WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

FACIAL EXPRESSION DETECTION AND MUSIC RECOMMENDATION (SENSE-BOT)

AKHILESH SINGH[1], ABHISHEK KUSHWAHA [2], ANUJ KUMAR VERMA [3], AAKASH PRAJAPATI, [4]VINEET SHRIVASTAVA [5]

[1] [2] [3] [4] Computer Science & Engineering
RAJ KUMAR GOEL INSTITUTE OF TECHNOLOGY, GHAZIABAD
[5] Guide, RAJ KUMAR GOEL INSTITUTE OF TECHNOLOGY, GHAZIABAD

ABSTRACT: -

In the contemporary realm of multimedia and technology, music serves as a vital and therapeutic element, enhancing individuals' lives with entertainment. Despite the evolution of music player features, the manual effort required for playlist creation remains a challenge. This research addresses this gap by proposing an innovative music recommendation system that automates playlist creation based on the user's emotional state, leveraging facial expression recognition technology. By capturing users' emotional states during interactions, the system aligns with the idea that emotions can be expressed through various channels, including facial expressions. The impact extends to enhancing personalized music recommendations by aligning users' musical preferences with their emotional states, making music listening more engaging and personalized. Furthermore, the system addresses the monotonous task of manual playlist creation, providing a dynamic and automated alternative.

Key Words: Emotion Recognition, Linear classifier, Facial Landmark Extraction, SVM Classification.

INTRODUCTION: -

Facial expression detection is a fascinating field that has gained significant attention in recent years due to its potential to understand human emotions. In this paper, we explore the use of facial expression detection technology in the context of music recommendation. Specifically, we aim to answer the following questions: How can facial expression detection are used to understand the emotional state of an individual? What are the key technologies and algorithms involved in facial expression detection? And, how can the detected facial expressions are used to recommend music based on the individual's emotional state? By

addressing these questions, we hope to shed light on the potential of facial expression detection technology in enhancing music recommendation systems, which can ultimately lead to a more personalized and enjoyable music listening experience for individuals. In the contemporary era, where technology seamlessly intertwines with various aspects of daily life, the realm of multimedia and entertainment has witnessed unprecedented advancements. With the proliferation of multimedia technologies, music players have evolved, offering features ranging from traditional playback controls to sophisticated functionalities such as genre classification and variable playback speed.

The motivation for this research stems from the recognition of music's pivotal role in enhancing individuals' lives and the desire to streamline the process of playlist creation. Despite the plethora of features in modern music players, the task of manually creating playlists remains time-consuming and burdensome. In response to this challenge, the present work proposes a cutting-edge approach that harnesses facial expression analysis to automate the generation of music playlists based on users' emotional states.

The human face serves as a rich source of information about an individual's mood, making it a compelling input for an automated music recommendation system. By employing facial recognition technology and leveraging the advancements in computer vision, the proposed system aims to extract emotional cues directly from users' facial expressions captured through a camera. This real-time extraction of emotional features enables the system to dynamically categorize users' moods, eliminating the need for manual playlist curation.

This research is poised to revolutionize the landscape of music recommendation systems by introducing a novel methodology that integrates facial expression recognition. As music has the profound ability to evoke emotions, this groundbreaking system seeks to create an immersive and personalized listening experience, aligning the playlist with the ever-changing emotional states of the user. The subsequent sections delve into the technical intricacies of the proposed system, outlining its methodology, experimental design, and potential implications for the broader field of affective computing in multimedia applications.

2. LITERATURE SURVEY

Renuka R Londhe et al. [1] introduced a paper focusing on changes in facial curvatures and corresponding pixel intensities, utilizing Artificial Neural Networks (ANN) for emotion classification. The authors also proposed varied playlist approaches. In a similar vein, Zheng et al. [2] categorized facial feature extraction into appearance-based and geometric-based methods, emphasizing essential facial points. Nikhil et al. [3] determined user mindsets through facial expressions, aiming to reduce time complexity in music playlist creation. The system captured user images through a webcam, converting them to binary format and utilizing Haar Cascade technologies. Zeng et al. [4] delved into advances in human affect recognition, particularly in audio/visual computing methods, offering a prototype for emotion categories. Parul Tambe et al. [5] automated interactions between users and a music player, learning preferences, emotions, and activities to provide song selections based on facial expressions. Jayshree Jha et al. [6] proposed an emotion-based music player using image processing, amalgamating various algorithms to enhance playlist creation. Anukritine et al. [7] devised an algorithm segregating emotions into categories and providing accurate audio information retrieval with reduced computational time. Aditya et al. [8] developed an Android application utilizing image processing for facial recognition, offering a customized music player based on user mood. A. Habibzad et al. [9] introduced a three-stage algorithm for facial emotion recognition, optimizing features for efficient classification. Prof. Nutan Deshmukh et al. [10] focused on a system capturing user emotions at intervals for an automated emotion-based music system.
Chang Liu et al. [11] incorporated Brain-Computer Interfaces (BCI) using EEG hardware to monitor cognitive states, while Swati Vaid et al. [12] reviewed EEG for estimating emotions, acknowledging its limitations in terms of portability and economics.

3. RELATED WORK:

Face Detection and Facial Expression Recognition System

Anagha S. Dhavalikar et al [8] proposed Automatic Facial Expression recognition system. In This system there are three phase 1. Face detection 2. Feature Extraction and 3. Expression recognition. The First Phase Face Detection are done by RGB Color model, lighting compensation for getting face and morphological operations for retaining required face i.e eyes and mouth of the face. This System is also used AAM i.e Active Appearance Model Method for facial feature extraction In this method the point on the face like eye, eyebrows and mouth are located and it create a data file which gives information about model points detected and detect the face the an expression are given as input AAM Model changes according to expression.

Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting

Yong-Hwan Lee, Woori Han and Youngseop Kim proposed system based on Bezier curve fitting [9]. This system used two step for facial expression and emotion first one is detection and analysis of facial area from input original image and next phase is verification of facial emotion of characteristics feature in the region of interest [8]. The first phase for face detection it uses color still image based on skin color pixel by initialized spatial filtering ,based on result of lighting compassion then to estimate face position and facial location of eye and mouth it used feature map After extracting region of interest this system extract points of the feature map to apply Bezier curve on eye and mouth The for understanding of emotion this system uses training and measuring the difference of Hausdorff distance With Bezier curve between entered face image and image from database.

Using Animated Mood Pictures in Music Recommendation

Arto Lehtiniemi and Jukka Holm et al [10] proposed system on animated mood picture in music recommendation. On this system the user interacts with a collection of images to receive music recommendation with respect to genre of picture. This music recommendation system is developed by Nokia researched center. This system uses textual meta tags for describing the genre and audio signal processing**.**

Human-computer interaction using emotion recognition from facial expression.

F. Abdat, C. Maaoui et al and A. Pruski et al. They proposed a system fully automatic facial expression and recognition system based on three step face detection, facial characteristics extraction and facial expression classification [11]. This system proposed anthropometric model to detect the face feature point combined to shi and Thomasi method. In this metod the variation of 21 distances which describe the facial feature from neutral face and the classification base on SVM (Support Vector Machine).

Emotion-based Music Recommendation By Association Discovery from Film Music.

Fang-Fei Kuo et al and Suh-Yin Lee et al. [12] With the growth of digital music, the development of music recommendation is helpful for users. The existing recommendation approaches are based on the users' preference on proposed system based on Bezier curve fitting [9]. This system used two steps for facial expression and emotion first one is detection and analysis of facial area from input original image and next phase is verification of facial emotion of characteristics feature in the region of interest [8]. The first phase for face detection it uses color still image based on skin color pixel by initialized spatial filtering ,based on result of lighting compassion then to estimate face position and facial location of eye and mouth it used feature map After extracting region of interest this system extract points of the feature map to apply Bezier curve on eye and mouth The for understanding of emotion this system uses training and measuring the difference of Hausdorff distance With Bezier curve between entered face image and image from database. © 2019, IRJET | Impact Factor value: 7.211 music. However, sometimes, recommending music according to the emotion is needed. In this paper, we propose a novel model for emotion-based music recommendation, which is based on the association discovery from film music. We

investigated the music feature extraction and modified the affinity graph for association discovery between emotions and music features. Experimental result shows that the proposed approach achieves 85% accuracy in average.

Moodplay: Interactive Mood-based Music Discovery and Recommendation

Ivana Andjelkovic et al and John O'Donovan et al [13] they proposed that a large body of research in recommender systems focuses on optimizing prediction and ranking. However, recent work has highlighted the importance of other aspects of the recommendations, including transparency, control and user experience in general. Building on these aspects, we introduce MoodPlay, a hybrid recommender system music which integrates content and mood-based iltering in an interactive interface. We show how MoodPlay allows the user to explore a music collection by latent affective dimensions, and we explain how to integrate user input at recommendation time with predictions based on a pre-existing user profile. Results of a user study (N=240) are discussed, with four conditions being evaluated with varying degrees of visualization, interaction and control.

An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions

Anukriti Dureha et al [14]. In this he proposed Manual segregation of a playlist and annotation of songs, in accordance with the current emotional state of a user, is labor intensive and time consuming. Numerous algorithms have been proposed to automate this process. However, the existing algorithms are slow, increase the overall cost of the system by using additional hardware (e.g. EEG systems and sensors) and have less accuracy. This paper presents an algorithm that automates the process of generating an audio playlist, based on the facial expressions of a user, for rendering salvage of time and labor, invested in performing the process manually. The algorithm proposed in this paper aspires to reduce the overall computational time and the cost of the designed system. It also aims at increasing the accuracy of the designed system. The facial expression recognition module of the proposed algorithm is validated by testing the system against user dependent and user independent dataset.

Enhancing Music Recommender Systems with Personality Information and Emotional States

Bruce Ferwerda et al and Markus Schedl et [15] al proposed that the initial research assumptions to improve music recommendations by including personality and emotional states. By including these psychological factors, we believe that the accuracy of the recommendation can be enhanced. The system gives attention to how people use music to regulate their emotional states, and how this regulation is related to their personality.

4. IMPLEMENTATION

The model integrates Haar Cascade, SoftMax classifier and convolutional neural network (CNN) for face recognition. Initially, the system uses Haar Cascade and SoftMax to detect faces in input images. It then crops the visible face and normalizes it to (48×48) pixel size. These facial expressions work for CNN. CNN takes the ideas and creates a list from 0 to 6 corresponding to emotions such as "anger," "disgust," "fear," "happiness," "interest," "surprise," and "none." Using CNNs to preprocess and analyze visual images can improve the model's ability to identify facial expressions.

Haar Cascade, SoftMax, and CNN Layered Architecture:

The input process is the first part of the CNN architecture where the user provides images/data to be recognized.

Haar cascade classifier

This is considered as an object detection method similar to Haar face recognition. Using Haar-like features with multiple features, the classifier uses features such as "line features" and "rectangles" to identify diagonals and edges around objects. Using composite images can reduce the computational cost. This algorithm processes the input image (usually a (24×24) window) pixel by pixel to analyze every Haar feature in the window. The training set represents positive and negative images, positive represents images with faces, and negative represents images without faces. The Haar stepped faceplate model is generally used for face detection. In the Haar Cascade layer, h(x) represents the category of the Haar feature extraction function and w represents the feature weight.

$$\text{Haar Cascade Output} = \sum h.wh.h(x)$$

SoftMax Classifier

The SoftMax classifier function is integral to converting a vector of K real values into a vector of K probabilities that sum up to 1. By transforming input values, regardless of their range, into probabilities between 0 and 1, it effectively categorizes them as such. This transformation ensures that small or negative inputs are converted into corresponding small probabilities, while large or positive inputs result in larger probabilities, all confined within the range of 0 to 1. In the formula provided, the SoftMax function calculates the probability distribution for each class, denoted by jth class. This mathematical operation is vital for generating accurate probabilistic outputs, particularly in classification tasks.

$$P(\text{Class}) = e^{z}_{j} / \sum_{k}^{K}{}_{=1} e^{z}_{k}$$

Convolutional Layer

The Convolutional Layer constitutes the second layer in the CNN architecture and is pivotal to the recognition process. This layer serves as the cornerstone of CNN, responsible for extracting essential features from images critical for their identification. Images are treated as pixel matrices in this layer, where convolution processes are performed using multiple filters. Through convolution, the filter matrix slides over the image pixel matrix, computing dot products to generate the convolved feature matrix. Representing the input to the CNN layer as xi, the weight matrix as Wi, and the bias term for the ith layer as bi, the output of the CNN layer is obtained via convolution operation followed by activation function σ.

$$Y = (W_i * X_i + b_i)$$

ReLU Layer

Following the Convolutional Layer, the ReLU Layer employs the Rectified Linear Unit function for normalization. ReLU calculates the maximum valued feature matrix from the input received from the convolutional layer. This non-linear operation replaces negative values in the feature map with zero, enhancing the clarity of significant features.

Pooling Layer

Next, the Pooling Layer reduces the dimensions of each activation map while preserving important features necessary for image identification. Though it reduces dimensionality, it ensures crucial image features are retained.

Fully Connected Layer

The Fully Connected Layer (FCL) serves as the final layer of the CNN model. This layer takes input from the pooling layer in flattened form and combines affine and nonlinear functions. Connected to filters of previous and subsequent layers, FCL ensures comprehensive feature analysis and classification.

Finally, the Output and Error Layer of CNN generate output for the given input image/data along with accuracy and error metrics. This layer employs a combination of Softmax or sigmoid functions and cross-entropy loss functions for calculating image recognition probability and error respectively.

Data Collection: A review was gathered from clients in view of 3 boundaries which are, 1.What kind of tunes would they need to pay attention to when they are blissful? 2. What sort of melodies would they need to pay attention to when they are miserable? 3. What kind of tunes could they need to pay attention to when they are furious? Gather a diverse dataset that includes information on user behavior, physiological responses, and contextual inputs. This dataset serves as the foundation for training the machine learning models.

Feature Extraction: Extract relevant features from the collected data, such as facial expressions, user interactions, and contextual information. These features will be used to train the emotion detection model.

Emotion Detection Model: Utilizing Neural Networks for Feature Detection involves the use of artificial neurons that, by tolerating the sign, establish connections among neurotransmitters. These connections form pathways, each considered as a route between two neurons. These pathways incrementally pass signals to recognize transferred images. The neurons engage in a meticulous process, checking against existing images to determine whether the signal follows a known path (indicating a matching or existing face) or diverges due to unfavorable lighting conditions, classifying the image as not matching any known faces.
In the scenario where a match is found, the signal continues along the path of the identified face and archives the image in a previous collection through landmark testing. Conversely, when the signal diverts its course, various criteria are examined. For instance, if the image contains large eyes, the system checks if it had images of faces with large or small eyes. If the image contains large eyes, the signal progresses towards all images with large eyes, conducting further assessments such as eye color. If the eyes are determined to be dark, additional qualifications, such as the distance between eyes, eyebrow color, lip types, and nose characteristics, are sequentially evaluated. This process allows for a detailed and iterative analysis, enabling the neural network to discern intricate facial features and patterns, contributing to effective feature detection.

Integration with Music Recommendation System: Integrate the emotion detection model with a robust music recommendation system. This system should consider user preferences, historical data, and real-time emotional states to generate personalized music recommendations.

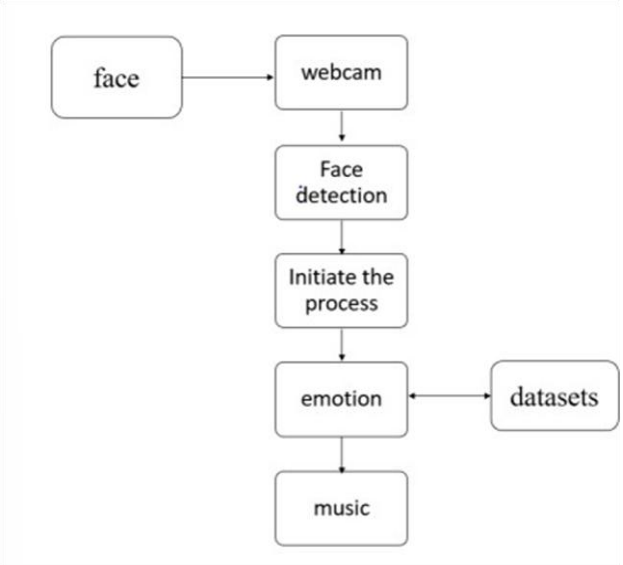


Fig.1. Application Architecture

5. Experimental results and discussions

Upon initiating the web application, users encounter a popup message prompting them to grant permission for camera access. Subsequently, the system engages in detecting the user's facial features. Following successful detection, the system employs advanced algorithms to recognize the user's emotional state. This recognition process encompasses a spectrum of emotions, including happiness, sadness, anger, or neutrality. The crucial aspect of this system lies in its ability to dynamically recommend a playlist of music based on the detected emotional state. For instance, if the system identifies a "happy" emotion, it suggests a playlist comprising lively and cheerful tunes. Similarly, when the detected emotion is "neutral," the system proposes a playlist tailored to that particular mood. This interactive experience empowers users to not only explore their emotional states but also engage with a personalized musical journey. The granular approach of linking emotions to music preferences enhances the overall user experience, making the system a versatile and intuitive platform. Throughout this process, utmost attention is given to user privacy and data security, emphasizing a commitment to responsible and secure technology integration. The seamless coordination of facial recognition, emotion detection, and personalized music recommendation creates a cohesive and user-centric application.

Fig.2. User's Emotion Detection (Fearful)

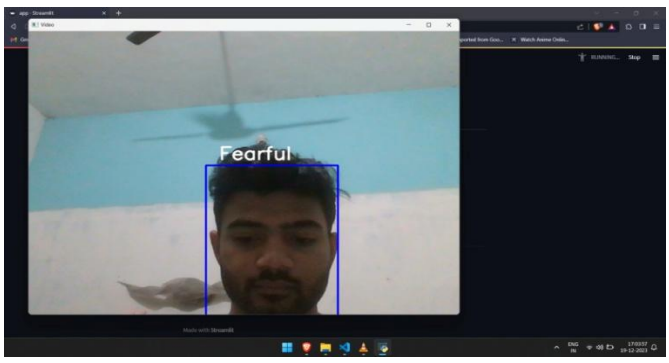


Fig.3 Sensing and recommending

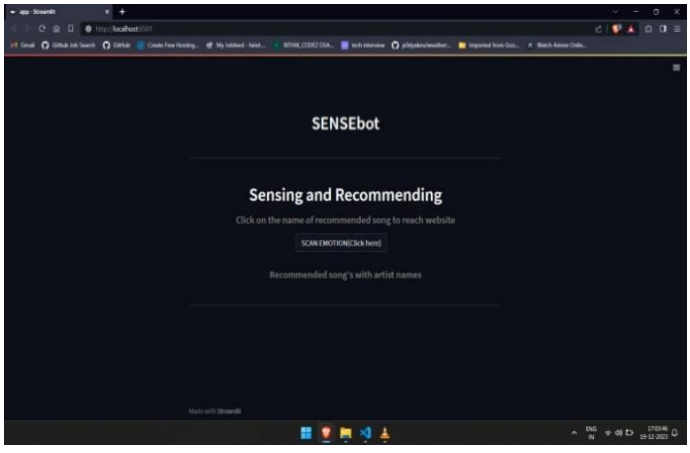
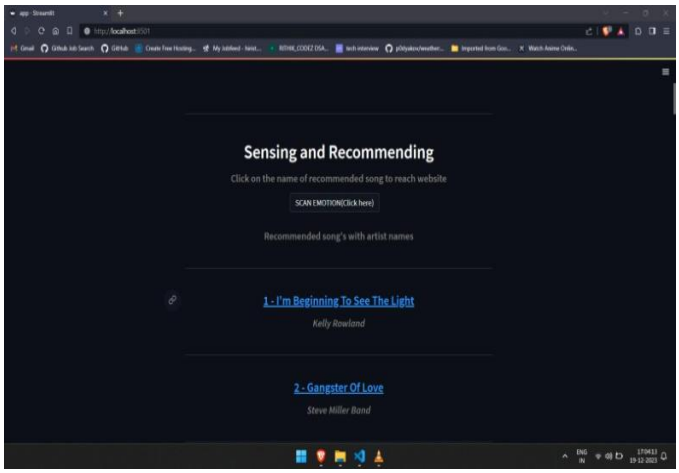


Fig.4. Recommended Music Playlist



6. FUTURE ENHANCEMENT

The music player based on facial recognition system is highly essential for all the person in modern day life ecology. This system is further enhanced with benefit able features for upgrading in future. The methodology of enhancement in the automatic play of songs are done by detection of the facial expression. The facial expression is detected by programming interface with the RPI camera. An alternative method, based on additional emotions which is excluded in our system as disgust and fear. On this emotion included to support the playing of music automatically.

CONCLUSIONS

In this project, we presented a model to recommend a music-based om the emotion based detected from the facial expression. This project proposed designed & developed an emotion-based music recommendation system using face recognition System. Music are 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. Thus, the proposed system presents Face based emotion recognition system to detect the emotions and play music from the emotion detected.

REFERENCES :

[1] Face Recognition Data, University of Essex, UK

[2] K. T. Talele, S. Kadam, A. Tikare, Efficient Face Detection using Adaboost, “IJCA Proc on International Conference in Computational Intelligence”, 2012.

[3] T. Mita, T. Kaneko, O. Hori, Joint Haar-like Features for Face Detection, “Proceedings of the Tenth IEEE International Conference on Computer Vision”, 1550- 5499/05 ©2005 IEEE.

[4] T. Ahonen, A. Hadid, M. Peitikainen , Face recognition with local binary patterns. “In Proc. of European Conference of Computer Vision”, 2004.

[5] M. A. Turk and A.P. Pentland, Face recognition using eigenfaces, “Proceedings of the IEEE”, 586- 591, 1991.

[6] J Lu, K. N. Plataniotis, A. N. Venetsanopoulos, Face recognition using LDA-based algorithms, “IEEE Neural Networks Transaction”, 2003.

[7] L. Wiskott, M. Fellous, N. Krger, and C. Malsburg,Face recognition by elastic bunch gra

[8] Anagha S.Dhavalikar and Dr. R. K. Kulkarni, “Face Detection and Facial Expression Recognition System” 2014 Interntional Conference on Electronics and Communication System (ICECS -2014).

[9]. Yong-Hwan Lee, Woori Han and Youngseop Kim, “Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting” 2013 16th International Conference on Network-Based Information Systems.

[10]. Arto Lehtiniemi and Jukka Holm, “Using Animated Mood Pictures in Music Recommendation”, 2012 16th International Conference on Information Visualisation.

[11]. F. Abdat, C. Maaoui and A. Pruski, “Human computer interaction using emotion recognition from facial expression”, 2011 UKSim 5th European Symposium on Computer

[12]. T.-H. Wang and J.-J.J. Lien, “Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D PoseEstimation,” J. Pattern Recognition, vol. 42, no. 5, pp. 962-977, 20

[13] Renuka R. Londhe, Dr. Vrushshen P. Pawar, “Analysis of Facial Expression and Recognition Based on Statistical Approach”, International

Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

[14] Anukriti Dureha "An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions":IJCA2014.

[15] Bruce Ferwerda and Markus Schedl "Enhancing Music Recommender Systems with Personality Information and Emotional States": A Proposal: 2014.