JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# Smart Emotion Based Music Player Using Python Image Processing

**Prof. Shruti Naik[1] , Tarush Wagh [2] , Parth Shinde[3] , Suhas Rathod[4] , Neha Shivarkar [5]**

[1] Asst. Professor, Department of Computer Engineering, JSPM's Narhe Technical Campus, Pune, Maharashtra, India.
[2,3,4,5] BE Students ,Computer Engineering, JSPM's Narhe Technical Campus, Pune, Maharashtra, India.

***Abstract :*** When predicting a person's emotions and mood, the face is crucial. Typically, a camera is used to extract human emotions. Numerous applications are being created based on the recognition of human emotions. A few examples of applications for emotion detection include business notification recommendation, e-learning, the diagnosis of mental illness and depression, and criminal activity. We create a prototype for a dynamic music recommendation system based on human emotions in this system that is being suggested. Songs for each emotion are taught based on human listening patterns. The emotion on a real person's face is recognised using an integration of feature extraction and machine learning techniques. Once the mood is determined from the input image, the appropriate music will be played to keep the users' attention. With this method, the application connects with user emotions and gives them a personal touch. In order to create music players that are emotion-based, our technology is focused on understanding human emotions using computer vision and machine learning approaches. We use CNN model architecture for music recommendation and emotion recognition in our experiments.

***IndexTerms* - BigData,Machine Learning, Neural Networks, CNN (Convolution Neural Network) .**

## 1. Introduction

Music has a significant positive impact on a person's life and is a vital source of enjoyment for music lovers and listeners. Due to the quick advancements in multimedia and technology, several music players have been developed in the present period with features like fast forward, reverse, variable playback speed, general classification, streaming playback with multicast streams, including volume modulation, etc. The user still needs to manually browse the song playlist and choose songs that suit their current mood and behavior, even though these skills may satisfy their basic needs.
By looking at their faces, you can locate and recognize people the quickest. No face-recognition algorithm will work without faces. Mood recognition based on emotion is one of the current concepts in several fields that offers solutions to many issues. For face identification, in addition to the usual challenges of obtaining facial images in uncontrolled contexts, such as different positions, different lighting, and different facial expressions, there are also different sound frequencies.
The database, which is used to compare the elements of sound frequency and facial features, is the most important component in any face and mood identification system. The database is built using the face's features, which are then saved there. Then, this database is used to evaluate the face and mood using a variety of algorithms. The basic goal of this system is to analyze an input face image to detect an individual's mood, and then use the results of that analysis to play an audio clip. Here, the input face image and the trained face image are compared using a face recognition approach. The suggested approach is simple to implement, efficient, and exact. This process yields accurate findings when compared to the present approach. Systems are essential in disciplines involving detection and recognition. This method therefore produces meaningful outcomes much more quickly than more traditional ones.

### 1.1. Problem Statement

Music listeners find it challenging to manually generate and group the play list when they have hundreds of tracks. Additionally, since songs are occasionally added but never played, taking up a lot of storage space and requiring manual song deletion by the user, it might be difficult to keep track of all the songs. Users are need to manually select tracks depending on interest and mood each time. When their playing habits change, users find it difficult to rearrange their playlists and play music. Currently, existing applications employ playlists and various mood categories to categorise music. Given that the developer is here also assuming the role of the user, it is difficult to say what kind of play list the user would prefer when they click on the play list created by the developer. Using the findings from this research, we may use a person's emotions to gauge their mood and then play music to match that mood.

## 2. LITERATURE REVIEW

Table no.2.1: Table of literature

| Sr.No | Author | Country | Name | Summary |
|---|---|---|---|---|
| I | Shlok Gilda, Husain Zafar, Chintan Soni and Kshitija Waghurdekar | India | Smart Music Player Integrating Facial Emotion Recognition and Music Mood Recommendation | Songs have long been a preferred form of expression for expressing and comprehending human emotions. We can greatly benefit from reliable emotion-based classification systems in helping us decipher their significance. The study of emotion-based music classification has not, however, produced the best outcomes. In this study, we introduce EMP, a cross-platform affective music player that makes music suggestions based on the user's current mood. By integrating emotion context reasoning capabilities into our adaptive music recommendation system, EMP offers smart mood-based music recommendations. Three components make up our music player: the Emotion Module, the Music Classification Module, and the Recommendation Module. Music classification, recommendation, and emotion modules. With an accuracy of 90.23 percent, the Emotion Module uses deep learning algorithms to discern the user's mood from an image of their face. |
| II | Swathi Swaminathan and E. Glenn Schellenberg | Canada | Current emotion research in music psychology | Because it expresses emotion and controls affect, music is, at least in part, a universal language. Music psychologists frequently investigate the connections between music and emotion. Here, we discuss three new findings: (a) the expression and perception of emotion in music; (b) the emotional effects of listening to music; and (c) determinants of musical preferences. |

| | | | | |
|---|---|---|---|---|
| III | Kyogu Lee and Minsu Cho | South Korea | Mood Classification from Musical Audio Using User Group-dependent Models | In this work, we present a music mood categorization system that takes a user's profile into account. We do this because we believe that a user's assessment of a song's mood is subjective and can change depending on factors like age or gender. To do this, we first establish a collection of general mood descriptors. Second, we create various user profiles based on the age and gender of the user. Then, we obtain musical samples for each group so that the statistical models can be trained separately. We demonstrate that both models have superior classification accuracy when the test data and the mood model are of the same sort, supporting our claim that the user profiles are crucial for mood perception. By using our method to automatically create play lists, we also show that taking into account the variation in user groups' perceptions of mood has a considerable impact on calculating musical similarity. |

# 3. METHODOLOGY

## 3.1 Supervised learning

A machine learning algorithm known as supervised machine learning trains the model using labelled data. In supervised learning, the algorithm is given the input data (also known as features) and the associated output data (also known as labels). In order to forecast the output variable for brand-new, unforeseen input data, the algorithm must learn a mapping function between the input and output variables.

The reason why supervised learning is named "supervised" is because a person oversees the algorithm during training and assigns the appropriate output labels to the input data. Based on the discrepancy between the algorithm's expected output and the actual output supplied by the human supervisor, the algorithm modifies its parameters. This procedure is repeated until the algorithm's accuracy is adequate.

## 3.2 Convolutional Neural Networks

ConvNets are made to handle data that is presented as numerous arrays, such as a color image made up of three 2D arrays that each hold the pixel intensities for the three different color channels. There are many different types of data modalities that take the shape of numerous arrays, including 1D signals and sequences for language, 2D visuals or audio spectrograms, and 3D video or volumetric images. Local connections, shared weights, pooling, and the utilization of several layers are the four fundamental concepts that underpin ConvNets, which exploit the characteristics of real signals. A typical ConvNet's architecture is divided into various stages. Convolutional layers and pooling layers make up the majority of the initial stages.

Each unit in a convolutional layer is connected to nearby patches in the feature maps of the preceding layer using a collection of weights called a filter bank. Convolutional layers are organized into feature maps. Following that, a non-linearity like a ReLU is used to process the local weighted sum's output. The same filter bank is shared by all units in a feature map.

# 4. SYSTEM ARCHITECTURE

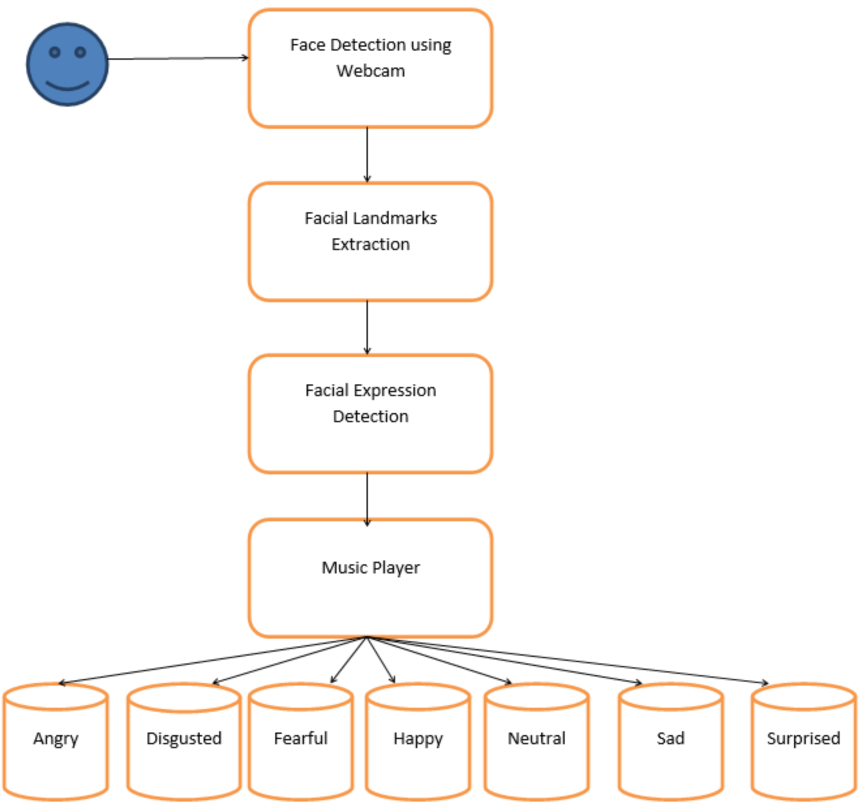
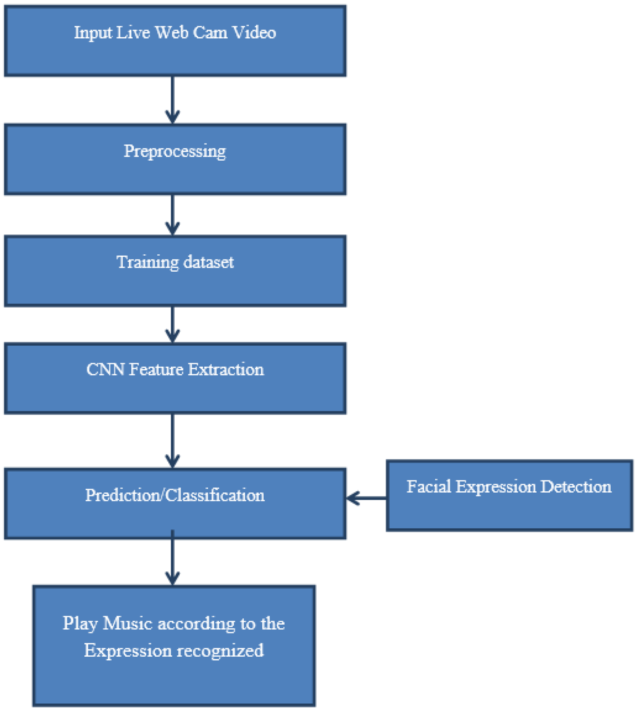



*Figure 1: System Architecture*

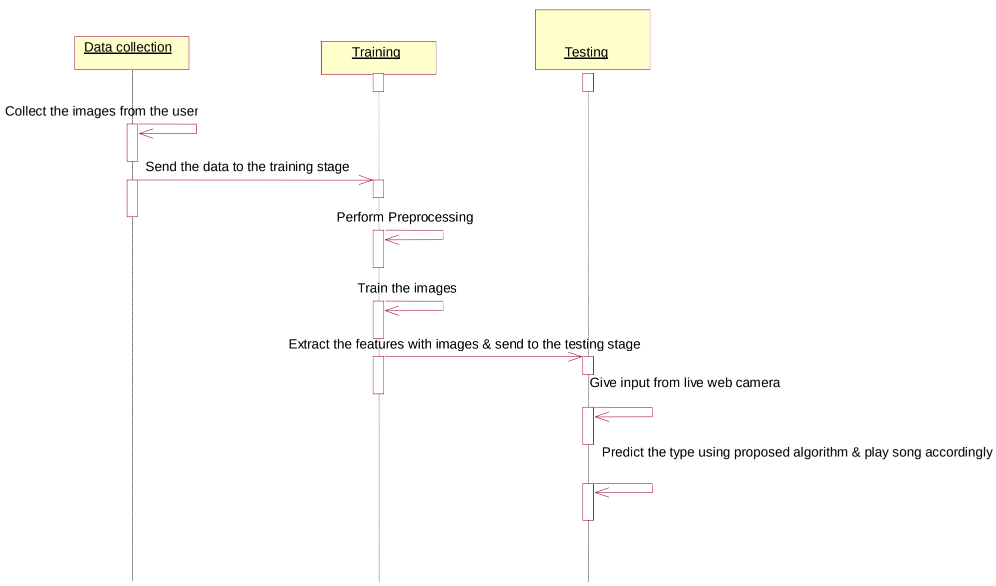
### 4.1 Data Flow Diagram

- The bubble chart is another name for the DFD. It is a straightforward graphical formalism that may be used to depict a system in terms of the data that is fed into it, the different operations that are performed on it, and the data that is produced as a result of those operations.
- The most crucial modelling tool is the data flow diagram (DFD). The system's component models are created using it. These elements include the system's operation, the data it uses, a third party that engages with it, and the way information moves through it.
- DFD demonstrates the information's flow through the system and the various changes that affect it. It is a graphical method for representing information flow and the changes made to data as it travels from input to output.
- DFD is another name for a bubble chart. Any level of abstraction for a system can be represented by a DFD. DFD can be divided into stages that correspond to escalating functional complexity and information flow.




### 4.3 Sequence Diagram

In the Unified Modelling Language (UML), a sequence diagram is a type of interaction diagram that demonstrates how and in what order processes interact with one another. It is a Message Sequence Chart construct. Event diagrams, event situations, and timing diagrams are other names for sequence diagrams.



# 5. DESIGNING ALGORITHMS

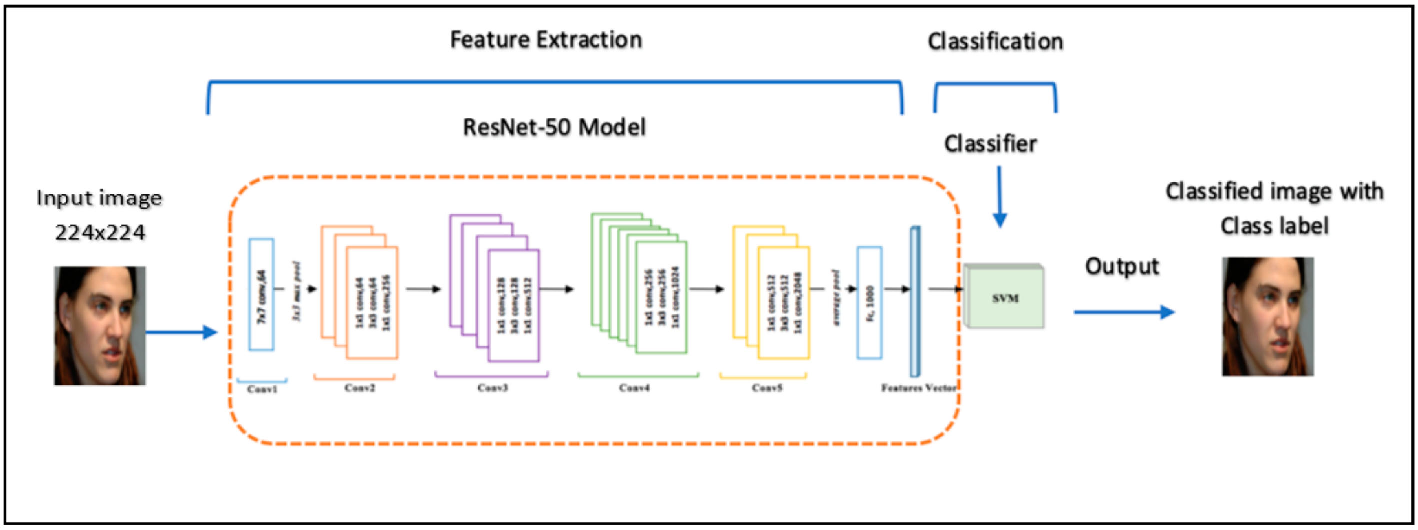
## 5.1. Convolutional Neural Network

The course 4's first module's goals are as follows:

Understanding the convolution process, understanding the pooling operation, and keeping in mind the terms used in convolutional neural networks (such as padding, stride, and filter) are all important.

Creating a convolutional neural network to classify photos into multiple categories.

The architecture of a typical ConvNet is broken down into different phases. The majority of the earliest stages are made up of convolutional layers and pooling layers. A filter bank is a group of weights that connects each unit in a convolutional layer to neighboring patches in the feature maps of the layer above. Feature maps are arranged using convolutional layers. The result of the local weighted sum is then processed using a non-linearity, such as a ReLU. All units of a feature map share the same filter bank.




There are three different sorts of layers in a typical neural network:

- The layer in which we provide input to our model is known as the input layer. The entire number of features in our data (or the number of pixels in the case of a picture) is equal to the number of neurons in this layer.
- Input Layer: The hidden layer receives the input from the input layer. Depending on our model and the volume of the data, there may be numerous hidden levels. The number of neurons in each hidden layer might vary, but they are typically more than the number of features. Each layer's output is calculated by multiplying the output of the layer below it by its learnable weights, adding learnable biases, and then computing the activation function, which makes the network nonlinear.
- Output Layer: After being passed into a logistic function like sigmoid or softmax, the output from the hidden layer is transformed into the probability score for each class.

## 6. DESIGN PROCESS

### 6.1. User Interface Design

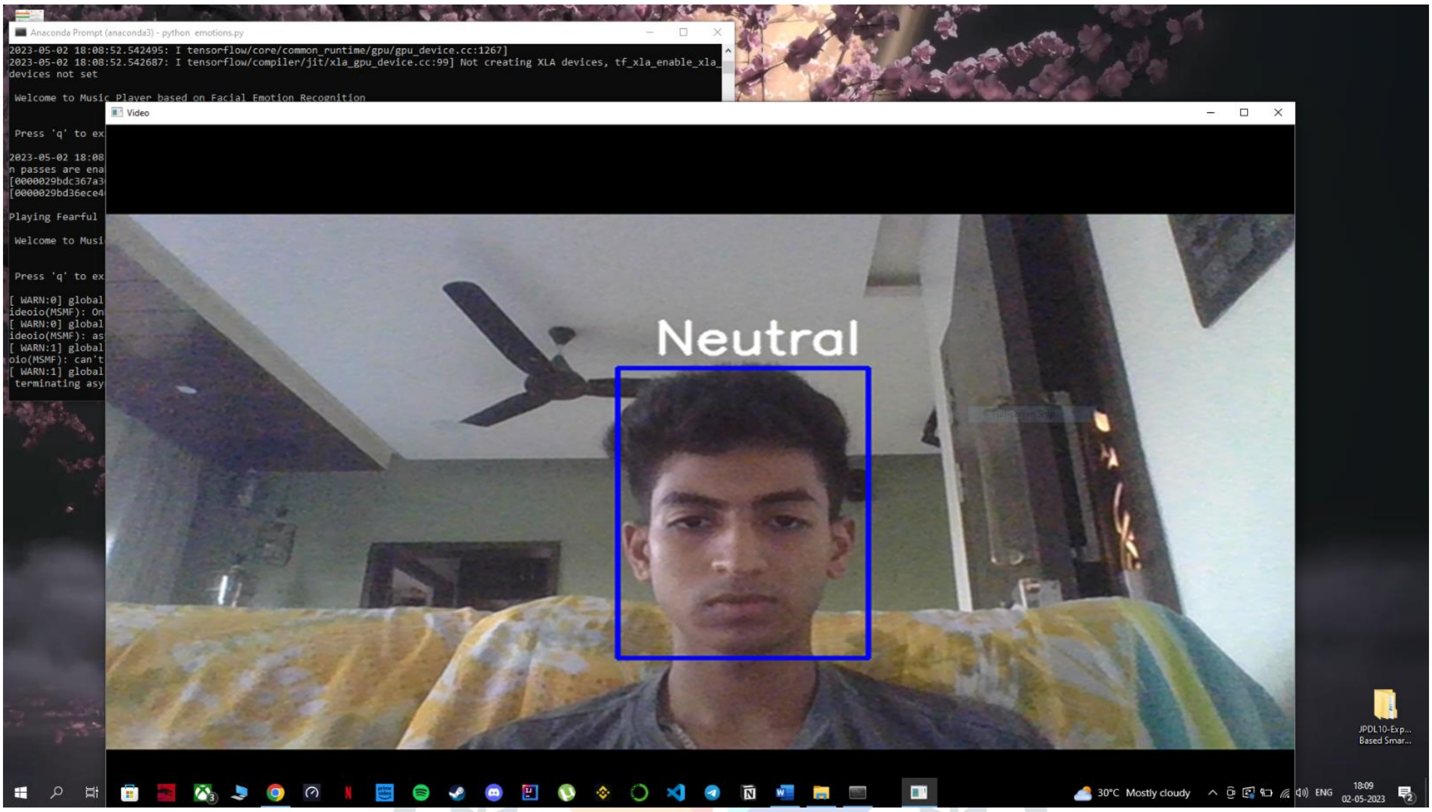



Fig-6.1.1:Emotion detection

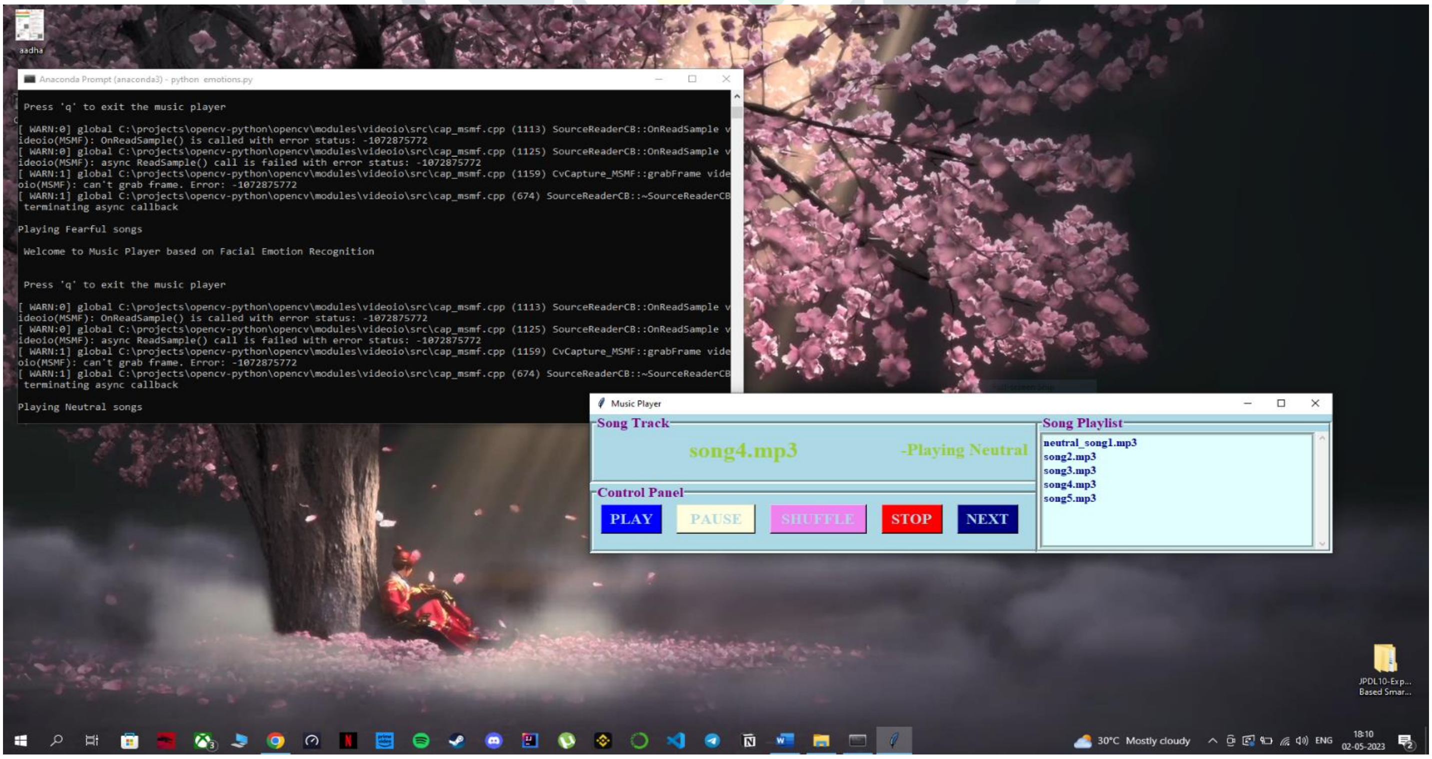



Fig-6.1.2: Playing songs according detected emotion

## 7. CONCLUSION

This program can be added to modern advanced music players as an extra function that recommends songs based on prior song history. The frequency with which the system suggests the song he needs would increase with the addition of a facial expression detecting system in the music player, increasing user happiness. This facial recognition approach can also be applied in a variety of other contexts, such as suggesting films and activities.

## 8. FUTURE SCOPE

We can gradually introduce learning into the application so that it gains knowledge from fresh data produced by the application. Based on user response, the programme will determine whether or not it made a correct prediction. The aforementioned procedure improves model quality and accuracy. To further improve the model's accuracy, we may also include new features like heart rate, which is partly correlated with emotional states in people. In order to anticipate emotions more accurately than the prior method, we can also take the environment into account. For instance, the programme must recognize gym-appropriate objects and motivational music if we are in the gym.

## 9. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to the Rajarshi Shahu College of Engineering at JSPM for providing us with the opportunity to prepare a project on the topic of "Smart Emotion Based Music Player Using Python Image Processing" and to our HOD Dr. N.A. Auti for providing us with the resources and time to investigate the topic. We truly appreciate Prof. Shruti Naik's assistance in guiding us through our research, which would have appeared challenging without her inspiration, consistent support, and insightful suggestions. Furthermore, without the cooperation, advice, and assistance of our friends and family, the complexity of this research paper would not have been possible.

## 10. REFERENCES

### 10.1. Journal Article

[1] N. Veeranjaneyulu, Jyostna Devi Bodapati, " Facial Emotion Recognition Using Deep Cnn Ba sed Features ISSN: 2278-3075, vol. 8, no. 7, May 2019.

[2] Mohammed Yasin, Akshit Tomar, Mrs R Amutha, "Real Time Emotion Based Music Player Using Image Processing", 2022

[3] A Machine Learning Based Music Player by Detecting Emotions ;S. Deebika;K. A. Indira;Jesline 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)

[4] B. Yang, J. Cao, R. Ni and Y. Zhang, "Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images", IEEE Access, vol. 6, pp. 4630- 4640, 2018.

[5] Zirujirao, FengchunQiao, NaimingYao, ZhihaoLi, HuiChen, Hongan Wang, An Ensemble of VGG Networks for Video-Based Facial Expression Recognition, First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia),2018

[6] Shlok Gilda, Husian Zafar, Chintan Soni and Kshitija Waghurdekar, "Smart Music Player Integrating Facial Emotion Recognition and Music Mood Recommendation", Conference: 2017 International Conference on Wireless Communications, Signal Processing and Networking, 2017

[7] A. S. Patwardhan, "Augmenting Supervised Emotion Recognition with Rule-Based decisionModel", arXiv, 2016.

[8] Swathi Swaminathan and Gelenn Schellenberg,"Current emotion research in music psychology",2015

[9] A. S. Patwardhan and G. M. Knapp, "Affect Intensity Estimation Using MultipleModalities,"Florida Artificial Intelligence Research Society Conference, May. 2014.

[10] A. S. Patwardhan and G. M. Knapp, "Multimodal Affect Analysis for Product Feedback Assessment," IIEE Annual Conference, Proceedings, Institute Engineers-Publisher, 2013.

[11] SE. Kahou, C. Pal, X. Bouthillier, P. Froumenty, C. Glehre,R. Memisevic, P. Vincent, A. Courville, Y.Bengio,

RC. Ferrari and M. Mirza. Combining modality specific deep neural networks for emotionrecognition in video. Proceedings of the 15th ACM on International conference on multimodalinteraction, 2013.

[12] F. D. Schönbrodt and J. B. Asendorpf, "The Challenge of Constructing Psychologically BelievableAgents," J. Media Psychology, vol. 23, pp. 100-107, Apr. 2011.