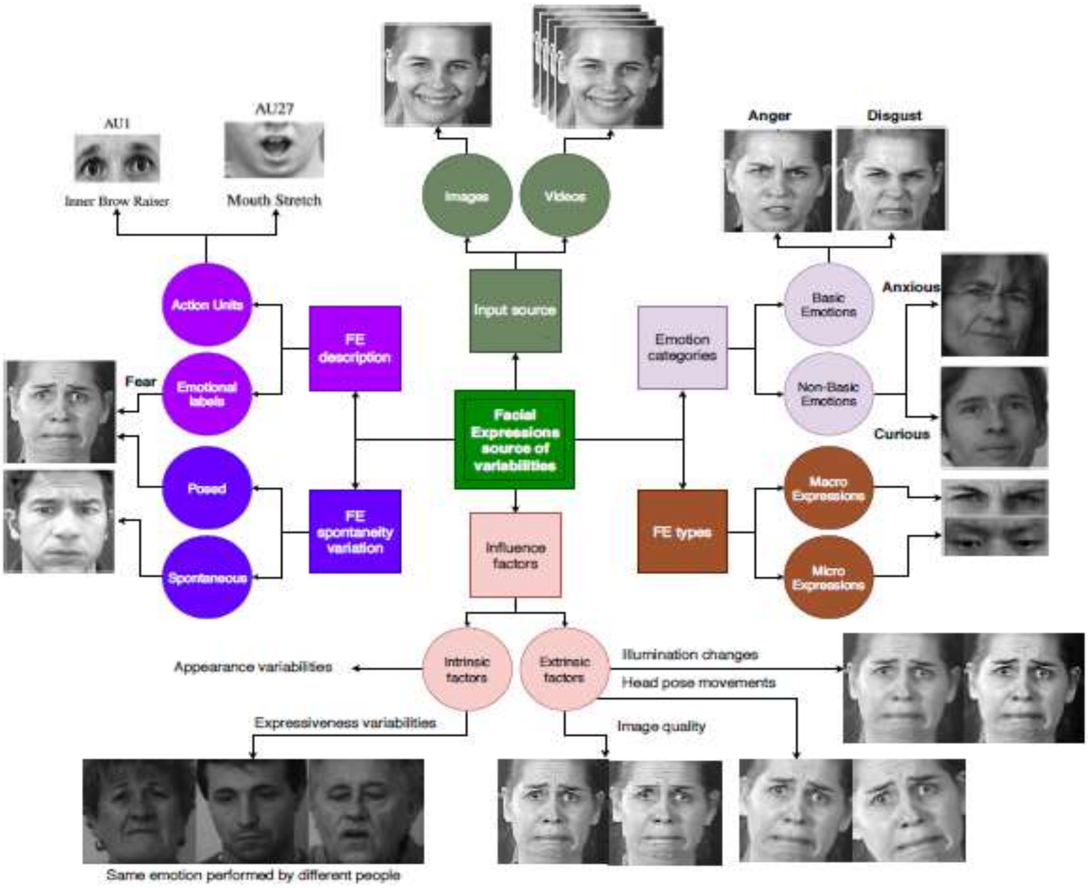

Figure.2: Schematic representation of potential variables and factors affecting facial expressions

There aren't many things that may damage the accuracy or speed with which facial expressions are recognised. Individual's real-world circumstance, where one's face is not always neutral, is the major source of variation in emotion.

3. Literature Survey

Expressions on a person's face convey a lot of information about how they're feeling or thinking. According to the results of a psychological study, just 7% of information is communicated by words, 38% through non-verbal cues including speech cadence and tone, and 55% through facial expressions[12]. As a result, FER is a subject that has been, and will continue to be, regarded as crucial to study for its theoretical and practical significance and its practical implications in everyday life. Emotional states, such as happiness, sadness, fear, etc., are communicated via facial expressions, and these states are triggered by a wide range of events. Anger, happiness, surprise, sadness, disgust, and fear are only some of the most common facial expressions, and they were all identified by Ekman (1992) as the six fundamental emotions. Numerous recent developments in fields like as computer vision, robotics, medicine, HCI, etc. have drawn a huge number of researchers to the study of automatic FER[13]. However, the FER system has several restrictions, such as the fact that conventional methods could struggle to deal with this volume of data in real time. The FER systems are also impacted by elements like as lighting, noise, occlusion, and fluctuations in the user's head position, all of which may lead to ambiguity and confusion. Many existing methods use custom-built features for FER, necessitating extra time and money in development and processing.

In this work, an architectural definition for a FER system to address the aforementioned research problems is offered. Within the framework presented here, taken into account are the six core emotions identified by Ekman (1992). The method incorporates the eyes, eyebrows, nose, and mouth characteristics to construct the FER model. This strategy suggests extracting these features from a reference picture of a neutral face using a geometric basis. Each component of the face takes use of the geometric structure and its characteristics [14]. The lips, for example, are mapped to the parabola and its attributes, the nose to the tetrahedron and its properties, and so on. Similarities in the neutral face's geometric parameters compared to those of other expressive faces are examined. Expression distortion is quantified mathematically. To further deal with the size, orientation, and translation invariant characteristic, the suggested FER model extracts the whole face and uses the idea of Interval Graph. A method is provided to compute the degree of deformation and the degrees of direction in the vertical, horizontal, and diagonal planes [15]. For the above-mentioned considerations, these values are used as feature vectors. Experiments are run on popular benchmark Datasets including FER2013, JAFEE, and CK+, and the resulting classification accuracies are tallied. Facial expressions are categorised using Support Vector Machine (SVM), Multi-Layer Perception (MLP), and Random Forest (RF) classifiers. The estimated expression of the pictures in the Dataset is determined to be accurate using the suggested method. The findings are compared to current practises in the field.

Figure.4. depicts the fundamental components of the FER system. In most cases, a fundamental FER programme will have two primary stages, which are known respectively as face abstraction and recognition of facial expressions. The primary objective of this study is to include current improvements on each of these metrics, namely, abstracting expressions and categorising facial emotion recognition based on FER behaviours. This review's primary objective is to discuss any new information on FER that has become available between the years 2013 and 2019, in particular.

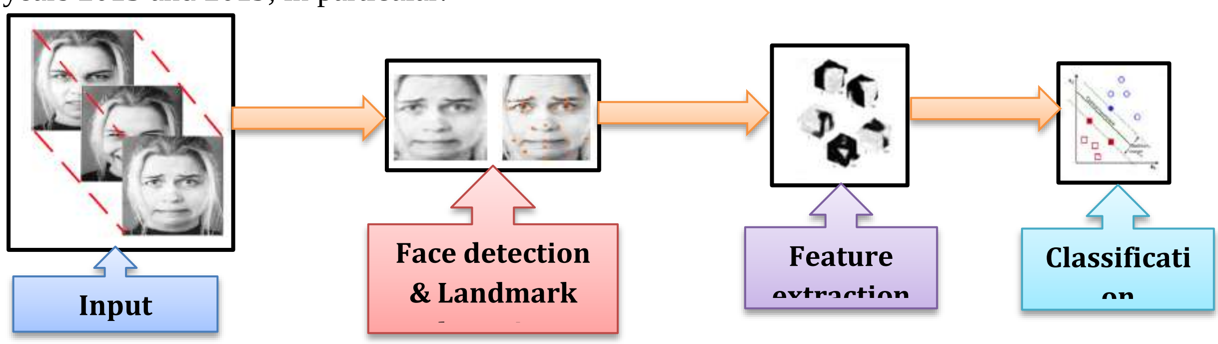

Figure.4:Basic FER System

In order to conduct a quantitative analysis of the different categorization algorithms, a number of FER investigations have been carried out[16]. Before we evaluate the work that was done on the FER, let's have a look at the specialist vocabulary that was discussed below. Even if the analysis of the research demands certain unique language, and these specific vocabularies have a key function to play, Figure.5. presents an illustration of the real-time FER system design.

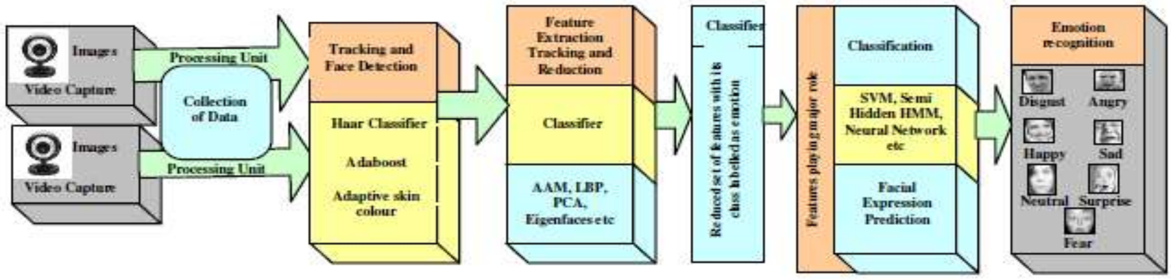

Figure.5: The system architecture of a real-time FER system

The combination of these two fundamental feelings will result in complex emotions[17]. There are a total of 22 different composite emotions that have been categorised, including the 7 fundamental emotions, the 12 composite emotions that are expressed the most often, and 3 extra emotions such as appeal, hatred, and adoration. Several instances of CE are shown in Figure.6., where a) the bottom and upper faces of AUs are shown (CK+ dataset photos). b). Spontaneous emotions (Facial images gathered from YouTube) c). Fundamental feelings (sad, fearful, and angry) d). Compound your feelings (sadly fearful, happily disgusted and happily surprised).

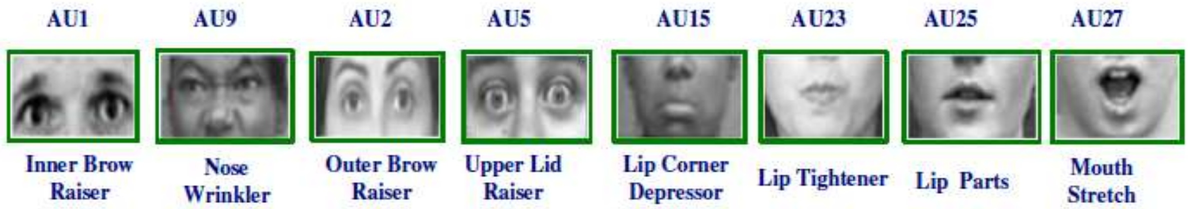

Figure.6: Examples of a Wide Range of Action Units and Expressions on the Face

4. Facial Feature Extraction Techniques For Static Images

The system must extract face elements like the lips, eyes, and so on in order to identify emotions from photos. Both frontal and non-frontal pictures can contribute to emotions. There are now two distinct face extraction techniques that may be used on still images: Characteristics-based geometry 2) Appearance-based techniques.

4.1 Geometric feature based methods

The geometrical characteristics of the eyes, mouth, and nose are computed using the feature-based or analytical technique. In this method, a feature vector of that image is formed from the face's contour and the locations of the various facial characteristics[18]. The size, configuration, direction, and placement of these organs all have an influence on the formation of facial expression. It is possible to arrange and coordinate the appearance of facial components such as the lips, nose, eyes, and eyebrows. In spatial approaches, face traits are connected to one another through spatial relationships[19]. Nevertheless, the majority of the time, capturing geometric elements calls for an intricate method that involves the identification of forms in phases [20]. That is challenging to adjust to in natural or real-time environments, which are subject to fluctuating conditions. In addition, graphical solutions for skin textural changes, such as wrinkles and furrows, that are essential for face expression modelling are often overlooked. Geometric characteristics derived from three separate extraction techniques, namely Active Shape Models (ASMs), Active Appearance Models (AAMs), and Scale-Invariant Feature Transforms (SIFT).

A. ASM

Timothy F. Cootes and colleagues have suggested a method for feature matching that centres on a mathematical approach known as an active shape model (ASM). ASM is a point distribution model (PDM) that assesses many form variations and a variety of approaches that aggregate the grey values across a set of key feature points. Additionally, ASM is an acronym for adaptive shape modelling[21]. The ASM role separation approach is characterised by 58 distinct face symbol features, which are shown in figure.7 as an overview. The ASM approach is divided into many distinct two-stage processes. models that were generated using the training data in the first form with numerous labeled landmark function points. Instead, localised texture models were constructed for each end of service that was specifically identified [22]. Second, an iterative search strategy might be used in order to modify the structure of the example in accordance with the two different construction models. As the FER nose, the work of Shbib et al. describes the geometric displacement that exists between the position of the expected ASM feature point and the mean ASM shape. In recent years, Cament and colleagues have created an improved version of ASM face recognition that they term active type and mathematical models (ASSM). This model, which has computational implications for FER, is known as ASSM.

B. AAM

The Active Appearance Model (AAM) was developed by Coos and colleagues in the year 2001. AAM successfully extends ASM by getting a comprehensive understanding of the nature and form of it. A simulation model is developed by AAM utilizing statistical data preprocessing in great detail. This is then followed by a relevant data test assessment using the mathematical model[23]. To identify the link between form and texture, AAM not only employs worldwide knowledge of shape and texture, but it also conducts a comparative study of the native customs of the surface. This is done under the auspices of ASM.

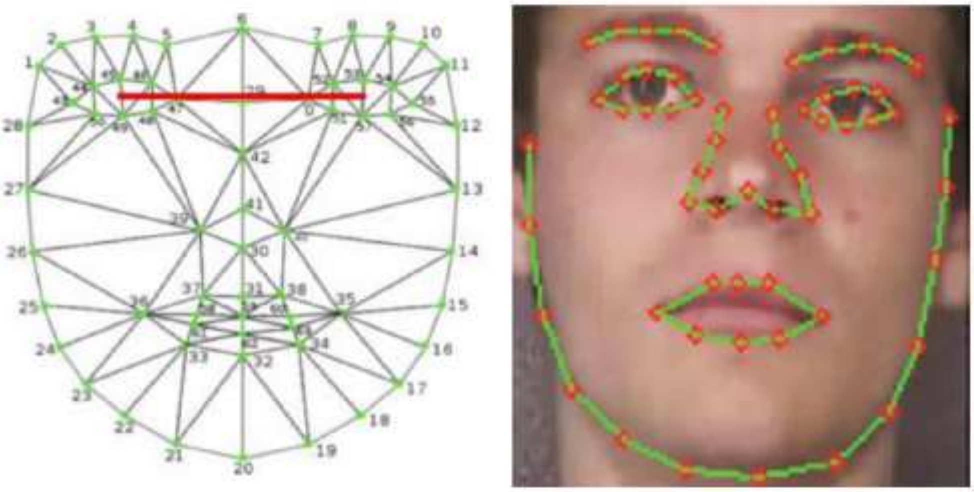

Figure.1: 58 facial landmarks are used in an ASM-based facial landmark detection system.
The differential AAM and variable learning were built so that a FER approach could be achieved. The difference that is assessed between AAM reference pictures and input images (such images with neutral voice) in order to evaluate differential-AAM functions[24]. There have also been developments in more complex AAM models, such as AAM-based Directed Gradient (HOG) histograms, dense-based AAM, and AAM-based regression. These models have been created. Investigating how well these newly produced AAM variations function against FER is an interesting piece of study.

C. SIFT

Scale-Invariant A transform function, also known as SIFT, has been proposed by David Lowe. This function, which is based on a local picture, is referred to as a matching descriptor. The characteristics of SIFT remain the same regardless of the dimensions, rotation, or translation of the images, and the lighting situation does not have a significant impact on these characteristics[25]. Figure.8 demonstrates how to get SIFT-descriptors at these core positions as primary marks for the facial markings in the major morphological sections of the face by using a feature extraction approach that is based on SIFT and was utilised in Berretti et al. An example picture that has 85 different face features and lines to find. In this particular scenario, the magnitude feature transformation approach was discriminated against by the SIFT-functional Extractor, which was utilised[26]. This method is able to make helpful judgements on the overall look. The multispectral picture was given a new transformation that was a scale-invariant function and was given the name GA-SIFT. This transformation was created by Li et al. using Geometric Algebra (GA).

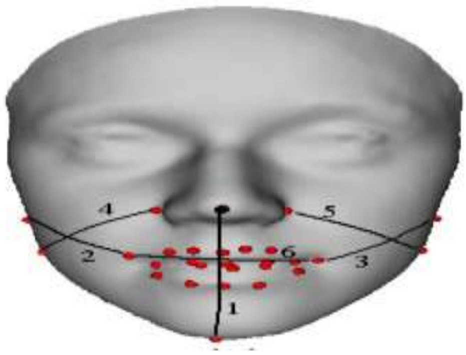

Figure.2:SIFT based feature extraction method

At first, an unique notion of spectral bands parts based on the GAtheory was established with definitions of both the spectral and spatial dimensions[27]. A scaled-down version of a multispectral picture had been created. It was discovered that GA-dependent volatility was evident in Gaussian near SIFT representations. The GA theory serves as the foundation for defining and cataloguing the featurepoints in a permanent manner.

D. Appearance-based methods

Appearance-based techniques propose the use of a full face image or distinct areas on a face photo to reflect the hidden properties of a face picture, especially those of subtle changes such as wrinkles and fractures[28]. These approaches are becoming more popular. Local binary patterns (LBP) and the wavelet representations by Gabor are two ways that are mostly accessible for knowledge extraction. Both of these methods are descriptor-based and appearance-based.

E. LBP

The Local Binary Pattern, often known as LBP, is a trained textural description operator that is capable of quantifying the picture and extracting information from the nearby texture [29]. The fundamental advantage of LBP operators is their strength, inventive capabilities, and grey invariance, which may be derived from the displacement, distortion, and lighting issues that are present in the picture. As a direct result of this, the LBP operator has a very solid reputation. Figure.9 presents an illustration of the LBP unit known as FER Extraction.

The extraction technique that is employed for the processing of LBP is monitored over the course of three primary steps. A photograph of the face was first cut up into numerous portions, and these parts did not overlap with one another. Second, the LBP histograms had been constructed for each and every machine in the lab. In prior study, they analysed the efficacy of dimensionality reduction tactics on FER practises by the LBP operator, such as biassed evaluations of local fisheries. This was done so that they could determine which strategies were most beneficial.

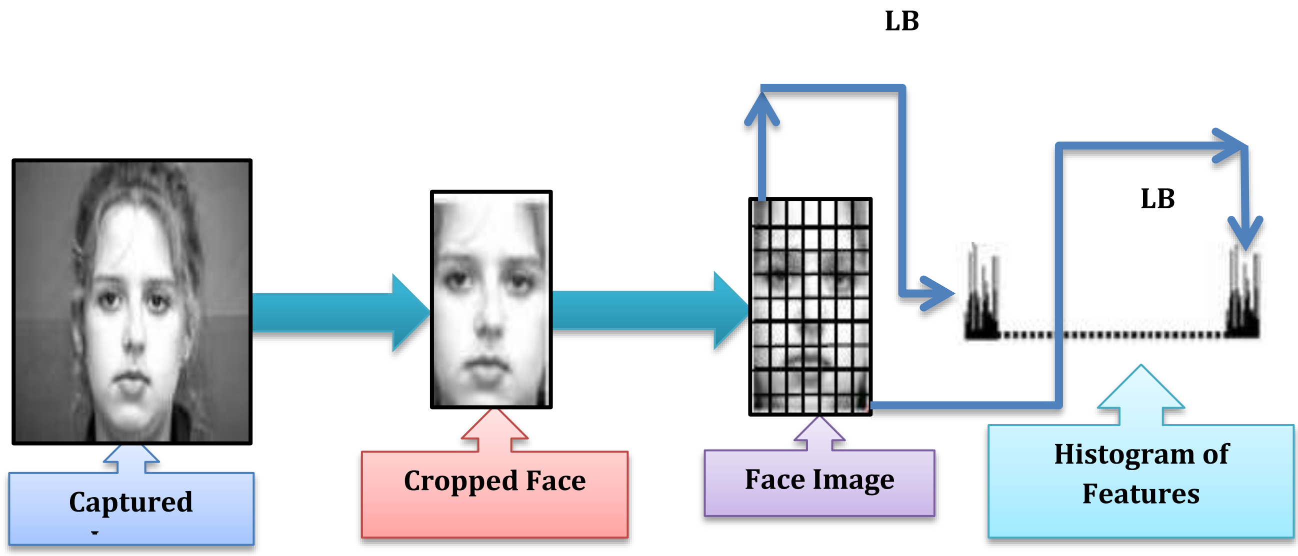

Figure.3: Feature extraction method using LBP

Some types of LBP operators found in the literature throughoutthe last years. Implementations of Local Binary Templates (LBP), also known as LBP-TOP, Three Orthogonal Planes (LBP), Geographical Directional Pattern (LDP), and Spatial Mapping are some examples of typical LBP implementations (LTP). For a three-dimensional FER, Li et al. devised a polytypic local multi-blockbinary pattern that is more often referred to as P-MLBP. P-MLBP combines 3D picture profile and textures to improve face look[30]. It comprises all of the main facial expression-dependent sections that correctly depict features.

F. Gabor

A common method of representing the characteristics of facial conversation is via the use of the Gabor wavelet representation. A filter series may be used to filter an image in more detail, and the consequences of the filtering can demonstrate the link between the gradient of a local pixel and the texture similarity of the image. The procedure for the Gabor wavelet representation that was used in order to get the face expression characteristic. It is able to detect touch motions on several scales and in multiple directions, and it had a little influence on the illumination changes. Figure.10 is a depiction of the Gabor wavelet that was used, which utilised a total of 18 Gabor kernels in three different sizes and six different orientations. A FER method was presented by Liu et al. in their work, which was based on the investigation of the Gabor wavelet functions and the core kernel of components (KPCA). Gu et al. carried out FER with the synthesis of radial coding and classification characteristics from the local Gabor, which contributed to the assumption that the processing speed increased. This filter was used as a local Gabor filter for bypassing the standard Gabor filter, which contributed to the assumption. The input pictures for this investigation were first processed using multi-scalar local Gabor filters, and then radial grids were used to encode the Gabor decomposition of the processed images. Owusu et al. have only lately built a neural-Adaboost FER system. They accomplished this by using the Gabor methods of extraction in order to get rid of vast quantities of the face features that indicate distinct facial forms of deformation.

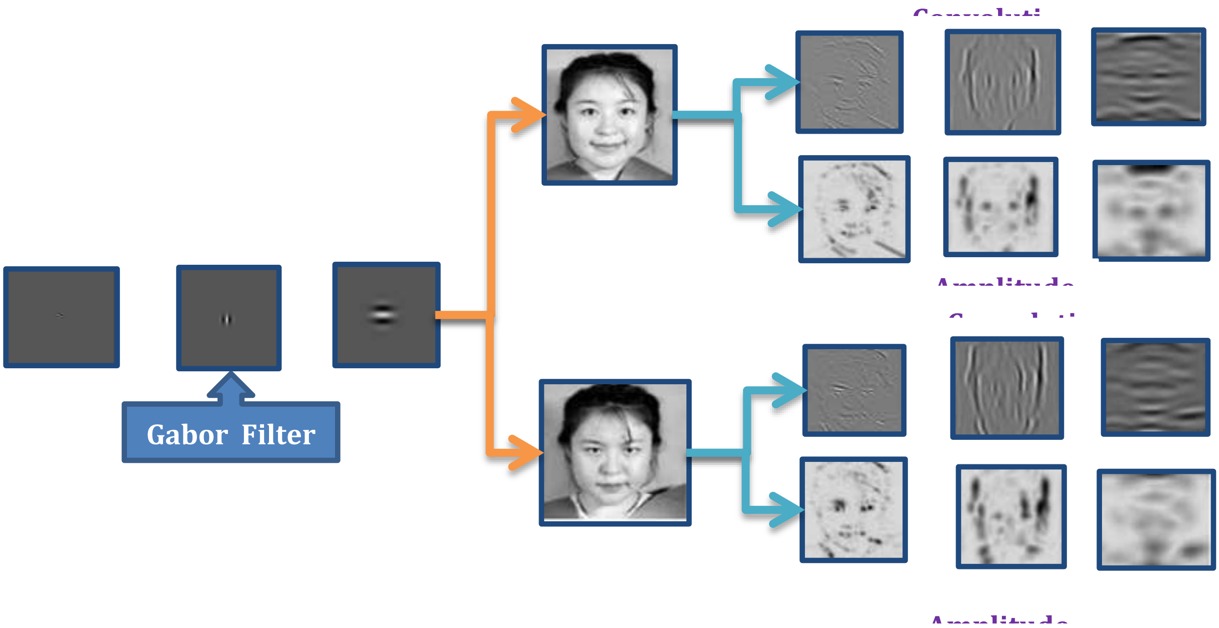

Figure.4: Examples of three different kernels for the Gabor wavelet representation

5. Results and Discussion

A variety of experimentations was done using JAFEE image datasets to evaluate the performance of the proposed technique. The suggested technique takes into account the input photos' eyes, eyebrows, noses, and mouths before employing facial landmarks to extract significant areas. Each of these components' geometric structures is mapped, and the separation and angle between their salient spots are calculated. These served as the input for the SVM classifier, which had an accuracy rate of 86% after being trained to recognize and classify facial emotions. In addition, the proposed FER system does classification using multilayer perceptron's and random forests, both of which are extensions of SVM classifiers. The Radial Basis Function (RBF), Sigmoid, Linear, Polynomial, and Lib SVM classifiers are employed as kernels in SVM along with other classifiers. The training technique also uses

MLP and RF models. The resampling technique is used by the RF tree to choose a random subset of the data from the Dataset.

The JAFFE Dataset's classification accuracy using LibSVM and RBF is shown in Table.1. It can beexamined how many variants there are for each facial expression using the confusion matrix. For instance, Tabl.1. shows that out of 30 surprise facial expressions, 22 are accurately identified as surprise, whereas 8 are misidentified as angry as well as joyful and disgust. The recommended approach successfully categorized 176 out of 213 images, with an overall accuracy of 82.65 percent. The classification accuracy for each facial emotion is shown in Table.1. Figure.11. shows accuracy using Lib SVM (RBF) on the JAFFE Dataset. The precision of JAFFE's classification using MLP is shown in Table.2. The accuracy of categorisation is highest for happy emotions and lowest for fearful ones. The overall classification accuracy is 88.31%. On the JAFFE Dataset, accuracy using MLP is shown in Figure.12. The classification precision of the RF technique is shown in Table.2. The accuracy is presently at 88.72 percent, a little rise. The classification accuracy has increased for expressions of happiness, neutrality, and surprise. On the JAFFE Dataset, accuracy using RF is shown in Figure.13.

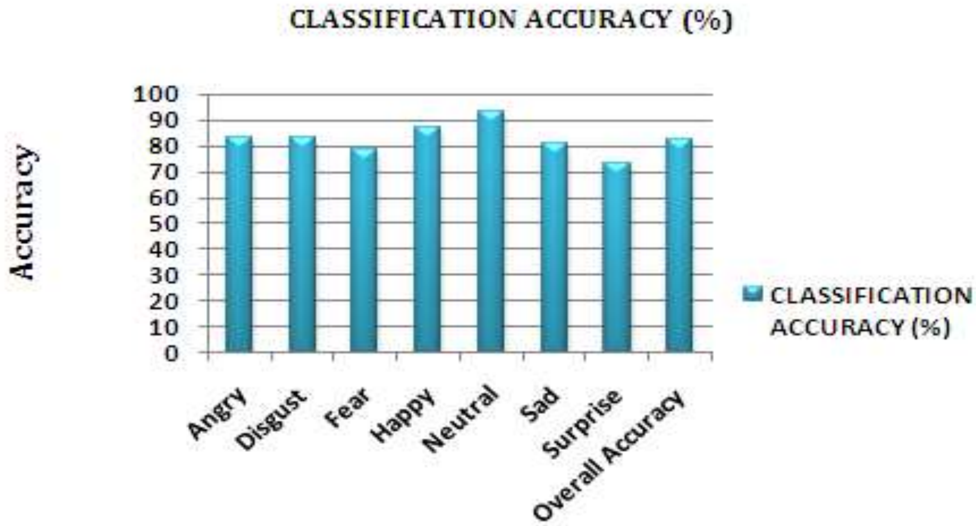

Figure.11: Accuracy using Lib SVM (RBF) on JAFFE Dataset

Table .1: Classification Accuracy using Lib SVM (RBF) on JAFFE Dataset

Facial Expressions	**Angry**	**Disgust**	**Fear**	**Happy**	**Neutral**	**Sad**	**Surprise**	**Classification Accuracy (%)**
Angry	25	0	3	0	0	2	0	83.33
Disgust	2	24	1	0	0	2	0	82.75
Fear	1	2	25	0	2	2	0	78.12
Happy	0	0	0	27	2	0	2	87.09
Neutral	0	1	0	0	28	1	0	93.33
Sad	0	1	0	0	3	25	0	80.64
Surprise	2	2	0	4	0	0	22	73.33
Overall Accuracy								82.65

Table.2: Classification Accuracy using MLP on JAFFE Dataset

Facial Expressions	**Angry**	**Disgust**	**Fear**	**Happy**	**Neutral**	**Sad**	**Surprise**	**Classification Accuracy (%)**
Angry	27	0	0	0	2	1	0	90.00
Disgust	2	26	1	0	0	0	0	89.65
Fear	1	2	26	0	2	1	0	81.25
Happy	0	0	0	29	0	0	2	93.54
Neutral	0	0	0	0	27	3	0	90.00
Sad	0	1	0	0	3	27	0	87.09
Surprise	2	0	0	2	0	0	26	86.66
Overall Accuracy								88.31

Table.3: Classification Accuracy using RF on JAFFE Dataset

Facial Expressions	**Angry**	**Disgust**	**Fear**	**Happy**	**Neutral**	**Sad**	**Surprise**	**Classification Accuracy (%)**
Angry	26	0	0	0	2	2	0	86.66
Disgust	2	25	1	0	0	1	0	86.20
Fear	1	0	28	0	2	1	0	87.50
Happy	0	0	0	29	2	0	0	93.54
Neutral	0	0	0	0	28	2	0	93.33
Sad	0	2	0	0	3	26	0	83.87
Surprise	0	0	0	3	0	0	27	90.00
Overall Accuracy								88.72

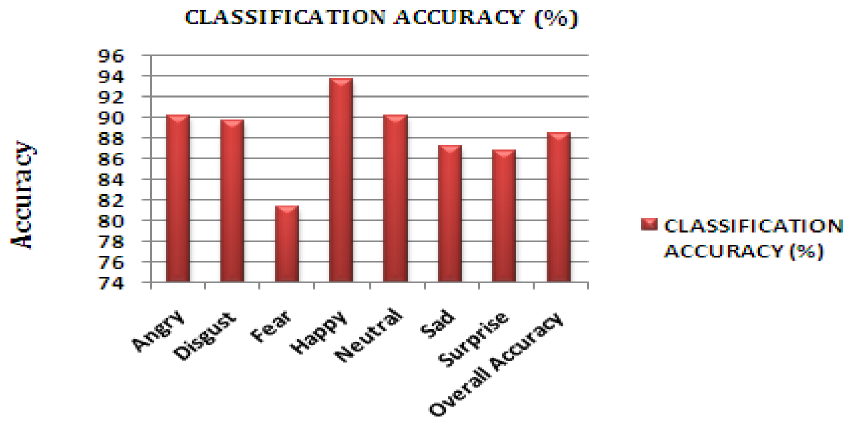

Figure.12: Accuracy using MLP on JAFFE Dataset

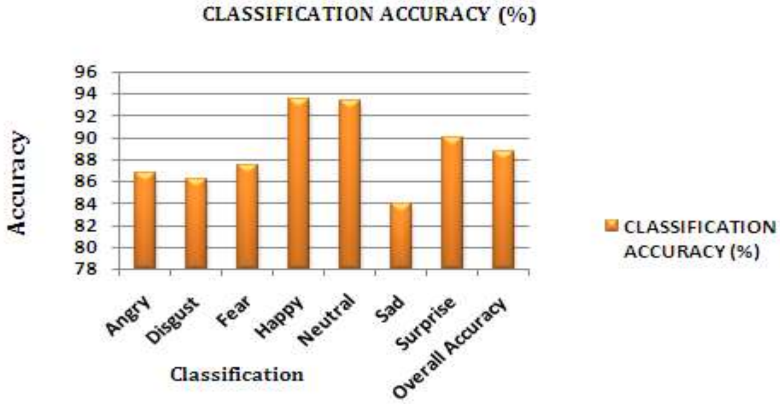

Figure.13: Accuracy using RF on JAFFE Dataset

The data above show that there is a high degree of accuracy in identifying neutral and joyful expressions, and a low degree of precision in identifying other emotions. The suggested FER system's performance on the JAFEE Dataset employing the Euclidian distance between all pairs of salient points within and across face components has so far produced positive results. Figure.14. shows a comparison of classification accuracy.

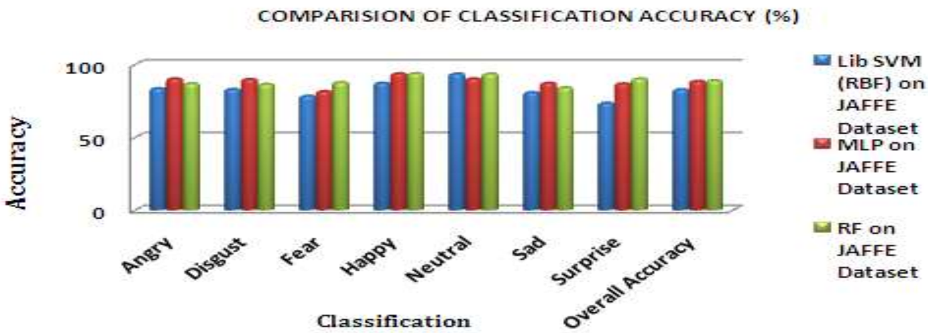

Figure.14: Comparison of Classification Accuracy

Averaging 86.56 percent classification accuracy, classification methods including LibSVM, MLP, and RF have been used. The results of the trials demonstrate that the extraction of image attributes is more computationally efficient and produces positive results.

6. Conclusion

This research investigates the problem of emotion identification using face motion data. The recognition performance, computing time, and comparison to the state-of-the-art performance are all evidence of the success of the suggested technique. The experimental findings also show that by taking facial motion characteristics into account, performance is much improved, and the results are encouraging when dealing with face registration problems. The findings reveal that when compared to point-based Gabor features, those based on patches perform better in terms of feature extraction, position preservation, recognition accuracy, and feature count. Different'salient' regions are associated with various emotions, but most of them are found in close proximity to the lips and eyes. The decision between point-based and patch-based characteristics does not seem to affect these'salient' regions for each emotion. The'salient' patches are dispersed throughout all scales, although they are most prominent at the larger ones. Out of four distances tested on the JAFFE and CK databases, DL2 fares the best. Anger, more than any other feeling, seems to play a role in this misunderstanding. Larger patch sizes are needed in the JAFFE database compared to the CK database.

References

1. Narayan T, Deshpande, Ravishankar S (2017). Face detection and Recognition using Viola-Jones algorithm and Fusion of PCA and ANN. Advances in Computational Sciences and Technology, 10(5):1173-1189.
2. Wagh P, Chaudhari J, Thakare R, Patil S (2015). Attendance System based on Face Recognition using Eigen Face and PCA Algorithms. In: IEEE International Conference on Green Computing and Internet of Things (ICGCloT), pp. 303-308.
3. Kumar GS, Prasad Reddy PVGD, Gupta S, Kumar RA (2012). Position Determination and Face detection using Image Processing Techniques and SVM Classifier. In: Proceedings of the International Conference on Information Systems Design and Intelligent Applications, pp. 227-236.
4. Kavinmathi B, Varadharajan E, Dharani R, Jeevitha S, Hemalatha S (2016). Automatic Attendance Management System using Face Detection. In: 2016 Online International Conference on Green Engineering and Technologies (IC-GET), Coimbatore, India, pp. 1-3, doi: 10.1109/GET.2016.7916753.
5. Zeng D, Veldhuis R, Spreeuwers L (2020). A survey of face recognition techniques under occlusion, IET Biometrics, https://doi.org/10.1049/bme2.12029.
6. Kahatapitiya K, Tissera D, Rodrigo R (2019). Context-Aware Automatic Occlusion Removal, In: IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, pp. 1895-1899, doi: 10.1109/ICIP.2019.8803141.
7. Srinivasan A, Balamurugan V, (2014), Occlusion Detection and Image Restoration in 3D Face Image, TENCON 2014 IEEE Region 10 Conference, Bangkok, Thailand, pp. 1-6, DOI: 10.1109/TENCON.2014.7022477.
8. Ma R, Mohamed A (2015), Image Processing Pipeline for Facial Expression Recognition under Variable Lighting, In https://web.stanford.edu/class/ee368/Ma_Mohamed.pdf.
9. Guo F, Yang J, Chen Y, Yao B (2018). Research on Image Detection and matching based on SIFT Features. In: 3rd International Conference on Control and Robotics Engineering (ICCRE), Nagoya, Japan, pp. 130-134, doi: 10.1109/ICCRE.2018.8376448.
10. Pal MK, Porwal A (2015). A Local Brightness Normalization (LBN) Algorithm for destriping Hyperion images. International Journal of Remote Sensing, 36(10): 2674-2696. https://doi.org/10.1080/01431161.2015.1043761
11. Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.
12. J. Kumari, R. Rajesh, and K. M. Pooja, “Facial Expression Recognition: A Survey,” ProcediaComput. Sci., vol. 58, pp. 486–491, 2015.
13. R. Lu, Y. Li, P. Yang, and W. Zhang, “Facial Expression Recognition Based on Convolutional Neural Network,” J. Phys. Conf. Ser., vol. 1757, no. 1, pp. 669–674, 2021.
14. E. Sariyanidi, H. Gunes, and A. Cavallaro, “Automatic analysis of facial affect: A survey of registration , representation and recognition,” vol. 8828, no. August 2013, pp. 1–22, 2014.
15. C. A. T. Iliou, “Features and classifiers for emotion recognition from speech : a survey from 2000 to 2011,” Springer Sci. Media Dordr., 2012.
16. L. Shuet al., “A Review of Emotion Recognition Using Physiological Signals,” Sensors, vol. 18, no. 2074, pp. 1–41, 2018.
17. W. Mellouk and W. Handouzi, “Facial emotion recognition using deep learning: Review and insights,” ProcediaComput. Sci., vol. 175, pp. 689–694, 2020.
18. B. Fasel and J. Luettin, “Automatic facial expression analysis: A survey,” Pattern Recognit., vol. 36, no. 1, pp. 259–275, 2003.

19. J. Goodfellowet al., “Challenges in Representation Learning: A report on three machine learning contests,” Int. Conf. Neural Inf. Process., no. 64, pp. 117–124, 2014.
20. X. Leng and J. Yang, “Research on improved SIFT algorithm,” J. Chem. Pharm. Res., vol. 6, no. 7, pp. 2589–2595, 2014.
21. D. Choudhary and J. Shukla, “Feature Extraction and Feature Selection for Emotion Recognition using Facial Expression,” 2020 IEEE Sixth Int. Conf. Multimed. Big Data, pp. 125–133, 2020.
22. S. S. Dlay, W. L. Woo, and J. A. Chambers, “Facial Expression Recognition using Local Gabor Gradient Code-Horizontal Diagonal Descriptor,” 2nd IET Int. Conf. Intell. Signal Processing, pp. 1–6, 2015.
23. Chin et al., “Enhanced Face Recognition Method Based On Local Binary Pattern and Principal Component Analysis For Efficient Class Attendance System,” IEEE Int. Conf. Signal Image Process. Appl. (IEEE ICSIPA 2019), pp. 23–28, 2019.
24. C. Mohammad and M. Refat, “Deep Learning Methods for Facial Expression Recognition,” 2019 7th Int. Conf. Mechatronics Eng., pp. 1–6, 2019.
25. Y. Bengio and P. Haffner, “Gradient-Based Learning Applied to Document Recognition,” Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
26. S. Li and W. Deng, “Deep Facial Expression Recognition : A Survey,” Int. Conf. Neural Inf. Process., pp. 1–25, 2018.
27. C. Pramerdorfer and M. Kampel, “Facial Expression Recognition using Convolutional Neural Networks: State of the Art,” Comput. Vis. Pattern Recognit., 2016.
28. S. Singh and F. Nasoz, “Facial Expression Recognition with Convolutional Neural Networks,” 2020 10th Annu.Comput.Commun.Work.Conf., pp. 324–328, 2020.
29. Saravanan, G. Perichetla, and D. K.S.Gayathri, “Facial Emotion Recognition Using Convolutional Neural Networks,” 2020 6th Int. Conf. Adv. Comput.Commun.Syst. ICACCS 2020, pp. 317–320, 2020.
30. Y. Wang, Y. Li, Y. Song, and X. Rong, “Facial expression recognition based on random forest and convolutional neural network,” Inf., vol. 10, no. 12, pp. 1–16, 2019.