Face Detection and Recognition Algorithms for Outdoor Environment

Bayisa Gutema

Prof. Yun Koo Chung (PHD)

A thesis Submitted to the Department of Computer Science and Engineering

School of Electrical Engineering and Computing

Office of Graduate Studies

Adama Science and Technology University

Presented in Partial Fulfillment of the Requirement for the Degree of Master's in Computer Science and Engineering

Adama

September 10, 2018

ACKNOWLEDGEMENT

First of all, let me give thanks to the gracious and merciful for God, every my work is possible only why god allow me to accomplish my work within every change and challenge in my life. My special thanks goes to my advisor, Prof. Yun Koo Chung (PhD) for his assistance, continuing encouragement and all-round support throughout the research. I would like to express my deepest gratitude to ASTU Post Graduate Coordinator Dr. Mesfin Abebe (PHD) for his unlimited support, Advice and Appreciation during my working of this thesis. No Words are found inadequate to express my bottomless gratitude to my mother, Workineh Ulfata, my father Gutema Ejersa, my older brother Adugna Gutema and my older sister Motu Gutema the source of my strength in every aspect of my life, your kind-hearted treatment and love and their endless support up to now, brought me to this day, and all my younger brother and Sister I thank you all my Family.

I do not forget to thanks all my teacher from Elementary to University teacher, you have great contribution to my coming here. Special thanks to my friends Guluma Abdissa for his goodness and his Positive advice as friends.

My all classmates, friends, and colleagues in and outside the University are all acknowledged for their friendship and sharing spiritual life together during my stay at Adama Science and Technology University. Last but not least I offer my special thanks for those involve directly or indirectly in this study and their names are not listed.

Approval Sheet

We, the undersigned, members of the Board of Examiners of the final open defense by ________________________________ have read and evaluated his thesis entitled "____"

and examined the candidate. This is, therefore, to certify that the thesis has been accepted in partial fulfillment of the requirement of the Degree of...

____________________	____________________	______________
Name of student	Signature	Date
____________________	____________________	______________
Advisor /Supervisor	Signature	Date
____________________	____________________	______________
Program chair	Signature	Date
____________________	____________________	______________
School Dean	Signature	Date
____________________	____________________	______________
Sch. of Grad. Studies Dean	Signature	Date

Declaration

I hereby declare that this M.Sc. thesis entitled ***Face Detection and Recognition algorithms for outdoor Environment*** has been compiled by me under the supervision of prof. Yun Koo Chung (PHD) at Department of Computer Science and Engineering ASTU, Adama, Ethiopia. This thesis contains no material which has been accepted for the award of any other degree or diploma, except where due reference is made. To the best of my knowledge, this thesis contains no material previously published or written by another person except where due reference is made in the text of the thesis.

Name: ____

Signature:

____

This MSc Thesis have been submitted for examination with my approval.

Name:

____

Signature:

____

Date of submission:

List of Figures

Table of Contents

Abstract

Face Detection and Recognition based research are the hottest and active research area of Computer Technology. It is the biometric identification of a human face, which matches the image against a library of known faces and one of the biometric individual identification for different purpose. Many studies are done on Face Detection and Recognition for long period of times. The accuracy of recognition under different Position, Illumination and Expression which called PIE under this study is the yet not fully solved .Face DCT based Face Detection and Recognition algorithms is designed for this study in order to fill the gap in Position ,illumination and Expression with respect to the accuracy.

The proposed solution is used Visual Studio 2015 with the integration of OpenCV IDE. The study used 1000 image from the collection of real-times system through camera and From ESSEC-EU Face Database from External Gallery. The image of different size are equalized and Training and Testing Process are follow it. The study also used 20 different expression at different expression. This increase the number of face label and it increase the probability of Face Recognition at different Position and expression.

Haar cascade classifier is used to identify either face or not. The proposed solution have high rate of accuracy when compared to existing related work research with respect to Position, Illumination and expression as well as it takes low training and Testing time. By using DCT algorithms with Haar Cascade Classifier the accuracy of recognized face is an average rate of 98%.It I also identified that the image input from real world through camera and from external database are also determine the rate of accuracy.

Keywords

Face Detection, Face Recognition, Facial expression, Biometric Identification, Feature Extraction, Discrete Cosine Transform, Linear Discriminate analysis, Principal Component Analysis, Artificial Neural Network ,Haar cascade Classifier

.

Chapter One

1. Introduction

1.1. Back ground of the Study

The mutual recognition of individuals is a natural and fundamental capability of most living creatures. Therefore most living systems have a very well-engineered system for the recognition of other living creatures. Early studies on face recognition in primates revealed a consistent neural activity in well identified areas of the brain, mainly involving the temporal sensory area. More recent research revealed that this is not the case, but many different brain areas are taken into play at different stages of face analysis and recognition [13].

Face Detection and Recognition based research are the hottest and active research area of Computer Technology. It is the biometric identification of a human face, which matches the image against a library of known faces and one of the biometric individual identification for different purpose. Face Detection and Recognition is widely applicable in biometric system, Security control, human identification, human computer interaction, view point recognition (extraction of the view rotation of faces), intelligent visual surveillance and immigration automated clearance system [9, 1-3]. In addition to this Face Detection and Recognition are have many theoretical value in subjects of pattern recognition, image processing, computer vision, machine learning, physiology, and so on, and it also has a high correlation with other biometrics recognition methods[4]

Commercial applications of automatic face recognition (AFR) as well as by the emergence of real-time processors, research on automatic recognition of faces has become very active and detection of face or group of face for face recognition is challenging task for many researcher [7].Combination of information such as race, age, and gender, obtained through independent analysis, improves recognition results [9] in 1990s.

Brunelli and Poggio classify face Recognition in to two [5]:

1. Geometric feature-based and

2. Appearances based [5].

Research related to Face Detection and Recognition under different condition especially in low light or complete darkness is proposed despite of high illumination detection [6]. Effectively utilization of the frontal and profile face images provides complementary information of the face is useful for pose-robust face recognition with assistance of 3D face shape models reconstructed from the gallery mugshot face images [10].H. Gan, N. Sang and R .Huang propose semi-supervised face recognition method for better classifier with both labeled and unlabeled data, in which semi-supervised linear discriminant analysis (LDA) and affinity propagation (AP) are integrated into a self-training framework [11]

S. K. Zhou et al identify three factors involved in face recognition [12], namely

I. Illumination
II. Pose and
III. Identity (PIE)

Viola-Jones face detection approach and Local Binary Patterns (LBP) which used AdaBoost algorithms are very fast and used for feature selection respectively [15].However it has large number of false positives. An automatically embedded process that can detect any frontal face in 3D face recognition systems, with face segmentation and surface curvature information using the hybrid interpolation method for 3D face landmark detection[16]

The idea of Face detection and Recognition is generated after Emerging of Computer Vision. Computer vision is an interdisciplinary field that deals with how computers can be made for gaining high-level understanding from digital images or videos [41]. From the perspective of engineering, it seeks to automate tasks that the human visual system can do. What distinguished computer vision from the prevalent field of digital image processing at that time was a desire to extract three-dimensional structure from images with the goal of achieving full scene understanding. Studies in the 1970s formed the early foundations for many of the computer vision algorithms that exist today, including extraction of edges from images, labeling of lines, non-polyhedral and polyhedral modeling, and representation of objects as interconnections of smaller structures, optical flow, and motion estimation. The next decade saw studies based on more rigorous mathematical analysis and quantitative aspects of computer vision. These include the concept of scale-space, the inference of shape from various cues such as shading, texture and focus, and contour models known as snakes. Researchers also realized that many of these mathematical concepts could be treated within the same optimization framework as regularization and Markov random fields [40].

When computer vision first started out in the early 1970s, it was viewed as the visual perception component of an ambitious agenda to mimic human intelligence and to endow robots with intelligent behavior. At the time, it was believed by some of the early pioneers of artificial intelligence and robotics (at places such as MIT, Stanford, and CMU) that solving the “visual input” problem would be an easy step along the path to solving more difficult problems such as higher-level reasoning and planning. In 1990 a lot of the previously mentioned topics continued to be explored, a few of them became significantly more active [41].

This research is proposed to develop the framework and algorithm for Face detection and recognition algorithm for outdoor Environment indifferent illumination, position, Expression, noise with possible Quality.

1.2. History of Computer Vision

Computer Vision make computers what human eye do. Even though this is not correct meaning it enables computer to understand image and video.

It is commonly accepted that the father of Computer Vision is Larry Roberts, who in his Ph.D. thesis (Cir. 1960) at MIT discussed the possibilities of extracting 3D geometrical information from 2D perspective views of blocks [43].Then later many researcher follow is his fundamental theories and Application many and studied computer vision in the context of the blocks world.

Later, researchers realized that it was necessary to tackle images from the real world. Thus, much research was needed in the so called ``low-level” vision tasks such as edge detection and segmentation. David Marr (1978) from MIT propose the major Milestone framework. He took a bottom-up approach to scene understanding. Low-level image processing algorithms are applied to 2D images to obtain the ``primal sketch” (directed edge segments, etc.), from which a 2.5 D sketch of the scene is obtained using binocular stereo. Finally, high-level (structural analysis, a priori knowledge) techniques are used to get 3D model representations of the objects in the scene.

This is probably the single most influential work in computer vision ever. Many researchers cried: ``from the paradigm created for us by Marr, no one can drive us out [23].

Nonetheless, more recently a number of computer vision researchers realized some of the limitation of Marr’s paradigm, and worked more on top-down and heterogeneous approach. Basically, the program of Marr is extremely difficult to carry out, but more important, for many if not most computer vision applications, it is not necessary to get complete 3D object models. For example, in autonomous vehicle navigation using computer vision, it may be necessary to find out only whether an object is moving away from or toward your vehicle, but not the exact 3D motion of the object. This new paradigm is sometimes called ``Purposive Vision” implying that the algorithms should be goal driven and in many cases could be qualitative [45].

From the history of computer vision, it because of the broad spectrum of potential applications, the trend has been the merge of computer vision with other closely related fields. These include: Image processing (the raw images have to be processed before further analysis).Photogrammetry (cameras used for imaging have to be calibrated. Determining object poses in 3D is important in both computer vision and photogrammetry). Computer graphics (3D modeling is central to both computer vision and computer graphics.

Computer vision studies the tools and theories that enable the design of machines that can extract useful information from imagery data (images and videos) toward the goal of interpreting the world. Basically computer vision has dual major goal.

I. From the biological science perspective, computer vision aims to come up with computational models of the human visual system.
II. From the engineering perspective, computer vision aims to build autonomous systems which could perform some of the tasks which the human visual system can perform (and even surpass it in many cases). Many vision tasks are related to the extraction of 3D and temporal information from time-varying 2D data such as obtained by one or more television cameras, and more generally the understanding of such dynamic scenes.

The properties and characteristics of the human visual system often give inspiration to engineers who are designing computer vision systems. Conversely, computer vision algorithms can offer insights into how the human visual system works

1.2.1. The importance of Computer Vision

Computer Vision as a field of research is notoriously difficult. Almost no research problem has been satisfactorily solved and yet on progress. One main reason for this difficulty is that the human visual system is simply too good for many tasks (e.g., face recognition), so that computer vision systems suffer by comparison. A human can recognize faces under all kinds of variations in illumination, viewpoint, expression, etc. In most cases we have no difficulty in recognizing a friend in a photograph taken many years ago. Also, there appears to be no limit on how many faces we can store in our brains for future recognition. There appears no hope in building an autonomous system with such stellar performance. The difficulties in autonomous system is that it is not flexible as human Vision system. In addition to that in autonomous system .The flexibility of vision and understanding is also determined with the ability and intention of the programmer and designer and this make computer vision not as flexible as human.

Two major related difficulties in computer vision can be identified rather than human knowledge are:-

How to

1. How to we distill and represent the vast amount of human knowledge in a computer in such a way that retrieved is easy?
2. How do we carry out (in both hardware and software) the vast amount of computation that is often required in such a way that the task (such as face recognition) can be done in real time.

1.3. History of Face Recognition and Detection

We may see one person and remember for long time or for some time. Among we remember an individual is face, facial appearance, Hair, and any other body. Traditionally our parent and a grandparent tell us the face of person to identify from other. The human face plays a very important role in our social interaction and in conveying a person's identity. Over years it happens to be one of the increasing popular trends in image analysis and processing. This kind of technology has many benefits which causes it to be used in several different areas which push more researchers into identifying some peculiar problem still existing in its current implementation as a biometric system. (Oxford English Dictionary, 2018) define Biometrics is simply refers to the term that encompasses the application of modern statistical methods to the measurements of biological objects .A lot of biometric technologies are emerging. Among these are:

- Voice recognition
- Fingerprint recognition
- Iris recognition
- face recognition and
- gait recognition

However these characteristics should satisfy the following properties:

- **Robustness:** The biometric should be sufficiently invariant. (Permanence) over a period of time and thus maintain a low intra-class variability.
- **Distinctiveness:** This indicates that biometric identifiers should differentiate (Uniqueness) any two persons and thus have large inter-class variability.
- **Availability**: Ideally, a biometric identifier should be possessed by every person (Universality).
- **Accessibility**: The characteristic should be easy to acquire (Collectability).

Since biometric system is essentially a pattern-recognition system, the need for data acquisition, pre-processing, data representation and decision making is required. This directs the extraction of physiological or behavioral characteristic from persons. Beforehand and using it as a template database to be used later for recognition. Biometric identifiers which are conceptually unique attributes are today portrayed as the panacea to verify person's identity. These are referred to as biometric identifiers which can be used for distinctive physiological and behavioral characterization of persons.

Face recognition is a task so common to humans, that the individual does not even notice the extensive number of times it is performed every day. Although research in automated face recognition has been conducted since the 1960's, it has only recently caught the attention of the scientific community.

Biometrics, which is based on authentication on the intrinsic aspects of a specific human being, appears as a viable alternative to more traditional approaches (such as PIN codes or passwords). Among the oldest biometric techniques is fingerprint recognition. This technique was used in China as early as 700 AD for official certification of contracts. Later on, in the middle of the 19^{th} century, it was used for identification of persons in Europe [42].

1.3.1. Human Face Recognition

When building artificial face recognition systems, it is essential for scientists to understand the architecture of human face recognition system. Focusing on the methodology of human face recognition system may be useful to understand the basic system. However, the human face recognition system utilizes more than that of the machine recognition system which is just 2-D data. The human face recognition system uses some data obtained from some or all of the senses; visual, auditory, tactile, etc. All these data is used either individually or collectively for storage and remembering of faces [41]. In many cases, the surroundings also play an important role in human face recognition system. It is hard for a machine recognition system to handle so much data and their combinations. The difficult is that it is hard for a human to remember many faces due to storage limitations. For a human face recognition system the important feature is its parallel processing capacity.

The issue "which features humans use for face recognition" has been studied and it has been argued that both global and local features are used for face recognition. It is harder for humans to recognize faces which they consider as neither "attractive" nor "unattractive". Both holistic and feature information are important for the human face recognition system. Studies suggest the possibility of global descriptions serving as a front end for better feature based perception. Chellappa et al (1995) define that if there are dominant features present such as big ears, a small nose, etc. holistic descriptions may not be used [30].

Hair, eyes, mouth, face outline have been determined to be more important than nose for perceiving and remembering faces. It has also been found that the upper part of the face is more useful than the lower part of the face for recognition. Also, aesthetic attributes (e.g. beauty, attractiveness, pleasantness, etc.) play an important role in face recognition; the more attractive the faces are easily remembered. For humans, photographic negatives of faces are difficult to recognize. But, there is not much study on why it is difficult to recognize negative images of human faces. Other researcher Bruce (1999) also conducted a study on the direction of illumination and showed the importance of top lighting [42]. He demonstrated that it is easier for humans to recognize faces illuminated from top to bottom than the faces illuminated from bottom to top.

1.3.2. The main use of Face Recognition

Face Recognition is used for different purpose. We can summarize these in to three primary face tasks.

1. Verification (authentication)
2. Identification (recognition))
3. Watch list

1. Verification

The verification task is aimed at applications requiring user interaction in the form of an identity claim, i.e. access applications. The verification test is conducted by dividing persons into two groups:

1. Clients: - people trying to gain access using their own identity.
2. Imposters:- people trying to gain access using a false identity, i.e. an identity known to the system but not belonging to them

The percentage of imposters gaining access is reported as the False Acceptance Rate (FAR) and a percentage of client rejected access is reported as the False Rejection Rate (FRR) for a given threshold.

2. **Identification**

The identification task is mostly aimed at applications not requiring user interaction, i.e. surveillance applications. The identification test works from the assumption that all faces in the test are of known persons. The percentage of correct identifications is then reported as the Correct Identification Rate (CIR) or the percentage of false identifications is reported as the False Identification Rate (FIR).

3. **Watch List**

The watch list task is a generalization of the identification task which includes unknown people. The watch list test is like the identification test reported in CIR or FIR, but can have FAR and FRR associated with it to describe the sensitivity of the watch list, meaning how often is an unknown classified as a person in the watch list (FAR).

1.3. 2. Face Detection

The aim of face detection is localization of the face in an image. In the case of video input, it can be an advantage to track the face in between multiple frames, to reduce computational time and preserve the identity of a face (person) between frames. Methods used for face detection includes: Shape templates, Neural networks and Active Appearance Models (AAM).In order to take Classification. The classification by used to bring the same things with the correspondents. In Face Detection and Recognition Classification is used

As a biometric system, a face recognition system operates in either or both of two modes:

(1) Face verification (or authentication), and

(2) face identification (or recognition).

Face verification involves a one-to-one match that compares a query face image against enrollment image whose identity is being claimed. It is try to match all face against individual face with targeted individual face.

Face identification involves one-to-one many matching that compares a query face against multiple faces in the enrollment database to associate the identity of the query face to one of those in the database. In some identification application, one just needs to find the most similar face. In watch list check for face identification in surveillance video, the requirement is more than finding most similar faces. A confidence level threshold is specified and all those face whose similarity scores above the thresholds are reported.

The performance of Face recognition system largely depends on a variety of factors such as illumination, facial pose expression, age span, hair, facial wear, and motion. Based on this face recognition may be divided in to two broad categories in terms of a user.

- Cooperative user scenarios and
- Non-cooperative user scenarios.

The cooperative case is encountered in applications such as computer login, physical access control, and e-passport, where the user is willing to be cooperative by presenting his/her face in a proper way (for example, in a frontal pose with neutral expression and eyes open) in order to be granted the access or privilege. In the non-cooperative case, which is typical in surveillance applications, the user is unaware of being identified. In terms of distance between the face and the camera near field face recognition (less than 1 m) for cooperative applications (e.g., access control) is the least difficult problem, whereas far field non cooperative applications (e.g., watch list identification) in surveillance video is the most challenging.

1.3.3. Machine Recognition of Face

Although studies on human face recognition were expected to be a reference on machine recognition of faces, research on machine recognition of faces has developed independent of studies on human face recognition. Chellappa et al. (1995) posited that, during 1970's, typical pattern classification techniques, which use measurements between features in faces or face profiles, were used [30]. During the 1980's, work on face recognition remained nearly stable. Since the early 1990's, research interest on machine recognition of faces has grown tremendously. The reasons may be;

- An increase in emphasis on civilian/commercial research projects.
- The studies on neural network classifiers with emphasis on real-time computation and adaptation,
- The availability of real time hardware,
- The growing need for surveillance applications.

The basic question relevant for face classification is that; what form the structural code (for encoding the face) should take to achieve face recognition. Two major approaches are used for machine identification of human faces; geometrical local feature based methods, and holistic template matching based systems. Also, combinations of these two methods, namely hybrid methods, are used. The first approach, the geometrical local feature based one, extracts and measures discrete local features (such as eye, nose, mouth, hair, etc.) for retrieving and identifying faces. One of the well-known geometrical-local feature based methods is the Elastic Bunch Graph Matching (EBGM) technique. The other approach, the holistic one, conceptually related to template matching, attempts to identify faces using global representations; Huang (1998). Holistic methods approach the face image as a whole and try to extract features from the whole face region.

In this approach, as in the previous approach, the pattern classifiers are applied to classify the image after extracting the features. One of the methods to extract features in a holistic system is applying statistical methods such as Principal Component Analysis (PCA) to the whole image. PCA can also be applied to a face image locally; in that case the approach is not holistic.

With the problem being a classification problem, training the face recognition system with images from the known individuals and classifying the newly coming test images into one of the classes is the main aspect of the face recognition systems. This problem seems to be easily solved by humans where limited memory can be the main problem; whereas the problems for a machine face recognition system are.

I. Facial expression change
II. Illumination change
III. Aging
IV. Pose change
V. Scaling factor (i.e. size of the image
VI. Frontal vs. profile
VII. Presence and absence of spectacles, beard, moustache etc.
VIII. Occlusion due to scarf, mask or obstacles in front

These all and other related problem and ambiguity are the solved with different approach and different algorithms.

1.4. Motivation

Currently in the era of digital human interest is increase with advancement of Technology through any fields of studies. Especially after generation and emerging of Computer Vision, image related research is highly motivated and hot research area. The development of image related research becomes the pioneer of Identification and detection. This identification and detection is also the fundamental of Recognition .The importance of Detection and Recognition become widely known and applicable after computer vision is used for face detection and Recognition. But the emergence of this technology does not mean that we are satisfied due to quality and per needed detected and recognized face.

Even though many face detection and recognition research and project exists over the globe it faced the challenge on quality, dynamicity, accuracy and reliability. Therefore it is needed to contribute some framework and algorithm for these interest fields of study.

1.4. Statement of Problem

Face detection and recognition are important for many reasons and thus most computer vision algorithms are dedicated to the task of facial feature analysis and recognition. But it is open area for research in terms of accuracy and reliability [17]. For most face detection algorithms faces are not tracked but the detection algorithm is run on each frame from a video or streaming camera. The more recent approaches to face detection and recognition are based on detecting facial landmarks i.e. nose, mouth, right eye and left eye. It can done by training several Convolutional Neural Networks (CNN) for detecting different facial landmarks and combine the results in the output network [18].

The other approach is probably the best performing and is based on a novel combination of 3D matching/alignment and deep learning. The face is converted into a 3D model by some sophisticated algorithms and based on the 3D model the algorithm performs a lot of normalization such as normalizing the pose of the face through piecewise affine warping before feeding the result to a nine layer deep neural network that extracts a deeper facial representation for reliable matching. Although many research are conducted on Face detection and recognition for more than two decade, the research related to outdoor Environment that have high illumination are not as enough as possible.

1.6. Research Question

From pioneer of Face Detection and Recognition different algorithms and methodology are used in order to add accuracy, efficiency and other parameters. This studies is find out the following research question:

RQ1. Is it possible to recognize at different expression.

RQ2. Is it possible to possible to improves quality of Face detection and recognition in high level of illumination

RQ3. Is it possible to increase the accuracy of face Detection for Outdoor Environment?

RQ4. What is the action taken with the algorithms when face is detected?

RQ5. How it possible to regulate equal quality of face detection and Recognition in the variance of day and night?

1.7. Research Methodology

The research done on methodology related publication research, Journal, Conference paper, surveys form different Digital Library and could include both present and historical information

On Face Detection and Recognition.

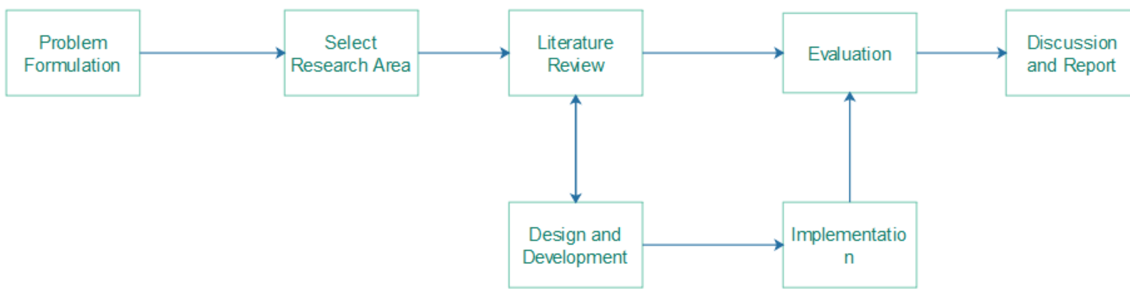

Figure 1. 1 Research methodology

1.8. Problem Formulation

This part identify the main limitation in the Current research related to face detection and Recognition and Find out the proposed and possible solution for identified problem. The research is focused on Illumination, Position and Expression or in general PIE per proposed research.

1.9. Literature Review

This is the important and necessary content where all possible reference and journal related to research is investigated and analyzed.

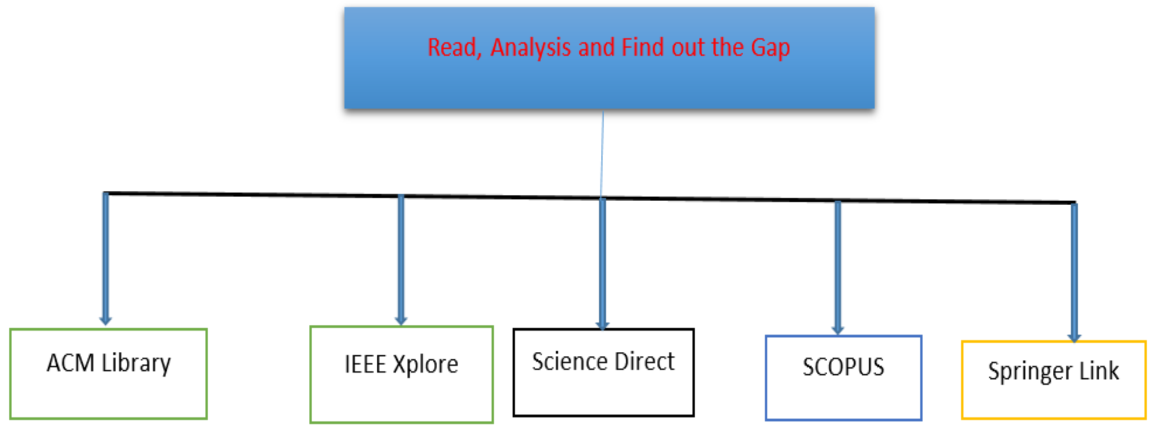

Figure 1. 2 Digital Library for proposed research

1.10. Design and Development

This is where the framework and workflow of the research is taken place. Pseudocode and flow chart for the system in the research is done here. The step and process in which Face Detection and Recognition take place are defined in Design and Development part of the research documentation. The proposed research is used with Visual Studio integrated with OpenCV C++ program. In this Face Detection and Recognition Design the process in which image Capturing, Storing to Recognition are designed and its developmental Process are taken as following Diagram.

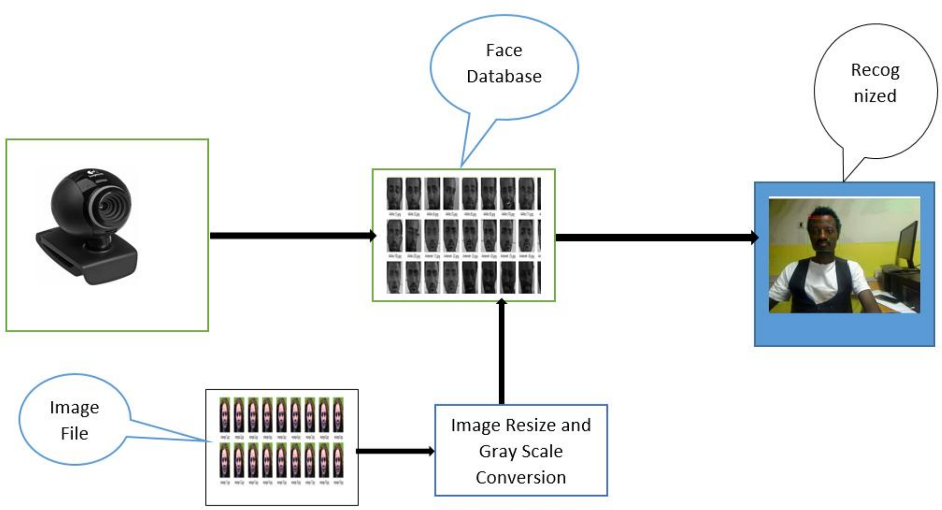

Figure 1. 3: Proposed System Design and Development

1.11. Objectives

1.11.1. General Objective

The general objective of this Study is to develop framework and algorithm for Face detection and Recognition for Outdoor Environment in different position, illumination and expression.

1.11.2. Specific Objective

- To decrease possible noise while Face recognition
- To identify and verification of recognized face in different expression.
- To increase quality of face detection at different illumination
- To develop framework and algorithm on quality of face detection
- To customize machines or computer as much as possible to understand face Detection and Recognition at different light
- To use different computation and algorithm for detection and recognition.

1.12. Scope and Limitation

The research will do on face Detection and Recognition for Outdoor Environment. It will find the way face at different position, expression and illumination are recognized. The research will be limited to analysis of newly developed frame work and review published Journal, and papers which are currently released.

1.13. Significance of the Study

Face Detection and Recognition is biometric identification of individual from another. It is used in forensics applications such as criminal identification and prison security .It also used as a strong potential to be widely adopted in Social application such as e-Banking, e-Commerce, and access control.

Due to the huge increase in the use of electronic transactions, e-Banking and e-Commerce are becoming the most important emerging application of biometrics. Credit card and smart card security, ATM security, check cashing and fund transfers; online transactions and web access are some of the examples of the applications.

The application of Face Detection and Recognition is not limited to only government organization. It can widely use in non-governmental organization. Face recognition minimizing victim trauma by where verifying Identify for court records .Since it is system face it is unbiased. So it is used as tool of justice. It give all report about victim and then simple to give judge.

The other Application of Face Detection and Recognition is for Security Purpose, Access control, comparing surveillance images to know terrorist. Thus Face Detection and Recognition can be used as tool of Counterterrorism. Face recognition can also work as a means of access control to ensure that only authorized individuals get into facilities like labs, boardrooms, bank vaults, training centers for athletes and other sensitive location.

1.14. Organization of Thesis

The thesis Contain Six Chapter

Chapter One is Introduction which describe Background of study, Statements of problem, Research methodology, Problem Formulation, Design and Development, Objective, Scope and Limitation and Application and Results. The Other Chapter are also organized, Chapter Two contains Literature Review and Related work, Chapter Three have also Methodology which include development tool, Chapter Four Describe the Proposed Solution, Chapter Five Describe the Implementation of proposed Solution and Finally Chapter Six Describe the Result, Discussion and Evaluation.

Chapter Two

2. Literature Review and Related Work

Face Detection and Recognition receives extensive attention recently. It is biometrics way of identification and this make it difficult to make machine or computer to take recognition as simple as possible. In order to make this simple different approach are taken.

2.1. Approaches to Face Detection and Recognition

2.1.1. Statistical Approaches

Statistical methods include template matching based systems where the training and test images are matched by measuring the correlation between them [52]. Moreover, statistical methods include the projection based methods such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), etc. In fact, projection based systems came out due to the shortcomings of the straightforward template matching based approaches; that is, trying to carry out the required classification task in a space of extremely high dimensionality[1].

2.1.2. Features-Based Approaches

Bobis et al. (1999) studied on a feature based face recognition system. They suggested that a face can be recognized by extracting the relative position and other parameters of distinctive features such as eyes, mouth, nose and chin. The system described the overall geometrical configuration of face features by a vector of numerical data representing position and size of main facial features. First, they extracted eyes coordinates [53]. Cagnoni and Poggi (1999) suggested a feature based approach instead of a holistic approach to face recognition. They applied the eigenface method to sub-images (eye, nose, and mouth) [53-54].

They also applied a rotation correction to the faces in order to obtain better results. Guan and Szu (1999) compared the performance of PCA and ICA on face images. They argue that, ICA encodes face images with statistically independent variables, which are not necessarily associated with the orthogonal axes, while PCA is always associated with orthogonal eigenvector [28].

2.1.3. Geometry Feature Based Methods

Geometry feature based methods uses the facial feature measures such as distance between eyes, ratio of distance between eyes and nose etc., but it is significantly different from the feature-based techniques that it constructs the topological graph using the facial features of each subject. Lades et al. (1997) presented dynamic link architecture for distortion invariant object recognition which employs elastic graph matching to find the closed stored graph.

Objects were represented with sparse graphs whose vertices were labeled with geometrical distances. In this system, individual faces were represented by a rectangular graph, each node labeled with a set of complex Gabor wavelet coefficients, called a jet. Only the magnitudes of the coefficients were used for matching and recognition. When recognizing a face of a new image, each graph in the model gallery was matched to the image separately and the best match indicated the recognized person. Wiskott et al. (1997) extended this system to handle larger galleries and larger variations in pose and to increase the matching accuracy [55]. Firstly, they use the phase of the complex Gabor wavelet coefficients to achieve an accurate location of the nodes and to disambiguate patterns which would be similar in the magnitudes of the coefficient. Secondly, they employ object adapted graphs, so that nodes refer to specific facial landmarks, called fiducial points. The correspondences between two faces can be found across large viewpoint changes.

Thirdly, a new data structure called the bunch graph was introduced which serves as generalized representation of faces by combining jets of a small set of individual faces. This allows the system to find the fiducial points in one matching process, which eliminates the need for matching each model graph individually. This also reduces computational effort significantly. It offers good performance of about 98% for FERET database. But the drawback in this feature matching approach is that it requires manual intervention to select the fiducial points in the facial image and it requires precise location of those points.

The multimodal person recognition system described Choudhury et al (1999) .This system consists of a face recognition module, a speaker identification module, and a classifier fusion module. It has the following characteristics:

1. The face recognition module can detect and compensate for pose variations; the speaker identification module can detect and compensate for changes in the auditory background;
2. the most reliable video frames and audio clips are selected for recognition;
3. 3D information about the head obtained through SFM is used to detect the presence of an actual person as opposed to an image of that person.

Chellappa et al (2001), also define a face verification system based on tracking facial features was presented. The basic idea of this approach is to exploit the temporal information available in a video sequence to improve face recognition. First, the feature points defined by Gabor attributes on a regular 2D grid are tracked. Then, the trajectories of these tracked feature points are exploited to identify the person presented in a short video sequence [29].

2.1.4. Other Statistical Approaches

Cevikalp et al. (2005) solved small sample size problem of LDA by optimizing a variant of Fisher's criterion. It searched for the optimal projection vectors in the null space of within class. So, to find the optimal projection vectors in the null space of it projects the face samples onto the null space of to generate common vectors for each class and then obtain the projection vectors by performing PCA on common vectors [30]. A new set of vectors, called as discriminative common vectors, are obtained by projecting face samples on the projection vectors. Thus each class is represented by a single discriminative common vector. Among two algorithms to extract the discriminant common vectors for representing each person in the training set of face database, one algorithm uses within-class scatter matrix of the samples in the training set while the other uses the subspace methods and the Gram-Schmidt orthogonalization procedures to obtain the discriminative common vectors. These discriminative common vectors are used for classification of new faces.

In Zhang et al (2004) the 2D PCA is based on 2D image matrices instead of 1D vector. The image covariance matrix is constructed directly from the original image. Let A denote the image with m rows and n columns. The image matrix is projected on to n-dimension column vector x as, y = Ax where y is an m dimensional column vector called the projected feature vector of the image A. The column vector x is chosen such that the generalized total scatter criterion is maximized. This means that total scatter of the projected vectors is maximized.

Li and Yuan (2005), in the recently proposed 2D-LDA showed the image is not reordered as a column vector. The process of projection of an image using 2D-LDA is given as follows. An image A of m rows and n columns is projected as y = Ax, where y is the projected vector and x is the projection vector. Optimal projection vectors are chosen when Fisher's criterion is maximized. The classification using nearest neighbor strategy is similar to the classification using 2D-PCA.

Hong and Jain (1998) showed that there is substantial related work in multimodal biometrics. They used face and fingerprint in multimodal biometric identification. Others have used face and voice. However, use of the face and ear in combination seems more relevant to surveillance applications. Turk and Pentland (1991) developed the traditional PCA algorithm used in the eigenfaces approach. This method computes the eigenvectors and eigenvalues of a covariance matrix and solves an eigenvalues system problem. This method performs no discrimination but instead the eigenvectors that correspond to the largest eigenvalues are used in as base vectors in a low-dimensional space.

Duda et al. (2001) tried several block-based transformations for feature selection and the K means algorithm to generate the code words for each codebook. They argue that the block based DCT coefficients produce good low-dimensional feature vectors with high recognition performance. This brings the possibility of performing face recognition directly on a DCT based compressed bit stream without having to decode the image. Eickeler et al. (1999) suggested a system based on Pseudo 2-D Hidden Markov Models (HMMs) and coefficient of the 2-D DCT as the features. A major advantage of their approach is that the system works directly on JPEG-compressed face images, i.e. it uses the DCT coefficients provided by the JPEG standard. Thus, it does not need any further decompressing of the image. Phillips (1999) applied SVM to face recognition. Face recognition is a K-class problem, where K is the number of known individuals; and SVM is a binary classification method. By reformulating the face recognition problem and reinterpreting the output of the SVM classifier, they developed a SVM-based face recognition algorithm. They formulated the face recognition problem in difference space, which models dissimilarities between two facial images. In difference space, they formulated the face recognition as a two class problem. The classes are; dissimilarities between faces of the same person and dissimilarities between faces of different people. By modifying the interpretation of the decision surface generated by SVM, they generated a similarity metric between faces, learned from examples of differences between faces.

2.1.5. Neural Network based Approaches

Artificial Neural Network (ANN) Bishop (1995), Mammone (1993), is a powerful tool for pattern recognition problems. The use of neural networks (NN) in faces has addressed several problems: gender classification, face recognition and classification of facial expressions. Kohonen (1988) outlined one of the earliest demonstrations of NN for face recall applications which was reported in Kohonen's associative map [32]. Using a small set of face images, accurate recall was reported even when input image is very noisy or when portions of the images are missing. A few NN based face recognition techniques are discussed in the following. Stonham (1984) reported a single layer adaptive NN (one for each person) for face recognition, expression analysis and face verification. A system named Wilke, Aleksander and Stonham's recognition devise (WISARD) was devised. It needs typically 200-400 presentations for training each classifier where the training patterns included translation and variation in facial expressions. One classifier was constructed corresponding to one subject in the database.

2.1.6. Feature-based approaches for face detection

Numerous methods have been proposed to first detect facial features and then to infer the presence of a face. Facial features such as eyebrows, eyes, nose, mouth and hairline are commonly extracted. Based on the extracted features, a statistical model is built to describe their relationships and to verify the existence of a face. One problem with these feature-based algorithms is that the image features can be severely corrupted due to illumination, noise and occlusion.

The systems based on the feature-based approach can be further divided into three groups.

A. Given typical face detection problem in locating a face in a cluttered scene, low-level analysis is executed which first deals with the segmentation of visual features using pixel properties such as gray scale and color. Because of the low-level nature, features generated from this analysis are ambiguous.

B. In feature analysis systems, visual features are organized into a more global concept of the face using face geometry information. Through feature analysis, feature ambiguities are reduced and locations of the face and facial features are determined.

C. The last group involves the use of active shape models. These models ranging from snakes, proposed in the late 1980s, to the more recent point-distributed models (PDM) have been utilized for the purpose of complex and non-rigid feature extraction and tracking such as eye pupil and lip tracking.

2.2. Mathematical Description and Some and Terminology in Face Detection and Recognition

2.2.1. Mathematical Based Operation and Modeling

1. Histogram Based Operation

An important class of point operations is based upon the manipulation of an image histogram or a region histogram.

A. Contrast stretching

Frequently, an image is scanned in such a way that the resulting brightness values do not make full use of the available dynamic range. This can be easily observed in the histogram of the brightness. If the image is intended to go from brightness 0 to brightness $2^B - 1$ then one generally maps the 0% value.

$$b[m,n] = [2^B - 1].\,\mathrm{a[m,n]} - \frac{\mathrm{minimum}}{\mathrm{maximum} - \mathrm{minimum}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots..(2{,}1)$$

This formula, however, can be somewhat sensitive to outliers and a less sensitive and more general version is given by

$$b[m,n] =: \begin{cases} 0 & a[m,n] \le P_{low\%} \\ [2^B - 1].\frac{\mathrm{a[m,n]-Plow\%}}{Phigh\% - Plow\%} & P_{low\%} \le a[m,n] < P_{high\%} \\ 2^B - 1 & a[m,n] \ge P_{low\%} \end{cases} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.2)$$

It is also possible to apply the contrast-stretching operation on a regional basis using the histogram from a region to determine the appropriate limits for the algorithm

B. Equalization

When one wishes to compare two or more images on a specific basis, such as texture, it is common to first normalize their histograms to a "standard" histogram. This can be especially useful when the images have been acquired under different circumstances. The most common histogram normalization technique is histogram equalization where one attempts to change the histogram through the use of a function $b = f(a)$ into a histogram that is constant for all brightness values.

C. Principal Component's analysis (PCA)

Using this method a subset of principal directions (principal components) can be found in a set of the training faces. Then the face images are projected into the space of these principal components and the feature vectors are obtained. Face recognition is performed by comparing these feature vectors using different distance measures. Basically, in the PCA-based face recognition method the eigenvectors and eigenvalues of the covariance matrix of the training data are calculated. PCA produces the optimal linear least-squares decomposition of a training set. Because the basis vectors constructed by PCA had the same dimension as the input face images, they were named eigenfaces [30].

PCA representation of an image of $X_0[m,n]$ of size $d_1 \times d_2$ pixel is interpreted as point in $R^{d1\times d2}$ space. If mean image of the N numbers of the Lexicographic ordered training image X_i is

$$\bar{m} = \sum_{i=1}^{N} X_i \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (2.3)$$

and the corresponding mean centered image is

$$w_i = x_i - m_i \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (2.4)$$

Then, a set of e_i 'scan be obtained which has the largest possible projection onto each of w_i's. The objective is to find a set of N orthonormal vector e_i for which the quantity

$$\lambda_{\, i=\sum_{k=1}^{N}(e_i^T w_{i)}} 2 \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (2.5)$$

is maximized with the orthonormal constraint given by

$$e_l^T e_{k=\delta_{\,1k}} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.6)\ )$$

The eigenvectors e_i and eigenvalues $\lambda_{\,i}$ of the covariance matrix $C = W^T W$ are calculated, where W is a matrix composed of the column vectors w_i placed side by side. The weight vector of the trained face images is treated as a feature vector or face descriptor and is calculated as

D. Image Acquisition

Image Acquisition process totally depends on the hardware system which may have a sensor that is again a hardware device [23]. A sensor converts light into electrical charges. The sensor inside a camera measures the reflected energy by the scene being imaged. The image sensor employed by most digital cameras is a charge coupled device (CCD).

E. Image preprocessing

Digital image processing is always an interesting field as it gives improved pictorial information for human interpretation and processing of image data for storage, transmission, and representation for machine perception. Image Processing is a technique to enhance raw images received from cameras/sensors placed on satellites, space probes and aircrafts or pictures taken in normal day-to-day life for various applications. This field of image processing significantly improved in recent times and extended to various fields of science and technology [23]. The image processing mainly deals with image acquisition, Image enhancement, image segmentation, feature extraction, image classification etc.

In image preprocessing, image data recorded by sensors on a satellite restrain errors related to geometry and brightness values of the pixels. These errors are corrected using appropriate mathematical models which are either definite or statistical models. Image enhancement is the modification of image by changing the pixel brightness values to improve its visual impact. Image enhancement involves a collection of techniques that are used to improve the visual appearance of an image, or to convert the image to a form which is better suited for human or machine interpretation.

Sometimes images obtained from satellites and conventional and digital cameras lack in contrast and brightness because of the limitations of imaging sub systems and illumination conditions while capturing image. Images may have different types of noise. In image enhancement, the goal is to accentuate certain image features for subsequent analysis or for image display [130]. Examples include contrast and edge enhancement, pseudo-coloring, noise filtering, sharpening, and magnifying. Image enhancement is useful in feature extraction, image analysis and an image display. The enhancement process itself does not increase the inherent information content in the data. It simply emphasizes certain specified image characteristics. Enhancement algorithms are generally interactive and application dependent. Some of the enhancement techniques are:

- Contrast image
- Noise Filtering
- Histogram modification

Preprocessing is a common names for operations with images at lowest level of abstraction both input and output are intensity. The aim of pre-preprocessing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing.

Four method of image pre-processing methods according to the size of pixel neighborhood that is used for calculation of a new pixel brightness:

- Pixel transformation
- Geometric transformation
- Pre-processing methods that used local neighborhood of processed pixel and
- Image restoration that requires knowledge about the entire image

Other classification of image pre-processing method exist. An Image preprocessing issued to correct degradation, but it is essential to know prioiri.

- Knowledge about the nature of degradation
- Knowledge about image acquisition devices, and condition under which image was obtained. The nature of noise (usually its spectral characteristics) is sometimes known.
- Knowledge about objects that are searched for an image ,which is simply the pre-processing very considerably
- If knowledge of object is not available in advance it can be estimated during processing.
  Pixel Transformation
- Brightness transformation modify the pixel brightness which depends on the properties of a pixel itself.
- Brightness corrections
- Gray scale transformations
- Brightness correction which consider Original brightness and Pixel position in the image

Position dependent Brightness Correction

Ideally the sensitivity of an image acquisition and digitization devices should not depends position in an image, but assumption is not valid in many practical cases.

Source of Degradation which uneven Sensitivity of light sensors and uneven object illumination

Systematic degradation can be suppressed

F. Face Detection

Given an arbitrary image, the goal of the face detection process is to determine whether or not there is any face in the image and, if present, return the image location and extent of each face. The challenges associated with the face detection process can be attributed to the following factors:

Occlusion: Faces may be partially occluded by other objects. In an image with a group of people, some faces may partially occlude other faces.

Facial expression: The appearances of faces are directly affected by a person's facial expression.

Pose: The images of a face vary due to the relative camera-face pose such as frontal, 45 degree, profile, upside down etc.

Illumination: When a face image is formed, factors such as lighting conditions and camera characteristics also a affect the appearance of a face.

2.2.2. Some of Operation and Terms used in Feature Extraction

A. Edge

The work was based on analyzing line drawings of the faces from photographs and aims to locate facial features. The work includes a line-follower implemented with the curvature constraint to prevent it from being distracted by noisy edges. Edge features within the head outline are then subjected to feature analysis using shape and position information of the face.

In an edge-based approach to face detection, edges need to be labeled and matched to a face model in order to verify correct detections. Sometimes this is accomplished by labeling edges as the left side, hairline or right side of a front view face and is followed by matching these edges against a face model, using the golden ratio for an ideal face. Further, edge-based techniques have also been applied for detecting glasses in facial images. Many different types of edge operators have been applied. The Sobel operator and Marr Hildreth edge operator are some of the operators used as edge detectors, a variety of first and second derivatives (Laplacian) of Gaussians have also been used in the other methods. For instance, a Laplacian of large scale was used to obtain a line drawing in and steerable and multi-scale orientation filters.

B. Gray-level analysis

Besides edge details, the gray information within a face can also be used as features. Facial features such as eyebrows, pupils and lips appear generally darker than their surrounding facial regions. This property is exploited to differentiate various facial parts. Several recent facial feature extraction algorithms search for local gray minima within segmented facial regions. In these algorithms, the input images are first enhanced by contrast-stretching and gray-scale morphological routines to improve the quality of local dark patches and thereby make detection easier. The extraction of dark patches is achieved by low-level gray-scale thresholding. Some algorithms make use of a weighted human eye template to determine possible locations of an eye pair. Local maxima, which are denied by a bright pixel surrounded by eight dark neighbors, are used to indicate the bright facial spots such as nose tips. The detection points are then aligned with feature templates for correlation measurements.

C. Feature Extraction

In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction.

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a feature vector). Determining a subset of the initial features is called feature selection. Features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

Feature extraction involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power, also it may cause a classification algorithm to over fit to training samples and generalize poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Many machine learning practitioners believe that properly optimized feature extraction is the key to effective model construction.

D. Training Set

Machine Learning basically means that you're training the machine to do something by providing set of training data's. Machine Learning have models/architectures, loss functions and several approaches that can be used to determine which would provide better image processing.

This research is done with training face image before Recognition. This is how to the algorithm and the machine is to become familiar with the existing face. If the image is not exist in the sample database it simply call it UNKNOWN FACE since it is not trained and this new face is not in Face Database.

2.3. Related Work

In the history of face detection and recognition different approach and different method are used to increase accuracy, quality and efficiency of an algorithm. Hailing Z. and Kin-Man L.et al were model human identity and aging variables simultaneously using Probabilistic Linear Discriminant Analysis (PLDA). PCA and ANFIS are used in variation involved in face recognition include illumination, Pose, and identity, facial expression, and hair style [26]. It is difficult even for human to recognize face correctly when illumination varies severely, since the same person appears to be very much different. They take the image from ORL database. They make image de-noised by Median filter before preprocessing and the accuracy of the proposed research is 96.666%.Hong P., Yaping Z., Liangzheng X.et al propose to represent face pattern with set of heterogeneous and complementary feature descriptors that include Generalized Haar-like (GH) descriptors Multi-Block Binary Patterns (MB-LBP) and Speed-up Robust Features (SURF) [27].This is make possible to detect face up-to 97.2% accuracy.

Viola-Jones algorithm provide quick and robust face detection. This algorithm has been widely implemented and deployed on diverse architectural platforms. It also used in Graphics Processing Unit (GPU) despite of its cost effectiveness. It is the work of [27]

Linear Discriminant Analysis (LDA) is well known dimension reduction in pattern recognition. It projects the original high-dimensional data onto a low-dimensional space. It creates combination of independent features which yields the largest mean differences between the desired classes. The goal of LDA is to maximize the between -scatter matrix measure. The Principal Component Analysis is also used to reduce dimension or unnecessary Parameters from face. [28] Combine novel method based on PCA image reconstruction and LDA for face recognition. The accuracy of PCA and LDA is the get the face recognition accuracy to 97.47%.Support Vector Machines (SVM) is then added to LDA and PCA increase the accuracy of face recognition to 97.74%.But it is limited by the size of Pixel which only about 24×24 pixel despite of increase the efficiency.

Face recognition under illumination variation must take into account the two factors of identity and illumination. Multidimensional PCA and Extreme learning machine is the new human face recognition algorithm based on bidirectional two dimensional principal component analysis

(B2DPCA) and extreme learning machine (ELM).It is based on curve let image decomposition of human faces and a sub-band that exhibits a maximum standard deviation is dimensionally reduced using an improved dimensionality reduction technique[31].

The DCT is a popular technique in imaging and video compression, which was first applied in image compression in 1974 by Ahmed et al. In 1992, the first international standard for image compression, known as the Joint Photographic Experts Group (JPEG), was stablished with the DCT encoder and decoder [22].

A MATLAB based Face Recognition System using Image Processing and Neural Networks

Novel approach for face recognition that derives from an idea suggested by Hjelmås and Low. They describes Face recognition with techniques of an image based approach towards artificial intelligence [22].

Since skin color in humans varies by individual, research has revealed that intensity rather than chrominance is the main distinguishing characteristic. The recognition stage typically uses an intensity (grayscale) representation of the image compressed by the 2D-DCT for further processing.

In the first stage, the D-DCT for each face image is computed, and feature vectors are formed from the discrete cosine transform (DCT) coefficients. The second stage uses a self-organizing map (SOM) with an unsupervised learning technique to classify vectors into groups to recognize if the subject in the input image is “present” or “not present” in the image database. If the subject is classified as present, the best match image found in the training database is displayed as the result, else the result displays that the subject is not found in the image database.

1. Discrete Cosine Transform (DCT)

The DCT algorithm is used for much application. One of application of DCT [20] is image compression which is used for recognition task such as face recognition. Face image have high correlation and redundant information which causes computational burden in terms of processing speed and memory. DCT transform image from spatial domain to frequency domain .Since lower frequency domain are visually significant in an image than higher frequencies, the DCT discards high-frequency coefficient and quantizes the remaining coefficient.

The 2D-DCT of a $M \times N$ matrix A is defined as

$$B_{pq=a_p a_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos(\pi(\frac{2m+1}{2m})P}) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (2.7)$$

Face Subspace Learning

From principal component analysis (PCA) [43] and Fisher's linear discriminant analysis, a dozen of dimension reduction algorithms have been developed to select effective subspaces for the representation and discrimination of face images. It has demonstrated that human faces, although usually represented by thousands of pixels encoded in high-dimensional arrays, they are intrinsically embedded in a very low dimensional subspace [37]. The using of subspace for face representation helps to reduce "the curse of dimensionality" in subsequent classification, and suppress variations of lighting conditions and facial expressions.

The earliest subspace method for face recognition is Eigenface [43], which uses PCA [23] to select the most representative subspace for representing a set of face images. It extracts the principal eigenspace associated with a set of training face images. Mathematically, PCA maximizes the variance in the projected subspace for a given dimensionality, de-correlates the training face images in the projected subspace, and maximizes the mutual information between appearance (training face images) and identity (the corresponding labels) by assuming that face images are Gaussian distributed. Thus, it has been successfully applied for face recognition. By projecting face images onto the subspace spanned by Eigenface, classifiers can be used in the subspace for recognition. One main limitation of Eigenface is that the class labels of face images cannot be explored in the process of learning the projection matrix for dimension reduction.

Fisherface finds class specific linear subspace. The dimension reduction algorithm used in Fisherface is Fisher's linear discriminant analysis (FLDA), which simultaneously maximizes the between-class scatter and minimizes the within-class scatter of the face data. FLDA finds in the feature space a low dimensional subspace where the different classes of samples remain well separated after projection to this subspace. If classes are sampled from Gaussian distributions, all with identical covariance matrices, then FLDA maximizes the mean value of the KL divergences between different classes. In general, Fisherface outperforms Eigenface due to the utilized discriminative information.

Although FLDA shows promising performance on face recognition, it has the following major limitations. FLDA discards the discriminative information preserved in covariance matrices of different classes. FLDA models each class by a single Gaussian distribution, so it cannot find a proper projection for subsequent classification when samples are sampled from complex distributions, for example, mixtures of Gaussians. In face recognition, face images are generally captured with different expressions or poses, under different lighting conditions and at different resolution, so it is more proper to assume face images from one person are mixtures of Gaussians. FLDA tends to merge classes which are close together in the original feature space. Furthermore, when the size of the training set is smaller than the dimension of the feature space, FLDA has the under sampled problem.

Subspace Learning—A Global Perspective

The key of face recognition lies in feature extraction and the design of the classifier, and PCA method is one of the most popular methods in feature extraction and the idea of using principal components to represent human faces was developed by Kirby and Sirovich and used by Turk and Pentland for face detection and recognition, namely "Eigenfaces" method. [22, 30]. PCA method has the characteristic of descending dimension and elimination of correlation.

Consider a training set with the following parameters: the training set is denoted by $A = (A_1, A_2, \dots, A_N)$ the training set has N images, j= 1, 2, · · · ,N, belonging to c classes, i= 1, 2, · · · , c, and the number of pixels in the image is n. The between-class scatter matrix, within-class scatter matrix, and total scatter matrix are defined as

$$\sum \quad S_b = \sum_{i=1}^{c} P_{w_i} \left(\mu_i - \mu_0 \right)\left(\mu_i - \mu_0 \right)T \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (2.8)$$

Where S_b the between-class scatter matrix is P_{w_i} is the priori probabilities of, in general

$P_{w_i} = (1/c)$. μ_i Is the mean vector of class w_i , $\mu_0 = (1/N) \sum_{j=1}^{N} A_j)$ is the mean vector of all the samples.

$$S_w = \sum_{i=1}^{C} p(w_i) S_i \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . . (2.9)$$

Where $S_w$the within-class is scatter matrix, $S_i = E\{(A - \mu_i)^T \; A \in w_i$ is the covariance matrix of w_i.

$$S_t = S_b + S_t = \sum_{j=1}^{N} (A_j - \mu_{0)} \big(A_j - \mu_0\big) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (2.10)$$

The between-class scatter matrix represents the scatter of class means μ_i around the overall mean μ_0 , and the within-class scatter matrix is the scatter of the samples around their respective class means μ_i .The PCA-based face recognition methods mainly use total scatter matrix or within-class scatter matrix to extract facial features. If the total scatter matrix is used as generating matrix, the optimal projection matrix is equal to computing the maximal eigenvalues and the corresponding eigenvectors of S_t and the optimal projection matrix $(X_1, X_2, \dots X_d)$ is the eigenvectors associated with the d largest generalized eigenvalues. Therefore,($X_1, X_2, \dots X_d$) describes the contribution of each eigenface in representing the input face image, then we can extract facial features through it [30].

2. **PCA and extreme learning machine**

It is a multilinear extension of principal component analysis (MPCA). MPCA is employed in the analysis of n-way arrays, i.e. a cube or hyper-cube of numbers, also informally referred to as a "data tensor". N-way arrays may be decomposed, analyzed, or modeled by linear tensor models such as CANDECOMP/Parafac, or multilinear tensor models, such as multilinear principal component analysis (MPCA), or multilinear independent component analysis (MICA), etc.

To improve speed and accuracy of a face recognition system various dimensionality reduction techniques have been developed. The paper (32] states that the use of linear discriminant analysis (LDA) to maximize the ratio of between-class scatter matrix and the within class scatter matrix for improved face recognition. An eigenspace based adaptive approach that utilizes a specific kind of genetic algorithm called evolutionary pursuit (EP) [31] and elastic bunch graph matching (EBGM) [114] have also been proposed research to generate an optimal set of projection axes. The use of kernel Hilbert space for ICA to adaptively generate nonlinear functions and to obtain a robust algorithm with regards to variations in source density, degree of non-Gaussianity, and presence of outliers is the work in [35].

3. 3D Face data extraction

Facial landmark localization is done mainly in two-dimensions (2D) and three-dimension (3D). Facial landmark is the term refers to key-points used to determine facial features in anthropometric investigations from [36-39]. Automatic facial detection systems combine the biometric performance of both approaches to create multidimensional systems (2D-3D), which minimize 3D system limitations in measuring biometric data under varying facial expression. 3D face meshes simplify facial landmark detection, and enhance (ICP) algorithm for mesh alignment and thus present substantial semantic information. Important landmarks in 3D biometric facial scans include nose tip, mouth corners, chin, and eye corners. Orientation, texture map, and posture determine the accuracy of locating landmarks. In the earlier years of anthropometric studies, 2D face biometrics data; width and height, formed the basis of face recognition research. However, 2D facial recognition suffered numerous limitations acknowledged in varying expression, pose, and illumination. 3D face recognition showed prospects of solving most of these limitations, and hence the increased interest in this field of research. From paper [30-32] it is possible to understand before landmark localization and registration, procedural steps involve synthesizing surfaces from 3D face meshes. Accuracy in high level 3D facial image processing and analysis typically depends on precision when segmenting the input image.

Data augmentation for face recognition

This is the research focused on accuracy, rapidity and convenience in the process of personal identification by face Recognition and proposed fast face recognition method based on fractal theory. This method is to compress the facial images to obtain fractal codes and complete face recognition with these code [3].

Fractal coding used to method for face recognition

1. Segmentation of range blocks and domain block

In the whole process of image segmentation, the range blocks and domain blocks can take any shape, but generally are rectangular. And the area of the domain block is usually larger than the range block to ensure that the corresponding mapping is a contraction transformation. The whole coding process is to find the fittest domain block for the range block and seek out the corresponding contraction mapping, contrast scaling and luminance shift.

2. Determination of contractive mapping

The matching process between domain blocks and range blocks occupies most of the time of fractal encoding. There are 8 basic affine transformations for each domain block and each range block has to compare with them one by one.

This research is work to reduce rate and time for face recognition based on Yale Database, Face recognition Technology (FERET) database and CMU PIE database.

Chellappa et al. (1995) posited that, during 1970's, typical pattern classification techniques, which use measurements between features in faces or face profiles, were used. During the 1980's, work on face recognition remained nearly stable. Since the early 1990's, research interest on machine recognition of faces has grown tremendously [29]. The reasons may be;

- An increase in emphasis on civilian/commercial research projects
- The studies on neural network classifiers with emphasis on real-time computation and adaptation
- The availability of real time hardware,
- The growing need for surveillance applications.

The basic question relevant for face classification is that; what form the structural code (for encoding the face) should take to achieve face recognition. Two major approaches are used for machine identification of human faces; geometrical local feature based methods, and holistic template matching based systems. Also, combinations of these two methods, namely hybrid methods, are used. The first approach, the geometrical local feature based one, extracts and measures discrete local features (such as eye, nose, mouth, hair, etc.) for retrieving and identifying faces. One of the well-known geometrical-local feature based methods is the Elastic Bunch Graph Matching (EBGM) technique. The other approach, the holistic one, conceptually related to template matching, attempts to identify faces using global representations; Huang (1998). Holistic methods approach the face image as a whole and try to extract features from the whole face region. In this approach, as in the previous approach, the pattern classifiers are applied to classify the image after extracting the features. One of the methods to extract features in a holistic system is applying statistical methods such as Principal Component Analysis (PCA) to the whole image. PCA can also be applied to a face image locally; in that case the approach is not holistic.

With the problem being a classification problem, training the face recognition system with images from the known individuals and classifying the newly coming test images into one of the classes is the main aspect of the face recognition systems. This problem seems to be easily solved by humans where limited memory can be the main problem; whereas the problems for a machine face recognition system are:

Discrete Cosine Transform (DCT)

The discrete cosine transform (DCT) is widely used in many signal processing applications. It is a Fourier-related transform similar to the DFT, which converts the signal from spatial domain to frequency domain [40,48].It is different from DFT by is that it transforms real-valued signals to real-valued DCT coefficients, which give DCT the advantage of convenience in applications. However, it is the energy compaction capability that makes DCT a popular and useful transform [20].

Chapter Three

3. Research Methodology

3.1. Overview

Face Detection and Recognition for Outdoor Environment is one of interesting Research and which need advanced software as well as Hardware. Because if the quality of hardware and the library of software in incomplete there could be possibility of missing the needed output. Haar Cascade classifier as Detector either face or not and Discrete Cosine Transform (DCT) are used in to do this study.

3.2. Development Tool

Microsoft Visual Studio

Microsoft Visual Studio is an integrated development environment (IDE) from Microsoft. It is used to develop computer programs, as well as web sites, web apps, web services and mobile apps. Visual Studio uses Microsoft software development platforms such as Windows API, Windows Forms, Windows Presentation Foundation, Windows Store and Microsoft Silverlight. It can produce both native code and managed code.

Visual Studio includes a code editor supporting IntelliSense (the code completion component) as well as code refactoring. The integrated debugger works both as a source-level debugger and a machine-level debugger. Other built-in tools include a code profiler, forms designer for building GUI applications, web designer, class designer, and database schema designer. It accepts plug-ins that enhance the functionality at almost every level—including adding support for source control systems (like Subversion and Git) and adding new toolsets like editors and visual designers for domain-specific languages or toolsets for other aspects of the software development lifecycle (like the Team Foundation Server client: Team Explorer). This study used Visual Studio 2015 as Developmental tool.

3.2.1. Selection of Tool

There are many developmental too for Face Detection and Recognition. This research is used Visual Studio for implementation of the proposed solution. The selection is consider library, platform, since it is free library. As an examples Visual studio integrated with OpenCV have the following advantage over MathLab [51].

1. **Free**: Large parts of OpenCV are free. We can use OpenCV freely in our commercial application by fixing some needed.

2. **Huge Optimized Library**: The collection of algorithms available in OpenCV dwarfs everything out there. The library is also optimized for performance.

3. **Platforms and devices**: You can obviously use OpenCV in your desktop app or as the backend of your web application.

4. **Big community:** There is a big community of developers (47,000 or so) that use and support OpenCV. It has been downloaded more than 9 Million times. Among this user most of them are Researcher.

The selection of Visual studio which integrated with OpenCv for this research is to use advantage of huge optimized library. Because Face Detection and Recognition is need

Visual Studio supports 36 different programming languages and allows the code editor and debugger to support (to varying degrees) nearly any programming language, provided a language-specific service exists. Built-in languages include C, C++, C++/CLI, Visual Basic .NET, C#, F# JavaScript, TypeScript, XML, XSLT, HTML, and CSS Support for other languages such as Python Ruby, Node.js, and M among others is available via plug-ins. Java (and J#) were supported in the past[47]. This study integrated Visual Studio with OpenCV C++.

The Selection of Microsoft Visual Studio is that it contain best library for Computer Vision as well as [51]

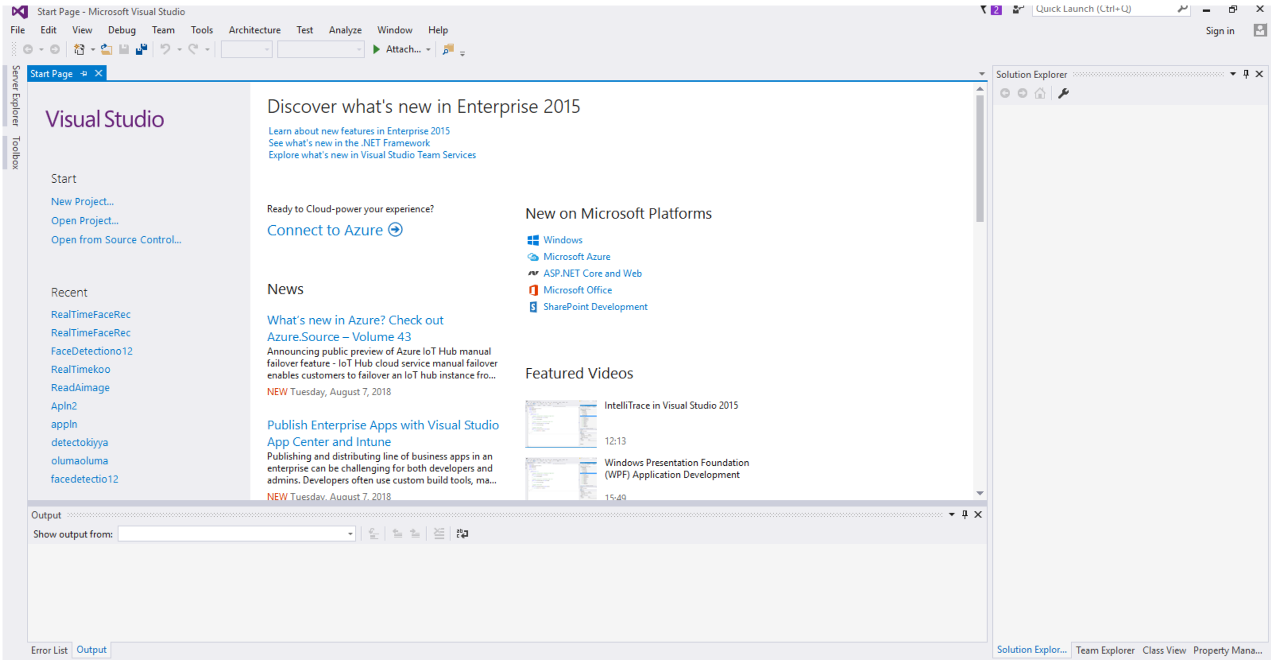

Figure 3. 1 Diagram of Visual Studio

The research is also used OpenCV library for Visual Studio 2015.

Computer Vision take different approach and algorithm to detect and recognize one face from other or to claim one individual. But this is only possible with algorithm and method to do so. According to M.H Yang et al [127] classification methods of face detection can categorized into four categories

1. Methods for a priori: - These methods is based on the rules used to model the knowledge of what makes a face. Typically, these rules represent relationships in facial features. Approaches invariant features. These approaches are based on structural features that exist even when the pose, the view or the illumination conditions vary, and use them to locate faces.
2. Methods based models:-Several standard models of faces are used to define a face model or models of facial features separately. The correlation between the image and the models is evaluated for the presence of face.
3. Methods for learning: - In contrast to methods based models, the models are learned from a set of training images which should allow characterizing the variability of the appearance of a face. These models are then used to learn the detection.

4. Method of Viola Jones: - based on a cascade simulated by simple descriptors called "Haar-like" descriptors. They can be calculated more efficiently by using an intermediate representation of the image called integral image. They also proposed a method for processing multi-stage classification which significantly reduces the duration of the audit while achieving almost the same accuracy compared to a classification algorithm single-phase which is certainly much slower and more complex.

Several references show the effectiveness of this system in terms of good detection, false alarms and speed. Marwa C. et al used method of

1. Integral Image: The integral image is defined as the summation of the pixel values of the original image. The value at any location (x, y).
2. Haar Features: Haar features are composed of either two or three rectangles. Face candidates are scanned and searched for Haar features of the current stage. The weight and size of each feature and the features themselves are generated using a machine learning algorithm from AdaBoost. he weights are constants generated by the learning algorithm. There are a variety of forms of features. Each Haar feature has a value that is calculated by taking the area of each rectangle, multiplying each by their respective weights, and then summing the results.
3. Classifier: A Haar classifier uses the rectangle integral to calculate the value of a Haar feature. The Haar classifier multiplies the weight of each rectangle by its area and the results are added together. Several Haar classifiers compose a stage. A stage accumulator sums all the Haar classifier results in a stage and a stage comparator compares this summation with a stage threshold. The threshold is also a constant obtained from the AdaBoost algorithm. Each stage does not have a set number of Haar features. Depending on the parameters of the training data individual stages can have a varying number of Haar features.
4. Cascade: The cascade eliminates candidates by making stricter requirements in each stage with later stages being much more difficult for a candidate to pass. Candidates exit the cascade if they pass all stages or fail any stage. A face is detected if a candidate passes all stages.

Similarly this research is also used Haar Cascade for face Detection and DCT for Face Recognition.

Summary

Generally Microsoft Visual Studio with OpenCV C++ have selected for this study. The reason for selection is Open CV or open Source Computer vision which have huge Visual libraries related to image processing.

Chapter Four

4.1. Proposed Solution

Many research are done on face detection and Recognition especially from 1970s .From this successive work done the aim is to improve quality, efficiency, effectivity and overall performance of Face Detection and Recognition. The model of Face Detection and Recognition for Outdoor Environment is like the following Diagram.

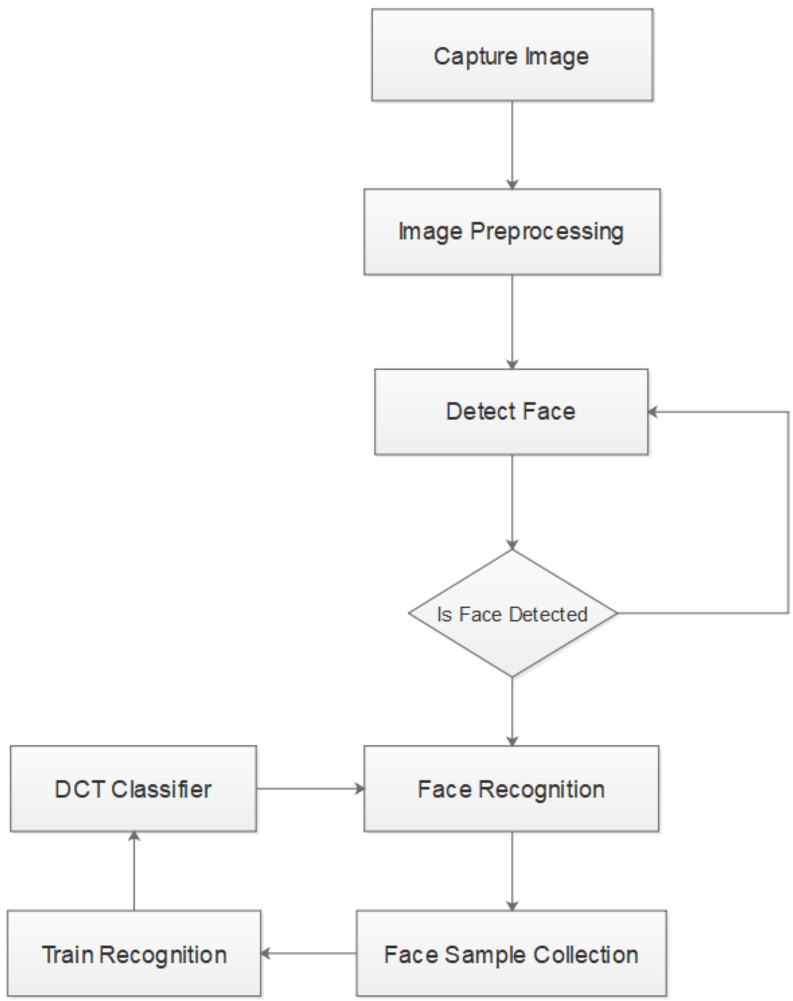

Figure 4. 1 Overall Diagram of proposed face Detection and Recognition Algorithms for Outdoor Environment

Different Face Detection and Recognition algorithm are determine the quality and effectivity as well as the accuracy of Detection and Recognition. The previous research work on automated face recognition has ignored the issue of just what aspects of the face stimulus are important for face recognition. This suggests the use of an information theory approach of coding and decoding of face images, emphasizing the significant local and global features. Such features may or may not be directly related to our intuitive notion of face features such as the eyes, nose, lips, and hair.

Principal Component Analysis (PCA) is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data. The purpose of PCA is to reduce the large dimensionality of the data space (observed variables) to smaller intrinsic dimensionality of feature space (independent variables), which are needed to describe the data economically. This is the case when there is a strong correlation between observed variables. The first principal component is the linear combination of the original dimensions that has the maximum variance

An important and largely unsolved problem in dimensionality reduction is the choice of the intrinsic dimensionality of the principal manifold. No analytical derivation of this number for a complex natural visual signal is available to date. To simplify this problem, it is common to assume that in the noisy embedding of the signal of interest (in our case, a point sampled from the face space) in a high-dimensional space, the signal-to-noise ratio is high. Statistically, that means that the variance of the data along the principal modes of the manifold is high compared to the variance within the complementary space. This assumption relates to the eigen spectrum - the set of the eigenvalues of the data covariance matrix.

Since many algorithm and different method are used in face Detection and Recognition this study is used DCT algorithm for Recognition purpose that is why the research is found to worked on improving DCT base Face Detection and Recognition. As face Detector Haar Cascade algorithms is used.

A problem in face recognition system is the curse of dimensionality. One of the way ay to reduce dimensionality of feature vector is Discrete Cosine Transform (DCT). The DCT converts high-dimensional face images into low-dimensional spaces in which more significant features are maintained. V.p.Vishwakarma et al states that DCT is a popular technique in imaging and video compression, which transforms signals in the spatial representation into a frequency representation

.The DCT of an image basically consists of three frequency components namely low, middle, high each containing some detail and information in an image [19-20].Discrete Cosine Transform is similar to Discrete Fourier Transform [DCT].

The forward 2D-DCT [20] of a $M \times N$ block image is defined as

$$c(u,v) = \alpha(u)\alpha(v)\sum_{N=1}^{M-1}\sum_{N-1}^{M=1} f(x,y)\cos\frac{\Pi(2x+1)u}{2M}\cos\frac{\Pi(2y+1)v}{2N} \dots\dots\dots\dots\dots\dots (4.1)$$

The inverse transform is defined as

$$f(x,y) = \sum_{N=1}^{M-1}\sum_{N-1}^{M=1} \alpha(u)\alpha(v)\cos\frac{\Pi(2x+1)u}{2M}\cos\frac{\Pi(2y+1)v}{2N} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (4.2)$$

Where

$$a_p = \{\tfrac{1}{\sqrt{M}} u = 0, \sqrt{\tfrac{2}{M}} u = 1,2, \dots M-1$$

$$a_q = \{\tfrac{1}{\sqrt{N}} v = 0, \sqrt{\tfrac{2}{N}} v = 1,2, \dots, N-1$$

and x and y are spatial coordinates in the image block, and u and v are coordinates in the DCT coefficients block.

The research is also used Haar Cascade as Face Detection algorithms.

Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real-time face detector. Historically, working with only image intensities (i.e., the RGB pixel values at each and every pixel of image) made the task of feature calculation computationally expensive. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. For example, let us say we have an image database with human faces. It is a common observation that among all faces the region of the eyes is darker than the region of the cheeks. Therefore a common Haar feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object (the face in this case).

The key advantage of a Haar-like feature over most other features is its calculation speed. Due to the use of integral images, a Haar-like feature of any size can be calculated in constant time (approximately 60 microprocessor instructions for a 2-rectangle feature).

Rectangular Haar-like features

A simple rectangular Haar-like feature can be defined as the difference of the sum of pixels of areas inside the rectangle, which can be at any position and scale within the original image. This modified feature set is called 2-rectangle feature. Viola and Jones also defined 3-rectangle features and 4-rectangle features. The values indicate certain characteristics of a particular area of the image. Each feature type can indicate the existence (or absence) of certain characteristics in the image, such as edges or changes in texture. For example, a 2-rectangle feature can indicate where the border lies between a dark region and a light region. This indicate the position of the face and identify face from other objects.

One of the contributions of Viola and Jones was to use summed-area tables, which they called integral images. Integral images can be defined as two-dimensional lookup tables in the form of a matrix with the same size of the original image. Each element of the integral image contains the sum of all pixels located on the up-left region of the original image (in relation to the element's position). This allows to compute sum of rectangular areas in the image, at any position or scale, using only four lookups. The computation of Rectangular face

4.2. Description of Framework Module

A. Capturing an Image

The proposed research is used either Logitech or Webcam of computer to capture face from real time object. Webcam is the normally the laptop frontal camera

Logitech Camera

Logitech has released its new Logitech BRIO 4K Pro Webcam with support for Windows Hello and other secure infrared-based facial recognition applications. This research is used Logitech Camera for detection and Recognition purpose [47].

The webcam is equipped with a second infrared LED and sensor, enabling users to combine it with the primary sensor and configured software to make the device Windows Hello certified so it can be used with the Microsoft security feature to login via facial recognition.

Windows Hello, a key biometric authentication feature of Windows 10, enables people to use their face, iris or fingerprint biometrics to unlock and gain access to Windows 10 devices.

“Microsoft is pleased that Logitech has chosen to bring enterprise-grade security with Windows Hello facial recognition to the BRIO webcam,” said Matt Perry, partner group program manager at Microsoft. “By qualifying BRIO with Windows Hello, Logitech is providing its consumer and enterprise customers with a Windows 10 login experience that is secure, password-free and seamless.”

From [129] Logitech Camera can be described as

- Recognized by Microsoft for Windows 10. Designed for Windows Hello (Windows 10)
- 1 second High speed recognition login your PC with just facing the Infrared camera. It's better to be plugged in to the PC’s built-in USB port (USB 3.0 recommended) directly to get enough data bandwidth. IR camera + RGB camera + Mic need full USB 2.0 data bandwidth to support work with windows hello. (when plugged on the USB hub or Docking Station it may get the "sorry" error when logging in unless they can supply enough data bandwidth)
- HD 720p RGB web cam with Dual Mic for skype ultra-sharp, FaceTime, professional quality video, streaming, webcasting and recording.
- Multi-user support. Identify users with faces even on shared computer such as family and group. Everyone can easily use account differently.
- Masquerade Detection by Infrared Cam with Depth Sensor. High-Security Biometrics. Masquerade by photos and images can be prevented.

Similarly this Face detection and Recognition for Outdoor Environment research is use Logitech camera with USB port and possible to capture high quality image to create image Database.

Figure 4. 2: Logitech Camera

In addition to this the image in Face Database bot Training face and testing face are taken from "ESSEX AC UN" Face Database .These face database "ESSEX AC UN" used for the same purpose. Computer Webcam is also used as image capture under this study.

In general the research take image or through two ways.

1. By Capturing from Real Video and
2. From External Face Database

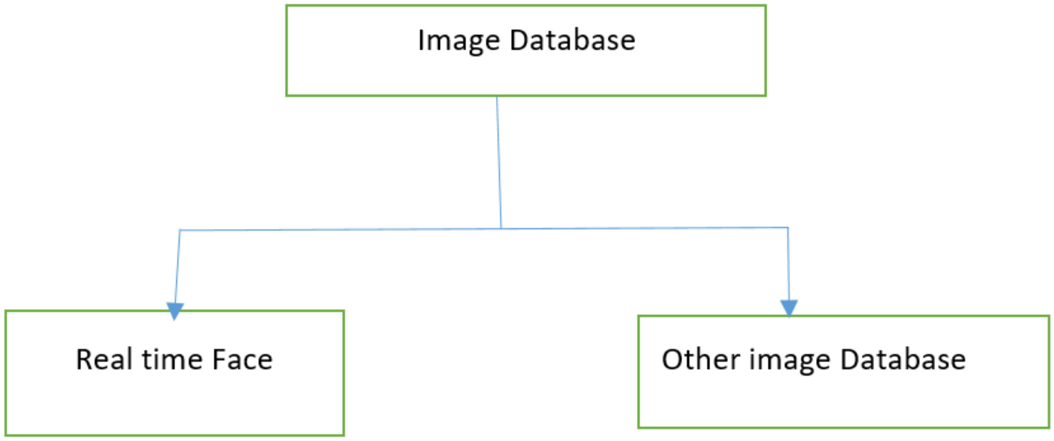

Figure 4. 3: Diagram of Possible creating Face Database

The captured image or the Face image database have an image of 512×512 pixels.

B. Face Detection

In the face Detection and Recognition face detection is used to identify either it is face or not. This research is used Haar Cascade as Face detection algorithm. It make some rectangular image on detected face. Otherwise it did not it. Haar features can be of various height and width. This research is used 16×16 $Pixel$. From the Haar feature applied to the face the sum of black pixel and sum of white pixel are calculated and
they are subtracted to get a single value. A single Haar-like feature, however, is far from enough to build a powerful nonlinear classifier. Cascade of classifiers makes the computation even more efficient, which follows the principles of pattern rejection [22]. Pattern Rejection make program to take only face image.

Haar features are calculated all over the image which will be almost 160000+ features per image. Summing up the entire image pixel and then subtracting them to get a single value is not efficient in real time applications. This can be reduced by using Ada boost classifier. Ada boost reduces the redundant features. The Integral Image is an algorithm for cost-effective generation of the sum of pixel intensities in a specified rectangle in an image of face in face Detection. It is used for rapid computation of Haar-like features. Calculation of the sum of a rectangular area inside the original image is extremely efficient, requiring only four additions for any arbitrary rectangle size. There are five types of Haar cascade [23].

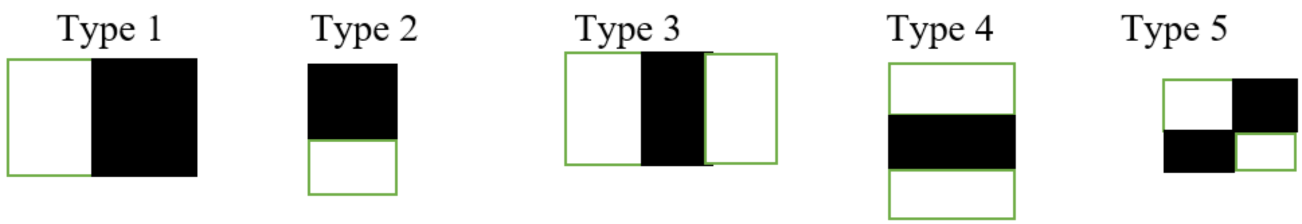

Figure 4. 4: Haar features in Viola Jones

Correct Face Position and Orientation

The Study did not contact any form of correct Position of an individual face either back or front as well as in any degree of inclination. Instead person who are going to register in Face Database are orient his/her position which can detected by camera. The face is detected only when the rea image is at correct position or frontal face is up to about180^0.

Face Recognition

This stage is where biometric human individual face is recognized. The individual identity in the face database is identified either it is face Database or not. If the face is not exist in the Database,

there should the report that face is no exist. Otherwise the real face is identified.

In this study the face Recognition is taken after training and Testing of an individual from database.

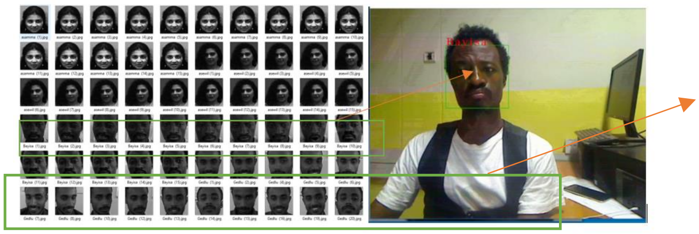

Figure 4. 5 Diagram show the recognized face

Histogram Matching

Histogram Equalization automatically determine the transformation that seeks to produce an output images that has uniform histogram [24]. In this proposed study it used to control brightness and Contrast of the Real-time object with respect to image Database in the database.

Face Sample Database

The collected image face Dataset is collected in the Face sample Database. It contains Training and testing database.

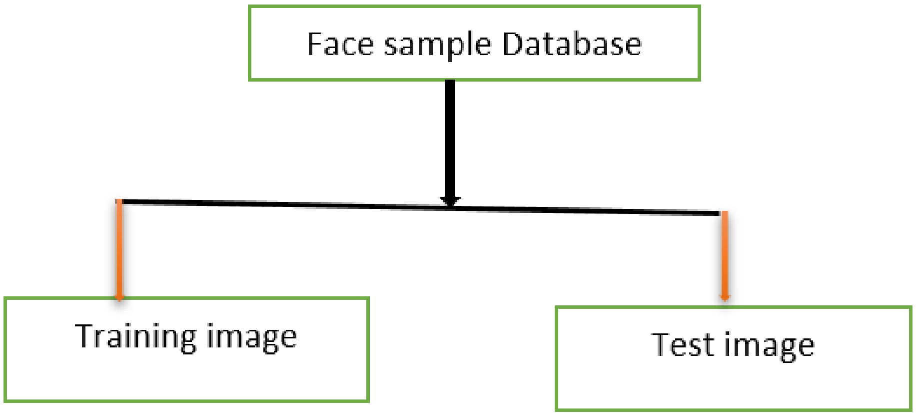

Figure 4. 6: Diagram of face Database

Chapter Five

5.1. Implementation of Proposed Solution

This chapter briefs the environment used in implementing face detection and Recognition for Outdoor Environment. It also describes the software and hardware integration for research implementation.

5.2. Image acquisition

An image capture (image acquisition the process of obtaining a digital image from a vision sensor, such as a camera. This capture used frame grabber, which captures single frames of video, converts the analogue values to digital, and feeds the result into the computer memory [24].The captured image is the stored in the Frame. Some of the image is stored in Image frame and other is stored in Test image. One part of an image is used for training and other for test. The machines or computer is learning every new image from the training and Test.

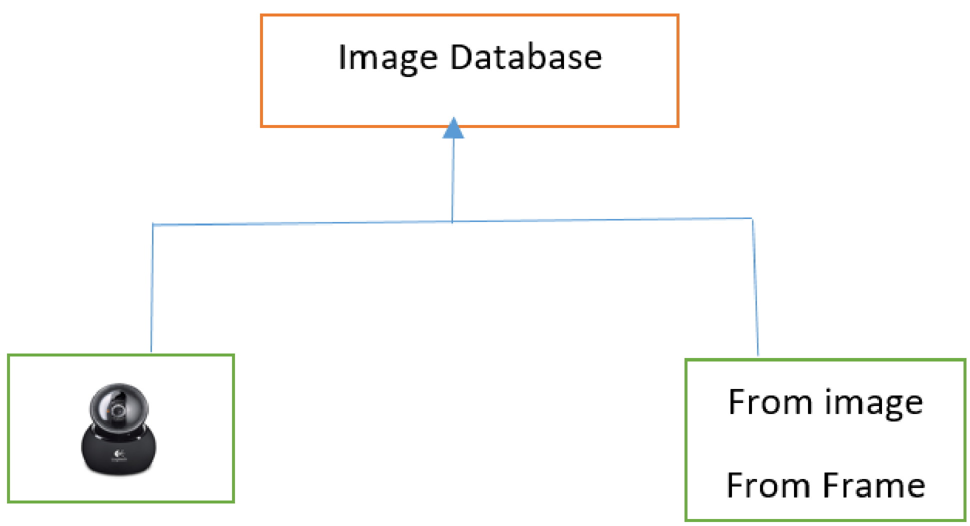

Figure 4. 7: Diagram show the proposed Study Image Capturing

This is OpenCv C++ code to take face image from Real-time

```
void takePicture() {

       int take = 0;
       int co = 0;
       string name;
       cout << "Enter Name: ";
       cin >> name;
       cout << "! click 'p' to capture Image (20 images)!";
       namedWindow("the recogined Person", 1);
```

This is take image from real time by webcam from computer or from different type of External camera. But this research is used Logitech Camera. The image captured is stored in frame. Additionally face image can added from external gallery.

The program is taken twenty image. Out of this images fifteen is for training and five image is for testing. We can also give additional image manually with equal number of an images.

5.3. Convert the image to gray-scale

A digital image usually contains both color information and luminance or grayscale. If you remove the color information, you are left with grayscale, resulting in a black and white image. Grayscale is an important aspect of images, and it is the only portion that is not removed; otherwise, a pure black image would result no matter what color information there is.

A digital image is composed of groups of three pixels with colors of red, green and blue (RGB), also called channels in digital imaging. Each channel also contains a luminance value to determine how light or dark the color is. To get a grayscale image, the color information from each channel is removed, leaving only the luminance values, and that is why the image becomes a pattern of light and dark areas devoid of color, essentially a black and white image.

```
std::vector<Rect> faces;
              UMat frame_gray;
              //pyrDown(frame, frame);
              cvtColor(frame, frame_gray, CV_BGR2GRAY);
```

We used Gray-Scale when

1. Signal to noise. For many applications of image processing, color information doesn't help us identify important edges or other features. There are exceptions. If there is an edge (a step change in pixel value) in hue that is hard to detect in a grayscale image, or if we need to identify objects of known hue. If we don't need color, then we can consider it noise.
2. Complexity of the code. If you want to find edges based on luminance AND chrominance, you've got more work ahead of you. That additional work (and additional debugging, additional pain in supporting the software, etc.) is hard to justify if the additional color information isn't helpful for applications of interest.
3. For learning image processing, it's better to understand grayscale processing first and understand how it applies to multichannel processing rather than starting with full color imaging and missing all the important insights that can (and should) be learned from single channel processing.
4. Difficulty of visualization. In grayscale images, the watershed algorithm is fairly easy to conceptualize because we can think of the two spatial dimensions and one brightness dimension as a 3D image with hills, valleys, catchment basins, ridges, etc. "Peak brightness" is just a mountain peak in our 3D visualization of the grayscale image.
5. Color is complex. Humans perceive color and identify color with deceptive ease.
6. Speed. With modern computers, and with parallel programming, it's possible to perform simple pixel-by-pixel processing of a megapixel image in milliseconds.

5.3. Histogram Equalizer

Histogram is the estimation of the probability distribution of a particular type of data. An image histogram is a type of histogram which offers a graphical representation of the total distribution of the gray values in a digital image [25].The Variation of illumination is equalized to

```
equalizeHist(frame_gray, frame_gray);
```

In a gray level image the probabilities assigned to each gray level can be given by the relation:

$$P_r(r_k) = \frac{n_k}{N} \quad 0 \le r_k \le 1, k = 0,1,2 \dots L-1 \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots .. (6.1)$$

Where r_k-The normalized intensity value

L- No. of gray levels in the image

n_k-No. of pixels with gray level r_k

N -Total number of pixels

$$\sum_{i}^{l-1} p(r_j) = 1$$

The cumulative probability of image pixel distribution$p_f(r_i)$ and the pixel gray value of the output images s_i, is then

$$p_f(r_i) = \sum_{j=0}^{i} p(r_j) \dots \dots \dots (6.2)$$

$$s_i = T(r_i) = \sum_{j=0}^{i} \frac{n_j}{N}, i = 0,1, \dots, L-1, 0 \leq r_i \leq 1 \dots \dots \dots (6.3)$$

The mean and standard Deviation of Pixel image is t

This research is used Histogram equalization method to equalize the variation in dark and light what we call control illumination.

Testing algorithms

```
void test() {
	char name[100];
	cout << "Test Image Name: ";
	cin.ignore();
	cin.getline(name, 100);
	Mat testImage = imread(name);
	Mat t;
	cvtColor(testImage, t, CV_BGR2GRAY);
	equalizeHist(t, t);
	string result = tester(t);
	putText(testImage, result, Point(testImage.cols/4, testImage.rows/ 3), 5, 2,
Scalar(11, 200, 11), 1);
	imshow("Test Result", testImage);
	waitKey();
	cout << result <<endl;
```

The test state is used to test the machine or the computer is understand ever image in the frame.

Chapter Six

6.1. Result, Evaluation and Discussion

In this chapter the result, Evaluation and Finding discussion in are introduced.

6.2. Result

The image is taken either from phot frame or real-time camera.

Figure 6. 1 The original image from Logitech camera and ESSEX AC UN Face Database

Figure 6. 2: The Resized and Gray-Scale converted images

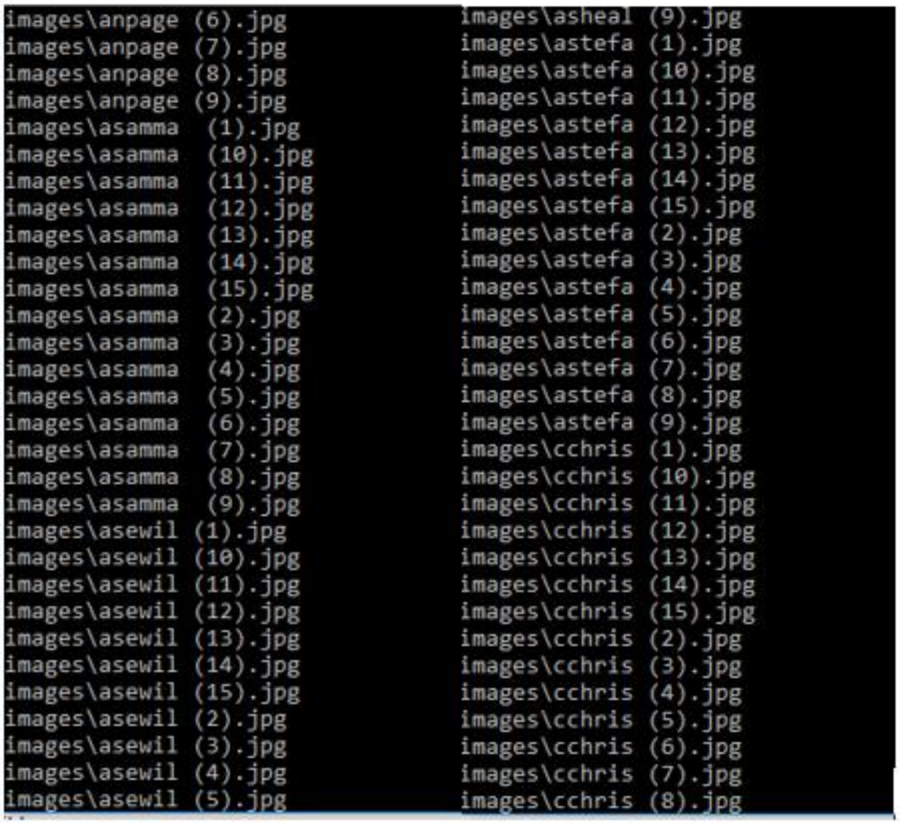
```
images\anpage (6).jpg
images\anpage (7).jpg
images\anpage (8).jpg
images\anpage (9).jpg
images\asamma  (1).jpg
images\asamma  (10).jpg
images\asamma  (11).jpg
images\asamma  (12).jpg
images\asamma  (13).jpg
images\asamma  (14).jpg
images\asamma  (15).jpg
images\asamma  (2).jpg
images\asamma  (3).jpg
images\asamma  (4).jpg
images\asamma  (5).jpg
images\asamma  (6).jpg
images\asamma  (7).jpg
images\asamma  (8).jpg
images\asamma  (9).jpg
images\asewil (1).jpg
images\asewil (10).jpg
images\asewil (11).jpg
images\asewil (12).jpg
images\asewil (13).jpg
images\asewil (14).jpg
images\asewil (15).jpg
images\asewil (2).jpg
images\asewil (3).jpg
images\asewil (4).jpg
images\asewil (5).jpg
images\asheal (9).jpg
images\astefa (1).jpg
images\astefa (10).jpg
images\astefa (11).jpg
images\astefa (12).jpg
images\astefa (13).jpg
images\astefa (14).jpg
images\astefa (15).jpg
images\astefa (2).jpg
images\astefa (3).jpg
images\astefa (4).jpg
images\astefa (5).jpg
images\astefa (6).jpg
images\astefa (7).jpg
images\astefa (8).jpg
images\astefa (9).jpg
images\cchris (1).jpg
images\cchris (10).jpg
images\cchris (11).jpg
images\cchris (12).jpg
images\cchris (13).jpg
images\cchris (14).jpg
images\cchris (15).jpg
images\cchris (2).jpg
images\cchris (3).jpg
images\cchris (4).jpg
images\cchris (5).jpg
images\cchris (6).jpg
images\cchris (7).jpg
images\cchris (8).jpg
```

Figure 6. 3: The snip shoot of implementation of training image code

Figure 6. 4: Diagram of training Images Frame

The Gray scale and equal pixels of 200 images are stored in Testing. Each individual images have 15 for training. The training and testing images are 512×512 pixel.

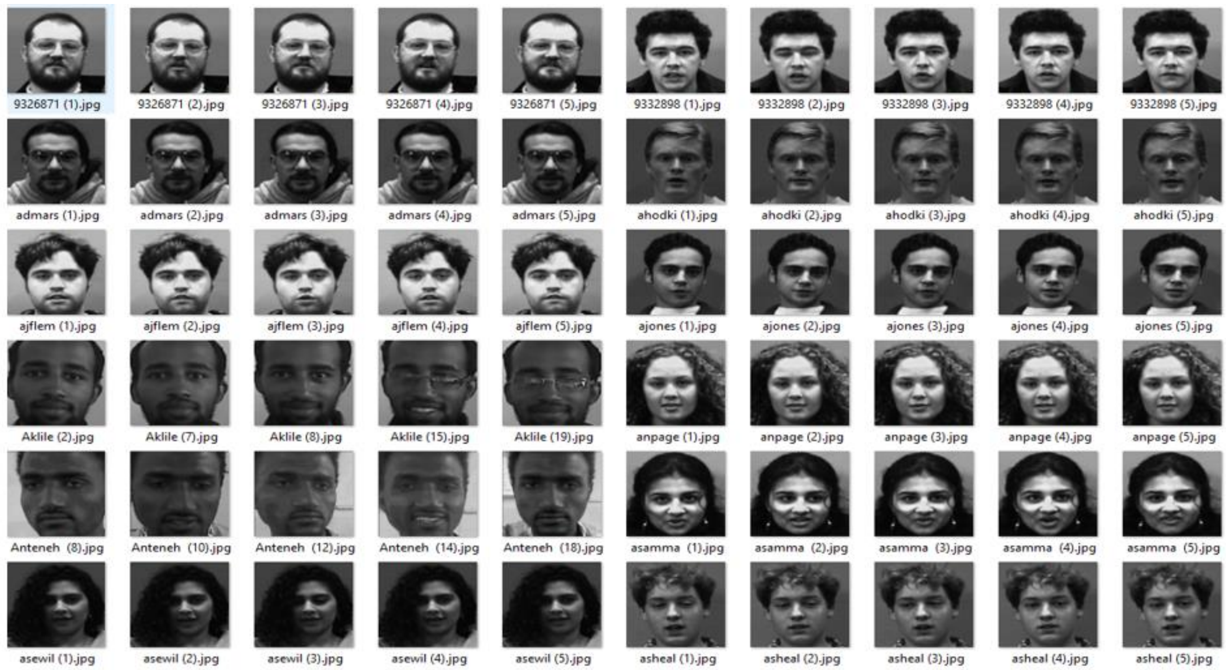

Figure 6. 5 The testing image Database

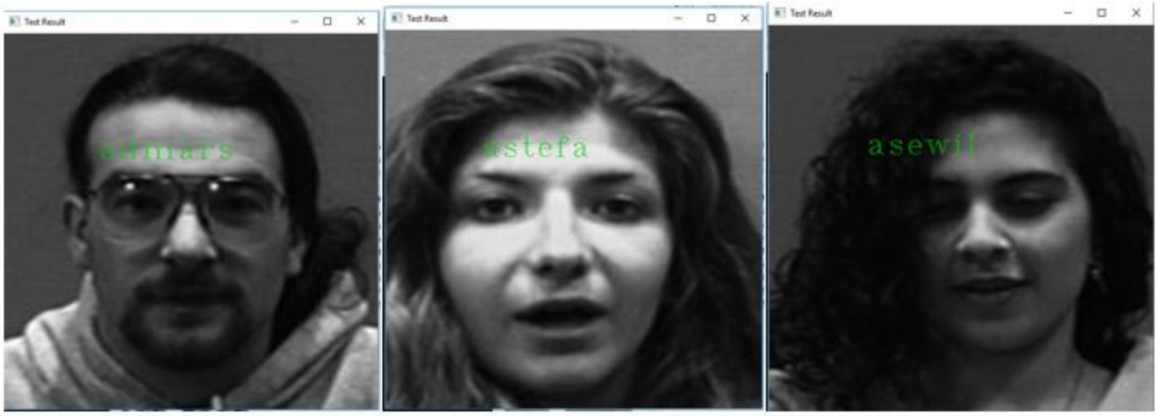

Figure 6. 6: Diagram show testing an image manually

Figure 6. 7: The image in training Frame before Resized and converted to gray scale

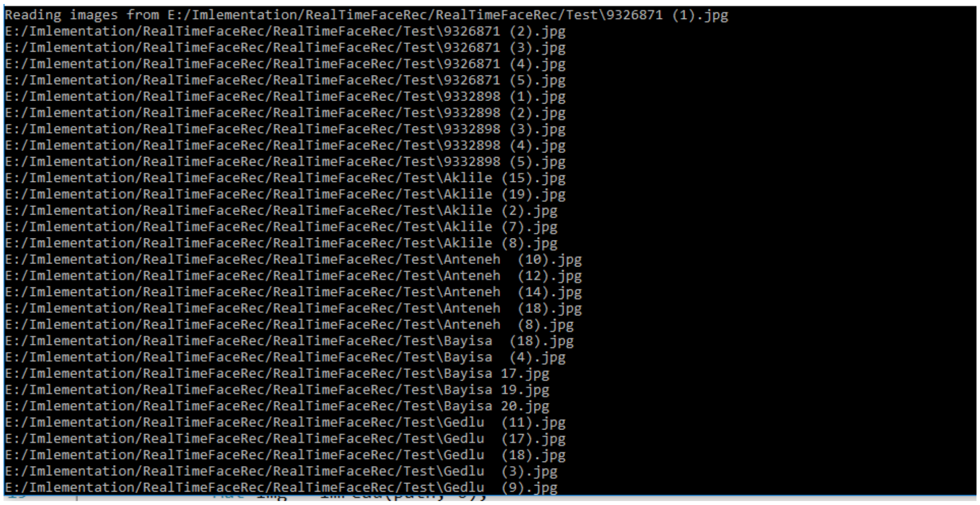

Figure 6. 8: The process of resize and Conversion to gray scale of test frame

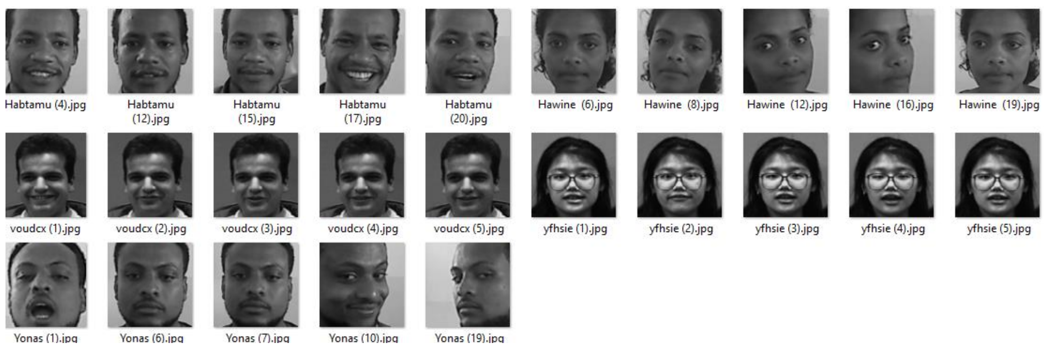

Figure 6. 9: The Resized and gray scale converted images

6.3. **Discussion**

Face Detection and recognition is the hot research yet. Face recognition is used for the identification and authentication of an individual. Different approach and Different algorithms are used to implement for this face Detection and Recognition. All approach and algorithms have different percentage of accuracy.

This research Face Detection and Recognition for outdoor Environment is proposed to improve the accuracy of face in different illumination, Positon and Expression.

This proposed research is take 1000 image from both real-time camera and ESSEX AC UN face database. The image from both two source is jpg extension. The Size of image Pixels is 512×512 that is more size when compared to other. The reason to take 1000 number of image which is

800 for Training and 200 for testing is to increase the number of cluster for an image. It have an implication to increase the number of labels of an images and increase the robustness of Recognition. Gan et al states that increase the number of individual in training and Testing have direct relationship with accuracy of Face Detection and Recognition [56].The selection of high number of an individual at different expression in this studies is also to increase number of labels and training for machines or computer. It have an implication to increase Detection and Recognition rate for this study.

Image Resizer is used to equalize the face image of both captured and from face database. It is worked on all individual image in both training and testing frame.

The manually testing in this proposed Research face is used to check the accuracy of the recognition. The face Detection used in this proposed research is used while we are taking image as input. It helps to identify either it is face or not.

Discrete Cosine Transform (DCT) algorithm is not prone to noise. The noise is reduced as much as possible.

Test images is help us to compare the accuracy and efficiency of the algorithms. The test image is similar to training image despite of number of image.

6.4. Evaluation

Accuracy shows the proportion of true results, both true positives and true negatives.

$$Precision = \frac{TP}{TP + FP} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 6.1$$

$$Fallout = \frac{TN}{TP + TN} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots .. 6.2$$

$$Recall = \frac{TP}{TP + FN} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 6.3$$

$$Specificity = \frac{TN}{FN + FP} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . 6.4$$

Where TP: True Positive are all samples that are classified as the correct class.

FP:False Positive are all samples that are classified as the wrong class.

TN: True Positive are all samples that are classified as unknown correctly.

FN: False Negative are all samples that are classified as unknown wrongly.

From equation (6.1) the Face recognizer compare the images of all individual with respect to all 1000 image in face database. The True Positive is then the accurate precision values that approached to accurate of Recognizing. The proposed algorithm accuracy is 99.9%.This does not mean that it often accurate under all condition. It depend on the quality of camera and image in Training and Testing face. If the image face in both training and testing are taken with not high quality camera and face database, the percent of False Positive (FP) increase and the degree of accuracy decrease. But the average face recognition at different quality of face is 98% of accuracy.

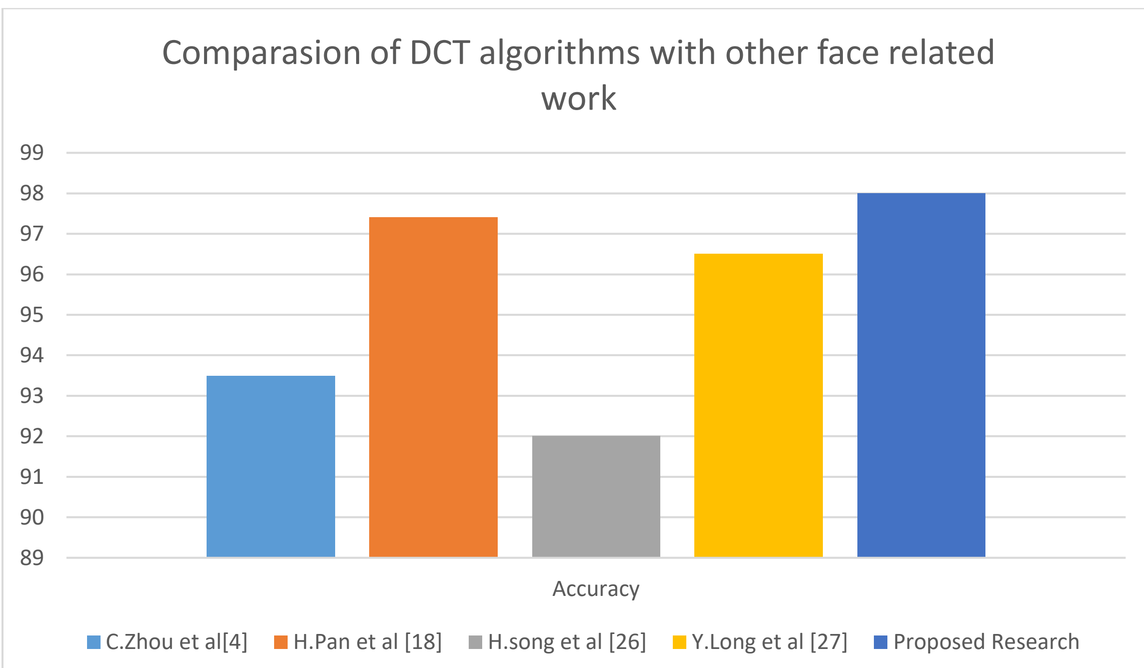

Fig 6.10: Average accuracy of face Detection and Recognition for Proposed DCT algorithms.

From the diagram it is possible to understand the proposed studies better to use as face detector.

6.5. **Conclusion**

The Face Recognition and Detection is the most important and hottest research area computer Vision. Even though many research and project done on this research yet it is accuracy, speed and different degree prone to noise are on progress. [1, 2, 4, 6, 7, 19] are the research worked on Face detection and Recognition. But it is limited on Illumination, Position and Expression. This proposed research DCT for Face Detection and Recognition try to improve in specially speed of image processing and increase the degree of correctness or accuracy. The algorithm is try to detect and recognize at different illumination, Position and expression. Haar Cascade Detector to identify either face or not is used, the detected face are marked with rectangular Haar classifier. Then we can store the detected face image in our Face database if needed and similarly it can used to detect and recognize the real time Video images.

The accuracy of the propose system is 100% if both training and testing in Face Database is high quality, But the average recognition accuracy for all 1000 images of 512×512 Pixel is 98%.

Discrete Cosine Transform algorithms take image of Form jpeg or jpg extension. The Face image use for training and testing are taken both from real time face and from stored Face database.

6.6. **Future work**

For more than two decade different approach and algorithms are used for face detection and Recognition. Each studies are almost works in terms of accuracy, But no more studies related to face Detection and Recognition is worked accuracy with respect to speed to be recognized. My recommendation for Researcher the parameter of Detection and Recognition is not only Accuracy and Error. The speed of Recognition is also very necessary in terms of timeliness.

The problem with Face Detection and Recognition is no more studies are consideration is when the input data are not accurate. This is mean that if the images in training and testing face have high noise it is difficult to recognize .So further studies are needed to work

Reference

1. Y. Zhu , C. Zhu , X. Li , Improved Principal Component Analysis and Linear regression classification for face recognition, Signal Processing(2017) pp 1-21
2. Charles C.-F. Or,Hugh R. Wilson ,Face recognition: Are viewpoint and identity processed after face detection, Vision Research Volume 50, Issue 16, 21 July 2010, PP1581-1589
3. *J.J. Lv, X.H.Shao, Jia-shuiHuang, Xiang-Dong Zhou, Xi Zhou* ,Data augmentation for face recognition, Neurocomputing 230 (2017) 184–196
4. Changjun Zhou,Lan Wang, Qiang Zhang, Xiaopeng Wei,Face recognition based on PCA image reconstruction and LDA ,Optik 124 (2013) 5599–5603
5. R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Trans. Pattern Anal. Mach. Intell. (1993) pp 1-21
6. M. ZHOU,H.LIN,S.S.YOUNG, and, J. YU, Hybrid sensing face detection and registration for low-light and unconstrained conditions, Applied Optics Vol. 57, Issue 1. pp. 69-78 (2018)
7. Kamran Etemad and Rama Chellappa, Discriminant analysis for recognition of human Opt. Soc. Am. A/Vol. 14, No. 8/August 1997,pp 1224-1733
8. A. Pentland, B. Moghaddam, T. Starner, and M. Turk,''View-based and modular eigen spaces for face recognition,''in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition(1994)
9. Y. Zhou, S.T. Zhou, Z.Y.Zhong, and H.-G. Li, A de-illumination scheme for face recognition based on fast decomposition and detail feature fusion, OPTICS EXPRESS 1294 (Optical Society of America, 2013).
10. D .Zeng, S. Long, J. Li, and Qijun, A Novel Approach to Mugshot Based Arbitrary View Face Recognition (Journal of the Optical Society of Korea Vol. 20, No. 2, April 2016), pp. 239-244
11. H.Gan, N.Sang, and R. Huang," Self-training-based face recognition using, Vol. 31, No. 1 / January 2014 / J. Opt. Soc. Am. A pp 1-7
12. Semi-supervised linear discriminant analysis and affinity propagation" (Vol. 31, No. 1 / January 2014 / J. Opt. Soc. Am. A)
13. S. K. Zhou and R. Chellappa," Image-based face recognition under illumination and pose variations" (Vol. 22, No. 2 /February 2005 /J. Opt. Soc. Am. A 217) pp. 218-229

14. D. Cui- Huan, Z. Hong, L. Li-ming, L. Jie, H. Xiang-yang," Face detection in video based on AdaBoost algorithm and skin model" (August 2013).
15. Viola-Jones Face Recognition object Detection algorithm. Proceedings of the 2001 IEEE CVPR 2001
16. H. Boukamch, M. Halleka, F. Smach, M. Atri," Automatic landmark detection and 3D Face data extraction" in the Journal of Computational Science 21 (2017) pp. 340–348
17. DeepFace:-Closing the gap to human level performance in face recognition.(Conference on Computer Vision and Pattern Recognition (CVPR).
18. T. Ahonen, A. Hadid, and M. Pietik ̈ainen," Face Recognition with Local Binary Patterns"(CCV 2004, LNCS 3021), pp. 469–481.
19. H. Pan,, Y. Zhu, L.Xia, Efficient and accurate face detection using heterogeneous feature descriptors and feature selection, Computer Vision and Image Understanding 117 (2013) 12–28
20. V.p.Vishwakarma, S. Pandey, and M.N. Gupta, 'A Novel Approach for Face Recognition Using DCT Re-scaling for Illumination Normalization' 15th International Conference on Advanced Computing and Communication.
21. X.M. Wand, CH. Huang, G.Y. Ni, J.G. Liu, 'Face Recognition Based on Face Gabor Image and SVM
22. Navpreet Kaur Review of Face Recognition System Using MATLAB International Journal of Computer Science Trends and Technology (IJCST) – Volume 4 Issue 3, May - Jun 2016, pp 30-33
23. Stan Z. Li Anil K. Jain Handbook of Face Recognition pp 294-320
24. N. T. Deshpande , Dr. S. Ravishankar , Face Detection and Recognition using Viola-Jones algorithm and Fusion of PCA and ANN, Advances in Computational Sciences and Technology ISSN 0973-6107 Volume 10, Number 5 (2017) pp. 1173-1189
25. https://www.encyclopedia.com/computing/dictionaries-thesauruses-pictures-and-press-releases/image-captureRetrivedSeptember4,2018
26. R. Krutsch and D. Tenorio, Freescale Semiconductor, Histogram Equalization, Rev. 0, June 2011
27. H. Song, S.Lee, J.Kim, and K. Sohn ,Three-dimensional sensor-based face recognition, APPLIED OPTICS pp 677-687

28. Y.Long ,Fan Zhu, L. Shao, J, Han ,Face recognition with a small occluded training set using spatial and statistical pooling, Information Sciences 000 (2017)pp 1–11

29. Li, B. and Chellappa, R. (2001), “Face verification through tracking facial features, Journal of the Optical Society of America A Vol. 18, Issue 12, pp. 2969-2981 (2001)

30. C. Zhou , L. Wang, Q. Zhang, X. Wei Face recognition based on PCA image reconstruction and LDA, Optik 124 (2013) pp 5599–5603

31. A.A. Mohammed , R. Minhas, Q.M. Jonathan Wu, M.A. Sid-Ahmed ,Human face recognition based on multidimensional PCA and extreme learning machine,Pattern Recognition 44(2011) 2588-2597

32. J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, Face recognition using LDA-based algorithms, IEEE Transactions on Neural Networks 14 (1) (2003) 195–200.

33. C. Liu, H. Wechsler, Evolutionary pursuit and its application to face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (6) (2000) 570–582

34. L. Wiskott, J.M. Fellus, N. Kruger, C. VonDerMalsburg, Face recognition by elastic bunch graph matching, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (7) (1997) 775–779.

35. F.R. Bach, M.I. Jordan, Kernel independent component analysis, Journal of Machine Learning Research 3 (2002) 1–48.

36. Hamdi Boukamcha , Mohamed Hallek , Fethi Smach , Mohamed Atri Automatic landmark detection and 3D Face data extraction, Journal of Computational Science 21 (2017) 340–348

37. M. Romero, N. Pears, Landmark localization in 3D face data, 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance vol 1 (2009).

38. A. Mian, M. Bennamoun, R. Owens, An efficient multimodal 2D-3D hybrid approach to automatic face recognition, IEEE Trans. Pattern Anal. Mach. Intell. Mach. Intell. 29 (no. 11) (2007) 1927–1943.

39. O. Celiktutan, B. Sankur, 2D automatic face landmarking, 2008 IEEE 16th Signal Processing, Communication and Applications Conference (2008)

40. https://en.wikipedia.org/wiki/Computer_vision Retrieved June 4, 2018.

41. Richard Szeliski, Computer Vision: Algorithms and Applications (2010)pp 10-11

42. R. Heindl. System und Praxis der Daktyloskopie und der sonstigen technischen Methoden der Kriminalpolizei. De Gruyter, Berlin, 1922.
43. Y. Aloimonos (ed.), Special Issue on Purposive and Qualitive Active Vision, CVGIP B: Image Understanding, Vol. 56 (1992).
44. D. Marr, ``Vision: A Computational Investigation into the Human Representation and Processing of Visual Information'', Freeman, San Francisco (1982).
45. L. Roberts, ``Machine perception of 3D solids", Chapter 9 in J. T. Tippett, et al. (eds), Optical and ElectroOptical Information Processing, MIT Press, pp. 159-197 (1965).
46. M-H Yang, D. J. Kriegman et N. Ahuja, Detecting Faces in Images : A Survey, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 34-58, Jan. 2002
47. Marwa C., Haytham B., Fatma E. S., Mohamed A. , Rached T., Software, Hardware for Face Detection, International Conference on Control, Engineering & Information Technology (CEIT'13)Proceedings Engineering & Technology - Vol.3, pp. 212-215, 2013
48. https://www.biometricupdate.com/201702/logitech-releases-4k-webcam-with-ir-facial-recognition-windows-hello-supportRetrivedAgust6,2018
49. https://opencv.org/about.htmlRetrivedAugust08,2018
50. Gaurav Kumar, Pradeep Kumar Bhatia"A Detailed Review of Feature Extraction in Image Processing Systems" 2014 Fourth International Conference on Advanced Computing & Communication Technologies., Electronic ISBN: 978-1-4799-4910-6
51. https://www.learnopencv.com/opencv-c-vs-python-vs-matlab-for-computer-vision/RetrivedSeptember18,2018
52. Guan, A. X. and Szu, H. H. (1999), "A Local Face Statistics Recognition Methodology beyond ICA and/or PCA", IEEE, Kansas City, pages 1016-1027.
53. Cox, I. J., Ghosn, J., and Yianilos, P. N. (1996) "Feature-based face recognition using mixture distance", In Proceedings, IEEE Conference on Computer Vision and Pattern
54. Cagnoni, S. and Poggi, A. (1999) "A Modified Modular Eigenspace Approach to Face Recognition", IEEE, Kansas City, pages 490-495.
55. Wiskott, L., Fellous, J. M., Krüger, N. and von der Malsburg, C. "Face Recognition by Elastic Bunch Graph Matching", IEEE, pages 129-132.

56. H. Gan, N. Sang, and R. Huang, Self-training-based face recognition using semi-supervised linear discriminant analysis and affinity propagation, Vol. 31, No. 1 / January 2014 / J. Opt. Soc. Am. A, pp 1-6

Appendix

OpenCV C++ Program for training

```
        // read all the training images
        readImages("images/", trainingImages, trainLabels, trainLabelIndex);

        //get all the dct of the images
        getAllDCT(trainingImages, features);
        saveDCT("dct.txt");
}

void Evaluate() {

        // read all the training images
        readImages("Test/", testImages, testLabels, testLabelIndex);

        for (Mat testImage : testImages) {
                resultLabels.push_back(tester(testImage));
        }
        double counter = 0;
             cout << "      " << "ACTUAL" << "          " << "PREDICTED" << endl;
        for (int i = 0; i < resultLabels.size(); i++) {
                cout <<"      " <<resultLabels[i] << "          " <<
testLabels[testLabelIndex[i]] <<endl;
                if (resultLabels[i] == testLabels[testLabelIndex[i]]) {
                        counter++;
```

The performance of an algorithms is measured

```
double s = testImages.size();
       double k = (counter / s);
       cout << "     Accuracy :    " <<  k * 100 << "% " << endl;
```

```
To Test image we used
```

```
oid test() {
       char name[100];
       cout << "Test Image Name: ";
       cin.ignore();
       cin.getline(name, 100);
       Mat testImage = imread(name);
       Mat t;
       cvtColor(testImage, t, CV_BGR2GRAY);
       equalizeHist(t, t);
       string result = tester(t);
       putText(testImage, result, Point(testImage.cols/4, testImage.rows/ 3), 5, 2,
Scalar(11, 200, 11), 1);
       imshow("Test Result", testImage);
       waitKey();
       cout << result <<endl;To capture an image from camera
```

```
VideoCapture capture(0);
       UMat frame;

       while (capture.read(frame))
```

```
	{
		std::vector<Rect> faces;
		UMat frame_gray;
		//pyrDown(frame, frame);
		cvtColor(frame, frame_gray, CV_BGR2GRAY);

		equalizeHist(frame_gray, frame_gray);

To Detect Face

face_cascade.detectMultiScale(frame_gray, faces, 1.1, 2, 0 | CV_HAAR_SCALE_IMAGE,
Size(20, 20), Size(200, 200));

		for (size_t i = 0; i < faces.size(); i++)
		{
			UMat theImag(frame_gray, faces[i]);
			Mat theImage = theImag.getMat(1);
			resize(theImage, theImage, Size(512, 512));
			imshow("some", theImage);
			String label = tester(theImage);

			rectangle(frame, Point(faces[i].x, faces[i].y), Point(faces[i].x +
faces[i].width, faces[i].y + faces[i].height), Scalar(7, 201, 7), 1);
			putText(frame, label, Point(faces[i].x, faces[i].y), 5, 1.2,
Scalar(10, 00, 250), 1);
		}
		//-- Show
		imshow(window_name, frame);

		int c = waitKey(10);
		if ((char)c == 'c') { return; }

	}

	waitKey();
}

void takePicture() {

	int take = 0;
	int co = 0;
	string name;
	cout << "Enter Name: ";
	cin >> name;
	cout << "! click 'p' to capture Image (20 images)!";
	namedWindow("the recogined Person", 1);

	VideoCapture capture(0);
	UMat frame;

	while (capture.read(frame))
	{

		std::vector<Rect> faces;
		UMat frame_gray;
		//pyrDown(frame, frame);
```

```
            cvtColor(frame, frame_gray, CV_BGR2GRAY);

            equalizeHist(frame_gray, frame_gray);
```

When DCT Algorithms is applied

```
tring tester(Mat testImage) {
       Mat testCof;

       getDCT(testImage, testCof);

       double maxMatch = 0;
       int i = 0;
       int index = 0;
       for (Mat cof : features) {

              double maxVal = correlation(testCof, cof);

              if (maxVal > maxMatch) {
                     maxMatch = maxVal;
                     index = i;
              }
              i++;
       }
       //cout << maxMatch << endl;
       if (maxMatch > 0.7) {
              //cout << index << endl;
              //cout <<trainLabelIndex[index] << endl;
              return trainLabels[trainLabelIndex[index]];
       }
       else {
              return "UNKNOWN";
       }

}

double correlation(Mat &image1, Mat &image2) {

       // convert data-type to "float"
       Mat im_float1;
       image1.convertTo(im_float1, CV_32FC1);
       Mat im_float2;
       image2.convertTo(im_float2, CV_32FC1);

       int n_pixels = im_float1.rows * im_float1.cols;

       // Compute mean and standard deviation of both images
       Scalar im1_Mean, im1_Std, im2_Mean, im2_Std;
       meanStdDev(im_float1, im1_Mean, im1_Std);
       meanStdDev(im_float2, im2_Mean, im2_Std);

       // Compute covariance and correlation coefficient
       double covar = (im_float1 - im1_Mean).dot(im_float2 - im2_Mean) / n_pixels;
       double correl = covar / (im1_Std[0] * im2_Std[0]);

       return correl;
}
```

```
void getDCT(Mat theImage, Mat &maxCof) {
	theImage.convertTo(theImage, CV_32FC1, 1.0 / 255);


	int height = theImage.size().height;
	int width = theImage.size().width;
	Mat dctImage = theImage.clone();
	Mat d = Mat::zeros(theImage.size(), theImage.type());
	//cout << theImage.type() << endl;


	for (int r = 0; r< height; r += 16) {
		for (int c = 0; c < width; c += 16) {

			// get the dct of the block
			dct(dctImage(Rect(c, r, 16, 16)), d(Rect(c, r, 16, 16)));

			double minv, maxv;
			// find the maximum cofficient in the block
			minMaxLoc(d(Rect(c, r, 16, 16)), &minv, &maxv);
			maxCof.push_back(maxv);
		}
	}

}

void getAllDCT(vector<Mat> &images, vector<Mat> &cof) {

	for (Mat theImage : images) {
		cout << "cool" << endl;


		int height = theImage.size().height;
		int width = theImage.size().width;
		Mat dctImage = theImage.clone();
		dctImage.convertTo(dctImage, CV_32FC1, 1.0 / 255);
		Mat d = Mat::zeros(dctImage.size(), dctImage.type());
		Mat maxCof;

		for (int r = 0; r< height; r += 16) {
			for (int c = 0; c < width; c += 16) {

				// get the dct of the block
				dct(dctImage(Rect(c, r, 16, 16)), d(Rect(c, r, 16, 16)));

				double minv, maxv;
				// find the maximum cofficient in the block
				minMaxLoc(d(Rect(c, r, 16, 16)), &minv, &maxv);
				maxCof.push_back(maxv);
			}
		}

		cof.push_back(maxCof);

	}
```

OpenCV

OpenCV (Open Source Computer Vision Library) [129] is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. Being a BSD-licensed product, OpenCV makes it easy for businesses to utilize and modify the code.

The OpenCV has more than 2500 optimized algorithms libraries, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc. OpenCV has more than 47 thousand people of user community and estimated number of downloads exceeding 14 million. The library is used extensively in companies, research groups and by governmental bodies.

Along with well-established companies like Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota that employ the library, there are many startups such as Applied Minds, VideoSurf, and Zeitera, that make extensive use of OpenCV. OpenCV's deployed uses span the range from stitching streetview images together, detecting intrusions in surveillance video in Israel, monitoring mine equipment in China, helping robots navigate and pick up objects at Willow Garage, detection of swimming pool drowning accidents in Europe, running interactive art in Spain and New York, checking runways for debris in Turkey, inspecting labels on products in factories around the world on to rapid face detection in Japan.

It has C++, Python, Java and MATLAB interfaces and supports Windows, Linux, Android and Mac OS. OpenCV leans mostly towards real-time vision applications and takes advantage of MMX and SSE instructions when available. A full-featured CUDA and OpenCL interfaces are being actively developed right now. There are over 500 algorithms and about 10 times as many functions that compose or support those algorithms. OpenCV is written natively in C++ and has a template interface that works seamlessly with STL containers [129].This is the reason why selected for this study.

Figure 3. 2: Diagram of openCV

OpenCV (Open Source Computer Vision Library) is released under a BSD license and hence it's free for both academic and commercial use. It has C++, Python and Java interfaces and supports Windows, Linux, Mac OS, iOS and Android. OpenCV was designed for computational efficiency and with a strong focus on real-time applications. Written in optimized C/C++, the library can take advantage of multi-core processing. Enabled with OpenCL, it can take advantage of the hardware acceleration of the underlying heterogeneous compute platform.

Acronyms

ACRONOYMS	Abbreviation
CV	Computer Vision
PSO	
JPEG	Joint Photographic Experts Group
JPG	Joint Photographic Group
DCT	Discrete Cosine Transform
AFR	Automatic Face Recognition
3D	3-Dimension
FAR	False Acceptance Rate
FRR	False Rejection Rate
CIR	Correct Identification Rate
FIR	False Identification Rate
NNM	Neural Network Model
AAM	Active Appearance Model
EBGM	Elastic Bunch Graph Matching
PCA	Principal Component Analysis
CNN	Convolutional Neural Networks
PIE	Position, Illumination and Expression
E-COMMERCE	Electronic Commerce
ATM	
LDA	Linear Discriminant Analysis
SVM	Support Vector Machine
ANN	Artificial Neural Network
charge coupled device	CCD
PLDA	Probabilistic Linear Discriminant Analysis
GH	Generalized Haar-like
MB-LBP	Multi-Block Binary Patterns
SURF	Speed-up Robust Features
GPU	Graphics Processing Unit
B2DPCA	bidirectional two dimensional principal component analysis
ELM	Extreme Learning Machine
FLDA	Fisher’s linear discriminant analysis
MPCA	Multilinear Extension of Principal Component analysis
EP	Evolutionary Pursuit
EBGM	Elastic bunch graph matching
DFT	Discrete Fournier Transform
LED	Light Emitting Diode
RGB	Red, Green, Blue
TP	True Positive
TN	True Negatives
FP	False Positive

FN	False Negative
IR	Infra Ray
AP	Affinity Propagation