International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362
IJARST

# Efficient Way To Recognize The Facial Emotion Using Deep Convolutional Neural Network

**Mr. S. T. Saravanan**
Computer Science and Engineering (*JUTUH*)
Sphoorthy Engineering College (*JNTUH*)
Hyderabad, India
kannansts@gmail.com

**Dr.Subba Rao Kolavennu**
Computer Science and Engineering (*JUTUH*)
Sphoorthy Engineering College (*JNTUH*)
Hyderabad, India
Profrao99@gmail.com

**Borampet.Bhavani**
Computer Science and Engineering (*JUTUH*)
Sphoorthy Engineering College (*JNTUH*)
Hyderabad, India
19n81a05j3bhavani@gmail.com

**Koti.Veena**
Computer Science and Engineering (*JUTUH*)
Sphoorthy Engineering College (*JNTUH*)
Hyderabad, India
veenakoti14@gmail.com

**Nishitha Madhavaram**
Computer Science and Engineering (*JUTUH*)
Sphoorthy Engineering College (*JNTUH*)
Hyderabad, India
nishithamadhavaram12345@gmail.com

**1.ABSTRACT**
The rapid growth of artificial intelligence has contributed a lot to today's world. As traditional algorithms failed to meet human needs in real-time, Machine learning and deep learning algorithms have received great success in different applications like classification systems, pattern recognition, etc. Emotion is an important factor which helps us to determine the human thoughts, behaviour, and feeling. An emotion recognition system can be built by utilizing the benefits of deep learning and different applications like feedback analysis, face unlocking, etc. This can be implemented with good accuracy. The main focus of this work is to create a Convolutional Neural Network (CNN) model that classifies 5 different types of human facial emotions. The model is trained, tested, and validated using the manually collected image data set.
***Keywords* — Convolutional Neural Network (CNN),Machine learning,Emotion Recognition.**

## *2.INTRODUCTION*

Facial expressions play a vital role in identifying the feelings of humans because it corresponds to emotions.In many cases(roughly in 55% of cases) [1], the facial expression is a nonverbal way of expressing emotions, and it can be considered as concrete evidence which is used in lie detection[2].

The current approaches mainly focus on facial investigation keeping background intact and hence gives rise to a lot of unnecessary and misleading features that mislead CNN training process. This manuscript focuses on five vital facial expression classes reported, which are anger, sad, happy, feared, and surprised[3]. The FERC algorithm presented in the current manuscript aims for expressional examination and to characterize the given image into these five major emotion classes.

Facial expression detection can be described by using two major approaches. The first one is distinguishing expressions [4] that are identified with an explicit classifier, and the second one is making characterization dependent on the extracted facial features [5]. In the facial action coding system (FACS) [6], action units are used as markers of expressions. These AUs were discriminable by facial muscle changes.

## *EXISTING SYSTEM*

All the other methods have just gone for a combined approach of removing background and face expression detection in a single CNN network. If we solve both issues separately then it reduces complexity and also the tuning time. Zao [7] have achieved maximum accuracy but at the cost of 22 layers neural network. Training such a large network is a very time-consuming job. Compared to existing

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

methods, only this method has key-frame extraction method, whereas others have only gone for the last frame. Jung [8] tried to work with fixed frames which make the system not so efficient .In every other method only 10 folds were used but here we can use 25 folds because of small network size. This method is much faster, compared to all the other methods and also gives output with higher accuracy.

| TITLE | NUMBEROFMOOD | KEYFRAME | NSIZE | ACCURACY | FOLD |
|---|---|---|---|---|---|
| FERC | 5 | Edge based | 8 | 96 | 25 |
| Zaoetal | 6 | Last frame | 22 | 99.3 | 10 |
| jungetal | 7 | Fixed frame | 4 | 91.44 | 10 |

Table. 1 Comparison table with the similar methods reported in the literature

## *PROPOSED SYSTEM*

Some unique features of this method is skin tone-based feature and Hough transform for circles-in-circle filters. The skin tone is a pretty fast and robust method of preprocessing the input data. These are the new functionalities which makes this the most preferred method for mood detection.The model also has comparable training accuracy and validation accuracy which convey that the model is having the best fit and is generalized to the data. The model uses an Adam optimizer to reduce the loss function.It can even classify the faces with makeup and facial hair.

## *3. LITERATURE SURVEY*

Facial expressions are the common signals for all humans to convey the mood. There are many attempts made to prepare an automatic facial expression analysis tools [9] as it is used in many fields like robotics, driving assist systems,medicine, and lie detector [10–12]. Since the twentieth century, Ekman et al. [13] defined 7

basic emotions which are anger, happy, fear, sad, contempt [14], disgust, and surprise. In a recent study, Sajid et al. found out the impact of facial asymmetry in estimating age [15]. They found that right-face asymmetry is better when compared to the left-face asymmetry. The big issue with face detection is face pose appearance. Ratyal et al. gave the solution for variability in the facial pose appearance. The three-dimensional pose invariant approach was used by them[16, 17]. There are various issues like excessive makeup [18] pose and expression [19] which are solved using these convolutional networks. Recently, researchers have made extraordinary accomplishment in facial expression detection [20–22], which led to the advancement of research, in the field of facial expression. Also, the development in the 2 areas which are computer vision [23] and machine learning [24] makes emotion identification much more accurate and accessible. The recognition of facial expression is growing rapidly. The applications are human–computer interaction [25], drunk driver recognition [26], psychiatric observations [27] and lie detector [28].

## *4. IMPLEMENTATION*

Convolutional neural network (CNN) is the most widely used method for analyzing images. CNN is very different from a multi-layer perceptron (MLP) as CNN have hidden layers. These hidden layers are called convolutional layers. This method is based on a two-level CNN framework. The first level is background removal [29], which is used to extract emotions from an image, as shown in Fig. 1. The conventional CNN network module is used to extract the primary expressional vector (EV). The expressional vector (EV) is generated by tracking down relevant and important facial points. EV is related directly to changes in expression. The EV is obtained with the help of a basic perceptron unit. This unit is applied to a background-removed face image. In the proposed model, the last stage is non-convolutional perceptron layer. Every convolutional layer receives the input data (or image)then it transforms it, and then sends it as output to the next level. This transformation is called convolution operation, which is shown in Fig. 1. Every convolutional layer used is capable of pattern detection. Four filters were used in each convolutional layer. The input image sent to the first-part CNN which is used for background removal, consists of shapes, textures, edges, and objects along with the face. The circle detector, edge detector, and corner detector filters are used at the starting point of the convolutional layer 1. Once the face has been detected, the second-part CNN filter detects facial features, such as eyes, lips, ears, nose, and cheeks. The edge detection filters used in this layer which is shown in Fig. 1. The second-part CNN consists of layers with 3 × 3 kernel matrix, e.g., [0.25, 0.17, 0.9; 0.89, 0.36, 0.63; 0.7, 0.24, 0.82]. These numbers are selected between 0 and 1 initially then these numbers are optimized for the EV detection. We used minimum error decoding to optimize filter values. Once the filter is tuned by the supervisory learning, it is then applied to the face in which the background is removed which is the output image of the first-part CNN, for detection of different facial parts like eye, nose, lips, ears, etc. To generate the EV matrix, all 24 different facial features are extracted. The EV feature vector is the values of normalized Euclidian distance between each face part.

International Journal For Advanced Research
In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362
IJARST

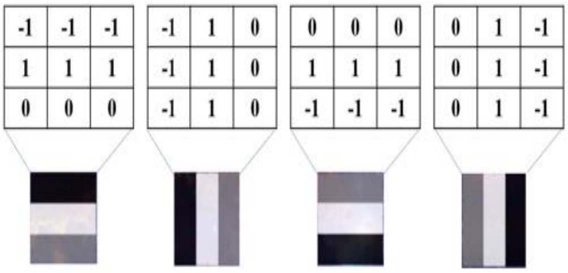


Fig. 1 Horizontal and Vertical edge detector filter which is in matrix form used at layer 1 of background removal CNN (first-part CNN).

### 4.1 Background removal

Once we get the input image, skin tone detection algorithm [30] is applied then human body parts from the image are extracted. This skin tone-detected output image is a binary image and is used as the feature, for the first layer of background removal CNN which is also referred to as the 1st part CNN in this manuscript. This skin tone detection depends on the range between 85 and 140, the Cr value should be between 135 and 200. The set of values mentioned in the above line was chosen by the trial-and-error method and worked for almost all of the skin tones available. We found that the skin tone detection algorithm has very low accuracy if the input image is grayscale. To improve accuracy CNN also uses a filter called the circles-in-circle filter. This filter operation uses Hough transform values for each circle detection. To maintain uniformity irrespective of the type of input image, the Hough transform was always used as the second input feature to background removal CNN.

### 4.2 Convolutional Filter

As shown in Fig. 2 for each convolution operation, the entire image is divided into overlapping 3 × 3 matrices, and then the corresponding 3 × 3 filter is convolved over each 3 × 3 matrix obtained from the image. The sliding and taking dot product operation is known as 'convolution' and hence the name 'convolutional filter' is given. During the convolution, the dot product of both 3 × 3 matrices is computed and stored at a corresponding location, e.g., (1,1) at the output, as shown in Fig. 2. The entire output matrix is calculated. Then this output is passed to the next layer of CNN for another round of convolution. simple perceptron is the last layer of face feature extracting CNN which tries to optimize values of exponent and scale factor depending upon deviation from the ground truth.

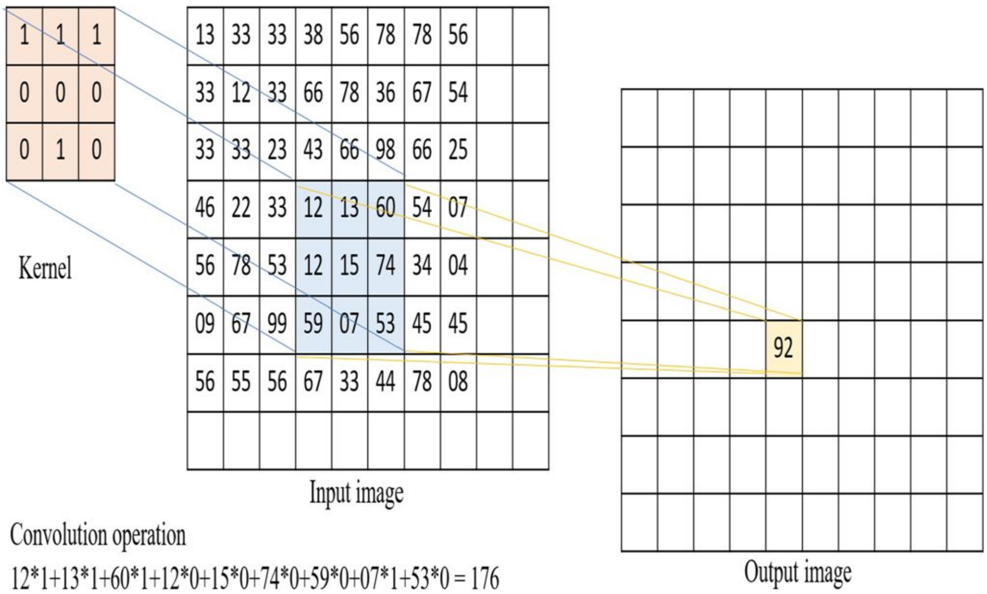


**Fig. 2** The Convolution filter operation is done with 3 × 3 kernel. Each pixel from input image and its eight of the neighboring pixels are multiplied with corresponding value in kernel matrix, and then finally, all multiplied values are added to achieve the final output .

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

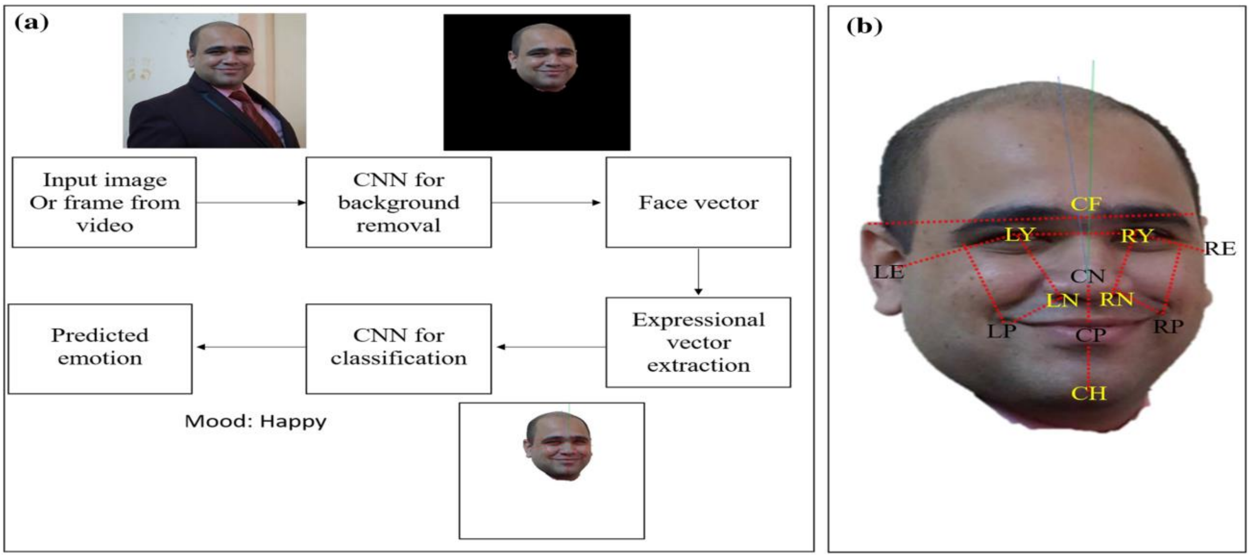


**Fig. 3** a) The Block Diagram of FERC. The input image is (taken from the camera or) extracted from the video. The input image is then passed to the first-part of CNN for background removal. After background removal, a facial expressional vector (EV) is generated. Another CNN (the second part of CNN) is applied with the supervisory model obtained from the ground-truth database. At last, emotion from the current input image is detected. b )Facial vectors are marked on background-removed face. Here, nose (N), eyes (Y), lip (P), forehead (F) are marked using edge detection and also nearest cluster mapping. The positions like left, right, and center are represented using the L, R, and C, respectively.

## *5. CONCLUSIONS*

This method uses CNN and supervised learning (feasible due to big data). The main advantage of this algorithm is that it works with different kinds of orientations (less than 30°) due to the unique 24-digit long EV feature matrix. Another advantage is the background removal which is useful in accurately determining the emotions. This could be the starting step, for emotion-based applications like lie detector and mood-based learning for students, etc. Facial hair causes a lot of issues. But it can be solved here as we are training this algorithm with images having facial hair too. The accuracy of our algorithm is very high. Another limitation is shadow on the face. It is overcome by automated gamma correction on images by assuming facial symmetry. The ideal number of images was found out to be in the range of 2000–10,000 for this to work properly with high accuracy.

## *REFERENCES*

[1] MeteMehrabian A (2017) Nonverbal communication. Routledge, London

[2] Bartlett M, Littlewort G, Vural E, Lee K, Cetin M, Ercil A, Movellan J (2008) Data mining spontaneous facial behavior with automatic expression coding. In: Esposito A, Bourbakis NG, Avouris N, Hatzilygeroudis I (eds) Verbal and nonverbal features of human–human and human–machine interaction. Springer, Berlin, pp 1–20

[3] Russell JA (1994) Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. Psychol Bull 115(1):102

[4] Gizatdinova Y, Surakka V (2007) Automatic detection of facial landmarks from AU-coded expressive facial images. In: 14th International conference on image analysis and processing (ICIAP). IEEE, pp 419–424

[5] Liu Y, Li Y, Ma X, Song R (2017) Facial expression recognition with fusion features extracted from salient facial areas. Sensors 17(4):712

[6] Ekman R (1997) What the face reveals: basic and applied stud ies of spontaneous expression using the facial action coding system (FACS). Oxford University Press, New York

[7] Zhao X, Liang X, Liu L, Li T, Han Y, Vasconcelos N, Yan S (2016) Peak-piloted deep network for facial expression recognition. In: European conference on computer vision. Springer, pp 425–442

[8] Jung H, Lee S, Yim J, Park S, Kim J (2015) Joint fne-tuning in deep neural networks for facial expression recognition. In: Proceedings of the IEEE international conference on computer vision. pp 2983–2991
[9] Zafar B, Ashraf R, Ali N, Iqbal M, Sajid M, Dar S, Ratyal N (2018) A novel discriminating and relative global spatial image representation with applications in CBIR. Appl Sci 8(11):2242
[10] Ali N, Zafar B, Riaz F, Dar SH, Ratyal NI, Bajwa KB, Iqbal MK, Sajid M (2018) A hybrid geometric spatial image representation for scene classifcation. PLoS ONE 13(9):e0203339
[11] Ali N, Zafar B, Iqbal MK, Sajid M, Younis MY, Dar SH, Mahmood MT, Lee IH (2019) Modeling global geometric spatial information for rotation invariant classifcation of satellite images. PLoS ONE 14:7
[12] Ali N, Bajwa KB, Sablatnig R, Chatzichristofs SA, Iqbal Z, Rashid M, Habib HA (2016) A novel image retrieval based on visual words integration of SIFT and SURF. PLoS ONE 11(6):e0157428
[13] Ekman P, Friesen WV (1971) Constants across cultures in the face and emotion. J Personal Soc Psychol 17(2):124
[14] Matsumoto D (1992) More evidence for the universality of a contempt expression. Motiv Emot 16(4):363
[15] Sajid M, Iqbal Ratyal N, Ali N, Zafar B, Dar SH, Mahmood MT, Joo YB (2019) The impact of asymmetric left and asymmetric right face images on accurate age estimation. Math Probl Eng 2019:1–10
[16] Ratyal NI, Taj IA, Sajid M, Ali N, Mahmood A, Razzaq S (2019) Three-dimensional face recognition using variance-based registration and subject-specifc descriptors. Int J Adv Robot Syst 16(3):1729881419851716
[17] Ratyal N, Taj IA, Sajid M, Mahmood A, Razzaq S, Dar SH, Ali N, Usman M, Baig MJA, Mussadiq U (2019) Deeply learned pose invariant image analysis with applications in 3D face recognition. Math Probl Eng 2019:1–21
[18] Sajid M, Ali N, Dar SH, Iqbal Ratyal N, Butt AR, Zafar B, Shafique T, Baig MJA, Riaz I, Baig S (2018) Data augmentation-assisted makeup-invariant face recognition. Math Probl Eng 2018:1–10
[19] Ratyal N, Taj I, Bajwa U, Sajid M (2018) Pose and expression invariant alignment based multi-view 3D face recognition. KSII rans Internet Inf Syst 12:10
[20] Xie S, Hu H (2018) Facial expression recognition using hierarchical features with deep comprehensive multipatches aggregation convolutional neural networks. IEEE Trans Multimedia 21(1):211
[21] Danisman T, Bilasco M, Ihaddadene N, Djeraba C (2010) Automatic facial feature detection for facial expression recognition. In: Proceedings of the International conference on computer vision theory and applications, pp 407–412. https://doi.org/10.5220/0002838404070412
[22] Mal HP, Swarnalatha P (2017) Facial expression detection using facial expression model. In: 2017 International conference on energy, communication, data analytics and soft computing (ICECDS). IEEE, pp 1259–1262
[23]Kong SG, Heo J, Abidi BR, Paik J, Abidi MA (2005) Recent advances in visual and infrared face recognition—a review. Comput Vis Image Underst 97(1):103
[24] Xue Yl, Mao X, Zhang F (2006) Beihang university facial expression database and multiple facial expression recognition. In: 2006 International conference on machine learning and cybernetics. IEEE, pp 3282–3287
[25] Kim DH, An KH, Ryu YG, Chung MJ (2007) A facial expression imitation system for the primitive of intuitive human robot interaction. In: Sarkar N (ed) Human robot interaction.IntechOpen, London
[26] Kumar KC (2012) Morphology based facial feature extraction and facial expression recognition for driver vigilance. Int J Comput Appl 51:2
[27] Ernst H (1934) Evolution of facial musculature and facial expression. J Nerv Ment Dis 79(1):109
[28] Hernández-Travieso JG, Travieso CM, Pozo-Baños D, Alonso JB et al (2013) Expression detector system based on facial images. In: BIOSIGNALS 2013-proceedings of the international conference on bio-inspired systems and signal processing
[29] Cowie R, Douglas-Cowie E, Tsapatsoulis N, Votsis G, Kollias S, Fellenz W, Taylor JG (2001) Emotion recognition in human– computer interaction. IEEE Signal Process Mag 18(1):32
[30] Hsu RL, Abdel-Mottaleb M, Jain AK (2002) Face detection in color images. IEEE Trans Pattern Anal Mach Intell 24(5):696