iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.59264

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 III Mar 2024- Available at www.ijraset.com

Human Face Emotion in 3D Using Machine Learning

Dr. Ravi Kumar[1], V. Mahesh[2], T. Ajay[3] N. Keerthi[4]
UG Student, Department of Computer Science & Engineering, CMR College of Engineering & Technology, Hyderabad, India

Abstract: ***Emotion detection, which is an effortless task for humans, is complex to perform on machines. The interaction between human beings and computers are very natural if computers can be able to perceive and respond to human non-verbal communication such as emotions. This project targets to develop a multi-modal emotion detection web app that utilizes advanced computer vision, speech, and natural language processing (NLP) techniques to accurately identify and track human emotions in images, and video streams. Automatic emotions recognition is based on face expressions is a engaging discipline, which has several areas such as safety, health and in human interaction interfaces. Utilizing neural networks in conjunction with feature e traction techniques enables the recognition of various facial emotions such s happiness, sadness, anger, fear, surprise, and neutrality through analysis of facial expressions.. So it is very important to detect these emotions on the face. Human beings are can capable to produce thousands of face reactions during communication that can be vary in Depth, Strength, and Significance. We are using Convolutional Neural Network(CNN), to extract features from image to detect emotions and RNN and classify facial emotions. CNN is very effective for emotions recognition task. They can extract features from input image, and a real time video and then use these features to train a classifier. Facial expression analysis software like Face Reader is ideal for collecting this emotion data. The software automatically analyzes the expressions Joyful, Unhappy, Furious, surprised, scared, disgusting, and Neutral. For Speech emotion Support Vector Machine (SVM) is used. At present, mostly the existing emotion recognition algorithms use the extraction of two-dimensional facial features from images to perform facial emotion prediction. But the rate of facial recognition of two-dimensional facial feature is not optimal, by using the effective algorithms a two-dimensional image and video from the input end constructs a3D face model from the output end. Therefore, by using 3D facial information we can estimate the continuous emotion of the dimensional space.***

Keywords: Convolutinal Neural Network, Feature extraction, Facial emotion recognition.

I. INTRODUCTION

Facial Expression Recognition (FER) and Facial Emotion Recognition (FER) are computer based technologies that play crucial roles in understanding human emotions through facial expressions. The process involves three main phases: Face detection, facial landmark detection, and facial expression/emotion classification. Initially, faces are detected within images or video frames, followed by the identification of key facial expression landmarks to understand facial structure. Finally, mathematical algorithms analyze facial movements to classify expressions into emotion states such as happiness, sadness, anger, surprise, fear, disgust, or neutrality. FER is a subset of emotion recognition, which encompasses various modalities including verbal expressions, body language, and gestures. These technologies find applications in diverse fields such as human-computer interaction[1], psychological research, and market analysis, enhancing our understanding of emotional responses in different contexts.

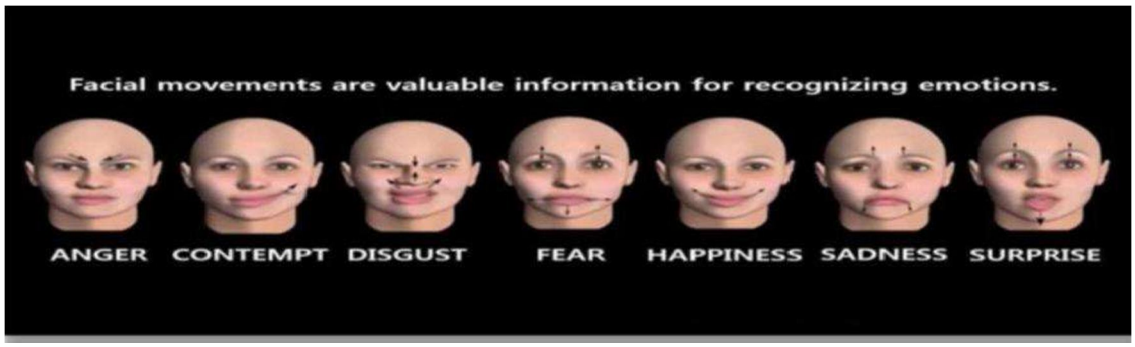


Fig 1 Types of Emotions

©IJRASET: All Rights are Reserved

We know Human face expressions are important for communicating emotions and intentions during interactions. These expressions involve subtle changes in key facial features such as eyebrows, eyes, mouth, cheeks, nose, jaw, and chin. For instance, raised eyebrows might signify surprise, while a smile typically indicates happiness. Dilated pupils can suggest excitement, while a clenched jaw may reveal anger or frustration. Facial emotions and their analysis play an important role in non-verbal communication. It makes oral communication more efficiently and conducive to understanding the concepts [2]. Face detection is thefirst step of locating or detecting face(s) in a video or only one image in the FER process. The images do not consist of faces only, but instead present with complex backgrounds.

II. LITERATURE SURVEY

Meng Wang[3] In the realm of 2D-to-3D image conversion methods, those employing human operators have shown remarkable success, at the expense of being time-consuming and costly. Conversely, automatic techniques, which typically rely on a predetermined 3D scene model, have yet to achieve comparable quality due to their reliance on assumptions easily invalidated in real-world scenarios. In this, we propose a novel approach centered on "learning" the 3D scene structure. We introduce a simplified and computationally efficient version of our recent 2D-to-3D image conversion algorithm. By leveraging a repository of 3D images, either in stereopairs or image+depth pairs, we identify k pairs whose photometric content closely matches that of a 2D query awaiting conversion. Subsequently, we fuse the corresponding depth fields from these pairs and align the resulting fused depth with the original 2D query. Unlike our prior work, we quantitatively validate this simplified algorithm using a dataset comprised of images and depth information captured by Kinect, comparing its performance against the Make3D algorithm. Although not without flaws, our findings demonstrate the capability for utilizing online repositories of 3D content to facilitate effective 2D-to-3D image conversion.

Navuluri Sainath[4] The primary objective of the paper is to identify the emotion conveyed by speakers during speech. Emotion detection has become increasingly important in contemporary times. Speech characterized by fear, anger, or joy tends t exhibit higher and wider pitch ranges, while emotions like calmness or neutrality typically have lower pitch ranges. Detecting emotions in speech is valuable for enhancing human –machine interactions. These models were trained to recognize a range of emotions including calmness, neutrality, surprise, happiness ,sadness, anger, fear, and disgust. The study yielded an accuracy of 86.5%, which was successfully replicated during testing with input audio samples.

Ioannis Hatzilygeroudis[5] Emotions constitute an innate and important aspect of human behavior that colors the way of human communication. The accurate analysis and explanation of the emotional content of human facial expressions is essential for the deeper understanding of human behavior. Although a human can detect and interpret faces and facial expressions naturally, with little or no effort, accurate and robust facial expression recognition by computer systems is still a great challenge. The evalution of a person's face characteristics and the recognition of its emotional states are considered to be very challenging and difficult tasks. The main difficulties come from the non-uniform nature of human face and variations in conditions such as lighting, shadows, facial poseand orientation. Deep learning approache have been examined as a stream of methods to achieve robustness and provide the necessary scalability on new type of data. In this , we examine the implementation of two known deep learning approaches (GoogLeNet and AlexNet) on facial expression recognition, more specifically the recognition of the existence of emotional content.

III. METHODOLOGY

A. *Algorithms*

The algorithms used are CNN a machine learning algorithm and RNN algorithm. These two are useful to detect the face emotion.

B. *CNN*

It is a deep learning algorithm. The Convolutional Neural Network (CNN) algorithm typically involves several key steps:

1) *Input:* Receive an input image. Images are usually represented as multidimensional arrays of pixel values.
2) *Convolution:* Apply a collection of convolutional filters to the input image. Each filter extracts certain aspects from the image bycomputing dot products between the filter and local area of the input.
3) *Activation:* Apply an activation function (like ReLU - Rectified Linear Unit) to introduce non-linearity, allowing the network tolearn difficult in patterns in the data.
4) *Pooling:* Perform pooling (often max pooling) to decrease the magnitude of the feature maps, making the network morecomputationally efficient and reducing overfitting.

5) *Flattening:* Flatten the 2D pooled feature maps into a 1D vector, which can be fed into a fully connected neural network.
6) *Fully Connected Layers:* Pass the flattened vector through one or more fully connected layers to learn difficulty in patterns in thedata.
7) *Output:* Generate the final output, often through a function for classification tasks, which gives the probability distribution overdifferent classes.
8) *Loss Calculation:* Calculate the loss between the predicted output and the actual target labels
9) *Backpropagation:* Use an optimization algorithm (e.g., Stochastic Gradient Descent) to update the weights of the network,minimizing the loss and improving the model's performance.
10) *Repeat:* Iterate through the steps above (especially steps 2-9) multiple times (epochs) on the training dataset to improve themodel's performance.

C. RNN

It is a deep learning algorithm. A Recurrent Neural Network (RNN) is a deep learning model that is trained to process and convert a sequential data input into a specific data output.The Recurrent Neural Network (RNN) algorithm is designed to handle sequential data by maintaining an internal state/memory. Here are the key steps involved in an RNN algorithm:

1) *Input:* Receive a sequence of input data. Each element in the sequence is typically represented as a vector. Initialization: Initialize the internal state/memory of the RNN. This is usually a vector of zeros or a learnable parameter.
2) *Recurrent Step:* For each element in the input sequence, update the internal state/memory of the RNN and generate an output. Thisis done using a set of weights that are shared across all time steps.
3) *Output:* Generate an output for each element in the input sequence. The output can be based on the internal state/memory of theRNN at that time step.
4) *Loss Calculation:* Calculate the loss between the predicted output sequence and the actual target sequence.
5) *Backpropagation Through Time (BPTT):* Use an optimization algorithm (e.g., Gradient Descent) to update the ranges of the RNN,minimizing the loss and improving the model's performance. BPTT is a variant of backpropagation designed for sequential data
6) *Repeat:* Iterate through steps 3-6 multiple times (epochs) on the training dataset to improve the model's performance. a deep learning algorithm. It is 19 layers deep Convolutional Neural Network algorithm. It is an extension of VGG16. VGG19 (Visual Geometric Group) is used for classifying images. To use it, we have to import Keras function and TensorFlow function. Because, it can easily assign weights with other frameworks.

D. Implementation of Block Diagram

Firstly, we have created modules for each phase.

Modules:

The various intents the face and signature classifier is experienced on are:

1) *Upload Image:* If no image is uploaded then a page will be redirected to the same screen.
2) *Detect Emotions:* User has to upload image then the system will predict the emotion using trained model.
3) *3D Modeling:* After predicting emotion the image will be send to BFM model which will make it into 3d.

> Data Analysis:
- Image converting into grayscale using CV2
- Resizing gray Image to 48x48 pixels
- Reshaping image using Numpy
-

>Face Detection:
- Detecting faces using “haarcascade_frontalface_default.xml”.
- Cropping image to face shape.

> Emotion Analysis:
- Loading model structure using json file.
- Predicting emotion using CNN trained model (model.h5)

> 3D Model :
- Loading BFM model for making 3d face using torch, aspose-3d and mediapipe.

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

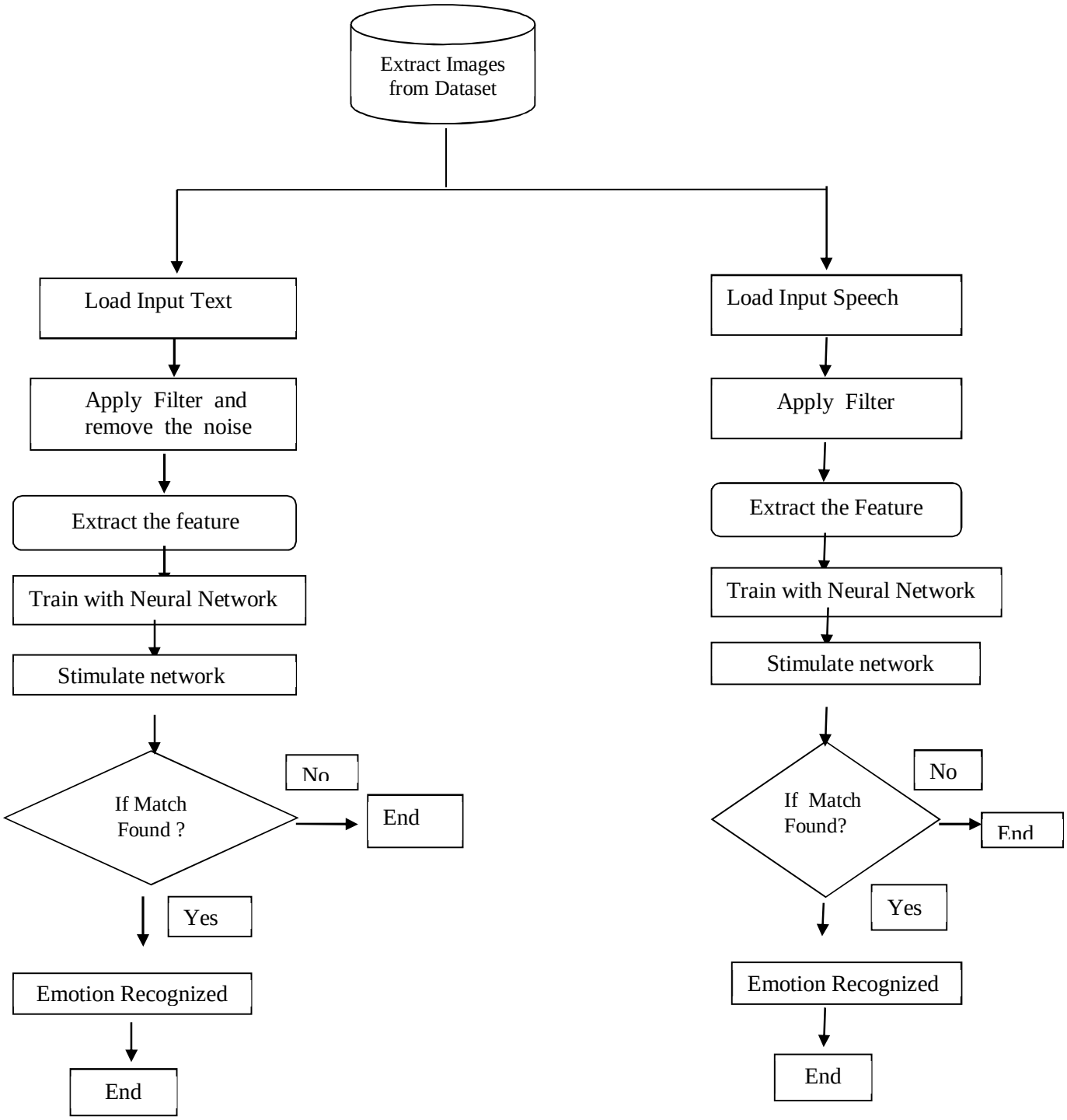


Fig 1: Model Architecture

©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 III Mar 2024- Available at www.ijraset.com

IV. RESULTS AND DISCUSSION

A. Figures

To run the modules for execution and get output screen a web page is created. When we click on it the execution starts andwe get the screen.



Fig 3: Welcome Page

To identify the Image Emotion we compare the uploaded image and Image in dataset. Then it shows the result .

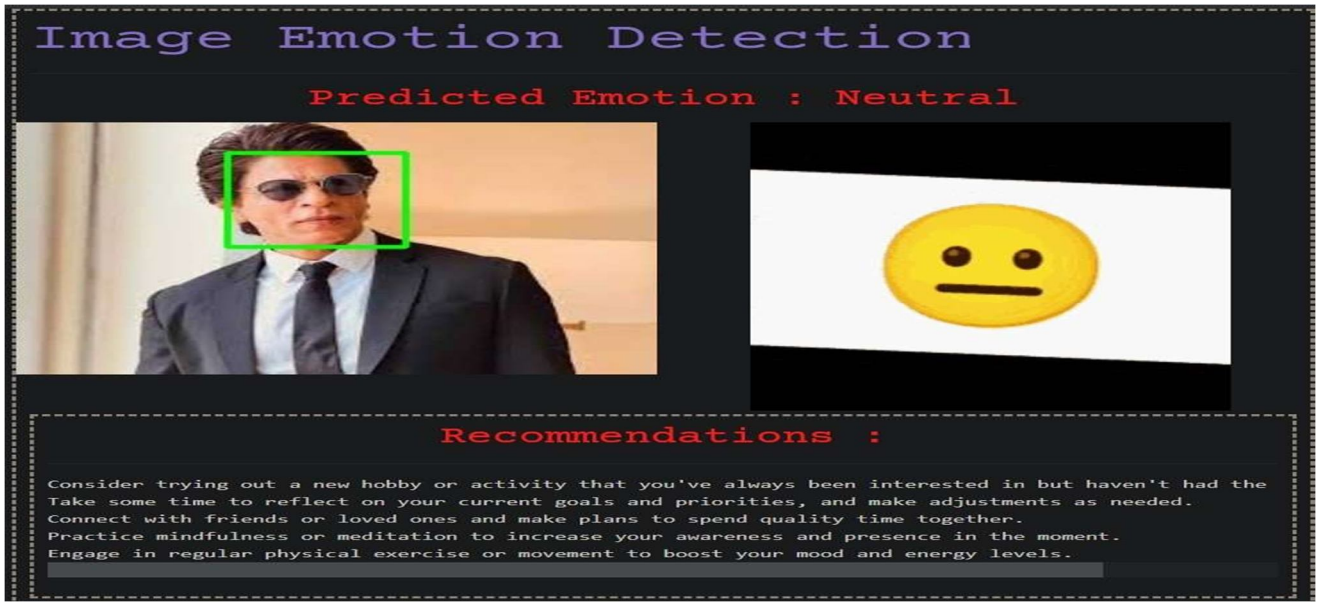


Fig 4: Image Emotion Detection



Fig 5 Real Time Emotion Detection

©IJRASET: All Rights are Reserved

To identify the Emotion through the speech then we upload the speech by audio file then it shows the results as an emotion.



Fig 5: Speech Emotion Detection.

To identify the Emotion through the text we enter the text and then by the text it shows the emotion as results.

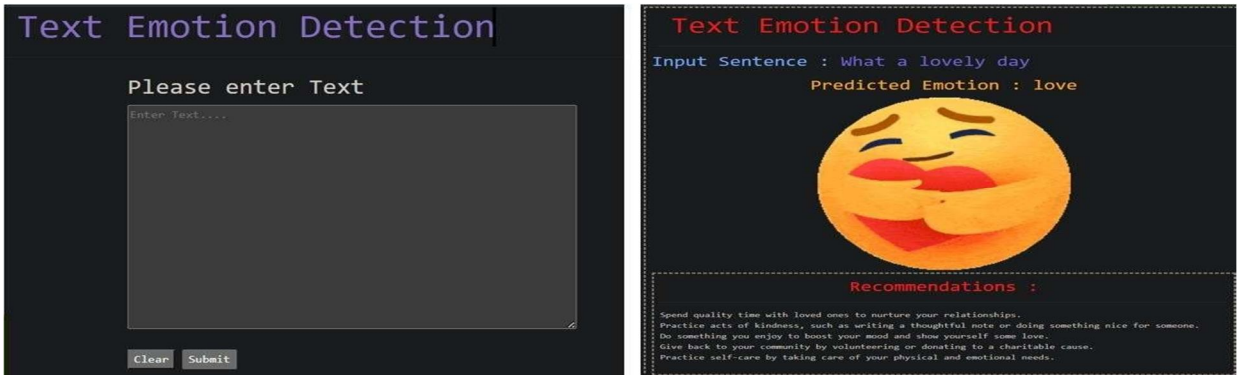


Fig 6: Text Emotion Detection

V. CONCLUSION

In conclusion, this project on multi-modal emotion detection has shown that combining various deep learning models can accurately detect emotions from facial expressions. By using different models like CNNs, RNNs, and LSTM, we achieved a high accuracy in identifying emotions such as anger, disgust, fear, happiness, sadness, and surprise. Additionally, the project goes beyond typical emotion detection systems by creating a 3D file of the input image. This innovative feature allows for a more immersive and interactive experience for users, enabling them to visualize and explore the emotions detected in a more engaging and dynamic way. The potential applications of this project are numerous, including in fields such as psychology, market research, and human computer interaction. Overall, this project represents a significant advancement in the field of emotion detection, offering a more comprehensive and interactive approach to understanding and analyzing human emotions.

REFERENCES

[1] Abadi, Martín & Barham, Paul & Chen, Jianmin & Chen, Zhifeng & Davis, Andy & Dean, Jeffrey & Devin, Matthieu & Ghemawat, Xiaoqiang. (2016). TensorFlow: A system for largescale machine learning.
[2] Sarraf, A. (2023). Binary Image Classification Through an Optimal Topology for Convolutional Neural Networks.
[3] Abu, Mohd Azlan & Indra, Nurul Hazirah & Abd Rahman, Abdul & Sapiee, Nor & Ahmad, Izanoordina. (2019). A study onImage Classification based on Deep Learning and Tensorflow.

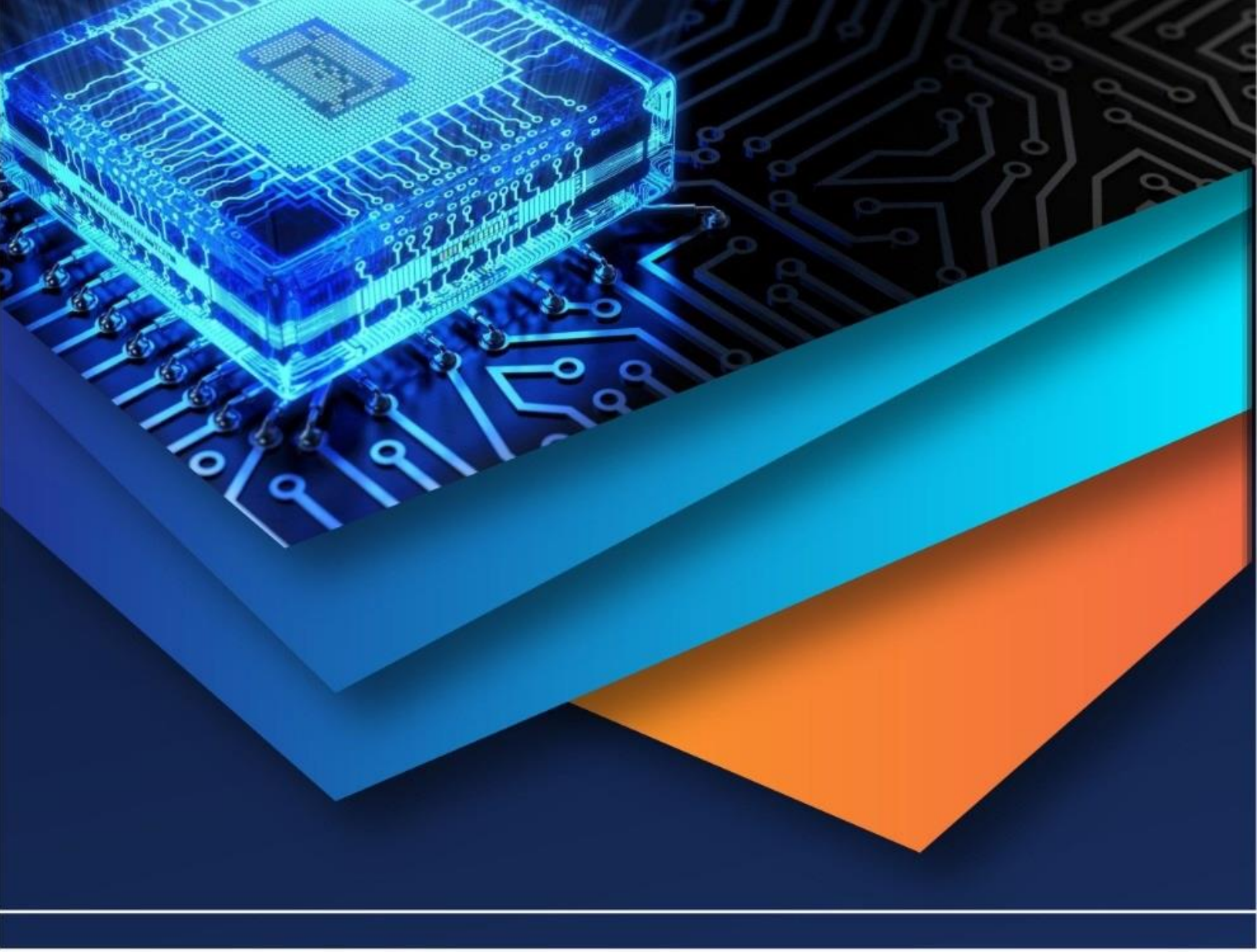
[4] A Lekdioui, Khadija; Ruichek, Yassine; Messoussi, Rochdi; Chaabi, Youness; Touahni, Raja (2017). [IEEE 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP) - Fez, Morocco (2017.5.22-2017.5.24)] 2017 face-regions., (), 1–6. doi:10.1109/ATSIP.2017.8075517

[5] Mostafa Mohammadpour, Seyyed Mohammad. R Hashemi,Shahrood, Hossein Khaliliardali , Mohammad. M AlyanNezhadi (2017) “Facialemotionrecognitionusingdeepconvolutionalnetworks”.EEE4thIInternationalConferenceonKnowledge-Based Engineering and Innovation (KBEI).doi: 10.1109/KBEI.2017.8324974.

[6] R. K. Madupu, C. Kothapalli, V. Yarra, S. Harika And C. Z. Basha, "Automatic Human Emotion Recognition System using Facial Expressions with Convolution Neural Network," 2020 4th International Conference on Electronics,Communication and Aerospace Technology(ICECA),2020,pp.1179-1183,doi:10.1109/ICECA49313.2020.9297483.

[7] DanielC.Tozadore, CaetanoM.Ranieri,Guilherme V.Nardari,RoseliA.F.Romero,VitorC.Guizilini,"EffectsofEmotionGrouping for Recognitionin Human-Robot Interactions", Intelligent Systems(BRACIS) 2018 7th Brazilian Conference on, pp. 438-443,2018.

[8] Adam H.M. Pinto, Caetano M. Ranieri, Guilherme Nardari, Daniel C. Tozadore, Roseli A.F. Romero, "Users' Perception Variance in Emotional Embodied Robots for Domestic Tasks", Robotic Symposium 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education



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)