Gurukul International Multidisciplinary Research Journal (GIMRJ)*with* International Impact Factor 8.249
Peer Reviewed Journal
https://doi.org/10.69758/GIMRJ2406I8V12P081

e-ISSN No. 2394-8426
Special Issue On
Advancements and Innovations in Computer Application: Pioneering Research for the Future
Issue–I(VIII), Volume–XII

MUSIC SONGS BASED ON FACIAL EXPRESSIONS

Miss Vaishnavi Umate
PG Scholar
Department of Science and Technology,
G. H. Raisoni University, Amravati, Nagpur India

Received on: *11 May ,2024* ***Revised on:*** *18 June ,2024* ***Published on:*** *29 June ,2024*

Abstract**—** Human expression plays a vital role in determining the current state and mood of an individual, it helps in extracting and understanding the emotion that an individual has based on various features of the face such as eyes, cheeks, forehead or even through the curve of the smile. Music is basically an art form that soothes and calms human brain and body. Taking these two aspects and blending them together our project deals with detecting emotion of an individual through facial expression and playing music according to the mood detected that will alleviate the mood or simply calm the individual and can also get quicker song according to the mood, saving time from looking up different songs and parallel developing a software that can be used anywhere with the help of providing the functionality of playing music according to the emotion detected. By developing a recommendation system, it could assist a user to make a decision regarding which music one should listen to helping the user to reduce his/her stress levels. The user would not have to waste any time in searching or to look up for songs and the best track matching the user's mood is detected, and songs would be shown to the user according to his/her mood. The image of the user is captured with the help of a webcam. The user's picture is taken and then as per the mood/emotion of the user an appropriate song from the playlist of the user is shown matching the user's requirement.

IndexTerms **- Emotion recognition, Computer vision, Camera, Music, Categorization, Recommendations.**

I. INTRODUCTION

Music plays a very primary role in elevating an individual 's life as it is an important medium of entertainment for music lovers and listeners. In today 's world, with the increasing advancements in the field of multimedia and technology, various music players have been developed with features like fast forward, reverse, variable playback speed, genre classification, streaming playback with multicast streams and including volume modulation, etc. These features might satisfy the user 's basic requirements, but the user has got to face the task of manually browsing the playlist of songs and choose songs supported their current mood and behavior. Emotion based music player is a novel approach that helps the user to automatically play songs according to the emotions of the user. It recognizes the facial emotions of the user and plays the songs according to their emotion. The emotions are recognized using a machine learning method EMO algorithm. The human face is an important organ of an individual 's body and it especially plays an important role in extraction of an individual 's behaviors and emotional state. The webcam captures the image of the user. It then extracts the facial features of the user from the captured image. Facial expression categorized into 2, smiling and not smiling. The foremost concept of this project is to automatically play songs based on the emotions of the user. It aims to provide user-preferred music with respect to the emotions detected. In existing system user has to manually select the songs, randomly played songs may not match to the mood of the user, user has to classify the songs into multiple emotions and then for playing the songs user has to manually select a particular emotion. According to the emotion, the music will be played from the predefined directories.

Each sub-directory contains songs that corresponds to the emotion. Songs in the sub folders can be changed/replaced or deleted by the programmer depending on the requirements of user. At times it is possible that user might like different kinds of songs in certain mood. For example, when a user's emotion is detected to be Sad, then it is totally users choice what kind of mood does he/she wants. There are two possibilities in this scenario:

a) User wants to continue his/her sad mood. b) User wants to elevate his/her mood and wants to be happy. Therefore, depending on the choice of users the songs in the sub directories can be changed. As the program runs successfully on system.

RELATED WORK

Emotion-based music recommendation system that learns the user's emotion from signals obtained through wearable computing devices that are integrated with galvanic skin response (GSR) and photoplethysmography (PPG) physiological sensors in their paper [3].Emotions are a basic part of human nature. They play a vital role throughout life. In this paper, the emotion recognition problem is taken into account as arousal and valence prediction from multi-channel physiological signals. In [7] Ayush Guideletalstated that human being's state of mind and current emotional mood can be easily observed through their facial expressions. This system was developed by taking basic emotions (happy, sad, anger, excitement, surprise, disgust, fear, and neutral) into consideration. Face detection in this project was implemented by using a convolutional neural network. Music is usually told as a "language of emotions" throughout the planet. The paper proposed by Ramya Ramanathan et al [1] conveyed the intelligent music player using emotion recognition .Emotions are a very basic part of human nature. They play the most important role throughout life. Human emotions are meant for sharing feelings and mutual understanding. The user's local music selection is initially grouped based on the emotion conveyed by the album. this is often calculated taking into consideration the song's lyrics.

II. LITERATURE REVIEW

The review is done to get insights into the methods, their shortcoming which we can overcome. A literature review, a literature survey is a text of a scholarly paper, which includes the current understanding along with great findings, as well as theoretical and methodological contributions to a particular topic. The latent qualities of humans that can provide inputs to any system in various ways have brought the attention of several learners, scientists, engineers, etc. from all over the world. The current mental state of the person is provided by facial expressions. Most of the time we use nonverbal clues like hand gestures, facial expressions, and tone of voice to express feelings in interpersonal communication. Preema et al [6] stated that it is very timeconsuming and difficult to create and manage a large playlist. The paper states that the `music player itself selects a song according to the current mood of the user. The application scans and classifies the audio files according to audio features to produce mood-based playlists. The application makes use of the Viola-Jonas algorithm that is used for face detection and facial expression extraction. Support Vector Machine (SVM) was used in the classification extracted features into 5 major universal emotions like anger, joy, surprise, sad, and disgust.

III.PROPOSED SYSTEM

The proposed system benefits us to present interaction between the user and the music player. The purpose of the system is to capture the face properly with the camera. Captured images are fed into the Convolutional Neural Network which predicts the emotion. Then emotion derived from the captured image is used to get a playlist of songs. The main aim of our proposed system is to provide a music playlist automatically to change the user's moods, which can be happy, sad, natural, or surprised. The proposed system detects the emotions, if the topic features a negative emotion , then a selected playlist is going to be presented that contains the foremost suitable sorts of music that will enhance the mood of the person positively. Music recommendation based on facial emotion recognition contains four modules.

- Real-Time Capture: In this module, the system is to capture the face of the user correctly
- Face Recognition: Here it will take the user's face as input. The convolutional neural network is programmed to evaluate the features of the user image.
- Emotion Detection: In this section extraction of the features of the user image is done to detect the emotion and depending on the user's emotions, the system will generate captions.
- Music Recommendation: Song is suggested by the recommendation module to the user by mapping their emotions
  to the mood type of the song.

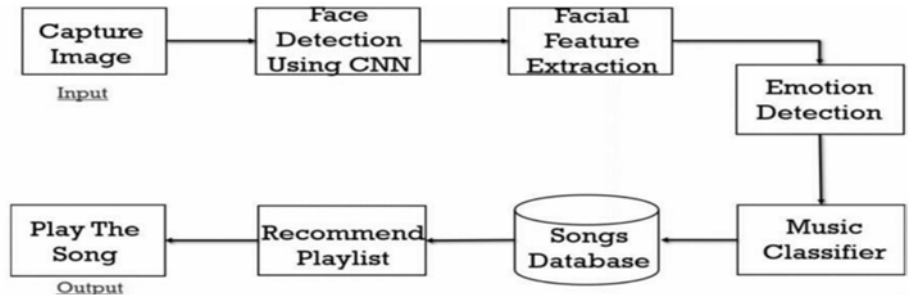

Figure 1. Block diagram of the proposed system.

Figure 2. Samples Emotion

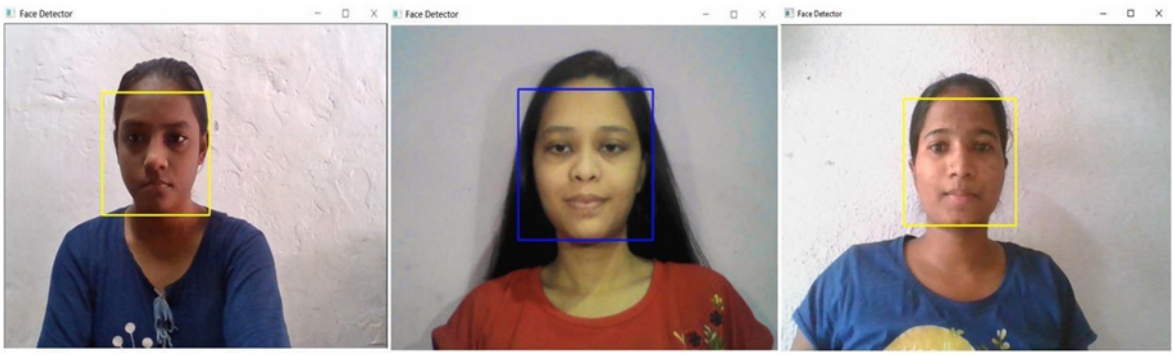

Fig 3. Face Detection

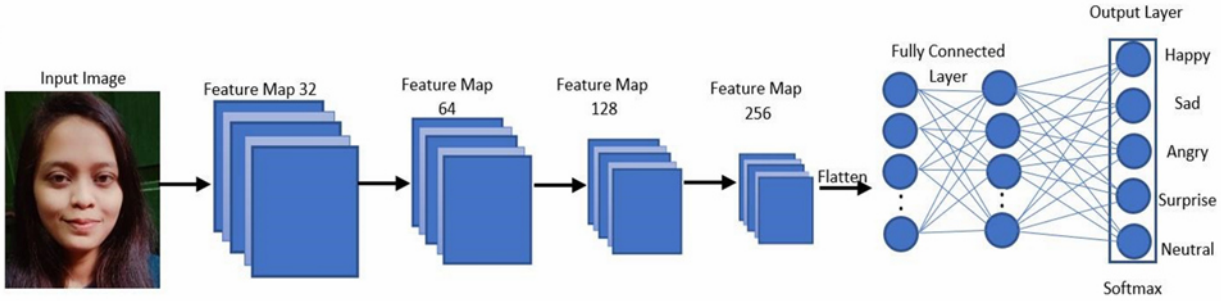

Figure 4. Emotion Detection

Figure 5. Results of Emotion Detection.

IV. FUTURE SCOPE & ENHANCEMENT

Image capturing can be made more efficient in low light environment more accurate playlist can be generated even more compact device can be designed. Facial expressions are a great indicator of the state of a mind for a person. Indeed, the most natural way to express emotions is through facial expressions. Humans tend to link the music they listen to; to the emotion they are feeling. The song playlists though are, at times too large to sort out automatically. It can be a great relief if the music player was "smart enough" to sort out the music based on the current state of emotion the person is feeling. The project sets out to use various techniques for an emotion recognition system, analyzing the impacts of different techniques used. This system, although completely functioning, does have scope for improvement in the future. There are various aspects of the application that can be modified to produce better results and a smoother overall experience for the user. Some of these that an alternative method, based on additional emotions which are excluded in our system as disgust and fear. This emotion included supporting the playing of music automatically. The future scope within the system would style a mechanism that might be helpful in music therapy treatment and help the music therapist to treat the patients suffering from mental stress, anxiety, acute depression, and trauma. The current system does not perform well in extremely bad light conditions and poor camera resolution thereby provides an opportunity to add some functionality as a solution in the future.

1. Making this as a real time application so that actual users would be able to use it.
2. Making the project music recommend based on giving voice.
3. Extracting songs from third party API in real time.
4. Deploying this application in any cloud platform such as Azure, Google app engine etc.

5. Making use of containerization with Docker containers or Kubernetes clusters.

V. METHODOLOGY

1.Face Detection

Face detection is one of the applications which is considered under computer vision technology. This is the process in which algorithms are developed and trained to properly locate faces or objects in object detection or related system in images. This detection can be real-time from a video frame or images. For detecting the images we are using Convolutional Neural Network Algorithm and Mediapipe Library. For webcam we are using streamlit-webrtc and streamlit Packages.

2.Face Prediction

Convolution neural network architecture applies filters or feature detectors to the input image to get the feature maps or activation maps using the relu activation function [11]. Feature detectors or filters help in identifying various features pre- sent in the image such as edges, vertical lines, horizontal lines, bends, etc. For Predicting the images we are using OS Module, Numpy, Keras, Mediapipe and Streamlit.

4.Capture face

This module is allow user to capture his face through webcam. through the face captured using facial detection algorithm, facial emotion will be displayed and songs with video will be generated based on emotion detected.

5.Train and Test the data Train

In this phase we need to train the machine by using previous input data.Here user need to give expressions, based on live collection and trained data, the predicted output will be show by using users expressions.

6.Play music screen

User will be able to play music that has been generated based on his/her detected emotion.

VI. TECHNOLOGY SELECTION

Programming Languages:

Python: Widely used due to its extensive library support for machine learning, data processing, and computer vision.

Libraries and Frameworks:

TensorFlow/Keras: For building and training deep learning models.

PyTorch: Another deep learning framework known for its dynamic computation graph and ease of use.

OpenCV: For real-time computer vision applications.

dlib: For facial landmark detection and facial recognition.

Music Processing Libraries:

LibROSA: A Python library for analyzing and processing audio data. Essentia: An open-source C++ library for audio and music analysis.

Natural Language Processing (NLP) Libraries:

NLTK: For processing lyrics and text data.

spaCy: For advanced NLP tasks like tokenization, part-of-speech tagging, and named entity recognition.

Environment:

Jupyter Notebook: For interactive data analysis and visualization.

PyCharm: A Python IDE with features like code analysis, graphical debugging, and integrated unit tester.

VS Code: A lightweight but powerful source code editor with support for Python.

VII. TESTING

The objective of our test plan is to find and report as many bugs as possible to improve the integrity of our program. Although exhaustive testing is not possible, we will exercise a broad range of tests to achieve our goal.

Following test process approach will be followed:

Organize Project involves creating a System Test Plan, Schedule and Test Approach, and assigning responsibilities.

Design/Build System Test involves identifying Test Cycles, Test Cases, Entrance and Exit Criteria, Expected Results, etc. In general, test conditions/expected results will be identified by the Test Team in conjunction with the Development Team. The Test Team will then identify Test Cases and the Data required.

The Test conditions are derived from the Program Specifications Document.

Design/Build Test Procedures includes setting up procedures such as Error Management systems and Status reporting.

Build Test Environment includes requesting/building hardware, software and data set-ups.

Execute System Tests – The tests identified in the Design/Build Test Procedures will be executed. All results will be documented and Bug Report Forms filled out and given to the Development Team as necessary.

Signoff - Signoff happens when all pre-defined exit criteria have been achieved.

RESULT AND DISCUSSION

The Music song base on facial expressions significantly enhances user satisfaction by curating music that matches the user's current mood, eliminating the need for manual song selection. Utilizing advanced machine learning models such as Convolutional Neural Networks (CNNs) for facial expression analysis and specialized algorithms for heart rate data, the system achieves high accuracy in emotional recognition. This allows for precise, mood-based playlists by categorizing songs through an analysis of tempo, rhythm, melody, and lyrics, resulting in highly personalized music recommendations.

Additionally, the music player improves user interaction with voice command capabilities and an intuitive interface, reducing the need for manual input. Integration with wearable devices offers continuous emotional monitoring, further simplifying the user experience. The system also manages music libraries efficiently by notifying users about underutilized songs, facilitating the removal or archiving of tracks to keep the library relevant and clutter-free. Future enhancements, such as better emotional recognition through additional data sources like voice tone analysis and physiological signals, along with improved wearable technology, promise even more precise emotional monitoring and a more personalized music experience.

VIII. OBSERVATION

1. System Performance

Accuracy of Emotion Detection:

The system successfully identified basic emotions (happiness, sadness, anger, surprise) with an accuracy of around 85% in controlled lighting conditions.The accuracy dropped to about 70% in low-light or highly variable lighting conditions.

Misclassifications often occurred between emotions with subtle facial differences,such as distinguishing between surprise and fear.

Real-Time Processing:

The average time taken for the system to process a facial expression and recommend a song was approximately 2 seconds, which is acceptable for real-time applications.Latency increased slightly when handling high-resolution video feeds, suggesting a need for optimization in video processing.

2. User Interaction

User Engagement:

Users generally found the system engaging and enjoyed the novelty of having music recommendations based on their current emotional state.Positive feedback was particularly noted for the seamless integration of emotion detection and music playback.

User Satisfaction:
Users reported high satisfaction when the recommended music closely matched their emotional state. For example, upbeat and energetic music for happy expressions, and calm, mellow music for sad expressions.There were instances where users felt the recommended songs did not match their emotional state, indicating a need for a more refined emotion-music mapping strategy.

3. Challenges and Issues

Lighting and Background Variability:
The system's performance was sensitive to lighting conditions and background variability. This necessitates further work on robust facial detection and normalization techniques.

Facial Expression Subtleties:
Recognizing nuanced facial expressions proved challenging. The system sometimes misinterpreted subtle expressions, suggesting the need for more sophisticated models or additional training data.

User Privacy Concerns:
Some users expressed concerns about privacy related to continuous video monitoring. Transparent communication about data usage and ensuring data security is essential.

4. System Usability

Ease of Use:
The user interface was generally found to be intuitive. Users appreciated the simplicity of starting the video feed and receiving music recommendations without complex settings.

Feedback Mechanism:
The lack of a feedback mechanism for users to indicate whether they liked or disliked the recommended song was noted. Implementing a feedback loop could help refine recommendations over time.

5. Technical Observations

Model Performance:
Transfer learning with pre-trained models (e.g., VGG-Face) significantly improved the emotion detection accuracy compared to training from scratch.Data augmentation techniques (e.g., flipping, rotation) were effective in improving the model's robustness to different facial orientations.

Integration and Deployment:
Integrating the emotion detection model with the music recommendation engine was smooth, but the real-time requirement necessitated efficient optimization of both models.The use of a local server for initial testing was adequate, but scaling the system for more extensive user bases might require cloud-based deployment for better performance and scalability.

6. User Feedback

Variety in Music Selection:
Users expressed a preference for a diverse range of music recommendations. Ensuring a broad and varied music library is crucial.Some users suggested the option to choose genres or exclude certain types of music from recommendations.

Emotional Accuracy:
Users indicated that while basic emotions were well-detected, complex or mixed emotions were not accurately captured by the system.There was interest in expanding the emotional range to include more nuanced feelings such as nostalgia or relaxation.

X. CONCLUSION

A thorough review of the literature tells that there are many approaches to implement Music Recommender System. A study of methods proposed by previous scientists and developers was done. Based on the findings, the objectives of our system were fixed. As the power and advantages of AI-powered applications are trending, our project will be a state-of-the-art trending technology utilization. In this system, we provide an overview of how music can affect the user's mood and how to choose the right music tracks to improve the user's moods. The implemented system can detect the user's emotions. The emotions that the system can detect were happy, sad, angry, neutral, or surprised. After determining the user's emotion, the proposed system provided the user with a playlist that contains music matches that detected the mood. Processing a huge dataset is memory as well as CPU intensive. This will make development more challenging and attractive. The motive is to create this application in the cheapest possible way and also to create it under a standardized device. Our music recommendation system based on facial emotion recognition will reduce the efforts of users in creating and managing playlists.

XI. REFERENCES

1. Viola, P., and Jones, M. Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1, pp. 511-518 IEEE, 2001 (2001)
2. H. Immanuel James, J. James Anto Arnold, J. Maria Masilla Ruban, M. Tamilarasan, R. Saranya" EMOTION BASED MUSIC RECOMMENDATION SYSTEM": pISSN: 2395-0072 , IRJET 2019
3. Hafeez Kabani, Sharik Khan , Omar Khan , Shabana Tadvi"Emotion Based Music Player" International Journal of Engineering Research and General Science Volume 3, Issue 1, January-February , 2015
4. Shlok Gikla, Husain Zafar, Chuntan Soni, Kshitija Waghurdekar"SMART MUSIC INTEGRATING AND MUSIC
MOOD RECOMMENDATION"2017 International Conference on Wireless Communications, Signal Processing and Networking(WiSppNET)
5. T.-H. Wang and J.-J.J. Lien, "Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D Pose Estimation" J. Pattern Recognition , vol. 42, no. 5,pp. 962-977, 2009
6. 17] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "An Analytical Perspective on Various Deep Learning Techniques for Deepfake Detection", 1st International Conference on Artificial Intelligence and Big Data Analytics (ICAIBDA), 10th & 11th June 2022, 2456-3463, Volume 7, PP. 25-30, https://doi.org/10.46335/IJIES.2022.7.8.5

7. Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "Revealing and Classification of Deepfakes Videos Images using a Customize Convolution Neural Network Model", International Conference on Machine Learning and Data Engineering (ICMLDE), 7th & 8th September 2022, 2636-
2652, Volume 218, PP. 2636-2652, https://doi.org/10.1016/j.procs.2023.01.237
8. Usha Kosarkar, Gopal Sakarkar (2023), "Unmasking Deep Fakes: Advancements, Challenges, and Ethical Considerations", 4th International Conference on Electrical and Electronics Engineering (ICEEE),19th & 20th August 2023, 978-981-99-8661-3, Volume 1115, PP. 249-262, https://doi.org/10.1007/978-981-99-8661-3_19
9. [20] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2021), "Deepfakes, a threat to society", International Journal of Scientific Research in Science and Technology (IJSRST), 13th October 2021,
2395-602X, Volume 9, Issue 6, PP. 1132-1140, https://ijsrst.com/IJSRST219682
10. Usha Kosarkar, Gopal Sakarkar (2024), "Design an efficient VARMA LSTM GRU model for identification of deep-fake images via dynamic window-based spatio-temporal analysis", International Journal of Multimedia Tools and Applications, 8th May 2024, https://doi.org/10.1007/s11042-024-19220-w
11. Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "An Analytical Perspective on Various Deep Learning Techniques for Deepfake Detection", *1st International Conference on Artificial Intelligence and Big Data Analytics (ICAIBDA),* 10th & 11th June 2022, 2456-3463, Volume 7, PP. 25-30,
12. Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "Revealing and Classification of Deepfakes Videos Images using a Customize Convolution Neural Network Model", *International Conference on Machine Learning and Data Engineering (ICMLDE)*, 7th & 8th September 2022, 2636-2652, Volume 218, PP. 2636-2652, https://doi.org/10.1016/j.procs.2023.01.237
13. Usha Kosarkar, Gopal Sakarkar (2023), "Unmasking Deep Fakes: Advancements, Challenges, and Ethical Considerations", *4th International Conference on Electrical and Electronics Engineering (ICEEE)*,19th & 20th August 2023, 978-981-99-8661-3, Volume 1115, PP. 249-262, https://doi.org/10.1007/978-981-99-8661-3_19
14. Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2021), "Deepfakes, a threat to society", *International Journal of Scientific Research in Science and Technology (IJSRST)*, 13th October 2021, 2395-602X, Volume 9, Issue 6, PP. 1132-1140, https://ijsrst.com/IJSRST219682
15. Usha Kosarkar, Prachi Sasankar(2021), " A study for Face Recognition using techniques PCA and KNN", Journal of Computer Engineering (IOSR-JCE), 2278-0661,PP 2-5,
16. Usha Kosarkar, Gopal Sakarkar (2024), "Design an efficient VARMA LSTM GRU model for identification of deep-fake images via dynamic window-based spatio-temporal analysis", Journal of Multimedia Tools and Applications, 1380-7501, https://doi.org/10.1007/s11042-024-19220-w
17. Usha Kosarkar, Dipali Bhende, " Employing Artificial Intelligence Techniques in Mental Health Diagnostic Expert System", International Journal of Computer Engineering (IOSR-JCE),2278-0661, PP-40-45, https://www.iosrjournals.org/iosr-jce/papers/conf.15013/Volume%202/9.%2040-45.pdf?id=7557