# Human Emotion Detection And Classification Using Convolution Neural Network

Komala K[1], Jayadevappa D[2], Shivaprakash G[3]

[1]*Visvesvaraya Technological University, Belagavi, India*
[2]*JSS Academy of Technical Education, Bengaluru, VTU, India*
[3]*Ramaiah Institute of Technology, Bengaluru, India*

*Email: [1]komala1724@gmail.com, [2]djayadevappa@jssateb.ac.in, [3]shivaprakash@msrit.edu*

***Abstract – Facial expression is the non-verbal communication that provides some related-information about the emotion of a person. Detecting and recognizing human emotion is a big challenge in computer vision and artificial intelligence. The main objective of this paper is to develop a robust technique that can detect and identify human emotions such as anger, sorrow, happiness, surprise, fear, disgust and neutral in real-time. In this paper, Viola–John's algorithm has been proposed to detect the emotions more accurately. This algorithm is used for tracking the emotions in real time. In this paper, the real-time images are captured, and then features are extracted from the face images. Image enhancement and gradient detection are used after extracting the features of face images, then multiple feature extraction like GLCM, LBP and PCA are applied. All extracted features of face images are combined and compared with different databases to get accurate emotion state. CNN classifier is used to achieve conditional detection of accurate facial expressions.***

***Keywords – Facial Emotion Recognition, Gray-level Co-occurrence matrix, Linear binary pattern, Principal Components Analysis***

## 1. INTRODUCTION

The face expression in humans is one of the most important and powerful tools for communication. Facial expressions are very expensive. There is only 7% of the total significance of the effect of the message in the verbal component of the message, 38% of the total signal in tone and 55% in portrayed. Extraction of feature is very widely used in surveillance, biometric, psychiatric, military and Human-computer interface (HCI). In all the above, facial image is used to identify the emotional state in humans. In the human population, there are seven basic emotions and they are angry, sad, happy, surprise, fear, disgust and neutral. The above emotional states can be identified from human facial expressions.

Identifying the emotion of the human being is not an easy task it varies from person to person. Several scholars have made an effort to recognize human facial features like age, sex and feelings for machine learning. Security, Human-computer interaction, detection of different human emotions etc. are challenging problems. Human's show their feelings using facial expressions that are also part of communication. It is very easy to recognize the emotions and facial expressions for human but for machines, emotion recognition is a challenging task. There are various factors that play a vital role in developing an emotion recognition system. The important step in any facial expression is to detect the face

accurately, extract the feature and classify them. A wide variety of methodologies have been used in this field to detect human emotions.

A general process of emotion detection has four stages. Face detection, Preprocessing, feature extraction and face recognition. Viola-Jones(VJ) algorithm is used to extract features for identifying emotion images, then it can categorize human emotion using Haar and K-Nearest Neighbor (KNN) [1-2]. The mapping of extracted facial images with databases is the major problem because of analysis of lips and eye features and 2-dimensional nature of an image, in order to overcome the above problems, the extracted images can be analyzed using Region of interest (ROI) [4]. Facial expression recognition can be done using Statistical unsupervised technique ICA(independent component analysis) and genetic algorithm. Genetic algorithm is the feature optimizing technique that implements a system to recognize and predict facial emotions [3]. About 55% effect of overall facial emotions is proved that contributed to social interactions [5].

The main aim of this paper is to develop an accurate and robust method for Face Emotion recognition using CNN. This is enforced by developing a system which is capable of classifying an image into one of the eight basic emotions.

## 2. PROPOSED SYSTEM

A Unique method is described for facial emotion recognition systems in this paper. There are 3 stages: face recognition, feature extraction and emotion classification.

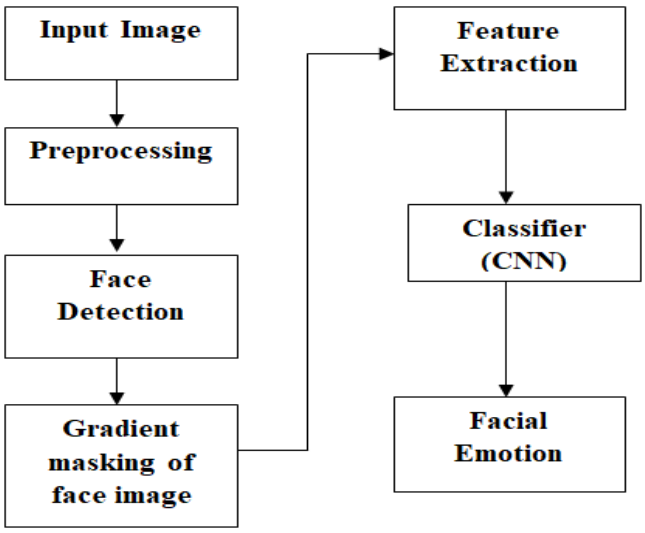


Fig1.Proposed block diagram

The dataset contains 215 images with resolution of 256*256 pixels and seven facial emotions. This dataset contains features such as Light intensity, background and size of face indicating real world situations [12-13].

### *2.1 Pre-processing*

This step is used to remove unwanted images such as noisy image, blur image and shadow image and also smoothens the images from video using Gabor filter. The resultant images are used to extract facial expressions.

*2.2 Face Detection*

VJ face detection technique is used to identify the facial image. Using the V-J technique Haar-wavelet concept is used to detect faces by developing Integral image.

There are four stages in VJ technique; they are Haar, Integral image, AdaBoost and Cascade structure.

Figure 2 shows different types of Haar features. Haar feature is applied on input images to check whether there are any facial images present in the captured image. Haar feature can be calculated by adding of whole image pixel and then are subtracted to get a single value. If the obtained value is more in that area, then it indicates the presence of a human face in the image.

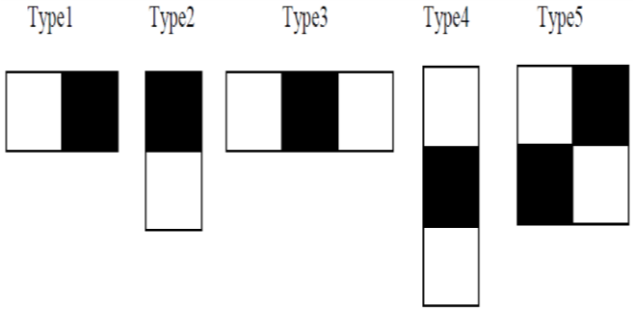


Fig2. The different type of Haar features

The Integral-Image technique is used to evaluate the summation of pixel beyond and left to the corresponding pixels in a particular rectangle of an image as shown in fig3.

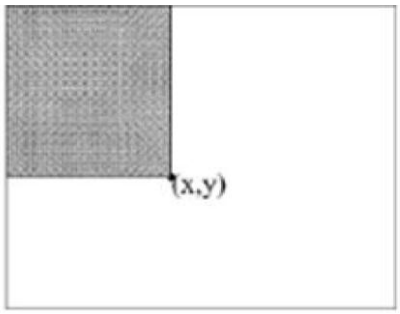


Fig3.Integral image

AdaBoost is used to generate strong classifiers from weak classifiers. This can reduce false positive detection rate and also reduce the redundant features $w_c$.

$$w_c = \begin{cases} 1: identified\ feature \\ 0: not\ identified\ any\ feature \end{cases} \quad (1)$$

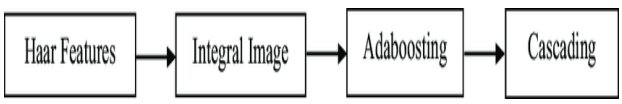


Fig4.Stages of VJ face detection

Cascade structure is used to remove the false positive images and also used to check the presence of face in a particular portion of an image.

*2.3 Feature Extraction*

Feature Extraction plays a vital role in the FER system by improving the accurateness of facial emotion images. There are various extraction techniques such as Local binary pattern (LBP), Gray-level co-occurrence matrix (GLCM), Gray Level Weight Matrix (GLWM), Traditional Gabor Filter (TGF) and DBWP which are used for classification of texture. In this paper, two feature extraction techniques such as GLCM and LBP are proposed for texture classification.

2.3.1 Gray-level co-occurrence matrix (GLCM)

GLCM is an analytical texture technique that creates a matrix by spatial relationship of pixels with the dependence matrix of gray-level. After creation of the matrix, statistical measures are extracted by approximating recurrent pixel pairs using precise value in an image. This is a square matrix with Ng dimension, where Ng is the number of gray image levels. Element [i, j] of the matrix is generated by calculating the number of times an occurrence of value i of neighboring pixel with value j of pixel, then the whole matrix is divided by the total number of such relationships [3]. Thus, every entry is considered as probability that a neighbor pixel with value i will occur near pixel of value j. GLCM can be defined as

$$P^{\theta}(x,y) = Pr(I(p_1) = x \wedge I(p_2) = y \wedge \| p_1 - p_2 \| = d) \quad (2)$$

Where P is the probability with positions p1 and p2 in the grayscale image I.

*2.3.2 Local binary pattern (LBP)*

LBP is an efficient-texture analysis operator which marks the image pixels by thresholding each neighboring pixel and then results in a binary number. Owing to power discrimination and computationally simple [4], LBP is the popular approach in real-time applications for its robustness to monotonically changes in gray-scale by change in light intensity in fig.4.

| | | |
|---|---|---|
| $i_0$ | $i_1$ | $i_2$ |
| $i_3$ | $i_4$ | $i_5$ |
| $i_6$ | $i_7$ | $i_8$ |

(a)

| | | |
|---|---|---|
| 1 | 2 | 4 |
| 128 | 0 | 8 |
| 64 | 32 | 16 |

(b)

Fig5.LBP

*2.4 Principal Components Analysis*

PCA is used to extract features from face images and resized face images. It is also used to reduce dimensionality of face image. It is mathematical approach which converts set of correlated N face images into a set of Eigen faces images.

A set of face images are in training, then it is denoted with large Eigen values by the greatest Eigen faces for better approximation of the face. After result of Eigen faces, each face image can be indicated by a linear permutation of Eigen faces, then it symbolized as vectors. The input image features are matched with standard features in dataset for recognition.

*2.5 Classifiers*

There are 3 different types of classifiers; they are KNN, CNN and SVM for classification of textures.

i.KNN classifier is a method to classify objects based on k. k implies the number of adjacent neighboring sample in feature space. If k=1, then the object is allocated to its adjacent neighbor class [11]. The neighbors are occupied from accurately classified set of known objects. This is taken as training set for the classifier.

ii. CNN is the most used ANN algorithms in CV. CNN contains of a series of convolutional layers, the output which is related only to local regions in the input. This is done by sliding-filter, or weighted-matrix, over the input and at every point calculating the convolution product between the input and filter [5].

## 3. EXPERIMENTAL RESULTS

In this paper, proposed work is implemented using Matlab. For this, input datasets is taken from Kaggle and KDEF databases [12-13] on facial expression detection. This dataset consists of 215 images with 7 facial emotions such as happy, sad, surprise, angry, disgust, fear and neutral. These datasets can be trained by loading the datasets

Initially preprocessing is used to remove the unwanted images and also smoothens the images from input datasets using Gabor filter. For FER, VJ face detection technique is used to identify each facial emotion in input dataset images. After FER, image features are extracted using GLCM, LBP method. To distinguish the texture information of image,GLCM and LBP technique is used to improve the classification performance. PCA is the feature selection technique that is used to extract features from face images and resized face images. It is used is used to reduce dimensionality of face data, the whole dataset is separated for training and classification. Out of number of face images, 75% of the face images are trained and remaining of 25% is tested using CNN classifier by predicting seven emotions.

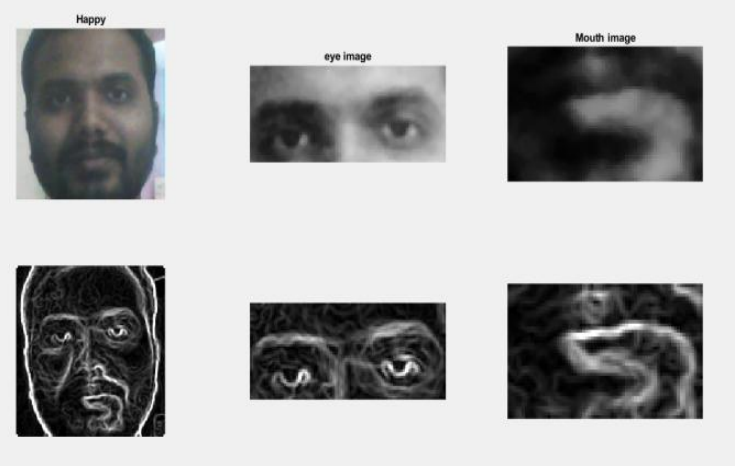


Fig6.Different face emotions result

The above figure 6 show face interested region and its corresponding gradient image. The classification of emotion is displayed above the input image.

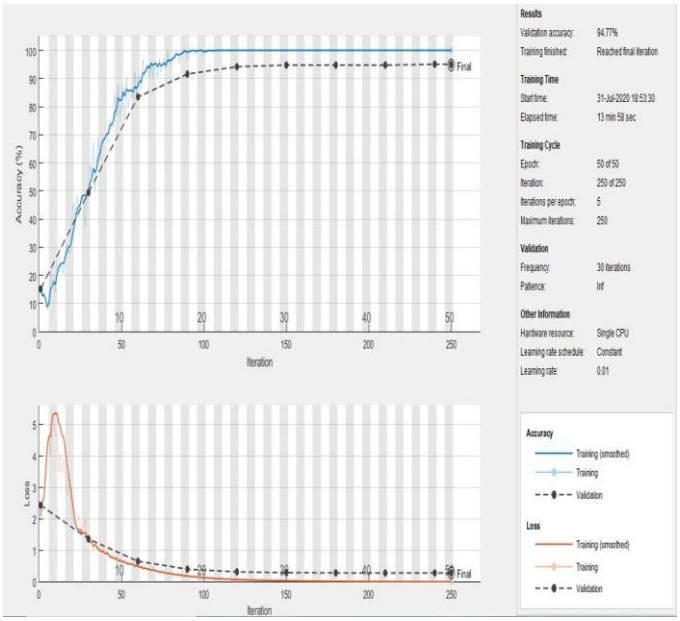


Fig7.Accuracy and log-loss plot of CNN

The confusion Matrix of the Facial Emotion Detection is constructed in Table1 using the complete Merged Image.

Table1.Confusion matrix of CNN

| | Anger | Disgust | Fear | Happy | Sad | Surprise |
|---|---|---|---|---|---|---|
| Anger | 75.00 | 0.00 | 0.00 | 16.67 | 0.00 | 0.00 |
| Disgust | 0.00 | 83.33 | 8.33 | 0.00 | 16.67 | 0.00 |
| Fear | 0.00 | 0.00 | 75.00 | 0.00 | 0.00 | 0.00 |
| Happy | 0.00 | 0.00 | 0.00 | 83.33 | 0.00 | 0.00 |
| Sad | 25.00 | 16.67 | 0.00 | 0.00 | 83.33 | 0.00 |
| Surprise | 0.00 | 0.00 | 16.67 | 0.00 | 0.00 | 100.0 |

Table 2 shows accuracy outcome of CNN is comparatively good than other classifiers.

Table2.Comparative result of different classifiers

| Methods | CNN | ANN | KNN | SVM |
|---|---|---|---|---|
| Accuracy | 94.77% | 75.00% | 66.23% | 65.74% |

## 4. CONCLUSION

FER is the important improvement in the area of CV and AI.The proposed work is used to identify the facial appearance from the image and classify different feelings. For FER, is used to identify each facial emotion in input dataset images, then histogram base image enhancement is applied to enhancing brightness of image followed by LPB, GLCM and PCA based feature extraction which extracted details from face images.PCA is used to reduces dimensionality of face data, The whole system is trained and classified using CNN classifiers for final emotion recognition. Accuracy of this work is about 94.77% is achieved for six emotions. In future, emotion recognition can be developed for real time.

## 5. REFERENCES

[ 1] Meijing Li, X. Yu, K. H. Ryu, S. Lee, and N. Theera- Umpon, "Face recognition technology development with Gabor, PCA and SVM methodology under illumination normalization condition," Clust. Computer. vol. 21, no. 1, pp. 1117–1126, Mar. 2018.

[ 2] L. Shen, H. Wang, L. D. Xu, X. Ma, S. Chaudhry, and W. He, "Identity management based on PCA and SVM," Information System Front., vol. 18, no. 4, pp. 711–716, Aug. 2016.

[ 3] Borovykh A, Bohte S, and Oosterlee CW, "Conditional time series forecasting with convolutional neural networks," preprintarXiv: 170304691, Sept. 2018.

[ 4] Vamshi Krishna Gudipati, Oindrila Ray Barman, Mofica Gaffoor, Harshagandha, Abdelshakour Abuzneid, "Efficient Facial Expression Recognition Using Adaboost and Haar Cascade Classifiers," IEEE Annual Connecticut Conference on Industrial Electronics, Technology and Automation (CT-IETA), 14-15 Oct. 2016.

[ 5] Anchal Garg, Rohit Bajaj, "Facial expression recognition & classification using hybridization of ICA, GA, and Neural Network for Human-Computer Interaction," Journal of Network Communications and Emerging Technologies Vol. 2, Issue 1, May 2015.

[ 6] Namrata Mahajan, Harshad Mahajan, "Emotion Detection Algorithm," International Journal of Electrical and Electronics Research, Vol. 2, Issue 2, pp: 56–60, April 2014.

[ 7] Ihor Paliy, Face Detection Using Haar - like Features Cascade and Convolutional Neural Network, IEEE International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), 19-23 Feb. 2008.

[ 8] Govind U. Kharat, Sanjay V. Dudul, "Emotion Recognition from Facial Expression Using Neural Networks," IEEE Conference on Human System Interactions, 25-27 May 2008.

[ 9] Abidin, T. and Perrizo, W. SMART-TV, "A Fast and Scalable Nearest Neighbor Based Classifier for Data Mining," Proceedings of ACM SAC-06, Dijon, France,. ACM Press, New York,NY, pp. 536-540, April 23-27, 2006.

[ 10] Caifeng Shan, Shaogang Gong and Peter W. McOwan, "Robust Facial Emotion Recognition Using Local Binary Patterns," IEEE International Conference on Image Processing, 14th Sept. 2005.

[ 11] Bouchra Abboud, Franck Davoine, Mo Dang, "Facial Emotion recognition and synthesis based on an appearance model," Vol.19, Issue 8, pp. 723-740, Sept. 2004.

[ 12] Yongsheng Gao, Maylor K,. H. Leung, Siu Cheung Hui, and Mario W. Tananda, "Facial Emotion Recognition from Line-Based Caricatures," IEEE-PART A: Systems And HumanS, VOL. 33, NO. 3, MAY 2003.

[ 13] M. Nehru and S. Padmavathi, "Illumination invariant face detection using viola jones algorithm," 4th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 1-4, 2017,

[ 14] K. Dang and S. Sharma, "Review and comparison of face detection algorithms," 7th International Conference on Cloud Computing, Data Science and Engineering, Confluence, Noida, pp. 629-633, 2017.

[ 15] Debishree Dagar, Abir Hudait, H. K. Tripathy, M. N. Das, "Automatic Emotion Detection Model from Facial Expression," International Conference on Advanced Communication Control and Computing Technologies, (ICACCCT), 25-27 May 2016.

[ 16] Ioanna-Ourania and George A. Tsihrintzis, "An improved Neural Network based Face Detection and Facial Expression classification system," IEEE international conference on Systems Man and Cybernetics, 2004.

[ 17] Sreevatsa n. A. N, Sat hish Kumar. K. G, Rakeshsharma. S and Mohd. Mansoor Roomi, "Emotion Recognition from Facial Expressions: A Target Oriented Approach using Neural Network," ICVGIP, Proceedings of the Fourth Indian Conference on Computer Vision, Graphics & Image Processing, 2004.

[ 18] Sreevatsa n. A. N, Satish Kumar. K. G, Rakeshsharma . S and Mohd. Mansoor Roomi. "Emotion Recognition from Facial Expressions: A Target Oriented Approach using Neural Network," ICVGIP, Proceedings of the Fourth Indian Conference on Computer Vision, Graphics & Image Processing, 2004.

[ 19] Zhang Z., Lyons M., Schuster M., Akamatsu S., "Comparison between Geometry- Based and Gabor Wavelets-Based Facial Expression Recognition Using Multi- Layer Perceptron," Proceedings of International Conference on Automatic Face and Gesture Recognition, pp. 454-459, 1998.

[ 20] Kobayashi H., Hara F., "Facial Interaction between Animated 3D Face Robot and Human Beings," Proceedings of International Conference on Systems, Man, Cybernetics, 732-737, 1997.

[ 21] Padgett C., Cottrell G.W., "Representing Face Images for Emotion Classification," Proceedings of Conference on Advances in Neural Information Processing Systems, pp. 894-900, 1996.

[ 22] Zhao J., Kearney G., "Classifying Facial Emotions by Back propagation Neural Networks with Fuzzy Inputs," Proceedings of Conference on Neural Information Processing, pp. 454-457, 1996.

[ 23] Ekman P., Huang T. S., Sejnowski T. J., Hager J. C., "Final Report to NSF of the Planning Workshop on Facial Expression Understanding, "July 30 to August 1, 1992

[ 24] Kobayashi, H., Hara, F., "The Recognition of Basic Facial Expressions by Neural Network," Proceedings of International Joint Conference on Neural Network, pp. 460-466, 1991.
[ 25] T. Kanade, "Computer Recognition of Human Faces, Birkhauser, Basel, 10.1007/978-3-0348-5737-6, Germany, 1973.