A Machine Learning Approach to Facial-Based Ethnicity Classification

by

Hajra Mehbub Momin

Submitted in fulfilment of the academic

requirements for the degree of

Doctor of Philosophy

in the

School of Mathematics, Statistics and Computer Science,

University of KwaZulu-Natal

Durban, South Africa,

2017

UNIVERSITY OF KWAZULU-NATAL

FACULTY OF SCIENCE AND AGRICULTURE

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

This is a thesis in which chapters are written as a set of discrete research papers, with an overall introduction and final discussion, where one (or all) of the chapters have already been published. Typically these chapters will have been published in internationally recognized, peer-reviewed journals and conference papers.

Signed: ________________________________

Hajra Mehbub Momin

Date: ....

As the candidate's supervisor, I have approved the dissertation for submission

Signed: ________________________________

Prof. Jules-Raymond Tapamo

Date: ....

UNIVERSITY OF KWAZULU-NATAL

FACULTY OF SCIENCE AND AGRICULTURE

DECLARATION 1 - PLAGIARISM

I, **Hajra Mehbub Momin** declare that

1. The research reported in this thesis, except where otherwise indicated, and is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    - Their words have been re-written but the general information attributed to them has been referenced.
    - Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed: ______________________________________

Hajra Mehbub Momin

UNIVERSITY OF KWAZULU-NATAL

FACULTY OF SCIENCE AND AGRICULTURE

DECLARATION 2 - PUBLICATIONS

DETAILS OF CONTRIBUTION TO PUBLICATIONS that form part and/or include research presented in this thesis (including publications in preparation, submitted, in press and published and giving details of the contributions of each author to the experimental work and writing of each publication)

1. Hajra Momin and Jules R. Tapamo, "Automatic Detection of Face and Facial Landmarks for Face Recognition", *Proceeding of the International Conference on Signal Processing, Image Processing and Pattern Recognition (SIP 2011)*, South Korea, pp. 244-253, December 2011.

2. Hajra Momin and Jules R. Tapamo, "A comparative study of face based ethnicity classification using colour spaces", *EURASIP Journal on Image and Video Processing*, (to be submitted).

3. Hajra Momin and Jules R. Tapamo, "Gabor Filter Based Texture Analysis of Facial Components for Ethnicity Classification", *Appl. Math. Inf. Sci.* vol 10, No 6, pp.2255-2265, November 2016.

4. Hajra Momin and Jules R. Tapamo, "Ethnicity Classification using fusion of Colour and Texture Features", *Pattern Analysis and Applications*, (to be submitted).

The author's contributions in each of the papers are as follows:

1^{st} Author : Literature review, design and implementation of algorithms, write-up of the paper.

2^{nd} Author : Giving ideas, providing advice, discussing issues on models and algorithms, proof-reading manuscripts.

Signed: ____

Hajra Mehbub Momin

To My Family,

Specially to my Husband,

Mehbub Momin
My Mother and Father

Contents

List of Figures

List of Tables

List of Abbreviations

ANN	Artificial Neural Networks.
ASM	Active Shape Model.
BIF	Biologically Inspired Features.
BIM	Biologically Inspired Model.
DET	Detection Error Trade-off.
DNA	Deoxyribo Nucleic Acid.
EBGM	Elastic Bunch Graph Matching.
EER	Expected Error Rate.
EPC	Expected Performance Curve.
FAR	Failure to Acquire Rate.
FER	Failure to Enroll Rate.
FPR	False Positive Rate.
FRR	False Rejection Rate.
FTC	Failure to Capture.
GDA	General Discriminant Analysis.
GEI	Gait Energy Image.
GMM	Gaussian Mixture Model.
GWT	Gabor Wavelet Transformation.
HCI	Human Computer Interaction.
HMM	Hidden Markov Model.

HTER	Half Total Error Rate.
ICA	Independent Component Analysis.
IDT	Inductive Decision Tree.
KCFA	Kernel Class-dependent Feature Analysis.
KFDA	Kernel Fisher Discriminant Analysis.
KNN	K-nearest neighbour.
KPCA	Kernel Principal Component Analysis.
LBP	Local Binary Patterns.
LCP	Local Circular Patterns.
LDA	Linear Discriminant Analysis.
LFA	Local Feature Analysis.
LVC	Learned Visual Codebook.
ML	Machine Learning.
MLP	Multilayer Perceptron.
MM-LBP	Multi-scale Multi-ratio Local Binary Pattern.
MORPH	Craniofacial Longitudinal Morphological Face Database.
NB	Naive Bayesian.
OGMs	Oriented Gradient Maps.
PCA	Principal Component Analysis.
PDF	Probability Density Function.
PIN	Personal Identification Number.
RBF	Radial Basis Function.
ROC	Receiver Operator Characteristics.
ROI	Region of Interest.
SVM	Support Vector Machine.
TPR	True Positive Rate.

WER	Weighted Error Rate.
WLD	Weber Local Descriptors.

Abstract

The determination of ethnicity of an individual can be very useful in a face recognition and person identification system in general. The face displays a complex range of information about identity, age, sex, race as well as emotional and intentional state. It is commonly assumed that the biological unit of human classification is the ethnic group, with hereditary physical features making up the group classification, based on the qualities such as the skin colour, the build, the head shape, the hair, the face shape, and the blood type.

In this thesis, the aim is to investigate methods and techniques to perform ethnicity classification of face images. Automatic face-based ethnicity classification has various applications in human computer interaction, surveillance, video and image retrieval, database indexing, and can give helpful insight for face recognition and identification. Since biometric systems have to deal with very large databases, it can be a good idea to partition the face database according to the ethnicity of a person. In addition, this has the potential to significantly improve the search speed, efficiency and accuracy of biometric systems.

Automatic face and landmark detection on images is very important for face recognition, face identification and for ethnicity classification. This study presents an approach for detecting face and facial features such as the eyes, the nose and the mouth in gray-scale images. In addition, the study makes use of thresholding and connected component labelling algorithms in order to detect a face and extract features that characterize this face.

This study investigates three different feature methods for the ethnicity classification of face images. A new ethnicity classification based on skin colour is proposed. Skin colour is one of the most important features in the human face. The skin colour differs from individual to individual belonging to different ethnic groups and from people across different regions. For instance, the skin colour of people belonging to White, Asian and Black groups is different from one another and extended from white to yellow to dark brown. Based on this different colour spaces are used to create a feature vector representing a given face image. A second feature model based on textures is proposed. Gabor filters are used to extract texture features. Thirdly, a combination of colour and texture features are used to further improve the ethnicity classification accuracy.

Four different classifiers, namely K-Means clustering, Naive Bayesian (NB), Multilayer Perceptron (MLP) and Support Vector Machine (SVM), were used to test the effectiveness of the automatic characterization of ethnicity by using the proposed features models. The ethnic groups considered were Asian, Indian, White and Black. Extensive experiments demonstrate that our models achieve very good results, confirming the consistently overwhelming performance of Asian classification. The proposed models also achieve very good classification results for different ethnic groups when compared with existing models.

Acknowledgement

It gives me great pleasure in expressing my gratitude to all those people who have supported me and made their contributions in making this thesis possible. First and foremost, I must acknowledge and thank The Almighty Allah for blessing, protecting and guiding me throughout this period. I could never have accomplished this without the faith I have in the Almighty.

I express my profound sense of reverence to my supervisor and promoter Prof. Jules Raymond Tapamo, for his constant guidance, support, motivation and untiring help during the course of my PhD. His in-depth knowledge on a broad spectrum of different topics has been extremely beneficial to me. He has given me unlimited freedom during my research, and his pleasant disposition towards me is much appreciated. His simplicity, sense of endeavour, and research skills have inspired me forever.

Most of all, I would like to thank Prof. Kesh Govinder (the Dean and Head of the School of Mathematics, Statistics and Computer Science), Prof. Henry Mwambi (the Academic Leader of Research and Higher Degrees) and Prof. Jacek Banasiak for their great support given to the submission of this thesis. Without their support my work might not have ever come to fruition. I cannot adequately express how thankful I am.

This thesis would not have been possible without the emotional support of my family: a very special thank you to my husband, for your continued and unfailing love, support and understanding that underpins my persistence in the postgraduate career. You have made the completion of this thesis possible. A special thanks goes to my dear mother for her undying love, care and affection. Thank you to my father Khwaja Momin for his constant encouragement and helpful advice. A special thanks to my brothers who are always giving me the necessary support that I require.

I would also like to thank all my colleagues, Soren Greenwood and Jay Moodley for their technical support and, Laila and Shereen who have helped me in a number of ways. I am pleased to express my gratitude to all the faculty and staff members of the School of Mathematics, Statistics and Computer Science, in the University of KwaZulu-Natal, for their valuable co-operation.

Finally, I thank all those who have helped me directly or indirectly in the successful completion of my thesis. Anyone whom I have missed in this acknowledgement is also thanked.

Chapter 1

Introduction

1.1 Background

Currently security is a priority for individual use, institutions, and organizations. Numerous government agencies are constantly improving their security frameworks as an after-effect of the expansion in terrorist assaults. The most important asset of the business is information, hence security plays a vital part in maintaining the secrecy of important information. Traditionally, different authorization techniques are used in the day-to-day life. For example, passwords are utilized in order to gain access to the online banking, cell phone or office systems. Signatures are widely utilized for personal authentication and electronic cards are used to gain access into buildings and systems. However this kind of authentication technique has numerous constraints. Securing information using a password has become less and less dependable on the grounds that they can be easily cracked, especially when individuals use their birth date or nick-name as their password or even forget their password. Electronic cards can be easily stolen, while signatures can also be imitated. These constraints can make these techniques less dependable and inefficient [1–4].

Biometrics is rapidly turning into another option for passwords and cards. Iris, finger print, face, and other biometric features are progressively being used for authentication, and are difficult to be stolen, borrowed and practically impossible to forget. These advances in authentication use the intrinsic aspects of a particular individual, and hence appear as a viable alternative to more traditional approaches, such as Personal Identification Number (PIN) codes or passwords [2]. Biometrics offers automated techniques of identity verification on the principle of measurable physiological or behavioral characteristics such as fingerprint, face, iris, signature, palm, hand geometry, or speech [3]. These technologies are being used in security banking systems, mobile phones, immigration, health, authentication systems, and many other applications [5].

A substantial amount of research on different biometric advancements has been presented in literature. The human face has been a subject of interest to a growing number of research endeavours in recent times. The reason is that the face is the most natural and acceptable for identity authentication [6]. Automatic identification of the human face is a very challenging research topic and, has received increased consideration in recent years. However, human faces do not only provide the identities, but also the demographic information, such as gender, age and ethnicity, which are also vital for human face perception. When we take a glimpse at an individual's face, we not only perceive who it is, but also process other information about the person, such as the expression, ethnicity, shape of the eyes and nose, age, gender and charm [7]. In the course of recent decades, a great deal of effort has been committed in the areas of the biological, psychological, and cognitive sciences in order to discover how the human brain perceives, represents and remembers faces. Similarly, it is hoped that the computer can automatically understand these messages from facial images [8].

Among the demographic attributes, ethnicity generally remains invariant throughout an individual's life time and significantly supports face identification and recognition systems. Therefore, automatic facial ethnicity classification has received increasing attention in recent years. Furthermore, it has promising potential to be applied for improvement in amongst others, database indexing, Human Computer Interaction (HCI), video/image retrieval, and surveillance.

1.2 Motivation

The importance of face recognition is appealing on the grounds that the face recognition system does not require the cooperation of the individual compared to many other systems. A face has unique features for human identification. Even identical twin's faces are somewhat not the same [9]. The human face has the most common characteristic that is utilized by people in order to distinguish or perceive other individuals. Face recognition has distinct advantages because of its non-contact process. Face images can be captured from a distance without touching or interaction with the person being identified [10]. In addition, face recognition serves crime deterrent purposes because face images that have been recorded and archived can later help to identify a person. It has become an important research area in many real-life machine vision applications which includes amongst others, access control, bank card identification, mug shots searching, security monitoring, surveillance systems, passports, suspect tracking and investigation, and human computer interaction.

Human facial images provide the demographic information, such as the ethnicity, gender and age of an individual. Ethnicity and gender additionally play an important role in face-related applications such as face recognition and identification [11, 12]. Humans without much of a stretch see demographics through facial appearance such as skin colour, texture, contour information, location and distance between facial features, and geometric features (such as eyes, nose, lips, face shape and hair). They do not recollect all the features (all pixel values in face image) of the face but only general features such as skin colour, texture, and contour information. Yet they are able to identify demographic information efficiently and successfully. This influences the belief that all image data is not necessary and that only a couple of selected features can prompt effective ethnicity classification which can speed up face classification.

As known from anthropometric, physiological, and psychological studies, the human brain stores a part-based representation of the face. These components play an important role not only in detection and recognition but additionally in gender, age, and ethnicity classification [13, 14]. What makes face detection and identification a difficult task are different facial expressions, pose, occlusion, vague and age. [10]. However, components usually vary less under pose changes than the image pattern of the entire object. A component based approach may be more vigorous against partial occlusions than a global approach [15]. Extending the concept of facial features to ethnicity raises a number of fascinating issues, particularly concerning the characteristics of face components, and its impact on the performance of the computer in recognizing and characterizing ethnicity.

When there is an interaction with other individuals, visual information plays a critical part. Humans can distinguish the ethnicity of a person by merely taking a glance at skin colour. Skin colour is one of the most imperative element of the human face. It is not merely a shift from individual to individual belonging to different ethnic groups, it is also across different regions. For instance, the skin colour of persons belonging to White, Asian and Black groups range from white to yellow to dark brown. Motivated by this, we investigated ethnicity classification by utilizing skin colour. Moreover, features such as the eyes, nose, mouth, hair and the face shape are distinct for different ethnic groups. The details of the physical facial characteristics of the three major ethnic groups namely Asian, White and Black are presented in [7]. These features can likewise play an important role in the ethnicity classification.

A biometric system needs to manage a very large database for face recognition and identification. Therefore, it is a good idea to partition the database before performing recognition or identification tasks. When a database that exceeds its size is considered, the performance of many state-of-the-art face recognition methods deteriorates rapidly. A successful face classification algorithm can significantly boost the performance of other applications for example face recognition and identification. Face recognition process will be greatly enhanced if the database of face images is separated according to ethnic groups. Thereafter instead of searching for the individual in the entire single database a search can be carried out for the images in the appropriate ethnicity database. It will likewise enhance the search speed and efficiency of the retrieval system [16].

1.3 Problem Statement

Numerous results on automatic face detection and face recognition have been published [9, 15, 17–19]. However, very few attempts have been made to perform human ethnicity classification. Retrieval of images from the human face database based on demographic characteristics such as gender and race can be very use-full in wide-ranging applications such as face detection and identification, but it remains poorly studied. The face shows a complex range of information about identity, age, sex, race as well as emotional and attentional state.

In this thesis, the aim is to investigate methods and techniques to perform ethnicity classification based on face images by utilizing diverse facial features. Skin colour stands out among the most important facial features in classifying images in different ethnic groups. People from different ethnicity groups have distinct skin colour which serve as an essential feature in ethnicity classification.

Facial components such as the eyes, nose and mouth are the most essential parts in a face image, but there are difficulties in recognizing appropriate facial components. These parts have the most imperative information about the person's age, gender and ethnicity. They play a vital role in face detection and recognition as well as in ethnicity classification. In spite of the fact that there is public literature on face recognition that utilizes facial components [20–22], there is a need to study these facial components and their impact on the performance of computer-based ethnicity classification. People of different ethnic groups are unique in relation to one another by geographic location and distance between facial features (face components) [23–25].

While a single facial element can give promising results in ethnicity characterization, there is still room for enhancing the accuracy rate by using different facial features. A combination of features can significantly improve the ethnicity classification task when compared to a single facial feature. For the most part in literature, face images are classified into two ethnic classes either Asian and non-Asian or White and Black.

1.4 Thesis Objective

The main aim of this research work is to solve the difficult task of robust ethnicity classification in a simple setup.

The specific objectives of this thesis are:

- To discuss how to detect face and facial components (the eyes, nose and mouth) from gray-scale images in frontal, left and right face view and moreover, to extract facial components from images with different facial expression such as smiling, sadness, when the eyes are closed or when the mouth is open.
- To design and develop a robust ethnicity classification model using distinctive facial features.
- To compare and contrast the facial features (colour and texture) on the basis of performance on ethnicity classification. To compare the fusion of facial features results with a single facial feature. Furthermore, to compare the implemented technique with the techniques that are available in public literature.

1.5 Contribution of Thesis

The major contributions of this thesis include the following:

A segmentation approach used to extract the face region from the image is introduced. After detecting the face in an image the following stage is to extract facial features such as the eyes, the nose and the mouth from gray-scale images using binarization and a connected component labelling algorithm in order to tackle distinctive face related issues. The proposed method detects facial landmarks from a frontal, left and right face view. In addition, it detects facial landmarks in the expressive images of a high complexity.

An exceptionally straightforward and efficient strategy is introduced for ethnicity classification of face images using RGB, HSV and YCbCr colour space models. K-means and NB classifiers are used to classify an image into proper ethnicity. The colour space component that produces the best result is evaluated. The presented method is very simple yet offers promising results.

A Gabor filter based texture analysis of facial components for ethnicity classification is presented. In this approach, firstly the facial components for example the eyes, the nose and the mouth are identified in the face image. A Gabor filter is then applied to each individual face component in order to capture the texture information and a feature vector is created for all face components. We study the contribution of various facial components in the context of ethnicity classification. We additionally answer the question "is it conceivable to recognize the ethnicity of a person considering facial components? Four different machine learning techniques were used and compared to evaluate results.

The fusion of facial features was also considered (this includes colour and texture features) for ethnicity characterization. Colour and texture features are consolidated for ethnicity classification in order to investigate whether combining more features would improve the precision of an ethnicity classification task. A comparison of the colour, texture and combined features are discussed in detail.

1.6 Limitations of Models Proposed

- Face detection methods using the connected components labelling algorithm presented in this work have two limitations. One, is that the input image should have a dark background and the other being that that it should have only one face in an image.

- Facial components that were extracted have some limitations. For instance, incorrect components are extracted for images containing beards and spectacles. The problem with such categories of face images is that more than one connected component is merged into a single component, which yields an automatic removal of one or more targeted facial components.

1.7 Thesis Outline

1. Chapter 2 presents a detailed background and a review of related work on face landmark detection and ethnicity classification.

2. In Chapter 3 face detection and face feature extraction methods are discussed. Binarization and connected component labelling algorithms used for face detection and feature extraction are discussed in detail.

3. Chapter 4 presents an ethnicity classification of face images using different colour space components. K-means and Bayesian classification algorithms are utilized to classify images into appropriate ethnicities.

4. Chapter 5 discusses the ethnicity classification of a face image using Gabor features of face landmarks. The performance comparison of four machine learning techniques for the proposed model is discussed. The level to which every facial component contributes to the ethnicity classification problem is likewise examined.

5. In Chapter 6, a combination of colour features and Gabor features are used to increase the performance of the ethnicity classification method. MLP is then used as a classifier.

6. A summary of work done and result obtained, conclusions and possible future works are presented in Chapter 7.

Chapter 2

Literature Review

2.1 Introduction

The term biometrics originates from the Greek words bios, which means life, and metrics, which means measure [6]. Biometrics can be characterized as measurable physiological and/or behavioural characteristics that can be used to check the identity of an individual. There are numerous biometrics being used today and a scope of biometrics that are still in the early phases of advancement [26]. In practice, the process of identification and authentication is the ability to verify and confirm an identity. A biometrics system works by acquiring biometric information from a person, extracting feature sets and comparing it with the template set in the database. Depending on the application context, the identity of a person can be determined in two ways namely verification and identification [27]. In the former, the person to be recognized presents a claim; which is either accepted or rejected. In the latter, a person is recognized without her/him claiming to be identified. In the literature, however, verification and identification are used interchangeably for biometrics recognition [27]. At present, biometric systems are used in ATMs, computers, cell phones, security establishments, credit cards, health and social services.

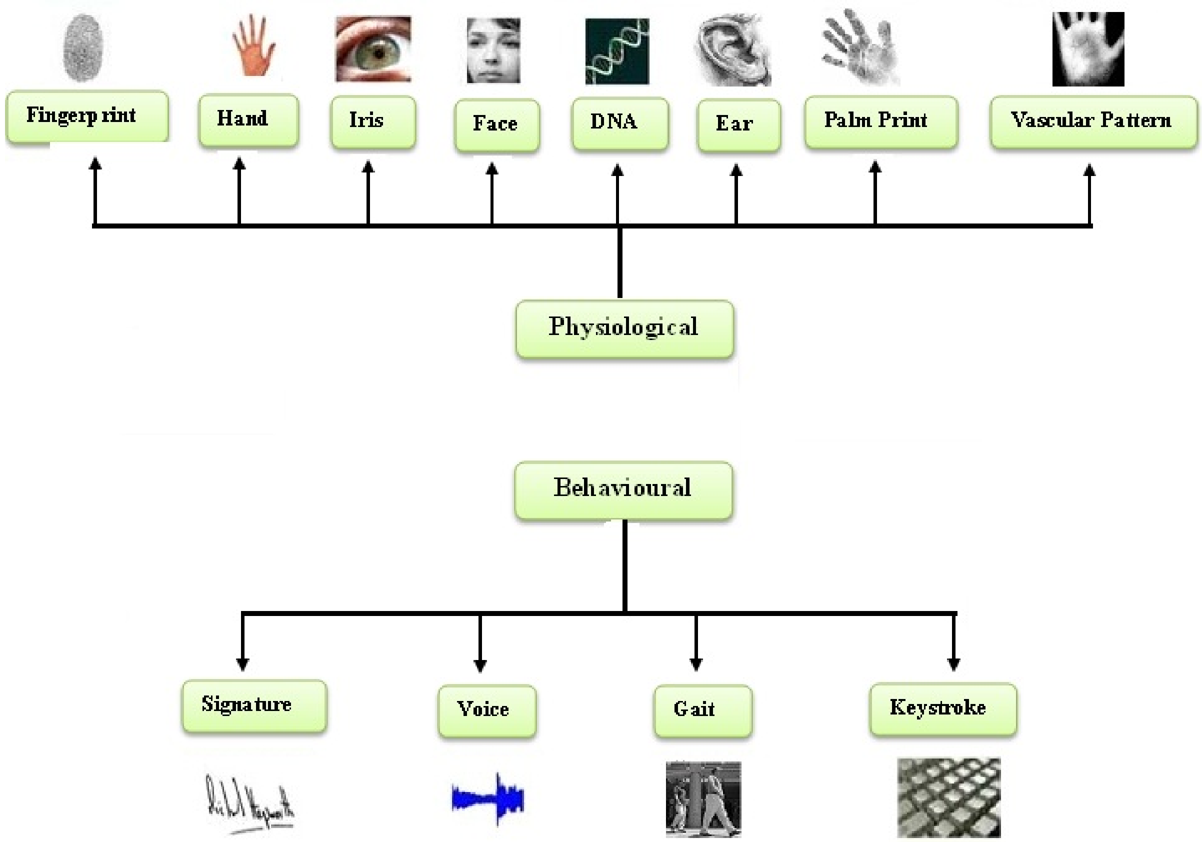


Figure 2.1: Example of different biometric technologies

2.2 Biometric Techniques

Currently, there are various biometric techniques that are either widely used or under intensive investigation. The biometric techniques include, fingerprint, iris, hand geometry, signature, facial geometry, voice, palm-print, retina, vein structure, ear form, DNA, odour, keyboard strokes, and gait [28]. Examples of these biometric techniques are shown in Figure 2.1. Each of them has different accuracy, and cost. Biometric technologies are separated into two classes: the first is based on physiological traits and the second is based on behavioural traits.

2.2.1 Physiological Biometric Technique

Characteristics, which can be measured passively on a part of the body at some point in time, are referred to as physiological biometrics [5, 29]. They are as follows:

Fingerprint

The fingerprint is one of the oldest and popular biometric technology that is used in the identification of an individual. It has been used in criminal investigations for more than 100 years. The validity of fingerprints as a basis for personal identification is thus well-established. A fingerprint is the pattern of ridges and valleys on the tip of a finger and is used for personal verification or for the identification of people [30]. The fingerprint-based recognition method is the most widespread and a dependable technique and is currently the most important biometric technology, due to its outstanding features of universality, permanence, uniqueness, accuracy and low cost. It has been empirically determined that the fingerprints of identical twins are distinctive [31].

Fingerprint identification has a number of advantages which makes it a popular method of identification in settings ranging from police stations to secured facilities. One great advantage of the fingerprint identification is that it is widely accepted in the legal community, among the law enforcement, and the general public. In addition, fingerprint identification is widely perceived as highly accurate and very reliable, since the statistical chance of two people on earth having identical fingerprints is very low. Fingerprints also tend to remain consistent throughout life, unless someone experiences an accident or works in an industry which has caustic or hot materials which ultimately damage the fingerprints [32]. Although the fingerprint biometric technology has many benefits, it also has some limitations. First, the device which captures the fingerprints also captures a picture of the dirt, grease, and contamination that is found on the finger. Therefore, in some areas, there are chances that the fingerprint identification can be rejected by the system. Firstly, the rejection can take place if a worker has a mark or some other contaminants on his or her finger.

Secondly, today's scanners still cannot recognize if the fingerprint is of a real finger or an artificial one, therefore, it is possible to trick the scanner by using a gelating print mold over a real finger. A third disadvantage of these devices is the extent of the damage when somebody does manage to steal the identity information. Thus, if someone's fingerprint data is compromised, it may not only compromise network security but would also comprise their bank account and their car. [33].

Face

The human face is one of the most well-known biometrics utilized for person identification or verification. It has a unique feature for human identification. Face recognition for its simple use and non-intrusion has made it one of the most important biometric recognition techniques that can be used [34]. Regardless of the fact that there are more reliable biometric recognition techniques, for example the fingerprint and iris recognition, these techniques are intrusive and their success depends exceptionally on user cooperation, since the user must position their eye in front of the iris scanner or put their finger in the fingerprint device. On the contrary, the face recognition is non-intrusive since it is based on images recorded by a distant camera, and can be extremely successful regardless of the possibility that the user is not aware of the presence of the face recognition system [35].

Personal identification based on facial images is considered the friendliest among all biometrics [35]. A large number of algorithms have been presented for face recognition. Such algorithms can be separated into two classes, namely geometric feature-based and appearance-based [36]. The geometry feature based methods include: Active Shape Model (ASM) [37, 38], Elastic Bunch Graph Matching (EBGM) [39, 40] and Local Feature Analysis (LFA) [41, 42]. Appearance-based methods include: Eigenfaces [43, 44], Fisherfaces [45, 46], Independent Component Analysis (ICA) [47, 48], Kernel Principal Component Analysis (KPCA) [49, 50], Kernel Fisher Discriminant Analysis (KFDA) [51], General Discriminant Analysis (GDA) [52, 53], Neural Networks [54, 55], and SVM [50, 56].

Challenges in Face Biometrics

The human face is an extremely difficult example to identify and perceive in that, although all faces have a similar structure, at the same time numerous environmental and personal factors influence the facial appearance.

The primary issue of face recognition is the enormous variability of the recorded images that are due to illumination conditions, pose, facial expressions, different hairstyle, use of cosmetics, presence of glasses, and beard [57]. Images of the same individual taken at various circumstances, may sometimes exhibit more variability due to the previously mentioned factors (intra-personal variability), than images of different individuals due to gender, race, age and individual variations (extra-personal variability). Another significant parameter in face recognition is aging. A robust recognition system should be able to recognize an individual even after many years, especially in mug-shot matching forensic applications [35].

Recent public facial recognition benchmarks have shown that in general, the identification performance diminishes linearly in the logarithm of the quantity of individuals in the gallery database. Likewise, from a demographic perspective, it was found that the recognition rates for males were higher than that of females, and that the recognition rates for older people were higher than that of younger people. These tests additionally uncovered that while the best recognition techniques were fruitful on extensive face databases recorded in well-controlled situations, their performance truly deteriorated in uncontrolled situations, which were mainly because of variations in illumination and head rotations. Such variations have turned out to be one of the most serious issues of face recognition systems [35].

Iris

The iris is the coloured ring that surrounds the pupil and contains easily visible yet complex and distinct combinations of corona, pits, filaments, crypts, striations, radial furrows and more [58]. The iris is a highly protected internal organ of the eye and iris patterns have a high level of irregularities. The iris is known as the ”Living password” because of its unique and random features. The iris is with the individual all the time and cannot be stolen or faked. In that capacity it makes iris recognition a highly valuable biometric technique [59, 60].

In 1936, ophthalmologist Frank Burch developed the concept of using iris patterns as a technique to recognize an individual. In 1985 Drs. Leonard Flom and Aran Safir, ophthalmologists, proposed the idea that no two irises are similar, and were granted a patent for the iris identification concept in 1987 [61].

The complex iris texture carries very distinctive information that is valuable for personal recognition. The precision and speed of currently deployed iris-based systems is promising and supports the feasibility of large-scale identification systems [60]. In 1994, Dr Daugman introduced his automated iris recognition algorithm. In 1995, this iris prototype became available as a commercial product [62]. He utilized the two-dimensional (2D) Gabor wavelet to transform iris images into iris codes. Thereafter a hamming distance is calculated for the purpose of iris identification. This iris recognition method has turned out to be reliable and robust. A few different methods have likewise been created for iris recognition. They include amongst others, the ICA approach by Huang [63], texture analysis using multi-channel Gabor filtering and Wavelet Transform [64], and partial iris for recognition that uses a one-dimensional approach [65].

Hand Geometry

Hand geometry refers to the geometric structure of the hand that is composed of the lengths of the fingers, the widths of the fingers, and the width of a palm[66]. One of the widely used applications of the hand geometry technology in companies is its use for time and attendance purposes in association with the time clocks [67]. The benefits of a hand geometry system are that it is a relatively simple method that can utilize low resolution images and can provide high efficiency with great user acceptance [68]. The main disadvantage of this technique is its low discriminative capability - it is very difficult for a hand geometry-based biometric system to achieve a very high identification accuracy, especially for a large population [33]. However, environmental factors such as dry weather or dry skin do not appear to adversely affect the authentication accuracy of hand geometry-based systems. On the other hand jewellery like rings may pose challenges in extracting accurate hand geometry data [69].

Retina

The human retina is a thin tissue made out of neural cells that is situated in the posterior portion of the eye. Due to the complex structure of the vessels that supply the retina with blood, each person's retina is unique [70]. The retina comprises of millions of photoreceptors whose function is to gather the light rays that are sent to them, and transform that light into electrical pulses that travel through the optic nerve into the brain, which then converts these pulses into images. The two distinct types of photoreceptors that exist within the retina are called rods and cones [71]. The idea for retinal identification was first conceived by Dr. Carleton Simon and Dr. Isodore Goldstein and was published in the New York State Journal of Medicine in 1935. In 1950, Dr Paul Tower discovered thateven among identical twins the blood vessel patterns of the retina are unique and different [71]. The retina has certain advantages as it is highly reliable because of the fact that no two individuals have a similar retinal pattern: an extremely low false positive and false negative rate. The main disadvantages of retinal recognition are that the process is not user friendly, in that the subject being scanned must be near the camera optics, and that it has a high equipment cost [72].

Ear

The significance of the ear in establishing identity was introduced by Imhofer (1906), an ear specialist from Prague. He found that in a set of 500 ears he only required four attributes to uniquely distinguish them. Alfred Iannarelli (1989) has made a noteworthy contribution in this field. He has made two large scale ear identification studies. The primary study compared 10,000 ears drawn from an arbitrarily chosen test sample in California and the second study inspected identical twins and trip-lets (who have identical genetic makeup). In both the studies, all inspected ears were found to be unique, though identical multiple birth siblings had similar, but not identical, ear structures [73].

The ear is made up of standard features including the helix and antihelix, the lobe and a u shaped notch known as the intertragic notch between the ear hole and the lobe [74]. The ear has the four properties for a potential biometric namely universality, uniqueness, permanence and collectability. It may be added that, though use of the ear is still new in the perpetually growing field of biometrics, it is already demonstrating its grit and is on the verge of emerging as a major passive biometric tool [73]. The main advantage of ear biometrics is that it has a rich and stable structure that is preserved from birth into old age [75].

Palmprint

The palmprint is the area between the wrist and the fingers. Palmprint features such as ridges, principal lines, wrinkles, minutia points, singular points and texture can be used for personal identification or verification [27]. The palms of the human hands contain larger areas of patterns of ridges and valleys than the fingerprints. As a result, palmprints are expected to be more distinctive than fingerprints[69].

The palmprint verification or identification systems employ either high resolution or low resolution images. The high resolution system uses high resolution images and can be used in forensic applications, for example, criminal detection. Singular points, minutia points and ridges are used as features. While the low resolution system uses low resolution images, it is more suitable for civil and business applications for example, access control. Wrinkles, texture and principal lines are used as features [27]. Palmprint recognition has the potential of achieving a high accuracy rate, especially when all the features of the hand such as geometry, principal lines, ridge and valley features, and wrinkles are combined to build a system [76]. Palmprint verification techniques can be fundamentally partitioned into four categories: (1) line based [77, 78], (2) texture based [79], (3) orientation based [80] and (4) appearance based [81, 82].

Vascular Pattern

Vascular Pattern Recognition, is also commonly referred to as Vein Pattern Authentication. The system uses near-infrared light to reflect or transmit images of the blood vessels of a hand or finger [83]. Different vendors use different parts of the hand, palms, or fingers, but depend on a similar technique. Researchers have confirmed that the vascular pattern of the human body is unique to a specific individual and does not change as the individual ages. This technology is difficult to forge, is contact-less, and is capable of one-to-one and one-to-many matching. Vascular patterns are difficult to recreate since they are inside the hand and, for some methodologies, blood needs to flow in order to enroll an image [83].

Fujitsu has developed palm vein pattern authentication technology that uses vascular patterns as personal identification data [84]. In [85], they used a finger vein pattern for personal identification. Kumar and Prathyusha used hand vein images and knuckle shape information for personal authentication [86]. Watanabe used a vascular pattern of the palm for personal identification [87].

DNA Sequence

The proposed biometric is Deoxyribo Nucleic Acid (DNA) which helps in the secure transmission of enormous amounts of data. It has a structure capable of saving enormous amounts of data and also is a unique feature among individuals. DNA provides high security level, long term stability, user acceptance, is intrusive and has 0% Failure to Enroll Rate (FER) [88]. The advantage of DNA as a biometric system is that it is highly accurate. The chance of two individuals sharing the same DNA profile is less than one in a hundred billion. However DNA biometric systems have some weaknesses in that it is easy to steal a piece of DNA from unsuspecting subjects that can be subsequently abused for an ulterior purpose. In addition, DNA cannot be done in real time, and is intrusive in that a physical sample must be taken, while other biometric systems only use an image or a recording [4].

2.2.2 Behavioural Biometric

Behavioural traits, which are actively learned or acquired over time, are referred to as behavioural biometrics. They are as follows.

Signature

Signature verification systems verify the identity of the user by using the distinctive behavioural features of a signature (such as speed, pressure and stroke order), instead of a straightforward crosscheck of one signature against another [27]. Signature verification can be classified into two categories: off-line and online. An off-line signature verification system uses signatures that are already written onto a template which is later captured by a CCD camera or a scanner. The signatures are verified using image properties and geometric features of signatures. On-line signature verification involves the capturing of both geometric and dynamic features such as the pressure of pen tips, the velocity along signature path, the time taken for signature, and the trajectory profiles of the signatures [27, 60]. The advantage of this biometric is that it is already an acceptable form of personal identification. It can therefore be easily incorporated into existing business processes, such as credit card transactions. A signature is a behavioural biometrics trait that changes over the course of time and is influenced by the physical and emotional conditions of the signatories [33]. Different types of signature verification methods and performance evaluation can be found in [89].

Voice/Speaker

Voice is a combination of physical and behavioural biometric characteristics. The vocal characteristics of humans are totally determined by the vocal tract, mouth, nasal cavities, and the other speech processing mechanisms of the human body [90]. There are two types of voice/speaker recognition systems: text-dependent (fixed-text) and text-independent (free-text). The text-dependent systems require a user to speak a particular predetermined phrase or sentences, that usually contains the same text as the training data. The text-dependent systems for the most part perform better than text-independent systems because the foresight of what is said can be exploited to align speech signals into more discriminate classes [4]. These days, text-dependent systems based on the Hidden Markov Model (HMM) using Gaussian or multi-Gaussian distributions, are more prominent [91]. In text independent speaker verification, the users are not restricted to any fixed or prompted phrases. They have the freedom to say whatever they want. To represent the expected freedom of utterances, diverse strategies have been proposed such as: long-term statistics and multi-dimensional autoregressive [92], and vector quantization [93].

The main challenges with voice-based recognition are that the speech features may not be sufficiently unique to permit an identification of an individual from a large population. In addition, voice features are sensitive to a number of factors such as the emotional and the physical state of the speaker as well as background noise. Moreover, some people seem to be extraordinarily skilled in mimicking an individual's voice [90].

Gait

The way in which one walks has been shown to vary between individuals. This was first recorded by Murray et al. in 1964, where it was found that amongst a small group of participants, each exhibited their own unique movement pattern [94]. Gait recognition is an emerging biometric technology which involves the process of identifying an individual purely through the analysis of the way they walk. While research is still underway, it has attracted interest as a method of identification because it is non-invasive and does not require the subject's cooperation [95]. While gait has several attractive properties as a biometric (it is unobtrusive and can be done with simple instrumentation), there are several confounding factors such as variations due to footwear, terrain, fatigue, injury, and the passage of time [96].

The first gait recognition approach into computer-vision started in the early nineties, with Niyogi and Adelson [95]. Interest in gait as a biometric gradually increased over the years, with Defense Advanced Research Projects Agency (DARPA) that established the human identification at a distance programme in order to encourage and support research into gait and other non-contact biometrics that could be observed from a distance [97]. The disadvantage of this biometric is that it may not remain invariant, particularly over a long period of time, due to major injuries involving joints or brain, changes in body weight, or due to inebriation [4].

Keystroke

It has been suggested that individuals have a characteristic way of typing on a keyboard [4]. Keystroke dynamics refers to the process of analyzing the way the user types on digital devices (such as the computer keyboard, the mobile phone, or a touch screen panel). The main motivation behind this is that keystroke dynamics biometrics is economical and can be easily integrated into the existing computer security systems with minimal alteration and user intervention [98].

The main advantage of this biometric is that unlike other biometric systems which may be expensive to implement, keystroke dynamics is almost free and the only equipment required is the keyboard. Furthermore, the recognition based on typing rhythm is not intrusive, which makes it quite applicable to computer access security as users already use their keyboards to type [99].

2.2.3 Biometric Qualities

There are several biometric characteristics being used in various application systems. Each biometric technology has its own advantages and disadvantages. Jain et al. recommend that a useful biometric system will have seven qualities [4]. They are as follows:

- Universality: every potential user has the modality;
- Uniqueness: the modality satisfactorily differentiates between any two users;
- Permanence: the modality profile remains relatively consistent over time;
- Collectability: the modality samples are easy to detect and acquire;
- Performance: the modality is robust and functional within a range of operational and environmental factors;
- Acceptability: the degree to which users are willing to accept and use the modality; and
- Circumvention: how susceptible the modality is to spoof attacks and identity fraud.

Of these seven fundamental characteristics, uniqueness and permanence are most integral to biometric performance evaluations. A brief comparison of different biometrics techniques concurring to seven parameters is shown in Table 3.1 [4, 100]. The biometrics technique that should be used for a given application relies on the requirements and characteristics of that application, and the properties of the biometrics technology. Fingerprint and iris based techniques are more precise than the voice-based technique. Nevertheless, in a phone banking application, the voice-based technique might be preferable as the bank can integrate it seamlessly into the existing telephone system. Likewise, fingerprint and face-based techniques are more suitable in criminal investigation than a signature or voice based technique [4].

Table 2.1: Comparison of biometric technologies (H=High, M=Medium, L=Low)

Biometrics Modality	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention
Face	H	L	M	H	L	H	L
Fingerprint	M	H	H	M	H	M	H
Hand Geometry	M	M	M	H	M	M	M
Keystrokes	L	L	L	M	L	M	M
Hand Vein	M	M	M	M	M	M	H
Iris	H	H	H	M	H	L	H
Retinal Scan	H	H	M	L	H	L	H
Signatures	L	L	L	H	L	H	L
Voice Print	M	L	L	M	L	H	L
Face Thermogram	H	H	L	H	M	H	H
Odour	H	H	H	L	L	M	L
DNA	H	H	H	L	H	L	L
Gait	M	L	L	H	L	H	M
Ear	M	M	H	M	M	H	M

2.3 Biometric Performance Metrics

Distinctive measurements (metrics) can be utilized to rate the performance of a biometric system. The most common biometric performance metrics are [69, 101]:

1. False Positive Rate (FPR): usually refers to the probability (the percentage of times) at which

the system incorrectly declares that a biometric sample belongs to the claimed identity when the sample actually belongs to a different subject (impostor).

2. False Rejection Rate (FRR): also called a type I error, is an empirical estimate of the probability at which the system fails to recognize claimed identity when the sample actually belongs to the genuine user.

3. True Positive Rate (TPR): It is defined as $1 - FRR$.

4. Expected Error Rate (EER) : The EER refers to the point at which genuine and imposter error rates are nearest to zero. The rate at which FPR is equivalent to FRR. The lower the EER, the more accurate the system is considered to be.

Besides these errors, biometric systems can encounter another type of errors as well. Failure to Capture (FTC) rate or Failure to Acquire Rate (FAR) denotes the probability that the system fails to capture a biometric input when presented correctly. This type of error typically occurs when the device is not able to locate a biometrics signal of sufficiently good quality [69]. Weighted Error Rate (WER) is referred to as the weighted sum amongst FRR and FPR. FER is the proportion of users that cannot be successfully enrolled in a biometrics system. This necessitates the design of robust and efficient user interfaces that can assist users during enrollment [69].

The performance of a biometric system is also depicted visually using different charts [101].

- Receiver Operator Characteristics (ROC): This curve is a graphical plot that is created by plotting the TPR and the FPR at various threshold values. This curve displays how the TPR changes with respect to the FPR and vice-versa. Alternatively, the ROC curve also plots the FRR versus FPR [102].

- Detection Error Trade-off (DET) Curve: A DET curve is similar to the ROC curve with the exception that the axes are often scaled non-linearly in order to highlight the region of the error rates of interest. Commonly used scales include the normal deviate scale and the logarithmic scale [101].

- Expected Performance Curve (EPC): While the ROC and the DET curves are a posteriori assessment charts, an EPC is an a priori chart. It has three variations: (1) Using FPR as a performance criterion and report performance in FRR; (2) Using FRR as a performance criterion and report performance in FPR; (3) Using WER as a performance criterion and report performance in Half Total Error Rate (HTER), where HTER is defined as the average of FPR and FRR [101].

In many classification experiments, the overall accuracy is used to report the performance [7, 12, 103–107]. The accuracy of the system measured was based on the percentage of correctly classified images. The percentage is determined via the following equation

$$Accuracy = \frac{Number\ of\ correctly\ classified\ Images}{Total\ Number\ of\ images} \tag{2.1}$$

Another performance measure used in classification and machine learning provides a finer level of detail. A confusion matrix, or matching matrix, looks at both the predicted class and given class of an instance and places it into the correct position of the table. In our experiments, we used accuracy and confusion matrix as performance measures.

2.4 Biometric Application

There are several biometric technologies that are in use in various applications. Wherever security is involved, the use of biometric technologies can help make operations, transactions and everyday life both safer and more convenient. Law enforcement and forensic departments have been using biometric technology for many years already. It is used mainly for the identification of criminals, surveillance, corpse identification, parenthood determination, missing children, terrorist identification. [4]. Automated fingerprint recognition systems are deployed around the world for criminal identification. North America has also deployed facial recognition surveillance systems in casinos for some time in order to spot known cheaters. More recently, the facial recognition surveillance system have been used at major events such as the 2001 Super Bowl in Tampa, Florida, and in the 2002 Winter Olympics at Salt Lake City [108].

Banking and financial services represent enormous growth areas for biometric technology, with many deployments currently functioning and with pilot projects announced frequently. Some applications in this area are account access, ATM, online banking, credit card, physical access control, PC/network access, e-commerce, electronic data security, cellular phone, time and attendance monitoring [4, 109]. Hand geometry, iris and facial recognition have been deployed at ATMs in North America, Europe, and Asia. The JPMorgan Chase Bank allows some customers to access accounts by speaker recognition [108]. Even firms that offer internet shopping are considering biometric technologies to authorize various types of transactions.

A key area of application for biometric technology is travelling and immigration. The use of biometrics to control the travellers crossing the national or state borders is increasing, especially in regions with high volumes of travellers or illegal immigrants, in passport control, in air travel, etc. [4, 109].

There are many applications of biometrics in the government sector, such as national ID cards, voter ID cards, driver's licenses, social security, employee authentication, and military programmes [4]. The Office of Legislative Counsel of the United States (U.S.) House of Representatives is installing the iris recognition system to secure confidential files and documents. Likewise, the Department of Defense (DOD), the Department of Energy, and the Department of Justice, are also adopting similar technologies. More than half of the nuclear power plants in the U.S. employ biometric hand geometry systems [108, 110].

Different countries are using biometrics in licensing and voting processes. For example, Mexico uses the facial recognition technology to check voter rolls for duplicates in its national elections. Costa Rica, Brazil, the Dominican Republic, Panama and Italy all use fingerprints to verify voters at the polling stations [108]. Biometric technologies are being used in countries as diverse as India, South Africa, Philippines and Spain as the basis for the national identification system, eliminating the need for multiple identification mechanisms. In India, the implementation of the biometrics-based Multipurpose National Identity Card will replace all other forms of identification and enable citizens to access public services and subsidies on food, energy and education that now suffer from major pilferage [110, 111].

Although the public's attention is primarily focused on the use of biometrics in critical infrastructure and security applications, biometric technologies such as iris recognition, the fingerprint, and the hand geometry modalities have gained traction in health care applications. This includes securing medical facilities and equipment, the protection and management of confidential medical records and patient identification [4, 109].

2.5 Related Work

2.5.1 Automatic Face and Facial Landmark Detection

In previous years, face recognition has attracted considerable attention from various disciplines and has seen a tremendous growth in research. Face recognition has been widely applied in security systems, surveillance systems, credit card verification and other applications. Particularly, face detection is an important part of face recognition and is the initial step of automatic face recognition. However, face detection is not straightforward since it has numerous variations of image appearance, such as occlusion, image orientation, pose (front, non-front), facial expression, and illuminating condition. Automatic face and facial features detection have become important due to applications such as face identification, face recognition, and facial expression analysis [57].

Different methodologies have been published thus far for the detection of face and facial features and are presented in [112, 113] as follows.

- *Geometric Approach:* The method is based on face geometrical configuration and encodes human knowledge of what constitutes a typical face (usually, the relationships between facial features). The top-down knowledge-based approach assumes a different face model at different coarse-to-fine scales. By using the facial components as well as positional relationship between them one can locate the face with ease. Face detection can often be achieved by detecting geometrical relationships among facial organs, because they are simple, straightforward and efficient. The strategy involves the computation of a set of geometrical features such as the nose width and length, the position of the mouth, the chin shape from the face image that need to be recognized. This set of features is then compared with the features of known individuals. A suitable metric such as Euclidean distance (finding the closest vector) can be used to find the closet match.

  Jeng et al. [18] proposed a useful geometrical face model and an efficient facial feature detection approach, which depends on the fact that human faces are constructed in the same geometrical configuration and could accurately detect facial features. This is more especially in the eyes, even when the images have complex backgrounds such as bad lighting conditions, or have skewed face orientations, and facial expressions. Some of the work that uses this approach was reported by Yang and Huang [114–117]. Majumder et al. [118] used facial geometry to find the nose, the mouth and the eyes positions in an image. In [119], facial features are extracted using the geometric relation between the face and the eyes. The advantage of using geometrical features as a basis for face recognition is that recognition is possible even at low resolutions and with noisy images (images with many disorderly pixel intensities). In spite of the fact that the face cannot be seen in detail its overall geometrical configuration can be extracted for face recognition. The technique's disadvantage is that the automated extraction of the facial geometrical features is extremely difficult. Automated geometrical feature extraction-based recognition is additionally very sensitive to the scaling and rotation of a face in the image plane [120].

- *Feature Invariant Approaches:* The bottom-up feature-based approach searches the image for a set of facial features and groups them into face candidates based on their geometrical relationship. Though this approach can be easily extended to multiple views, it is unable to work well under different image conditions because the image structure of the facial features varies too much to be robustly detected by the feature detectors. Leung et al. [121] reported work using this approach. He used multi-orientation and multi-scale Gaussian derivative channels to identify facial components. The vector of filter responses is learnt for 5 points on the face, namely both pupils, both nostrils and the nose lip intersection. The bank of filter channels is then applied to an image and the best matches (using a dot product metric) are held as the candidate's eyes, and both nostrils. The facial components are combined into conceivable face candidates. This is accomplished by selecting all feature points over a specific threshold and pairing them up. Each pair is then used to characterize small elliptical regions where the remaining 3 feature points are likely to occur. In the event that further feature points are found inside the predicted regions, then a face candidate is formed.

Some of the work that used this method is reported by Yow and Cipola [112, 122]. A more recent feature based approach is introduced by Hamouz et al. [123], who utilize a bank of Gabor channels to scan for 10 facial components (eye centers, eye corners, nostrils and mouth corners). All features are modeled using a Gaussian Mixture Model (GMM) of feature responses. Any triplet of feature detections, with an acceptable spatial orientation, produces a face location hypothesis. These face candidates are then normalized by using a relative transformation and afterward tested by utilizing a SVM region classifier.

- *Template Matching Methods:* In this approach, a template model of the entire face is constructed and used to search the image. Several standard patterns are stored to describe the face as a whole or the facial elements independently. Typically, a template face model is scanned across a target image at various scales and then requested to classify each subregion as face or non-face. In [124, 125] the template matching method is used for face localization and detection by computing the correlation of an input image to a standard face pattern.

Different researchers have attempted to solve this pattern recognition problem by modelling the distribution of human faces or more commonly by applying well known machine learning techniques to differentiate between face and non-face distributions, for example, face detection with eigenface [19], neural network [126], and SVM [127]. An implementation of template matching called Correlation Templates utilizes an entire bank of fixed sized templates to identify facial features in an image [120]. By using several templates of different fixed sizes, faces of different scales are recognized. The other implementation of template matching is to use a deformable template [128]. Instead of using several fixed size templates, they used a deformable template (which is non-rigid), thereby changing the size of the template hoping to detect a face in an image.

- *Colour-based or Texture Based Approach:* Colour and texture are two essential modalities in many image processing tasks, that range from remote sensing to medical imaging, robot vision, face recognition, etc. Skin detection plays a vital role in face detection and identification. Some research results show that human skin colours cluster in a small region only in the RGB colour space instead of the HIS colour space. Human skin colours differ more in brightness than in colours; and every texture is distinctive and distinguishable from one another. Therefore, the normalized RGB or texture models are considered to be capable of characterizing the human face with less variance in colour or texture.

Recently, many intensive studies have been reported in literature [129–131]. A face detection algorithm based on skin colour segmentation is proposed in [132] and [133]. In [134] they used skin colour information to extract the face and facial features from an image. Ben Jemaa and Khanfir [135] likewise used skin colour information for detecting faces in images. They make use of chrominance components to detect the features in the image. Mauricio et al. [136] presented a technique for face detection using region clustering, shape analysis, and edge detection and landmark detection using depth relief curves and surface curvature information.

2.5.2 Component Based Face and Facial Landmark Detection

Component based automatic face recognition has gained a lot of attention in the past fifteen years. A component-based object detection system is one that searches for an object by looking for its identifying components rather than the entire object. An example of such a system is a face detection system that finds a face when it locates a nose, a pair of eyes and a mouth in the proper configuration [137, 138]. Rakhmadi et al. [139] presented a connected component labelling algorithm for detecting face in digital images. Huang et al. presented a SVM based face recognition system which breaks down the face into a set of components that are interconnected by a flexible geometrical model [140]. They used a 3D morphable model to generate 3D face models from only two input images. Fonou Dombeu and Tapamo [141] proposed an algorithm, based on image thresholding and connected component labelling, that extracts facial components from the gray-scale images. They used a convex hull algorithm to validate the detected components. Poggio et al. successfully used a component-based approach for person detection in a cluttered environment [138].

In [20] Heisele et al. introduced a component-based face recognition method and compared it with two global face methods, in regards to the robustness against the pose changes. Components were initially extracted from the face, and then features were extracted from each component and combined in one single features vector. Furthermore, SVM was used in the component-based and global methods to classify faces. The component-based methods outperformed the global methods. However, the component-based classifier was still performing poorly with the pose changes. Perona et al. detected 5 features on the face: the left eye, the right eye, the nose/lip junction, the left nostril and the right nostril using random labelled graphs. They assumed that the feature detectors are fallible [121]. In [21] they presented a component-based framework for face detection and identification. They described an algorithm which automatically learns two separate sets of facial components for the detection and identification tasks. In experiments they compared the detection and identification systems to standard global approaches. They concluded from the experimental results that a component-based approach is better than the global approaches. Agushinta et al. presented a face component extraction by using a segmentation method for a face recognition system. From the experiment, they concluded that the determination of face components and face component distances can be used to identify a face as a subsystem of a face recognition system [22].

2.5.3 Ethnicity Classification

Ethnicity identification presents yet another challenge in face processing. The determination of the ethnicity of an individual can be very useful in a face recognition and identification system. Numerous pieces of research on automatic face detection and recognition have been presented in the public literature. However, very few attempts have been made to perform human ethnicity classification. The most common ethnic groups are shown in Figure 2.2. Table 2.2 gives the physical facial characteristics of the three major ethnic groups in terms of the skin colour, the eyes, the nose, the lip structure, the build, the head shape, the hair, and the face shape [7].

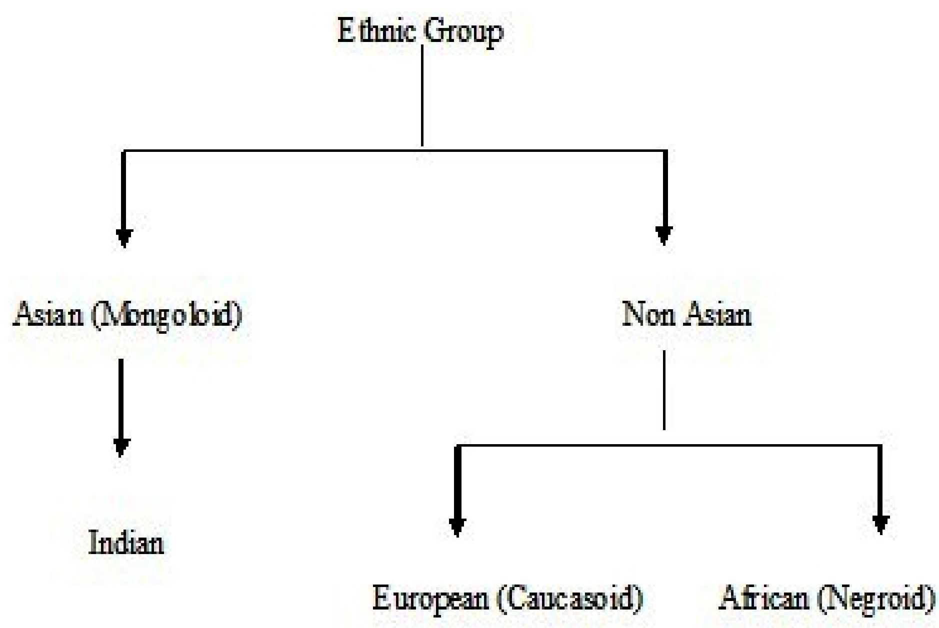


Figure 2.2: Major ethnic groups.

Some effort has been made on the ethnicity categorization area over the last two decades. In [103], the authors have classified images into Asian and non-Asian. They used Principal Component Analysis (PCA) and ICA for feature generation and features extraction respectively. Then, they used SVM for the training process and in combining different SVM classifiers to some new classifiers, they called it ”Three to One” or ”321”, which is explained in details in [103]. This improved the classification rate. In [142], real time face detection, gender and ethnicity classification were proposed. The authors used three types of rectangular filters to extract features. The images are classified into two classes, Asian and non-Asian, using SVM and boosted classifiers. In [143], a fusion of multi-view gait is used for the determination of ethnicity. Gait Energy Image (GEI) is used to analyse the recognition power of gait for ethnicity. However the highest classification rate

Table 2.2: The physical facial characteristics of the three major ethnic groups.

Ethnic Group	Facial Characteristics
Asian (Mongoloid)	Eyes: narrow, epicanthic fold. Nose: low, average width. Lips: average fullness. Face Shape: short with flat, projected cheekbones. Hair: thick and straight or slightly wavy, thin facial hair. Skin: yellowish.
European (Caucasoid)	Eyes: double eyelid, exposed tear trough, large. Nose: prominent, high bridge, narrow. Lips: thin, tight. Face Shape: center of face juts outward, wedge shaped, long face. Hair: wavy or curly, thick body and facial hair. Skin: light or brown.
African (Negroid)	Eyes: large with exposed tear trough. Nose: broad, low. Lips: thick, stretched mouth. Face Shape: long head. Hair: tight curls or heavy waves. Skin: dark (high melanin quantity).

they achieved was 84%.

Gutta et al. [12] considered ethnicity categorization as a multi-class problem including Caucasian, African, Oriental and Asian origin categories. For the categorization, based on the intensity faces, they adopted a hybrid classifier that consisted of two parts, namely an ensemble of Radial Basis Function (RBF) and Inductive Decision Tree (IDT). The achieved accuracy rate was 94%. Xiaoguang and Jain presented a Linear Discriminant Analysis (LDA) based scheme for two classes (Asian and non-Asian) for the task of classifying ethnicity. They reported 96.3% accuracy in classifying 131 non-Asian faces and 132 Asian faces [11].

Manesh et al. proposed a two class ethnicity classification problem, namely Asian and non-Asian by using an appearance-based technique to decide the confidence of various facial regions using a SVM classifier. They reported a 0.0261% error rate with normalized faces which used the eye and the mouth positions [144]. Hosoi et al. used the Gabor Wavelet Transformation (GWT) and retina sampling to extract key facial features, and then used SVM for ethnicity classification. They reported an overall accuracy of 94% for ethnicity estimation. They classified images into Asian, European and African [7]. Zhang and Wang investigated the ethnicity discriminality of both 2D and 3D face features. They make use of the Multi-scale Multi-ratio Local Binary Pattern (MM-LBP) method, which is a multimodal method for ethnicity classification. They claimed that their system accuracy is 99%. However, their method has a poor performance for face expression variation [107].

Lyle et al. [145] classified images into gender and ethnicity using gray-scale pixel intensities and periocular texture computed by Local Binary Patterns (LBP) as features. They used SVM as a classifier. They reported 91% accuracy for the ethnicity classification. Ai et al. [105] used LBP and Haar wavelets as features and the classifier they used was Adaboost. They had classified images into Asian and non-Asian and reported an accuracy of 97%.

In [146], the authors classified images into three Chinese minority groups using two types of features: LDA-based algebraic features and elastic model-based geometric features. They reported an accuracy of 79% using algebraic features and a 90.95% accuracy with geometric features with K-nearest neighbour (KNN) and C5.0 as classifiers. Guo et al. [104] used a Craniofacial Longitudinal Morphological Face Database (MORPH) II database to detect how gender and age affect the ethnicity classification. They used Biologically Inspired Features (BIF) based on Gabor filters. SVM was used as a classifier.

Lin et al. presented a method for recognizing age, ethnicity and gender with facial images. Gabor filter banks and Adaboost learning are consolidated to extract key facial features of each pattern. Thereafter they used a Gabor + Adaboost features-based SVM classifier to recognize the face image of each pattern. They classified images into Yellow vs White, Yellow vs Black and

White vs Black and achieved 89.05%, 91.48% and 90.55% accuracy respectively [116]. In [147], Xie et al. presented a novel approach for ethnicity classification tasks on large scale databases. They used Kernel Class-dependent Feature Analysis (KCFA) combined with facial colour as features for ethnicity classification. A new design of multiple filtered responses of the Kernel Class-dependent Feature Analysis was used for ethnicity classes. They classified images into Asian, African American and Caucasian. Their method achieves an accuracy of 95% for Asian, and 97% for African-American and Caucasian ethnic groups.

In [148], a fusion scheme that uses block-based uniform local binary patterns and Haar wavelet transform to combine local and global features are used for ethnicity identification. The images are classified into three ethnic groups namely European, Oriental and African. Accuracy obtained for European is 96.92%, Oriental is 95.38% and for African is 96.61%. Zhong et al. presented an efficient fuzzy 3D face ethnicity classification algorithm. The algorithm comprises of two stages, namely, learning and mapping. In the learning stage, the Learned Visual Codebook (LVC) technique is used to learn the visual codes for both the Eastern and Western individuals. From these codes two distance measures are learned, namely, a merging distance and a mapping distance. Using the merging distance, they learnt the Eastern and Western codes which are based on the visual codes. In the mapping stage, they computed the probabilities for every 3D face mapped to Eastern and Western individuals using the mapping distance. The membership degree is determined by their defined membership function. Ratios of Eastern individuals categorized into the Eastern area are 75.59%, 80.16% and 74.85% respectively in the three circumstances (neutral, small and large). The ratios of Western individuals categorized into the Western area are 89.65%, 79.05% and 77.20% respectively in the three circumstances (neutral, small and large) [106].

In [149], the authors proposed ethnicity and gender identification from silhouetted face profiles using a computer vision technique for ethnicity identification. Shape context-based matching was employed for classification. The images are classified into Black (B), East and Southeast

Asian (ESEA), South Asian (SA) and White (W). They achieved an average accuracy of 71.66% for ethnicity classification. Toderici et al. classified face images into Asian and non-Asian using 3-D face recognition techniques. The average accuracy they achieved is 99.1% for Asian and 98.4% for non-Asian [150]. Husain and Enas used two types of features namely the cranio-facial features (the eyes, the nose and the mouth) of the faces and the variance of the skin colour together, to further improve the accuracy of ethnicity classification. For classification, five ratios have been calculated as a mathematical relation between features using four ethnic groups selected from FG-NET, the CPIR database and the database that was collected by the user. The system achieves an accuracy of 82% [151].

Tofiq and Shafagat introduced a race and ethnicity recognition method using portrait photographs. The reference image is formed based on selected geometric points of the face and a special algorithm for calculating the characteristic parameters of the images available in the database. Furthermore, the original image is compared with the reference image and thus, the ethnic group of the original image is determined [23]. Huang et al. presented an approach of ethnicity classification from 3D face models by combining both boosted local texture and shape features. A novel local descriptor, namely Local Circular Patterns (LCP) is presented. They reported an accuracy of 95.50% for gender classification and 99.60% for the ethnicity classification task using the FRGC v2.0 dataset; and 95.60% for gender and 97.42% for ethnicity classification task using the BU-3DFE dataset [152].

In [153], the authors developed a framework to classify race from facial images into three categories namely Arabic, Asian, and Caucasian. This was done by transforming the face images to a lower dimension using PCA forwarded to a classifier constructer using feed foreword neural network with different parameters such as different inputs (PCA components 40, 50, and 60), layers, neurons and training algorithms. The accuracy they achieved for race classification was 83.5%. Demirkus et al. proposed algorithms for face and session soft biometric feature extraction from surveillance videos with low resolution and uncontrolled illumination, in order to achieve gender and ethnicity classification. They used pixel intensity values and Biologically Inspired Model (BIM) features for soft biometry computation and the SVM as a classifier. For ethnic classes, they achieved an accuracy of 85% for Asian vs. others, 86.5% for African vs. others and 78.5% for Caucasian vs. others [154]. Hadid and Pietikäinen presented the ethnicity classification task into Asian and non-Asian from face videos using manifold learning. They used four different methods for ethnicity classification namely, in the use of pixels +SVM+fusion an accuracy of 96.6% accuracy was achieved; the use of LBP+SVM+fusion achieved a 97% accuracy; the use of spatiotemporal achieved 99.2% accuracy; and finally, the use of manifold learning achieved an accuracy of 100% [155].

Ding et al. presented a method for ethnicity classification into Asian and non-Asian faces using the fusion of boosted local texture and shape features extracted from 3D face models. The proposed methods make use of the Oriented Gradient Maps (OGMs) to highlight local geometry and texture variations of human faces, and then adopts Adaboost to learn a compact set of discriminative features highly related to the ethnicity properties for classification. The performance achieved is 98.3% in separating Asians from non-Asians [156]. In [157], the authors used mixtures of experts on the ethnic classification of human faces. The mixture of experts consists of ensembles of RBF, IDT and SVM. By using the ERBF/DT approach, their experimental results yield an average accuracy rate of 92% on the ethnicity classification task.

In [158], two different modalities of human faces namely range and intensity are used for ethnicity classification using the SVM. They utilized range information, containing a 3D shape of the face object, for ethnicity identification. A fusion scheme is also developed by integrating the range and intensity to identify the ethnicity from facial scans. Buchala et al. used PCA to analyze different properties of faces, such as gender, age, ethnicity, and identity. Using LDA they estimated the encoding powers, with respect to different properties, of the components obtained by

PCA. Images were classified into three categories namely Caucasian, African and East Asian and an accuracy of 81.67% was obtained [159]. Han et al. presents a generic framework for automatic demographic race estimation from a given face image. They extracted demographic informative features from the commonly used BIF, and then employed a hierarchical approach consisting of between-group classification, and within-group regression [160]. In [161], the authors proposed an approach for race estimation from real-life face images acquired in unconstrained conditions. BIF are extracted from the normalized face image, including both the central face region and the surrounding context region. SVM classifiers are then used to predict the race of a subject.

In [162], Muhammad et al. explored and compared the performance of local descriptors for ethnicity classification from face images. They used two types of local descriptors in their study: LBP and Weber Local Descriptors (WLD). For the Asian group, they reported an accuracy of 99.47 %, for the African 98.99%, for the Hispanic 96.83%, for the Middle Eastern 100% and for the European 100%. In our study we used only four facial components for ethnicity classification. The features that we have used are much fever compared to the features used in literature. Hence the speed of the algorithm will be better. A very good accuracy was achieved with the limited features: for Asian it is 99.49%, for Indian 82.69%, for White 92.3% and for Black 72.3%. We also summarize the previous works on ethnicity classification in terms of the database used, feature extraction methods, race groups, and classifier. The results that were achieved are presented in Table 2.3. Different features used for this purpose includes gray-scale pixel intensities (used directly or represented in terms of PCA eigenvectors), Gabor wavelets, LBP, LDA, geometric features and Haar wavelets among others. The classifiers of choice are Adaboost (along with variants of boosting), Neural Networks, LDA, SVM, MLP and KNN. While each of the classifiers has its own advantages and limitations, SVM appears to be the most popular choice for gender and ethnicity classification due to its relatively high accuracy and generalizing ability.

Table 2.3: Summary of research done in ethnicity classification.

Reference	Features	Classifier	Race Groups	Database	Recognition
[142]	Rectangular Features	Boosted Classifier and SVM	Asian, non-Asian	Faces from WWW	22.6% error rate
[11]	Range and Pixel Intensity	SVM	Asian, non-Asian,	Asian PF01, NLPR, Yale, AR	98%
[12]	Gray-scale Pixel Intensities	Neural networks using SVM and DT's	Caucasian, South Asian, East Asian , African	FERET	92%
[107]	MM-LBP	AdaBoost	Asian, White	FRGC v2.0	99.5%
[103]	PCA	SVM	Asian, non-Asian	FERET	82.5%
[143]	MPCA	Weighted sum rule and majority vote rule	Asian, American	Database collected by author	84%
[144]	Gabor	SVM	Asian, non-Asian	CAS-PEAL, FERET	96%
[7]	Gabor wavelet transform with retina sampling	SVM	Asian, European, African	HOIP dataset	Asian-96% European - 93% African-94%

[145]	LBP	SVM	Asian , non-Asian	FRGC Face dataset	91%
[105]	LBP,Haar like features	Adaboost	Asian, non-Asian	FERET,PIE	97%
[146]	LDA, PCA, Geometric feature	KNN, C5.0	Minority ethnic group in Chinese race: Tibetan, Uighur, Zhuang	Database collected by user	79% with algebric features and 90.95% with geometric features
[104]	Biologically inspired features	SVM	Black and White	MORPH II	Black-98.3% White-97.1%, Hispanic-74.2%, Asian-59.5%, Indian-6.90%
[162]	LBP, WLD	minimum distance classifier (City-block, Euclidean, and Chi-square.)	Asian, Black Hispanic, Middle Eastern, White	FERET	96.84% for LBP and 95.85% for WLD,

[116]	Gabor filter banks and Adaboost	SVM	Yellow, White and Black	FERET	Yellow vs White-89%, Yellow vs Black-91.48%, White vs Black-90.55%
[147]	KCFAwith facial colour	KNN and SVM	Asian, African American, Caucasian	MBGC Database and database collected by user	96%
[148]	Block-based uniform LBP and Haar wavelet	KNN	European, Oriental and African	GUFD, PUT, PDA, Fanar, CUHK, HKU, MUCT	European-96.92%, Oriental-95.38% and African-96.61%
[106]	LVC method	Merging and mapping distance	Eastern, Middle Eastern, Western	FRGC2.0 3D Face Database	74%
[149]	Shape context	KNN	Black, East and Southeast Asian, South Asian, White	FCollected by user	71.66%

[150]	Facial-structure-based distance function	KNN, MDS, Waveletes	Asian, White	FRGC2.0	99%
[151]	Cranio-facial and skin colour features	Ratios	African, Asian, White, Middle Eastern	FG-NET, CPIR database and database collected by user	82%
[152]	Boosted local texture and shape features	Local circular pattern	Asian, White	FRGC v2.0 and BU-3DFE	99.60% on FRGC v2.0 database and 97.42% on BU-3DFE database
[153]	PCA	Feed forward neural network	Arabic, Asian, and Caucasian	FERET, Database collected by author	83.5%
[154]	Pixel intensity-based features, BIM features	SVM	Asian, African and Caucasian	Database collected by author	Asian vs others-85%, African vs others-86.5% and Caucasian vs others-78.5%

[156]	Boosted local texture and shape features	Adaboost	Asian and non-Asian	FRGC v2.0 and BU-3DFE	98.3% on FRGC v2.0 database and 97.88% on BU-3DFE database
[157]	RBF, IDT	SVM	Caucasian, South Asian, East Asian, and African	FERET	92%
[158]	Range and intensity	SVM	Asian, non-Asian	Database from University of Notre Dame and Michigan State University	2.0% error rate
[159]	PCA	Euclidean measure	Caucasian, East Asian, and African	FERET	81.67%
[160]	BIF	Euclidean measure	Caucasian, East Asian, and African	FERET	81.67%

[161]	BIF	Binary decision tree	White and Black	MORPH II and PCSO database	99.1% on MORPH II and 98.7% on PCSO database

2.6 Machine Learning Techniques

Machine Learning (ML) is about learning to make predictions from examples of desired behaviour or past observations. Learning methods have found numerous applications in performance modelling and evaluation (for example [163–167]). ML is constructing computer programmes that develop solutions and improve with experience. ML techniques are preferred in situations where engineering approaches like hand-crafted models simply cannot cope with the complexity of the problem. ML techniques assist to solve problems which cannot be solved by enumerative methods or calculus-based techniques. It has applications in an incredibly wide variety of application areas, from medicine to advertising, from military to pedestrian. The ML system effectively "learns" how to estimate from a training set of completed projects [168, 169]. The basic ML system is presented in Figure 2.3.

ML techniques are broadly classified into three categories [169, 170]. These are:

- Supervised Learning: A training set of examples with the correct labels is provided and, based on this training set, the algorithm generalizes to respond correctly to all possible inputs. This is called learning from exemplars.
- Unsupervised Learning: No labels are provided. Instead, the algorithm tries to identify similarities between inputs so that inputs that have something in common are categorized

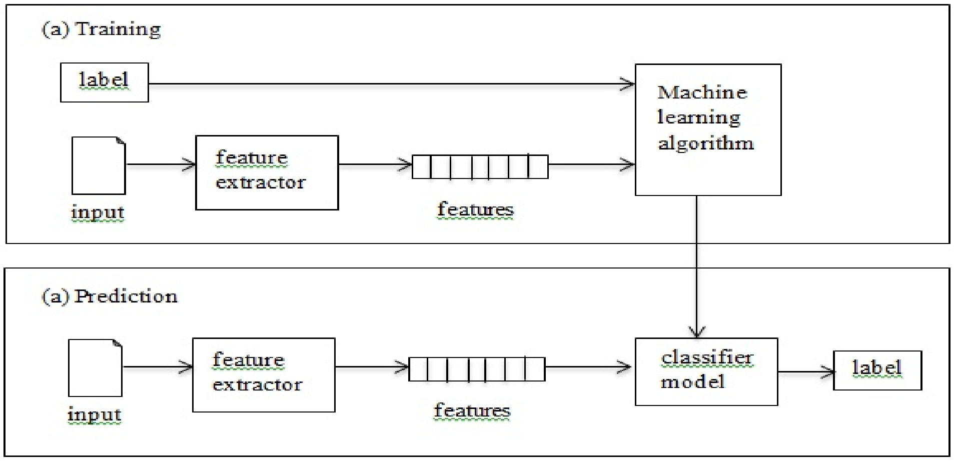


Figure 2.3: Basic machine learning system.

together.

- Reinforcement Learning: This is between supervised and unsupervised learning. The algorithms gets told when the answer is wrong, but does not get told how to correct it. It has to explore and try out different possibilities until it works out how to get the correct answer.

ML techniques used in this study are discussed below.

2.6.1 K-means Clustering

K-means is the clustering algorithm used to determine the natural spectral groupings present in a data set [171]. It [172, 173] is a non-hierarchical unsupervised clustering algorithm that classifies input data points in multi-dimensional space into multiple classes based on the distance from each other (typically the Euclidian distance between the cluster centers and the candidate vector). In other words, given $x_1,\ldots,x_N \in \mathbb{R}^d$ and $k \in \{1,2,\ldots,N\}$, the objective is to minimize the following function

$$W(\omega_1,\omega_2,\ldots,\omega_k;\mu_1,\ldots,\mu_k) = \sum_{i=1}^{k}\sum_{x_j\in\omega_i} \|x_j-\mu_i\|^2 \tag{2.2}$$

where the clustering $\{\omega_1, \omega_2, \ldots, \omega_k\}$ is a partition of $\{x_1, \ldots, x_N\}$ and $\mu_1 \ldots, \mu_k$ are the representatives of the clusters $\omega_1, \ldots, \omega_k$. The K-means algorithm finds locally optimal solutions with respect to the clustering error. It is a fast iterative algorithm that has been used in many clustering applications. It is a point-based clustering method that starts with the cluster centers that are initially placed at arbitrary positions and then proceeds by moving each step at the cluster centers in order to minimize the clustering error. K stands for the number of clusters. Distance measures that can be used are : euclidean, city-block, cosine, correlation and hamming. K-means computes centroid clusters differently for the different supported distance measures [174].

This clustering algorithm, in its standard formulation consists mainly of four steps that are briefly described below [175]:

1. Choose the number of clusters (k) and randomly select the initial cluster centers.

2. Generate a new partition by assigning each data point to the nearest cluster center.

3. Re-compute the cluster centers by using the current cluster memberships.

4. Repeat steps 2 and 3 until a distance convergence criterion is met.

The K-means clustering algorithm is used by Chitade and Katiyar [171] to classify satellite images into five classes. Zhuolin et al. used the K-means clustering algorithm for face and action recognition [176]. In [177], the authors used the K-means clustering algorithm for face recognition. At each clustering step, the test and training faces are projected to a discriminant space and the projected training data are partitioned into clusters. Gutta et al. also used the K-means clustering algorithm for the ethnicity classification of face images [12].

2.6.2 Naive Bayesian

Bayesian probability, is one of the major theoretical and practical frameworks for reasoning and decision making under uncertainty. NB Classifier is a simple probabilistic classifier based on Bayes theorem. It builds a probability model on the category description for all feature vectors in the training set. During the testing, it computes the posterior probability of a feature vector belonging to a particular class [178].

The problem of classification of face images can be solved by using a Bayesian classifier. If C_i represents the i^{th} ethnic class; we can then have n ethnic classes $C_1, C_2 \ldots, C_n$. Suppose there are 3 ethnic groups namely Asian, White and Black, then $n = 3$. Each face image is represented by a feature vector $X = [X_1, \ldots, X_j]$. If the class of a face with a feature X has been found, then the probability that an image belongs to that particular class C_i is given by the posterior probability $P(C_i|X)$ of that class C_i given the feature vector X. One can write Bayes theorem in the form [179, 180]:

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \tag{2.3}$$

where $P(C_i|X)$ is the posterior probability of the feature vector $P(X|C_i)$ is a likelihood Probability Density Function (PDF) for feature vector X given that the image belongs to class C_i and $P(C_i)$ is a prior probability of the class C_i. The $P(X)$ is given by:

$$P(X) = \sum_{i=1}^{n} P(C_i)P(X|C_i) \tag{2.4}$$

which ensures the sum of predicted occurrence probabilities:

$$\sum_{i=1}^{n} P(C_i|X) = 1 \tag{2.5}$$

The prior probability $P(C_i)$ is simply taken as

$$R_i / \sum_{j=1}^{N} R_j \tag{2.6}$$

Where R_i is the number of images of a ethnic race type C_i. To implement Equations 2.3 and 2.4, we need PDF's $p(X|C_i)$ for all race types C_i. The most common types of distributions are Gamma, Exponential, Lognormal, Poisson, Weibull, Chi-Square and Beta [181, 182]:

- Gamma Distribution: The gamma probability density function with parameters $\alpha \geq 0$ and $\lambda \geq 0$ is

  $$f(x;\alpha,\lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$$

  The quantity $\Gamma(\alpha)$ is called Gamma function and is given by: $\Gamma(\alpha) = \int\limits_0^\infty e^{-x} x^{\alpha-1} dx$

- Exponential Distribution: The exponential distribution is given by:

  $$f(x;\lambda) = \lambda e^{-\lambda x} \; ; x \geq 0 \; ; \lambda > 0$$

- Lognormal Distribution : The lognormal distribution is a probability density function of a random variable whose logarithm is normally distributed and is given by:

  $$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(lnx-\mu)^2}{2\sigma^2}}$$

- Poisson Distribution: The Poisson distribution is given by:

  $$f(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

- Weibull Distribution: The Weibull density for $\alpha > 0$ and $\beta > 0$ is given by:

  $$f(x;\alpha,\beta) = \alpha\beta^{-\alpha} x^{\alpha-1} e^{[-(\frac{x}{\beta})^\alpha]}$$

Detailed information about different distributions can be found in [181, 182]. In our method, however, we assume that each class C_i has a Gaussian distribution with a mean of μ_{C_i} and a standard deviation of σ_{C_i} which is defined as

$$p(X|C_i) = \prod_{j=1}^{n} P(x_j|C_i)p(C_i) \tag{2.7}$$

Where $X = (x_1, \ldots, x_p)$ and

$$p(x_j|C_i) = \frac{1}{\sqrt{2\pi\sigma_{C_i}}} e^{-\frac{(x_j-\mu_{C_i})^2}{2\sigma_{C_i}}} \tag{2.8}$$

After getting the PDF for all race types, one can use Equation 2.3 to calculate the occurrence probability of any race type at known indicator variables. If the posterior occurrence probability is highest for vector X into a race type C_i then vector X belongs to C_i race.

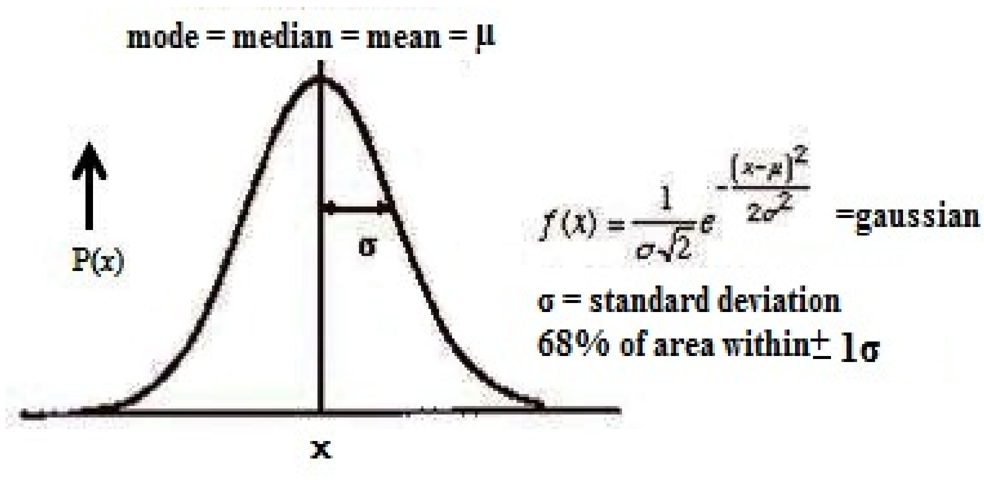


Figure 2.4: Gaussian Probability Distribution.

Figure 2.4 presents Gaussian probability distribution. It is also called a bell shaped curve or normal distribution. The total area under the curve is normalized to one. Note the bell shape of the curve and that its end/tail does not touch the horizontal axis below [183, 184].

Chengjun Liu used the Bayesian Discriminating Feature method for face detection. They have then used Bayes classifier to classify the images into two classes namely face class and non-face class where the face class contains only faces and the non-face class contains all other objects [165]. Khlifia and Mohamed proposed an approach to combining distance tangent, K-means algorithm and Bayesian network for image classification. First, they used the technique of tangent distance to calculate several tangent spaces representing the same image. Second, they cut the image into different blocks. For each block, they computed a vector of descriptors. They then used K-means to cluster the low-level features including colour and texture information to build a vector of labels for each image. They then finally applied five variants of Bayesian networks classifiers to classify the image of faces by using the vector of labels [180].

Hamid et al. [179] have used Bayesian classification theorem for the colour image segmentation. Their experiment result demonstrated that the Bayesian classifier is ideal as far as classifier accuracy is concerned when compared to the other two methods namely SVM and KNN classification methods. This was demonstrated by classifying images into foreground and background using colour features. When the image is exposed to uneven lighting conditions the result demonstrated that the Bayes classifier classify images better when compared to SVM and KNN classification methods. Kim et al. presented a gender classification task of discriminating between images of the faces of men and women. They used Gaussian process classifiers (GPCs) which are Bayesian kernel classifiers for the gender classification of face images [185]. In [166], the authors proposed a technique for direct visual matching of images for the purpose of face recognition and image retrieval, using a probabilistic measure of similarity, based primarily on a Bayesian analysis of image differences. Huang et al. proposed Bayesian estimation framework to deal with the patch similarity for predicting the gender from the facial images. They used a hybrid Bayesian framework to marginalize over the feature patches to determine the classification decision [167].

2.6.3 Multilayer Perceptron

A MLP is a feed forward Artificial Neural Networks (ANN) model that maps sets of input data onto a set of appropriate output. A MLP consists of multiple layers of simple, two state, sigmoid processing elements (nodes) or neurons that interact by using weighted connections. The MLP network has the lowest layer consist of an input layer, any number of intermediate layers or hidden layers and an output layer [186, 187]. Figure 2.5 presents a perceptron network model which contains three layers. This network has an input layer (on the left), one hidden layer (in the middle) and an output layer (on the right). The input layer has three neurons, middle layer has four neuron and output layer has two neurons. There is one neuron in the input layer for each predictor variable [187].

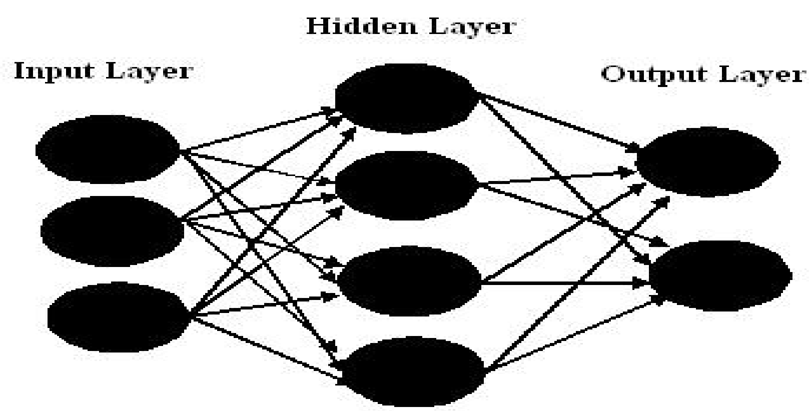


Figure 2.5: A Multilayer Perceptron Neural Network Model.

All neurons in each layer are fully connected to the neurons in an adjacent layer while there is no interconnection within a layer. In MLP, all nodes and layers are arranged in a feed-forward manner. Weights measure the degree of correlation between the activity levels of the neurons that they connect [186].

Let w_{ij} be the connection between a node i in the input layer and a node j in the hidden layer. The output y_{ku} of the j^{th} node in the hidden layer corresponding to the input vector X_u is given by:

$$y_{ju} = g(\sum_{i=1}^{d} w_{ij} x_{ui}); \quad j = 1, \ldots, h \tag{2.9}$$

where x_{ui} is the i^{th} component of the input vector X_u and $g(x)$ is the sigmoid function defined as

$$g(x) = \frac{1}{1+e^{-x}}. \tag{2.10}$$

Similarly , the k^{th} output y_{ui} of the output layer is

$$y_{ku} = g(\sum_{j=1}^{h} w_{jk} x_{ju}); \quad k = 1, \ldots, m \tag{2.11}$$

There are a few issues involved in designing and training an MLP network such as [187]:

- Selecting the number of hidden layers to use in the network;
- Deciding how many neurons to use in each hidden layer;
- Finding a globally optimal solution that avoids local minima;
- Converging to an optimal solution in a reasonable period of time; and
- Validating the neural network to test for over-fitting.

One or two hidden layers are adequate to tackle any nonlinear complex problem. For almost all problems, one hidden layer is sufficient. Two hidden layers are required for modeling data with discontinuities such as a saw tooth wave pattern. Using two hidden layers hardly improves the model, and it might present a greater risk of converging to a local minima. There is no hypothetical explanation behind utilizing more than two hidden layers. Three layer models with one hidden layer are suggested. The unnecessary increasing of a hidden layer may cause an increase in the complexity of the network [187–189].

One of the most important characteristics of a perceptron network is the number of neurons in the hidden layer(s). The nodes in the hidden layer are calculated as [188, 189]:

$$a = (attributes + classes)/2 \tag{2.12}$$

Tamura et al. [190] used a neural network model based on the MLP, and used a back-propagation algorithm to classify gender from face images. They used multiple resolutions face images (from 32-by-32 to 16-by-16 and 8-by-8 pixels). They conducted an experiment on 30 test images which shows that their network is able to determine gender from face images of 8-by-8 pixels. They reported an average error rate of 7%. In [191], the problem of face recognition using the MLP and the fuzzy MLP is addressed. In both cases, the Gabor wavelet transforms of the input face images are taken for four scales and six orientations. This feature vector is used as input to the MLP for classification. Sankar et al. presented a neural network model based on the MLP, and used the back-propagation algorithm, which is capable of the fuzzy classification of patterns. The model converts numerical and linguistic inputs to linguistic terms, and provides output decisions in terms of class membership values [186].

2.6.4 Support Vector Machine (SVM)

The SVM is a relatively new ML method originally used for binary classification [192, 193]. It has also been used for multi-class classification [194, 195]. It is one of the most used nonparametric supervised classifiers available today. It provides a novel means of classification using the principle of structural risk minimization. The decision surface is a weighted combination of elements of the training set. These elements are called support vector and characterize the boundary between the two classes. The goal of the SVM is to create a model which predicts the target estimation of data instances in the test set which are only given the attributes [7, 196]

.

Classification is done using training and testing data which consist of some data instances. The input to the SVM algorithm is a set (x_i, y_i) of labeled training data , where x_i is the data and y_i = -1 or 1 is the label.

$$x_i \in R^n, y_i \in \{-1, 1\}(i = 1, \ldots, l) \tag{2.13}$$

When we assume that a definite and separable hyperplane exists, this hyperplane can be expressed as

$$(w.x) + b = 0 \tag{2.14}$$

By applying to the Quadratic Programming Problem and solving, the decision function can be stated as:

$$f(x) = \sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b \tag{2.15}$$

Here, $K(x,x_i)$ is a kernel function, and it is used to determine to which class the input data x belongs through $f(x)$. There are several kernel functions that have been used in the literature such as

Polynomial kernel: $K(x,x_i) = (x_T.x_i + 1)^d$

Guassian kernel: $K(x,x_i) = exp(-x - x_i)/\sigma^2$

RBF kernel: $K(x,x_i) = exp - \gamma \|x - x_i\|^d, \gamma > 0$

Pearson VII function based universal kernel:

$$K(x,x_i) = 1/[1+(2(x-x_i)\sqrt{2^{(1/\omega)}}-1/\sigma)^2]^\omega \tag{2.16}$$

Hosoi et al. used the SVM for ethnicity classification of face images. In [103], the authors used a SVM for classifying a frontal face into Asian or non-Asian. They used the SVM for a training process and combined different SVM classifiers to some new classifiers, which improved the classification rate to a new level. Shrawan and Shubhamoy proposed the effect of various kernels and feature selection methods on the SVM performance for detecting email spams. The interaction of four Kernel functions of SVM i.e. "Normalised Polynomial Kernel (NP)", "Polynomial Kernel (PK)", "Radial Basis Function Kernel (RBF)", and "Pearson VII Function-Based Universal Kernel (PUK)" with three feature selection techniques i.e. "Gain Ratio (GR)", "Chi-Squared", and "Latent Semantic Indexing (LSI)" have been tested on the "Enroll Email Data Set" [196]. Saeed and Ehsan used the SVM method for sea target detection using HSV colour space. The SVM is utilized for classification and colour features are applied to train the SVM [197].

2.7 Experimental Environment

2.7.1 Data Set

The images used in this research project are taken from five databases, the ORL face database [198], the Asian face image database PF01 [199], the MORPH database [200], MUCT face database [199] and the Indian face database [201]. The ORL and the Asian face image database were chosen for component detection experiments for several reasons. These databases contain images with different scales, orientations and facial expressions. These databases were used to conduct the experiment with different orientations (left and right profile) and facial expressions (neutral, sad, happy, mouth open, surprised and eyes closed).

MORPH database is a collection of facial images taken from public records. The images were taken under real-world conditions and not by researchers under controlled conditions. The MORPH public release used for this research work is from Album 2. MORPH Album 2 for ethnicity estimation on faces has been used in previous works [104]. The data was captured under real-world conditions and not in a research environment. Results from the MORPH should provide a more accurate measure of performance of the proposed methods in real-world applications. Most of the images in the MORPH face database are from the ethnic group of White and Black. There are very less Asian face images in the MORPH database hence the Asian and Indian face database was used in combination with the MORPH face database. Additionally the MUCT face database was used to check the performance of the proposed method. The MUCT database consists of face images with different lighting conditions, ages, facial expressions, and ethnicities. Most of the images in the database are from the White and Black ethnic groups.

2.7.2 Software Development Environment

The proposed method is experimented in the working platform of MATLAB R2010a with the system configuration, Intel(R) Core i3-4005U CPU @ 1.70 GHz with 4GB RAM. The tool box used in MATLAB were image processing toolbox, Statistics and Machine Learning Toolbox and various other functions that are used in the software. The Weka 3.6.11 data mining software was used for ethnicity classification.

2.8 Conclusion

Biometrics has a long history and is widely used in our day-to-day life. Among the main advantages of biometrics is security, which shows how accurately and attentively people treat their private information. There are many applications and solutions of biometrics technology that are utilized as part of security systems. Despite the fact that the biometrics security system still has many concerns, for example physical privacy, information privacy and religious objections; users cannot deny the fact that this new technology will improve our lives.

We have presented an extensive survey of ethnicity classification algorithms using face images which uses different facial features. Some of the best algorithms are still too computationally expensive to be applicable for real-time processing, but this is likely to change with the up-coming improvements in computer hardware. Also, most of the literature addresses the problem of ethnicity classification into two classes namely Asian and non-Asian and some addresses the problem into three classes namely Asian, White and Black.

In this research work, ethnicity classification is done by using the colour and texture of facial features. The fusion of colour and texture features is also considered for accuracy improvement. The study of how different facial components (the eyes, the nose and the mouth) contribute to the ethnicity classification task is also discussed in detail. The methods used are very simple and efficient.

Chapter 3

Automatic Detection of Face and Facial Landmarks

3.1 Introduction

Face and facial landmark detection has been a topic of extensive research for several decades due to the important role in a number of applications such as face detection, face recognition and facial expression analysis. Face detection in an image is the first and key step in face recognition. Different face detection algorithms have been proposed in the literature. More information about the subject could be found in [15, 17, 142, 202]. Detailed literature of face detection and facial feature extraction is discussed in sections 2.5.1 and 2.5.2.

Face detection using a colour-based approach is important due to its invariance to face scaling, poses and facial expression. A face detection algorithm based on complexional segmentation and eyes location is presented in [203]. The algorithm establishes a proper colour model in the concrete colour space. In [203], the authors distinguish the facial regions and the non-facial regions using the facial complexion segmentation approach. Once a face is detected in an image, the next step is to extract facial features from the image. These features generally include eyes, mouth, nose and chin.

Terrilon et al. used a skin colour model based on the Mahalanobis metric and a shape analysis based on invariant moments to automatically detect human faces in two-dimensional natural scene images. First, they performed the colour segmentation of an input image by thresholding in a perceptually plausible hue-saturation colour space where the effects of the variability of human skin colour and the dependency of chrominance on changes in illumination are reduced. They they grouped regions of the resulting binary image which had been classified as face candidates into clusters of connected pixels [130]. In [204], the authors proposed an algorithm that creates a skin colour model for the face region detection from colour images.

In this research work, face detection and facial features extraction from gray-scale images using a connected component labelling algorithm is proposed. Section 3.2 gives a detailed description of the proposed face detection model. Section 3.3 gives a detailed description of the process of a face landmark localization. Section 5.5 presents experimental results followed by a conclusion in section 3.5.

3.2 Face Detection Process

In this section, face region extraction process is discussed using the segmentation approach from an input face image. The two constraints imposed on the input image are presented. The image considered should:

- Have a dark background; and
- Not have more than one face.

Figure 3.1 presents a flow diagram of the face detection process. The segmentation algorithm that has been presented can be divided into two main modules. The first module will convert the gray-scale input image into a binary image. The second module locates the face in the binary image and then inserts a bounding box around that face.

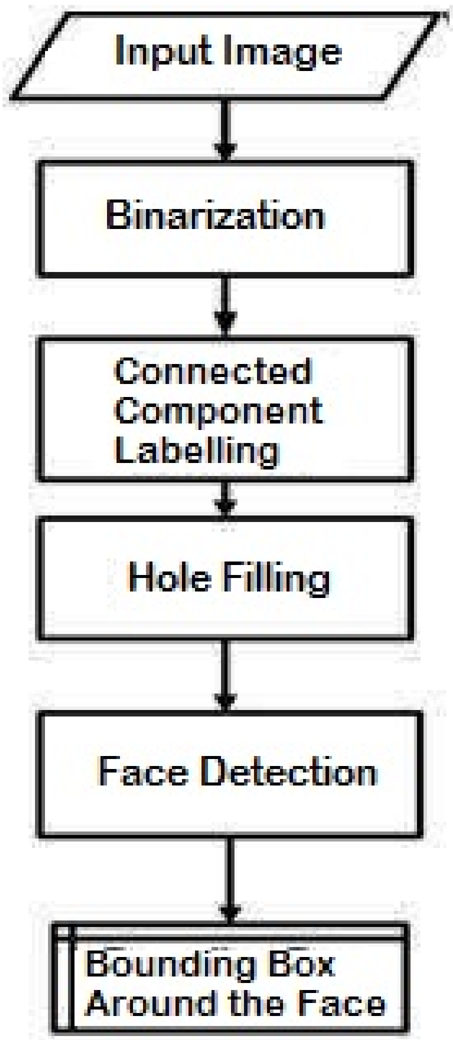


Figure 3.1: Face detection process.

3.2.1 Binarization and Connected Component Labelling

Numerous binarization algorithms have been presented in the literature. The binarization algorithm based on the standard deviation and the mean of the input image as described in [141, 205] is used in this study. An input image I with m rows and n columns is defined as:

$$\begin{aligned} I &= \{(i, j, x_{ij}) | 0 \leq i \leq n-1, \\ & 0 \leq j \leq m-1, 0 \leq x_{ij} \leq g-1\} \end{aligned} \tag{3.1}$$

where n, m and g are positive integers. The threshold T for an input image can be computed as follows:

$$t = k_1\sigma + k_2\mu \tag{3.2}$$

The value of k_1 and k_2 should be chosen from 0 to 2 depending on the resolution quality [141, 205]. In our work, we have established that $k_1 = 0.65$ and $k_2 = 0.9$ give the optimal solution. μ and σ are the mean and the standard deviation respectively. Formally, binarization transforms a gray level image I into a binary image I_B defined as:

$$I_B(i, j) = \begin{cases} 1 & \text{if I(i,j)} \geq T \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

The face will appear on the foreground of the binary image as shown in Figure 3.2(b) after the binarization process. Connected component labelling algorithm using 8-connectivity will then be applied on the resulting image in order to single out the main face components. Given the binary image, I_B, connected component labelling consists of partitioning I_B as follows:

$$I_B = B \cup (\cup_{i=1}^{n} F_i) \tag{3.4}$$

Where B is the background and F_i, for $i = 1, 2, \ldots, n$, are the connected components belonging to the foreground. All pixels in each component have the same label. Different algorithms for connected component labelling have been presented in the literature [206, 207]. Figure 3.4(a)

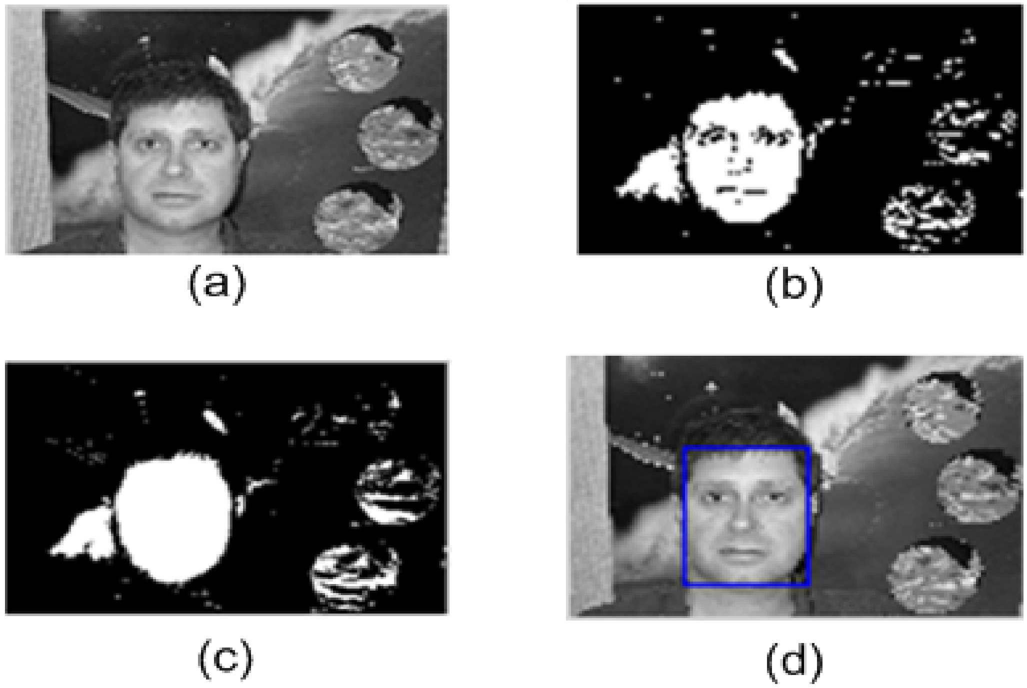


Figure 3.2: Different phases of face detection. (a) Original image (b) Binarized image (c) Resulting image after the hole filling process. (d) Detected face with rectangular boundaries.

and 3.4 (b) show some examples of connected components of binary images. A post-processing, consisting of hole-filling inside the faces, is then performed as shown in Figure 3.2(c). After applying the hole filling algorithm there may be more than one blob in an image. A blob with the greatest surface area is then chosen. The final outcome comprises of a bounding box put around the largest blob, which is most likely the face detected as shown in Figure 3.2(d). Figure 3.2 shows step-by-step the results of face detection when using the proposed method.

3.3 Face Landmark Localization

After detecting a face in an image, the next phase will involve finding the landmarks for solving face related problems. Most common landmarks in a face are the eyes, the nose and the mouth. To extract these landmarks, the face image is first divided into two parts namely the upper and the lower part with the help of masking. From the upper part, the eyes are extracted and from the lower part the nose and the mouth are extracted.

3.3.1 Concept of Masking

A basic concept in image processing is that of applying a mask to an image. The concept comes from the image processing operation of convolution, but is used in a general sense in image analysis as a whole. A mask is a small matrix whose values are called weights. The application of a mask to an input image yields an output image of the same size as the input [208, 209].

A binary mask exclusively has pixel intensity values of "0" (background) or "1" (object). Masking images allows the image processor to discriminate a Region of Interest (ROI); these may also be referred to as areas of interest or objects. A second reason to mask images is to remove noise from the images. In our face landmark detection process, the face is first divided into two parts: the upper and lower parts as shown in Figure 3.3(e) and (f). This is done by the use of a binary mask. The mask used is shown in Figure 3.3(c) and (d) for the eyes, the nose, and the mouth respectively. The result of the application of those masks on the original image is shown in Figure 3.3(e) to (f). A square mask was used for the eyes and an oval mask was used for the nose and the mouth. The square mask used for the upper part in order to obtain a specific area of interest i.e. the eyes, and an oval mask was used for the lower part in order to remove the unwanted noise from the chin area. Dividing images into two parts helps in locating facial parts from face images as compared to the whole image.

The proposed algorithm firstly searches and eliminates all the components with pixels touching the outer border of the image. Then, the components that remain in the inner face space are most likely face components. Thereafter, the size of each remaining component is computed as its total number of pixels. Based on the fact that, in the inner face space, the biggest components are most probably the eyes, the mouth, and the nose the remaining components are tested. The two biggest components from each masked image are then selected. As shown in Figures 3.4(a) and 3.4(b) the biggest components will be the eyes, the mouth, and the nose. Bounding boxes are then

Figure 3.3: ROI detection. (a) Detected face in input image (b) cropped face region (c) mask for upper part (eyes) (d) mask for lower part (nose and mouth) (e) masked image for upper part of image (f) masked image for lower part of image.

placed around the eyes, the mouth, and the nose on the original image as shown in Figure 3.4(e).

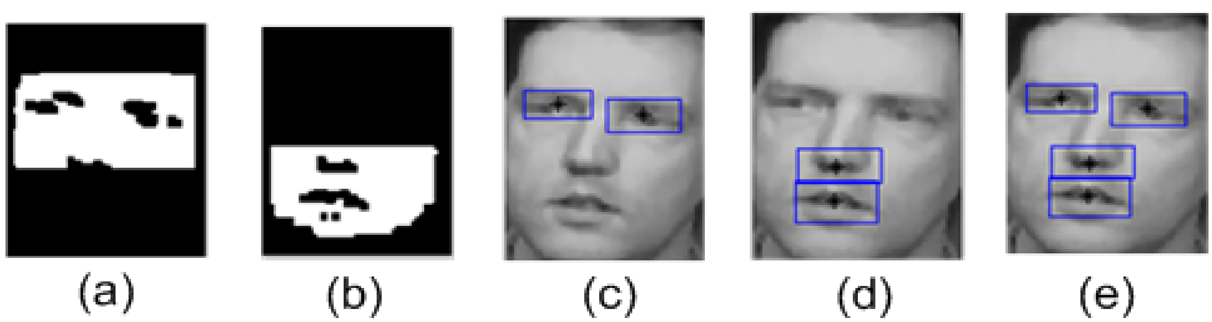


Figure 3.4: (a) and (b) labelled image (c) and (d) detected bounding box for the component (e) detected components on single image.

3.3.2 Proposed Algorithm

Given an image I, the ROI containing the face previously detected, and the proposed method to detect the landmarks are summarised in $Algortithm 1$.

In Algorithm 1 $ExtractBoundingBox(CC)$ extracts the bounding box around the connected component CC. The output of the proposed algorithm is a bounding box around each of the fol-

Algorithm 1 Automatic Detection of Face Components

Require: I ▷ Image Region of Interest containing the face

Ensure: *LEye, REye, Nose, Mouth*, C_{le}, C_{re}, C_n, C_m ▷ Bounding boxes representing the left eye, the right eye, the nose and the mouth and their centroids.

Split I into two regions, upper ($Part_1$) and lower($Part_2$) parts.

Binarize $Part_1$ and $Part_2$ using suitable threshold values of k_1 and k_2. ▷ We have used $k_1 = 0.65$ and $k_2 = 0.9$ for both $Part_1$ and $Part_2$

Calculate, using 8-connectivity, the connected components, remove the components which are touching the outer border of the image, to obtain $CPart_1 = (CPart_{1i})_{i=1,2,...,p}$ and $CPart_2 = (CPart_{2i})_{i=1,2,...,q}$ of relevant connected components of $Part_1$ and $Part_2$ respectively

Sort $CPart_1$ and $CPart_2$ in increasing order of sizes of components.

if $p < 2$ **then** ▷ The two eyes have not been detected which may be due to the hairs that are covering the eye, or that the eye got connected to the border during the binarization process.

  Extract the second eye component based on human facial geometry.

else $LEye = ExtractBoundingBox(CPart_{1,p-1})$, $REye = ExtractBoundingBox(CPart_{1,p})$

end if

if $q < 2$ **then** If only one component is detected, the second one is computed depending on the component already detected and if we don't find any valid component in this part then we can extract the mouth and the nose of the person depending on the component detected in the upper part of the face. ▷ In most cases it is due to the beard

else $Nose = ExtractBoundingBox(CPart_{2,q-1})$, $Mouth = ExtractBoundingBox(CPart_{2,q})$

end if

$C_{le} = Centroid(CPart_{1,p-1})$ ▷ Left eye centroid

$C_{re} = Centroid(CPart_{1,p})$ ▷ Right eye centroid

$C_n = Centroid(CPart_{2,q-1})$, ▷ Nose centroid

$C_m = Centroid(CPart_{2,q})$ ▷ Mouth centroid lowing face components: the two eyes, the nose, the mouth and centroids of face components. Figure 3.5 (d) show an example of detected face components.

3.4 Experimental Results

The proposed method was tested on images from the *ORL face database* [198]. The ORL database contains 400 images of 40 different subjects and has 10 images per person in different orientations, scales and facial expressions. The images are gray-scale images of size 112 x 90. The detection rate of face components in a frontal, a left, and a right view is presented in Table 3.1. In the frontal view, the detection rates of the left eye and the right eye are nearly 96%, the nose 92%, and the mouth over 88%. In the left view, the ratio of detection of the nose, the mouth, and the right eye is over 90% and the left eye is over 75%. In the right view, the ratio of detection of the left eye and the nose is over 90 % and for the right eye and the mouth is over 80%. Figure 3.6 shows some results of detected components of faces in different orientations. The detection rate of face components in different facial expressions are presented in Table 3.2. The detection rate of the left eye and the right eye for a smiling facial expression is 95%, and the nose and the mouth is 75%. The detection rate for a sad facial expression is 90% for the left and the right eye, 80% for the mouth and 75% for the nose. The detection rate for the closed eyes has no effect on the detection rate that is used by our method. Therefore, even if the eyes are closed, it will still be detected. The detection rate for the open mouth is 92% for the left and the right eye, 85% for the nose and 80% for the mouth. Examples of detected facial components in different facial expressions are shown in Figure 3.7.

Moreover, the proposed method was tested on the Asian face image database PF01 [199]. The PF01 database contains the true-colour face images of 103 people having 17 images per individual. Images in this database are normal images and images of systematic variations for illumination, pose and expression. The detection rates of face components in a frontal, left, and right view is presented in Table 3.1. The detection rates for the left and the right eye, the nose and the mouth are 98%, 96%, 87%, and, 92% respectively. In the left view, the detection rate for the left eye and the mouth are 95%, for the right eye it is 96% and for the nose it is 89%. In the right view, the detection rate for the left eye is 94%, for the right eye it is 95%, for the nose it is 90% and for the mouth it is 96%. The detection rates of face components in different facial expressions are presented in Table 3.2. The different facial expressions in the PF01 database are happy, surprised, irritated and the closed eye. The detection rate for a happy facial expression is above 94% for the left and right eyes, 71% for the nose and 89% for the mouth. The detection rate for a surprised facial expression is, above 94% for the left eye, the right eye, and the mouth, and 89% for the nose. The detection rate for an irritated facial expression is, 94% for the left eye, 96% for the right eye, 86% for the nose and 98% for the mouth. The detection rate for a closed eye is 95% for the left eye, 96% for the right eye, 94% for the nose and 98% for the mouth.

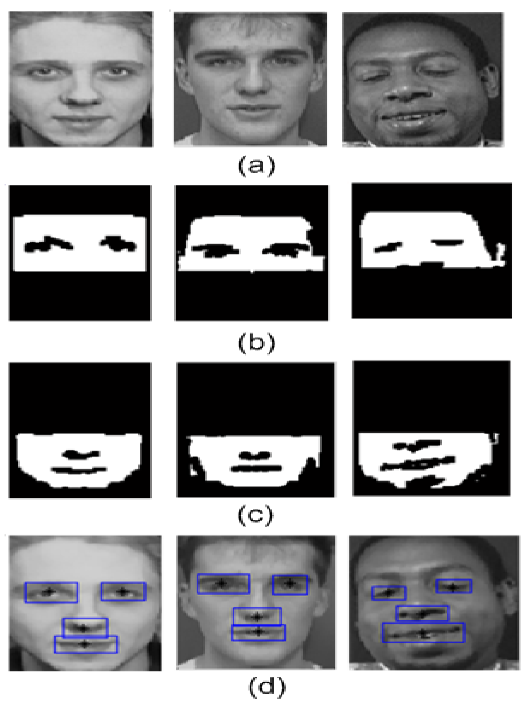


Figure 3.5: (a) input image (b) connected component in upper half part (c) connected component in lower half part (d) detected component with bounding box.

Figure 3.6: Faces in different orientations with bounding boxes around the detected components.

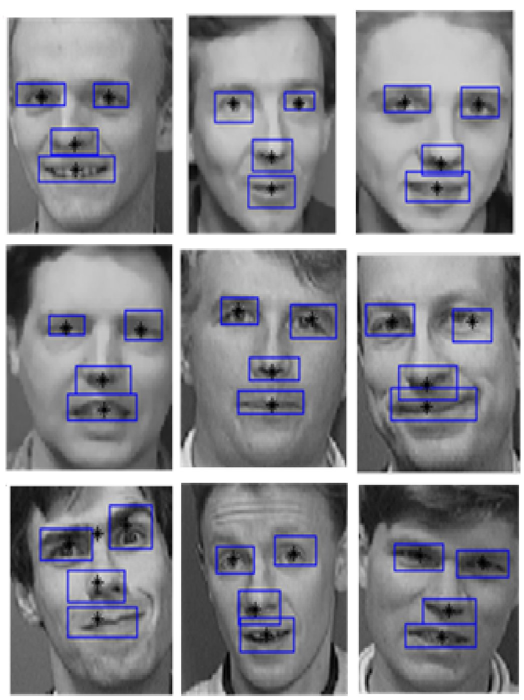


Figure 3.7: Faces with different facial expressions and bounding boxes around the detected components.

Table 3.1: Detection rate (%) of facial component in the frontal, left, right view.

Face Orientation	ORL face database				PF01 database			
	Left Eye	Right Eye	Nose	Mouth	Left Eye	Right Eye	Nose	Mouth
Front View	96	96	92	88	98	96	87	92
Left View	75	96	90	93	95	96	89	95
Right View	92	82	93	87	94	95	90	96

Table 3.2: Detection rate (%) of facial component for different facial expression.

ORL face database					PF01 database				
Facial Expression	Left Eye	Right Eye	Nose	Mouth	Facial Expression	Left Eye	Right Eye	Nose	Mouth
Smiling	95	95	75	75	Happy	95	94	71	89
Sad	90	90	75	80	Surprised	96	95	89	94
Closed eyes	96	96	87	85	Irritated	94	96	86	98
Mouth Open	92	92	85	80	Closed eyes	95	96	94	98

3.5 Conclusion

A fully automatic method was designed. The proposed method detects a face and facial landmarks in a frontal, a left and a right face view. In addition, it detects the expressive images of high complexity. The complexity of the expression was presented by closed/semi-closed eyes, smiling, sad, a variety of mouth appearances including an open and tight mouth. The proposed segmentation approach extracts the entire face region by combining thresholding and connected component labelling. The system locates facial features, which are the eyes, the mouth, and the nose in an image. Additionally, it also does not require any manual intervention by users, such as manually creating a feature location model.

Extensive experiments demonstrate that the proposed model has achieved very good results with face images from frontal, right and left view and with changes in facial expressions whereas in [141] they did not use images with different facial expressions. They reported 89.1% overall accuracy where as we achieved 90% accuracy for ORL face database and 93.58% for PF01 face database. The number of images they used for their experiments were much fewer. In [210], the authors reported an accuracy of 91% for the facial component extraction using rectangular measure. In [211], the authors reported an accuracy of 89.25% for facial component extraction. In our case we achieved 90.70% accuracy. The presented approach is simple and fast for detecting facial components. However, there are some challenging cases, such as faces with beard and specs, for which it has a poor performance. In future, the method can be extended in order to handle faces which have those challenges.

Chapter 4

A Comparative Study of Face-Based Ethnic Classification using Colour Space Models

4.1 Introduction

The human face conveys a large amount of visual information. By looking at someone's face, information processes about that person includes age, gender, ethnicity and emotions. During the past decade, research involving face-based ethnicity classification has emerged, and grown rapidly. A detailed review on some progress made can be seen in section 2.5.3.

Skin colour is one of the significant features of the human face. It not only varies from individual to individual belonging to different ethnic groups but also across different regions. For example, the skin colour of people belonging to European, Asian and African groups range between white, yellow and dark. There are several colour spaces that have been used in in early work of skin detection, such as RGB, normalized RGB, HSV, YIQ, YCbCr, HSI and CIELAB [212]. A wide variety of studies and different approaches regarding the skin colour modelling, segmentation and recognition have been proposed. In [132], an algorithm is used to segment skin colour regions using a combination of colour and edge information in colour images. They used a skin colour model based on the Bayesian decision rule for minimum cost and non-parametric density estimation to detect the presence of skin colour in an input image. The detected skin-coloured regions are then refined by using the homogeneity property of the human skin.

Hashem used a technique which combines two methods for face detection in order to achieve good detection rates. The two methods are the face skin detection using HSV colour space and the back propagation neural networks [133]. Yoo et al. used hue and Cr colour information to find the skin colour pixels in an image. They also used a range of thresholds acquired from red and blue components in normalized RGB colour space to eliminate non-skin colour pixels. Post-processing is used to eliminate such noises by a morphological operator. Moreover, the algorithm performs temporal filtering to remove skin-colour pixels that immediately appear from frame to frame. This is done by using an object tracking process to act as a memory for collecting skin-colour objects obtained from the previous frame to guide the next frame [213]. Hasan and Misra established a system which can identify specific hand gestures. They used local brightness of each block of the hand gesture image and calculated the local brightness of each divided block after applying coloured segmentation operation by using the HSV colour model [214].

To our knowledge, no work on ethnicity classification with colour spaces has been introduced in public literature. In this chapter, ethnic classification of the human face is proposed by using different colour spaces. First, images are classified into two ethnic groups being Asian and non-Asian and then into three ethnic groups being Asian, White and Black. The aim is to determine which ethnicity group an individual is likely to belong to. The classification of the images is done by using the HSV, RGB and YCbCr colour space. The mean values are calculated for the hue, saturation, value, red, green, blue, Y, Cb and Cr colour components. A feature vector is created for each face image. The K-means classification algorithm and NB algorithm are then used for classification of face images into proper ethnic groups.

During this study, a comparative investigation of ethnicity classification by using RGB, HSV and YCbCr colour spaces are presented. Section 4.2 describes tools and methods used for ethnic classification. The techniques used for the segmentation of the images are discussed in section 4.3. Section 4.4 presents the experimental results and is followed by a conclusion in section 4.5.

4.2 Features Extraction

4.2.1 Colour Models

Colour adds very meaningful information for vision. Humans use colour to differentiate and separate an object. It is hence an interesting approach to perform segmentation using colour information. Based on the fact that the colour is the feature that is used for ethnicity classification, three colour models are discussed in the sections that follow.

RGB Colour Space

RGB is the most commonly used colour space for storing and representing digital images, because images captured by a camera are normally provided as RGB. This is an additive colour system based on the trichromatic theory. It is a hardware-oriented model and is well known for its colour-monitor CRT display purposes. RGB colour space represents three primary colours: red, green and blue, respectively. To decrease the dependence on lighting, RGB colour components are normalized so that the sum of the normalized components is united ($r + g + b = 1$). RGB is very common, and is being used in virtually every computer system as well as television, video. The bottom corner, when red = green = blue = 0 is black, while the opposite top corner, where red = green

= blue = 255 (for an 8 bit per channel display system), is white. RGB is frequently used in most computer applications since no transform is required to display information on the screen. RGB has been a most popular and common choice for skin-detection and has been used by[215–217]. In our experiment different combinations of red, green and blue components of input images are used for ethnicity estimation.

HSV Colour Space

HSV colour space is represented by three components namely Hue (H), Saturation (S) and Value (V). The HSV space describes colour as *Hue* which is the property of a colour that varies in passing from red to green; *Saturation* being the property of a colour that varies in passing from red to pink and *Value* (also called intensity or lightness or brightness) which is the property that varies in passing from black to white. Hue varies from 0 to 1 on a circular scale i.e. the colours represented by H=0 and H=1 are the same. Saturation varies from 0 to 1, 1 representing 100 percent purity of the colour. H and S scales are partitioned into 100 levels and the colour histogram is formed by using H and S. The hue component describes the colour itself and is depicted in the form of an angle between [0,360] degrees. 0 degrees means red, 120 green 240 blue, 60 degrees is yellow and 300 degrees is magenta. The saturation component signals how much the colour is polluted with the colour white. The HSV colour space has been used by Brown et al. [216], Garcia and Tziritas [218]

The RGB image can be converted into an HSV image by using

$$H = \arccos \frac{1/2(R-G)+(R-B)}{\sqrt{(R-G)^2+(R-B)(G-B)}} \tag{4.1}$$

Saturation is

$$S = 1 - 3\frac{min(R,G,B)}{R+B+G} \tag{4.2}$$

and Value is

$$V = 1/3(R+G+B) \tag{4.3}$$

YCbCr Colour Space

This is yet another hardware-oriented model. However, unlike the RGB space, here the luminance is separated from the chrominance data. This colour space consists of brightness (Y) and two colour difference components (Cb and Cr). The Y value represents the luminance (or brightness) component which is calculated as a weighted sum of RGB values, where as the chrominance (Cr and Cb) component is calculated by subtracting the luminance component from R and B values. The YCbCr colour space is one of the most popular choice for skin detection and has been used by Hsu et al. [195], Chai and Bouzerdoum [219], and Chai and Ngan [220].

The RGB image can be converted into YCrCb colour space as follows

$$Y = 0.299R + 0.587G + 0.114B \tag{4.4}$$

$$Cr = R - Y \tag{4.5}$$

$$Cb = B - Y \tag{4.6}$$

Figure 4.1 presents an example of an RGB image converted into red, green, blue, HSV, hue, saturation, value, YCbCr, Y, Cb, and Cr colour components using equations [4.1-4.6] for the three different ethnic groups of White, Black and Asian.

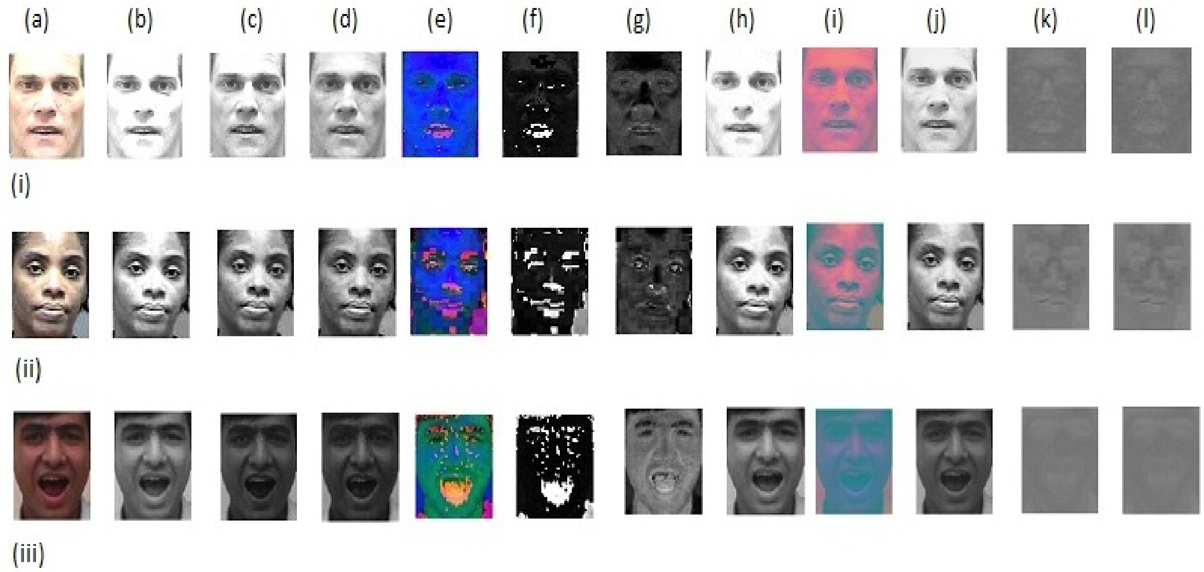


Figure 4.1: (a) RGB Image (b) Red colour component image (c) Green colour component image (d) Blue colour component image (e) HSV image (f) Hue colour component image (g) Saturation colour component image (h) Value colour component image (i) YCbCr Image (j) Yellow colour component image (k) Cb colour component image (l) Cr colour component image for (i) White (ii) Black and (iii) Asian.

4.2.2 Features Computation and Combination

Given a set of n face images. Each image is cropped and resized to 75×70. The RGB colour space is considered first. For each image, the mean value for the *red*, *green* and *blue* components are calculated, and stored in μ_r, μ_g, and μ_b using:

$$\mu = \frac{\sum_{i=1}^{n} X_i}{N} \tag{4.7}$$

The RGB image model can be converted into an HSV image model. It is then possible to calculate the mean of *hue*, *saturation* and the *value* of each cropped image, which is subsequently stored in μ_h, μ_s, and μ_v. The RGB image model can be converted into the YCbCr image model. It is then possible to calculate the mean of Y, Cb and Cr for each cropped image which is then, stored in μ_y, μ_{cb}, and μ_{cr}. The result is nine different mean colour components of each face image which is represented in the following vector

$$M = (\mu_r, \mu_b, \mu_g, \mu_h, \mu_s, \mu_v, \mu_y, \mu_{cb}, \mu_{cr}).$$

Different combinations of these components of this vector will be extracted to probe the combination that gives the optimal ethnic classification.

4.3 Ethnicity Classification

This section presents a method for identifying the ethnicity of a person from an input face image that uses different colour spaces. Two machine learning algorithms K-means clustering and NB are used for ethnic classification of face images as explained in section 2.6.1. The K-means clustering algorithm with k=2, is used for classifying face images into the two groups Asian and non-Asian; and k=3 to classify images into Asian, White and Black. Euclidian distance is used as similarity measure. A pseudo-code of K-means clustering is shown in Algorithm 2.

Algorithm 2 K-means Clustering Algorithm

Require: A set, $X \subset \mathbb{R}^d$, of faces features ($X = \{x_1, x_2, \ldots, x_n\}$); an integer k
Ensure: A partition of points into clusters $\omega_1, \omega_2, \ldots \omega_k$, along with a center μ_j for each cluster, so as to minimize function W defined in equation 2.2

```
1: Choose randomly k prototypes {μ_1^0, μ_2^0, ..., μ_n^0}
2: n ← 0
3: n ← n + 1
4: for j=1 to k) do
5:     ω_j^n ← {x ∈ X | ∀k ≠ j, f(x, μ_j^{n−1}) ≤ f(x, μ_k^{n−1})}
6:     compute μ_j^n from ω_j^n
7: end for
8: if ω_j^n ≠ ω_j^{n−1} and n ≤ n_0 then
9:     goto 3
10: end if
```

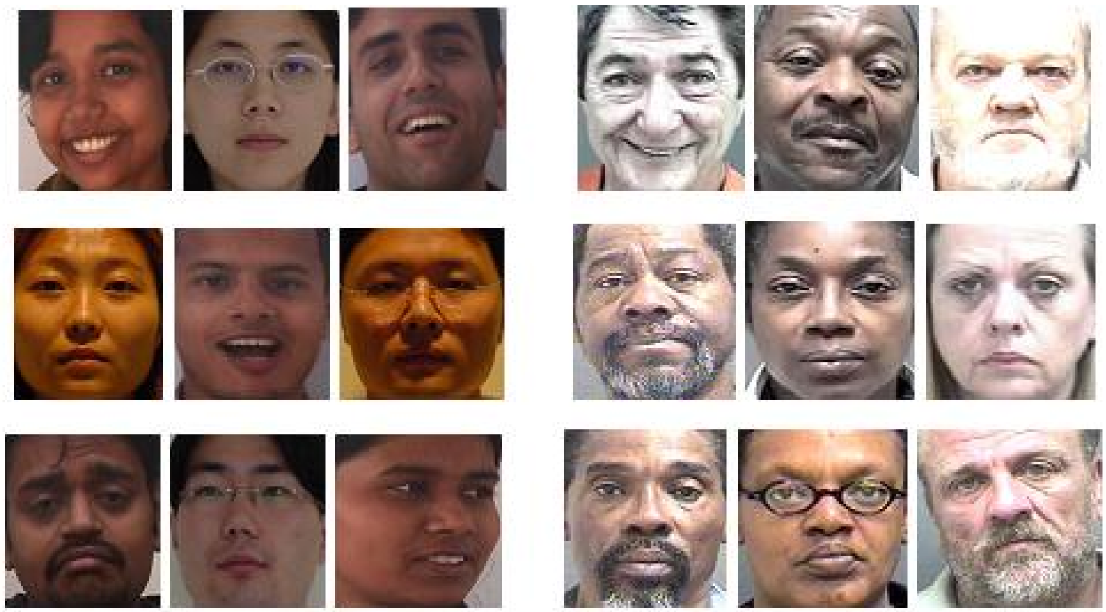
4.4 Experimental Results

The data-set used for experimentation is a combination of three different face data-sets. Together there are a total of 1013 images with different facial expressions, illumination and orientations that are used. They are first separated into two ethnic groups, Asian and non-Asian. The Asian group is composed of images from an Asian face database [199] and an Indian face database [201]. Most of the Asian faces are of Chinese, Korean and Indian origin whereas the Indian face database contains only Indian faces. The non-Asian group is composed of images from the MORPH database [200]. Most of the non-Asian faces are of White and Black ethnic groups. Table 4.1 gives a detailed description of the images that have been used. The non-Asian group is further separated into White and Black subgroups in order to help classify images into an additional three ethnic groups being Asian, White and Black. Sample images from the three databases are shown in Figure 4.2.

Table 4.1: Database description.

Face Database(Total Images = 1013)		Number of Images	Size of face Area	Variations Included
Asian (401 Images)	Asian PF01	317	75X70	Normal, Illum, Expr
	Indian Face Database	93	75X70	Pose, Illum, Expr
non-Asian (612 Images)	MORPH Database (White)	305	75X70	Age, Illum, Expr
	MORPH Database (Black)	307	75X70	Age, Illum, Expr

Figure 4.3 shows a sample of a K-means plot for ethnicity classification into Asian and non-Asian ethnic groups. Images are separated into two clusters. Figure 4.4 exhibits a sample of a K-means plot for ethnicity classification of images into Asian, White and Black ethnic groups. Furthermore, the NB algorithm is used to classify the images into appropriate ethnic groups. A Gaussian distribution is used to calculate a pdf for all race types. Figure 4.5 depicts a Gaussian distribution for two ethnic groups being Asian and non-Asian, by using a blue colour component. Figure 4.6 presents a Gaussian distribution for three ethnic groups being Asian, White and Black, through the use of a green colour component.



(a) (b)

Figure 4.2: Examples of face image (a) Asian (b) non-Asian.

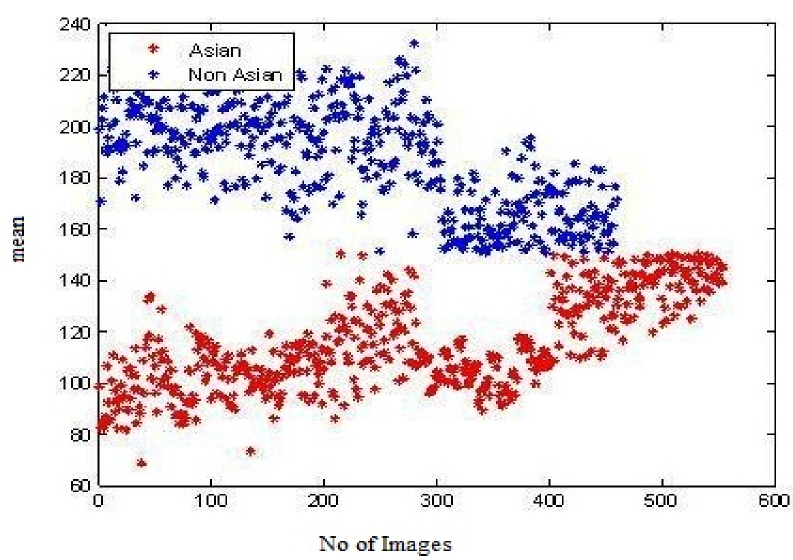


Figure 4.3: K-means classification output for two ethnic groups Asian and non-Asian.

Tables 4.2 to 4.4 show the ethnicity classification which uses the K-means clustering algorithm and NB algorithm for Asian and non-Asian ethnic groups. This is done by using the RGB, the HSV and the YCbCr colour components features computed in section 4.2.2. Tables 4.6 to 4.8

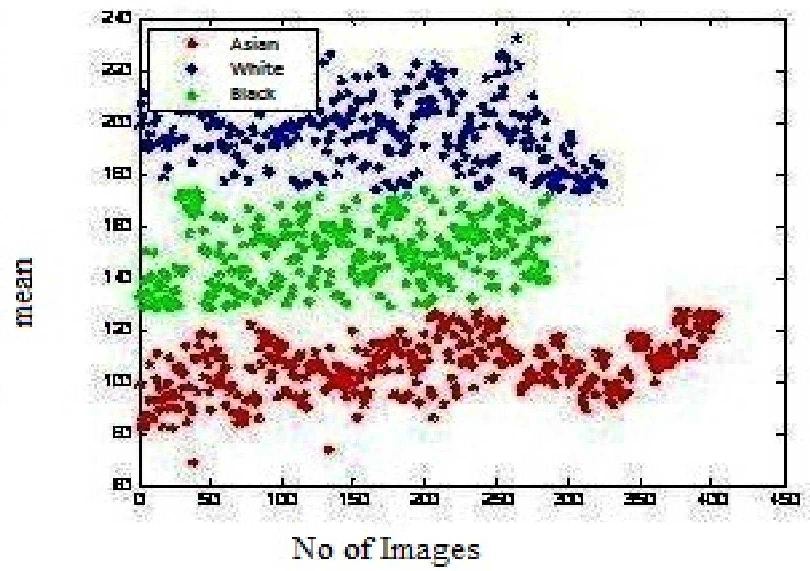


Figure 4.4: K-means classification output for three ethnic groups Asian, White and Black.

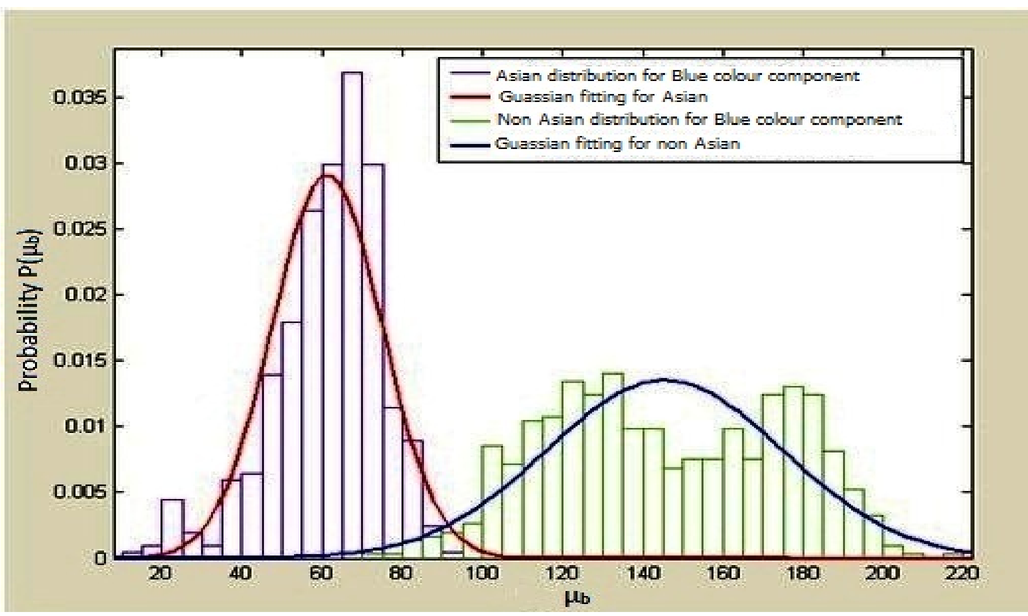


Figure 4.5: Guassian distribution for two ethnic groups using Blue colour component.

portray the ethnicity classification for the three ethnic groups of Asian, White and Black and via the K-means clustering and NB algorithm. This is done through the use of the RGB, the HSV and the YCbCr colour components features computed in section 4.2.2.

The K-means clustering and NB classification algorithms are then used to classify images

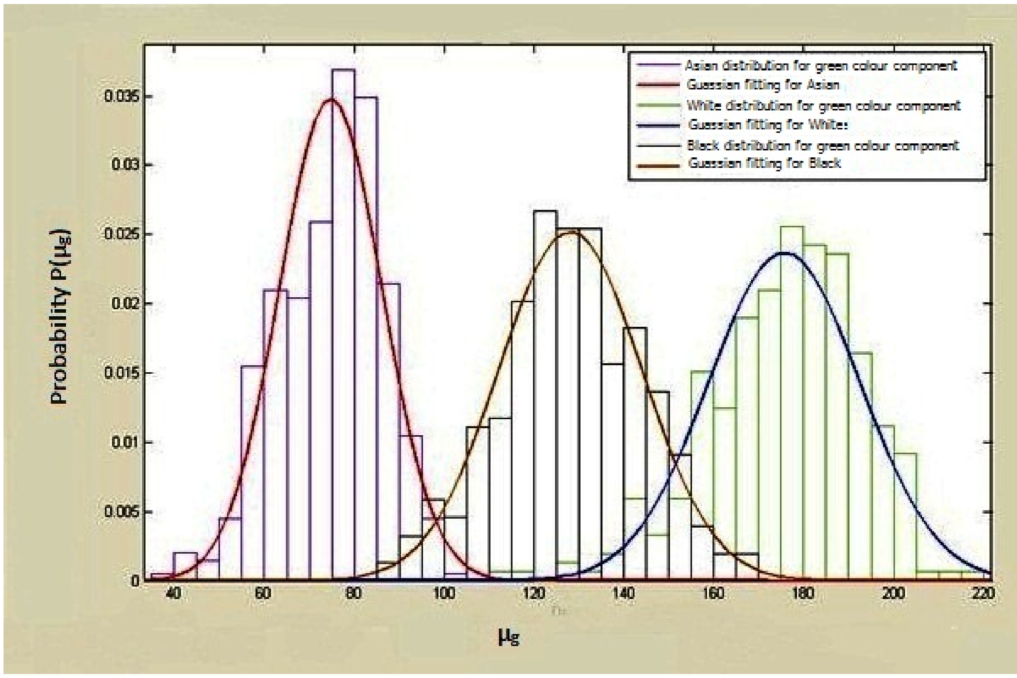


Figure 4.6: Guassian distribution for three ethnic groups using green colour component.

into Asian and non-Asian and subsequently into Asian, White and Black. 10-Fold cross validation is then carried out to evaluate the performance of the proposed algorithm. Tables 4.2 shows the classification rate of face images into Asian and non-Asian ethnic groups by using the RGB colour space through the application of K-means clustering and NB algorithms. Table 4.6 shows the classification into the Asian, White and Black ethnic groups. The blue colour component in the RGB colour space provides the best results when compared to the red and green colour components. When the RGB colour space is compared with other two colour spaces, it is evident that the RGB colour space gives the most reliable results.

Tables 4.3 and 4.7 show the ethnic classification that uses the HSV colour space. It can be seen that the value colour component delivers good results when compared to hue and saturation. Conversely, when hue, saturation and value colour components are combined, the results are even better. Tables 4.4 and 4.8 show the ethnicity classification rate using the YCbCr colour space. The colour component Cb and Cr does not contribute to the ethnicity classification but the Y colour component gives a promising result. Therefore, it can be concluded that blue, value and Y colour components consistently give very good results. The NB algorithm shows good results when compared to the K-means clustering algorithm. To improve the accuracy further, these colour components were combined and then classified using NB algorithm. The results of this classification are shown in Tables 4.5 and 4.9. It is found that combined colour components improves the accuracy of ethnicity classification. The accuracy for the RGB colour space colour outperforms that of HSV and YCbCr. YCbCr shows the poorest performance among the three colour spaces considered. Figure 4.7 shows a comparison graph of ethnicity classification using the three colour spaces.

The ROC curve is a graphical plot that is created by plotting TPR and FPR at various threshold values. The performance of a biometric system is also depicted visually on the curves. The ROC curves display the way in which the TPR varies with respect to the FPR and vice-versa [102]. Figure 4.8 shows the resulting ROC curves for the three colour space components of RGB, HSV and YCbCr respectively. From the curve, it is clear that in the RGB colour space, the blue component gives the best results. The value and yellow colour components give promising results in the HSV and YCbCr colour spaces.

Table 4.2: Ethnicity classification rate (%) into two ethnic groups using RGB colour space.

Colour Component	K-means		Naive Bayesian	
	Asian	non-Asian	Asian	non-Asian
Red Colour Component	100	75	94	94.3
Green Colour Component	100	85.42	99.50	97.71
Blue Colour Component	**100**	**90.35**	**99.71**	**98.52**
RGB Colour Component	100	85.62	99.50	96.89

Table 4.3: Ethnicity classification rate (%) into two ethnic groups using HSV colour space.

Colour Component	K-means		Naive Bayesian	
	Asian	non-Asian	Asian	non-Asian
Hue Colour Component	88.02	74.34	87.03	78.26
Saturation Colour Component	91.77	93.62	90.77	94.28
Value Colour Component	100	75.65	94.26	94.44
HSV Colour Component	**100**	**92.48**	**98**	**95.42**

Table 4.4: Ethnicity classification rate (%) into two ethnic groups using YCbCr colour space.

Colour Component	K-means		Naive Bayesian	
	Asian	non-Asian	Asian	non-Asian
Yellow Colour Component	100	83.33	98.50	97.38
Cb Colour Component	63.84	81.37	50.37	91.10
Cr Colour Component	63.84	81.37	50.37	91.10
YCbCr Colour Component	**100**	**87.58**	**98.25**	**98.85**

Table 4.5: Ethnicity classification rate (%) into two ethnic groups using a combination of three colour spaces.

Colour Component	Naive Bayesian	
	Asian	non-Asian
BVY Colour Component	99.50	97
RGBHSVY Colour Component	100	97.22
RGBHSVYCbCr Colour Component	**100**	**97.22**

Table 4.6: Ethnicity classification rate (%) into three ethnic groups using RGB colour space.

Colour Component	K-means			Naive Bayesian		
	Asian	White	Black	Asian	White	Black
Red Colour Component	93.01	94.42	78.50	97.51	94.28	80.78
Green Colour Component	100	90.49	89.25	99.50	92.13	86.97
Blue Colour Component	**100**	**87.21**	**92.83**	**99.75**	**90.49**	**89.25**
RGB Colour Component	100	90.81	88.59	99.50	92.78	86.97

Table 4.7: Ethnicity classification rate (%) into three ethnic groups using HSV colour space.

Colour Component	K-means			Naive Bayesian		
	Asian	White	Black	Asian	White	Black
Hue Colour Component	85.03	61.31	37.13	89.3	54.1	30.6
Saturation Colour Component	69.82	87.21	25.73	91.5	74.4	51.5
Value Colour Component	93.51	94.42	78.17	93.26	93.11	79.80
HSV Colour Component	**100**	**91.47**	**90.55**	**97**	**92.78**	**87.94**

Table 4.8: Ethnicity classification rate (%) into three ethnic groups using YCbCr colour space.

Colour Component	K-means			Naive Bayesian		
	Asian	White	Black	Asian	White	Black
Yellow Colour Component	99.75	90.81	87.94	98.50	92.78	87.62
YCbCr Colour Component	**100**	**91.7**	**90.5**	**99.50**	**92.13**	**91.20**

Table 4.9: Ethnicity classification rate (%) into three ethnic groups using a combination of colour spaces.

Colour Component	Naive Bayesian		
	Asian	White	Black
BVY Colour Component	99.25	92.78	86.97
RGBHSVY Colour Component	100	92.78	88.27
RGBHSVYCbCr Colour Component	**100**	**92.78**	**89.25**

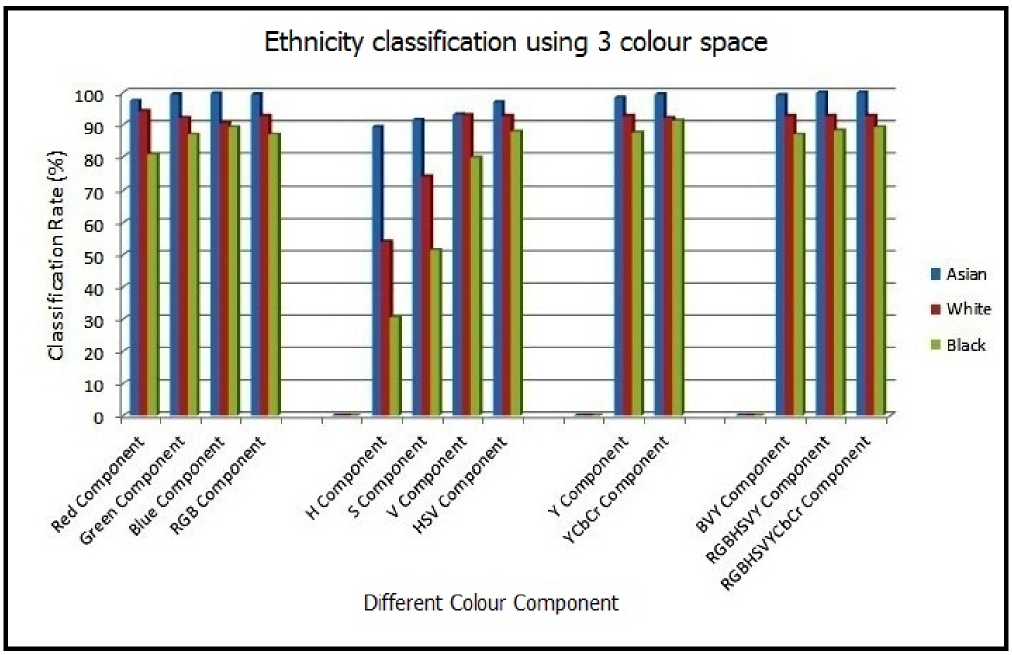


Figure 4.7: Graph of ethnicity classification using 3 colour spaces and a combination of 3 colour spaces.

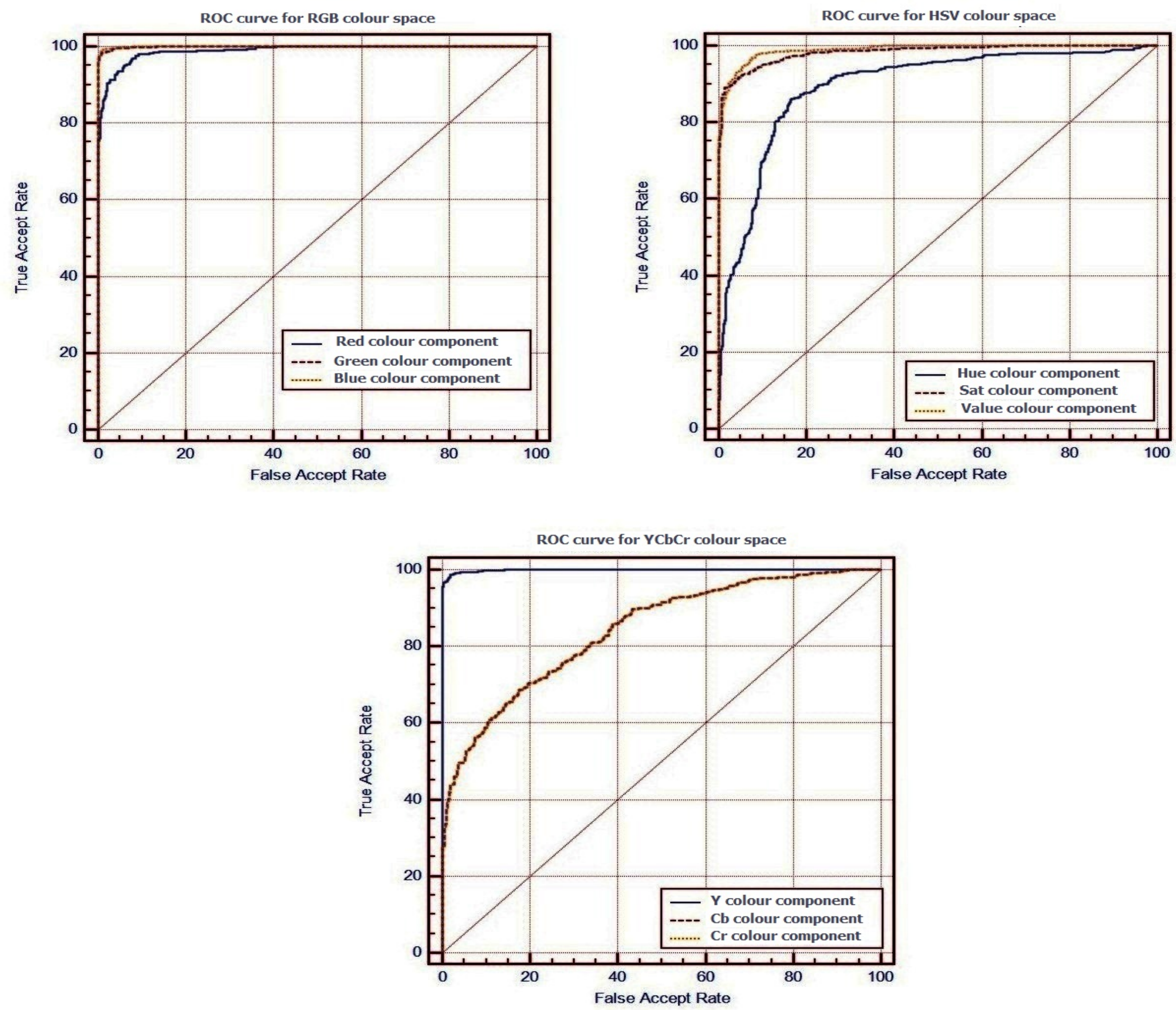


Figure 4.8: ROC curve for RGB, HSV and YCbCr colour components.

4.5 Conclusion

A simple and efficient method of ethnicity classification which is based on face images with high complexity has been presented in this study. The complexity of these images used has various characteristics namely: illumination condition, facial expressions, different orientations, and the presence or absence of facial hair and spectacles. Experiments were carried out using 1013 images. The presented method classifies face images into two ethnic groups, namely Asian and non-Asian, and with accuracies of 99.75% and 98.52% respectively. It then classifies face images further into three ethnic groups namely Asian, White and Black and produced accuracies of 99.50%, 92.13% and 91.20% respectively.

It is evident that the RGB colour space gives better results when compared to the HSV and the YCbCr colour space. Another important finding is that hue, saturation, Cb and Cr do not contribute significantly to the ethnicity classification. The blue colour component of the RGB colour space, the value component of the HSV colour space and the yellow component of the YCbCr colour space are more effective in ethnicity classification. For future work we aim to combine geometric, texture and colour features to further improve the accuracy of ethnicity classification.

Chapter 5

Gabor Filter-Based Texture Analysis of Facial Components for Ethnicity Classification

5.1 Introduction

Component-based face recognition and identification has been extensively studied over the past two decades. Human face components such as the eyes, the mouth and the nose are the most significant features of the face and are therefore said to be the most important characteristics in deciding the ethnicity [7]. Face processing is a difficult task, primarily because of the inherent variability of the image formation process in terms of photometry, image quality, geometry, disguise, change, and occlusion. These challenges are discussed in detail in [34, 221].

Ethnicity classification algorithms based on facial components such as the eyes, the nose and the mouth and the Gabor filter are introduced in this chapter. Four machine learning techniques are used for comparing the results. The following three combinations of ethnic groups are considered in the classification:

1. Two ethnic groups: Asian and non-Asian;
2. Three ethnic groups: Asian, White and Black; and
3. Four ethnic groups: Asian, Indian, White and Black.

The rest of the chapter is organized as follows. In section 2.5.3, previous studies related to ethnicity classification are presented. Section 5.2 discusses materials and methods employed in ethnic classification. Section 5.3 describes feature vectors computation of facial components using the Gabor Filter Transformation (GWT). Section discusses the classifiers used to explore the effectiveness of the proposed model. Experimental results are presented in section 5.5 and a conclusion is presented in section 5.6.

5.2 Materials and Methods

5.2.1 Gabor Filters

Over the past two decades Gabor filters have been receiving significant interest. Gabor filters have been successfully used in various computer vision applications. It has been specially used for texture segmentation due to its appealing simplicity and ability to localize joint spatial frequency. A number of authors have applied Gabor filters to various image analysis applications, such as texture analysis [222], face recognition [223], vehicle detection [224], iris recognition [225], and handwritten number recognition [226].

The 2D Gabor filter can be represented as a complex sinusoidal signal modulated by a Gaussian kernel function [227] defined as

$$g(x,y) = s(x,y)w_r(x,y) \tag{5.1}$$

where $s(x,y)$ is a complex sinusoidal, known as the carrier and defined as:

$$s(x,y) = e^{j(2\pi(\mu_0 x + \nu_0 y + p))} \tag{5.2}$$

with parameters μ_0 and ν_0 defining the spatial frequency of the sinusoidal, in cartesian coordinates. This spatial frequency can also be presented in polar coordinates with magnitude F_0 and direction θ_0 as

$$F_0 = \sqrt{\mu_0^2 + \nu_0^2} \tag{5.3}$$

$$\theta_0 = \tan^- 1 \begin{pmatrix} \mu_0 \\ \nu_0 \end{pmatrix} \tag{5.4}$$

we then have the complex sinusoid rewritten as

$$s(x,y) = e^{j(2\pi F_0((x\cos\theta_0 + y\sin\theta_0) + p))} \tag{5.5}$$

The function $w_r(x,y)$ is a 2-D Gaussian shaped function, known as the envelope which is defined as

$$w_r(x,y) = K^{-\pi(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2)} \tag{5.6}$$

where (x_0, y_0) is the peak function, a, b are scaling parameters of the Guassian, and the $_r$ represents a rotation operation.

$$(x-x_o)_r = (x-x_0)\cos\theta + (y-y_0)\sin\theta \tag{5.7}$$

and

$$(y-y_o)_r = (x-x_0)\sin\theta + (y-y_0)\cos\theta \tag{5.8}$$

Considering $x_0 = y_o = 0$ and $a = b = \sigma$ then, 2-D Fourier transform of the Gabor function will be

$$g(x,y) = Ke^{(-\pi a^2 (x-x_0)_r^2) + (y-y_0)_r^2)} e^{j(2\pi F_0 (x\cos\theta_0 + y\sin\theta_0)) + P)} \tag{5.9}$$

where K is the magnitude of the Gaussian envelope.

θ is the Rotation angle of the Gaussian envelope.

F is Spatial frequencies of the sinusoidal carrier in polar coordinates or inversely, the wavelength. It can also be expressed in cartesian coordinates as (μ_0, ν_0).

P is the Phase of the sinusoidal carrier.

σ represents the standard deviation of the Gaussian envelope.

Wavelets with a large wavelength will react to gradual changes in image intensity. Those with short wavelengths will react to sharp edges and bars. As per the Nyquist sampling theory, a signal containing frequencies higher than half of the sampling frequency cannot be reconstructed totally. Thus, the upper limit frequency for a 2D image is 0.5 cycles/pixel, whereas the low limit is 0. However, a valuable frequency band of face images is much smaller. It is observed that useful information is mainly contained in the low frequency domain when tested on sample face images in the frequency domain [228]. Hence, $F_{max} = 0.49$ is selected for feature extraction.

In our experiments five different frequencies and eight orientations were employed to extract the texture feature from facial components. Filter's parameter selected are , $\sigma = 1$, P=0, K=1,

$F \in \{0.13, 0.22, 0.31, 0.40, 0.49\}$ and $\theta \in \{0, \pi/8, 2\pi/8, 3\pi/8, 4\pi/8, 5\pi/8, 6\pi/8, 7\pi/8\}$. The Gabor filter bank composed of 40 channels was then created. The created filter set is then applied to the input facial image, by convolving the face image with each Gabor filter from this set. The resulting Gabor responses were then combined into a feature vector. The process is described in the subsequent section.

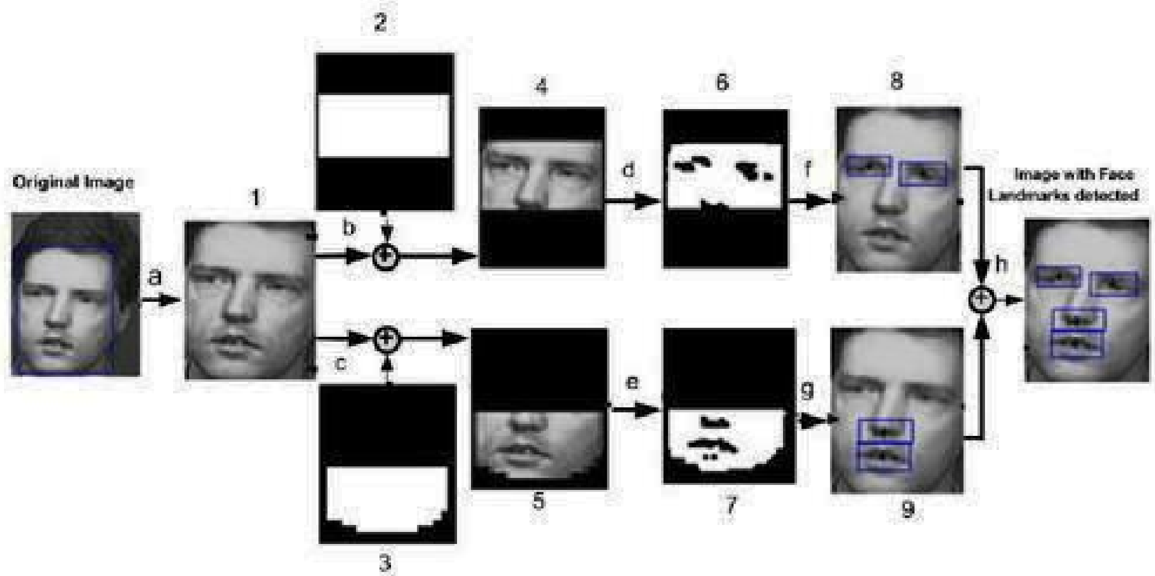


Figure 5.1: Face Landmarks Detection Process: letters (a-h) represent transformations of face region, and numbers (1-9) represent the results of the transformations. 1 is the region of interest of the face extracted; 4 and 5 are the upper and the lower part of the face which was extracted using masks 2 and 3 respectively; 6 and 7 are the binarized version of 4 and 5 using processes d and e respectively; f and g extract the eyes, the mouth and the nose from 6 and 7 and produce 8 and 9; h merges the results from 8 and 9 which represents a face with all landmarks detected.

5.3 Feature Computation of Facial Components using GWT

This section describes how features were extracted from the face image. The face feature extraction is initialized by the detection of components (eye, mouth, nose) using the algorithm described in Chapter 3. Figure 5.1 summarizes the process of landmark detection. Gabor features used for classification are extracted using GWT on each facial component. Gabor filter with five different frequencies and eight orientations are used in our experiments. For each component, the mean value for five frequencies and eight orientations were computed and stored in feature vectors $(\mu_{1\_Le} \ldots \mu_{40\_Le})$ for the left eye, $(\mu_{1\_Re} \ldots \mu_{40\_Re})$ for the right eye, $(\mu_{1\_No} \ldots \mu_{40\_No})$ for the nose, and $(\mu_{1\_Mo} \ldots \mu_{40\_Mo})$ for the mouth. Similarly, for each component, the standard deviation is computed with five different frequencies and eight orientations and stored in different feature vectors, $(\sigma_{1\_Le} \ldots \sigma_{40\_Le})$ for the left eye, $(\sigma_{1\_Re} \ldots \sigma_{40\_Re})$ for the right eye, $(\sigma_{1\_No} \ldots \sigma_{40\_No})$ for the nose, and $(\sigma_{1\_Mo} \ldots \sigma_{40\_Mo})$ for the mouth. In total, there will be 160 means and 160 standard deviations, since there are 4 facial components and each has 40 channels. For each face image, the means vector is then represented as

$$\mu_{face} = (\mu_{1\_Le} \ldots \mu_{40\_Le}, \mu_{1\_Re} \ldots \mu_{40\_Re}, \mu_{1\_No} \ldots \mu_{40\_No}, \mu_{1\_Mo} \ldots \mu_{40\_Mo}). \tag{5.10}$$

Similarly, the standard deviations vector is represented as

$$\sigma_{face} = (\sigma_{1\_Le} \ldots \sigma_{40\_Le}, \sigma_{1\_Re} \ldots \sigma_{40\_Re}, \sigma_{1\_No} \ldots \sigma_{40\_No}, \sigma_{1\_Mo} \ldots \sigma_{40\_Mo}). \tag{5.11}$$

A face was then represented by a feature vector of 320 components as

$$V_{face} = (\mu_{face}, \sigma_{face}) \tag{5.12}$$

It is also important to note that, each face component was characterized by a vector of 80 features (40 means and 40 standard deviations) as

$$V_{LeftEye} = (\mu_{1\_Le} \ldots \mu_{40\_Le}, \sigma_{1\_Le} \ldots \sigma_{40\_Le}). \tag{5.13}$$

$$V_{RightEye} = (\mu_{1\_Re} \ldots \mu_{40\_Re}, \sigma_{1\_Re} \ldots \sigma_{40\_Re}). \tag{5.14}$$

$$V_{Nose} = (\mu_{1\_No} \ldots \mu_{40\_No}, \sigma_{1\_No} \ldots \sigma_{40\_No}). \tag{5.15}$$

$$V_{Mouth} = (\mu_{1\_Mo} \ldots \mu_{40\_Mo}, \sigma_{1\_Mo} \ldots \sigma_{40\_Mo}). \tag{5.16}$$

Figure 5.2 shows the framework of the Gabor Feature extraction process.

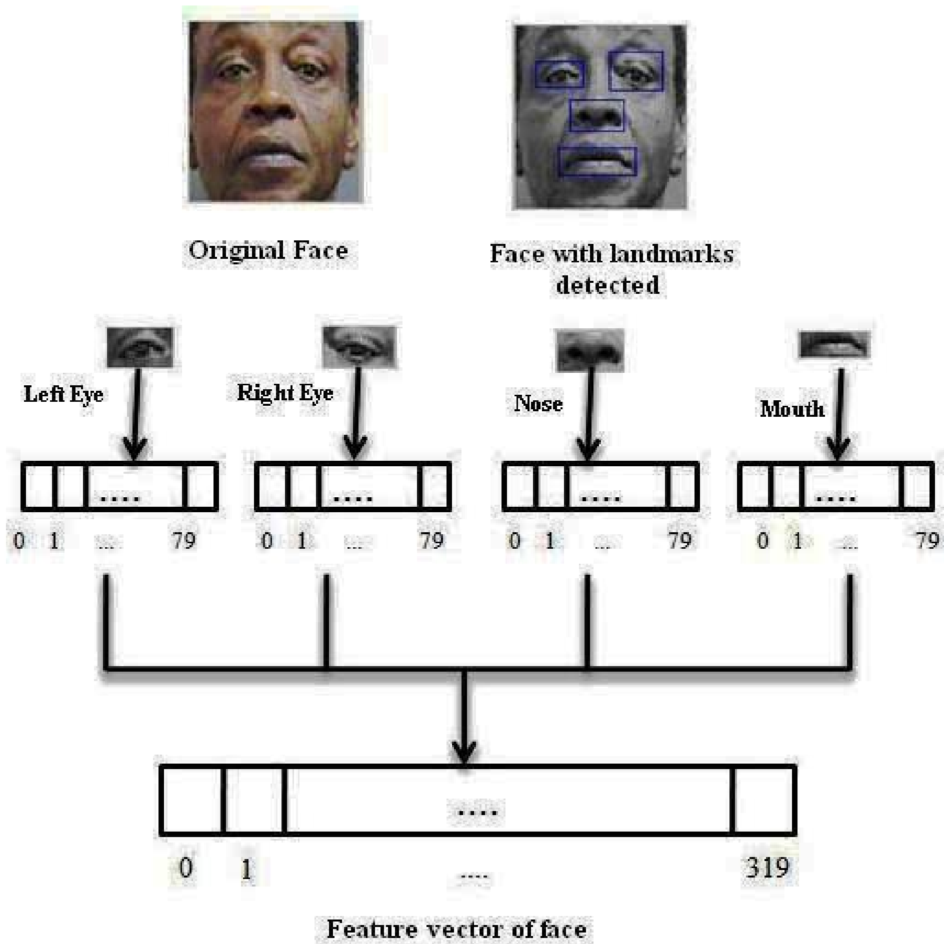


Figure 5.2: Framework of Gabor Feature extraction process

5.4 Ethnicity Classification

Four classifiers, namely K-means, NB, MLP and SVM were used to classify features extracted from the faces.

K-means Clustering

The K-means clustering algorithm is explained in detail in section 2.6.1. It classifies input data points into multiple classes based on their inherent distance from each other (typically the Euclidian distance between the cluster centres and the candidate vector). In our experiment, a K-means algorithm with $k = 2$ for classifying face images into either Asian or non-Asian ethnic group was used. A further k value of 3 was used for Asian, White and Black groups classification, and $k = 4$ for Asian, Indian, White and Black ethnic groups.

Naive Bayesian

The NB algorithm was used to classify the images as explained in section 2.6.2. Gaussian distribution is used to calculate a pdf for all race types. Once the pdf is calculated the images are classified into different ethnic groups.

Multilayer Perceptron (MLP)

Once the feature vector is created using Gabor wavelet transformation, these vectors then act as input nodes to MLP. In our method, MLP model (a) for two ethnic groups Asian and non-Asian, has a feature vector with an input layer consisting of 320 attributes (nodes), a hidden layer with 161 nodes, and an output layer with two nodes (Asian and non-Asian). For a single component the MLP contained an input layer with 80 nodes, a hidden layer with 41 nodes and an output layer with two nodes. (b) For three ethnic groups the input and hidden layer will be identical to those used for two ethnic groups and the output layer contains three nodes (Asian, White and Black). (c) For four ethnic groups the MLP has a hidden layer with 162 nodes, an input layer with 320 nodes, and the output layer will have four nodes (Asian, Indian, White and Black). For a single component the MLP contained an input layer with 80 nodes, a hidden layer with 42 nodes and an output layer with four nodes.

Support Vector Machine

The SVM algorithm is used to classify images into appropriate ethnic groups. It is widely acknowledged that a key factor in a SVM's performance is the choice of the kernel. Therefore, experiments are performed with various kernels (polynomial, Gaussian, RBF and PUK) in order to compare their performance. In our case, PUK provides good results when compared to other kernels. Therefore, we used a Pearson VII function based universal kernel which is defined in section 2.16. The parameters used in the experiments for SVM include the complexity parameter C which is taken as 1 (C=1) which controls the trade-off between margin maximization and error minimization. A PUK kernel with parameters $\omega = 1$ and $\sigma = 1$ is used for ethnicity classification. The results of the techniques used for ethnicity classification are presented in Tables 5.1 to 5.3.

5.5 Experimental Results

The data-set used for experimentation is the combination of 3 different face databases. A total of 511 images with different facial expressions, illumination and orientations were incorporated into the study. The Asian group comprises of images from an Asian face database [199] and an Indian face database [201]. A total of 52 images from an Indian face database and 199 images from an Asian face database were used for experimentation. The non-Asian group is composed of images from the MORPH database [200]. A total of 260 images were used from this database.

The entire available data-set is randomly partitioned into a training set (in-sample) and a test set (out-of-sample). Cross-validation is utilized to precisely depict the predictive performance of the machine learning techniques. Cross-validation is a re-sampling technique which makes use of multiple random training and test sub-samples. A well accepted method is N-fold cross validation, in which you randomize the data-set and create N equal size partitions. In our study, 10-fold cross validation is used to assess or compare the performance of the proposed models. The available sample is randomly divided into 10 sub-samples of size n/10. For testing the model, a single sub-sample is held as the validation data out of the 10 sub-samples and rest of the nine sub-samples are used as training data. The cross-validation process is repeated 10 times (the folds), with each of the sub-samples used exactly once as the validation data. The outcomes from the folds then can be averaged to deliver a single estimate. The advantage of cross-validation is that all observations have been used for training and validation. Each observation is used exactly once for validation. The cross-validation analysis will yield valuable insights on the reliability of the machine learning technique with respect to sampling variation.

Machine learning algorithms such as the K-means clustering algorithm, NB, MLP and SVM are used in our experiments. The feature vector that is created using Gabor filters are fed into each of the chosen classifiers. Images are classified into the two ethnic groups of Asian and non-Asian. The images are further classified into the three ethnic groups of Asian, Black and White and then into the four ethnic groups of Asian, Black, Indian, and White, using different facial components.

In our study, four facial components (the left eye, the right eye, the nose and the mouth) were used for ethnicity classification. Table 5.1 shows ethnicity classification results for each of the separate facial component and for each classifier. All four classifiers presented similar results. Facial components as the nose and the mouth shows better results when compared to the left and the right eye. Table 5.2 shows the ethnicity classification for three different ethnic groups. In this case, a MLP algorithm delivered best results when compared with three other classifiers and the SVM showed the poorest results. In this classification, again the nose and the mouth gave the best results when compared to the left and the right eye. A combination of all four components gave better results as opposed to single components. Images are further classified into the four ethnic groups of Asian, Indian, White and Black. In this case as well, MLP gave the best results when compared to the other three classifiers. The SVM delivered the poorest results. In this classification the nose and the mouth produced better results as opposed to the left and the right eye. All four components were combined to improve the performance. The results are presented in Table 5.3. Figure 5.3 shows a graph of ethnicity classification using different facial components and different ML techniques.

Table 5.1: Ethnicity Classification rate (%) using different classifiers for Asian and non-Asian

	Classification rate		
Face Component	**Asian**	**non-Asian**	**Combined**
K-means Clustering			
Left Eye	98.8	95.76	97.26
Right Eye	100	97.69	**98.82**
Nose	99.6	97.69	**98.63**
Mouth	99.6	97.69	**98.63**
All Components	**99.6**	**99.23**	**99.41**
Naive Bayesian Classifier			
Left Eye	96.01	95.76	95.89
Right Eye	97.21	97.3	97.26
Nose	98	98.46	**98.23**
Mouth	99.6	98.46	**99.02**
All Components	**99.6**	**99.23**	**99.41**
Multilayer Perceptron Classifier			
Left Eye	96.81	95.53	96.67
Right Eye	99.2	97.69	**98.43**
Nose	98.4	97.69	98.04
Mouth	99.2	98.46	**98.82**
All Components	**99.6**	**99.23**	**99.41**
SVM Classifier			
Left Eye	98.4	97.69	96.28
Right Eye	99.6	97.3	**98.43**
Nose	98.4	97.69	98.04
Mouth	99.6	97.69	**98.63**
All Components	**99.6**	**98.46**	**99.02**

Table 5.2: Ethnicity Classification rate (%) using different classifiers for Asian, White and Black

Face Component	**Classification rate**			
	Asian	**White**	**Black**	**Combined**
K-means Clustering				
Left Eye	98	80.76	46.92	80.62
Right Eye	100	80	55.38	83.56
Nose	**99.6**	**83.84**	**83.07**	**91.38**
Mouth	99.6	76.92	63.84	**84.73**
All Components	99.6	80.76	80	**89.82**
Naive Bayesian Classifier				
Left Eye	96.42	82.3	40.76	78.86
Right Eye	98.82	83.07	46.92	79.84
Nose	96.82	76.92	83.07	**88.54**
Mouth	99.6	73.84	63.84	**83.95**
All Components	**99.6**	**80**	**75.38**	**88.45**
Multilayer Perceptron Classifier				
Left Eye	97.21	83.07	48.46	81.21
Right Eye	99.2	86.92	50.76	83.75
Nose	98.4	85.38	76.92	**89.62**
Mouth	99.2	88.46	60	**86.49**
All Components	**99.6**	**81.53**	**80.76**	**90.21**
SVM Classifier				
Left Eye	98.4	87.69	34.61	79.45
Right Eye	99.6	89.23	41.53	82.19
Nose	**98.4**	**92.3**	**72.3**	**90.21**
Mouth	99.6	77.69	60.76	**84.14**
All Components	99.6	90.76	54.61	85.91

Table 5.3: Ethnicity Classification rate (%) using different classifiers for Asian, Indian, White and Black

Face Component	Classification n rate				
	Asian	**Indian**	**White**	**Black**	**Combined**
			K-means Clustering		
Left Eye	100	48.07	80	44.61	75.53
Right Eye	100	80.76	80	56.15	**81.8**
Nose	99.49	42.3	85.38	83.84	**86.1**
Mouth	98.99	50	75.38	64.61	79.25
All Components	**99.49**	**73.07**	**82.3**	**80.76**	**87.67**
			Naive Bayesian Classifier		
Left Eye	95.97	63.46	83.07	43.07	75.92
Right Eye	97.98	86.53	84.61	48.46	**80.82**
Nose	91.95	63.46	77.69	84.61	**83.56**
Mouth	96.98	69.23	74.61	63.07	79.84
All Components	**95.97**	**78.84**	**81.53**	**73.84**	**84.93**
			Multilayer Perceptron Classifier		
Left Eye	**98.99**	63.46	89.23	36.92	77.1
Right Eye	98.49	92.3	90.76	38.46	80.62
Nose	96.48	50	86.15	86.41	**86.1**
Mouth	97.48	65.38	89.23	58.46	**82.19**
All Components	**98.99**	**86.53**	**79.23**	**79.23**	**87.67**
			SVM Classifier		
Left Eye	99.44	50	88.46	39.23	76.32
Right Eye	99.49	82.69	90	41.53	**80.62**
Nose	98.49	0.038	92.3	72.3	**80.62**
Mouth	99.49	21.15	80.76	60	76.71
All Components	**99.49**	**69.23**	**92.3**	**55.38**	**83.36**

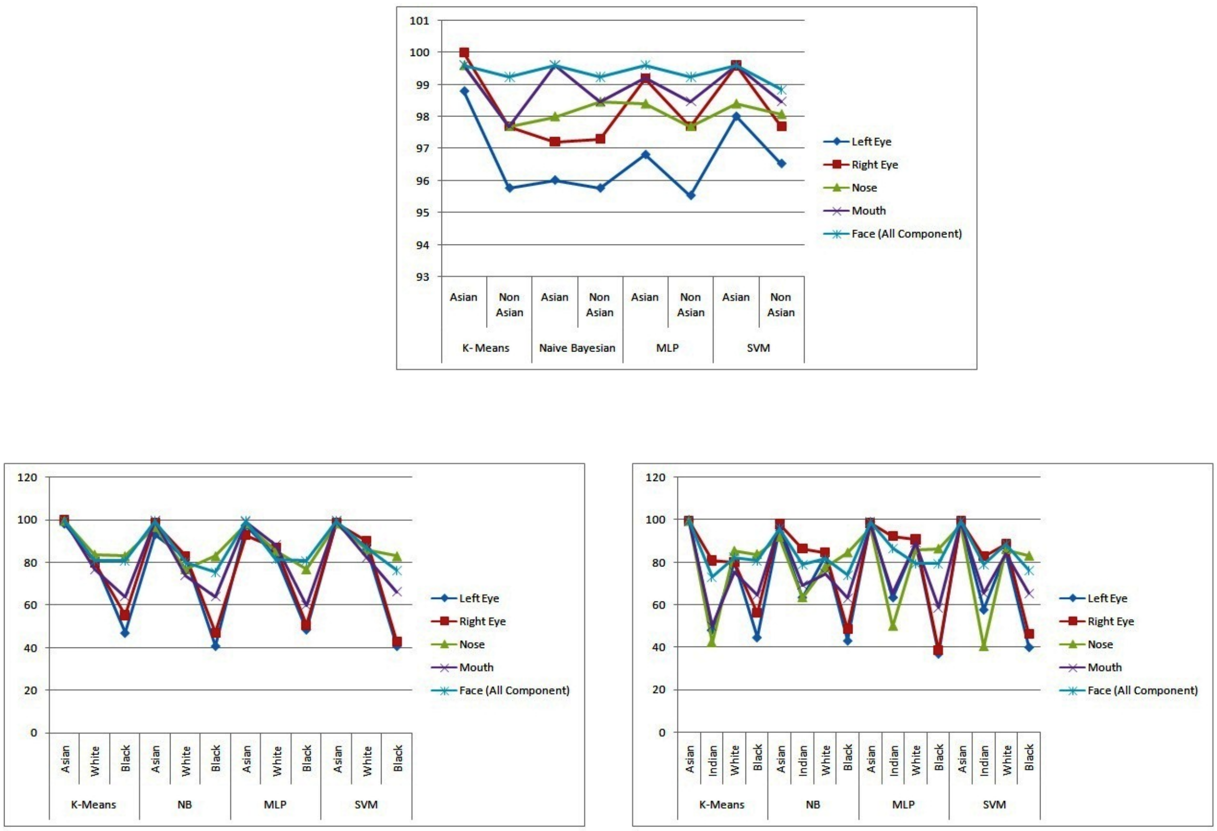


Figure 5.3: Graph of ethnicity classification using different facial components and different machine learning algorithms for different ethnic groups.

5.6 Conclusion

A method for automatic ethnicity classification based on the texture analysis of facial components is presented. The face images are considered in a complex context. The complexity of the images is presented in terms of different illumination conditions, facial expressions and orientations. The method has been tested by using three configurations of classes: (1) two different ethnic groups: Asian and non-Asian, (2) three ethnic groups: Asian, White and Black, and (3) four ethnic groups: Asian, Indian, White and Black. This was done by using different facial components such as the left and the right eye, the nose and the mouth and by combining all these components. The texture information of facial components is captured using a bank of multi-channel 2D Gabor filters. The K-means clustering algorithm, NB, MLP and SVM are used to classify images into the correct ethnic classes.

From the experiments it is clear that the single facial component contributes greatly in ethnicity classification. Instead of the whole face image, important facial components are used for the classification. The proposed method has achieved a high degree of accuracy, that is for Asian and non-Asian being 99.60%, for Asian, White and Black being 90.21%, and for Asian, Indian, White and Black being 87.67%. From our experiments, we observed that mostly the nose and the mouth gave superior results as compared to the left and the right eye. This prompted a hypothesis that "the mouth and the nose" could be a strong element that characterizes the face among various races. However, when all four facial components are considered in combination, better results are attained as opposed to a single component. The MLP delivers good results as compared to other ML algorithms used whereas the SVM offers the poorest results when compared to the other three classifiers. In future, we would like to extend our work by including additional features such as colour or geometric features with the intention of further improving the accuracy of ethnic classification.

Chapter 6

Ethnicity Classification using fusion of Colour and Texture Features

6.1 Introduction

The determination of ethnicity of an individual can be very useful in face recognition and identification systems. Many research results on automatic face detection and face recognition have been published. Very few attempts however have been made to perform human ethnicity classification. Research in identifying human characteristics for example ethnicity, age and gender through visual cues is progressing at a very quick pace.

There are three levels of fusion in biometrics. Ross and Jain in [100] described the three possible levels of fusion. These levels include:

1. Fusion at the Feature Extraction Level: Here, different features are extracted from the data acquired from each sensor. The features computed from different biometric indicators are

then combined into a single vector.

2. Fusion at the Matching Score Level: At that level, each system provides a similarity score indicating the distance between the input feature vector and the template vector. The similarity scores are combined to ascertain the accuracy of the claimed identity.

3. Fusion at the Decision Level: Here, each system provides a similarity score per biometrics trait indicating the distance between the input feature vector and the template vector, which is individually classified into the two classes: accept or reject. A majority vote scheme is used to make the final decision.

In this chapter, fusion at feature extraction level is employed. Two types of features are extracted from the face images in order to identify the ethnicity of a person and to improve the accuracy. The fusions of these features are used for ethnicity classification. Since colour and texture features have shown to provide good results, these features are combined to improve the accuracy of ethnic classification.

The colour and texture features are combined and a single feature vector is created for the face image. After a fusion of feature vectors, a decision mechanism is employed to classify the image into an ethnic group. The MLP classification algorithm is then used to classify images into proper ethnic groups.

The rest of the chapter is organized as follows. Section 6.2 explains the methodology used for ethnicity classification. Section 6.3 describes the ethnicity classification process. Section 6.4 provides the experimental results and is followed by section 6.5 which provides concluding remarks.

6.2 Methodology

The proposed methodology describes a method of classifying the images into their ethnic groups. Figure 6.1 presents a flow diagram of the methodology used in ethnicity classification. For ethnicity classification two different features (colour and Gabor features) are used and their fusion is used for ethnicity classification. First, images are classified using colour features and then using Gabor features. Second is the fusion of the two features which is followed by a comparison of the results. The feature extraction process is explained in detail in the subsection below.

First colour features are extracted for input face images as explained in section 4.2, then texture features are extracted from the input face image as explained in section 5.3. Figure 6.2 shows the feature extraction process using (a) Gabor filters, (b) colour features and (c) a combined feature vector.

6.3 Ethnicity Classification

In this section, a method for identifying the ethnicity of a person from an input face image is presented. The MLP algorithm is used for ethnicity classification. The ethnicity classification is carried out as follows:

1. Using the Colour Feature
2. Using the Gabor Feature
3. Using the Fusion of Colour and Gabor Features

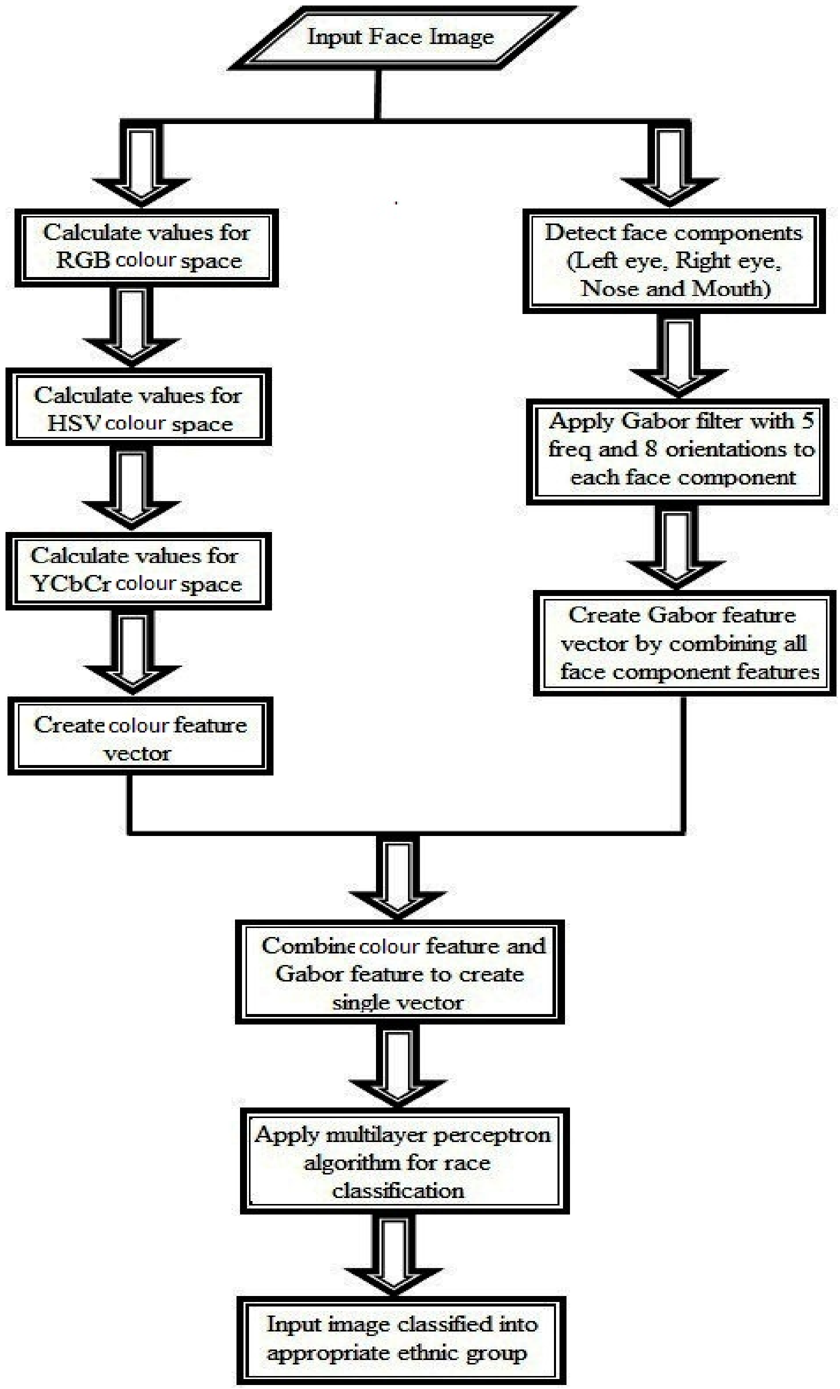


Figure 6.1: Flow diagram of ethnicity classification.

Multilayer Perceptron

Multilayer perceptron neural networks are trained to classify the input images according to their ethnicity. Once the feature vectors are created using a combination of colour features and the Ga-

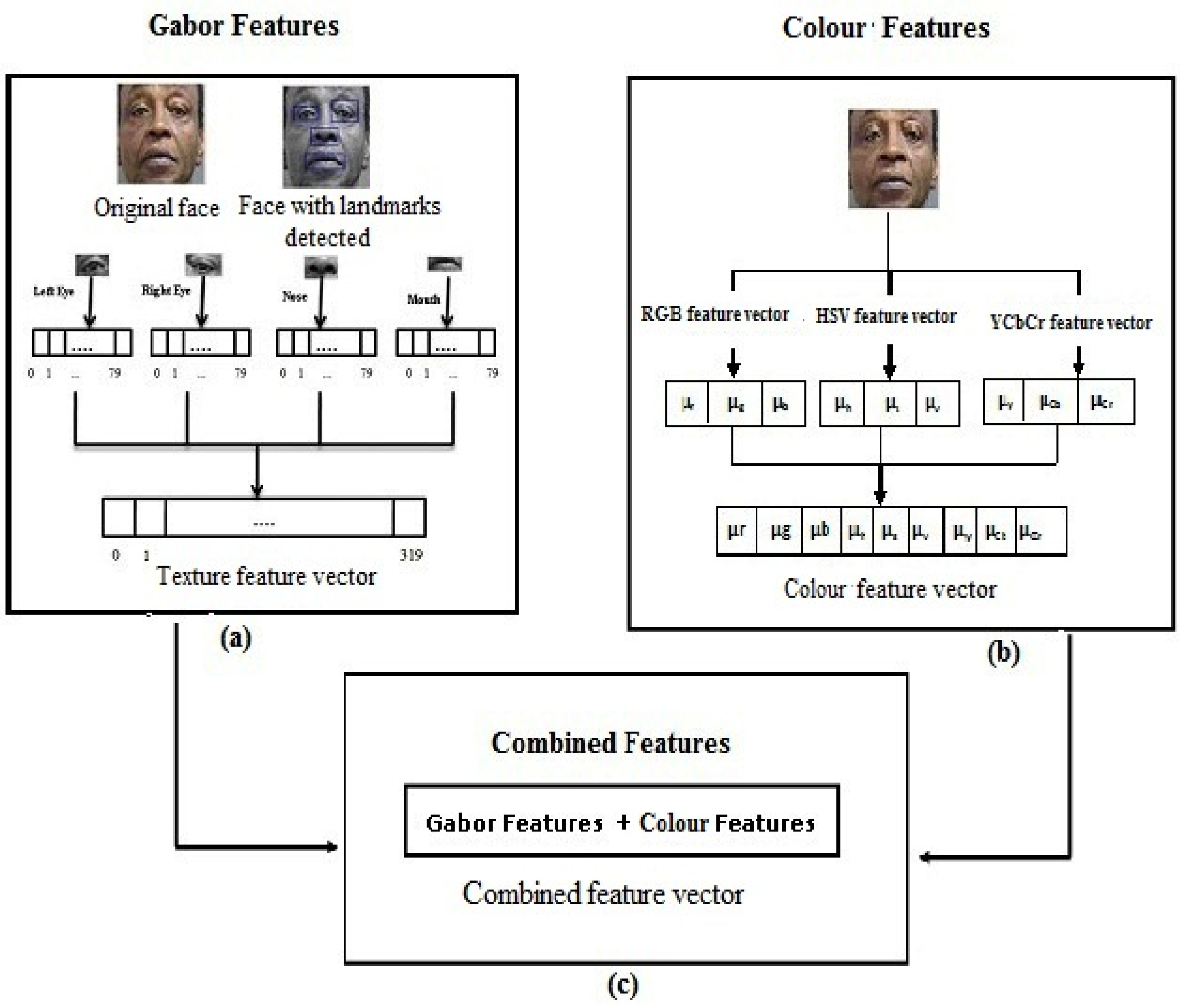


Figure 6.2: Feature fusion.

bor features, they will be employed for ethnicity classification. The MLP is then used to classify the images into proper ethnic groups.

We consider ethnicity classification into three classes:

- Asian-non-Asian
- Asian-White-Black
- Asian-Indian-White-Black

Figure 6.3 shows the MLP model (a) for two ethnic groups, Asian and non-Asian. This model contains an input layer with 11 attributes, a hidden layer with six nodes, and an output layer with two nodes. The nodes in the hidden layer are calculated using equation 2.12. The 11 attributes comprise nine colour components being red, green, blue, hue, saturation, value, yellow, Cb, Cr, and texture features mean and standard deviations. The output layer with two nodes is Asian and non-Asian. Figure 6.3 (b) shows the MLP architecture for three ethnic groups namely, Asian, White and Black. This model contains an input layer with 11 attributes, a hidden layer with seven nodes, and an output layer with three nodes, for Asian, White and Black. Figure 6.3 (c) shows the MLP architecture of four ethnic groups, Asian, Indian, White, and Black. This model contains an input layer with 11 attributes, a hidden layer with seven nodes, and an output layer with four nodes for Asian, Indian, White and Black groups.

6.4 Experimental Results

We used two different data-sets for the evaluation of our method. The first data-set is a combination of the Asian face database [199], the Indian face database [201] and the MORPH database [200]. The second data-set is a combination of the MUCT face database [229], an Asian face database and an Indian face database. In the first data-set a total of 52 images from an Indian face database, a total 199 images from an Asian face database and a total of 260 images from the MORPH database was used. A combined total of 511 images from these databases were used. In the second data-set a total of 150 images of the black ethnic group and 150 of the white ethnic group were used from the MUCT face database. A total of 300 images were used from the MUCT face database, 175 from an Asian face database and 52 images from an Indian face database. A combined total of 512 images were used from these databases.

The images are first classified into Asian and non-Asian groups using colour features, Ga-

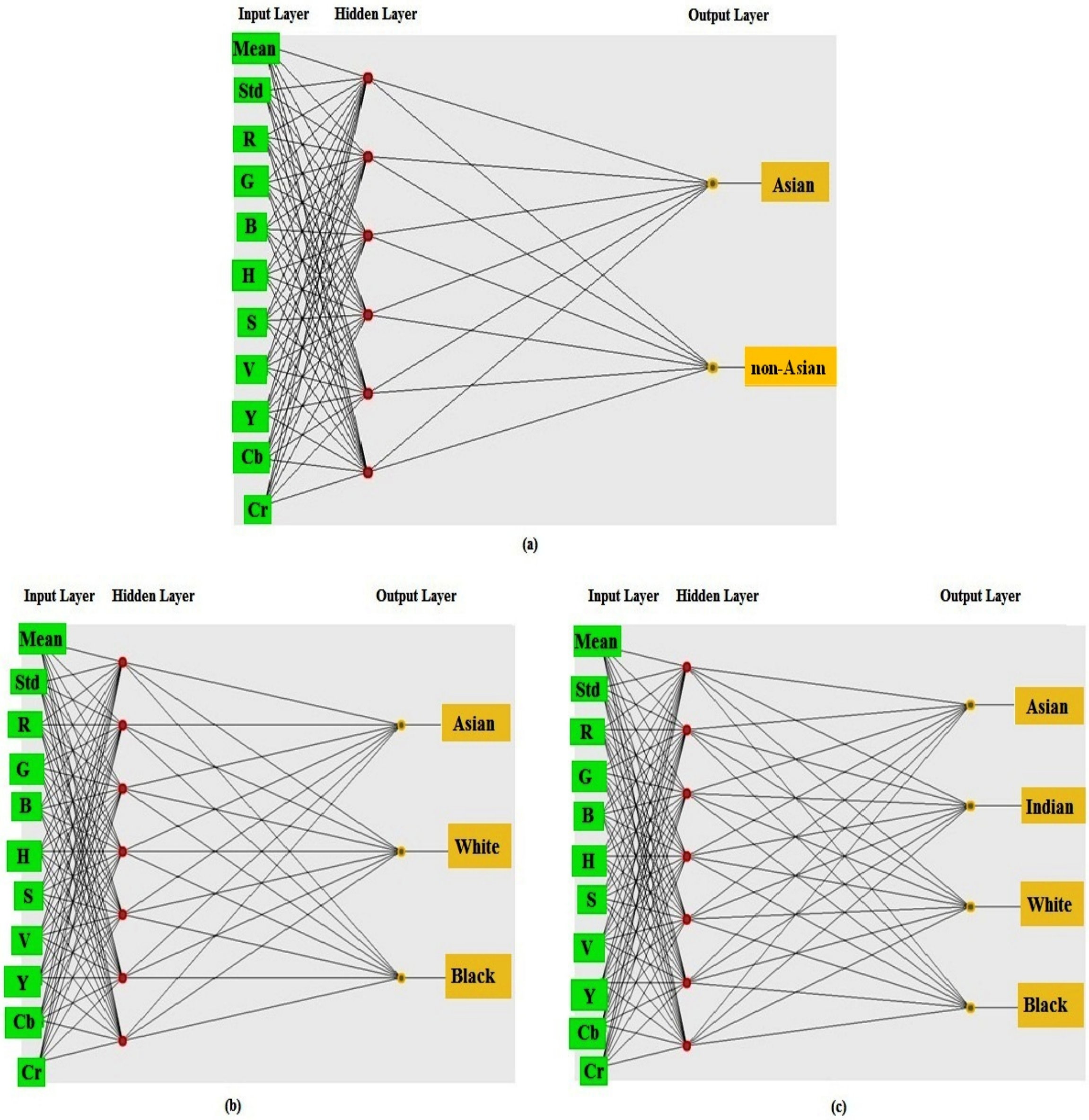


Figure 6.3: Multilayer perceptron architecture for (a) Asian and non-Asian (b) Asian, White and Black (c) Asian, Indian, White and Black.

bor features and then the fusion of both the features. The results of the classification are shown in Table 6.1 for two different data-sets. From Table 6.1 it is clear that the colour and combined features provide almost the same classification results whereas the Gabor features provide a little less classification rate. Secondly, images are classified into Asian, White and Black ethnic groups. The experimental results are shown in Table 6.2. In this case, the colour feature outperforms that of the combined features for the first data-set but the combined features reveal the highest results for the second data-set. Finally, the images are classified into the four ethnic groups of Asian, Indian, White and Black. The results of the classification are shown in Table 6.3. This case as well the combined features gave better results when compared to the use of a single colour or Gabor features.

Based on the results of experiments that took place using two different data-sets, the percentage of Asian and non-Asian ethnic classification using colour features is 99.60%, and 99.78%; using Gabor features is 99.41%, and 97.26%; and using the fusion of both is 99.60%, and 99.78% for each data-set respectively. For three classes namely Asian, White and Black, the percentage outcomes were 96.86% , 90.21%, and 93.58% respectively using colour, Gabor and feature fusion for the first data-set and 97.05%, 91.78%, and 98.31% respectively using colour, Gabor and feature fusion for the second data-set. Lastly, for four classes namely Asian, Indian, White and Black the percentage outcomes were 94.91% , 87.61%, and 95.49% respectively, using colour, Gabor and feature fusion for the first data-set; and 95.06%, 81.4%, and 97.15% respectively using colour, Gabor and feature fusion for the second data-set. This method achieved a high degree of accuracy for Asian, 98.49%, for Indian, 91.53%, for White 98.07% and for Black 93.84% using the first data-set; and 98.28% for Asian, 98.66% for Indian, 94.66% for White, and 94.23% for Black using the second data-set.

Figure 6.4 (a) shows a comparison graph of different features (colour, Gabor, and combined) for the two race classifications of Asian and non-Asian; Figure 6.4 (b) shows a comparison graph for Asian, White and Black; Figure 6.4 (c) shows a comparison graph for Asian, Indian, White, and Black. From these graphs, it is evident that colour features provide better results when compared to Gabor filter features. The combination of these features provides better results as opposed to single features.

We have classified our first data-set into five subsets to assess how the classification rate varies according to the number of images. Table 6.4, Table 6.5 and Table 6.6 show different datasets and their results for Asian, non-Asian; for Asian, White and Black; and for Asian, Indian, White and Black. Figure 6.5 shows a graphical representation of how ethnicity classification varies when the size of the dataset is altered.

Table 6.1: Ethnicity classification rate (%) into two classes using different features.

Features	MORPH+Asian+Indian face database			MUCT+Asian+Indian face database		
	Asian	non-Asian	Combined	Asian	non-Asian	Combined
Colour Features	100	99.23	99.60	99.42	100	99.78
Gabor Features	99.6	99.23	99.41	98.28	96.66	97.26
Combined Features	100	99.23	**99.60**	99.42	100	**99.78**

Table 6.2: Ethnicity classification rate (%) into three classes using different features.

Features	MORPH+Asian+Indian face database				MUCT+Asian+Indian face database			
	Asian	White	Black	Combined	Asian	White	Black	Combined
Colour Features	100	94.61	93.07	**96.86**	98.28	98	94	97.05
Gabor Features	99.6	81.53	80.76	90.21	97.14	88	89.33	91.78
Combined Features	100	89.23	91.53	95.10	98.28	97.33	98.66	**98.31**

Table 6.3: Ethnicity classification rate (%) into four classes using different features.

Features	MORPH+Asian+Indian face database					MUCT+Asian+Indian face database				
	Asian	White	Black	Indian	Combined	Asian	White	Black	Indian	Combined
Colour Features	96.48	93.84	96.09	92.30	94.91	98.28	98	91.33	84.61	95.06
Gabor Features	86.53	79.23	79.23	86.53	87.67	97.14	86.66	79.33	19.23	81.4
Combined Features	98.49	98.07	93.84	91.53	**95.49**	98.28	98.66	94.66	94.23	**97.15**

Table 6.4: Ethnicity classification rate (%) into two classes using combined features and different dataset sizes.

Dataset size	Classification Rate(%)		
	Asian	non-Asian	Combined
100	100	100	100
200	100	100	100
300	100	100	100
400	100	99.5	99.75
511	100	99.23	99.60

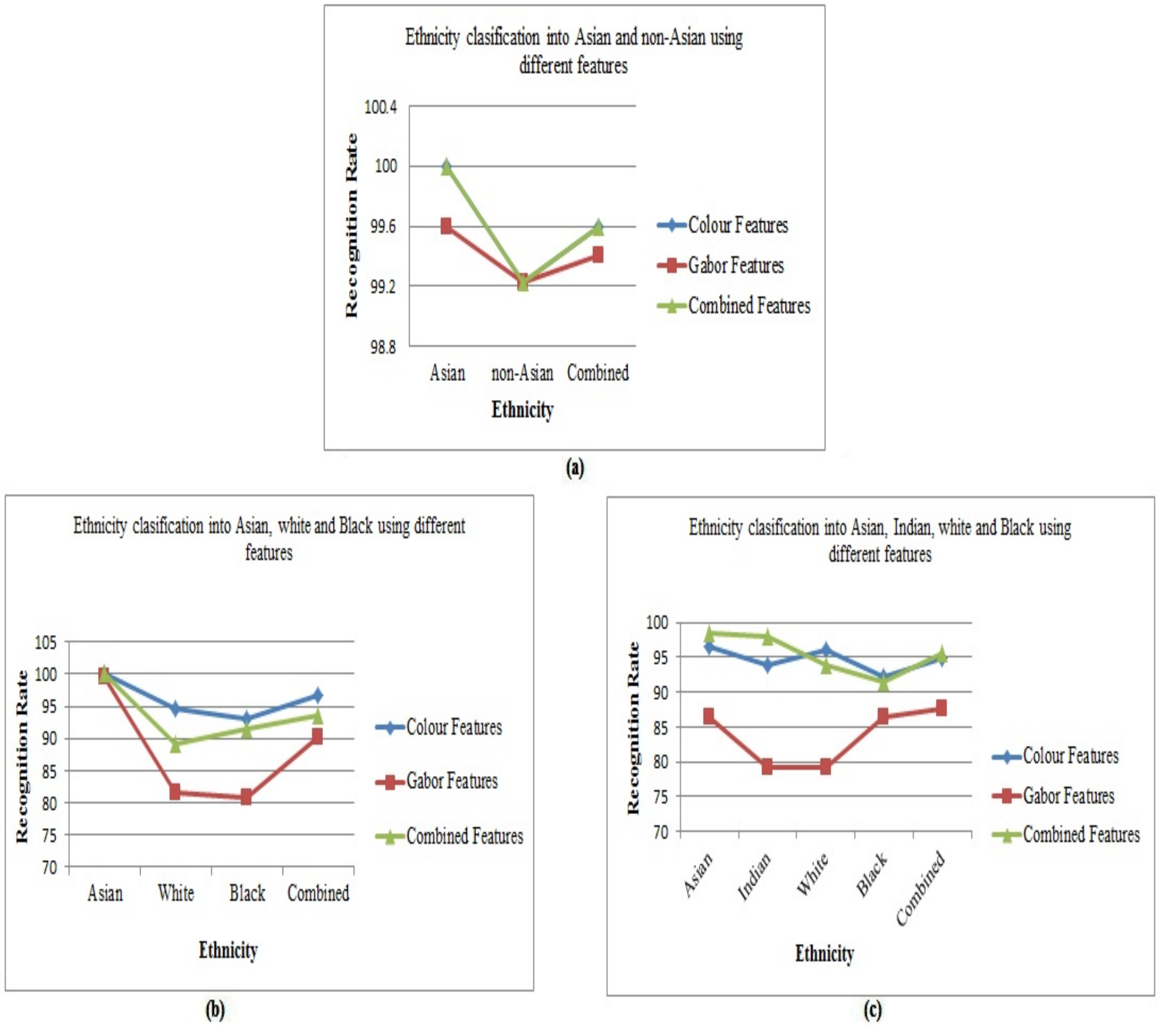


Figure 6.4: Graph of ethnicity classification using different features (a) Asian and non-Asian, (b) Asian, White and Black (c) Asian, Indian, White and Black.

Table 6.5: Ethnicity classification rate (%) into three classes using combined features and different dataset sizes.

Dataset size	Classification Rate (%)			
	Asian	White	Black	Combined
100	100	100	100	100
200	100	96.66	96.66	98
300	99.16	97.77	93.33	97
400	99	94	94	96.5
511	100	89.23	91.53	95.10

Table 6.6: Ethnicity classification rate (%) into four classes using combined features and different dataset sizes.

Dataset size	Classification Rate (%)				
	Asian	White	Black	Indian	Combined
100	100	100	96.66	100	99
200	100	96.36	96.36	100	98
300	100	94.11	97.64	100	97.66
400	99.28	90	96.36	97.5	95.75
511	98.49	91.53	93.84	98.07	95.49

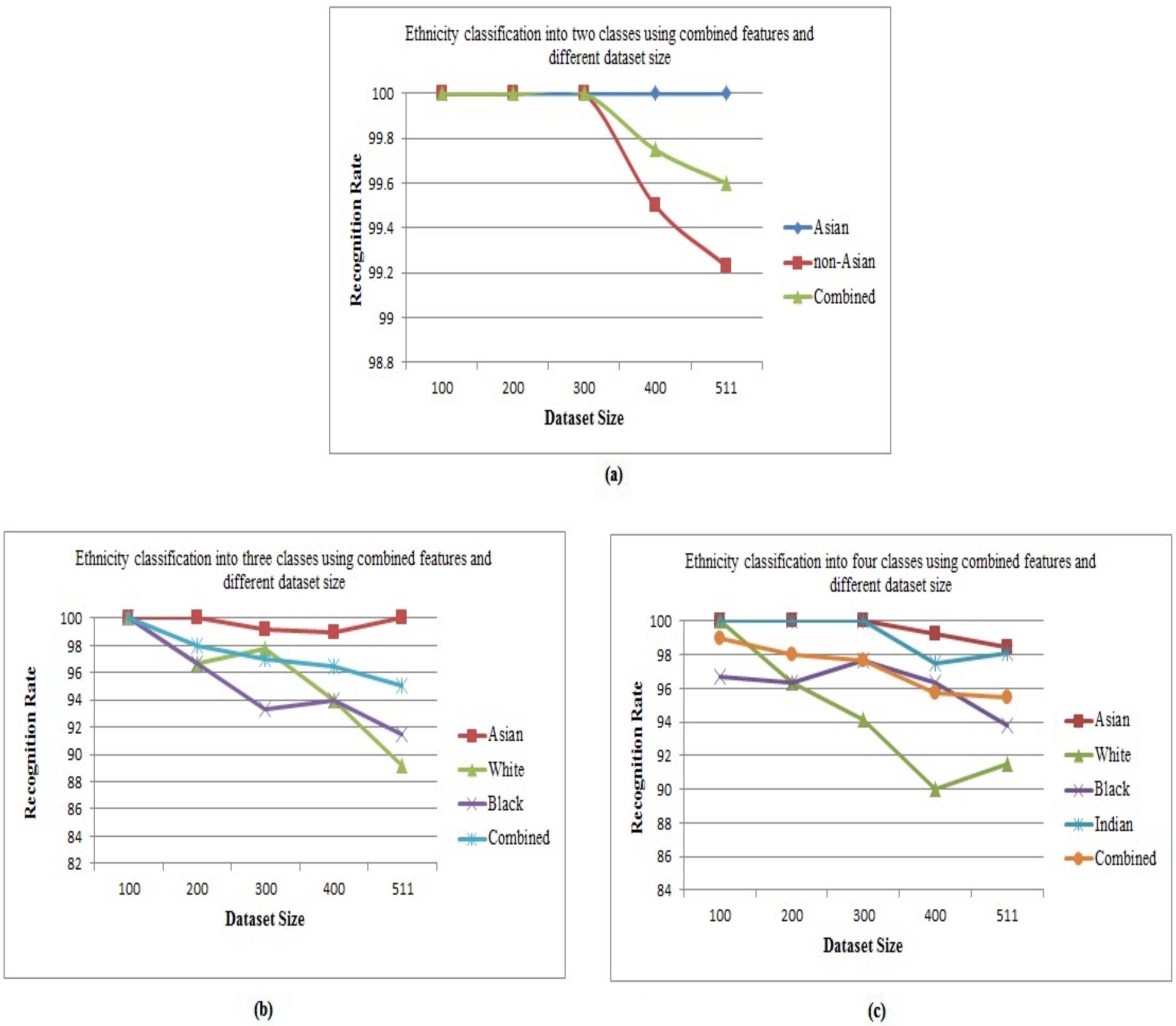


Figure 6.5: Graph of ethnicity classification using different dataset sizes for (a) Asian and non-Asian, (b) Asian, White and Black (c) Asian, Indian, White and Black.

6.5 Conclusion

Ethnicity classification of face images using colour, Gabor and a fusion of colour and Gabor features are presented in this chapter. Experiments were carried out on two different data-sets. The proposed method classifies the images into four different ethnic groups namely Asian, Indian, White and Black by using colour features, Gabor features and then using a combination of both. Interestingly, the proposed method of classifying the images by using Gabor wavelets, which were expected to outperform colour features, performed poorly on the ethnic classification task. Colour features provided better results as compared to Gabor features when compared individually.

This research shows that a combination of features provides better results when compared to the singular colour or Gabor features when classified into three and four ethnic groups. The feature fusion provides the same result as colour features when classified into two ethnic groups but provides better results when compared to the Gabor filter. To further improve accuracy, more features can be used for the ethnicity estimation.

Chapter 7

Conclusion and Future work

7.1 Summary of Work

In this thesis, a framework for face detection, component-based face features extraction and ethnicity classification using colour features, Gabor features and geometric features are presented. Firstly, a comprehensive literature survey of state-of-the-art of face detection methods, feature extraction methods, and different methods used for ethnicity classification was carried out. The challenges addressed, and the novel techniques designed and investigated include:

- An algorithm was presented using a segmentation approach which extracts the face region by using thresholding and connected component labelling. The segmentation algorithm first converts the gray-scale input image into a binary image, then locates the face in the binary image and puts a bounding box around that face. This algorithm is very useful for different face related applications like face feature extraction and face recognition.

- A simple image processing operation on facial images was applied to automatically and more accurately detect facial components. Segmentation and a connected component labelling algorithm were used to extract key facial components: the eyes, the mouth, and the nose.

The presented approach automatically detects facial components at a high accuracy since the face image was divided into two parts which simplifies the landmark detection process. The components are detected in different orientations and different facial expressions.

- Three different types of features were used and compared for ethnicity classification. The first features used were based on different colour spaces. The RGB, the HSV and the YcbCr colour space were used to classify the input face images into the different ethnic groups of Asian, White and Black by using a K-Means clustering algorithm and an NB algorithm. The best colour space for feature extraction was the RGB colour space. This algorithm is very simple yet provides promising results.

- Texture features were used for ethnic classification of input face images using Gabor filters. The face images were classified into four different ethnic groups, namely Asian, Indian, White and Black. The question considered in this work was which parts of the face hold reliable ethnicity information for machine-based classification. Texture features were computed for different facial components such as the left eye, the right eye, the nose and the mouth. Extensive experiments demonstrated that mostly the nose and the mouth provide good classification results when compared to the left and the right eyes. Different ML techniques, K-Means Clustering, NB classifier, MLP and SVM were used for ethnicity classification and the results were compared.

- A fusion of texture and colour features was used for ethnicity classification. Extensive experiments demonstrated that the fusion of features provides better results when compared to the single colour or Gabor features.

7.2 Comparative Study

A comparative study of the above proposed approaches is presented below. Table 7.1 shows a comparison of ethnicity classification into two race groups namely, Asian and non-Asian. In [11], they presented a LDA based scheme for two classes (Asian and non-Asian) for the task of ethnicity classification. They classified 132 Asian faces and 131 non-Asian faces and reported an accuracy of 96.3% . In [142] real time face detection and classification into gender and ethnicity was proposed. They extracted features using three types of rectangular filters. The images are classified into two classes being Asian and non-Asian using SVM and boosted classifier. Ou et al. used PCA for feature generation and ICAfor feature extraction and classified a frontal face image into Asian and non-Asian. They reported an accuracy rate of 82.5% [103]. Lyle et al. used periocular region images using gray-scale pixel intensities and periocular texture computed by LBP as features for gender and ethnicity classification and SVM as classifier [145] . They achieved a 91% accuracy for the ethnicity classification. Our method however achieved a very good accuracy rate of 99.60%, 99.41% and 99.60% using colour, Gabor and combined features respectively.

Table 7.2 shows a comparison of ethnicity classification into the three race groups of Asian, White and Black. Gutta et al. reported an accuracy rate of 92% for an ethnicity classification task into Caucasian, Asian, Oriental and African origins categories [12]. Xie et al. classified images into Asian, African-American and Caucasian ethnic groups. Their method achieves an accuracy of 95% for Asian, 97% for Caucasian and 97% for African-American [147]. Hosoi et al. used GWT and retina sampling to extract key facial features, and then used SVM for ethnicity classification. They classified the images into Asian, European and African and reported an overall accuracy of 94% [7]. The highest accuracy achieved in the literature for the three race groups are 96%. However, our method achieves an accuracy rate of 98.31%. We achieved 97.05% using colour features, 91.78% using Gabor features and 98.31% using combined features.

Table 7.3 shows a comparison of ethnicity classification into four race groups, namely Asian, White, Black and Indian. To our knowledge less literature is available for ethnicity classification into four groups. The one that is reported in the literature has limited classification results for Asian and Indian. Our method however achieves a 95.06% accuracy rate using colour features, 81.4% using Gabor features and 97.15% using combined features. The proposed approach based on colour, texture and combined features achieved better results for ethnicity classification.

Table 7.1: Ethnicity comparison in two groups Asian and non-Asian from literature.

Reference	**Features**	**Classification Rate**
[11]	range and pixel intensity	96.3%
[142]	Rectangular Features	22.6% error rate
[103]	PCA	82.5%
[144]	Gabor	96%
[145]	LBP	91%
[146]	LDA, PCA, Geometric feature	79% with algebraic features and 90.95% with geometric features
Our Method	Colour Features	99.60%
	Gabor Features	99.41%
	Combined Features	99.60%

Table 7.2: Ethnicity comparison in three groups Asian, White and Black from literature.

Reference	**Features**	**Classification Rate**
[12]	Grayscale Pixel intensities	92%
[7]	Gabor wavelet transform with retina sampling	Asian-96% European-93% African-94%
[147]	KCFA with facial colour	96%
[106]	LVC method	74%
[154]	Pixel intensity-based features, BIM features	85% for Asian vs. others, 86.5% for African vs. others and 78.5% for Caucasian vs. others
[159]	PCA	81.67%
Our Method	Colour Features	97.05%
	Gabor Features	91.78%
	Combined Features	98.31%

7.3 Future Work

The face detection method presented in this work has two limitations, one being that the input image should have a dark background and second that it should have only one face in an image.

Table 7.3: Ethnicity comparison in four groups Asian, Indian, White and Black from literature.

Reference	Features	Classification Rate
[104]	Biologically inspired features	Black-98.3% White-97.1%, Hispanic-74.2%, Asian-59.5%, Indian-6.90%
Our Method	Colour Features	95.06%
	Gabor Features	81.4%
	Combined Features	97.15%

In future, these problems can be solved by combining a connected components labelling algorithm and face skin colour from images to detect a face in an image. Facial components that were extracted have some limitations. For instance, incorrect components are extracted for images containing beards and spectacles. The problem with such categories of face images is that more than one connected component are merged into a single component, which yields an automatic removal of one or more targeted facial components. In future, this particular problem could be solved by using geometric approaches.

Features used for classification were confined to local-appearance based texture features and colour features. Key point features as well as shape features might encode more reliable information for classification. Geometric features such as the height and width of each facial component along with distances, can also be used for better ethnicity classification. Classification methods used in this work are machine-learning based, and therefore more insight into the problem may be provided by attempting a statistical classification method.

7.4 Conclusion

A facial based ethnicity classification system was presented. The face and facial components were first extracted from the face image using a connected component labelling algorithm. The presented method achieves better results when compared to the methods reported in the literature.

The reason is that the images are divided into two parts namely the upper and the lower part. Searching for two facial components in each part is simple when compared to searching for all parts together. For ethnicity classification, colour features, texture features, geometric features and a combination of colour and texture features were used. The classification was done using four different machine learning techniques namely K-Means clustering, Naive Bayesian classifier, MLP and SVM.

Of the three colour spaces (RGB, HSV and YCbCr) that were used for ethnicity classification, the RGB colour space provides better results when compared to HSV and YCbCr colour space. Another important finding is that hue, saturation, Cb and Cr channels do not contribute significantly to the classification of ethnicity. The blue colour component of the RGB colour space, the value component of HSV colour space and yellow component of the YCbCr colour space are more effective in ethnicity classification. Colour features provided more promising results which may be due to the fact that skin colour is one of the most important features of the human face for ethnicity identification. The skin colour differs from one ethnic group to another. One can identify the ethnicity of a person just by looking at the skin colour.

The performance of the ethnicity classification approach on facial texture images was evaluated. From the experiments, it is clear that a single facial component contributes greatly in ethnicity classification which is due to the fact that different ethnic groups have different physical facial characteristics. It is observed that among the four selected areas (the left eye, the right eye, the nose and the mouth), mostly the nose and the mouth gave superior results when compared to the left and the right eye. The nose and mouth component provided better results as it is seen in face components that the nose and the mouth region is very different for different ethnic groups when compared to the eyes which look similar. For example, the nose of the Black ethnic group is broad, for the Asian group it has an average width and for the White ethnic group it is narrow. The mouth for the Black ethnic group is stretched (with thick lips), for the Asian group it has average lips and for the White ethnic group it has thin lips. When combining the four facial regions, performance is improved, suggesting that each area has its own impact.

From our experiments, it is observed that colour features perform better classification when compared to texture features. Fusion of colour and texture features for ethnicity classification was also considered. This research shows that a combination of features provides better results when compared to the single colour or texture features. Within these experiments, a conclusion could be drawn with regard to the best classification method. MLP provided the best results as opposed to SVM which appears to be the weakest. In terms of application, the proposed system will be very helpful in video and image retrieval, database indexing, and can provide useful clues for face recognition and identification.

Bibliography

[1] U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain, "Biometric cryptosystems: issues and challenges," in *Proceedings of the IEEE*, vol. 92, pp. 948–960, IEEE, 2004.

[2] B. Dorizzi, S. Garcia-Salicetti, and M. A. Mellakh, "Multimodality in biosecure: Towards an evaluation protocol on virtual multi-modal databases," in *13th European Signal Processing Conference*, pp. 1–4, IEEE, 2005.

[3] V. Matyáš and Z. Říha, "Biometric authentication security and usability," in *Advanced Communications and Multimedia Security*, pp. 227–239, Springer, 2002.

[4] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 4–20, IEEE, 2004.

[5] D. Dessimoz, J. Richiardi, C. Champod, and A. Drygajlo, "Multimodal biometrics for identity documents," *Forensic Science International*, vol. 167, no. 2, pp. 154–159, 2007.

[6] "Biometric history." http://www.biometrics.gov/Documents/BioHistory.pdf, 7 Aug, 2006.

[7] S. Hosoi, E. Takikawa, and M. Kawade, "Ethnicity estimation with facial images," in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 195–200, IEEE, 2004.

[8] H. Lin, H. Lu, and L. Zhang, "A new automatic recognition system of gender, age and

ethnicity," in *The Sixth World Congress on Intelligent Control and Automation*, vol. 2, pp. 9988–9991, IEEE, 2006.

[9] K. S. Yoon, Y. K. Ham, and R.-H. Park, "Hybrid approaches to frontal view face recognition using the neural network," in *International Conference on Neural Networks*, vol. 3, pp. 1359–1362, IEEE, 1997.

[10] X. Ding and C. Fang, "Discussions on some problems in face recognition," in *Advances in Biometric Person Authentication*, pp. 47–56, Springer, 2005.

[11] X. Lu and A. K. Jain, "Ethnicity identification from face images," in *Defense and Security*, pp. 114–123, International Society for Optics and Photonics, 2004.

[12] S. Gutta, H. Wechsler, and P. J. Phillips, "Gender and ethnic classification of face images," in *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 194–199, IEEE, 1998.

[13] P. Vuilleumier, J. L. Armony, J. Driver, and R. J. Dolan, "Distinct spatial frequency sensitivities for processing faces and emotional expressions," *Nature Neuroscience*, vol. 6, no. 6, pp. 624–631, 2003.

[14] N. P. Costen, D. M. Parker, and I. Craw, "Effects of high-pass and low-pass spatial filtering on face identification," *Perception & Psychophysics*, vol. 58, no. 4, pp. 602–612, 1996.

[15] B. Heisele, T. Serre, M. Pontil, and T. Poggio, "Component-based face detection," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR*, vol. 1, pp. I–657, IEEE, 2001.

[16] J. Lu and K. N. Plataniotis, "Boosting face recognition on a large-scale database," in *International Conference on Image Processing*, vol. 2, pp. 109–112, IEEE, 2002.

[17] M. H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[18] S. H. Jeng, H. Y. M. Liao, C. C. Han, M. Y. Chern, and Y. T. Liu, "Facial feature detection using geometrical face model: an efficient approach," *Pattern Recognition*, vol. 31, no. 3, pp. 273–282, 1998.

[19] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, 1990.

[20] B. Heisele, P. Ho, J. Wu, and T. Poggio, "Face recognition: component-based versus global approaches," *Computer Vision and Image Understanding*, vol. 91, no. 1, pp. 6–21, 2003.

[21] B. Heisele, T. Serre, and T. Poggio, "A component-based framework for face detection and identification," *International Journal of Computer Vision*, vol. 74, no. 2, pp. 167–181, 2007.

[22] R. Dewi Agushinta, A. Suhendra, S. Madenda, and H. Suryadi, "Face component extraction using segmentation method on face recognition system," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 2, no. 2, pp. 67–72, 2011.

[23] T. Kazimov and S. Mahmudova, "About a method of recognition of race and ethnicity of individuals based on portrait photographs," *Intelligent Control and Automation*, vol. 5, pp. 120–125, 2014.

[24] G. Paul, *Ethnic Identity and the Perception of Facial Profile Aesthetics : A Preliminary Study*. PhD thesis, The University of Western Ontario, 2013.

[25] M. J. K. Leslie G. Farkas and C. R. Forrest, "International anthropometric study of facial morphology in various ethnic groups/races," *The Journal of Craniofacial Surgery*, vol. 16, no. 4, pp. 615–646, 2005.

[26] J. Ashbourn, *Biometrics: Advanced Identity Verification: the Complete Guide*. Springer, 2000.

[27] A. Mir, S. Rubab, and Z. Jhat, "Biometrics verification: a literature survey," *International Journal of Computing and ICT Research*, vol. 5, no. 2, pp. 67–80, 2011.

[28] M. Gamassi, M. Lazzaroni, M. Misino, V. Piuri, D. Sana, and F. Scotti, "Accuracy and performance of biometric systems," in *Proceedings of the 21st IEEE Instrumentation and Measurement Technology Conference (IMTC 04)*, vol. 1, pp. 510–515, IEEE, 2004.

[29] S. Hossain and G. Chetty, "Human identity verification by using physiological and behavioural biometric traits," *International Journal of Bioscience, Biochemistry and Bioinformatics*, vol. 1, no. 3, pp. 199–205, 2011.

[30] H. Faulds, "On the skin-furrows of the hand," *Nature*, vol. 22, p. 605, 1880.

[31] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*. Springer Science & Business Media, 2009.

[32] Mridula and Priyanka, "A review on classification of fingerprint images," *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, vol. 9, no. 3, pp. 61–66, 2014.

[33] L. Hong, *Automatic Personal Identification Using Fingerprints*. PhD thesis, Michigan State University, East Lansing, MI, USA, 1998.

[34] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–741, 1995.

[35] F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis, "Face recognition," in *Encyclopedia of Multimedia*, pp. 235–243, Springer, 2006.

[36] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003.

[37] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 6, pp. 681–685, 2001.

[38] C. Tirkaz and S. Albayrak, "Face recognition using active appearance model," in *IEEE 17th Signal Processing and Communications Applications Conference*, pp. 940–943, April 2009.

[39] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775–779, 1997.

[40] M. Hanmandlu, D. Gupta, and S. Vasikarla, "Face recognition using elastic bunch graph matching," in *Applied Imagery Pattern Recognition Workshop: Sensing for Control and Augmentation (AIPR)*, pp. 1–7, IEEE, 2013.

[41] P. S. Penev and J. J. Atick, "Local feature analysis: A general statistical theory for object representation," *Network: Computation in Neural Systems*, vol. 7, no. 3, pp. 477–500, 1996.

[42] Z.-m. Qian, P.-y. Su, and D. Xu, "Face recognition based on local feature analysis," in *International Symposium on Computer Science and Computational Technology (ISCSCT'08)*, vol. 2, pp. 264–267, IEEE, 2008.

[43] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[44] M. Slavković and D. Jevtić, "Face recognition using eigenface approach," *Serbian Journal of Electrical Engineering*, vol. 9, no. 1, pp. 121–130, 2012.

[45] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[46] H. Sahoolizadeh and Y. A. Ghassabeh, "Face recognition using eigen-faces, fisher-faces and neural networks," in *7th IEEE International Conference on Cybernetic Intelligent Systems (CIS)*, pp. 1–6, IEEE, 2008.

[47] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Transactions on Neural Networks*, vol. 13, no. 6, pp. 1450–1464, 2002.

[48] I. Marques and M. Graña, "Face recognition with lattice independent component analysis and extreme learning machines," *Soft Computing*, vol. 16, no. 9, pp. 1525–1537, 2012.

[49] K. I. Kim, K. Jung, and H. J. Kim, "Face recognition using kernel principal component analysis," *IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 40–42, 2002.

[50] S. Meshgini, A. Aghagolzadeh, and H. Seyedarabi, "Face recognition using Gabor filter bank, kernel principle component analysis and support vector machine," *International Journal of Computer Theory and Engineering*, vol. 4, no. 5, pp. 767–771, 2012.

[51] M. H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods," in *Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 215–220, IEEE, 2002.

[52] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Computation*, vol. 12, no. 10, pp. 2385–2404, 2000.

[53] B. Vesnicer, J. Z. Gros, N. Pavešić, and V. Štruc, "Face recognition using simplified probabilistic linear discriminant analysis," *International Journal of Advanced Robotic Systems*, vol. 9, pp. 1–10, 2012.

[54] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98–113, 1997.

[55] P. Latha, L. Ganesan, and S. Annadurai, "Face recognition using neural networks," *Signal Processing: An International Journal (SPIJ)*, vol. 3, no. 5, pp. 153–160, 2009.

[56] K. Jonsson, J. Kittler, Y. Li, and J. Matas, "Support vector machines for face authentication," *Image and Vision Computing*, vol. 20, no. 5, pp. 369–375, 2002.

[57] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[58] A. Das, *Iris Recognition in Eigen Space*. PhD thesis, Jadavpur University Kolkata, 2012.

[59] S. L. Baird, "Biometrics" security technology": It is important for students to understand that technology can be used as part of a solution to a problem. (resources in technology)," *The Technology Teacher*, vol. 61, no. 5, p. 18, 2002.

[60] S. Viriri, *Bi-Modal Biometrics Authentication Based on Iris and Signature*. PhD thesis, University of KwaZulu-Natal, Durban, South Africa, 2010.

[61] L. Flom and A. Safir, "Iris recognition system." http://www.google.com/patents/US4641349, Feb. 3 1987. US Patent 4641349.

[62] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1148–1161, 1993.

[63] Y. P. Huang, S. W. Luo, and E. Y. Chen, "An efficient iris recognition system," in *International Conference on Machine Learning and Cybernetics*, vol. 1, pp. 450–454, IEEE, 2002.

[64] Y. Zhu, T. Tan, and Y. Wang, "Biometric personal identification based on iris patterns," in *15th International Conference on Pattern Recognition*, vol. 2, pp. 801–804, IEEE, 2000.

[65] Y. Du, B. Bonney, R. Ives, D. Etter, and R. Schultz, "Analysis of partial iris recognition using a 1d approach," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*, vol. 2, pp. 961–964, IEEE, 2005.

[66] M. Bača, P. Grd, and T. Fotak, *Basic Principles and Trends in Hand Geometry and Hand Shape Biometrics*. INTECH Open Access Publisher, 2012.

[67] N. Science and T. Council(NSTC), "Hand geometry.." Biometrics.gov. NSTC Subcommittee on Biometrics, 27 March 2006.

[68] M. Golfarelli, D. Maio, and D. Malton, "On the error-reject trade-off in biometric verification systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 786–796, 1997.

[69] A. Ross and A. K. Jain, "Human recognition using biometrics: an overview," in *Annales Des Télécommunications*, vol. 62, pp. 11–35, Springer, 2007.

[70] R. Mahajan, T. Gupta, S. Mahajan, and N. Bawa, "Retina as authentication tool for covert channel problem," *World Academy of Science, Engineering and Technology*, no. 56, pp. 153–158, 2009.

[71] M. F. Zibran, "Eye based authentication: Iris and retina recognition," tech. rep., Technical Report# 2011-04, Department of Computer Science, University of Saskatchewan, Canada, 2009.

[72] S. Sukumaran and M. Punithavalli, "Retina recognition based on fractal dimension," *International Journal of Computer Science and Network Security*, vol. 9, no. 10, pp. 66–70, 2009.

[73] R. Purkait, "Ear biometric: An aid to personal identification," *Anthropology Today: Trends, Scope and Applications*, vol. 3, pp. 215–218, 2007.

[74] D. J. Hurley, M. S. Nixon, and J. N. Carter, "Force field feature extraction for ear biometrics," *Computer Vision and Image Understanding*, vol. 98, no. 3, pp. 491–512, 2005.

[75] L. Yuan, Z. Mu, and Z. Xu, "Using ear biometrics for personal recognition," in *Advances in Biometric Person Authentication*, pp. 221–228, Springer, 2005.

[76] N. Duta, A. K. Jain, and K. V. Mardia, "Matching of palmprints," *Pattern Recognition Letters*, vol. 23, no. 4, pp. 477–485, 2002.

[77] L. Zhang and L. Zhang, "Characterization of palmprints by wavelet signatures via directional context modeling," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 3, pp. 1335–1347, 2004.

[78] C. C. Han, H. L. Cheng, C. L. Lin, and K. C. Fan, "Personal authentication using palmprint features," *Pattern recognition*, vol. 36, no. 2, pp. 371–381, 2003.

[79] A. Kong, D. Zhang, and M. Kamel, "Palmprint identification using feature-level fusion," *Pattern Recognition*, vol. 39, no. 3, pp. 478–487, 2006.

[80] A. W.-K. Kong, D. Zhang, and G. Lu, "A study of identical twins palmprints for personal verification," *Pattern Recognition*, vol. 39, no. 11, pp. 2149–2156, 2006.

[81] T. Connie, A. T. B. Jin, M. G. K. Ong, and D. N. C. Ling, "An automated palmprint recognition system," *Image and Vision computing*, vol. 23, no. 5, pp. 501–515, 2005.

[82] S. Ribaric and I. Fratric, "A biometric identification system based on eigenpalm and eigenfinger features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, pp. 1698–1709, 2005.

[83] J. Zok, "Vascular pattern recognition." Biometrics.gov. NSTC Subcommittee on Biometrics, 7-Aug, 2006.

[84] V. S. S. V. H. Kawai and V. A. Wakabayashi, "Business expansion of palm vein pattern authentication technology," *Fujitsu Sci. Tech. J*, vol. 41, no. 3, pp. 341–347, 2005.

[85] D. Mulyono and H. S. Jinn, "A study of finger vein biometric for personal identification," in *International Symposium on Biometrics and Security Technologies (ISBAST)*, pp. 1–8, IEEE, 2008.

[86] A. Kumar and K. V. Prathyusha, "Personal authentication using hand vein triangulation and knuckle shape," *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 2127–2136, 2009.

[87] M. Watanabe, “Palm vein authentication,” in *Advances in Biometrics:Sensors, Algorithms and Systems*, pp. 75–88, Springer, 2008.

[88] G. Zayaraz, V. Vijayalakshmi, and D. Jagadiswary, “Securing biometric authentication using DNA sequence and naccache stern knapsack cryptosystem,” in *International Conference on Control, Automation, Communication and Energy Conservation (INCACEC)*, pp. 1–4, IEEE, 2009.

[89] R. Plamondon and S. N. Srihari, “Online and off-line handwriting recognition: a comprehensive survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 63–84, 2000.

[90] J. P. Campbell Jr, “Speaker recognition: A tutorial,” *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1437–1462, 1997.

[91] M. F. BenZeghiba and H. Bourlard, “User-customized password speaker verification using multiple reference and background models,” *Speech Communication*, vol. 48, no. 9, pp. 1200–1213, 2006.

[92] C. Montacié, P. Deléglise, F. Bimbot, and M.-J. Caraty, “Cinematic techniques for speech processing: temporal decomposition and multivariate linear prediction,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-92)*, vol. 1, pp. 153–156, IEEE, 1992.

[93] F. K. Soong, A. E. Rosenberg, B. H. Juang, and L. R. Rabiner, “Report: A vector quantization approach to speaker recognition,” *AT&T Technical Journal*, vol. 66, no. 2, pp. 14–26, 1987.

[94] M. P. Murray, A. B. Drought, and R. C. Kory, “Walking patterns of normal men,” *The Journal of Bone & Joint Surgery*, vol. 46, no. 2, pp. 335–360, 1964.

[95] L.-F. Liu, W. Jia, and Y.-H. Zhu, “Survey of gait recognition,” in *Emerging Intelligent Computing Technology and Applications. With Aspects of Artificial Intelligence*, pp. 652–659, Springer, 2009.

[96] J. E. Boyd and J. J. Little, “Biometric gait recognition,” in *Advanced Studies in Biometrics*, pp. 19–42, Springer, 2005.

[97] R. D. Seely, *On a Three-dimensional Gait Recognition System*. PhD thesis, School of Electronics and Computer Science, University of Southampton, 2010.

[98] P. S. Teh, A. B. J. Teoh, and S. Yue, “A survey of keystroke dynamics biometrics,” *The Scientific World Journal*, vol. 2013, 2013.

[99] F. Monrose and A. D. Rubin, “Keystroke dynamics as a biometric for authentication,” *Future Generation Computer Systems*, vol. 16, no. 4, pp. 351–359, 2000.

[100] A. Ross and A. Jain, “Information fusion in biometrics,” *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2115–2125, 2003.

[101] N. Poh, C. H. Chan, J. Kittler, J. Fierrez, and J. Galbally, “Description of metrics for the evaluation of biometric performance.” http://www.beat-eu.org/, 2012.

[102] R. M. Bolle, J. H. Connell, S. Pankanti, N. K. Ratha, and A. W. Senior, “The relation between the roc curve and the cmc,” in *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 15–20, IEEE, 2005.

[103] Y. Ou, X. Wu, H. Qian, and Y. Xu, “A real time race classification system,” in *IEEE International Conference on Information Acquisition*, pp. 378–383, IEEE, 2005.

[104] G. Guo and G. Mu, “A study of large-scale ethnicity estimation with gender and age variations,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 79–86, IEEE, 2010.

[105] Z. Yang and H. Ai, "Demographic classification with local binary patterns," in *Advances in Biometrics*, pp. 464–473, Springer, 2007.

[106] C. Zhong, Z. Sun, and T. Tan, "Fuzzy 3D face ethnicity categorization," in *Advances in Biometrics*, pp. 386–393, Springer, 2009.

[107] G. Zhang and Y. Wang, "Multimodal 2d and 3d facial ethnicity classification," in *Fifth International Conference on Image and Graphics (ICIG'09)*, pp. 928–932, IEEE, 2009.

[108] U. G. A. Office, "Using biometrics for border security," November 2002.

[109] "Book-understanding biometrics, griaule biometrics." http://www.griaulebiometrics.com/en-us/book/understanding-biometrics/introduction/applications-2.

[110] X. Giné, J. Goldberg, S. Sankaranarayanan, P. Sheerin, and D. Yang, "Use of biometric technology in developing countries." Banking the World: Empirical Foundations of Financial Inclusion, Cambridge, MA: MIT Press, 2013.

[111] I. P. Commission, "Performance evaluation of targeted public distribution system," 2005.

[112] K. C. Yow and R. Cipolla, "Feature-based human face detection," *Image and Vision Computing*, vol. 15, no. 9, pp. 713–735, 1997.

[113] E. Hjelmås and B. K. Low, "Face detection: A survey," *Computer Vision and Image Understanding*, vol. 83, no. 3, pp. 236–274, 2001.

[114] G. Yang and T. S. Huang, "Human face detection in a complex background," *Pattern Recognition*, vol. 27, no. 1, pp. 53–63, 1994.

[115] J. Wang and T. Tan, "A new face detection method based on shape information," *Pattern Recognition Letters*, vol. 21, no. 6, pp. 463–471, 2000.

[116] C. Lin and K.-C. Fan, "Human face detection using geometric triangle relationship," in *15th International Conference on Pattern Recognition*, vol. 2, pp. 941–944, IEEE, 2000.

[117] K.-W. Wong, K.-M. Lam, and W.-C. Siu, "An efficient algorithm for human face detection and facial feature extraction under different conditions," *Pattern Recognition*, vol. 34, no. 10, pp. 1993–2004, 2001.

[118] A. Majumder, L. Behera, and V. K. Subramanian, "Automatic and robust detection of facial features in frontal face images," in *13th International Conference on Computer Modelling and Simulation (UKSim)*, pp. 331–336, IEEE, 2011.

[119] S. Chaudhari, A. Kale, K. S. Kinage, and S. Bhirud, "Face feature detection and normalization based on eyeball center and recognition," in *2nd International Conference on Future Computer and Communication (ICFCC)*, vol. 3, pp. 503–507, IEEE, 2010.

[120] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 10, pp. 1042–1052, 1993.

[121] T. K. Leung, M. C. Burl, and P. Perona, "Finding faces in cluttered scenes using random labeled graph matching," in *Fifth International Conference on Computer Vision*, pp. 637–644, IEEE, 1995.

[122] H. P. Graf, T. Chen, E. Petajan, and E. Cosatto, "Locating faces and facial parts," in *Proc. First International Workshop Automatic Face and Gesture Recognition*, pp. 41–46, 1995.

[123] M. Hamouz, J. Kittler, J.-K. Kamarainen, and P. Paalanen, "Affine-invariant face detection and localization using gmm-based feature detector and enhanced appearance model," in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 67–72, IEEE, 2004.

[124] I. Craw, D. Tock, and A. Bennett, "Finding face features," in *Computer Vision ECCV'92*, pp. 92–96, Springer, 1992.

[125] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic face identification system using flexible appearance models," *Image and Vision Computing*, vol. 13, no. 5, pp. 393–401, 1995.

[126] O. Bernier, M. Collobert, R. Feraud, V. Lemaire, J.-E. Viallet, and D. Collobert, "Multrak: A system for automatic multiperson localization and tracking in real-time.," in *International Conference on Image Processing (ICIP)*, pp. 136–140, 1998.

[127] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 130–136, IEEE, 1997.

[128] A. L. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates," *International Journal of Computer Vision*, vol. 8, no. 2, pp. 99–111, 1992.

[129] L. Fan and K. K. Sung, "A combined feature-texture similarity measure for face alignment under varying pose," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 308–313, IEEE, 2000.

[130] J.-C. Terrillon, M. David, and S. Akamatsu, "Detection of human faces in complex scene images by use of a skin color model and of invariant fourier-mellin moments," in *Fourteenth International Conference on Pattern Recognition*, vol. 2, pp. 1350–1355, IEEE, 1998.

[131] X. Tao, H. Yun, and F. Chongxi, "Automatic facial feature detection and its application in image coding.," *Journal of Tsinghua University,*, vol. 39, no. 5, pp. 81–84, 1999.

[132] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color and edge information," in *Seventh International Symposium on Signal Processing and Its Applications*, vol. 1, pp. 525–528, IEEE, 2003.

[133] H. F. Hashem, "Adaptive technique for human face detection using HSV color space and neural networks," in *National Radio Science Conference (NRSC)*, pp. 1–7, IEEE, 2009.

[134] L. Zhi Fang, Y. Zhi sheng, A. Jain, and W. Yun qiong, "Face detection and facial feature extraction in color image," in *Fifth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA)*, pp. 126–130, IEEE, 2003.

[135] Y. B. Jemaa and S. Khanfir, "Automatic local Gabor features extraction for face recognition," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 3, no. 1, 2009.

[136] M. P. Segundo, L. Silva, O. R. P. Bellon, and C. Queirolo, "Automatic face segmentation and facial landmark detection in range images," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 40, no. 5, pp. 1319–1330, 2010.

[137] L. Shams and J. Spoelstra, "Learning Gabor based features for face detection," in *Proceedings of World Congress in Neural Networks. International Neural Network Society*, pp. 15–20, 1996.

[138] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 4, pp. 349–361, 2001.

[139] A. Rakhmadi, M. Rahim, A. Bade, H. Haron, and I. Amin, "Loop back connected component labeling algorithm and its implementation in detecting face," *World Academy of Science, Engineering and Technology*, vol. 64, pp. 105–110.

[140] J. Huang, V. Blanz, and B. Heisele, "Face recognition using component-based SVM classification and morphable models," in *Pattern Recognition with Support Vector Machines*, pp. 334–341, Springer, 2002.

[141] J. F. Dombeu and J.-R. Tapamo, "Validation of detected facial components for and accurate face recognition," in *Proceedings of the 18th Annual Symposium of Pattern Recognition Association of South Africa (PRASA), Pietermaritzburg, South Africa*, pp. 41–46, 2007.

[142] G. Shakhnarovich, P. Viola, and B. Moghaddam, "A unified learning framework for real time face detection and classification," in *Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 14–21, IEEE, 2002.

[143] D. Zhang, Y. Wang, and B. Bhanu, "Ethnicity classification based on gait using multi-view fusion," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 108–115, IEEE, 2010.

[144] F. S. Manesh, M. Ghahramani, and Y.-P. Tan, "Facial part displacement effect on template-based gender and ethnicity classification," in *11th International Conference on Control Automation Robotics & Vision (ICARCV)*, pp. 1644–1649, IEEE, 2010.

[145] J. R. Lyle, P. E. Miller, S. J. Pundlik, and D. L. Woodard, "Soft biometric classification using periocular region features," in *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pp. 1–7, IEEE, 2010.

[146] D. Xiao-dong, W. Cun-rui, L. Xiang-dong, L. Zhi-Jie, W. Jun, and Z. Hai-long, "Ethnic features extraction and recognition of human faces," in *2nd International Conference on Advanced Computer Control (ICACC)*, vol. 2, pp. 125–130, IEEE, 2010.

[147] Y. Xie, K. Luu, and M. Savvides, "A robust approach to facial ethnicity classification on large scale face databases," in *IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pp. 143–149, IEEE, 2012.

[148] S. H. Salah, H. Du, and N. Al-Jawad, "Fusing local binary patterns with wavelet features for ethnicity identification," in *Proceedings of IEEE International Conference on Signal Image Process*, vol. 21, pp. 416–422, 2013.

[149] U. Tariq, Y. Hu, and T. S. Huang, "Gender and ethnicity identification from silhouetted face profiles," in *16th IEEE International Conference on Image Processing (ICIP)*, pp. 2441–2444, IEEE, 2009.

[150] G. Toderici, S. M. Omalley, G. Passalis, T. Theoharis, and I. A. Kakadiaris, "Ethnicity and gender based subject retrieval using 3D face recognition techniques," *International Journal of Computer Vision*, vol. 89, no. 2-3, pp. 382–391, 2010.

[151] M. A. H. Abdullah and E. W. Abood, "Race classification using craniofacial features from colored face images," *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 3, no. 6, pp. 138–143, 2014.

[152] D. Huang, H. Ding, C. Wang, Y. Wang, G. Zhang, and L. Chen, "Local circular patterns for multi-modal facial gender and ethnicity classification," *Image and Vision Computing*, vol. 32, no. 12, pp. 1181–1193, 2014.

[153] A. Ahmad, A. Ashraf, and W. Salameh, "Arabic race classification of face images," *International Journal of Computers & Technology*, vol. 4, no. 2, pp. 234–239, 2013.

[154] M. Demirkus, K. Garg, and S. Guler, "Automated person categorization for video surveillance using soft biometrics," in *SPIE Defense, Security, and Sensing*, vol. 7667, pp. 76670P–76670P–12, International Society for Optics and Photonics, 2010.

[155] A. Hadid and M. Pietikäinen, "Demographic classification from face videos using manifold learning," *Neurocomputing*, vol. 100, pp. 197–205, 2013.

[156] H. Ding, D. Huang, Y. Wang, and L. Chen, "Facial ethnicity classification based on boosted local texture and shape descriptions," in *10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, pp. 1–6, IEEE, 2013.

[157] S. Gutta, J. R. Huang, P. Jonathon, and H. Wechsler, "Mixture of experts for classification of gender, ethnic origin, and pose of human faces," *IEEE Transactions on Neural Networks*, vol. 11, no. 4, pp. 948–960, 2000.

[158] X. Lu, H. Chen, and A. K. Jain, "Multimodal facial gender and ethnicity identification," in *Advances in Biometrics*, pp. 554–561, Springer, 2005.

[159] S. Buchala, N. Davey, T. M. Gale, and R. J. Frank, "Principal component analysis of gender, ethnicity, age, and identity of face images," vol. 7, Citeseer, 2005.

[160] H. Han, C. Otto, X. Liu, and A. Jain, "Demographic estimation from face images: Human vs. machine performance," 2014.

[161] H. Han and A. Jain, "Age, gender and race estimation from unconstrained face images," *Dept. Comput. Sci. Eng., Michigan State Univ., East Lansing, MI, USA, MSU Tech. Rep.(MSU-CSE-14-5)*, 2014.

[162] G. Muhammad, M. Hussain, F. Alenezy, G. Bebis, A. M. Mirza, and H. Aboalsamh, "Race classification from face images using local descriptors," *International Journal on Artificial Intelligence Tools*, vol. 21, no. 05, p. 1250019, 2012.

[163] E. Kiciman, D. Maltz, and J. C. Platt, "Fast variational inference for large-scale internet diagnosis," in *Advances in Neural Information Processing Systems*, pp. 1169–1176, 2008.

[164] I. Rish, M. Brodie, and S. Ma, "Accuracy vs. efficiency trade-offs in probabilistic diagnosis," in *Eighteenth national conference on Artificial intelligence*, (Menlo Park, CA, USA), pp. 560–566, American Association for Artificial Intelligence, 2002.

[165] C. Liu, "A bayesian discriminating features method for face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 725–740, 2003.

[166] B. Moghaddam, T. Jebara, and A. Pentland, "Bayesian face recognition," *Pattern Recognition*, vol. 33, no. 11, pp. 1771–1782, 2000.

[167] H. C. Shih, C. Y. Chuang, C. L. Huang, and C. H. Lin, "Gender classification using bayesian classifier with local binary patch features," in *IEEE 4th International Conference on Nonlinear Science and Complexity (NSC)*, pp. 45–50, IEEE, 2012.

[168] A. Beygelzimer, J. Langford, and B. Zadrozny, "Machine learning techniques reductions

between prediction quality metrics," in *Performance Modeling and Engineering*, pp. 3–28, Springer, 2008.

[169] S. Marsland, *Machine learning: an algorithmic perspective*. Chapman & Hall/CRC Machine Learning & Pattern Recognition, CRC press, 2014.

[170] H. D. III, *A Course in Machine learning*. CIML, 2012.

[171] A. Z. Chitade and S. Katiyar, "Colour based image segmentation using k-means clustering," *International Journal of Engineering Science and Technology*, vol. 2, no. 10, pp. 5319–5325, 2010.

[172] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297, Oakland, CA, USA., 1967.

[173] M. Trosset, "Representing clusters: K-means clustering, self-organizing maps, and multidimensional scaling," tech. rep., Technical Report 08-03, Department of Statistics, Indiana University, Bloomington, IN, 2008.

[174] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[175] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[176] Z. Jiang, Z. Lin, and L. S. Davis, "Class consistent k-means: Application to face and action recognition," *Computer Vision and Image Understanding*, vol. 116, no. 6, pp. 730–741, 2012.

[177] M. Kyperountas, A. Tefas, and I. Pitas, "Face recognition via adaptive discriminant clustering," in *15th IEEE International Conference on Image Processing (ICIP)*, pp. 2744–2747, IEEE, 2008.

[178] H. Meng, K. Appiah, A. Hunter, and P. Dickinson, "FPGA implementation of naive bayes classifier for visual object recognition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 123–128, IEEE, 2011.

[179] H. Rahimizadeh, M. Marhaban, R. Kamil, and N. Ismail, "Color image segmentation based on bayesian theorem and kernel density estimation," *European Journal of Scientific Research*, vol. 26, no. 3, pp. 430–436, 2009.

[180] K. Jayech and M. A. Mahjoub, "Clustering and bayesian network for image of faces classification," *International Journal of Advanced Computer Science and Applications, Special Issue on Image processing and Analysis (IJACSA)*, pp. 35–44, 2011.

[181] D. Vose, "Fitting distributions to data." `http://www.vosesoftware.com/vosesoftware/whitepapers/Fitting\%20distributions\%20to\%20data.pdf`, 2010.

[182] V. Ricci, "Fitting distributions with r," *Contributed Documentation available on CRAN*, vol. 96, 2005.

[183] E. W. Weisstein, "Normal distribution." `http://mathworld.wolfram.com/NormalDistribution.html`, 2002.

[184] M. I. Ribeiro, "Gaussian probability density functions: Properties and error characterization," *Institute for Systems and Robotics, Lisboa, Portugal*, 2004.

[185] H.-C. Kim, D. Kim, Z. Ghahramani, and S. Y. Bang, "Appearance-based gender classification with gaussian processes," *Pattern Recognition Letters*, vol. 27, no. 6, pp. 618–626, 2006.

[186] S. K. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 683–697, 1992.

[187] W. Sibanda and P. Pretorius, “Novel application of multi-layer perceptrons (MLP) neural networks to model HIV in south africa using seroprevalence data from antenatal clinics,” *International Journal of Computer Applications*, vol. 35, no. 5, pp. 26–31, 2011.

[188] G. Panchal, A. Ganatra, Y. Kosta, and D. Panchal, “Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers,” *International Journal of Computer Theory and Engineering*, vol. 3, no. 2, pp. 332–337, 2011.

[189] S. Karsoliya, “Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture,” *International Journal of Engineering Trends and Technology*, vol. 3, no. 6, pp. 713–717, 2012.

[190] S. Tamura, H. Kawai, and H. Mitsumoto, “Male/female identification from 8×6 very low resolution face images by neural network,” *Pattern Recognition*, vol. 29, no. 2, pp. 331–335, 1996.

[191] D. Bhattacharjee, D. K. Basu, M. Nasipuri, and M. Kundu, “Human face recognition using fuzzy multilayer perceptron,” *Soft Computing*, vol. 14, no. 6, pp. 559–570, 2010.

[192] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 2000.

[193] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[194] K. Crammer and Y. Singer, “On the algorithmic implementation of multiclass kernel-based vector machines,” *The Journal of Machine Learning Research*, vol. 2, pp. 265–292, 2002.

[195] R. L. Hsu, M. Abdel Mottaleb, and A. K. Jain, “Face detection in color images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696–706, 2002.

[196] S. K. Trivedi, S. Dey, and P. Shikhar, "Effect of various kernels and feature selection methods on svm performance for detecting email spams," *International Journal of Computer Applications*, vol. 66, no. 21, 2013.

[197] S. Mirghasemi and E. Banihashem, "Sea target detection based on SVM method using HSV color space," in *IEEE Student Conference on Research and Development (SCOReD)*, pp. 555–558, IEEE, 2009.

[198] AT&T, "The ORL database of faces." `www.uk.research.att.com/facedatabase.html`, 2002.

[199] I. multimedia Lab, "Asian face image database PF01," technical report, Pohang University of Science and Technology, San 31, Hyoja-Dong, Nam-Gu, Pohang, 790-784, Korea.

[200] K. Ricanek Jr and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," in *7th International Conference on Automatic Face and Gesture Recognition (FGR)*, pp. 341–345, IEEE, 2006.

[201] V. Jain and A. Mukherjee, "The indian face database," 2002.

[202] D. Marius, S. Pennathur, and K. Rose, "Face detection using colour thresholding and eigen-image template matching EE368: Digital image processing project," 2003.

[203] Q. Zhang and Z.-J. Liu, "Face detection based on complexional segmentation and feature extraction," in *International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06)*, pp. 293–296, IEEE, 2006.

[204] C. T. Yuen, M. Rizon, W. S. San, and M. Sugisaka, "Automatic detection of face and facial features," in *Proceedings of the 7th WSEAS International Conference on Signal Processing, Robotics and Automation (ISPRA '08)*, vol. 8, pp. 230–234, 2008.

[205] G. Ritter and J. Wilson, *Handbook of Computer Vision Algorithms in Image Algebra*. Second Edition CRC Press, 2001.

[206] Y. Yang and D. Zhang, "A novel line scan clustering algorithm for identifying connected components in digital images," *Image and Vision Computing*, vol. 21, no. 5, pp. 459–472, 2003.

[207] L. Di Stefano and A. Bulgarelli, "A simple and efficient connected components labeling algorithm," in *International Conference on Image Analysis and Processing*, pp. 322–327, IEEE, 1999.

[208] "Binary image analysis." `https://courses.cs.washington.edu/courses/cse576/book/ch3.pdf`.

[209] A. Chau, E. Oji, and J. Walters, "The detection of faces in color images: EE368 project report," tech. rep., Deptartment of Electrical Engineering, Stanford University.

[210] Y. Gizatdinova and V. Surakka, "Automatic localization of facial landmarks from expressive images of high complexity," 2008.

[211] B. Heisele, T. Serre, and T. Poggio, "A component-based framework for face detection and identification," *International journal of computer vision*, vol. 74, no. 2, pp. 167–181, 2007.

[212] T. W. Yoo and I. S. Oh, "A fast algorithm for tracking human faces based on chromatic histograms," *Pattern Recognition Letters*, vol. 20, no. 10, pp. 967–978, 1999.

[213] T. Sawangsri, V. Patanavijit, and S. Jitapunkul, "Face segmentation using novel skin-color map and morphological technique," in *Proceedings of World Academy of Science, Engineering and Technology*, vol. 2, pp. 41–44, Citeseer, 2005.

[214] M. M. Hasan and P. K. Mishra, "HSV brightness factor matching for gesture recognition system," *International Journal of Image Processing (IJIP)*, vol. 4, no. 5, pp. 456–467, 2010.

[215] L. M. Bergasa, M. Mazo, A. Gardel, M. Sotelo, and L. Boquete, "Unsupervised and adaptive gaussian skin-color model," *Image and Vision Computing*, vol. 18, no. 12, pp. 987–1003, 2000.

[216] D. A. Brown, I. Craw, and J. Lewthwaite, “A SOM based approach to skin detection with application in real time systems.,” in *Proceedings of the British Machine Vision Conference*, vol. 1, pp. 491–500, Citeseer, 2001.

[217] T. S. Caetano and D. A. C. Barone, “A probabilistic model for the human skin color,” in *11th International Conference on Image Analysis and Processing*, pp. 279–283, IEEE, 2001.

[218] C. Garcia and G. Tziritas, “Face detection using quantized skin color regions merging and wavelet packet analysis,” *IEEE Transactions on Multimedia*, vol. 1, no. 3, pp. 264–277, 1999.

[219] D. Chai and A. Bouzerdoum, “A bayesian approach to skin color classification in YCbCr color space,” in *TENCON 2000 Proceedings*, vol. 2, pp. 421–424, IEEE, 2000.

[220] D. Chai and K. N. Ngan, “Locating facial region of a head-and-shoulders color image,” in *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 124–129, IEEE, 1998.

[221] A. Samal and P. A. Iyengar, “Automatic recognition and analysis of human faces and facial expressions: A survey,” *Pattern Recognition*, vol. 25, no. 1, pp. 65–77, 1992.

[222] T. P. Weldon, W. E. Higgins, and D. F. Dunn, “Efficient Gabor filter design for texture segmentation,” *Pattern Recognition*, vol. 29, no. 12, pp. 2005–2015, 1996.

[223] K.-C. Chung, S. Kee, and S. Kim, “Face recognition using independent component analysis of Gabor filter responses,” in *IAPR Workshop on Machine Vision Applications*, pp. 331–334, Citeseer, 2000.

[224] Z. Sun, G. Bebis, and R. Miller, “Improving the performance of on-road vehicle detection by combining Gabor and wavelet features,” in *IEEE 5th International Conference on Intelligent Transportation Systems*, pp. 130–135, IEEE, 2002.

[225] C. C. Tsai, J. S. Taur, and C. W. Tao, "Iris recognition using Gabor filters optimized by the particle swarm technique," in *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pp. 921–926, IEEE, 2008.

[226] Y. Hamamoto, S. Uchimura, M. Watanabe, T. Yasuda, Y. Mitani, and S. Tomita, "A Gabor filter-based method for recognizing handwritten numerals," *Pattern Recognition*, vol. 31, no. 4, pp. 395–400, 1998.

[227] J. R. Movellan, "Tutorial on Gabor filters," *Open Source Document*, 2002.

[228] L. Shen, L. Bai, and M. Fairhurst, "Gabor wavelet and general discriminant analysis for face identification and verification," *Image and Vision Computing*, vol. 25, no. 5, pp. 553–563, 2007.

[229] S. Milborrow, J. Morkel, and F. Nicolls, "The MUCT landmarked face database," *Pattern Recognition Association of South Africa*, 2010. `http://www.milbo.org/muct`.