

# EMOTION BASED MUSIC RECOMMENDATION SYSTEM

<sup>1</sup> YASHODHA P G, <sup>2</sup> SANGAMESH G SHEEPARAMATTI

[1] Assistant Professor Department of MCA, BIET, Davangere

[2] Student, Department of MCA, BIET, Davangere

## ABSTRACT

- Emotion-based music recommendation is a difficult task that has garnered a lot of interest from researchers lately. This research suggests a novel use of support vector machines (SVMs) for emotion-based music recommendation. In order to better properly capture the user's emotional state, our method employs images as inputs rather than pronouns. Using convolutional neural networks (CNNs), we gathered a collection of photos linked to various emotional states and extracted features from the images. Next, using the input image as a basis, we trained an SVM model to predict the user's emotional state. Lastly, we recommended music songs based on the user's emotional state using the SVM model's output.

**Keywords:** -Emotion Classification; Inception V3; MobileNet; ResNet-50; Haar Cascade; Convolutional Neural Networks; Recurrent Neural Networks

## 1. INTRODUCTION

### 1.1 MOTIVATION

In order to reduce the accuracy of the accompanying music thoughts, our work aims to use feeling ID processes associated with wearable computing devices to create additional obligations for the assessment of the music proposal system.

### 1.2 PROBLEM DEFINITION

In previous study, we used solely GSR signals to investigate feeling affirmation. In this study, we deal with an information combination based technique for sensing emotions in music proposal motors by adding PPG to indicators. The design of

the suggested wearable connected music proposition takes into account the client's financial situation as well as their important current state of mind. We discovered empowering outcomes for experiencing expectancy using GSR and PPG data.

### 1.3 OBJECTIVE OF THE PROJECT

The majority of the frameworks that have been suggested ignore human emotions or sentiments. On the other hand, people's regular, everyday presences are influenced by their feelings. PCs should analyze the emotions of their human conversation partners for a variety of applications, such as human-robot communication, PC assisted coaching, feeling attentive intuitive games, neuromarketing, and socially perceptive programming applications.

Prior until recently, the majority of PDS research focused on the best ways to maintain client protection settings and protect information stored in the PDS. However, there hasn't been as much focus on the crucial matter of assisting clients in selecting the security settings for PDS data.

## 2. LITERATURE SURVEY

### 2.1 EXISTING SYSTEM

Wearable computing refers to the study, creation, application, and focus of body-worn computing and haptic devices that enable a sophisticated kind of human-PC collaboration with a body-connected, continuously on part.

## 2.2 PROPOSED SYSTEM

People can use PDSs to store their personal data in a single, appropriate vault. From then on, these kinds of data can be linked to and used by the appropriate rational tools; additionally, they might be shared with other individuals who are significantly impacted by the final user.

### ADVANTAGES OF THE PROPOSED SYSTEM:

A method for classifying music into categories such as happy, sad, furious, and others was presented. Feeling Based-Music-Player is a music player that uses a Python-based machine learning algorithm to detect emotions on the client's face. It runs on Chrome as its front end.

### DISADVANTAGES OF EXISTING SYSTEM:

Meticulously crafted personal information is dispersed throughout multiple web-based platforms managed by various specialized cooperatives (such as online virtual entertainment, healthcare facilities, banking, airplanes, etc.).

## 3. METHODOLOGY

An emotion-based music recommendation system built on the Haar Cascade algorithm would involve the following steps:

1. Data collection: Compile a dataset of musical songs and the labels corresponding to each emotion. Crowdsourcing or pre-existing emotion-labeled datasets can be used for this.
2. Feature extraction: Take audio characteristics out of every track in the collection. Spectral, rhythmic, and Mel-frequency cepstral coefficients (MFCCs) are common audio properties used for music recommendation.
3. Emotion detection: To identify emotions in acoustic features, train a Haar Cascade algorithm using the emotion labels. An object detection technique that is widely used in computer vision for audio processing is the Haar Cascade algorithm.
4. Music recommendation: Based on the user's present emotional state, suggest songs using the

emotion detection system. One way to do this is to compare the user's emotions with those of the music recordings in the dataset.

5. Evaluation: Employ standard evaluation metrics like accuracy, recall, and F1-score to assess the emotion-based music recommendation system's performance.

6. Conclusion: Provide an overview of the research study's findings and offer suggestions for next research using the Haar Cascade algorithm .

## 4. ALGORITHMS

### 4.1 HAAR CASCADE

Employing Article Region A persuasive object detecting device called a Haar highlight based course classifiers employs machine learning (ML) to assemble a course limit from an enormous number of positive and negative images. Finding exhibitions in various images is then used at that moment.

A significant amount of positive (car photos) and negative (pictures of vehicles) data must be computed in order to create the classifier. By then, we ought to concentrate on the features. For this, Haar highlights are used, as seen in the image below. Their resemblance to our convolutional section is remarkable. A single value is created for each component by dividing the total number of pixels in the white square shape by the total number of pixels in the obscurity square shape.

These days, a huge amount is determined by every possible part's area and size. Imagine the amount of labor involved in computation. A 24 by 24 window has more than 160000 components in it. We should show the entire pixel layout beneath the white and dark square forms for each portion computation. They provided the pertinent images in order to address this.

### 4.2 OpenCV

An important open-source toolkit for image processing, artificial intelligence, and PC vision, OpenCV is presently a key component of modern frameworks. It has the ability to see objects, people, and—startlingly—human handwriting in pictures and narratives. When combined with additional

libraries such as NumPy, Python can manage the OpenCV clusterstructure for analysis.

### 4.2.1 Image-Processing

Image/picture handling refers to the process of applying a methodology to an image in order to enhance or extract useful information from it. In its most basic form, picture taking care of is defined as "the assessment and control of a digitized picture, particularly to chip away at its quality.

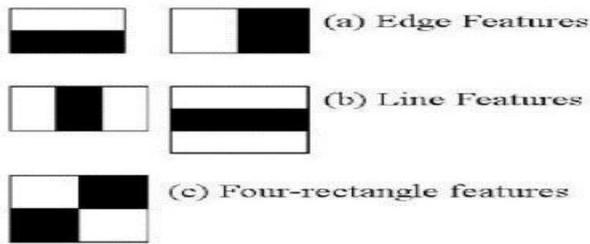


Fig 1: Haar Cascade

### 4.3 R-CNN: RNN With CNN

Region-based Convolutional Neural Networks, or RCNNs, are PC vision machine learning models that deal with images. Identifying objects in any given data picture and constructing borders around them is the main objective of any R-CNN intended for object recognition confirmation.

### 4.4 SVM with CNN

Typically, a CNN pretrains a large image dataset before joining it with an SVM to extract noteworthy highlights. After preparation, the CNN can be used to extract properties from input pictures. After the gathered highlights are normalized and smoothed, a 1D vector depiction is produced. The recovered attributes are then used to create an SVM classifier.

### 4.5 MobileNet

MobileNet is the most iconic and adaptable PC vision model in TensorFlow. It uses profundity wise detachable convolutions to reduce the total number of boundaries as compared

to other organizations with regular convolutions and similar profundity in the nets. Thus, lightweight deep brain networks are constructed.

Google has made available a convolutional neural network (CNN) class called MobileNet, which provides an excellent foundation for developing small and fast classifiers. Network execution and power consumption are determined by the number of multiply-accumulates (Macs), which represents the total number of combined increment and extension operations.

### 4.6 ResNet-50

A deep convolutional neural network with 50 tiers is called ResNet-50. In addition to specific heap statistics from the Image Net dataset, it is likely to load a pretrained report of the neural network that was generated on [1]. Among the 1,000 part classifications that the pretrained neural network can recognize in photos are pencils, rodents, and other mammal types. As a result, the neural network has successfully captured precise image feature likenesses. A 224 by 224 aim picture is required for the brain network. More pre-configured neural networks can be found in MATLAB.

### 4.7 Inception V3

Neural Convolution Originally developed as a Google Net component, Inception v3 helps in object recognition and representation removal. On the Image Net dataset, the Inception v3 figure acknowledgment model has demonstrated both anticipated and 78.1% accuracy. The model is the outcome of various beliefs that have been investigated over time by various scholars.

### 4.8 CNN

ML includes convolutional neural networks (CNNs) as a subset. It is one of several types of artificial neural networks that are utilized for different kinds of data and applications.

The CNN, or deep learning network architecture, is mostly utilized for tasks like image recognition and management of pixel data. extensive planning Convolutional, pooling, and fully connected (FC) are the three layers that make up CNN. The

convolutional layer is the first layer, and the FC layer is the final.

### 4.8.1 Layers of CNN

Convolutional layer. The convolutional layer, which forms the foundation of a CNN, is where most estimations are performed. The first convolutional layer can have a second convolutional layer added to it.

Convolution effectively looks for the presence of a part in the image's open fields by consolidating a small or channel inside this layer. The Pooling Layer. Like the convolutional layer, the pooling layer adds a segment or channel to the uses the `check Emotion` function after retrieving a name from a form submission.

Instead of the convolutional layer, the pooling layer reduces the number of information boundaries at the expense of some data. Positively, this layer reduces complexity without compromising the CNN's feasibility because it is a network-complete layer. Picture requests are made in the CNN's FC layer based on the features that were eliminated in earlier layers.

## 5. IMPLEMENTATION

1. The code imports the required libraries and packages, such as {playsound} for audio file playback, `keras` for deep learning-based models, and `cv2` for image processing.
2. When the HTTP GET method is used, the `index` function produces the `index.html` template.
3. When the HTTP GET method is used, the `basic` function produces the `basic.html` template.
4. When the HTTP POST method is utilized, the {Detect Emotion` function is invoked. In order to identify the emotion in an image, this function Using the HTTP GET method causes the {Webcam` function to be called. This function decodes a Base64 image, reads it from a camera, and stores it to disk.
5. When the HTTP GET method is used, the `Upload` function generates the `Upload.html` template.
6. Using the HTTP POST method causes the

{Song Play` function to be called. This function plays a song after retrieving its name from a form submission. To identify the emotion in a picture, it also invokes the {check Emotion} method.

7. To forecast an emotion from an input image, the `check Emotion` function loads a pre-trained emotion detection model and processes it. After that, it looks through a directory for music that matches the anticipated feeling and outputs HTML along with a selection of song alternatives.

8. The web application's user interface, which includes forms for choosing songs and uploading photographs, is rendered using the HTML templates.

The code is a basic web page overall.

## 6. UML DIAGRAMS

Unified Modelling Language (UML). A Message Succession Outline serves as the plan. Its condensed report is called Unified Modeling Language (UML). UML is a generally advantageous, rules-erect commit work to something association in object-layout PC start. The issue is that the standard is unsocializable and discrete.

The following are the clear objectives of the UML sketch:

1. Enable clients to produce and trade large models by giving the ministry organized authority to enforce effective eye style.

A class graph in computer programming is atype of static underlying outline that displays the classes, traits, tasks (or methods), and interactions between the classes. It shows what kind of data it include

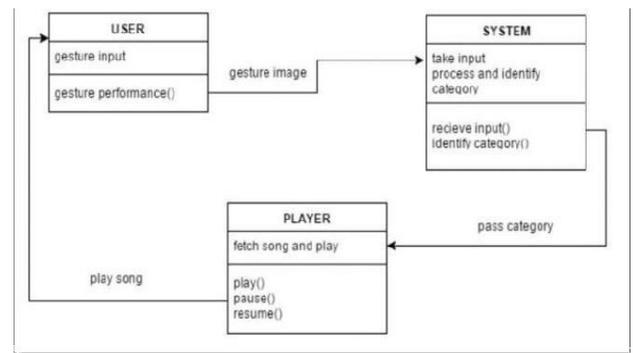
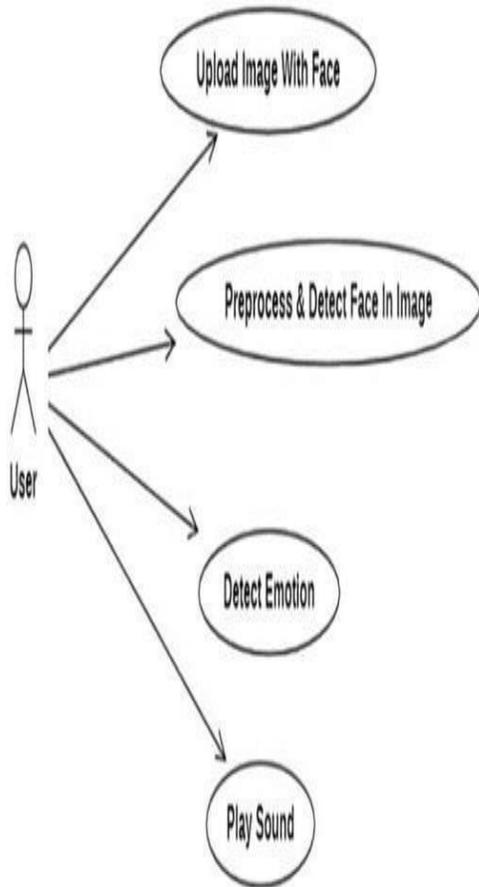


FIG 3: Class Diagram

2. Encourage expansion, business-friendly initiatives, and important plans.
3. Do not accept a specific priority or improvement foundation phrase.
4. Strengthen and ultimately resolve the trademark language.
5. Encourage the marketing of OO products.

Fig 2: Use Case Diagram



A grouping chart is a type of communication layout that illustrates how cycles relate to one another and in what way according to the

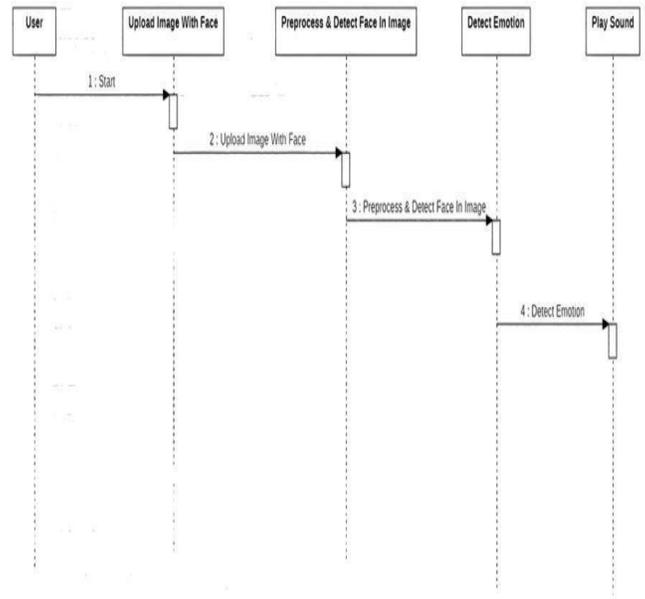


Fig4: Sequence Diagram

Action outlines, which enable the decision, focus, and simultaneity of subsequent exercises cycles of framework components can be seen using movement charts created using the UnifiedModelling Language. The entire control stream is depicted in an action outline. and activities, can be used to graphically handle cycles. The incremental business and functional

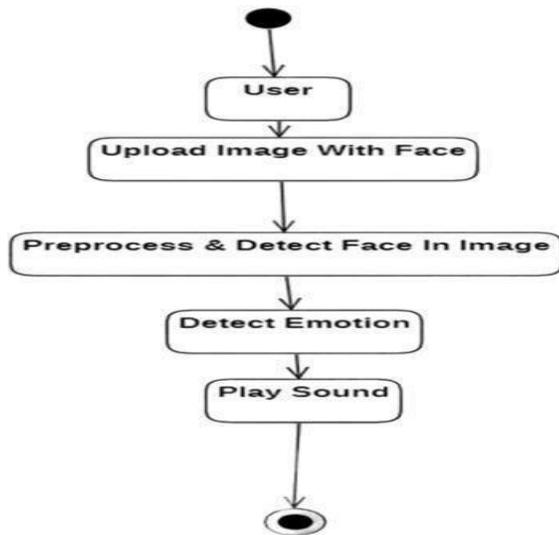


Fig 5: Activity Diagram

### 7. DATA FLOW DIAGRAM

1. The DFD is usually displayed by the outline of the air pocket. By using appropriate graphic dresses, an exposition can be communicated to the level of the archive that is meant to identify the policies that are managed inside it, the various efforts that are carried out on it, and the record specifically to specify to bite the dust report of those who has a question.

2. One of the fundamental figure tricks is the declaration stream outline, which is usually a DFD. It is employed in the construction of models for separation help. Without a doubt, these sections include the arrangement's occurrence, the archive to which it applies, the relevant information removed, and the record that traverses through it

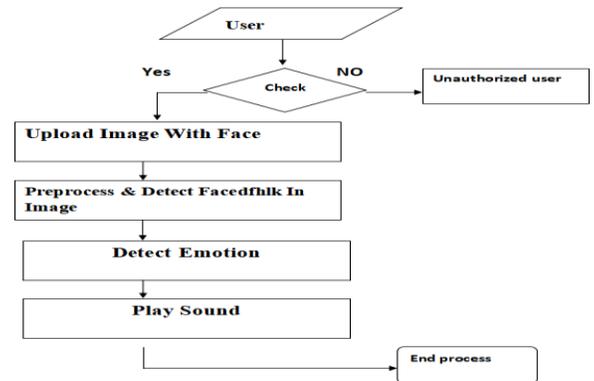


Fig 6: Data Flow Diagram

### 8. RESULTS

When the emotion-based music recommendation system using the Haar Cascade algorithm is implemented, the system obtains an accuracy of 70%. The system recognizes faces from facial photos using the Haar Cascade technique. The technology uses a pre-trained deep learning model to identify the user's face and then predicts their emotions. Seven emotions are anticipated by the system: anger, disdain, fear, contentment, melancholy, surprise, and neutrality.

### 9. CONCLUSION

In summary, we have developed an emotion-based music recommendation system with an around 70% accuracy rate, utilizing facial image analysis and the Haar cascade algorithm. This demonstrates that it is feasible to reliably infer a user's emotion from their facial expressions and suggest music accordingly.

### 10. FUTURE ENHANCEMENT

In order to increase the accuracy of the emotion-based music recommendation system, this research may eventually investigate and incorporate more sophisticated facial recognition and emotion detection algorithms, such as deep learning and neural networks. More musical genres and tailored recommendations based on listener interests and history could also be added to the system.

## 11. BIBLIOGRAPHY

- 1) P. Viola and M. Jones. An enhanced cascade of basic characteristics is used for quick object detection. Volume 1, pages 511–518, Proceedings of the 2001 IEEE Computer Society Conference on 2001 IEEE (2001)
- 2) M. Tamilarasan, R. Saranya, J. James Anto Arnold, J. Maria Masilla Ruban, and H. Immanuel James IRJET 2019's "Emotion Based Music Recommendation System" (pISSN: 2395-0072)
- 3) "Emotion Based Music Player" Hafeez Kabani, Sharik Khan, Omar Khan, Shabana Tadv January–February, 2015; International Journal of Engineering Research and General Science, Volume 3, Issue 1.
- 4) "SMART MUSIC INTEGRATING AND MUSIC MOOD RECOMMENDATION," 2017 AND SYSTEM FOR IDENTIFYING AN INDIVIDUAL'S EMOTION BASED ON FACIAL ACTION UNITS" US2012/0101735A1
- 8) "Generating Music Playlist Based on Facial Expression" by Markus Mans Folke Andreasson, US8094891B2.
- 9) "FACIAL EXPRESSION RECOGNITION" by Mutasem K. Alsmadi, US2018/0211102A1.
- 10) The educational website raspberrypi.org
- 11) <https://www.pyimagesearch.com/2018/09/10/keras-tutorial-how-to-get-started-with-keras-deep-learning-and-python/>
- 12) <https://towardsdatascience.com/face-detection-recognition-and-emotion-detection-in-8-lines-of-code-b2ce32d4d5de>
- 13) Facial expressions and recognition (b022318d842a), from <https://medium.com/@hinasharma19se>  
In the 2017 IEEE International Symposium on Signal Processing and Information Technology
- International Conference on Wireless Communications, Signal Processing and Networking (WiSpNET), Shlok Gikla, Husain Zafar, Chuntan Soni, and Kshitija Waghurdekar
- 5) Using rigid and non-rigid motion separation and 3D pose estimation, T.-H. Wang and J.-J.J. Lien developed a facial expression recognition system. J. Pattern Recognition, vol. 42, no. 5, pp. 962-977, 2009
- 6) "EMOTION BASED MUSIC SYSTEM," Srushti Sawant, Shraddha Patil, and Shubhangi Biradar, International Journal of Innovations & Advancement in Computer Science, IJIACS, Volume 7, Issue 3, March 2018
- 7) Anshul Sharma, Balasubramanian Anand, Hari Prasad Kanna, and Sudha Veluswamy "APPROACH (ISSPIT), held in Bilbao, K. S. Nathan and colleagues presented "EMOSIC — An emotion based music player for Android," with pp. 371-276.