MusicBox – Face your new mood

Gangarapu Tharun Kumar
Student
tharunthammy600@gmail.com

Koppolu Leela Krishna
Student
leelakrishnakoppolu@gmail.com

Priyanka Bharti
Assistant Professor
REVA University

Konduru Rupesh
Student
rupesh06082000@gmail.com

Nukala Naveen Teja
Student
naveentej.nukala@gmail.com

Abstract:

In recent years, Because of its promise in the fields of computer vision and human-computer interface, facial emotion detection has gotten a lot of interest. There have been several techniques and applications developed and realized as a result of the ongoing research in this sector. This paper offers a music recommendation application for expression identification that can detect the user's expressions and automatically play a song that improves the user's current mood. We did some research into how music might improve people's feelings in the near term, to collect information and supply clients with a selection of music tracks that are effective at elevating end user moods. The proposed application is developed using Flask, HTML, CSS and JavaScript that identifies expressions, and if the individual is experiencing a sad feeling, a certain song will be played to help him feel better. If the identified emotion is good, on the other hand, suitable music will be played to amplify the pleasant feelings. We effectively developed the suggested Music recommendation application utilizing the Haar-Cascade Classifier method and the CNN.

Keywords - Emotion detection, music recommendation application, Haar-cascade classifier, CNN, Flask, HTML, CSS.

I. Introduction:

Individual communication is an important component of daily living. It communicates precise details and millions of pieces of information between humans, whether through words, tone, or emotion [1]. The easiest way to understand individuals in conversation is to look at their faces and bodies. Face expression, in particular, is a kind of nonverbal communication and is one of the most essential aspects of human communication. The activity of more than one face muscle or skin is characterized as a facial expression. These activities aim to communicate a person's facial emotions, such as happiness, sadness, neutral, anger, and surprise [1]. Facial detection and expression detection are two distinct concepts that have recently sparked attention in domains such as digital image processing, computer vision and pattern recognition [5]. Face detection is a two-step procedure that involves locating faces in a photograph or video, regardless of whether or not

they belong to a person. The program then forms a box around every face in the image, regardless of mood, alignment, face position, brightness, or complex backgrounds Emotion detection is one of the most common facial expressions and the most common way of assessing a person's feelings. Expression detection is now widely employed in a variety of applications, including Tracking, image investigation, multimedia indexing, citizen application, security, and adaptive computer-human interfaces in multimedia environments are all examples of smart card applications [15]. Furthermore, emotion detection is a more reliable and cost-effective technique to learn what customers think about their items. Most businesses and stores employed traditional marketing methods such as advertising, sale points, client feedback, and price in the past. These approaches may be effective in some circumstances, but they are costly, time-consuming, and potentially unreliable.

II. Literature Survey:

Facial expressions are regarded as the most significant aspect of human communication because they allow us to discern the intentions of others. People often derive the emotional states of others, such as happiness, anger, sadness, and surprise, through two major channels: speech tone and facial expression [1].

Learned a lot about how to use the Haar-cascade classifier to detect the faces from the images from the paper published by Viola and Jones who have developed the Haar-cascade classifier [2].

According to Mehrabian [5], nonverbal communication accounts for two-thirds of all human communication, with facial expression accounting for the majority of this percentage. While verbal communication only conveys one-third of emotion. As a result of its academic significance and economic potential, face emotion recognition is currently one of the most important disciplines in computer vision. Looking back at the studies on this subject from its start, we can see that the methods utilized by researchers may be split into two categories: traditional and neural networks-based methods.

Meanwhile, the researchers used deep learning approaches to lessen the reliance on face physics models and replace it with end-to-end learning models. There are other deep learning models, but the Convolutional Neural Network (CNN) is the most extensively used model in the field of face recognition. Have gone through a paper published on Convolutional Neural Networks [3], where we have learned a lot about CNN and how to detect the user’s mood in real-time.

Flask is a simple and lightweight Python web framework that provides handy tools and functionalities for constructing Python online applications [4]. We have gone through a paper on Flask and python technology to know about how to create an online application.

According to D. Mazinanian and N. Tsantalis[11], learned the syntax and new techniques to do frontend styling to the HTML templates.

According to Sharma, Aakanksha [12], and an online site- www.w3schools.com [13], tried to learn about HTML and its tags to create beautiful templates for our application.

III. *Objectives:*

The main objective of the project is to capture a person's feelings through the expressions on their face. A music player is meant to capture human emotion using the web camera interface accessible on computing platforms. The program captures the user's data and image, then refines it using image segmentation and image processing algorithms. Then it takes characteristics from a target human's face and tries to figure out what emotion he or she is feeling. The purpose of the project is to lift the user's spirits by playing music that is suited for the scenario.

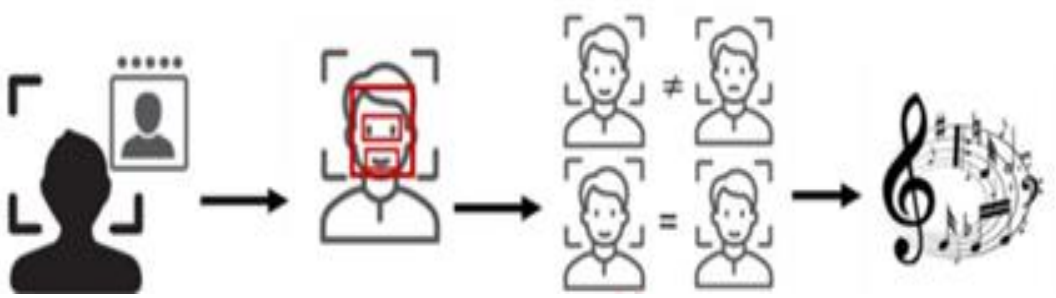
By capturing the user's image, the user's needs can be met. The best way of communication has existed since ancient times. Facial expression recognition is the most well-known form of expression analysis.

IV. Methodology:

The suggested application is a music recommendation software that can automatically recognize emotions. The photos that will be used as input to the proposed application are captured using a camera, as shown in Figure 1, after which it is delivered to the expression detector, which classifies them into one of five categories: "Happy," "Neutral," "Sad," “Surprise” and "Angry."

Figure - 1

Songs will be played based on the expression of emotion identified. Users will be able to listen to party tunes if the user expression is classified as happy, for example. We utilized a Kaggle FER dataset with 24,000 pictures representing four different emotions: happy, neutral, angry, fear, surprise, and sadness. The phases of the suggested approach are depicted in Figure 2.



User Image Capture Face Detection Emotion Detection Play music

Figure -2

A. Image Capturing

As a first stage in the proposed method, we use a built-in laptop webcam to capture an image of the user's face. One face in the frontal position against a continuously illuminated backdrop is required for the recommended method to successfully analyze a face image. Also, nothing on the user's face, such as spectacles, should hinder the detecting process. We allow the user to retake the image to acquire a picture that fits the proposed application requirements if the original image fails to do so.

B. Face Recognition

The system will recognize the face using the Viola-Jones approach (Haar-cascade classifier) after taking the image. This algorithm is recognized as one of the earliest real-time object recognition foundations. Simply said, Viola-Jones scans images and detects the characteristics of the face in the image using a sub-window. The image is reduced to merely containing the face once it has been identified to increase the proposed system's performance. The flowchart in figure-3 depicts the entire procedure.

C. Emotion Detection

The end user's emotion must next be determined. To accomplish so, we use CNN, a Tensor Flow technique that is well-known for facial emotion detection. Finally, we are able to recognize facial expressions.

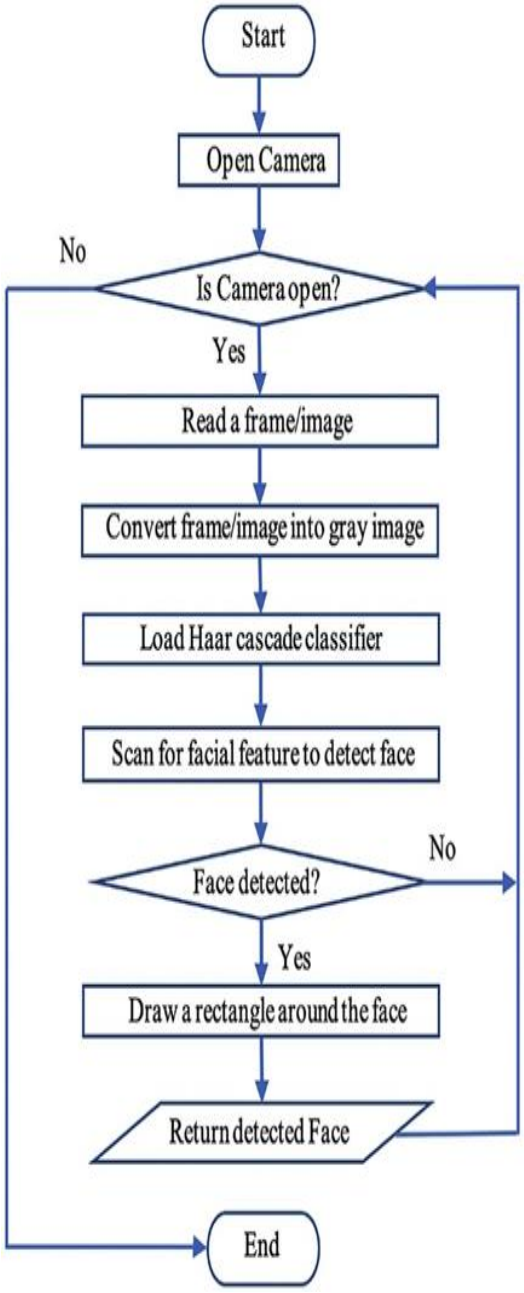


Figure – 3

D. Playing the appropriate Emotion enhancement song

The proposed system will play the appropriate song based on the identified emotion. Because there are four emotions, we have four different emotion-based songs database with carefully selected music samples. The party songs will be active for happy emotions, while the happy songs will be activated for neutral emotions. For unpleasant emotions such as sadness and anger, happy songs will be played to improve the user's mood.

V. Experimental Results:

We were able to effectively construct a music recommendation application that uses the Haar-Cascade classifier with the help of CNN to recognize the emotions expressed in the suggested work and depending on the obtained picture, to imply an appropriate song will be played, with the goal of to enhance the user's experience. Figures 4, 5, and 6 illustrate a snapshot of the proposed system and how it generates a suitable list depending on the detected mode.

Figure – 4

We have trained the model with a 24,000 images dataset and tested it with 6,000 images. Finally, we got an accuracy of around 69% as in figure – 7.

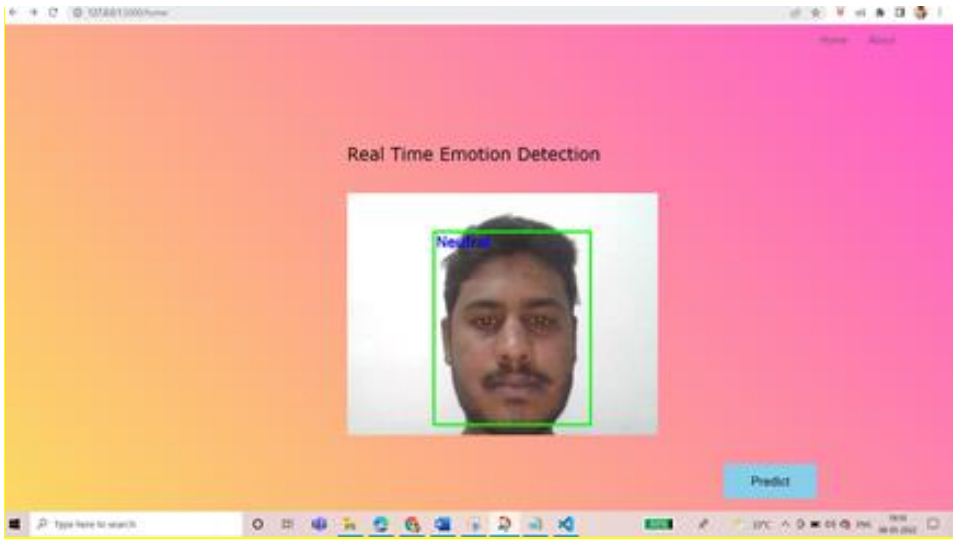


Figure – 5

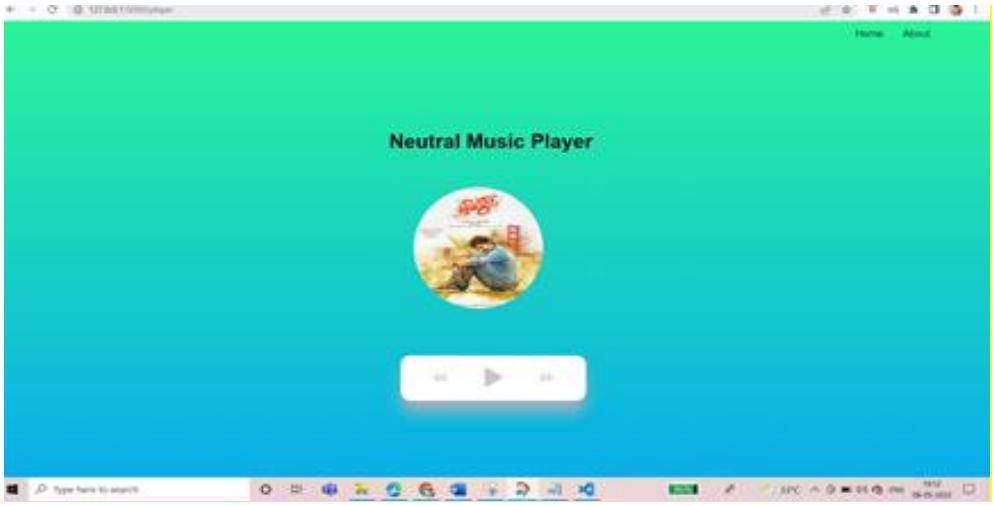


Figure - 6

Figure - 7

Confusion matrix as shown in figure – 8.

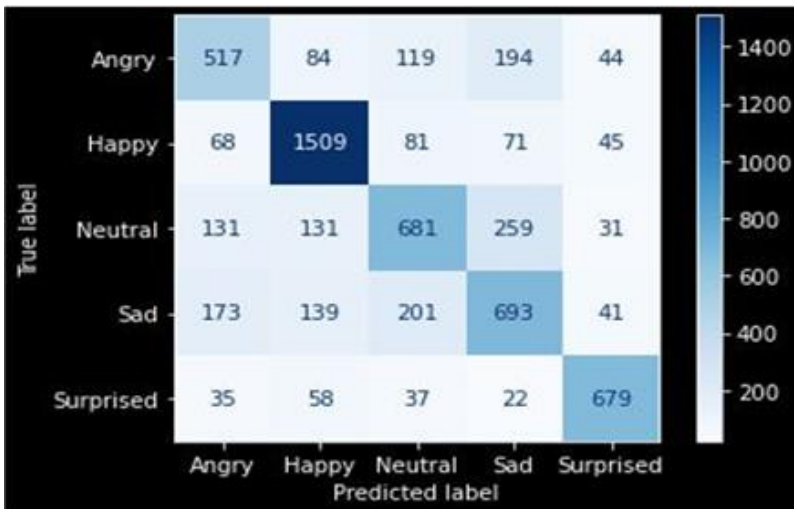


Figure - 8

VI. Conclusion:

We present an outline of how songs can affect a end user's mood and how playing a song automatically can improve a user's mood in this paper. We also demonstrated approaches for detecting emotions. We employed MobileNetV1 Tensor Flow and the viola-Jonze Algorithm to develop the proposed system. The technology in place was able to detect the user's emotions. The algorithm was able to distinguish happy, sad, neutral, and angry emotions. After assessing the user's emotion, the suggested system created a playlist for the user that included music that matched the mood. The song was played to boost the end user's emotion after study on what sort of song is suited for each mood.

VII. References:

[1] B. C. Ko, "A Brief Review of Facial Emotion Recognition Based on Visual Information," sensors, vol. 18, p. 401, 2018

[2] P. Viola, M. Jones: Rapid Object Detection using a Boosted Cascade of Simple Features, Conference Paper in Computer Vision and Pattern Recognition, 2001, Vol.2.

[3] R. Chauhan, K. K. Ghanshala and R. C. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), 2018, pp. 278-282, doi: 10.1109/ICSCCC.2018.8703316.

[4] Aslam, Fankar & Mohammed, Hawa & Lokhande, Prashant. (2015). Efficient Way Of Web Development Using Python And Flask.. International Journal of Advanced Research in Computer Science. 6.

[5] A. Mehrabian, "Communication without words," Communication theory, pp. 193-200, 2008.

[6] J. James Anto Arnold, H. Immanuel James, J. Maria Masilla Ruban, R. Saranya, M. Tamilarasan.

"EMOTION BASED MUSIC RECOMMENDATION SYSTEM": p-ISSN: 2395-0072, IRJET 2019

[7] M Uma and S Matilda Florence "Music Recommendation System based on User Facial Expressions", 2020

[8] S. Lukose et al "Music player based on voice signal emotion identification", 2017

[9] Sharik Khan, Hafeez Kabani, Omar Khan, Shabana Tadvi" Emotion Based Music Player" International Journal of Engineering Research and General ScienceVolume3, Issue1, January-February, 2015

[10] Yaser Khan Chris Huybreqts Jaeyoun Kim Thomas C Butcher "Real-Time Emotion Recognition From Audio Signal" Report-2021

[11] D. Mazinanian and N. Tsantalis, "An Empirical Study on the Use of CSS Preprocessors," 2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER), 2016, pp. 168-178, doi: 10.1109/SANER.2016.18.

[12] Sharma, Aakanksha. (2018). Introduction to HTML (Hyper Text Markup Language) - A Review Paper. International Journal of Science and Research (IJSR). 7. 1337-1339.

[13] Kumar Raja, D. R., Hemanth Kumar, G., Basha, S. M., & Ahmed, S. T. (2022). Recommendations based on Integrated Matrix Time Decomposition and Clustering Optimization. *International Journal of Performability Engineering*, *18*(4).

[14] Patil, K. K., & Ahmed, S. T. (2014, October). Digital telemammography services for rural India, software components and design protocol. In *2014 International Conference on Advances in Electronics Computers and Communications* (pp. 1-5). IEEE.

[15] https://towardsdatascience.com/face-detection-recognition-and-emotion-detection-in-8-lines-of-code-b2ce32d4d5de

[16]https://medium.datadriveninvestor.com/review-on-mobilenet-v1-abec7888f438

[17] Sudha Veluswamy, Hariprasad Kanna, Balasubramanian Anand, Anshul Sharma “METHOD AND APPARATUS FOR RECOGNIZING AN EMOTION OF AN INDIVIDUAL BASED ON FACIAL ACTION UNITS” US2012/0101735A1

[18] Markus Mans Folke Andreasson ”GENERATING MUSIC PLAYLIST BASED ON FACIAL EXPRESSION” US8094891B2