iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 Issue: VI Month of publication: June 2020

DOI: http://doi.org/10.22214/ijraset.2020.6239

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VI June 2020- Available at www.ijraset.com

Smart Home Security using Facial Recognition and Unusual Event Detection

Akshay H A[1], Lakshman Sharma K P[2], Menan Deep[3], Nuthan G S[4], Asha P. N[5]
[1, 2, 3, 4, 5]Dept. of Information Science and Engineering, Sapthagiri College of Engineering Karnataka, India

Abstract: Privacy and Security are two universal rights and to ensure that in our daily life we are secure, and lot of research is going on in the field of home security where IOT is the turning point for the industry. At home, security of door lock is one amongst the important aspects to be looked upon as many times the door lock fails to supply the specified security. Facial recognition is a well-established process in which the face is detected and identified out of the image. In our proof of concept of a smart door a live HD camera on the front side of setup attached to a display monitor connected to camera to show who is standing in front of the door, also the whole system will be able to give voice outputs by processing text them on Raspberry Pi ARM processor used and shows the answers as output on the screen. The algorithm used for face popularity is Local Binary Pattern Histogram (LBPH)

Keywords: Local Binary Pattern Histogram, internet of things, Raspberry Pi

I. INTRODUCTION

In past few decades, the significant efforts in the field of tracking and moving object detection have been done to make the following applications robust, reliable and efficient. They are

A. Video Surveillance,
B. Authentication System,
C. Media Production and Biological Metric System, etc.

Unusual activities and behavior detection in crowded areas is challenging as identifying the identity of the suspect and his motive is a critical task. Face detection is more challenging because of some unstable characteristics, for example, glasses and beard will impact the detecting effectiveness. Moreover, different kinds and angles of lighting will make detecting face generate uneven brightness on the face, which will have an influence on the detection process. An intensive study of OpenCV platform and its in-built libraries has been conducted to generate a code, which does correct and reliable facial recognition with new and efficient use of hardware. This proposed system also acts as a home security system for both Person detection and provide security for door access control by using facial recognition for the home environment. The human body is identified as an intruder within a home environment achieved by capturing live video from web camera and processing will be done on captured video frames. The web camera to capture the series of images as soon as the person press switch. The advantage of this system is for accessing the door is that face detection and recognition are performed by using face detection technique and the entire face recognition is completed by pressing single and tiny push button switch. But the challenges such as dynamic background, illumination change, shadow, Camouflage and occlusion can produce hurdles in the improvement of these applications. In the conventional target tracking approaches the high resolution (HR) videos are used to extract the exact contour and shape features of target. These approaches work on high resolution frames and require more computational cost.

II. LITERATURE SURVEY

Computer scientists have made ceaseless efforts to replicate cognitive video understanding abilities of human brains onto autonomous video surveillance systems. As video surveillance cameras become ubiquitous, there is a surge in studies on unusual event detection in surveillance videos. Nevertheless, video content analysis in public scenes remained a formidable challenge due to intrinsic difficulties such as severe inter-object occlusion in crowded scene and poor quality of recorded surveillance footage. Moreover, it is nontrivial to achieve robust detection of unusual events, which are rare, ambiguous, and easily confused with noise. A door access system managed employing a webpage is proposed in [4]. Principal element Analysis (PCA) is employed for face recognition and uses Raspberry Pi and ZigBee to implement Wireless sensor Networks (WSN) for the communication between the Raspberry Pi and alternative modules.

The system identifies the guests and notifies it to the owner via SMS/ email. The owner will management the access through the webpage or portable. The proposed framework uses a combination of HAAR cascade and Locally Binary Patterns Histogram (LBPH) for feature extraction and the Support Vector Machine (SVM) algorithm for face classification. A large dataset of a crowded office environment in Middle East is collected and used for evaluation of the proposed model. The experimental results show that the proposed method has acceptable results for face detection in complex surveillance scenarios.

In the technical paper “A Novel Algorithm to Predict and Detect Suspicious Behaviors of People at Public Areas for Surveillance Cameras” [2] proposes an algorithm for identifying unusual events more accurately. Suspicious activities seriously endanger at public areas and personal security. There are millions of video surveillance systems used in public areas, such as streets, prisons, holy sites, airports, and supermarkets. It is essential to investigate the detection and recognition of suspicious activities contents from surveillance video. The common suspicious activities at public areas with an aspect of security are fighting, running, leave luggage and run, put an unusual packet in somewhere like a dustbin and leave. The paper focuses on the recognition of suspicious activity and aim to find a method that can automatically detect suspicious activity using computer vision methods. Complex background, illumination changes and different distances between the human and the camera have made this topic very challenging, especially in the case of real-time applications. The GMM method is adopted to produce candidate regions whose has suspicious activity of motion features extracted from the magnitude information of Optical Flow, and this method is known as Suspicious Activity Region Detector (SARD). Experimental results on several benchmark datasets have demonstrated the robustness of the proposed framework over the state-of-the-arts in terms of both detection accuracy and processing speed, even in crowded scenes.

A. *Real-time Unusual Events Detection Using Video Surveillance System for Enhancing Security*

The proposed approach is used for detecting unusual events such as overcrowding, fight for the low-resolution video particularly in an ATM.

1) Limitations

a) Proposes an approach that is confines only to close spaces.

b) We aim to build a system which is capable of functioning in a large area.

B. *A Novel Algorithm to Predict and Detect Suspicious Behaviors of people at Public Areas for Surveillance Cameras*

This paper provides an approach to detect unusual events based on suspicious activity region detector (SARD). It works in 2 steps. First, it searches for a region of unusual event and next distinguishes between violent and non-suspicious events.

1) Limitations

a) This paper uses Gaussian Mixture Model (GMM) which gives rise to noisy images.

b) This system also had drawbacks such as false alarms.

Face recognition has been the best choice for the problem of biometrics and it has a various type of applications in the society. An efficient face recognition system can be of great help in forensic sciences, identification for law enforcement, authentication for banking and security system, and giving preferential access to authorized users i.e. access control for secured areas etc. A real-time door lock access system by face recognition system. The algorithm used here is Local Binary Patterns Histograms (LBPH), based on HAAR Feature-based Cascade Classifiers is presented in the project. These techniques are used for face detection. The HAAR Feature-based cascade classifier also efficiently handles situations where the face of a person is being covered or not visible clearly.

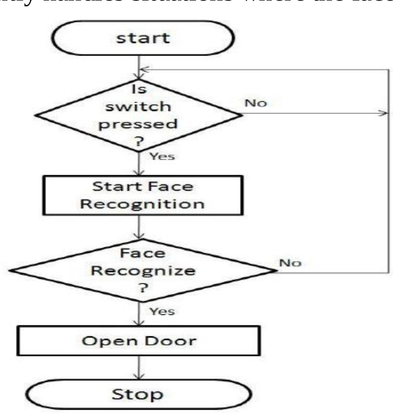


Figure 1: Data flow diagram of Door Lock Access by using Face Recognition

Figure shows the procedural view person detection process. In this system first in authentication module is having small switch button when the person presses the switch button then the implemented code is the call. Then by using the camera the system captures the image sequence and compare these images with the database. If the image is a match then in application module get the signal to unlock the door After unlocking the door, the SMTP is to send alert to authorized personnel.

III. FACE RECOGNITION MODULE

Here we use different software in order to maximize the efficiency of the module. These are explained detailed given below.

1) *Python IDE:* This has high level data structures and effective approach toogject oriented programming.
2) *OpenCV:* This is a library of programming functions mainly amed at reak-time computer vision. Here we are using image processing module that includesseveral static libraries,etc. We include libraries such as Viola-Jones or HAAR classifier, LBPH (Lower Binary Pattern histogram) face recognizer, Histogram of oriented gradients (HOG)

Image Processing module

Purpose of image processing

The purpose of image processing is divided into 5 groups. They are:

a) *Visualization*- Observe the objects that are not visible.
b) *Image Sharpening and Restoration*- To create a better image.
c) *Image Retrieval*- Seek for the image of interest.
d) *Measurement of Pattern*– Measures various objects in an image.
e) *Image Recognition*– Distinguish the objects in an image.

A. HAAR Classifier

This object detection framework is to provide competitive object detection rates in real-time like detection of faces in an image. A human can do this easily, but a computer needs precise instructions and constraints. To make the task more manageable, Viola–Jones requires full view frontal upright faces. Thus, in order to be detected, the entire face must point towards the camera and should not be tilted to either side. While it seems, these constraints could diminish the algorithm's utility somewhat, because the detection step is most often followed by a recognition step, in practice these limits on pose are quite acceptable the characteristics of Viola–Jones algorithm which make it a good detection algorithm are:

1) *Robust* – Very high detection rate (true-positive rate) & very low false-positive rate always.
2) *Real Time* – For practical applications at least 2 frames per second must be processed.
3) *Face detection only (not recognition)* - The goal is to distinguish faces from non-faces (detection is the first step in the recognition process).

All human faces share some similar properties. These regularities may be matched using HAAR Features.

a) *A Few Properties Common To Human Faces*
   i) The eye region is darker than the upper-cheeks.
   ii) The nose bridge region is brighter than the eyes.
b) *Composition Of Properties Forming Matching Facial Features*
   i) *Location and Size:* eyes, mouth, bridge of nose
   ii) *Value:* oriented gradients of pixel intensities

B. Histogram of oriented gradients (HOG)

Histogram of oriented gradients (HOG) is a feature descriptor used to detect objects in computer vision and image processing. The HOG descriptor technique counts occurrences of gradient orientation in localized portions of an image - detection window, or region of interest (ROI).

Implementation of the HOG descriptor algorithm is as follows:

1) Divide the image into small connected regions called cells, and for each cell compute a histogram of gradient directions or edge orientations for the pixels within the cell.
2) Discretize each cell into angular bins according to the gradient orientation.
3) Each cell's pixel contributes weighted gradient to its corresponding angular bin.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VI June 2020- Available at www.ijraset.com

4) Groups of adjacent cells are considered as spatial regions called blocks. The grouping of cells into a block is the basis for grouping and normalization of histograms.
5) Normalized group of histograms represents the block histogram. The set of these block histograms represents the descriptor.

IV. UNUSUAL EVENT DETECTION MODULE

In general, any object tracking system includes four main building blocks to automate the surveillance system

1) Weapons Detection: We are using CCN Algorithm
2) More Than Two Person Detection: We are using CCN and HAAR cascade Algorithm
3) Helmet and Mask Detection: We are using CCN Algorithm
4) Object identification

Implementing the CNN: The computer vision module uses a CNN to scan the incoming camera frames for firearms. The implementation of the CNN is based on Google's open source software library for machine learning, called Tensor flow. Tensor flow allows the creation of graphs, where each graph represents a mathematical model, to facilitate numerical computations necessary for machine learning algorithms. The graphs consist of nodes, each one representing a mathematical operation. Multidimensional arrays, called tensors, are used as graph edges to communicate the data between the different nodes. The CNN is defined, therefore, as a collection of nodes, where a tensor is given as input and another tensor returned as output from the last node of the graph

The network architecture loosely follows the architecture proposed by the TF Learn's image classification example, which uses the CIFAR-10 dataset. As a result, the input tensor was a 32x32 (width x height) RGB image, or 3 channels, giving a total of 32x32x3 = 3072 nodes. Initially, the dimensions were kept as 32x32 and were later adjusted to 80x45. In order to extract the features that we are trying to detect from the input tensor, we will use a combination of convolution and max-pooling steps:

(1) - 32 filter convolution => (2) – 2 stride max-pooling => (3) - 64 filter convolution => (4) – 64 filter convolutions => (5) – 2 stride max-pooling Every convolution layer applies a number of filters (either 32 or 64), each of which is a square window with a 3x3 pixels size. For instance, the first convolution layer takes as input the incoming tensor and summarizes it with 32 filters of 3x3 smaller neural networks. Every maxpooling step is a reduction operation, where each rectangular window of size 2x2 pixel is represented by its maximum value. As a result, each max-pooling layer reduces the dimensions to a quarter. The complete network architecture is shown in the figure below:

By applying this routine, the incoming nodes can be summarized using fewer resources. Each convolution splits the images into tiles and attempts to summarize them with a small neural network. Each max-pooling step performs an efficient down-sampling by keeping the most interesting bits (with respect to graph weights).

After the network architecture has been defined, with each training step a regression layer is applied and the weights of the model refined. With every training step, the weights are either increased or decreased based on the network's performance. The amount of change is defined as the learning rate of the network and set at 0.001. With a larger learning rate, the network is trained 71 more quickly but may result in overfitting. In to avoid over-fitting, a percentage of the data is dropped during each training step. Each node of the network has an equal probability of being dropped, which is defined as the dropout rate of the network. The dropout rate is initially defined as 0.50 and later adjusted to 0.15.

The output tensor from the last max-pooling layer is summarized by a fully-connected layer of 512 nodes. These 512 nodes hold the summarized features from the incoming camera frame. The SoftMax, or normalized exponential, activation function is used to generate a binary classification output from the 512 nodes. The complete network architecture is shown in the figure below:

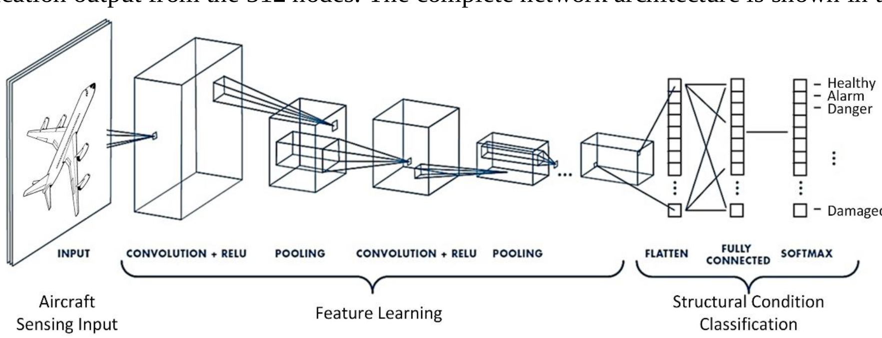


Figure 2: Neural Network Architecture

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VI June 2020- Available at www.ijraset.com

Convolutional Neural Network (CNN) — Deep Learning

In neural networks, Convolutional neural network (ConvNets or CNNs) is one of the main categories to do images recognition, images classifications. Objects detections, recognition faces etc., are some of the areas where CNNs are widely used.

CNN image classifications take an input image, process it and classify it under certain categories (E.g., Dog, Cat, Tiger, Lion). Computers sees an input image as array of pixels and it depends on the image resolution. Based on the image resolution, it will see h x w x d (h = Height, w = Width, d = Dimension). E.g., An image of 6 x 6 x 3 array of matrix of RGB (3 refers to RGB values) and an image of 4 x 4 x 1 array of matrix of grayscale image.

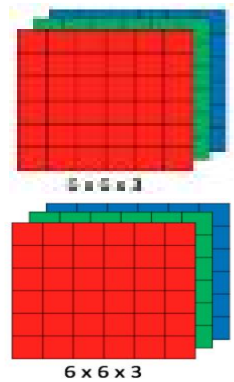


Figure 3: Array of RGB Matrix

Technically, deep learning CNN models to train and test, each input image will pass it through a series of convolution layers with filters (Kernels), Pooling, fully connected layers and apply SoftMax function to classify an object with probabilistic values between 0 and 1.

V. HAAR CASCADE FACE LANDMARKS

Object Detection using HAAR feature-based cascade classifiers is an effective object detection method proposed by Paul Viola and Michael Jones in their paper, “Rapid Object Detection using a Boosted Cascade of Simple Features” in 2001. It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images.

Here we will work with face detection. Initially, the algorithm needs a lot of positive images (images of faces) and negative images (images without faces) to train the classifier. Then we need to extract features from it. For this, HAAR features shown in below image are used. They are just like our convolutional kernel. Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle.

Now all possible sizes and locations of each kernel is used to calculate plenty of features. (Just imagine how much computation it needs? Even a 24x24 window results over 160000 features). For each feature calculation, we need to find sum of pixels under white and black rectangles. To solve this, they introduced the integral images. It simplifies calculation of sum of pixels, how large may be the number of pixels, to an operation involving just four pixels. Nice, isn’t it? It makes things super-fast.

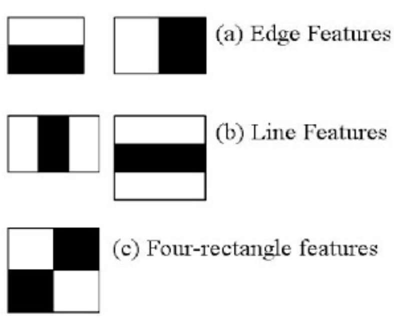


Figure 4: Extraction of features

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VI June 2020- Available at www.ijraset.com

But among all these features we calculated, most of them are irrelevant. For example, consider the image below. Top row shows two good features. The first feature selected seems to focus on the property that the region of the eyes is often darker than the region of the nose and cheeks. The second feature selected relies on the property that the eyes are darker than the bridge of the nose.
But the same windows applying on cheeks or any other place is irrelevant. So how do we select the best features out of 160000+ features? It is achieved by **Adaboost**.

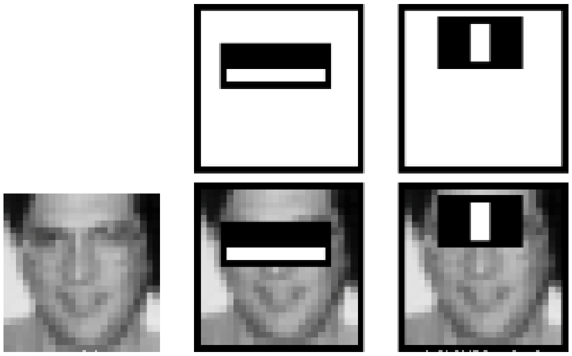


Figure 4: Eye feature extraction

For this, we apply each and every feature on all the training images. For each feature, it finds the best threshold which will classify the faces to positive and negative. But obviously, there will be errors or misclassifications. We select the features with minimum error rate, which means they are the features that best classifies the face and non-face images. (The process is not as simple as this. Each image is given an equal weight in the beginning. After each classification, weights of misclassified images are increased. Then again same process is done. New error rates are calculated. Also new weights. The process is continued until required accuracy or error rate is achieved or required number of features are found).
Final classifier is a weighted sum of these weak classifiers. It is called weak because it alone can't classify the image, but together with others forms a strong classifier. The paper says even 200 features provide detection with 95% accuracy.
HAAR-cascade Detection in OpenCV

Here we will deal with detection. OpenCV already contains many pre-trained classifiers for face, eyes, smile etc. Let's create face and eye detector with OpenCV.

First, we need to load the required XML classifiers. Then load our input image (or video) in grayscale mode.

```
import numpy asnp
import cv2
face_cascade=cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
eye_cascade=cv2.CascadeClassifier('haarcascade_eye.xml')
img=cv2.imread('sachin.jpg')
gray=cv2.cvtColor(img,cv2.COLOR_BGR2GRAY)
```

Now we find the faces in the image. If faces are found, it returns the positions of detected faces as Rect(x,y,w,h). Once we get these locations, we can create a ROI for the face and apply eye detection on this ROI (since eyes are always on the face !!!).

```
faces=face_cascade.detectMultiScale(gray,1.3,5)
for(x,y,w,h) infaces:
img=cv2.rectangle(img,(x,y),(x+w,y+h),(255,0,0),2)
roi_gray=gray[y:y+h,x:x+w]
roi_color=img[y:y+h,x:x+w]
eyes=eye_cascade.detectMultiScale(roi_gray)
for(ex,ey,ew,eh) ineyes:
cv2.rectangle(roi_color,(ex,ey),(ex+ew,ey+eh),(0,255,0),2)
cv2.imshow('img',img)
cv2.waitKey(0)
cv2.destroyAllWindows()
```

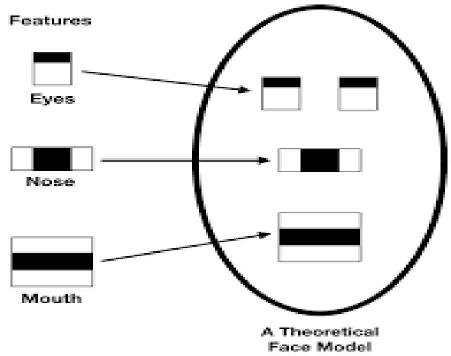


Figure 5: Theoretical Face Model

VI. CONCLUSION

The motive of this project is Unusual Event Detection and Door Automation. Here, the suspicious or dangerous activities are expected to be detected and face recognition for Door Automation System. Image Processing is the domain chosen for this project as it is one of the most rapidly growing technologies in the recent times.

In the Literature Survey, many papers have been referred related to the Unusual Event Detection and Door Automation System using Image Processing. These papers discuss algorithms for scenarios such as suspicious activity detection in public areas, unusual events in an ATM and Automated Vehicular Surveillance systems which encompasses concepts that are related to this project. The algorithms and technologies have been analyzed to select the appropriate technologies that perform comparatively better with regard to this project. The technologies identified are Rolling Average Background Subtraction along with Morphological operations for Unusual Event Detection and Local Binary Patterns Histograms (LBPH), based on HAAR Feature-based Cascade Classifiers for face detection. The Rolling Background Subtraction for Unusual Event Detection is used as the objects that are detected can be dynamic and the detection can be done using a low-resolution camera. The HAAR Feature-based cascade classifier that is used for face detection efficiently handles situations where the face of a person is being covered or not visible clearly. It has higher performance and greater accuracy compared to algorithms such as PCA or eigenfaces.

REFERENCES

[1] S Rashmi, K.R Vinothini, “Real Time Unusuall Event Detection Using Video Surveillance System for Enhancing Security”, Online International Conference on Green Engineering and Technologies, 2015.

[2] Abouzar Ghasemi, C.N, Ravi Kumar, “A NOVEL ALGORITHM TO PREDICT AND DETECT SUSPICIOUS BEHAVIORS OF PEOPLE AT PUBLIC AREAS FOR SURVEILLAVE CAMERAS”, International Conference on Intelligent Sustainable Systems, 2017.

[3] Vikas Tripathi, Durga Prasad Gangodkar, Vivek Latta, and Ankush Mittal, “Robust Abnormal Event Recognition via Motion and Shape Analysis at ATM Installations”, Journal of Electrical and Computer Engineering, Volume 2015.

[4] Saleem Ulla Shariff, MaheboobHussain, Mohammed Farhaan Shariff, “Smart unusual event detection using low resolution camera for enhanced security”, 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 17-18 March 2017.

[5] Sharayu Sadashiv Phule, Sharad D. Sawant, “Abnormal activities detection for security purpose unattained bag and crowding detection by using image processing”, 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), 15-16 June 2017.

[6] Hoda Qezavati, Babak Majidi, “Partially Covered Face Detection in Presence of Headscarf for Surveillance Applications”,2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA).

[7] Zakaria Charouh, Mounir Ghogho, International University of Rabat, “Improved Background Subtraction-based Moving Vehicle Detection by Optimizing Morphological Operations using Machine Learning”,2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), Vancouver, BC, 2018.

[8] Sebastian Salazar-Colores, Eduardo Cabal-Yepez, “A Fast Image Dehazing Algorithm Using Morphological Reconstruction” IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 28, NO. 5, MAY 2019

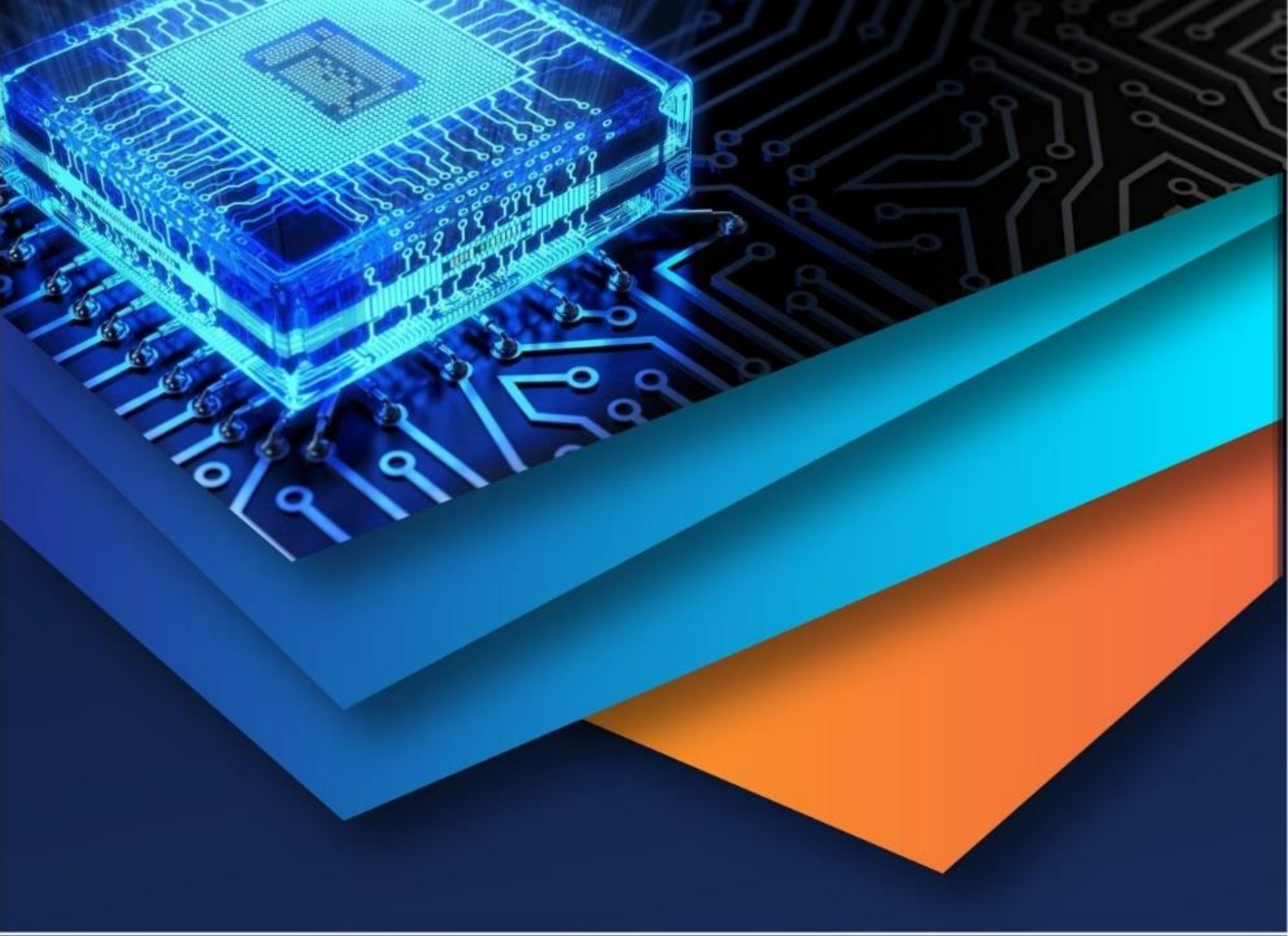
[9] Maddula J N V Sai Krishna Asrith, K Prudhvi Reddy, “Face Recognition and Weapon Detection from Very Low-Resolution Image”, 2018 International Conference on Emerging Trends and Innovations in Engineering and Technological Research (ICETIETR).

[10] Jignesh J. Patoliya, Miral M. Desai, “Face detection-based ATM security system using embedded Linux platform”, 2017 2nd International Conference for Convergence in Technology (I2CT), 7-9 April 2017.

[11] Francesco G.B. De Natale, Giulia Boato, “Detecting morphological filtering of binary images”, IEEE 2017 Transactions on Information Forensics and Security.

[12] S. Shriram, Swastik B. Shetty, Vishnu Prasad P. Hegde, KCR Nisha, Dharmambal V, “Smart ATM Surveillance System”, 2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT].

[13] Daniela D’ Auria, Fabio Persia, Fabio Betteni and Sven Helmer, “Predicting and Preventing Dangerous Events Via Video Surveillance Using a Robotic Platform”,2019



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY