M.H. SABOO SIDDIK COLLEGE OF ENGINEERING
8, Saboo Siddik Road, Byculla, Mumbai - 400 008
Department of Computer Engineering

Project Report
On

MUSIC AND MOVIE RECOMMENDATION SYSTEM BASED ON EMOTION DETECTION

By
FERNANDES RENITA CHRISTOPHER (312015)
SHAIKH BILAL AHMED M AKRAM (312044)
SHAIKH KAMRAN ASIF (312045)
SHAIKH NIDA ABDUL MAJEED (312051)

Guided By
Ansari Fatima Anees

University of Mumbai

2022-2023

M.H. SABOO SIDDIK COLLEGE OF ENGINEERING
8, Saboo Siddik Road, Byculla, Mumbai -400 008.
This is to certify that,

Roll no	Name
312015	**FERNANDES RENITA CHRISTOPHER**
312044	**SHAIKH BILAL AHMED M AKRAM**
312045	**SHAIKH KAMRAN ASIF**
312051	**SHAIKH NIDA ABDUL MAJEED**

Of Third year (T.E Semester VI) degree course in Computer Engineering, have completed the specified project report on,

MUSIC AND MOVIE RECOMMENDATION SYSTEM BASED ON EMOTION DETECTION

As a partial fulfilment of the project work in a satisfactory manner as per the rules of the curriculum laid by the University of Mumbai, during the Academic 2022-2023.

_____
Internal Guide

_____ _____
Internal Examiner External Examiner

_____ _____
Head of Department Principal

Mini Project Approval

This project report entitled ***MUSIC AND MOVIE RECOMMENDATION SYSTEM BASED ON EMOTION DETECTION*** **by FERNANDES RENITA CHRISTOPHER, SHAIKH BILAL AHMED M AKRAM, SHAIKH KAMRAN ASIF, SHAIKH NIDA ABDUL MAJEED** is approved for the degree of Computer Engineering.

EXAMINERS

1.

2.

SUPERVISORS

1.

2.

Date:

Place: Mumbai

ACKNOWLEDGEMENT

We would like to express our gratitude and appreciation to our parents for motivating and encouraging us throughout our career.

We wish to express our sincere thanks to our Principal Dr. Ganesh Kame, M.H. Saboo Siddik College of Engineering for providing us all the facilities, support, and wonderful environment to meet our project requirements.

We would also take the opportunity to express our humble gratitude to our Head of Department of Computer Engineering Dr. Zainab Pirani for supporting us in all aspects and for encouraging us with her valuable suggestions to make our project successful.

We are highly thankful to our internal project guide Ansari Fatima Anees whose valuable guidance helped us understand the project better, their constant guidance and willingness to share their vast knowledge made us understand this project and its manifestations in great depths and helped us to complete the project successfully.

We would also like to acknowledge with much appreciation the role of the staff of the Computer Department, especially the Laboratory staff, who gave the permission to use the labs when needed and the necessary material to complete the project.

We would like to express our gratitude and appreciate the guidance given by other supervisors and project guides, their comments and tips helped us in improving our presentation skills. Although there may be many who remain unacknowledged in this humble note of appreciation, there are none who remain unappreciated.

FERNANDES RENITA CHRISTOPHER,
SHAIKH BILAL AHMED M AKRAM,
SHAIKH KAMRAN ASIF,
SHAIKH NIDA ABDUL MAJEED.

Table of Content

List of Tables

List of Figures

Abstract

We suggest a novel method for automatically selecting music and movies based on facial expression. The majority of the current methods include manually playing music, wearing wearable computers, or categorizing based on auditory attributes. We suggest changing the manual sorting and playing instead. For emotion identification, we have employed a convolutional neural network. Python libraries are used to provide recommendations for music and movies. The computational time required to produce the findings and the overall cost of the system are both likely to be reduced by our proposed approach, boosting the system's overall accuracy. The dataset of 8025 images is used to test the system. An internal camera is used to record facial expressions. On the input face photos, feature extraction is done to identify emotions including happiness, sadness, disgust, and neutrality. By determining the user's current feeling, a music and movie playlist is created automatically. Compared to the algorithm in the existing literature, it performs better in terms of computational time.

Introduction and Motivation

1.1 Introduction

Music and Movie has always been a crucial component of human culture and has the ability to elicit emotional experiences from listeners. As digital Music and Movie consumption grows, so does the demand for personalized Music and Movie recommendations that adapt to individual tastes and moods. Current Music and Movie recommendation systems rely on user listening history and collaborative filtering, but they frequently fail to capture the emotional context in which the Music and Movie is listened to.

Emotion detection techniques have been introduced in Music and Movie recommendation systems to fill this void. These systems can provide recommendations tailored to the user's current emotional state by analyzing the emotional content of a song. This method has the potential to improve the overall experience of Music and Watching Movies by allowing users to discover new Music and Movies that relate to their emotions and mood.

The purpose of this research paper is to examine the effectiveness of Music and Movie recommendations based on emotion detection. It will review the research on emotion detection techniques and their use in Music and Movie recommendation systems. It will also assess the accuracy with which different emotion detection algorithms predict the emotional content of a song. Finally, the paper will present the findings of a user study intended to compare the efficacy of emotion-based Music and Movie recommendations to traditional recommendation methods. Overall, the goal of this research is to help develop more personalized and emotionally intelligent Music and Movie recommendation systems.

1.2 Aim

The aim of this study is to contribute to the development of more customized and emotionally intelligent music and movie recommendation systems that can enhance the user experience while offering more memorable experiences and relevant recommendations.

1.3 Objectives

The purpose of this research paper is to look into the impact of emotion-based music and movie recommendation systems. The paper specifically intends to:

- Examine the existing literary works on emotion detection techniques and their application in music and movie recommendation systems.
- Compare the accuracy with which different emotion detection algorithms predict the emotional content of a media piece.
- Through a user study, contrast the performance of recommendation systems that take emotions into account with more conventional recommendation techniques.
- Contribute to the development of more personalized and emotionally intelligent music and movie recommendation systems.
- Provide insights into the potential of emotion detection techniques in improving user experience and providing

1.4 Motivation

The motivation for this research study stems from the huge popularity of personalized recommendations in the world of digital media. As more people consume media online, it is becoming extremely relevant to provide recommendations that are customized to the user's requirements and tastes. Traditional music and movie recommendation systems have relied on user viewing history and collaborative filtering, but frequently fail to consider the emotional context in which the media is consumed.

The goal of this research paper is to provide insights into the potential of emotion detection techniques in improving the effectiveness of music and movie recommendation systems. These systems can provide more personalized and meaningful recommendations by taking into account users' emotional needs, contributing greatly to a more satisfying media consumption experience. The findings of this study can be used to help develop more personalized and emotionally intelligent recommendation systems, which will lead to a more engaging and satisfying media consumption experience for users.

Problem Statement

2.1 Problem Statement

The advancement of technology has led to a massive amount of data being generated every day. As a result, it has become increasingly challenging to filter and select information that suits a person's interests and preferences. Recommendation systems are a popular solution to this problem. Music and movie recommendation systems are commonly used to help users discover new content that aligns with their interests.

In recent years, researchers have explored the use of facial expressions as a means of enhancing the accuracy of these recommendation systems. Facial expressions can convey a person's emotional state and provide insights into their preferences. The proposed system aims to use facial expression analysis to recommend music and movies to users based on their emotional response.

The challenge lies in developing an accurate facial expression recognition system that can detect emotions accurately and recommend content accordingly. The proposed system will use machine learning algorithms to analyze facial expressions and recommend music and movies based on the emotions expressed by the user. The system's success will depend on its ability to accurately recognize facial expressions and recommend content that aligns with the user's emotional state. Overall, this project aims to develop a novel music and movie recommendation system that can enhance user experience by incorporating facial expressions as a new data source.

Literature Review

3.1 Survey of Literature

Reviewing the methods is done to learn about their shortcomings and how to fix them. Many students, scientists, engineers, and other professionals from all over the world have been interested in the latent abilities of individuals that can supply inputs to any system in a variety of ways. Facial expressions can reveal a person's present mental condition. In interpersonal communication, we frequently convey our feelings through hand gestures, face expressions, and voice tones.

Using emotion recognition, the study put out by Ramya Ramanathan et al. [1] described an intelligent music player. Human nature includes emotions at its most fundamental level. They have the most significant role in life. Human emotions are intended to be shared and understood by others. Initial grouping of the user's local music collection is done according to the mood the album evokes. This is frequently determined by taking into account the lyrics of the music. The methods for detecting human emotions that are available for use in creating music players that are emotion-based, the method a music player uses to detect human emotions, and the best way to use the proposed system for emotion detection are all addressed in detail in this paper. It additionally offers a brief idea about our systems working, playlist generation, and emotion classification.

Preema et al. [2] stated that making and managing a big playlist takes a lot of time and effort. According to the report, the "music player itself chooses a tune based on the user's present mood. To create playlists depending on moods, the application analyses and categorises audio files according to audio attributes. The Viola-Jonas method, which is used for face detection and facial emotion extraction, is utilised by the programme. Anger, joy, surprise, sadness, and disgust were the five main universal emotions that Support Vector Machine (SVM) was used to classify.

According to Radhika et al. [3], manual playlist segregation and song annotation based on a user's emotional state is a time-consuming and labor-intensive operation. There have been many algorithms suggested to automate this process. Unfortunately, the currently used algorithms are slow, use extra hardware (such EEG structures and sensors), raise the system's overall cost, and have substantially lower accuracy. In order to save time and labour costs associated with executing this procedure manually, the article introduces an algorithm that automatically creates an audio playlist based on a person's facial expressions. The algorithm described in the study aims to cut down on both the system's overall computational time and cost. It also tries to improve the system design's correctness. Comparing the system's facial expression recognition module to a dataset that is both user-dependent and user-impartial serves to validate the system.

Swati Vaid and coauthors reviewed EEG - Electroencephalography (EEG) is a branch of medicine that uses the electrical activity of brain cells to create recordings of the brain. The brain's cells record the electrical activity of the neurons contained therein. An approximation is created based on the neurons' recorded activity, and the person's emotion is inferred from that analysis. Although the method outlined above achieves the goal of stimulating brain activity, it falls short in terms of portability and affordability [4].

Nikhil et al. [5] use of facial expression to infer the user's mindset. People frequently convey their emotions through facial expressions, hand gestures, and changes in voice or tone, but they mostly do it through their faces. A music player that is based on emotions makes use of less time. People typically have a lot of music on their playlist. Random song play does not make the user feel better. With the use of this system, users can have songs played automatically based on their mood. With the use of the web camera, the user's image is recorded, and the images are preserved. The RGB to binary conversion of the images comes first. A feature-point detection method is the term used to describe this approach of displaying the data. The Haar Cascade technology offered by Open CV can also be used for this procedure. Java was used in the development of the music player. It handles the database and plays the song in accordance with the user's mood.

Research on numerous developments in human affect recognition was conducted by Zeng et al. [6] He concentrated on several methods for handling recordings of affective emotions in audio and/or visual form. The study offers a thorough analysis of audio/visual computing techniques. The effect is regarded as a prototype for various emotion states, including joy, sorrow, fear, rage, disgust, and surprise. In order to aid in the identification of emotions, this paper discussed the difficulties in developing automatic, spontaneous affect recognizers. Also, it pointed out several issues with uni-modal posed emotion recognition that had been overlooked or avoided.

Brain-Computer Interfaces, or BCIs, are used in a system that was described by Chang Liu et al. BCI uses equipment to transmit signals to the processing systems. The person's cognitive state of mind is observed using EEG devices. The system's disadvantage is that they continuously rely on the user's brain for input in order to conduct the classification. The signals from the user's brain are continuously monitored and processed by an algorithm based on MID, which then uses these signals to actively monitor and generate the emotions the user is currently experiencing [7].

The goal of Prof. Nutan Deshmukh et al. [8] was to develop a system that uses a camera to gather the user's emotion and then automates the outcome using an emotion detection algorithm. The user's mood may not remain the same after some time; it may or may not change. This algorithm records the user's mood after every predetermined interval of time. The proposed method, which was superior to earlier existing algorithms and lowers the cost of creating, generates an emotion-based music system in an average computed estimation of 0.95 to 1.05 seconds.

In a cross-database experiment [9], it was shown that using raw features and logistic regression produced the best results when testing the RaFD (Radboud Faces Database) database and the mobile pictures dataset. Using the CK+ dataset as a training set, the accuracy was 66% and 36%, respectively. The experiment's accuracy for the SVM (Support Vector Machine) was dropped from 89% due to the new features (distance and area). More effectively than SVM and a number of other algorithms, the developed approach adapted the findings from the training set to the testing set.

Patra et al. study [10] describes a rudimentary method for classifying the mood of Hindi music by utilising straightforward variables that are derived from the audio. The 10-fold cross validation method produced an average accuracy of 51.56% for the MIREX (Music Information Retrieval Evaluation eXchange) mood taxonomy.

Moreover, an article by Puri et al. [11] states that the current music recommendation study is the outcome of the description of musical resources. It is asserted that present research suffers from a lack of systematic study of user behaviour and needs, a low level of feature extraction, and a single evaluation metric. Situation was shown to be a key component in the personalised music suggestion system. Lastly, it was determined that the accuracy of the recommendation findings was significantly decreased when the weights assigned to all contextual factors were equal.

3.2 Drawbacks of the Existing System

Due to the limited availability of images in the used image collection, the system is still unable to accurately record all emotions. For the classifier to produce accurate results, the image that it receives must be captured in a well-lit environment. The classifier can accurately estimate the user's sentiment if the image quality is at least higher than 320p.

3.3 Research Gap

Work cited	Study	Algorithm & Dataset used	Drawbacks and Research gap
[1]	Ramya Ramanathan et al.	CNN Model was used. 28,709 48x48 pixel grayscale portraits of faces. Seven emotions detected.	Bad light issues in camera resulted in wrong detection of emotion
[2]	Preema et al.	Viola Jonas Method, Support Vector machine(SVM)	Not suitable for large datasets.
[3]	Radhika et al.	Naive Bayes' classification is used. sentiment intensity metrics are used in architecture.	two databases used; user's database focusing on sentiment values and music database focusing on four emotions and 3.0 MB as maximum size of the song.
[4]	Swati Vaid et al.	Electroencephalography (EEG)	It falls short in terms of portability and affordability.
[5]	Nikhil et al.	Feature Point Detection, Haar Cascade Technology	High false-positive detection. hence less accurate
[6]	Zeng et al.	CNN Model was used.	Difficulties in developing automatic, spontaneous affect recognizers.

[7]	Chang Liu et al.	Brain Computer Interface (BCI)	Continuously rely on the user's brain for input in order to conduct the classification.
[8]	Prof. Nutan Deshmukh et al.	K means and K-Nearest algorithm was used.	Records the user's mood after every predetermined interval of time i.e.0.95 to 1.05s
[9]	Raut et al.	RaFD (Radboud Faces Database) database and the mobile pictures dataset, CK+ dataset, SVM algorithm	The accuracy dropped by 89% after using the SVM algorithm.
[10]	Jingye Zhang	Deep Residual Network was used and 28,709 48*48 grayscale images and 6 emotions detected	Average accuracy for the SVM algorithm was quite low.
[11]	Harshal Gunjal et al.	K means and K-Nearest algorithm was used.	Getting input through surveys based on choice and website is not adequate.

Proposed System

4.1 Introduction

A method for music and movie recommendations is suggested in the project. The system is built around the user's emotions. In our lives, music plays a huge role. To relieve tension, people listen to music. Movies play a significant role in our lives. Everyday, people watch movies as a stress reliever and educational tool. The main issue is that consumers occasionally struggle to select appropriate music or films based on their mood or emotion. Each form of emotion is possible, including happiness, sadness, disgust and neutrality. The user will be redirected from home page to the page where he/she will be given the choice to select the mode of emotion detection; whether he/she wants to choose the emotion manually or by switching the camera on and allowing facial recognition. based on the mode selected by the user, the most accurate facial expression is recognized and the mood is detected. The system will then fetch the appropriate song from the database and display a list of the recommended music and movie to the user.

4.2 Flow Chart

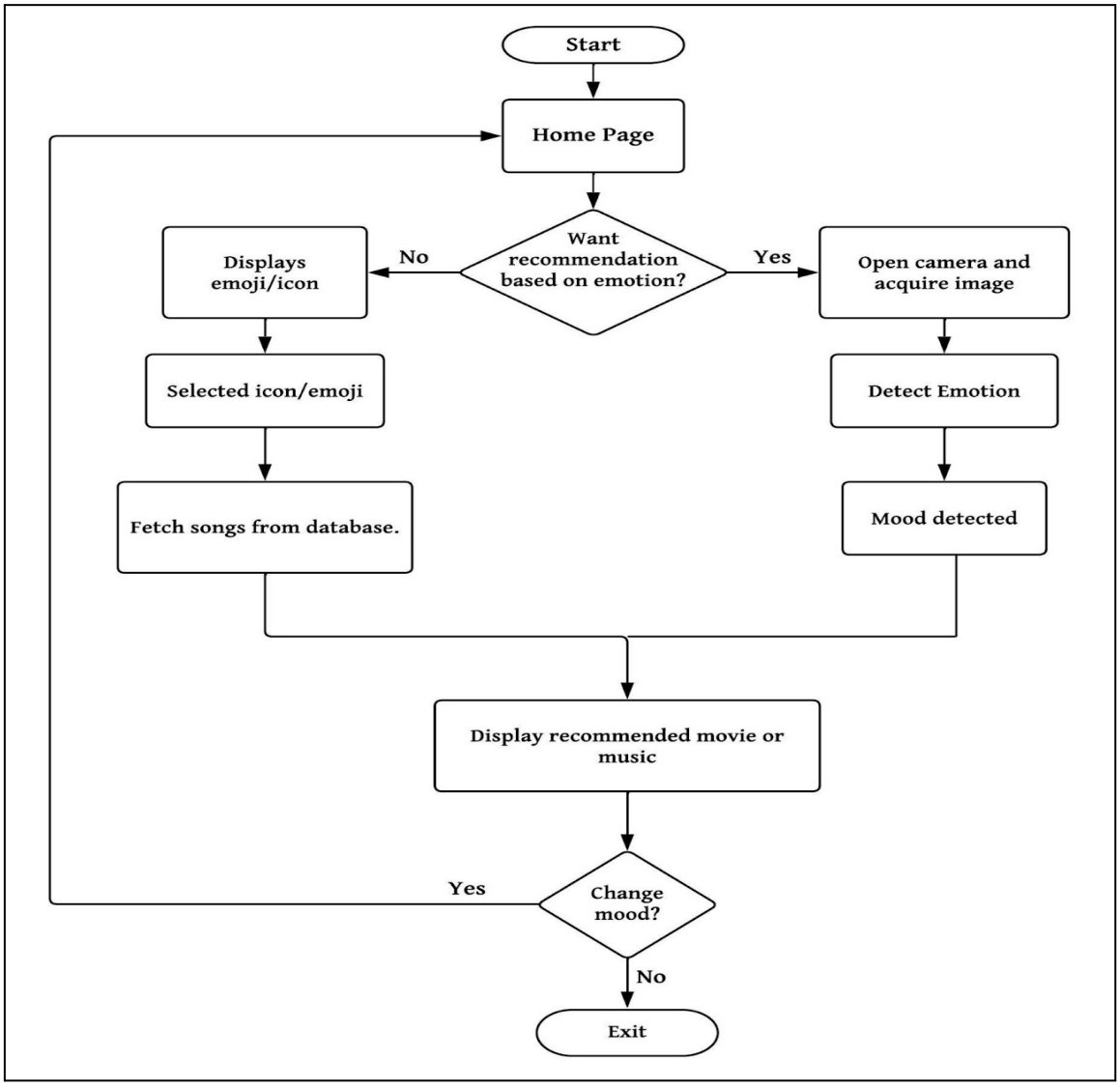

Fig. 1 Flowchart

4.3 Methodology

The methodology for this project on music and movie recommendation based on emotion detection using CNN MobileNet architecture consists of the following steps:

Data collection: A dataset of music and movie clips with emotional labels is collected from publicly available sources. The dataset contains a wide range of emotions, such as happiness, sadness, neutrality and disgust. There are a total 10,700 images out of which 8025 images are used for training and 2675 images are used for validation.

Face Detection: Real time face detection is done with the help of Haar Cascade algorithm which helps in detecting objects in images, irrespective of their scale in image and location. This algorithm is not so complex and can run in real-time.

Data preprocessing: The collected data will be preprocessed to extract relevant features such as Mel-frequency cepstral coefficients (MFCC), chroma features, and spectral features. These features will be used as input to the CNN MobileNet architecture.

CNN MobileNet architecture: The CNN MobileNet architecture is used to train an emotion detection model. MobileNet is a lightweight deep learning model that is designed for mobile devices and has been shown to perform well on image and audio classification tasks. We adapted the architecture to accept audio inputs and train the model on the preprocessed dataset.

Evaluation: The performance of the emotion detection model is evaluated using appropriate metrics such as accuracy and precision. The model is tested on a separate test set to evaluate its generalization performance. This will help in determining the effectiveness of the recommendation system.

Integration into recommendation system: Once the emotion detection model is trained and validated, it is integrated into a music and movie recommendation system. The system will use the emotion detection model to provide personalized recommendations based on the user's current emotional state.

User study: A user study is conducted to compare the effectiveness of emotion-based recommendations with traditional recommendation methods. The study is to evaluate the user experience and satisfaction with the recommended media.

The proposed methodology aims to provide insights into the effectiveness of CNN MobileNet architecture for music/movie emotion detection and its integration into a personalized recommendation system. The study aims to contribute to the development of more personalized and emotionally intelligent recommendation systems for music and movies. Here we have compared two CNN architectures which are MobileNet and GoogleNet and we saw that MobileNet gained an accuracy of 79% and GoogleNet gained an accuracy of 65%.

Technology Used

5.1 Hardware Specification

System	Intel Core i3, i5, i7, and 2 GHz Minimum
RAM	8GB or above
Disk	256 GB or above
Graphic card	1 GB or above

5.2 Software Requirements

Operating System	Windows 7, 10, or Higher Versions
Platform	Pycharm

5.3 Programming Language Used

Model Training	Python
Front End	HTML , CSS
Back End	Flask Framework

Conclusion

&

Future Scope

6.1 Conclusion

A detailed analysis of the available literature reveals that there are various ways to put the Music and Movie Recommender System into practice. A review of approaches put forth by earlier researchers and developers was conducted. The goals of our system were fixed in light of the findings. Our project will be a cutting-edge application of a trendy technology, since the power and benefits of AI-powered apps are on the rise. In this system, we give a general explanation of how music and movies might impact a user's mood as well as advice on how to pick the best songs and films to lift a user's spirits. The technology in place is capable of identifying the user's emotions. The system could distinguish between happy, sad, neutral, and disgust. The accuracy gained by MobileNet architecture was 79% and the accuracy gained by GoogleNet architecture was 65%. Hence, it can be clearly seen that MobileNet architecture outperforms GoogleNet architecture. The suggested approach presented the user with a playlist of music matches that corresponded to the user's emotion after identifying it. Memory and CPU usage increase as a result of processing a large dataset. Development will become more difficult and appealing as a result. The goal is to develop this application as affordably as feasible and on a common platform. Our face emotion-based music recommendation system will lessen users' playlist creation and management efforts.

6.2 Future Scope

Even though this system is fully functional, there is still room for future development. The application can be altered in a number of ways to improve user experience overall and output better results. Some of these use an alternate approach based on extra emotions like disgust and fear that are not considered in our methodology. This feeling included backing the automated playing of music. The system's future plans include designing a mechanism that could aid in the treatment of individuals who are experiencing mental stress, anxiety, acute depression, and trauma through music therapy.

References

[1] Ramya Ramanathan, Radha Kumaran, Ram Rohan R, Rajat Gupta, and Vishalakshi Prabhu, an intelligent music player based on emotion recognition, 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017. https://doi.org/10.1109/CSITSS.2017.8447743

[2] Preema J.S, Rajashree, Sahana M, Savitri H, Review on facial expression-based music player, International Journal of Engineering Research & Technology (IJERT), ISSN-2278-0181, Volume 6, Issue 15, 2018.

[3] CH. sadhvika, Gutta.Abhigna, P. Srinivas Reddy, 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.

[4] PVaid S, Singh P and Kaur C 2015 EEG signal analysis for BCI interface: A review In2015 fifth international conference on advanced computing & communication technologies 143-147 IEEE

[5] Gupte A, Naganarayanan A and Krishnan M Emotion Based Music Player-XBeats International Journal of Advanced Engineering Research and Science

[6] Zeng Z, Pantic M, Roisman GI and Huang TS 2008 A survey of affect recognition methods Audio, visual, and spontaneous expressions IEEE transactions on pattern analysis and machine intelligence 31 39-58

[7] Liu C, Xie S, Xie X, Duan X, Wang W and Obermayer K 2018 Design of a video feedback SSVEP-BCI system for car control based on improved MUSIC method In2018 6th International Conference on Brain-Computer Interface (BCI) 1-4 IEEE

[8] Rabashette MD, Tale MR, Hinge MA, Padale MK, Chavan MR and Deshmukh N Emotion Based Music System

[9] Raut, Nitisha, "Facial Emotion Recognition Using Machine Learning" (2018). Master's Projects. 632. https://doi.org/10.31979/etd.w5fs-s8wd

[10] Patra, Braja & Das, Dipankar & Bandyopadhyay, Sivaji. (2013). Automatic Music Mood Classification of Hindi Songs.

[11] Puri, Raghav & Gupta, Archit & Sikri, Manas & Tiwari, Mohit & Pathak, Nitish & Goel, Shivendra. (2020). Emotion Detection using Image Processing in Python.