

MUSIC RECOMMENDATION USING EMOTION DETECTION

SOFTWARE REQUIREMENT SPECIFICATION

08-07-2022

JITHIN MATHEW - RET19CS107

JOANNE JUNEIE - RET19CS108

JOHN JAIS - RET19CS111

JOSHNA ANN SALU - RET19CS115

Table of Contents

1. Introduction	1
1.1 Purpose	1
1.2 Scope	1
1.3 References	1
1.4 Overview	2
2. General Description	3
2.1 Product Perspective	3
2.2 Product Functions	3
2.3 User Characteristics	3
2.4 General Constraints	3
3. Specific Requirements	4
3.1 External Interface Requirements	4
3.1.1 User Interfaces	4
3.1.2 Hardware Interfaces	4
3.1.3 Software Interfaces	4
3.2 Functional Requirements	4
3.2.1 Emotion Detection	4
3.2.1.1 Introduction	4
3.2.1.2 Inputs	4
3.2.1.3 Processing	4
3.2.1.4 Outputs	4
3.2.1.5 Error Handling	5
3.2.2 Music Player	5
3.2.2.1 Introduction	5
3.2.2.2 Inputs	5
3.2.2.3 Processing	5
3.2.2.4 Outputs	5
3.3 Use Cases	5
3.3.1 Use Case	5
3.4 Non-Functional Requirements	6
3.4.1 Performance	6
3.4.2 Reliability	6
3.4.3 Security	6
3.4.4 Portability	6
3.5 Inverse Requirements	6
3.6 Design Constraints	6
4. Analysis Models	7
4.1 Architecture Diagram	7
4.2 Sequence Diagram	8

1. Introduction

The aim of our project is to develop a music recommendation system using emotion detection to help users find better playlists that suit their mood. The emotion of the person is taken as input to display the corresponding music playlist.

A picture of the user showing an emotion is given as input. The face is detected and proceeds to recognize the emotion. According to the emotion extracted from the input picture, the related playlist is displayed. The emotions that will be recognized will be happy and sad. The two emotions have a separate playlist. Datasets are downloaded from kaggle. From the given playlist the user will select a song. The song selected will be played using a music player.

1.1 Purpose

The purpose of the project is to create a web application that plays music on the detection of the emotion and facial recognition.

1.2 Scope

The software product: Music Recommendation using Emotion Detection.

The software is developed for the users to enjoy music depending on their moods. The software is developed to identify two kinds of emotions: happy and sad. Each of these emotions have a set of songs. The goal of the product is to display the set of songs depending upon the emotion of the user. If the user is happy and desires to listen to happy songs the product should be able to display happy songs. If the user is sad and wants to listen to sad songs to relive their sadness, the product should be able to identify the emotion and display songs corresponding to the emotion.

1.3 References

- [1] Shambhavi Bhardwaj; Ashutosh Dixit; Kshitij Yadav; Gokul Rajan V
IEEE paper , A Study on Human Emotion Detection, 2021
Available: <https://ieeexplore.ieee.org/abstract/document/9404755> .
- [2] Charvi Jain; Kshitij Sawant; Mohammed Rehman; Rajesh Kumar
IEEE paper , Emotion Detection and Characterization using Facial Features, 2018.
Available: <https://ieeexplore.ieee.org/document/8710406> .
- [3] Ahlam Alrihail; Alaa Alsaedi; Kholood Albalawi; Liyakathunisa Syed
Music Recommender System for users based on Emotion Detection through Facial Features, 2019. Available: <https://ieeexplore.ieee.org/document/9073556>
- [4] Karthik Subramanian Nathan; Manasi Arun; Megala S Kannan
EMOSIC — An emotion based music player for Android, 2017. Available:
<https://ieeexplore.ieee.org/document/8388671> .

1.4 Overview

The remaining sections of this document provide a general description, including characteristics of the users of this project, the product's software, and the functional and data requirements of the product. General description of the project is discussed in section 2 of this document.

Section 3 gives the functional requirements, data requirements and constraints and assumptions made while designing the music recommendation system. It also gives the user viewpoint of the product. Section 3 also gives the specific requirements of the product. Section 3 also discusses the external interface requirements and gives detailed description of functional requirements.

2. General Description

The resolution of the images in the dataset has to be higher in order to train the images to predict emotion. The music player will only function if a valid emotion is extracted from the emotion detector, specifically happy and sad. If no emotion is detected the music player will cease to run.

2.1 Product Perspective

The project Music recommendation system using emotion detection uses Viola Jones algorithm for efficient extraction of emotion from the face of the user. The edge features, line features and four sided features are the three Haar like features implemented to predict emotions. Only the frontal view is accepted to predict the emotion.

2.2 Product Functions

The software is designed to detect the face of the user and recognise the emotion. The emotion detected will be sent to the music player as an input. Depending on the input, the corresponding playlist will be displayed. The user can select a song and play using the music app. The music app functions to play, pause, stop, change the song to previous or next.

2.3 User Characteristics

The web application provides direct access to songs based on the user's current emotion at that specific instance of time without the user having to explicitly type the emotion. We are expecting our users to use our product for speedy listening to music without having to search for playlists composed of songs of that particular emotion.

2.4 General Constraints

The images have to be of higher resolution for the recognizer to train and predict the correct result. Limited availability of higher resolution image dataset.

3. Specific Requirements

3.1 External Interface Requirements

3.1.1 User Interfaces

The UI is designed such that users of any demographic can use the system easily. The software's user interface is designed such that an image can be captured with the click of a button and the image will be displayed along with the emotion. Based on the emotion detected, a music player will pop up which can be used to select the song to be played.

3.1.2 Hardware Interfaces

The system requires a webcam of good quality so that the image captured is of high resolution so that the system can detect the image easily.

3.1.3 Software Interfaces

The system can run on any device with a compatible version of python installed.

3.2 Functional Requirements

3.2.1 Emotion Detection

3.2.1.1 Introduction

In this phase the emotion of the person is detected by the system. Here after taking the picture of the person, the image is processed. After processing, the particular emotion of the user is returned. This returned value is given to the music player.

3.2.1.2 Inputs

Input required is only the image of the person. This image is passed to the system for processing its emotion.

3.2.1.3 Processing

We have already trained the system to detect different types of emotions using Haar based classifier(Cascade Classifier) based on Viola Jones Algorithm. For that different datasets of emotions were also provided. The input image is processed based on the algorithm we used and with the help of datasets provided and trained, the input image is checked for similarities. According to the similarities found out the emotion is returned.

3.2.1.4 Outputs

The output is the input image with the particular emotion returned.

3.2.1.5 Error Handling

If the input image is not of a person, the output will be the same image itself with the emotion not returned.

3.2.2 Music Player

3.2.2.1 Introduction

In this phase, the list of songs according to the emotion detected by the emotion detector is displayed. We can play any of the song displayed from the list.

3.2.2.2 Inputs

Here the input is passed to the player by the emotion detector with help of a text file. The music player reads this text file and displays the corresponding list of songs according to the emotion.

3.2.2.3 Processing

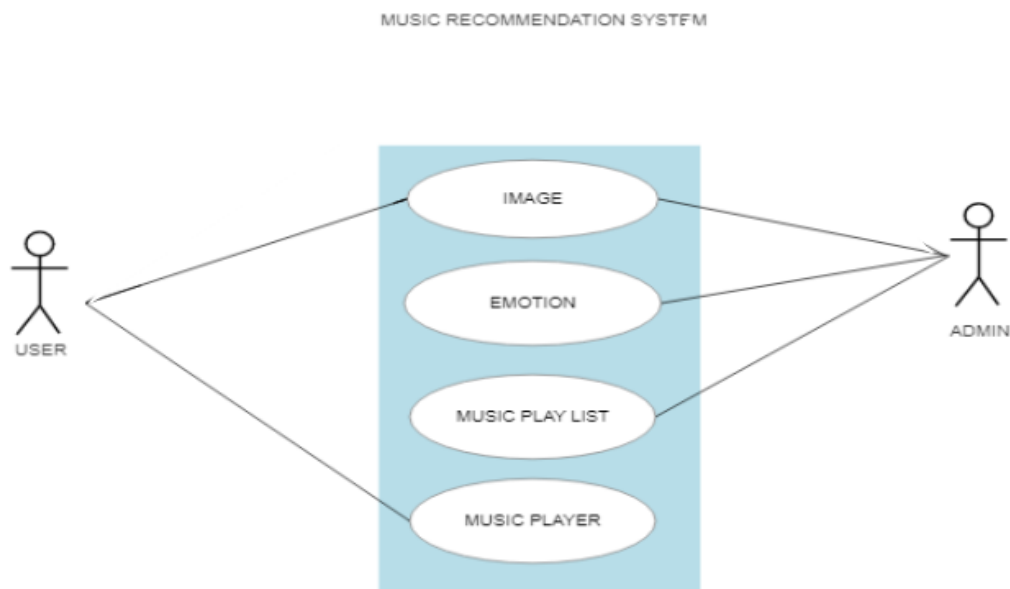
The emotion read from the text file is displayed on the search bar of the player and when we click the button the list of songs will be displayed. We can play any of the songs displayed according to our likes and enjoy the music. The music player also include the option to play, pause, stop and skip the music according to our needs.

3.2.2.4 Outputs

The output is the list of songs based on the emotion.

3.3 Use Cases

3.3.1 Use Case



The user has access to take the image and to choose the music from the displayed playlist. The admin has access to image, detection of emotion and displays the playlist according to the result emotion.

3.4 Non-Functional Requirements

3.4.1 Performance

The performance of the program depends upon the hardware components of the system used. The resolution of the images captured and the quality of the webcam also factors in.

3.4.2 Reliability

Providing an accurate and large dataset improves the accuracy of the trained model. Good lighting conditions for the image to be captured is optimal for ensuring the reliability of the system.

3.4.3 Security

The privacy of the user will not be violated when using the system. Data received will not be shared or publicized as the customer's privacy is of the utmost importance.

3.4.4 Portability

The system can be executed in any system with a compatible version of python provided the dataset and the required dependencies are installed.

3.5 Inverse Requirements

The emotion can be extracted only from the human face, no other feature can be used to detect the emotion.

Only the frontal view of a human face can be used to predict emotion, any other perspective view of the human face will not be processed.

3.6 Design Constraints

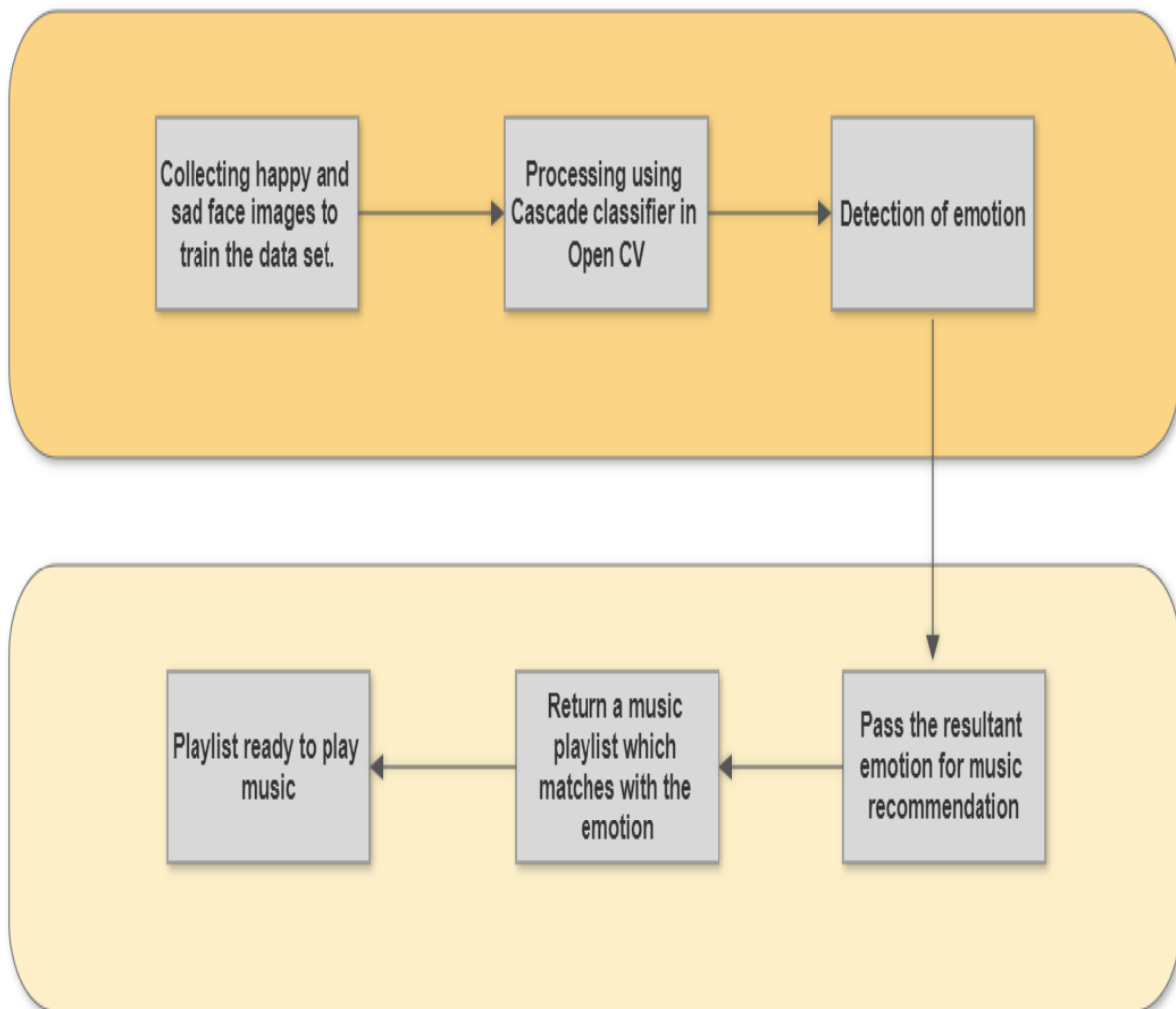
- There are no memory requirements.
- The product must be stored in such a way that allows the client easy access to it.
- Response time for loading the product should take no longer than five minutes.
- A general knowledge of basic computer skills is required to use the product.

4. Analysis Models

4.1 Architecture Diagram

Dataset with 100+ images of happy and sad emotions were collected for the training process. These datasets were processed using the Haar cascade classifier to predict the emotion with maximum accuracy.

The emotion predicted is passed into the music player to display the most suitable music playlist which can then be played by the user.



4.2 Sequence Diagram

The user first gives an image as input, the server checks if it's a valid input input(here checks if it is a human face). The corresponding music playlist along with the emotion predicted is displayed to the user.

