iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XII **Month of publication:** December 2022

DOI: https://doi.org/10.22214/ijraset.2022.48387

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 10 Issue XII Dec 2022- Available at www.ijraset.com

A Robust Emotion Detection and Music Recommendation System using Mini Xception CNN

R Venkat Krishna, V. Vijayalakshmi, M Vignesh Balaji
[1, 3]*Department of Networking and Communication, SRM Institute of Science and Technology, Kattankulathur*
[2]*Assistant Professor, Department of Networking and Communication, SRM Institute of Science and Technology, Kattankulathur*

Abstract: ***Due to the emerging developments in artificial intelligence(AI) and machine learning(ML) Technology various systems are developed in recent days that late the human emotions and real time aspects of human psychology detection. Facial recognition based music recommendation system (MRS) is a interesting area of research where he plays an important role in handling the psychology patients. Face recognition system is extensively applied in security systems surveillance systems fault identification etc. Based on Emotion of the human, the music recommendation needs to be provided to analyse the phycology changes with the patients. The proposed approach is focused on considering the constraints available with the facial recognition system in existing frameworks such as deep feature extraction processing delay need to be reduce the hair using deep convolutional neural network(DCNN) architecture based Mini exception algorithm is developed. The system considered FER-2013 image dataset that contains 35000 face images with automated labels that would be helpful for the presented approach to identify the emotion class accurately. The Mini exception algorithm used in CNN layers act as a lightweight system compared with various states of approaches. The proposed system removes the barrier between the existing frameworks and achieved the accuracy of 92%. The recommended music is derived from the music database and further mapped with respect to the algorithm result.***

Keywords: Music, Songs, Facial recognition, Recommendation system, Mini- Xception, Haar cascade classifier.

I. INTRODUCTION

Human emotion is a unique one that cannot be easily predictable. The Real Emotion of the human being is expressed through facial expressions Christopher human feeling can be reflected by many kinds of actions. In order to identify the human emotions and recommending the music related to the emotion is a demanding field of study in artificial intelligence [1].

Music recommendation systems are often helpful to suggest the users recommended music patterns in automated ways and using deep learning algorithms. Music recommendation systems also relieve humans' stress during prolonged work and various environments. Identifying the human emotion is a highly important in a recent days where scientific innovations are created every day to identify the Real Emotion present with the human. Some of the human motions are categorised such as happy neutral angry contempt discussed scared etc. Facial expression is the first sign of human emotion where any changes in the real emotion of the human can be reflected in the facial expressions such as eyes nose mouth eyebrows etc. Because of evaluation of recent Technology such as neural networks mission learning algorithm and deep learning algorithms detection of human emotion systems increasing. Various music recommendation data sets are available where 2013 consists of 35000 images with labelled emotions. The face data set consists of various emotions identified by the changes occurring in the facial skeleton [3]. Deep convolutional neural network algorithms are utilised in facial emotional recognition systems where the classification of facial emotion such as happy sad discussed are identified by the recommendation algorithm accurately. The presented paper considers various constraints present with the facial recognition system such that the feature extraction and the similarity problem are overcome by accurately developing Mini exception model based facial recognition system [4]. Human emotions are unique one. Various facial features are associated with the reflection of emotion shown in the face. The real emotions create various changes in the human body. Physiological changes such as skin temperature get varied using the emotional affect.

The recent evaluations in affective computing enable us to develop prediction models on human emotions [5].

1) The proposed system considers FER 2013 face dataset that provides 35000 images of various emotions which is being labelled by the dataset.
2) Deep convolutional neural network (DCNN) architecture is created and the data set is divided into training and testing sets.
3) Lightweight architecture is implemented in order to extract the emotions of the human security.
4) The advantage of proposed approach using Mini-xception algorithm through CNN identify the emotions accurately by considering various layers of process with numerous iterations.

©IJRASET: All Rights are Reserved

The rest of the paper is formulated as making detailed literature study in Section II. The system tool selection, problem identifications are discussed in Section III. The system architecture, detailed system design steps are discussed in Section IV. The rest of the paper is concluded with future enhancement.

II. BACKGROUND STUDY

H. -G. Kim et al., (2019) presented a system using deep residual bidirectional recurrent neural networks[1]. Facial expressions are dynamically changing in time. In order to improve the performance of the system, ensemble approach is implemented. Dynamic spectrogram is used; long term and short term spectrograms are implemented. Recurrent neural network considers various features and continuously change the performance of the system based on dynamic changes.

D. Wang et al., (2021) describes context aware music recommendation system through low-dimension dense network. Music is considered as part of the life style, where intrinsic features need to be considered while developing automated systems. Using convolution neural network (CNN) algorithm, facial emotion are recognized[2]. Context aware framework also considers the interaction between the systems.

W. Gong et al., (2021) propose a deep music recommendation algorithm[3] based on dance motion analysis and evaluate it through quantitative measures. For quantitative evaluation, this work implements a LSTM-AE based music recommendation method which learns the correspondences between motion and music. In experiments, the two methodologies are compared and the motion analysis based methods outperform their rival by large margins. This work also proposes a quantitative measure of accurately recommended music genre. The proposed motion analysis based method achieves a recommendation accuracy of 91.3% using late fusion of joint and limb features.

I. *Agrawal et al., (2021)* presented a system in which several convolutional neural network topologies are discussed. Emotion identification of humans using m a facial expression database[5] is being discussed here. The proposed approach consider 35000 images of facial images and that is drastically reduces the computation time and availability using convolutional neural network architecture using time efficient hybrid model. Implemented in which the real time constants are considered. The presented emotion recognition system provides the idea of making the data preparation from the Global data set. Cross validation process is highly important for making the validation process accurately.

The author *S. Begaj et al., (2020)* discusses the emotion identification system using multi emotion facial expression data set full stop using convolutional neural network architecture various emotions of humans are identified with respect to the facial features derived with presented approach. The study of facial expression is a challenging task since the Real Emotion of the humans are very critical to predict[6].

Z. Rzayeva et al., (2019) The author considered ohn-Kanade and RAVDESS datasets. Using convolution neural network architecture facial emotions are detected. Many constrains are there when considering facial expression identification.[7] Face features are identified through unique metrics. Convolution neural network architecture consists of sequentially connected layers such as input layer, convolution layers that act as filter, output layer etc.

Considering various existing frameworks, it is clear that convolution neural network architecture is used in many cases. Facial expressions are derived through Haar cascade algorithm and HOG histograms of gradients algorithms frequently. The Haar cascade model act as the standard method for face objects detection.

III. SYSTEM DESIGN

A. *Problem Analysis*

1) The face emotion recognition using images having certain levels of limitations. The face images are captured with the environment of specific elimination that could provide the features uniquely. While detecting the facial features the pixels of the images are considered as an important parameter for pattern recognition.

2) In case of luminous changes with different environment the features is directly affected by the certain environmental changes. Facial expression detection need for the more extra features for making the decision accurately. Inspite of having facial skeleton, changes occurs with eyes, nose, mouth etc.

3) Facial emotions recognitions are mapped. Any changes in human emotion directly impact the facial organs such as eyes, nose, mouth, eyebrows etc.

4) These changes are accurately predicted by the systematic approach automatically for the evaluated with various levels of facial emotion data bases to identify the accurate emotion.

5) Image based the facial expressions need extra attention in terms of adding up the physiological parameters that could be extended as a future research.

B. FER2013 Dataset

The FER (facial emotion recognition) dataset is collected from various volunteers with approximately 30000 facial changes are recorded. The images are captured as RGB images, with random emotions are recorded and properly labelled. This dataset is publicly available for analysis purposes. The different expressions of size 48x48 is mainly labelled. Emotions such as 0 as Angry, 1 as Disgust, 2 as Fear, 3 as Happy, 4 as Sad, 5 as Surprise, 6 as Neutral are recorded. Keeping the standard dataset as a base, the analysis is performed.

IV. METHODOLOGY

A. System Architecture

The system architecture of proposed emotion recognition system using Mini-Xception is depicted here. The system process contains reading the input image, resizing the image, converting the image into grayscale are initially done. The normalized image is further processed to detect the facial objects.

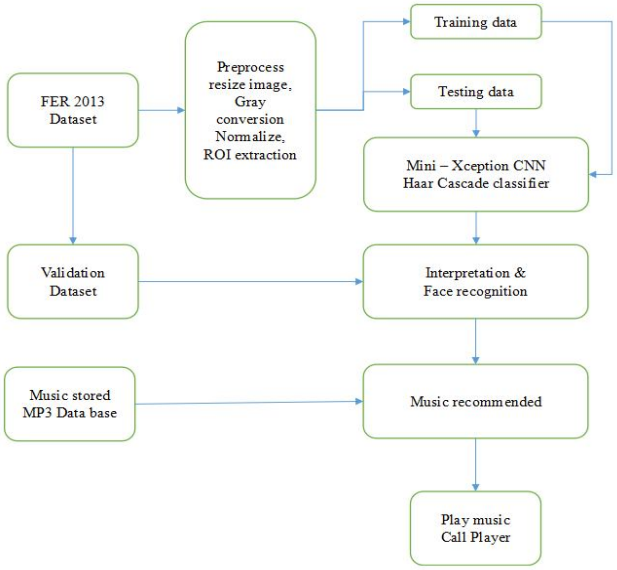


Fig. 1. System architecture of Proposed MRS through Mini-Xception model

Fig. 1. Shows the system architecture of proposed MRS using Mini-xception model.

B. Haar Cascade model

The Haar Cascade model detects the face objects such as eyes, nose, mouth etc. Irrespective of scale, image location, the algorithm considers the real time images and segment the facial parts accurately. Various traces of multiple image samples are utilized in the Haar cascade classifier to identify the similar pattern present in it. These features are saved as .XML file. The pixel regions that matches the pattern of facial objects are highlighted by the boundary. The line of boundary is modelled using voila Jones algorithm. The high speed computing system adopted by the Voila jones algorithm accurately segment the facial parts.

Once the facial parts are segmented, the dataset is divided into training set and testing set. A part of the training images (say 20%) is utilized as testing data.

C. Deep CNN

Deep convolutional brain network (ConvNet or DCNN) is the powerful technique in deep learning frameworks regularly utilized in PC vision and video handling applications. The DCNN comprises of info layer, yield layer and secret layer. Numerous pretrained networks are used in deep learning applications. (Visual math bunch network) VGG Net in CNN model contained pre-prepared layers used for preparing the picture datasets. The CNN with VGG Net go about as the vigorous design for achieving better exactness in picture arrangement process.

The design comprises of different degrees of channels used to separate the important elements and relative spatial highlights. The information picture is passed to flowed convolution channels to extricate the remarkable property from the pictures.

VGG Net is the powerful deep learning convolutional brain network engineering that upholds subsequent 16 layers. The sequenced include extraction process is used for characterization. The info pictures hold the pixel values in the scope of 0-255. In the pre-handling stage, the mean worth of the information picture is killed concerning the whole mean worth assessed from the Picture Net preparation dataset.

The CNN architecture consists of input layer, convolution layer and fully connected layer. In between the fully connected layer and convolution layer Soft-max layers are used to reduce the fitting problem in CNN [8]-[13].

D. Mini-Xception model

As the convolution neural network architecture is used for deep analysis of inputs to identify the unique patterns as fast as possible. The robust structure is still used in many recent evaluations with improved layer formation. The xception model considers improved Convolution 2D layer with Batch normalization associated with each layers. The max-pooling layer is provided at the end of the mini batch normalization process. The depth wise separable convolution provides better accuracy.

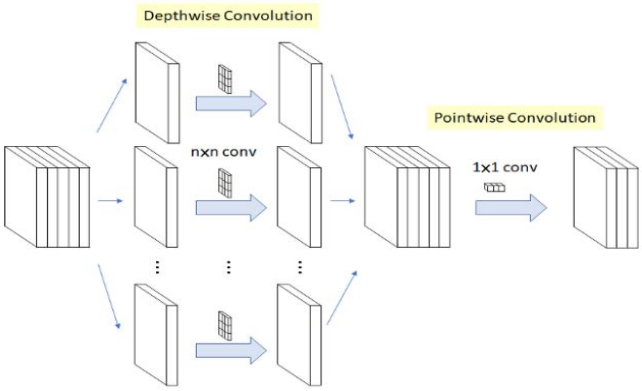


Fig 2. Mini-Xception architecture

Fig.2. shows the Mini-Xception architecture.

1) The Depth wise seperable convolution provides spatial changes that proides deeper comparison of values with the training images.

2) Comparing with conventional neural network architecture the number of connections are improved and lighter than that of other network models. The presence of nonlinearity in the input images are improved.

3) The inception module is a unique small tunable architecture placed wherever required in the convolution neural network. The number of epochs depends upon the complexity of the input data.

V. RESULTS AND DISCUSSIONS

A. Emotion Detection

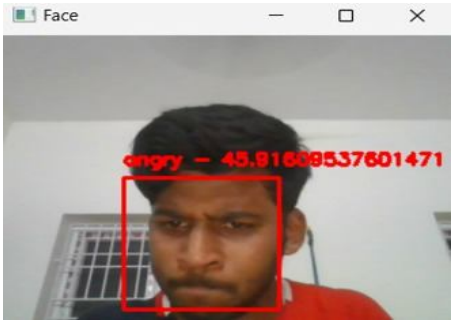


Fig 3. Emotion detection – Angry

Fig. 3 shows the emotion detected result Angry with 45.916% detection rate is identified.

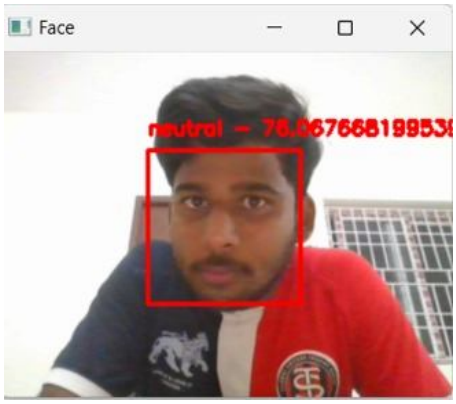


Fig 4. Emotion detection Neutral

Fig 4. Shows the emotion detected as neutral. The face objects are initially detected through haar cascade algorithm. The bounding box is created with the Haar cascade method further the Mini-xception model perform deeper analysis. Hence the face expression is identified accurately.

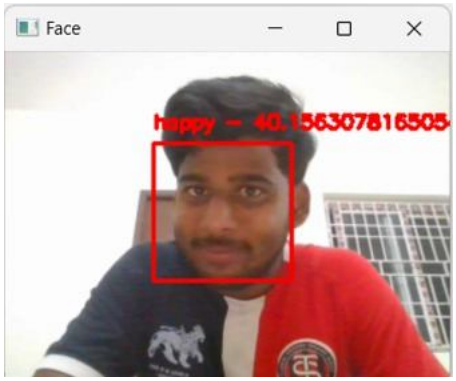


Fig. 5. Emotion detection – Happy

Fig.5. Shows the Emotion detection as happy. The Detection rate is identified within 405 itself.

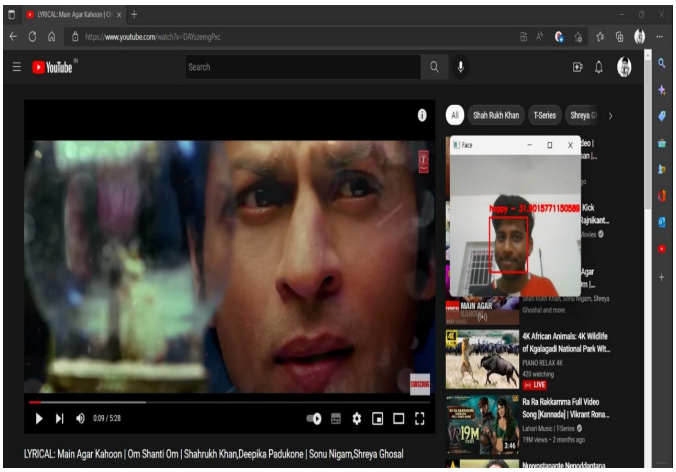


Fig 6. Music recommendation result

Fig 6. Shows the result of music recommendation system using proposed Mini-xception CNN. Based on te detected emotion the music meta data is compared and recommended through smart recommendation system.

VI. CHALLENGES

The major challenge persist with the proposed model is that the usage of large dataset. More image dataset and related processing consumes more graphical processing unit (GPU) space. Hence the system performance is being degraded. The sampling and scalability f input images before processing need to be done. Feature extraction using multiple class is recommended for deep extraction and improved accuracy

VII. CONCLUSION

Human emotions are unique. Each person expresses the emotions through various forms of external hints. Face expressions, Skin temperature changes, body language changes are frequently occurring things in emotional affect. The recent evaluations of affective computing and artificial intelligence technology improved the conventional method of emotion detection and Phycology analysis. The presented study is focused on developing such scenario. Music recommendation system with Mini-Xception CNN is implemented here. The proposed approach achieved accuracy of 92% with reduced error rate upto 0.00125. Further through presented system is improved by evaluating deep focus on feature extraction process, multiple feature extraction and deep cascaded analysis using improved Xception models, etc.

REFERENCES

[1] H. -G. Kim, G. Y. Kim, and J. Y. Kim, "Music Recommendation System Using Human Activity Recognition From Accelerometer Data," in IEEE Transactions on Consumer Electronics, vol. 65, no. 3, pp. 349-358, Aug. 2019, DOI: 10.1109/TCE.2019.2924177.

[2] D. Wang, X. Zhang, D. Yu, G. Xu, and S. Deng, "CAME: Content- and Context-Aware Music Embedding for Recommendation," in IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 3, pp. 1375-1388, March 2021, DOI: 10.1109/TNNLS.2020.2984665.

[3] W. Gong and Q. Yu, "A Deep Music Recommendation Method Based on Human Motion Analysis," in IEEE Access, vol. 9, pp. 26290-26300, 2021, DOI: 10.1109/ACCESS.2021.3057486.

[4] K. Chen, B. Liang, X. Ma, and M. Gu, "Learning Audio Embeddings with User Listening Data for Content-Based Music Recommendation," ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2021, pp. 3015-3019, DOI: 10.1109/ICASSP39728.2021.9414458.

[5] I. Agrawal, A. Kumar, D. Swathi, V. Yashwanthi and R. Hegde, "Emotion Recognition from Facial Expression using CNN," 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC), 2021, pp. 01-06, DOI: 10.1109/R10-HTC53172.2021.9641578.

[6] S. Begaj, A. O. Topal and M. Ali, "Emotion Recognition Based on Facial Expressions Using Convolutional Neural Network (CNN)," 2020 International Conference on Computing, Networking, Telecommunications & Engineering Sciences Applications (CoNTESA), 2020, pp. 58-63, DOI: 10.1109/CoNTESA50436.2020.9302866.

[7] Z. Rzayeva and E. Alasgarov, "Facial Emotion Recognition using Convolutional Neural Networks," 2019 IEEE 13th International Conference on Application of Information and Communication Technologies (AICT), 2019, pp. 1-5, DOI: 10.1109/AICT47866.2019.8981757.

[8] Y. -H. Cheng, P. -C. Chang and C. -N. Kuo, "Convolutional Neural Networks Approach for Music Genre Classification," 2020 International Symposium on Computer, Consumer and Control (IS3C), 2020, pp. 399-403, DOI: 10.1109/IS3C50286.2020.00109.

[9] A. Arora, A. Kaul, and V. Mittal, "Mood Based Music Player," 2019 International Conference on Signal Processing and Communication (ICSC), 2019, pp. 333-337, DOI: 10.1109/ICSC45622.2019.8938384.

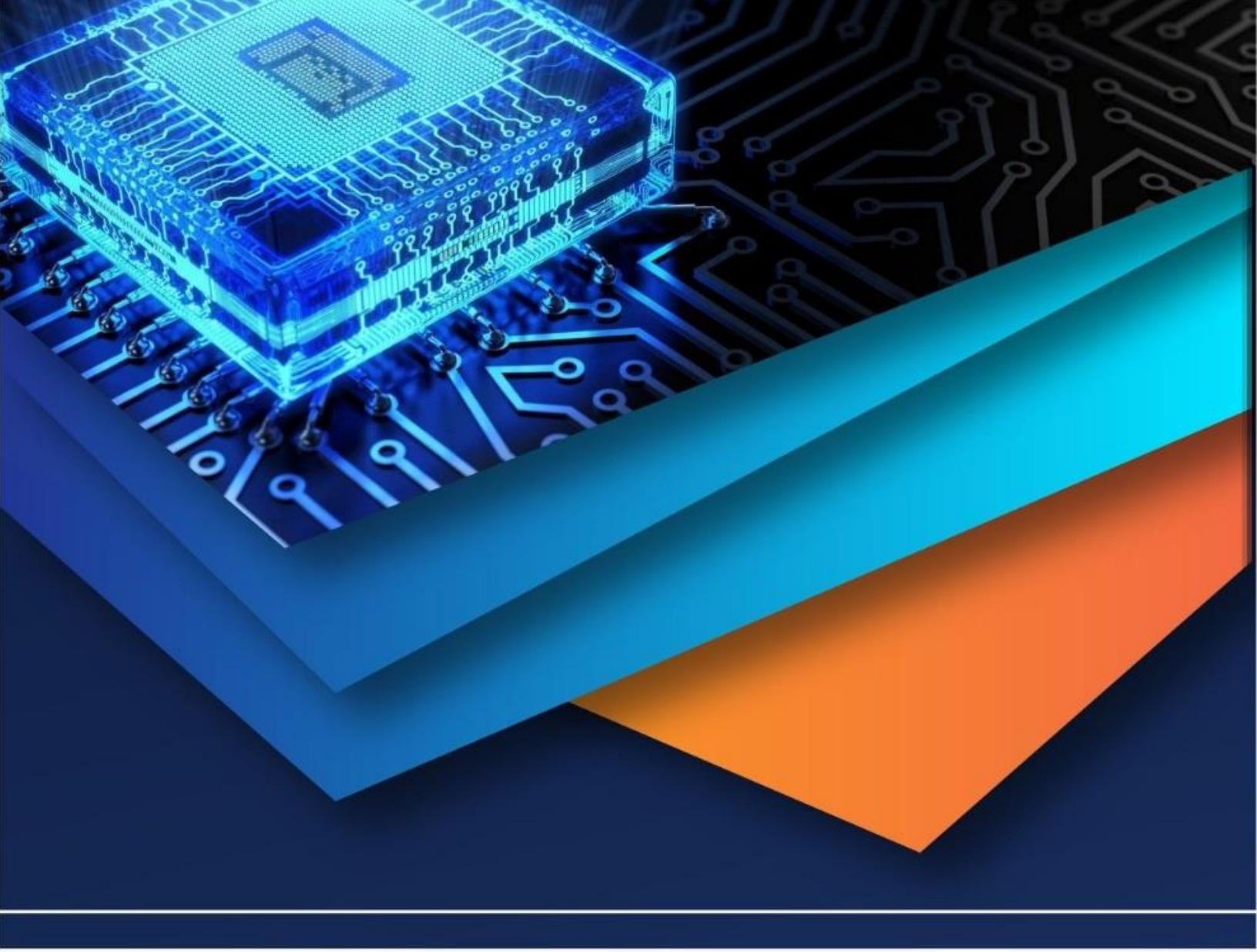
[10] S. L. P and R. Khilar, "Affective Music Player for Multiple Emotion Recognition Using Facial Expressions with SVM," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC), 2021, pp. 622-626, DOI: 10.1109/I-SMAC52330.2021.9640706.

[11] S. Muhammad, S. Ahmed, and D. Naik, "Real Time Emotion Based Music Player Using CNN Architectures," 2021 6th International Conference for Convergence in Technology (I2CT), 2021, pp. 1-5, DOI: 10.1109/I2CT51068.2021.9417949.

[12] A. Patel and R. Wadhvani, "A Comparative Study of Music Recommendation Systems," 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), 2018, pp. 1-4, DOI: 10.1109/SCEECS.2018.8546852.

[13] G. Yamaguchi and M. Fukumoto, "A Music Recommendation System based on Melody Creation by Interactive GA," 2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2019, pp. 286-290, DOI: 10.1109/SNPD.2019.8935654.

©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)