

---

# Smart Attendance Monitoring System Using Computer Vision

---

*Submitted in fulfillment of the academic requirements for the degree of  
Master of Science in Engineering (Computer Eng.)*

*By*

Louis MOTHWA  
Student No. 217080040

*Under the supervision of:*

Prof. Jules-Raymond TAPAMO  
&  
Dr Temitope MAPAYI



University of Kwa-Zulu Natal

**Examiner's Copy**

August 2019

**UNIVERSITY OF KWAZULU-NATAL, COLLEGE  
OF AGRICULTURE, ENGINEERING AND  
SCIENCE DECLARATION**

The research described in this thesis was performed at the University of KwaZulu-Natal under the supervision of Prof. Jules-Raymond Tapamo and . I hereby declare that all materials incorporated in this thesis are my own original work except where the acknowledgement is made by name or in the form of reference. The work contained herein has not been submitted in part or whole for a degree at any other university.

Signed:.....

Name: Louis Mothwa

Date: August 2019

As the candidate's supervisor, I have approved this thesis for submission.

Signed:.....

Name: Prof. Jules-Raymond Tapamo

Date: August 2019

As the candidate's co-supervisor, I have approved this thesis for submission.

Signed:.....

Name: Dr Temitope Mapayi

Date: August 2019

I, Louis Mothwa, declare that:

- (i) The research reported in this thesis, except where otherwise indicated, is my original research.
- (ii) This thesis has not been submitted for any degree or examination at any other university.
- (iii) This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
- (iv) This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then
  - a) Their words have been re-written but the general information attributed to them has been referenced;
  - (b.) Where their exact words have been used, then their writing has been placed inside quotation marks, and referenced.
- (v) Where I have reproduced a publication of which I am an author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.
- (vi) This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

# UNIVERSITY OF KWAZULU-NATAL, COLLEGE OF AGRICULTURE, ENGINEERING AND SCIENCE DECLARATION 2 - PUBLICATIONS

I, Louis Mothwa, declare that the following publications from part of this dissertation.

1. **L. Mothwa**, J.R. Tapamo, and T. Mapayi, Conceptual Model of the Smart Attendance Monitoring System Using Computer Vision, in *Proceedings of the 2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, ISBN 978-1-5386-9385-8, Las Palmas de Gran Canaria, Spain, 26-29 November 2018, pp. 229 - 234, November 2018 .
2. **L. Mothwa**, J.R. Tapamo, and T. Mapayi, Machine Learning Approach to Attendance Monitoring , Journal Article (in preparation).

Signed:.....



# Acknowledgements

First I would like to thank God for giving me strength, wisdom, and understanding throughout my work.

This work was supported and funded by the Armaments Corporation of South Africa (Armsco). I am grateful to have this opportunity to fulfill my dreams.

I thank my Supervisors Prof J.R. Tapamo and Dr. T. Mapayi for their support, dedication, patience, and endless insight. This feat could not have been possible without your assistance. I am grateful to the University of KwaZulu-Natal students who offered to appear in the live experiments conducted to collect data to produce training and testing dataset's used in this dissertation.

Lastly, I would like to thank my family and friends, for their emotional support and endless motivation, which encouraged me to be determined till the end.

## Abstract

Monitoring of student's attendance remains the fundamental and vital part of any educational institution. The attendance of students to classes can have an impact on their academic performance. With the gradual increase in the number of students, it becomes a challenge for institutions to manage their attendance. The traditional attendance monitoring system requires considerable amount of time due to manual recording of names and circulation of the paper-based attendance sheet for students to sign their names. The paper-based attendance recording method and some existing automated systems such as mobile applications, Radio Frequency Identification (RFID), Bluetooth, and fingerprint attendance models are prone to fake results and time wasting. The limitations of the traditional attendance monitoring system stimulated the adoption of computer vision to stand in the gap. Student's attendance can be monitored with biometric candidate's systems such as iris recognition system and face recognition system. Among these, face recognition have a greater potential because of its non-intrusive nature. Although some automated attendance monitoring systems have been proposed, poor system modelling negatively affects the systems. In order to improve success of the automated systems, this research proposes the smart attendance monitoring system that uses facial recognition to monitor student's attendance in a classroom. A time integrated model is provided to monitor student's attendance throughout the lecture period by registering the attendance information at regular time intervals. Multi-camera system is also proposed to guarantee an accurate capturing of students. The proposed multi-camera based system is tested using a real-time database in an experimental class from the University of KwaZulu-Natal (UKZN). The results show that the proposed smart attendance monitoring System is reliable, with the average accuracy rate of 98%.

# Contents

<b>Declaration of Authorship</b>	<b>i</b>
<b>List of Publications</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Abstract</b>	<b>v</b>
<b>Contents</b>	<b>vi</b>
<b>List of Figures</b>	<b>viii</b>
<b>List of Tables</b>	<b>x</b>
<b>List of Algorithms</b>	<b>xi</b>
<b>List of Acronyms</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background	1
1.2 Overview of Computer Vision Systems and Applications	3
1.2.1 Image Acquisition	3
1.2.2 Image Processing	4
1.2.3 Images Analysis	4
1.3 Motivation	4
1.4 Problem Statement	5
1.5 Research Aim and Objectives	6
1.6 Scope Of The Study And Limitations	6
1.7 Contributions	7
1.8 Overview	7
<b>2 Literature Review</b>	<b>8</b>
2.1 Introduction	8
2.2 Attendance Monitoring System Technologies	9
2.2.1 RFID Authentication System	9
2.2.2 Mobile Smart Attendance System	9
2.2.3 Bluetooth Based Attendance Monitoring System	10
2.2.4 Fingerprint Attendance Monitoring System	10
2.2.5 Attendance Monitoring Systems Using Face Recognition	12
2.2.6 Multiple-Camera Positioning And Collaboration	15
2.3 Face Recognition	17
2.3.1 Face Detection	17
2.3.2 Image Pre-Processing	20
2.3.3 Feature Extraction And Classification	21
2.3.4 Global Feature Extraction And Classification Methods	21
2.3.5 Local Feature Extraction And Classification Methods.	23

---

2.4	SUMMARY . . . . .	24
<b>3</b>	<b>System Modelling And Design</b>	<b>26</b>
3.1	Introduction . . . . .	26
3.1.1	Collection the Training Dataset . . . . .	27
3.1.2	System Architecture . . . . .	29
3.1.2.1	The Graphic User Interface (GUI) . . . . .	32
3.1.2.2	Face Recognition-Components . . . . .	33
3.1.2.3	Back End . . . . .	35
3.2	Summary . . . . .	39
<b>4</b>	<b>Methods and Techniques</b>	<b>40</b>
4.1	Introduction . . . . .	40
4.2	Face Detection . . . . .	40
4.2.1	Multi-camera Architecture . . . . .	45
4.3	Image Pre-processing . . . . .	46
4.4	Feature Extraction . . . . .	49
4.4.1	Principle Component Analysis (PCA) for Eigen Features . . . . .	49
4.4.2	Fisherface Features . . . . .	52
4.4.3	The Local Binary Patterns . . . . .	53
4.5	Classification . . . . .	55
4.5.1	Euclidean Distance Euclidean Distance (ED) . . . . .	55
4.5.2	Support Vector Machine . . . . .	56
4.5.3	Naive Bayes Classifier . . . . .	58
4.5.4	Backpropagation Neural Network . . . . .	59
4.6	Summary . . . . .	61
<b>5</b>	<b>Experimental Results And Discussion</b>	<b>62</b>
5.1	Introduction . . . . .	62
5.2	Experimental Setup . . . . .	62
5.3	Conclusion . . . . .	74
<b>6</b>	<b>Conclusions and future work</b>	<b>75</b>
6.1	Summary of the Work . . . . .	75
6.2	Limitations of the System . . . . .	76
6.3	Future Works . . . . .	76
	<b>Bibliography</b>	<b>77</b>
	<b>Appendix</b>	<b>88</b>

# List of Figures

1.1	Computer vision history time-line [102] . . . . .	2
1.2	Generic architecture of computer vision system using face recognition . . . . .	3
2.1	A generic representation of the RFID system [93] . . . . .	9
2.2	Example of the student attendance management system using fingerprint [8] . . . . .	11
2.3	The overview of a face detection system [84] . . . . .	12
2.4	DBN multi-camera face recognition [7] . . . . .	13
2.5	The conceptual model of the three-fold attendance management system [65] . . . . .	14
2.6	Face recognition system and motion detection denoted by red lines [65] . . . . .	14
2.7	Computation of the target-point angle [64] . . . . .	15
2.8	Face side detected and fused using a CHM [42] . . . . .	16
2.9	Calibration between two cameras, with measured angles and distance from the target object [83] . . . . .	17
2.10	The flow diagram of the face recognition system adopted from [72]. . . . .	17
2.11	Input image or video for face location and detection . . . . .	18
2.12	The Artificial Neural Network (ANN) in face detection classification algorithms [88] . . . . .	19
2.13	The series of Boosting stages [102] . . . . .	19
2.14	Pose angle estimation [67]. . . . .	21
3.1	The interaction of cameras with the system . . . . .	26
3.2	The conceptual area-coverage view enclosed by three static cameras . . . . .	27
3.3	Collection of data during registration of faces. Student faces are registered once by a system Administrator . . . . .	28
3.4	The three components of the Smart Attendance Monitoring System connected to one another . . . . .	30
3.5	The use case diagram of the system . . . . .	31
3.6	Detention and recognition of faces captured from multiple camera . . . . .	32
3.7	The system user interface . . . . .	32
3.8	Form to register students information . . . . .	33
3.9	Email Window . . . . .	33
3.10	The flow diagram of the smart attendance monitoring system . . . . .	34
3.11	The internal Work flow of our smart attendance monitoring system . . . . .	35
3.12	The Relational Entity Diagram . . . . .	36
3.13	Background model behind attendance computation . . . . .	37
3.14	The UML class diagram for the smart attendance monitoring system . . . . .	38
4.1	Rectangular red mark showing the detected faces. The faces are extracted and cropped . . . . .	40

---

4.2	(a) Haar-features are represented with a rectangle pattern. (b) The sum of the pixels inside the rectangle D can be calculated as : $4 + 1 - (2 + 3)$ [50]. . . . .	41
4.3	Face detection with Haar cascade without eye dection . . . . .	43
4.4	Validation of face detection with eye detection . . . . .	43
4.5	Dataset without false detection images. The data is enhanced with histogram equalizer, bilateral filter, and elliptical cropping [73] . . . . .	45
4.6	Live scenarios captured from the UKZN classroom. The scenes shows the importance of positioning multiple cameras in a classroom to create a multi-view environment . . . . .	46
4.7	Pre-processing stages . . . . .	47
4.8	The three phases of histogram equalizer adopted in this work . . . . .	48
4.9	Normal cropped image to elliptical cropped image. . . . .	49
4.10	Summary of the PCA process . . . . .	50
4.11	The training images consisting of $n$ of images represented in the form of a vector . . . . .	50
4.12	a) Normal images, and b) average mean images . . . . .	51
4.13	P Selected eigenfaces such that $P < M$ feature. The selected feature represents the entire training data . . . . .	51
4.14	Each face in the training set can be represented by a weighted sum of $k$ eigenfaces and the mean face. . . . .	52
4.15	General computation of the LBP [5] . . . . .	54
4.16	LBP with different neighborhood radius [46] . . . . .	54
4.17	Face divided into regions with histogram for every region . . . . .	55
4.18	RBF function used to separate non-linear separable data mapped in to feature [92] . . . . .	58
4.19	Schematic of a three layered neural network [82] . . . . .	59
5.1	The sample of training images. . . . .	63
5.2	Detection and recognition of students in a classroom. . . . .	63
5.3	Students faces are recognized from a different position in a classroom . . . . .	64
5.4	Recognition results from different cameras . . . . .	64
5.5	Detection rate in the Multi-camera . . . . .	65
5.6	Classification of feature using ED . . . . .	66
5.7	The RBF kernel value . . . . .	67
5.8	Classification of feature using Support Vector Machine Support Vector Machine (SVM) . . . . .	68
5.9	Classification of feature using Naive Bayes (NB) . . . . .	69

# List of Tables

5.1	Face detection rate from differently positioned cameras at a fixed time . . . . .	65
5.2	The trained results of the BPNN with different number of neurons in the middle layer . . . . .	70
5.3	Learning rate against error rate of the BPNN . . . . .	71
5.4	Comparison of the performance of the proposed method to existing methods using similar feature extraction and classification algorithms . . . . .	72
5.5	Example of register of students attendance for the whole lecture period . . . . .	73

## LIST OF ALGORITHMS

1	Collecting training images of the students . . . . .	29
2	Computing email address to inform about results . . . . .	37
3	Detecting faces with an Adaboost trained cascade classifier [50] . . . . .	42
4	Eye detection . . . . .	44
5	Histogram Equalization [100] . . . . .	48



# List of Acronyms

<b>AMS</b>	Attendance Monitoring System
<b>ANN</b>	Artificial Neural Network
<b>CVS</b>	Computer Vision System
<b>ED</b>	Euclidean Distance
<b>ICA</b>	Independent Component Analysis
<b>LDA</b>	Linear Discriminant Analysis
<b>LBP</b>	Local Binary Patterns
<b>NB</b>	Naive Bayes
<b>PCA</b>	Principle Component Analysis
<b>RFID</b>	Radio Frequency Identification
<b>SVM</b>	Support Vector Machine

# 1 | Introduction

## 1.1 Background

An Attendance Monitoring System Attendance Monitoring System (AMS) validates whether a person has attended a meeting/lecture or not. Attendance monitoring system can be used to confirm that the right people are in the right place at the right time [14].

Recent studies have shown that there is a correlation between student's academic performance and punctuality to classes [41, 27]. Therefore, academic institutions need to be equipped with a proper way of managing attendance in classrooms and examination venues.

Attendance management can be carried out in many ways. This includes lecturers invigilating students, circulation of attendance register for students to indicate their presence by appending their signature, or automated attendance management systems. However, many attendance monitoring systems are prone to problems such as, false results, time wasting, cheat, and use of expensive materials [87]. As a result computer vision is being explored to build a better attendance management systems.

Humans are visual creatures [102]. Almost every human depends on sight to interpret the world. Most information and discoveries are products of the visual faculty. The interpretation of visuals makes the world to be informative and advances intelligence. Although sight is not the only channel to acquire information, it is regarded as the greatest source for collecting information and grow

the knowledge. It is for the same reason that scientific instruments such as microscope were created, as supplementary tools to view small objects that cannot be seen with the human naked eyes.

In the 1970's [31, 32], when the concept of computer vision was new, it was defined as the visual perception of images and videos emulating the human sight and human intelligence [102, 88]. Computer vision can also be defined as the machines that surrogate human vision's ability. Computer vision provides computers with the vision associated with human perception capabilities, understanding and interpretation of information [102].

Evidence of the pioneering of computer vision can be traced in the early 1970's as shown in Figure 1.1, in which it was believed by only few specialists of artificial intelligence and robotics when they were trying to solve the visual input problems [102].

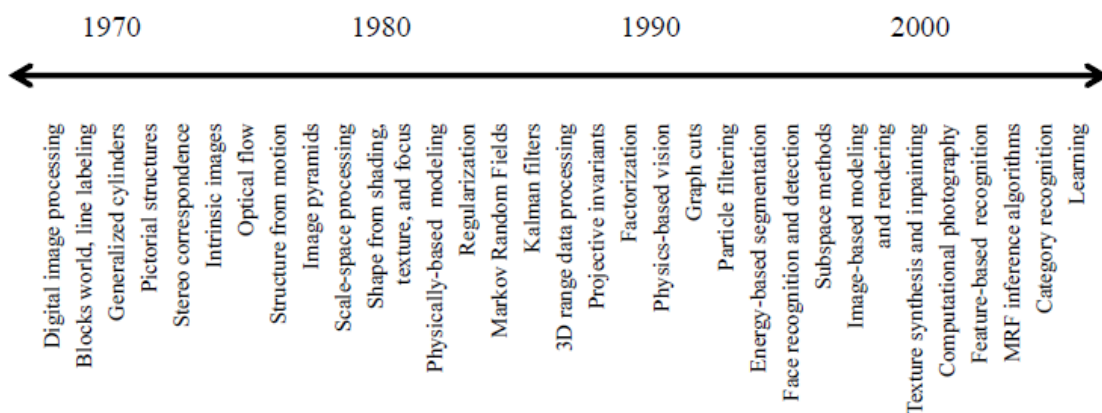


FIGURE 1.1: Computer vision history time-line [102]

The field of computer vision has matured through the years [33, 35]. In 1996 an undergraduate student, Marvin Linking was assigned a task to connect the camera with the computer and get the computer to interpret what it saw [16]. Computer Vision attributes and applications such as surveillance, biometrics, motion and object detection, machine inspection, expression detection, and object recognition can be integrated to advance the students attendance management system. In [59], machine learning and computer vision approach was used to adopt facial recognition system to automatically monitor students attendance in a classroom. Advantages of using these technologies are: automated system reduces the need for human labour, increases productivity,

can enable the processing of huge amount of data faster, can generate safe storage, and can be made more appropriate to real-time [102].

## 1.2 Overview of Computer Vision Systems and Applications

The major building blocks of computer vision applications are image acquisition, image processing and image analysis. Most computer vision systems are built from this protocol. Figure 1.2 illustrate the Computer Vision System (CVS) where the training and recognition process are shown.

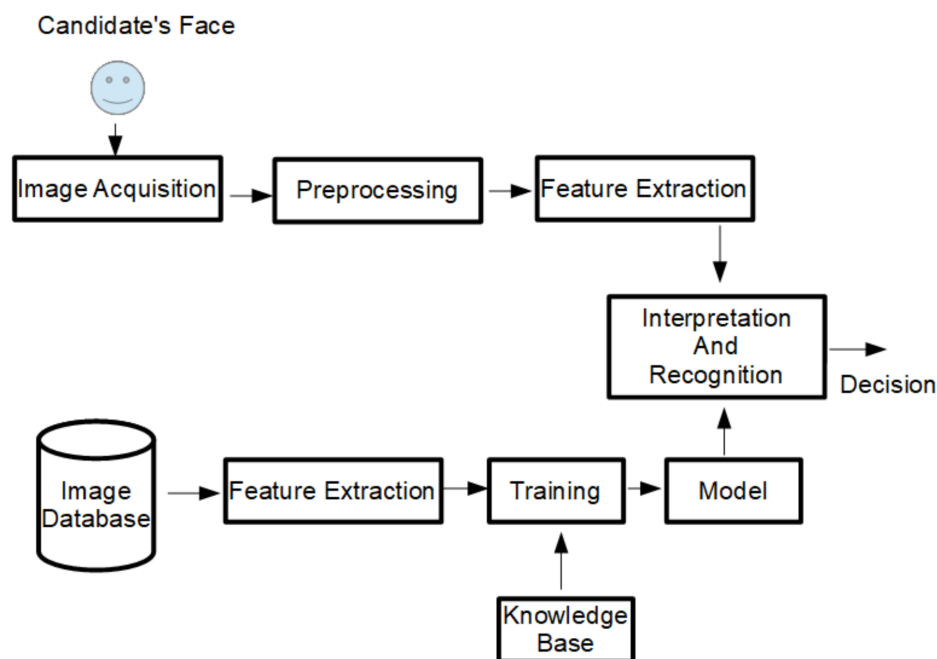


FIGURE 1.2: Generic architecture of computer vision system using face recognition

### 1.2.1 Image Acquisition

Image acquisition is the process of locating and capturing the desired image. In the case of class attendance monitoring this process will further to search faces in a given frame. This process is mainly focused on locating the object of interest. In this project images will be captured using cameras and converted into a grayscale image for further processing and analysis [100, 113, 31, 106].

### 1.2.2 Image Processing

The second phase of computer vision is image enhancement. Most computer vision algorithms such as computational photography and object recognition require data to be pre-processed in order to be enhance the phenomenon targeted in the image [102]. Images acquired from the first stage can contain some defects. Hence, the imperfections in the raw image have to be reduced for better analysis of the image.

Images can be affected by exposure, color balancing, image noise, unbalanced light, rotations, low sharpness, and blurriness [100]. Defect can be caused by the quality of the camera or some environmental effects such as a drastic change in light. The major challenge is that most images captured always have noise that cannot be analyzed with naked eyes. Therefore, it is important to always enhance an image before processing [100, 9, 102].

### 1.2.3 Images Analysis

Image analysis involves feature extraction and feature classification of a given data. In feature extraction, the best features representing an image are selected [9], while feature classification clusters homogeneous feature separately [102]. These features can be used to distinguish the images. Features can also be used for interpretation, classification, recognition and decision making between images. Object recognition performance heavily depends on feature extraction and classification algorithms used [100].

## 1.3 Motivation

Recent studies have shown that students attendance correlates with the academic improvements [53]. The moment students are in a class listening to a lecturer, writing notes, and participating on the lecturers lessons, leaves knowledge in their minds. Therefore, developing a reliable and accurate attendance monitoring system to record student's presence through the entire lecture, could be of a great help to enhance the quality of academic deliveries.

The use of facial recognition has the potential to improve the automated attendance management systems. It is well documented that face recognition has improved significantly [85, 87, 68, 22, 52, 19]. One of the most important characteristics of facial recognition based attendance

recording is their non-invasive nature. They do not require physical interaction between the user and the system. Face serves as a human primary identification [59]. The main advantage of a face recognition model is that it can detect and recognize many faces within a short period of time.

On the other hand, with the advanced and effective feature extraction algorithms and classification available today, the chances of false positives and negatives results are slim [102]. The nature of cameras has also advanced significantly in terms of quality for capturing distinctive attributes of an image [49].

This dissertation leverages facial recognition system to monitor the presence of students in a classroom. Whilst the students are recognized, the system automatically populates the register with student's attendance information. This work adopts multi-camera architecture to effectively detects faces from different angles and to reduce facial occlusion. The proposed facial recognition model will use time-interval technique to periodically update the attendance recording information.

## 1.4 Problem Statement

Students can acquire more information and understand lectures better when they attend classes. Knowledge is enriched by listening. Engaging in a particular meeting can equip a person with information pertaining to the theme of the meeting. It has been established that students who do not punctually attend to their lectures showed a poor academic performance than those who attend classes [40]. This implies that the chances of students failing are less if they attend their classes. Hence, there is a need for proper attendance management systems.

One of the key factors about the attendance is the time. A considerable amount of time is consumed in classes, when teachers call names of students, trying to figure out who is or isn't present. Several frauds may go undetected without the use of an efficient system. In addition, it is disturbing to circulate a register to collect student signature, with an extra risk of some students signing for the absentees. Also, a dubious student may present a false identification and write the exam for a friend who may not have prepared for the exam. This renders the traditional paper-based method inefficient for student's attendance monitoring [85, 90].

Existing automated modalities such as mobile application, RFID and Bluetooth based-systems are also prone to false results, intrinsic models, time wasting, and expensive instruments [21, 90]. Hence, the need to improve the attendance monitoring system is necessary. True authentication and identification are important features for attendance recording.

Although, facial recognition applications can be susceptible to false positives and false negatives, invariant lighting, facial occlusion, and drastic change in facial expression [87, 85, 68]. A well-modeled face recognition system, can improve the attendance management, due to its non-intrinsic nature [21]. This non-intrinsic nature makes facial recognition biometric system to be essentially preferred when compared to other forms of biometrics.

## 1.5 Research Aim and Objectives

The main aim of this research is to design and implement an automated smart model that employs real-time face recognition to monitor student attendance without time wasting and physical interaction with the system. In order to achieve the main goal, the following specific objectives must be carried out:

- Propose a conceptual model for a smart attendance monitoring system that uses face recognition to monitor student's attendance during their lectures.
- Design a multi-camera system architecture for effective face detection, with reduced effects of face occlusion.
- Provide time integrated model that can update student's attendance information using time intervals, to track the availability of students throughout the attendance period.
- Demonstrate the suitability of the different feature extraction techniques investigated in this study.

## 1.6 Scope Of The Study And Limitations

This section elaborates the limitations of the investigations performed in this research.

The video and images dataset of students in a classroom is very rare online. The proposed model was tested with a private database of students in a classroom.

## 1.7 Contributions

The main contributions of this research are:

- Survey on techniques and existing systems for automatic attendance monitoring system.
- Proposition of a real-time smart attendance monitoring model that is able to monitor students attendance throughout the entire lecture.
- The design of the multi-camera architecture for efficient detection of students in a classroom and reduction of facial occlusion.
- Experimentation of the real-time attendance monitoring system.

## 1.8 Overview

Chapter 2 reviews the state-of-art of the facial recognition process. It elaborates on the contextualization of existing attendance management systems. Chapter 3 presents the system model and the Graphic User Interface (GUI) of the proposed model. In chapter 4, materials, algorithms, and methods such as face detection, image pre-processing, and feature extraction are presented. A full discussion on the experimental conditions, the results of the experiments from different classification algorithms, is presented in chapter 5. Chapter 6 draws the conclusion and presents recommendations for future work. Consent for the use and publication of images captured in the classroom is presented as appendix at the end of the dissertation.



## 2 | Literature Review

### 2.1 Introduction

Student attendance to lecture can have an impact on their academic's results [22]. Monitoring the attendance of students to classes is an important and challenging task. In fact, absenteeism and attendance punctuality have always been a problem in meetings. Students randomly miss classes and without a proper system in place, it is difficult to perform a reliable inventory of the class attendance. Some students dodge classes and let their friends sign the register on their behalf [27].

Inappropriate attendance monitoring applications can also contribute to attendance management problems. Unfortunately, most educational institutions are still using paper-based attendance monitoring method. This method is found tedious, inconvenient, time-wasting, and a waste of paper material [65, 85].

As the number of students increases throughout the years, it is becoming a challenging task for institutions to manage the attendance register [85]. Absenteeism has been identified as one of the roots of poor pass rate in institution of higher learning [30].

Laws and protocols encourage students to change their pattern of attending classes, but not efficient enough to curb student absenteeism from classes. Several attendance management systems have been proposed in the literature [85, 87, 68, 22, 52, 19].

## 2.2 Attendance Monitoring System Technologies

Recently, different automated technologies have been developed to manage attendance of students [85]. The following section presents some initiatives.

### 2.2.1 RFID Authentication System

Radio Frequency Identification (RFID) is a wireless technology communication [93], which uses electronic waves between the reader and the tag to authenticate stored data. It is precisely defined as the communication between the tags, antenna, and the chip/card, through electromagnetic signals [96]. Figure 2.1 illustrates the operation of the RFID system.

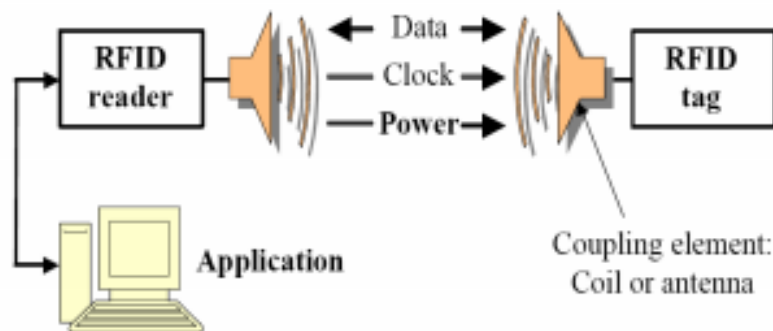


FIGURE 2.1: A generic representation of the RFID system [93]

In [96], a RFID system supplemented with fingerprint recognition model was created to monitor student's attendance. Students had to click their tags on the RFID reader to be authenticated. Fingerprint is used on mobile phones application to confirm the availability of students in a classroom. Subsequently, the register is marked positive when the RFID tag code matches with fingerprint biometric template. The system sounds robust, especially the fact that it uses fingerprint to get the identity of students. The problem with the technology is that mobile phones are unreliable, because they can get lost, lose power and get broken.

### 2.2.2 Mobile Smart Attendance System

Mobile phone technologies have a significant influence in our lives. A Multi-functional model system was suggested in [62], to manage the attendance recording of students. The system consisted of a set of functionalities such as email notification, school grade checking, weather,

health service, friendship, family affair, and emotion detection. The health attribute was created for students to express their health information on the App. The social attribute was created for social information. Students could report school bullying, violence, and any form of abuse. The system also catered for domestic issues or family matters such as poverty and care, to be addressed through the application. The application helped students to express their feelings better without fear and shyness.

Students accessed the mobile application installed on their phones using their identity number password as verification. The application collected and validated the presence of students in a classroom using a Global Position System (GPS) tracking [75]. Despite the fact that mobile phones are expensive and unreliable, the proposed algorithm saved time of calling out student's names when recording attendance [52]. It also helped students to express their private matters confidentially [62].

### **2.2.3 Bluetooth Based Attendance Monitoring System**

A comparison between two monitoring attendance techniques was conducted in [1]. The first approach was Bluetooth based, to create a communication between a control unit and the cell phones. The application installed on student's phones exchanged information with the control system via Bluetooth. The student ID (Identity number) was detected through Bluetooth when students got close a certain designed range next to the classrooms. Thus, the identified student numbers were saved on an excel format as the register.

The second approach was an off-line system, which used data access between card reader and the card [1]. Identical information about students was saved in a memory card. The availability of students was captured through the click of the card to the card reader. The second approach was found better than the first because Bluetooth-based systems can encounter connection problems.

### **2.2.4 Fingerprint Attendance Monitoring System**

The fingerprint model uses the minutiae for identification. It featured with the ridge ending, bifurcation, and short ridge [8]. These features are unique for everyone. Figure 2.2 presents the

student management system using fingerprint.

Some discussion about the fingerprint model which operated on a mobile phone to manage student's attendance was presented in [8]. Students marked the register by scanning their fingers on the fingerprint scanner embedded on mobile phones and the results were recorded and stored in the application database. To validate student's location, the application operated via a school Wi-Fi signal which is located within the circumference of the classroom. However, not every student has a smart phone. Moreover, fingerprint systems in general are regarded expensive because of installation of hardware and fingerprint reader, and students has to physically enter their fingerprints on the machine.

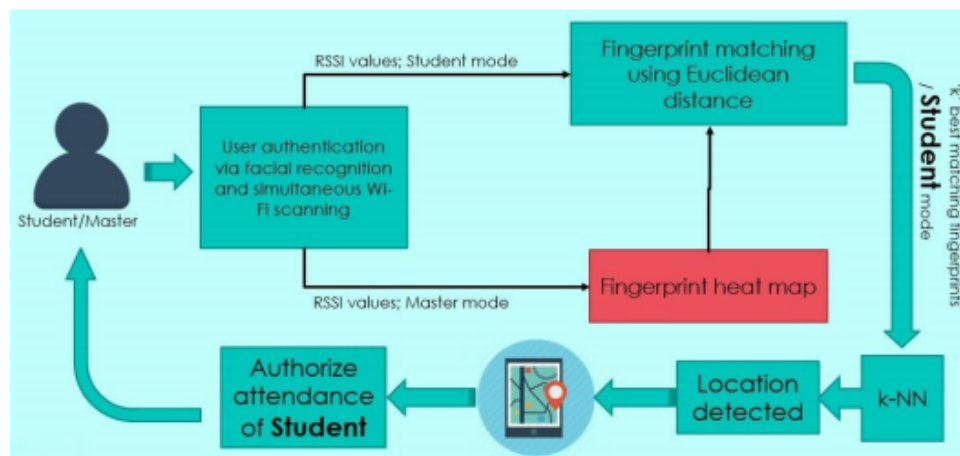


FIGURE 2.2: Example of the student attendance management system using fingerprint [8]

In [98], a payroll salary was determined by using a wireless transmitted data. The system automatically calculated the overall time workers spent in the working environment. Employees had to scan their fingerprints when they enter the working premises and when leaving after work. Unlike the paper-based signing register, the system was made to reduce fraudulent information. The system computed the difference between time-in and time-out and produced the overall time workers were on duty.

### 2.2.5 Attendance Monitoring Systems Using Face Recognition

The attendance monitoring system models are presented in different ways. In [84], facial recognition attendance management model was supplemented with audio and gender classification to validate gender of a student as to whether the recognized student is a male or female. Figure 2.3 shows students detected and recognized in a classroom.

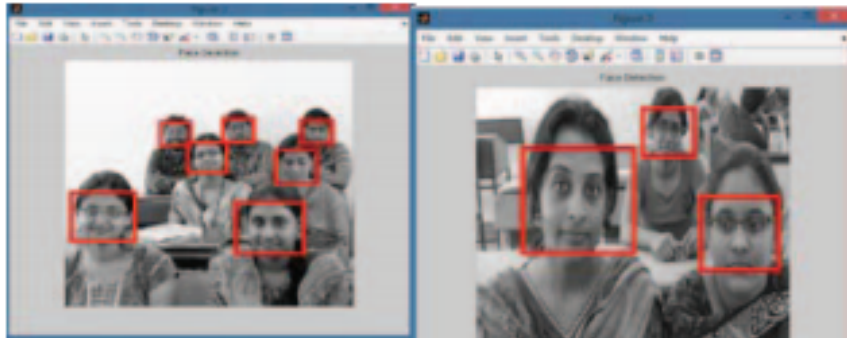


FIGURE 2.3: The overview of a face detection system [84]

After recognition an audio system is used to read names of available students for cross-checking. Although the system was real-time, the specific time of recognition was not clearly demonstrated. Students can go to a classroom just to register their presence, and leave before the end of the lecture. The efficiency of the attendance monitoring system can depend on the time which the system recognizes students in a classroom.

Unlike the previous study. Nirmalya et al. [59] modelled the student's attendance recording system that collects the attendance information at clock-in and clock-out.

Similarly, Le et al. [7] designed a multi-camera face recognition architecture that can collect recognition information using a time-slice topology. The topology was designed to capture data at some pre-defined time fragments.

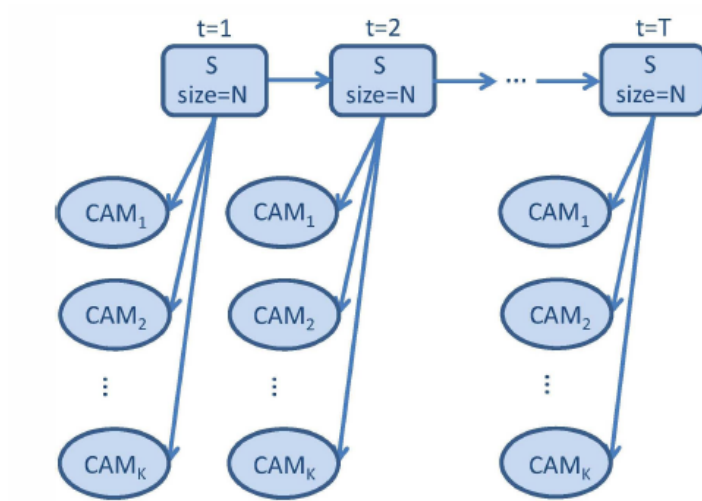


FIGURE 2.4: DBN multi-camera face recognition [7]

Figure 2.4 shows a multi-camera system integrated with the Dynamic Bayesian Network (DBN) for facial recognition. The architecture and positions of cameras are not demonstrated. Poor planning of the model can affect the system negatively [109]. However, the DBN technique was used to create data communication between different connected cameras.

A real-time automated student attendance monitoring system was conceptualized in [65]. Unlike previously mentioned system, this model is built using a single camera with connected computer vision algorithms such as face detection and recognition, motion analysis, and behavioral analysis as shown in Figure 2.5

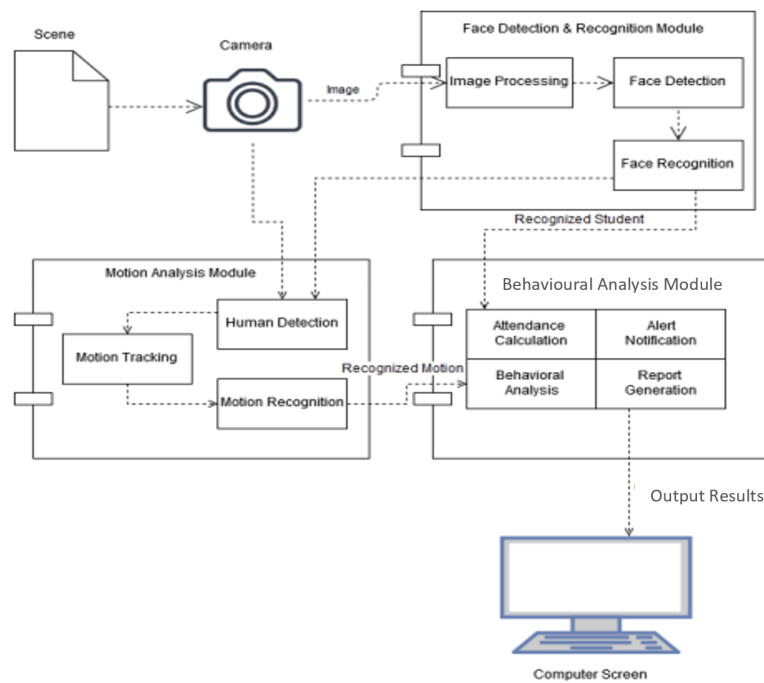


FIGURE 2.5: The conceptual model of the three-fold attendance management system [65]

Furthermore, Internet of Things(IoT) is implemented for connectivity and data transmission from stage to stage. Attendance information is acquired from the face recognition results. Motion detection analysis and behavioral detection is used to analyze movement of students in a classroom as shown in Figure 2.6

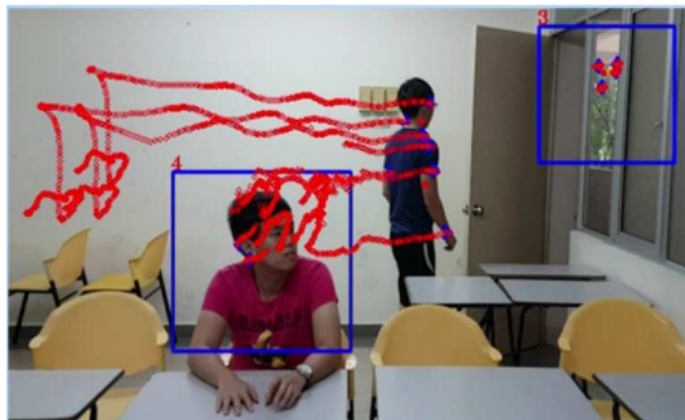


FIGURE 2.6: Face recognition system and motion detection denoted by red lines [65]

Security and confidentiality must be considered when creating a student's attendance management system. An anti-spoofing face recognition algorithm was proposed in [60]. The system tracked the unique number of eye blinks in a period for each person. The detection and recognition of

the iris provided strength to identify students. The model obtained a success ratio of 94,8%, excluded persons with eye glasses. The proposed algorithm was tested on a real time mobile video and it couldn't find the pupil centers. Further effort can be directed towards improving eye location algorithm to get accurate eye points [60].

### 2.2.6 Multiple-Camera Positioning And Collaboration

Yong et al. [109] mentioned that multi-camera technology is the key to reduce detection and recognition challenges. Multi-camera architecture provides more comprehensive data in object tracking and multi-view scenes than a single camera. It can help to reduce occlusion, chaotic scenes, and can cope in drastic change in light. However, application with many cameras can introduce more challenges. The systems have to account for collaboration between cameras, information fusion, and object matching. Nevertheless, a great system model can improve the computation ability of cameras.

The concept of full-view coverage was proposed in [64]. The basic idea was to detect faces from different directions. Regardless of the position and direction the face turned, it was always in the view point. Figure 2.7 shows the calculated camera view angle.

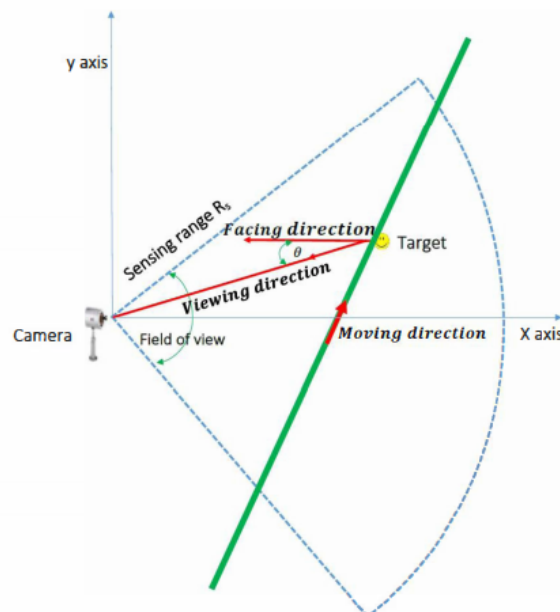


FIGURE 2.7: Computation of the target-point angle [64]



The special case of  $K$  – coverage is investigated on sensor cameras network. The angle between the view point of the target object heavily depended on the sensing quality of the cameras. Cameras moved and detected faces between the calculated view-point range. The idea of a full camera view coverage ensures efficient detection and recognition of the target object [45, 64].

Harguess et al. [42] presented a solution that can deal with the above limitations for working with frontal and full-view faces. The proposed multi-camera model detected and recognized half-face or side-view faces. The system implemented the Cylinder Head Model (CHM) to track faces in multiple cameras. The system detected half parts of the faces and merge them to identify a common face as shown in Figure 2.8.

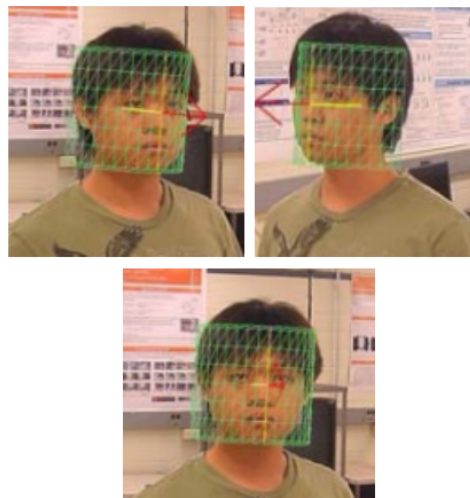


FIGURE 2.8: Face side detected and fused using a CHM [42]

It can be stated that this system requires enough storage to store the amount of training (full and half view images) data. The system can also be improved by clearly modelling the architecture of camera positioning.

Park et al. [83] proposed a technique that used Coaxial-Concentric motion camera and the static PTZ camera to achieve distance invariant face detection and recognition. The cameras were connected and were able to exchange information. Figure 2.9 below shows the schematic of the designed system.

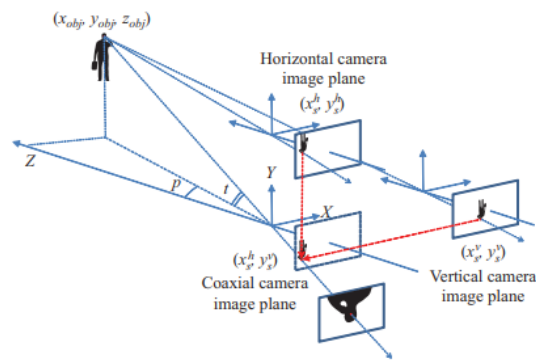


FIGURE 2.9: Calibration between two cameras, with measured angles and distance from the target object [83]

The system was tested by tracking 50 people at different distances ranging from 6 to 12 meters. The PTZ camera managed to detect good resolution images even from a greater distance.

## 2.3 Face Recognition

Face recognition has proven to be a good candidate for attendance monitoring and has a room for improvement [65, 24, 80, 84, 6]. Figure 2.10 shows stages of face recognition from face detection to classification.

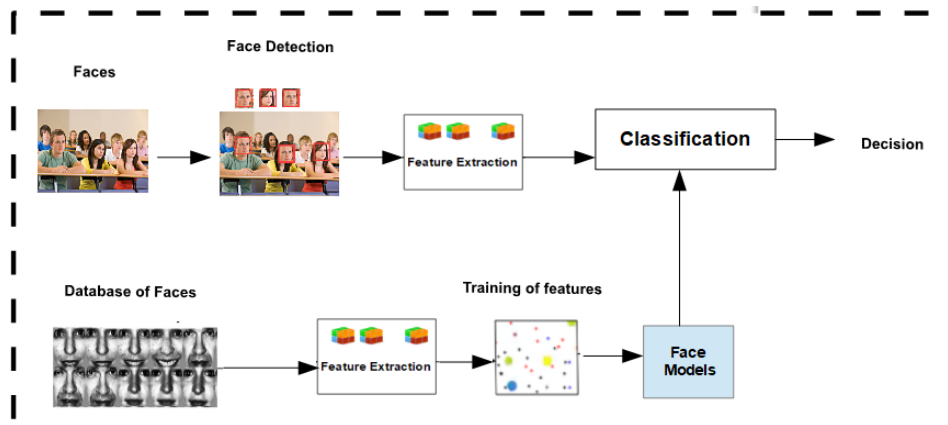


FIGURE 2.10: The flow diagram of the face recognition system adopted from [72].

### 2.3.1 Face Detection

Face detection is the first stage of facial recognition [77]. Reliability in face detection remains an issue, with challenges such as multi-view appearance, illumination invariant system, pose

estimation, quality of the camera, and resolution of the images. The detection process must be designed in a manner that the cameras would be able to capture every targeted face.

Algorithms such as multi-view appearance, Viola and Jones algorithm [106], Facial landmarks detection algorithm [113], histogram face recognition technique [5], were used for face detection.

Yang et al. [111] categorized the detection methods as feature-based, template-based and appearance [102]. Feature-based technique attempts to locate face features and face components, such as the nose, mouth, and eyes. It considers the classification of the right position of the face components [10, 11]. Figure 2.11, shows facial tracking process in image frame and video.

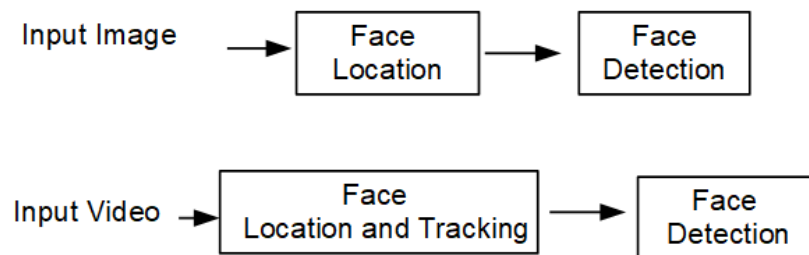


FIGURE 2.11: Input image or video for face location and detection

Facial Landmark detection technique is used to track facial feature points such as the edges of the eyes, nose, lips, and the cheek muscles [10]. In real world, this technique is used in mobile phones, laptops, security surveillance, and facial expression recognition. Barro et al. [11] proposed the Head Pose Estimation (HPE) technique. The algorithm was applied for facial landmark alignment, to locate the angle of the head pose. The head motion was tracked using the Kalman filter. The HPE algorithm uses an assembled regression trees to align the facial landmarks components in both two and three dimensions. Facial features were detected even in partially occluded environment.

Unlike facial detection algorithms which do not classify the detected frames as to whether they contain faces or not, the Artificial Neural Network ANN used strong classification layers to classify face images in the frames [88]. ANN also applied multiple retinal connected networks to examine small windows of the image, and decided whether faces exist in each window or not.

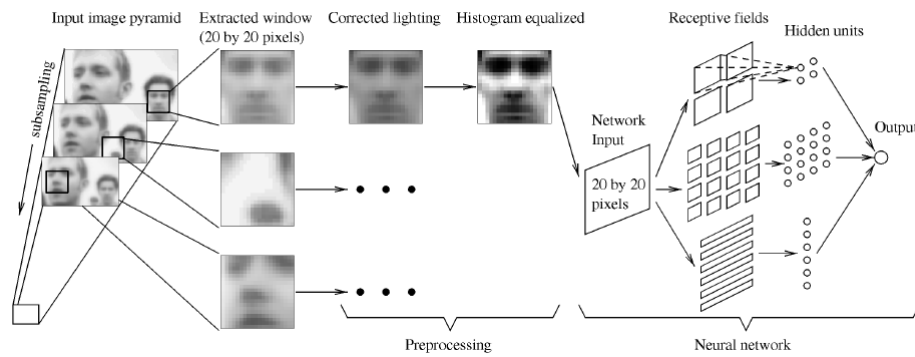


FIGURE 2.12: The ANN in face detection classification algorithms [88]

Similarly, in [107] boosting model has been used to examine false positive images and classified them separately. The Boosting technique is an attribute of the Viola-Jones face detection algorithm and it can detect faces fast and in real-time. Figure 2.13 shows the boosting classification process.

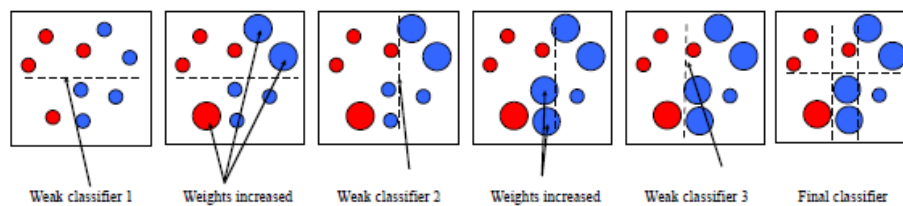


FIGURE 2.13: The series of Boosting stages [102]

The Viola and Jones face classification struggles to detect face images in an environment concentrated with light. Therefore, it is important to observe constant light environment.

Olegs et al. [78] used the local binary pattern based on face localization with the neural network to locate the position of the face. Detecting more features and multiple face components simultaneously helped to strengthen the model and to reduce false positive detection. ANN was employed to classify the face detected in frames.

Hashen [43] presented the technique that combined HSV (hue, saturation, value) color space and backpropagation artificial neural networks to detect faces. A high detection rate with less non-face images was achieved.

Detecting faces from a distance can affect the performance of detection algorithm and the recognition accuracy significantly [83]. Moreover, faces detected from a distance can experience low resolution which can negatively affect the pre-processing and recognition accuracy [71]. Jesorsky et al. [56] implemented the Hausdorff distance technique to locate and detect human faces. The method employed the shape comparison approach to detect the quality of image resolution. Edge-based and gray-scaled images were produced and were invariant to drastic change in light.

### 2.3.2 Image Pre-Processing

Face recognition can perform badly when poor quality data is used. Images have to be enhanced to prepare them for feature extraction and classification. Image enhancement accounts for image processing algorithms, such as image segmentation, smoothing, cropping and gray-scaling [28].

Dharavath et al. [28] proposed a processing strategy to modify images before feature extraction. Histogram Equalization was applied to normalize non-uniform contrast on the images. The Cumulative Distribution Function (CDF) was also used at each image pixel value to equalize the pixel values over gray levels. The contrast before and after processing was improved. Since the histogram equalizer can add great noise on the images [9], to de-noise the Gaussian noise; a Low Pass Filter (LPF) was applied to reduce high frequency information. After feature extraction the recognition results for processed data was 92.62% compared to 62.77% recognition accuracy of raw data.

Galigekere et al. [37] proposed an interpolation based approach to equalize variation of the illumination in the images. The adopted approach allowed for both synthesis and analysis of images in different light conditions.

In [112], a model used the Sparse Low-rank Component Base Representation (SLCR) to solve uniform light images was proposed. The sparse representation classification (SRC) struggled with low quality data and noise. The SLCR model was designed to solve the problems of SRC and CFR classification by creating a dictionary of low-rank images with non-low-rank images. The features extracted from the dictionary was normalized, it contained features from the low-rank component and the non-low-rank images.

The facial recognition process is usually done through two dimensional faces. This can limit the system because the whole face, which is the blue print of the face cannot be entirely analyzed [67]. Lu and Jain proposed a 3D face recognition model in [67]. The 3d face structure was achieved from different poses through facial segmentation, face masking and pose angle quantization [67]. Figure 2.14 illustrates face detection from different dimensions.



FIGURE 2.14: Pose angle estimation [67].

The model focuses on automatic points feature extraction and Head Pose Estimation (HPE) in the presence of large pose variation. The extracted features are used for face alignment in the three-dimensional space. Thus, the fully automated three 3D facial recognition model was developed and evaluated [67].

### 2.3.3 Feature Extraction And Classification

Feature extraction is the process of extracting image characteristics for further classification, analysis and identification. There are two types face feature extraction techniques, global and the local feature extraction methods. Both techniques work well depending on the state of data used and the type of experiment conducted [2].

### 2.3.4 Global Feature Extraction And Classification Methods

The global features such as Principle Component Analysis Principle Component Analysis (PCA) [104], Linear Discriminant Analysis Linear Discriminant Analysis (LDA) [12], Independent Component Analysis Independent Component Analysis (ICA), Discrete Random Transformation DRT, and the Gabor Wavelets are among the most used feature extraction techniques used in the field of computer vision and image processing. Global descriptors can speculate the whole

image with a vector, and include shape descriptors, texture features and contour representation [2].

Face recognition based on Independent Component Analysis ICA [101] deals with generalization of high-order statistic, because information can be found in high-order interrelation than in the second-order. The ICS produced feature vectors that has a high statistical independence [15].

Guoyi et al. [115] applied the matrix analysis technique proposed in [63] to model facial recognition and facial expression detection. The technique is not different from the PCA and LDA feature extraction methods. It uses non-negative constraints [63] to perform facial expression recognition with an accuracy of 66.2%.

In [70], a Neural Network for Sammons Projection [91], a nonlinear discriminant analysis (NDA) network, Linear Discriminant Analysis LDA network, nonlinear projection (NP-SOM) based on the Kohonens Self-Organizing Map, and Principle component analysis PCA were applied. The Sammons algorithm lacked the ability to project new dimensional data from the old training set. Hence, the unsupervised backpropagation model has been adopted to train the feedback neural network to overcome the limitation of the Sammons algorithm. The behavioral results of each model were studied and recorded according to the dataset used. The common rhythm about the models was that they all used an adaptive learning model technique, which enabled the system to adapt to change in environment [70].

Instead of using the distance classifier, Agarwal et al.[4] applied artificial neural network because of its ability to learn from observed data. PCA was used to extract the Eigen features and was also used as a dimensionality reduction tool [104]. The data was expressed as feature vectors, known as eigenvectors [4]. The ANN clustered faces that belong together (of the same person), as the positive examples for a person's network. Faces were trained by being fed into the ANN network layers. The number of networks used were equal to the number of entities used to train the model i.e. 30 layers for 30 different people (faces). The results in [4], outperformed the compared k-means method, and the Fuzzy C-means method. A recognition rate of 97,18% was achieved using the ORL face database [4].

LDA is the statistical technique, which performs the class specific dimensional reductions. LDA maximizes the ratio between the between-class matrix and within-class scatter [2, 12, 81]. The technique is feasible; some classes cluster tightly together, while different classes are as far away as possible, from each other in the lower dimensional representation [12, 81]. LDA is the most robust and dominant algorithm for feature extraction in appearance-based methods [101, 55]. Large dataset is recommended for extraction of good feature [58]. The LDA method also uses PCA for obtaining the tolerable data dimension [12]. Drastic change in light, different pose, image occlusion, and trivial change expression, are factors that limit the recognition accuracy of the face recognition process. Hence, this can cause the process of face recognition to perform poorly. Although, there is a need to thoroughly pre-process the data before using it, LDA can cope with unprocessed data [12].

### 2.3.5 Local Feature Extraction And Classification Methods.

Local features is robust against monotonic shadows and drastic change in light [2]. The LBP and the histogram features were the common local features types used.

Nazari and Mahammad [76] designed the model that combines the global and local Gabor feature types. The face image was divided into the four equal sub-regions. Due to the trivial change of the face structure, the local Gabor feature was used to each sub-region and to obtain invariant features. Contrarily, the global Gabor features were extracted from the whole face as a unit. The features were then combined to form a single feature vector. PCA was used to reduce the data dimension, while the KNN and the multi-SVM were applied to perform feature classification. Positive results were observed against independent global Gabor feature extraction method and the G-2DFD feature-fusion face recognition [76].

Gabor filter is one of the high performing methods in for texture segmentation [108, 26]. Wang et al. [108], applied the multi-channel Gabor filter to extract the feature vectors which represents the texture features. This generates data with huge dimentionality. To deal with this, Pulse Couple Neural Network (PCNN) was used to classify images and eventually segment the texture images. The multi-PCNN increased the processing speed and lowered the dimensional space.



The common problem in image processing is that extracted features are vulnerable to change in light and expression [9]. Therefore, it is important to use a strong edge and structured features to reduce the effects of change in facial structures, and illumination conditions. HOG and the SURF features were applied to the AT and T [73] database, for face recognition [108]. The texture features extracted with the two algorithms were robust against monotonic light. As with the LBP, the HOG features have large inter-class variance and a small intra-class variance [97]. The combination of the HOG, Spatial Histogram, and the normalized SIFT techniques produced the invariant key-point descriptor [29, 26]. The histogram provided the local feature that is translation invariant, and thus the texture analysis became robust against change in light.

The LBP model is a simple model to compute. It is recommended because of the texture detection ability. It expresses the texture information including the light and dark points, the edges and the distribution results [66]. LBP is invariant to rotation and light reflection, meaning that the features are less affected by bad lighting changes [66]. LBP has tolerance to monotonic illumination, which makes it invariant to monotonic gray changes [66, 5, 79].

A comparative study between the combination of the local and global features has been proposed in [39]. The local features were computed using the Multi-Scale Block Local Binary Pattern (MB-LBP). The global features were constructed with PCA. Global features were integrated using the lower dimensional algorithm. Different classification models such as the Mahalanobis cosine distance, Euclidean distance, and the cosine distance were used to predict the similarity between the trained and the testing data [39]. The MB-LBP produced high recognition accuracy using the Mahalanobis cosine distance. The recognition accuracy of MB-LBP increased with an increase in training images.

## 2.4 SUMMARY

The literature evidences that the solution to the traditional paper-based attendance management procedure is the automated attendance monitoring systems. The afore-mentioned problems of the paper-based methods can be addressed by different technologies such as mobile technology, Radio Frequency Identification (RFID), card system, and biometrics systems such as fingerprint,

iris, and voice recognition. Automated attendance monitoring algorithms can collect information easily and at the faster pace. Among the automated attendance monitoring algorithms, facial recognition biometric provided a better attendance management system [85, 61].

# 3 | System Modelling And Design

## 3.1 Introduction

In this research, a multi-camera system has been chosen to enable a large-scale view. Three cameras are installed in a classroom with a fixed/observed light. Detected images are stored as the training dataset for face recognition. An administrator operates the graphic user interface of the system. Figure 3.1 shows the system architecture of the attendance monitoring system.

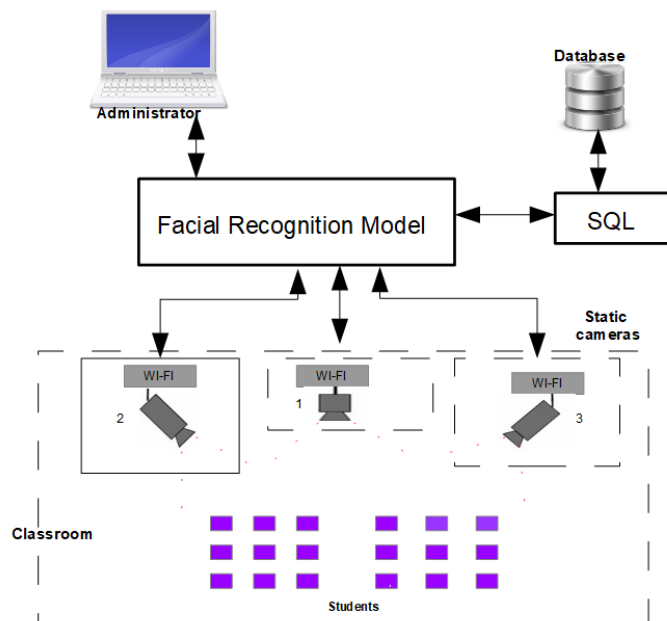


FIGURE 3.1: The interaction of cameras with the system

Cameras are connected to a computer, as shown in Figure 3.1, and exchange data through the system using wireless connection. Scenes of the classroom are captured at different time intervals and sent to the recognition system where faces are extracted.

### 3.1.1 Collection the Training Dataset

Cameras are positioned from three different positions, the right front, left front and the middle front in a classroom as shown in Figure 3.2. The cameras are static and focused on the positioned area. Each area is represented with respect to angles  $\beta, \alpha, \theta$  and each angle enclose an area of 90.

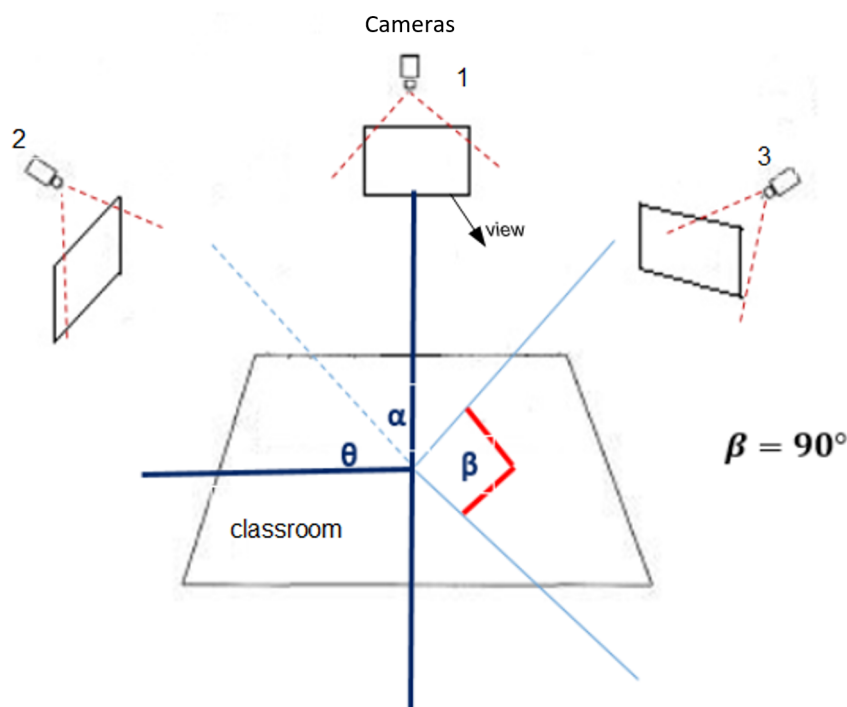


FIGURE 3.2: The conceptual area-coverage view enclosed by three static cameras

The overall view of the three cameras  $C(x)$  can be computed with the equation ?? below;

$$C(x) = \sum(\beta_x + \alpha_x + \theta_x) = 270eq3k1 \quad (3.1)$$

During the stage of image registration, for each student, collected face images differs from each other with pixels, pose, expression and light condition [9, 38]. In Figure 3.3 the cameras capture frames of images and save them to the system directory where faces are detected. Multiple faces detected from the same person are recorded with differences to a certain degree with respect to a chosen threshold  $\tau$ . Pictures are stored only when the resolution  $\Phi$  is above a chosen threshold.  $D(x, Y)$  represents the difference between image  $x$  and image  $Y$ .

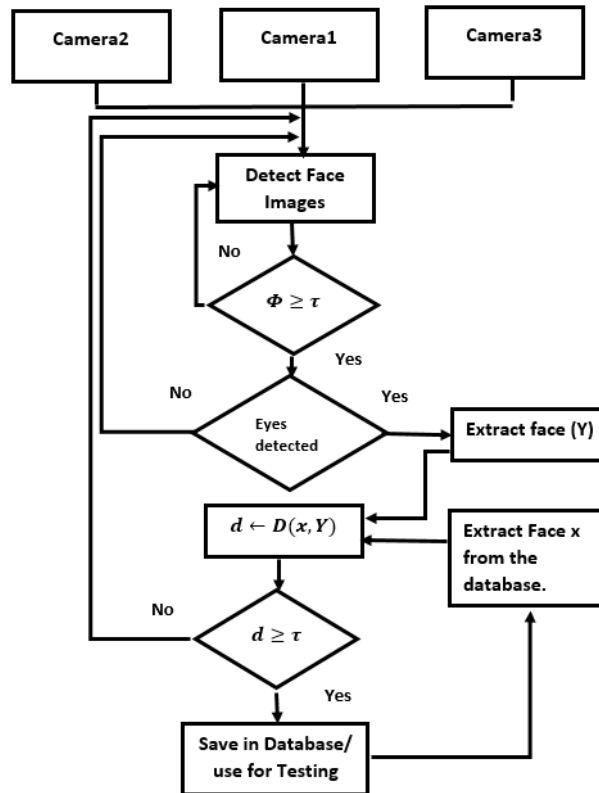


FIGURE 3.3: Collection of data during registration of faces. Student faces are registered once by a system Administrator

The performance of the facial recognition system depends heavily on the type of data being collected. Therefore, it is crucial to collect reliable data to build the training dataset. It is a good practice to train the system with different data of the same person [9] e.g. pose and different facial expression. If limited variety of data is used to train the system, this can deteriorate the recognition accuracy. Figure 3.3 is further explained in Algorithm 1.

---

**Algorithm 1** Collecting training images of the students

---

**Input:** : Database of enrolled images  $Y$ ; Threshold  $\tau$ .**Output:** :  $Y$  plus 15 different training images of learned student.

```
1: NbrImages = 0 // Number of images learned
2: for  $NbrImages \leq 15$  do
3:   Capture an image(x)
4:   Similarity =  $D(x, Y)$ 
5:   if similarity  $\geq \tau$  then
6:      $Y = Y + \{x\}$ 
7:      $SNbrImages = NbrImages + 1$ 
8:   end if
9: end for
```

---

There are many strategies that can be employed to get a good training image such as, sliding and rotating the image a bit, shifting the images by few pixels, and adding the mirror faces to the training faces. To manage data from multiple cameras, object matching and data fusion fusion technique is adopted.

### 3.1.2 System Architecture

The architecture of the proposed system comprises of three components; the front end, face recognition software and the back end. The front end part is made of cameras and the GUI. The engine of the system that performs task such as adding students, viewing results, managing data, and storage components (Database, and other information system) can be accessed through the graphic user interface. All the three components of the system are connected and share data as shown in Figure 3.4.

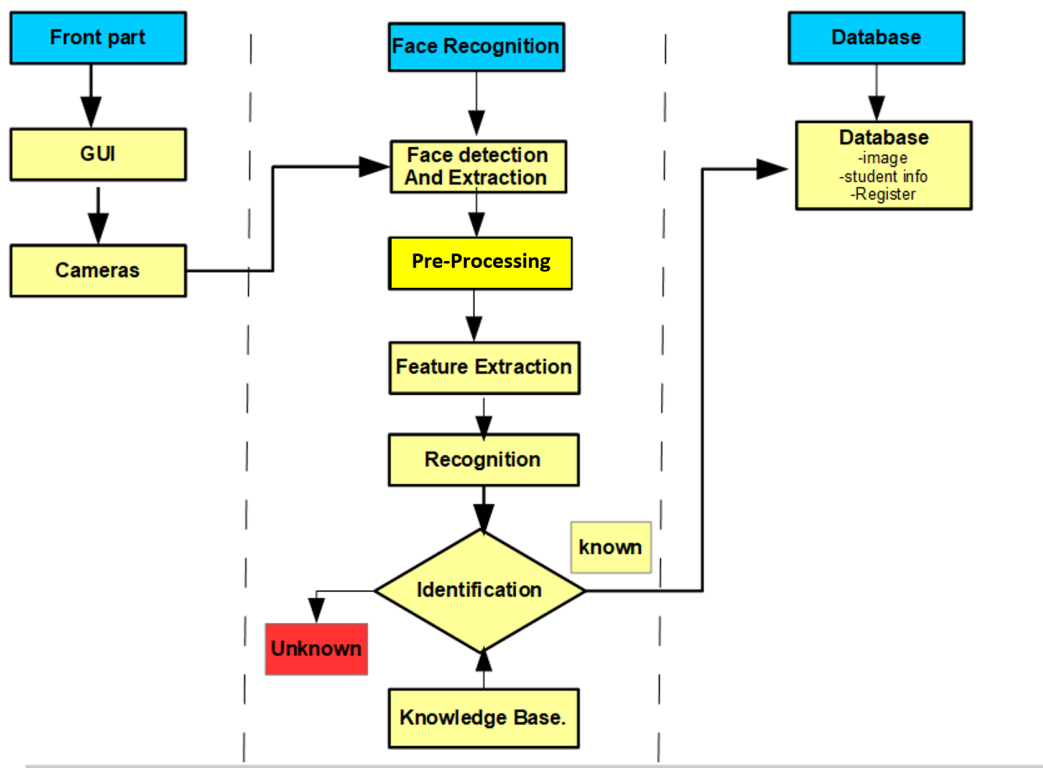


FIGURE 3.4: The three components of the Smart Attendance Monitoring System connected to one another

Figure 3.5 explains the functionality of the smart attendance monitoring system from the user's perspective. It demonstrates how users interact with the system. Students do not operate with the system physically. Faces of students are detected and recognized from a distance. The administrator is responsible to interact with the GUI. Lecturers and students can view attendance information when it is sent to them via email.

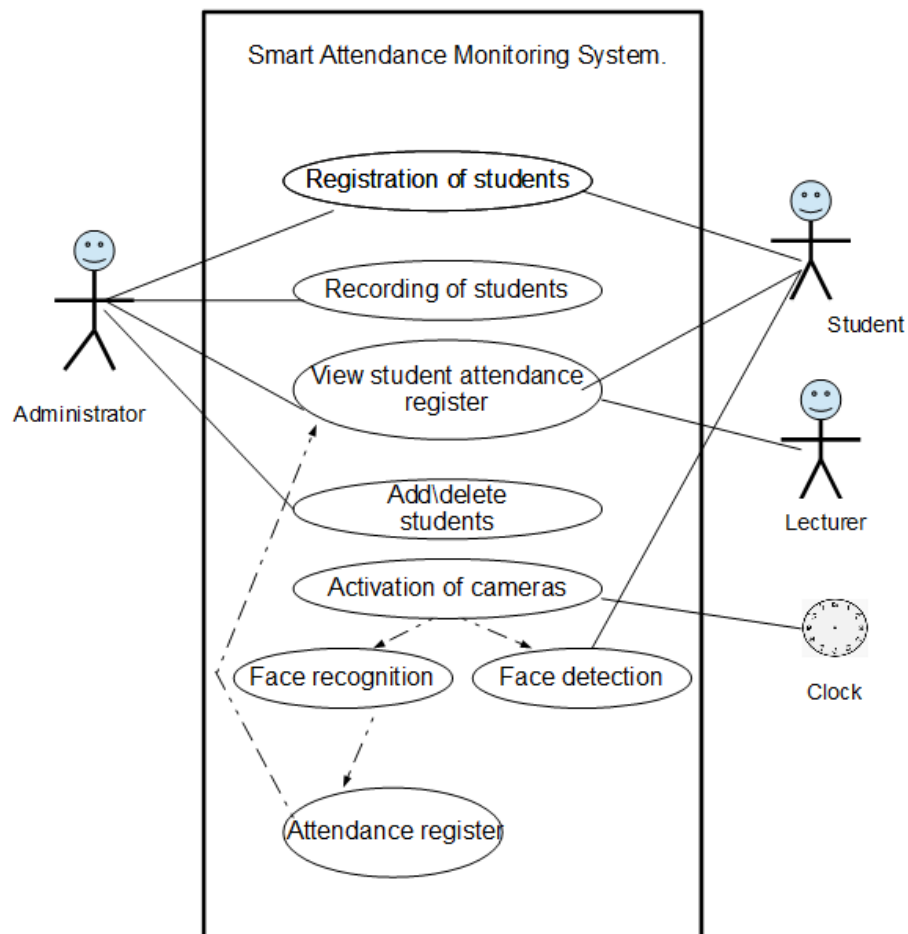


FIGURE 3.5: The use case diagram of the system

In Figure 3.6 the effectiveness of multiple cameras is measured by counting the number of students detected and recognized by each camera. Cameras capture snapshots of images at a particular time, and faces are detected from the captured frames. Extracted faces are labeled, and saved respectively in directories according to cameras. For each camera, there is a representative directory with detected face images  $img_{i1}, img_{i2}, \dots, img_{ip}$  from camera one and  $img_{n1}, img_{n2}, \dots, img_{nq}$  from camera two. Before counting detected images from each camera, detected faces are pre-processed and features are extracted. A training data is loaded and detected faces from different database are recognized. Recognized results of images from each directory are saved separately. The detection rate for each camera is judged by the number of images detected and recognized. Since the training images are saved by labels, the system can validate when one person is recognized twice, from different cameras and database.



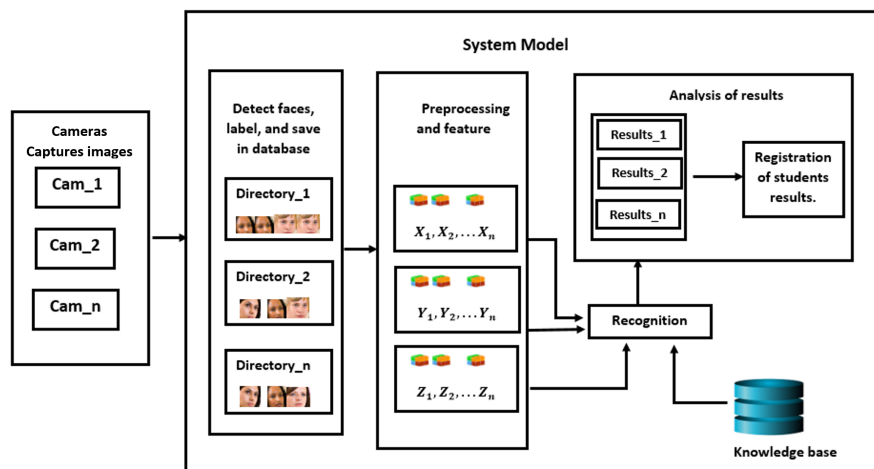


FIGURE 3.6: Detention and recognition of faces captured from multiple camera

### 3.1.2.1 The Graphic User Interface (GUI)

The first page of the GUI comprises interface to all important features of the system such as face detection, adding students in a database, face recognition as displayed in Figure 3.7. Time and date are synchronized with the attendance registration of students.

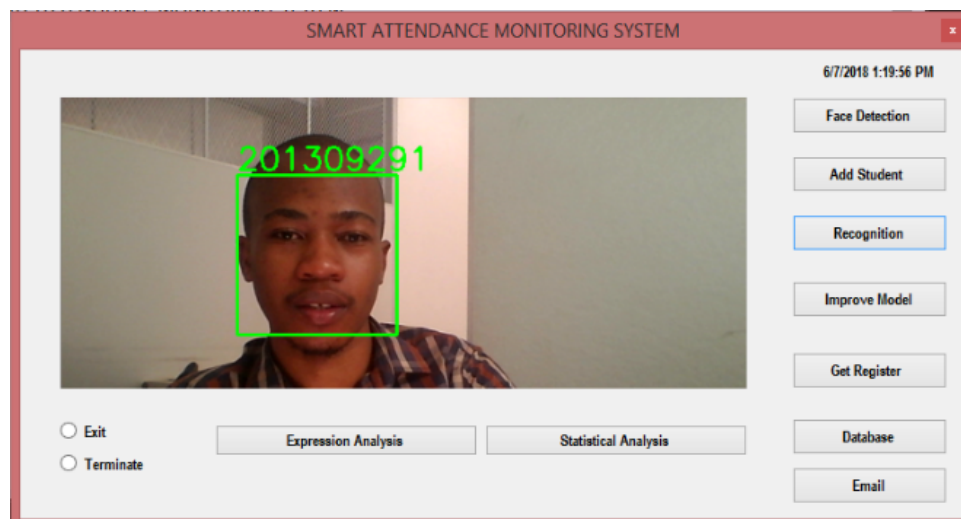


FIGURE 3.7: The system user interface

Figure 3.8 shows the student registration window. Student details such as student number, student name and surname and the student images are uploaded through the window. The process of adding students to the system is deployed only at the beginning when students register. Otherwise, it is done to update data and when adding new students.

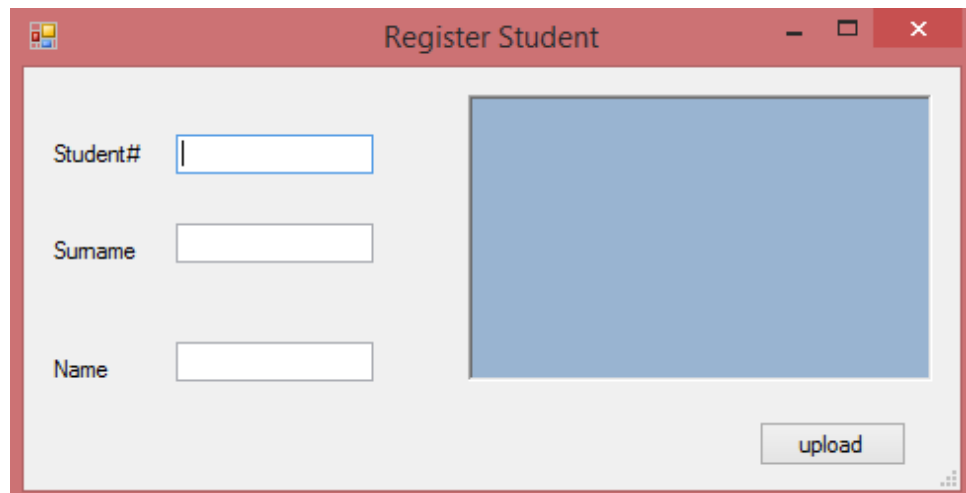


FIGURE 3.8: Form to register students information

Another important feature from the GUI is the email window, (See Figure 3.9). The system is designed to automatically send messages to students who are not punctual to their lectures. The system automatically calculates the attendance percentage of students. If the percentage is lower than expected attendance threshold, an email is generated, and sent to students and course coordinator to inform them about the attendance status of students.

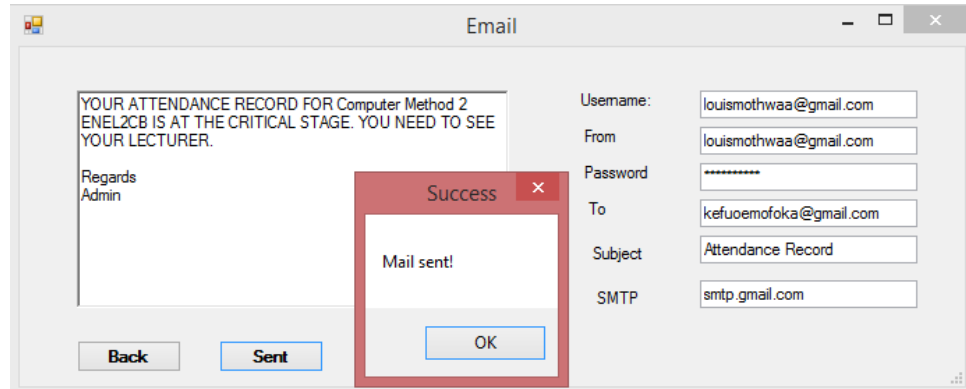


FIGURE 3.9: Email Window

### 3.1.2.2 Face Recognition-Components

Figure 3.10 illustrates the flow diagram of activities of the system.

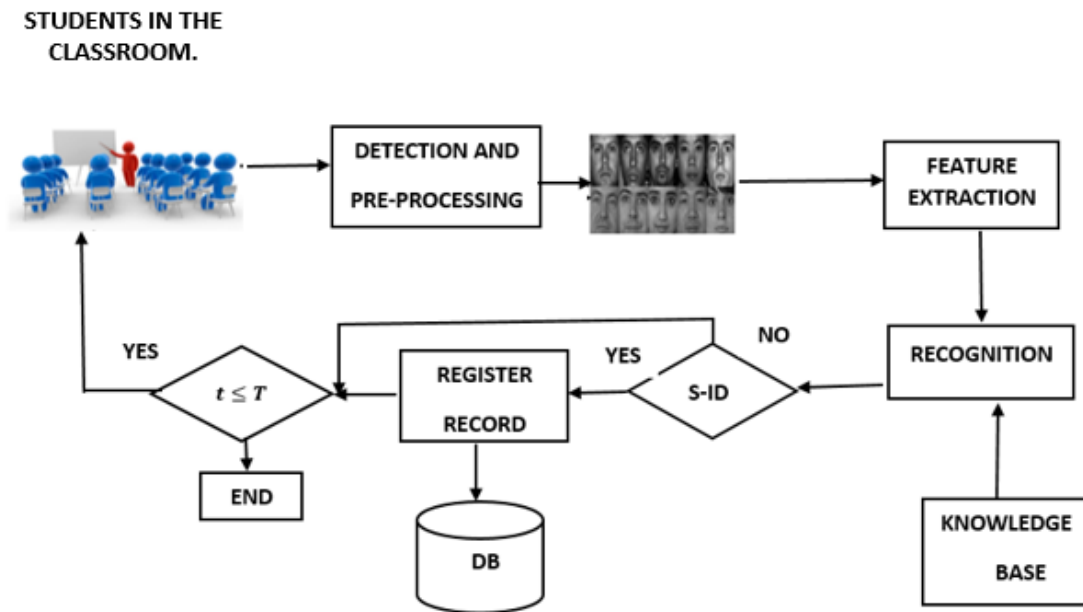


FIGURE 3.10: The flow diagram of the smart attendance monitoring system

Functions of Figure 3.10 are presented below;

- Face detection: Student faces are detected using the cameras and the adopted face detection algorithm.
- Pre-processing: Faces are enhanced to remove the image distortions and prepares them for feature extraction and classification. Training and the testing images go through the same pre-processing technique. Enhanced data can improve the recognition accuracy [9].
- Features extraction: Three feature extraction algorithms are adopted to extract features from images. Features extracted are then used in combination with data in the knowledge base for recognition and identification of images.
- Database: Training images and information about the students is saved in the database.
- Time intervals: Information recording is timely and sequentially performed via predefined time intervals  $\Delta t$ . If the starting of the lecture time is  $t_0$  and the lecture lasts  $T$  time, information will then be recorded at each  $t_i$  defined as;

$$t_i = t_0 + \Delta t \times i \quad (3.2)$$

where  $i = 0, 1, \dots, n - 1$  and  $n = T/\Delta t$ . At each  $t_i (i = 0, 1, \dots, n - 1)$  the register is updated with attendance information from the beginning to the end of the lecture, as show in Figure 3.11

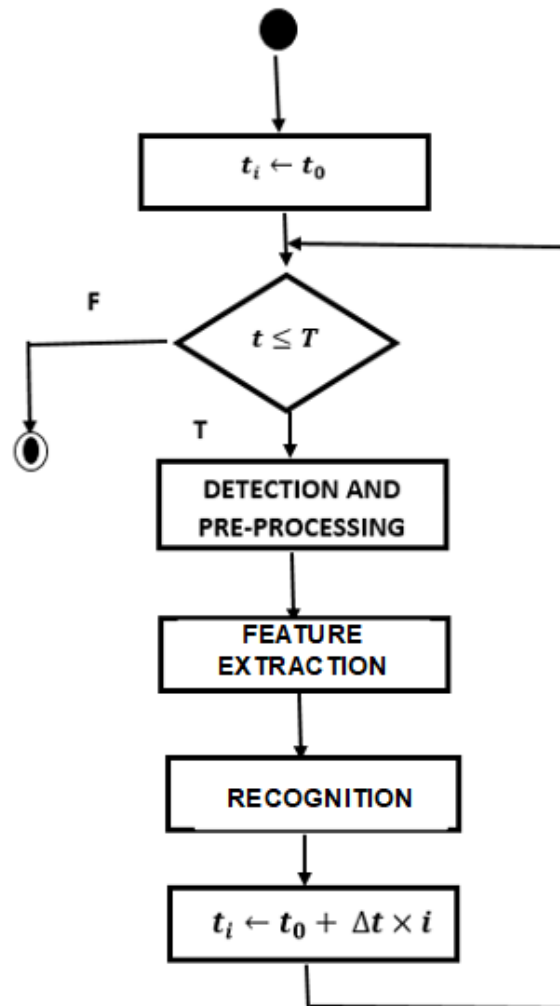


FIGURE 3.11: The internal Work flow of our smart attendance monitoring system

Recording student's attendance information in time intervals can be used to check late coming, dodging classes, and hence improving students punctuality to the classroom.

### 3.1.2.3 Back End

The system and the GUI is connected to the database and data can be entered to and accessed from the database. Training images for each student are saved in a database. To collect information for the first time, student images are captured by an administrator.

Figure 3.12 summarizes the designed database for the attendance management application. For each student, the department, school, lecturer, classroom, time and the student information such as student number, names and course enrolled is extracted to provide a reliable information for the students register.

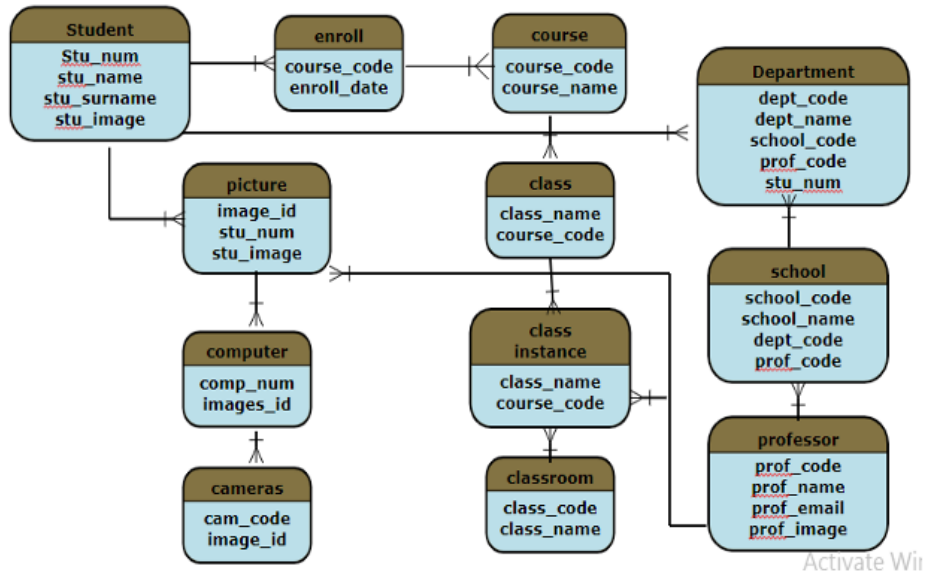


FIGURE 3.12: The Relational Entity Diagram

After identification of students faces the results are saved in a database as a register. The system is equipped with warning module that calculates the average attendance time for each student.

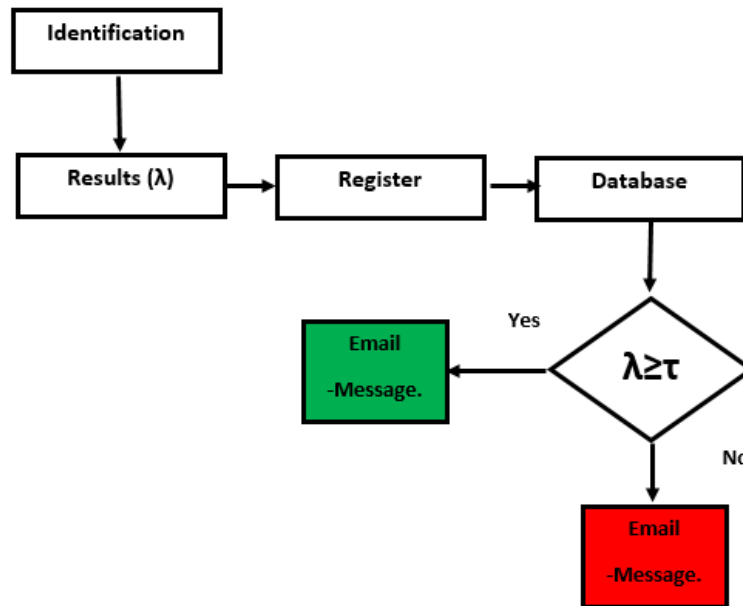


FIGURE 3.13: Background model behind attendance computation

If the overall time of attendance is less than the threshold  $\tau$ , warning is generated and sent to lecturers and students, otherwise it is sent to students to view and confirm with results on the register (See Figure 3.13 and Algorithm 2). Let  $\lambda_T$  be the number of occurrence/times students has attended lectures out of total lectures allocated for a module for a specific period (T). The total attendance is denoted by (A) as shown in Algorithm 2.

---

**Algorithm 2** Computing email address to inform about results

---

**Input:**  $\tau$  threshold,  $T$  Time

**Output:** Rates of attendance (A)

```

1: for all  $\lambda_T \in \theta_T$  do
2:    $A = \frac{\lambda_T}{\theta_T} \times 100$  //attendance score
3:   if  $A \leq \tau$  then
4:      $\leftarrow$  Generate Warning
5:   else
6:      $\leftarrow$  Email : Generate attendance confirmation
7:   end if
8: end for
  
```

---

The database consists of 12 relational tables linked together to achieve data communication. Unified Modelling Language (UML) class diagram in Figure 3.14 gives the structure of the

database used in the system. *LectureTime()* and *LectureDate()* are the functions that returns time and date of the module for the class *classroom*, *CouresName()* provides the name of the course and module attended.

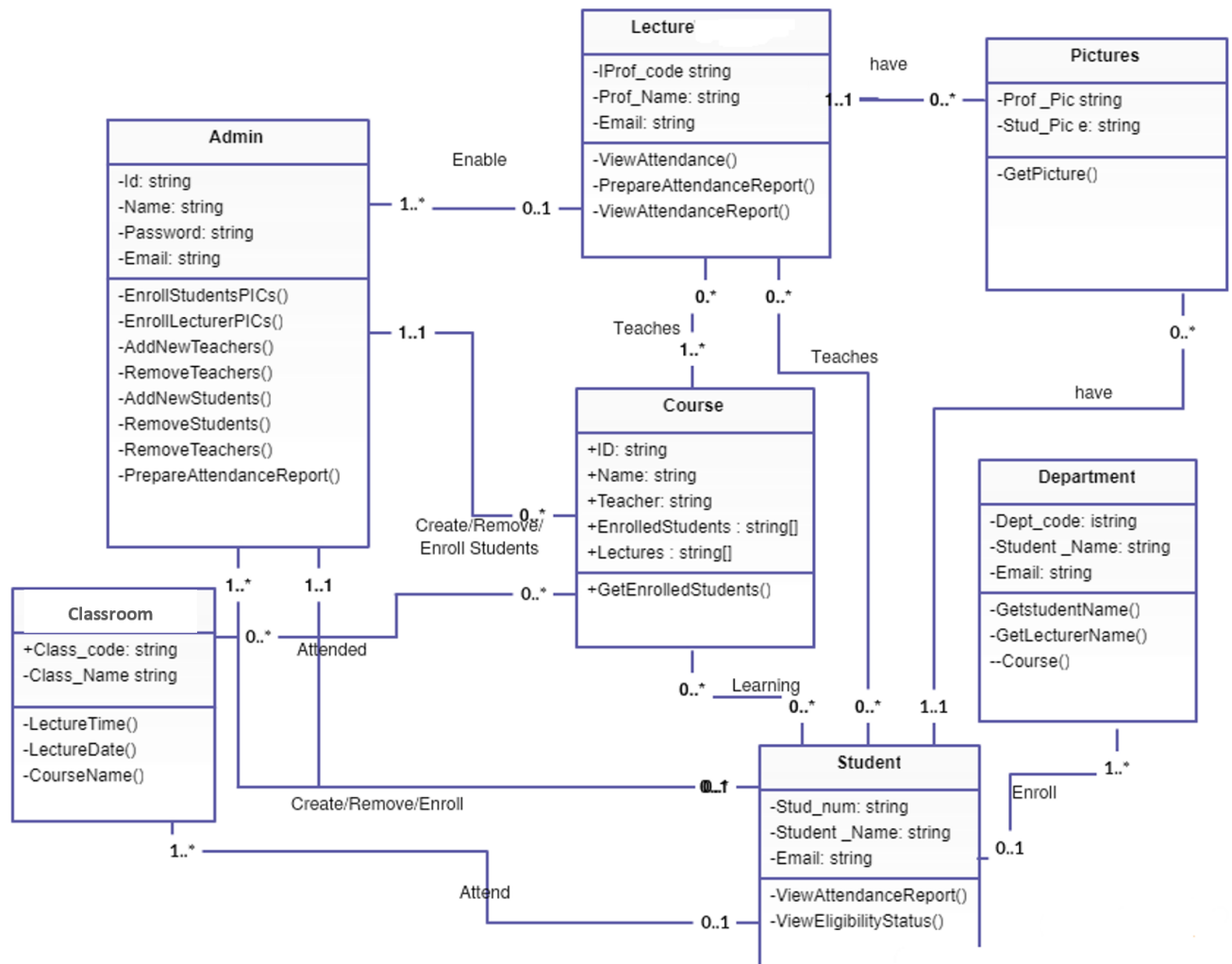


FIGURE 3.14: The UML class diagram for the smart attendance monitoring system

It can also be seen from Figure 3.14 that the administrator has control over the functions of the system. The department, course, classroom and lecturer give additional information about students and they make it easy to track the module students have attended. Students and lectures pictures are stored in a database as the training data.

## **3.2 Summary**

A system architecture that comprises of three components; the front end, face recognition software and the back end, was proposed in this chapter. The conceptual area-coverage view enclosed by three static cameras has also been designed to capture the facial images for the system. A Multi-camera model has been presented to illustrate how cameras captures images of student from different positions. The next chapter entails the process of face detection, feature extraction and feature classification algorithms.



# 4 | Methods and Techniques

## 4.1 Introduction

This chapter presents the processes and algorithms used to build the smart attendance monitoring system, the face detection algorithm, pre-processing techniques, three feature extraction algorithms and four classification techniques adopted to achieve the student's attendance monitoring system.

## 4.2 Face Detection

The Viola- Jones object detection technique has been adopted for detection of faces in real-time [54, 102, 9] (See Figure4.3). The detection algorithm comprises of the Haar-features, Adaboost, Integral image, and the cascading classifier [50, 107, 106].

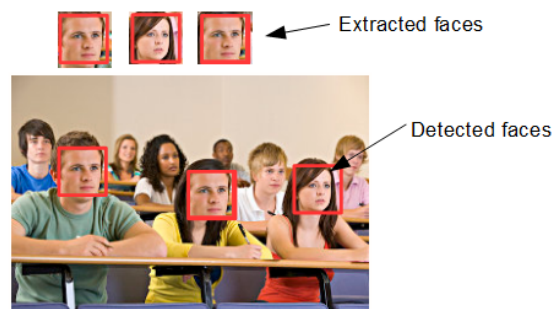


FIGURE 4.1: Rectangular red mark showing the detected faces. The faces are extracted and cropped

- Haar –like rectangle features

Haar-like features can be classified in three categories: edge features, linear feature, and central and diagonal features. The eigenvalues of the template are defined by the difference

of the black matrices from the white matrices. The Haar eigenvalues reflects the change of the grayscale image [48], as shown in Figure 4.2 (a).

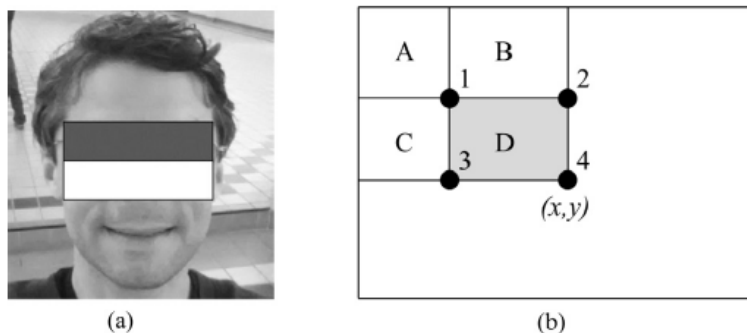


FIGURE 4.2: (a) Haar-features are represented with a rectangle pattern. (b) The sum of the pixels inside the rectangle D can be calculated as :  $4 + 1 - (2 + 3)$  [50].

- Integral Image

The computation of the integral image at the location  $(x, y)$ , is equal to the sum of all the pixel in the upper left corner of the point  $(x, y)$  as shown in Figure 4.2 (b). Let  $\Gamma(x, y)$  be the integral image of a point  $(x, y)$  and  $I(x, y)$  be the grayscale value of any pixel  $(x', y')$  in the integral image [48], then:

$$\Gamma(x,y) = \sum_{x' \leq x, y' \leq y} I(x', y') \quad (4.1)$$

- AdaBoost

The core idea of the AdaBoost algorithm is to train different weak classifiers and assemble these weak classifiers to form a strong classifier. The algorithm judges whether an object is face or not by comparing the eigenvalues of the input image with the threshold obtained from trained optimal weak classifiers [48]. Algorithm 3 presents face detection using the Viola Jones Adaboost algorithm.

---

**Algorithm 3** Detecting faces with an Adaboost trained cascade classifier [50]

---

**Input:** : An  $M \times N$  grayscale  $I$  and  $L$ -layer cascade of shift

**Output:** :  $\rho$ , set of windows declared positive by the cascade.

```

1: Parameter: a window scale multiply by  $c$ 
2: Set  $\rho [i, i + e - 1] \times [j, j + e - 1] \subset I : e = 24c^k K, k \in N$ 
3: for  $l = 1$  to  $L$  do
4:   for every window in  $\rho$  do
5:     Remove the windowed image's mean and compute its standard deviation  $\sigma$ 
6:     if  $\sigma > 1$  then
7:       Divide image by  $\sigma$ 
8:       Compute image features required by the shifted classifier at layer  $l$ 
9:       if Cascade's  $l - th$  layer predict negative then
10:        Discard this windows from  $\rho$ 
11:       end if
12:     else
13:       Discard this windows from  $\rho$ 
14:     end if
15:   end for
16: end for
17: Return  $\rho$ 

```

---

Haar face detection algorithm is susceptible to false positives and false negatives. It can detect face-like objects as faces, and thus providing false detection as shown in Figure 4.3 . The combination of face detection algorithm with eye detection algorithm has been employed to make face detection effective. It is rare that both eye and face detection can miss-detect faces.

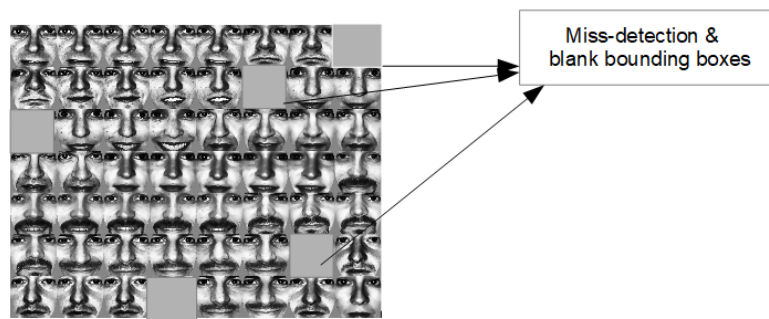


FIGURE 4.3: Face detection with Haar cascade without eye detection

Figure 4.4 shows some face miss-detected boxes around the detected face. As shown in Figure 4.4 in order to validate face detection it is necessary to combine it with eye detection. Therefore, the frame is rejected. Haar cascade eye detection algorithm has been implemented to validate whether the object detected has eyes or not. This strengthens the detection technique, and reduces the chances of false detection.

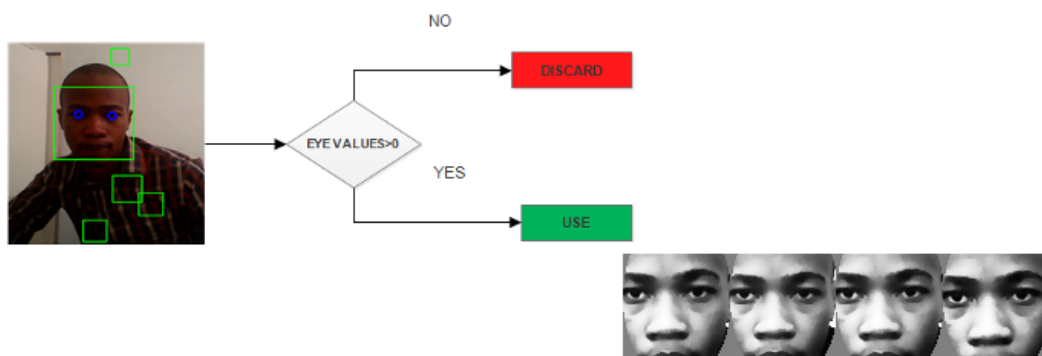


FIGURE 4.4: Validation of face detection with eye detection

The combination of face and eye detection is robust and effective. The method that was implemented in this work detects faces with no chance of false positives. All the bounding boxes with no faces detected are discarded as show in Figure 4.4.

In Algorithm 4, quality of the image resolution ( $\Phi$ ) is checked before images could be used. If image resolution is less than the threshold ( $\tau$ ), the image is not considered. The Mean Square Error (MSE) is adopted to calculate the distance between the eye region and compared to a given threshold [25]. Algorithm 4 is uses the MSE to check whether both eyes are detected and it also checks the quality of the image resolution.

---

**Algorithm 4** Eye detection

---

**Input:** : Image  $I$  of size  $M \times M$ ,  $\tau$  = Threshold,  $\Phi$  = Image Resolution**Output:** : Region of eyes detected in the face image.

```

1: Left eye:  $E_{yL}$ , Right eye:  $E_{yR}$ 
2: for all  $I$  detected do
3:   if  $\Phi_i \geq \tau$  then //check resilution of each image i.
4:     Initialize : classifier (Haar cascade)
5:     if  $E_{yL} > 0$  then //Get the left eye center
6:       Left eye detected
7:     end if
8:     if  $E_{yR} > 0$  then //Get the right eye center
9:       right eye detected
10:    end if
11:    if (  $MSEs(E_{yL}, E_{yR}) \leq \tau$  ) then
12:      Face : eyes detected
13:    end if
14:  end if
15:  else
16:    Discard image
17: end for

```

---

Including the eye detection to confirm whether the detected object is a face or not can decrease detection rate. This is because the person can arbitrarily close one eye or two eyes. Detection reliability drops when eyes are closed, and when the person has worn eyeglasses. Nevertheless, the use of eye detection algorithm in face detection cannot be ignored in this research, as it helps to scale and align the face upright, also reduces the problem of face rotation. The Haar left and right eye detection and the Haar eye-glass detection algorithm is used to locate eyes in a face.

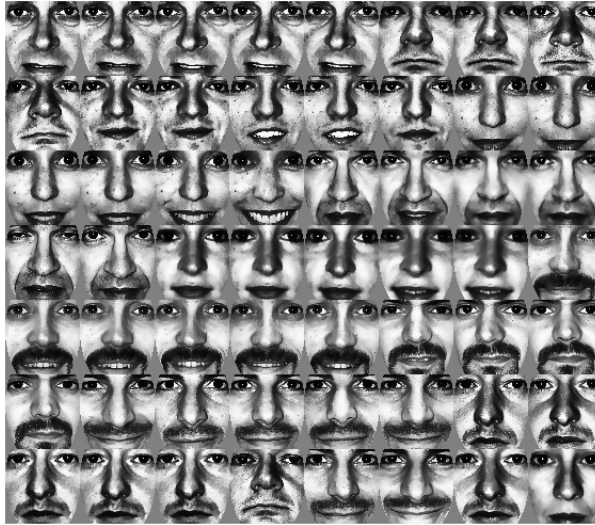


FIGURE 4.5: Dataset without false detection images. The data is enhanced with histogram equalizer, bilateral filter, and elliptical cropping [73]

Figure 4.5 shows images detected with the adopted technique in Algorithm 4. Eye detection can also help to scale images upright and provides the system with non-rotating images.

#### 4.2.1 Multi-camera Architecture

A multi-camera architecture system is adopted in this study. Cameras are installed in a classroom to create a multi-view environment and to reduce occlusion [7]. The designed system can detect frontal images from different directions. Three cameras were used to collect face images from the front, from the left, and from the right as shown in Figures 3.2 and 4.6. This is important because it can enhance the detection of faces. The effective detection of faces can thus improve facial recognition accuracy rate [7].

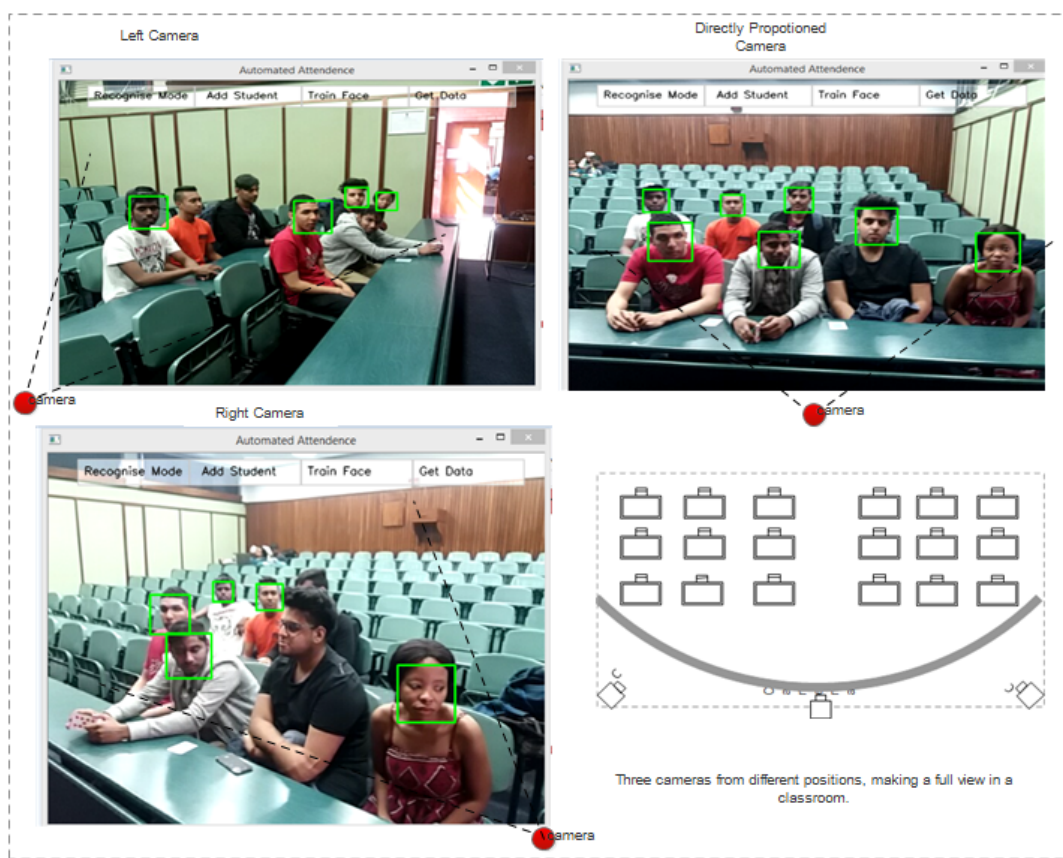


FIGURE 4.6: Live scenarios captured from the UKZN classroom. The scenes shows the importance of positioning multiple cameras in a classroom to create a multi-view environment

On Figure 4.6 (a), it can be seen that the left camera is able to detect faces that are on the left side, and Figure 4.6 (c) shows students detected from the right hand side by a camera focused on the right. The frontal camera Figure 4.6 (b), is capable of detecting faces that are directly positioned on the front. The frontal camera always detects most faces as the students spend most of their time focusing on the lecturer or the projector.

### 4.3 Image Pre-processing

Image pre-processing in this study entails performing geometrical and photo-metrical enhancement of the image [102, 9]. The main operators used are cropping, gray-scaling, rotation, Histogram Equalizing, smoothing, and elliptical cropping.

- Cropping and grayscale transformation: The process of face recognition using the eigen features works with the gray-scaled images. After detecting the images, they are cropped to the size of the face. This helps to reduce the foreign features that can be extracted with the face characteristics around the frame. The frame is reduced to the size of the face perspective [9].
- Rotation: Computing the distance between the two detected eyes and the angle of 180 degrees, helps to align and scale the image upright. To prepare the recognition process, images have to be scaled and aligned upright as shown in Figure 4.7.

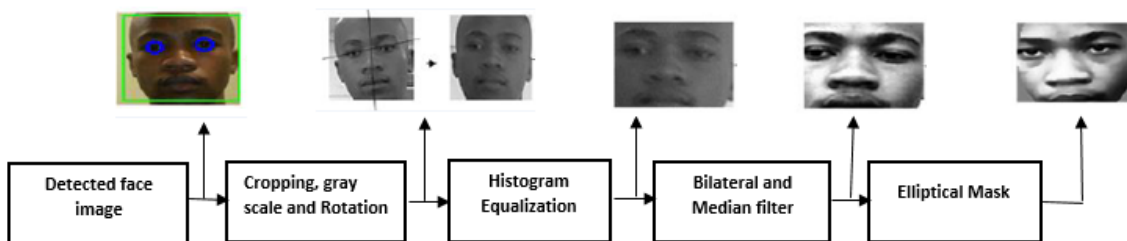


FIGURE 4.7: Pre-processing stages

- Histogram equalization ; Histogram equalizer is applied to balance contrast on the images. The entire image contrast has to be uniform for the effective feature extraction and image recognition. The algorithm is first applied to the left part of the face, then to the right part of the face, and lastly to the entire face image. It helps to reduce the light variation of the images (See Algorithm 5 and Figure 4.8) [9].



**Algorithm 5** Histogram Equalization [100]**Input:** :  $I$  of size  $J \times W$  with  $K$  grayscale**Output:** : Image with equalized contrast,  $I_E$ 


---

```

1: for all  $i = 0, \dots, k - 1$  do
2:    $H[k] = 0$ 
3: end for
4: for all pixels  $p \in I$  do //compute histogram
5:    $H[p] += 1$ 
6: end for
7:  $H_c[0] = H[0]$ 
8: for  $p = 1, \dots, k - 1$  do //Compute CDF
9:    $H_c[p] = H_c[p - 1] + H[p]$ 
10: end for
11: for  $p = 1, \dots, k - 1$  do //Normalize CDF and remap pixels
12:    $T[p] = \text{round} \frac{k-1}{JW} \times H_c[p]$ 
13: end for
14: return  $I_E = T[k_p]$ 

```

---

In Figure 4.8 , (a) presents the whole image. The left and right-parts of the face can experience invariant light balances. Figure 4.8 (b) and (c) are the results histogram equalizer to each part separately standardizes the brightness and makes the image edges to become visible. Subsequently in Figure 4.8 (d) contrast is adjusted on the combined parts of the image.



FIGURE 4.8: The three phases of histogram equalizer adopted in this work

- Smoothing : Bilateral filter is applied to reduce noise on the images. Bilateral filter is good at smoothing and reducing noise on the image while keeping the edges sharp. The filter is

used to overcome the heavy pixel noise. [9].

- Elliptical mask: Figure 4.9, (a) shows a cropped image that contains the part of the neck and other non-face portions . Figure 4.9 (b) shows the corners of the frame with masks used to remove the overlapping features around the face. Figure 4.9 (c) displays the results of the application of elliptical mask that has removed some hair, neck and some remaining background on the image. To mask an image, the black-filled or gray-filled ellipse is created around the cropped image to shade the corners [9].

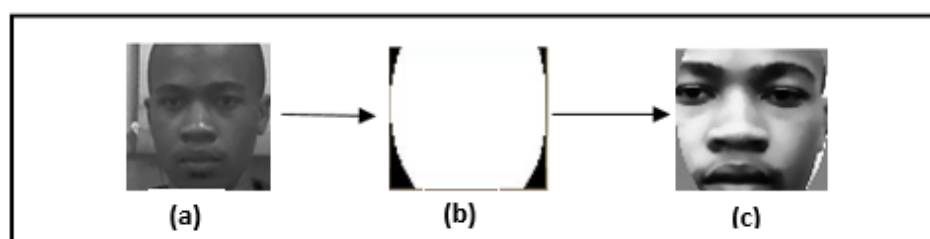


FIGURE 4.9: Normal cropped image to elliptical cropped image.

## 4.4 Feature Extraction

A relational database to systematically store the training database has been created using PCA, LDA and LBP. Features are extracted from images saved in the database. Face recognition algorithms are applied to select optimal features from each sample [103, 114].

### 4.4.1 Principle Component Analysis (PCA) for Eigen Features

PCA is used to extract eigen feature vectors in the in the face images. A face image is represented in terms of linear combination of eigenfaces. Faces are then approximated by selecting the best M-eigenfaces, which has the large eigenvalues. The best M-eigenfaces span an M-dimensional subspace which is called “face space” of possible images. The algorithm is valuable because of its speed, simplicity, learning capacity, and relative insensitivity to small gradient changes in the face images [9, 103, 13, 17, 2, 39]. High dimensional is the primary problem in image representation. For example, two images  $j$  and  $z$  with 100 pixels each spans a dimensional  $m = j \times z$ ,  $100 \times 100$  which is 10000 dimensional pixels space. What amount of dimensions would suffice for large dataset? To resolve the problem, PCA accounts for the useful information by making a decision

based on the variance in data. The process of calculating eigen features using PCA to perform face recognition is illustrated in Figure 4.10 and subsequent steps.

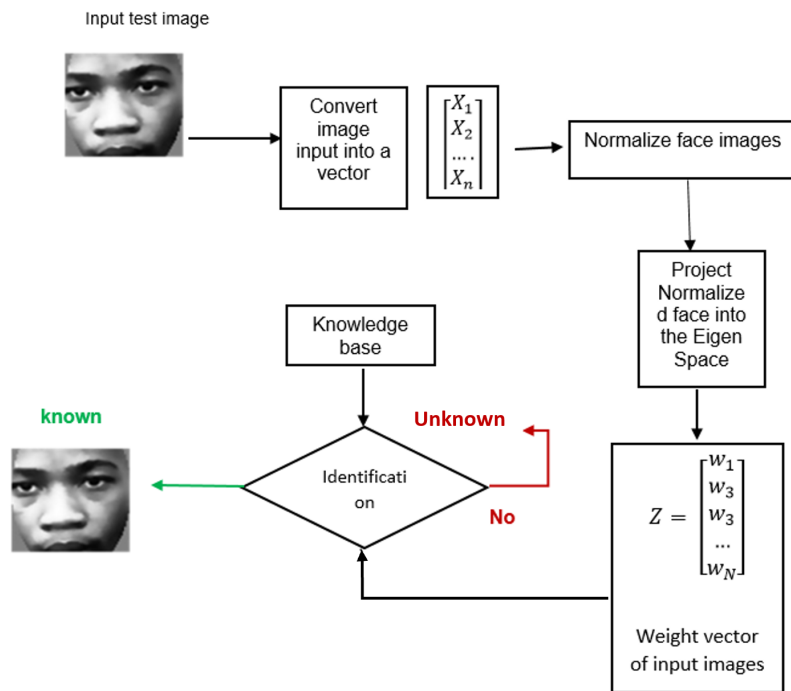


FIGURE 4.10: Summary of the PCA process

Let  $k = x_1, x_2, \dots, x_n$ , be a set of the random vector as shown in Figure 4.11



FIGURE 4.11: The training images consisting of  $n$  of images represented in the form of a vector

The average image  $K$  is calculated as;

$$K = \frac{1}{N} \sum_{i=1}^N X_i \tag{4.2}$$

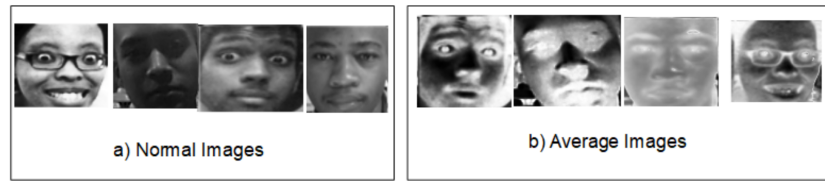


FIGURE 4.12: a) Normal images, and b) average mean images

then normalize each face  $X$  as

$$Q = X - K \tag{4.3}$$

The covariance between two sets of data reveals how much sets correlates [4]. Now, compute the eigenvector  $u_z$  and the eigenvalue  $\lambda_z$  of the covariance matrix  $C$  as

$$C = \frac{1}{N} \sum_{i=1}^N Q_i Q_i^T \tag{4.4}$$

$$= \mathbf{A} \mathbf{A}^T$$

where  $\mathbf{A} = [Q_1, Q_2, \dots, Q_N]$

The number of eigenvectors for covariance matrix can be reduced. The  $M$ -eigenfaces of the highest eigenvalue are selected to produce a complete basis for the face space. A new face  $J$  is transformed into its eigenface components [4], as

$$\mathbf{W}_J = U_K^T (J - X) \tag{4.5}$$

for  $J = 1, \dots, N$



FIGURE 4.13:  $P$  Selected eigenfaces such that  $P < M$  feature. The selected feature represents the entire training data

The weight from a vector  $Z^T = [W_1, W_2, \dots, W_{N'}]$  describes the contribution of each eigenface in the representation of the input face image, treating eigenfaces as basis set.

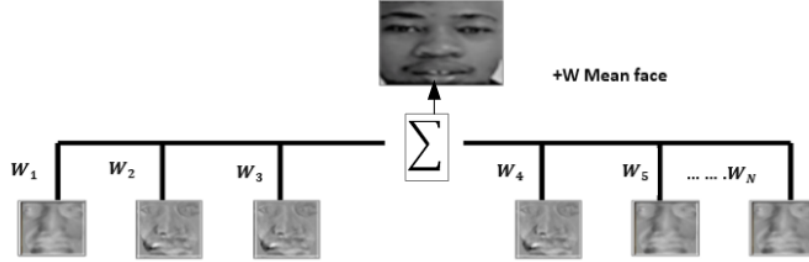


FIGURE 4.14: Each face in the training set can be represented by a weighted sum of  $k$  eigenfaces and the mean face.

#### 4.4.2 Fisherface Features

The PCA eigenvector method has almost similar computation method as the LDA technique [13]. However, LDA method shows a high error rate . In face recognition, LDA is used to perform the Fisherfaces features and for dimensionality reduction via the linear projection while preserving linear variation. In order to find the most discriminant features, that distinguish the best between classes, the Fisherface technique maximizes the ratio of the within-class to between-class scatter [13, 34]. In addition, the computation of projecting data from the high-dimensional space to the low-dimensional feature space makes the features to be robust against change in light direction and facial expression [44, 17].

The optimal projection  $W_{opt}$  maximizes the ratio of the determinant of the between-class scatter matrix of the projected sample to the determinant of the within-class scatter matrix of the projected sample [13], and can be define as,

$$W_{opt} = argmax \frac{|W^T S_B W|}{|W^T S_W W|} \quad (4.6)$$

a) where  $S_B$  is the between-class scatter,

$$S_B = \sum_{i=1}^c |N_i| (\mu_i - \mu)(\mu_i - \mu)^T \quad (4.7)$$

b)  $S_w$  represents the within-class scatter,

$$S_W = \sum_{i=1}^c \sum_{x_j \in X_i} (x_j - \mu_i)(x_j - \mu_i)^T \quad (4.8)$$

where  $c$  is the total number of the training and testing samples,  $\mu_i$  defines the mean of the class  $x_i$ , and  $N_i$  denote the number of the sample in the class  $x_i$ . If  $S_w$  not singular, the optimal projection  $W_{opt}$  is chosen as the matrix with orthonormal columns. Therefore,  $w_i | i = 1, 2, \dots, m$  is the set of generalized eigenvectors of  $S_B$  and  $S_w$  corresponding to the  $m$  largest generalized eigenvalues  $\lambda_i | i = 1, 2, \dots, m$ , i.e.  $S_B w_i = \lambda_i S_w w_i$ ,  $i = 1, 2, \dots, m$  [13].

#### 4.4.3 The Local Binary Patterns

The local features are less prone to the drastic change in light effect and image rotation [47]. LBP summarizes the local structure in an image by comparing each pixel with its neighborhood. Let  $\Gamma$  be a neighborhood of grayscale image of the size  $3 \times 3$  pixels with a center pixel  $P_c(x, y)$  having intensity  $P_c$  as threshold. Each Pixel in the neighborhood is then compared with the threshold  $P_c$  so that a binary code of  $P_c(x, y)$  can be generated sequentially. If the intensity pixel in the neighborhood  $P_i(x, y)$  is greater than  $P_c$ , the corresponding code is assigned 1, otherwise 0 as shown in Equation 4.9 and Figure Figure 4.15 [51, 46].

$$\Gamma_{(x_c, y_c)} = \sum_{i=0}^7 S(P_i - P_c) 2^i \quad (4.9)$$

where  $S_{\odot}$  is defined as,

$$S(\xi) = \begin{cases} 1 & \xi > 0 \\ 0 & \xi \leq 0 \end{cases} \quad (4.10)$$

The decimal number generated forms the Local Binary code and is used to label the given pixel [5]. Figure 4.15 shows an example of the generation of such code.

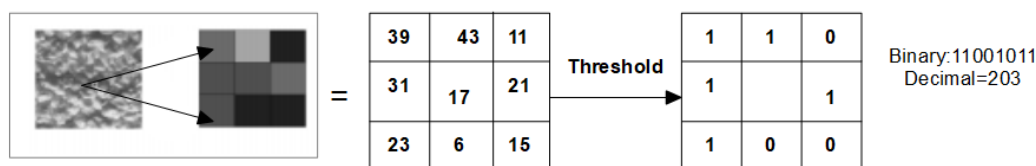


FIGURE 4.15: General computation of the LBP [5]

The LBP operator can be extended from the  $3 \times 3$  block to any neighborhood with a circle having a radius  $R$  and  $k$  samples equally spaced on it, as shown in Figure 4.16. The other pixels in the neighborhood can be denoted with intensities  $P_i \in \{P_0, P_1, \dots, P_{k-1}\}$  and the intensity of the center pixels is denoted as  $P_c$ ;

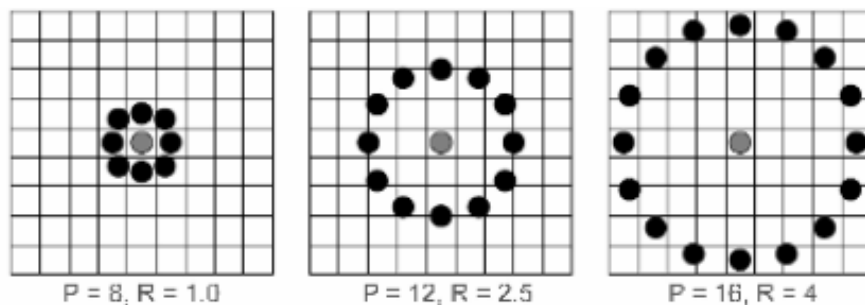


FIGURE 4.16: LBP with different neighborhood radius [46]

After calculating LBP for each pixel, the feature vector of the image can be calculated. A modified LBP operator called uniform pattern is used. The number of bit-wise transitions from 1 to 0 or 0 to 1 defines the pattern. The LBP is called uniform if its uniformity measure is less or equal to 2. For example, patterns 11111111 (0 transition), 01111100 (1 transition) and 110000111 (2 transitions) are uniform, while 10001000 (3 transitions) and 11010011 (4 transitions) are not uniform. For dimesionality reduction, histogram equalizer is used to reduce the image features from 256-dimensional decimal to 59-dimensional histogram. The reduced histogram feature contains information about the local pattern. Each uniform pattern of a histogram uses a separate bin, and one separate bin for all non-uniform patterns. The 8-bit binary number consists of 58 uniform patterns, therefore 58 bins is used and one bin for all non-uniform patterns. The global description of the face image is obtained by linking together all the regional histogram. The overall value of the LBPH  $H(k)$  can be presented in histogram as

$$H(k) = \sum_{i=0}^n \sum_{j=1}^m f(LBP_{P,R}(i, j), k), k \in [0, k] \quad (4.11)$$

where  $P$  is the sample points and  $R$  is the radius. Figure 4.17 shows the transformation from the original image to the a feature histogram.

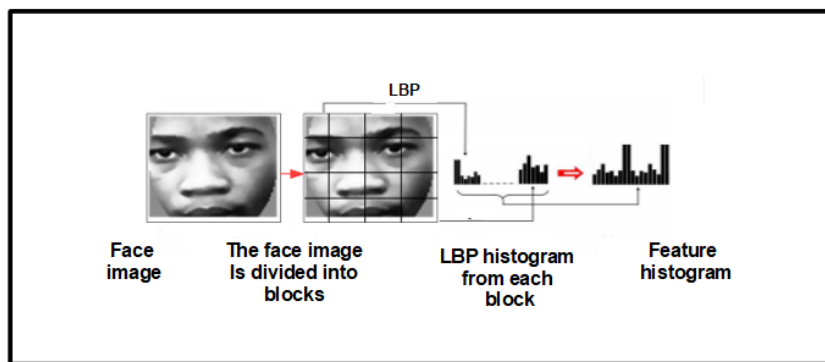


FIGURE 4.17: Face divided into regions with histogram for every region

## 4.5 Classification

A classifier is a function that distinguish data patterns [100]. Different classifiers such as Euclidean distance ED, Support Vector Machine (SVM), Naive Bayesian (NB) and backpropagation neural network (BPNN) are adopted in this research.

### 4.5.1 Euclidean Distance ED

Let  $X = (x_1, x_2, \dots, x_n)$  and  $Y = (y_1, y_2, \dots, y_n)$  be two feature points, the ED between  $(X, Y)$  is given by

$$D(X, Y) = \sqrt{\left(\sum_{i=1}^n |x_i - y_i|\right)^2} \quad (4.12)$$

where  $i = 1 \dots n$  are n corresponding eigenvalues. A face is identified when  $D(X, Y)$  is less than a certain threshold value [69].



### 4.5.2 Support Vector Machine

SVM aims to discover the ideal hyperplane separating classes by focusing on the training instances positioned at the bottom of the class descriptors. These instances of training are called vectors of assistance. Training instances are discarded other than support vectors [46, 105, 92]. For a binary classification problem. Let  $\{x_i, y_i\}, i = 1, 2, \dots, N$  be the training data, and  $y_i \in \{-1, +1\}$ , where  $N$  denotes the number of training samples  $y_i = +1$  for class  $w_1$  and  $y_j = -1$  for class  $w_2$ . Suppose the two classes can be separated linearly. This implies that at least one hyperplane defined by a vector  $w$  with a bias  $b$  can be found that can separate classes without error:

$$f(x) = wx + b = 0 \quad (4.13)$$

To find a hyperline,  $w$  and  $b$  can be estimated in a way that  $y_i(wx_i + b) \geq +1$  for  $y_i = +1$  and  $y_j(wx_j + b) \leq -1$  for  $y_j = -1$ . The two class equations can be combined to give Equation 4.14

$$y_i(wx_i + b) - 1 \geq 0 \quad (4.14)$$

The purpose of SVM is to obtain a hyperline that produce the maximum margin between classes. Therefore support vectors must be defined [105]. The support vectors are found between the two hyperplanes which are parallel to the optimal and are given by

$$(w \cdot x_i + b) = \pm 1 \quad (4.15)$$

If the hyperline parameters  $w$  and  $b$  are re-scaled, the margin can be given as  $\frac{2}{\|w\|}$ . The optimal hyperline can be found by solving the following optimization problem: Minimize  $\frac{1}{2}\|w\|^2$ . Subject to  $y_i(wx_i + b) - 1 \geq 0, i = 0, 1, \dots, M$

The Langrangian formulation can be applied to translate the above problem:

Maximize  $\sum_{i=1}^N \sigma_i - \frac{1}{2} \sum_{i,j=1}^N \sigma_i \sigma_j y_i y_j (x_i, x_j)$ . Subject to  $\sum_{i=1}^N \sigma_i y_i$  and  $\sigma_i \geq 0, i = 1, 2, \dots, N$ . where  $\sigma_i$  are Langrange multipliers. The ideal hyperline discriminant function becomes:

$$f(x) = \sum_{i \in s} \sigma_i y_i (x x_i) + b \quad (4.16)$$

where  $s$  is a subset of training samples that related to the non-zero Lagrange multipliers.

The Support Vector Machine maps the input vector  $x$  into a high dimensional feature space to generalize the above technique to non-linear discriminating features and then builds the ideal separating hyperplane in that space. The optimal hyperplane can be calculated as a decision surface [46, 105, 92].

$$f(x) = \text{sign}\left(\sum_i \sigma_i y_i k(x, x_i) + b\right) \quad (4.17)$$

where  $k(x, x_i) = \Phi(x_i)^T \Phi(x)$  is the predefined kernel function. Radial basis function is adopted and defined as follows;

$$k(x, x_i) = \exp(-\gamma_i \|x_i - x\|^2), \gamma > 0 \quad (4.18)$$

where  $\gamma = 0.125$ . The coefficients  $\sigma_i$  and  $b$  in Equation 4.18, can be derived as

$$\max \left[ \sum_i \sigma_i - \frac{1}{2} \sum_{i,j} \sigma_i \sigma_j y_i y_j k(x_i, x_j) \right] \quad (4.19)$$

$$\text{s.t. } \sum_i \sigma_i y_i = 0$$

$$0 < \sigma_i < C, \text{ for all } i$$

The parameter  $C = 2$  represents the trade off between minimizing the training set error and maximizing the margin, and  $\sigma$  is assigned different values from 0.1 to 0.8. Since SVM is a binary classifier, one-against-one approach, which is a pairwise method was used to extend for m-class face recognition. Hence,  $m(m-1)/2$  SVM classes are needed for training [46]. Figure 4.18 shows the non-linear separable data mapped into a feature using the RBF function [92].

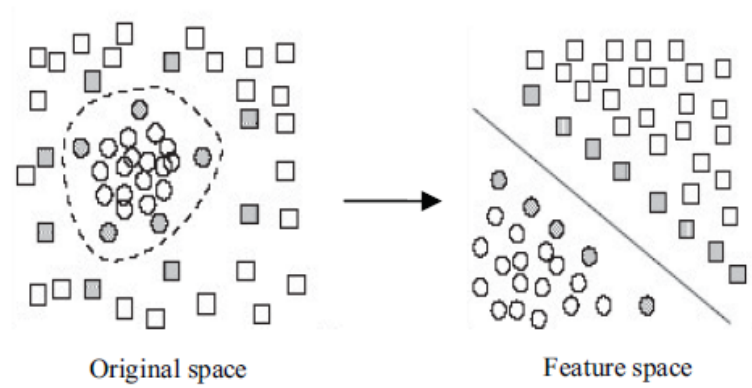


FIGURE 4.18: RBF function used to separate non-linear separable data mapped in to feature [92]

### 4.5.3 Naive Bayes Classifier

Naive Bayes classifier is one of the popular and simplest classifier [36, 94]. It can be used for feature classification in the process of face recognition. A given face is assigned to one of the  $k$  subjects, namely  $w_1, w_2, \dots, w_k$ , on the basis of a vector  $V = [v_1, v_2, \dots, v_T]$  associated with the face image. Feature vectors  $V$  is a T-dimension observation drawn arbitrarily from the class conditional PDF  $p(v, w_l)$ , where  $w_l$  is the class to which the feature vector belongs [86]. Naive Bayes decision can minimize the Bayes risk, which is expected value of the loss function [86]. The Bayes computation assigns a face to the subject  $w_l$  for which the conditional risk is minimum.

$$R(w_l|V) = \sum_{m=1}^k L(w_l, w_m)p(w_m|V) \quad (4.20)$$

where  $L(w_l, w_m)$  is the loss caused in choosing the subject  $w_l$ , when a real subject is  $w_m$ , and where  $p(w_m|V)$  is the subsequent function. The Bayes decision rule can simplify the MAP decision rule to the loss function [86, 57, 20], by assigning the input face denoted by  $F$  to the subject if

$$P(w_l|V) > P(W_m|V) \forall \neq e \quad (4.21)$$

Bayes Theorem can be used to compute the posterior PDF as

$$p(w_l, V) = \frac{p(V|w_m)p(w_m)}{p(V)} \quad (4.22)$$

Since features extracted from the class images are acquired from the orthogonal polynomial functions. Considering the given class, features can be regarded as independent. Therefore, the PDF of features given the class of face images can be written as [86];

$$p(V|w_m) = \prod_{r=1}^T p(v_r, w_m) \quad (4.23)$$

From the training face images, the class conditional mean ( $\mu$ ) of the feature elements can be estimated as

$$\mu_{rm} = \frac{1}{\lambda_{tr}} \sum_{k=1}^{\lambda_{tr}} v_{rk}^m \quad (4.24)$$

#### 4.5.4 Backpropagation Neural Network

Backpropagation neural network compares the output of the network and calculates an error measure based on sum of square differences [100]. Figure 4.19 shows a schematic of the backpropagation network consisting of three layers: input layer, hidden layer and output layer.

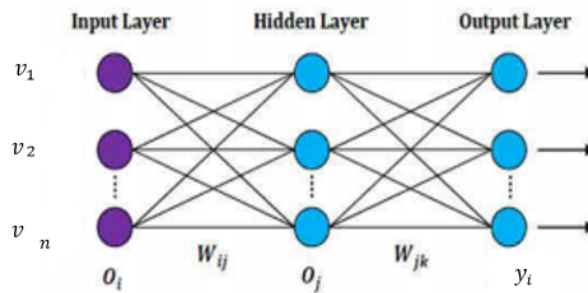


FIGURE 4.19: Schematic of a three layered neural network [82]

The input layer is fed with sample  $V = v_1, v_2, \dots, v_n$ . Each layer has a weighted connections,  $w_{ij}$  represents the weight from a unit  $j$  in one layer to a unit  $i$  in the first layer [82, 110, 18]. The input layer  $i$  consists of the input nodes and the output layer  $y$  consists of the output nodes. The hidden layer  $j$  creates an intersection between the input layer and the output layer. Each node

in the network is connected to another node, and each link has a weight associated with it. The weighted sum of all the input values forms an output value [82, 100, 89]. The output is expected to correspond with the associated images from the database.

Let  $v^i$  be the training set,  $y^i$  be an actual outputs and  $w^i$  be the desired output. The algorithm can be summarized with the following steps [110, 89, 3].

- 1) The weights  $w_{ij}^{[l]}$  and the threshold  $\gamma_j^l$  are randomly initialized.
- 2) After feeding the training dataset  $I_P$  and the output dataset  $O_p$  use Equation 4.25 to compute the output of all layers

$$y_{jp} = f\left(\sum_{i=1}^{N1} w_{ij}^{[l+1]} y_{ip}^{[l]} + \gamma_j^{l+1}\right) \quad (4.25)$$

- 3) Calculate the square error in each layer as follows;

$$Err_{jp}^{[l]} = f'(y_{jp}^{[L]})(d_p - y_{jp}^{[L]}) \quad (4.26)$$

in the  $i$ 'th hidden layer ( $i = L - 1, L - 2, \dots, i$ )

$$Err_{jp}^{[l]} = f'(y_j^{[l]}) \sum_{k=1}^{N_{L+1}} Err_{kp}^{l+1} w_{ij}^{l+1} \quad (4.27)$$

- 4) Use Equations 4.28 and 4.29 to calculate change in the weights between the output and the input

$$\gamma_{ij}^{[l]}(n+1) = \gamma_i^{[l]}(n) + n \cdot Err_{jp}^l \quad (4.28)$$

$$\gamma_{ij}^{[l]}(n+1) = w_{ij}^{[l]}(n) + n \cdot Err_{jp}^l \cdot y_{ip}^{[l-1]} \quad (4.29)$$

- 5) If the mean –square error is more than the threshold, go back to step 2 and repeat the process, otherwise stop and print the weights value. In our experiment, the sigmoid function was used as the activation function.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4.30)$$

## 4.6 Summary

This chapter discussed different static multi-camera architecture used for effective detection of the faces. It also discussed the facial detection techniques, feature extraction algorithms, and classification methods used in this study. The next chapter presents the experimental result and discussion.

# 5 | Experimental Results And Discussion

## 5.1 Introduction

In order to validate the system, it is important to measure the performance. This chapter validates the detection accuracy of multi-camera system. It discusses the relevant factors in choosing systems parameters and testing the feature classification algorithms against different feature dimension.

## 5.2 Experimental Setup

A sample of 35 students with 40 face images each were produced from University of kwaZulu Natal (UKZN) classroom. The system detects and recognizes face images from students of different skin tone (Blacks, Whites, and Indians) in a classroom. The data produced was under a homogeneous background with a constant lighting. Three static cameras placed in different positions are used for detection of images. Features are extracted with [PCA](#), [LDA](#), and Local Binary Patterns ([LBP](#)), and classified using Euclidean Distance [ED](#), Support Vector Machine [SVM](#), Naive Bayes Naive Bayes ([NB](#)) and Backpropagation Neural Network BPNN.

Figure 5.1 presents a sample of photos from the database produced from students in a classroom.



FIGURE 5.1: The sample of training images.

Figure 5.2 and Figure 5.3 show scenes of students in a classroom detected from different angles and positions. Each camera was positioned according to the structure presented in Figure 3.2. However, faces that faced/focused down were hardly detected. It shows that the multi-camera system can help to detect many faces from different angles.



FIGURE 5.2: Detection and recognition of students in a classroom.

Figure 5.3 shows students sitting in different positions with arbitrary arrangements while focusing on the lecturer. Therefore, the multi-camera provides efficient face detection despite the sitting positions of students. Some students in Figure 5.2 were not detected, yet in Figure 5.3 were able to be captured. This shows that students trivially focus their faces everywhere around the



classroom, hence multiple positioned cameras helps for detection of faces from any angle in a classroom.



FIGURE 5.3: Students faces are recognized from a different position in a classroom

The recognition results of student’s face detected by each camera are presented in Figure 5.4. As described in Figure 3.6, detected faces from captured frames are saved in respective directories according to the cameras. Cameras captured snaps of frames from different positions at the same time. The results for each camera are compared with all the testing images of 35 students available in a classroom. The made up testing images are called “virtual camera”.

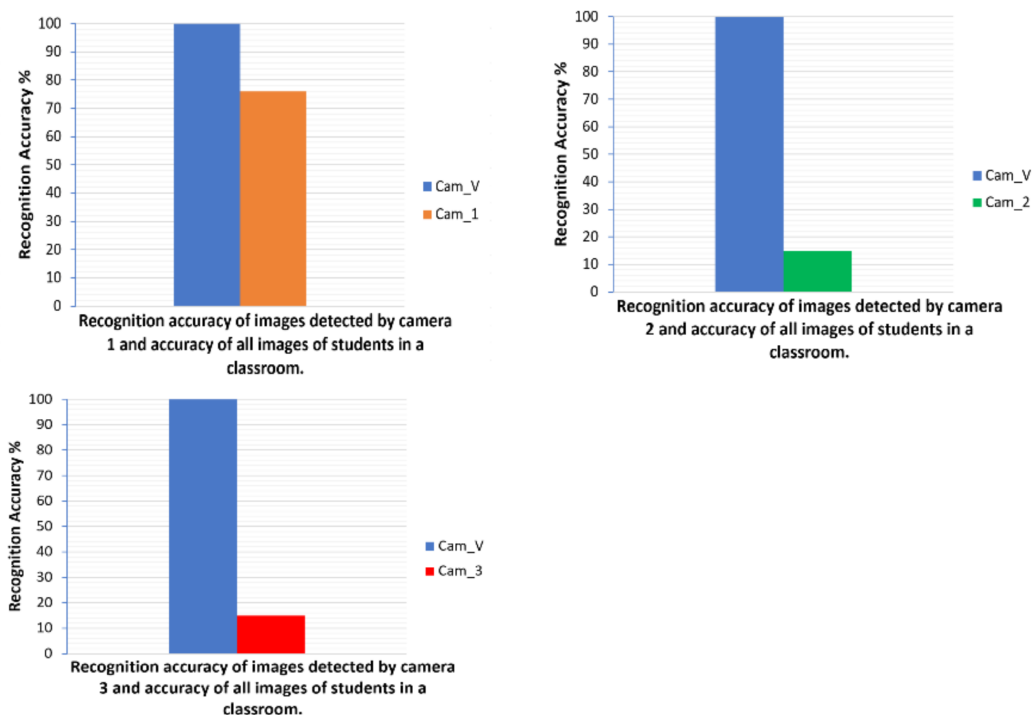


FIGURE 5.4: Recognition results from different cameras

The recognition accuracy of 76% is obtained using testing images from camera 1, and 15% from camera 2 and camera 3. The results shows that the combination of the cameras positioned according to the proposed architecture can capture students in a greater view.

In Table 5.1, facial detection was evaluated from the recognition results in Figure 5.4. From the recognized faces we ware able to track number of faces detected by each camera. Detected face were respectively saved in different databases and label according to the cameras. Camera 1 is directly positioned to the front of the classroom, it detected more images than other two cameras with a score of 85%. Camera 2 and 3 each scored the same average of 25% the combination of all cameras were able to detect all the faces available.

TABLE 5.1: Face detection rate from differently positioned cameras at a fixed time

Camera	Detection rate %
Cam1	85
Cam2	25
Cam3	25
Cam1,2, and 3	100

In Figure 5.5, facial detection was evaluated from different positions with different combination of cameras. Results from different cameras were fused to investigate the effect of multi-camera positions.

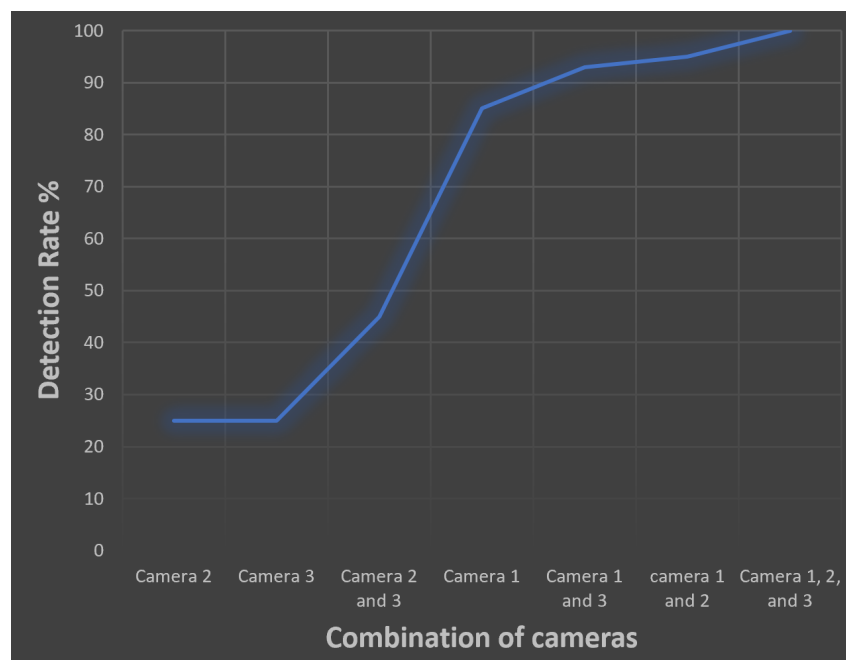


FIGURE 5.5: Detection rate in the Multi-camera

A high detection rate of 100% was achieved from the combination of all cameras. Camera 2 and camera 3 have less detection rate because they only can detect images when students faces are subjected sideways, or when students look sideways in the class room as shown in Figure 4.6. The use of multiple cameras creates multiple views and thus increases the detection and recognition accuracy.

Pre-processing methods such as, gray-scaling, cropping, re-sizing, elliptical masking, median filter, bilateral filter and histogram equalization, were applied to modify the face images. Images were cropped to the face perspective and re-sized to  $128 \times 128$ , which makes the total feature of 16384. Subsequently features were extracted and classified using the adopted classification algorithms. images were labeled and divided into the training and the testing data.

Figure 5.6 presents the results of PCA, LDA, and LBP using Euclidean distance as a classification technique. The performance is judged based on the number sample dimensions used. Results show that accuracy increase with an increase in the number of samples, with LBP algorithm score of 91%, 88% was obtained from LDA, and 85% from PCA.

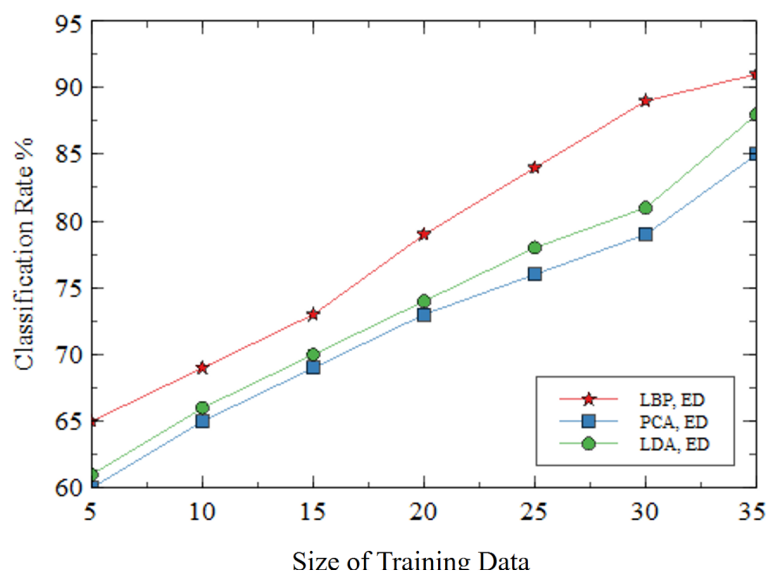


FIGURE 5.6: Classification of feature using ED

In Figure 5.7, PCA, LDA and LBP were used as feature extraction algorithms and the Radial Basis function as the kernel function to train the SVM classifier. The parameter value  $\sigma$  of the RBF can affect the classification accuracy significantly. Different values for  $\sigma$  were selected from 0.1 – 0.8 with a scale of 0.2 to test the RBF. The experiment obtained high accuracy when the

value of sigma was selected to be 0.1, with the accuracy of 90% for **LBP**, 87% for **LDA** and 85% for **PCA** as shown in Figure 5.7. The results shows that when the RBF kernel value increases the classification accuracy is constantly decreasing.

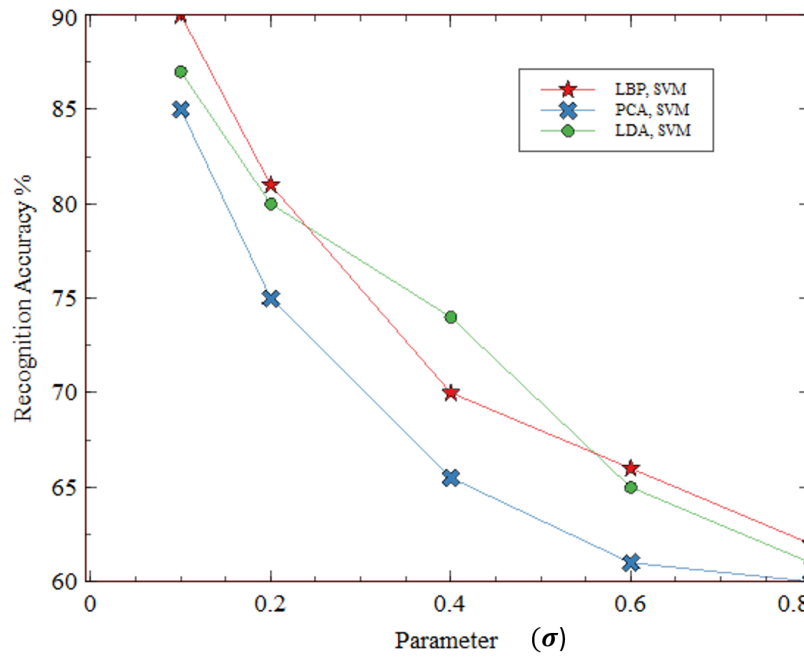


FIGURE 5.7: The RBF kernel value

The recognition accuracy produced can be affected by the number of the training samples and the type of classification method used. Different size of training dataset was used with the **SVM** classification algorithm. Figure 5.8 shows that **LBP** produced high recognition accuracy of 96% and **PCA** obtained 95% with **SVM**, which is higher than the accuracy of **LBP** and **PCA** with **ED** classifier in Figure 5.6.

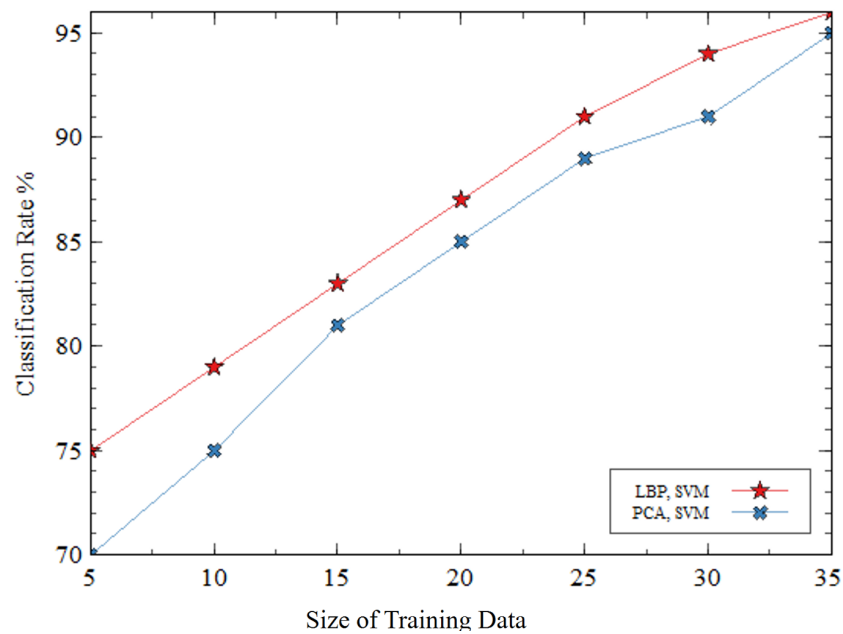


FIGURE 5.8: Classification of feature using Support Vector Machine SVM

The dimension of the features space grows and becomes sparse when we keep adding training data. It becomes much more easy to find a separable hyper-line because the likelihood that a training sample lies on the wrong side becomes infinitely small with an increase in number of samples [23].

The classification of features has been performed using the Naive Bayes. The NB classifier is simple, fast and known to perform remarkably well with a huge size dataset. The experiment is based on the effect of increasing the number of features to examine the classification of the naive Bayes technique. The experiment conducted produced results shown in Figure 5.9.

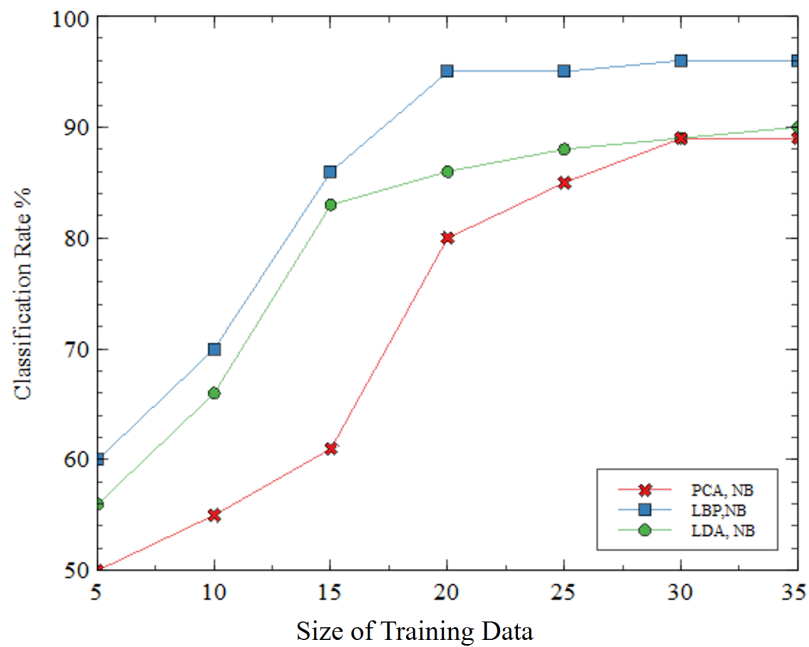


FIGURE 5.9: Classification of feature using Naive Bayes (NB)

As with SVM and ED the classification accuracy increased as the size of the training dataset increases, when using NB classifier. The accuracy of LBP with NB, and PCA with NB remained constant from 30 to 35 training sample with the score of 96% for LBP, 89% for PCA. The accuracy of LDA is increasing with a score of 90% at 90 number of sample. Naive Bayes produced the same classification accuracy with SVM using LBP features and over performed SVM and ED when using PCA and LDA.

An experiment was carried using Back-Propagation Neural Network (BPNN) Classifier. Images were pre-processed and resized to  $68 \times 68$  and cropped to a face perspective magnitude as shown in Figure 4.5. Training data consists of 35 images from each student, and 5 images from each class were used for testing. The layered BPNN consists of 30 neurons in the input layer, 10 neurons in the output layer, and different number of neurons in the hidden layer [82]. The back propagation error is distributed along the network to rearrange and correct weights of the training data [82, 3, 99]. To improve the training speed and classification accuracy the momentum of 0.5 and the learning rate of 0.001. Table 5.3 present the results of the experiment.

TABLE 5.2: The trained results of the BPNN with different number of neurons in the middle layer

Number of nuerons in the hidden layer	Execution Time (In seconds)	Number of iterations (Number of times the BPNN trained)	PCA + BPNN	LBP + BPNN
15	10.78	27651	80%	89%
20	10.65	25342	85%	89%
25	10.01	24123	89%	96%
30	9.88	23324	96%	98%
35	9.72	23211	93%	97,5%
40	9.5	20991	91.5%	96%
45	8.9	20623	91%	95%
50	8.5	20113	90%	94%

The network produced less outcomes when the neurons were below 35. When dealing with large dataset, small number of neurons in hidden layer struggle to convey the correct information from the input layer to the out layer [95]. The accuracy of 98% for LBP and 96% for PCA was produced when number of neurons were 35. The results gradually deteriorate as the number of neurons was increasing after above 35. The iterations of training the model reduce with an increase in the number of neurons and less time was also considered.

Table 5.3 shows that the back propagation neural network can be influenced by parameters used, such as the learning rate and error rate. Smaller learning rate can allow the model to learn optimal set of weights. However small learning rate costs the model with a significantly longer time to train.

Varying learning rate from 0.3 to 0.9 and error rate from 0.5 to 3.5 have been used to test to train the model. Increasing the learning rate can decrease the training period, however high error rate can stabilize the model. Table 5.3 suggests that effective value for learning rate and error rate are 0.4 and 0.5 respectively with a high recognition accuracy of 98%. The predictions of the model improves with the decreasing value of error and training cycles also reduces when the mean square error decreases.

TABLE 5.3: Learning rate against error rate of the BPNN

Error Rate								
Learning Rate		3.5	3	2.5	2	1.5	1	0.5
	0.9	69%	67%	67%	63%	69%	53%	56%
	0.8	65%	65%	63%	57%	73%	53%	55%
	0.7	67%	66%	67%	63%	66%	51%	53%
	0.6	65%	65%	65%	63%	67%	51.3%	67%
	0.5	65%	66%	65%	69%	69%	50%	85%
	0.4	<b>69.5%</b>	<b>70%</b>	<b>76%</b>	<b>86%</b>	<b>96.9%</b>	<b>97.5%</b>	<b>98%</b>
	0.3	66%	67%	67.5%	69.6%	78%	81%	97%

Table 5.4 presents the comparison between the existing studies and the proposed system. Although different databases have been used, the experiments were conducted with similar algorithms. False positive rates that were encountered during the experiments are presented. The proposed system produced better results especially through the classification of the LBP features.



TABLE 5.4: Comparison of the performance of the proposed method to existing methods using similar feature extraction and classification algorithms

Reference	Year	Method	Database	Accuracy	False Positive rate
[21]	2013	LBP+ED, and PCA + SVM	NITW - Database	95% , 95%	25%, 51%
[21]	2013	PCA + Bayes	NITW - Database	94%	52%
[85]	2017	PCA, LDA +ED	ORL [73]	66%, 83.57%	Refer to [85]
[74]	2015	LDA	ORL[73]	88%	NA
[74]	2015	PCA + BNN	ORL [73]	89.5%	NA
[3]	2018	LBPH, multi-KNN, and BPNN	ORL [73]	98%	NA
<b>Proposed</b>	2019	PCA, LDA, and LBP + ED	Independent	85%, 88%, 91%	25% , 15%, 10%
	2019	PCA, LDA, and LBP + Bayes	Independent	89%, 90%, 96%	21%, 10%
	2019	PCA, LBP+ SVM	Independent	95%, 96%	15%, 10%
	2019	PCA, LBP + BPNN	Independent	96%, 98%	0%

The effect of false recognition from multiple cameras and feature extraction algorithms does not degrade the positive results significantly.

This work emphasized the approach to effectively detect faces so that there could be an improved periodical facial recognition results during the lecture as presented in Table 5.5.

Table 5.5 shows the register with three time intervals, early minutes, second tier of the lecture, and the last minutes of the lecture. Each time fragment lasts 15 minutes. P on the register denotes present/available and A denotes absent. The Code of the module and date is shown at the top of the register. A complete attendance register for the whole lecture period from first 15 minutes of the lecture to last 15 minutes. The register report when the students has come late to class and when they left early. If the students does not attend the lecture at all, the register report NO for late come and NO left early, and marks the student absent. It can be seen

that student Dube KJ left the classroom early because he was not present after the first tier 0 – 15 minutes as shown in Table 5.5. Ramasila J and Kundi P were not present for the first 15 minutes and they were recorded as late comers. If a student does not attend or miss two tiers (30 minutes) and only attend for 15 minutes, the student is regarded as absent (See student Hendri P on the register). The importance of this attendance monitoring system is to monitor student's attendance throughout the lecture classroom.

TABLE 5.5: Example of register of students attendance for the whole lecture period

<b>Module: ENEL2CA H1 Date: 22/02/2017</b>							
<b>Student Number</b>	<b>Names</b>	<b>Time (m)</b>			<b>Late Come</b>	<b>Left Early</b>	<b>Status (0-45)</b>
		<b>0-15</b>	<b>16-30</b>	<b>31-45</b>			
217080041	Muzzi L	P	P	P	NO	NO	P
217080042	Smith K	P	P	P	NO	NO	P
217080143	Jones K	P	P	P	NO	NO	P
217087044	Dube KJ	P	A	A	NO	YES	A
217180442	Kundi P	A	P	P	YES	NO	P
217180542	Lanka C	A	A	A	NO	NO	A
217084542	Mohamed k	A	A	A	NO	NO	A
217188849	Naido K	P	P	P	NO	NO	P
217180148	Zulu N	P	P	P	NO	NO	P
217181142	Don P	P	P	P	NO	NO	P
217180042	Zitha B	P	P	P	NO	NO	P
217182142	Ramasila J	A	P	P	YES	NO	P
217180542	Hendri P	P	A	A	NO	YES	A
217185642	Nevar T	A	A	A	NO	NO	A
217185842	Kane J	P	P	P	NO	NO	P
217185942	Faroh BP	P	P	P	NO	NO	P
217184142	Smith G	P	P	P	NO	NO	P
217185142	Mac M	P	P	P	NO	NO	P
217185442	Jobs F	P	P	P	NO	NO	P
217185542	Dlomo D	P	P	P	NO	NO	P

### **5.3 Conclusion**

We have proposed a system that ensures a face detection and recognition model that guarantee accurate class attendance monitoring. The multi-camera system has demonstrated ability of the system to efficiently capture students in a classroom. Detection of face images were performed from different cameras positioned at different angles. The experiment shows that increasing the size of the training samples and choosing the right parameters values can improve the classification accuracy. As shown in the comparison with other existing classification techniques, BPNN outperformed other classification techniques with the recognition accuracy of 98%. The periodical recording of student's presence enables the proposed system to manage the availability of student's in a classroom from the beginning of the lecture till the end of the lecture.

## 6 | Conclusions and future work

### 6.1 Summary of the Work

This research has been centered on the development of the smart attendance monitoring system. A need for such arises from challenges schools and institutions are facing, to efficiently monitor students attendance throughout the entire lecture.

The proposed approach presents a unique face recognition architecture and combines several computer vision methodologies to achieve the smart attendance monitoring system. Face recognition process is inherently a challenge due to factors such as the dynamic structure, including face detection, feature extraction, and data classification.

The research proposed a multi-camera architecture supports the detection of faces by locating face images from different directions in a classroom. The multi-camera application helps to reduce occlusion; as frontal images were efficiently detected from different angles. Furthermore, an algorithm for collecting images was designed by computing the difference between the prior image and the saved image. This helped to collect various training data to make the feature learning process robust. Additionally, eye detection has been designed to reduce false positive detection and data redundancy.

In order to validate the presented algorithms, features were extracted using PCA, LDA, and the LBP. Different classifiers such as, SVM, Euclidean distance, Naive Bayesian and back propagation neural network were implemented and compared to improve the recognition accuracy. The

algorithms used were tested against different dataset sizes. The maximum recognition accuracy of 98% was obtained using the LBP and back propagation neural network classifier.

Time integration was proposed to periodically populate the register with attendance information. The recognition process took place at each interval and the results were recorded in the attendance register. The system collected attendance of each student from the beginning of the lecture till the end. The system also calculated the amount of time each student has attended, and if the attendance is below a certain threshold the student and his/ her lecturer will be informed.

The results showed that the presented system has an ability to capture and recognize students from different positions in a lecture, and therefore liable to be adopted as a robust students attendance monitoring system.

## 6.2 Limitations of the System

The capturing ability of cameras used in this experiment decreases as the distances increases, cameras with good lens and megapixels can be used to optimize the capturing and detection performance.

## 6.3 Future Works

Some future works that can be considered are the following:

1. Tracking the attendance of students during the course of the whole module and analyze the correlation between student's attendance and their academic performance, can be important to inform policies.
2. The system can be extended to detect and recognize mood/expression to analyze concentration of students during the classroom, and correlate them with the final results.

# Bibliography

- [1] M. A. Abas, Auji, and M. Dahlui. “Attendance management system (AMS): Comparison of two different approaches”. In: *2017 International Conference on Engineering Technology and Technopreneurship (ICE2T)*. 2017, pp. 1–7.
- [2] Abdulla Amin Aburomman and Mamun Bin Ibne Reaz. “Ensemble of binary SVM classifiers based on PCA and LDA feature extraction for intrusion detection”. In: *AIEEE, dvanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*. 2016, pp. 636–640.
- [3] Mohannad A Abuzneid and Ausif Mahmood. “Enhanced Human Face Recognition Using LBPH Descriptor, Multi-KNN, and Back-Propagation Neural Network”. In: *IEEE Access* 6 (2018), pp. 20641–20651.
- [4] Mayank Agarwal et al. “Face recognition using principle component analysis, eigenface and neural network”. In: *IEEE International Conference on Signal Acquisition and Processing*. 2010, pp. 310–314.
- [5] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen. “Face description with local binary patterns: Application to face recognition”. In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 28.12 (2006), pp. 2037–2041.
- [6] Verónica Almeida et al. “Automatic age detection based on facial images”. In: *2nd IEEE International Conference on Communication Control and Intelligent Systems (CCIS)*. 2016, pp. 110–114.
- [7] Le An, Mehran Kafai, and Bir Bhanu. “Face recognition in multi-camera surveillance videos using dynamic Bayesian network”. In: *Sixth IEEE International Conference on Distributed Smart Cameras*. 2012, pp. 1–6.

- 
- [8] S Anand et al. “Attendance monitoring in classroom using smartphone and Wi-Fi fingerprinting”. In: *IEEE Eighth International Conference on Technology for Education*. 2016, pp. 62–67.
- [9] Daniel Lélis Baggio. *Mastering OpenCV with practical computer vision projects*. Packt Publishing Ltd, 2012.
- [10] Ruchika Banerjee et al. “Use of facial landmarks for adaptive compression of videos on mobile devices”. In: *IEEE 10th International Conference on Communication Systems & Networks*. 2018, pp. 320–327.
- [11] Jilliam Maria Diaz Barros et al. “Fusion of Keypoint Tracking and Facial Landmark Detection for Real-Time Head Pose Estimation”. In: *IEEE Winter Conference on Applications of Computer Vision*. 2018, pp. 2028–2037.
- [12] Peter N. Belhumeur, João P Hespanha, and David J. Kriegman. “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection”. In: *IEEE Transactions on pattern analysis and machine intelligence* 19.7 (1997), pp. 711–720.
- [13] Peter N Belhumeur, Joao P Hespanha, and David J Kriegman. “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection”. In: *European Conference on Computer Vision*. Springer. 1996, pp. 43–58.
- [14] Michael J Bennett. *Time and attendance system and method therefor*. US Patent 5,550,359. Aug. 1996.
- [15] Farooq Ahmad Bhat and M Arif Wani. “Elastic bunch graph matching based face recognition under varying lighting, pose, and expression conditions”. In: *IAES International Journal of Artificial Intelligence (IJ-AI)* 3.4 (2016), pp. 177–182.
- [16] Margaret A. Boden. *Mind as machine: A history of cognitive science*. Oxford University Press, 2008.
- [17] Fatma Zohra Chelali, A Djeradi, and R Djeradi. “Linear discriminant analysis for face recognition”. In: *IEEE International Conference on Multimedia Computing and Systems, 2009. ICMCS'09*. 2009, pp. 1–10.
- [18] Chang-Shian Chen et al. “Development and application of a decision group Back-Propagation Neural Network for flood forecasting”. In: *Journal of hydrology* 385.1-4 (2010), pp. 173–182.
- [19] Jun-Horng Chen et al. “Improving accuracy of facial recognition systems for distant targets”. In: *IEEE 2nd Global Conference on Consumer Electronics*. 2013, pp. 92–93.

- [20] Y Chien. “Pattern classification and scene analysis”. In: *IEEE Transactions on Automatic Control* 19.4 (1974), pp. 462–463.
- [21] Shireesha Chintalapati and MV Raghunadh. “Automated attendance management system based on face recognition algorithms”. In: *IEEE International Conference on Computational Intelligence and Computing Research*. 2013, pp. 1–5.
- [22] Danish Ali Chowdhry et al. “Smart security system for sensitive area using face recognition”. In: *IEEE Conference on Sustainable Utilization and Development in Engineering and Technology*. 2013, pp. 11–14.
- [23] Carlton Chu et al. “Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images”. In: *NeuroImage* 60 (Dec. 2011), pp. 59–70.
- [24] Marcus Credé, Sylvia G Roch, and Urszula M Kieszczyńska. “Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics”. In: *Review of Educational Research* 80.2 (2010), pp. 272–295.
- [25] Tiziana D’Orazio et al. “An algorithm for real time eye detection in face images”. In: *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*. Vol. 3. IEEE. 2004, pp. 278–281.
- [26] Navneet Dalal and Bill Triggs. “Histograms of oriented gradients for human detection”. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol. 1. 2005, pp. 886–893.
- [27] Stephen Devadoss and John Foltz. “Evaluation of factors influencing student class attendance and performance”. In: *American Journal of Agricultural Economics* 78.3 (1996), pp. 499–507.
- [28] Krishna Dharavath et al. “Impact of image preprocessing on face recognition: A comparative analysis”. In: *IEEE International Conference on Communications and Signal Processing*. 2014, pp. 631–635.
- [29] Shan Du and Rabab Ward. “Face recognition under pose variations”. In: *Journal of the Franklin Institute* 343.6 (2006), pp. 596–613.
- [30] Joyce L Epstein and Steven B Sheldon. “Present and accounted for: Improving student attendance through family and community involvement”. In: *Taylor & Francis, The Journal of Educational Research* 95.5 (2002), pp. 308–318.



- [31] Martin A Fischler and Robert A Elschlager. “The representation and matching of pictorial structures”. In: *IEEE Transactions on computers* 100.1 (1973), pp. 67–92.
- [32] Martin A Fischler and Oscar Firschein. *Readings in computer vision: issues, problem, principles, and paradigms*. Elsevier, 2014.
- [33] Martin A Fischler et al. *The vision problem: Exploiting parallel computation*. Tech. rep. SRI Inter nation Menlo Park CA Artificial Intelligence Center, 1989.
- [34] Ronald A Fisher. “The use of multiple measurements in taxonomic problems”. In: *Annals of eugenics* 7.2 (1936), pp. 179–188.
- [35] Francois Fleuret and Donald Geman. “Coarse-to-fine face detection”. In: *International Journal of computer vision* 41.1-2 (2001), pp. 85–107.
- [36] Nir Friedman, Dan Geiger, and Moises Goldszmidt. “Bayesian network classifiers”. In: *Machine learning* 29.2-3 (1997), pp. 131–163.
- [37] Vishnukumar Galigekere and Gutemberg Guerra-Filho. “An interpolation based approach for lighting variation in image recognition”. In: *19th IEEE International Conference on Image Processing*. 2012, pp. 1901–1904.
- [38] Athinodoros S. Georghiades, Peter N. Belhumeur, and David J. Kriegman. “From few to many: Illumination cone models for face recognition under variable lighting and pose”. In: *IEEE transactions on pattern analysis and machine intelligence* 23.6 (2001), pp. 643–660.
- [39] GN Girish, Shrinivasa Naika CL, and Pradip K Das. “Face recognition using MB-LBP and PCA: A comparative study”. In: *Computer Communication and Informatics (ICCCI), 2014 International Conference on*. IEEE. 2014, pp. 1–6.
- [40] Michael A Gottfried. “Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach”. In: *American Educational Research Journal* 47.2 (2010), pp. 434–465.
- [41] Steven E Gump. “The cost of cutting class: Attendance as a predictor of success”. In: *College Teaching* 53.1 (2005), pp. 21–26.
- [42] Josh Harguess, Changbo Hu, and JK Aggarwal. “Fusing face recognition from multiple cameras”. In: *IEEE Workshop on Applications of Computer Vision*. 2009, pp. 1–7.
- [43] Hassan Fahmy Hashem. “Adaptive technique for human face detection using HSV color space and neural networks”. In: *IEEE Radio Science Conference, 2009. NRSC 2009. National*. 2009, pp. 1–7.

- 
- [44] Erwin Hidayat et al. “A comparative study of feature extraction using PCA and LDA for face recognition”. In: *Information Assurance and Security (IAS), 2011 7th International Conference on*. IEEE. 2011, pp. 354–359.
- [45] Teh Wei Hsiung and Shahrizat Shaik Mohamed. “Performance of iris recognition using low resolution iris image for attendance monitoring”. In: *IEEE International Conference on Computer Applications and Industrial Electronics*. 2011, pp. 612–617.
- [46] Deng-Yuan Huang et al. “Face Recognition Based on Discriminating Features by Histogram Matching Method”. In: *17* (2009), pp. 1033–1038.
- [47] Di Huang et al. “Local binary patterns and its application to facial image analysis: a survey”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 41.6 (2011), pp. 765–781.
- [48] Jing Huang, Yunyi Shang, and Hai Chen. “Improved Viola-Jones face detection algorithm based on HoloLens”. In: *EURASIP Journal on Image and Video Processing* 2019.1 (2019), p. 41.
- [49] IEEE. “IEEE Approved Standard for Camera Phone Image Quality (CPIQ) - Corrigendum 1”. In: *P1858-2016Cor1D2, December2016* (2017), pp. 1–29.
- [50] Peter Irgens et al. “An efficient and cost effective FPGA based implementation of the Viola-Jones face detection algorithm”. In: *HardwareX* 1 (2017), pp. 68–75.
- [51] Usiholo Iruansi, Jules R Tapamo, and Innocent E Davidson. “Insulator condition using local binary patterns combined with support vector machines”. In: *AFRICON, 2017 IEEE*. IEEE. 2017, pp. 21–25.
- [52] Md Milon Islam et al. “Development of smartphone-based student attendance system”. In: *IEEE Region 10 Humanitarian Technology Conference*. 2017, pp. 230–233.
- [53] Zakirul Islam et al. “Relation of Academic Performance and Class Attendance of Undergraduate Medical Students and Reasons of Absenteeism”. In: *Eastern Medical College Journal* 2520-5048 (2017), pp. 10–16.
- [54] Anil K Jain and Stan Z Li. *Handbook of face recognition*. Springer, 2011.
- [55] Mohammed Javed and Bhaskar Gupta. “Performance comparison of various face detection techniques”. In: *International Journal of Scientific Research Engineering & Technology (IJSRET) Volume 2* (2013), pp. 019–0027.

- [56] Oliver Jesorsky, Klaus J Kirchberg, and Robert W Frischholz. “Robust face detection using the hausdorff distance”. In: *International conference on audio-and video-based biometric person authentication*. Springer. 2001, pp. 90–95.
- [57] PP Jonker. “Statistical pattern recognition: a review”. In: *IEEE Transactions on pattern analysis and machine intelligence* 22.1 (2000), pp. 4–37.
- [58] Meina Kan et al. “Side-Information based Linear Discriminant Analysis for Face Recognition.” In: *BMVC*. Vol. 11. 2011, pp. 125–1.
- [59] Nirmalya Kar et al. “Study of implementing automated attendance system using face recognition technique”. In: *IACSIT Press, International Journal of computer and communication engineering* 1.2 (2012), p. 100.
- [60] M. Killioğlu, M. Taşkıran, and N. Kahraman. “Anti-spoofing in face recognition with liveness detection using pupil tracking”. In: *2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*. 2017, pp. 000087–000092.
- [61] BG Anil Kumar et al. “A Bluetooth low energy based beacon system for smart short range surveillance”. In: *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology*. 2016, pp. 1181–1184.
- [62] Nicholas D Lane et al. “A survey of mobile phone sensing”. In: *IEEE Communications magazine* 48.9 (2010).
- [63] Daniel D Lee and H Sebastian Seung. “Learning the parts of objects by non-negative matrix factorization”. In: *Nature* 401.6755 (1999), p. 788.
- [64] Chaoyang Li, Andrew Rosen, and Anu G Bourgeois. “On k-full-view-coverage-algorithms in camera sensor networks”. In: *2nd IEEE International Conference on Computer and Communications*. 2016, pp. 2218–2222.
- [65] Jian Han Lim et al. “Automated classroom monitoring with connected visioning system”. In: *IEEE Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*. 2017, pp. 386–393.
- [66] Kai Liu and Seungbin Moon. “Robust dual-stage face recognition method using PCA and high-dimensional-LBP”. In: *Information and Automation (ICIA), 2016 IEEE International Conference on*. IEEE. 2016, pp. 1828–1831.
- [67] Xiaoguang Lu and Anil K Jain. “Automatic feature extraction for multiview 3D face recognition”. In: *null*. IEEE. 2006, pp. 585–590.

- [68] Samuel Lukas et al. “Student attendance system in classroom using face recognition technique”. In: *IEEE International Conference on Information and Communication Technology Convergence*. 2016, pp. 1032–1035.
- [69] MD Malkauthekar. “Analysis of Euclidean distance and Manhattan distance measure in Face recognition”. In: (2013).
- [70] Jianchang Mao and Anil K Jain. “Artificial neural networks for feature extraction and multivariate data projection”. In: *IEEE transactions on neural networks* 6.2 (1995), pp. 296–317.
- [71] Tomasz Marciniak et al. “Influence of low resolution of images on reliability of face detection and recognition”. In: *Multimedia Tools and Applications* 74.12 (2015), pp. 4329–4349.
- [72] MathWorks. *Face Recognition*. 1994. URL: <https://www.mathworks.com/discovery/face-recognition.html>.
- [73] Kresimir Delac Mislav Grgic. *FACE RECOGNITION HOMEPAGE*. 2007. URL: <http://www.face-rec.org/databases/> (visited on 2018).
- [74] Rajath Kumar MP, KM Aishwarya, et al. “Artificial neural networks for face recognition using PCA and BPNN”. In: *TENCON 2015-2015 IEEE Region 10 Conference*. IEEE. 2015, pp. 1–6.
- [75] Gitty N Nasserbakht and Mitra Nasserbakht. *Automated mobile system*. US Patent 8,311,513. Nov. 2012.
- [76] Sara Nazari and Mohammad-Shahram Moin. “Face recognition using global and local Gabor features”. In: *21st IEEE Iranian Conference on Electrical Engineering (ICEE)*. 2013, pp. 1–4.
- [77] A Neethu, Kamal Bijlani, et al. “People count estimation using hybrid face detection method”. In: *IEEE International Conference on Information Science (ICIS)*. 2016, pp. 144–148.
- [78] Olegs Nikisins and Modris Greitans. “Local binary patterns and neural network based technique for robust face detection and localization”. In: *IEEE BIOSIG-Proceedings of the International Conference of the Biometrics Special Interest Group*. 2012, pp. 1–6.
- [79] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”. In: *IEEE Transactions on pattern analysis and machine intelligence* 24.7 (2002), pp. 971–987.

- [80] Kennedy Okokpujie et al. “Development of a facial recognition system with email identification message relay mechanism”. In: *IEEE International Conference on Computing Networking and Informatics*. 2017, pp. 1–6.
- [81] Rutuparna Panda and Manoj Kumar Naik. “EBFS-Fisher: An efficient algorithm for LDA-based face recognition”. In: *Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on*. IEEE. 2009, pp. 1041–1046.
- [82] Dhiren Pandit and Jayesh Dhodiya. “PCA and LDA method with neural network for primary diagnosis of genetic syndrome”. In: *algorithms* 2.10 (2015).
- [83] Unsang Park et al. “Face tracking and recognition at a distance: A coaxial and concentric PTZ camera system”. In: *IEEE Transactions on Information Forensics and Security* 8.10 (2013), pp. 1665–1677.
- [84] S Poornima et al. “Attendance monitoring system using facial recognition with audio output and gender classification”. In: *IEEE International Conference on Computer, Communication and Signal Processing*. 2017, pp. 1–5.
- [85] Anshun Raghuvanshi and Preeti D Swami. “An automated classroom attendance system using video based face recognition”. In: *2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology*. 2017, pp. 719–724.
- [86] SM Mahbubur Rahman, Shahana Parvin Lata, and Tamanna Howlader. “Bayesian face recognition using 2D Gaussian-Hermite moments”. In: *EURASIP Journal on Image and Video Processing* 2015.1 (2015), p. 35.
- [87] Hemantkumar Rathod et al. “Automated attendance system using machine learning approach”. In: *IEEE International Conference on Nascent Technologies in Engineering*. 2017, pp. 1–5.
- [88] Henry A Rowley, Shumeet Baluja, and Takeo Kanade. “Neural network-based face detection”. In: *IEEE Transactions on pattern analysis and machine intelligence* 20.1 (1998), pp. 23–38.
- [89] David E Rumelhart, Geoffrey E Hinton, Ronald J Williams, et al. “Learning representations by back-propagating errors”. In: *Cognitive modeling* 5.3 (1988), p. 1.
- [90] M. Sajid, R. Hussain, and M. Usman. “A conceptual model for automated attendance marking system using facial recognition”. In: *Ninth International Conference on Digital Information Management (ICDIM 2014)*. 2014, pp. 7–10.

- [91] John W Sammon. “A nonlinear mapping for data structure analysis”. In: *IEEE Transactions on computers* 100.5 (1969), pp. 401–409.
- [92] Maizura Mohd Sani, Khairul Anuar Ishak, and Salina Abdul Samad. “Evaluation of face recognition system using Support Vector Machine”. In: *IEEE Student Conference on Research and Development (SCOReD)*, 2009, pp. 139–141.
- [93] Nurbek Saparkhojayev and Selim Guvercin. “Attendance Control System based on RFID-technology”. In: *IJCSI International Journal of Computer Science Issues* 9.3 (2012), pp. 1694–0814.
- [94] Abuobayda M Shabat and Jules-Raymond Tapamo. “A comparative study of the use of local directional pattern for texture-based informal settlement classification”. In: *Journal of Applied Research and Technology* 15.3 (2017), pp. 250–258.
- [95] Imran Shafi et al. “Impact of varying neurons and hidden layers in neural network architecture for a time frequency application”. In: *2006 IEEE International Multitopic Conference*. IEEE. 2006, pp. 188–193.
- [96] Tarun Sharma and SL Aarth. “An automatic attendance monitoring system using RFID and IOT using Cloud”. In: *IEEE Online International Conference on Green Engineering and Technologies*. 2016, pp. 1–4.
- [97] Chang Shu, Xiaoqing Ding, and Chi Fang. “Histogram of the oriented gradient for face recognition”. In: *Tsinghua Science and Technology* 16.2 (2011), pp. 216–224.
- [98] Benfano Soewito et al. “Smart mobile attendance system using voice recognition and fingerprint on smartphone”. In: *International Seminar on Intelligent Technology and Its Applications*. 2016, pp. 175–180.
- [99] Mithila Sompura and Vinit Gupta. “An efficient face recognition with ANN using hybrid feature extraction methods”. In: *International Journal of Computer Applications* 117.17 (2015).
- [100] Milan Sonka, Vaclav Hlavac, and Roger Boyle. *Image processing, analysis, and machine vision*. Cengage Learning, 2014.
- [101] Zahraddeen Sufyanu et al. “FEATURE EXTRACTION METHODS FOR FACE RECOGNITION”. In: *International journal of applied engineering research (IRAER)* 5 (Oct. 2016), pp. 5658–5668.
- [102] Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

- [103] Matthew A Turk and Alex P Pentland. "Face recognition using eigenfaces". In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91*. 1991, pp. 586–591.
- [104] Mehmet Türkan, Ibrahim Onaran, and A Enis Çetin. "Human face detection in video using edge projections". In: *14th IEEE European Signal Processing Conference*. 2006, pp. 1–5.
- [105] A Tzotso. "A support vector machine approach for object based image analysis". In: *Proceedings of OBIA* (2006).
- [106] Paul Viola and Michael Jones. "Rapid object detection using a boosted cascade of simple features". In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol. 1. 2001, pp. I–I.
- [107] Paul Viola and Michael J Jones. "Robust real-time face detection". In: *Springer, International journal of computer vision* 57.2 (2004), pp. 137–154.
- [108] Minqin Wang et al. "Unsupervised texture Image segmentation based on gabor wavelet and multi-PCNN". In: *IEEE Second International Symposium on Intelligent Information Technology Application*. Vol. 2. 2008, pp. 376–381.
- [109] Y. Wang, K. Lu, and R. Zhai. "Challenge of multi-camera tracking". In: *2014 7th International Congress on Image and Signal Processing*. 2014, pp. 32–37.
- [110] Lin-nan Yang et al. "A prediction model for population occurrence of paddy stem borer (*Scirpophaga incertulas*), based on Back Propagation Artificial Neural Network and Principal Components Analysis". In: *Computers and Electronics in Agriculture* 68.2 (2009), pp. 200–206.
- [111] Ming-Hsuan Yang, David J Kriegman, and Narendra Ahuja. "Detecting faces in images: A survey". In: *IEEE Transactions on pattern analysis and machine intelligence* 24.1 (2002), pp. 34–58.
- [112] Shicheng Yang and Ying Wen. "A novel SRC based method for face recognition with low quality images". In: *IEEE International Conference on Image Processing (ICIP)*. 2017, pp. 3805–3809.
- [113] Zhanpeng Zhang et al. "Facial landmark detection by deep multi-task learning". In: *European Conference on Computer Vision*. Springer. 2014, pp. 94–108.
- [114] Jia Zhijuan et al. "A research of face feature extraction based on MB-LBP". In: *IEEE Power Engineering and Automation Conference*. Vol. 3. 2011, pp. 69–72.

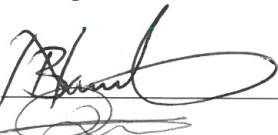
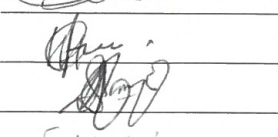
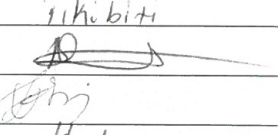
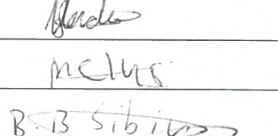
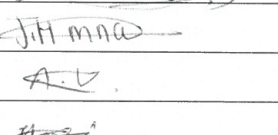
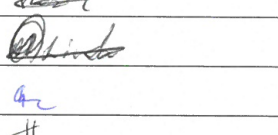
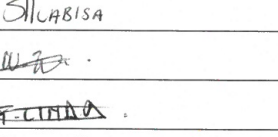
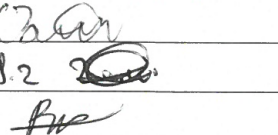
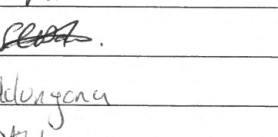
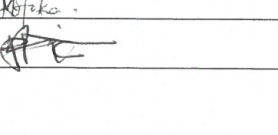





- 
- [115] Ying Zilu and Zhang Guoyi. “Facial expression recognition based on NMF and SVM”. In: *2009 IEEE International Forum on Information Technology and Applications*. 2009, pp. 612–615.



# Appendix

This section gives the consent of students for use and publication of images capture in classroom scenario experiments.

## Consent for the use and publication of images captured in the classroom scenario experiments

Names	Signature
1) BONGANI NXUMALO	
2) TABALLO MATHAMELA	
3) MUZI LUBISI	
4) AYODEJI BAMISAYE	
5) IOKELO KUBITI	
6) NATHAN SIBANDA	
7) VISUAL JOBRAT	
8) STEELE NAIDOO	
9) Melusi Mthunzi	
10) Busiswa Sibing	
11) Jabulle Mawango	
12) Andree Mkhwanazi	
13) Sabelo Sueri	
12) Mkhululi Linda	
13) Gabriel Duzi	
14) Sandi Hlabisa	
15) Wendy Ntunga	
16) Linda Thuthukani	
17) Cebokasi Zikalala	
18) Siculo Zwane	
19) Phumla Mthuthwa	
20) Cibise Spindo	
21) Pellen Mdungana	
22) Kefuwe Mofaka	
23) Kundi P	

**Consent for the use and publication of images captured in the classroom scenario experiments**

- 24) Lanka C
- 25) Jones E
- 26) Hendri P
- 27) Ramawale
- 28) Zitha B
- 29) Zulu M
- 30) Dan P
- 31) Smith E
- 32) Mac M
- 33) Plawo D
- 34)
- 35)

- Jh
- Jc
- H.P
- Ramawale
- Z. Mawugwe
- M. Zulu
- DP
- Smith
- Mac M
- Plawo D
- 
-