iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** II **Month of publication:** February 2025

DOI: https://doi.org/10.22214/ijraset.2025.67020

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue II Feb 2025- Available at www.ijraset.com

Emotion Based Music Recommendation System

Khushali Pandav[1], Neramaye N Deshpande[2]
[1]*Student, Department of Computer and Information Science, Nagindas Khandwala Collage , Mumbai, Maharashtra, India*
[2]*Assistant Professor, Department of Computer and Information Science, Nagindas Khandwala Collage, Mumbai, Maharashtra, India*

Abstract: ***The integration of emotion state and music recommendation systems holds a high potential for enhancing experiences through mapping emotion state with preference for music. In this work, an emotion state-dependent music recommendation system is proposed, utilizing face expression recognition for determination of emotion state and real-time suggesting of music in relation to it. With Convolutional Neural Networks (CNNs) and the FER-2013 dataset, emotion categories including happiness, sadness, anger, surprise, and neutrality are determined. A dynamic module then retrieves personalized playlists of music in relation to detected emotion, offering real-time, adaptable recommendations to users. Computer vision techniques such as Haar Cascade classifiers are utilized for face feature extraction and emotion classification through machine learning algorithms. Outcomes are represented through confusion matrices, classification reports, and emotion distribution plots, providing an insight into performance of the system. The proposed model introduces a new direction for personalized consumption of music, enhancing emotion well-being through music and opening doors for future emotion AI breakthroughs, with implications in mental wellness and entertainment.***

Keywords: Emotion Recognition, Convolutional Neural Networks, FER-2013 Dataset, Music Playlist, Face Detection, Real-time Suggestion, Computer Vision, Feature Extraction.

I. INTRODUCTION

Music plays an extremely important role in enhancing one's life because it is an excellent source of delightment to fans of music and audiences alike. With today's reality of instant digital media and technology, there are players that exist that have playback with fast forward backward, differentiated playback pace, i.e., seeking and compression of time, on-demand playback, streaming playback with multicast streams, volume adjustment, category, and so on. Since the alternatives mentioned above fulfill only the user's basic need, dynamic browsing of the playlist of music and selecting tracks that go according to his mood and attitude are to be done by the user. In this approach, extra effort for the selection of music that displays the moods of the user is avoided by forecasting the emotions of the user and recommending songs according to them. Face capture and emotion detection/recognition give users an edge in surmising how a person is feeling using deep learning. A perfect method for finding or trying to know your emotions is through a check on your facial expression. Because each human face comprises unique demographics, the model can analyze and distinguish any emotion sooner, reducing calculation time. Thus, face expression is the most successful technique to anticipate emotion, allowing us to offer suitable music. At a time when digital music technology is advancing with each passing day, there is an increasing need to develop a personalized music recommendation system that would recommend songs to the users. It is a unique task to make recommendations based on the immense data available over the internet. E-commerce giants such as Amazon and eBay give personalized recommendations to customers based on their tastes and history, and Spotify and Pandora rely on Machine Learning and Deep Learning algorithms to suggest relevant recommendations. Some research has been conducted on personalized recommendation of music to suggest songs on the basis of predilection of the user. The motive of this essay is to improve the mood of the user by listening to music that fulfills the requirement of the user and by taking snapshot of the user. Recognition of facial expression has long been regarded as the best form of estimation of the expression known to humans. Facial grimaces are the most effective way for people to elucidate or regulate the emotion, frame of mind, or thoughts that another person is pursuing to suggest. In some situations, mood shifts can help patients overcome despondency and melancholy. Most of the health risks can be prevented by expression analysis, and efforts can be made to improve the user's mood .

II. RESEARCH OBJECTIVES

1) To create a system that can detect and understand a user's emotions through their facial expressions.
2) To build a recommendation engine that matches music to the emotional mood of the user.
3) To develop a live, interactive interface that updates music suggestions based on real-time emotional feedback.
4) To design an easy-to-use interface that encourages users to interact while offering immediate music recommendations.

©IJRASET: All Rights are Reserved

III. LITERATURE REVIEW

Athavale. A, et. al. developed Emo Player, innovative software that automatically streams tracks on the basis of mood of the user. It detects the mood of the user through face expression and streams tracks accordingly. The mood identification is done through a machine learning technology known as the SVM algorithm. The outer part can possibly be the basic part of an individual body, and it also helps much to expose the feelings and actions of the individual. The camera snaps a snapshot of the user. It extracts the face of the user out of the snapshot taken by it. Facial expression is of two natures like smiling or non-smiling. Depending on the impression, music square measure contests prearranged directories.

Athavle. M, et. al. (2021) the paper describes a novel technique to computerized music selection based on face emotion recognition, with the aim of escalating efficiency and user satisfaction.The system detects emotions from facial expressions using CNNs in real time, unlike in previous approaches. Recommendations of ideal music based on the observed emotions are provided using the Pygame and Tkinter packages. A few of the key advantages are lesser processing times, lesser cost of the system, and therefore higher accuracy. The method will be verified by using the FER2013 dataset and proving its caliber in distinguishing happiness, sadness, and neutrality. This technique outperforms current algorithms in terms of refining efficacy by designing music playlists automatically based on the user's emotional state. This work has marked an important step toward exploiting face expression recognition for personalized music experiences, with higher user engagement and contentment.

Carolin. T, (2022) the system here suggested identifies feelings using a Support Vector Machine
technique with supervised learning of a novel form of model. The Olivetti face dataset is employed here to train faces, which have 400 faces along parameters or values. The face is held by attributes of the user or extracts them from an already taken image. The process of training is triggered by random values and iterates through to where estimated values coincide with values of the model training model is set up to test its performance and effectiveness upon its use in factual circumstances.

Florence. M, & Uma. M, (2020) train set that is employed to train the classifier of this study is called CohnKanade Extended (CK+). It contains 593 facially coded sequences of 123 individuals. The categorization is engaged to describe what is conveyed by the subject's expression. It also has intimate to 200 images of the HELEN dataset that were employed to train the classifier. In addition to images, there was also one .txt file corresponding to each image that contained the cluster defining 164 landmark points on that image

Mahadik. A, et. al. (2012) much of the research is ongoing in the area of computer vision machine learning where machines are trained to make estimates of diverse types of feelings of humans moods. Machine learning implies a variety of advancements for human feeling recognition them is mobile-Net using Keras, that built very small trained model and boosted
Android-ML integration.

More. I, et. al. the most challenging task of music listening based on our mood is choosing the exact tune, which can be solved using pragmatic CNN techniques that classify users' moods reliably. The issues that the Facial Expression Recognition system should identify are face detection and positioning in chaotic photos, facial feature extraction, and expression assortment. The successful model classifies, after training, emotions into angry, happy, sad, and neutral with many variations.

Samuvel. D, et. al. (2020) there exist various technologies, which can already identify facial emotion. However some systems suggest music. The overall purpose of the article is to develop a system that will recommend tracks on the emotional state of how the user is feeling, through face-based emotional ratings.Emotion recognition could also be integrated on robots to interpret sentiments without requiring humans interference.

Sarah. S, et. al. (2020) it was stated in this paper that the emotional state of the user is detected by the usage of the Google Mobile Vision SDK. The detected emotion state is then fed to the Expression -X algorithm which prioritizes the music according to the fed emotion value and prepares a playlist tailored to the emotional state of the user. The forecast should be tough to reach 100%, as everyone asseverates emotions differently; however, it did get a 70-75% success rate in detecting the correct mood state of the user and delivering the proper set of music recommendations after its repeated testing.

IV. METHODOLOGY

We built our Convolution Neural Network model on an application of Kaggle dataset. Database is FER2013 that is split into two sets of train and test dataset. It has 24176 in train dataset and 6043 in test dataset. It has 48x48 gray scale face images in the dataset. It has one label of face out of five, i.e., happy, sad, angry, surprise, and neutral, within FER-2013.

The faces are enrolled automatically in such a manner that they're approximately in the same position in each photograph and occupy approximately the same portion of space. There're posed and unposed headshots in photos in FER-2013, and both of them're grayscale and 48x48 in resolution.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

FER-2013 dataset was collected through harvesting output of a Google search for each emotion and emotion synonyms. FER systems trained with an unbalanced dataset can function well for dominant ones such as happy, sad, angry, neutral, and surprised but not well for under-represented ones such as disgust and fear.

Usually, the issue is resolved using weighted-SoftMax loss by weighting loss on each of the emotion classes supported by its ratio within training sets. However, that kind of weighted-loss scheme is dependent on utilizing a SoftMax loss function, where that function is reported to have a strong bias to make features of diverse classes separated without considering intra-class compactness. A good mechanism for overcoming the challenge of SoftMax loss is through employing an auxiliary loss to train a neural network. For managing missing and Outlier values, we have utilized a loss function by namecategorical crossentropy. A loss function is employed within one iteration to calculate the value of error treating missing values and outlying values, we have employed a loss function of categorical crossentropy.

A. Face Detection

Face detection is one of the methods of computer vision technology. It is how algorithms and training is created and trained to detect faces or objects within an object detection or like system within images accurately. It is possible to detect real-time within a video frame or within an image. The face detection utilizes such classifiers, algorithms that tag a face (1) or is not face (0) within an image. Images are arranged to train face detection to have accuracy. OpenCV utilizes two classifiers, LBP (Local Binary Pattern), and Haar Cascades. It is also an use of use of a Haar classifier within face 11 detection where a classifier is trained using predefined variable face data, hence can detect variable faces accurately. The face detection is aiming to detect face within the frame by filtering out outside noises and other issues. It is a process within a use of use of machine learning where a cascade function is trained using a set of input files. It is supported by use of use of the Haar Wavelet process within an observation of analyzing pixels within an inner part of an image into squares through function. It utilizes use of use of machine learning processes within reaching out to have good amount of accuracy out of what is called "training data"

B. Feature Extraction

While performing feature extraction, employ pre-trained network, sequential model, as random feature extractor. Forward pass your input image through it up to predefined layer, and employ its result as our feature. Initial convolutions of Convolutional network extract top-level feature out of taken image, hence employ fewer filters. While advancing deeper, make more deeper layers, and for that double/triple previous layer's filter.

Filters of deeper layer have more feature but computationally very expensive. By doing that, have utilized strong, discriminative feature learned by Convolution neural network. Model result will be feature map, intermediate form for all subsequent layers up to first one. Load your input image for which you need to visualize its feature map to know, for its classification, what feature had prominent role. Feature map is produced through use of Filter or feature detectors on input image or feature map result of previous layers. Visualizing feature map will unveil information regarding inner representations for particular input of each Convolutional layer of model.

C. Emotion Detection Figure

Convolution neural network Architecture. Convolution neural network architecture makes use of feature detectors or filters of input image to accept feature maps or activation maps using Relu activation function. The feature detectors or feature extractors are used to search for diverse features of an image like edges, vertical lines, horizontal lines, bends, etc. Pooling is done on feature maps to achieve translation invariance. Pooling is expected on the fact that whenever we alter the input by a little amount, then the pool of output remains same.

Any of the poolings can be employed like min, average, and max. But max-pooling will give better performance compared to min and average poolings. All of our inputs are flattenings and providing these flattenings to a deep neural network which are sent to the class of object. 12 Classification of an image will be binary or multi-class classification to recognize a digit or to recognize among diverse apparel items. Neural networks are like a black box, and Neural Network's learned feature is non-interpretative. Thus, actually we pass an input image then the CNN model will give out the result. Emotion is identified using loading of a model trained using a weight using CNN. Once, whenever we take real-time image using a user then that image passed to pre-trained model of CNN, then classify the emotion and label it using adding label in an image.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

D. Music Recommender Module

Songs Database We have developed Hollywood songs' database. Everyone is aware that part of changing our mind set is also accomplished by music. So, if a user is sad, our system will provide him or her that type of music playlist that will motivate him or her and through that, of course, will make him or her happy with its mind set.

Music Playlist Recommendation Real-time mood of a user is recognized by utilizing emotion module. It will label like Happy, Sad, Angry, Surprise, and Neutral. Once the feeling is recognized, the system will employ it to fetch a set of corresponding songs of that feeling. The recognized feeling is translated in the music_dist dictionary to its corresponding CSV file, i.e., songs/angry.csv for anger or songs/happy.csv for happiness. The function music_rec() will employ respective CSV of recognized feeling, importing respective columns like Name, Album, and Artist. After filtering out, 15 of the tracks are taken and sent out as custom recommendations to that person, keeping that person's emotional situation in mind.

E. System Architecture

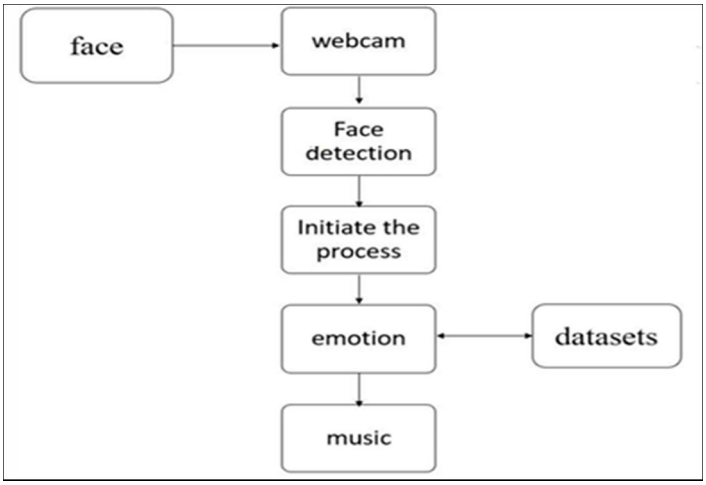


Fig: Flow diagram of proposed system

System Architecture describes overall model structure under development. It is composed of sequence of steps in the system.

Step-1: Load the Dataset

Import the dataset i.e., FER-2013, face expression in terms of pixels (csv format) and have been labelled with respective emotion labels.

Step-2: Pre-Processing

The pre-processing is performed through importing ‘label encoder’ from sklearn.

The same can involve reshaping of pixel values in terms of image arrays and then normalize them.

Step-3: Feature Extraction

Extract desired features of the dataset such as respective emotion and respective pixels.

Step-4: Feature Selection

Select respective features of the dataset that model will use such as sad, happy and neutral emotion.

Step-5: Train & Test Split

Split the data into train and test.The train set is 80%, test set is 20%.

Step-6: Model Application

Utilize trained model for predicting emotion from user image .The model will produce predicted emotion label

Step-7: Input Image User
User takes image through a web cam integration in order to predict emotion and suggest a respective song according to predicted emotion.

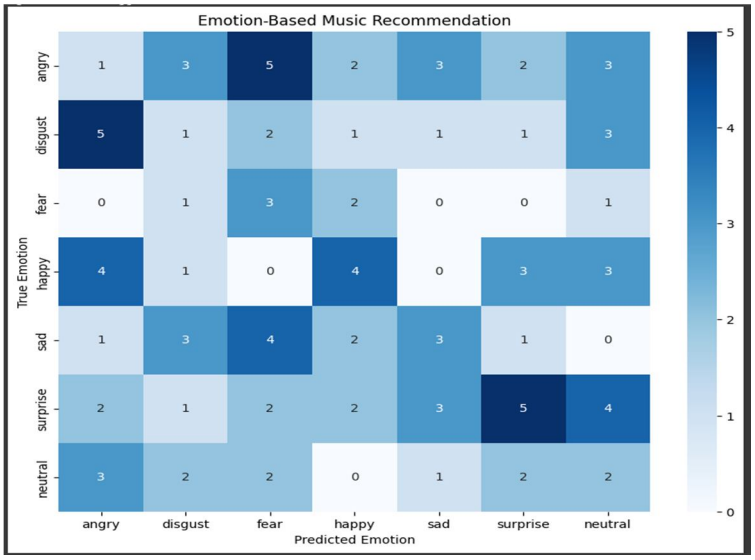
Step-8: Deep Learning Techniques
In Deep learning model we used CNN techniques.

Step-9: Model Evaluation
Analyze the performance of the emotion detection model and overall music recommendation system.

Step-10: Predicting Music
Based on the emotion predicted, the system queries the music database. After filtering, 15 songs are extracted and returned as personalized recommendations for the individual.

V. RESULTS

A. Confusion Matrix



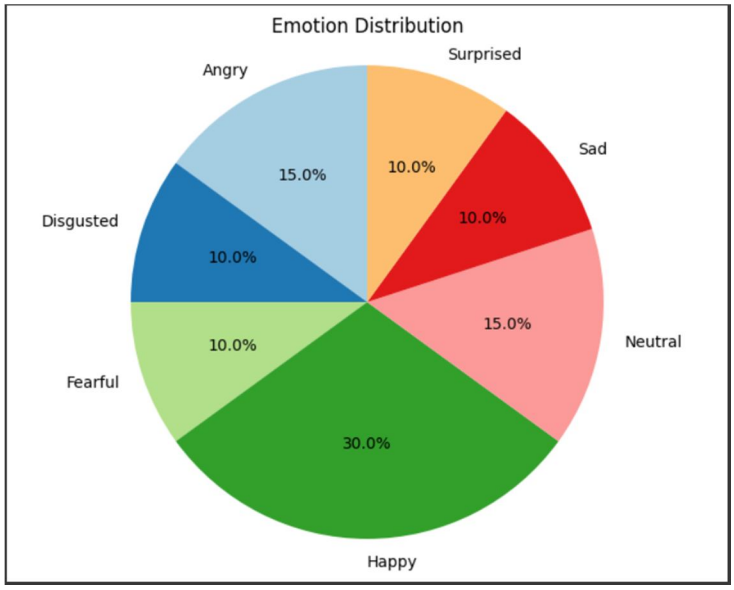
The provided code is an illustration of creating and plotting a confusion matrix for testing an emotion-based music recommendation system performance.It first creates seven emotional tags: angry, disgust,
fear, happy, sad, surprised, neutral. The lists of true_emotions and predicted_emotions are filled with samples drawn at random from these labels to simulate a real scenario (you'd use actual values in practice, not samples drawn at random). The confusion_matrix function in the module of sklearn.metrics then compares with accuracy with which predicted emotions map onto actual ones. The generated confusion matrix is then printed with a heatmap, produced with the use of the seaborn library, in which intensity of colors represent correct and incorrect prediction occurrences.

B. Classification Report

	precision	recall	f1-score	support
angry	0.14	0.09	0.11	23
disgust	0.08	0.04	0.06	23
fear	0.32	0.46	0.38	13
happy	0.14	0.17	0.15	12
sad	0.07	0.20	0.10	5
surprise	0.00	0.00	0.00	16
neutral	0.00	0.00	0.00	8
accuracy			0.12	100
macro avg	0.11	0.14	0.11	100
weighted avg	0.11	0.12	0.11	100

A classification report to evaluate an emotion-predicting model's performance with regard to music tracks.Emotion labels have been assigned using seven categories of angry, disgust, fear, happy, sad, surprise, and neutral.The lists of true_emotions and predicted_emotions have been filled with random samples of these labels in an attempt to simulate a real case (which in practice, one'd use actual data for). The function classification_report in the module sklearn.metrics generates key statistics such as precision, recall, and F1-score for each emotion label, and a complete report of model performance is printed out. Target_names is an optional parameter in the function that will cause the report to include these statistics for each emotion label. This classification report can then be utilized to evaluate how well an emotion prediction model is working and develop an idea about areas for improvement.

C. Emotion Distribution in Pie Chart



Visualization of predicted emotion distribution in a pie chart for emotion_dict is a function to map numerical labels (0 to 6) to specific emotion, for example, "Angry," "Happy," "Sad," etc. predicted_emotions is a list of predicted emotion for a collection of samples (e.g., songs). It then lists occurrences of each emotion in predicted_emotions in a list comprehension. It then plots a pie chart representing distribution of predicted emotion using matplotlib with labels for emotion name and distribution in terms of percentage displayed using autopct='%1.1f%'. plt.axis('equal') is used to make a circle out of the pie chart, and plt.show() to plot it. This visualization helps in simply estimating how frequently a certain emotion occurs in prediction.

VI. CONCLUSION

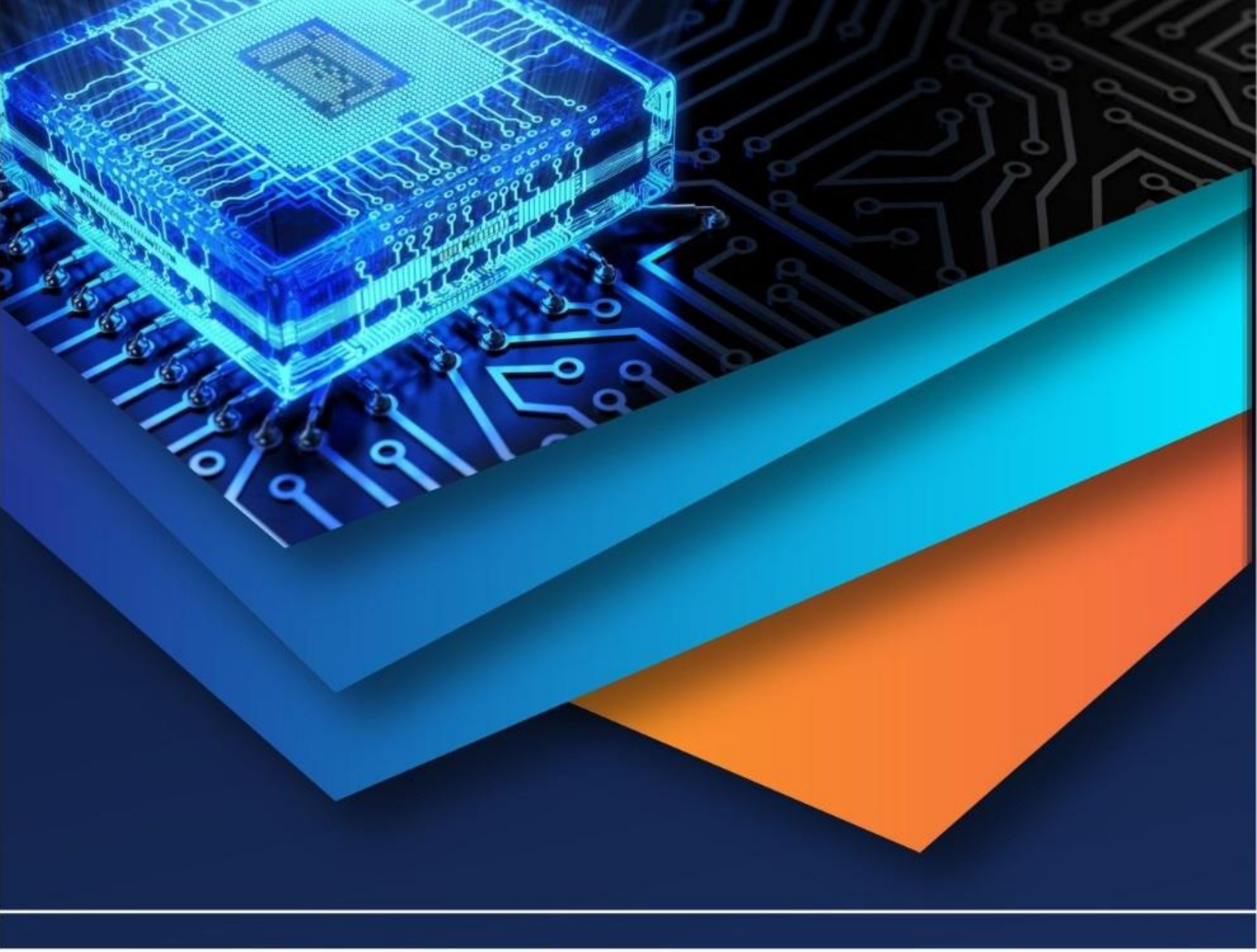
Emotion music recommendation seamlessly fuses emotion recognition via facial expression and personalized music recommendation to maximize the user's enjoyment. By leveraging deep learning algorithms such as Convolutional Neural Networks (CNNs) for emotion recognition and mapping onto a personalized playlist, real-time, dynamic recommendations resonate with the state of emotion in the user. Computer vision integration with databases of music creates an interactive and intuitive platform for personalized music recommendation in relation to the state of mind in the user, offering a new model for consumption of music. Not only can such a system enable improvement in the state of emotion in the user through music, but it can lay the basis for future breakthroughs in emotion AI, with potential application in a variety of sectors such as mental wellbeing and user-focussed entertainment. As technology continues to evolve, it can become increasingly sensitive and attuned, offering a personalized service specifically designed for individual emotion needs.

REFERENCES

[1] P. V Shitole, O. Kawade, P. Kamble, H. Chavan, and A. Athavale, "Emotion based Music Player Department of Computer Engineering." [Online]. Available: www.ijsrd.com
[2] M. Athavle, D. Mudale, U. Shrivastav, and M. Gupta, "Music Recommendation Based on Face Emo-tion Recognition," Journal of Informatics Elec-trical and Electronics Engineering, vol. 02, no. 018, pp. 1–11, 2021, doi: 10.54060/JIEEE/002.02.018.
[3] T. J. Carolin, "EMUSIC USING SUPPORT VECTOR MACHINE LEARNING ALGORITHM," 2022, doi: 10.22159/ijet.2022v10i1.46865.
[4] S. Metilda Florence and M. Uma, "Emotional Detection and Music Recommendation System based on User Facial Expression," in IOP Conference Series: Materials Science and Engineering, IOP Publishing Ltd, Sep. 2020. doi: 10.1088/1757-899X/912/6/062007.

[5] A. Mahadik, V. Bharathi Jagan, S. Milgir, V. Kavathekar, and J. Patel, "Mood based music recommendation system," 2021. [Online]. Available: https://www.researchgate.net/publication/352780489
[6] I. More, V. Shirpurkar, Y. Gautam, and N. Singh, "Melomaniac-Emotion Based Music Recommendation System." [Online]. Available: www.ijariie.com
[7] D. J. Samuvel, B. Perumal, and M. Elangovan, "Music recommendation system based on facial emotion recognition," 3C Tecnología_Glosas de innovación aplicadas a la pyme, pp. 261–271, Mar. 2020, doi: 10.17993/3ctecno.2020.specialissue4.261-271.
[8] M. V. Manoj Kumar, B. S. Prashanth, S. Sarah, Md. Kashif, and H. R. Sneha, "Expression X: Emotion Based Music Recommendation System," J Comput Theor Nanosci, vol. 17, no. 9, pp. 3958–3963, Dec. 2020, doi: 10.1166/jctn.2020.8996.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)