IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, June 2022
Impact Factor: 6.252

Facial Emotion Recognition with Music Recommendation

Muskan Jamadar[1], Sakshi Jadhav[2], Rutuja Dame[3], Mayuri Patil[4], Sujata Patil[5], Prof. R. S. Nejkar[6]
Students, Department of Computer Science and Engineering[1,2,3,4,5]
Assistant Professor, Department of Computer Science and Engineering[6]
Sanjeevan Engineering and Technology Institute, Panhala, Maharashtra, India

Abstract: *Facial emotion recognition is advancement in computer vision and machine literacy and with the help of this computing technology it's easy to identify mortal emotion through images also. In this paper we propose fashion call facial emotion recognition with music recommendation using Convolutional Neural Network (CNN). The FER is grounded on three corridor. The first part removes the background from the picture, the alternate part concentrates and maps the facial point vector birth, and the third part recommends music grounded on prognosticated emotion. So, to train the images online database is taken from the Kaggle and consequently the feelings are labelled with 96 of delicacy. Further, grounded on emotion vaticination music or audio song will be recommended from the database.*

Keywords: Facial Emotion Recognition

I. INTRODUCTION

Facial Emotion Recognition is grounded on Deep literacy. Deep literacy is a branch of Machine literacy which is fully grounded on Artificial Neural Network as neural network is going to mimic the mortal brain. One of the main corridors of Neural Network is Convolutional Neural Network (CNN) which comes under deep literacy. CNN are made up of neurons. A CNN is neural network that has one or further convolutional layers and are used substantially for image processing, bracket, segmentation and also other bus identified data. The main advantage of CNN is that it automatically detects the important features without any mortal supervision. This is why CNN would be an ideal result to computer vision and image bracket problems. The facial emotion recognition is a process of detecting mortal feelings and facial expressions. We concentrate on five essential facial expressions which are wrathfulness, sad, happy, stressed, and surprised. This design aims for expressional examination and to characterize the given image into these five essential feelings. Haarcascade Algorithm is substantially used to identify Mortal face and helps in background junking of a input image. After junking of background only face vector is taken into consideration. Grounded on the vaticination of emotion through face vector music will be recommended from database.

1.1 Convolutional Neural Network (CNN)

Convolutional Neural Network is a Deep literacy algorithm which can take in an input image assign significance (learnable weights and impulses) to colorful aspects/ objects in the image and be suitable to separate one from the other. A CNN is a type of artificial neural network used in image recognition and processing, that's especially designed to reuse pixel data. Images contain data of RGB combination. The computer does not see an image, all it sees is an array of figures. Color images are stored in 3- dimensional array. The first two confines correspond to the height and range of the image (the number of pixels) The last dimension corresponds to the red, green and blue colors present in each pixel. Three layers of CNN: There are three types of layers in CNN

1. Convolutional Layer: Convolutional Layer a typical neural network each input neuron is connected to the coming retired subcaste. In CNN only a small region of the input subcaste neurons connects to the neuron hidden subcaste
2. Pooling Layer: Pooling Subcaste The pooling subcaste is used to reduce the dimensionality of the point chart. There will be multiple activation and pooling Players inside the retired subcaste of the CNN.
3. Fully - Connected layer: form the last many layers in the network. The input to the completely connected

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, June 2022
Impact Factor: 6.252

subcaste is the affair from the final Pooling of convolutional subcaste, which is smoothed. and also fed into the completely connected subcaste. This is how CNN works with image processing in facial discovery.

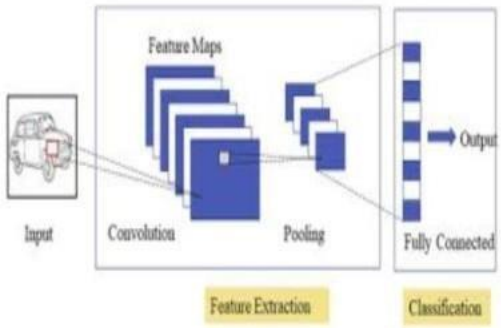

Fig.1 CNN Layers

1.2. Haar Cascade Algorithm

Haar cascade algorithm is used to identify faces in an image or a real time videotape. Haar cascade is a Machine literacy grounded approach where a lot of positive and negative images are used to train the classifier

Working of Haar cascade in image processing A Haar- suchlike point considers conterminous blockish regions at a specific position in discovery window, Sum up the pixels intensities in each region and calculates the difference between these totalities. This difference is also used to classify subsections. of image. Positive images contain the images which want our classifier to identify. Negative images don't contain the object we want to descry. With our design directory structure reviewed, we will move on to apply our Haar cascade in real time with Open CV.

A. Frontal Face Detector

A facial recognition system uses biometric to map facial features from a image or video. It compares the information with database known faces to find a match.

II. LITERATURE REVIEW

Facial expression is a common signal for all humans to convey the mood, facial emotion recognition using Convolutional Neural network (FERC) (1) Ninand Mehendale said that with the recent advancement in Computer Vision and machine learning it is. possible to descry emotion from images but the methodology of the same is relatively tricky and may beget to be the time- consuming process. Meharabian A Nonverbal Communication, Routledge, London (2). Mehrabian said that indeed through our Society Subtly discourage the Verbal expression of emotion utmost of us, the presumably Conferring to our places no way the less manage to express likes dislikes status differences personalities as well weakness in non-verbal ways using Oral expression gesture posture and movements the book was published in 2017. [3] Bartlett M, Little worth G, Viral E, Lee K, Cetin M, Ercil A, Magellan J (2008) Says that the computer vision field has advanced to the point that we are now able to begin to apply automatically facial expression recognition system to import research questions in behavioral science. The Machine perception lab at UC san Diego has developed as a system based on machine learning for fully automated detection of 30 actions from the facial action coding system (FACS). 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.[4]. Russell JA (1994) is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. Psychol Bull 115(1):102 Russell says that Emotion is universally recognized from facial expressions. A review of the methods used in that research raises questions of its ecological, convergent, and internal validity forced-choice response format, within subject design, preselected photographs of posed facial expression and other feature o method are each problematic. When they are altered less supportive or non-supportive results occurs.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, June 2022
Impact Factor: 6.252

III. IMPLEMENTATION

3.1 Modules Description

This is the 4-step process. In the first, we load the XML file for detecting the presence of faces and then we retain our network with our images on 5 different categories. After that, we import the label_image.py program and set up everything in real time. The modules are as follows:

1. Retraining the Network- TensorFlow Image.
2. Classifier Implementation of HAAR CASCADES.
3. Importing the Re-trained Model and Detection of Facial Emotion
4. Music Recommender.

Step 1: Retraining the Network-TensorFlow Image Classifier.
We're going to produce an image classifier that identifies whether a person is happy, sad and so on. And also show this textbook on the OpenCV window. This step will correspond of several sub-steps We need to first produce a directory named images. In this directory, produce 5 or 6sub-directories with names like happy, sad, angry calm, and neutral. Now fill these directories with separate images by downloading them from the internet., in "Happy" directory, fill only those images of a person who are happy. Now run the "face-crop.py" program. To retrain the network, we will be using Mobile net Model which is relatively fast and accurate.

Step 2: Implementation of HAAR CASCADES.
We'll be using the "Frontal Face Alt" classifier for detecting the presence of face in the webcam. This train includes with this depository. Next, we've the task to load this train, which can be set up in thelabel.py program. Everything can be set with thelabel.py program. In this process, background of an image or videotape is completely removed and only face vector is considered as a farther input.

Step 3: Importing the Re-trained Model and Detection of Facial Emotion.
For detecting the emotion, we've put everything under a train (say "label_tf. py") from where we can get everything. It'll open a new window of Open CV and also identifies the facial expression.

Step 4: Music Recommendations
It suggests music by rooting different facial feelings of a person.

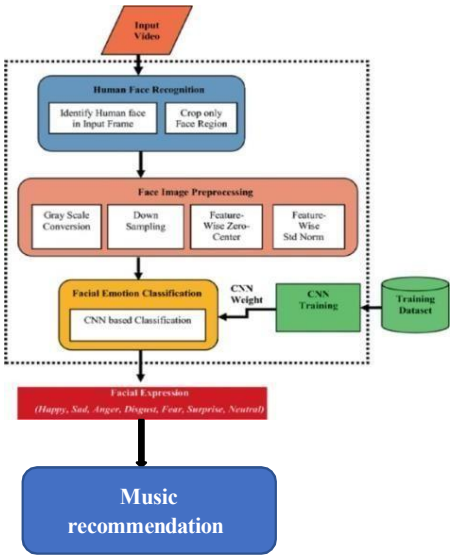

Fig. 2. System Architecture

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, June 2022
Impact Factor: 6.252

IV. METHODOLOGY

Convolutional neural network (CNN) is the most popular way of assaying images. CNN is different from a multi-layer perceptron (MLP) as they've hidden layers, called convolutional layers. The proposed system is grounded on a two-position CNN frame. The first position recommended is background junking, used to prize feelings from an image, the conventional CNN network module is used to prize primary expressional vector (EV). The expressional vector (EV) is generated by tracking down applicable facial points of significance of the input image fed to the first- part CNN (used for background junking) generally consists of shapes, edges textures, and objects along with the face. The edge sensor, circle sensor, and corner sensor pollutants are used at the launch of the convolutional subcaste 1. Once the face has been detected, the alternate- part CNN sludge catches facial features, similar as eyes, cognizance, lips, nose, and cheeks. The edge discovery pollutants used in this subcaste. The alternate- part CNN consists of layers with 3 × 3 kernel matrix. These figures are named between 0 and 1 originally. These figures are optimized for EV

4.1. Train the Model

We've trained the model using Convolutional network for accurate vaticination. This trained model to each and every emotion will assign the markers emotion for each and every emotion generates an n- dimensional array of pixels of the test images. We've taken online dataset from Kaggle and Created Which Contains past a Csv train too. All training & testing After prosecution this csvfile is one formerly divided into two different lines for training and other for testing Captured the image is honored the feelings.

4.2. Face Detection and Image Processing

In this module, we will be using the “Frontal Face Alt” classifier for detecting the presence of face in the webcam. This train includes with this depository. Next, we've the task to load this train, which can be set up in thelabel.py program. Everything can be set with thelabel.py program. In this process, background of an image or videotape is completely removed and only face vector is considered as a farther input. Real- time videotape is captured using the camera at the rate of 30 frames per second(fps). The frames are in BGR (Blue Green Red) format. It's converted into greyscale format which makes computing_easy.

Fig.3. Conversion RGB image to gray scale image

4.3. Capturing and Recognizing Emotion

Further web cam gets enabled and videotape streaming gets started. Once the videotape streaming gets started face will be detected through web cam after preprocessing. After the discovery of the face, it'll label the mortal feelings according to its probability generation. Further it quantifies the emotion that how important percent the person is happy, angry, sad or stressed. After this process, it'll be diverted to the music recommendations.

4.4. Music Recommendations

It suggests music by rooting different facial feelings of a person. The song’s will be played from the detected emotion. The feelings are assigned to every song. When the emotion is transferred the separate song and the feelings are numbered are arranged and assigned to every song still, we can use numerous kinds of models to recommend because of their delicacy and we're using a fisher face that contains the PCA and LDA algorithms so it gives the delicacy better than other algorithms. And for the sound medium we're using the palm sound and with frequency and duration it the generally used python library for introductory sound playing for the medium. attained are being compared the values that are present as a threshold.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, June 2022
Impact Factor: 6.252

V. HARDWARE AND SOFTWARE DETAILS

Processor Intel CORE i5 processor with minimal GHz speed. RAM Minimum 4 GB. Hard Fragment Anaconda It's a free and open- source distribution of the Python and R programming languages for data wisdom and machine literacy related operations (large- scale data processing, prophetic analytics, scientific computing), that aims to simplify package operation and deployment. Package performances are managed by the package operation system conda. The Anaconda distribution is used by over 6 million druggies, and it includes further than 250 popular data wisdom packages suitable for Windows, Linux, and MacOS.Spyder Spyder (formerly Pydee) is an open- source cross- platform integrated development terrain (IDE) for scientific programming in the Python language. Spyder integrates NumPy, SciPy, Matplotlib and Python, as well as other open- source software. It's released under the MIT license. Spyder is extensible with plugins, includes support for interactive tools for data examination and embeds Python-specific law quality assurance and soul-searching instruments, similar as Pyflakes, Pylint and Rope. It's available cross-platform through Anaconda, on Windows with WinPython and Python(x, y), on macOS through MacPorts, and on major Linux distributions similar as Arch Linux, Debian, Fedora, Gentoo Linux, open SUSE and Ubuntu.

VI. RESULT AND DISCUSSION

To dissect the performance of the algorithm, extended Cohn – Kanade expression dataset was used originally. Dataset had only 486 sequences with 97 apers, causing delicacy to reach up to 45 outside. To overcome the problem of low efficiency, multiple datasets were downloaded from the Internet and also author's ow filmland at different expressions were included. As the number of images in dataset increases, the delicacy also increased. We kept 70 of 10K dataset images as training and 30 dataset images as testing images. In all 25 duplications were carried out, with the different sets of 70 training data each time. Eventually, the error bar was reckoned as the standard divagation. the optimization of the number of layers for CNN. For simplicity, we kept the number of layers and the number of pollutants, for background junking CNN (first- part CNN) as well as face point birth CNN (the alternate- part CNN) to be the same. In this study, we varied the number of layers from 1 to 8. We set up out that maximum delicacy was attained around. It wasn't veritably intuitive, as we assume the number of layers is directly commensurable to delicacy and equally commensurable to prosecution time. Hence due to maximum accuracy attained with 4 layers; we named the number of layers to be 4. The prosecution time was adding with the number of layers, and it wasn't adding significant value to our study. the number of filters optimization for both layers. Again, 1 – 8 filters were tried for each of the four- subcaste CNN networks. We set up those four filters were giving good delicacy. Hence, FERC was designed with four layers and four filters. As an unborn compass of this study, experimenters can try varying the number of layers for both CNN singly. Also, the vast quantum of work can be done if each subcaste is fed with a different number of flters. This could be automated using waiters. Due to computational power limitation of the author, we didn't carry out this study, but it'll be largely appreciated if other experimenters come out with a better number than 4(layers),

VII. CONCLUSION

FERC is a new way of facial emotion discovery that uses the advantages of CNN and supervised literacy (doable due to big data). The main advantage of the FERC algorithm is that it works with different exposures (lower than 30°) due to the unique 24- number long EV point matrix. The background junking added a great advantage in directly. determining the feelings. FERC could be the starting step, for numerous of the emotion- grounded operations similar as lay sensor and also mood- grounded literacy for scholars, etc.

REFERENCES

[1]. Facial Emotion Recognition using CNN (FERC), Springer Nature Switzerland AG 2020, Ninad Mehendale.
[2]. Mehrabian A Nonverbal communication. Routledge, London Mehrabian, published in 2017.
[3]. Bartlett M, Little worth G, Viral E, Lee K, Cetin M, Ercil A, Movellan J (2008) 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.
[4]. Russell JA (1994) Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. PsycholBull115(1):102Russel

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, June 2022
Impact Factor: 6.252

[5]. Little wort G, White hill J, Wu T, Fasel I, Frank M, Movellan J, Bar-Shan C, Gong S, McOwan PW (2009) Facial expression recognition based on local binary patterns: a comprehensive study. Image Vis Compute 27(6):803

[6]. Caltech Faces (2020) http://www.vision.caltech.edu/html- files/archive.html. Accessed 05 Jan 2020

[7]. The CMU multi-pie face database (2020) http://ww1.multipie.Org/. Accessed 05 Jan 2020 NIST mugsho identification database (2020) https://www.nist.gov/itl/iad/image-group/resources/biometric-special- databases-and-software. Accessed 05 Jan 2020

[8]. 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