iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** IV **Month of publication:** April 2022

DOI: https://doi.org/10.22214/ijraset.2022.41121

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 IV Apr 2022- Available at www.ijraset.com

Real-Time Passenger Safety System Using Facial Emotion Recognition

Anuraj Srivastava[1], Anusha Kukreja[2], Satya Bhanu Khichi[3]
[1, 2, 3]*School of Computer Science and Engineering (SCOPE), VIT, Vellore, Tamil Nadu, India*

Abstract: ***In this paper, a "REAL-TIME PASSENGER SAFETY SYSTEM USING FACIAL EMOTION RECOGNITION" is presented using the help of Neural Network Models. Our facial expressions keep on changing, and we have the intelligence of understanding the meaning of several of them. Broadly, we can classify the facial expression as - Happy, Sad, Angry, Scared, Surprised, Disgusted and Neutral. So, the major aim in this project is to show the mood of the person in front of the camera. And this requires that we provide our computer with the intelligence required to do so like our brains. Our brains have neural networks which are responsible for all kinds of thinking (decision-making, understanding) that we do, and we try to develop these neuron capabilities artificially called artificial neural network. Moreover, we will be using the concept of these neural networks to build the application which can recognize and inform about the expression on the face of the passenger. I will also work on adding a feature which notifies close contacts, in case of negative emotions like fear, sad etc.***
Keywords: Emotion Recognition, Facial Expression, CNN, Deep Learning, Computer Vision

I. INTRODUCTION

Between Verbal & Non-Verbal form of communication, facial expression is a form of non-verbal communication but it plays a pivotal role. It expresses the human perspective or filling & his or her mental situation. Facial expressions play a very important role in human communication The human face is the richest source of emotions. Therefore, a system that detects emotions from facial expressions would be widely applicable, especially in the field of safety and security.

II. LITERATURE SURVEY

This project concerns about recognition system through identifying the facial expression of human being. It has often been said that the eyes are the "window to the soul". This statement may be carried to a logical assumption that not only the eyes but the entire face may reflect the "hidden" emotions of the individual. The human face is the most complex and versatile of all species. For humans, the face is a rich and versatile instrument serving many different functions, it serves as a window to display one's own motivational state. This makes one's behaviour more predictable and understandable to others and improves communication. A quick facial display can reveal the speaker's attitude about the information being conveyed. Research in psychology has indicated that at least six emotions are universally associated with distinct facial expressions. The six principal emotions are happiness, sadness, surprise, fear, anger, and disgust. Several other emotions and many combinations of emotions have been studied but remain unconfirmed as universally distinguishable. In [1], Machine recognition of facial expression by correctly integrating audio and video data is discussed while weighting matrix is used a matric. The issues with this approach are that sometimes emotion Misclassification happens. While in [2], The system was developed to operate as a user-independent system, based on physiological signal databases obtained from multiple subjects. vector machine was adopted as a pattern classifier, it's drawback was that they could not employ non- linear or chaotic analyses, as they usually require long- term monitoring of signals as they used Vector as a metric. In [5], the approach where Distance measures extracted from 3D face vectors was used on the BU-3DFE database. The drawback was that almost all of the methods developed to use the 2D distribution of facial features as inputs into a classification system, and the outcome is one of the facial expression classes. They differ mainly in the facial features selected and in the classifiers used to distinguish between the different facial expressions. Whereas, in [8], A Human Facial Expression Recognition Model based on Eigen Face Approach is discussed where Face Detection is done using Hue-Saturation Value. Finally, in [9], Haar Classifier based method for face detection LBP based feature extraction method Dimensionality is reduced using Principal Component Analysis. Its limitation was The proposed methodology is limited to classify frontal image only. However, rotation of face or occlusions degrades the performance of the system. Thus, on the basis of the above previously done works and researches, the idea to use CNN model for classification of facial emotion was discussed. It was further implemented as a program which uses live feed to capture the face and detect the emotion in real-time. Moreover, in case of a negative emotion, an e-mail is also sent by the system to the concerned authorities asking for help ASAP.

©IJRASET: All Rights are Reserved

III.PROPOSED WORK DETAILS

A. Objects of Proposed System

1) *Field of Invention*: The present invention relates to systems and methods for automatic facial recognition, including one-to-one and one-to-many correlations. More particularly, it relates to a sys- tem for correcting pose or lighting prior to determining recognition.
2) *Discussion of Related Work*: Automated facial recognition systems, used to identify individuals from images of faces, have existed for some time. There are several different types of facial recognition systems. In all such systems, a newly acquired image is compared to one or more stored images. Generally, facial recognition systems can be separated into two categories: authentication systems and identification system. Generally, authentication systems provide more uniform images than for identification systems. Often, in an identification system, the subject is not cooperative and lighting conditions are poor. Therefore, a need exists for a system which allows modification of images to correct for differences in lighting or pose, in either identification systems or authentication systems.

B. Summary of Proposed Work

The project substantially overcomes the deficiencies of the prior work by providing a facial recognition system which processes images to correct for lighting and poses prior to comparison. According to one aspect of the proposed system, the images are corrected for lighting and pose by using shape information. The system processes two-dimensional image of a face to create a three-dimensional image of the face. The three-dimensional image is manipulated to change the pose and lighting characteristics. Finally, the modified three-dimensional image is converted back to a two-dimensional image prior to processing for recognition.

As per another aspect of the invention, the three-dimensional image is manipulated to be facing forward and with diffuse light from the front. A facial recognition system compares a newly acquired image of a face to images of faces in a database to determine a match. The newly acquired image includes one or more two-dimensional images. The three-dimensional images are stored in a database for later comparison with another image. The three-dimensional images are converted into two-dimensional images before being stored in a database. An iterative process is used to create a three-dimensional image from the original two-dimensional image. An initial shape is used with data from the two-dimensional image to create a three-dimensional shape. The three-dimensional shape is iteratively adjusted to match the original image. At each iteration, a two-dimensional image is rendered from the three-dimensional image

C. Process Flow of the Proposed System

This project aims to predict the emotions of the person by his/her facial expression and swap the appropriate emotion title enclosed in a rectangular box around the face. This will continue on a real-time live feed.

- *Step 1: Preparing the dataset*
  It involves the preparation of dataset upon which the learning algorithm will work. We will be using different datasets of faces available on internet example CK+ dataset and one of our own and applying it to convolution neural networks. We will be training the data to recognize different emotions.
- *Step 2: Detecting Real-time passengers' faces*
  This is done by HAAR cascade function in OpenCV After detecting the faces the image is converted to greyscale and is resized to the same size as the images in a dataset.
- *Step 3: Learning to Recognise Emotions*
  This step involves training the program to recognise and differentiate between the emotions. Training is done by Convolution Neural Network, a deep learning algorithm and will be on the layers on the dataset of faces.
- *Step 4: Swapping emoticons in place of faces*
  This is the final step and it involves placing the appropriate emoticon over the face of the person according to his/her emotion. The HAAR cascade function returns the coordinates of the faces detected and these coordinates can be used to place the emoticon at the exact place.
- *Step 5: Notifying Taxi Company in case of an emergency*
  When the software detects Fear continuously for 10 seconds, an auto-mail is generated and sent to the company notifying them, the passenger(s) are scared or are in trouble.

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

D. Architecture of Proposed Work

In this project, we will be making the expression recognition system using the DCNN (deep convolution neural network). This approach has been proven to be a quiet much better than CNN. Convolutional Neural Networks (CNN) is composed of two basic layers, respectively called convolutional layer (C layer) and subsampling layer (Slayer). Different from general deep learning models, CNN can directly accept 2D images as the input data, so that it has a unique advantage in the field of image recognition. 2D images are directly inputted into the network, and then convoluted with several adjustable convolutional kernels to generate corresponding feature maps to form layer C1. A classic CNN model is shown as in the below figure.

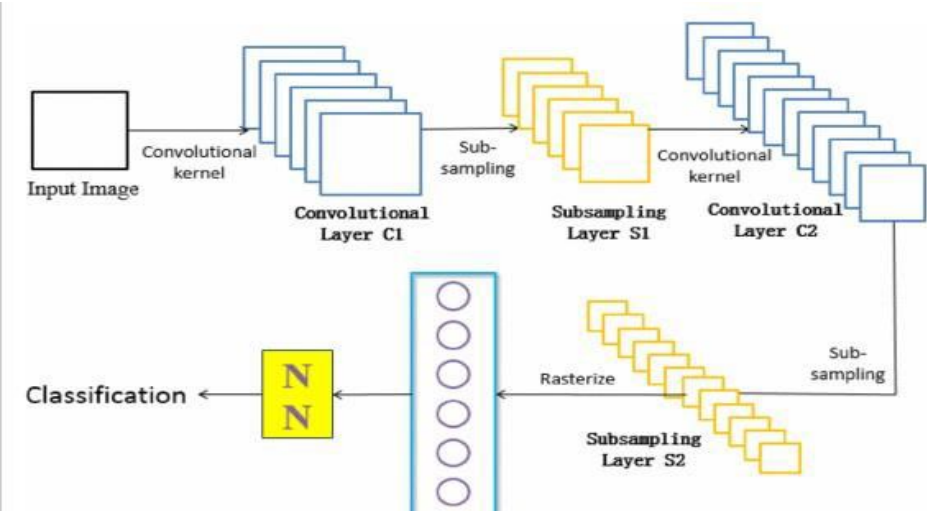


Fig. 1 CNN model - architecture of layers

2D images are directly inputted into the network, and then convoluted with several adjustable convolutional kernels to generate corresponding feature maps to form layer C1. Feature maps in layer C1 will be subsampled to reduce their size and form layer S1. Normally, the pooling size is 2×2. This procedure also repeats in layer C2 and layer S2. After extracting enough features, the two-dimensional pixels are rasterized into 1D data and inputted to the traditional neural network classifier.

E. Algorithms Used

1) Face Detection Algorithms: For static images, Viola-Jones, which achieves fast and reliable detection for frontal faces will be used.

2) Feature Extraction Technique: After the pre-processing step, the feature extraction technique is applied to the face image to extract the features from the image. In this work approximation image, Gabor local binary pattern features method is applied on the face image for extracting the features as discussed below.

a) *AIG BP (Approximation image Gabor Local Binary Pattern):* Initially, on the face images, bi-level wavelet decomposition method has been applied, which transforms face images into approximation images. Then, on approximation images, Gabor filter has been applied, which restricts distortion and noise present at a distinct location in the image up to a certain extent and also provides robustness against brightness and contrast of images.

b) *Bi-Level Wavelet Decomposition*: Wavelet decomposition method is a time-frequency signal analysis method. It can be used to decompose a face image into many sub-band images with different spatial resolution, frequency characteristic and directional features. Approximation coefficients contain the lowest frequency components and details coefficients contain the highest frequency component of an image. Instead of considering the entire coefficients, only approximation coefficients are taken for further procedure.

c) *Gabor Filters*: In the fields of computer vision, pattern recognition and image processing, Gabor filter have a large number of applications. 2D Gabor filter is a selective filter in terms of frequency and orientation. Gabor filter response hasn't been disturbed by noise and distortion exists at different locations due to accuracy in time-frequency localization. Gabor function has cosine (real part) and sine (imaginary part) functions of the sinusoidal plane. In the proposed work, for further processing, only the real parts of the Gabor function are to be considered. The experimental analysis of the imaginary part also proves the same.

d) Principles of Local Binary Pattern: The outcome of the Gabor filter has been applied to LBP (local binary pattern) in which face image is first split into small regions from which local binary patterns also known as histograms are being extracted. Different LBP histograms extracted from each of the face images are finally concatenated into single feature histogram, which forms the representation of the face image. Using AIG BP method for the feature extraction process, when we applied LBP technique along with wavelets and Gabor filters, then the computational time of overall process gets reduced, which increases the efficiency of the system and also tends to improve the overall performance of the system.

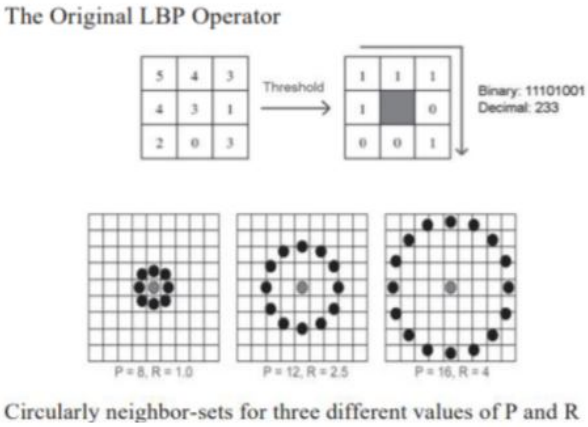


Fig. 2 LBP- Local Binary Pattern

IV. EXPERIMENTAL ANALYSIS

Python platform is used to perform all the simulations operations. The performance is measured by classification rate (CR). The classification rate is considered as the percentage of correctly classified face images with the overall face images in the testing set

Face databases used for experimentation The evaluation of the proposed system has been done by using a number of publicly available face databases having sufficient amount of male and female images with their facial expressions taken in conditional as well as unconditional Environments. Face databases used in this work are:

1) FERET
2) INDIAN FACE
3) AR FACE

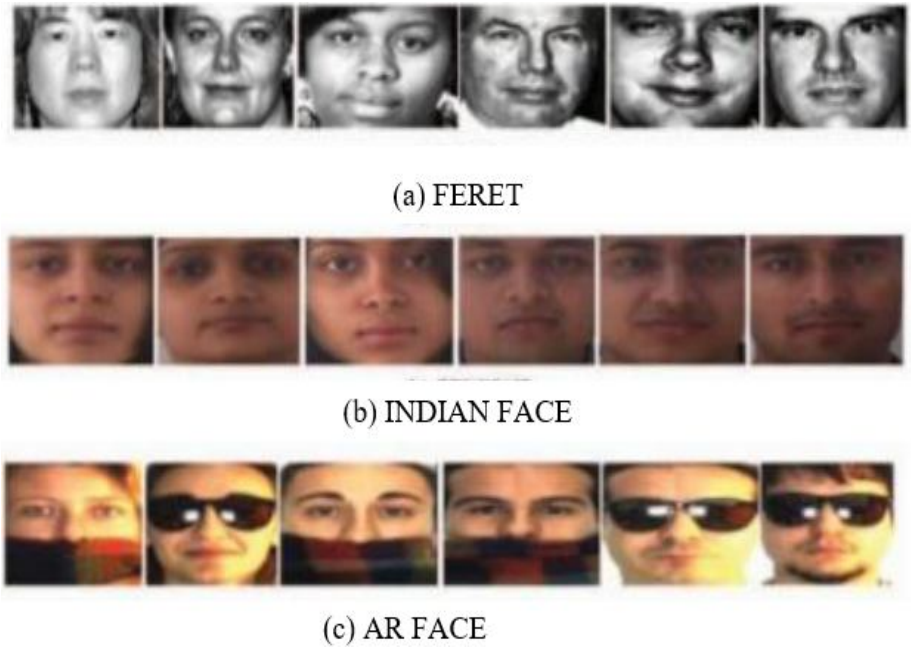


Fig. 3 Standard face image databases

A. Performance Measures

The proposed work used 2-fold cross-validation in which the overall dataset is split into two group sets, let say (D1) and (D2) of equal size. The proposed system is trained with a group set (D1) and testing has been done on group set (D2). As both the training and testing group sets are large in size, the idea of 2 cross-validations helps the testers to get an accurate conclusion in short span of time. During the training phase, the proposed system uses 50% of total images and the rest of the images used in the testing phase and each sample image is used for both training and testing during each fold. With two-fold cross-validation, computation time gets reduced, which is an important necessity in real systems. The purpose of the proposed system is to do correct classification the face images as a male/female with their emotions as happy/sad.

The system is developed with the aim of real-life deployment and hence the criterion suits the purpose, i.e., Classification rate (CR %), which is to be considered to evaluate system performance. Classification rate concludes the accuracy of the system by showing a number of images which is correctly classified. The classification rate gives result in the percentage of correctly classified face images from a total number of face images in the testing set.CR% of a system is defined by

$$CR\% = M1/M2 * 100$$

Where M1 denotes the number of correctly classified face images and M2 is the total number of face images in the testing set.

B. Experimental Results

In this proposed system we have two classifications results i.e. gender classification and emotion detection. With the captured face image, different results are obtained with a number of publicly available databases as there are variations in expressions and poses of the face images. FERET is used by most of the researchers due to good quality images. Also, INDIAN FACE database contains images of a bright homogeneous background. AR FACE database contains occluded face images i.e. person images in this database wears scarf and sunglasses and it is used for cross-database validation.

1) *Testing the performance for normal faces images:* Considering the classification of gender, the proposed system has achieved 94.98 % of classification rate (CR) with the FERET database. With the Indian face database, the system has obtained 90% of the classification rate. Whereas for the emotion classification the proposed system has achieves79.33 % of classification rate (CR) with the FERET database. With the Indian face database, the system has achieved 80 % of the classification rate.

TABLE I
CLASSIFICATION RATE OF GENDER AND EMOTION CLASSIFICATION

Dataset	Classification Rate (CR %)	
	Gender classification	Emotional Classification
FERET	94.98	79.33
INDIAN FACE	90	80

2) *Testing the performance for cross-database*: Occlusion on face images usually occurs when a person wears sunglasses, suffers an injury on faces, covers his/her face with scarf or hand, or puts a mole on his/her face. To get cross-database performance, the system uses non-occluded face images for training (FERET and INDIAN FACE, in our case), but testing is done for gender classification on occluded images (i.e. AR Face database, in our case). Moreover, the system is independent of person, i.e. face images used for training and testing are of a different set of people. With the FERET database, the proposed work gives accuracy up to 58.52% and with INDIAN FACE database, the system gives an accuracy rate of up to 52.68%. TABLE 2 indicates the cross-database performance with standard databases.

TABLE II
CROSS DATABASE CLASSIFICATION RATE FOR NON-OCCLUDED IMAGES

Dataset	Classification Rate (CR%)
Training/ Testing	AR Face Dataset
FERET	58.52
INDIAN FACE	52.68

©IJRASET: All Rights are Reserved

V. IMPLEMENTATION

A. Working Screenshots

1) Emotion Detected: Happy.

Action – None.

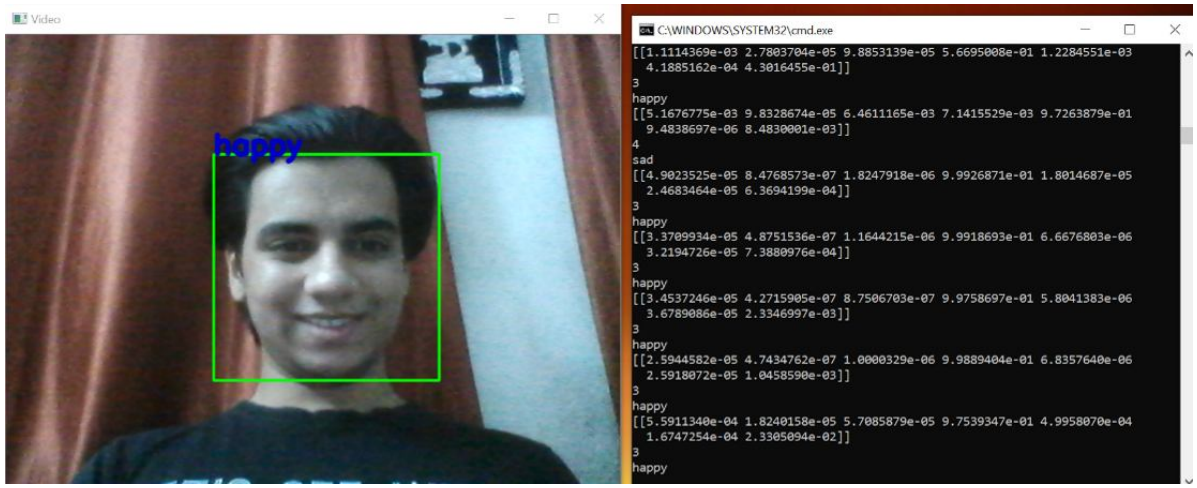


Fig 4: Screenshot of Happy facial emotion

2) Emotion Detected: Fear.

Action- E-mail Sent.

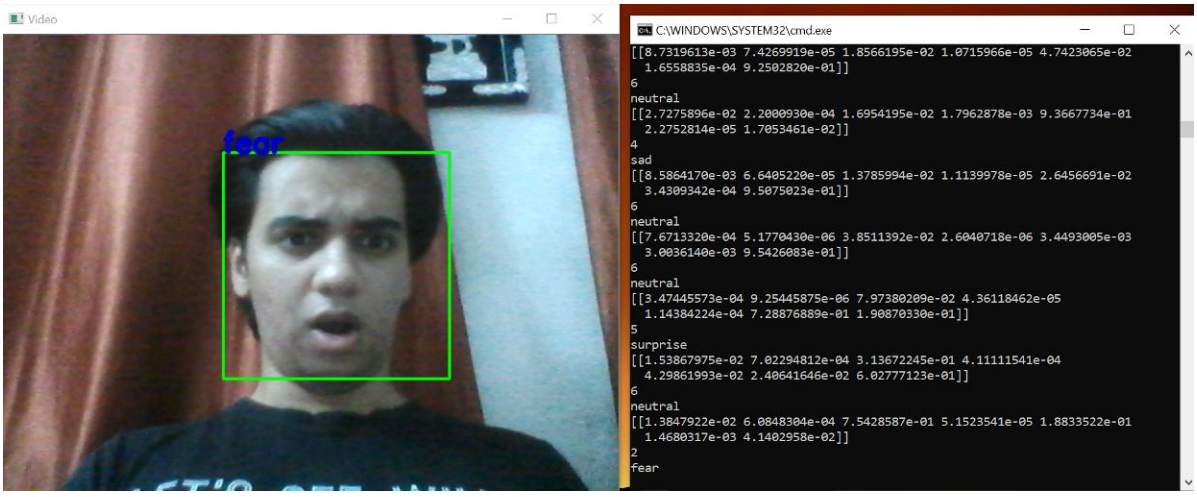


Fig 5: Screenshot of Fear facial expression

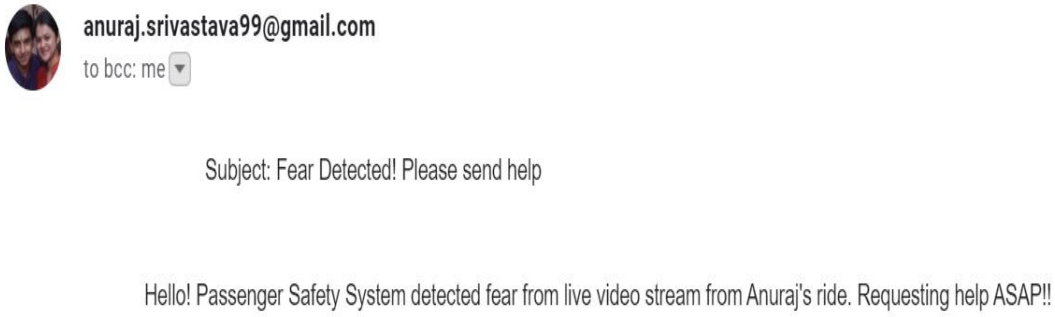


Fig 6: E-mail Screenshot

VI. CONCLUSIONS

Emotion recognition technology has come a long way in the last twenty years. Today, machines are able to automatically verify identity information for secure transactions, for surveillance and security tasks, and for access control to buildings etc. These applications usually work in controlled environments and recognition algorithms can take advantage of the environmental constraints to obtain high recognition accuracy. However, next generation face recognition systems are going to have widespread application in smart environments -- where computers and machines are more like helpful assistants.

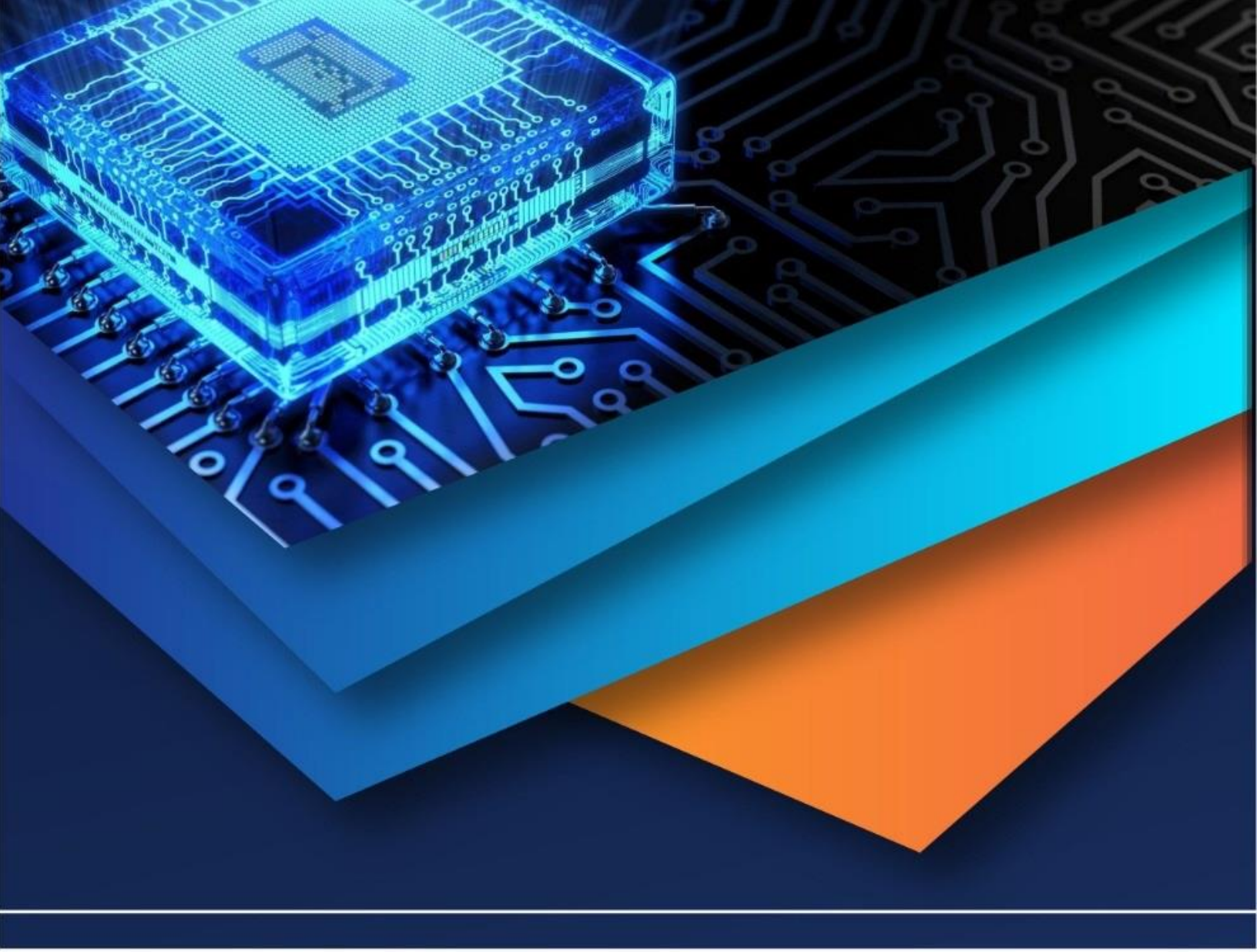
To achieve these goal computers must be able to reliably identify nearby people in a manner that fits naturally within the pattern of normal human interactions. They must not require special interactions and must conform to human intuitions about when recognition is likely. This implies that future smart environments should use the same modalities as humans, and have approximately the same limitations. These goals now appear in reach -- however, substantial research remains to be done in making person recognition technology work reliably, in widely varying conditions using information from single or multiple modalities.

And hence, putting these goals to a better application and ensuring the security of passengers can create a better business scope of Taxi companies, building trust between passengers and drivers.

REFERENCES

[1] Adelson, E. H., and Bergen, J. R. (1986) The Extraction of Spatio- Temporal Energy in • Human and Machine Vision, Proceedings of Workshop on Motion: Representation and • Analysis (pp. 151-155) Charleston, SC; May 7-9.
[2] Brunelli, R. and Poggio, T. (1993), Face Recognition: Features versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):1042-1052.
[3] Ethem Alpaydin. Introduction to Machine Learning, second edition.2010.
[4] Christopher Bishop. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., 2006.
[5] Baron, R. J. (1981). Mechanisms of human facial recognition. International Journal of Man-Machine Studies, 15:137-178.
[6] Charles Darwin. The expression of the emotions in man and animals. AMS Press,1972
[7] Wallace V. Friesen Paul Ekman. Facial action coding system. A Human Face, 2002.
[8] Bichsel, M. (1991). Strategies of Robust Objects Recognition for Automatic Identification of Human Faces. PhD thesis, Eidgenossischen Technischen Hochschule, Zurich.
[9] Jang., J., Sun, C., and Mizutani, E. (1997) Neuro-Fuzzy and Soft Computing. Prentice Hall

©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)