|  |  |
| --- | --- |
|  | **DAYANANDA SAGAR UNIVERSITY**  **KUDLU GATE, BANGALORE – 560068** |



**BACHELOR OF TECHNOLOGY**

In

**COMPUTER SCIENCE AND TECHNOLOGY**

**Capstone Project Phase-II Report**

**(Symphony)**

By

**Aishwarya Shetty- ENG19CT0001**

**Apeksha Prabhu- ENG19CT0006**

Under the supervision of

**Prof. Suparna**

Assistant Professor

Computer Science and Technology

**DEPARTMENT OF COMPUTER SCIENCE & TECHNOLOGY,**

**SCHOOL OF ENGINEERING**

**DAYANANDA SAGAR UNIVERSITY,   
Bangalore – 560068**

**(2021 – 2022)**

Logo

Description automatically generated with medium confidence

School of Engineering

Department of Computer Science & Technology

Kudlu Gate, Bangalore –560068

Karnataka, India

**CERTIFICATE**

This is to certify that the Phase-II project work titled **“Symphony”** is carried out by **Aishwarya Shetty (ENG19CT0001), Apeksha Prabhu (ENG19CT0006)** students of Bachelor of Technology in Computer Science and Technology at the School of Engineering, Dayananda Sagar University, Bangalore in partial fulfillment for the award of degree in **Bachelor of Technology** in **Computer Science and Technology**, during the year **2021-2022**.

|  |  |  |
| --- | --- | --- |
| **Prof Suparna** | **Dr. M Shahina Parveen** | **Dr. Uday Kumar Reddy** |
| Assistant/Associate/ Professor  Dept. of CS&T,  School of Engineering  Dayananda Sagar University  Date: | Chairman CST  School of Engineering  Dayananda Sagar University  Date: | Dean  School of Engineering  Dayananda Sagar University  Date: |

**Name of the Examiner** **Signature of Examiner**

1.

2.

<

**DECLARATION**

We, **Aishwarya Shetty (ENG19CT0001), Apeksha Prabhu (ENG19CT0006)** are students of the seventh semester B.Tech in **Computer Science and Technology**, at School of Engineering, **Dayananda Sagar University**, hereby declare that the phase-II project titled **“Symphony”** has been carried out by us and submitted in partial fulfilment for the award of degree in **Bachelor of Technology in Computer Science and Technology** during the academic year **2021‑2022**.

|  |  |
| --- | --- |
| **Students :** | **Signature :** |
| **Name1:** Aishwarya Shetty  **USN :** ENG19CT0001 |  |
| **Name2:** Apeksha Prabhu  **USN :** ENG19CT0006 |  |
| **Place : Bangalore**  **Date :** |  |
|  |  |
|  |  |

**ACKNOWLEDGEMENT**

It is a great pleasure for us to acknowledge the assistance and support of many individuals who have been responsible for the successful completion of this project work.

First, we take this opportunity to express our sincere gratitude to **School of Engineering & Technology**, **Dayananda Sagar University** for providing us with a great opportunity to pursue our Bachelor’s degree in this institution.

We would like to thank **Dr. Uday Kumar Reddy.** **Dean**, **School of Engineering & Technology**, **Dayananda Sagar University** for his constant encouragement and expert advice. It is a matter of immense pleasure to express our sincere thanks to **Dr. M Shahina Parveen,** **Department Chairperson**, **Computer Science, and Technology**, **Dayananda Sagar University,** for providing the right academic guidance that made our task possible.

We would like to thank our guide **Prof. Suparna**, **Assistant Professor**, **Dept. of Computer Science and Technology Engineering**, **Dayananda Sagar University**, for sparing her valuable time to extend help in every step of our project work, which paved the way for smooth progress and the fruitful culmination of the project.

We are also grateful to our family and friends who provided us with every requirement throughout the course. We would like to thank one and all who directly or indirectly helped us in the Project work.

**TABLE OF CONTENTS**

Page

CERTIFICATE .............................................................................................................. iii

DECLARATION ........................................................................................................... iv

ACKNOWLEDGEMENT ............................................................................................. v

LIST OF ABBREVIATIONS ........................................................................................ vi

LIST OF FIGURES ………........................................................................................... vii

ABSTRACT ............................................................................................................ viii

CHAPTER 1 INTRODUCTION………................................................................. 1

CHAPTER 2 PROBLEM DEFINITION ………..................................................... 5

CHAPTER 3 LITERATURE SURVEY................................................................... 5

CHAPTER 4 PROJECT DESCRIPTION.................................................................... 5

4.1. PRELIMINARIES…………………………………………………………. 5

CHAPTER 5 REQUIREMENTS ................................................................................ 6

5.1. FUNCTIONAL REQUIREMENTS ............................................................ 6

5.2. NON FUNCTIONAL REQUIREMENTS.................................................... 6

CHAPTER 6 METHODOLOGY AND IMPLEMENTATION.................................. 7

CHAPTER 7 TESTING AND CONCLUSION ........................................................... 9

REFERENCES... .......................................................................................................... 10

**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| AI | Artificial Intelligence |
| ML | Machine Learning |
| CV | Computer Vision |
| OpenCV | Open - Source Computer Vision Library |
| KDEF | Karolinska Directed Emotional Faces |
| VGG | Visual Geometry Group |
| CNN | Convolutional Neural Network |
| CK | Cohn Kanade |
| CK+ | Extended Cohn - Kanade |
| RaFD | Radboud Faces Database |
| MIREX | Music Information Retrieval Evaluation eXchange |
| ANN | Artificial Neural Networks |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| Fig. No. | Description of the figure | Page No. |
| 3.1 | Proposed Block Diagram | 5 |

**Abstract**

A user's facial expressions can reveal his or her level of emotion. These expressions can be obtained from the system's camera's live feed. In the area of computer vision (CV) and machine learning (ML), a lot of research is being done to train machines to recognize different human emotions or moods. Machine learning offers a variety of methods for detecting human emotions.

A review of existing music systems revealed that many music applications rely on the user's past listening choices rather than recommending songs based on their current emotion. The goal of this project is to identify emotions in human faces using real-time data and to suggest songs according on those emotions.

Music is a great unifier. It binds us despite our differences in ages, backgrounds, languages, interests and levels of income. Due to its accessibility and ability to be used alongside daily activities, travel, sports, and other activities, music players and other streaming apps are in high demand. Digital music has emerged as the main form of consumer content that many young people are looking for because to the quick growth of mobile networks and digital multimedia technology.

Music is frequently used by people as a tool for mood control, specifically to improve mood, boost energy, or soothe tension. Additionally, listening to the correct music at the right moment can help with mental wellness. So, music and feelings in people are closely related.

As a result, the proposed system is an interactive platform for suggesting music depending on user’s present emotional state. This also could be a great feature to be incorporated in existing music player applications.

Verified By:

Guide Name:

Date:

Signature:



1. **INTRODUCTION**

Numerous studies conducted over the past few years have shown that music has an impact on people's emotions, actions, and cognitive processes. One of the most crucial purposes of music, according to researchers studying the reasons people listen to music, is its relationship with our emotion. Two of music's most important applications are its ability to raise participants' levels of self-awareness and mood. It has been demonstrated that emotional states and personality traits are closely related to musical preferences.

A wide range of applications, including smart card applications, surveillance, image database investigation, criminal, video indexing, civilian applications, security, and adaptive human-computer interfaces in multimedia environments, use emotion detection as the most important technique at the moment.

Thanks to developments in digital signal processing and other effective feature extraction algorithms, automated emotion detection in multimedia attributes like music or movies is growing quickly. This system can play a significant role in many potential applications, such as music entertainment and human-computer interaction systems. We offer a technology for emotion identification that recommends music based on user moods as determined by their facial expressions. The proposed system can identify a person's emotions, and if the person is down, a playlist of the cheeriest, musically-related songs will be played.

Additionally, if the emotion is positive, a specific playlist of songs from different musical genres will be provided to reinforce the positive feelings.

Facial expression recognition is implemented using convolutional neural networks, which have a 95% accuracy rate.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Department of Computer Science & Technology, SOE, DSU Page **9** of **19**



**2. PROBLEM DEFINITION**

Most of the existing music applications do not take user’s facial expression as an input parameter to recommend songs. That drawback inspired us to develop an application that focuses on the same. The application we are developing will use **AI** to detect user's facial expression in **real-time** and recommend songs solely based on that.

Though we think that focusing on user’s emotion would make our application unique, we are aware that its not the only contributing factor involved in building a software product. There are a lot of corner cases to be taken care.

One major case is that not all users are expressive. Even if they are expressive, sometimes due to bad lighting, the application might not be able to detect the user’s face. In such cases the user can select the mood manually, just like how we select genre or artist in other widely used music applications.

The application works for four expressions as of now. Those are **happy**, **sad**, **angry** and **neutral**. These are the most common emotion expressed by people in real life. We are planning to add more if our application produce accurate results for the existing emotions.

With accuracy, it is also our priority to provide good user experience with attractive user interface.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Department of Computer Science & Technology, SOE, DSU Page **10** of **19**



**3. LITERATURE REVIEW**

1. In a particular system [1], Anaconda and Python 3.5 softwares were used to test the functionality and Viola-Jones and haar cascade algorithms were used for face detection. Similarly, KDEF (Karolinska Directed Emotional Faces) dataset and VGG (Visual Geometry Group) 16 were used with CNN (Convolution Neural Network) model which was designed with an **accuracy of 88%**, for face recognition and classification that validated the performance measures.

2. Another system [2] used Python 2.7, OpenSource Computer Vision Library (OpenCV) & CK (Cohn Kanade) and CK+ (Extended Cohn-Kanade) database which gave approximately **83% accuracy.**

3. It was observed in a cross-database experiment [3] that raw features worked best with Logistic Regression for testing RaFD (Radboud Faces Database) database and Mobile images dataset. The **accuracy achieved was 66% and 36%** respectively for both using CK+ dataset as a training set. The additional features (distance and area) reduced the accuracy of the experiment for SVM (Support Vector Machine) from 89%.

4. There has also been research done on the Music Recommendation System. According to one such research [4], a preliminary approach to Hindi music mood classification has been described, that exploits simple features extracted from the audio. MIREX (Music Information Retrieval Evaluation eXchange) mood taxonomy gave an **average accuracy of 51.56%** using the 10-fold cross validation.

5. Renuka R Londhe et al. [5] proposed a paper which focused on the study of changes in the curvatures of the face and the intensities of the corresponding pixels. The author used Artificial Neural Networks (ANN), which was used to classify the emotions. The author also proposed various approaches for a playlist.

6. Zheng et al. [6] proposed two significant categories for facial feature extraction, which included Appearance-based feature extraction and geometric based feature extraction, which included extraction of some essential points of the face such as mouth, eyes, and eyebrows.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Department of Computer Science & Technology, SOE, DSU Page **11 of 19**



**4. PROJECT DESCRIPTION**

The project being developed is a music application named **Symphony**. Symphony’s aim is to play music based on user’s facial expression and this feature makes symphony stand out when compared to existing music applications. 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.

The music application is in the form of a website developed using **Flask**, a web application framework written in Python. And for database, it uses **Sqlite3**. The database is used to store user account details such as email, name and password which are required fields.

On logging in, the user will be provided with two options, either to select his or her mood manually or to let the webcam detect it. After the application detects the user’s mood, it automatically plays the song based on the detected mood.

There are several techniques to detect face, but we will use the **haar cascade** algorithm because its not so complex and can run in real-time. Haar cascades were first introduced in 2001, and it was one of the most popular object detection algorithms in **OpenCV**.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Department of Computer Science & Technology, SOE, DSU Page **12** of **19**



**5. REQUIREMENTS**

**5.1 FUNCTIONAL REQUIREMENTS:**

The solution to the problem statement is to develop a web application. Preliminary steps in building the application:

1. Authentication (Login/Signup) using **Flask** and frontend using **HTML**, **CSS**, **JavaScript**.
2. Detect the user’s mood using the **Haar Cascade algorithm** through web cam if user choses real time mood detection otherwise the input is taken from the user.
3. The song is played on Youtube based on the user’s facial expression detected by the algorithm.

* **Software Requirements:**

1. Operating System: Windows/Linux
2. Python IDE
3. Python Latest version
4. Ci/Cd Tools - Docker, Azure
   1. Github Actions

* **Hardware Requirements:**
  1. A Desktop/Laptop with Stable Internet Connection

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Department of Computer Science & Technology, SOE, DSU Page **13** of **19**



**5.2 NON-FUNCTIONAL REQUIREMENTS:**

* **Usability** : The application shall be user-friendly and doesn't require any guidance. In other words, the application will be easy to learn and use.
* **Reliability**: The application will not have any unexpected failure. In order to avoid any failure occurrence, the specifications have been respected and followed correctly. Any component can be modified to correct faults, improve performance or other attributes, or adapt to a changing environment, because of the development methodology followed.
* **Security –** Pylint & CodeQL as the primary tool of choice to perform vulnerability checking

Access permissions for the particular system information may only be changed by the system’s data administrator.

* **Performance** - Monitor/test location coordinate data transmission latency.
* **Correctness of Output** - Hence maintainability is ensured by timely monitoring of correct working of the system.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Department of Computer Science & Technology, SOE, DSU Page **14** of **19**



**CHAPTER 6** **METHODOLOGY AND IMPLEMENTATION**

Real-time mood recognition is the main goal of the programme known as the mood-based music

recommendation system. It is a prototype for a brand-new product that has two primary modules:

music suggestion and facial expression recognition/mood detection.

1)Module for detecting mood ,The following two sections make up this module:

● Face Recognition: the capacity to locate faces in any input picture or frame. The

discovered faces' bounding box coordinates are the output. Initially, the Python library

OpenCV was taken into consideration for this purpose. However, integrating it with an

android app was a difficult issue, thus the Java FaceDetector class was taken into

consideration. This library counts the number of faces present in a given picture and

recognises the faces of individuals in a Bitmap visual object.

● Mood Recognition Emotions on the face are categorized as pleased, angry, sad, neutral,

surprised, fearful, or disgusted. The typical Keras module of Python was utilized for this

purpose, however it was discovered via the survey that this method works slowly when

linked with Android apps and requires a lot of time to train and evaluate. Therefore, the

CNN architectural model for image classification and mobile vision called MobileNet

was employed. Other models exist as well, but what sets MobileNet apart is how little

computing resources it needs to operate or use transfer learning on. This means that it

may be used without sacrificing the accuracy of the findings on mobile devices,

embedded systems, and computers with poor computing efficiency or GPUs.Lightweight

deep neural networks are constructed using depth-wise separable convolutions. The FER

2013 dataset and the MMA Facial Expression Recognition dataset from Kaggle were

combined to create the dataset used for training. The grayscale photos in the FER 2013

dataset were 4848 pixel images. Images with various characteristics might be found in the

MMA Facial Expression Recognition dataset. To get an even larger dataset with 40,045

training photos and 11,924 testing images, all of these images were transformed to match

the FER 2013 dataset's images. To train and test our model for seven classes—happy,

angry, neutral, sad, surprise, fear, and disgust—we used MobileNet and Keras. It was

trained for 25 epochs, and we got an accuracy of about 75%.

2)Music Recommendation Module

On Kaggle, a dataset of songs labeled according to mood was discovered for Hindi and English.

In order to save, retrieve, and query this song's data as needed by users, research was done to

find a reliable cloud storage platform. There were alternatives like AWS, Google Cloud, etc., but

they were rejected since they were expensive and offered very little storage for free. After then,

research was done on open-source streaming services like Restream.io, Ampache, etc., but once

more, these systems were web-based, utilized for live YouTube broadcasting, and were only

intended for private use. After much investigation (and due to time restrictions), Firebase was

decided to be a backend server. It only has one click to link with an Android app, and its free

plan offers 5GB of storage.However, because services like user inquiries, server upgrades, etc.

are only available as part of a subscription plan, it was decided to restrict the project's scope. The

songs' mp3 files were manually added to Firebase storage and connected in the Real Time

database according to the songs' mood and language (for filters).

3)Integration

The trained MobileNet model was saved as an.p file, which was then converted to a.tflite file

using TensorFlow Lite Converter in order to integrate these two modules in an Android

application. It accepts a TensorFlow model as input and produces an output with the.tflite

extension for a TensorFlow Lite model. The size of the tflite file is anticipated to be between 20

and 25 Megabytes (MB), which was the targeted amount since the MobileNet model is

employed. The.tflite file and labels.txt file were placed in an assets folder that was made in

Android Studio. The model's class labels are contained in the labels.txt file. The correct

procedures for loading the model, starting the interpreter, and getting the results were all

developed.

A Firebase project was made, and MP3 files were added to the storage area. These songs are

included in the real-time database area according to language and mood. After that, Android

Studio was connected to the Firebase database. The tflite model methods were connected with

the songs on Firebase, and an appropriate user interface for the Android application was made.

The programme was then examined for defects and, if any, fixed.



Fig:Displays the System Architecture

The system architecture diagram shows the interactions, limitations, and boundaries between

components as well as the general layout of the software system. The user's Android app will

open to the home screen, which has three buttons: take pictures, utilize emoji, and play music.

When the "take snap" button is clicked, the camera opens and the user snaps the photo. This

image is used as the input for a facial recognition algorithm. An appropriate error message is

presented to the user if no face is discovered or if numerous faces are detected. The image is sent

as input to the mood detection module after successful single face detection.

The user is shown the detected mood, and the "play tunes" button then becomes active.

The user can choose and play a song from the playlist that is appropriate for the identified mood,

as illustrated in Fig. A screen with five emojis will appear if the user clicks the "use emoji"

button, as seen in Fig. Any emoji may be clicked by the user to access the corresponding playlist.

The user only needs to touch the back button to leave the app.



Fig:Shows the data flow of the system

**CHAPTER 7 RESULTS AND CONCLUSION**

Accurately identifying human emotion or mood is challenging since each person has distinctive

facial traits. But it may be recognised to some extent with the right facial expressions.

The device's camera needs to have a greater resolution. Python-based recognition It takes a lot of

work to develop this system since several technologies must be combined with the programme.

The users and music enthusiasts of this system stand to benefit greatly. Even if this system may

still benefit from refinement, the primary goal of this project is to play music that corresponds to

a person's emotions, and as of now, that goal has been met to some extent. However, not every

task in this development sector can be stated to be flawless. The system was able to identify

happy, sad, angry, neutral, or shocked emotions. 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 use 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 tasks. We have picked up a lot of knowledge about the

subject of development and learnt a lot of new things. We anticipate success from this.

**REFERENCES**

[1] 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.

[2] 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 emotion-specified 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.

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

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

[5] Londhe RR and Pawar DV 2012 Analysis of facial expression and recognition based on

statistical approach International Journal of Soft Computing and Engineering 2.

[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] Ketki R. Kulkarni and Sahebrao B. Bagal ,“Facial expression recognition”, 2015 International

Conference on Information Processing (ICIP).

[8] V. R. Ghule, A. B. Benke, S. S. Jadhav, and S.A. Joshi, “Emotion Based Music Player Using

Facial Recognition,” vol. 5, no. 2, p. 7, 2007. International Journal of Innovative Research In

Computer and Communication, 2017.

[9] Gokul Krishnan K, Parthasarathy M, Sasidhar D and Venitha E, "Emotion detection and music

recommendation system using machine learning" in 2018 International Journal of Pure and

Applied Mathematics, 2018, vol. 119, pp. 1487-1498.

[10]Shlok Gilda, Husain Zafar, Chintan Soni and Kshitija Waghurdekar, “Smart Music Player

Integrating Facial Emotion Recognition and Music Mood Recommendation”, presented at the

IEEE WiSPNET 2017 conference.