FACIAL EMOTIONS BASED MUSIC CLASSIFICATIONUSING DEEP LEARNING ALGORITHM

Karan.S[1], Pradeesh.E[2], Jeba Stanley.J

[1] *S.Karan, Computer Science And Engineering, N.S.N College Of Engineering & Technology, Tamilnadu, India*

[2] *E.Pradeesh , Computer Science And Engineering, ,N.S.N College Of Enggineering & Technology, Tamilnadu, India*

[3] *Mr.J.*Jeba Stanley, *Department of Computer Science And Engineering, Supervisor Assistant Professor, N.S.N College Of Enggineering & Technology, Tamilnadu, India*

ABSTRACT

Facial emotion recognition is the process of detecting human emotions from facial expressions. The human brain recognizes emotions automatically, and software has now been developed that can recognize emotions as well. This technology is becoming more accurate all the time, and will eventually be able to read emotions as well as our brains do. AI can detect emotions by learning what each facial expression means and applying that knowledge to the new information presented to it. Emotional artificial intelligence, or emotion AI, is a technology that is capable of reading, imitating, interpreting, and responding to human facial expressions and emotions. Facial expression is an effective way for humans to communicate since it contains critical and necessary information regarding human affective states. It is a critical part of affective computing systems that aim to recognize and therefore better respond to human emotions. Automatic recognition of facial expressions can be an important component in human-machine interfaces, human emotion analysis, and decision making. As a result, facial expression recognition has become a prominent research topic in human-computer interaction, as well as in the fields of image processing, pattern recognition, machine learning, and human recognition. In this project, we will implement the techniques to automatically detect facial parts using HAAR CASCADES algorithm and classify the emotions using deep learning algorithm. To recognize emotion using the correlation of the facial feature sequence, a deep neural network for emotion recognition based on Convolutional neural network is proposed. The first layer of the deep neural network is the CNN layer, which is used to mine the context correlation in the input facial feature sequence. The second layer is the full-connect layer, which is used to integrate information and act as the major role of the classifier. And present playlist of songs which is suitable for his current mood using K-Nearest Neighbor classification algorithm. In testing side, would supply a test image whose expression it desires to recognize. This test image would be matched with facial databases to play music based on recognized emotions. Finally provide emotion-based music player with improved recognition rate.

Keyword: - *Face detection, Emotion recognition, Webcam, HAAR classification, Music Playlist.*

1. INTRODUCTION

AI (artificial intelligence) is the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions) and self-correction. Particular applications of AI include expert systems, speech recognition and machine vision. AI can be categorized in any number of ways, but here are two examples. The first classifies AI systems as either weak AI or strong AI. Weak AI, also known as narrow AI, is an AI system that is designed and trained for a particular task. Virtual personal assistants, such as

Apple's Siri, are a form of weak AI. Strong AI, also known as artificial general intelligence, is an AI system with generalized human cognitive abilities so that when presented with an unfamiliar task, it has enough intelligence to find a solution. The Turing Test, developed by mathematician Alan Turing in 1950, is a method used to determine if a computer can actually think like a human, although the method is controversial.

2. LITERATURE SURVEY

[1] MULTI-REGION ENSEMBLE CONVOLUTIONAL NEURAL NETWORK FOR FACIAL EXPRESSION RECOGNITION

Facial expressions play an important role in conveying the emotional states of human beings. Recently, deep learning approaches have been applied to image recognition field due to the discriminative power of Convolutional Neural Network (CNN). Here we first propose a novel Multi-Region Ensemble CNN (MRE-CNN) framework for facial expression recognition, which aims to enhance the learning power of CNN models by capturing both the global and the local features from multiple human face sub-regions. Second, the weighted prediction scores from each sub-network are aggregated to produce the final prediction of high accuracy. Third we investigate the effects of different sub-regions of the whole face on facial expression recognition. Our proposed method is evaluated based on two well-known publicly available facial expression databases: AFEW 7.0 and RAF-DB, and has been shown to achieve the state-of-the-art recognition accuracy.

[2] INTELLIGENT EMOTION DETECTION SYSTEM USING FACIAL IMAGES

Facial image recognition has seen many developments in the recent past. The human face plays a prodigious role for automatic recognition of emotion and the interaction between human and computer for some real time applications. The emotions are effectively changeable happenings, so in real life application, detection of emotion is a very challenging task. An emotional state of a person may influence concentration, task solving and decision making skills and as such the main vision is to make systems able to recognize and influence human emotions in order to enhance the productivity and effectiveness of humans in performing any task. Here we have surveyed one such intelligent system to recognize human facial expressions automatically. The system has been segmented into four phases. The first phase involves pre-processing and face detection. The second phase involves segmentation and feature point extraction. The third phase involves data reduction method which reduces the dimension of feature points of the facial components obtained by feature point extraction. The fourth phase involves the emotion classification wherein the face is classified based on the emotions it exhibit.

[3] I KNOW HOW YOU FEEL: EMOTION RECOGNITION WITH FACIAL LANDMARKS

Classification of human emotions remains an important and challenging task for many computer vision algorithms, especially in the era of humanoid robots which coexist with humans in their everyday life. Currently proposed methods for emotion recognition solve this task using multi-layered convolutional networks that do not explicitly infer any facial features in the classification phase. In this work, postulate a fundamentally different approach to solve emotion recognition task that relies on incorporating facial landmarks as a part of the classification loss function. To that end, to extend a recently proposed Deep Alignment Network (DAN), that achieves state-of-the-art results in the recent facial landmark recognition challenge, with a term related to facial features.

[4] FACIAL EMOTION RECOGNITION: A SURVEY AND REAL-WORLD USER EXPERIENCES IN MIXED REALITY

Extensive possibilities of applications have made emotion recognition ineluctable and challenging in the field of computer science. The use of non-verbal cues such as gestures, body movement, and facial expressions convey the feeling and the feedback to the user. This discipline of Human–Computer Interaction places reliance on the algorithmic robustness and the sensitivity of the sensor to ameliorate the recognition. Sensors play a significant role in accurate detection by providing a very high-quality input, hence increasing the efficiency and the reliability of the system. Automatic recognition of human emotions would help in teaching social intelligence in the machines. Its presents a brief study of the various approaches and the techniques of emotion recognition.

[5] EXPNET: LANDMARK-FREE, DEEP, 3D FACIAL EXPRESSIONS

A deep learning-based method for estimating 3D facial expression coefficients. Unlike previous work, this process does not relay on facial landmark detection methods as a proxy step. Recent methods have shown that a CNN can be trained to regress accurate and discriminative 3D morphable model (3DMM) representations, directly from image intensities. By foregoing facial landmark detection, these methods were able to estimate shapes for occluded faces appearing in unprecedented in-the-wild viewing conditions. On those methods by showing that facial expressions can also be estimated by a robust, deep, landmark-free approach. ExpNet CNN is applied directly to the intensities of a face image and regresses a 29D vector of 3D expression coefficients. To propose a unique method for collecting data to train this network, leveraging on the robustness of deep networks to training label noise.

3. EXISTING SYSTEM

There are numerous areas in human–computer interaction that could effectively use the capability to understand emotion. The problem of face detection can be viewed as a problem of binary classification of image frame as either containing or not containing a face. In order to be able to learn such a classification model, we first need to describe an image in terms of features, which would be good indicators of face presence or absence on a given image. The existing approach is generally involves two tasks: The first is for extracting ASM motion based a pyramid ASM model fitting method and the second for the projected motion classification obtained by applying Adaboost classifiers. After the segmentation of face candidates, 68 feature points in each face are then extracted using ASM fitting technique. The system then line up three extracted feature points, eyes and nose part, to the mean shape of ASM, and ignore the other portion of the ASM against the mean face shape of ASM to estimate the geometrical dislocation information between current and mean ASM points coordinates. Then, facial expressions recognition is the obtained based on this geometrical motion using Adaboost classifier. And also extracting features using viola jones. The features that Viola and Jones used are based on wavelets. Wavelets are single wavelength square waves (one high interval and one low interval). In two dimensions, a square wave is a pair of adjacent rectangles - one light and one dark.

3.1 Disadvantage

- Provide large number of features points from facial images
- Emotions may be wrongly classified
- Complexity is high
- Difficult to implement in real time environments

4. PROPOSED SYSTEM

In this project, a novel emotion recognition system based on the processing of physiological signals is presented. This system shows a recognition ratio much higher than chance probability, when applied to physiological signal databases obtained from tens to hundreds of subjects. The system consists of characteristic face detection, feature extraction and pattern classification stages. Although the face detection and feature extraction stages were designed carefully, there was a large amount of within-class variation of features and overlap among classes. In order to detect Emotion from an image, used frontal view facial images. If computers can understand more of human emotion, we can make better systems to reduce the gap of human computer interaction. To handle the emotion recognition problem from arbitrary view facial images. The facial region and others part of the body have been segmented from the complex environment based on skin color model. Thus, in this project showed some differences between different color models that are used to implement the system and which color model can be used where. Another aspect is to extract facial parts from the face. And for that used HAAR CASHCADES to detect the eye and lips region from a face and then by the help of CNN classification detected emotion from those features. From the positioning of mouth and eyes, tried to detect emotion of a face. The proposed system tries to provide an interactive way for the user to carry out the task of creating a playlist. The working is based on KNN mechanisms carrying out their function in a pre-defined order to get the desired output. The classified expression acts as an input

and is used to select an appropriate playlist from the initially generated playlists and the songs from the playlists are played.

4.1 ADVANTAGES

- User friendly application
- Can be used in real time environments with voice alert
- Reduce number of features are extracted
- Easily classify multiple emotions

5. PROPOSED FRAMEWORK

A facial emotion-based music player is a software application that uses facial recognition technology to detect the emotions of the user and selects music that matches their mood. The idea is to create a personalized music experience that responds to the user's emotional state, making it more engaging and enjoyable. In this architecture, display the overall proposed system to classify the facial emotions based on camera capturing and play the music based on recognized emotions.

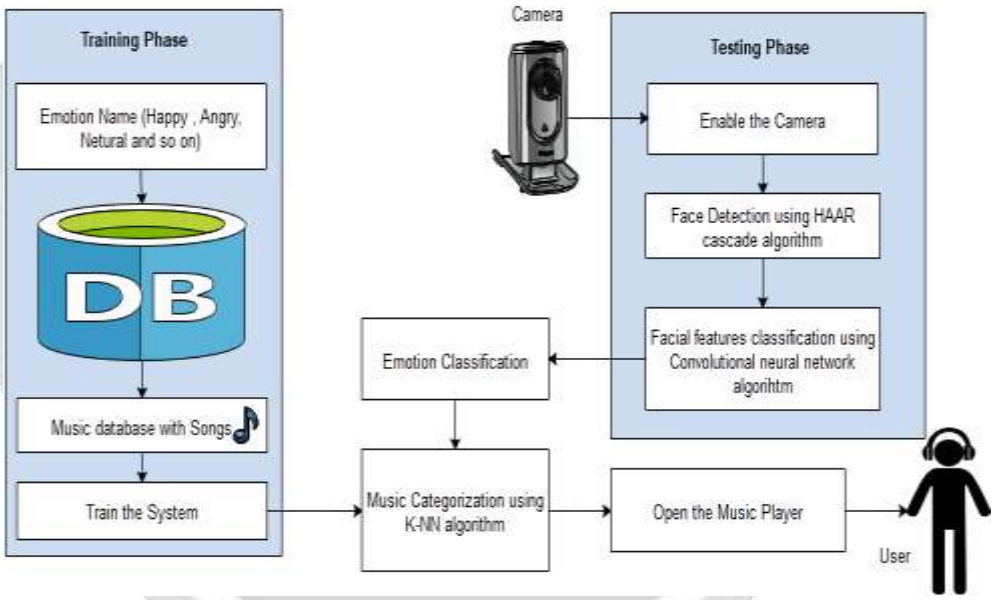

Fig -1: System Architecture

6. SYSTEM IMPLEMENTATION

Emotion recognition is the process of identifying and interpreting the emotions expressed by an individual through various means, such as facial expressions, vocal tone, body language, and speech content. It involves analysing and understanding the emotional state of an individual in real-time or retrospectively. There are various techniques and technologies used for emotion recognition, such as machine learning algorithms, deep learning models, and natural language processing. These techniques are used to analyse and interpret the emotional content of human speech, text, and images. Emotion recognition has various applications, including improving human-computer interaction, developing personalized marketing campaigns, detecting and preventing crime, and improving mental health diagnosis and treatment. It has also gained significant attention in the field of artificial intelligence and robotics, where the ability to recognize and respond to human emotions can enhance the quality of interactions between humans and machines.

6.1 FACIAL IMAGE ACQUISITION

In this module, capture the face image or upload the datasets. The uploaded datasets contain 2D face images. In face registration can identify the faces which are captured by web camera. Then web camera images known as 2D images. Admin can be train the face images with multiple emotions. And also train the music player based on languages.

6.2 PREPROCESSING

In this module, perform the pre-processing steps such as gray scale conversion, invert, and border analysis, detect edges and region identification. The Grayscale images are also called monochromatic, denoting the presence of only one (mono) color (chrome). The edge detection is used to analyze the connected curves that indicate the boundaries of objects, the boundaries of surface markings as well as curves that correspond to discontinuities in surface orientation.

6.3 FACIAL FEATURES EXTRACTION

In this module implement HAAR cascades which are an algorithm employed the computer technology that determines the locations and sizes of human faces in arbitrary (digital) images. It detects facial features and ignores anything else, such as buildings, trees and bodies. Face detection can be regarded as a more general case of face localization. In face localization, the task is to find the locations and sizes of a known number of faces (usually one).

6.4 EMOTION CLASSIFICATION

In this module analyze on the expression recognition for testing facial images. For a testing facial image, we first extract the facial features and then perform the questionnaire estimation, where CNN classifier is used for this purpose. After obtaining the question results, we synthesize facial feature vectors based on testing facial feature vector and use them as the model predictors of the positive model. Finally, the model response corresponding to the expression class label vector is calculated and the expression category of the testing facial image can be obtained based on it.

6.5 MUSICAL CLASSIFICATION

If the emotion is positive means, play happiest songs which are stored in database. Using KNN algorithm to classify the music based on emotions classified by previous modules. K-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The k-NN algorithm is among the simplest of all machine learning algorithms. Both for classification and regression, it can be useful to assign weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. For example, a common weighting scheme consists in giving each neighbor a weight of 1/d, where d is the distance to the neighbor. The neighbors are taken from a set of objects for which the class (for k-NN classification) or the object property value (for k-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. Based on neighborhood values, music are classified and played in emotional database.

7. CONCLUSION

In this project proposed support vector machine algorithm for emotion recognition. Considering an expressive face as a superposition of a neutral face with expression component, we proposed an algorithm to decompose an expressive test face into its building components. For this purpose, we first generate grids for captured face using HAAR Cascade algorithm. Knowing that the face component of the test face has sparse representation in the face database and the expression part can be sparsely represented using the expression database; we decompose the test face into these feature vectors. The elements of the test face along with the vectors are then used for face and expression recognition. For this purpose, the separated components are sparsely decomposed using vectors while the grouping structures of the vectors are enforced into the sparse decomposition. The experimental results on both databases showed that the proposed method achieves competitive recognition performance compared with the state-of-the-art methods under same experimental settings and same facial feature. Based on their emotions, play the songs to recover from depression. In this project we can be implemented the system to using image processing techniques to detect the faces from camera capturing. Then efficiently track the faces and to provide bounding boxes on face images. Finally set the distance limits to identify whether the person is

near to the system or not. And also calculated the person constant seeing conditions. This system can be useful to all aged peoples in various applications such as gaming applications, project works and so on.

8. FUTURE ENHANCEMENT

In future we can extend the system to implement various face detection algorithms to improve the accuracy of the system and implement in different scenarios. We can also be implemented in various types of monitors.

9. REFERENCES

[1]. Fan, Yingruo, Jacqueline CK Lam, and Victor OK Li. "Multi-region ensemble convolutional neural network for facial expression recognition." International Conference on Artificial Neural Networks. Springer, Cham, 2018.

[2]. Chang, Feng-Ju, et al. "Expnet: Landmark-free, deep, 3d facial expressions." 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). IEEE, 2018.

[3]. Kulkarni, Aishwarya, et al. "Intelligent emotion detection system using facial images." Int. Res. J. Eng. Technol (2015): 2395-0072.

[4]. Tautkute, Ivona, Tomasz Trzcinski, and Adam Bielski. "I know how you feel: Emotion recognition with facial landmarks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018.

[5]. Mehta, Dhwani, Mohammad Faridul Haque Siddiqui, and Ahmad Y. Javaid. "Facial emotion recognition: A survey and real-world user experiences in mixed reality." Sensors 18.2 (2018): 416.