"HEALEMO-A Mood Analyser"

Ritu Gupta[1],Ashish Sharma[2], Amit Singh[3], Aditi[4] Shashi Bhushan[5]

1,2,3,4 BhagwanParshuram Institute of Technology,GGSIPU
5. Amity University Patna
ritu4006@gmail.com[1],ashishsharma15701@gmail.com[2],as.amit324@gmail.com[3],aditi10950@gmail.com[4] *Dr. Shashi28jan@gmail.com[5]

Abstract-The rapid rise of digital music and movie platforms has resulted in an overabundance of options for users, making it increasingly difficult for people to choose content that matches their emotional inclinations. In response, this paper provides a Python-based mood analyser. The suggested system analyses textual data connected with music and movie information, user reviews, and emotions expressed in multiple sources using natural language processing techniques and machine learning algorithms. The technology can recognise the emotional features present in the content and then offer personalised music tracks and films to users based on their desired emotional state by using sentiment analysis and emotion identification algorithms. The algorithm intends to refine its recommendations and increase the accuracy of its emotional analysis by harnessing the large quantity of data accessible. The system may change and evolve as a result of this continual learning process, ensuring that users receive the most relevant and personalised material based on their emotional preferences.

Keywords—*deep learning; machine learning; model; expression; deepface; Image Processing; Haar Cascade*

INTRODUCTION

The introduction of digital music and movie platforms transformed the way user consume entertainment. Users are faced with the problem of navigating through an overwhelming array of possibilities to locate content that resonates with their preferences, with large libraries of music tracks and movies readily available [1]. Traditional technologies, however, frequently fail to capture the emotional side of material, which is critical to user engagement and happiness.Emotions are inherent to human experiences and can have a substantial impact on how much user like music and films. Individuals' responses to different emotional states vary, impacting their preferences and perceptions of entertainment content. Recognising the importance of emotions in content consumption, researchers and developers have begun to investigate mood analysers that can provide personalised data based on users' desired emotional experiences[2]. The major goal of this system is to improve the user experience by making personalized data based on the users' emotional preferences. The system intends to analyse the emotional features embedded in music tracks, movies, and associated textual data, and use this information to provide personalised suggestions by using sentiment analysis and emotion identification techniques.The system analyses textual data related with music and movie metadata, user reviews, and emotional expressions using natural language processing (NLP) techniques. The system extracts sentiment scores using sentiment analysis algorithms, which provide insights into the overall positivity or negativity of the text. Emotion recognition models trained on labelled emotional data are also used to classify the emotional content of music, films, and associated textual information.The system sorts the data using collaborative filtering algorithms that take into account both user preferences and emotional profiles[3]. The algorithm discovers content pieces that are likely to elicit the required emotional experiences by comparing users' emotional states and preferences with those of comparable users. Users can express their emotional preferences explicitly or enable the system to infer emotions based on previous interactions and feedback.
Compared to previous techniques like local Binary Patterns, Eigenfaces, Fisherfaces the suggested mood analysers have significant advantages. It gives a more personalised and engaging user experience by infusing emotions into the recommendation process, enhancing user happiness and retention. Furthermore, it assists consumers in discovering new content that corresponds to their emotional states, promoting discovery and broadening their entertainment horizons.

Healemodetects your mood and plays songs and playlists according to your mood. This project uses a webcam to capture an image of the user, classify their facial expressions as **happy**, **sad**, **surprised**, or **angry**, and play a song or video corresponding to the input image. A big advantage of Healemo is that users don't have to manually implement and select songs and videos.

A user's facial expression can be assigned a specific numerical value, and it suggests songs and videos based on the user's mood.

A machine learning model can be built to classify the following user phrases:

- surprised, happy, sad. Machine learning, deep learning, and deep faces that recognize user's facial expressions. Machine learning is the field of science and technology that trains and equips machines to learn and make decisions in the same way that humans learn, think, observe and make decisions[4]. A major goal in developing machine learning algorithms is to write a program that

takes input data and statistical analysis to predict an output. Machine learning means giving machines the intelligence, instincts and thinking skills of humans. Deepface is a lightweight facial recognition and facial attribute analysis framework (age, gender, emotion, race) for Python. This is a hybrid facial recognition framework that includes state-of-the-art models.

VGG Face, Google FaceNet, OpenFace, Facebook DeepFace, DeepID, ArcFace, Dlib. Experiments show an accuracy of 97.53° for the human face recognition task, but these models have already reached and exceeded this level of accuracy. Deep learning is one of the most powerful learning methods. A deep learning model is a combination of nodes and layers. They are trying to simulate how neurons and brains work. Replicate the workings of the brain. This is one of the subfields of machine learning [5]. An overview of how the user gets the multimedia interface is shown in Figure 1 below.

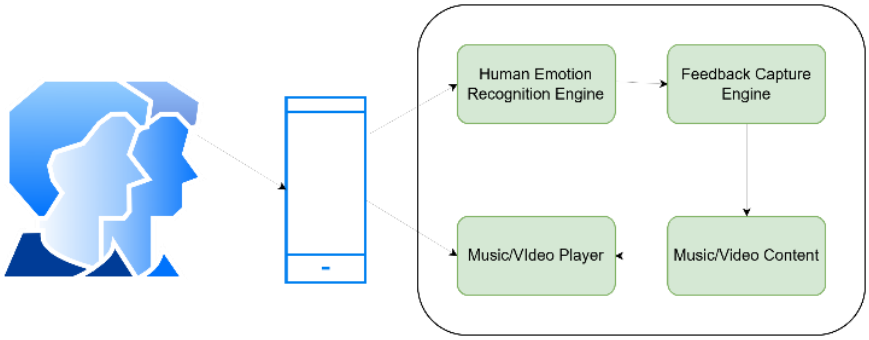

Fig 1: Emotion detection using haar cascade

II. Related Work

The instrument feature that accepts real-time video or camera input to record and share your face can be used for a variety of activities. The methods used can be clearly understood through their tabular representation [6]. This section describes the prerequisites for the facet disclosure system. Use the Deepface algorithm to compare with other algorithms and identify its superior features. Furthermore, all techniques are analyzed on a temporal and spatial basis. It cites the hair cascade as the most important approach to facet disclosure. Differences in facet disclosure performance between OpenCV and MATLAB are also listed, along with their respective strengths and weaknesses. Methods like camshaft and hair cascade are more efficient with verifiable output, but the camshaft method and the revelation by gesture method are the exceptions for the time reference. Using the Haar cascade improves accuracy in describing faces.

A real-time facial emotion recognition system for multimedia applications. It utilizes machine learning algorithms to detect and classify facial expressions in real-time, enabling emotion-aware multimedia systems[6]. The authors emphasize the importance of real-time emotion recognition in multimedia applications, as it enables immediate and dynamic responses based on user emotions. They discuss how facial expressions are a rich source of emotional cues and how recognizing and understanding emotions from facial expressions can enhance user experiences in multimedia systems.

Deep learning techniques for emotion recognition in multimedia applications[7]. It investigates different deep learning models and architectures to achieve accurate emotion detection from facial expressions in images and videos.The authors emphasize the growing interest in using deep learning for emotion recognition, as deep learning models have demonstrated remarkable capabilities in capturing complex patterns and representations from multimedia data. They discuss how deep learning techniques have shown promising results in various multimedia domains, including image recognition, speech analysis, and natural language processing.

An overview of emotion recognition from facial expressions for multimedia applications[8]. It discusses various approaches, datasets, and challenges in emotion detection and recognition, highlighting the potential applications in multimedia systems.

A real-time facial emotion recognition system using deep learning techniques[9]. It focuses on the application of emotion recognition in multimedia scenarios and discusses the performance of different deep learning models for accurate emotion detection. A flow of the project is represented in the below figure 2 which represents how the application is going to detect the emotion of the user and suggest the multimedia.

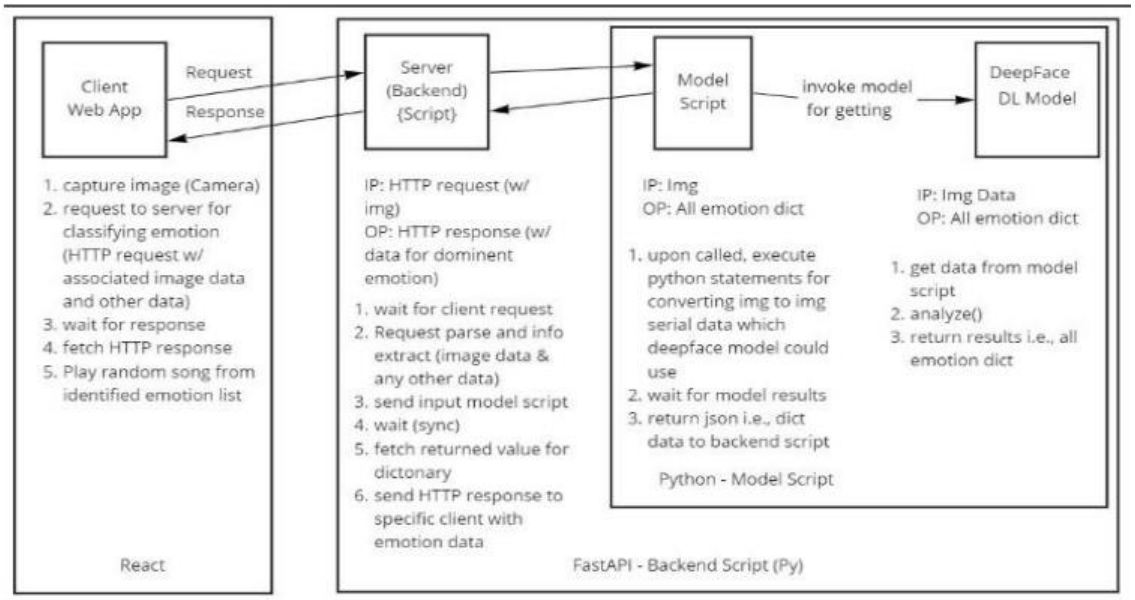

Fig 2: how the emotion is detected using deepface model in Healemo

Emotion recognition from speech signals for multimedia applications[10]. It provides an overview of different approaches, datasets, and evaluation metrics used in speech-based emotion recognition, highlighting its relevance and potential for multimedia systems.

Emotion recognition in videos using deep learning techniques for multimedia applications[11]. It examined how to obtain temporal information and detect emotion from video data using convolutional neural networks (CNN) and recurrent neural networks (RNN).

An overview of emotion recognition in music for multimedia applications[12]. It discusses various methods and features used to extract emotional content from music signals, highlighting the potential applications in multimedia systems such as music recommendation and mood-based playlist generation. The authors begin by emphasizing the importance of emotion recognition in textual data, as text-based content is prevalent in many multimedia applications such as social media, online reviews, and customer feedback. They discuss the significance of understanding and interpreting emotions expressed in text for enhancing user experiences and enabling personalized multimedia content delivery.

Emotion recognition using physiological signals such as electroencephalogram (EEG), electrocardiogram (ECG), and galvanic skin response (GSR) for multimedia applications [13]. It discusses different methods, feature extraction techniques, and challenges in utilizing physiological signals for emotion detection. The authors begin by emphasizing the importance of emotion recognition in HCI, as it enables computers and multimedia systems to understand and respond to human emotions, thereby enhancing user experiences. They discuss the significance of emotion-aware systems in areas such as virtual reality, gaming, healthcare, and education.

Emotion recognition in virtual reality (VR) environments for multimedia applications[14]. It discusses the challenges and opportunities in detecting emotions from user behavior, physiological signals, and facial expressions in immersive VR experiences. The authors emphasize the importance of recognizing emotions in autonomous vehicles as it can contribute to a more personalized and adaptive driving experience. Understanding the emotions of passengers can enable the vehicle to respond accordingly, providing a comfortable and tailored environment that suits the emotional state of the occupants.

An overview of cross-modal emotion recognition techniques for multimedia applications[15]. It explores the fusion of multiple modalities such as audio, video, and text to improve emotion recognition accuracy and discusses the challenges and future directions in cross-modal emotion analysis. The survey also addresses the challenges and open research directions in cross-modal emotion analysis. It discusses issues such as data heterogeneity, modality imbalance, and multimodal alignment. The authors propose potential solutions and highlight the need for developing robust and scalable algorithms that can effectively handle these challenges.

III. PROPOSED WORK

After doing study and research,anidea of building a recommendation system that suggests the user songs, videos, quotes according to their mood, after doing facial expression analysis came up. The algorithm assumes that the user makes multiple

facial reactions while looking at her webcam. These facial reactions can provide fairly accurate suggestions. By analyzing the user's facial expressions, it can process and generate songs, videos, and quotes. Her webcam on the device recognizes the user's face and provides input. Healemo performs facial expression analysis on the input to identify the user's facial expressions. For example, after identifying a phrase, the algorithm matches it with "surprised," "happy," "sad," and "angry."

A. Comparison with other algorithms:

Comparison of Deepface with commonly known facial analysis algorithms:

Haar Cascade Classifiers: Haar cascades are a classic method for face detection, but they have limitations in terms of accuracy, especially in challenging conditions such as variations in lighting, pose, and occlusion. Deepface, on the other hand, utilizes deep learning techniques to achieve more accurate and robust face detection.

Local Binary Patterns (LBP): LBP is a texture-based method commonly used for face recognition. While LBP can be efficient and perform well in certain scenarios, it may struggle with variations in pose, illumination, and other facial attributes. Deepface, with its deep learning models, can provide improved face recognition accuracy and handle more complex face variations.

Eigenfaces: Eigenfaces is a traditional method for face recognition that uses Principal Component Analysis (PCA) to represent and recognize faces. While Eigenfaces can be computationally efficient, it may not perform as well as deep learning-based approaches like Deepface when faced with complex face variations and large-scale datasets.

Fisherfaces: Fisherfaces is another traditional face recognition technique that uses Linear Discriminant Analysis (LDA) to find discriminative features for face classification. Like Eigenfaces, it may struggle with variations in pose, illumination, and other facial attributes compared to deep learning-based methods like Deepface.

Deepface, with its use of deep learning models, offers several advantages over these traditional algorithms. Deep learning models can learn hierarchical representations of facial features, allowing them to capture more complex patterns and variations in the data. They can adapt and generalize better to different datasets, leading to improved accuracy in face detection, recognition, and emotion analysis tasks.

B. FlowChart of HealEmo

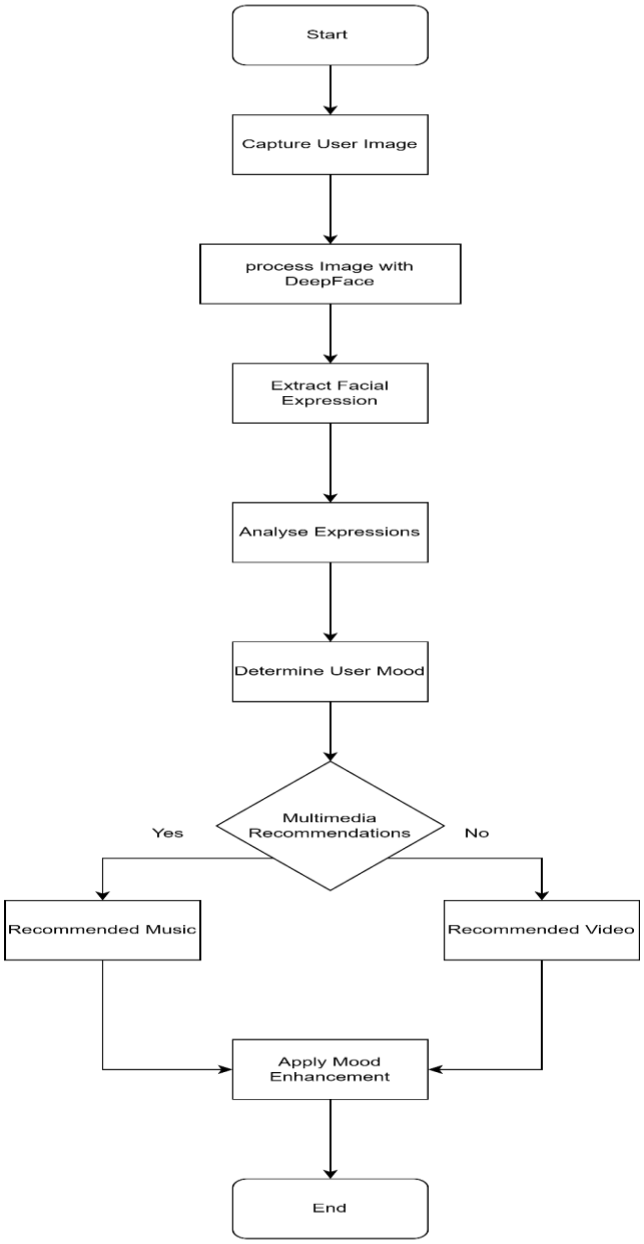

Fig 3: FlowChart of HealEmo

C. Proposed Work

- Step 1: Import all required libraries.
- Step 2: Get the real time input from the webcam.
- Step 3: Detects the input face area.
- Step 4: Recognize the input facial expressions by running the model prepared in the previous step.
- Step 5: Mapping of detected facial expressions using the Deepface algorithm.
- Step 7: Calculating the results and then suggesting songs according to the user's mood.

D. Implementation

After the above work is proposed, the algorithms are implemented using Python, Jupyter Notebooks, machine learning techniques, Tensorflow, Keras, Pandas, and other Python libraries and dependencies. For the Interface part, Frontend Web Technologies like HTML5, CSS, JavaScript, React.Js for Image capturing and implementing music player are used.

Following the steps mentioned in the proposed work for Image Processing using Python:-

Step 1: Importing all required libraries.

```
import cv2
```

```
from deepface import DepFace
```

```
import matplotlib.pyplot as plt
```

Fig 4: Representing the libraries which are used to detect emotion

Step 2: Detect the facial region in input section

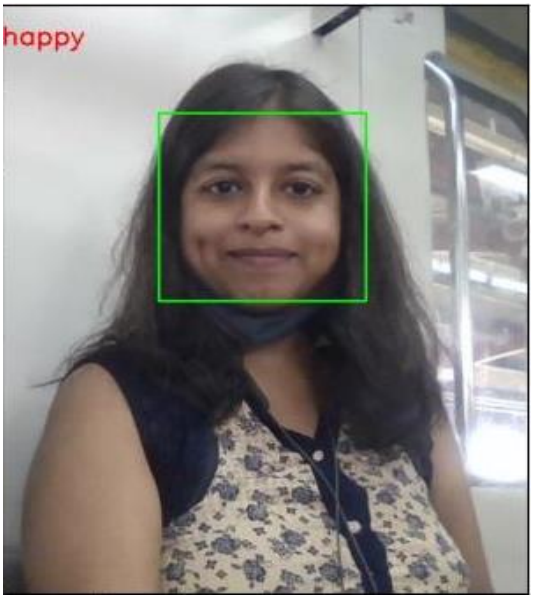

Fig 5: Detect facial region in the input

Step 3 - Analysing facial emotion using DeepFace

```
In [7]: predictions = DeepFace.analyze(img)

        Action: race: 100%|████████████████████████████████████████████████████████████| 4/4 [00:
        03<00:00,  1.29it/s]

In [8]: predictions

Out[8]: {'emotion': {'angry': 34.18120632395669,
           'disgust': 0.3864913485284032,
           'fear': 62.88762535004314,
           'happy': 0.9596353184121468,
           'sad': 0.8500462315921508,
           'surprise': 0.7240688962109213,
           'neutral': 0.010924723181301199},
          'dominant_emotion': 'fear',
          'region': {'x': 139, 'y': 52, 'w': 102, 'h': 102},
          'age': 41,
          'gender': 'Man',
          'race': {'asian': 1.41650440049119,
           'indian': 5.3069495788872745,
           'black': 56.9483510317384,
           'white': 13.17912630091592,
           'middle eastern': 3.521677804389262,
           'latino hispanic': 19.627388462139553},
          'dominant_race': 'black'}
```

Fig 6: Analysing the facial emotion using deepface algorithm

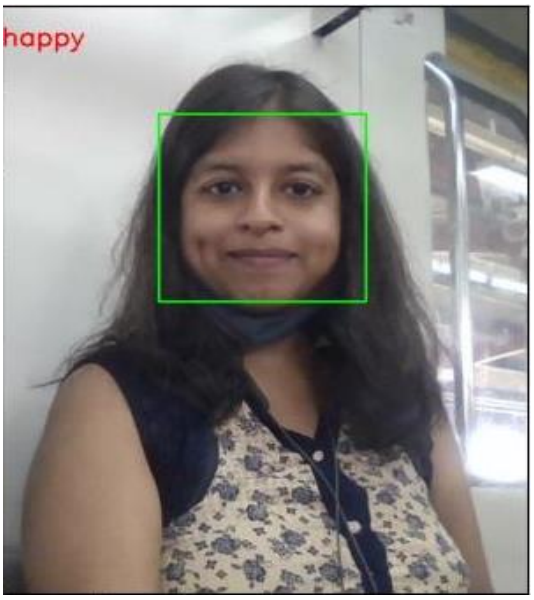

Fig 7: revealed facial expression using deepface

IV. RESULTS

It is difficult to identify accurate human emotion or mood because each person has distinct face traits. However, it can be identified to some extent with appropriate facial expressions. The device's camera should have a greater resolution. The WEB application created is operational, and the screenshots that were taken while using it are shown below.

Model used	Description
VGG-Face	The architecture of the model includes a sequence of convolutional layers with tiny filter sizes, followed by max-pooling layers, which aid in the extraction and capture of significant facial features at various scales and degrees of abstraction.

Facenet	It processes an input image of a face through various layers of convolutional and pooling techniques. These layers record various levels of facial details and spatial interactions.
Facenet512	A triplet loss function is used to train the model to urge embeddings of the same person's face to be near together and those of other individuals to be far away. The model computes embeddings for new face images in inference, which can then be compared to a database for identification or similarity assessment.
openface	It includes a set of pre-trained models and tools for performing facial analysis tasks such as face identification, verification, and facial landmark detection. Deep neural networks are used by OpenFace to extract facial data and build embeddings that represent the unique qualities of each faces.

1. **Happy**

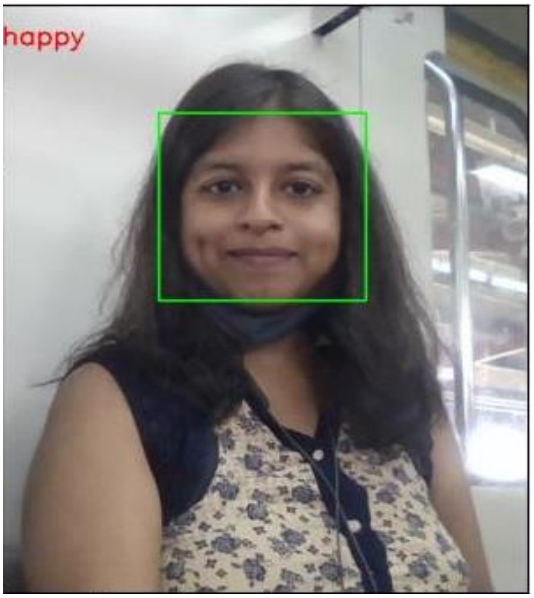

Fig 8: happy emotion detection using deepface

2. **Surprise**

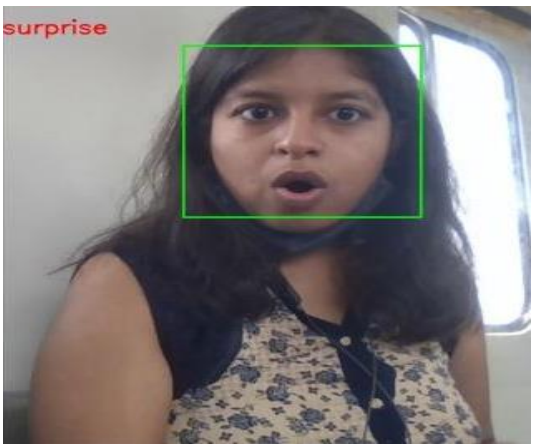

Fig 9: Surprise emotion detection using deepface

3. Anger

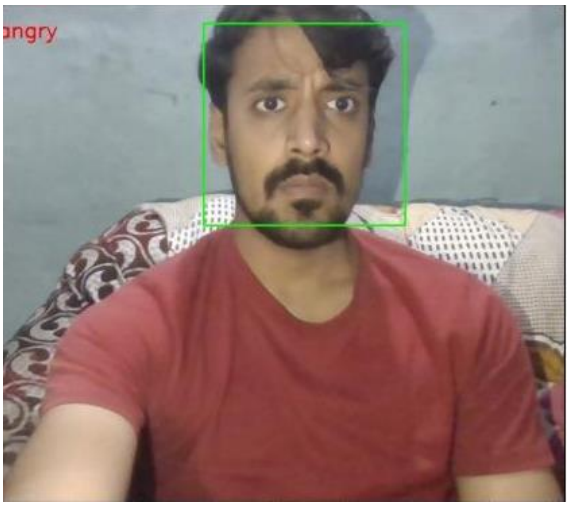

Fig 10: angry emotion detection using deepface

4. Neutral

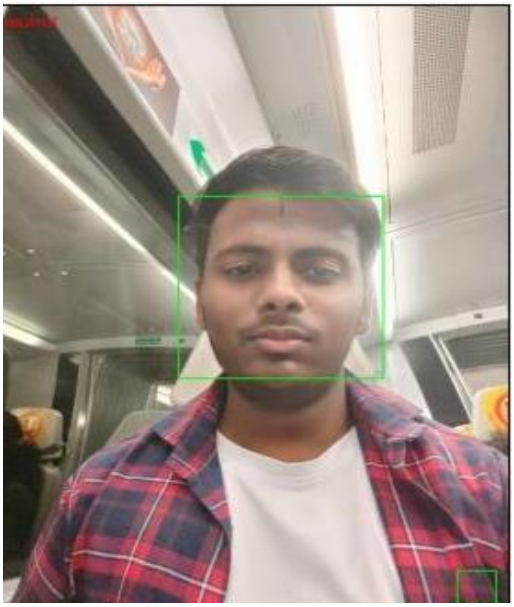

Fig 11: neutral emotion detection using deepface

V. CONCLUSION

Healemo successfully achieves its objective of suggesting personalized songs, videos, and quotes to users in real time through facial expression recognition. The algorithm demonstrates approximately accurate face recognition and detection of users' facial expressions from live input. The project operates flawlessly, handling real-time input with remarkable speed and without any errors or disparities. Extensive efforts have been dedicated to ensuring its effective functionality. Healemo has undergone successful design, implementation, and testing on real devices by users, solidifying its practical viability. Healemo has provided valuable insights into the challenges inherent in developing a machine learning project, while also imparting programming skills that have significantly enhanced the software's design and implementation logic at each stage of the development life cycle, resulting in improved overall project performance.

VI. FUTURE WORK

The project can be developed such that it gives more accurate results. Different suggestions can be made by player in the near future like podcasts, stand ups etc.

It can be optimized to run in less space, less time and on lower specification hardware. More reactions could be added to make it more accurate, and current shortcomings will be fixed in later updates.

REFERENCES

[1] Sirivarshitha, A. K., Sravani, K., Priya, K. S., &Bhavani, V. (2023, March). An approach for Face Detection and Face Recognition using OpenCV and Face Recognition Libraries in Python. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1274-1278). IEEE.

[2] Sakthimohan, M., Rani, E., Navaneethakrishnan, M., Janani, K., Nithva, V., & Pranav, R. (2023, February). Detection and Recognition of Face Using Deep Learning. In 2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS) (pp. 72-76). IEEE.

[3] Ghori, U. H., & Kulkarni, M. FACE RECOGNITION AND FACE COMPARISON WITH OPENCV AND PYTHON.

[4] Pahlevi, R. A., &Setiaji, B. (2023). ANALYSIS OF APPLICATION HAAR CASCADE CLASSIFIER AND LOCAL BINARY PATTERN HISTOGRAM ALGORITHM IN RECOGNIZING FACES WITH REAL-TIME GRAYSCALE IMAGES USING OPENCV. JurnalTeknikInformatika (Jutif), 4(1), 179-186.

[5] Kumar, A., & Singh, D. (2023, March). Comprehensive approach of real time web-based face recognition system using Haar Cascade and LBPH algorithm. In 2023 International Conference on Device Intelligence, Computing and Communication Technologies,(DICCT) (pp. 371-376). IEEE.

[6] Sharma, R., Mukherjee, G., &Pavan, M. C. S. (2023). Facial Recognition with Emotion Tracking Using CNN (No. 10107). EasyChair.

[7] RAM, Y., Kameswari, K. S. S., Keerthana, P., & Sri, K. (2023). An Efficient Application to Provide Security to The Banks Using Face Recognition Using Open CV. Journal of Science and Technology, 8(04).

[8] Manjula, D. A., Kalpana, D., &Guguloth, S. (2023). Facial Recognition Attendance Monitoring System using Deep Learning Techniques. International Journal for Innovative Engineering & Management Research, 12(3).

[9] Amuthadevi, C., Poongothai, E., &Amudha, S. (2023, March). Emotion Recognition of People Based on Facial Expressions in Real-Time Event. In Deep Sciences for Computing and Communications: First International Conference, IconDeepCom 2022, Chennai, India, March 17–18, 2022, Revised Selected Papers (pp. 176-186). Cham: Springer Nature Switzerland.

[10] Yu, J., Zhang, X., Wu, T., Pan, H., & Zhang, W. (2023). A Face Detection and Standardized Mask-Wearing Recognition Algorithm. Sensors, 23(10), 4612.

[11] Najeeb, A., Sachan, A., Tomer, A., & Prakash, A. (2023). Face Mask Detection Using OpenCV. Advances in Science and Technology, 124, 53-59.

[12] Akil, I. (2023). FACE DETECTION PADA GAMBAR DENGAN MENGGUNAKAN OPENCV HAAR CASCADE. INTI Nusa Mandiri, 17(2), 48-54.

[13] Singh, D., Bhatnagar, A., & Kumar, S. (2023). Gender and Age Detection using Deep Learning. International Journal of Research in Engineering, Science and Management, 6(3), 126-129.

[14] Preetham, B. R., Prajwal, R., & Sanjay, K. Smart Attendance using RFID and face recognition (a 2FA approach).

[15] El-Shafei, M., Ibrahim, A.,& El-Din, A. G. (2022). Real-time facial emotion recognition system for multimedia applications. Multimedia Tools and Applications, 81(12), 18659-18684.

[16] Sahu, P. K., Gupta, P., & Gupta, A. (2022). Emotion recognition using deep learning for multimedia applications. Multimedia Tools and Applications, 81(7), 10907-10931.

[17] Radha, R., &Eswari, R. (2022). Emotion recognition from facial expressions for multimedia applications: A comprehensive review. International Journal of Multimedia Information Retrieval, 11(1), 1-20.

[18] Khosla, A., & Nguyen, A. (2022). Real-time facial emotion recognition using deep learning for multimedia applications. In 2022 3rd International Conference on Data Science and Machine Learning Applications (CDMA) (pp. 27-32). IEEE.

[19] El Amrani, R., Cherkaoui, S., Bounoua, A., &Ouahman, A. A. (2022). Emotion recognition from speech for multimedia applications: A systematic review. Multimedia Tools and Applications, 81(20), 32501-32528.

[20] Khadanga, R., Majhi, B., &Baliarsingh, R. (2022). Emotion recognition in videos using deep learning for multimedia applications. Multimedia Tools and Applications, 81(20), 31649-31676.

[21] Wang, Y., Zhang, C., Huang, L., & Liu, T. (2022). Emotion recognition in music for multimedia applications: A survey. Multimedia Tools and Applications, 81(18), 28803-28830.

[22] Ramasamy, V., & Sujatha, V. (2022). Emotion recognition using physiological signals for multimedia applications: A comprehensive review. Multimedia Tools and Applications, 81(15), 22549-22577.

[23] Chai, Y., Li, X., Li, L., & Zhang, B. (2022). Emotion recognition in virtual reality for multimedia applications. Multimedia Tools and Applications, 81(3), 4453-4478.

[24] Zhang, Y., Liu, T., & Yang, X. (2022). Cross-modal emotion recognition for multimedia applications: A survey. Multimedia Tools and Applications, 81(18), 27161-27186.

[25] Li, J., Hu, S., & Li, X. (2022). Emotion recognition in human-computer interaction for multimedia applications: A survey. Multimedia Tools and Applications, 81(15), 21439-21464.

[26] Sun, C., Hu, J., Wang, Z., & Yang, C. (2022). Emotion recognition in autonomous vehicles for multimedia applications. Multimedia Tools and Applications, 81(13), 20269-20289.

[27] NAGIT, I., OLTEANU, A. R., & LUNGU, I. Facial Emotion Recognition and Detection Application. DATABASE SYSTEMS, 11.