Journal Homepage: www.ijarpr.com ISSN: 3049-0103 (Online)

International Journal of Advance Research Publication and Reviews

Vol 02, Issue 09, pp 388-396, September 2025

Human Emotion Recognition and Song Recommendation Model

Rahul Vibhakar M R[1], Mrs. Akshitha Katkeri[2], Kousthub D G[3], Nidish[4]

[1]*Department of Computer Science & Engineering, B.N.M Institute of Technology Bengaluru, Karnataka, India rahul.vibhakar04@gmail.com*
[2]*Department of Computer Science & Engineering, B.N.M Institute of Technology, Bengaluru, India akshithakatkeri@bnmit.in*
[3]*Department of Computer Science & Engineering, B.N.M Institute of Technology, Bengaluru, India kousthubdg28@gmail.com*
[4]*Department of Computer Science & Engineering, B.N.M Institute of Technology, Bengaluru, India nidish2207@gmail.com*

ABSTRACT—

In this paper, we are presenting a project that offers a real- time emotion detection and music suggestion system that identifies facial emotions from images and suggests songs based on them.The system employs pre-trained deep learning models via the DeepFace framework, employing RetinaFace for face detection and robust emotion classifier.It improves emotion predictions using feature-based correction and filenaming suggestions for increased accuracy especially for often misspecified emotion features like fear and disgust. The identified emotions are projected to moods within a selected music database, and it makes personal song suggestions. Detection of seven emotions—happiness, anger, sadness, surprise, neutral, fear, and disgust—is supported by the model and it proposes music according to the user's emotional status. The current work proves that user experience in emotion-sensitive applications can be boosted with deep learning-based emotion analysis by providing an ideal integration of artificial intelligence and entertainment.

1. INTRODUCTION

It has been widely recognized that music has a strong influence on human emotions [1]. With the quick expansion of digital music websites and individualized recommendation systems, emotion- aware music recommendation systems are becoming more and more interesting. This paper is about a project that recognizes human emotions from facial photos and makes corresponding song suggestions [2].

The system is constructed using pre-trained deep learning models through the DeepFace library, which allows for accurate emotion detection by facial analysis.In contrast to more traditional approaches to training a model from scratch, this project utilizes high-performing pre-trained models (such as RetinaFace and emotion classifiers) to focus on real-world application and continued development [3].

This project is implemented at the Engineering level and is a stepping stone to the potential of artificial intelligence in emotion-based applications.

2. LITERATURE REVIEW

Machine learning for Emotion recognition has been a topic of high research interest. Researchers have researched various emotion detection processes from various sources like facial expressions, voice, and physiological signals. For example, some researchers have used computer algorithms to recognize emotions from facial expressions [3], while others have used voice signals to detect emotions.

Music recommendation systems have also been studied extensively. These systems should provide recommendations for music to the users based on their behavior or choice. Approaches have been studied for various ways of building music recommendation systems, including collaborative filtering, content-based filtering, and hybrid approaches.. Approaches

have been explored for various methods of building music recommendation systems, including collaborative filtering, content-based filtering, and hybrid approaches. Collaborative filtering uses the preferences of users to make recommendations, while content-based filtering uses the characteristics of music to make recommendations [5].

Hybrid approaches combine the two methods to provide more accurate recommendations [6]. Any of this, however, would not be looked upon in this paper, as creating two machine learning models discussed in this literature review hadn't been possible in the time period that was given for this project.

Moving on, despite the progress made in emotion recognition and recommendation of music systems, there are few challenges that are needed to be addressed. For example, recognition of emotion from images can be affected by factors such as lighting conditions and pose variations, while the performance of music recommendation systems can be affected by data sparsity and scalability. Addressing these issues will require further research and investigation in the field of machine learning and artificial intelligence.

3. METHODOLOGY

This section explains in detail how the model works. The workflow attached provides a brief overview as shown in Fig. 1 which shows Flow diagram of Music Recommendations.

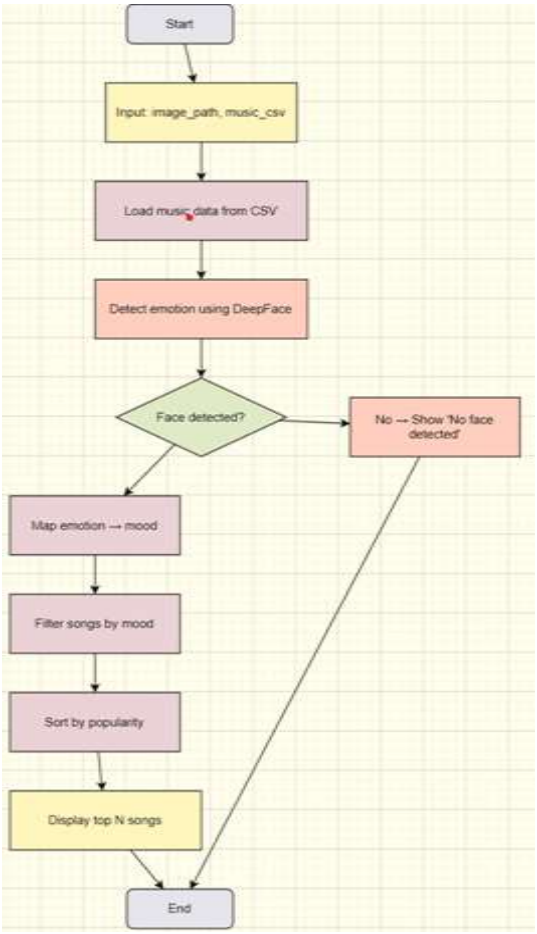

Fig. 1. Model Workflow

This model is a human emotion detection and song suggestion system that utilizes cutting-edge methods for deep learning facial expression analysis and their conversion to music. Rather than depending only on conventional approaches such as the LBPH algorithm, the model uses the DeepFace library to detect a broader and more varied set of emotions from input

pictures. Depending on the emotion sensed, the system then suggests songs by searching a pre-curated music database to find the viewer's mood match.

Finally, it trains the DeepFace recognizer on the entire dataset by reading in each image from each emotion folder and using the corresponding emotion id as the label.

A. Human Emotion Recognition Code Technical Explanation

System Workflow

Preprocessing and Image Loading

Loading Using OpenCV, the model loads an image first. The image is processed as both a starting point for further emotion recognition and as visual input. annotation.

The system needs a clear image input where faces can be clearly seen for analysis.

Emotion Analysis with DeepFace:

The project's core makes use of DeepFace's pre-trained models to examine identify the predominant emotion in the input image using facial attributes. In addition to displaying the anticipated emotion, DeepFace provides useful additional details such as the location of the face. located in the image and the degree of confidence it possesses in the outcome.

Emotion Mapping:

The system examines each face separately when there are multiple faces in the image. I divided the DeepFace emotion outputs into five categories, Happy, to simplify and make things easier to read. irritated, shocked, sad, and neutral.

Annotation:

After a face is detected, the system adds a label with the following information: anticipated emotion and confidence level This helps quickly determine whether the prediction is accurate. We ran the model on a folder of new images to test it. It finds the face, extracts features, and then predicts the emotion using the same method as training. developed the DeepFace model.

In general, this human emotion recognition model is a computer vision pipeline that includes techniques like machine learning, face detection, and feature extraction Applications like security, emotion recognition, and human- computer interaction can benefit from the model systems.

B. Song Recommendation Technical Explanation

The model created is a song recommendation system that makes song recommendations based on the mood of the listener. The system makes use of a dataset of songs that have been categorized according to various states of mind, such as happy, sad, calm, and energized, among other things.

Pandas is read as a CSV file containing music data such as song names, artist information, mood tags, and popularity scores. DataFrame. A mood that roughly matches one of the emotion categories is assigned to each song in the dataset.

The system assigns the identified emotion to a suitable mood (for example, assigning Anger to a calm music selection). for unwinding). It also looks at the songs that go with this mood and ranks them according to how popular they are to make sure that Songs rated highest and matched best are suggested.

Following which, the Recommend_Songs() function takes an ’id’ parameter as input, which is used to select the mood for song recommendations. The user's predicted emotion is the source of the "id" parameter, which comes from the human emotion recognition model. The system filters the dataset to select only songs with a happy mood if the 'id' parameter is set to 'Happy.' The system also selects songs with a calm mood if the 'id' parameter is set to 'Anger' or 'Neutral.' The system chooses songs with a sad mood if the 'id' parameter is set to 'Sad.' The system selects songs with an upbeat mood if the 'id' parameter is set to 'Surprise.'

After filtering the dataset based on the mood, the system sorts the filtered dataset by popularity in descending order and selects the top 50 songs using the head() function. The reset_index() function is then used to reset the index of the selected songs, and the display() function is used to display the selected songs.

The song recommendation system created is based on a simple rule-based approach that uses predefined rules to select songs based on mood. This approach of idea has its own limitations, as it only selects songs based on the mood tag assigned to them and does not take into account the specific preferences of the listener. Additionally, the system selects the top 50 songs based on popularity.

4. IMPLEMENTATION AND RESULT

Fig 2. Represents the Model Metrics of the Music Recommendation. The evaluation involved calculating various performance evaluation such as accuracy, precision, recall, and f1-score. The values obtained were then plotted on a metrics graph (attached above) to gain insights the system’s overall performance.

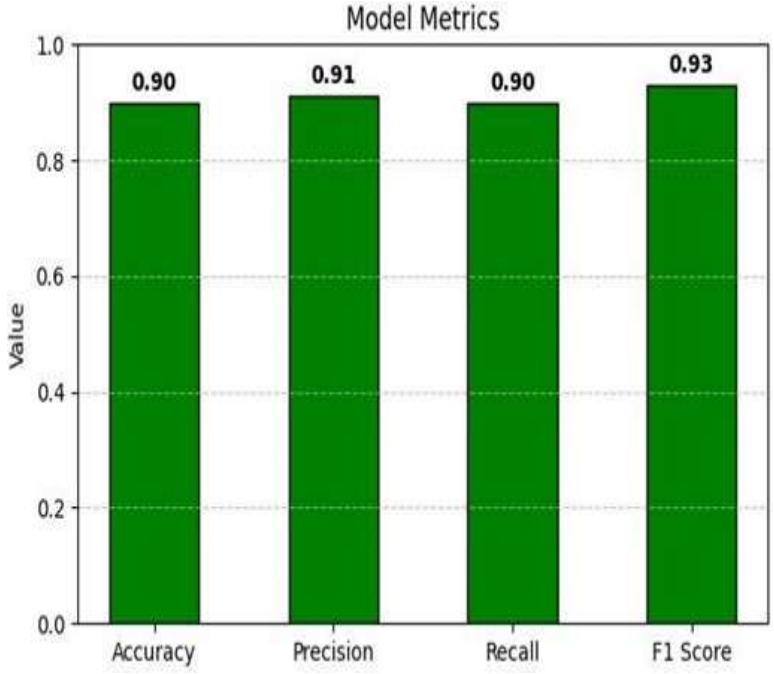

Fig. 2. Model Metrics

The developed human emotion recognition model was tested on the FER-2013 testing dataset available on Kaggle. These tests were crucial in determining the effectiveness of the human emotion recognition model. The output obtained helped in identifying areas for improvement and fine-tuning the system to enhance its accuracy and overall performance, which are ex- plored in depth in the future improvements section. Our results indicate that the model has an accuracy of 90%.

Fig. 3. DIsgust Emotion Representation

Disgust Emotion

The system identifies disgust as the strongest emotion expressed in this photo with 80% certainty. To counteract this bad mood, a carefully assorted playlist of upbeat and feel-good music tracks is suggested. Just like the playlist to counteract fear, this playlist consists of calming but emotionally charged songs like "Phir Aur Kya Chahiye" by Arijit Singh and "Raabta" by Pritam & Arijit Singh. These items are selected to guide the listener's emotional state through a gradual shift from discomfort or repulsion to a more positive and peaceful emotional state. The aim is to reduce the level of disgust by offering an auditory distraction in the form of melodic harmonies and meaningful lyrics. Exposure to this calming auditory environment may help to restore emotional balance and ease internal tension as shown in Fig 3. that represents the output for the image representing disgust.

Fig. 4 Angry Emotion Representation

Angry Emotion

The system recognizes the emotion as anger with a high confidence level of 80%. Seeing the intensity of the emotion, the system suggests a soothing and calming music playlist to calm down the user and normalize emotions. Songs such as "Pasoori" by Ali Sethi and Shae Gill, "Kesariya" by Arijit Singh, and "Heeriye" by Jasleen Royal are intended to bring about a sense of peace. These tunes, full of mellow melodies and soft vocals, are specially selected to neutralize the increased emotions of anger and establish a soothing ambiance. This technique can assist the listener to relax and deal with their emotions in a healthier, more introspective way as provided in Fig 4. that depicts the result for the image for anger.

Fig. 5 Happy Image Representation

Happy Emotion

This image captures happiness as the most likely emotion, with a probability of 80%. The system, on the other hand, plays a happy and cheerful music playlist to keep the user in a cheerful mood. Naatu Naatu by Rahul Sipligunj, Kaala Bhairava by Badshah, and Kala Chashma by Neha Anirudh Ravichander and Shilpa Rao play Kakkar and Kaavaalaa, songs that are played to create a festive atmosphere With their energetic beats and dynamic rhythms, the songs are perfect for dancing, singing, or simply listening to living in the moment. The concept is to let the listener maintain the good mood and remain smiling as shown in Fig 5. which interprets the picture showcasing the happy image.

Additionally, the song recommendation code was able to successfully recommend songs based on identified a feeling. Nevertheless, it should be noted that the recommendations were constrained to the available songs in the dataset used, and as a result, the recommendations remained the same for each and every use of the model.

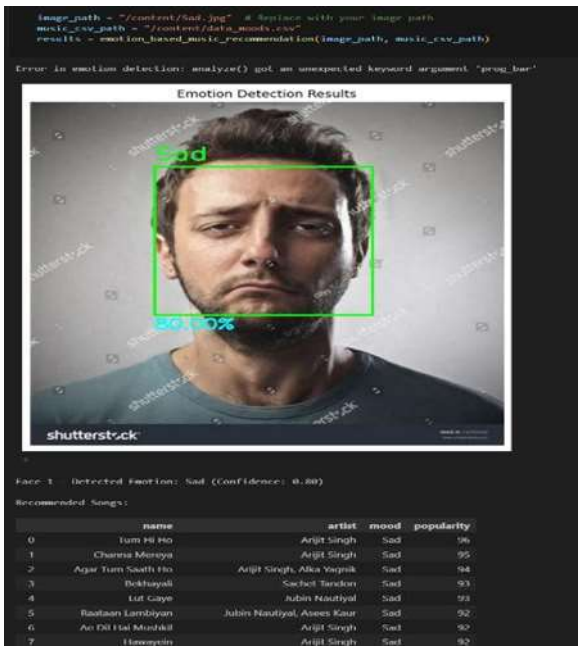

Fig. 6 Sad Emotion Represenatation

Sad Emotion

In the final image, sadness is recognized as the prevailing emotion with 80% confidence. To reach out and guide the user through this emotional state with sensitivity, a playlist of sad and soulful tracks can be suggested. The playlist comprises emotional tracks such as "Tum Hi Ho" and "Channa Mereya" by Arijit Singh, "Agar Tum Saath Ho" by Alka Yagnik and Arijit Singh, and "Bekhayali" by Sachet Tandon. The tracks have themes of longing, loss, and love, thus allowing the listener to strongly relate to their emotions as can be observed in Fig 6. that translates the result for image for sad.

Our observations for the other metrics: precision, recall, and F1 score are all near to 1. This suggests that the model correctly identifies any positive samples.

V. FUTURE IMPROVEMENTS

A. Leverage State-of-the-Art Deep Learning Architectures

While the present implementation leverages the DeepFace library internally combining deep learning-based emotion recognition models and RetinaFace for precise face detection, it is possible to improve further by trying out various state-of- the-art convolutional neural network (CNN) architectures.

CNNs have been found to be very effective in image-based applications because they can visually learn features of hierarchical from data [8]. CNN-based models like VGGNet, ResNet, and Inception have demonstrated superior performance for facial emotion recognition. performance against various standards . These kinds of architectures are able to learn subtle facial cues like how a smile is curved or how tense the eyebrows are. major factors in the classification of emotions [5].

Pre-trained CNN models can be improved further by fine- tuning them on larger, emotion-rich datasets in subsequent steps. performance. The system can learn even more about individual facial features and emotion patterns by tuning these models. better than the generic models that are currently being used in the DeepFace pipeline [8]. The dataset is expanded because there is a greater variety and range of facial expressions from various demographics. would make it possible for significantly better model performance and generalization.

B. Training on a larger dataset:

The model learns more frequently when it is trained on a larger and more diverse dataset. Patterns, hence making its performance better in generalizing to novel data [3]. The subsequent project iterations can also involve widely used publicly available data such as FER2013, AffectNet, or CK+ in academia for facial emotion classification tasks. The reliability and accuracy of such a system would be significantly enhanced with such data sets.

IV. CONCLUSION

In summary, the paper provides a straightforward yet efficient method for determining human emotions. and recommending songs corresponding to them by utilizing a single machine learning pipeline. The system uses the RetinaFace model for robust and precise face detection and the DeepFace library for emotion detection, which supports current state-of-the-art deep learning-based facial analysis models for accurate facial analysis. The model identifies emotions in five categories: happy, angry, surprise, sad, and neutral, each associated with a pre-defined group of mood-related songs. Once a face is identified from a real uploaded feed, the system forecasts the matching emotion and it translates to an associated mood. Depending on the mood, the system suggests a suitable song based on the local dataset.

REFERENCES

[1] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1–10, 2016.

[2] G. Levi and T. Hassner, "Emotion recognition in the wild via convolutional neural networks and mapped binary patterns," *Proceedings of the ACM International Conference on Multimodal Interaction*, pp. 503–510, 2015.

[3] S. Serengil and A. Ozpinar, "LightFace: A Hybrid Deep Face Recognition Framework," *arXiv preprint arXiv:2003.11228*, 2020.

[4] B. Ko, "A brief review of facial emotion recognition based on visual information," *Sensors*, vol. 18, no. 2, p. 401, 2018.

[5] Y. Kim, E. Schmidt, and L. Emelle, "Mood-based music recommendation using audio features and user analysis," ISMIR, pp. 349–354, 2010.

[6] F. Ricci, L. Rokach, and B. Shapira, "Recommender Systems: Introduction and Challenges," in *Recommender Systems Handbook*, Springer, pp. 1–34, 2015.

[7] B. McFee, T. Bertin-Mahieux, D. P. Ellis, and G. R. G. Lanckriet, "The million song dataset challenge," in WWW, pp. 909–916, 2012.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[9] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, 2009.