BioGecko

Vol 12 Issue 03 2023

ISSN NO: 2230-5807

Emotion Detection using Facial Features and Recommending Music to Enhance the Mood

Mrs. V. Himabindu[1], Dr. V. Anantha Krishna[2], Salikuti Nikitha[3], Rahela Shoheb[4], Baikadi Naveena[5]

[1]Assistant Professor, Computer Science and Engineering, Sridevi Women's Engineering College Hyderabad, India.
[2]Professor, Computer Science and Engineering, Sridevi Women's Engineering College Hyderabad, India.
[3,4,5]Computer Science and Engineering, Sridevi Women's Engineering College, BTech IV Year Hyderabad, India.

Abstract
Music is a form of art, which is known for having a connection with the listener's emotions. It has the power of lifting up one's mood. During our fast-paced life style, everyone goes through a change in moods within seconds. People listen to music according to their emotions. Therefore, our project aims on building a music recommendation system based on facial features. This is done through Conventional Neural Network algorithm. Facial expressions are captured using an in-built camera. An input face image is taken. Feature Extraction is done to detect various emotions such as happy, sad, angry, and neutral. A music playlist is generated automatically by identifying the real-time mood of the user. The proposed system would turn out to be more efficient than the existing systems. It will be both time and cost efficient.

Keywords: Convolutional Neural network (CNN), Feature Extraction, Facial Expression, Emotion Detection, Image Processing, Music Playlist.

1. INTRODUCTION

The classification of human emotions can be done as: happy, sad, angry, neutral, surprise, etc. Other emotions such as cheerful (a variant of happy) and contempt (a variant of disgust) can be categorized under this group of emotions.

Numerous music players have been created with capabilities like skip, loop, local and streaming playbacks due to rapid improvements in multimedia and technology. Although they serve the user's fundamental needs, the user still has to browse through the music playlist or manually search for the song according to his present state of mind. With the evolution of digital music technology, the evolution of a customized music recommendation system which recommends music is needed for users.

Emotion Recognition is a technique that allows analyzing the emotions on an input face of the user using advanced facial emotion processing. Recognition facial emotions and fetching the music that enhance the mood of the user based on the emotion. Deep learning algorithms proved to be very handy for pattern recognition and classification. Thus, they can be used for mood detection too.
Therefore, our project fixates on building an efficient music recommendation system using facial recognition techniques. Thus, the purpose of our proposed system is to recognize facial emotions of the user and recommend songs accordingly.

Existing system:
The existing system involves playing music manually. The two major existing approaches for customized music recommendation are: collaborative filtering and content-based filtering. Collaborative filtering is a technique which assumes that the users who have similar choices in the

past are likely to have the same choices in the future too and recommends music based on it. Content-based filtering is a technique that recommends music based on the characteristics the users have interacted with in the past like loudness, tempo, genre, etc.). When a user listens to a song for more than 30 seconds, the existing technique uses that information and it recommends a similar type of song to the user. Existing systems recommend a song based on the user's previous feedback on that particular song, downloading the songs, when a user gives a like to that song, song skipping behavior, etc.

Disadvantages Of Existing System:

1. The existing approaches involve playing music manually.
2. Songs are played irrespective of the user's mood.
3. The existing systems cannot recommend music which is dissimilar to the music a user has interacted with in the past.
4. Maximum music apps use collaborative filtering and content-based filtering.
5. The popularity bias problem: popular songs (which have high ratings, likes, etc.) get a lot of exposure and are recommended more compared to the less popular ones.

Problem statement:

To develop an efficient music recommendation system based on the real-time mood of the user through a web camera using deep learning algorithms.

Proposed System:

The motive of the system is to capture the facial expressions with the web camera and feed those images into the Convolutional Neural Network. An emotion is derived from the captured image and the song playlist is generated.

The main purpose of the project is to provide a music playlist to users automatically to change their mood, which can be happy, sad, neutral, or surprised. Our proposed system detects the emotion of the user and if it is detected to be sad, then a song that will enhance the mood of the person positively will be played automatically.

1. Real-Time Capture: the system will capture the facial expressions of the user correctly using webcam in laptop or mobile.
2. Face Recognition: The user's facial expression is taken as an input. CNN is programmed to evaluate the facial features of the input image.
3. Emotion Detection: Features are taken from the input image to recognize the emotion of the user and depending on the user's emotions, the system will generate the song playlist.
4. Music Recommendation: The recommendation module will suggest the songs to users by mapping their emotions to the type of songs.

Advantages Of Proposed System:

1. The proposed system suggests music by extracting different facial emotions of the user.
2. It doesn't require the recommendation techniques of the existing system.
3. Our facial expression recognition system can identify up to four different types of emotions in real-time and does so anonymously, to ensure the privacy of the users at all times.

System Design Architecture:

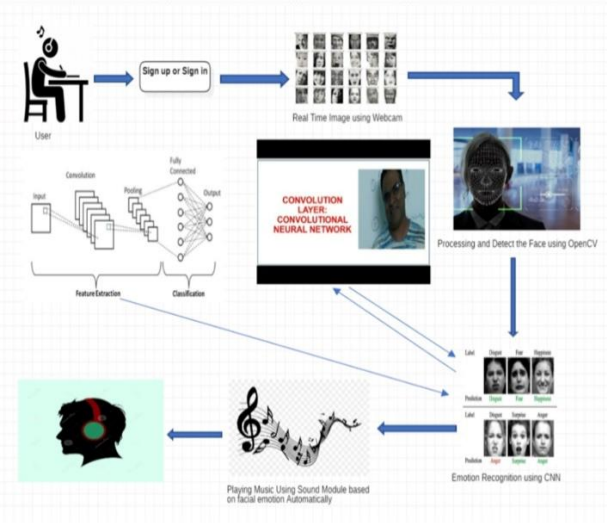
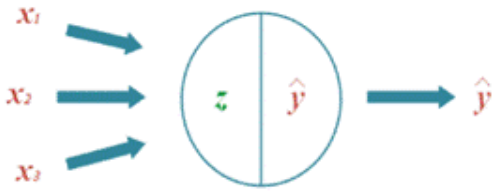


Figure: Architectural view of Emotion Detection for Music Recommendation

Datasets:
The system is using 4 datasets to detect the individual facial emotions as;
Happy, Sad, Angry, Neutral.

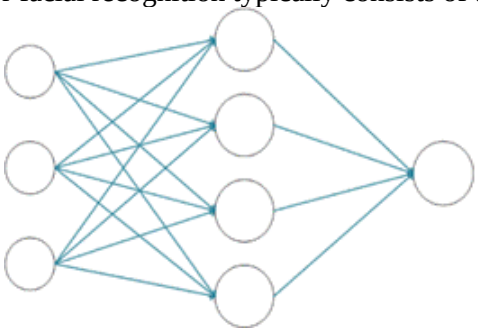
Algorithm:



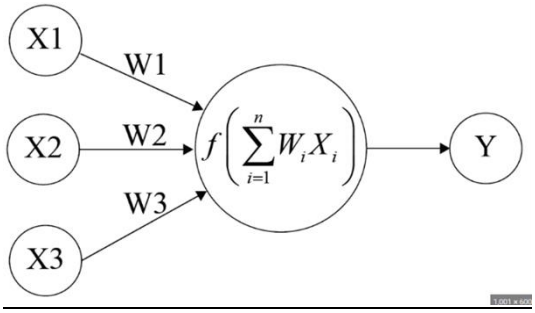
$$z = w_1x_1 + w_2x_2 + w_3x_3$$
$$\hat{y} = g(z)$$

CNN Algorithm: The Convolutional Neural Network (CNN) is a popular algorithm used for facial recognition. It is a type of Deep Learning algorithm that is capable of recognizing patterns in images and can be trained to detect and identify special facial features.

The CNN algorithm for facial recognition typically consists of the following steps:



1. **Data Processing:** The first step is to preprocess the input images to ensure that they are of the same size and have the same orientation. This is important because CNNs require consistent inputs to learn the features.
2. **Convolutional Layers:** The convolutional layers are the core of the CNN algorithm. They consist of a set of filters that convolve over the input image, extracting features at different scales and orientations. Feature maps are generated as the output, which capture the presence of specific features in the input image.



3. **Pooling Layers:** The pooling layers down sample the feature maps to reduce their dimensionality and extract the high priority features. Max pooling is an ultimate pooling operation, that selects the maximum value in each window.
4. **Fully Connected Layers:** The output from the convolutional and pooling layers is taken. It is used for the final classification.
5. **Training:** The CNN algorithm is trained using an abundant dataset of labeled facial images. The training process involves optimizing the weights of the filters to minimize the difference between the predicted and real labels.
6. **Testing:** Once CNN has been programmed, it can be used to classify new facial images. The input image is imported through the neural network, and the output probabilities for every class are evaluated. The final prediction is the class with the soaring probability.

Modules:
Modules in our project are designed to analyze facial expressions in real-time, detect emotions, and suggest appropriate music to enhance the user's mood. The modules used to run this project are as follows:

1. **Real-time face capture:** This module captures the real-time video feed from the user's in-built camera and processes it to detect facial features.
2. **Feature Extraction:** This module detects the facial features such as eyes, eyebrows, nose, mouth and cheeks and preprocesses it using OpenCV.
3. **Emotion Detection:** This module uses the CNN algorithm to analyze the facial expressions and detect the user's emotion, such as happiness, sadness, anger or surprise.
4. **Music Recommendation:** Based on the detected emotion, the module recommends music to enhance the user's mood.
5. **User Interface:** This module provides a user-friendly interface for users to interact with the system, including options to start the real-time face capture process.

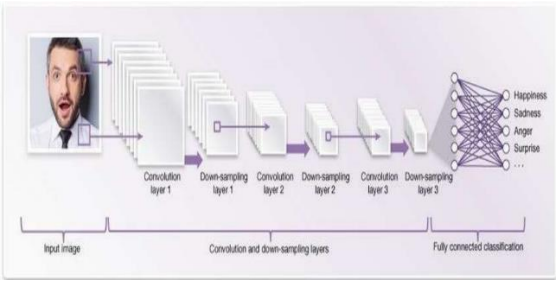
Overall, these modules combine Computer Vision and Machine Learning techniques to provide real-time emotion detection and music recommendation to enhance the user's mood based on their facial expressions.

Module description:
Real-Time Face Capture: The real-time face capture process in the emotion detection module typically involves the following steps:

User capture Image

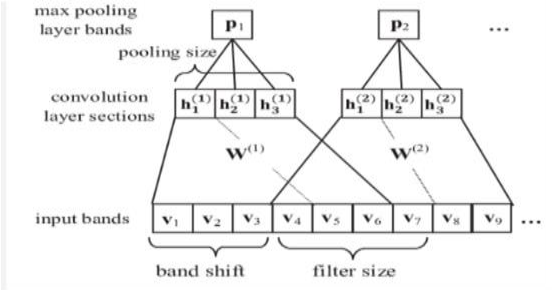
1. **Accessing the camera:** This module accesses the camera on the user's device, such as a webcam.
2. **Capturing frames:** This module captures a series of frames, typically at the rate of 15-30 frames per second, from the camera feed. Each frame is a still image that represents a moment in time.

Feature Extraction:
Facial feature detection is an important component of real-time face capture for emotion detection. The process of detecting facial features typically involves the following steps:



1. **Image Preprocessing:** This module preprocesses the captured frame to enhance the contrast and remove noise from the image, which can improve the accuracy of feature detection.
2. **Face Detection:** This module uses OpenCV to detect the face in the image. This typically involves analyzing the color and texture of pixels in the image and using Machine Learning algorithms to identify facial features.
3. **Landmark Detection:** Once the face is detected, the module identifies specific points on the face, called landmarks, which correspond to different facial features. Landmark detection is usually performed using Machine Learning models, such as Convolutional Neural Networks (CNNs), which have been trained on large datasets of facial images.
4. **Feature Extraction:** After landmarks have been detected, the module extracts special features, like the position of the eyes, nose, mouth, and eyebrows. These features are then used to analyze the user's facial expression and detect their emotion.

Emotion Detection:
Emotion detection from facial features is typically achieved using Machine Learning algorithms that analyze extracted facial features and classify them into different emotional states. This is done using Deep Learning algorithms, such as Convolutional Neural Networks (CNNs), which can learn to detect and classify emotions directly from the raw input data.



These algorithms are trained on large datasets of labeled images or video frames with known emotional states, and learn to identify patterns and features in the data that are associated with different emotions.

The input data consists of facial features extracted from the user's face, such as the position and shape of the mouth, eyebrows, and eyes. The algorithm then uses these features to predict the emotional state of the user, such as happiness, sadness, anger, or fear.

The accuracy of emotion detection depends on various factors like the quality of input data, the particular algorithm and technology employed, and the diversity and size of the training dataset used to train the algorithm.

Music Recommendation:

The emotion-music module is a software component that uses the output of an emotion detection module to recommend music that is appropriate for the user's emotional state. The module typically takes as input the emotional state detected by the emotion detection module and uses this information to opt music that matches the user's mood.

The specific approach in this module is using Machine Learning algorithms to analyze the emotional content of music and match it to the user's emotional state. These algorithms may use various features of music, such as the tempo, rhythm, and pitch, to predict the emotional content of the music and recommend songs that match the user's mood.

The output of the emotion-music module may be exhibited to the user through a Graphical User Interface that displays the recommended music.

1. When an Angry face is detected

2. When a Sad face is detected

3. When a Neutral face is detected

4. When a Happy face is detected

Overall, the emotion-music module provides a personalized music experience for the user that is tailored to their emotional state. The module can help users to manage their emotions and enhance their mood through the power of music.

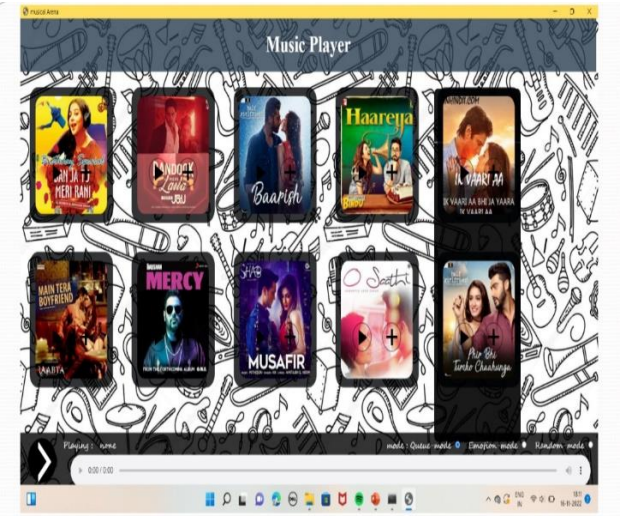
User Interface:
The user interface module is a software component that provides an interface for users to interact with the system or application.

The specific design and functionality of the user interface module may differ based on the application and the target audience. Some applications may have a simple and intuitive interface, while others may have a more complex interface that requires specialized training or knowledge to use.

The user interface module may interact with other modules in the system, such as the emotion detection module, or the music recommendation module, to display output or receive input from the user. For example, the module may display the user's emotional state detected by the emotion detection module.

Overall, the user interface module plays a crucial role in providing users with a seamless and intuitive experience when using a system or application. The module can help to improve user satisfaction and engagement, and can make the system more accessible and usable for a wider range of users.

Implementation:
The application is displayed on the screen.



Testing:
Testing is a process in which point each reasonable mistake or proneness in a work product is found by conducting a sequence of various tests. It creates an opportunity to check the functionality of a finished product. It is the procedure in which a software is exercised to guarantee that the software system meets allure necessities and user requirements and does not forsake in a repugnant manner.
There are various types of testing strategies. Few of them are:

Unit Testing: Unit Testing is a technique in which individual components (units) of a software system are separated from the rest and tested to verify that each unit's performance meets its requirements.

Integration Testing: Integration testing is a testing methodology that verifies whether the interactions between different components or modules of a system are working as expected or not.
Functional Testing: Functional test provides a demonstration that functions tested are available as specified by the system documentation, user requirements and, business and technical requirements.
System Testing: System testing is a technique in which a quality assurance team evaluates the interaction of different units of an application in a fully-integrated system or application.

2. RESULTS

Evaluation metric			Result
Accuracy of emotion detection			**85%**
User satisfaction with recommended music			4/5
Effectiveness of recommended music in enhancing the mood			73%
Emotion	Correct recommendations (%)	Incorrect recommendations (%)	Failure conditions
Happy	80%	20%	Wearing masks
Sad	75%	25%	Different skin colours
Angry	70%	30%	Limited facial expressions
Neutral	85%	15%	Multiple people in the frame

Once the system captures the facial Expression of an individual it will play the music based on the facial emotion automatically without any external action. The accuracy of our application is 80%.

3. CONCLUSION

The passion acknowledgment of the facial features fed into the trained model is an ultimate facet concerning this project. The main aim search out uses the facial emotion detecting feature. The suggested approach aims to advance the music recommendation system by using facial recognition techniques to generate a music playlist. The proposed system can detect various emotions: Neutral, Happy, Sad, Surprise, Fear, Disgust. For example, when an angry emotion is detected, the sounds that are pleasant, harmonized will be played to enhance the mood of the user.

Future scope:
For the future enhancement, we can extend to apply GAN, SVM, Resnet. We can also add as many songs as we can by using the song server.

4. REFERENCES

1. Convolutional Neural Networks (CNN) With Tensor Flow by Sourav from Edureka [https://www.youtube.com/watch?v=umGJ30-15_A]
2. Deep Learning Simplified by Sourav from Edureka [https://www.youtube.com/watch?v=dafuAz_CV7Q&list=PL9ooVrP1hQOEX8BKD plfG86ky8s7Oxbzg]
3. P. Melville and V. Sindhwani, “Recommender systems,” in Encyc. of mach. learn. Springer, 2011, pp. 829–838.
4. N. Sebe, I. Cohen, T. S. Huang et al., “Multimodal emotion recognition,” Handbook of Pattern Recognition and Computer Vision, vol. 4, pp. 387– 419, 2005.

5. R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 10, pp. 1175–1191, 2001.
6. D. Ayata, Y. Yaslan, and M. Kamasak, "Emotion recognition via galvanic skin response: Comparison of machine learning algorithms and feature extraction methods," IU J. of Elect. & Elect. Eng., vol. 17, no. 1, pp. 3129–3136, 2017.
7. P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions." Am. Assoc. for Adv. of Sci., 1983.
8. I.-h. Shin, J. Cha, G. W. Cheon, C. Lee, S. Y. Lee, H.-J. Yoon, and H. C. Kim, "Automatic stress-relieving music recommendation system based on photoplethysmography-derived heart rate variability analysis," in IEEE Int. Conf. on Eng. in Med. and Bio. Soc. IEEE, 2014, pp. 6402–6405.
9. S. Nirjon, R. F. Dickerson, Q. Li, P. Asare, J. A. Stankovic, D. Hong, B. Zhang, X. Jiang, G. Shen, and F. Zhao, "Musicalheart: A hearty way of listening to music," in Proc. of ACM Conf. on Emb. Netw. Sens. Sys. ACM, 2012, pp. 43–56.
10. H. Liu, J. Hu, and M. Rauterberg, "Music playlist recommendation based on user heartbeat and music preference," in Int. Conf. on Comp. Tech. and Dev., vol. 1. IEEE, 2009, pp. 545–549.