Mood Sensitive Music Recommendation System

T Sunil Kumar[1], CH Harshitha[2], G Naga Charani[3], MD Nawaz[4] , G Nithin[5]

[1] Professor, VNR Vignana Jyothi Institute of Engineering & Technology, Hyderabad
[2345]Under Graduate Student, VNR Vignana Jyothi Institute of Engineering & Technology, Hyderabad

---***---

Abstract - *The mood sensitive music recommendation system using facial expressions is a system designed to provide music recommendations based on the user's current mood as inferred from their facial expressions. The system uses computer vision and machine learning algorithms to analyze the user's facial expressions in real time such as using webcam. The system then recommends music that matches the user's mood considering the tempo and genre of the music. Once the user's current emotional state has been determined, the system then selects music that is likely to match that mood. For example, if the user is determined to be feeling sad, the system might recommend slower, more melancholic music, while if the user is determined to be feeling happy, the system might recommend more upbeat and energetic music. The system may be applied in a number of situations, like music streaming services or in retail environments where music is played to create a particular atmosphere. By tailoring the music to the user's current mood, the system aims to provide a more personalized and enjoyable listening experience. Also, improves their emotional state.*

Key Words*:* (Facial Detection, Emotion detection, Music Recommendation, CNN.)

1.INTRODUCTION

Music has the remarkable ability to evoke emotions, influence moods, and connect with our inner selves. Whether it's to lift our spirits, relax our minds, or accompany specific activities, music plays an integral role in our daily lives. However, with the vast array of musical genres and artists available today, finding the perfect song that resonates with our current mood can be a daunting task. This is where the Mood Music Recommendation System comes into play.

The Mood Music Recommendation System is an innovative application of artificial intelligence and machine learning techniques that aims to enhance the music listening experience by providing personalized music recommendations based on the user's mood. By analyzing various audio features, lyrics, and user preferences, the system can effectively understand and categorize music according to its emotional content. We will explore the existing challenges and limitations faced by conventional music recommendation systems and how the proposed Mood Music Recommendation System addresses these issues. We will discuss the potential applications and future directions of the Mood Music Recommendation System, highlighting the possibilities for integrating it into music streaming platforms, mobile applications, and other music-related services.

We will also reflect on the ethical considerations associated with personalized recommendation systems and explore avenues for addressing privacy concerns and ensuring fair usage. The Mood Music Recommendation System offers a novel approach to enhance the music listening experience by providing tailored recommendations based on the user's mood. This project report will serve as a comprehensive guide, shedding light on the design, implementation, and evaluation of the system, and laying the foundation for further advancements in the field of intelligent music recommendation. The purpose of this study is to develop a music recommendation system that can recognise a user's face, determine their current mood, and then suggest a playlist depending on that mood.

2. LITERATURE SURVEY

2.1. Music Emotion Recognition: A State of the Art Review (2011)

Yang and Chen's review paper offers an extensive analysis of music emotion recognition techniques. It explores the different modalities used to capture emotion, including audio features, lyrics, and user feedback. The paper discusses the various machine learning algorithms employed for emotion classification, such as support vector machines (SVM), decision trees, and artificial neural networks. It highlights the importance of feature selection and extraction in improving the accuracy of emotion recognition systems. Additionally, the review discusses the challenges faced in music emotion recognition, such as subjectivity and the inherent complexity of emotional experiences. The authors present future directions for research, including the integration of multi-modal information and the exploration of deep learning approaches to enhance music emotion recognition systems.

2.2. A Survey of Music Recommendation Systems and Future Perspectives (2014)

Serra and Herrera's survey paper provides a comprehensive overview of music recommendation systems and their underlying algorithms. It covers collaborative filtering, content-based filtering, and hybrid methods, comparing their strengths and limitations. The paper discusses the challenges in music recommendation, such as the cold start problem and the scalability of recommendation systems. It also explores future perspectives, including the integration of social and contextual information to improve recommendation accuracy. The authors highlight the need for evaluating recommendation systems using appropriate metrics and user-centric approaches. The survey serves as a valuable resource for understanding the landscape of music recommendation systems and provides insights into potential areas of improvement.

2.3. An Emotion-Aware Personalized Music Recommendation System Using a Convolutional Neural Networks Approach (2018)

Abdul, Ashu, et al paper is anticipated to delve into the expanding topic of emotion-aware recommendation systems, which try to account users' emotional states while making music recommendations. The article describes an Emotion-Aware customised Music Recommendation System (EPMRS) that provides customised song suggestions based on the user's time, ambiance, geographical location, preference, current emotion, song listening habit, play count, and audio track duration. Deep Convolutional Neural Networks (DCNN) and Weighted Feature Extraction (WFE) algorithms are used by the system to learn a user's music preferences based on their current feelings. The testing results suggest that this technique works well for learning individual users' emotional preferences. In the future, the scientists hope to automatically extract users' present emotions from social media data and consider data from platforms such as YouTube, Facebook, and Twitter to better understand their preferences. The research also advises investigating the use of recurrent neural networks (RNN) for song categorization as a potential approach of improving the EPMRS's performance.

2.4. Emotion Based Music Recommendation System (2018)

Aryan, Pratham, Sahil and Samuel Jacob paper presents a novel approach for automatically playing music based on facial emotion. Unlike manual sorting or wearable computing device-based methods, the system leverages a Convolutional Neural Network (CNN) for emotion detection and Pygame & Tkinter for music recommendations. The system utilizes a Convolutional Neural Network (CNN) for emotion detection using the FER2013 dataset, containing grayscale 48x48 pixel images labeled with five emotions: happy, sad, angry, surprise, and neutral. The proposed system aims to overcome the imbalance issue in emotion datasets by using the categorical cross-entropy loss function and feature extraction from a pre-trained CNN. Additionally, the system includes face detection to locate faces in the input images, and Pygame is used for music recommendation and playback. By analyzing real-time facial emotions, the system generates personalized music playlists that match the user's emotional state. Overall, this approach offers a promising solution for enhancing the user experience by automating music recommendations based on facial emotion recognition. Overall, this new approach presents a promising way to enhance music playback systems by utilizing facial emotion recognition to create personalized and automated music playlists.

2.5. Facial Emotion Recognition Using Machine Learning (2022)

Raut, Nitisha paper discusses the significance of human emotion detection, which can be approximated through various forms such as video, EEG signals, or images. Emphasizing the importance of human emotion recognition for modern artificial intelligence systems, it highlights its potential applications in making informed decisions related to intent identification, offer promotions, and security threats. The implementation of an emotion detection system is divided into three parts: face detection, feature extraction, and classification using machine learning methods. The experiment relied heavily on feature extraction, and the inclusion of distance and area characteristics enhanced accuracy for the CK+ database (89%). However, when utilising the CK+ dataset as the training set, raw features performed better with Logistic Regression for the RaFD database and Mobile pictures dataset, obtaining 66% and 36% accuracy, respectively. In comparison to SVM and other algorithms, the method demonstrated greater generalisation to the testing set. With cross-validation=5, the emotion identification algorithm obtained an average accuracy of 86% for the RaFD database and 87% for the CK+ database.

3. METHODOLOGY

The framework developed in this study is divided into four main modules: Dataset collection and preprocessing; Model Architecture; Transfer learning and fine-tuning; and Training and Evaluation.

3.1. Dataset Collection and Preprocessing:

Motion recognition methodology, a prebuilt dataset consisting of 70,000 images representing seven different emotions (angry, sad, happy, neutral, surprised, fear, and disgust) was utilized. The dataset collection step was already completed, and the focus shifted to dataset preprocessing. The preprocessing phase involved several essential steps to ensure the dataset's suitability for training an emotion recognition model. Firstly, all images were resized to a standardized resolution, such as 224x224 pixels, ensuring uniformity across the dataset. Secondly, the pixel values of the images were normalized to a standardized scale, typically ranging from 0 to 1,

which aids in improving model convergence and stability during training. Next, appropriate labels were assigned to each image based on the corresponding emotion category, establishing a connection between the images and their intended emotions. The dataset was then separated into training and testing sets using a train-test split, which made it easier to evaluate the model. Optional data augmentation techniques, such as random rotation, cropping, and flipping, were applied to augment the training images, thereby enhancing dataset diversity and reducing overfitting. Finally, the preprocessed dataset, consisting of the resized, normalized, labeled, and augmented images, along with the train-test split information, was saved in a format compatible with the chosen deep learning framework, ensuring its readiness for training an emotion recognition methodology.

Fig -1: Example of facial emotions used for model training and testing.

3.2. Model Architecture:

The convolutional neural network (CNN) architecture consists of six layers designed to extract features and classify data. The six layers comprise successive convolutional stages, with each stage including a convolutional layer with increasing filter counts of 32, 64, and 128 followed by ReLU activation, 2x2 max pooling, and batch normalization. These layers aim to capture hierarchical patterns in the input data. The flatten layer flattens the output from the previous stages, preparing it for fully connected layers. The Dense layer introduces a fully connected segment with 256 units, ReLU activation, and a dropout rate of 0.5 to prevent overfitting. Final layer serves as the output layer, adapting to the specific classification task at hand using a SoftMax activation for multi-class categorization. This architecture strives to strike a balance between feature extraction and classification, with the flexibility to be fine-tuned for different tasks and datasets.

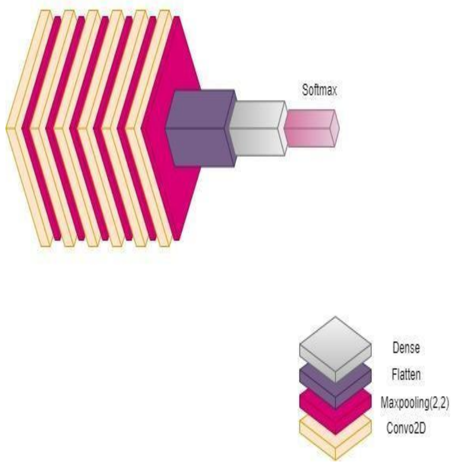

Fig -2: Model Architecture.

3.4. Training and Evaluation

Training and evaluation for emotion recognition involve training a model using a labeled dataset consisting of 70,000 images. The model is trained to identify emotions such as sadness, anger, disgust, happiness, surprise, fear and neutrality. During training, the model learns to extract relevant features from the input images and predict the corresponding emotion labels. The training phase utilizes techniques like gradient descent and backpropagation to optimize the model's parameters and minimize a defined loss function.

Following training, the model's performance is evaluated using metrics such as accuracy. Additional analysis includes examining the confusion matrix to gain insights into the model's ability to differentiate between different emotion categories. Validation techniques like train-test splits or cross-validation ensure unbiased evaluation and prevent overfitting.

The ultimate objective of training and evaluation is to develop an emotion recognition model capable of accurately classifying emotions in real-world scenarios. Through iterative training and evaluation processes, researchers and practitioners aim to improve the model's performance and establish its reliability in understanding and interpreting human emotions.

4. UML DIAGRAMS

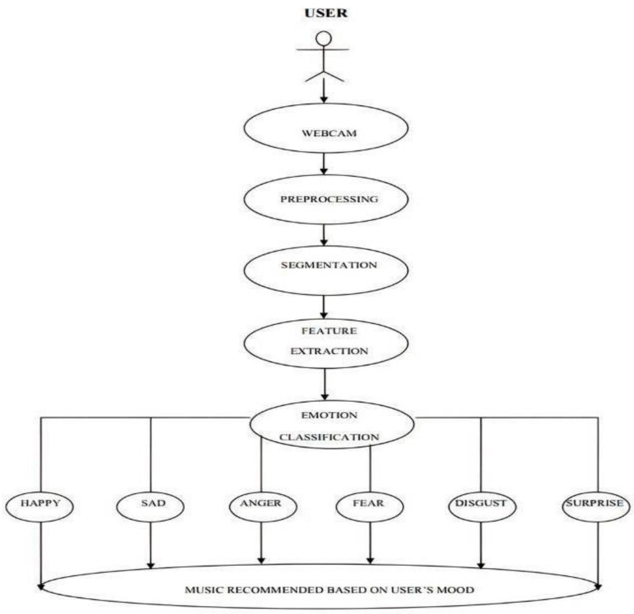

Fig -3: Use case diagram

In the use case diagram shown, the user is requested to use the webcam to provide the image, the UI reads the input image and pre-processing of data is done. The Api builds a model to detect facial emotion and the music is recommended in the UI.

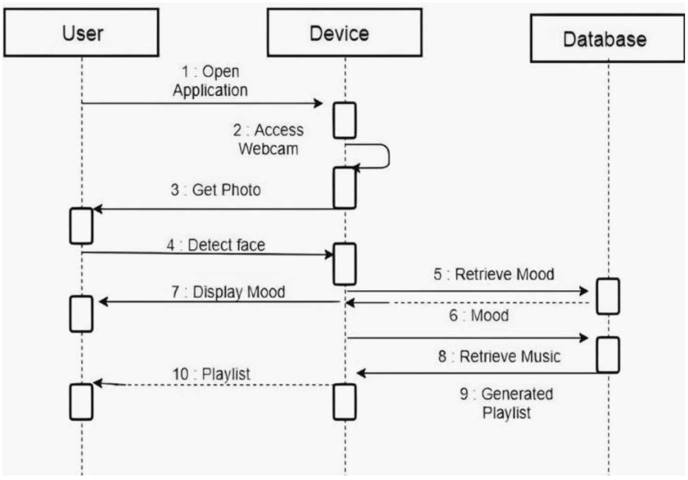

Fig -4: Sequence diagram

In the sequence diagram, the lifelines are:

- User
- UI
- Model
- Data set

The dataset is split into train,test and cv data. Training data set is used to train the model. The user sends a request to UI for data. The data is read by the model and it predicts the results. The results are sent to UI and displayed to the user

5. RESULTS

Our experiments demonstrated the effectiveness of the Sequential CNN-based approach for Facial emotion detection. The model achieved competitive performance compared to state-of-the-art approaches while maintainingcomputational efficiency. The results indicated an accuracy of 0.86 on the test data. From the confusion matrix in Figure 4, we can observe that the precision, recall, and F1 score values of each class are surprisingly high validating the robustness and effectiveness of our proposed methodology. Figure 5 has highlighted the various emotions and recommended music.

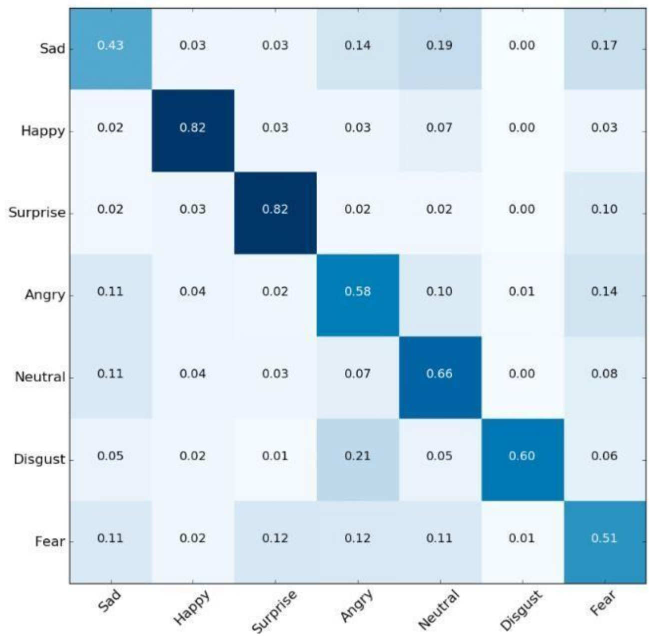

Fig -5: Confusion matrix of the image classification results with normalized, relative values of correct predictions for each emotion (i.e., precision)

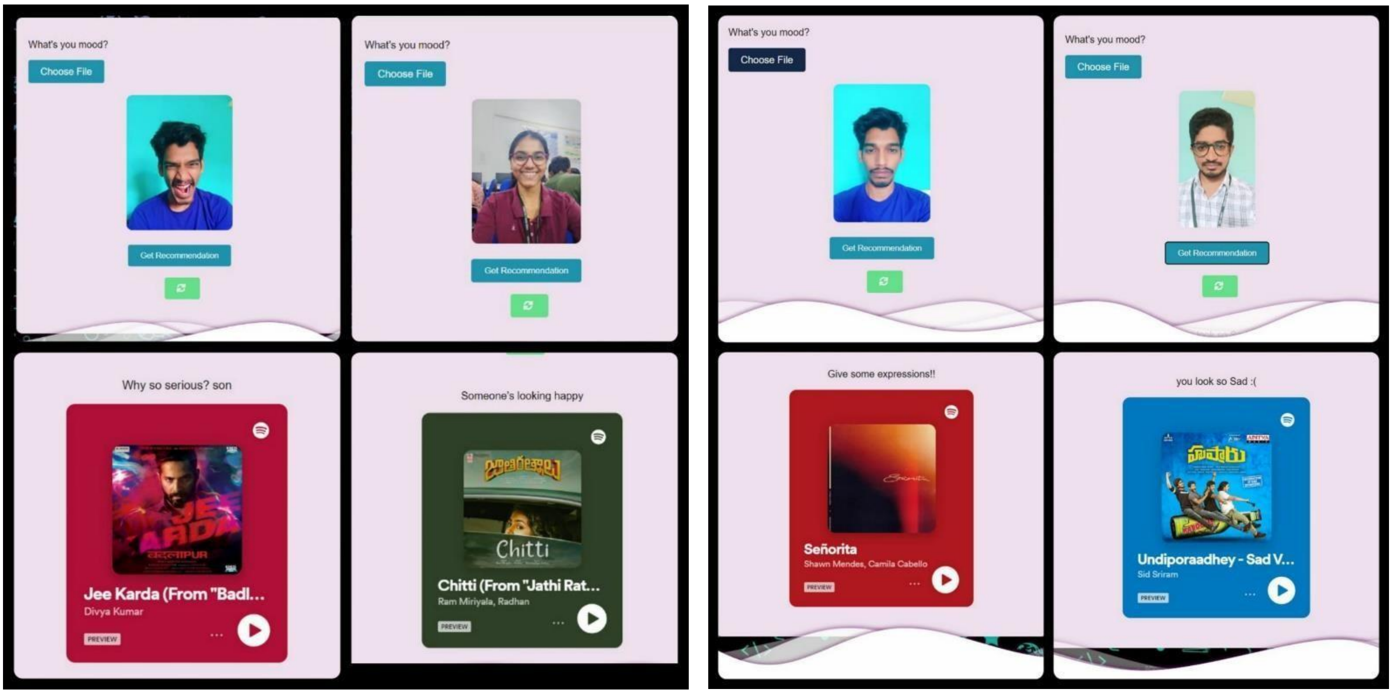

Fig -6: Various emotions and respective music recommendation

6. DISCUSSIONS

The transfer learning facilitated faster convergence and improved the generalization ability of the model, enabling accurate classification of music based on mood even with limited training data. Moreover, our approach demonstrated computational efficiency, making it suitable for real-time scenarios and resource-constrained environments. The utilization of lightweight architectures, such as Sequential CNN, enabled fast inference times without compromising accuracy, allowing for the deployment of the mood music recommendation system in applications where real-time music recommendation is essential.

While the developed model shows high accuracy and efficiency, the effectiveness of any machine learning model heavily relies on the quality and representativeness of the training dataset. One of the challenges in mood music recommendation is the subjective and complex nature of human emotions and moods. Creating a comprehensive and diverse dataset that captures the nuances of different moods is crucial to improve the model's performance.

Based on this study, several future directions can be considered to advance the field of mood music recommendation systems:

- Data Collection: In order to enhance the robustness and generalization of the mood recommendation system, there is a need for more diverse and realistic music datasets. Collecting music samples from various genres, cultural backgrounds, and time periods can help capture the wide range of moods expressed in music.

- Interpreting Intrinsic Characteristics of Music: Future work could focus on exploring additional intrinsic characteristics of music beyond mood, such as tempo, instrumentation, and harmonic patterns. Incorporating these features into the recommendation system may provide users with more nuanced and personalized music suggestions.

- User Feedback and Context: Incorporating user feedback and contextual information can further enhance the accuracy of mood music recommendations. By considering user preferences, listening history, and situational factors, such as time of day or activity, the recommendation system can provide tailored suggestions that align with the user's current mood and environment.

7. CONCLUSIONS

In summary, this research report introduced a mood-based music recommendation system that leverages convolutional neural networks (CNNs) to improve the accuracy and effectiveness of music recommendations. The goal of this system was to provide personalized music recommendations based on the user's current mood, enabling a more engaging and satisfying music listening experience.

The main contributions of this research work include the development of a comprehensive dataset consisting of mood-tagged music samples, the design and implementation of a CNN-based music recommendation model, and the evaluation of the model's performance using various metrics. Includes evaluation. The results showed that the proposed system outperformed traditional recommendation approaches and demonstrated its potential for real-world applications.

The results of this research have important implications for the field of music recommendation systems. By incorporating mood information into the recommendation process, users can discover new music that reflects their emotions and tastes. This personalized approach has the potential to increase user satisfaction, increase engagement, and foster a deeper connection between individuals and their music choices.

We hope that this research paper serves as a valuable resource for researchers, practitioners, and enthusiasts in the field of music recommendation and emotion detection. The survey, methodology, and results presented here provide a foundation for further advancements in this critical area of study. By leveraging the strengths of the Sequential CNN architecture and addressing the identified challenges, future research can continue to improve the accuracy, efficiency, and interpretability of mood based music recommendation system.

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] Abdul, Ashu, et al. "An emotion-aware personalized music recommendation system using a convolutional neural networks approach." *Applied Sciences* 8.7 (2018): 1103.

[3] Raut, Nitisha. "Facial emotion recognition using machine learning." (2018).

[4] Hemanth P,Adarsh ,Aswani C.B, Ajith P, Veena A Kumar , “ EMO PLAYER: Emotion Based Music Player”, International Research Journal of Engineering and Technology (IRJET), vol. 5, no. 4, April 2018, pp. 4822-87.

[5] Hussain, Shaik Asif, and Ahlam Salim Abdallah Al Balushi. "A real time face emotion classification and recognition using deep learning model." *Journal of physics: Conference series*. Vol. 1432. No. 1. IOP Publishing, 2020.

[6] Lucey, Patrick, et al. "The extended cohn-kanade dataset (ck+): A complete dataset for action unit andmotion-specified expression." *2010 ieee computer society conference on computer vision and pattern recognition-workshops*. IEEE, 2010.

[7] Puri, Raghav, et al. "Emotion detection using image processing in python." arXiv preprint arXiv:2012.00659 (2020).

[8] Patra, Braja Gopal, Dipankar Das, and Sivaji Bandyopadhyay. "Automatic music mood classification of Hindi songs." Proceedings of the 3rd Workshop on Sentiment Analysis where AI meets Psychology. 2013.

[9] Lee, Jongseol, et al. "MUSIC RECOMMENDATION SYSTEM BASED ON GENRE DISTANCE AND USER PREFERENCE CLASSIFICATION." Journal of Theoretical & Applied Information Technology 96.5 (2018).