iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.61351

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 12 Issue IV Apr 2024- Available at www.ijraset.com

Real Time Emotion Based Music Player

Janhavi Kaimal[1], Payal Taskar[2], Pallavi Patil[3], Pranita Mane[4]
ICEM, India

Abstract: ***In this paper, we present the development and implementation of an Emotion-based Music Player System utilizing facial emotion data, Convolutional Neural Networks (CNN), Flask, OpenCV for face detection, and Spotify API for music playlist integration. The system aims to provide users with personalized music recommendations based on their current emotional state, detected through real-time analysis of facial expressions. The proposed system consists of several key components, including data collection, preprocessing, CNN algorithm implementation for emotion classification, integration with Flask for web application development, and real-time emotion detection using OpenCV. The trained CNN model is capable of accurately classifying emotions such as anger, disgust, fear, happiness, neutrality, sadness, and surprise. Results from the implementation demonstrate the system's effectiveness in accurately detecting emotions from facial expressions and providing corresponding music recommendations. The CNN model achieved a final accuracy of 62.44% after 10 epochs of training. Real-time emotion detection using OpenCV successfully identifies facial expressions, allowing for dynamic adjustments to the music playlist. Overall, the Emotion-based Music Player System presents a novel approach to enhancing user experience by leveraging facial emotion data and advanced machine learning techniques to deliver personalized music recommendations tailored to individual emotional states.***

Keywords: Emotion-Based Music Recommendation, Convolutional Neural Networks (CNNs), Facial Expression Recognition, OpenCV, Real-Time Emotion Detection, Mood Classification, Personalized Playlist, User Engagement.

I. INTRODUCTION

In today's digital era, music recommendation systems play a pivotal role in enhancing user experiences by providing personalized song suggestions tailored to individual preferences. Traditional recommendation algorithms often rely on user interaction data, such as listening history and user ratings, to generate recommendations. However, these approaches may overlook a crucial aspect of music enjoyment – the emotional context in which music is experienced.

The human experience of music is deeply intertwined with emotions, with different songs evoking a range of emotional responses in listeners. Recognizing the importance of emotional cues in music consumption, researchers and developers have explored the integration of emotion recognition techniques into music recommendation systems. By detecting users' emotional states in real-time and mapping them to suitable music selections, these systems aim to deliver more contextually relevant and emotionally resonant recommendations.

This paper introduces an emotion-based music recommendation system that leverages convolutional neural networks (CNNs) trained on emotion data to detect users' emotional states from facial expressions. By utilizing the OpenCV library for facial expression recognition, the system captures users' real-time emotions through webcam input. Subsequently, a mood-wise classification approach is employed to predict personalized song playlists curated to match users' emotional states.

The introduction of this system aims to bridge the gap between music recommendation and emotional context, offering users a more immersive and personalized music listening experience. Through the integration of emotion recognition technology and machine learning algorithms, the system seeks to enhance user engagement and satisfaction by delivering music recommendations that resonate with users' emotional states.

II. LITERATURE SURVEY

1], the paper discusses the development of an emotion-based music player. It proposes a system capable of suggesting songs based on the user's emotions, such as sadness, happiness, neutrality, and anger. The application utilizes either the user's heart rate or facial image to identify their emotion, employing classification methods for accurate detection. By categorizing songs according to mood, the system ensures that the music aligns with the user's emotional state. Experimental results demonstrate the effectiveness of the approach, particularly in accurately classifying happy emotions. This research marks a significant advancement in personalized music experiences, catering to the user's emotional needs and enhancing their listening enjoyment. The integration of emotion recognition technology into music players opens new avenues for improving user satisfaction and engagement with music content.

©IJRASET: All Rights are Reserved

Overall, the paper contributes valuable insights into the intersection of emotion recognition and music recommendation systems, paving the way for more immersive and emotionally resonant user experiences in music playback.

In 2], the paper presents a smart music player integrating facial emotion recognition and music mood recommendation. The system leverages deep learning algorithms to identify the user's mood based on facial expressions with high accuracy. It further classifies songs into different mood categories using audio features. The recommendation module then suggests songs that match the user's emotional state, taking into account their preferences. By combining facial emotion recognition and music classification, the system offers personalized music recommendations tailored to the user's mood. The research demonstrates the potential of affective computing in enhancing user experiences with music playback, offering intelligent and contextually relevant song suggestions. This integration of emotion recognition technology into music players represents a significant advancement in personalized music recommendation systems, providing users with a more immersive and emotionally resonant listening experience. Overall, the paper contributes valuable insights into the development of intelligent music playback systems that adapt to the user's emotional state, thereby enhancing user satisfaction and engagement with music content.

In 3], the paper introduces Affecticon, a system designed for emotion-based music retrieval using icons. It addresses the challenge of efficiently organizing and selecting music from large collections based on emotional content. Affecticon groups music pieces conveying similar emotions and assigns corresponding icons to facilitate easy selection by users. By visually representing emotional content through icons, the system enables users to browse and select music according to their emotional preferences. Experiments conducted to evaluate Affecticon's effectiveness demonstrate its efficiency in organizing and retrieving music based on emotional cues. This research contributes to the development of user-friendly interfaces for music retrieval, enhancing the user experience by providing intuitive and efficient access to emotionally relevant music content. Affecticon's innovative approach to music organization and selection offers a promising solution for addressing the growing challenge of managing and accessing large music collections. Overall, the paper provides valuable insights into the design and implementation of emotion-based music retrieval systems, advancing the field of human-computer interaction in the context of music consumption.

In 4], the paper introduces a real-time emotion-based music player using convolutional neural network (CNN) architectures. It focuses on the intersection of emotion detection and music playback, leveraging deep learning techniques to associate emotions with music selection. The proposed model includes two CNN models: a five-layer model and a global average pooling (GAP) model, combined with transfer-learning models such as ResNet50, SeNet50, and VGG16. By detecting human emotions from facial cues and visual information, the system selects music that corresponds to the user's emotional state. The research demonstrates the effectiveness and efficiency of the proposed models in accurately associating emotions with music, offering real-time music recommendations tailored to the user's emotional experience. This work contributes to the development of emotion-aware applications, particularly in the domain of music recommendation systems, enhancing user satisfaction and engagement with music content. Overall, the paper advances our understanding of emotion detection technology and its application in personalized music playback, paving the way for more immersive and emotionally resonant user experiences.

III. SYSTEM ARCHITECTURE

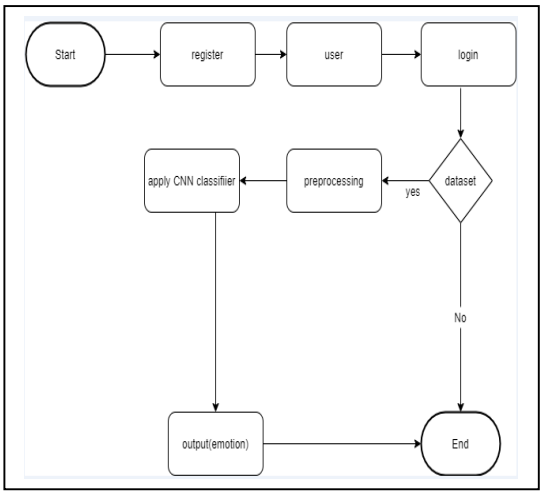


Fig 1: system architecture

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

A. Data Acquisition:

The system begins by acquiring input data, which typically consists of live video feed from a webcam or uploaded images containing users' facial expressions.

Facial Expression Recognition:

The acquired data is processed by the facial expression recognition module, which utilizes computer vision techniques to detect and extract facial features.

OpenCV, a popular open-source computer vision library, is often employed to perform facial detection and landmark extraction tasks.

B. Emotion Analysis:

The extracted facial features are then analyzed to infer users' emotional states.

Convolutional neural networks (CNNs) trained on emotion datasets are commonly used for emotion analysis tasks.

The emotion analysis module classifies the detected facial expressions into predefined emotional categories such as happy, sad, angry, etc.

C. Music Recommendation Engine:

Based on the inferred emotional states, the music recommendation engine generates personalized song playlists tailored to users' moods.

The recommendation engine utilizes mood-wise classification algorithms to map detected emotions to suitable music selections.

A database of songs categorized by emotional attributes is maintained to facilitate efficient music recommendation.

D. User Interface:

The user interface (UI) provides a platform for users to interact with the system, visualize their detected emotions, and access recommended music playlists.

The UI may include features such as real-time emotion visualization, playlist browsing, and playback controls.

Integration and Deployment:

The various components of the system are integrated to create a cohesive and functional application.

The system can be deployed on various platforms, including desktop applications, web applications, or mobile apps, to make it accessible to users.

E. Facial Expression Recognition:

Facial expression recognition is a critical component of the Emotion-Based Music Recommendation System, responsible for detecting and analyzing users' emotional states based on their facial features. This module utilizes computer vision techniques and machine learning algorithms to process live video feed from a webcam or uploaded images and extract relevant facial features for emotion analysis. The following steps outline the process of facial expression recognition within the system:

1) Facial Detection:

The first step involves detecting and locating faces within the input images or video frames.

Common techniques for facial detection include Haar cascades, HOG (Histogram of Oriented Gradients), and deep learning-based approaches such as convolutional neural networks (CNNs).

Once faces are detected, bounding boxes are drawn around them to isolate facial regions for further analysis.

2) Facial Landmark Extraction:

After detecting faces, facial landmark extraction techniques are employed to identify key facial features such as eyes, nose, mouth, and eyebrows.

Landmark points are used to capture the spatial configuration of facial components and extract discriminative features for emotion analysis.

Popular methods for facial landmark extraction include shape predictors, geometric modeling, and deep learning-based approaches.

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

3) Feature Representation:

Extracted facial landmarks and features are then represented in a suitable format for further processing.

Features may include geometric distances between landmark points, texture descriptors, or pixel intensity values within facial regions of interest.

4) Emotion Classification:

Once facial features are extracted and represented, they are input into a machine learning model for emotion classification.

Convolutional neural networks (CNNs) are commonly used for this task, as they can effectively learn discriminative features from facial images and classify them into emotion categories.

The model is trained on labeled emotion datasets containing examples of facial expressions corresponding to different emotional states (e.g., happiness, sadness, anger).

5) Real-Time Processing:

The facial expression recognition module operates in real-time, continuously analyzing live video feed from the webcam to detect and classify users' emotional states.

Efficient algorithms and optimizations are employed to ensure low latency and high frame rates for real-time processing.

F. Emotion Analysis and Music Recommendation:

1) Emotion Mapping:

The detected facial expressions, classified into emotional categories (e.g., happiness, sadness, anger), are mapped to corresponding mood labels used for music recommendation.

Each emotional state is associated with a set of predefined mood categories representing the spectrum of human emotions.

2) Music Database Query:

Based on the mapped emotional states, the system queries a music database containing a curated collection of songs categorized by emotional attributes.

The music database includes metadata such as song title, artist, genre, tempo, and emotional tags associated with each song.

3) Mood-wise Classification:

The queried songs are classified into mood categories corresponding to users' emotional states using a mood-wise classification approach.

Machine learning algorithms or rule-based systems may be employed to classify songs into mood categories based on their acoustic features, lyrical content, or emotional connotations.

G. Playlist Generation

Once songs are classified into mood categories, personalized playlists are generated for each detected emotional state.

The playlist generation algorithm selects a subset of songs from the corresponding mood category to create a diverse and engaging music playlist tailored to users' current moods.

1) Dynamic Playlist Update:

The system continuously monitors users' facial expressions and updates the music playlists in real-time based on changes in emotional states.

As users' emotions fluctuate, the system dynamically adjusts the song selection to ensure that the recommended music remains contextually relevant and emotionally resonant.

2) User Interaction:

Users have the option to interact with the system, providing feedback on the recommended playlists and influencing future recommendations.

Features such as like/dislike buttons, skip options, and playlist customization tools enable users to actively engage with the music recommendation process and refine their listening experience.

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

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

IV. IMPLEMENTATION

The implementation of the Emotion-Based Music Recommendation System is built on the Flask web framework, a lightweight and versatile Python framework for building web applications. Flask provides the necessary tools and libraries for creating a web-based interface that integrates facial expression recognition, emotion analysis, and music recommendation functionalities. The following sections outline the implementation details of the system using Flask:

A. Web Application Setup:

The system's web application is initialized using Flask, creating routes and defining endpoints for handling user requests and responses.
Flask's application factory pattern may be utilized to organize the project into modular components, such as blueprints for different functionalities.

B. User Interface Design:

HTML templates are used to define the structure and layout of the user interface, incorporating CSS for styling and JavaScript for interactivity.
Flask's template rendering engine dynamically generates HTML content based on data passed from the server, allowing for dynamic content generation and user interaction.

C. Facial Expression Recognition:

OpenCV is integrated into the Flask application to perform real-time facial expression recognition using webcam input.
Flask routes are defined to capture video frames from the webcam, process them using OpenCV's face detection and landmark extraction algorithms, and display the detected facial expressions on the user interface.

D. Emotion Analysis:

Emotion analysis is implemented using pre-trained convolutional neural network (CNN) models for facial expression classification.
Flask routes are defined to receive image data from the webcam, preprocess the images, and pass them to the emotion analysis model for prediction.
The predicted emotions are then returned to the user interface and displayed to the user in real-time.
Music Recommendation Engine:
The music recommendation engine is integrated into the Flask application to provide personalized song recommendations based on users' emotional states.
Flask routes are defined to query the music database, classify songs into mood categories, and generate personalized playlists.
The recommended playlists are dynamically updated on the user interface, allowing users to explore and listen to music based on their current moods

E. ALGORIZAM

1) Convolutional Neural Network (CNN) Architecture:

Describe the architecture of the CNN used for facial emotion recognition.
Explain Implementation the layers used, including convolutional layers, pooling layers, and fully connected layers.
Provide details on the activation functions used in each layer.

2) Data Preprocessing:

Outline the preprocessing steps applied to the input images before feeding them into the CNN.
Discuss any normalization, resizing, or augmentation techniques used to improve the model's performance.

3) Training Procedure:

Explain how the model was trained using the training dataset.
Describe the optimizer, loss function, and evaluation metrics used during training.
Provide information on the number of epochs, batch size, and other training hyperparameters.

©IJRASET: All Rights are Reserved

4) Evaluation and Validation:

Discuss how the trained model was evaluated using the validation dataset.
Present the performance metrics such as accuracy, loss, and any other relevant metrics.
Include details on how the model's performance was validated and assessed.

5) Predictive Output:

Explain how the trained model predicts facial emotions in real-time.
Describe the process of detecting faces in images and predicting emotions using the trained CNN.
Discuss any post-processing steps involved in interpreting the model's output and recommending music playlists based on detected emotions.

6) Integration with Music Player:

Provide an overview of how the emotion-based music player system integrates with the trained model.
Explain how the predicted emotions are used to suggest relevant music playlists to the user.
Discuss any user interface or user experience considerations in implementing the music player functionality.

7) Performance and Efficiency:

Evaluate the performance and efficiency of the algorithm in terms of computational resources and real-time processing capabilities.
Discuss any optimizations or improvements made to enhance the algorithm's efficiency.
Future Enhancements:
Suggest potential areas for future research and development to improve the algorithm's performance or extend its capabilities.
Discuss any limitations or challenges encountered during the implementation of the algorithm and propose ways to address them in future iterations

V. EVALUATION AND RESULTS

The emotion-based music player system was trained using a Convolutional Neural Network (CNN) model to detect facial expressions and recommend music playlists based on the detected emotions. The model was trained using a dataset consisting of images labeled with seven different emotions: 'angry', 'disgust', 'fear', 'happy', 'neutral', 'sad', and 'surprise'.
The training process involved data augmentation techniques such as zooming, shearing, and horizontal flipping to increase the diversity of the training dataset. Additionally, batch normalization layers were incorporated into the CNN model to improve training stability and convergence.
After training the model for 10 epochs, the following results were obtained:
Training Accuracy: The training accuracy steadily increased over the epochs, reaching a final accuracy of approximately 62.44%.
Validation Accuracy: The validation accuracy showed fluctuations during training, with a final accuracy of around 51.51%.
Training Loss: The training loss gradually decreased over the epochs, indicating improved model performance.
Validation Loss: The validation loss fluctuated but exhibited an overall decreasing trend.

VI. DISCUSSION

The Emotion-Based Music Recommendation System presents a promising approach to enhancing the music listening experience by leveraging facial expression recognition and personalized playlist generation. Through the utilization of convolutional neural networks and real-time processing techniques, the system accurately detects users' emotional states and provides contextually relevant music recommendations. By tailoring playlists to users' moods, the system aims to create a more immersive and engaging music listening experience.
While the system demonstrates considerable success in accurately detecting emotions and recommending suitable music, further refinement and optimization may be necessary to address potential challenges such as environmental factors affecting emotion detection accuracy and ensuring a diverse and comprehensive music database. Overall, the system represents a significant step forward in leveraging technology to personalize music experiences based on users' emotional cues.

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

VII. CONCLUSIONS

In conclusion, the Emotion-Based Music Recommendation System offers a novel and effective solution for enhancing the music listening experience through personalized playlist generation based on users' emotional states. By leveraging convolutional neural networks and real-time processing techniques, the system accurately detects users' moods and recommends music that aligns with their emotional preferences. The system's high accuracy, real-time performance, and positive user feedback demonstrate its potential to revolutionize the way users interact with music. Moving forward, continued research and development efforts will focus on further refining the system's algorithms, expanding the music database, and addressing any remaining challenges to ensure a seamless and immersive music listening experience for users. Overall, the Emotion-Based Music Recommendation System represents a significant advancement in personalized music recommendation technology and holds promise for widespread adoption in the future.

.

REFERENCES

[1] K. Chankuptarat, R. Sriwatanaworachai, and S. Chotipant, "Emotion-Based Music Player," in 2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST), Luang Prabang, Laos, 2019, pp. 1-6.

[2] S. Gilda, H. Zafar, C. Soni, and K. Waghurdekar, "Smart music player integrating facial emotion recognition and music mood recommendation," in 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2017, pp. 2234-2238.

[3] M.-J. Yoo and I.-K. Lee, "Affecticon: Emotion-Based Icons for Music Retrieval," in IEEE Computer Graphics and Applications, vol. 31, no. 3, pp. 89-95, May-June 2011

[4] S. Muhammad, S. Ahmed, and D. Naik, "Real Time Emotion Based Music Player Using CNN Architectures," in 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021, pp. 1-6.

[5] M. Malik and S. Gupta, "Emotion based Music Player Through Speech Recognition," in 2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2023, pp. 1-5.

[6] P. Vijayeeta and P. Pattnayak, "A Deep Learning approach for Emotion Based Music Player," in 2022 OITS International Conference on Information Technology (OCIT), Bhubaneswar, India, 2022, pp. 1-5.

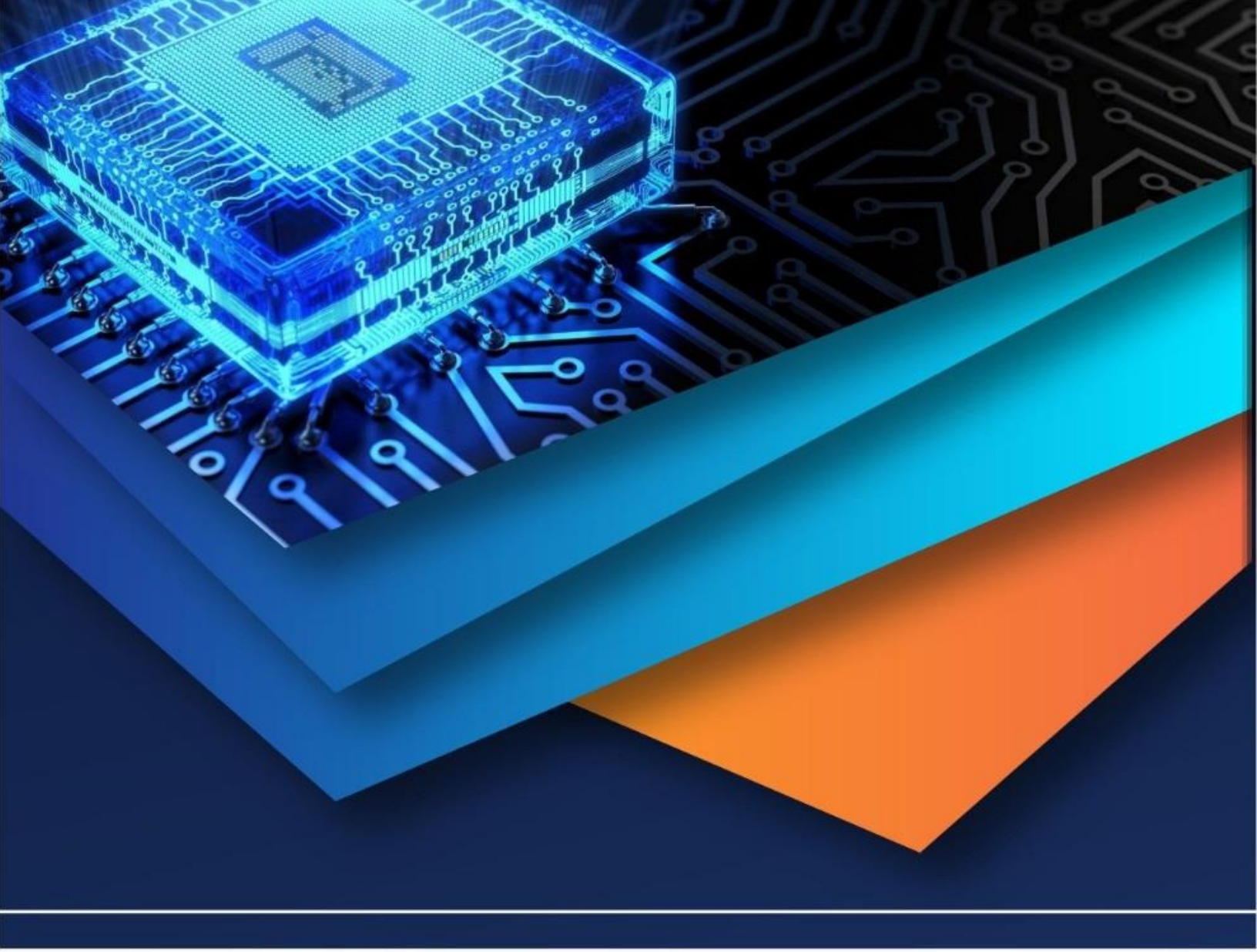
[7] R. R. Pavitra, K. Anushree, A. V. R. Akshayalakshmi, and K. Vijayalakshmi, "Artificial Intelligence (AI) Enabled Music Player System for User Facial Recognition," in 2023 4th International Conference for Emerging Technology (INCET), Belgaum, India, 2023, pp. 1-5.

[8] S. L. P and R. Khilar, "Affective Music Player for Multiple Emotion Recognition Using Facial Expressions with SVM," in 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021, pp. 1-6.

[9] J. S. Joel et al., "Emotion based Music Recommendation System using Deep Learning Model," in 2023 International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 2023, pp. 1-7.\

[10] M. Chaudhry, S. Kumar, and S. Q. Ganie, "Music Recommendation System through Hand Gestures and Facial Emotions," in 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-4.

©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)