

Republic of Yemen
Ministry of High Education
& Scientific Research
Al-Rayan University
Faculty of Higher Studies



SMART SECURITY SYSTEM FOR A BANK BASED ON EYE GESTURE RECOGNITION

Thesis submitted to Al-Rayan University to complete the requirements of
obtaining a Master's degree in Information Technology

By

Raed Awadh Bakunah

Supervisor

Dr. Saeed Mohammed Baneamoon

1441 / 2020 هـ

Republic of Yemen
Ministry of High Education
&Scientific Research
Al-Rayan University
Faculty of Higher Studies



SMART SECURITY SYSTEM FOR A BANK BASED ON EYE GESTURE RECOGNITION

**Thesis submitted to Al-Rayan University to complete the requirements of
obtaining a Master's degree in Information Technology**

By

Raed Awadh Bakunah

Supervisor

Dr. Saeed Mohammed Baneamoon

2020 م / 1441 هـ

Approval of the Scientific Supervisor

I certify that the master's dissertation titled, (**SMART SECURITY SYSTEM FOR A BANK BASED ON EYE GESTURE RECOGNITION**) submitted by the student **Raed Awadh Bakunah** has been completed in all its stages under my supervision and so I nominate it for discussion.

Proofreader: **Associate Professor Dr. Saeed Mohammed Baneamoon**

Signature:

Date: **30 / 09 / 2020**

Approval of the Proofreader

I certify that the master's dissertation titled, (**SMART SECURITY SYSTEM FOR A BANK BASED ON EYE GESTURE RECOGNITION**) submitted by the student Raed Awadh Bakunah has been linguistically reviewed under my supervision and has become in scientific style and clear from linguistic errors and for that I sign.

Proofreader: Abdullah Amer Al- Kathiri

Academic Title: Assistant Professor

University: Al-Rayan University

Signature: 

Date: 29 / 9 /2020

Acknowledgment

I would like to thank my supervisor Dr. Saeed Mohammed Baneamoon for his guidance, patience, willingness and assistance during the writing of my thesis. I also thank Dr. Mohammed Abdullah Bamatraf and Dr. Abdullah Aldhaibani for their advice and assistance.

Abstract

Bank safety is one of the most important safety units, which must represent a top level of security. The bank becomes the target of criminals, therefore, this research proposes a new human-security system interaction based on blink gesture recognition for smart bank security. In order to recognize the blink gestures of the employee to reduce crimes in the banks. A different processes implemented in the blink gestures recognition analyzing, these processes are face and facial landmark detection, Eye aspect ratio (EAR) and amplitude and duration of blink. In this research, the Soukupova algorithm is developed to be able to detect the voluntary blink gestures of motion of eyelids and also to design a prototype that alerts the employee that the system accepted his interaction. The using web camera and successfully implemented a prototype proposed system is tested for the gestures of the blink with two eyes successful, It is robustness against face resolution, varying illumination and facial expressions. The proposed system is evaluated by a number of experiments for security in bank from robbery which is having a detection accuracy rate more than 96% which, is efficient compared with other related approaches. Overall, the proposed system increases the efficiency of security systems and provides security to the money, customers and employees.

الخلاصة

يعتبر البنك من أهم وحدات الأمان التي يجب أن يمثل أعلى مستوى من الأمان. حيث أصبح البنك هدفاً للمجرمين ، لذلك يقترح هذا العمل تصميم نظام أمن بنكي تفاعلي ذكي من أجل الحد من جرائم السطو في البنوك، وذلك عن طريق أدراك وتمييز إيماءات الرمش الطوعية (التغميز) لجفون العين (Blink Gesture Recognition)، وذلك من خلال المرور بعدة عمليات وهي كشف الوجه (Face detection) وملامح الوجه (Facial Landmarks) ثم قياس درجة فتوحية العينين ومدة بقاء العين مفتوحة او مغلقة بواسطة حساب معامل (Eye Aspect Ratio). في هذا البحث تم تطوير لوغاريتمية سوكويفا (Soukupova) لتكون قادرة على تمييز التغميز الطوعي لجفون العين (Blink Gesture)، بالإضافة لتصميم نموذج اولي (simple Prototype) يقوم بتبئية الموظف بان النظام قد تعرف ايمائه. جرب نظام النموذج الاول بنجاح للإيماء العين باستخدام كاميرا وب، واثبت بانه قوي مع مختلف الشروط كدقة الصورة، الإضاءة وكذلك التعبيرات الوجهية و غيرها. قيم النظام المقترح بعدد من التجارب محققا معدل كشف للإيماء بنسبة اكثر من 96%، وهى نسبة فعالة مقارنة مع المنهجيات دات الصلة. بشكل عام، النظام المقترح ممكن ان يعمل بالتوازي مع أنظمة الأمان الاخرى المستخدمة في البنوك، مما يؤدي لزيادة فعالية نظم الأمان البنكيه والحفاظ على المقتنيات الثمينة داخل البنوك، بالإضافة الى جلب الأمان الى الموظفين والعملاء.

The Discussion Committee Decision

Based on the decision of the President of the University No. () in the year () regarding the nomination of the committee for discussing the master's thesis entitled, (**SMART SECURITY SYSTEM FOR A BANK BASED ON EYE GESTURE RECOGNITION**) for the researcher **Raed Awadh Bakunah**. We, the head of the discussion committee and its members, acknowledge that we have seen the aforementioned scientific thesis and we have discussed the student in its contents and what related to it.

Chairman of the Committee


Associate Professor Dr. Saeed Mohammed Baneamoon

Signature: 

Committee member

Associate Professor

Dr. Naziha Mohammed Al-Aidroos

Signature: 

Committee member

Assistant Professor

Dr. Mohammed Abdullah Bamatraf

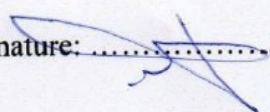
Signature: 

Table of Contents

Acknowledgment	5
Abstract.....	6
الخلاصة	7
Table of Contents	9
List of Tables	12
List of Figures	13
List of Abbreviations.....	15
CHAPTER ONE: INTRODUCTION.....	16
1.1 Introduction.....	2
1.2 Motivation.....	2
1.3 Problem Statement	3
1.4 Objectives	3
1.5 Scope and Limitations	3
1.6 Research Approach.....	4
1.6.1 Problem Identification.....	4
1.6.2 Analysis of Current Techniques	4
1.6.3 Proposed Method	5
1.6.4 Procedure and Experiments	5
1.6.5 Results and Evaluation	5
1.7 Research Contributions.....	5
1.8 Structure of Thesis.....	6

CHAPTER TWO: BACKGROUND	8
2.1 Introduction.....	9
2.2 Computer Vision and Image processing.....	9
2.3 Vision-based Human Computer Interaction	10
2.4 The Gesture.....	11
2.4.1 Vision based hand gesture recognition.....	12
2.4.2 Vision based eye and eyelid blink gesture recognition.....	12
2.4.3 Vision based Facial expressions recognition.....	14
2.4.4 Vision based multimodal gestures recognition.....	16
2.5 Blinking (eyelids motion)	17
2.5.1 Types of Blinking	18
2.5.2 Vision-based Blink Detection.....	18
2.5.3 Blink detection methods.....	18
2.6 Chapter Summary.....	19
CHAPTER THREE: LITERATURE REVEIW.....	20
3.1 Introduction.....	21
3.2 Related Works.....	21
3.3 Critical Evaluation of Existing Approaches	25
3.4 Remarks	27
3.5 Chapter Summary.....	28
CHAPTER FOUR: PROPOSED METHOD & IMPLEMENTATION	29
4.1 Introduction.....	30
4.2 Proposed Method.....	30
4.2.1 Facial Landmark Technique	30

4.2.2	Eye Aspect Ratio (EAR) Technique	31
4.2.3	Duration and Amplitude Techniques	36
4.2.4	Proposed Method Methodology.....	37
4.3	Component and Tools.....	38
4.3.1	Laptop and Cameras.....	38
4.3.2	Python with Open CV and Dlib Libraries	39
4.4	Implementation	40
4.4.1	Procedure.....	41
4.4.2	Experimental Methodology	41
4.4.3	Experiments Details	45
4.5	Results and Performance of the system.....	54
4.5.1	Real-time capabilities.....	55
4.5.2	Recognition performance	55
4.6	Comparison of Results with Previous Works	57
4.7	Remarks	58
4.8	Chapter Summary.....	58
	CHAPTER FIVE: CONCLUSION AND FUTURE WORK.....	59
5.1	Conclusion	60
5.2	Limitation of The Proposed Approaches	60
5.3	Future Work.....	60
	References.....	62
	List of Publications.....	67

List of Tables

Table No.	Title	Page
3.1	Critical evaluation of Relevant Approaches	25
4.1	The mean values of ear for complete closed, normal open max openness of the eyes	45
4.2	Processing time of frame and mean frame rate for detection without detection and baseline corresponds to displaying the video without any processing by same camera.....	46
4.3	Gesture processing time for number of frames.....	46
4.4	Frame rate and processing time with / without face and landmarks detection and also the baseline corresponds to displaying the video on the screen without any processing	55
4.5	Possible decisions of gesture output	56
4.6	Accuracy of the Amplitude gestures recognition system averaged over all four participants in low resolution.....	56
4.7	Accuracy of the Duration gestures recognition system averaged over all four participants in low resolution.	56
4.8	The distribution of the system performance and accuracy, in deferent conditions states: direction views of camera, glasses and low lights.....	56
4.9	The distribution of the system performance and accuracy for the duration gestures, in deferent conditions states: glasses, without glasses and low lights	57
4.10	The comparison of proposed system with Gupta and Kumar works.....	57

List of Figures

Figure No.	Title	Page
1.1	Research approach	4
2.1	Eye predefined gestures.....	13
2.2	Eight gaze orientation categories gazes in right side	13
2.3	The six landmark points of each eye	16
2.4	Multiple modalities human-computer interaction system	17
2.5	Natural blink	18
3.1	Critical Evaluation of Existing Approaches	27
4.1	Visual representation of 68 key facial Landmarks.....	31
4.2	Six unique coordinates of the eye	32
4.3	Open and closed eyes with landmarks <i>pi</i> automatically detected....	33
4.4	The right blink is detected but the left blink is missed.....	33
4.5	A red box demonstrates the scanning time window.....	34
4.6	The precision-recall curves.....	35
4.7	Proposed blink gesture	37
4.8	Proposed method methodology.....	38
4.9 (a)	Core 2 Duo laptop 4G RAM	38
4.9 (b)	HD 720p web camera	38
4.10	Detection of face	41
4.11	Visual representation of 68 key Facial landmark points	42
4.12	Visual representation of landmark points of each eye	42

Figure No.	Title	Page
4.13	Real signal of complete blink gesture.....	43
4.14	Experimental flowchart of the system.....	44
4.15	EAR of 4 participants for open, closed and max openness of the eyes	45
4.16 (a, b, c)	Samples of complete signal of blink gesture.....	47
4.17	Gesture time for number of frames and number of frames for complete time of gesture signal	47
4.18	Gesture's signal shape and values of the HT, LT and S LT	49
4.19	participants how are doing the amplitude and duration gestures	49
4.20	Gestures stays detected at different regular luminous.....	50
4.21 (a, b, c, d)	Lost and misinterpreted duration gesture detected in unregularly luminous intensity conditions	50
4.22 (a, b)	The detection of amplitude gesture is in stable detection until in unregularly lights	51
4.23 (a, b, c)	Face orientation: camera put it in different directions side, down and up views	52
4.24 (a, b)	Duration gesture in glasses states	52
4.25(a, b, c, d)	Amplitude gestures in glasses states	53
4.26 (a, b)	Participant put his hand on his face	53
4.27 (a, b)	Participant put his hand on his mouth.....	54
4.28	Participant put his hand on his nose	54
4.29	The comparison of proposed system with Gupta and Kumar works	58

List of Abbreviations

EAR	Eye Aspect Ratio
SMS	Short Message Service
HCI	Human Computer Interaction
SVM	Support Vector Machine Algorithm
LED	Light Emitting Diode
ATM	Automatic Teller Machine
ARM	Advanced RISC Machines processor
GSM	Global System for Mobile Communication
GPS	Global Positioning System
RFID	Radio Frequency identification
LAN	Local Area Network
OTP	One-Time Password
UIN	Unique Identify Number
CPU	Central Processing Unit
RAM	Random Access Memory
USB	Universal Serial Bus
4GB	Four Gigabyte
SCR	Silicon Controlled Rectifier

CHAPTER ONE: INTRODUCTION

1.1 Introduction

Bank is one of the most important safety units, which must represents a top level of security. The bank become the target of criminals, because it has increased number of branches, which have expensive jewellery, important documents and cash.

In the last years, there are increased numbers of robberies and crimes of banks which, occurred in cities during daylight hours when the employees were subjected to force, violence or gunpoint [1].

So there is a strong need to develop technique to enable the employee to activate the alarm system, and help to inform the local security centre or the nearest police station in order to abort the crime in timely manner, easy capture of suspects and keeps the expensive life of all employees and customers safe.

By the rapid development of computer vision, image processing and video intelligent technologies, it's now possible to integrate human gestures recognition techniques with video surveillance systems to achieve smart security system.

Therefore, in this thesis will propose a new human-security system interaction that based on blink gesture detection technique in order to recognize blinks gesture of the employee, therefore inform local security and the nearest police station to the rapid intervention, without the knowledge, feeling and attention by suspects, attackers and accomplices, by using sophisticated computer vision technologies.

1.2 Motivation

The current banking security systems have some weaknesses that related to early detection of crimes, so the money, life of employees, customers are exposed to risk. It would be useful to have better solutions to fill the gaps in these systems. Thus the People's life and money will be safe and also easily capturing of suspects in a timely manner.

1.3 Problem Statement

Most current bank security systems usually use video surveillance system, which is basically used to provide evidence after crimes, and it cannot do real-time early warning.

Standard banks policy is to avoid violence as much as possible, so they will normally hand over the money and try to obey to the robber's demands [2]. In addition the bank employees, and customers cannot ask for assistance (Especially at gunpoint or in the case of hands are tied), therefore they cannot contact the police in a timely manner, causing the escape of the suspects.

For these previous reasons, it is a necessary to solve this problem, and make a system to analyze video stream in order to inform the local security official and police.

1.4 Objectives

The objectives of this research relate to reduce crimes in the banks through design system that increase the efficiency of security systems and to provide security to the money, customers and employees. In more detail, the objectives are as follows:

1. To develop the **Soukupova and Cech** method [3], in order to detect the blink gestures of voluntary movements of eyelids, by proposing specific amplitude and duration values for gesture, and classify the gestures by compare them with measured values . To do that first, detecting the face in video stream, then extract eye landmarks from facial landmarks, after that calculating the eye aspect ratio values.
2. To design a prototype that alerts the employee that the system accepted his interaction, and also notified the local security and the nearest police station, by sending SMS messages.

1.5 Scope and Limitations

This study will focus on developing vision-based security interaction system for a bank based on eyelids' blink recognition technique that by analysing the videos stream which captured using web camera. The proposed system will let user to interaction with it,

in order to recognize user's gestures. This study was implemented in room conditions with four participants of my family. The prototype to be designed is based on web camera.

1.6 Research Approach

Figure 1.1 shows all steps of the proposed approach. In more detail, the steps are as follows:

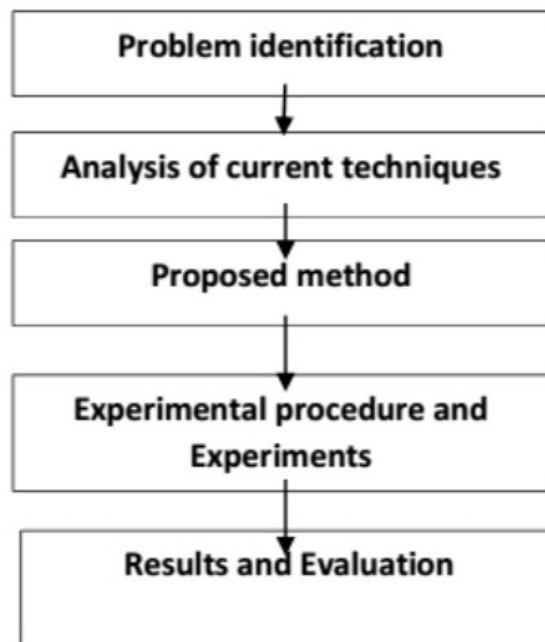


Figure (1.1) research approach.

1.6.1 Problem Identification

This step will present the problem which is the weaknesses in the banks' security systems that may lead to crimes. Also will find a gap that the researchers did not give enough solution about. Now the question is, how can we reduce this weakness and support the security systems?

1.6.2 Analysis of Current Techniques

In this step will be presented the analyzing of the most used vision-based methods and vision-based gesture techniques which are related to the human computer interaction and security systems.

1.6.3 Proposed Method

This step will investigate all techniques that are used in a proposed method in order to recognize the gesture. These techniques are facial landmark technique, eye aspect ratio technique and duration & amplitude of blink techniques.

1.6.4 Procedure and Experiments

In this step, first, an initial training, recording and collecting the captured videos for participants who perform the gestures. Then, writing the codes and implementing the experiments.

1.6.5 Results and Evaluation

In this step, first we will measure and record the results and calculate the accuracy then, evaluate the performance of the whole system and comparing them with Gupta et al. and Kumar et al. [4] studies [5].

1.7 Research Contributions

The approaches of the security systems that based on hand gesture recognition techniques face some challenges such as easy to identify and observe by suspicious persons. To solve this problem, it is proposed a new approach which based on the recognition of eyelid motion to detect blink gestures using face detection, facial landmarks detection and estimating of the eye opening state. To do that we have to conduct several experiments in order to determine if system would be able to recognize genuine gesture and reject random others. The particular method has become possible through the popularization of all cameras and would be feasible and secure in different conditions. This system can work in parallel with other security systems and can also be applied to all important sites such as museums, and other institutions. The detail's contributions of this thesis are as follows:

- 1) Proposed a new accurate approach for security system based on detection of eyelid blink voluntary gesture using face detection to localize face, facial landmarks detection to

localize eyelid and applying eye aspect ratio technique in order to estimate the eye opening state. This approach is hard to identify and to be observed by suspicious persons.

2) Developed system using duration and amplitude parameters of the blink to detect eyelid gesture as a pattern of eye aspect ratio values in temporal window of 30fps videos stream frames as an input, achieving real-time performance and accurate recognition. This system can work in parallel with other Bank security systems and can also be applied to all important sites such as museums.

3) Built a rule that generates different types of duration and amplitude gestures to interact with the system. One of these types is chosen to use with the appropriate site as other important institutions where people and officials need help, assistance and support.

1.8 Structure of Thesis

This thesis consists of five chapters which will cover the designing of vision-based gesture interaction security system. Here is an overview of the content of each presented chapter:

Chapter One: this chapter introduces the problem, gives an overview about the study and describes the needs of vision-based gesture interaction security system in the banks and other important sectors. This chapter also discusses the scope, the limitation, the significance, the motivation of the study and its objectives.

Chapter Two: this chapter covers the background of relevant information about image's components, image processing and computer vision techniques. Moreover, this chapter represents some vision-based methods which are used the gesture recognition and detection techniques.

Chapter Three: This chapter covers the literature review which is the previous related works that they been done before. Moreover, this chapter represents the used techniques of these works in order to understand the study more.

Chapter Four: This chapter explains the details of the selected methodology and techniques that we are going to use in the study. Moreover this chapter discusses all implementation steps which are Procedures, experiments' details and evaluation process.

Chapter Five: This chapter discusses the conclusion and future works to improve this study.

CHAPTER TWO: BACKGROUND

2

2.1 Introduction

Bank security systems are custom-designed of systems that based on each bank's needs, allowing the security specialists to completely secure the institution's vulnerable points. There are two types of security systems which used in banks the first one is informatics security against **cybercriminals** to ensure the transaction works safe and prevents any activity that could cause lost to the bank and its clients or customers and not disclose to anybody with no authorization only legal staff and legal bank's customer can operate any of the tasks as in Pradhan system [6]. The security of information banking systems also contains two components: computer security and network security. **Computer security** is considered as an autonomous system, which is provided by means of operating systems and applications and built-in computer hardware. **Network security** involves the protection of data at the time of transmission over communication lines and protection from unauthorized remote access to the network as in [7]. The second one is against individual criminals to ensure that the highly-priced possessions and cash in locker rooms, ATM machines and also the life of customers and employees are in safe as in Khera et al. [8], Verma et al. [9] Tarief et al. [10], Archana et al [11], Georis [2] and B. Saranraj et al. [12] works. In this thesis we focused on security systems which based on the interaction between an employee and a system as in Gupta et al [4], Kumar et al [5]works.

2.2 Computer Vision and Image processing

An image is a two-dimensional matrix of pixels. Every image consists of a set of pixels. Pixels are the raw building blocks of an image. A pixel as the "color" or the "intensity" of light that appears in a given place in image. A color model (**channel**) is a way of describing colors. These are usually represented as tuples of numbers, typically as three or four color components. An RGB model is used most often, whereby each color is encoded as a composition of a red, green and blue channel [13].

Image processing deals with extracting low-level information from images by treating them as signals. In fact, image processing is often regarded as a subfield of signal

processing and uses a large amount of the same techniques .This kind of processing can be used to detect edges or perform blurring [13] [14].

Computer Vision is a subfield of computer science that deals with techniques for acquiring, processing, analyzing, and understanding digital images in order to extract numerical and symbolic information in order to make decisions This usually involves finding more high-level information than image processing, although image processing techniques are often used before applying computer vision techniques [15] [16].

2.3 Vision-based Human Computer Interaction

Human computer Interaction (HCI) can be described as the point of communication between the human and a computer. Commonly used input devices include the following: keyboard, computer mouse, trackball, touchpad and a touch-screen. Every new device can be seen as an attempt to make the computer more intelligent and make humans able to perform more complicated communication with the computer. This has been possible due to the result oriented efforts made by computer professionals for creating successful human computer interfaces. As the complexities of human needs have turned into many fields and continues to grow, so the need for complex programming ability and intuitiveness are critical attributes of researchers to survive in a competitive environment, they have been incredibly successful in easing the communication between computers and human. This success helped to facilitate and better human control over human monitoring systems as computer commands [17] [18] [19], and activity Recognition [20] [21] [22] controlling and navigating devices [23] [24]. However, a shift towards a user friendly environment has driven them to revisit the focus this area. The abstract from previous context that is to make computer understand human language and develop a user friendly environments as human computer interaction (HCI). Finally the Human-computer interaction is a technology which lets computer and machine hear, see, and feel. At present, the field of human-computer interaction has entered the stage of multimodal interaction which make a computer understand the speech, facial expressions and human gestures as in [25]. Since the gestures are perceived through vision techniques, it is a subject of great interest for recent computer vision researchers.

2.4 The Gesture

The Many researchers had tried to define gestures but their actual meaning is still arbitrary. Gesture is defined as the motion of the body that is intended to communicate with other agents [26]. For a successful communication, a sender and a receiver must have the same set of information for a particular gesture. As per the context of the project, gesture is defined as an expressive movements of body parts, which have a particular messages, to be communicated precisely between a sender and a receiver.

A gesture is scientifically categorized into two distinctive categories: dynamic and static. A dynamic gesture is intended to change over a period of time whereas a static gesture is observed at the spurt of time. A waving hand means goodbye is an example of dynamic gesture and the stop sign is an example of static gesture [19] [27]. The possible location for origin of the human gestures or body gestures can be from any part of our body, but the most of the gesture-based studies that are carried out involve the gestures that have originated from the hand , face, head, eyes, eyelids and tongue [19] [28] [6] [23] [18] [17],because most normal people transfer the messages through this locations. To understand a full message, it is necessary to interpret all the static and dynamic gestures over a period of time, this process is called gesture recognition.

Gesture recognition is the process of recognizing and interpreting a stream continuous sequential gestures from the given set of input data. The main entire point of the gesture recognition, is to teach a machine to learn different gestures made by humans, and to interpret the same, which means to make the sensible conversations between a human and a machines possible. **Due to** most fields of gesture recognition techniques belong to the computer vision which belong to computer science, **so it is** a hot topic today and carried out a wide range of community **as a result**, modern researches of the control of computers changed from standard peripheral devices to remotely commanding computers and machines through speech, emotions and human or body gestures [25]. The following texts gives some background information about gestures systems and an overview of the methods currently used that based on gesture recognition techniques, which developed

solutions to meet the requirements of various applications that related to vision-based interaction between human and machine.

2.4.1 Vision based hand gesture recognition

Recently, the use of human movements, especially hand gesture, has become an important part of human–computer interaction (**HCI**), which serves as a motivating force for research in modeling, analyzing, and recognizing the hand gestures. For example, vision commands computer is proposed based on hand gesture recognition technique using six static and eight dynamic hand gestures. This technique use transfer learning approach which include hand shape recognition using (**VGG16**) retrained model on a large dataset and tracing of detected hand using skin color algorithm in a frame then converting the data to commands [19]. On the other hand, human-security interaction systems is proposed for security purpose and deaf people. This systems uses techniques to recognize the special hand gesture and Sign Language of the parsons. This type systems configured to processes such as image acquisition, skin color information for recognizing hand gesture which is obtained from the arm region of the hand, background removal, canny edge detection, contour detection, feature extraction, image classification, training and testing of data [4] [5].

2.4.2 Vision based eye and eyelid blink gesture recognition

The states of human eyes have proved to be an important factor in determining intend and level of attention of a person towards his surroundings, so they contain a lot of useful information. Due to eye gaze, eye blink and an eyelids movements are the main method of collecting the visual information and plays an important role in human-computer, human system interaction and detection of state of human as drowsiness, fatigue and others, so several vision-based gesture recognition studies have developed various methods, techniques and systems to meet the requirements of various applications. For example, eye gesture recognition system is introduced in order to interact with portable devices. The system combined techniques from image processing, computer vision and pattern recognition to track and analyze the sequence of eyes gaze directions (left, right, up, down and middle directions) in the video stream to recognize a predefined gesture. This system

support four number of eye predefined gestures as in Figure 2.1, and was implemented using OpenCV library [18]. Moreover an eye movement recognition method is proposed based on Convolutional Neural Network Techniques (CNN) [29]. It built dataset for eye gaze movement contains of 19200 images, the images were divided into eight gaze orientation categories (up, down, left, right, up and left, up and right, down and left, down and right) gazes as in Figure 2.2. The activities of this type of CNN methods usually consist of data collection, dataset building, training and model testing and finally recognition, in addition this method can be used in cumbersome interface devices for human computer interaction (HCI) and have better performance than SVM methods.

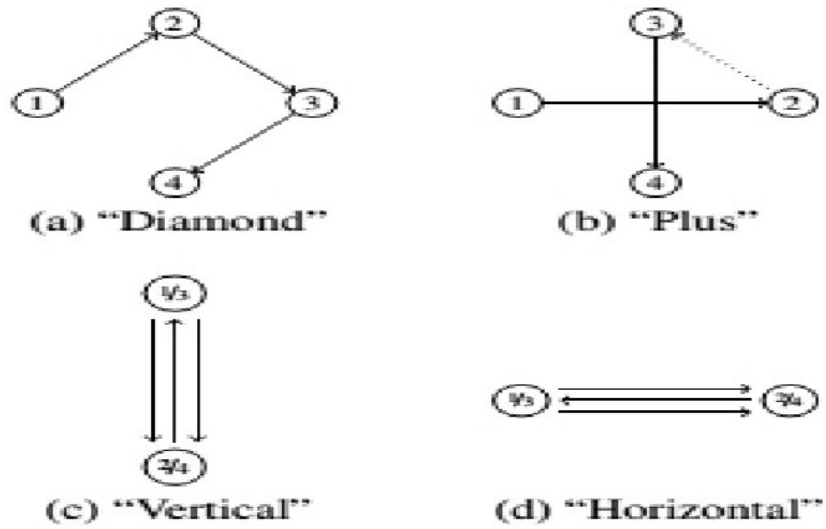


Figure (2.1): Eye predefined gestures

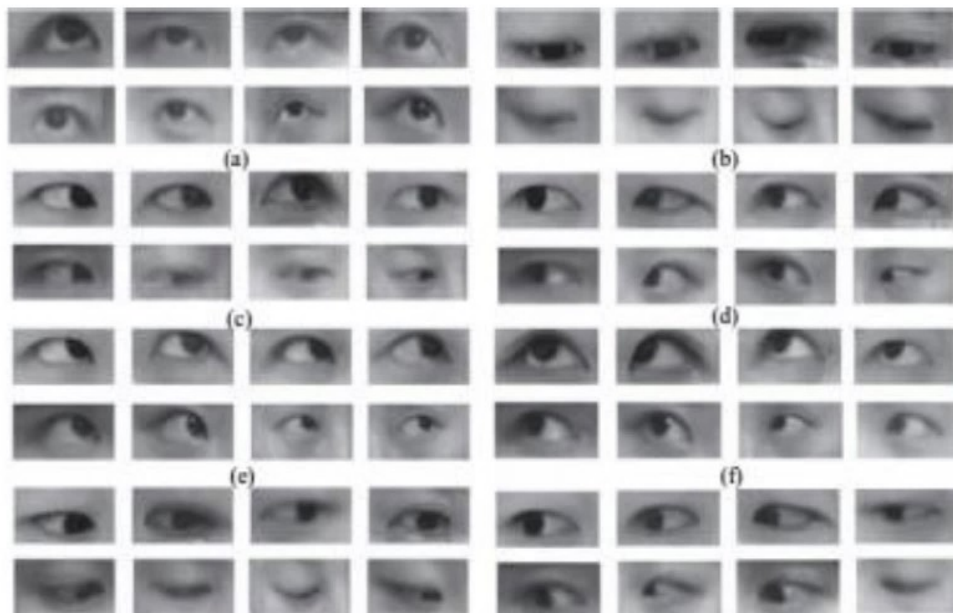


Figure (2.2): Eight gaze orientation categories gazes

But regarding to blink of eyes, interaction control commands computer vision system is designed based on blink gesture recognition technique [17]. The system used to help people with disabilities to become a part of the information society by detecting the voluntary eye blinks gestures and interpret the estimation of eye blink duration and sequence of blinks as control commands in real time. The methodology of this technique includes face detection, eye-region extraction, eye-blink detection and eye-blink classification. On the other hand, drowsy, driver monitoring and accident prevention system is presented based on monitoring the changes in the eye blink duration [30]. The system defined three states for the driver drowsiness: blink duration is less than 400ms is awake, bigger than 800ms is sleeping and between them is drowsy. Also method employed in vehicles and helped in reducing the amount of fatigue related accidents by alarming the driver [31] is presented, based on determining the eye blinking rate. The system defined 8-10 blinks per minute for normal driver's state and inform the driver if this value is decreased.

This types of methods are including face detection, eye detection, and eye blink detection and at the most they used viola-jones and neural networks algorithms to detect the face and eyes. The methods can also detect the state of the eye via a single frame without requiring additional frames. In addition most systems which are based on this methods rely on comparison between the detected blink rate and duration with predefined threshold values to detect the level of human's state as in fatigue and drowsiness states.

2.4.3 Vision based Facial expressions recognition

The face, including the eyes, eyebrows, nose, jawline, mouth and tongue plays an important role in visual communication. By looking at them, human can automatically extract many nonverbal messages, such as humans' identity, intent and emotion. In computer vision, the automatically localization, extraction and detection of those facial information are called facial expressions recognition. There are many facial analysis methods built up on accurate detection of these facial expressions. For example, proposed vision control interaction system is proposed based on Tongue expression gesture technique to help any people with suffering movement disability who face challenges in their daily

life activities. The system controls predefined devices through four movements of the tongue, namely left, right, top, and down. The system used Viola-Jones detection algorithm to detect the face and mouth and it used the analysis of Coordinate system CSA to detect and trace the movements of tongue [23]. In other studies methods to detect facial expressions based on facial landmarks gesture technique to localize the key points of eyebrows, nose, mouth, and jawline, eyes and eyelid contours. On the other hand, a technique to detect eye blinks is designed. This method estimates the facial landmark positions and extracts the vertical distance between eyelids for each video frame of an input video as in Equation (2.1) [32].

$$d = \sqrt{(p_2.x - p_1.x)^2 + (p_2.y - p_1.y)^2} \quad (2.1)$$

Where d is the vertical distance, p_1 and p_2 are landmark points.

This equation assumed that the distance between upper and lower eyelids is fixed when the eyes are open, and approaches to zero when the eyes are closing. Whereas a simple real-time algorithm to detect eye blinks in a video is designed [3] [33]. It is called **EAR SVM classifier** which is trained by 300-VW dataset. This algorithm used to estimate the eye opening state by calculating the ear value for every frame from six landmarks points of each eye as in Equation (2.2). One point in each corner of the eye (p_1 , p_4), two points on upper eyelid (p_2 , p_3) and two on lower eyelid (p_5 , p_6) as in Figure 2.3. After the landmarks detected in the image, the Eye Aspect Ratio (**EAR**) value is calculated and averaged for both eyes. The **ear** value is fixed when the eyes are open, and approaches to zero when the eyes are closing.

$$EAR = \frac{|p_2-p_6| + |p_3-p_5|}{2|p_1-p_4|} \quad (2.2)$$

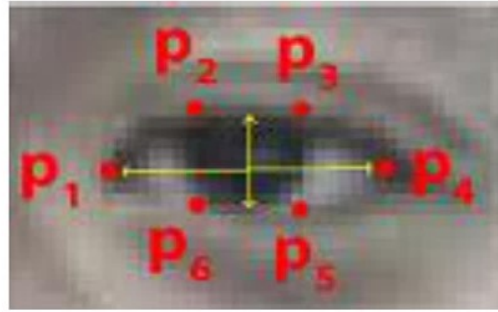


Figure (2.3) The six landmark points of each eye.

This Active Shape Models method was robust to low image quality and faces are not always frontal to the camera, the expression is not always neutral, people are often emotionally speaking or smiling or wear glasses. For these reasons this approach is better than the Viola-Jones approaches. An authenticity algorithm is designed to generate a unique signature for an individual which when stored in the user account information database will limit the creation of fake or multiple accounts [6]. The system based on Face landmarks expression gesture, Smile expression gesture and Encryption techniques. It consist of the following steps: Smile elimination, Feature extraction and Digital signature. So after generates and extracts 68 point facial landmark features, convert them to a highly compressed and secure digital signature by using some OpenCV 3.0 library classifiers.

2.4.4 Vision based multimodal gestures recognition

Multimodal gestures is merge of some gesture techniques to enable a multi-modal interaction between users and system which make system understand the speech, facial expressions and others human gestures. So by fusing information obtained from different modalities of human-scale multimedia environment, the system can better understand the user's need and discern false information. For example, a multiple modalities human-computer interaction system is introduced [25]. The system based on face, body gesture and speech recognition techniques to monitor the user facial behavior, body gesture and spatial location by using cameras. After that an extended human-computer interaction system is introduced based on head pose, eye gaze and body gestures techniques to monitor and tracking the movements' of user's head , eyes and body gesture using web camera [28]. This types of system allow users in front of monitor or big screen to interact with games,

PowerPoint presentation and more by using static pose as pointing gesture, where user stretches his arm and point at a target location in screen, and dynamic motion pattern as waving hands and changing direction of eye's gaze as in Figure 2.4. The gesture recognition of the systems include Data collection, Data preprocessing, Feature extraction and Recognition.



Figure (2.4) multiple modalities human-computer interaction system.

From previous studies, we vasisgated different methods of human computer interaction systems. Now we will highlight on detection of Eyelids movements (**blinking**) method which is our focus in this research.

2.5 Blinking (eyelids motion)

Blinking is defined as a rapid closing and reopening the eyelids of a human eye (eyelids motion). Each individual has a little bit different pattern of blinks which differs in the speed of closing and opening, a degree of squeezing the eye. The main blink parameters are duration, rate, and amplitude. Blink Duration is the period of time when the eyes were in a closed state, it lasts approximately 100-400ms it depends on people's mood. Blink Rate (blink frequency) is the number of blinks per given timeframe, and it is defined as the blink count in a minute. The natural blink rate during rest was 17 blinks/min, it increases to 26 during conversation, and it decreases to 4.5 while reading. Finally the Blink Amplitude which describes the degree of openness of eye, it differs between people and it also depends on mood and head pose [34] [33] [31]. Figure 2.5 presented the shape of normal blink.

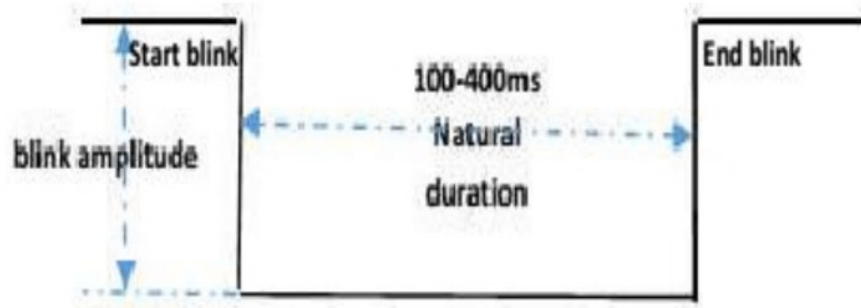


Figure (2.5) Natural blink

2.5.1 Types of Blinking

The involuntary one, which include: **(1)** Spontaneous blink which is done without any obvious external stimulus. It happens while breathing or digesting. **(2)** Reflex blink, which is caused by contact with the cornea, fast visual change of light in front of the eye, sudden presence of near object or by a loud noise. **The voluntary one**, which implements under the control of the individual [17].

2.5.2 Vision-based Blink Detection

The states of human eyes contain a lot of useful information, and the research on recognition of unit eye has become a hotspot in human activity recognition. The movements of eye and eyelids are the main method of collecting the visual information, and play an important role in Human-Computer Interaction (**HCI**). Eye and eyelid recognition not only have research significance in the field of human-computer interaction, but also have extensive application prospect and profound significance in activity recognition psychology, medicine, design, and military field [29].

2.5.3 Blink detection methods

Many of different methods have been developed for blink detection that they estimate eye state as either open or closed, or track eye closure events so the output can be either **recognition** of open or closed eye, or **detection** of blink. These techniques using only a camera. They use either single image characteristics to recognize the eyes, or they use several subsequent frames from a video sequence to detect the blink. They usually have

three stages: a face detection, an eye region detection and a blink detection or recognition. The Viola-Jones algorithm is mostly used for face detection.

The single image processing methods are based on skin color segmentation, edge detection and a parametric model fitting to find the eyelids. The traditional image processing methods for computing blinks typically involve some combination of eye localization, thresholds to find the **whites** of the eyes and determining if the white region of the eyes disappears for a period of time as in [31].

Other methods use template matching, where templates with open and closed eye images are learned and a normalized cross correlation coefficient is computed for an eye region of each image. The Template matching based methods use an eye image model that is compared with the segments of the input image. The matching part of the input image is determined by correlating the template against the current image frame [17]. Most works which based on eye blink technique, rely on the blink rate and duration parameters to determine the states of human, and human gestures.

This research rely on blink amplitude and duration in order to detect the special eye blinks gesture of the employee which, is Difficult to observe by criminals, suspects and accomplices. So that the security man of bank and nearby police station can be informed about the start of crime, just by recognizing the special eyes gesture made by the employee in front of the cameras. In this work, a vision-based system for detection of blinks gesture is presented, together with its implementation as a human-security system interaction for important reigns like banks.

2.6 Chapter Summary

This chapter introduces an idea about an image, computer vision, image processing and gestures technologies, and it also introduces global look about gesture recognition system which, based on visions techniques. Moreover the chapter highlights human interaction systems and focus on eye and eyelids blink interaction systems which, are the main of this research.

CHAPTER THREE: LITERATURE REVEIW

3.1 Introduction

This chapter presents some studies that related to security systems for banks. It presents used methods and techniques, their advantages and disadvantages and also critical analysis and comparison between these methods and techniques.

3.2 Related Works

The related work can be categorized into two types; Employee without help gesture which is not depending on the gesture of employee and employee with help gesture which is depending on gestures of employees as shown in Figure 3.1.

(**B. Georis, 2004**) presented a real time platform for semantic video interpretation applied to bank agency monitoring. The proposed system is a multi-camera platform, which recognizes user predefined scenarios, such as bank attack scenarios. These scenarios are modelled by domain experts using a back and forth process and based on a representation language. In order to address bank monitoring issues, the system has been improved based on two evaluation types. First, a repair stage guided by a careful technical evaluation has been performed at each level of the interpretation chain. As a consequence, the robustness obtained was sufficient enough to recognize all scenarios of interest. Second, an end-user evaluation has helped the experts to improve the scenario models to adapt them to real life situations. The system modelled a large set of scenarios to take into account the variety of bank robberies to prevent any drama [2].

(**Yongxiang et al, 2009**) designed a Bank Intelligent Video image processing and Monitoring Control System, which rely on the OpenCV computer vision library. The system has a strong adaptability for the varieties in the monitoring region. It is able to ensure that background images which are based on static areas refresh will be updated in time. And by reforming the temporal and spatial characteristics analysis method, it profits from the robustness against noise of background method. Thus, accurate detection of unusual actions is guaranteed. This system also fills the blank of real time alarm monitoring system in banks, greatly improving the safety of surveillance area and providing strong protection for bank [35].

(**Tarief et al, 2011**) introduced a design of simple security system for museums and banks, it is based on the fact that any human being radiates an infrared radiation. The system consists of antenna working at IR frequency band, Programmable System on Chip and a Silicon Controlled Rectifier (SCR). The output of the SCR will activate a load which could be a buzzer or Light Emitting Diode (LED) to give either a sound or a light alarm, respectively. If any human being approached valuable protected item, the antenna collects the infrared radiation and gives an analogue output equivalent to this radiation. This system can activate all time and has immediate response, no apparent components, no interference with daily business [10].

(**Archana et al, 2017**) introduced system that comes with detect the ATM theft from Robberies, and then it reduce the problem with existing technique that used in the society. When the theft occurs, PIR sensor is used to observe the human motion, then alarm is produced beep sound from ATM machine. This system uses ARM controller based on embedded system to process real time data collected using the PIR sensor. Once the alarm is produce the beep sound, the DC Motor is close the ATM door, Stepper motor is used to leak the gas inside the ATM to bring the thief into unconscious stage, Camera continuously sends the videos to the computer for future surveillance. after that the system send message to the nearest police station and corresponding bank through the GSM and in next second send the videos to concerned Bank and as well as police station through the GPS. The system used to prevent the robbery from easily caught out [11].

(**Saranraj et al, 2020**) introduced security system based on biometric and authentication techniques to guarantee better security in ATM exchanges in order to prevent the robbery to take cash until if he caught the valid ATM card holder, client IDs and passwords. The proposed system consists of ARDUINO Controller, finger print sensor, RFID smart Tag per user and RFID Reader. To enhancing security in this system, Fingerprint sensor are used and also the ATM room is kept under surveillance for the whole day. First, the individual need to put the ATM card in the ATM reader. The user whoever entering ATM room need to put the RFID TAG in the RFID Reeder and the user afterward need to keep their unique finger impression. When the RFID and the unique mark are matching the client can perform further exchanges. This framework gives high proficiency

and maintains a strategic distance from the illicit exchanges. It is profoundly solid for security related issues [36].

(**Verma et al, 2013**) designed system for validating, monitoring and controlling the security at bank locker rooms. It includes a biometric system which are responsible for the security of the main door of the locker room and the system also includes a RFID system to provide access of the locker room area to only authorize people. To monitor the unauthorized people in the locker room area, a passive infrared sensor is fixed. In case of any unauthorized motion the picture from the camera will be mailed to security officials and the alarms will be on to inform the local security [9].

(**Hongjie et al, 2011**) introduced a design technique to recognize the illegal modification of criminals in Automatic Teller Machine (**ATM**) for purpose to obtaining username and password. The design is when user enters and leaves the self-service Bank, the system automatically obtains two images, and compares them, if there are differences in the images, then alarms are triggered. The authors used image recognition and image matching processing techniques of computer vision technology. This design reduced the problem of low efficiency in the real-time monitoring system especially when there are many monitoring points [37].

(**Khera et al, 2014**) introduced security system based on image processing technique. The system focus on the safety of the bank locker rooms by detecting and controlling unauthorized motion. The system save the images whenever the motion will be detected, so the images will can be used in future for investigation, then the system communicate the image data continuously to the remote location control rooms using web based monitoring through Local Area Network (**LAN**). Finally the warning text short message service (**SMS**) is sending to the operator using GSM technique. The authors addressed the problem with current manually supervised security system is that if the robbery occurs then the banks are not been able to identify the robbers due to lack of proof [8].

(**Blessed Joshua, 2016**) introduced security system based on three levels of authentication in order to secure the locker of bank. In this system the author used RFID, Pattern recognition, Bluetooth and GSM techniques. The algorithm of the system starts

when the user uses the RFID card which is read by the RFID reader and the user has to draw the specified pattern registered in the server during the first registration. After that the pattern is checked by Camera-OPEN CV image processing step. If the pattern is matched, then a message is sent to the user's mobile with an OTP and the user is asked to type the OTP in the android mobile app , after this the OTP + UIN is sent through Bluetooth and it is checked in the bank's server. If it is matched the locker opens or alert message is sent to the manager by GSM module. The system made the access to the bank lockers and obtaining information of the financial institution and customer by others is difficult [38].

(**Kumari et al, 2019**) introduced locker security system based on Face Recognition, voice recognition and GSM techniques, which may be utilized in Banks, Security Offices and financial institution for giving protection to highly-priced possessions. In this system, only the legal character can get admission to the treasured such things as cash and jewels from locker moreover this system is reliable and inexpensive [39].

From all the previous studies, it is clear that there is an intense interest to wards direction of robust and safety security systems with good performance, but the used techniques in these systems are not depending on employees to detect the criminals, when he is feeling and watching the start of crime. At this state the employee stay arms folded so he cannot do anything especially at gunpoint.

Now we will present some human-security system interaction studies that based on hand gestures techniques.

(**Gupta et al, 2015**) presented a hand gesture analysis for human-security system interaction. The system is used for the banking security and security purposes like companies or personal secured places. This system uses techniques to recognize the special hand gesture of the employees so that nearby police can be informed about the robbery. The authors configured the system for recognition the hand gesture to five processes of techniques such as image acquisition, skin color information for recognizing hand gesture which is obtained from the arm region of the hand, background removal, canny edge detection and contour detection. The authors used skin color and canny edge detector algorithms which are implemented in OpenCV library [4].

(Kumar et al, 2018) proposed a hand gesture recognition system based on Sign Language, the Process of it have eight steps of techniques (image acquisition, skin color, removal of background, canny edge detection, Feature extraction, image classification , training and testing of data). The authors collected dataset of 100 images for 15 hand gestures, and they used SVM classifier to train and test the dataset in order to recognize the gestures. This system can be used for security purpose and also for deaf people [5].

The systems in these studies are depending on hand gesture techniques of employees to detect the criminals. These techniques when used in banks easily identifiable and observed by suspicious persons, criminals or accomplices that leads to doubt and escape them in timely manner or exposing employees, customers and money to risk. Moreover hand gesture techniques are ineffective in state when the hands are tied.

3.3 Critical Evaluation of Existing Approaches

The previous studies is evaluated and presented the Strengths and weaknesses of their used techniques as shown in Table 3.1 and Figure 3.1.

Table (3.1) Critical Evaluation of Existing Approaches

Techniques		Author	Year	Strength	Weakness
human behavior recognition	Computer vision	B Georis	IEEE2004	The system modelled a large set of scenarios to take into account the variety of bank robberies to prevent any drama.	1) This systems focus on the behavior of attacker and robberies only. 2) Most of these systems provide the safe to lockers, valuables and not for employee and customer Safety.
Refresh background image based on OpenCV	computer version and pattern recognition	Yongxiang Wu	IEEE2009	The system also fills the blank of real time alarm monitoring system in banks	
An infrared radiation of human	electronic	Tarief M. F. Elshafiey	IEEE2011	The system can activate all time and has immediate response, no apparent components, no interference with daily business	

Table (3.1) Continued					3) The employee cannot do anything if he is seeing a crime or robbery in his bank (Arms folded).
Sensor , Controllers, GSM and GPS technology	Electronic and communication	Archana	IEEE2017		
biometric and authentication techniques	security	B. Saranraj et al.	IEEE(2020)	This framework gives high proficiency and maintains a strategic distance from the illicit exchanges. It is profoundly solid for security related issues.	
image recognition and image matching processing	Computer vision	Hongjie et al.	IEEE2011	The system reduced the problem of low efficiency in the real-time monitoring system especially when there are many monitoring points.	
biometric and RFID systems technique	Security	Amit Verma et at	2013IEEE	The security can be used in other highly restricted areas also like private offices, laboratories etc.	
RFID, Pattern recognition, Bluetooth and GSM	Biometric , Computer vision and Communication	Blessed Joshua	2016 IEEE	It made the access to the bank locker and obtaining information of the financial institution and customer is difficult.	
Face Recognition, voice recognition and GSM	Computer vision ,electronic and communication	Kumari et al.	International Journal of Scientific Research and Review ISSN No.: 2279- 543XVolume 07, Issue 03, March 2019 UGC Journal No.: 64650	This system is reliable and inexpensive.	
Motion of background image using MATLAB and GSM technique	Computer vision and communication	Khera et al	IEEE2014	low cost, stable link, large coverage area, and low communication cost	

Table (3.1) Continued					
Hand gesture recognition	Computer vision	Gupta et al.	IEEE2015	The system is also used for security purposes like companies or Personal secured places	Easily identifiable and observed by suspicious persons and attackers or accomplices.
Hand gesture recognition based on Sign Language	Computer vision	Kumar et al	IEEE 2018	The system can be used for security purpose and also for deaf peoples	And ineffective in the state that hands are tied.

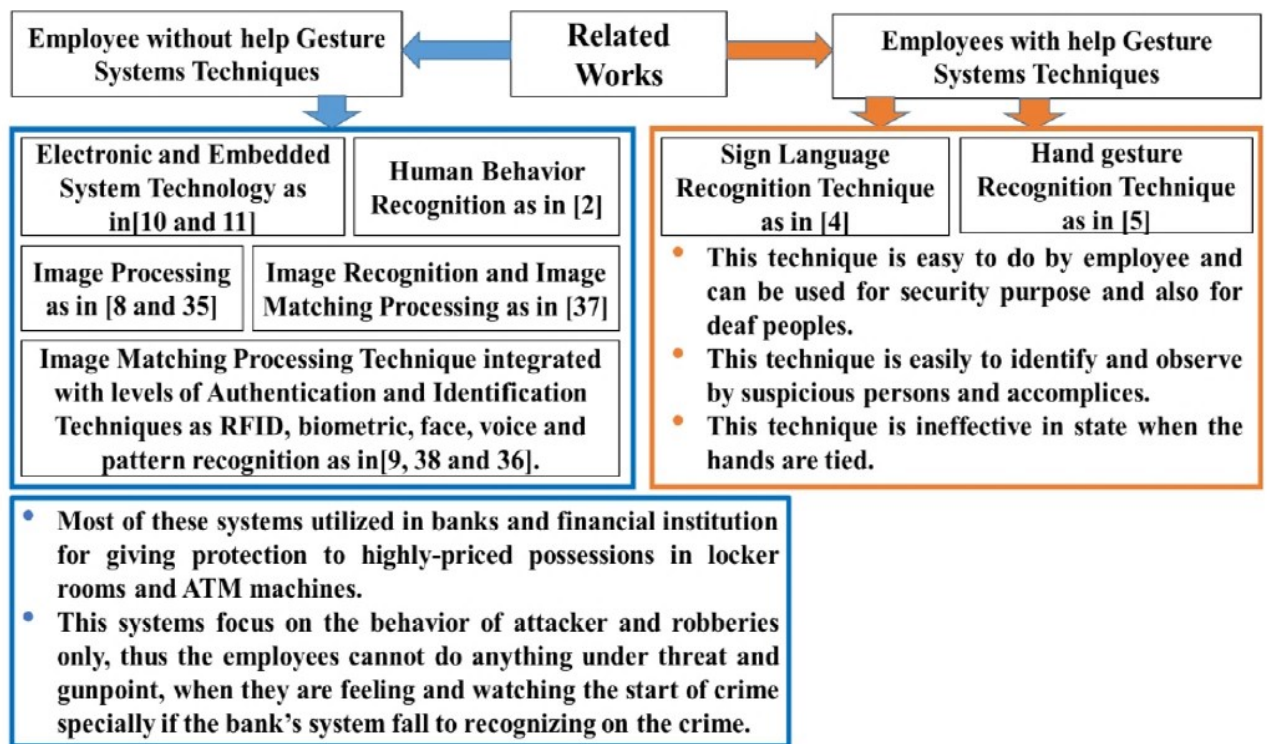


Figure (3.1) Critical Evaluation of Existing Approaches

3.4 Remarks

This research is based on blink gesture recognition technique which is difficult to be observed by the criminals, so that the security man of bank and nearby police station can be informed about crime just by recognizing the special blink gesture made by the employee in front of the cameras and without any knowledge, feeling and attention by any suspects and accomplices. In addition, this research is not eliminates the previous studies, but it can

work in parallel with them, and add a new feature, and also fills the blank of real time alarm monitoring systems and providing security to banks which has never used before.

3.5 Chapter Summary

The analysis of literature review had broadened the scope of security systems issues in banks and can conclude that security systems play an important role in the bank institutions. The information which is collected from this chapter is used as a guidance to propose new method for developing new security system.

CHAPTER FOUR: PROPOSED METHOD & IMPLEMENTATION

4.1 Introduction

Most current systems have some gaps because they used techniques and methods which, are not efficient. So, to enhance this problem, a new method is proposed as described in details in this chapter.

4.2 Proposed Method

The proposed method relies on duration and amplitude of blink to recognize the gesture. This method is fast, efficient, and easy to implement by using python with OpenCV and dlib libraries. This method based on important computer vision techniques: are Facial landmark, Eye aspect ratio (EAR), Duration and amplitude of blink.

4.2.1 Facial Landmark Technique

The face plays an important role in visual communication. By looking at the face, human can automatically extract many nonverbal messages, such as humans' identity, intent, and emotion. In computer vision, the facial information gained through the facial landmark locations (key points) can provide important information for human and computer interaction, entertainment, security surveillance, and medical applications. These points are describing the unique location of a facial component as shown in Figure 4.1. The facial landmark detection is the process of localizing the key facial points on a face, including the eyes, eyebrows, nose, mouth, and jawline. The facial landmark detection algorithms aim to automatically identify the locations of the facial key landmark points on facial images or videos.

Due to facial images and videos undergo to varying facial expressions, head poses, illuminations, and facial occlusion. So some algorithms can be able to handle with these challenging conditions, while some others may fail. It is usually assumed that face is already detected and given for most of the existing facial landmark detection algorithms. The detected face would provide the initial guess of the face location and face scale [40].

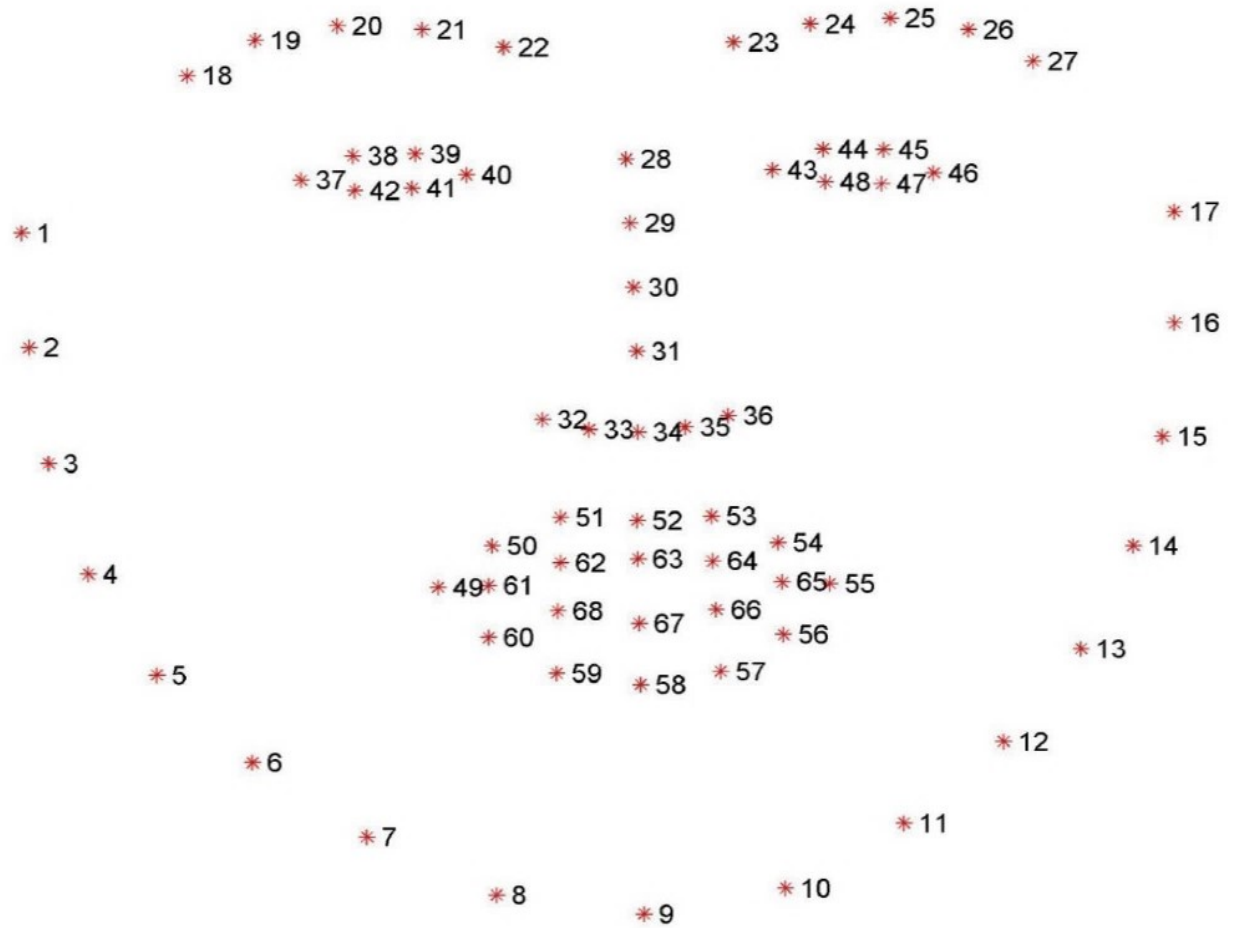


Figure (4.1) Visual representation of 68 key facial landmarks

4.2.2 Eye Aspect Ratio (EAR) Technique

This technique was proposed by (Soukupova et al) [3]. The eye aspect ratio is simple calculation based on the ratio of distances between facial landmarks of the eyes each eye has 6 unique coordinates, 2 on the top of the eye, 2 on the bottom and 1 on each side horizontal edge of the eye as shown in Figure 4.2. The **ear** gives us a singular value, relating the distances between the vertical eye landmark points (p2, p3, p6, p5) to the distances between the horizontal landmark points (p1, p4). Six (6) coordinates are named as P1 to P6 starting from the left edge of the eye and going clockwise with the remaining coordinates. The **ear** is mostly constant when eye is open and getting close to zero while closing an eye as shown in Figure 4.2.



Figure (4.2) six unique coordinates of the eye. The ear is mostly constant when eye is open and getting close to zero while closing an eye

a) Soukupova Method Details [33]

The eye blink is a fast closing and reopening of a human eye. Each individual has a little bit different pattern of blinks. The pattern differs in the speed of eyelid movement, in a degree of squeezing the eyes and in a blink duration. The eye blink lasts approximately 100-400ms. Soukupova et al. method exploited state-of-the-art facial landmark detectors to localize the eyes and eyelid contours. From the landmarks detected in the image with face, Soukupova derived the eye aspect ratio (EAR) that is used as an estimate of the eye openness state.

This approach is supervised thus training process is needed. An SVM classifier is learned on training dataset. Since the per-frame EAR may not necessarily recognize the eye blinks reliably, a classifier that takes a larger temporal window of a frame into account is trained. For every video frame, the eye landmarks are detected. The eye aspect ratio (EAR) between height and width of the eye is computed by Equation (4.1).

$$EAR = \frac{||p2-p6||+||p3-p5||}{2||p1-p4||} \quad (4.1)$$

Where p1 to p6 are the 2D landmark locations.

The EAR is mostly constant when an eye is open and is getting close to zero while closing the eye. It is partially person and head pose insensitive. Eye aspect ratio of the open eye has a small variance among individuals and it is fully invariant to a uniform scaling of the image and in-plane rotation of the face. Since eye blinking is performed by both eyes synchronously, the EAR of both eyes is averaged. An example of an EAR signal over several frames in the video sequence is shown in Figures 4.3 and 4.4.

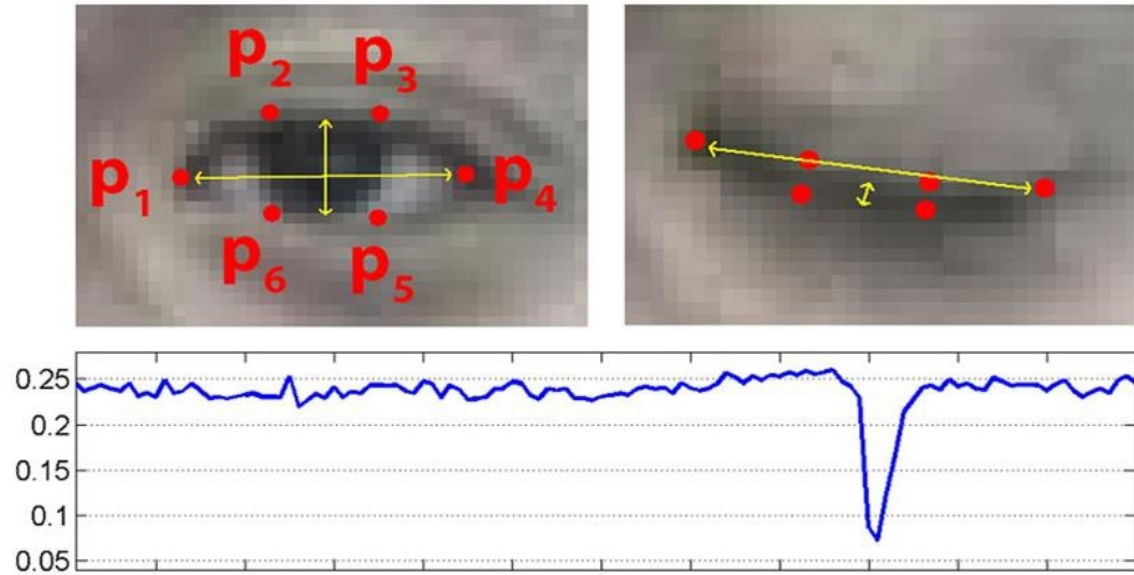


Figure (4.3) Open and closed eyes with landmarks p_i automatically detected. The eye aspect ratio EAR in Eq. (4.1) plotted for several frames of a video sequence. A single blink signal is present.

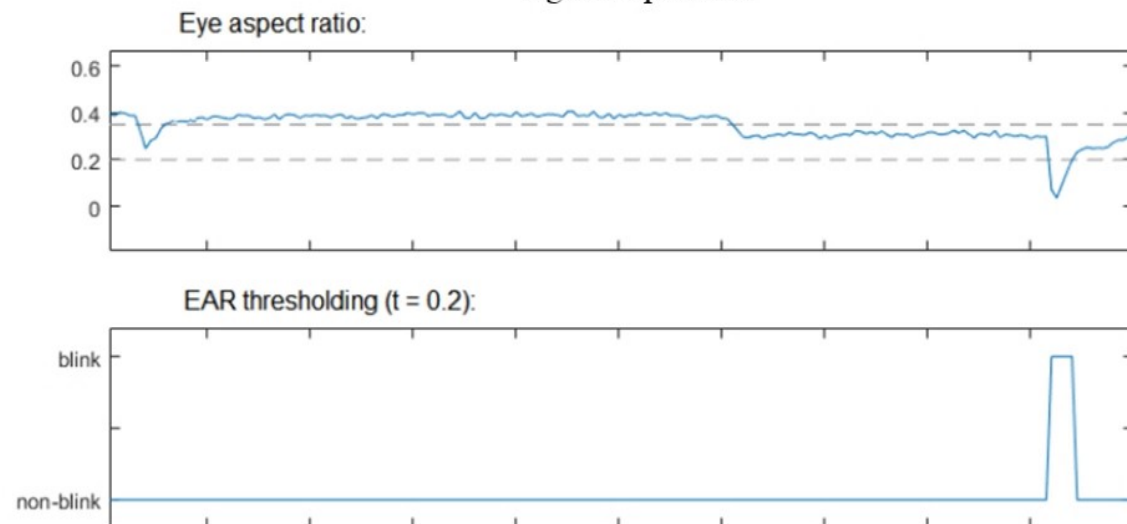


Figure (4.4) an example pointing to how difficult is to set a correct threshold while thresholding the EAR curve. If the threshold is set to 0.35 the left blink will be detected but in right half of a plot there would be long false positive. If the threshold is set to 0.2 then the right blink is detected but the left blink is missed.

b) Eye blink detection using Support Vector Machine classifier

It generally does not hold that low value of the EAR means that a person is blinking. A low value of the EAR may occur when a subject closes his eyes intentionally for a longer time or performs a facial expression, yawning, etc., or the EAR captures a short random fluctuation of the landmarks. Therefore, the classifier that takes a larger temporal window of a frame as an input. Due to a normally blink length being from 100 ms to 400 ms, so the approximately 430 ms can have a significant impact on a blink detection. Further A single blink is present. A red box demonstrates the scanning time window. Values are given for 30fps videos. Thus, for example having 30fps video, for each frame a 13-dimensional feature is gathered by concatenating the EARs of its ± 6 neighboring frames as shown in Figure 4.5.

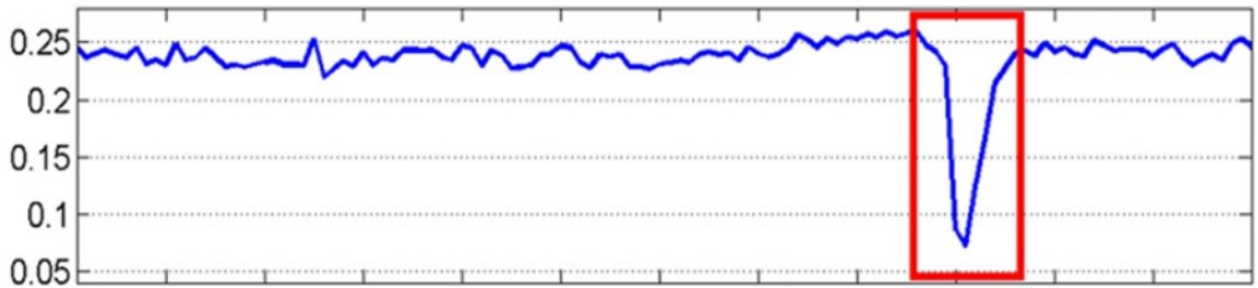


Figure (4.5) the eye aspect ratio EAR plotted for several frames of a video sequence. A single blink is present. A red box demonstrates the scanning time window.

A linear SVM classifier (called EAR SVM) is trained from manually annotated sequences. Positive examples are collected as ground-truth blinks, while the negatives are those that are sampled from parts of the videos where no blink occurs, with 5 frames spacing. Additionally, to avoid only eye closing or only eye opening being detected as a positive blink, the starts and ends of the ground-truth blinks are considered as negatives exemplified for training the SVM classifier. While testing, a classifier is executed in a scanning-window fashion. A 13-dimensional feature is computed and classified by EAR SVM for each frame except the beginning and ending of a video sequence. The values are proportionately recalculated for different frame rates than 30 fps. This algorithm runs online, it means that the blink is detected immediately after the blink's end.

c) EAR SVM eye blink detector

The SVM classifier uses a 13-dimensional features for all frame rates. Therefore time windows must be interpolated to be 13-dimensional for different frame rates than 30 fps. The experiment with EAR SVM is done by the authors in a cross-dataset fashion using all the three blink datasets. It means that the SVM classifier is trained on the **Eyeblink8** and **ZJU** and tested 1 on the **Silesian** and all three combinations are alternated. The testing is done using sliding window with a step of a single frame. Besides testing the EAR SVM method that is trained to detect the specific blink pattern, and compared with a simple baseline method, EAR thresholding, which only thresholds the EAR in Eq. (4.1) values. The precision-recall curves shown in Figure 4.6 of the EAR thresholding and EAR SVM classifier were calculated by spanning a threshold of the EAR and SVM output score respectively.

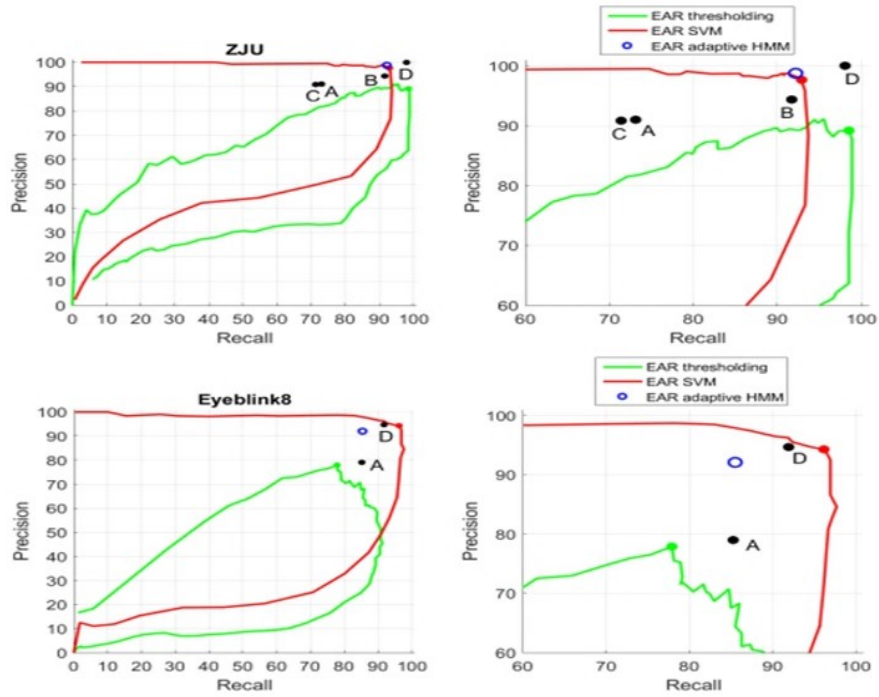


Figure (4.6) the precision-recall curves

Notice, the precision-recall curves are not monotonic due to the matching between the ground-truth and detection blinks involved in the evaluation, So it does not hold that with increasing recall it decreases precision and vice versa. For easier comparison the operational points with the highest F1 score were extracted from the precision recall curves as in Equation (4.2).

$$F1 = 2 \frac{precision \cdot recall}{precision + recall} \quad (4.2)$$

d) Datasets for blink detection

i. ZJU Dataset

The ZJU dataset is consisting of 80 short videos of 20 subjects. Each subject has 4 videos: with and without glasses, 3 videos are frontal and 1 is an upward view. The 30fps videos are of size 320×240 px. An average video length is 136 frames and contains about 3.6 blinks in average. An average IOD is 57.4 pixels. The subjects do not perform any noticeable facial expressions. They look straight into the camera at close distance, almost do not move, do not either smile nor speak. A ground truth blink is defined by its beginning frame, peak frame and ending frame.

ii. Eyeblick8 Dataset

This dataset is more challenging. It consists of 8 long videos of 4 subjects that are smiling, rotating head naturally, covering face with hands, yawning, drinking and looking down probably on a keyboard. These videos have length from 5000 to 11000 frames, also 30fps, with a resolution 640×480 pixels and an average IOD 62.9 pixels. The videos contain about 50 blinks on average per video. Each frame belonging to a blink is annotated by half-open or close state of the eyes.

iii. Silesian Dataset

The Silesian dataset is recorded in 100 fps in a resolution of 640×480 pixels. It consists of 5 videos of men at close distance in front of the camera. The average IOD is 146.7 pixels. The subjects almost do not move and they are recorded from frontal position all the sequence. Each video lasts about two minutes, more accurately and have length from 8000 to 16000 frames, so 56 blinks on average per video sequence.

4.2.3 Duration and Amplitude Techniques

The blink duration is the duration of complete blink in a given time interval, another interesting blink parameters is an amplitude which describes the degree openness of eyelids

of eye, it differs between people. There are two types of voluntary gestures that can be configured by using blink parameters of eyelids are amplitude and duration gestures.

Duration gestures which are the changes in duration time of blink, the eyelids of eye stay closed in predefined period of time. **Amplitude gestures** which are the changes in duration and amplitude of blink, the openness of eyelids are changing differently from normal, it increase in predefined period of time and decrease in predefined period of time. In this type there is no need to completely close the eye.

The proposed gestures of our system interaction had specific and sufficient duration and amplitude thresholds values, in order to avoid any trigger of a system by the spontaneous eye blink of human as shown in Figure 4.7.

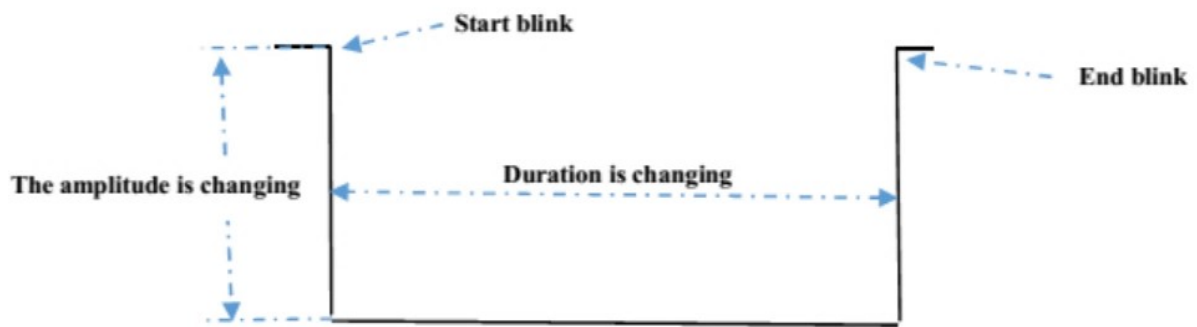


Figure (4.7) the proposed blink gesture

4.2.4 Proposed Method Methodology

The methodology of the proposed method include Face Detection, Facial Landmarks Detection and gesture classification steps as shown in Figure 4.8. Before detecting the facial structure points of eyes in the face region, firstly localizing the face in stream video. Then detecting the blink gesture, by measuring the current value of **eye aspect ratio (EAR)**, which is ratio of distances between facial landmarks of the eyes and **compares** this value with predefined thresholds. After that measuring the duration and an amplitude of the current blink and compare them with predefined thresholds, in order to classifying and recognize the true gesture. Finally alarm is activated, inform the local security and send SMS message to nearby police station about suspected in bank.

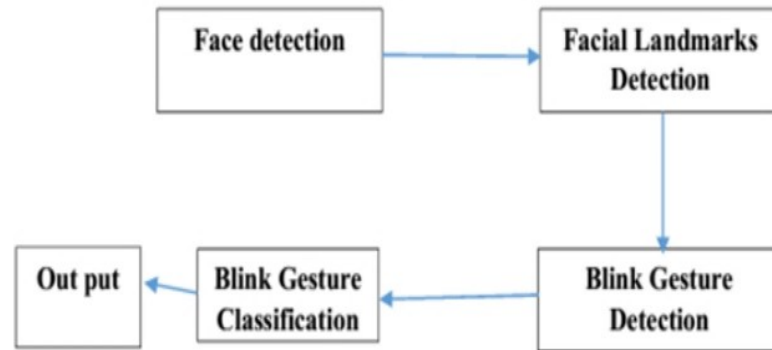


Figure (4.8) Proposed method methodology

4.3 Component and Tools

There are two types of tools which are used in this work hardware such as computer and cameras and software such as programming languages, editors and more.

4.3.1 Laptop and Cameras

- a) **HP laptop computer:** (64-bit Windows 10, Intel Core 2 Duo CPU at 2.13 GHz processor 4G RAM) as shown in Figure 4.9.a.
- b) **Logitech C270 HD web camera :**(Sensor Resolution: 720p and speed rate: 30fps) as shown in Figure 4.9.b
- c) **Text Editor** as Microsoft Word Editor.



(a)



(b)

Figure 4.9. (a) Core 2 Duo laptop 4G RAM (b) HD 720p web camera

4.3.2 Python with Open CV and Dlib Libraries

A general purpose programming language called **Python** was started by Guido van Rossum that became very popular very quickly, mainly because of its simplicity and code readability. In this language the programmer is able to express his ideas in fewer lines of code without reducing or damaging the readability. Python is slower when compared to languages such as C/C++. That said, Python can be easily extended with C/C++, which allows us to write computationally intensive code in C/C++ and create Python wrappers that can be used as Python modules. The obtained two advantages are: first, the code is as fast as the original C/C++ code (since it is the actual C++ code working in Background) and second, it is easier to code in Python than C/C++. The simple, intuitive syntax allows you to focus on learning the basics of deep learning, rather than spending hours fixing crazy compiler errors in other languages [16] [38].

OpenCV (open source Computer Vision Library) is mainly used in some advanced image processing, feature detection and tracking, motion analysis, object segmentation and recognition etc. The Platform MATLAB, will be installed under the directory of OpenCV. Its focus on real-time image processing and includes patent-free implementations of the latest computer vision algorithms. Open CV now supports a multitude of algorithms related to Computer Vision and Machine Learning. Open CV supports a wide variety of programming languages such as C++, **Python**, Java, etc., and this is available on different platforms including Linux, Windows and Android. Due to the OpenCV is released under a BSD license and its source codes are open, the developers can modify the codes or add new classes into the library, so it is used in **academic projects** [41].

Open CV-Python is the Python API for Open CV, combining the best qualities of the Open CV C++API and the Python language. Open CV-Python is a library of python bindings designed to solve Computer vision problems. For the original Open CV C++ implementation Open CV-Python is a Python wrapper. Open CV-Python makes use of Numpy, which is a highly optimized library for numerical operations with MATLAB-style syntax. All the Open CV array structures are converted to and from Numpy arrays. This

method also makes it easier to integrate with all other libraries that make use Numpy such as SciPy and Matplotlib [13].

Dlib is a general purpose cross-platform software library. It is written in the programming language C++. Dlib is a modern C++ toolkit containing machine learning, image processing algorithms and tools for creating complex software in C++ to solve real world problems and there are using of number of its tools from **python applications**. Dlib is used in both industry and academia in a wide range of domains including robotics, embedded devices, mobile phones, and large high performance computing environments. Due to the Dlib is open-source software released under a Boost Software License, so it is using in any application, free of charge [42] [43].

4.4 Implementation

The main parts of the work is programmed by using libraries of python program language except the detection of face and facial landmarks which is programmed by using C++ Dlib library. The OpenCV and imutils libraries are used for capturing, collecting and processing of images of videos stream. The processing time of the videos stream is computed by using imutils library and the time for processing one frame using our proposed method is not exceed 200 ms, these times already include a processing time taken for finding face, facial landmarks and gesture. Thus the proposed algorithms run in about 5-6 fps while using laptop (64-bit Windows 10, Intel Core 2 Duo CPU at 2.13 GHz processor 4G RAM). The recording of participants' videos is done using standard (640 × 480,30fps web camera) and Logitech C270 web camera. The text is written in Microsoft Word Editor.

A real-time gesture detector prototype has been prepared. It runs in python and uses imutils to capture images from Logitech C270 web camera. A one beep sound when a subject repeated the true gesture tow times in front of the camera, So that he can know that the system has accepted his interaction's request.

The implementation activities of this research consist of procedure, experimental methodology and experiments details.

4.4.1 Procedure

The experimental procedure for each participant consisted of four parts: initial training, input of tow predefined eye gestures on a laptop and interaction with a security application running on the laptop. In the first three parts, are recorded videos from the web cameras of the laptop. During the fourth task, the recognition was carried out in real time. In more details:

- 1) Initial training, so four participants will record videos who are perform the predefined gestures.
- 2) The USB web camera was fixed at distance about 70 cm away from the participant's face, so participant must keep the distance to the camera as stable as possible.
- 3) Each participant was asked to blink sometimes (normal blinks and gesture blinks, alternately).
- 4) Finally python, OpenCV and Dlib code will written to perform face and facial landmark detection to recognize the gesture in video streams.

4.4.2 Experimental Methodology

The first section to build an eye blinks gesture of the system, is localizing the face by applying a pre-trained HOG + Linear SVM object detector algorithm in a given frame of a video stream as shown in Figure 4.10. This algorithm is implemented in dlib library.

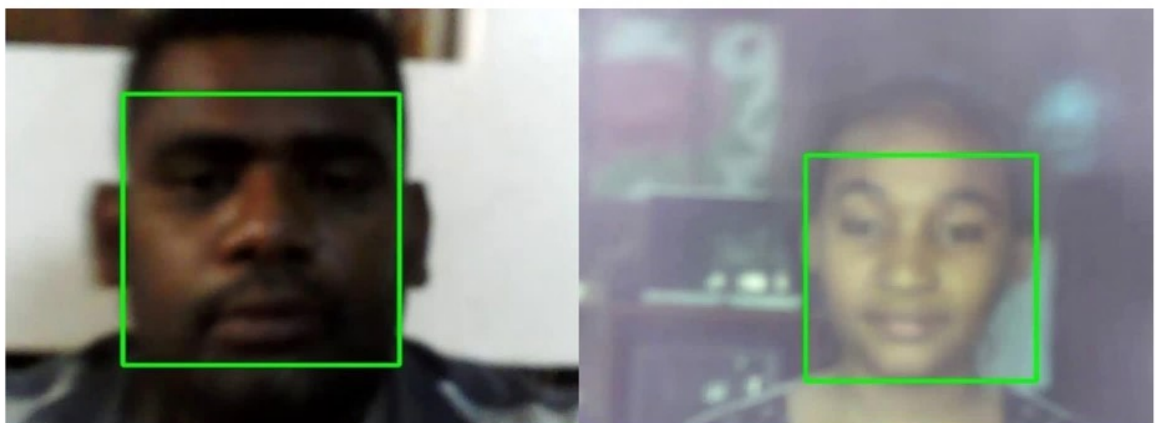


Figure (4.10) detection the face

Then facial landmark detector is used, to perform facial landmark detection of 68 key facial coordinates, in order to localize the eyes landmarks in face region as shown in Figure 4.11. This detector is an implementation of One Millisecond Face Alignment with an Ensemble of Regression Trees in [44] which, is included in the dlib library.

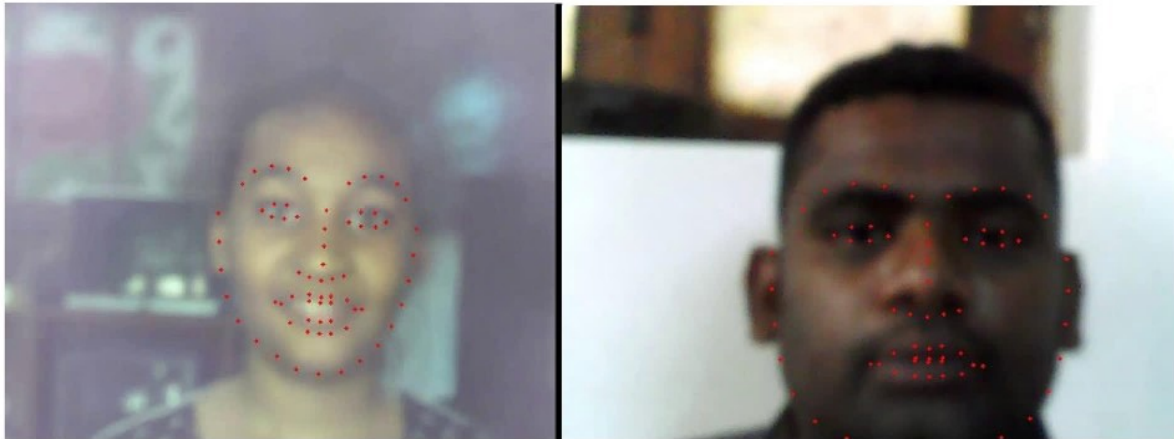


Figure (4.11) Visual representation of 68 key Facial landmark points.

Once we have the facial landmarks for both eyes as shown in Figure 4.12, the Soukupova et al. method [3] is used to measure eye aspect ratio (EAR), which is ratio of distances between the six facial landmark key coordinates points of each eye. The EAR value for both eyes is calculated and then the mean is taken. The EAR is calculated for each consecutive frame and a thresholds for EAR is set in the code.



Figure (4.12) Visual representation of landmark points of each eye.

The second section is to determine when the blink is considered to be a gesture, to do that we first set two threshold values of ear that must be set in code, higher one for open eye (HT) and lower one for closed eye (LT). After that we measure the real detected ear value and compare it with LT and HT threshold values, if the comparison is true, then we determine the amplitude and duration of the detected blink and compare it with predefined thresholds.

The duration is the number of sequence frames when eyes are in closed state or when value of ear = \leq LT value for period of time equal TFseq. When TFseq is predefined threshold must be set in the code. The duration gesture is different change in duration as in natural blink, so the duration gesture is detected when the value of ear fall to LT threshold and stay in it predefined period of time equal TFseq, then it rise again to value when the eyes are in normal open state. The amplitude of blink is openness of the eyes, and the amplitude gesture is different changes of the openness as in natural blink. so amplitude gesture detected when ear value change from low to high or high to low in TFseq predefined period of time (number of frames) passing through LT and HT threshold values. Figure 4.13 shows the signal of complete blink gesture with LT, HT and TFseq thresholds.

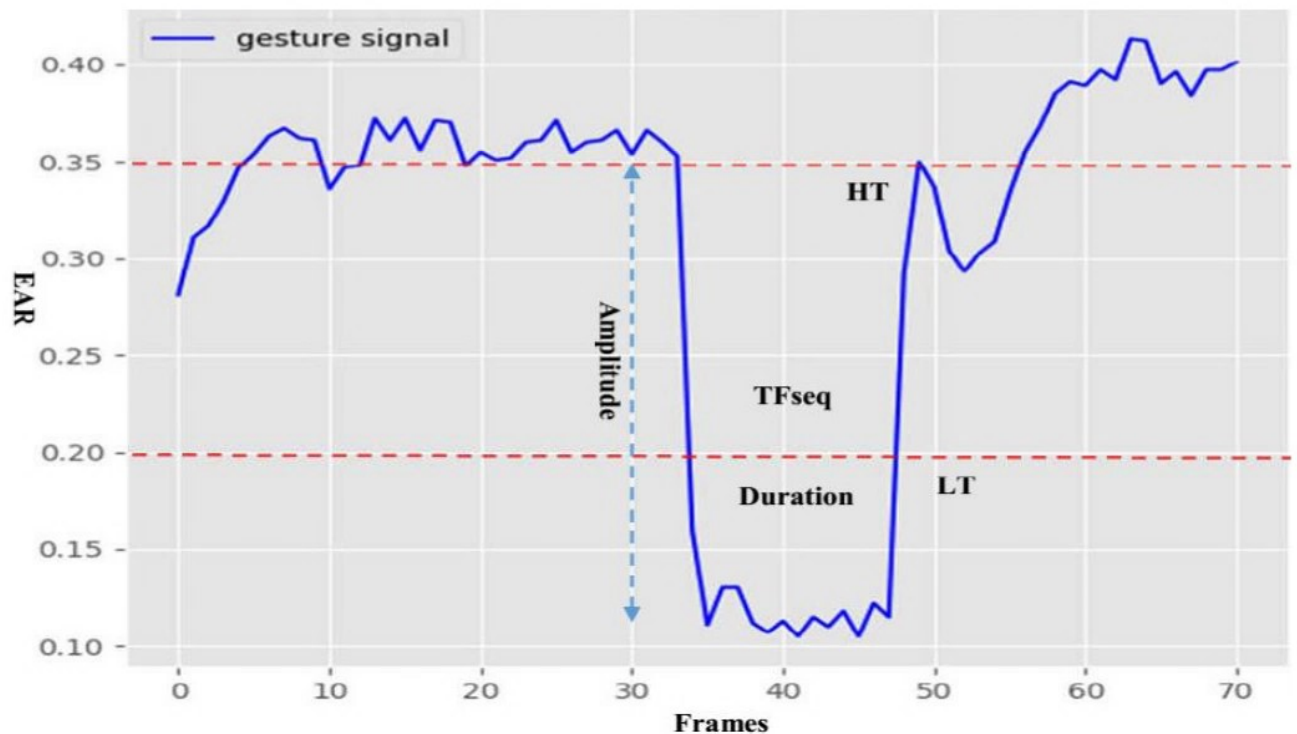


Figure (4.13) Real signal of complete blink gesture

Once the comparison is true and matched, the **alarm is activated** and the system also sends warning message to the security officials and nearest police station. This section is performed by a code on python language. The flowchart of the experimental methodology is shown in Figure 4.14.

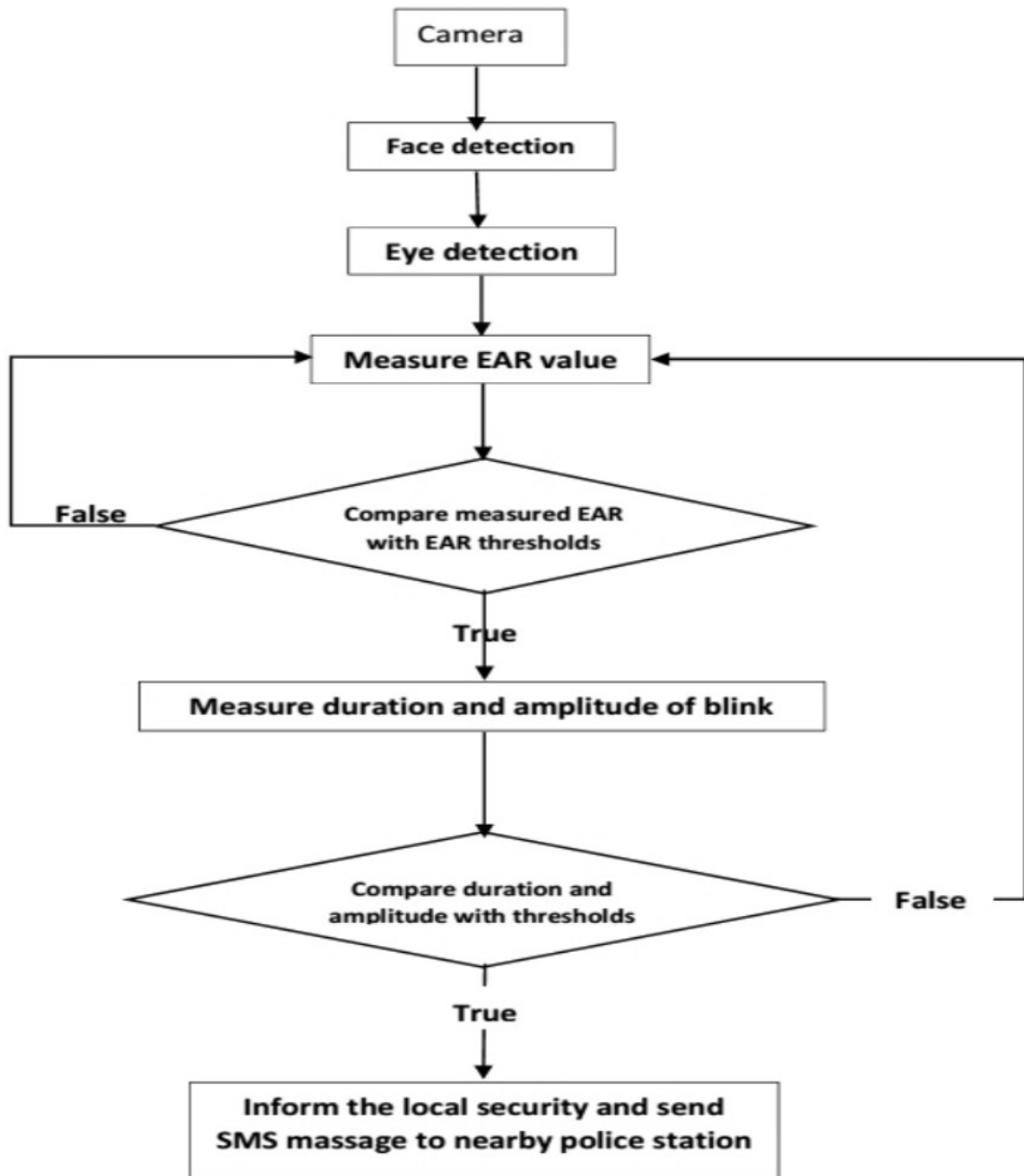


Figure (4.14) Experimental flowchart of the system

4.4.3 Experiments Details

In this part, we presented all experiments that related to gesture types, how to create them, the time needed to process them, and also the conditions that affect the accuracy of their recognition.

a) Participant vs. EAR

The EAR values for 4 different people were recorded. The ear values for completely closed, normal open and maximum open of the eyes were measured. For accuracy the ear was calculated for several times and the mean value was recorded. Table 4.1 shows the mean values. Using the values a Graph was plotted as shown in Figure 4.15.

Table (4.1) the mean values of ear for complete closed, normal open and max openness of the eyes.

participants	EAR average for eyes closed	EAR for normal eyes open	EAR Values of max eye openness
Participant 1	0.104	0.345	0.425
Participant 2	0.092	0.323	0.389
Participant 3	0.072	0.313	0.382
Participant 4	0.103	0.294	0.358

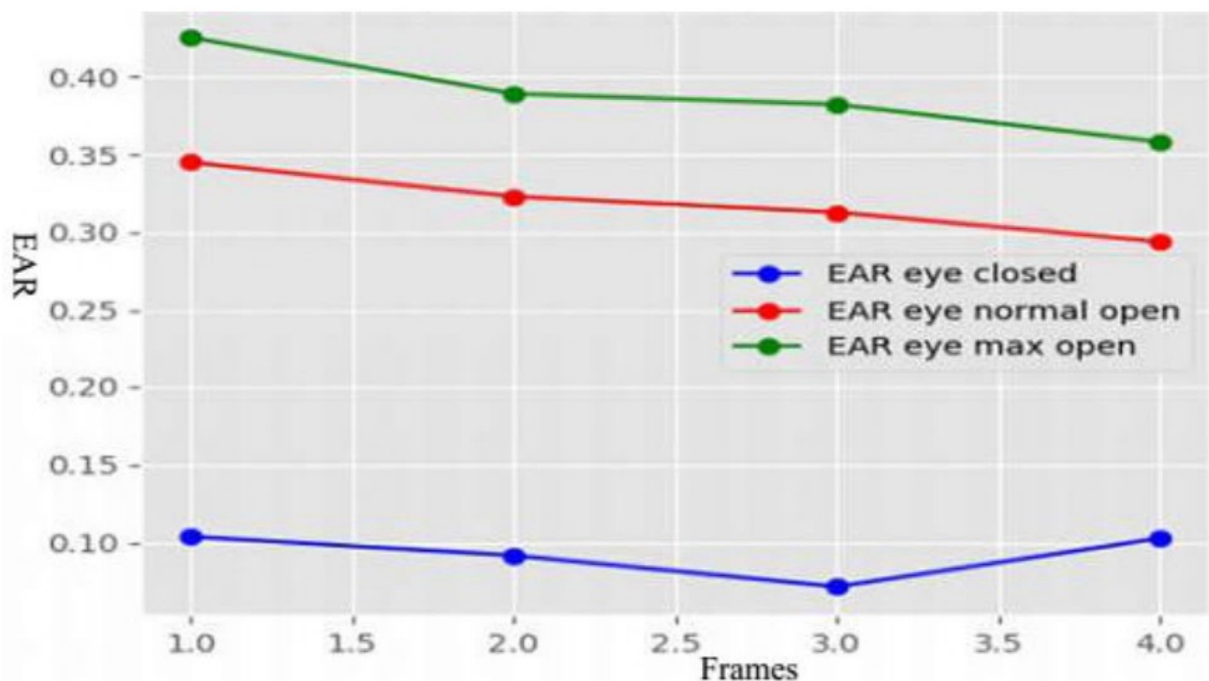


Figure (4.15) EAR of 4 participants for open, closed and max openness of the eyes

b) Gesture's processing Vs. Frames

In Table 4.2 the maximum processing time of frame and mean frame rate for recognition of face, landmarks and gesture from real video streams, that captured by 30fps web camera was measured, and the processing time and frame rate without any appearance of a human face was measured. Also we measured the baseline corresponds to displaying the video without any processing by same camera.

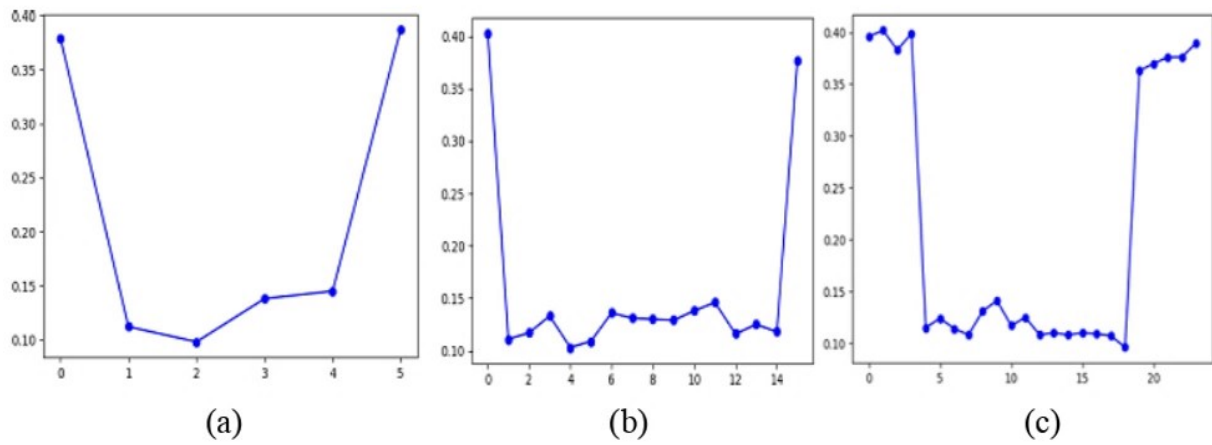
Table (4.2) Processing time of frame and mean frame rate for detection, without detection and baseline corresponds to displaying the video without any processing by same camera.

	Baseline	Without human face	Detection
Mean frame rate	27.3 fps	5.9 fps	5.3 fps
Processing time	(0 – 37.2)ms	(110 - 142) ms	(125 – 200) ms

In Table 4.3 the time of complete gesture for different number of frames was measured to find the most appropriate number of frames to set in the code. Figure 4.16 (a), (b) and (c) show some signals' samples of complete gesture for different number of frames and the appropriate number of frames must be set in the code. Using values in Table 4.3 a Graph was plotted as shown in Figure 4.17.

Table (4.3) Gesture processing time for number of frames.

TFseq to set in code	Number of Frames for gesture	Processing time of Gesture
4	6	1.257s
5	7	1.467
6	8	1.668s
8	9	1.879s
12	16	3.343s
*13	*15	*3.137s
*14	*16	*3.353s
15	17	3.553s
16	18	3.718s
18	20	4.186s



Figure(4.16) (a), (b) and (c) samples of complete signal of blink gesture.

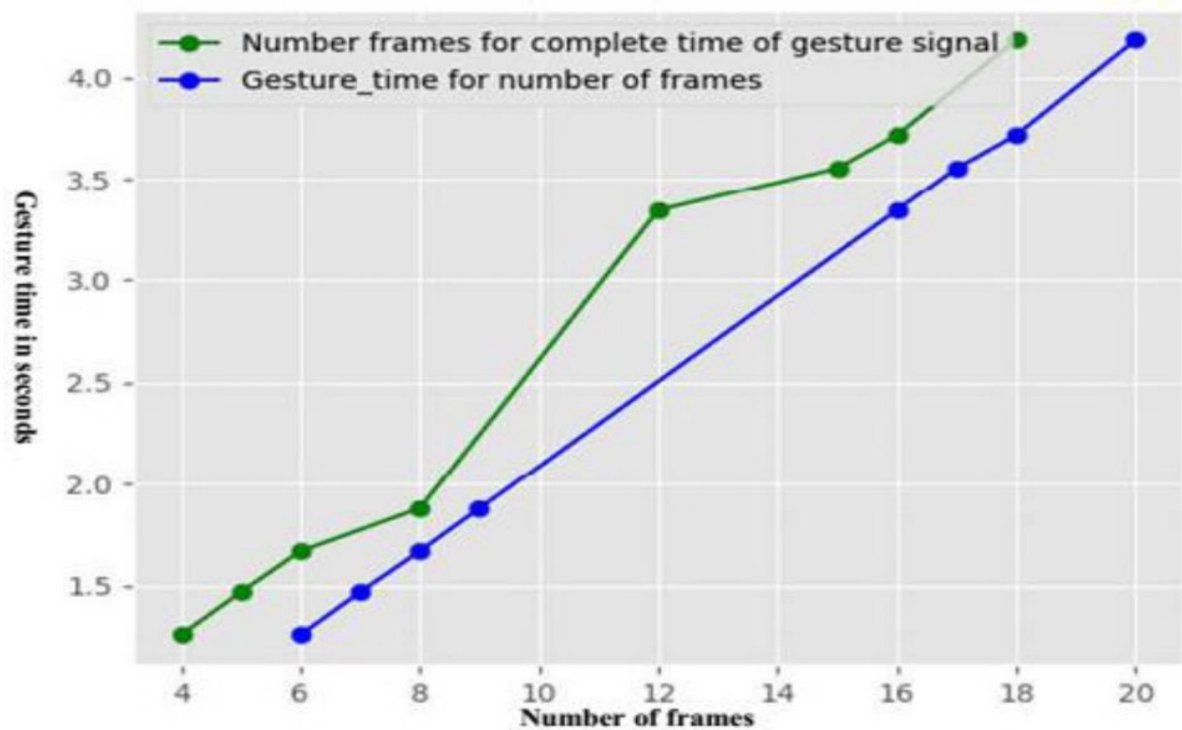


Figure (4.17) Gesture time for number of frames and number of frames for complete time of gesture signal.

c) Gestures Classification vs. Thresholds

There are different types of gestures that can be configured by using threshold values of ear and frames (time) as in the following: First, by using values from Table 4.1 and the plotted Graph in Figure 4.18 we calculated the individual lower thresholds (LT), individual higher thresholds (HT) and common thresholds for all participants. The HT threshold must be a value of ear that bigger than the value when the eyes in normal open state. The LT threshold must be a value of ear that is smaller than value when the eyes are in normal open

state or closed to a value when eyes in closed state. Then by some calculations we found the common HT and LT thresholds as computed by Equations (4.2) (4.3). Second, from Table 4.1 we can see that, the minimum **ear** value of normal open state for participant 4 is **0.294** and the maximum **ear** value for closed state for participant 1 is **0.104**. Hence by using equation (4.2), the **LT** threshold was set to (**0.1989 = 0.2**) in the code. The maximum **ear** value of the normal open state for participant 1 is 0.345 and the minimum **ear** of the maximum openness of the eyes for participant 4 is 0.358. Hence by using Equation (4.3), the **HT** threshold was set to **0.351** in the code.

$$LT = \frac{\text{Big ear value in closed state} + \text{small ear in normal open state}}{2} \quad (4.2)$$

$$HT = \frac{\text{Big ear in normal open state} + \text{small ear in max openness state}}{2} \quad (4.3)$$

Finally, the signal of the **amplitude gesture** was detected, when the value of **ear** rise to **HT** and filing to LT in predefined period of time (number of Sequential frames = TFseq), then it rise again to **HT**. In state of amplitude gesture there is no need to completely close eyes, so the LT threshold can be shafted to be between values when eyes are in closed state and normal open state. Figure 4.14 shows the **HT**, **LT** and **Shafted LT** thresholds (**SLT**). But according to the signal of **duration gesture** is detected, when the value of ear filing to **LT** threshold and stay in it predefined period of time (TFseq threshold), then rise again to value when the eyes are in normal open state. Figure 4.19 show some participants how are doing the amplitude and duration gestures.

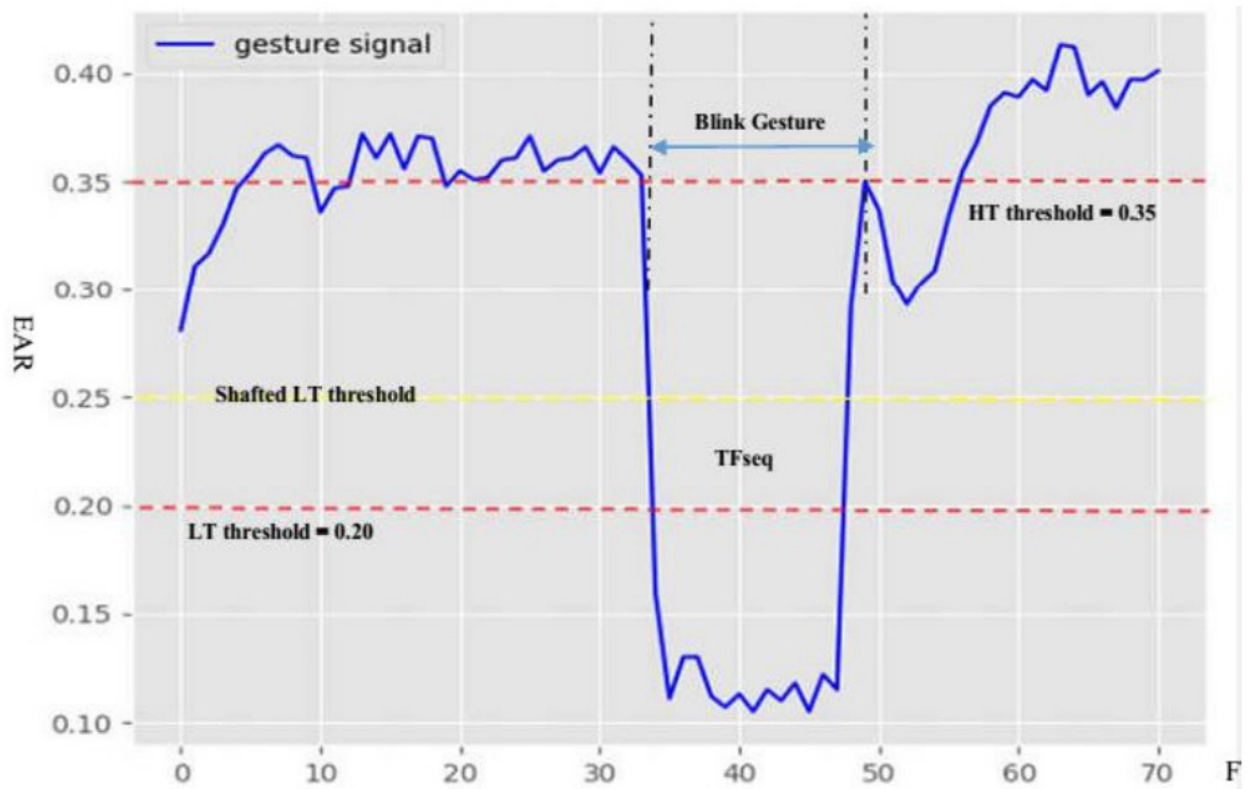


Figure (4.18) Gesture's signal shape and values of the HT, LT and S LT

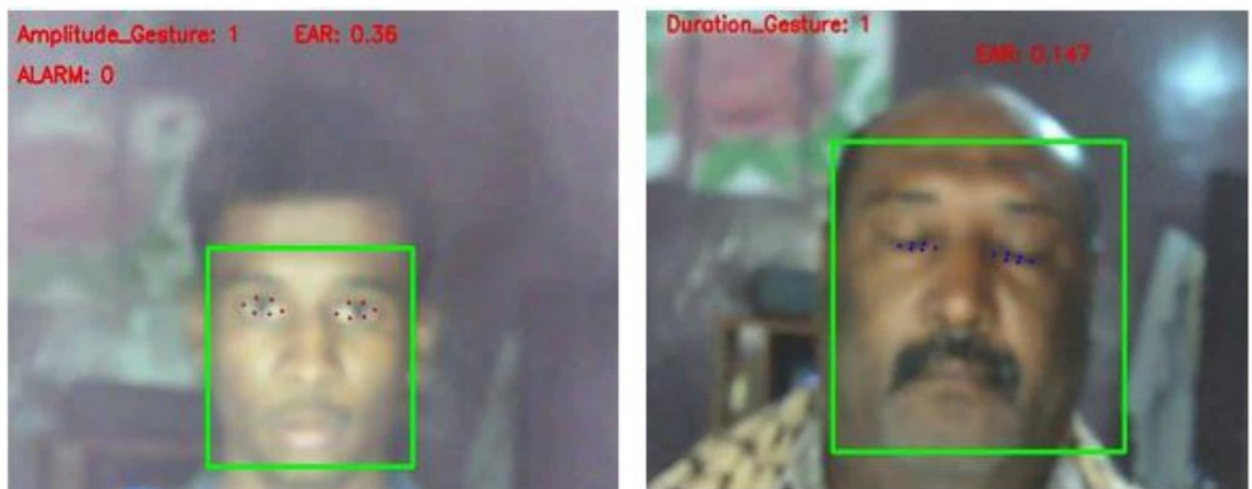


Figure (4.19) participants how are doing the amplitude and duration gestures

d) Gestures vs. Different conditions

Experiments were done with one participant at different luminous intensities, head orientation, wear glasses and occlusions.

i. Luminous intensity

We noticed that the gesture stay detected at all different regular luminous levels even in low lights as shown in Figure 4.20.

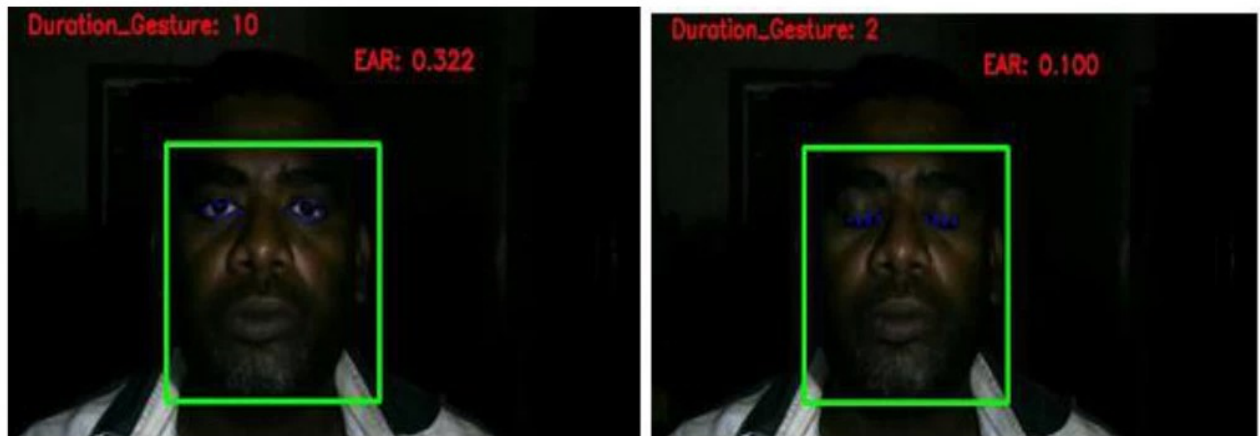
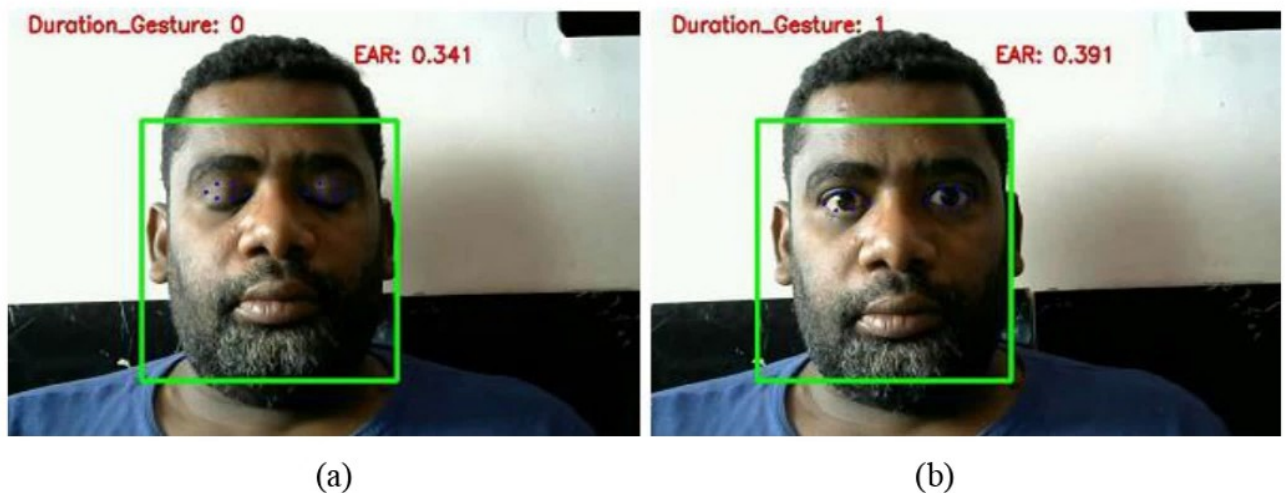


Figure (4.20) Gestures stays detected at different regular luminous.

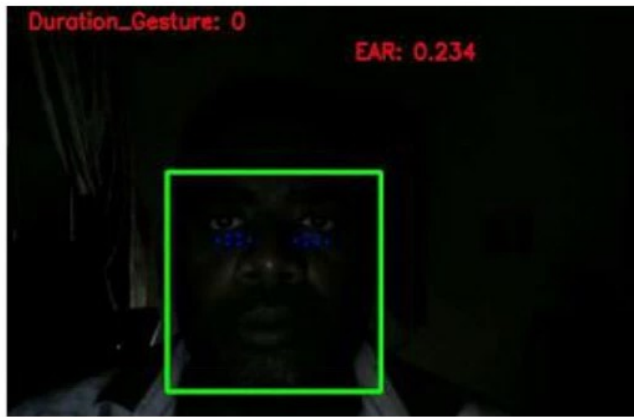
But while luminous intensity is unregularly, for example when lights at one side or very low, the misinterpreted detection occurs because, the predictor of facial landmark considers the cover of eye (upper or lower lid) as an open eye nevertheless eyes are actually closed as in Figure 4.21(a) and (c). In other side due to loss of face detector, the eye landmarks is also lost, therefore lost detection is occurred in duration gesture as in Figure 4.21 (d).



(a)

(b)

Figure (4.21) (a) and (b) Lost and misinterpreted duration gesture detected in unregularly luminous intensity conditions.



(c)



(d)

Figure (4.21) (c) and (d) Lost and misinterpreted duration gesture detected in unregularly luminous intensity conditions.

Due to the amplitude gesture depend on amplitude characteristic of the blink, the misinterpreted detection is small as shown in Figure 4.22 (a) and (b).



(a)



(b)

Figure (4.22) (a) and (b) the detection of amplitude gesture is in stable detection until in unregularly lights.

ii. Head Orientation

An experiment was done in different views of face orientation upper, down and sided views as shown in Figures 4.23 (a), (b) and (c). The amplitude gesture was stay detected until the face is not always frontal to the camera, but it may lose discriminability for out of face rotations.

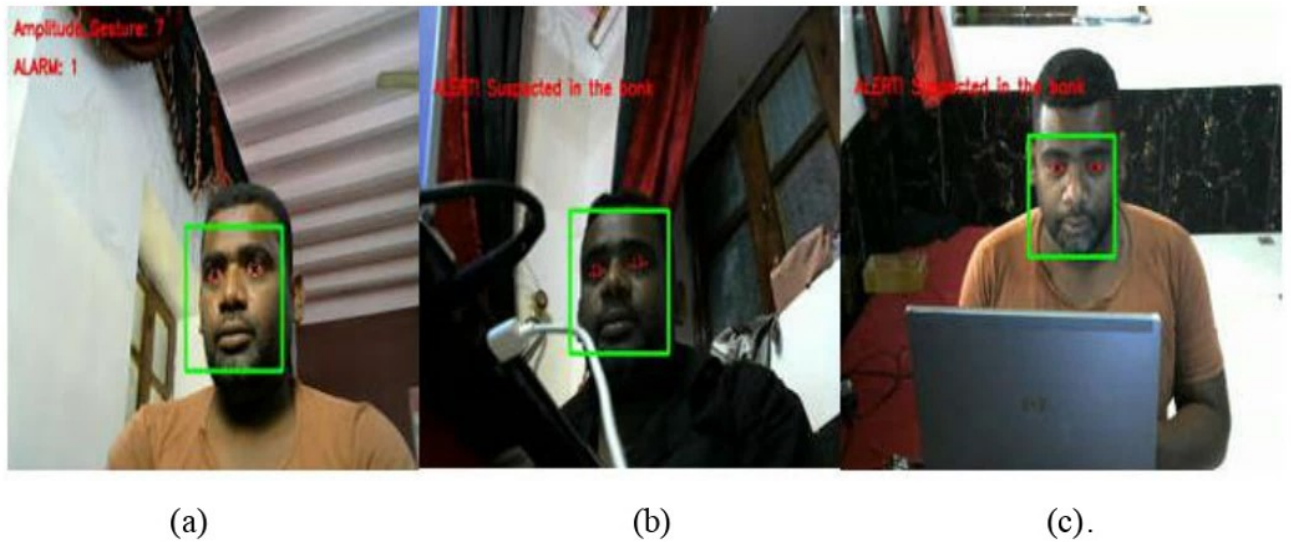


Figure (4.23) (a), (b) and (c) Face orientation: camera put it in different directions side, down and up views.

iii. Glasses and Occlusions

In this experiment, the blink gestures were detected with good accuracy when the participant was wearing glasses as in Figure 4.24 (a) and (b) for duration gesture and also for amplitude gesture as shown in Figure 4.25 (a), (b), (c) and (d).

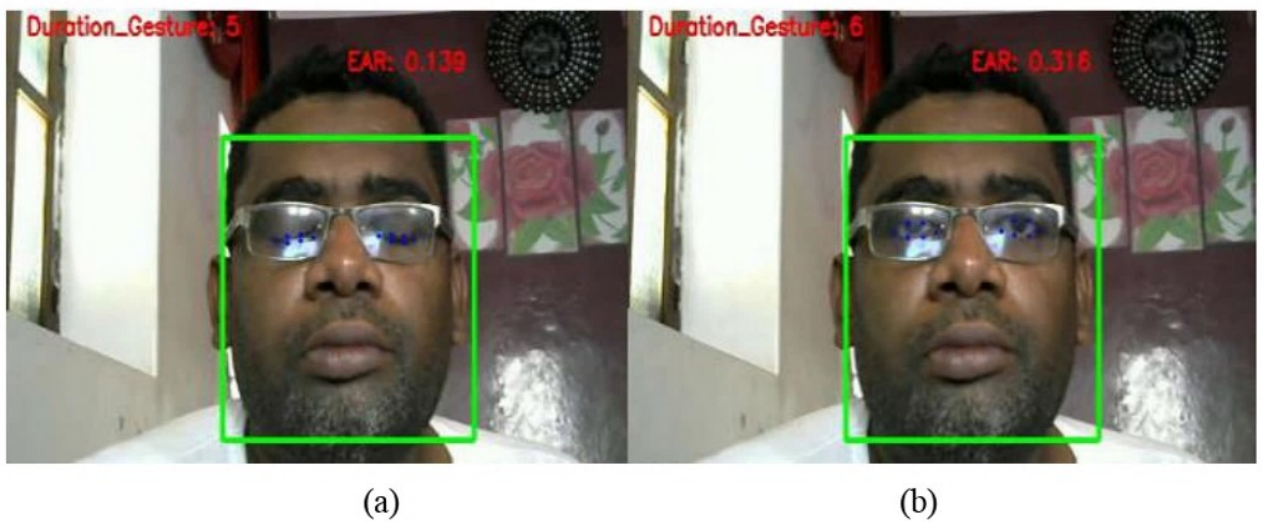


Figure (4.24) (a) and (b) Duration gesture in glasses states



(a)

(b)

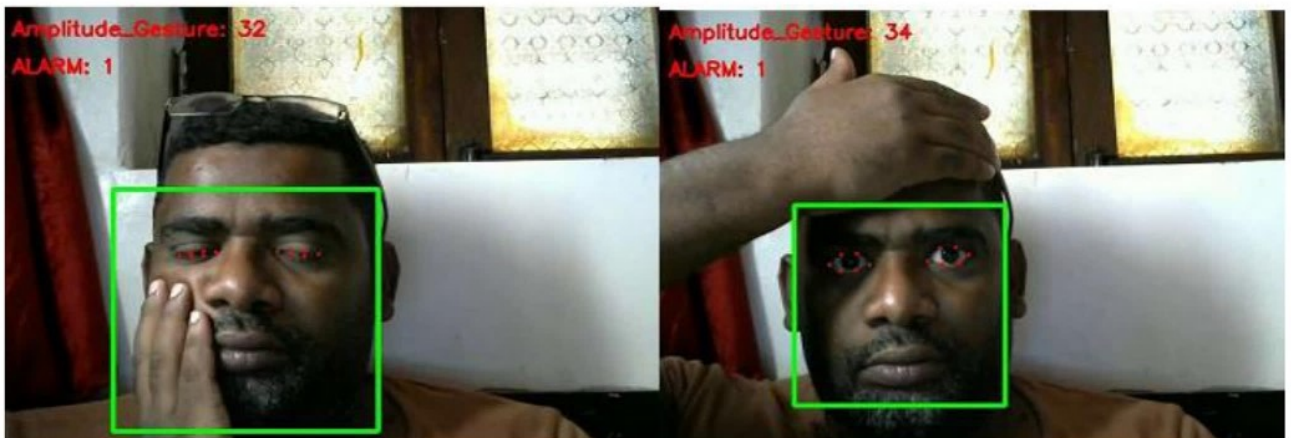


(c)

(d)

Figure (4.25) (a), (b), (c) and (d) Amplitude gestures in glasses states

In otherwise due to occlusions may not be distracting for the landmark detector, so the gestures are also detected when the participant put his hand on face as shown in Figure 4.26 (a), (b) and mouth as shown in Figure 4.27 (a), (b) or nose as shown in Figure 4.28.



(a)

(b)

Figure (4.26) (a), (b) participant put his hand on his face.

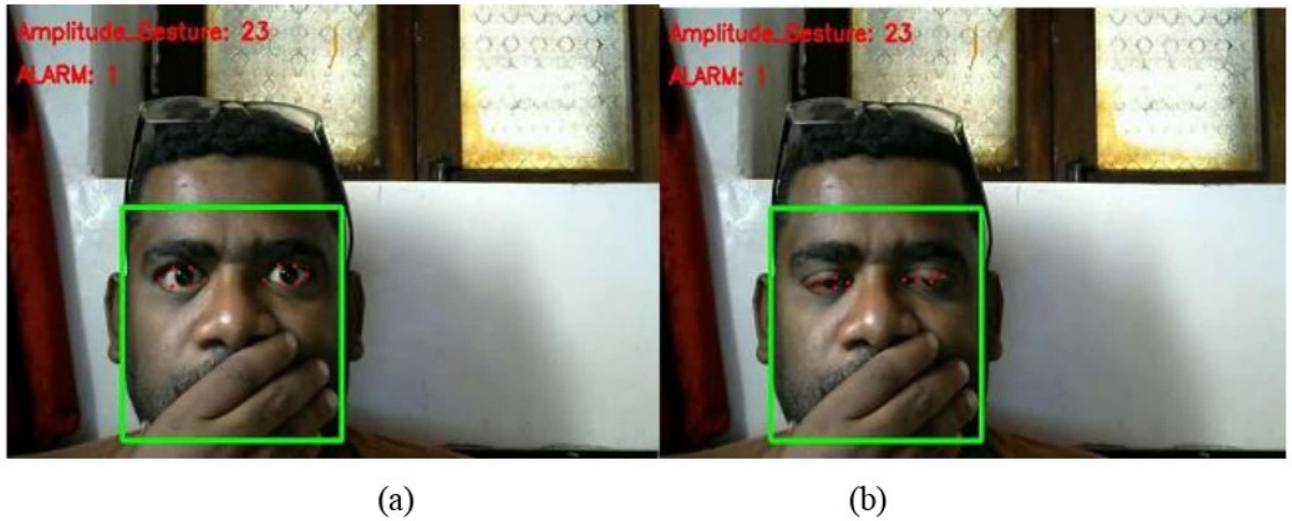


Figure (4.27) (a), (b) participant put his hand on his mouth.

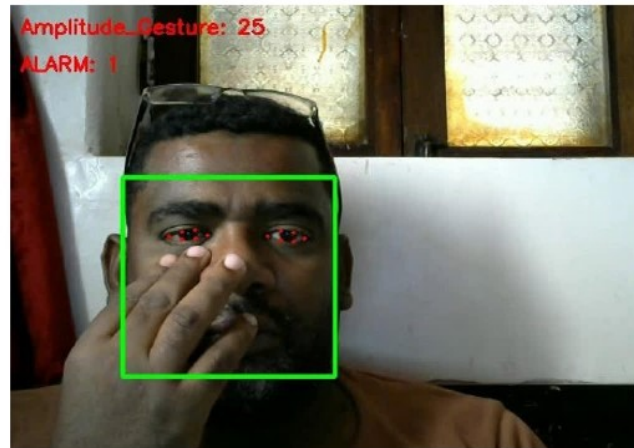


Figure (4.28) participant put his hand on his nose.

However, the amplitude gesture is detected with good accuracy better than duration gesture in states: unregularly lights, wearing glasses, face orientation, speaking and in occlusions.

4.5 Results and Performance of the system

The proposed system for Blink gesture detection was tested using Intel Core 2 Duo CPU at 2.13 GHz processor 4G RAM on the 30 fps sequences frames from the USB Logitech 270 Pro webcam and Jewa web cameras. The size of the input images sequences were equal to 480, 640 pixels. Testing of the system took place in a room illuminated by 2 fluorescent lamps and daylight from two windows. The recording and analysing of videos were implemented by python program language.

4.5.1 Real-time capabilities

First, analysing whether the system was able to process the video and recognize gestures in real-time. To do this, we measured the processing time for each frame and computed the mean frame rate. As can be seen from Table 4.4, the time for individual frame without any detection processing of face, landmarks and gesture is between 110ms, 142ms, and the mean frame rate was 5.9 Fps. If the system performs face and landmarks detection, the processing time up between 125ms, 200ms. So the mean frame rate drop to 5.3 Fps. While the baseline corresponds to displaying the video on the screen without any processing was closed to frame rate of the camera around 27 fps.

Table (4.4) frame rate and processing time with / without face and landmarks detection and also the baseline corresponds to displaying the video on the screen without any processing.

	Baseline	Without a face	Detection
Mean frame rate	27. 3 fps	5.9 fps	5.3 fps
Processing time	(0 – 37.2)ms	(110 - 142) ms	(125 – 200) ms

4.5.2 Recognition performance

The detection of the blink gesture was done in real time with the camera speed of 30 fps. Two kinds of errors were identified: false detection (the system detected gesture when it was not present) and missed gesture (a present gesture that was not detected by the system). The possible decisions distribution of the gesture detector output is presented as shown in Table 4.5.

Table (4.5) possible decisions of gesture output

Possible decision		Actual state	
		Gesture present	Gesture not present
Result of detection	Gesture detected	TP	FP
	No gesture detected	TN	FN

The correctly detected of the gestures are denoted as True Positives (TP), the false detections are denoted as False Positives (FP) and the missed gestures are denoted as True Negatives (TN). Based on these parameters, measure of the system performance was introduced the accuracy as in Equation (4.4) [17].

$$Accuracy = \frac{TP}{TP + FP + TN} \quad (4.4)$$

This measure was used to assess the system robustness in detecting the blink gestures as well as for the overall system performance. The tests of the system in low and good resolution of cameras and others conditions were also performed and the results are presented as in Tables 4.6, 4.7, 4.8 and 4.9.

Table (4.6) Accuracy of the Amplitude gestures recognition system averaged over all four participants in low resolution.

videos	Amplitude Gesture Total GT	Detected Gesture	TN	FP	Accuracy
Participant 1	19	18	1		94.7%
Participant 2	3	3			100%
Participant 3	6	5	1		83.3%
Participant 4	10	8	2		80%
Averaged Accuracy					89.5

Table (4.7) Accuracy of the Duration gestures recognition system averaged over all four participants in low resolution.

videos	Duration Gesture Total GT	Detected Gesture	TN	FP	Accuracy
Participant 1	19	17	2		89.47%
Participant 2	6	6			100%
Participant 3	16	15	1		93.75%
Participant 4	10	9	1	1	82%
Averaged Accuracy					91.3%

Table (4.8) The distribution of the system performance and accuracy, in deferent conditions states: direction views of camera, glasses and low lights.

Deferent conditions states of videos	Amplitude Gesture Total(GT)	Detected Gesture (DG)	TN	FP	Accuracy
Down view	10	7	3	1	64%
Side view	10	8	2		80%
middle view	40	39	1	0	97.5%
Middle and glasses	20	19	1	0	95%
Middle and low light	20	20	0	0	100%

Table (4.9) The distribution of the system performance and accuracy for the duration gestures, in deferent conditions states: glasses, without glasses and low lights.

States of video	Duration gesture Total(GT)	Detected gesture	FN	FP	Accuracy
Middle	20	20	0	0	100%
glasses	20	19	1	0	95%
Low light	20	20	0	0	100%

4.6 Comparison of Results with Previous Works

The used techniques of most previous works do not depend on employee to detect the criminals, when he is feeling and watching the start of crime. The employee in other works use hand gesture technique to interact with system. this technique it used in banks easily identifiable and observed by suspicious persons, criminals or accomplices that leads to doubt and escape them in timely manner or exposing employees, customers and money to risk.

The proposed system combines facial landmark, eye aspect ratio and amplitude and duration of computer vision techniques to recognize the blink gestures which is never before and also it is hard to observe. It achieves a real-time gesture recognition accuracy more than 96%. Table 4.10 and Figure 4.29 show some comparison of proposed work with Gupta and Kumar works. The accuracy is plotted as shown in Figure 4.29.

Table (4.10) The comparison of proposed system with Gupta and Kumar works.

Works	Used technique	Observe by suspicious & accomplices	Difficulty level by employee	Accuracy
Gupta	Hand gesture recognition	Easily	Easily	95.7%
Kumar	Hand gesture recognition	Easily	Easily	94.5%
Proposed system	Amplitude blink gesture recognition	Difficult	Little Difficulty	96.5%
	Duration blink gesture	Difficult	Easily	98.3%

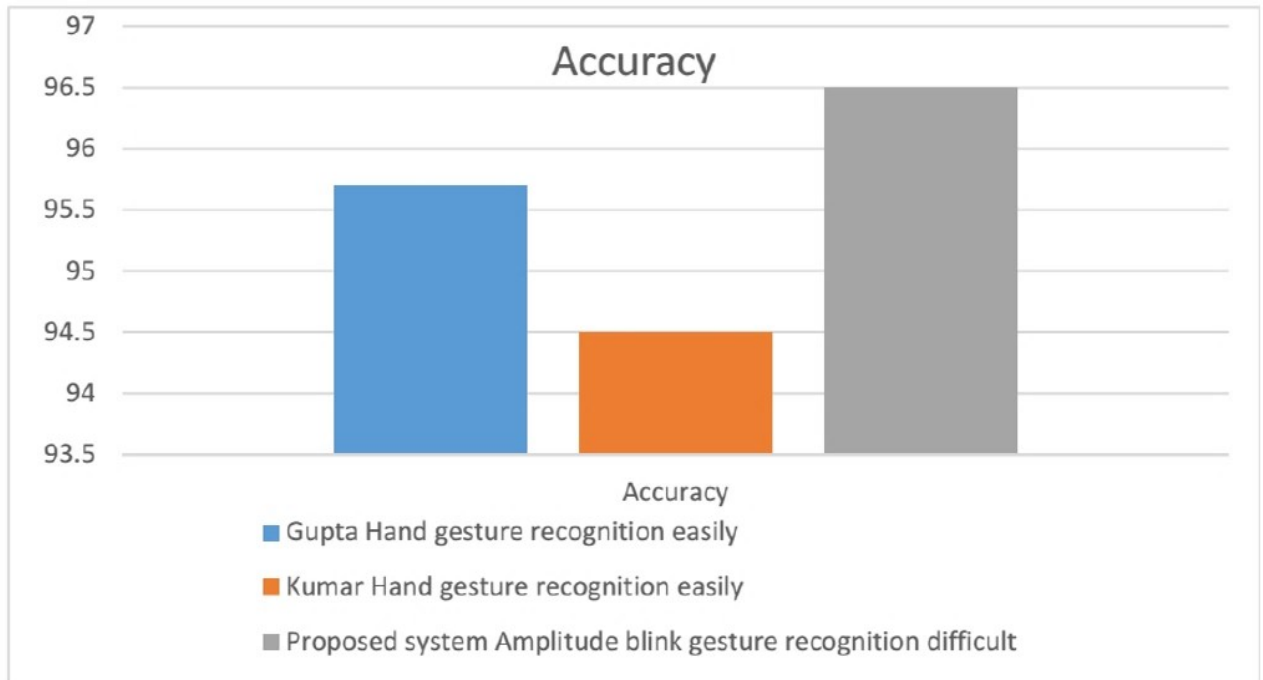


Figure (4.29) The comparison of proposed system with Gupta and Kumar works.

4.7 Remarks

There are some remarks that noticed while experiments as in the following:

- 1) All experiments of this work is implemented in door, with four participant which, perform about 20 blinks, so to obtain a stable accuracy of the results more participants must be needed.
- 2) Most misinterpreted detection occurs in duration gesture, because it is effected by some conditions.
- 3) The frame rate drops to **5.3 fbs** in detection state, that is because low performance and the high load of the used computer.

4.8 Chapter Summary

This chapter vasisgated all techniques that used in proposed method and presented the different types of tools which are used in this work. Finally different experiments are implemented to get on the results and to measure the performance of all system and compared them with previous related works.

CHAPTER FIVE: CONCLUSION AND FUTURE WORK

5.1 Conclusion

In this work security interaction system based on the blink gesture recognition technique is proposed. The employee is pre-trained about how and what gesture to make in case of robbery in bank before joining at the bank. This proposed system combines facial landmark, eye aspect ratio and amplitude and duration of computer vision techniques to recognize the blink gestures. The implemented real time blink gesture recognition security system has an accuracy of more than 96% which, is better as compared with the existing results of related works in [4]and [5].

5.2 Limitation of The Proposed Approaches

The limitation related with the system are as follows:

- 1) The failure of the face detection would directly lead to the failure detection of eye facial landmarks, so the performance of system may be affected by the face detector accuracy.
- 2) The system loss discriminability when the face move to be not frontal to the camera. The solution might be to define the EAR in 3D.
- 3) To get top performance of the system, every participant should have his own LT and HT thresholds, so every participant must has his own sub-system unit.

5.3 Future Work

The limitations described previously can be addressed as future work. Further, future works to improve this human-security system interaction are as follows:

- 1) Propose more blink gestures, which can be used to handle with more security purposes and other interaction intelligent systems as controlling vehicles and also helps the disabled people for start, stop and changing speed of their wheelchairs and cars.

- 2) Duration blink gesture can be integrated with Amplitude blink gesture in one system in order to obtain robust and safety security system with good performance.
- 3) Design human- security systems based on gesture of motion of nose and Eye brows.

References

- [1] Royanews. "جرائم السطو المسلح على البنوك في الأردن في العام 2018".(2018) Retrieved Dec, 27, 2018, from: <https://royanews.tv/news/171426>.
- [2]Georis. B, Maziere. M, Bremond. F and Thonnat. M.(2004) "A video interpretation platform applied to bank agency monitoring" (pp. 46-50). IEEE Intelligent Distributed Surveillance Systems, London, UK.
- [3] Soukupov. T and. Cech. J.(2016) " Real-Time Eye Blink Detection using Facial Landmarks," 21st Computer Vision Winter WorkshopLuka`Cehovin, Rok Mandeljc, Vitomir`Struc (eds.)Rimske Toplice, Slovenia, February 2016. Retrieved from <https://vision.fe.unilj.si/cvww2016/proceedings/papers/05.pdf>
- [4] Gupta. A, Kumar.Y and Malhotra. S (2015)"Banking security system using hand gesture recognition," International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE), Noida, (pp. 243-246).
- [5] Kumar. S, Yadav.G, H. P. Singh. H. P, Malhotra. S and Gupta. A., (2018) "SSVM Classifier and Hand Gesture Based Sign Language Recognition, " International Conference on Intelligent Circuits and Systems (ICICS), Phagwara, pp. 456-460.
- [6] Pradhan. C, Banerjee. D, Nandy. N and Biswas. U,(2019). "Generating Digital Signature using Facial Landmlark Detection," International Conference on Communication and Signal Processing (ICCSP), Chennai, India, (pp. 0180-0184).
- [7] Nyrkov. P, Abramova. V, (2018) "Technologies of Safety in the Bank Sphere from Cyber Attacks, " 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), Moscow, Russia, (pp. 102-104).
- [8] Khera. N and Verma. A, (2014). "Development of an intelligent system for bank security," 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence), Noida, (pp. 319-322).
- [9] Verma. A, (2013) "A Multi Layer Bank Security System," International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), Chennai, (pp. 914-917).

- [10] Elshafiey. T. M. F,(2011). "Design and Implementation of a Museum and Banks Security System Using Antenna as IR Proximity Sensor and PSoC Technology," IEEE Symposium on Wireless Technology and Applications (ISWTA), Langkawi,(pp. 156-161).
- [11] Archana. K, Bhaskara. P (2017), "To Enhance the Security for ATM with the help of Sensor and Controllers", 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, (pp. 456-460).
- [13] Rosebrock. Adrian (2016). "Practical Python and OpenCV", 3rd Edition, pyimagesearch.com.
- [14] Wikipedia, "Image processing," [Online]. Available: https://en.wikipedia.org/wiki/Digital_image_processing.
- [15] wikipedia, "computer vision," (2017). [Online]. Available: https://en.wikipedia.org/wiki/Computer_vision.
- [16]Rosebrock. Adrian (2017). "Deep Learning for Computer Vision with Python", 1st Edition, PyImageSearch.com.
- [17]lak. A. K (2011). "Eye-blink detection system for human-computer interaction," Universal Access in the Information Society 11(4) Springer, (pp. 1-11).
- [18] Vaitukaitis. V and Bulling. A (2012). "Eye Gesture Recognition on Portable Devices," UbiComp '12: Proceedings of the ACM Conference on Ubiquitous Computing,(pp. 711-714).
- [19] Hussain. S (2017). "Hand Gesture Recognition Using Deep Learning," International SoC Design Conference (ISOCC), IEEE,(pp. 48-49).
- [20] Karakostas. A, Lazarou. I, Meditskos.G, Stavropoulos. T. G, Kompatsiaris.I and Tsolaki. M (2015). "Sensor-based in-home monitoring of people with dementia using remote web technologies," International Conference on Interactive Mobile Communication Technologies and Learning (IMCL), Thessaloniki,(pp. 353-357).
- [21] Guesgen. H. W (2013). "Human Behavior Recognition Technologies: Intelligent Applications for Monitoring and Security", New Zealand: Massey University.

- [22] Fauzi. C, Sulistyo. S and Widyawan (2018). "A survey of group activity recognition in smart building," 2018 International Conference on Signals and Systems (ICSigSys), IEEE, Bali, (pp. 13-19).
- [23] Gupta. A, Chitranshi. G, Sharma.V and Singh. H (2017). "Controlling devices using facial expression identification and recognition for assisting physically disabled,". 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, Noida, (pp. 574-576).
- [24] Ravindu. H. M, Bandara. T, Priyanayana. K. S, Buddhika. A. G, Jayasekara. P, Chandima.D. P , Gopura. R. A. R. C,(2020). " An Intelligent Gesture Classification Model for Domestic Wheelchair Navigation with Gesture Variance Compensation," Applied Bionics and Biomechanics, Hindawi, (pp. 1-11).
- [25] Zhao. R, Wang. K, Divekar. R, Rouhani. R, Su. H, Ji. Q,(2018). "An immersive system with multi-modal human-computer interaction," Automatic Face & Gesture Recognition (FG 2018) 13th IEEE International Conference, Xi'an, China, (pp. 517-524).
- [26] MEENA. S, (2011). "A Study on Hand Gesture Recognition Technique," National Institute Of Technology, Roukela, INDIA.
- [27] Shanthakumar.V. A, Peng. C, Hansberger. J, Cao. L, Meacham. S. C and V. Blakely. R, (2020). "Design and evaluation of a hand gesture recognition approach for real-time interactions, " Multimed Tools Appl 79, (pp.17707–17730).
- [28] Wang. K, Zhao. R and Ji. Q,(2018). "Human Computer Interaction with Head Pose, Eye Gaze and Body Gestures," 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, (pp. 789-789).
- [DOI 10.1109/FG.2018.00126](https://doi.org/10.1109/FG.2018.00126)
- [29] Cheng. Bing, Zhang. Chao, Ding. Xiaojuan, Wu. Xiaopei, (2018). "Convolutional Neural Network Implementation for Eye Movement Recognition based on Video," the 33rd youth Academic Annual Conference of Chinese Association of Automation(YAC), Nanjing, China.,(pp.179-184).

- [30] Danisman. T, (2010)."Drowsy Driver Detection System Using Eye Blink Patterns,"(2018). In International Conference on Machine and Web Intelligence, (pp. 230 - 233). Algiers, Algeria Retrved from: <https://hal.archives-ouvertes.fr/hal-00812315/document>.
- [31] Haq. Z. A and Hasan. Z,(2016). "Eye-blink rate detection for fatigue determination", 1st India International Conference on Information Processing (IICIP),(pp. 1-5).
- [32] S. Al-gawwam.(2018). "Robust Eye Blink Detection Based on Eye Landmarks and Savitzky–Golay Filtering". Information, 9(4), 93; Retrved from: <https://doi.org/10.3390/info9040093>
- [33] Soukupov'a. T ",(May 2016). "**Eye-Blink Detection Using Facial Landmarks**",(Diploma Thesis) Center for Machine Perception, Technical University Technick'a 2, 16627 Prague 6, Czech Republic, Prague.
- [34] Michal. Ševčík, et al (2017). " **Eye Blink Detection Using Webcam**," (Bachelor Thesis), Slovak University of Technology in Bratislava - Faculty of Informatics and Information Technologies.
- [35] **Wu, Yongxiang et al.**, (2009). "Research on bank intelligent video image processing and monitoring control system based on OpenCV," 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication, Hong Kong,(pp. 211-214).
- [36] Saranraj. B, Sri Priya. N,(2020). "ATM Security System Using Arduino", 2020 6th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India , (pp. 940-944,).
- [37] Hongjie et al,(2011). "Automatic Alarm System for Self-Service Bank Based on Image Comparation,". International Symposium on Computer Science and Society, Kota Kinabalu, (pp. 48-50).
- [38] Joshua. B, Grace. J and A. M. George. A. M,(2016). "Open CV Pattern Based Smart Bank Security System With Theft & Identification Using Android," 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Bangalore,(pp. 254-257).

- [39] Kumari. P, (2019). "High protection bank locker security system using live image and voice authentication," International Journal of Scientific Research and Review / UGC Journal No.: 64650, vol. Volume 07, no. Issue 03, (pp. 2616-2619), March 2019.
- [40] Wu. Y,(2018). "Facial Landmark Detection: A Literature Survey," IJCV, Springer, (pp. 1-28).
- [41] OpenCV team,(2020). "OpenCV (Open Source Computer Vision Library) " , [Online]. Available: <https://opencv.org/about/>.
- [42] Rosebrock. A,(2017). "Eye blink detection with OpenCV, Python, and dlib," 24 4 2017. [Online]. Available: to: <http://www.pyimagesearch.com>.
- [43] Dlib. teams, "Dlib C++ library," Dlib.net, 14 12 2019. [Online]. Available: Dlib.net.
- [44] Kazemi,(2014). "One Millisecond Face Alignment with an Ensemble of Regression Trees," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (pp. 1867-1874).

List of Publications

Bakunah, R. A., Baneamoon, S. M., (2020). **An Intelligent Security System for a Bank based on Amplitude of Blink Gesture Recognition.** *Journal of Emerging Technologies and Innovative Research (JETIR)*, 7(9), (pp.504-513), September-2020, Available: <http://www.jetir.org/papers/JETIR2009368.pdf>