International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 3, May-June-2020, ISSN (Online): 2395-566X

Emotion Recognition Using Machine Learning Techniques and Edge Detection

Manisha Dey
Department of ECE,
CMR Institute of Technology,
Bengaluru, India

Raunak Sinha
Department of ECE
CMR Institute of Technology,
Bengaluru, India

Ritwik Karan
Department of ECE
CMR Institute of Technology,
Bengaluru, India

Associate Prof. Naveen Kumar G N
Department of ECE
CMR Institute of Technology,
Bengaluru, India

***Abstract* – Applied Emotions are natural states related with the tangible framework invited on by neurophysiological changes diversely associated with thoughts, notions, lead responses and level of bliss or frustration. Sentiments can be perceived as outrage, disgust, dread, bliss, dismal, shock and unbiased. Programmed outward appearance acknowledgment framework has numerous applications including, however not constrained to, human conduct seeing, direct checking of the impacts of an amusement substance or items on the crowd or client, location of mental issue, and engineered human articulations. Acknowledgment of outward appearance by PC with high acknowledgment rate is as yet a difficult errand. So, we propose a canny strategy utilizing Gabor channels for pre-preparing and CNN (Convolutional neural network) for arrangement.**

***Keywords* – Emotions, Gabor filter, CNN, channels.**

I. INTRODUCTION

"Machine Learning will robotize occupations that many people thought must be finished by persons"- Dave Waters. Feeling is a psychological state which includes a ton of practices, activities, contemplations and emotions. The book, "The Expression of the Emotions in human and Animals" was composed by Charles Darwin in 1969, after distinguishing the feelings in various gatherings of individuals regardless of the social contrasts. Ekman and Friesen grouped six passionate articulations to be all inclusive: astonishment, disgust, pity, satisfaction, and distress. . After this there was no stopping, many authors have researched on this subject and have paved the way towards modern day research work.

In [1], there is a distinctive correlation between different strategies, for example, PCA, LEM, Gabor Filter, Local Binary, Pattern, Neural Network, ICA, LDA and SVM with the assistance of suitable Datasets to recognize human facial feelings. Feelings have been perceived based on exactness and computational time. In any case, the most well-known disadvantage was in term of acknowledgment rate or timing.
. So they concluded on proposing combination of two or more techniques for achieving the most accurate recognition rate. Success of methods depends on pre-processing on the images because of illumination changes and how easily the features can be extracted. One of the techniques in the above-mentioned paper, Principal Component Analysis (PCA) accompanied by Viola and Jones algorithm face expression has been seen being used in [2]. This technique proposed utilizes Viola and Jones calculation and principal component analysis that attempts to coordinate a picture concerning to expression of the face. The proposed method consists of an image to be resized to N x N, changed into grayscale, identifying facial feature points, edge detection, mapping of image etc.

In [3], there has been a conversation on identification of feelings in profoundly undermined boisterous condition. The methodology includes expulsion of clamor from the picture by the Wiener Filter. Nonetheless, results show that the commotions are not expelled appropriately and in fact they are inserted to the picture record so for evacuating the rest of the dot clamor Mutation based microorganisms Foraging streamlining strategies must be utilized. The Viola Jones classifier technique gives great outcomes for distinguishing faces however it requires some investment for the recognition and it doesn't generally give exact outcomes.

In [4], skin shading based following procedure is observed. It is regularly perceived that a static skin shading model is utilized which is gotten either from a disconnected arrangement of library pictures or the initial hardly any edges of a video stream for a similar reason. These models henceforth are not strong to light changes or imaging conditions. So in this paper there is a proposition of a versatile skin shading model dependent on the Gaussian blend model to make up for the issues brought about by the evolving conditions. At first Expectation amplification calculation has been utilized to assess the number and loads of skin shading bunches. Further the model adjusts to changes in imaging condition by progressively changing the model boundaries utilizing spatial and fleeting imperatives. The technique has been

supposed to be working viably in following hand and face areas.

In [5], there is a conversation of utilizing profound learning strategy (convolutional neural systems) to recognize the key seven human feelings: outrage, nauseate, dread, bliss, bitterness, shock and impartiality. The Kaggle (Facial Expression Recognition Challenge) and Karolinska Directed Emotional Faces dataset has been utilized. They have investigated the VGG-16 and ResNet50 models for perceiving facial feelings utilizing convolutional neural systems. The help vector machine multiclass classifier has been utilized as the gauge, which had a precision execution of 31.8%.

Further there is additionally a work of group and move learning procedures to use the exactness execution. Accordingly, the precision utilizing outfit learning was 67.2% and with move learning was 78.3%.

[6], discusses how the prior FER approaches don't completely consider and use the highlights of facial component and muscle developments, which speak to critical static and dynamic just as geometric qualities of facial feelings. Along these lines, it proposes a way to deal with make up for this confinement by utilizing patch-based 3D Gabor highlights, choosing the significant fixes, and performing patch coordinating activities accordingly giving "notable" separation highlights. Because of the thought of facial component and muscle developments, the trial results show a high right acknowledgment rate (CRR), promising outcomes under face enrollment mistakes, noteworthy execution upgrades, and furthermore quick preparing time. When contrasted with the presentation of different methods the outcomes affirm that the proposed approach accomplishes the most noteworthy CRR on the JAFFE database and is additionally among the Cohn-Kanade (CK) database top entertainers on the.

[7], has recommended a completely mechanized, multistage framework for constant acknowledgment of facial feelings. It utilizes facial movement to describe monochrome frontal perspectives on outward appearances and furthermore works viably in powerful and jumbled scenes It can perceive the six general feelings related with one of a kind outward appearance, to be specific joy, trouble, nauseate, shock, dread, and outrage. A spatial proportion layout tracker calculation is utilized to find the countenances. A constant usage of a hearty angle model is utilized to decide the optical progression of the face. The feeling acknowledgment framework at that point does a normal of facial speed data over distinguished locales of the face and takes proportions of this found the middle value of movement to offset unbending head movement. The henceforth decided movement marks are then ordered utilizing Support Vector Machines (SVM) as either non expressive or as one of the six fundamental feelings.

In [8], presents an interesting profound neural system engineering for facial feeling acknowledgment. The proposed approach comprises of two convolutional layers and afterward followed by max pooling and four initiation layers. The Inception layers are utilized in order to expand the profundity and width of the system while not adding to the computational multifaceted nature. It is a solitary segment engineering which accepts facial pictures as the info and does feeling characterization into both of the seven fundamental feelings. This methodology was tried on seven notable openly accessible databases. The outcomes were good and confirmed that the proposed organize is superior to a few different procedures.
The proposed approach depends on Convolutional Neural Network (CNN) where the system learns without anyone else the important highlights that map the information picture to the relating marks or feelings. The chart is conveyed to confirm the viability of the framework.

II. METHODOLOGY

1.Gabor filter

The Gabor channel, named after Dennis Gabor, is a straight channel used for surface assessment, feature extraction, edge detection which infers that it basic examinations whether there are a specific repeat content in the image in express headings in a restricted region around the point or locale of investigation. The traits of explicit cells in the visual cortex of certain all around advanced animals can be approximated by these channels. These channels have been seemed to have perfect restriction properties in both spatial and repeat territory and thusly are fit for surface division issues. Gabor channels are unprecedented classes of band pass channels, i.e., they license a particular 'band' of frequencies and reject the others.

$$g(x,y,\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right)\exp\left(i\left(2\pi\frac{x'}{\lambda}+\psi\right)\right) \quad (1)$$

Real part:

$$g(x,y,\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right) \quad (2)$$

And imaginary part:

$$g(x,y,\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right)\sin\left(2\pi\frac{x'}{\lambda}+\psi\right) \quad (3)$$

where

$$x' = x\cos\theta + y\sin\theta \quad (4)$$

and

$$y' = -x\sin\theta + y\cos\theta \quad (5)$$

In the above equation,
λ — Wavelength of the sinusoidal component.
Θ — The orientation of the normal to the parallel stripes Gabor function.
Ψ — The phase offset of the sinusoidal function.

σ—The sigma/standard deviation of the Gaussian envelope
ɣ — The spatial aspect ratio and specifies the ellipticity of the support of Gabor function.

The proposed filter is applied on the original image as shown in figure 1, we obtain another image which is filtered image as shown in figure 2.

In Figure 2, the filtered image is shown which is the result of application of gabor filter with the given parameters on the originla image as shown in figure 1.:
(x,y)= (8,8), σ =0.5, θ = π / 4, λ=2.8, ɣ=0.7, Ψ= 0.3

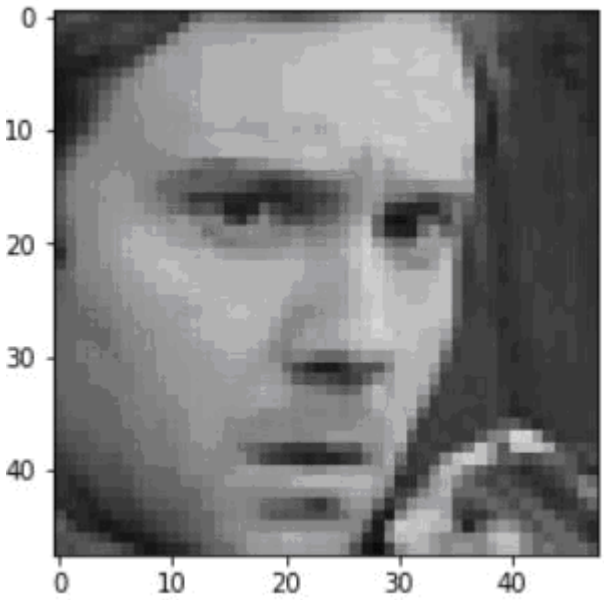

Fig .1. Original image from dataset.

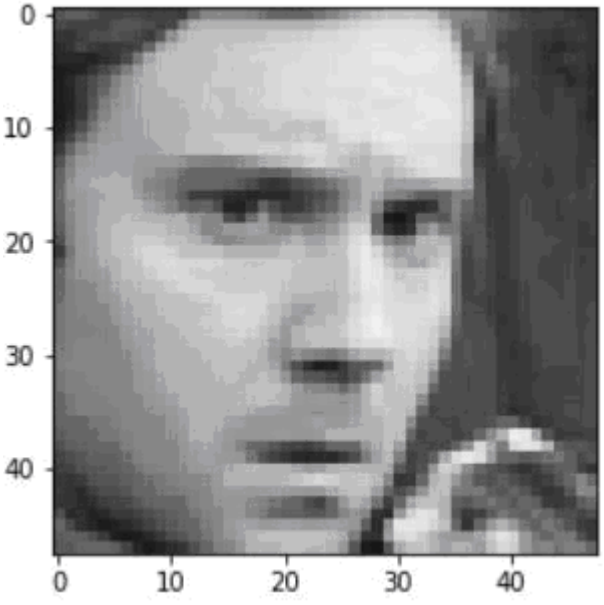

Fig 2. Filtered image.

2. Dataset Used
For this situation we are utilizing FER2013 database of pre-edited, 48-by-48-pixel grayscale pictures of appearances each marked with one of the 7 feeling classes: angry, disgust, fear, happy, sad, surprise, and neutral. The errand is to sort each face dependent on the feeling appeared in the outward appearance in to one of seven classifications (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral).

0: - 4593 pictures Angry
1: - 547 pictures Disgust
2: - 5121 pictures Fear
3: - 8989 pictures Happy
4: - 6077 pictures Sad
5: - 4002 pictures Surprise
6: - 6198 pictures Neutral

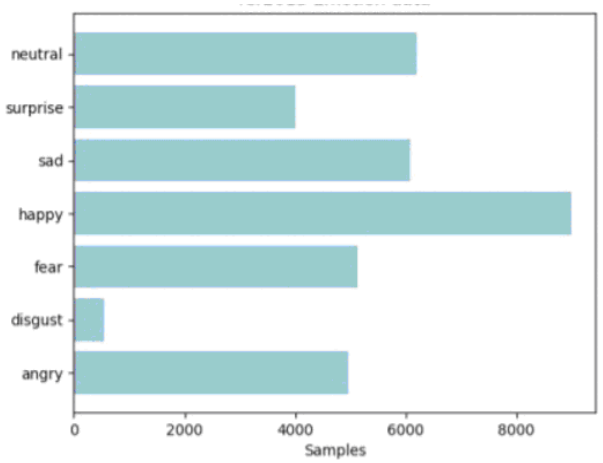

Fig. 3. Distribution of pictures.

Fig .4. Some pictures from dataset.

3. CNN Architecture
The block diagram of the convolutional neural network is shown in figure 5. Here the picture is taken as an information then different various layers are included for extraction of highlight lastly a completely associated layer is available to order the 7 feelings for example angry, disgust, fear, happy, neutral, sad and surprise.

International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 3, May-June-2020, ISSN (Online): 2395-566X

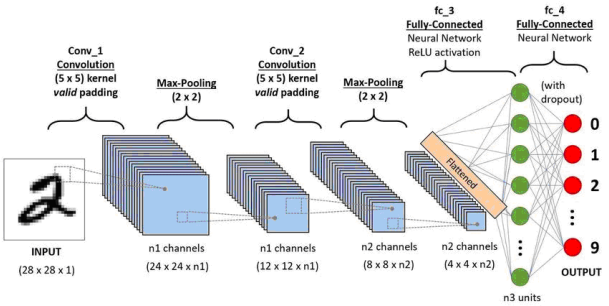

Fig .5. Block diagram of CNN

.

From Table 1, we can picture the CNN design which will be the utilizing to arrange the feelings. In the principal stage, we applied a convolution of a 5 × 5 channel on the picture. Next, another convolutional system will be applied to the 5 × 5 channel size, and in the subsequent stage, utilizing the Max Pooling capacity, we will lessen the size to 24 × 24 × 64. Once more, a convolution of a 5 × 5 channel size with a back to back channel of 5 × 5 size is applied and in the following stage utilizing the Max Pooling capacity, we will decrease the size to 12 × 12 × 128. After this again two successive convolution layers of channel size 3 × 3 is applied and utilizing Max Pooling capacity the size is diminished to 6 × 6 × 256, utilizing the Flatten work, all information is changed over to a vector of the size 9216. At that point, it is changed over to a vector of length 128, and toward the end, it is decreased to seven, which is the 7 classes of passionate states.

Rectified Linear Unit (relu) function:

Y=max(0,x)
(6)

Softmax Function:

$\sum_i X_i = 1 \qquad x_i = [0,1]$ (7)

Table I: CNN Architecture

Layer type	Details	Output Shape
Conv	Conv(5x5)	48,48,64
Conv	Conv(5x5)	48,48,64
Batch_Normalization		48,48,64
MaxPooling	Pool size (2,2)	24,24,64
Conv	Conv(5x5)	24,24,128
Conv	Conv(5x5)	24,24,128
Batch_Normalization	----------	24,24,128
MaxPooling	----------	12,12,128
Conv	Conv(3x3)	12,12,256
Conv	Conv(3x3)	12,12,256
Batch_Normalization	----------	12,12,256
MaxPooling	Pool size (2,2)	6,6,256
Flatten	Flatten to a vector	9216
Dense	----------	128
Batch_Normalization	----------	128
Activation	relu	128
Dropout	0.2	128
Dense	Output classes	7
Activation	Softmax	7

III. TRIAL RESULTS

For the outcomes, the fer2013 database is utilized. The database contains 35887 facial feeling pictures that incorporate seven passionate conditions of the face. Table 2 shows the particular of the reenactment equipment. From Table 3, we can make an examination from the current philosophy and the proposed one. The current technique begins from 10% and rises upto 90% after 25 epochs though our philosophy gets an increasing speed from the beginning point for example it begins from 36% and ascends till 94% after 25 epochs. In this way, it very well may be denoted a superior outcome.

Table II: Hardware specification

Processor	Intel® Core™ i5-7200U CPU @2.50Ghz
RAM	8.00 GB (7.89 GB usable)
System type	64-bit Operating System, x-64 based processor
Processor Graphics	Intel® HD Graphics 620
GPU	NVIDIA GeForce 940MX

Table III: Comparison accuracy of proposed method and existing method

Epoch	CNN method accuracy	Gabor + CNN method
1	0.1050	0.3669
10	0.5138	0.8607
15	0.7348	0.9208
20	0.8343	0.9396
25	0.9006	0.9469

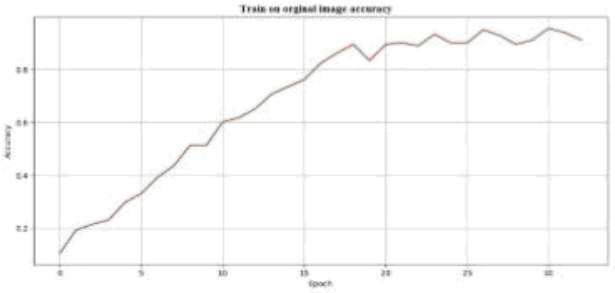

Fig .6. Accuracy of Simple CNN method.

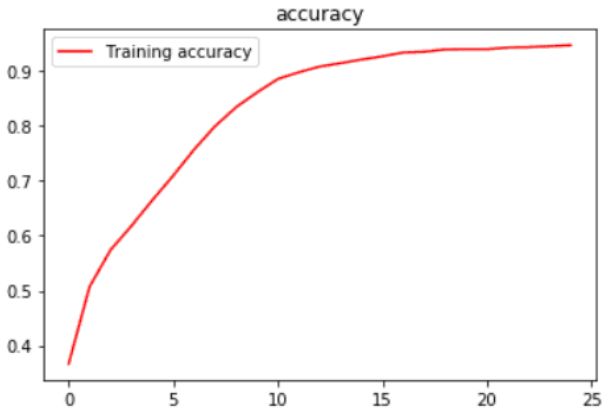

Fig. 7. Accuracy of proposed method.

IV. CONCLUSION

The outcomes unmistakably show improvement in both speed and precision. Because of the utilization of gabor channels on the pictures in the pre-preparing stage the extraction of highlight turned out to be simple for the system as it utilizes edges for making sub-highlights. Edge identification works extraordinary in the situation where there is different appropriated edges and gabor channel is one of them. By utilizing gabor channel with convolutional neural system the order was improved with speed.

REFERENCES

[1]. Vaibhav Kumar J. Mistry, Mahesh M. Goyani "A writing review on Facial Expression Recognition utilizing Global Features " International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-4, April 2013
[2]. Samiksha Agrawal, Pallavi Khatri "Outward appearance Detection Techniques: Based on Viola and Jones calculation and Principal Component Analysis " 2015 Fifth International Conference on Advanced Computing and Communication Technologies
[3]. Kanchan Lata Kashyap, Sanjivani Shantaiya "Commotion Removal of Facial Expression Images utilizing Wiener Filter" National Conference on Emerging Trends in Computer Science and Information Technology (ETCSIT) 2011 Proceedings distributed in International Journal of Computer Applications® (IJCA)
[4]. Reza Hassanpour, Asadollah Shahbahrami, And Stephan Wong, "Versatile Gaussian Mixture Model for Skin Color Segmentation", procedures of World Academy of Science, Engineering and Technology Volume 31 July 2008 issn 1307-6884.
[5]. Alexandru Savoiu and James Wong "Perceiving Facial Expressions Using Deep Learning"
[6]. Ligang Zhang and Dian Tjondronegoro "Outward appearance Recognition Using Facial Movement Features" IEEE Transactions on Affective ComputingOctober 2011
[7]. Keith anderson and peter w. Mcowan "a real-time automated system for the recognition of human facial expressions" ieee transactions on systems, man, and cybernetics—part b: cybernetics, vol. 36, no. 1, february 2006.
[8]. Ali Mollahosseini, David Chan, and Mohammad H. Mahoor1 "Going Deeper in Facial Expression Recognition utilizing Deep Neural Networks"