International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 05 | May 2024 www.irjet.net p-ISSN: 2395-0072

SoulSound: Enhancing Musical Therapy through Facial Expression Recognition with Machine Learning

Padmavati K[1], Amruth Kumar D K[2], Ganavi S M[3], Keerthi K C[4], Sourabha Y N[5],

[1]*Assistant Professor, Dept. of Computer Science & Engineering, Mandya, Karnataka, India*
[2,3,4,5]*Student, Dept. of Computer Science & Engineering, GMIT, Mandya, Karnataka, India*

Abstract - Musical therapy is a type of treatment that uses music to help people improve their physical, emotional, cognitive, and social functioning. It is a non-invasive treatment that has been shown to have beneficial effects on mental health. Facial recognition technology has grown in prominence as a tool for analyzing human emotions and behavior's in recent years. We are investigating the possibility of facial recognition technology in musical therapy. In this project, we propose approach to musical therapy that uses facial recognition technology to create personalized playlists for individuals based on their emotional state. We use a convolution neural network (CNN) to analyze facial expressions and classify them into different emotional states, such as happiness, sadness, or anger. Based on the emotional state detected, we recommend music that is likely to have a positive effect on the individual's emotional well-being. A voice assistant is also integrated into the system to provide additional support and guidance to the users. The system is designed to be accessible to people of all ages and abilities, including those with disabilities. Our approach has the potential to improve the effectiveness of musical therapy by providing personalized recommendations based on the individual's emotional state. Moreover, our project provides a convenient and accessible platform for individuals to receive musical therapy without the need for in-person visits to a therapist.

***Key Words*: Musical Therapy using Facial expression, Emotion Based Music Detection, Terapeutic Music Intervention, Music Driven Facial Analysis, Mood Enhancement.**

1. INTRODUCTION

The humans being has natural intelligences to watch a face and guess their emotional state. This intelligences, if copied by an electronic computer - desktop, robotic automaton or a movable machinery - will have useful applications in the real world. Music is seen as a good medium of emotional communique. Emotional expression is seen as an important treatment idea. In every form and technical literature, it is linked to mental welfare when inhibition of feelings seems to have a role in different diseases, along with physical sicknesses.

Using traditional music therapies, where the therapist previously needed to manually analyze the patient's mood using various observing techniques that didn't involve any computer machinery or specific intelligent systems. After determining the patient's mood, the therapist picks songs that could calm their mood and emotional state. This process was laborious, more time consuming, and an individual often encountered the problem of arriving at a suitable song list. Even if they identify the user's mood, their choice of songs for creating a playlist simply selects songs that reflect the user's current feelings and does not try to improve their mood in any way.

Therefore, if the user is sad, they are presented with a list of songs with sad emotions that could worsen their mood and potentially lead to sadness. So, the system proposed in this paper can identify the user's emotions from their facial expressions. It will then offer the user a playlist of songs, ensuring that the user may feel better.

2. LITREATURE SURVEY

1."Smart Music Player Integrating Facial Emotion Recognition and Music Mood Recommendation"

Authors: Shlok Gilda, Husain Zafar, Chintan Soni, Kshitija Waghurdekar.

Methodology: The methodology encompasses the utilization of deep learning algorithms for facial emotion recognition and the extraction of acoustic features for music classification. These processes are integrated to map the user's mood to recommended songs, enhancing the personalized music listening experience. The study leverages datasets comprising facial expressions and labeled songs to train artificial neural networks for accurate emotion identification and music mood classification. The evaluation of the system's performance involves testing on a substantial dataset of images and songs to achieve high accuracy levels. Future enhancements include exploring the system's performance with all seven basic emotions, incorporating songs from diverse languages and regions, and integrating user preferences through collaborative filtering for further refinement of the recommendation system.

2."Smart Space with Music Selection Feature Based on Face and Speech Emotion and Expression Recognition"

Authors: Jose Martin Z. Maningo, Argel A. Bandala, Ryan Rhay P. Vicerra, Elmer P. Dadios, Karla Andrea L. Bedoya.

Methodology: The methodology of the research paper involves developing an embedded system using a Raspberry Pi 3, camera, microphone, and speakers to select appropriate music based on user emotion. The system comprises three main features: Face Recognition for emotion detection, Speech Recognition for extracting speech components, and Emotion Classification for identifying displayed emotions. The Facial Expression Database used is the Extended Cohn-Kanade Dataset, while the Speech Emotion Database is the RAVDESS database. Face Emotion Recognition involves image processing, feature extraction, and training the algorithm using a dataset split into training and test sets. The system captures user images and processes them for emotion recognition. Speech input is recorded and classified using SVM algorithms. The system runs in a controlled Smart Space environment with sufficient lighting and minimal noise. The study recommends gathering more data samples for training, improving system efficiency, integrating additional modalities for emotion recognition, and using better components for enhanced accuracy.

3."Musical Therapy using Facial Expressions"

Authors: Pranjul Agrahari, Ayush Singh Tanwar, Biju Das, Prof. Pankaj Kunekar.

Methodology: The methodology of the study "Musical Therapy using Facial Expressions" involves a multi-step approach to implement the proposed facial expression-based music therapy system. Firstly, facial expressions are extracted from images using image processing techniques and machine learning algorithms. Multiple images are captured from a webcam to predict emotions accurately, with blurred images being averaged to reduce errors. Histogram equalization is applied to enhance image contrast, followed by cropping and conversion to grayscale for foreground isolation. For music emotion recognition, a music dataset is obtained and feature extraction is performed using MIR packages like LibROSA and PyAudio Analysis. These packages analyze musical features such as tempo, rhythm, and timbre, along with lyrics, for clustering using unsupervised learning with the k-means algorithm. Centroids are manually set for efficient execution, and songs are labeled with descriptors for search optimization post-facial emotion recognition. Machine learning algorithms are utilized for emotion classification, with the development of an emotion API to provide output based on the input to the system.

4."Musical Therapy through Facial Expression Recognizer using Machine Learning"

Authors: Shruti Aswal, Sruthilaya Jyothidas, S Shankarakrishnan.

Methodology: The methodology of the study involved the development and implementation of a novel approach to musical therapy through the integration of convolutional neural networks (CNNs) and facial emotion recognition technology. Real-time video data capturing facial expressions was analyzed using CNNs to classify emotional states such as happiness, sadness, or anger. This emotional state classification was then used in conjunction with a music recommendation algorithm to provide personalized playlists tailored to the individual's mood, enhancing the therapeutic experience. Furthermore, a voice assistant component was integrated into the system to offer additional support and guidance, ensuring inclusivity for users of varying ages and abilities. By leveraging machine learning techniques, the study aimed to create a personalized and immersive musical therapy experience, with the ultimate goal of improving mental health and well-being through innovative and effective therapy sessions.

3. Objective

The goal is, like, creating this system thing that can, like, detect and classify different facial expressions such as happy, sad, mad, scared, chill, surprised, and grossed out in, like, real-time. By using, like, image stuff and smart algorithms, the goal is to automate the whole emotion thing, you know, so we can play the tunes for whoever based on how they're feeling. And, like, we also wanna figure out what kind of tunes folks like based on their faces, so we can make them playlists that match their vibes. So, with these smart algorithms, man, the plan is for the system to keep learning and get better at reading emotions, which will make music therapy even more awesome. Plus, we wanna make it easier for therapists by picking tunes based on expressions, so they don't gotta work so hard, you know? This study also wants to make therapy a better experience by making the interface cool and relaxing, so people are into it. At the end of the day, what we're really aiming for is to show that using facial expressions for music therapy is totally doable and actually works. It's a whole new way of helping folks with their mental health, bringing together tech and therapy in a rad way.

4. Problem Statement

The problem statement was addressed in the study on "Musical Therapy using Facial Expressions" is a need for a more precise and customized approach to music therapy within mental health treatment. Traditional methods of music therapy often involve therapists manually analyzing a patient's mood and selecting appropriate songs, which can be time-consuming and may not always lead to the most effective therapy outcomes. This manual process lacks automation and personalization, leading to challenges in providing tailored music therapy that aligns with the patient's emotional state. Moreover, the traditional approach may not always consider the nuances of individual emotions or provide a comprehensive understanding of the mental well-being of the patient. A more adaptive and personalized therapy solution is needed that can dynamically respond to the patient's emotional cues in real-time, enhancing the

effectiveness of music therapy sessions. The study aims to address these challenges by developing an intelligent system that integrates facial expression recognition technology with music therapy. By automating the process of emotion detection and music selection based on facial expressions, the system seeks to provide a more efficient, personalized, and engaging music therapy experience for patients. This innovative approach aims to overcome the limitations of traditional music therapy methods and improve the overall quality of mental health treatment through technology and therapeutic interventions.

5. Proposed Statement

The intended system can identify the facial statement of the user and supported on his/her facial tone extract the facial milestones, which would then be classified to acquire a unique emotion of the user. When the feeling has been classified the tunes corresponding the user's emotions would be shown to the user. In this intended system, we construct a model in suggestion of forceful music suggestion system depending on human feelings. Depending on each human listening outline, the tunes for each emotions are skilled. Composition of feature extraction and machine education methods, from the actual face the feeling are identified and once the feeling is produced from the contribution display, respective tunes for the exact spirit would be played to contain the users. In this advance, the request gets associated with human beliefs thus deliver a particular stroke to the users. Consequently our considered system intensify on recognizing the human feelings for developing feeling based musical player using computer sense and machine learning methods. For empirical effects, we practice openCV for feeling identification and musical endorsement.

6. Methodology

In the propose algorithm, face images and facial landmark detections are performing first for recognition of stress. A Convolutional Neural Nets (CNN) algorithm was used in the propose network, face images and expressions detected earlier are input in order to output stress recognition results.

The results of face recognition are consisted of students present in the class.

Step 1: Convolution

A convolution is a connected integration of two methods that shows you how one method transforms the other.

Step 2: Implement the RLU (Revised Linearity Unit)

In this step, the correctional function is used to increase nonlinearity on CNN. The data set consists of different objects which are not linear to each other. Under this function, the grouping of information can be seen as a linear problem, even though it is a non-linear problem.

Step 3: Pooling

Spatial invariance is a term that does not affect the neural network's ability to detect its particular features when discovering an item in the data collection. Pooling helps CNN to detect swimming pools, such as max and minimum pools, for example.

Basing on the weather forecast, detect the faces and student's attendance will be taken.

7. MODULE DESCRIPTION:

Data Collection Module

A survey is being collected from users based on 3 parameter which are, 1. What type of song would they wants to listen to when they was happy? 2. What kind of song would they wants to listen to when they is sad? 3. What type of song would they like to listen to if they are angry?.

Emotion Extraction Module

The picture of the user is captured with the help of a camera/web cam. When the photo is taken, the frame of the picture captured from webcam feed is converted to a grayscale image to better the completion of the classifier, which is used to identify the face existing in the picture. When the conversion is done, the picture is sent to the classifier algorithm which, using the feature extraction methods, can extract the face from the frame of the webcam feed. From the face extracted, individual features is got and sent to the trained network to detect the emotion expressed by the user. These pictures will be used to educate the classifier so that when an absolutely new and strange set of pictures are shown to the classifier, it is able to extract the position of facial landmarks from those pictures based on the knowledge that it already had acquired from the training set and return the coordinates of the new facial landmarks that it detected. The network is educated using extensive data set. This is used to realize the emotion being expounded by the user.

Audio Extraction Module

After the emotion of the user is extracted the music/audio based on the emotion expounded by the user is shown to the user, a list of songs based on the emotion is shown, and the user can enjoy any song he/she likes. Based on the regularity that the user would listen to the songs are shown in that order.

Emotion - Audio Integration Module

The emotions that are extracted for the songs are stored, and the songs based on the emotions are shown on the web page. For example, if the emotion or the facial feature is classified under the word happy, then songs from the happy database are shown to the user.

8. SYSTEM ARCHITECTURE:

The overall software features concern with define requirements and establish the high level of system. In architectural design, various web pages and their connections are identified and are designing. The major software components identify and decompose into processing modules and conceptual data structures and interconnections amongst modules are identified. Following modules identify in proposed system!

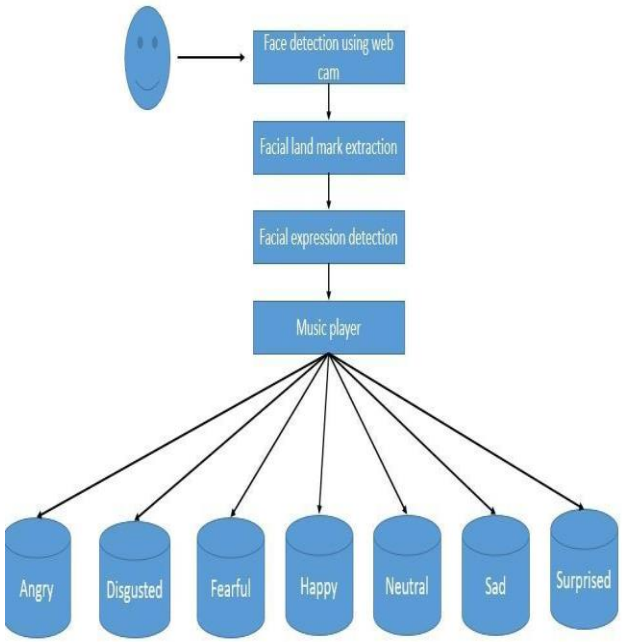

Fig:- System Architecture

9. Application

- Online Mental Health Platforms
- Mood Tracking and Self-Care
- Support for Special Needs Individuals
- Clinical Therapy Settings
- Stress Management Programs
- Special Education and Developmental Disorders
- Rehabilitation Programs
- Senior Care and Dementia Management
- Pain Management
- Rehabilitation and Recovery
- Stress Reduction in Workplaces

10. Advantages

- Reduce Anxiety
- Reduce Depression
- Reduce Stress
- Improve Emotional Regulation & Processing
- Promote Feelings of Safety & Security
- Improve Self-Esteem & Sense of Identity
- Real-Time Feedback
- Enhanced Engagement
- Accessible Mental Health Support

CONCLUSION

In this project, we have discussed that how our proposed system recommend music based on facial expression using machine learning algorithms. The proposed system is also scalable for recommending music based on facial expressions by using techniques after collecting data. The system is not having complex process to recommend music that the data like the existing system. Proposed system gives genuine and fast result than existing system. Here in this system we use machine learning algorithms to recommend music based on real time facial expression.

FUTURE WORK

- Personalization: Implement a feature that allows users to create personalized playlists based on their specific emotional needs and preferences.
- Real-time Feedback: Incorporate a feedback mechanism for users to rate the effectiveness of the recommended music in improving their mood.
- Expanded Music Database: Enhance the music database by including a wider range of genres and songs to cater to diverse emotional states and preferences.
- Machine Learning Algorithms: Explore advanced machine learning algorithms for more precise emotion detection and music recommendation, such as deep learning models or ensemble methods.

References

[1] Shlok Gilda, Husain Zafar, Chintan Soni, Kshitija Waghurdekar,"Smart music player integrating facial emotion recognition and music mood recommendation", Date of Conference: 22-24 March 2017, India.

[2] A. Fathallah, L. Abdi, and A. Douik, “Facial Expression Recognition via Deep Learning,” in 2017 IEEE/ACS 14th International Conference on Computer Systems and

Applications (AICCSA), Hammamet, 2017, pp. 745–750, doi: 10.1109/AICCSA.2017.124

[3] H. Bahuleyan, "Music Genre Classification using Machine Learning Techniques," arXiv:1804.01149 [cs, eess], Apr. 2018.

[4] A. Elbir, H. Bilal Çam, M. Emre Iyican, B. Öztürk, and N. Aydin, "Music Genre Classification and Recommendation by Using Machine Learning Techniques," in 2018 Innovations in Intelligent Systems and Applications Conference (ASYU), Oct. 2018, pp. 1–5, doi: 10.1109/ASYU.2018.8554016.

[5] S. L. Happy and A. Routray,"Automatic facial expression recognition using features of salient facial patches," in IEEE Transactions on Affective Computing, vol.6, no. 1, pp. 1-12, 1 Jan.-March 2015.

[6] Rahul ravi, S.V Yadhukrishna, Rajalakshmi, Prithviraj, "A Face Expression Recognition Using CNN & LBP",2020 IEEE.

[7] Anuja Arora, Aastha Kaul, Vatsala Mittal,"Mood based music player",Date of conference : 2019, India

[8] S.Deebika, K.A.Indira, Jesline, "A Machine Learning Based Music Player by Detecting Emotions", 2019 IEEE

[9] Karthik Subramanian Nathan, manasi arun, Megala S kannan," EMOSIC — An emotion-based music player for Android,2017 IEEE

[10] H. I. James, J. J. A. Arnold, J. M. M. Ruban, M. Tamilarasan, And R. Saranya, "Emotion Based Music Recommendation System," Vol. 06, No. 03, P. 6, 2019.