IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205037 |

Real Time Analysis of Facial Emotion Using Convolutional Neural Network

Dr. A. Prakash, Aryan Baranwal, Suraj Kumar Sah, Sudhansu Jha, Sirish Chaudhary

Professor, Dept. of ISE, Faculty of Engineering and Technology, Jain (Deemed-to-be-University), Bangalore, India

UG Student [4], Dept. of ISE, Faculty of Engineering and Technology, Jain (Deemed-to-be-University), Bangalore, India

ABSTRACT: In this paper the major focus is on Facial Emotion Recognition and how it is resembles our Emotions. Emotions are viable part of a human as every thought or action of a human can be defined as his/her emotion which they depict and by analyzing it we can get to know a lot about one individual. The facial emotion recognition gives in detail account of a person's behavior and their thoughts. Our research majorly focuses on recognizing a person's emotion with respect to the analysis of their face and also in this paper we have conducted various phases of preprocessing and dealt with the change in accuracy with respect to it. In our model the major feature is that we are using the combination of extraction of features through the Local Binary Pattern (LBP) along with the help of region based Oriented Fast and rotated Brief (ORB) also we have used this in terms of Convolutional Neural Network (CNN). Our mechanism is working on the model of classification that is purely based on segmented layers of CNN which we have used to classify emotions in the terms of mental and emotional to study in depth. Our model majorly performs its routine action totally based on its segmented layers that is four layers of CNN along with two feature classifiers.

KEYWORDS: FER, Neural Network, Naïve Bayes, DL, ML, NLP, Feature Extraction, Facial Expression.

I. INTRODUCTION

The facial expressions of an individual and the emotions conveyed via them play a significant role in non-verbal communication. Facial emotion plays a vital role in the realm of oral communication, significantly enhancing its effectiveness and potency. It is also the primary determinant in matters of etiquette, mindset, psychology, wrongdoing, and so forth. The model lacks attention towards factors that are not of primary importance, such as gender, origin, ethnicity, racial complexity, and facial complexity. The methodology employed in this study is sufficiently robust to detect and analyze emotions, irrespective of the aforementioned component. The first step in Facial Emotion Recognition involves the detection of the face, which can be in the form of a collection of photos, a video, or a single image. The image may either include a prominent facial element or it may encompass a broader visual composition inclusive of a background. One crucial aspect to be taken into account in face identification is the process of isolating facial images from their backdrop, thereby extracting the face from the overall image collection. The identification of facial features holds significant importance in both research and security domains. It can be utilized for the recognition of certain facial characteristics, as well as in the automation of camera systems. The primary field of study that can be utilized encompasses the realm of recognition and indexing.

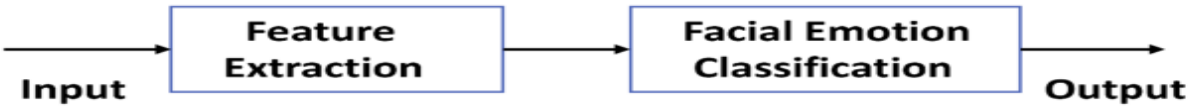

Figure 1:-. Simple Block Based View of Facial Emotion Recognition

In this research the major thing that we have done is model creation for Facial Emotion Recognition with the help of CNN. The approach that we are going to use is capable enough of performing tasks at real time which i.e., means it is dynamic and can get its input from the webcam and then work over it to generate the expected result.

The major considerations of the proposed method are:

i. CNN method for recognizing facial features with extraction carried out with LBP, ORB and CNN.
ii. The proposed method has four-layer 'Conv-Net' for better recognition.
iii. Indulging different datasets to increase the detection accuracy.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205037 |

II. LITERATURE SURVEY

Several studies in the field of facial emotion recognition have been carried out since many years. After pandemic of Corona the increase in research of Facial Emotion Recognition was seen. As we know for any of the experiment there is some loophole and there is a room for research. The objective of this research is to create a new system which gives better accuracy for the pre-existing datasets and analyzes them in efficient manner. In existing methodologies some lack in identifying gray scale images while some has feature lack in pixel-based evaluation. In [1], Compound approach of facial emotion detection consisting of two- stage of recognition, the author of the paper has mentioned about how one stage and two stage model can enhance the data recognition in the meantime. Similar to the context in [2], a combined approach of CNN and LSTM has been used to improvise the accuracy and extract facial expression. [3]. In the paper as mentioned the feature vector extraction is useful in extraction through which the error is reduced and the mechanism supports the range of data to be improvised.

As of experiment with some advantages there exists some deficiency. In the field of facial emotion recognition, the major hindrance that we came across and that drawn our attention includes the accuracy of certain model, their learning capability and approach of machine learning, there in detail exploration of machine learning, the factor of lacking in recognizing actions or behavior's. Also, majority of the model do lack the performance over different datasets. After referring to the articles and journals the final result is that there is need of impact identification technology which makes the perfect use of the dataset in providing accuracy in results.

Paper Name	Author	Objective	Conclusion
Multi-Modal Emotion Recognition from Speech and Facial Expression Based on Deep Learning	Linqin Cai; Jiangong Dong; Min Wei	Multi-scale Kernel Convolutional Block to increase accuracy based on images and speech.	It used CNN along with LSTM to provide accuracy and detect features.
Feature Vector Extraction Technique for Facial Emotion Recognition Using Facial Landmarks	Alwin Poulose, Jung Hwan Kim and Dong Seog Han	Reducing classification error and to achieve 99.96% accuracy in classification.	A vector extraction technique that combines the facial image pixel values with facial landmarks.
Two-stage Recognition and Beyond for Compound Facial Emotion Recognition	Zhiyuan Zhang; Miao Yi; Juan Xu; Rong Zhang; JianPing Shen	Symmetrical emotion labels enhancing and robust results.	Compound emotion recognition technique with elimination of noisy data.
Facial emotion recognition based on LDA and Facial Landmark Detection	Lanxin Sun; JunBo Dai; Xunbing Shen	Use of dimension reduction technique of supervised learning and minimizing intra class variance.	The comparative analysis of accuracy of emotion recognition in which facial landmark gave best possible result.
Emotion Detection with Facial Feature Recognition Using CNN & OpenCV	Sarwesh Giri; Gurchetan Singh; Babul Kumar; Mehakpreet Singh; Deepanker Vashisht	Human- computer domain interaction (HCI) based research and demonstration of the emotion detection.	This approach was constructed using comprehensive research based on CNN algorithm with conjunction of Keras, TensorFlow and retraining principles.

Table 1: Analysis of Previous works.

III. PROPOSED METHODOLOGY AND DISCUSSION

When we are talking about machine learning the thing we should start with is deep neural network and the major key player in the field that is Convolutional neural network which is one of the form of deep neural network that has it usage in processing of vision based images and also in machine vision i.e., computer vision. With a lot of features some

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205037 |

deficiency also exists the major deficiency that we found during our research the contradiction in cases of gleaming contradiction along with location-based contradiction which exists mostly in the CNN. In contrast to the methods that we are having in the meantime our method stands out with a structure that is segmented and is arranged in sequential fashion. Our approach performs the task in segments and they all are in sequenced manner i.e., beginning with modeling of the data after that analysis there after prediction based on the data and at last producing a final output which is performed and have been through these all segments.

The major distinctions of our proposed methodology are:
i. Use of LBP in extracting the features.
ii. Differences in the field of Mathematical Modeling.
iii. In depth architecture with layered approach.

A. LBP (Local binary Pattern)

Local Binary pattern (LBP) is the word that is used in defining the properties of picture on local range or level which means it defines the properties of picture which include appearance features of an image. The advantage that is associated with LBP is the rotating factor and deviation in case of gray scaling. The LBP is one of the major classifying factors that we are considering in the deployment of this machine learning model. As it is simple to use and can serve for long term it is one of the most common techniques that is been used in detection of objects and expressions of faces for extracting required features. The initialization for the LBP begins with the pixel value which is created as a 8 bit binary number following 3*3 neighboring range of the pixel and it does for each pixel. Later on after getting all into segments by binary operator it is turned to the corresponding decimal where the LBP gets result from binary based image patches. The equation that we have considered for calculation for the LBP in our model is:

(x_c, y_c):

LBPP, R $(X_c, Y_c) = \sum_{n=0}^{7} 2^n \, S(in - ic) \, , S(x) = \{^{1, x \geq 0}_{0, x < 0}$

where, i_c=gray value of center pixel,
i_n=gray value of neighboring pixel of i_c,
P=8 maximum of 8 neighbors of enter pixel.
R = 1 for selected box of 3 × 3

The value of Local Binary Pattern (LBP) is being determined in the given scenario by tracking the available bins in a clockwise manner within the radial pattern. The subsequent methodology is founded on the categorization of the input derived from the code generated by LBP. This approach will thereafter be utilized as the input for the subsequent phase. The Local Binary Pattern (LBP) algorithm has robustness in its feature extraction capabilities, demonstrating effectiveness even in the presence of light disturbances or variations in pixel values. After careful consideration of the aforementioned advantages, it can be determined that this particular solution is very suitable for real-time detection tasks. In the below figure here we have used (P, R) that is to represent the neighbors of P which are the sample points that are considered in the radical which has an Radius R and in this we can clearly observe that there is even spacing. In this there is symmetric arrangement inside the circle that we can also observe. After considering LBP we have considered the LBP following uniform pattern that means that either it can be 0 – 1 or exist between 1-0 which makes the whole pattern uniform in case of a string that constitutes of circular bits. Considering example: 000000, 001100, 111000 10101100 we can see these all are patterns which are uniform.

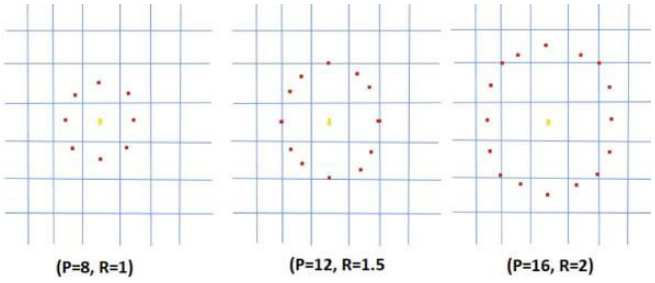

Figure 2: Three samples of extended version of LBP

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205037 |

The operations of LBP generate a histogram representation of picture which can be labeled as C_l (x, y) and can be explained as:

$$Si = \sum x, y_I (C_l (x, y) = i), i=0,1,\ldots..,n-1$$

Where, the different labels number is denoted by n.

$$I (A) = \{^{1,A\ is\ true}_{0,A\ is\ false}$$

B. Region Based Orb

The traditional ORB uses too much feature points in extractions to resolve the challenge what we are going to use is Region Based ORB. We update the ORB algorithm with base of dividing the region-based feature points in reference to the total point to be considered for feature extractions. The measure that we are going to take is as follow:

i. Starting with the division of the image into a square matrix that is X*Y of equal size. In this X and Y denotes the row and column respectively. The Scattered Points can be denoted as (S_1, S_2……., S_{X*Y})

ii. The threshold is set as T_h

$$T_h = \frac{n}{XY}$$

Where, the total number of points considered is denoted by n.

i. Recognition of features is done on every possible region and the considerations here are that in case we have a greater number of features than our threshold than we consider threshold as the number of feature else we have threshold lesser than we again degrade the threshold and call the process again and repeat it.

ii. The non-maximal suppression approach is used to choose the best feature points, if we are having lesser value for n than the points considered as feature point.

iii. The number of points that are considered as the feature points is necessary to be lesser so that each of the region can be covered and feature can be classified out of them.

C. Analysis of CNN Model

In the approach that we are proposing consists of 4 different layers of convolution which has additional two layers of classifiers which are additional. In order to create a convoluted feature map, we have fused the features from vector along with the features from LBP. The layers of pools are using a static method for the conversion of activation in the phase when weighting of convolutional layer is there for the training purpose. We have used the rectified layer unit (ReLU) so that there is not any type of linearity in the whole network and which won't affect the operational field of our respective convolutional layer. The efficiency of convolutional layers is determined as the loops. For efficiency add on we use back training in our model that is our output calculation is sent back to the training dataset which has no loss during the phase of back propagation. During the max pooling or convolution layer there are some chances of some layers to be lost in the scenario. The generated result is in spatial dimension which is nonlinear in nature and went through nonlinear down process. While during the pooling there is reduction in spatial size in which there is reduction of factor and computation that provides the required control over the fitting. At the last we again remap the two connected layers into a Pool feature map which again forms multi-dimensional structure after that it is converted into single dimension which is a vector and this vector is going to act as the feature vector in future operations. To continue the process of classification what we do is that we start treating the vector obtained as fully connected layer.

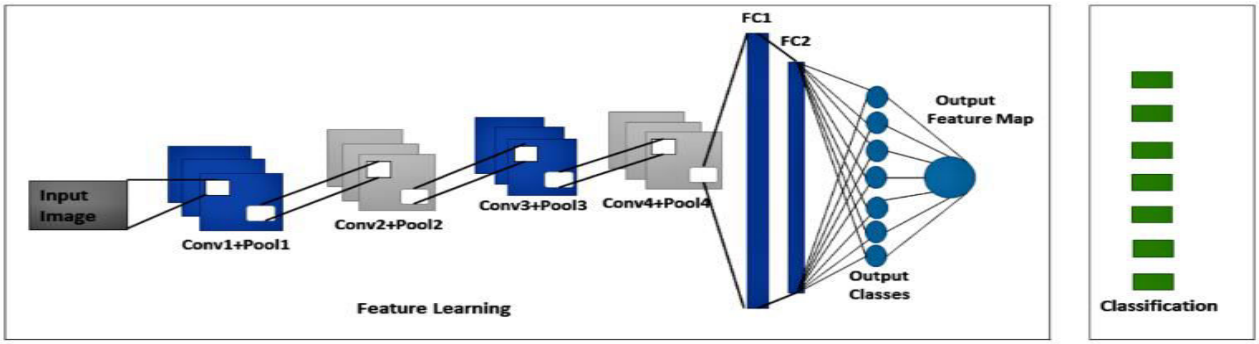

Figure 2: Graphical representation of proposed method for FER.

i. Rectification of Proposed CNN: In this phase we are not only considering rectification of our proposed method in spite we are also considering the rectification approach which automatically replaces and instead of the dataset used in

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205037 |

the pre training approach it gets real time data and replace the previously existing layers with new layers that is generated after use of the model. We are doing this on the level of Kernel of pre training model which uses the technique of back propagation to send base layer back to the set of connected layers. The rectification parameters that are taken into consideration involves the resultant size of the convolutional layer are padding, stride, batch size, filter, sliding windows and the learning rate parameter. Padding in the case adds zeros to the bounds of the input. Stride controls over height parameters. In our approach our layers only have the task of training the model in the detailed feature classification which is based on block essence and in the end, we get the 7 emotions classification as a result of our model from our connected layers.

ii. Pipeline for proposed CNN.: Our pipeline architecture for the CNN has convolution along with pooling layers which are add on and in numbers we can have four layer of convolution and additional layers of two which are fully connected. The layer that we are considering as part of our sequenced structure contains the ReLU layer along with normalization that is batch based and a layer of dropout which completes our structure of fully connected layer.

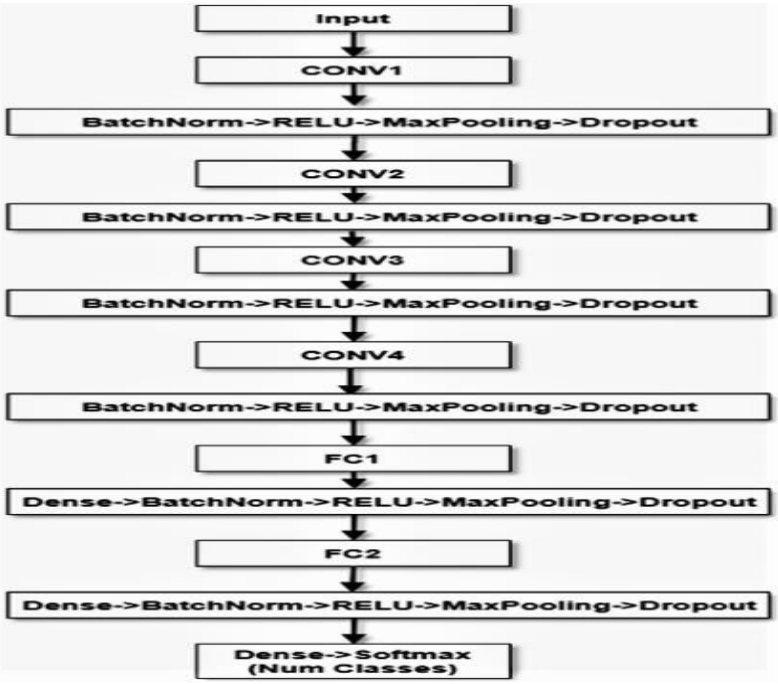

Figure 3: Architecture of Proposed CNN.

IV. RESULTS

As we have used composite dataset for the feature performance, we obtained an accuracy above of 95 % and valid accuracy of 91%, which would be enhanced after several experiment approaches. To measure the performance of the system the results are measured using three metrics: accuracy rate, specificity rate and sensitivity rate. Truly-positive (TP), Truly-negative (TN), Falsely-positive (FP) and Falsely-negative (FN) are used to compute the characteristics.

$$\text{Accuracy} = \frac{TN + TN}{TP + FP + FN + TN},$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}‘$$

$$\text{Specificity} = \frac{TN}{TN + FP},$$

In our proposed model the total run time is of;
$O(n^4 + n^2) \leftrightarrow O(n^4) \quad \because \forall n \geq 1 \mid n^4 + n^2 \leq 2n^4$

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205037 |

V. CONCLUSION

We have used the indulging factors from LBP and ORB for betterment of CNN model. The dataset used were CK+, JAFFE and FER2013 all as gray scale images. In this we have found seven emotions and classified according to it.
In the analysis we have found it on average and has some look around to be improved in future with a greater number of experiments and datasets. The implementation majorly focused on the classifiers and the 4 layers which were working better with FER2013 dataset and provided quite adequate results in the whole.
Furthermore we are aiming to focus on mental issues of man or woman also with the emotion recognition mechanism and for that also we are training our model and indulging it to the practice.

REFERENCES

1. Zhang, M. Yi, J. Xu, R. Zhang and J. Shen, "Two-stage Recognition and Beyond for Compound Facial Emotion Recognition," 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), 2020, pp. 900-904, doi: 10.1109/FG47880.2020.00144.
2. L. Cai, J. Dong and M. Wei, "Multi-Modal Emotion Recognition From Speech and Facial Expression Based on Deep Learning," 2020 Chinese Automation Congress (CAC), 2020, pp. 5726-5729, doi: 10.1109/CAC51589.2020.9327178.
3. A. Poulose, J. H. Kim and D. S. Han, "Feature Vector Extraction Technique for Facial Emotion Recognition Using Facial Landmarks," 2021 International Conference on Information and Communication Technology Convergence (ICTC), 2021, pp. 1072-1076, doi: 10.1109/ICTC52510.2021.9620798.
4. C. Dalvi, M. Rathod, S. Patil, S. Gite and K. Kotecha, "A Survey of AI-Based Facial Emotion Recognition: Features, ML & DL Techniques, Age-Wise Datasets and Future Directions," in IEEE Access, vol. 9, pp. 165806-165840, 2021, doi: 10.1109/ACCESS.2021.3131733.
5. Ebenezer Owusu, Jacqueline Asor Kumi, Justice Kwame Appati, "On Facial Expression Recognition Benchmarks", Applied Computational Intelligence and Soft Computing, vol. 2021, Article ID 9917246, 20 pages, 2021. https://doi.org/10.1155/2021/9917246
6. Kumar Arora T, Kumar Chaubey P, Shree Raman M, Kumar B, Nagesh Y, Anjani PK, Ahmed HMS, Hashmi A, Balamuralitharan S, Debtera B. Optimal Facial Feature Based Emotional Recognition Using Deep Learning Algorithm. Comput Intell Neurosci. 2022 Sep 20;2022:8379202. doi: 10.1155/2022/8379202. PMID: 36177319; PMCID: PMC9514924.
7. Küntzler T, Höfling TTA, Alpers GW. Automatic Facial Expression Recognition in Standardized and Non-standardized Emotional Expressions. Front Psychol. 2021 May 5;12:627561. doi: 10.3389/fpsyg.2021.627561. PMID: 34025503; PMCID: PMC8131548.
8. Khan, Amjad Rehman. (2022). Facial Emotion Recognition Using Conventional Machine Learning and Deep Learning Methods: Current Achievements, Analysis and Remaining Challenges. Information. 13. 268. 10.3390/info13060268.
9. S. Giri et al., "Emotion Detection with Facial Feature Recognition Using CNN & OpenCV," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2022, pp. 230-232, doi: 10.1109/ICACITE53722.2022.9823786.
10. L. Sun, J. Dai and X. Shen, "Facial emotion recognition based on LDA and Facial Landmark Detection," 2021 2nd International Conference on Artificial Intelligence and Education (ICAIE), 2021, pp. 64-67, doi: 10.1109/ICAIE53562.2021.00020.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details