iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60217

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Music Recommendation System by Detection of Emotions

Prof. Y. A. Dhumale[1], Siddhesh Sambe[2], Om Gavhane[3], Mohit Choudhary[4], Aditya Jagtap[5]
[1, 2, 3, 4, 5] *Dept. Of Computer Science & Engineering, Prof. Ram Meghe Institute of Technology & Research, Amravati, Maharashtra, India*

Abstract: ***Expressing emotions through facial expressions is a fundamental aspect of human communication, offering insight into an individual's mood and well-being. The subtle nuances of the eyes, cheeks, forehead, and even the curve of a smile convey a wealth of emotional information. Music, as an art form, holds a unique power to relax and comfort both body and mind. Our project capitalizes on the synergy between these two elements by developing a system capable of discerning emotions from facial expressions and selecting appropriate music in response. This innovative approach not only aids in calming individuals and improving their mood but also streamlines the music selection process, saving valuable time. Furthermore, we aim to create software that offers this functionality universally, enabling users to access emotional music responses anytime, anywhere.***
Keywords: Recommendation System, Emotion, Predictive Model, Facial Expression.

I. INTRODUCTION

Music is a major source of entertainment for listeners and music lovers, it plays a crucial part in improving people's lives. Many music players with features like fast forward, reverse, variable playback speed, genre classification, streaming playback with multicast streams and including volume modulation, etc., have been developed in the modern world due to the increasing advancements in multimedia and technology. These capabilities may meet the user's fundamental needs, but it will be up to the user to actively browse the playlist and select songs that fit their current behaviour and mood. An innovative method that assists the user in automatically selecting music depending on their feelings is the emotion-based music player.

It interprets the user's facial expressions to determine the appropriate tune to play. The EMO algorithm, a machine learning technique, is used to identify the emotions. A vital organ of the human body, the face is particularly useful for interpreting a person's emotions and behaviours. The user's image is captured via the webcam. From the acquired image, it then extracts the user's facial traits. Two categories of facial expression exist: smiling and not smiling. The primary idea behind this project is to play music automatically depending on the user's emotions. With regard to the emotions identified, it seeks to deliver the user's preferred music. The current approach requires the user to actively choose the songs; songs that are played at random might not fit the user's mood. The user must first categorize the songs into different emotions, and then they must manually choose a certain feeling before the song can be played. The music will be played from the predefined folders based on the emotion.

There are music in each subdirectory that fit the mood. The programmer has the ability to alter, replace, or remove songs within the subfolders based on the user's specifications. Users may occasionally enjoy different types of music depending on their mood. When a user's emotion is identified as sadness, for instance, the user has complete control over the mood they wish to be in. In this case, there are two possible outcomes: a) The user wishes to keep feeling depressed. c) The user wishes to feel happier and to improve their mood.

II. LITERATURE REVIEW

In their article, Yusuf Yaslan et al. suggested an emotion-based music recommendation system that determines the user's mood based on data from wearable computing devices combined with physiological sensors such as photoplethysmography (PPG) and galvanic skin response (GSR) [3].

Ayush Guidel et al. claimed in [7] that facial expressions are a simple way to read a person's mental state and present emotional mood. Basic emotions (happy, sad, angry, excited, surprised, disgusted, fearful, and neutral) were taken into consideration when developing this system. In this research, a convolutional neural network was used to accomplish face detection.

The intelligent music player that uses emotion recognition was described in the study that was proposed by Ramya Ramanathan et al. [1]. A fundamental aspect of human nature is emotion.

They have the most significant role in life. Human emotions are designed to facilitate comprehension and sentiment sharing. The album's emotional content is the first factor used to categorize the user's local music selection. This is frequently computed with the lyrics of the song in mind.

As a labor-intensive and time-consuming task, CH Radhika et al. [8] recommended manual playlist segregation and song annotation based on the user's current emotional state. Many algorithms had been suggested to automate this process. But current algorithms are slow, require more hardware (such as sensors and EEG structures), add to the system's total cost, and have far worse accuracy.

III. PROPOSED METHODOLOGY

In this project we proposed facial expression recognition using CNN algorithm and we are implementing the model in music recommender system. Working of Music based recommender system is as given below. As shown in following figure, any user will be able to register in our system. After registration he/she will be able to log in into the system and open camera, and get songs based on emotions . User can set their preferences by their own. Our proposed model will open camera, and customer will submit their emotions. The model will capture face image, pre-process the image, extract features and classify into the matching emotion category as per the features of the image. As we are using CNN algorithm, there is no need to use any other algorithm to extract image features. CNN extract features and then classify the image into appropriate emotion category.

The extracted emotion will decide the preference of the user. Emotion will be sent to music recommendation model. The model will find out user preferences based on profile and emotion and recommend songs accordingly.

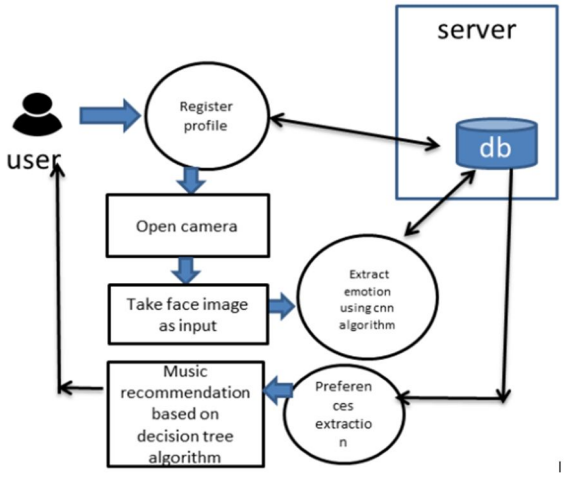


Fig.1 Admin Panel

IV. ALGORITHMS USED

CNN:

Convolutional Neural Networks (CNNs or ConvNETs) are Deep Learning algorithms that process images, assign importance to objects in the image using learnable weights and biases, and can differentiate images from each other.

A convolutional neural network is essentially a neural network that uses a convolution layer and pooling layer. The convolutional layer convolves into a smaller area to extract features, while the pooling layer picks the data with the highest value within an area.

They require less pre-processing in comparison to other classification algorithms and are able to learn filters and characteristics.

A Music Recommendation System using a Convolutional Neural Network (CNN) focuses on identifying emotions in music by analysing audio features. The process involves collecting a dataset of music tracks with emotional labels such as happy, sad, angry, or relaxed. The audio tracks are then pre-processed using techniques like Mel-frequency cepstral coefficients (MFCCs) to extract emotional content.

The CNN architecture is specially made to analyse the audio features taken from music tracks. Unlike regular CNNs that work on images, audio CNNs use one-dimensional convolutional layers to capture patterns in the audio features for emotion detection.

During the training process, the Convolutional Neural Network (CNN) is taught with the pre-processed audio data and emotion labels. The network is trained to associate input audio features with specific emotional categories or dimensions. To achieve this, the network's parameters (weights and biases) are adjusted using methods such as backpropagation and gradient descent.

Detecting Emotions: After training the CNN, we can use it to identify emotions in unfamiliar music tracks. The characteristics of the input track are inputted into the trained network, which then predicts the emotional content of the music. This forecast can be in the form of specific labels (such as happy or sad) or continuous values that represent emotional aspects (like excitability or sentiment).
Integrating Emotional Detection into Music Recommendations: One possibility is to incorporate the results generated by an emotion detection model into a music recommendation system. In this setup, the detected emotions would serve as supplementary criteria, in addition to factors such as user preferences, listening habits, and contextual details. By factoring in the user's emotional state or preferences, the recommendation system can provide suggestions for music tracks that are well-suited to the user's mood.
Assessing and Tweaking: We can assess the effectiveness of the emotion detection model and recommendation system by looking at metrics like accuracy, precision, recall, and user satisfaction. Once we have this information, we can make adjustments and enhancements to the model and system in an ongoing process to improve their performance.
In general, using a CNN-based method for a Music Recommendation System that detects emotions provides a data-driven approach to tailor music suggestions according to users' emotional conditions, which could improve the user's interaction and satisfaction with the system.

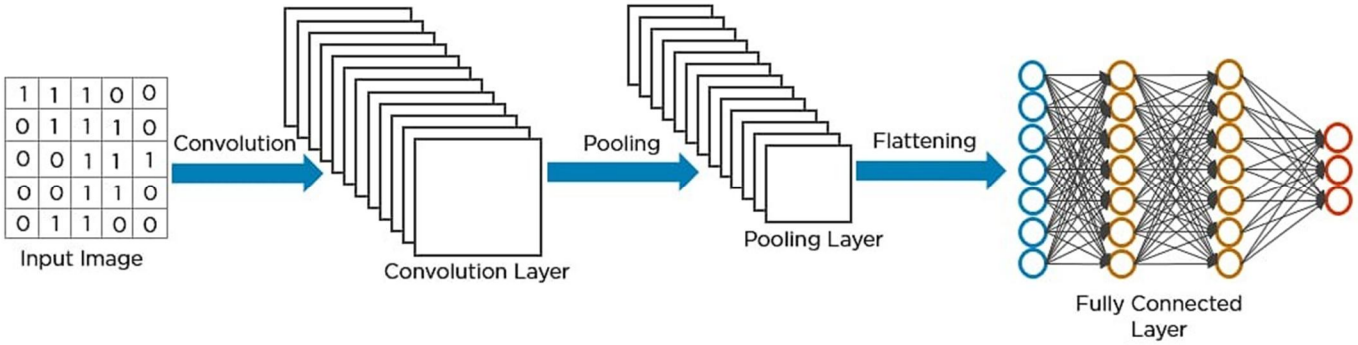


Fig.2 Convolutional Neural Network

A. Facial expression classification algorithm based on CNN

The high performance of computers has made deep learning popular across a wide range of industries. In the paper, the CNN k-based deep neural network is used to tackle the expression categorization problem. Higher dimensional expression features are output to the lower convolution layer after each convolution layer's upper expression features have been twisted by a learnable convolution kernel. The following is the expression feature graph calculating formula.

$$G_i = f(G_i * W_i + B_i)$$

In formula (1), Gi is the output of i-th layer neuron, Gi-1 is the input of i-1-th layer I neuron; Wi is the weight vector of convolution core of i-th neuron, Bi is the offset vector, and f is the activation function. To some extent, the pool layer can keep the feature scale and reduce the feature map dimension. The pooling formula is:

$$G_i = Pooling(G_{i-1})$$

where pooling is the pooling function and GI is the lowest sampling layer. The activation function is a crucial component of a neural network that retains and maps data while also adding nonlinear variables to better mimic the structure of a human brain network. It is also known as a multi-layer perceptron because the whole connection layer, which links all of the neurons in the layer to achieve the goal of synthesizing the retrieved features, uses the following calculation method.

$$F(x) = f(x * W + B)$$

Where, F(x) is the fully connected layer, f usually called the activation function, W usually called the weight vector, and B can be called the offset vector. SoftMax function is usually used to solve multi classification problems.

In this paper, SoftMax function is used to map the output of multiple neurons to a value from 0 to 1, indicating that the probability of expression image x belonging to category j is

$$P\{y^{(i)} = j \mid x^{(i)}, \theta\} = \frac{e^{\theta_j^T x^{(i)}}}{\sum_{i=1}^{k} e^{\theta_j^T x^{(i)}}}$$

In formula (3), $P\{y^{(i)} = j \mid x^{(i)}, \theta\}$ is the probability of expression image x corresponding to each label category, and θ is defined as parameter to be fitted.

V. RESULT ANALYSIS

Facial emotion-based music recommendation algorithms present an interesting avenue for more individualized listening experiences. Through the analysis of your facial expressions, the system is able to customize music recommendations to match your present state of happiness and mental health. Imagine being depressed and getting a playlist of upbeat songs, or frustrated and getting suggestions for relaxing music. But there are still difficulties. Music and emotions are subjective, and facial recognition technology isn't perfect. To improve emotion recognition precision, take into account the complexity of human emotions, and protect user privacy, more research is required. Notwithstanding these challenges, face emotion-based music suggestion shows potential for the development of personalized music exploration in the future.

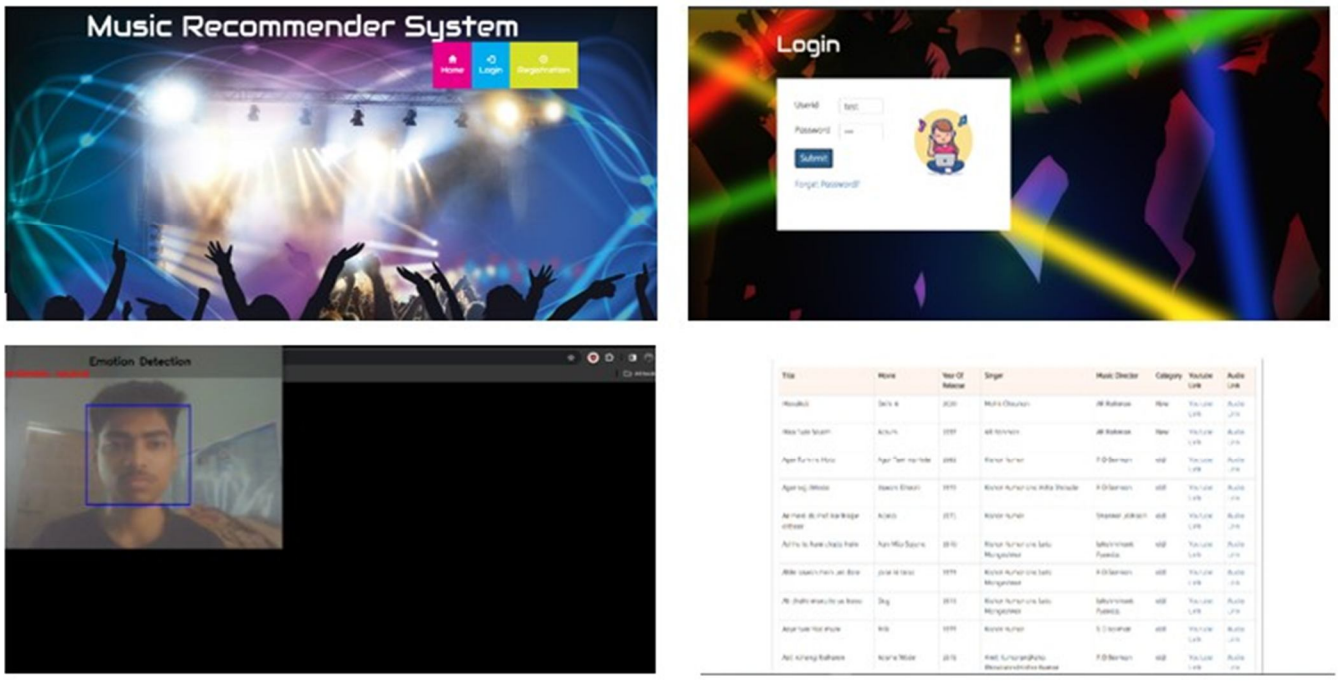


Fig 3 : Result

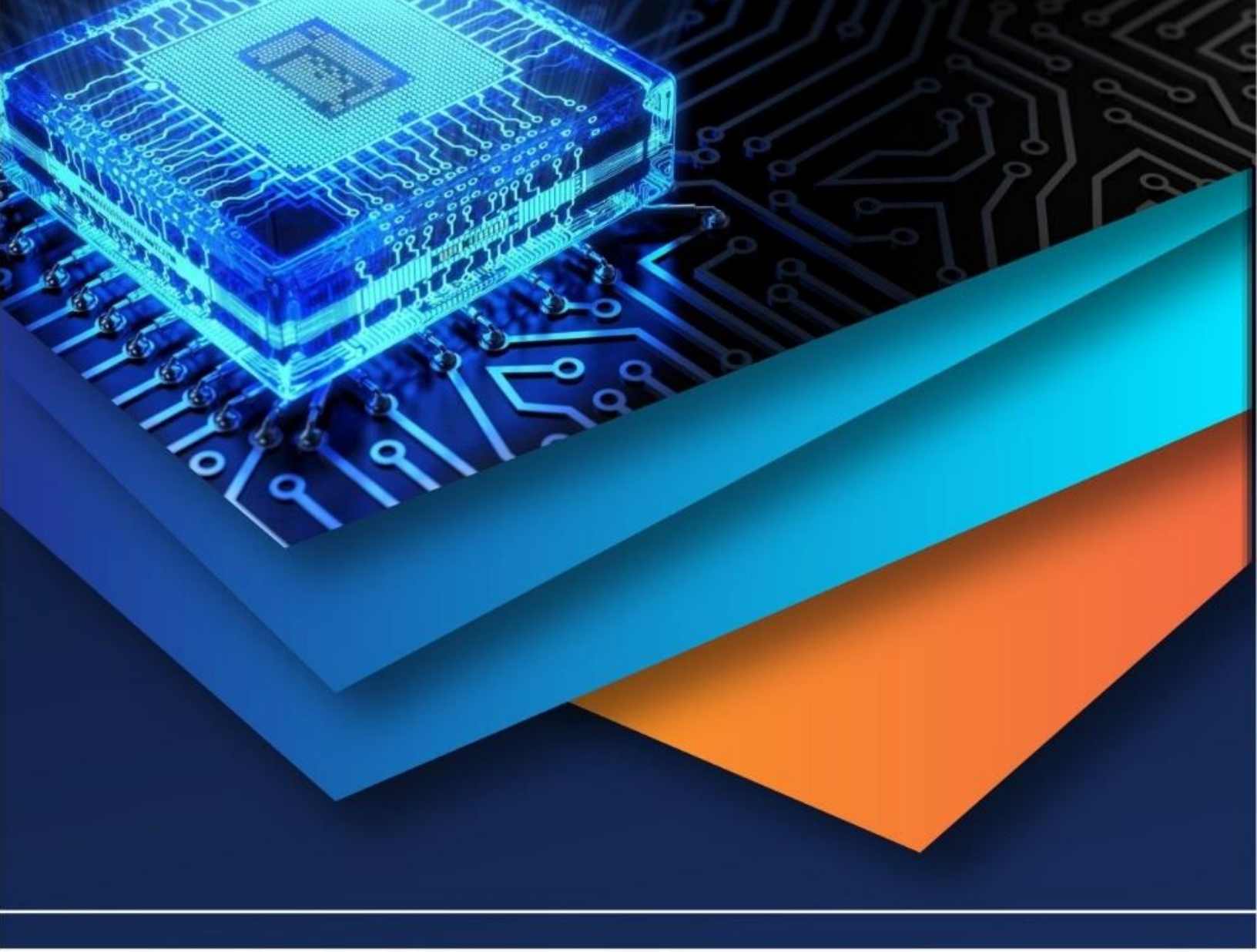
VI. CONCLUSION

An inventive method of improving user experience and engagement is a face emotion recognition-based music recommendation system. The system can precisely identify and analyse users' emotional states in real-time by employing sophisticated face analysis techniques. This enables the system to provide personalized music recommendations that align with users' interests and moods. The suggestion process is further improved by using contextual variables like activity, location, and time of day. This guarantees that the music is appropriate for the user's present circumstances and surroundings.

Furthermore, the integration of user feedback methods facilitates the system's ongoing adaptation and enhancement of its recommendations, ultimately resulting in a more personalized and fulfilling listening experience for users.

REFERENCES

[1] Ramanathan, Radha, Kumaran, Ram Rohan R, Rajat Gupta, and Vishalakshi Prabhu, "An intelligent music player based on emotion identification," 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions, 2017. 10.1109/CSITSS.2017.8447743.
[2] https://www.ijraset.com/research-paper/movie-recommendation-system-using-machine-learning3
[3] Kamasak, wearable physiological sensors to create an emotion-based music recommendation system M. Athavle along with others. Online ISSN: 2582-7006 Artificial Intelligence Conference International (ICAI-2021) IEEE Transactions on Consumer Electronics, vol. 14, no. 8, May 2018, Journal of Informatics, Electrical and Electronics Engineering (JIEEE) A2Z Journals, sensors.What is the doi for TCE.2018.2844736?
[4] https://www.academia.edu/50214622/Music_Recommendation_Based_on_Face_Emotion_Recognition
[5] https://www.researchgate.net/publication/354855186_Music_Recommendation_Based_on_Face_Emotion_Recognition/fulltext/615121e6522ef665fb619ae1/Music-Recommendation-Based-on-Face-Emotion-Recognition.pdf
[6] https://www.irjmets.com/uploadedfiles/paper/issue_1_january_2024/48173/final/fin_irjmets1704818691.pdf
[7] https://www.academia.edu/50214622/Music_Recommendation_Based_on_Face_Emotion_Recognition/
[8] AYUSH Guidel, Birat Sapkota, Krishna Sapkota, Music recommendation by facial analysis, February 17, 2020
[9] Emotion-based music recommendation system, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad; International Journal of Emerging Technologies and Innovative Research (JETIR), Volume 7, Issue 4, April 2020; CH. Sadhvika, Gutta.Abigna, P. Srinivas reddy.
[10] https://ijarsct.co.in/A10827.pdf



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY