iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62014

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 V May 2024- Available at www.ijraset.com

Music Beats - A Music Recommendation System using Facial Expressions

Sakshi Shukla[1], Prince Prasad[2], Amit, Nitish Kumar[3], Nakul Panchal[4]
Computer Science Department, IMS Engineering College, Ghaziabad

Abstract: ***Interest in the nexus between music recommendation systems and affective computing has grown recently. This article examines the state of the art in music recommendation systems that use facial expression analysis to improve customer satisfaction. Facial expression detection has been incorporated into computer vision and machine learning systems, opening up new possibilities for personalized music curating.***

The basic theories and techniques supporting facial expression analysis in the context of emotional computing are first examined in this study. It explores the developments in emotion identification technology, emphasizing important datasets, algorithms, and their advantages and disadvantages.

Additionally, this research assesses how well these systems capture user preferences while resolving issues with scalability, accuracy, and privacy. It talks about the effects on user satisfaction and engagement of using facial expressions as a channel for feedback in recommendation systems.

In order to improve and develop facial expression-based music recommendation systems, this paper concludes by outlining future directions and possible research routes and highlighting the necessity of multidisciplinary cooperation between computer science, psychology, and musicology.

To sum up, this paper offers a thorough analysis of the current state-of-the-art in facial expression-driven music recommendation systems, outlining their innovations, difficulties, and potential directions for growth of multidisciplinary cooperation between computer science, psychology, and musicology.

Keywords: facial expression analysis, music recommendation, machine learning, deep learning, personalized recommendation s.

I. INTRODUCTION

What smartphone should I get? What movie should I see this weekend? Where should my family and I spend the upcoming holidays? What books should I bring with me for my extended vacation? – These are just a few instances of the frequent indecision for which we ask friends and acquaintances for advice. Sadly, almost all of us have found that, despite their best efforts, those well-meaning recommendations are frequently ineffective because other people's tastes do not always coincide with our own. These recommendations frequently have biases as well. The other unusual options we have include becoming decision science experts and experimenting with the intricate theories, surfing the web, and waste hours in confusing reviews and suggestions, hire a specialist, follow the crowd, or just listen to our hearts instead of wasting hours reading through the bewildering reviews and recommendations. The main point is that it can be quite difficult to pinpoint a specific recommendation for the things that we might find interesting. Having a personal advisor to guide us and recommend the best course of action whenever we need to make a decision would be very beneficial. Fortunately, we have one of these in the form of the recommender system (RS).

People frequently use their facial expressions to convey their feelings. People's moods have long been recognized to change in response to music. The ability to identify the emotion that a person is expressing, as well as play music that fits the user's mood, can gradually ease the user's anxiety and have a pleasant overall effect. The project's goal is to use a person's facial expressions to convey the emotions. The webcam interface for computer systems is used by a music player to record human emotion. The program takes a picture of the user, then uses image processing and segmentation methods to extract information from the target person's face and someone is attempting to convey a feeling. By playing music that fits the user's needs and projecting the user's image, the initiative seeks to lighten the user's mood.

Facial expression is the finest means by which the people can deduce or evaluate the emotion, sentiment, or thoughts that someone else is attempting to convey. Occasionally, changing one's mood can also aid in overcoming difficult circumstances like despair and depression numerous health related issues can be avoided with the use of expression analysis.

Risks can be minimized, and actions can be performed to assist in improving a user's mood.

©IJRASET: All Rights are Reserved

As digital signal processing and other efficient feature extraction methods become more advanced technologically, automated the field of emotion recognition in multimedia properties, such as music or movies, is expanding quickly, and this system has a lot of potential uses, including music entertainment and human-computer interaction systems. We employ expressions on our faces. To develop an emotion detection recommender system that can identify user emotions and provide a list of suitable songs. The suggested system recognizes a person's feelings; if the person is experiencing a negative emotion, then a specific playlist with the most appropriate music genres to lift his spirits will be displayed. And in the event that the feeling is pleasant, a particular playlist featuring various kinds of music that will inflate the positive emotions of the user.

II. LITERATURE REVIEW

In recent years, a lot of research has been done on music recommendation systems with the goal of offering users personalized music recommendations. Collaborative filtering, which suggests music based on a user's past listening history as well as the listening histories of users who are similar to them, was one of the first methods used in music recommendation systems. Collaborative filtering can't reliably gauge a user's mood or offer tailored suggestions depending on how they're feeling at the moment, though. In a number of disciplines, including computer vision, psychology, and neuroscience, facial expression analysis has been used to infer an individual's emotional state from their facial expressions. Numerous studies have looked into the application of facial expression analysis to music recommendation systems in recent years.

A facial expression analysis-based music recommendation system was created in a study by Yang et al. (2014) to identify the user's emotional state and offer tailored music recommendations. The system used a collaborative filtering algorithm to provide personalized music recommendations and a deep belief network to learn the relationship between facial expressions and emotional states.

A multidisciplinary music recommendation engine that combined physiological signals and analysis of facial expressions was created in a different study by Yang et al. (2016). The system used a physiological signal-based model to measure the user's physiological responses, such as skin conductance and heart rate, and a model based on deep learning to identify and categorize facial expressions. Then, to provide tailored music recommendations, the system combined content-based and collaborative filtering techniques in a hybrid fashion.

A facial expression-based emotion recognition model was utilized in a more recent study by Li et al. (2020) to create a personalized music recommendation system. Convolutional neural networks (CNNs) were used by the system to identify and categorize facial expressions, and k-nearest neighbor algorithms were used to provide user-specific music recommendations based on their emotional states.

These studies show how algorithms that use machine learning can be used to create personalized music recommendations based on a user's state of mind, and they also highlight the potential of facial expression analysis in music recommendation systems.

With the aid of the AdaBoost algorithm, Xuan Zhu et al. developed an integrated music recommendation system technique that includes automatic music genre classification, music emotion classification, and music similarity query. They also proposed a new tempo feature called (LMFC), which has timbre features that improve the efficiency of music classification. In this process, an interface is implemented where the user needs to press two buttons in order for the music to be played in accordance with his selection of musical genres. Additionally, there is another feature where the user-selected songs are available for the user to play a song of their choice. This approach' primary objective is to segregate the enormous music databases into distinct selections based on the given genre.

Ziyang Yu and colleagues (2010) suggested There are a few fundamental steps to achieving these goals because big data is growing at a rapid pace and deep learning is made possible by these technologies, which combine automatic music recommendation algorithms with convolutional neural networks' micro-expression recognition capabilities to suggest music based on an individual's emotions. First, all expressions in pictures should be gathered. Next, the images should be pre-processed. Lastly, the images should be classified. After this was implemented, the recognition rate was 62.1%. Here, the user's current emotions will be used to identify the expressions rather than their listening history. There are occasionally a few issues to be resolved following this.

However, more research is still needed in this field, especially to create robust and accurate models for facial expression analysis and to assess how well these systems work in practical situations.

In this paper, we propose a novel system for personalized music recommendations that makes use of facial expression analysis. The system uses a collaborative filtering algorithm for songs suggestion and a deep learning-based model for facial expression analysis. We demonstrate the efficacy of the proposed system in accurately predicting the user's emotional state and providing personalized music recommendations by evaluating it against a dataset of facial expression images and music preference ratings.

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

III. METHODOLOGY

1) *Problem Definition and Goal Setting:* Simply state what the system's goals are. Establish the precise objectives, such as raising user satisfaction through emotion-based music recommendations, boosting user engagement, or offering personalized recommendations.
2) *Data Collection and Preprocessing:* Gather a varied collection of facial expressions representing a range of emotional states. We use Kaggle dataset for this process like Spotify dataset, mood dataset after that Annotation and Cleaning of dataset in which we apply matching emotional labels to the data and carry out preprocessing operations like face detection, alignment, and normalization.
3) *Face detection:* One application that falls under the category of computer vision technology is face detection. This is the procedure that creates and trains algorithms to identify faces or objects in photos correctly for object detection or related systems. Real-time detection from an image or video frame is possible. These classifiers are used in face detection, and they are algorithms that identify whether a given image contains a face (1) or not (0). To increase accuracy, classifiers are trained to recognize faces in a large number of images. Two types of classifiers are used by OpenCV: Haar cascades and LBP (Local Binary Pattern). When a Haar classifier is used for face detection, it is trained using pre-defined, varying face data, allowing it to recognize various faces accurately. Reducing background noise and other distractions is the primary goal of face detection in order to identify the face within the frame. The method is based on machine learning, and the cascade function is trained using a collection of input files. The Haar Wavelet technique, which divides image pixels into squares, is supported by operate .This makes use of "training data" and machine learning techniques to extract a high degree of accuracy.



Fig. 1 Example Of Emotion

4) *Feature extraction:* We use the pre-trained network, which is a sequential model, as an arbitrary feature extractor when we extract features. letting the input image continue until it reaches the predetermined layer, at which point we take the layer's outputs and use them as our features. Use only a few filters because the initial layers of a convolutional network extract high-level features from the captured image. We increase the number of filters to twice or three times the dimension of the filter from the previous layer as we create deeper layers. Although they require a lot of computation, filters in the deeper layers gain more features. In order to do this, we made use of the reliable, discriminating features that the convolution neural network had learned. Feature maps, which are an intermediate representation for every layer above the first, will be the model's outputs. To find out which features were used to classify the image, load the input image that you wish to see the feature map for. Applying filters or feature detectors to the input image or the feature map output of the previous layer yields feature maps. Visualization of feature maps will shed light on the internal representations for particular input for every Convolutional layer in the model.
5) *Emotion Detection:* Using the Relu activation function, convolution neural network design applies filters or feature detectors to the input picture to obtain feature maps or activation maps. Bends, vertical and horizontal lines, edges, and other characteristics that are present in the image can all be identified with the use of feature detectors or filters. The feature maps are then subjected to pooling for translation invariance. Pooling is predicated on the idea that the combined outputs remain constant when we make slight changes to the input. Any pooling from min, average, or max can be used. However, max-pooling outperforms min or average pooling in terms of performance. All of the input should be flattened before being fed into a deep neural network, which will then provide outputs for the class of the object.

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

The image's class will either be binary or multi-class, used for things like classifying different identical products or digit identification. The learnt features in a neural network cannot be interpreted, making neural networks akin to a black box. In other words, the CNN model receives an input image and outputs the results. To detect emotions, load the CNN-trained model that has been weight-trained. When we capture a user's real-time image, the pre-trained CNN model receives it, predicts the emotion, and labels the image accordingly.

6) *Result:* The findings show that the suggested system is capable of correctly predicting the user's emotional state and making tailored music recommendations that fit their disposition. Performance indicators including precision, recall, and F1 score are used to assess the system's efficacy in making tailored music recommendations and its accuracy in predicting the user's emotional state.

IV. PROPOSED SYSTEM OVERVIEW

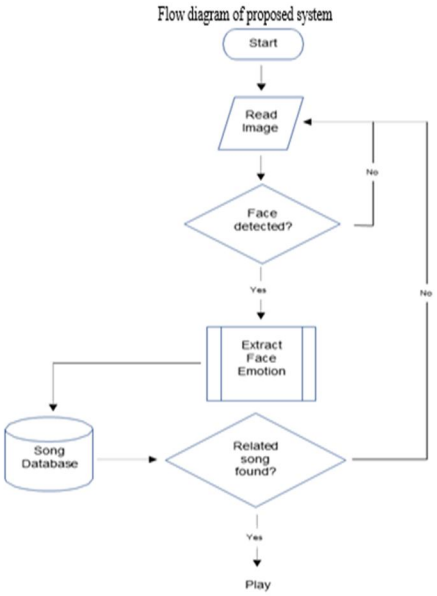


Fig. 2

1) Real-Time Capture: This module's system must accurately capture the user's face.
2) Face Recognition: In this case, the user's face will be entered. The user image's characteristics are assessed by the convolutional neural network according to its programming.
3) Emotion Detection: In this stage, the user image's features are extracted in order to hghjghj Identify the user's emotions. The system then generates captions based on the user's feelings.
4) Music Recommendation: Based on the user's emotions and the mood type of the music, the recommendation module makes a song recommendation.

V. CONCLUSION

A comprehensive analysis of the literature reveals that there are numerous ways to put the music recommender system into practice. Methods put out by earlier researchers and developers were examined. Our system's goals were set in stone after the results. Since AI-powered applications are becoming more and more popular, our project will be a cutting-edge example of a popular use of technology. We give a summary of how music can influence a user's mood in this system, along with advice on how to select the appropriate music tracks to lift users' spirits. Emotions of the user can be detected by the developed system. The emotions that the machine was able to identify were neutral, shocked, joyful, sad and furious.

The suggested algorithm identified the user's sentiment and then gave them a playlist with songs that matched their mood. Processing a large dataset requires a lot of memory and processing power. Development will become more appealing and harder as a result. The goal is to develop this application as cheaply as feasible while using a standardized platform. Our facial emotion recognition-based music recommendation system will lessen users' workloads when it comes to making and maintaining playlists.

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

We hope to create an easily installable Android app in the future that will provide an emotion-based music recommendation system for our phones. Additionally, we intend to create unique features that would offer quotes from notable figures based on the user's emotions. For example, if the user is identified as down, a song recommendation and a statement from the quoted figure will be shown to uplift the user.

The suggested algorithm identified the user's sentiment and then gave them a playlist with songs that matched their mood. Processing a large dataset requires a lot of memory and processing power. Development will become more appealing and harder as a result. The goal is to develop this application as cheaply as feasible while using a standardized platform. Our facial emotion recognition-based music recommendation system will lessen users' workloads when it comes to making and maintaining playlists.

We hope to create an easily installable Android app in the future that will provide an emotion-based music recommendation system for our phones. Additionally, we intend to create unique features that would offer quotes from notable figures based on the user's emotions. For example, if the user is identified as down, a song recommendation and a statement from the quoted figure will be shown to uplift the user.

VI. FUTURESCOPE

Exciting potential for music recommendation systems are ahead thanks to technological developments and a better understanding of consumer tastes. Here are a few possible applications for these systems in the future:

1) Multi-Modal Integration: Facial Expressions with Biometric Data: To gain a more complete knowledge of emotions, facial expressions can be combined with additional biometric data such as skin conductivity, heart rate, or EEG signals.

2) Contextual Awareness: Environmental environment: To improve the relevancy of music recommendations, take into account contextual indications like location, time of day, weather, or social environment.

3) Explainable AI in Recommendations: Increasing user trust and comprehension by creating systems that can explain the reasoning behind a particular music recommendation based on facial expressions or other data.

4) Personalized Emotional Journey: Dynamic Playlists: Putting together dynamic playlists that adjust to users' evolving emotions as well as their current emotional states, providing music that either complements or contrasts with their shifting moods.

5) AI Co-Creation and User Engagement: Collaborative Creation: Encouraging more engaging and customized experiences by enabling AI systems to work in tandem with users to create or curate music based on emotional cues.

6) Continuous Learning and Feedback: Using systems that actively learn from and adjust to continuous user feedback on music recommendations will help recommendations become more accurate over time.

REFERENCES

[1] Li, X., Zhang, W., Zhao, S., & Li, W. (2020). Personalized music recommendation based on facial expression recognition. IEEE Access, 8, 52454- 52462.

[2] Yang, C., Li, Z., Liu, X., Wang, X., & Jia, Y. (2016). Multimodal music recommendation based on facial expression and physiological signals. Multimedia Systems, 22(1), 59-68.

[3] X. Zhu, Y. Y. Shi, H. G. Kim, and K. W. Emm, “An integrated music recommendation system,” IEEE Trans. Consum. Electron., vol. 52, no. 3, pp. 917–925, 2006, doi: 10.1109/TCE.2006.1706489.

[4] Z. Yu, M. Zhao, Y. Wu, P. Liu, and H. Chen, “Research on Automatic Music Recommendation Algorithm Based on Facial Micro-expression Recognition,” Chinese Control Conf. CCC, vol. 2020-July, pp. 7257–7263, 2020, doi: 10.23919/CCC50068.2020.9189600

[5] Lee, J., & Lee, J. H. (2020). Music recommendation system based on deep learning and emotion recognition. In 2020 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 342-346). IEEE.

[6] Wang, D., Hao, Y., Yang, X., Li, B., & Li, Y. (2021). A novel deep learning-based music recommendation system integrating facial expression recognition. Neurocomputing, 453, 329-339.

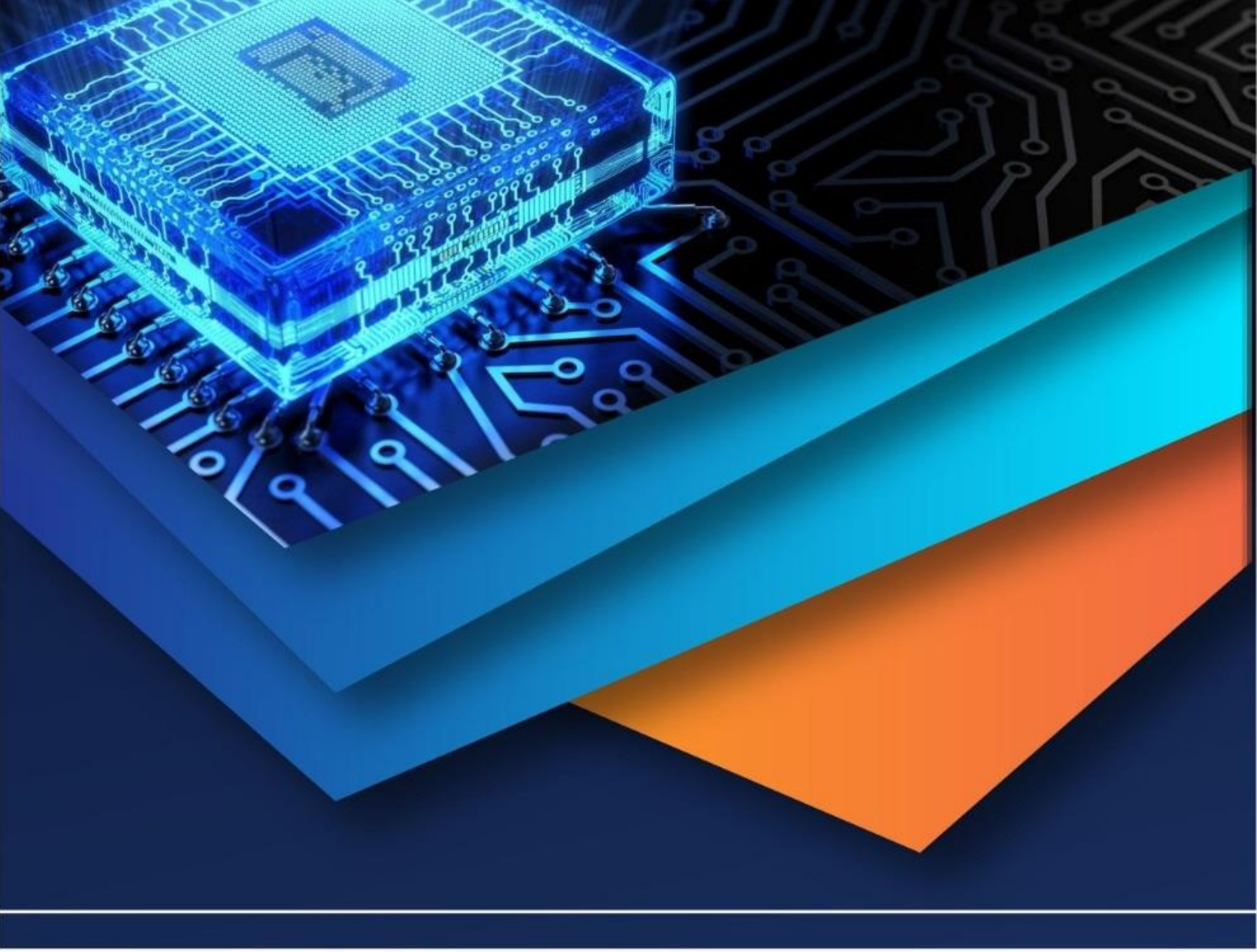
[7] Ramya Ramanathan, Radha Kumaran, Ram Rohan R, Rajat Gupta, and Vishalakshi Prabhu, an intelligent music player based on emotion recognition, 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017. https://doi.org/10.1109/CSITSS.2017.8447743

[8] CH. Sadhvika, Gutta . Abigna , P. Srinivas reddy, Emotion-based music recommendation system, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad; International Journal of Emerging Technologies and Innovative Research (JETIR) Volume 7, Is-sue 4, April 2020.

[9] Chang, S. W., Lee, S. H., & Kim, S. W. (2019). Music recommendation system based on facial expression recognition using convolutional neural network. Applied Sciences, 9(8), 1701.

[10] Lee, J., & Lee, J. H. (2020). Music recommendation system based on deep learning and emotion recognition. In 2020 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 342-346). IEEE.

©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)