Journal of IoT in Social, Mobile, Analytics, and Cloud (ISSN: 2582-1369)
www.irojournals.com/iroismac/

Automatic Detection and Classification of Human Emotion in Real-Time Scenario

Ashish Keshri[1], Ayush Singh[2], Baibhav Kumar[3], Devenrdra Pratap[4], Ankit Chauhan[5]

[1,2,3,5]ECE Departmen,NIET, G Noida, India
[4]Assistant Professor, ECE Departmen, NIET, G Noida, India

E-mail: [1]hkeshri00@gmail.com, [2]ayushsingh38269@gmail.com, [3]vaibhavdav99@gmail.com, [4]devpratap7188@gmail.com, [5]ankitchauhan95571684@gmail.com

Abstract

This work proposes the implementation of the idea of real-time human emotion recognition through digital image processing techniques using CNN. This work presents significant literacy calculations used in facial protestation for exact distinctive verification and acknowledgment that can effectively and capably see sentiments from the vibes of the client. The proposed model gives six probability values based on six different expressions. Large datasets are explored and investigated for training facial emotion recognition model. In support of this work, CNN using Deep learning model, OpenCV, Tensorflow, Keras, Pandas, and Numpy is used for digital computer vision procedures involved, and an lite experiment is conducted for various men and women of different age, race, and colour to descry their feelings and variations for different faces are found. This work is improved in 3 targets as face location, acknowledgment and feeling arrangement. Open CV library, and facial expression images dataset are used in this proposed work. Also python writing computer programs is utilized for computer vision (using webcam) procedures. To demonstrate ongoing adequacy, an investigation is directed for a very long time to distinguish their internal feelings and track down physiological changes for each face. The consequences of the examinations exhibit the idealizations in face investigation framework. At long last, the exhibition of programmed face detection and recognition are measured with very high accuracy and in real-time. This method can be implemented and is widely useful in various domains such as security, schools, colleges and universities, military, airlines, banking etc.

Keywords: CNN, Deep learning model, Face emotion recognition

DOI: https://doi.org/10.36548/jismac.2022.1.005
Received: 23.03.2022, received in revised form: 14.04.2022, accepted: 01.05.2022, published: 14.05.2022

1. Introduction

In general, there are mainly two powerful tools for an effective communication that is human feelings and facial expression. Only 8 percent of valuable information can be carried by a massage signal by linguistic component of any message ,which is influencing the original message. Hence, Facial expression is considered main tool among those accessible which is exceptionally expressive and assumes a pivotal part to get the sensation of every individual's figures plainly regarding an individual's character. The fundamental target of this undertaking is to foster an Automatic Facial Expression Recognition System which takes human facial pictures containing some appearance as info and perceive and group it into seven distinct demeanour class like Neutral, Angry, Disgust, Fear, Happy, Sadness, Surprise.

Figure 1. Facial Expression of Men and Women

In the current time an enormous exploration work is going in the field of computerized picture and picture handling, with the improvement of equipment and the development of big data, deep learning technology which learns the suitable pattern for the purpose itself has attracted the attention. Deep learning can overcome the technical limitations of existing machine learning where the performance is drastically deteriorated for complex problems by using deep neural networks to extract high level features appropriate to the given data. Convolutional Neural Networks (CNN) are developed to imitate human visual cognition processes in deep learning technology and has been widely applied to the field of image recognition and shows high performance [1]. Communication between a human and a computer is a typical pattern and intrinsic capacity to recognize among various appearances [2]. Until ongoing several computer vision issues were very difficult yet coming of current innovations has inconsequentially modified by the problem of changing light, by age, hair and various additions. The ongoing development in the inventions has brought about the application of artificial frames as they are suitable for understanding and understanding systems through the use of facial expressions to identify facial expressions. Later this is an

attempt to present a highly sophisticated machine development. computer communication utilizing profound learning or Convolution neural networks (CNN) model [3]. The deep learning models have awesome attributes of figure out the inside and out math and has huge interest as moment chip called Graphical handling unit (GHU) [4].

1.1 Convolution-Neural-Networks (CNN)

The advancement of CNN fundamentally affects the computer vision space as of late notwithstanding this it has a significant stage and capacity to perceive object [5]. This acquiring significance of the method was conceivable because of expanding worth of the computer also how much information accessible to frame the computer neural networks.

CNN models have matrix of array to distinguish line capacities or blueprint for a picture. This model has loads organized in cluster of qualities to frame and get wanted attributes [6]. Each model apportions a proper place to decide the control of picture to be received. The rates in the picture address a threshold value to which complication exertion varies, accordingly the item isn't set in gravestone with the area in the picture [7]. Now the question is how CNN is used in this project? It starts with the introduction of the model followed by cluster standardization layer and afterward various religious circles layers with ReLu as an enactment work, max pool layers, and dropouts to do advancing proficiently. You can likewise change the design by starting the layers of your decisions with various quantities of neurons and actuation capacities. After this, the model involving Adam as an enhancer, misfortune as absolute cross-entropy, and measurements as exactness as displayed in the underneath code are assembled. In the wake of ordering the model, at that point, the information is fit for preparing and approval. Here, the cluster size is considered to be 32 with 30 ages.

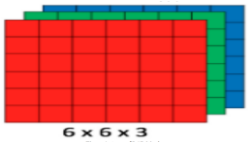

Figure 2. Red-Green Blue Matrix.

This CN network models are meant to prepare & test the planned framework. Here the pictures are contribution with a few layers as channels associated along with loads as wanted. It has Softmax work used to get and order the picture somewhere in the range of 0

&1. Henceforth the beneath fig. 2 shows the portrayal of CNN models handling in view of the information picture[8].

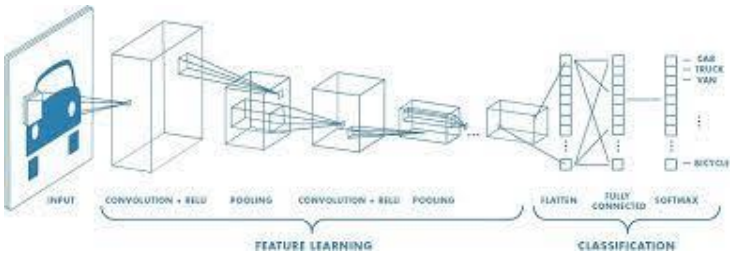

Figure 3. Neural networks having various convolutional neural layers

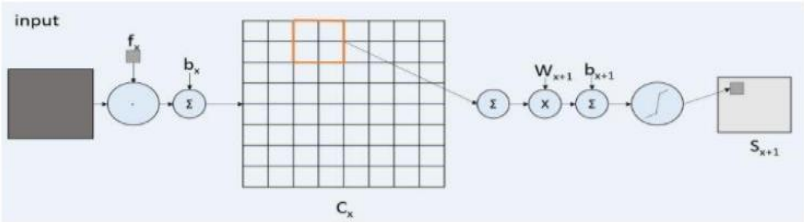

Figure 4. Processes of CN Network

The figure 4 shows a fundamental 1st layer of CNN where elements are extricated from I/P picture. In the wake of extricating the highlights they are passed to complication subcaste where loads capacities are duplicated to acquire the result of each subcaste. The numerical handling happens between picture grid, 2 data sources and portion esteems[9]. The got picture grid addressed with aspects and volume is given as (h x w x d) in figure 1and the channel as (fh x fw x d) and results of volume aspects are given as the function of (h-fh+1), x (w-fw+1), x(1). Further these subtleties of CN layers loads, and work esteems are portrayed in upcoming sections.

Primary advancement in CN networks are the local field of acumen and weight apportionment, sampling of time/space to segregate works & lessen the size of planning limits. The advantages of CN networks computation is that it dodges a sensible segment of limits and atypical learning and planning of the parametrized data. The proper use of these subsample structure on time/space can give adequacy, size & speed. Network segments and geology can be superb by and large and offer the extraordinary benefits of voice affirmation and picture dealing with. CN networks computation requires knowledge in project planning and ought to consistently see application practice bungles to find the best solution for a specific CNN application designing.[3].

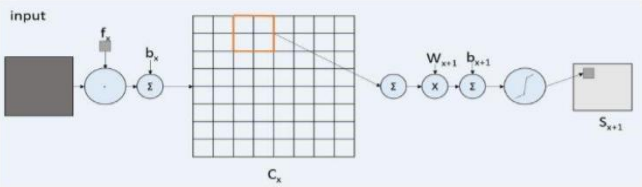

Figure 5. CN network-model stage.

2. Approach

The proposed sort out conveys in three consecutive strides as Face Detection. In the original step a camcorder is employed to catch the mortal face and distinguish the specific area of face by a jumping box facilitates for the face honored continuously. This procedure includes face identification exercising Haar-cascade discovery with (open CV) library. Two techniques such as Viola jones computation and Haar-Cascade are used to identify mortal face. The CNN model as displayed in figure 4 uses VGG 16 to coordinate the face from the data set and perceive with the name for the face honored. Faces are perceived from the information base and are varied with fete or distinguish the face through installing vectors.[10] Finally, the apparent human face is portrayed considering the appearances continuously as scorn , angry, fear, happy, fair, shock.

The organization design VGG-16 is worked with CNN model for colossal informational collection affirmation and plan. The arranged association model has honeycomb 3 x 3 layers where the two related layers having 4096 center points with SoftMax gathering. The area combined model histogram is used as open CV library for perceiving the human appearances[11]. The image pixels are perceived by putting down a boundary and the result is tended to in sort of a twofold number. In solicitation to play out this LBPH uses 4 limits like radio, neighbors, Grid X and Y. here the proportion button used describes the local range in round manner for the twofold model. The indirect area equal model tends to the amount of test centers as neighbors[12][13]. The X and Y parts of cross section portrays the amount of level and vertical course cells. The round matched model is outlined to the unquestionable locales with appropriate imprint to get the test data from machine. To play out the procedural delayed consequence of face area at first make a coordinator and name it and form two python records create_data.py and face_recognize.py then haarcascade_frontalface_default.xml library set and run absolute reports from Integrated advancement climate of python [8].

2.1 Suggested Practical Design

The Suggested Practical design includes its system and working is depicted in the figure given below. All individual utilitarian squares have their own particular activity The picture is caught utilizing camera which is preprocessed further to take care of the brain organizations. With the assistance of organization design to construct and prepare the dataset for ID and feeling order the caught picture is demonstrated [14][15]. The subtleties are expounded in consecutive strides as stream graph in figure.

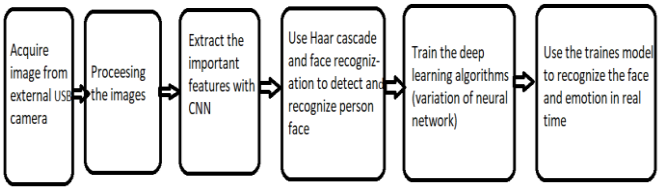

Figure 6. System flow of face spotting and recognition Classification

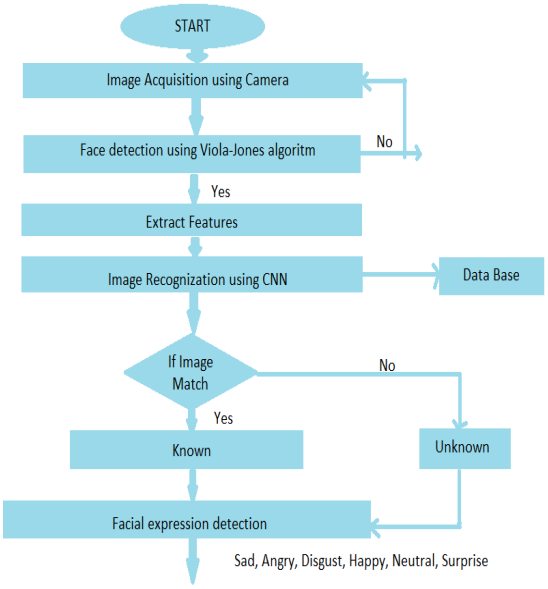

Figure 7. Face sensing and identification flow chart

2.2 Maintenance of the Integrity and of the Specification

The actual time face identification, acknowledgment and grouping is explained in the mentioned figure. The caught picture is first limited in a rectangular box and afterward changed over into double example to indicate it as component vector and put away in information base. The changed over pictures are prepared to match the information picture

and furthermore order the demeanor of facial elements as miserable, irate, cheerful, loathing and nonpartisan. In the preparation interaction, there are an aggregate of 7-stages which are addressed underneath as stacking dataset-preprocessing and increasing the information as element vectors and then building and gathering the plan model, then training the models and again storing the feature vector and at last validating the test models[16].

3. Implementation and Code

The functional design is tested in working topology on several different levels.

3.1 Creating a Sorted Dataset to Sort Human Emotion Using CN Network Layer

The dataset used in proposed work is tended to in the association. The KDEF educational record as in the mentioned figure had 4900 pictures where 1999 pictures were taken out as they were inside positions and arranged the dataset with 2901 pictures by separating and setting up the dataset in to 75 % and 25% individually [17][18].

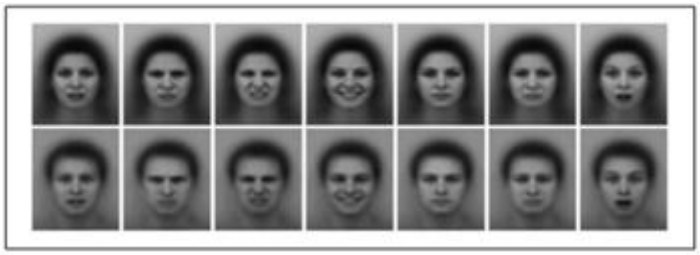

Figure 8. Sample dataset (KDEF) from the Kenya Dryland Education Fund (2019)

3.2 Coding done in Python Language using Anaconda.

```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns
import os

# Importing Deep Learning Libraries

from keras.preprocessing.image import load_img, img_to_array
from keras.preprocessing.image import ImageDataGenerator
from keras.layers import Dense,Input,Dropout,GlobalAveragePooling2D,Flatten,Conv2D,BatchNormalization,Activation,MaxPooling2D
from keras.models import Model,Sequential
from keras.optimizers import Adam,SGD,RMSprop
```

Figure 9. Code Implementation 1

Importing all the libraries necessary in this program. The Python 3.9.7 and Anaconda 3-11.0 has been used to obtaining the results.

- This block of code is written for displaying of the images.

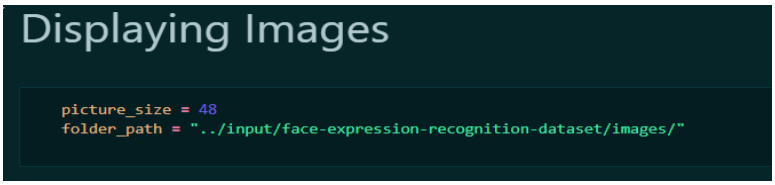

Figure 10. Code Implementation 2

- Making training and validating the data.

Figure 11. Code Implementation 3

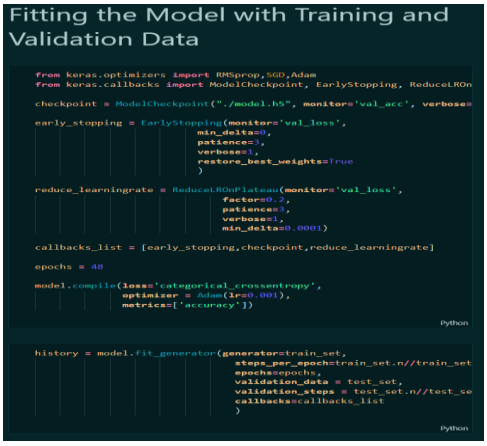

Figure 12. Code Implementation 4

Figure 13. Code Implementation 5

This block is of the main.py that will finally use the computer webcam recognize the facial expression in real time.

4. Experimental Results

For face detection a quick and powerful face locator carried out in OpenCV has been used. A well-designed face detector that is implemented in OpenCV has been used in the detection of human face.A real time video captured by the webcam of the computer, the human face detector now finds the accurate location of the image and differentiate the face or the region that has no face. An XML file has been designed to classify the face and non-face region. Haar Cascade classifier file of OpenCV is selected for the XML data detection. An array of frames is convoluted together from the webcam and human faces are detected. Now, there must be at least one human face should be there in the frame so that the system would identify the facial expression of that face.

Figure 14. Result 1 - Neutral

Figure 15. Result 2 – Happy

Figure 16. Result 3 - Surprise

5. Limitations and Future Works

The current methodology processes just the apex frame of the arrangements, nearly to most different methodologies where a video transfer is handled. Utilizing the peak outline emphasizes the differentiation between feelings, staying away from outlines where the articulations are more unpretentious, however such an impact may be adjusted by restricting the framework to a solitary casing rather than the movement data that could be extricated from the successions[19]. It would be feasible to utilize such movement data in a later form of the framework in more than one way. One potential technique would be like the one utilized for the constant application: conglomerating the forecasts from each edge, and afterward doling out the most usually anticipated name to the grouping. Another technique is stacking the includes consecutively and let the classifier handle very high dimensionality important elements.

The CK+ dataset doesn't appear to offer data on regardless of whether a portion of the members were wearing glasses. While testing the continuous application with glasses, when a model prepared without face location was utilized, it was seen that "disgust" was frequently identified rather than different feelings until the glasses were eliminated. This isn't the business as usual as facial impediments are for the most part difficult to deal with [10]. On the other hand, when tried with the best model found (coordinating face recognition), the glasses were not an issue any longer. Given this perception, it would be intriguing to test the model on members that current facial impediments, for example, headwear, glasses, or beard growth. Emotion detection of pilots and soldiers during training and war can be very useful for territory sovereignty. Detection of emotion can be help in Suggesting movies, songs and other content according to the present emotional condition of the consumer. This model is

found to be a new business tool for the Enhancement in the technique in gathering the audience reaction[20].

6. Conclusion

This proposed work intends to identify, perceive, and characterize human face progressively. In this continuous framework, the characterized articulations are addressed in seven states as displayed in above outcomes. Anaconda and python 3.5 programming are utilized to test the usefulness for facial expression discovery and viola jones algorithm is used. Also, VGG 16 and KDEF Dataset are utilized with convolution neural network model for face acknowledgment and grouping. OpenCV, Tensor flow, Keras, and Numpy are used in image processing. CNN model is used for the classification purpose. The seven looks referenced above show the changed states of an individual. Notwithstanding, the outcomes demonstrate that the organization engineering planned has preferable headways over existing calculations. This can be broadly used as an application in numerous areas, for example, instruction, industry, clinical and gadgets.

Acknowledgment:

The authors would like to thank Krish Naik who is cofounder of iNeuron, and Snehan Kekre who is instructor at Cousera. We would also like to thank laboratory at the NIET college for providing the infrastructure support and e-learning. We also want to thank our mentor for helping us in the completion of this project. Finally, we are thankful to all who volunteered on whom we have tested our model.

References

[1] Ken Nozaki, HisaoIshibuchi 1996. Adaptive fuzzy rule based Classification systems, IEEE transactions on Fuzzy systems, Vol.4, No.3. S. Miao, H. Xu, Z. Han and Y. Zhu, "Recognizing Facial Expressions using a Shallow Convolutional Neural Network," in IEEE Access, vol. 7, pp. 78000-78011, 2019

[2] Farid Ghareh Mohammadi, Mohammed Saniee Abadeh 2014. Image Steganalysis using abee colony based feature selection Algorithm. Engineering Applications of AI. Science Direct, pp. 35-43A. M. Yusuf, M. B. Mustaafa, M. and M. Malekzedeh, “emotion

recognition research: an analysis of research focus," Int. J. Speech Technol., vol. 0, no. 0, pp. 1–20, 2018.

[3] Liu, F L C Y, 2015. Improving Steganalysis by Fusing SVM Classifiers for JPEG Images. IEEE, pp. 185-190.

[4] Dharma, A. S. A. a. D., April 2018. Novel directional patterns and a Generalized Supervised Dimension Reduction System (GSDRS) for facial emotion recognition. Springer Link, 77(8), p. 9455–9488 .

[5] Ji, Y. Z. a. Q., May 2005. Active and Dynamic Information Fusion for Facial Expression Understanding from Image Sequences. IEEE, 27(5), p. 699–714.

[6] Dr. Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi 2020 J. Phys.: Conf. Ser. 1432 012087

[7] Mundher Al-Shabi, W. P. C. a. T. C., Aug 2019. Facial Expression Recognition Using a Hybrid CNN– SIFT Aggregator. ResearchGate(RG), pp. 139-149.

[8] CUDA Toolkit 10.1 Update 2 available from [2 September 2019]

[9] C. Szeigedy, V. Vanhouicke, J. Shleins, and Z. Woijna, "Rethinking the Inception Architecture for Computer Vision," 2014.

[10] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," Pattern Recognition, vol. 36, no. 1, pp. 259–275, 2003.

[11] Isha Talegaonkar, Kalyani Joshi, Shreya Valunj, Rucha Kohok, Anagha Kulkarni, (2019), Real Time Facial Expression Recognition using Deep Learning, (ICCIP)International Conference on Communication and Information Processing

[12] Dweepna Garg, Parth Goel, Sharnil Pandya (2018), A Deep Learning Approach for Face Detection using YOLO, Institute of Electrical and Electronics Engineers (IEEE).

[13] Hao Wang, S. W. a. B. F., August 2018. Facial expression recognition using iterative fusion of MO-HOG and deep features. The Journal of Supercomputing.

[14] Liu, F L C Y, 2015. Improving Steganalysis by Fusing SVM Classifiers for JPEG Images. IEEE, pp. 185-190.

[15] MalyalaDivya, R Obula Konda Reddy, C Raghavendra (2019), Effective Facial Emotion Recognition using Convolutional Neural Network Algorithm, (IJRTE) International Journal of Recent Technology and Engineering ISSN: 2277-3878, Volume-8 Issue-4

[16] Hao Wang, S. W. a. B. F., August 2018. Facial expression recognition using iterative fusion of MO-HOG and deep features. The Journal of Supercomputing.

[17] Ken Nozaki, HisaoIshibuchi 1996. Adaptive fuzzy rule-based Classification systems, IEEE transactions on Fuzzy systems, Vol.4, No.3.

[18] CUDA Toolkit 10.1 Update 2 available from <https://developer.nvidia.com/cuda-downloads> [2 September 2019]

[19] Hao Wang, S. W. a. B. F., August 2018. Facial expression recognition using iterative fusion of MO-HOG and deep features. The Journal of Supercomputing.

[20] Face Detection and Tracking using OpenCV. S.V.Viraktamath, Mukund Katti, Aditya Khatawkar, Pavan Kulkarni. 3, s.l.: SIJ, July-August 2013, The Standard International Journals (The SIJ) , Vol. 1, pp. 45-50. ISSN: 2321 – 2403