CHAPTER 1 INTRODUCTION

1.1 INTRODUCTION

Artificial intelligence, an extensive, prominent and imperative domain that has attracted a lot of researchers and programs in recent times. This particular domain has taken over the world in very short notice. It is incorporated in over daily life in the form of catboats, digital assistants like Syrian several other technology-based systems. One of the most prominent powers up of artificial intelligence is face recognition techniques. There are many existing systems that could recognize facial emotions. On the other hand, there are systems that recommend music. Bringing together, a system which will recommend music by recognizing the mood of the user from facial emotions is the overall concept described in the paper. Emotion recognition would have larger scope in the near future in fields like robotics for efficient sentimental analysis without the involvement of another human.

1.2 PROBLEM STATEMENT

In this proposed system, we develop a prototype in recommendation of dynamic music recommendation system based on human emotions. Based on each human listening pattern, the songs for each emotions are trained. Integration of feature extraction and machine learning techniques, from the real face the emotion are detected and once the mood is derived from the input image, then the text formatted message in Computer is converted to speech then respective songs for the specific mood would be played to hold the users. Therefore our projected system concentrate on identifying the human feelings for developing emotion based music player using computer vision and machine learning techniques.

1.3 OBJECTIVE OF THE PROPOSED RESEARCH

1. The system aims to help a persons who can’t speak but want to convey their messages to other peoples who are blind.
2. The system aims to automatically play the songs according to the emotions of the user.
3. It determines the extraction of an individual’s behaviours and emotional state.

1.4 METHODOLOGY OF THE PROPOSED RESEARCH

In the proposed system, we integrate computer vision and machine learning techniques for connecting facial emotion for music recommendation. We will use Jupyter notebook tool for coding. And python as a coding language. In this we consider real face input of human using webcam, then image processing techniques are performed on the input acquired image. The features from the input images are extracted using point detection algorithm and some emotion models are extracted from Kaggle dataset. The classification algorithm in OpenCV is used for training the input images for facial emotion detection. Based on the emotions detected the text message on model is converted to speech using Google Text-To –Speech API. And the music would be automatically played from the coding folder.

II. Literature Survey

2.1 Need for research:

Sometimes when we want to catch the exact expression of the user, doing manually may be difficult task when it comes to security purpose. So, we need a model that identity the exact emotion of the person. Example, during the lie detection test or at the airport security the person's expression matters a lot, by the expression only we can recognize if the person is involved in any suspicious activity. If his expression is not neutral then the lie detector test cannot be carried out because the output of that will not be accurate. Also, in airport security if the expression is fear, then that person can be held and an enquiry can be made on him. Also, in the banks or any public places this model can be implemented for the security purpose of the people.

2.2 Existing system:

[1]. The conventional CNN network module is used to extract primary expressional vector (EV). The expressional vector (EV) is generated by tracking down relevant facial points of importance. The music is played according the emotion that is recognized by the user from the database.

[2]. Har cascade Classifier for the face detection in images, SVM classifier for emotion recognition . Identification of human emotion based on human face emotions n has many applications in real life. It avoids the hectic work of choosing the song every time depending on the individual's mood.

[3]. Convolution Neural Network (CNN) to recognize facial emotion and detection of the face, Adaptive Music Player by use the of popular Learning Algorithm stochastic Gradient Descent (SGD) . The results obtained above are very promising. The high accuracy and quick response time of the application makes it suitable for most practical purposes.

[4] .Haar and Hog algorithm for face detection, Fisher face algorithm for emotion detection. Dataset Used, CK+ and Helen Datasets. Here different algorithms are used to detect and recognize human face and emotion.

[5] This model uses FER2013 data set for model training. This paper used Support Vector Machine for Face detection and recognition. Facial expression recognition is carried out by using convolutional neural network (CNN). The music song recommendation experiment in this paper is mainly carried out in PyCharm integrated development environment (Python version is 3.6.8.

[6] Active Shape Model (ASM) is a model-based method, which makes use of a prior model of what is expected in the image, and typically attempts to find the best match position between the model and the data in a new image. The paper used Distribution Model (PDM) is a shape description technique that is used in locating new instances of shapes in images. And Violo-jonce for detection of face

[7]. Viola-Jonze Algorithm to detect objects (eyes, face, nose, mouth) , Principle Component Analysis(PCA) to detect the user emotions. The implemented system will be able to detect the user emotions. After determining the user's emotion, the system provides the user with playlist that contains music matches that detected the mood.

[8]. Haar cascade Classifier for the face detection in images, SVM classifier for emotion recognition. Identification of human emotion based on human face emotions n has many applications in real life. It avoids the hectic work of choosing the song every time depending on the individual's mood.

[9] The system will be able to understand the facial emotion and provide the user with the desired song based on the emotion detected. To classify the emotions Convolutional Neural Network (CNN) is used in this paper.

[10]. Har cascade classifier for face detection. The proposed system is going to process image of facial expressions, recognize the actions the actions with respect to basic emotions and afterwards play music based on emotions.

[11] Har cascade algorithm, fisher face algorithm, PCA. In this project, music recommendation model it is based on the emotions that are captured in real time images of the user.

2.3 Proposed system:

The image is inputted to the webcam. The emotion of the user is being recognized by capturing the facial expression using OpenCV by identifying the Har features. Feature Extraction will take place using VGG16.The emotions are then classified into five different classes of emotions using a convolutional neural network. Then the emotion is being sent to server and matched with Kaggle dataset. The recognized expression is stored in database and the song which is related to that emotion is played from the music model.

CHAPTER 4 SOFTWARE REQUIREMENT SPECIFICATION

4.1 The Overall Description

In this project, we will present a model to recommend a music based on the emotion which is detected from the facial expression. This project proposed designed and developed an emotionbased music recommendation system using face recognition system. Music is the one that has the power to heal any stress or any kind of emotions. Recent development promises a wide scope in developing emotion-based music recommendation system. The proposed system will present facebased emotion recognition system to detect the emotions and play music from the emotion detected.

4.2 Product Perspective

This product is dependent and it is not self-contained because this model is dependent on user's facial expression which is an input. As this model comes under face detection when the model starts the face of human is detected then the emotion is recognized based on that suitable song is played. In the current system the other human being has to recognize the facial expression when to know, whether the other human being is lying or not, or what the other person wants to convey but our model will replace this system to recognize the facial expression and based on their mood a song is played.

3.1.1 System Interfaces

There are some software and hardware interfaces that are required for this model to work properly.

3.1.2 Interface:

Interface is a shared boundary across which two or more separate components. The monitor will act as interface in this project. Between the facial expression as a input and recognition the monitor will act as a interface.

3.1.3 Hardware Interfaces The system has no hardware interface requirements. Although it might not have now, as it only has a software which can be installed on Windows Systems it can be fed to ML based model and can run the model.

3.1.4 Software Interface
Application development Platform: Jupiter Notebook .
Programming Language Used: Python.
Database Used: Kaggle.
Libraries Used: Deep Face (For deep Learning Model Implementation), Open CV.

3.1.5 Communication Interfaces

The main communication protocol will be hypertext transfer protocol (HTTP). This will be used to transfer information back and forth form to client to server. HTTP and GET to use the send the information securely over the browser. Network communications capability will be used to connect to remote databases and/or web services for authentication and importing the user's information.

3.1.6 Memory Constraints

To install Jupyter Notebook-Minimum RAM required is -2GB .
To install Tensor Flow libraries and model-4GB.
Secondary Memory: Minimum Hard Disk space – 2 GB.

3.1.7 Operations

To install Jupiter notebook in windows.
To install some Tensor Flow models.
To install required libraries to run the python code.
To set up the PC/Desktop with good resolution camera.

3.4 Constraints

1.If we are providing the input facial expression through video, then the video quality should be good.
2.If we are providing the input in real-time then there should not be any shadow image in the background and the facial expression should be clear.
3.There should not be any background noise when the music was playing in background.
4.There should be some milliseconds of delay between the facial expression when expression is giving as input.

5. PC should reach the required constraints to run the model.

3.5 Assumptions and Dependencies

The image which is acquired by the camera must be human face. There are only seven expressions on which recognition is taking place. When user is giving a facial expression as input only the expression on the facial part is considered. There should be some time delay (in milliseconds to seconds) should be there from changing one expression to another.

4.2 Function

Display personal emotion and indicate an individual's intentions within a social situation.

4.3 Performance Requirements:

This model gives result for only one face at a time. If more than one face is detected than the accurate result will be display for the face that meets all the requirements correctly (more clearly visible face, face without presence of any hair on it etc.). For the correctly detected face the result will be obtained in some milliseconds to seconds.

4.5.1 Standards Compliance:

No compiler is used.

4.6 Software System Attributes

The availability of this model is user dependent and works on user's facial action. Hence it is available for different emotions.

4.6.1 Reliability:
The more accurate the deep learning model is built the more reliable the model becomes.
4.6.2 Availability:
The usage of lighter deep learning model will increase the availability. The lighter the model is lesser the time it will take to load the output.
4.6.3 Security:
The images stored in the database can be accessed only by the admin.so there is no chance of other people accessing the data.
4.6.4 Maintainability:
It is maintainable and its functionality is to ease for user to use.
4.6.5 Portability:
As a lighter deep learning model is used the storage space required is drastically reduced which increases the portability.

4.9 Software and Hardware requirements

Software Requirements:
➢ Database Used: Kaggle
➢ Libraries Used: Deep Face (For deep Learning Model Implementation) Open CV

Hardware Requirements:

➢ Processor
➢ Hard Disk
➢ Memory
➢ Laptops or computers with webcam

CHAPTER 5 DESIGN PHASE

5.1 ER-diagram

ER Diagram stands for Entity Relationship Diagram, also known as ERD is a diagram that displays the relationship of entity sets stored in a database. In other words, ER diagrams help to explain the logical structure of databases. ER diagrams are created based on three basic concepts: entities, attributes and relationships.ER Diagrams contain different symbols that use rectangles to represent entities, ovals to define attributes and diamond shapes to represent relationships.

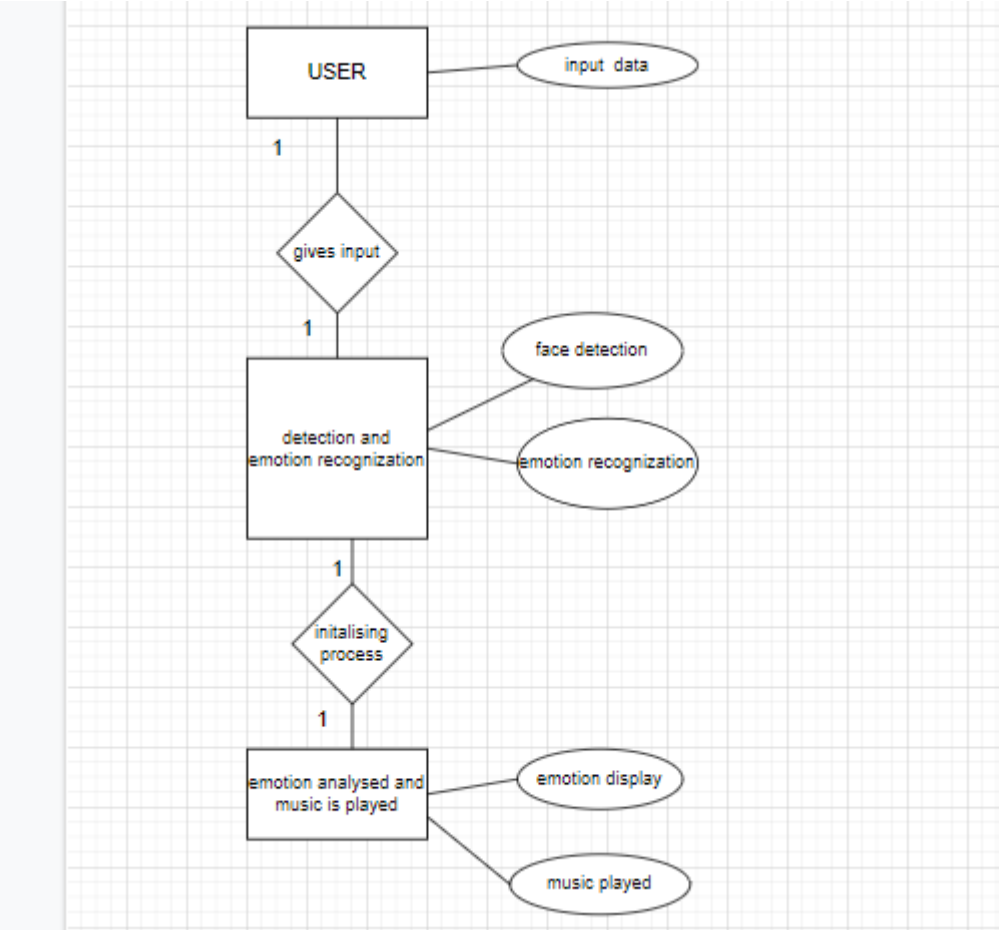


Fig 1. ER Diagram of FEB music recommendation system

5.2 ARCHITECTURAL DESIGN

Architectural design is a process for identifying the sub-systems making up a system and the framework for sub-system control and communication. The output of this design process is a description of the software architecture. Architectural design is an early stage of the system design process. It represents the link between specification and design processes and is often carried out in parallel with some specification activities. It involves identifying major system components and their communications.

The architectural design for emotion-based music recommendation system consists of various steps. Firstly, the data is collected from the captured image. It processes various technique including Tenser flow, music module, feature extraction detection is done by OpenCV, TensorFlow Keras Api, and emotion is classified by using Convolution Neural Network. And based on emotion music is played as output.

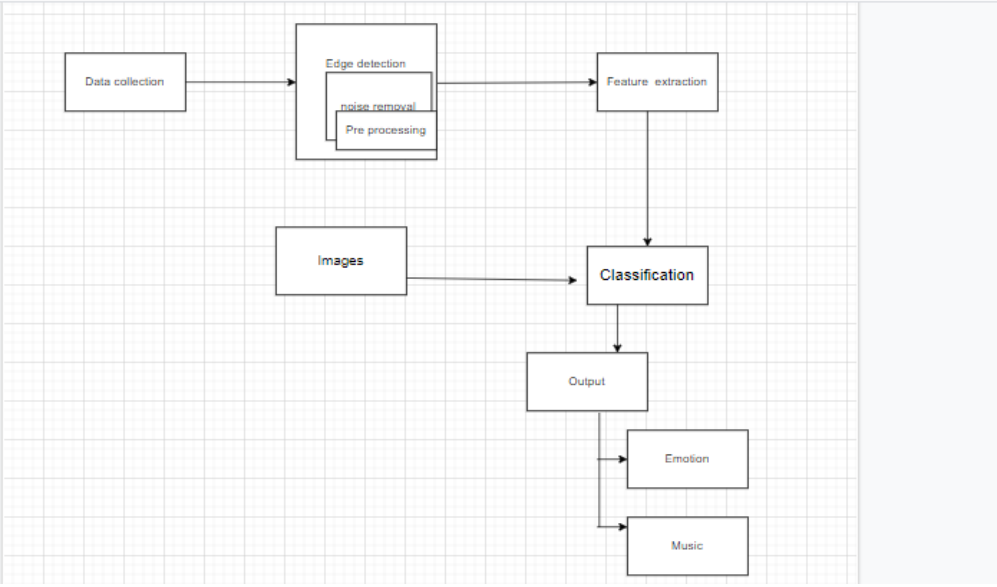


Fig 2. Architectural Designm of FEB music recommendation system

5.3 Data flow diagram

Data Flow Diagrams are used to graphically represent the flow of data in a business information system. DFD describes the processes that are involved in a system to transfer data from the input to the file storage and reports generation. Data flow diagrams can be divided into logical and physical. The logical data flow diagram describes flow of data through a system to perform certain

functionality of a business. The physical data flow diagram describes the implementation of the logical data flow.

The above data flow diagram from Fig. 2. Shows the flow of data for recognizing the emotion and based on that playing the song model.

Step1. The image is inputted to the webcam.

Step 2. The emotion of the user is being recognized by capturing the facial expression using OpenCV by identifying the Haar features.

Step 3. Feature Extraction will take place using VGG16

Step 4. The emotions are then classified into 7 different classes of emotions using a Convolution Neural Network.

Step 5. In this step the emotion is being sent to server and matched with Kaggle dataset.

Step 6. The recognized expression is stored in database and the song which is related to that emotion is played from the music model.

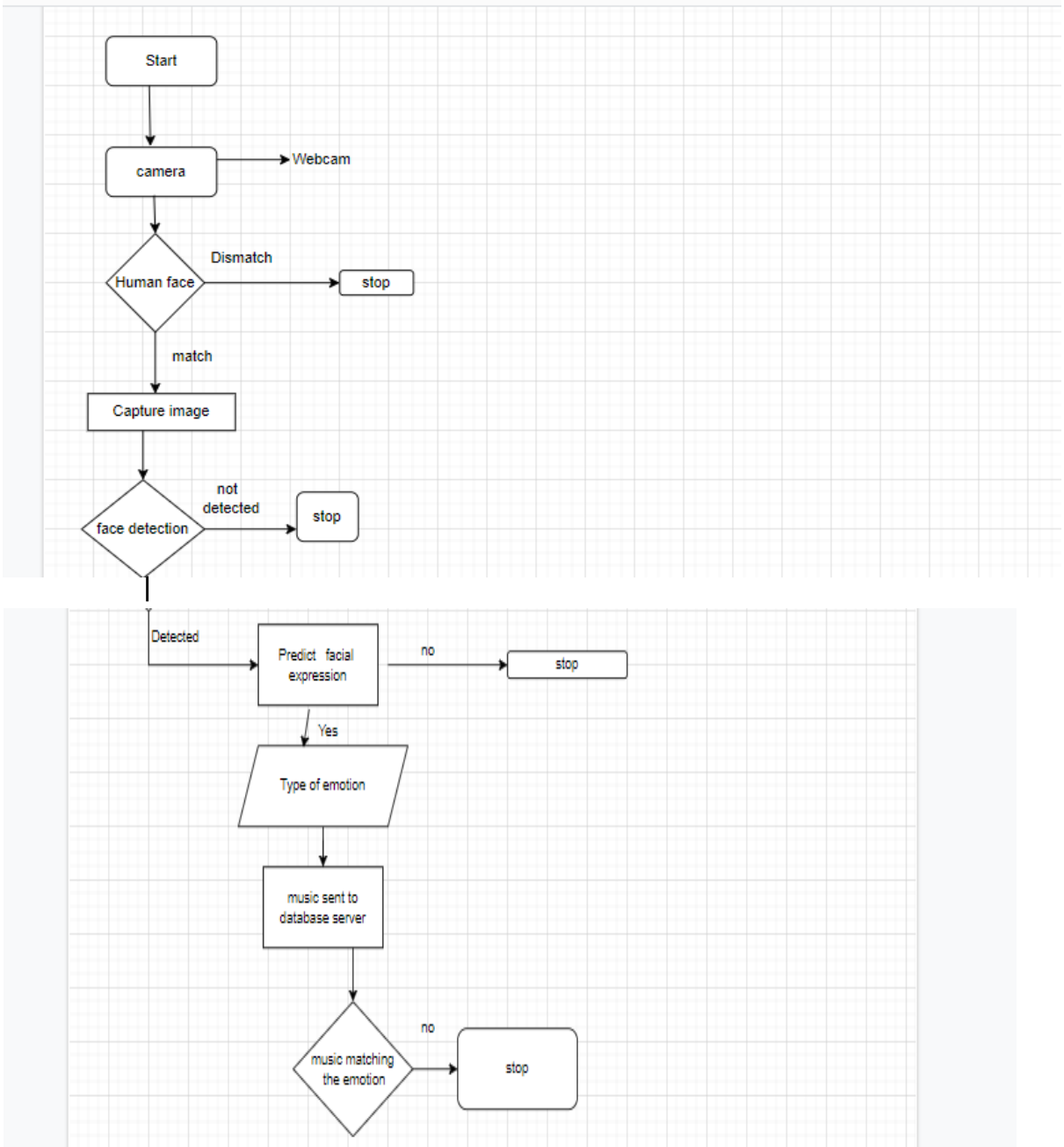


Fig 3. Data Flow Diagram of FEB music recommendation system

5.4.CLIENT SERVER MODEL

In client-server model, any process can act as Server or Client. It is not the type of machine, size of the machine, or its computing power which makes it server; it is the ability of serving request that makes a machine a server. A system can act as Server and Client simultaneously. That is, one process is acting as Server and another is acting as a client. This may also happen that both client and server processes reside on the same machine.

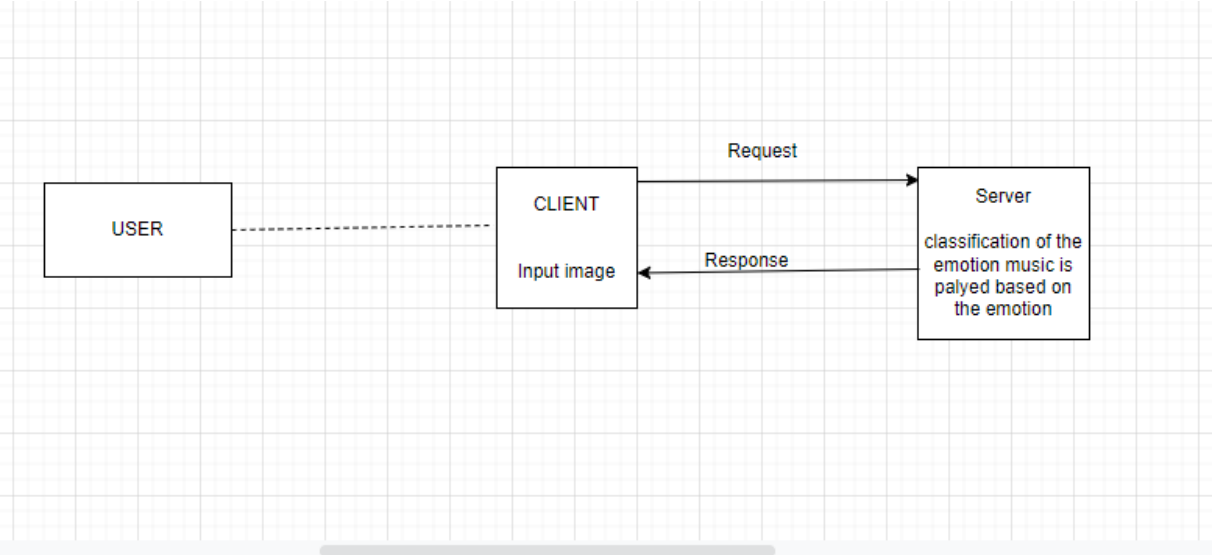


Fig 4. Client Server Model of FEB music recommendation system

5.5 Class diagram

In software engineering, a class diagram in the Unified Modelling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information

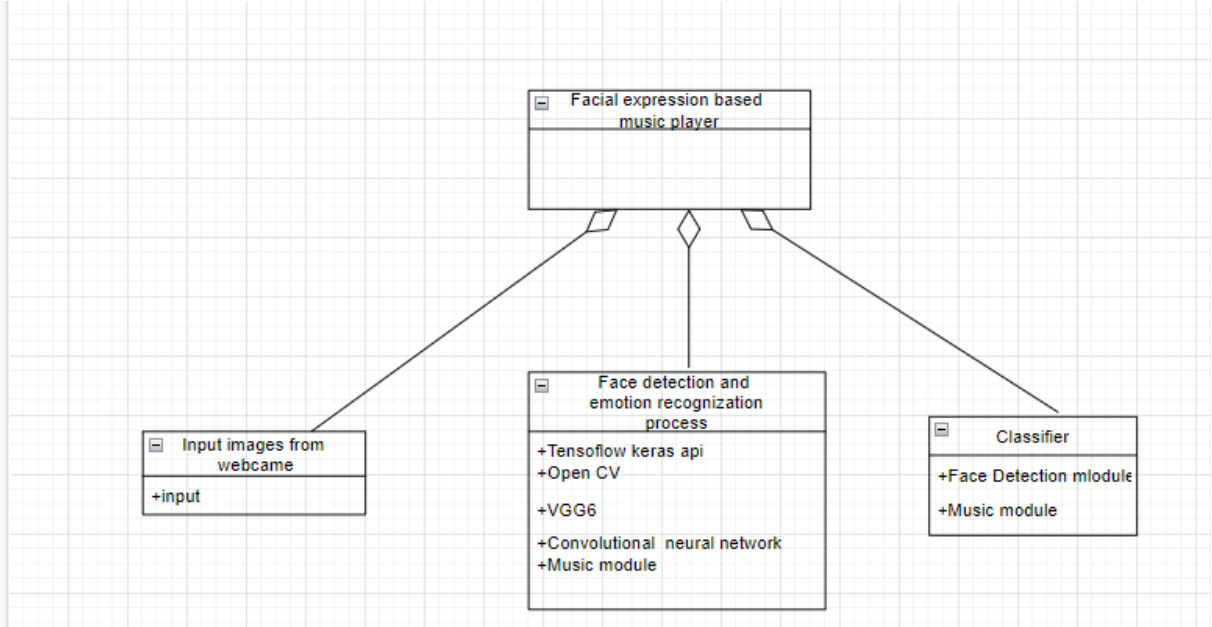


Fig. 5. Class Diagram of FEB music recommendation system

5.6 Advance state diagram

A state diagram is a graph whose nodes are states and whose directed arcs are transitions between states. It is used to represent the condition of the system or part of the system at finite instances of time. It's a behavioural diagram and it represents the behaviour using finite state transitions. State diagrams are also referred to as State machines and State-chart Diagrams.

Face Detection and Emotion Based Music Recommendation System

State names must be unique within the scope of state diagram. All objects in a class execute the state diagram for that class, which models their common behaviour. It provides a very detailed picture of how a specific symbol changes state.

AGGREGATION CONCURRENCY

State aggregation means collection of state diagrams one for each part and relationship. Transition for one object depends on another object that allows interaction between the state diagram. The below figure shows the Facial expression based music player for Input images from webcam where face detection and emotion recognition process are done.

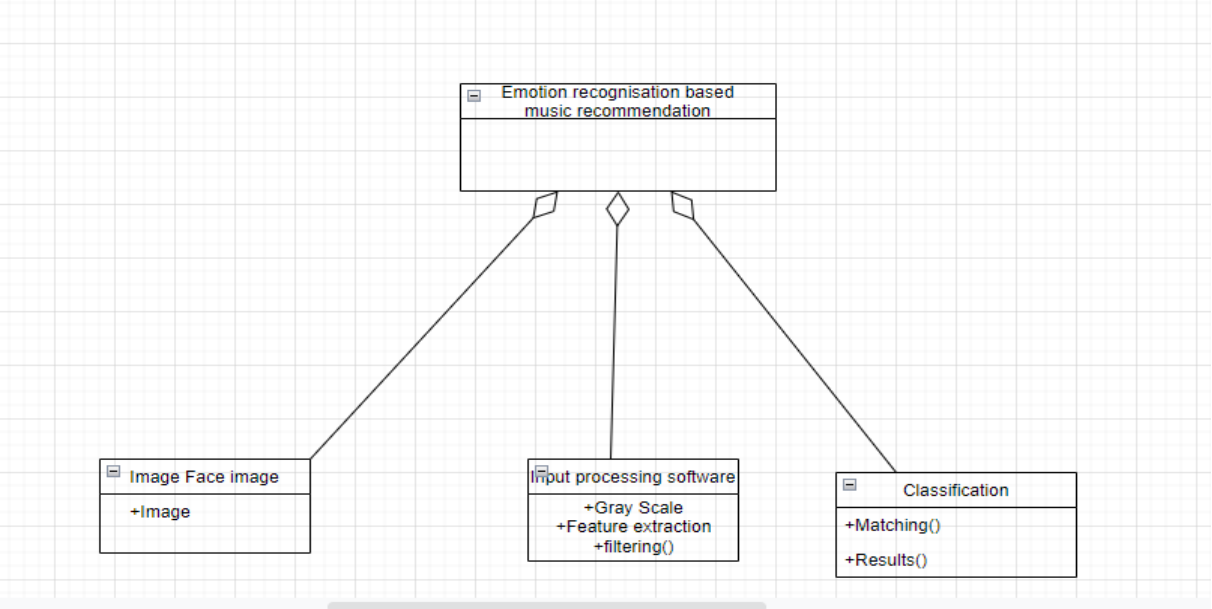
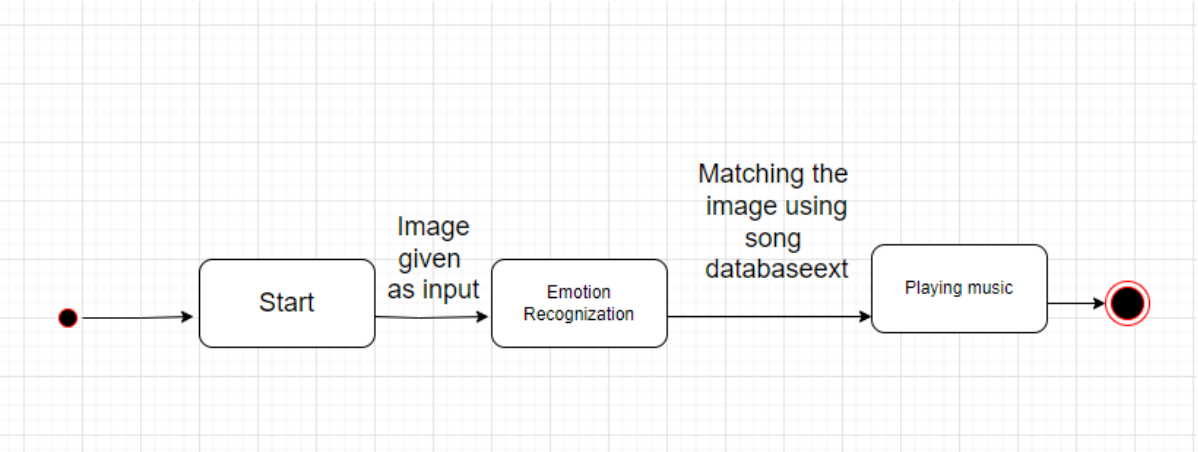


Fig. 5.6.1. Advanced Class Design of FEB music recommendation system



5.6.2. Advanced State Model of FEB music recommendation system

5.7 Use case diagram

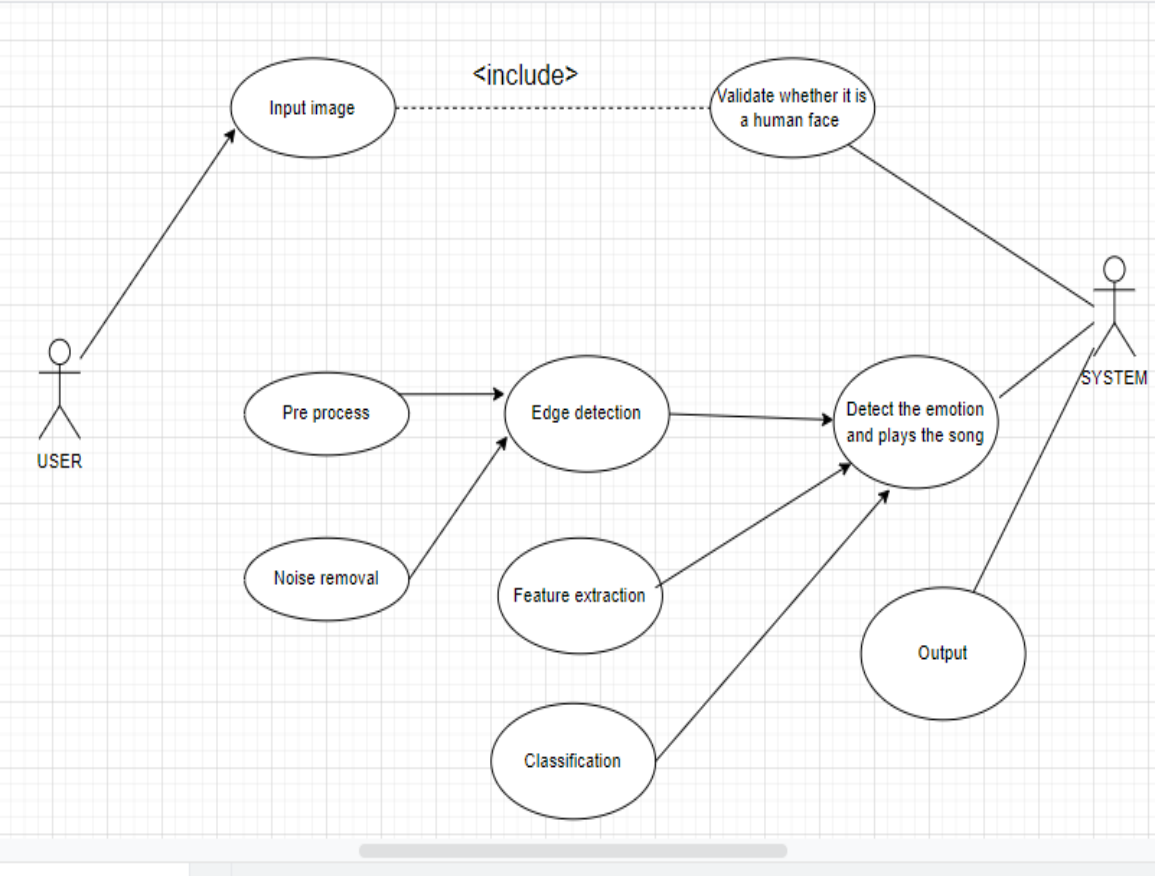


Fig 5.7. Use case Model of FEB music recommendation system

The following gives the description of use cases.

1.Use case: Input human face as input

User will give the human face as input to the model.

2. Actors: Users

Include: Validation

Summary: System will validate whether given image is Human face or not.

Actors: User

3.Use case: Classification

Summary: The System classifies the emotion

Actors: System

4. Use case: Recognition

Summary: Recognize the emotion of the face.

Actors: System

Generalization: Recognizing the expression of a face and playing the music according to the music.

5.8. SEQUENCE MODEL

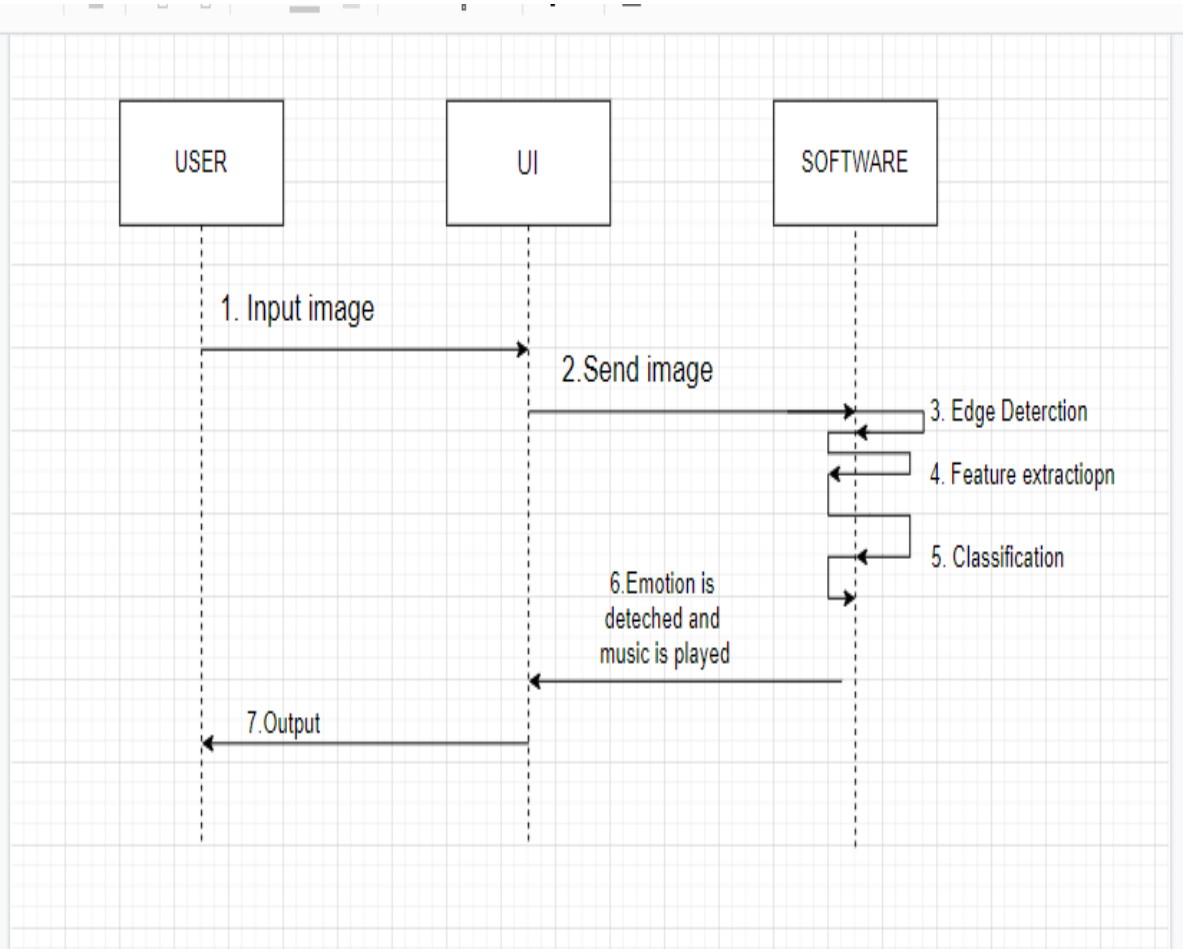


Fig. 5.2.2. Sequence model of FEB music recommendation system

Objects: User, UI, Software

Messages:

1.Input Image

2.Send Image

3.Edge Detection

4.Feature Extraction

5.Classification

6.Emotion is detected and music played

7.Output

CHAPTER 7 TESTING PHASE

7.1 System Setup:

➢ Hardware Requirements:

o Processor CORE i3 or high

o Hard Disk -200 GB

o Memory -8GB RAM

o Camera

➢ Software Requirements:

o Python version 3 and above

o PyCharm IDE

o Music player

o Python Packages(numpy,pip,keras,matplotlib,eel,os etc)

7.2 Types of Tests carried out:

➢ Unit testing

➢ Integrated testing

➢ Performance testing

7.3. Test Cases:

The test cases given below are for individual modules and test the important features ofthe module:

Test Case ID		Test Case Description		
01		Check for image capture status		
Steps	**Test data**	**Expected results**	**Observed results**	**Remarks**
Step 1	Check For Image Capture status	System should capture theuser's face as input	System captures the user's face as input	PASS

Test Case ID		Test Case Description		
02		Check for happy expression		
Steps	**Test data**	**Expected results**	**Observed results**	**Remarks**
Step 2	Check For happy expression	System should identify the captured expression ashappy and should play therespective music	System identifies the captured expression as happy and plays the respective music	PASS

Test Case ID		Test Case Description		
03		Check for sad expression		
Steps	**Test data**	**Expected results**	**Observed results**	**Remarks**
Step 3	Check for sad expression	System should identify the capturedexpression as sad and should play the respective music for sad	System identifies the captured expression as sad and plays the respective music .	PASS

Test Case ID		Test Case Description		
04		Check for angry expression		
Steps	**Test data**	**Expected results**	**Observed results**	**Remarks**
Step 4	Check for angry expression	System should identify the captured expression as angry and should play the respective music for sad	System identifies the captured expression as angry and plays the respective music .	PASS

Test Case ID		Test Case Description		
05		Check for neutral expression		
Steps	**Test data**	**Expected results**	**Observed results**	**Remarks**
Step5	Check for neutral expression	System should identify the captured expression as neutral and should play the respective music for neutral	System identifies the captured expression as neutral and plays the respective music	PASS

Test Case ID		Test Case Description		
06		No face detected		
Steps	**Test data**	**Expected results**	**Observed results**	**Remarks**
Step 6	No face detected	Song should not be played	No song is played	PASS

Test Case ID		Test Case Description		
07		Displaying the detected expression as text message in the console.		
Steps	**Test data**	**Expected results**	**Observed results**	**Remarks**
Step 7	Displaying the detected expression As text message in the console.	For respective expressions detected, emotion should be displayed as text message in console	For respective Expressions detected, emotion is displayed as text message in console	PASS

CHAPTER 9 APPLICATION

1. Helps in observing the pilots psychological condition and face before take-off of the plane.
2. Can be performed on a psychological disorder patient by a psychiatric doctor.

3. In Social Welfare Gathering information would be profitable in the case of person who can’t talk and the one who can’t see.

4. It is used in Driver Monitoring in Monitoring driver facial expressions while driving. 5. Used as a lie detector.

CHAPTER 10 CONCLUSION

In this model, we propose a music recommendation system based on user emotions which is converted to speech from the text formatted output in the model. The human face is given as input, from which facial emotion is detected and the text formatted emotion output is converted to speech. Based on the speech output music is played automatically. Thus our proposed system will provide good level of accuracy in detecting emotions of a face and providing the speech and music formatted output