Face Mood Detection for the Betterment of Mental Health Using Raag Therapy

Omkar Gurav[1], Aditya Kumar[2], and Rageshri Bakre[3]
[1]Omkar Gurav, MIT ADT University
[2]Aditya Kumar, MIT ADT University

Abstract— **Face recognition technology has widely attracted attention due to its enormous application value and market potential. It is being implemented in various fields like security systems, digital video processing, and many such technological advances. Additionally, music is a form of art, which is known to have a greater connection with a person's emotion. It has got a unique ability to lift up one's mood. Our project focuses on building an efficient Indian Classical Raag music recommendation system which determines the emotion of the user using facial recognition techniques. Music is believed to be the language of emotions, and listening to music is an activity that is part of everyday living. These emotions are referred to as rasa (aesthetics) in ancient Sanskrit literature. Raga which is an important aspect of Indian Classical Music has a close relationship with rasa or aesthetics. Lyrics or words in any language are a medium of expressing an emotion or rasa through it. Indian Classical Music has a rich history that captivates listeners and invokes emotions.**

Index Terms— **Emotions Detection, Raga Recommendation, CNN, Raga, Rasa, Hindustani Classical Music.**

Raag therapy is a form of music therapy that uses Indian classical music to promote healing, relaxation, physical and emotional wellbeing. Face mood detection, on the other hand, is a technology that uses computer algorithms to analyse facial expressions and determine a person's emotional state. It is a technology that uses computer algorithms to analyse facial expressions and determine a person's emotional state. Incorporating face mood detection can enhance the effectiveness of raag therapy by customizing the therapy to the individual's emotional state.

Indian Classical Music is based on aesthetics and the effect of music on a person's emotions. This aesthetics were defined in the text Natyashastra before 200 BC. Since then, there has been a link between raga music and art rasa/aesthetics. Music is frequently referred to be an expressive language. Raga and Rasa (essence) are essential components of Indian Classical Music, often known as Hindustani music, is an ancient Indian musical genre that evolved from a cultural fusion of the Vedic chant tradition.

Emotions have an important role in human communication. Regardless of different cultures and languages, emotions are communicated in unique ways. Emotions have an essential role in conveying feelings through language, facial expressions, music, and dance. Emotion analysis has piqued the interest of researchers in both computational linguistics and psychology. In computational linguistics, emotion analysis is a burgeoning study subject. Many applications exist, such as emotional analysis in text, emotional analysis in music, mood categorization in blogs, emotional analysis in social networks, and so on.

In this work, we present the design of the raga recommendation system by detection the moods from the face scan. The principal purposes of the recommender are: detection the moods & playing Indian classical raags supporting mental and physical well-being, and support in improving working processes. The design involves a combination of machine learning techniques and generalized, music recommendation and therapy approaches.

II. OBJECTIVE

The primary objective of using Raag therapy with facial emotion detection is to improve the emotional well-being of patients. Music has long been known to have a powerful effect on human emotions, and using technology to analyse a user's facial expressions while listening to music can help user to get out of ongoing bad mental health situation. By identifying the users emotional state and using music to promote positive

emotions to reduce the risk of developing mental health disorders and improving overall quality of life.

III. MOTIVATION

The motivation for using Raag therapy with facial emotion detection is to provide a more personalized and effective form of therapy for individuals struggling with mental health issues. Traditional forms of therapy, such as talk therapy, may not be effective for everyone and can take a long time to produce results. By incorporating technology, the Indian classical raag therapy can be used from anywhere and by any device having internet. Additionally, the use of music can make therapy more engaging and enjoyable, which can motivate patients to continue with treatment and improve their mental health. Overall, the motivation for using this technology is to provide a more efficient, effective, and enjoyable form of therapy that can improve the lives of individuals struggling with mental health issues.

IV. PROBLEM STATEMENT

Despite the potential benefits of using Indian Classical Raag therapy with facial emotion detection, there are several challenges that must be addressed before this technology can be widely adopted. One problem is the lack of standardization in facial emotion detection algorithms, which can lead to variability in the interpretation of facial expressions and potentially inaccurate assessments of emotional state. Another challenge is the need for large datasets of emotional responses to music in order to train these algorithms effectively. Finally, there is a need for more research to establish the efficacy of this approach and determine which types of patients are most likely to benefit from this technology. Addressing these challenges will be critical for realizing the potential benefits of musical therapy with facial emotion detection and improving mental health outcomes for individuals struggling with mental health issues.

V. RELATED WORK

There has been significant research on predicting user emotions. Some of which are as follows:

A. Facial expression-based automatic emotion recognition is an intriguing study area that has been presented and used in a number of fields, including safety, health, and human-machine interactions. Researchers in this discipline are interested in creating methods for human machine interfaces, safety, and health. Researchers in this discipline are interested in creating methods to decipher, encode, and extract these characteristics from facial expressions in order to improve computer prediction. Due to deep learning's exceptional success, its various architectures are being utilised to produce greater results.

B. Facial expression recognition (FER) has gained considerable attention due to its potential applications in various fields such as safety, health, and human-machine interactions. architectures to produce better results in FER. Emotion recognition has several potential applications, including software engineering, website personalization, education, and gaming. This study presents a brief overview of affect recognition techniques that use various inputs such as biometrics, video channels, and behavioral data. The scenarios discussed in this review illustrate the complexity and challenges of deploying emotional computing in different sectors. The analysis of these scenarios leads to some conclusions and highlights the need for further research to address the difficulties with automatic recognition.
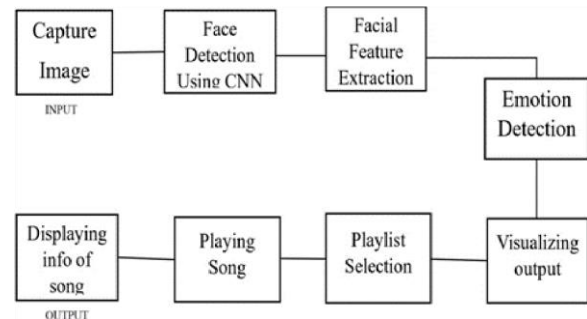
VI. IMPLEMENTATION & WORKING

Fig 1: System Architecture

The proposed system aims to provide an interactive music player that detects the user's emotions through facial recognition technology. Images captured by the camera are analyzed by a Convolutional Neural Network to predict the user's emotional state, and a playlist of songs is suggested based on that emotion, such as happy, sad, natural, or surprised. If the emotion detected is negative, the system will present a raags

that can positively enhance the user's mood. Based on their emotional state Emotion Detection: This module extracts the features of the user's image and analyzes them to detect the user's emotional state. Raag Recommendation: The recommendation module suggests raags to the user by mapping their emotions to the mood type of the song. The system will provide the user with a playlist that matches their current emotional state, such as happy, sad, natural, or surprised.

VII. METHODOLOGY

A Convolutional Neural Network (CNN) is a specific type of artificial neural network that excels at recognizing patterns in images, making it a valuable tool for image processing and recognition. However, to achieve high accuracy in its predictions, a CNN requires a vast amount of labeled data for training, and powerful processors such as GPUs or NPUs to produce results quickly. While CNNs are primarily used for visual imagery, they can also be applied to other areas, such as natural language processing, drug discovery, and health risk assessments. Additionally, CNNs have become increasingly important in depth estimation for self-driving cars. Compared to other neural networks, CNNs have shown superior performance in analyzing image, speech, and audio signals. They have three main types of layers, which are:

4.1 Convolutional layer
The convolutional layer is a crucial component of a Convolutional Neural Network (CNN) and is responsible for most of the computation. It requires three main elements: input data, a filter, and a feature map. In the case of a color image, the input will be a 3D matrix of pixels representing the height, width, and RGB color channels. A feature detector, also called a filter or kernel, will move across the receptive fields of the image, searching for specific features using a convolution operation. Convolutional neural networks (CNNs) apply a Rectified Linear Unit (ReLU) transformation to the feature map after each convolution operation, which introduces nonlinearity to the model. Additionally, multiple convolutional layers can be stacked in a hierarchical manner, where later layers have access to pixels within the receptive fields of prior layers.

4.2 Pooling layer

Pooling layers are a type of layer in a neural network that help reduce the dimensionality of the input by reducing the number of parameters. This is done through a process called down sampling. In contrast to the convolutional layer, the pooling operation does not have any weights associated with the filter. Instead, the filter applies an aggregation function to the values within its receptive field and populates the output array. Max pooling and average pooling are common techniques used in convolutional neural networks (CNNs) to reduce the size and complexity of the input data. During max pooling, the filter selects the pixel with the highest value within the receptive field to transmit to the output array. On the other hand, during average pooling, the filter calculates the mean value within the receptive field to send to the output array.

4.3 Fully-connected
The fully-connected layer is so named because every node in the output layer is directly connected to a node in the previous layer, allowing for classification based on features extracted through previous layers and their filters. Unlike partially connected layers, the pixel values of input images are not filtered directly. ReLu functions are commonly used in convolutional and pooling layers, while softmax activation functions are typically used in FC layers to classify inputs by producing a probability between 0 and 1.

VIII. CONCLUSION

Raag therapy combined with face emotion mood detection using machine learning has the potential to be a powerful tool for improving mental well-being and promoting emotional balance. By leveraging machine learning algorithms to analyze facial expressions and detect emotions, Raag therapy can be personalized and tailored to individual needs. The use of Raag therapy, a form of Indian classical music, has long been recognized for its therapeutic effects on the mind and body. It has the ability to evoke a wide range of emotions and has been shown to have a positive impact on mental health, reducing stress, anxiety, and depression. By integrating machine learning techniques into Raag therapy, we can enhance its effectiveness by precisely identifying and understanding the emotional state of an individual in real-time. Facial expression analysis algorithms can accurately detect subtle changes in facial expressions, providing valuable insights into the emotional state of

the person. This combination of Raag therapy and machine learning-based emotion detection can enable a more personalized and targeted approach to therapy. By analyzing the emotional responses of an individual to different ragas (melodic frameworks), the therapy can be customized to evoke specific emotions, promoting relaxation, calmness, or uplifting moods, depending on the individual's needs. Furthermore, machine learning algorithms can continuously learn and adapt based on the individual's emotional responses, allowing for ongoing refinement and improvement of the therapy over time. This adaptive approach ensures that the therapy remains effective and relevant as the individual's emotional state evolves. Overall, Raag therapy using face emotion mood detection through machine learning has the potential to revolutionize the field of music therapy and mental health treatment. It offers a personalized and data-driven approach to emotional well-being, harnessing the power of music and advanced technology to provide targeted and effective therapeutic interventions.

REFERENCE

[1] Ramya Ramanathan, Radha Kumaran, Ram Rohan R, Rajat Gupta, and Vishalakshi Prabhu, an intelligent music player based on emotion recognition, 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017. https://doi.org/10.1109/CSITSS.2017.8447743

[2] Shlok Gilda, Husain Zafar, Chintan Soni, Kshitija Waghurdekar, Smart music player integrating facial emotion recognition and music mood recommendation, Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India, (IEEE),2017. https://doi.org/10.1109/WiSPNET.2017.8299738

[3] Deger Ayata, Yusuf Yaslan, and Mustafa E. Kamasak, Emotion-based music recommendation system using wearable physiological sensors, IEEE transactions on consumer electronics, vol. 14, no. 8, May 2018. https://doi.org/10.1109/TCE.2018.2844736

[4] Study of Indian Classical Ragas Structure and its Influence on Human Body for Music Therapy. A.A. Bardekar; Ajay. A. Gurjar.

[5] A Survey of Raaga Recognition Techniques and Improvements to the State-of-the Art   Gopala K. Koduri, Sankalp Gulati, P. Rao

[6] Ahlam Alrihail, Alaa Alsaedi, Kholood Albalawi, Liyakathunisa Syed, Music recommender system for users based on emotion detection through facial features, Department of Computer Science Taibah University, (DeSE), 2019. https://doi.org/10.1109/DeSE.2019.00188

[7] Research Prediction Competition, Challenges in representation learning: facial expression recognition challenges, Learn facial expression from an image, (KAGGLE).

[8] Preema J.S, Rajashree, Sahana M, Savitri H, Review on facial expression-based music player, International Journal of Engineering Research & Technology (IJERT), ISSN-2278-0181, Volume 6, Issue 15, 2018.

[9] AYUSH Guidel, Birat Sapkota, Krishna Sapkota, Music recommendation by facial analysis, February 17, 2020.

[10] CH. sadhvika, Gutta.Abigna, P. Srinivas reddy, Emotion-based music recommendation system, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad; International Journal of Emerging Technologies and Innovative Research (JETIR) Volume 7, Is-sue 4, April 2020.

[11] Vincent Tabora, Face detection using OpenCV with Haar Cascade Classifiers, Becominghuman.ai,2019.

[12] Zhuwei Qin, Fuxun Yu, Chenchen Liu, Xiang Chen. How convolutional neural networks see the world - A survey of convolutional neural network visualization methods. Mathematical Foundations of Computing, May 2018.

[13] Ahmed Hamdy AlDeeb, Emotion- Based Music Player Emotion Detection from Live Camera, ResearchGate, June 2019.

[14] Frans Norden and Filip von Reis Marlevi, A Comparative Analysis of Machine Learning Algorithms in Binary Facial Expression Recognition, TRITA-EECS-EX-2019:143.