Real Time Facial Emotion Recognition to Annotate the Music in Players

Ragavee R, Ragul Krishnan R, Sri Varshini S ,Pavithra G

Department of Computer Science and Engineering, M.Kumarasamy College of Engineering , Karur - 639113.

ABSTRACT: As we all know that face is the index of the mind, face acts as both behavioral and emotional indicator of a person. Through face we can able to detect the mind state of an individual. The emotion of a person can be identified using the expression on the face of the individual. The process of identifying emotion of human using the facial expression as well as the verbal expression is called Emotion Recognition. In our project the major importance is given to the facial expression. The main objective of the project is to recognize the emotion of the people play music in the music player to keep their mind calm and make them to feel better. The face of an individual can be captured using the web camera that is attached to the laptop or computer. Then the captured image is processed and the emotion of an individual is recognized. Then a song is played according to the recognized emotion. For the emotion recognition process SVM classification is used. KNN algorithm is used to classify the music according to the recognized emotion. The three main motto of the project are 1) reduce the stress level of the people because many suicides are caused due to mental stress 2) to help visually impaired people as they can't able to play songs by their own 3) to keep our mind calm and to feel better at any situation. In this project, no datasets are used for the preprocessing of images and their emotion recognition.

KEYWORDS: Facial expression, Emotion Recognition, SVM classification, KNN algorithm, music classification

I. INTRODUCTION

The project involves the dominance of two major domains: 1) Image processing and 2) Machine Learning. Image processing can be characterized as the way toward investigating and controlling the computerized images. A few tasks are performed in the computerized images to get an improved image and to separate some helpful data from the images. Once in a while the advanced images can be accessible as caught images or images that are cut from the recordings. Numerous calculations are utilized for the processing of the images. The calculations are utilized to expel clamors from the images and to perform division of the images. The way toward isolating foundation objects from forefront object of an image is called division. At that point the component indicates are extricated perform different estimations. A portion of the utilizations of image processing are clever transportation framework, remote detecting, moving article following, guard reconnaissance, bio-restorative imaging methods and programmed visual investigation framework. Machine learning is a rising strategy in the cutting edge world and it is an utilization of man-made reasoning. The idea of machine learning enables the PCs to learn and to take choices by its own from the prior experience. It ought to be managed without intercession of the people just as pre-composed projects. Because of this, computerizations are performed in numerous organizations which upgrade the logical research. The real utilizations of machine learning strategies are budgetary administrations, showcasing and deals, government, human services, video observation, online client support, web index results and internet based life administrations. In this venture, the picture

preparing method is utilized to catch the essence of an individual and to process the picture for extricating the element focuses. At that point the feeling can be identified utilizing the extricated highlight focuses. SVM classification is utilized for this procedure. Machine learning systems are utilized for the grouping of music and playing of melodies dependent on the perceived feeling. KNN classification is utilized for the explanation of melodies in the players. The blend of picture handling and machine learning procedures gives improved productivity to the compelling usage of the framework. The continuous catching of pictures without utilization of any datasets diminishes the intricacy and time taken for the preparing.

II.LITERATURE SURVEY

Yingruo Fan, Jacqueline C.K. Lam and Victor O.K. Li[1] have proposed the Multi-Region Ensemble Convolutional Neural Network structure. Through this system, the learning intensity of CNN models has been upgraded. For the assessment of the proposed strategy two understood outward appearance Databases are utilized: AFEW 7.0 and RAF DB. The Acted Facial Expression in the Wild 7.0(AEFW 7.0) Database comprises of numerous video cuts which incorporates 773, 383 and 653 video cuts for preparing, approval and test separately. The Real-world Affective Faces Database(RAF DB) comprises of 12271 preparing tests and 3068 testing tests as pictures. The marked articulations in this strategy are: irate, appall, dread, upbeat, miserable, nauseate, astonishment and impartial. The three sub areas of face that taken into contemplations are: left-eye, nose and mouth. The normal precision of this technique is 76.73%. The framework has the favorable position that the last expectation exactness was improved by the data that is assembled from the various areas of facial pictures. This is the principle preferred standpoint of this strategy. The calculation cost is costly because of the use of lattice registering for the assurance of commitments of individual sub-systems.

Feng-Ju Chang, Anh Tuan Tran, Tal Hassner, Lacopo Masi, Ram Nevatia and Gerard Medioni[2] have proposed a profound learning strategy for the estimation of 3D outward appearance coefficients. This paper acquainted another method with model outward appearances straightforwardly from the picture powers. The proposed profound neural system was known as ExpNet. The subjective and quantitative correlations of outward appearances are made. It clarified about the portrayal of 3D pictures and their demeanors. It utilizes CK+ datasets and EmotiW 17 datasets. The Cohn-Kanade+ is an obliged dataset and it comprises of frontal pictures that taken from the research facility. It contains around 327 face video cuts that are utilized for the acknowledgment of seven feelings, for example, irate, hatred, disturb, dread, cheerful, trouble and shock. The Emotion Recognition in the Wild datasets contains video outlines that are gathered from 54 motion picture DVDs. I t contains around 383 face video cuts that are utilized for the acknowledgment of seven feelings, for example, outrage, nauseate, dread, upbeat, unbiased, pity and astonishment. For the acknowledgment of outward appearance the accompanying advances are included: speaking to 3D countenances and articulations, producing 3D articulation information, preparing ExpNet to anticipate articulation coefficients and evaluating articulations coefficient with ExpNet. The handling velocity of this framework is quicker than its choices. The drawback of this framework is here and there the forecasts are conflicting because of extraordinary survey conditions that are trying for milestone identification techniques. It contains normal exactness of about 88%.

Aishwarya Kulkarni, Ankita Shendge, Vaishnavi Varma and Nisha V.Kimmatkar [3] have proposed a shrewd framework for the programmed identification of human articulations. The framework is separated into four stages as: 1) pre handling and face recognition 2) division and highlight point extraction 3) information decrease and 4) feeling

order. The harsh sets hypothesis is utilized for the preprocessing of pictures. The preprocessing procedure is conveyed with the assistance of ImageNet datasets. The vital Component Analysis (PCA) is connected for information decrease process. The facial segments that taken into contemplations are: left eye, right eye, eyebrows, eyes, lips and nose. All the facial segments are extricated and prepared independently. The Support Vector Machine (SVM) is utilized for the grouping of articulations dependent on the extricated highlight focuses. Two sorts of arrangement plans are utilized in SVM grouping which incorporates: One Against One (OAO) and One Against All (OAA). Utilizing this two grouping plans six essential feelings are perceived as pursues: furious, sicken, dread, upbeat, tragic and shock. It gives viable examination of outward appearances. This framework gives the general exactness of about 95%.

Dhwani Mehta, Mohammad Faridul Haque Siddiqui and Ahmad Y. Javai [4] have proposed a framework for the programmed acknowledgment of human feelings which thus helps in the educating of social knowledge to the machines. Sensors are utilized for the precise location of feelings which builds the productivity and the dependability of the framework. This framework includes the cooperation of Augmented Reality. The blended reality gadget named Microsoft HoloLens (MHL) is utilized for the discovery of feelings. This paper analyzes the outcomes that are gotten from both MHL and webcam. Notwithstanding facial highlights it utilizes voice, forms of non-verbal communication dialects and motions for the acknowledgment of feelings. The means that are engaged with the acknowledgment of feelings are: highlight choice, include grouping, include extraction and classifiers. Three principle databases are utilized for the arrangement of feelings: 1) RGB Database 2) Thermal Database and 3) 3D Database. The RGB Database incorporates CK, JAFEE and MULTI PIE Databases that contains 593, 213 and 7,50,000 pictures individually. The absence of force names in RBG Databases makes burden for the analyses which thusly diminishes the productivity. The warm Database comprises of databases, for example, Imaging, Robotics and Intelligent System Database (IRIS) , National Institute of Standards and Technology (NIST) and Natural Visible and Infrared Facial Expressions (NVIE). The IRIS comprises of 30 pictures in which 28 are men and 2 are ladies. The NIST absolutely contains 1573 individual pictures in that 78 are ladies and remaining are men. At that point the NVIE comprises of absolutely 100 individual pictures. The warm database does not work legitimately amid the varieties in posture, age, scaling and temperature. A standout amongst the most utilized 3D databases is BU – 3DFE. It contains 100 individual pictures in which 56% are female and 44% are male. The downside of utilizing 3D Database is it isn't accessible in wealth which thus influences the precision of the framework. The feelings that are perceived utilizing this framework are: Happy, Sad, Angry, Surprise and Neutral. The exactness of this framework is 90%.

Pratik Gala, Raj Shah, Vineet Shah, Yash Shah and Sarika Rane [5] have structured a music player dependent on the feeling of a person. It takes out the procedure of manual playing of tunes which is a tedious procedure. Rather its plays a tune without anyone else from the locally accessible machine dependent on the perceived feeling. Viola Jones calculation and Principle Component Analysis are utilized for the face recognition process. The facial highlights that are thought about for the acknowledgment of feeling are: eyes pair, nose and mouth. The feelings can be named pursues: Happy, outrage, dismal, sicken and unbiased. In the preparation stage the picture dataset is utilized for the grouping of pictures dependent on feeling. The picture dataset comprises of pictures of different individuals and their different feeling states. The preparing of the framework pursues the accompanying advances: 1) Image upgrade 2) Face identification 3) Feature extraction and Emotion arrangement 4) Play music. In this framework the test pictures coordinates precisely with the pictures that are available in the preparation dataset. The exactness of the framework depends on the highlights that are separated from the pictures amid preparing. The normal precision of the framework is about 94%. The usefulness and the precision of the framework are high when contrasted and others. At once just a single feeling can be identified.

Ivona Tautkute, Tomasz Trzcinski and Adam Bielski [6] have proposed a strategy that utilizes multilayered convolution layer for the feeling acknowledgment of a person. They proposed another technique called Emotional DAN (Deep Alignment Network) for the accomplishment of milestone acknowledgment challenge. AffectNet datasets which is a biggest dataset is utilized in the preparation stage. It contains around 1,00,000 facial pictures that are gathered from web. In the preparation stage, CK+ and JAFEE datasets that contains around 180 pictures of various sexes are utilized. For the element extraction process, the whole face is taken and the patches in the face are disregarded. This framework perceived seven fundamental feelings: glad, tragic, irate, shock, sicken, dread and impartial. As a straightforward structure it classifies feelings as: positive, negative and nonpartisan. The exactnesses differ starting with one dataset then onto the next. The general exactness of this framework is 89.6%.

Out of the blue, Sagar Gupta and Ashutosh Vaish [7] have utilized the Kaze highlights for the feeling location process. The Kaze highlights are utilized for the obscuring of picture. The Gaussian Mixture Model (GMM) is developed after the extraction of Kaze highlights from the pictures. This framework accomplishes high exactness and low calculation cost contrasted with different frameworks. The feeling identification includes the accompanying advances: preprocessing, highlight extraction and order of pictures. The pictures with same shading and size are mulled over for the preprocessing stage. CK+ datasets are utilized for the preparation and testing of the pictures. The means that are engaged with the discovery of feelings are: preparing of pictures, grayscale transformation, extraction of highlights, development of GMM and recognition of feeling. A general exactness of the framework is 87.5%.

Byoung Chul Ko [8] has proposed a framework that perceives feeling dependent on outward appearance of a person. The means engaged with this methodology are: 1) face and facial segment recognition 2) highlight extraction and 3) feeling order. This framework utilizes either facial pictures or video outlines for the feeling discovery process. The datasets that are utilized by this framework are CK+, CE, DISFA, BU-3DFE, JAFEE, B+, MMI, BP – 4D Spontanous and KDEF which contains around 593 video succession, 5060 pictures, 1,30,000 stereo video outlines, 100 facial pictures, 213 pictures, 16,128 facial pictures, 2900 video grouping, 41 facial pictures and 4900 facial pictures. The feelings that are perceived by this framework are: cheerful, miserable, dread, furious, shock, disturb, joyfully dismal, cheerfully astounded, joyfully sickened, unfortunately frightful, tragically irate and tragically astonished. Use of extensive scale datasets and huge measure of memory makes the framework to set aside more effort for the acknowledgment of feeling.

Biao Yang, Jinmeng Cao, Rongrong Ni and Yuyu Zhang [9] have proposed a framework for the acknowledgment of feeling dependent on the weighted blend profound neural system. For the programmed extraction of highlights Weighted Mixture Deep Neural Networks (WMDNN) are utilized. Face discovery, revolution, amendment and information growth are the means that are engaged with the pre-processing procedure. Facial grayscale pictures and Local Binary Pattern (LBP) facial pictures are the two channels of facial pictures. The two channels are prepared by WMDNN.. The highlights are extricated from the entire face for the acknowledgment of pictures. The framework has the capacity of perceiving six fundamental outward appearances: upbeat, pitiful, outrage, appall, amazement and dread. Three primary datasets will be datasets are utilized in the pre-processing stage: CK+, JAFEE and Oulu – CASIA which contains normal correctnesses of 97%, 92.2% and 92.3% separately. The CK+ dataset comprises of 593 arrangements of 123 subjects. At that point the JAFEE dataset incorporates 213 examples of 10 Chinese females and the Oulu – CASIA contains 10800 examples from 80 subjects. The equivalent execution can be accomplished utilizing less demanding systems. Calibrating is required amid the inadequate gathering of information tests. The framework comprises of following strides for the feeling acknowledgment: 1) Pre preparing 2)

Feature extraction from grayscale facial pictures 3) Feature extraction from LBP facial pictures 4) Weighted combination of various yields. The primary goal of this framework is to accelerate the calculation by improving the system. The general precision of this framework is about 96.98%.

ZhanpengZhang, PingLuo, ChenChangeLoy and XiaoouTang [10] have proposed a framework that changed from feeling acknowledgment to relational connection expectation. Every one of the examples in the datasets must contains age, sex and articulations. In this framework, the head stances of an individual likewise thought about. Siamese model is utilized for the expectation of relational relations. Articulation in the Wild (ExpW) datasets are utilized in this framework that contains right around 90,000 web pictures. Seven fundamental feelings can be perceived utilizing this framework: furious, disturb, dread, glad, tragic, shock and impartial. Eight relational attributes can be perceived in this framework: overwhelming, aggressive, trusting, war, benevolent, included, expressive and guaranteed. Left profile, left, frontal, right, right profile are the represents that are utilized for the relational connection expectation. The general exactness of the framework is about 74.5%.

III. EXISTING SYSTEM

Numerous applications are accessible for the acknowledgment of feelings. The current framework includes the accompanying stages: 1) preparing stage 2) pre – handling 3) face discovery 4) highlight extraction 5) feeling acknowledgment. In the more seasoned systems, for the acknowledgment of human feelings numerous datasets are utilized. The preparation stage includes the utilization of these datasets for the accurate acknowledgment of feeling. The datasets may contain pictures, recordings of different subjects and their diverse feelings. In some cases the acknowledgment of feelings of people other than the people in the datasets is troublesome. The quantity of subjects in the datasets differs starting with one dataset then onto the next. A portion of the critical datasets is CK+, JAFEE, ISE datasets, and so forth. Likewise, the preparation period of the datasets devours additional time. The current framework utilized the Viola Jones calculation and dynamic shape demonstrates for the acknowledgment of feelings. The significant disadvantage of the current frameworks were 1) removes such huge numbers of highlight focuses for the acknowledgment of feelings 2) the multifaceted nature of the framework was high 3) the feelings might be wrongly perceived 4) hard to execute continuously condition. The execution of music player isn't such a great amount of powerful in the current framework.
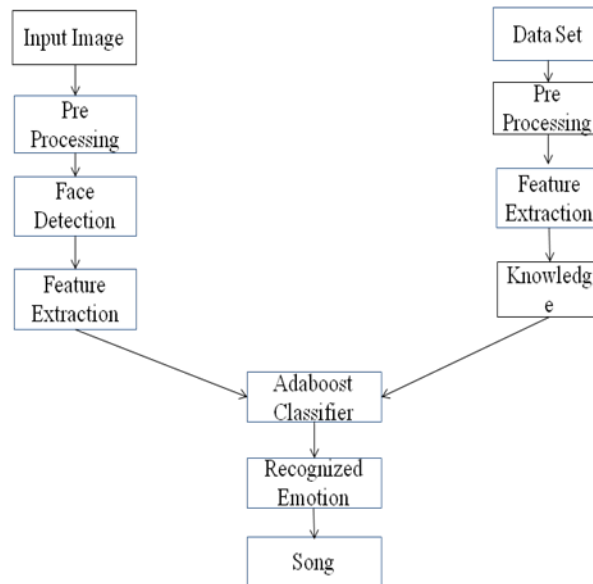
Fig 1: Architecture of existing system

IV.PROPOSED SYSTEM

In this project, the feeling of the individual is recognized without the use of any datasets. This framework includes the accompanying stages: 1) face discovery 2) pre preparing 3) highlight extraction 4) feeling acknowledgment 5) playing music. At first the substance of an individual is distinguished utilizing the web camera that is accessible in the framework. At that point the highlights are extricated from the eyes and lips of the caught face. In light of the extricated highlights the feeling of an individual can be perceived. The feelings are recognized utilizing the SVM classification. After the acknowledgment of feeling, a melody is played by the perceived feeling. The order of music is performed by KNN classification.
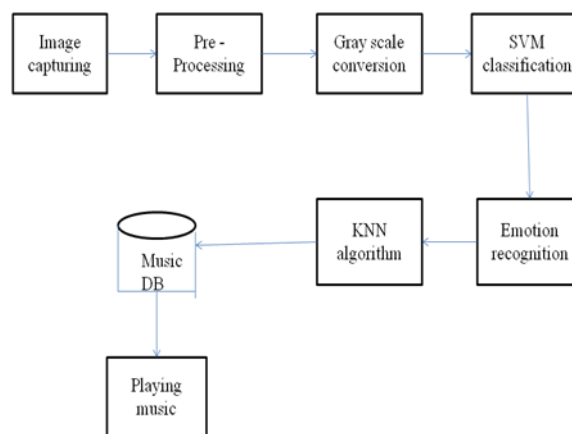
Fig 2: Architecture of proposed system

V. CONCLUSION

We proposed an ongoing feeling acknowledgment framework for the acknowledgment of feeling and play melodies as indicated by the perceived feeling. This framework utilizes the SVM arrangement and KNN calculation for the acknowledgment of feeling and explanation of music in players separately. As far as anyone is concerned, this is the principal continuous facial feeling acknowledgment framework which does not utilizes any datasets for the preparation of pictures. This framework perceives four essential feelings: glad, dismal, irate, rest and nonpartisan. This framework is very helpful to the outwardly tested individuals as they can't ready to play tunes by their own. As the future upgrade of this undertaking, we are intended to build up an android application utilizing this framework. As it is the principal feeling acknowledgment framework without the utilization of datasets, it can't be contrasted and other existing framework. The outcome investigation of the framework in examination with different frameworks that utilizes SVM grouping is as per the following:

TITLE	ALGORITHM	ACCURACY
Real Time Facial Expression Recognition in Video using Support Vector Machines	SVM classification	60.7%
A Survey and Real-World User Experiences in Mixed Reality	SVM classification	90%

"National Conference On Convergent Communication Technologies And Telemedicine Networking (NC²CT²N–2019)"
27th March 2019

Organized by

Department of ECE & BME,Vinayaka Missions Kirupananda Variyar Engineering College, India

A Survey and Real-World User Experiences in Mixed Reality	SVM classification	90%
Intelligent Emotion Detection System Using Facial Images	SVM classification	95%
Real time facial emotion recognition to annotate the music in players	SVM classification	95%

Table 1: Result analysis of proposed system

REFERENCES

[1] Multi-Region Ensemble Convolutional Neural Network for Facial Expression Recognition In: International Conference on Artificial Neural Networks, IEEE, 2018.

[2] Landmark-Free, Deep, 3D Facial Expressions In: IEEE Conference on Automatic Face and Gesture Recognition (FG), Xi'an, China, 2018.

[3] Intelligent Emotion Detection System Using Facial Images In: International Research Journal of Engineering and Technology (IRJET), 2018.

[4] Facial Emotion Recognition: A Survey and Real-World User Experiences in Mixed Reality In: Sensors, Vol. 18, Issue 2, Number 416, 2018.

[5] MoodyPlayer : A Music Player Based on Facial Expression Recognition In: International Research Journal of Engineering and Technology (IRJET), 2018.

[6] I Know How You Feel: Emotion Recognition with Facial Landmarks In: Computer Vision Foundation, IEEE, 2018.

[7] Employing kaze features for the purpose of emotion recognition In: Advances in Intelligent Systems and Computing book series (AISC, volume 710), 2018.

[8] A Brief Review of Facial Emotion Recognition Based on Visual Information In: Sensors 18(2):401, 2018.

[9] Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images In: IEEE Access PP(99):1-1, 2017.

[10] From Facial Expression Recognition to Interpersonal Relation Prediction In: International Journal of Computer Vision, 2017.