iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XI **Month of publication:** November 2022

DOI: https://doi.org/10.22214/ijraset.2022.47572

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 XI Nov 2022- Available at www.ijraset.com

Song Recommendation System Based on Real-Time Facial Expressions

Assist. Prof. Himanshi Agrawal[1], Ramesh Patil[2], Prathmesh Abitkar[3], Nikhil Vartak[4], Shatayush Thakare[5]
[1, 2, 3, 4]*Students,* [5]*Professor, Department of Computer Engineering, SKN Sinhgad Institute of Technology and Science, Kusgaon (BK), Pune*

Abstract: ***Traditional methods of playing music according to a person's mood required human interaction. Migrating to computer vision technology will enable the automation of the such system. This article describes the implementation details of a real-time facial feature extraction and emotion recognition system. One way to do this is to compare selected facial features from an image to a face database. Recognizing emotions from images has become one of the active research topics in image processing and applications based on human-computer interaction. The expression on the face is detected using a convolutional neural network (CNN) for classifying human emotions from dynamic facial expressions in real time. The FER dataset is utilized to train the model which contains approximately 30,000 facial RGB images of different expressions. Expression-based music players aim to scan and interpret data and build playlists accordingly. It will suggest a playlist of songs based on your facial expressions. This is an additional feature of the existing feature of a music player.***

Keywords: Emotion Recognition, Face Detection, Music recommendation, Deep Learning, Supervised Learning, Machine Learning.

I. INTRODUCTION

Music Plays an important role in a person's life it helps one to get relief from his/her lifestyle and is directly connected to his/her emotions or feelings. An emotion can be defined as a physiological and mental state which is subjective and private. It includes many actions, behaviour, thoughts, and feelings. In the field of computer vision, these emotions play an important role in various research purposes. Emotions often mediate and facilitate interactions between people. Therefore, understanding emotions often brings context to seemingly bizarre and complex social communications. Emotion can be recognized through a variety of means such as voice intonation, body language, and more complex methods such as electroencephalography (EEG). But a simpler and more practical way is to study facial expressions. There are seven types of human emotions shown to be universally recognized across different cultures: anger, disgust, fear, happiness, sadness, surprise, and contempt. Interestingly, even for complex expressions where a mixture of emotions could be used as descriptors, a cross-cultural agreement is still observed. emotion detection based on facial expressions is one of the current topics in the various fields which provide a solution to various challenges. Most music lovers users found themselves in a hectic situation when they do not find songs corresponding to their mood the situation. Again, there are increasing advancements in the field of multimedia and technology, with increasing features like fast forward, reverse, variable playback speed (seek & time compression), local playback, streaming playback with multicast streams, volume modulation, and genre classification. But these features satisfy the user's basic requirements, yet the user has to face the task of manually browsing through the playlist of songs and selecting songs based on his current mood and behaviour. So, there is a need for a system that can reduce human efforts of manually playing the song based on human mood

II. LITERATURE REVIEW

Lokesh Singh, the objective is to introduce the needs and applications of facial expression recognition. Between Verbal and Non-Verbal forms of communication facial expression is a form of non-verbal communication but it plays a vital role. It expresses the human perspective or feelings and also his or her mental situation big research is been made to enhance Human-Computer Interaction (HCI) over two decades. This paper includes the introduction of the facial emotion recognition system, its application, a comparative study of popular facial expression recognition techniques, and phases of the automatic facial expression recognition system. The author reviewed the paper related to this topic and gave its pros and cons. The Facial expression recognition system is also described. Zheng proposed two main categories of facial feature extraction. This includes appearance-based feature extraction and shape-based feature extraction, including extraction of some key points of the face such as the mouth, eyes, and eyebrows. Published papers were limited to recognizing facial expressions based on key facial points such as the mouth, eyes, and eyebrows.

©IJRASET: All Rights are Reserved

The techniques used for facial expression recognition were appearance-based feature extraction and shape-based feature extraction.
An approach by S. Matilda used webcams instead of images to capture emotions on a user's face. User emotions were categorized using the fish face algorithm. According to the paper, the only facial expressions captured were her two: happy and sad. The papers published by the author were limited to recognizing only two facial expressions of her, happy or sad.
Vinay paper. The emphasis is on extracting appearance and geometric features from the p input. These features are useful for examining face shape and pixel intensity. Support vector machines (SVMs) were used for this purpose. We achieved 60-70% accuracy on real-time images and 80-90% accuracy on static images. A published paper was used to recognize facial expressions with high accuracy based on still images, while real-time recognition was less accurate.
H Immanuel James, J James, Anto Arnold, J Maria Masilla and Ruban Sara. This post discusses human emotion detection for emotion-based music player development, approaches developed in available music players to detect emotions, methods followed by music players to detect human emotions and Focus on how to use it better. their emotion recognition system. It also briefly discusses playlist generation and sentiment classification. The published paper had music restrictions, limiting the set of recommended music. As an algorithm he used SVM, which also reduced accuracy.
Parul Tambe proposed a system for automating interactions between users and music players. The system runs and learns the user's preferred emotions and various activities, and recommends songs based on the user's current emotions. Published papers had limitations in recognizing facial expressions and songs. The algorithm used was CNN (Convolutional Neural Network).
Yading Song, Simon Dixon, and Marcus Pearce apply an SVM-based approach for classifying the music emotions based on tags of the Last.FM website. There are 4 emotions provided in this research; namely, happy, anger, sad, and relax.
Aayush Bhardwaj detect emotion using Electroencephalography (EEG) signals. EEG signals are that it detects real emotions arising straight from our mind and ignores external features like facial expressions or gesture. Hence EEG can act as a real indicator of the emotion depicted by the subject. Classify EEG signals into seven different emotions using independent component analysis (ICA) and machine learning techniques such as support vector machines (SVM) and linear discriminant analysis (LDA).
Londhe Focused on the study of the changes in curvatures on the face and intensities of corresponding pixels of images. Artificial Neural Networks (ANN) were used in the classification of extracted features into 6 major universal emotions like anger, disgust, fear, happy, sad, and surprise

A. Problems with the Present system

Music players now support a large number of songs in a playlist. Playlist songs play randomly regardless of your mood. Sometimes we want to play a song according to our mood for that, the song was to be searched manually from the playlist, which is time-consuming. Once you choose a song, the next song doesn't necessarily have the same genre or mood. Every time we were required to select songs according to our mood from the playlist. So time-consuming and frustrating. That's why you need an application that plays music on a whim, doesn't take long, and is easy to use. The proposed application is able to extract the user's emotion by capturing the facial feature of the user using a camera and thus will detect the user's emotion. The proposed is implemented on the basis of the Haar cascade and CNN algorithm. The image processing also plays an important role in the given input image various features are calculated. The input image of our model is a grayscale image with 48*48 pixels, which is a pre-processed image from the previous module. The CNN part can be considered as having several stages. We are implementing an efficient method that will be used for song recommendation from the playlist based on facial emotion.

B. Objectives of Proposed System

The key objectives of this project can be split into two parts, the emotion recognition phase of the user and music recommendation based on facial emotion. This project focuses on emotion recognition and music recommendations. The user first captures the image and then by using image processing techniques we get the fine-tuned image for feature extraction.

1) *Emotion Recognition:* In emotion recognition, the system detects the user's image through the camera and then the system classifies the image into one of the seven universal expressions -Happiness, sadness, anger, surprise, disgust, fear, and neutrality labelled in the FER2013 dataset. The training was performed using CNNs, a category of neural networks with proven productivity in image processing. The dataset was first split into training and test datasets, and then it was trained on the training set. The feature extraction process was not done on the data before feeding it into CNN.
2) *Music Recommendations:* After facial emotion detection, each song will be played from the playlist based on the user's mood such as sad, angry, chilled-out party, or happy.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

III. METHODOLOGY

A. System Architecture

It consists of the following steps:

1) Data Preparation
2) Feature Extraction using Deep CNN
3) Visualization
4) Training and Testing of Data
5) Identifying the emotions based on the facial expression
6) Based on the emotion opens the specific music file and play music that describes the emotion of the person

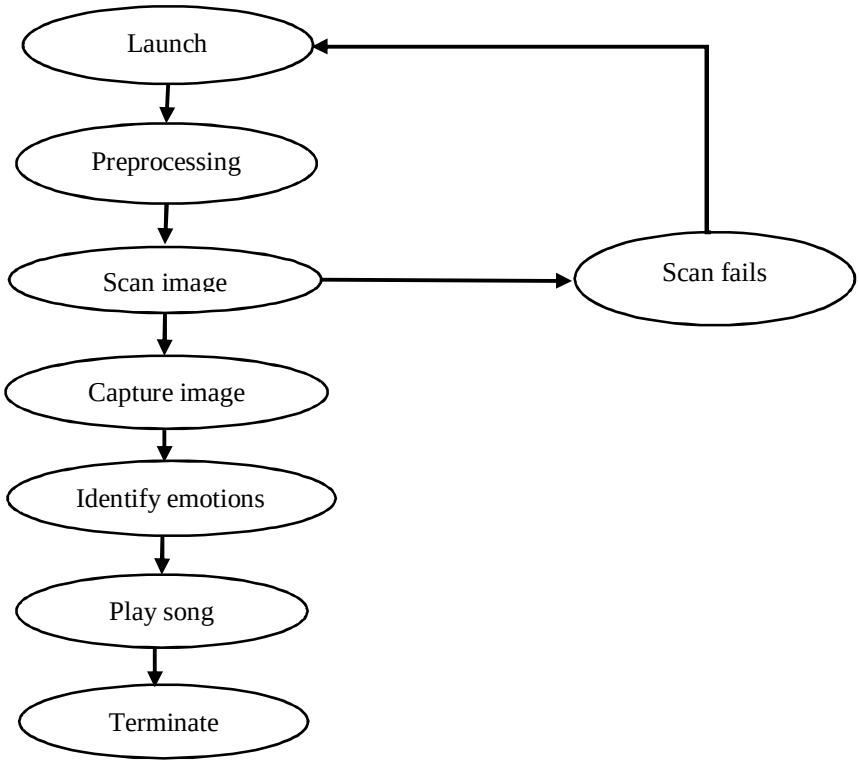


Fig1. System Methodology

B. Dataset

The dataset used to implement the FER system was the FER2013 dataset from Kaggle. The data set consists of 35,887 annotated images split into 3589 test images and 28709 training images. The dataset consists of 35,887 annotated images split between 3589 test images and 28709 training images. The dataset consists of 3589 additional private test images on which the final test was run during the challenge. The images in the FER2013 dataset are 48x48 in size and black and white. The FER2013 dataset contains images with different viewing angles, lighting, and scales. Table 1 shows the dataset description.

Emotions	No. of images
Angry	4593
Happy	8989
Sad	6077
Fear	5121
Surprise	4002
Neutral	6198
Disgust	547

Table1: Description of FER dataset

©IJRASET: All Rights are Reserved

C. Facial Expression Recognition Process

The FER process consists of three phases. The pre-processing phase consists of preparing the dataset in a format that works with generalized algorithms and produces efficient results. In the face detection phase, faces are detected from images captured in real time. The sentiment classification step consists of implementing a CNN algorithm to classify input images into one of seven classes. These stages are described using a flowchart in *Fig2.*



Fig2. Flowchart for facial emotion recognition

D. Pre-processing

FER input images may contain noise, lighting variations, size, and color. Several pre-processing operations were performed on the images in order to obtain accurate and fast results from the algorithm. The pre-processing strategy used is to convert the image to grayscale, normalize the image and resize it.

1) Normalization – Image normalization is performed to remove lighting variations and obtain an enhanced facial image from the image.
2) Gray scaling is done because color images are difficult to process with algorithms.
3) Resize - The image is resized to remove unwanted parts of the image. This reduces memory requirements and increases computation speed.

E. Face Detection

When frames enter the face detector, the system first applies a feature-based hair cascade classifier to detect human face regions, which OpenCV already implements as a built-in function. The regions detected by OpenCV are cropped from the original frame and converted to grayscale.

F. Emotion Classification

In this step, the system classifies images into one of seven universal representations (happiness, sadness, anger, surprise, disgust, fear, and neutrality) as identified in the FER2013 dataset. increase. The training was performed using CNNs, a category of neural networks with proven productivity in image processing. The dataset was first split into a training data set and a test data set, then trained on the training set. No feature extraction process was run on the data before it was input to the CNN.

The emotion classification step consists of the following phases:

1) *Data Split:* The dataset was split into three categories: training, public testing, and private testing, according to the label "usage" of the FER2013 dataset. We used the Training and Public Test sets to generate the model, and the Private Test set to evaluate the model.
2) *Training and Model Creation:* A neural network architecture consists of the following layers:
a) Convolutional Layer: In a convolutional layer, randomly instantiated learnable filters are slid or convolved on the input. This operation performs a dot product between the filter and each local extent of the input. The output is a 3D volume composed of multiple filters, also called feature maps.
b) Max Pooling: Pooling layers are used to reduce the spatial size of the input layer to reduce the input size and computational cost.
c) Fully Connected Layer: In a fully connected layer, each neuron in the previous layer is connected to an output neuron. The final output layer size corresponds to the number of bins into which the input image is classified.
d) Stack normalization: Stack normalization applies a transformation that speeds up the training process, bringing the mean activation closer to 0 and the activation standard deviation closer to 1.

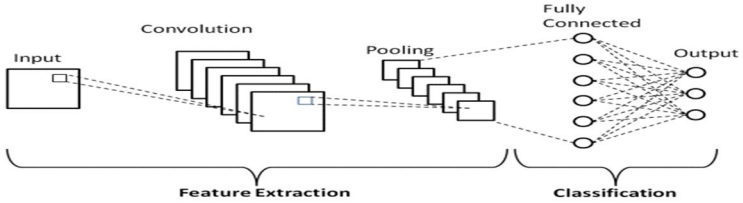


Fig3. CNN model layers.

3) *Model Evaluation:* The model generated during the training phase was evaluated on a validation set of 3589 images.
4) *Real-time Image Classification Using Models:* The concept of transfer learning can be used to recognize emotions in real-time captured images. The model generated during the training process consists of pre-trained weights and values that can be used to implement new facial expression recognition problems. FER is faster for real-time images because the generated model already contains the weights.

G. Feature Extraction

In this phase, all extracted features are calculated and determine the eyes, mouth, and nose location on a person's face. On basis of this calculation, face motion is detected.

H. Emotions Detection and Music Recommendation

By applying CNN classifier to the extracted features, the emotion Happy, Neutral and Sad are detected. Based on the user's mood like sad, angry, party relaxed, and happy the particular song from the playlist is played.

IV. CONCLUSION

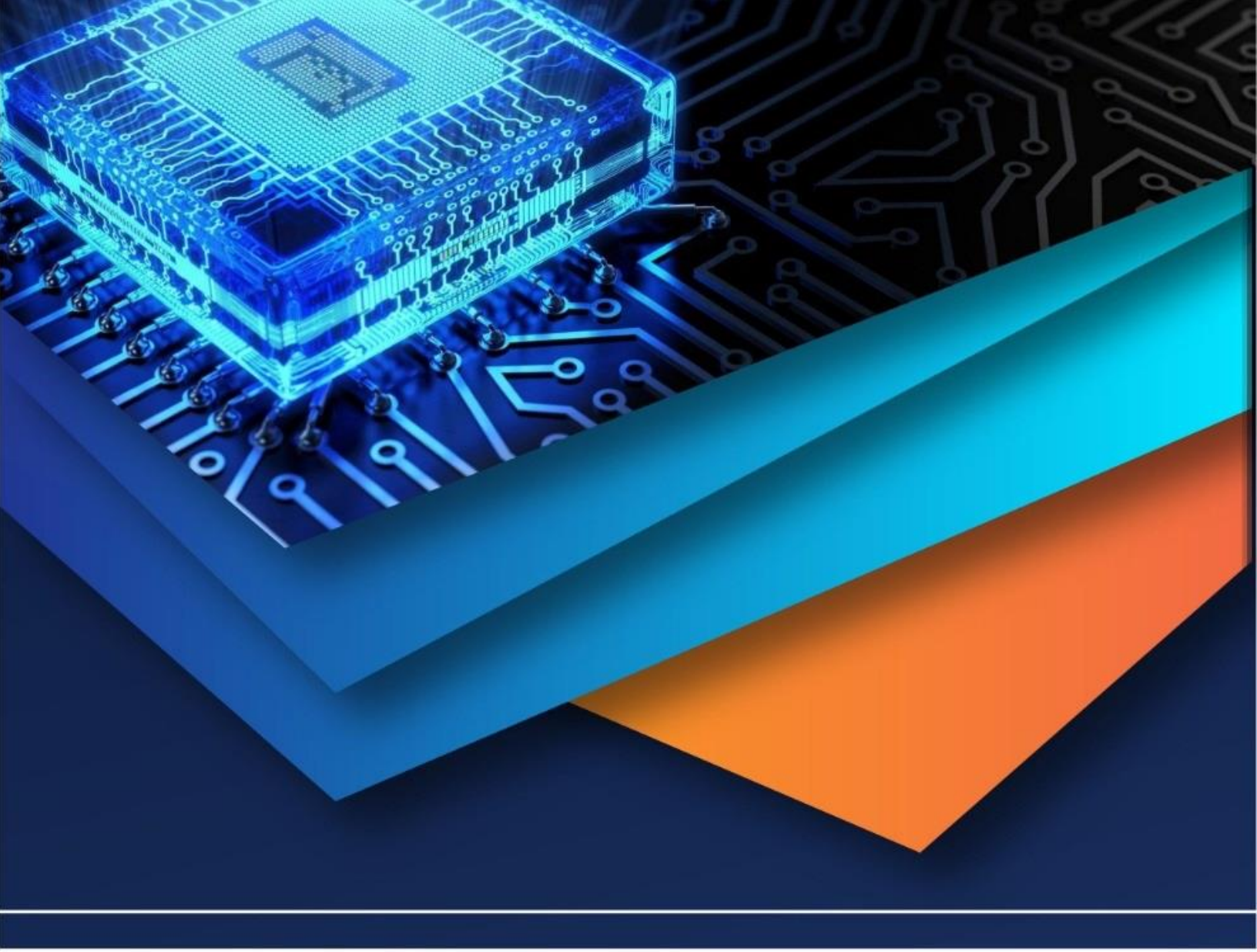
Facial features are one of the most powerful channels of emotion recognition. Convolutional Neural Networks (CNN) can be used as an emotion detection solution. As the accuracy of facial expression recognition technology improves, its scope of application will expand both in innovation and accuracy and in both changes in mechanical events and in the past. In this article, we propose an emotion detection model that recommends music based on the user's mood.

V. ACKNOWLEDGEMENT

The authors would like to thank the management and directors of the SKN Sinhgad Institute of Technology and Science, Lonavala for their support and guidance in providing the necessary support and guidance to complete this work.

REFERENCES

[1] Londhe RR and Pawar DV 2012 Analysis of facial expression and recognition based on statistical approach International Journal of Soft Computing and Engineering
[2] Florence, S. Metilda, and M. Uma. "Emotional Detection and Music Recommendation System based on User Facial Expression." IOP Conference Series: Materials Science and Engineering. Vol. 912. No. 6. IOP Publishing, 2020.
[3] Vinay p, Raj p, Bhargav S.K., et al. "Facial Expression Based Music Recommendation System" 2021 International Journal of Advanced Research in Computer and Communication Engineering, DOI: 10.17148/IJARCCE.2021.10682
[4] Zeng Z, Pantic M, Roisman GI and Huang TS 2008 A survey of affect recognition methods Audio, visual, and spontaneous expressions IEEE transactions on pattern analysis and machine intelligence
[5] James, H. Immanuel, et al. "EMOTION BASED MUSIC RECOMMENDATION SYSTEM." EMOTION 6.03(2019).
[6] ParulTambe, YashBagadia, Taher Khalil and Noor UlAin Shaikh 2015 Advanced Music Player with Integrated Face Recognition Mechanism International Journal of Advanced Research in Computer Science and Software Engineering
[7] Hui-Po Wang, Fang-Yu Shih,"Detect and Transmit Emotions in Online Chat using Affective Computing", National Tsing Hua University, Hsin Chu, Taiwan,2020.
[8] Yading Song, Simon Dixon, Marcus Pearce, "EVALUATION OF MUSICAL FEATURES FOR EMOTION CLASSIFICATION", University of London, ISMR, 2012.
[9] Aayush Bhardwaj ; Ankit Gupta ; Pallav Jain ; Asha Rani ; Jyoti Yadav "Classification of human emotions from EEG signals using SVM and LDA Classifiers", 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)
[10] Rahul Hirve1, Shrigurudev Jagdale2, Rushabh Banthia3, Hilesh Kalal4& K.R. Pathak5, "EmoPlayer: An Emotion Based Music Player", Imperial Journal of Interdisciplinary Research (IJIR) Vol-2, Issue-5, 2016
[11] F. Abdat, C. Maaoui and A. Pruski,"Human-computer interaction using emotion recognition from facial expression", 2011 UKSim 5th European Symposium on Computer Modelling and Simulation
[12] Dubey, M., Singh, P. L., "Automatic Emotion Recognition Using Facial Expression: A Review," International Research Journal of Engineering and Tech nology,2016.
[13] Duncan, Dan, Gautam Shine, and Chris English. "Facial emotion recognition in real time." Computer Science (2016): 1-7



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)