MUSIC RECOMMENDATION SYSTEM USING EMOTION DETECTION FROM FACIAL EXPRESSIONS

Bestin Lalu

10th March, 2022

Abstract

In this report, I have proposed the idea of a music recommendation system using feature extraction on facial images to identify the mood of the user and play music accordingly. It is frequently difficult for a person to choose which music to listen to from a vast array of available selections. Depending on the user's mood, there have been numerous recommendation frameworks accessible for concerns such as music, dining, and shopping. Our music recommendation system's main goal is to give consumers choices that are tailored to their tastes. The examination of the user's facial expression/emotion may lead to a better understanding of the user's present emotional or mental state. Music and videos are one area where there is a big opportunity to provide clients with a wide range of options based on their preferences and collected data. Humans are well-known for using facial expressions to indicate what they want to say and the environment in which they intend to say it. More than 60% of users believe that at some point in the future, the number of songs in their music collection will be so enormous that they will be unable to find the song they need to play. By building a suggestion system, it may be possible to aid a user in deciding which music to listen to, hence lowering the user's stress levels. The user would not have to waste time searching or looking up songs since the best track matching the user's mood would be recognised and songs would be displayed to the user based on his or her mood. With the use of a camera, the user's image is taken. The user's photo is captured, and then suitable music from the user's playlist is presented that matches the user's requirements, based on the user's mood/emotion.

1.0 Introduction

People's emotions are mostly expressed through their facial expressions. Music has long been recognised to change a person's mood. Capturing and identifying a person's emotion and showing relevant tunes that fit the person's mood can help to soothe the user's mind and provide a nice overall impression. The goal of the project is to capture a person's feelings through their facial expressions. Through the web camera interface available on computer systems, a music player is meant to record human expression. The program will take the user's image and then extract characteristics from the face of a target human being using image segmentation and image processing techniques in order to discern the emotion that the person is attempting to express. The project seeks to brighten the user's mood by playing tunes that meet the user's

criteria by collecting the user's photograph. Facial expression recognition has been the finest type of expression analysis known to humanity since ancient times. People's facial expressions are the best feasible approach to interpret or deduce the emotion, mood, or thoughts that another person is attempting to communicate. Mood change may also aid in the treatment of depression and melancholy in some circumstances. Many health concerns may be prevented with the use of expression analysis, and efforts can also be made to improve a user's mood.

2.0 Market/Customer/Business need Assessment

As of June 2019, 68 percent of individuals between the ages of 18 and 34 said they listened to music every day, and the majority of their previous generations said they did as well. Music is one of the biggest and evergreen businesses. It can be difficult to select which music to listen to from a vast array of available selections. There have been much advancement in the field of recommendation frameworks accessible for concerns such as music, dining, and shopping. Therefore, by using this technique, we aim to provide a recommendation system that identifies the mood of the user and plays music accordingly.

3.0 Target Specification

The proposed system/service will provide the users with some music choices according to the mood of the user from our database so that their mood can be boosted and they no longer have to choose which music to listen to. The aim of the algorithm is to provide the user with a better experience while listening to music. As for the input from the user, they just have to provide their facial image using the webcam and our algorithm can identify the rest and provide the relevant output.

4.0 External Search

The sources I have used as reference for analyzing the need of such a system for and how such a technique can be used to boost up user’s mood and provide a much better experience to the user while listening to music have mentioned below:

- https://iopscience.iop.org/article/10.1088/1757-899X/912/6/062007/pdf
- https://ieeexplore.ieee.org/document/9073556

4.1 Benchmarking

There are a lot of projects and research papers based on such a system using different methods available on websites like GitHub and IEEE. So such a system is well sought after and with improved accuracy can be applied on an application like Spotify or Apple Music etc. But this technique would hinder the working of the already applied recommendation systems. For a new application such a technique can be tried on and reviewed.

4.2 Applicable Patents

- Patent 1 - Music recommendation method and system

4.3 Applicable Constraints

- Data Collection from the vast music libraries.
- Continuous data collection and maintenance.
- Lack of technical knowledge for the user.
- Taking care of rarely bought products.
- Convincing the users to implement the system in their activities.

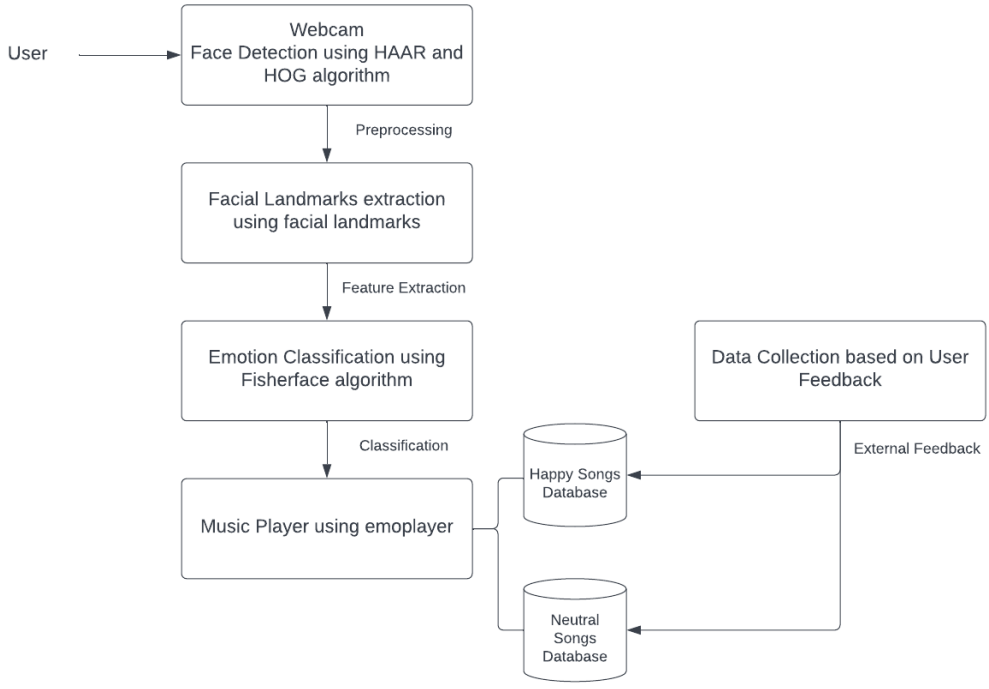
5.0 Business Opportunity

Since the above technique has only been used by only certain applications which have really positive reviews from their user base, this can be extended for small music businesses, not only as a music recommendation system, but also integrated into a music player. Therefore, there is a fair chance of this service being a great business opportunity. Every small music business that requires a continuous flow of music can and would want to opt for using this service in order to always get accurate recommendations. The emergence of every small business is thus a fairly great business opportunity for the service provided by us. As a result, the birth of every small business represents a rather significant commercial potential for the services we supply.

6.0 Final Product Prototype

The final product is a system that provides small music businesses with the user's mood and a list of songs from the provided database that they could use to provide recommendations to the user. The facial image is taken as input to the classifier. The network is trained with the help of CK extensive data set. This is used to identify the emotion being voiced by the user.

The collected frame of the image from the webcam feed is then transformed to a grayscale image to increase the performance of the classifier that is used to identify the face present in the picture after it has been shot. After the conversion is complete, the picture is given to the classifier algorithm, which uses feature extraction techniques to extract the face from the web camera feed's frame. Individual aspects of the face are retrieved and transmitted to the trained network to determine the emotion conveyed by the user once the face has been extracted. The HELEN dataset is used to train a classifier that is used to recognise or acquire facial landmarks from the user's face. There are almost 2000 photos in the HELEN collection. These images will be used to train the classifier so that when it is presented with a completely new and unknown set of images, it will be able to extract the position of facial landmarks from those images using the knowledge it gained from the training set and return the coordinates of the new facial landmarks it detected. The network is trained using the extensive data set provided by CK. This is used to determine the emotion the user is expressing. When this is recognised, the music player chooses a song that best suits the user's mood. The system's general goal is to improve the user's experience and, as a result, reduce some tension or brighten the user's mood. The user does not have to waste time searching or looking for songs, and the music player automatically detects and plays the best tune that matches the user's mood. With the use of a camera, the user's image is taken. The user's photo is captured, and then suitable music from the user's playlist is played, fulfilling the user's requirements, based on the user's mood/emotion.



7.0 Conclusion

The music industry is one of the best grossing and evergreen industries. So for businesses, there should be certain factors which can set them apart from the rest of the competition. Integration of such a system seems like a great opportunity to improve sales and help these businesses grow.

As a result, I've proposed that this system be used. This isn't a full-fledged plan, but it appears to be doable with a lot of hard work and effort.