Facial Expression Recognition based on Hybrid Featured CNN (FERHC) Algorithm

Rajan Saini
M.Tech Scholar
Department of Computer Science Engineering
Yamuna Institute of Engineering and Technology,
Yamunanagar, Haryana

Shainy Bakshi
Assistant Professor
Department of Computer Science Engineering
Yamuna Institute of Engineering and Technology,
Yamunanagar, Haryana

ABSTRACT

Human has various type of emotions by which his/ her mood can identify. There are three kinds of emotions such as basic, ambiguous, and complex type of emotions. Anger, joy, fear, sadness, neutral, surprise, and disgust are included in basic types of emotions. In the complex type of emotions pride, shame, anxiety, etc. are included. The ambiguous emotions are hope and compassion. The process which is used to detect human facial emotions is known as facial expression recognition. For a human, it is easy to understand the facial expression but for a machine it is difficult. The advancement in technology made expression recognition easier with the help of artificial intelligence. There are various applications of facial expression recognition systems such as user verification, mug shot matching, and user access control. There are several existing techniques of human facial expression recognition, but still, it is a challenging task. In the proposed work, an advanced learning-based CNN (Convolutional Neural Network) model with hybrid features is used to recognize human facial expressions. In images, some features are connection-oriented and some are edge-oriented, therefore hybrid methodology is used. The PCA (principal component analysis) and LBP (Local binary pattern) are used to extract the hybrid features from the image. For the classification of facial expression, a deep learning-based CNN model is used. Numerous fully connected layers are applied to the surface of one another in deep neural networks, for efficient recognition of complicated parts. The proposed model is evaluated on different parameters such as MSE, FAR, FRR, and accuracy.

Keywords

Face expression recognition, emotions, hybrid features using LBP+PCA, CNN (Convolutional Neural Network).

1. INTRODUCTION

Facial expression is amongst the strongest, automatic, and fundamental indicators of living creatures that communicate emotional reactions or expectations, irrespective of geographical boundaries, gender, or age, and it also has a broad array of applications, including healthcare delivery and supported transportation systems [1]. Face expression seems to be a behavioral means of expressing sentiment throughout the majority of instances [2], and can also be used as definite evidence to determine whether anyone is telling the whole truth or telling lie. Generally, the term emotion refers to something like a powerful, spontaneous psychological response that arises regarding what see or hear. Emotions are frequently assumed to be deliberately felt or deliberate. Aggressive behavior, psychiatric evaluation, and psychological changes all contribute to our feelings and emotions [3]. In table 1 several emotions of human and facial expressions are depicted.

Table 1. Various types of emotions with expressions [4]

Sr. No.	Emotion type	Expressions
1.	Happiness	• Smiling • Joy • Satisfaction • Gratification
2.	Anger	• Glaring • Frowning
3.	Sadness	• Lethargy • Quietness • Dampened mood • Crying
4.	Fear	• Pulling back chin • Widening the eyes
5.	Disgust	• Nose wrinkling • Upper lip curling
6.	Surprise	• Widening the eyes • Mouth opening • Raising the brows

The advancement in technology made expression recognition easier with the help of artificial intelligence. With the help of artificial intelligence, the detection of facial expressions becomes easy. The AI integrates emotions and helps to recognize the mood of humans [5]. The architecture of the facial emotion recognition system is presented in figure.1.

Various applications of the facial emotion recognition systems are User-verification, law-enforcement, Bank Security, and MNC (IT) sector [6]. In this work, hybrid features are extracted from the image because hybrid features are more helpful. The two different techniques will be combined for the hybrid feature extraction. The first technique is PCA (Principal Components Analysis) and the second is LBP (Local Binary Pattern).Feature extraction using LBP, and hybridization using LBP+PCA algorithm. It is a procedure of extracting data from an edge detector image of the feature set and utilizing that data for recognition or detection of the feature set.
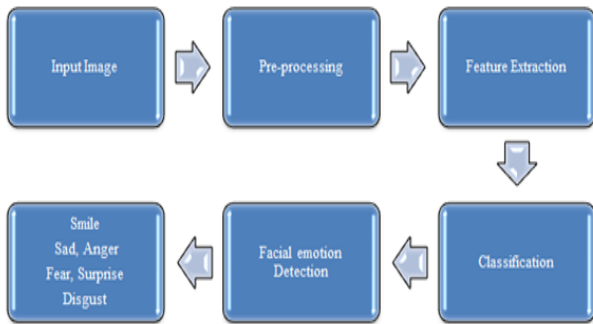
Fig1: A framework of Face Emotion Recognition System

These extracted feature data are also known as feature vectors. It is the most significant phase in implementing a detection system. Different categories of feature extracted methods are utilized to extract the FV (feature vector) from face emotion images. LBP feature extraction method is generating a binary number which is created for an individual image of pixels, where each-bit is assigned value-based on it is different from 1 of the image pixels at a defined radius. It is an easy version that labels individual face image pixel by a no. of derived from it is 3*3 matrix neighborhood. The nearest neighborhood face image pixels are assigned a binary number of one if larger than the center face image pixel and zero otherwise. Neighborhood image pixels are applying the PCA algorithm. The PCA algorithm searches the Eigenvectors (V) of a covariance matrix with the maximum or highest Eigenvalues (E) and then utilizes those to project the information into a new subspace of equal (==) or minimum dimensions.

We ask that authors follow some simple guidelines. In essence, we ask you to make your paper look exactly like this document. The easiest way to do this is simply to download the template, and replace the content with your own material.

2. LITERATURE REVIEW

Mehendale, et al. [7] introduced a unique facial expression assessment methodology based on deep convolutional neural. The facial expression CNN was built on 2 different convolutional neural networks (CNN): the first one eliminates the backgrounds from the image, while the focuses were specifically upon on retrieval of face extracted features. The expressional vector (EV) was utilized in the FERC framework to identify the 5 various forms of normal facial expressions. Information for supervision was taken from a repository comprising 10,000 photographs that had been saved (154 persons). Through the use of an EV with duration 24 data, it was feasible to appropriately emphasize the emotions with 96 percent accuracy. For each repetition, the final level of the neural network modifies the values and exponential levels, resulting in 2 different CNN. Using singular CNN, FERC varies from commonly used techniques, enhancing accuracy. Additionally, before the formation of EV, a new image editing approach was used to properly deal with many issues that come up.

According to **Gupta et al., [8]** emotions and feelings, are made up of a range of sub-emotions that were difficult to categorize.ResNet, an attentive component, and CNN, which provides spatial visibility to the system, were integrated with the suggested deep-learning system. The proposed framework was more effective at detecting real-world facial reactions.Mostly on the FER sample, the developed approach had produced satisfying performance and demonstrated effectiveness. By reaching better accuracy of 85.76 percent throughout the training process and 64.40 percent throughout

the test phase, the developed scheme beat established CNN-based models. Also on the FER database, the prediction design performed much better and might be employed in application scenarios. **Supta, S. R., et al. [9]** used both static images and real-time implementations for the recognition of facial movements. From the input image, the face and external components were first identified. To minimize the illumination effect, the pre-processing step was applied and for the identification of components contrast enhancement and object, sharpening techniques were used. The recognizable image characteristics throughout recognition algorithms then were extracted using HOG and grouped into a single functional unit. After that, a computation linear function had been used to classify expressions using the SVM. Besides seven simple facial movements, the conceptual framework was tested using the well-known Cohn-Kanade, and JAFFE facial expression databases. It was demonstrated that the developed FER method produces an accurate result. It presented that for the JAFFE and Cohn-Kanade repositories, the proposed FER method had an increase in accuracy to 97.62 percent and 98.61 percent, respectively. **Chen, Y., et al. [10]** analyzed that majority of current DL-based face expression recognition approaches did not take conceptual information into account and losing the ability to retrieve relevant attributes. (FES) Facial Expression Subsequent framework was proposed and new FER architecture for expression recognition was designed. The authors had established a novel division that produces a facial mask that focuses on the changing sections of the facial muscles. The suggested incorporating specific domain information was used as the learning guidelines for facial mask training. The median discrepancies were amongst externalizing behaviors and the associated emotional expressions. Additional testing on multiple facial features benchmark problems demonstrates that the suggested method outperforms the current state-of-the-art strategies. **Kim, S., et al. [11]** designed a DNN (deep neural network) based methodology for the recognition of facial expression. The authors had analyzed that there was no explanation regarding face emotion assumptions. The authors had discovered a correlation between Emotion mark and Facial Action Coding Units in the CK+ Dataset, based on the assumptions that facial expression was a mixture of facial muscles. The authors had introduced a framework that takes facial action units to describe the prediction performance of a CNN-based neural model. The convolutional neural network classifier was constructed on the CK+ Dataset and uses derived features to identify the feeling. The CNN neural model's derived characteristics and expression groups were used to classify the numerous facial action coding units in the interpretation framework. The interpretation technique had made well with the only attributes and expression groups collected through classification algorithm, according to proposed analysis.

Table 2. Various existing techniques for facial emotion recognition

Author's Name	Year	Proposed Method	Classification technique	Compared techniques	Results
Mehendale, et al. [7]	2020	CNN based face emotion identification system	Convolutional Neural Network	CNN, MLP	Accuracy= 96%
Gupta et al., [8]	2020	Deep self-attention network-based face emotion recognition system	CNN	CNN CNN+ Residual connection	Accuracy= 85.76%
Supta, S. R., et al. [9]	2020	Machine learning-based facial emotion expression detection system	Support Vector machine	SVM	Accuracy= 97.62%
Chen, Y., et al. [10]	2019	Deep learning-based prior network for the detection of facial emotions	CNN	PFN, FMG	Accuracy=98.06%
Kim, S., et al. [11]	2019	Deep model for the recognition of facial action	CNN	SVM, DNN model	Accuracy= 98.47%

Abbreviations: CNN (Convolutional Neural Network), MLP (Multilayer Perceptron), SVM (Support Vector Machine), DNN (Deep Neural Network).

3. EXISTING ISSUES AND GAPS

The purpose of face recognition would be to analyze or categories a provided emotional image into multiple expression forms, like Ekman's expression recognition classes such as fear, anger, sadness, disgust, and surprise [15]. To improve identification, feature vectors through one image to others is a complex and delicate process. The different existing methods of facial expression recognition with problems are depicted in Table 3.

Table 3. Various existing methods of facial expression recognition

Sr No.	Author's	Year	Research Gap
1.	Niu, B., et al.[12]	2021	• Facial expression recognition through static images only • High computational time • Need to explore more features to improve the accuracy of the system
2.	Mehendale, et al. [7]	2020	• Complex background generated more challenges
3.	Tcherkassof, et al. [13]	2020	• No guidance was provided for such visualization of multi-functional and behavioral expressions on some kind of one-to-one scaling
4.	Zhang, F., et al. [14]	2020	• Need to work on the classification of the image

4. RESEARCH METHODOLOGY

The research methodology steps are:
(i)   Acquisition
(ii) Preprocessing
(iii) Feature extraction
(iv) Classification

4.1 Face Image Acquisition:
Initially, download the face emotion dataset from the online UCI repository site. The research work has downloaded the 7 face emotions such as happy, sad, neutral, surprise, and other. All face emotions images are *.jpeg format. This proposed work has divided into two phases:
(i) Training and
(ii) Testing
Upload the face emotion images which have three-dimensional formats. After that, the uploaded image is converted into a three to two-dimensional image format. This conversion process is also called a grayscale image.

4.2 Face Image pre-processing:
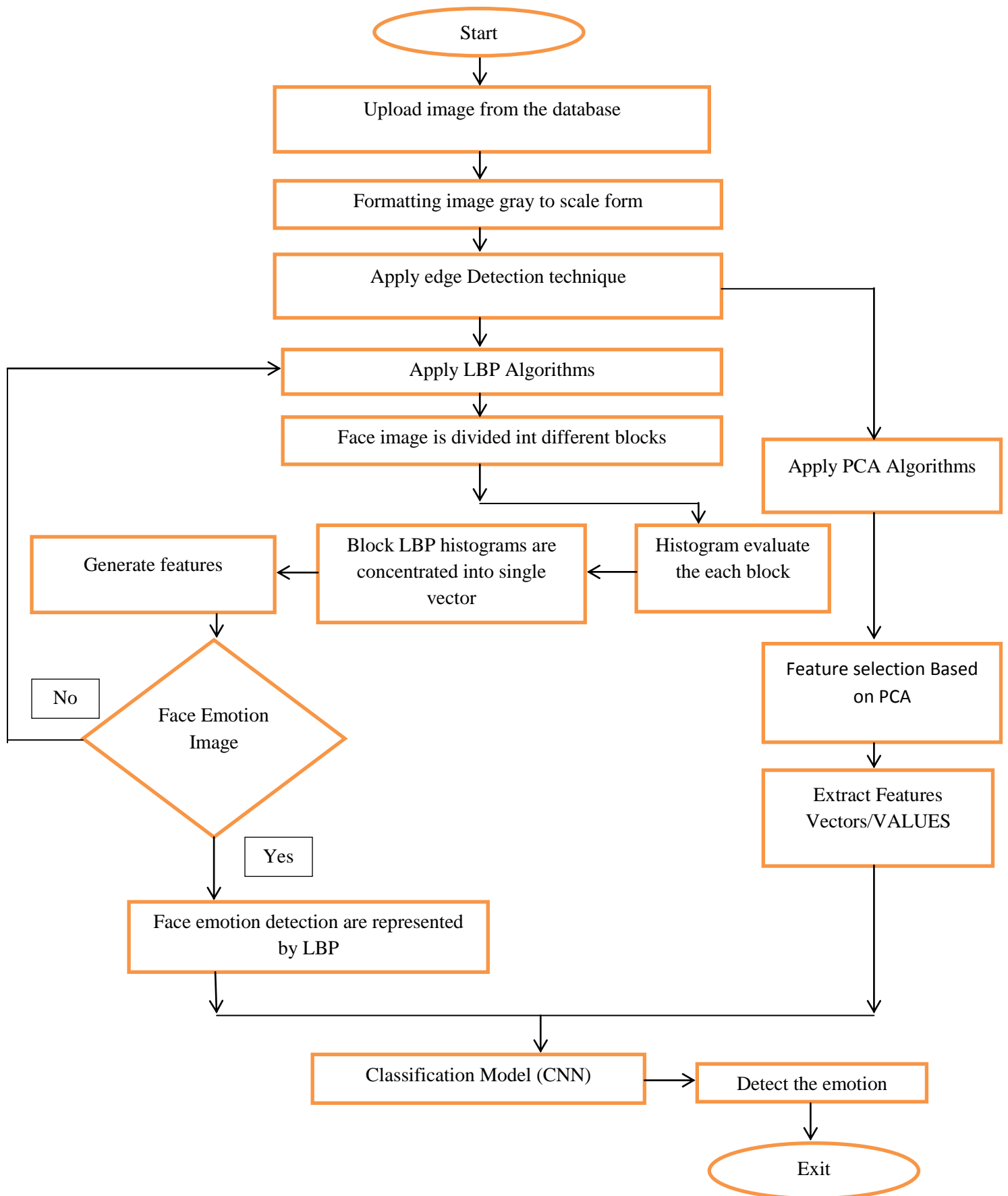This preprocessing process has shown three steps such as

Fig 2. Proposed Work

unwanted noise, filter, and edge detection. This phase is identified by the unwanted Salt & Pepper Noise in the train uploaded image. Then, the proposed model has introduced the filtration algorithm to remove the unwanted artificial noises in the facial emotion images. After that, the edge detection process is used as a Canny edge detector. The Canny edge detection is used to calculate the edges or regions in the facial areas.

4.3 Face Image Feature Extraction:
This process has defined the LBP and PCA algorithms. LBP stands for Linear Binary Pattern, PCA stands for Principal Component Analysis Algorithm. These algorithms are used to extract the feature sets. LBP algorithm is used to extract the feature vector on the right and left sides of the uploaded image. After that PCA algorithm is implemented to extract the optimized feature sets. The proposed work has introduced the hybridization algorithm to filter the feature vectors and values. Every image is passed through a face recognition method to extract more reliable feature sets. It improves the reliability and increases the speed, the PCA and LBP features are extracted from the face emotion region; normally, edge division is innovatively employed using the novel eigenvector and values to avoid the concentration of the feature sets.

4.4 Classification using CNN algorithm:
The deep learning (DL) based approach utilizes CNN to attain the automatic learning rule of the extracted face feature sets. CNN approach has been implemented to detect the emotional feature sets. This algorithm has implemented several layers such as Pooling, Conv, ReLu, Max Pooling Layers, and Input layers. These layers train and test the feature values and classify the features and evaluate the performance metrics such as accuracy rate, mean square error rate and compared to the existing ORB+LBP algorithm.

5. SIMULATION RESULT ANALYSIS
5.1 Dataset Description
This dataset contains 213 face images from 10 different Japanese female emotion subjects, and the individual image category has three /four illustrations of 7 face expressions. In this dataset, there are 6 basic facial expressions and 1 neutral expression. The face images are without color with a high image resolution of 256*256.
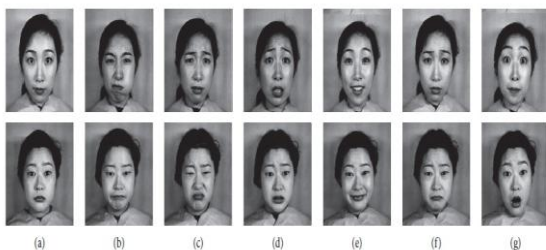

Fig 3: Dataset Different Face Emotion Images: (a) Neutral. (b) Anger. (c) Disgust. (d) Fear. (e) Happy. (f) Sad. (g) Surprise.

In research simulation, the complete 213 face emotion images (i) Anger = 30 face emotion images (ii) Disgust = 29 face emotion images (iii) Fear = 32 face emotion images (iv) Happy = 31 face emotion images (v) Neutral = 30 face emotion images (vi) sad = 31 face emotion images (vii) surprise = 30 face emotion images are utilized to calculate the research method. Some face emotion images from the jaffe [15] dataset are defined in Figure 3.

5.2 Mathematical Formula'
The research arithmetic formula is described as below:

5.2.1 MSE: MSE is defined as the estimator that measures the mean of the square of the difference between expected and actual values.

$$MSE = 1/m \sum_{k=0}^{n}(xi - xi^{\wedge})\mathbf{2} \qquad (i)$$

Here, MSE-Mean Square Error, m- number of total data points, x_i – expected values, $x_i ^$ - actual values.

5.2.2 FRR: It is determined as a percentage of the total number of false rejection instances divided by the total number of inspections performed by the system.

5.2.3 FAR: The FAR of a system is commonly calculated by dividing the number of erroneous acceptances by the quantity of identifying tries.

5.2.4 Accuracy: Accuracy is defined as the ratio of the addition of true negative prediction and true positive prediction to a total number of predictions. Because high accuracy necessitates simultaneously high performance and high correctness, it is defined as a collection of both kinds of observable errors (systematic and random).

$$Accuracy = \frac{True\ negative\ prediction\ + True\ positive\ prediction}{Total\ number\ of\ predictions} \qquad (2)$$

5.3 Result Analysis
This interface shows the knowledge domain, train the model using CNN classification. The knowledge domain is also called the training model. This model represents upload the multiple face emotion images (*.jpeg). The proposed model has applied the preprocessing steps such as noise, filter, and edge detection method. It develops the hybrid feature extraction method using LBP and PCA algorithms. After that, classify the face emotions such as happy, sad, surprise, and disgust and then evaluate the performance metrics such as MSE, FAR, FRR, Accuracy rate and compare them with the existing models.

Fig 4: Upload Image

Figure 4 shows the upload of the test face image. This uploaded image has a grayscale image pattern. This uploaded image is using the JAFFE face emotion dataset. This dataset has 7 different categories as Happiness, Sadness, Surprise, etc.

Fig 5: (i) Noise and (ii) Filter Image

Figure 5 represents the test image is converted into a grayscale format image. It reduces the dimensionality size from 3D to 2D. It has added artificial noise (Salt & Pepper). This noise is represented by the background and front-end noise information in the uploaded test image. After that, it developed the filter algorithm using the median filter. This

filtration algorithm has reduced the random noise and enhances the image quality, filtering is utilized, and here is analyzed for altering with kernel size 3*3 matrix.

Fig 6: Edge Image

Figure 6 defines edge detection using canny as a multi-stage method to identify the worldwide range of edges in face images. It extracts useful manageable data from the different visions of face images or objects and dramatically optimizes the number of information to be processed.
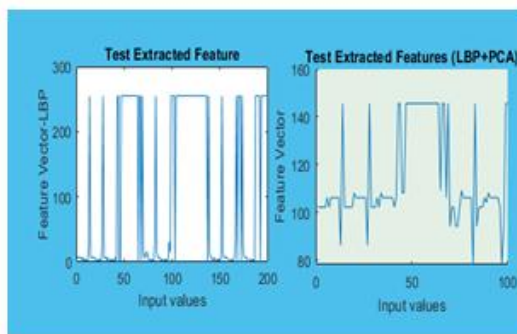

Fig7: Feature Extracted using LBP, and Hybridization using LBP and PCA algorithm

Figure 7 represents the feature extraction using LBP, and hybridization using the LBP+PCA algorithm. It is a procedure of extracting data from an edge detector image of the feature set and utilizing that data for recognition or detection of the feature set. These extracted feature data are also known as feature vectors. It is the most significant phase in implementing a detection system. Different categories of feature extracted methods are utilized to extract the FV (feature vector) from face emotion images. LBP feature extraction method is generating a binary number which is created for an individual image of pixels, where each-bit is assigned value-based on it is different from 1 of the image pixels at a defined radius. It is an easy version that labels individual face image pixel by a no. of derived from it is 3*3 matrix neighborhood. The nearest neighborhood face image pixels are assigned a binary number of one if larger than the center face image pixel and zero otherwise. Neighborhood image pixels are applying the PCA algorithm. The PCA algorithm searches the Eigenvectors (V) of a covariance matrix with the maximum or highest Eigenvalues (E) and then utilizes those to project the information into a new subspace of equal (==) or minimum dimensions.

The detection phase is using the CNN classification algorithm. This approach is the main concept of facial expression detection that often depends on analogic methods. This method helps and searches to normalize human facial efficiently while their interval need is a fraction of the existing used approaches. The method initiates with the detection of face emotions based on features in the form of color and structure of the human facial and that of the face background. CNN algorithm is used for classification with the help of different layers. After that, it evaluates the proposed system

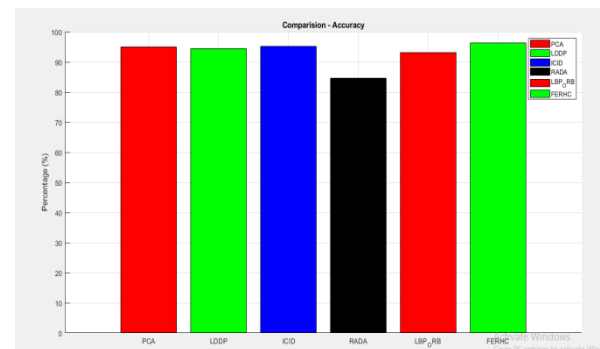performance metrics such as MSE, FAR, FRR, and accuracy rate.

Fig 8: Comparison

When the "COMPARISON" button is clicked and then shows the comparison graph in figure 8. This figure shows the bar graph plotting with the comparison between research and existing models parameter with accuracy rate (%). The research method has improved the accuracy rate compared to various types of methods like PCA, LDDP, ICID, RADA, LBP+ORB, and FERHC.

Table 4. Proposed Performance Metrics

Parameters	Values
Accuracy Rate (%)	96.3916
FAR	0.0086
FRR	0.0275
MSE	0.2834

Table 5. Comparison Analysis

Parameters	ICID	LDDP	PCA	RADA	LBP+ORB	FERCH
Accuracy (%)	95.10	94.50	95	84.60	93.20	96.39

Above table 4 below described the various performance metrics such as accuracy rate, FAR, FRR, and MSE rate. This proposed method has achieved an accuracy rate value is 96.39 percent, error types are FAR value is 0.0086, FRR value is 0.00275, and MSE value is 0.2834. After that table 5 shows the comparative analysis with various methods such as ICID, LDDP, PCA, RADA, LBP+ORB, and FERCH. This research model has reduced the error rates and improved the accuracy rate. The research model has improved the system performance and reduced the existing issues.

6. CONCLUSION AND FUTURE SCOPE

In this work, the basic facial expressions such as happiness, fear, sadness, etc. are recognized by the proposed system. A hybrid technique has been used for the extraction of essential features. The architecture of the facial expression recognition system with different steps is depicted. The two different techniques have been combined for the hybrid feature extraction. The first technique is PCA PCA (Principal

Components Analysis) and the second is LBP (Local Binary Pattern). The main objective of PCA is to determine the highest uncertainty patterns in high-dimensional data and expand the dataset around a new sub-space with the similar or fewer properties. LBP is used to find essential texture features which are very helpful in face detection and recognition system. A convolutional layer of the CNN model is the major aspect of the neural network which implements extraction of features using a mixture of differential equations functions, such as pooling layers and access points. There are multiple layers in CNN that provide different characteristics. The main advantages of the CNN model are detection of essential features automatically, provide highly accurate results, due to dense network it provides efficient results. Various existing methods of facial emotion recognition systems are surveyed. The facial expression recognition system is divided into 2 distinct categories, as per the literature. First is feature selection, followed by categorization. To distinguish distinct facial gestures, we can apply any classifier. However certain novel methods to facial feature extraction or depiction are required. After reviewing various existing techniques of facial expression recognition, it was observed that there are several issues in existing techniques. The problems of existing methods have been discussed. The most difficult part is integrating individual face recognition research into an elevated system that models individual interactions. The different challenges of facial expression recognition systems are aging, hyper spectral imaging, IRIS, poses, etc. IRIS is an integral component of face, but iris dimensionality reduction is critical in face biometrics research. Vision Monitoring and feature extraction after such a lengthy period of accessibility require maturity symmetry. Face Recognition Device gaining attribute in various age groups with different sets across statistics gathered with accuracy testing. The MATLAB simulation tool is used for the implementation of the proposed technique. For the proposed performance evaluation of the method, different metrics are used like MSE (Mean Square Error Rate), FAR, FRR (false acceptance and rejection rate), and accuracy rate. This proposed model has achieved a 96.39 per cent accuracy rate as compared to the existing methods.

7. REFERENCES

[1] Wang, W., Xu, K., Niu, H., & Miao, X. 2020 Emotion Recognition of Students Based on Facial Expressions in Online Education Based on the Perspective of Computer Simulation. Complexity, 2020.

[2] Mehendale, N. 2020. Facial emotion recognition using convolutional neural networks (FERC). SN Applied Sciences, 2(3), 1-8.

[3] "Emotion | Introduction to Psychology", Courses.lumenlearning.com, 2021. [Online]. Available: https://courses.lumenlearning.com/wsu-sandbox/chapter/emotion/. [Accessed: 22- May- 2021].

[4] "The 6 Types of Basic Emotions and Their Effect on Human Behavior", Verywell Mind, 2021. [Online]. Available: https://www.verywellmind.com/an-overview-of-the-types-of-emotions-4163976#:~:text=The%20emotions%20he%20identified%20were,shame%2C%20embarrassment%2C%20and%20excitement. [Accessed: 22- May- 2021].

[5] Liliana, D. Y. 2019, Emotion recognition from facial expression using deep convolutional neural network. In Journal of physics: conference series (Vol. 1193, No. 1, p. 012004). IOP Publishing.

[6] Borkar, N. R., &Kuwelkar, S. 2017, Real-time implementation of face recognition system. In 2017 International Conference on Computing Methodologies and Communication (ICCMC) (pp. 249-255).IEEE.

[7] Mehendale, N. 2020. Facial emotion recognition using convolutional neural networks (FERC). *SN Applied Sciences*, *2*(3), 1-8.

[8] Gupta, A., Arunachalam, S., &Balakrishnan, R. 2020. Deep self-attention network for facial emotion recognition. *Procedia Computer Science*, *171*, 1527-1534.

[9] Supta, S. R., Sahriar, M. R., Rashed, M. G., Das, D., &Yasmin, R. 2020, December. An Effective Facial Expression Recognition System. In 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE) (pp. 66-69). IEEE.

[10] Chen, Y., Wang, J., Chen, S., Shi, Z., &Cai, J. 2019, December. Facial motion prior networks for facial expression recognition.In 2019 IEEE Visual Communications and Image Processing (VCIP) (pp. 1-4).IEEE.

[11] Kim, S., & Kim, H. 2019, February. Deep explanation model for facial expression recognition through facial action coding unit.In 2019 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 1-4).IEEE.

[12] Niu, B., Gao, Z., &Guo, B. 2021. Facial Expression Recognition with LBP and ORB Features. Computational Intelligence and Neuroscience,.

[13] Tcherkassof, A., &Dupré, D. 2020. The emotion–facial expression link: evidence from human and automatic expression recognition. Psychological Research, 1-16.

[14] Zhang, F., Zhang, T., Mao, Q., &Xu, C. 2020. Geometry guided pose-invariant facial expression recognition. IEEE Transactions on Image Processing, 29, 4445-4460.

[15] M. Lyons, S. Akemastu, and M. Kamachi, "Coding facial expressions with gabor wavelets," in *Proceedings of the ;ird IEEE International Conference on Automatic Face and GestureRecognition*, pp. 200–205, Nara, Japan, April 1998.