AI based Emotion Aware Music Composition using Spotify App

N.Subbulakshmi
School of Computing
Kalsalingam Academy of Research and Education
Virudhunagar, India
subbulakshminammalwar@gmail.com

S. Ariffa Begum
School of Computing
Kalsalingam Academy of Research and Education
Virudhunagar, India
s.ariffabegam@klu.ac.in

Bayapureddy Lakshmireddy gari Archana
School of Computing
Kalsalingam Academy of Research and Education
Virudhunagar, India
99210041152@klu.ac.in

R Chandru
Electronics and Communication Engineering
Sri Ramakrishna Engineering College
Coimbatore, India
r.chandru@srec.ac.in

***Abstract*— Users' emotions and moods can be inferred from their facial expressions. The live video from the system camera reveals these expressions. Many researches have been carried out in the fields of machine learning (ML) and computer vision to teach robots to identify various human emotions or moods. Numerous techniques for identifying human emotions are provided by machine learning. Our proposed method is a music player which works the human mood-based and it provides song recommendations in terms of the real-time mood detection. This is now a capability that is linked with the standard music player application that come preinstalled on our phones. The objectives of this method are to study the user's look, forecast their facial expression, and suggest music based on mood. The discipline of music creation has changed dramatically as a result of technical improvements, providing new avenues for artists and composers to create captivating and emotionally stirring musical works. The concept of "Emotional-Aware Music Composition," a cutting-edge technique that employs artificial intelligence and machine learning to give compositions a profound emotional authenticity, is explored in this abstract. Emotionally aware music creation allows composers to produce music that not only transfers emotion but also deeply engages listeners by bridging the gap between human creativity and technological prowess. This multidisciplinary field uses components from psychology, music theory, and machine learning to achieve this.**

Keywords*— *Music Recommendation system, Artificial Intelligence, Machine Learning

I. INTRODUCTION

Emotions in humans can be widely classified as follows: fear, scorn, fury, surprise, sadness, happiness, and neutrality. Other emotions that fall under this group include contempt, which is a type of disgust, and cheerfulness, which is a variant of gladness. These emotions are quite subtle. There is very little movement of the facial muscles, so it is challenging to discern between variances because even a small alteration may alter the expression. Furthermore, because emotions are so context-dependent, many people may express their facial expression and the same feeling in several ways. Even though the lips and eyes are the only facial features that are seen and they express the most emotion, how these gestures are identified and classed remains a significant issue. Machine learning and neural networks have been effectively used for these purposes.

Machine learning algorithms can be used to determine mood because they have demonstrated to be highly useful in categorizing. In recent years, machine learning and deep learning based neural network technologies are used by the music recommendation system. It has the ability to capture facial expressions, recognize emotions, and create a playlist tailored to the user's mood. The face expression can be captured via a webcam or an Android app. To suggest music, we make use of libraries like pandas, openCV, etc. The goal of this project is to create a music player or recommendation system that can identify a user by face, ascertain their mood at the time, and then offer a playlist based on that mood.

Our intention is to facilitate interactions that gradually become more productive for each individual user. Our technology, the ADAPTIVE PLACE ADVISOR, seeks to assist users in choosing a location that suits their preferences—in this example, restaurants. Three innovative contributions are made by the ADAPTIVE PLACE ADVISOR. To the best of our knowledge, this is one of the few conversational natural language interfaces with a personalized, long-term user model, and the first personalized spoken dialogue system for suggestion. Secondly, it presents a new paradigm for obtaining, applying, and modeling user models. Thirdly, it serves as evidence of a decrease in the quantity of system-user contacts and the length of the dialogue required to identify a suitable item. The shortcomings of both strategies are addressed by combining personalized suggestion with discussion systems. The majority of dialogue systems respond to users in a similar way and don't save information learned from a single session for later use.

The Main objectives of this work includes:

Emo Player: Emo player, an emotion-based music player, is a cutting-edge method that assists the user in having music play automatically according to their feelings. [2] Sound Tree is a web-based music recommendation system that may be implemented as a web service or incorporated into an external online application. It makes use of people-to-people correlation by looking at historical user behavior, including previously downloaded or listened to music. [3]

Reel Time AI: For this system to function, the user must subscribe. After that, the user can post pictures of places with

2024 3rd International Conference on Sentiment Analysis and Deep Learning (ICSADL) | 979-8-3503-9615-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICSADL61749.2024.00015

979-8-3503-9615-7/24/$31.00 ©2024 IEEE
DOI 10.1109/ICSADL61749.2024.00015

lots of people, like restaurants, movie theatres, and shopping centres. The cheerful and sad moods are then recognized by the system. It interprets facial expressions to determine which ones represent happiness and which ones represent sadness and uses those facial interpretations to determine the situation's outcome. Music.AI - It takes the user's mood from the list of moods as input and makes song recommendations based on the mood they've chosen. It combines a collaborative filtering approach with models of content-based filtering. A number of characteristics, such as time, mood, learning history, and setting, are taken into account when choosing music [5].

Rest of the study is organized as follows. Section II Related Works. Section III describes Proposed method. Section IV focuses on Results and Discussions with data analysis. Finally, Section V concludes the filter.

II. Related Work

There are various techniques for analysing emotions through speech, body language, gestures, and facial expressions, among other methods. Numerous studies have been carried out using various methods to identify and categorize the users' physiological, behavioural, and emotional states as displayed on their faces. Several algorithms are applied to the pre-processed digital image of the face in order to extract features and classify it. It's easy but erroneous to classify music by lyrics alone. The main hindrance to this tactic is the language barrier, which restricts how tokens can be categorized. belonging to a linguistic group. Conversely, sentimental It is possible to classify the qualities and meaning of the music, utilizing acoustically related components of sound, such as pitch, rhythm, and tempo. This method aims to identify and extract feature vectors that embody the characteristics of a specific mood.

The older forms of identification, such as pattern, pin, or password, have been replaced in recent years by fingerprint and face recognition technologies. When utilizing facial recognition software, it is important to be able to distinguish between faces that are similar in appearance but differ in certain other ways. To distinguish between the individuals, features must be noticed; feature analysis can be done in a variety of ways.

1. Analysis of principal components 2. Analysis of linear discriminants 3. examination of independent components, etc. By grouping comparable features together using clustering technology, it is possible to prevent features from being too similar in all of the groups and prevent distinct features from falling under the same group. Eventually, by combining comparable qualities, it will be simple to extract distinct features and distinguish them from one another. A preliminary method for classifying the mood of Hindi music has been reported in one such study [11], which makes use of basic audio feature extraction.

Using the 10-fold cross-validation, the MIREX (Music Information Retrieval Evaluation exchange) mood taxonomy yielded an average accuracy of 51.56%. Furthermore, according to an article [10], the findings of current research on music suggestion are described from the standpoint of music resources. It is indicated that the current state of research has a single evaluation index, a poor degree of feature extraction, and a lack of systematic study on user behaviour and demands. Maintaining the Integrity of the Specifications.

According to another study [12], the idea behind their hybrid recommendation system technique will function after their model has received sufficient training to identify the labels. Brain waves and auditory characteristics of songs are automatically used by the process for automatically managing user preferences in the personalized music suggestion service. In their work, music genre classification difficulties are solved using a very short feature vector that is derived from low dimensional projection and pre-existing audio data.

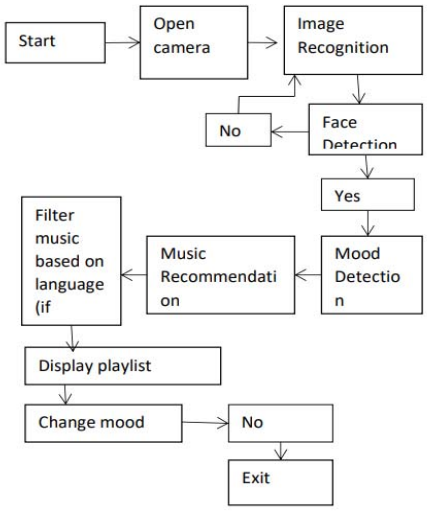

Figure. 1 System Architecture.

III. Proposed Method

The primary objective of the work is known as the mood-based music recommendation system is real-time facial expressions. It is a new product prototype with two primary modules: Recognizing facial expressions, detecting moods, and suggesting music.

This Module is divided into two parts:

Face Detection

The capacity to locate a face in any input image or frame is known as face detection. The bounding box coordinates of the faces that were found are the output. Initially, the OpenCV Python Library was taken into consideration. However, because it was going to be difficult to integrate it with an Android app, the Java Face Detector class was taken into consideration. This library counts of facial expressions and how it recognizes them as a visual object.

Mood Detection

Categorizing their expression into pleased, sad, angry, surprised, neutral, afraid, and disgusted categories. Conventional Python package (Keras) is used for this assignment, however the survey revealed that this method works slowly when connected with Android apps and

requires a significant amount of time for training and validation. Therefore, MobileNet—a CNN architecture model for mobile vision and image classification—was employed. While there are other models as well, MobileNet stands out due to its extremely low processing power requirements for operation and transfer learning applications. The training dataset was created by merging the MMA Facial Expression Recognition dataset [7] and the FER 2013 dataset [6] from Kaggle. The 48x48 pixel grayscale photos in the FER 2013 collection. Images with varying parameters could be found in the MMA Facial Expression Recognition dataset.

Music Recommendation Module

The collection of songs categorized by mood was discovered on Kaggle for two distinct emotions, such as Sad, Angry, Disgust, and so forth. It makes use of a pretrained model to recognize our facial expressions and suggest a playlist tailored to our moods, allowing us to improve or alter our feelings.

Integration - The trained Mobile Net model was saved as a.h5 file. The assets folder created in Android Studio had the labels.txt and tflite files. It can also be connected to the Spotify callback server, allowing us to play a playlist based on our mood.

These are a few of the pictures that were shot with our built Android application. Moods can be roughly classified as "Sad," "Angry," "Happy," "Neutral," "Surprise," or "Fear".

The user can also choose from a selection of emojis based on their mood if their mobile device lacks a camera. The model has been trained to identify the aforementioned emotions and comes with a built-in soundtrack for each one. Additionally, we have the option of choosing the language before listening to the suggested music.

The greatest movie advice system that has been suggested uses cosine similarity and cooperative filtering with singular value corruption to suggest a list. By including the content information of the photographs during the item similarity computations, the current study enhances those procedures. The suggested method suggests to drug users who are interested in stoner possibilities that have not yet been rated the best possible listing of photos. visually represents the likelihood of previously viewed photos by utilizing stoner and selfies that are meant to make stoners happy.

Music Recommendation

A movie advising machine is a complex way to explain a method that tries to predict your selected specifics based on your or people who are similar to you. To put it in layman's words, an advice gadget is one that is intended to predict and filter out specifics based on the user's behavior.

Content Based Filtering content material

Based on a user's effort content, grounded segmentation is a type of advisor machine that attempts to predict what a person might also like. In order to create recommendations, grounded filtering uses keywords and attributes that are assigned to objects in a database, such as details in an online business, and compares them to a profile of stoners. For example, a stoner may choose entertainment apps for their profile; additional features may be implicitly based on the apps that they have already connected. For instance, if the stoner installs another app that was posted by Science Us, the version must recommend details pertinent to this stoner in content- grounded recommender.

Score Point Matching

The connection score informs us of how closely the images above relate to the movie we provided; the closer the score is to 1, the more choices there are. The movie that our recommender framework suggested, Dark Catamount, might be enjoyed by the stoner. This is a peaceful place. Dietary etc. for the details Marvel Lady is a film idea that makes sense because it provides crucial information about a football player preparing a woman for separation. The endeavor that is returned is significant, and men need the teacher Quan Li Kou Sha, among other things.

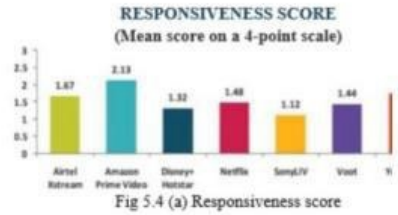

Fig 5.4 (a) Responsiveness score

Figure 2. Streaming Score

IV. Results and Discussions

A. Data Analysis

It is genuinely appropriate for information examination, artificial intelligence, and education. It is a free Jupiter tablet administration that does not require any setup. It also provides free access to functional cash safes in conjunction with the information vigil issue that was resolved successfully. The adviser structure recommended black catamount as a peaceful location and mentioned Sarkar's pailwaan viswasam, etc. for the information vigil for the info marvel lady project. Hereditary etc. for the information; the miracle worker who brought in the bundles and looked through the dataset completed the necessary dataset.

Rating System

A Recommendation system is a class of data sifting framework that looks to guess the standing or inclination a stoner would provide for a thing they are essentially utilized in attractive tasks factors influencing the trustability of stoner survey to-standing transformation are considered a certainty position is figured for conditions created from stoner surveys the certainty position is viewed as in the proposal articulation process utilization of the certainty

position increments vaticination delicacy and stoner fulfillment.

Packages:
1.Pandas (pd) :

One type of data sifting framework is a proposal or recommender framework that looks to guess the standing or inclination a stoner would provide for a thing they are essentially utilized in attractive tasks factors influencing the trustability of stoner survey to-standing transformation are considered a certainty position is figured for conditions created from stoner surveys the certainty position is viewed as in the proposal articulation process utilization of the certainty position increments vaticination delicacy and stoner fulfillment.

2.numpy (np) :

A core Python library for numerical computing is called NumPy. Large, multidimensional arrays and matrices are supported, as are mathematical operations on these arrays.

3.Seaborn (sns) :

Purpose: Seaborn is a data visualization library based on Matplotlib. It simplifies the process of creating informative and attractive statistical graphics.
4.train_test_split :

Purpose: train_test_split is a function from Scikit-Learn used for splitting a dataset into training and testing subsets.
5.Keras :

Python-based Keras is an open-source framework for deep learning. Deep neural networks, are intended to be trained using this interface, which is intended to be user-friendly and intuitive. Although Keras was initially created independently, it has since been incorporated into TensorFlow, one of the most widely used deep learning libraries, and is now a crucial component of TensorFlow's high-level API. Figure. 3, Figure.4, Figure 5, Figure. 6 and Figure.7 are denoted to represent the recommendation based on the moods, Sad, Neutral, Fearful, Happy and Surprised respectively.

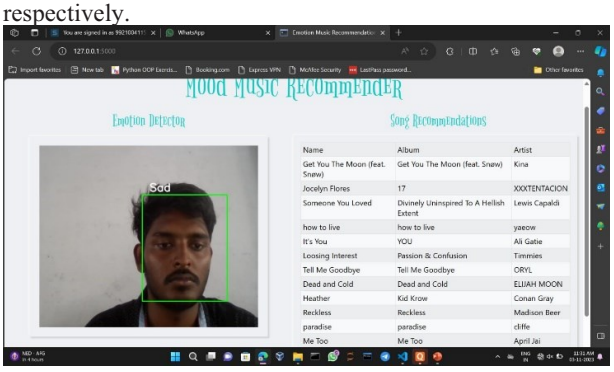

Figure 3. Recommending Songs for SAD Mood

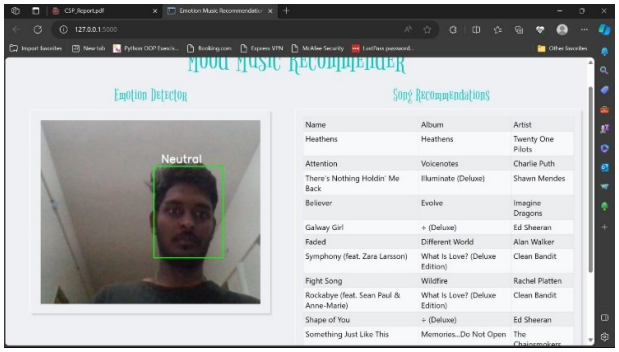

Figure 4. Recommending Songs for Neutral Mood

B. ADAM OPTIMIZATION

In the domain of deep learning and neural networks, the optimization approach known as "Adam" (short for Adaptive Moment Estimation) is in fact frequently utilized. It is a well-liked variation of stochastic gradient descent (SGD) that incorporates concepts from momentum and RMSprop.

Adam is known for its effectiveness in training deep neural networks. It adapts the learning rates for each parameter during training, making it robust to various types of data and architectures.

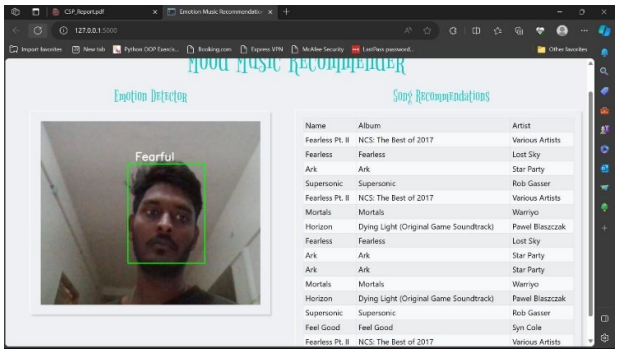

Figure 5. Recommending Songs for Fearful Mood

Adam optimizer and other optimization algorithms are tools used in the training phase of machine learning models for emotion-aware music composition. They help in optimizing the model's parameters to generate music with desired emotional characteristics based on the training data and desired outcomes.

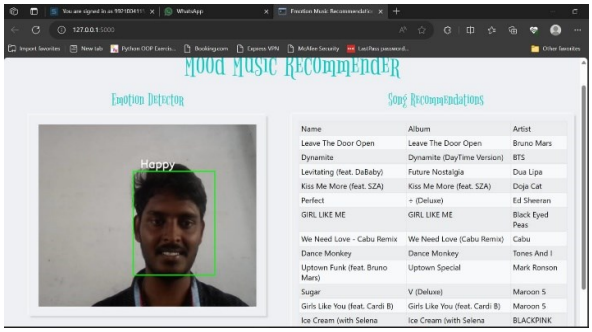

Figure 6. Recommending Songs for Happy Mood

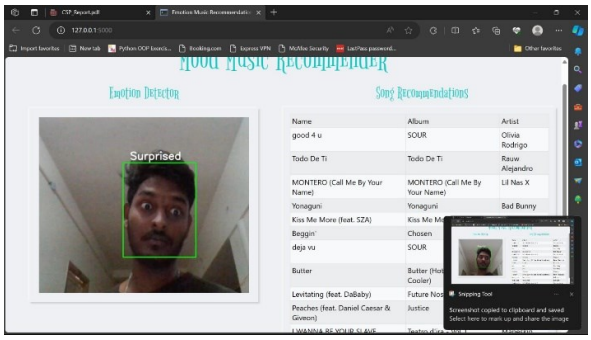

Figure 7. Recommending Songs for Surprised Mood

V. Conclusions

A machine learning model is used to train and recognize a variety of emotions that can be separated from everyone using specific facial expressions, despite the fact that human emotions are subtle and complicated. A person's facial expression is a way of determining their mood, and based on the determination, appropriate music is suggested. Our program, with accuracy of 75 percentage, which detects seven moods: disgust, anger, joyful, fear, surprise, sad, and neutral. The music selected by our Android app will thereafter reflect the identified mood. Apart from the proper recognition of disgust/fear moods, extra attributes like temperature of body or heart rate are considered in addition to only looking at facial expressions. Finding the right music to play can sometimes be a challenging when disgust or dread is present. Therefore, it can be thought of as a prospective area of concentration for our project. The overfitting of our trained model accuracy is sometimes challengeable for instance, such as the "disgust" emotion is frequently categorized as the "angry" mood since the physical features (eyebrows, cheeks) are comparable.

References

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

[2] Hemanth P,Adarsh ,Aswani C.B, Ajith P, Veena A Kumar, "EMO PLAYER: Emotion Based Music Player", International Research Journal of Engineering and Technology (IRJET), vol. 5, no. 4, April 2018, pp. 4822-87.

[3] Music Recommendation System: "Sound Tree", Dcengo Unchained: Sıla KAYA, BSc.; Duygu KABAKCI, BSc.; Işınsu KATIRCIOĞLU, BSc. and Koray KOCAKAYA BSc. Assistant : Dilek Önal Supervisors: Prof. Dr. İsmail Hakkı Toroslu, Prof. Dr. Veysi İşler Sponsor Company: ARGEDOR

[4] Tim Spittle, lucyd, GitHub, , April 16, 2020. Accessed on: [Online], Available at: https://github.com/timspit/lucyd

[5] A. Abdul, J. Chen, H.-Y. Liao, and S.-H. Chang, "An Emotion-Aware Personalized Music Recommendation System Using a Convolutional Neural Networks Approach," Applied Sciences, vol. 8, no. 7, p. 1103, Jul. 2018.

[6] Manas Sambare, FER2013 Dataset, Kaggle, July 19, 2020. Accessed on: September 9, 2020. [Online], Avalilable at: https://www.kaggle.com/msambare/fer2013

[7] MahmoudiMA, MMA Facial Expression Dataset, Kaggle, June 6, 2020. Accessed on: September 15, 2020. [Online],Available at: https://www.kaggle.com/mahmoudima/mma-facial- expression

[8] Dr. Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi, "A real time face emotion classification and recognition using deep learning model", 2020 Journal. of Phys.: Conf. Ser. 1432 012087

[9] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotionspecified expression," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 2010, pp. 94-101, doi: 10.1109/CVPRW.2010.5543262.

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

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

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

[13] Kaufman Jaime C., University of North Florida, "A Hybrid Approach to Music Recommendation: Exploiting Collaborative Music Tags and Acoustic Features", UNF Digital Commons, 2014.

[14] D Priya, Face Detection, Recognition and Emotion Detection in 8 lines of code!, towards data science, April 3, 2019.