University of Benghazi

Faculty of Information Technology

Department of Computer Sciences

A Fusion of Features Exactions to Recognize Human Ear Biometrics

By

Bothaina Farag Mohammed Gargoum

Supervisor

Dr. Ahmed Lawgali

This Thesis was Submitted in Partial Fulfilment of the Requirements for Master's Degree in Computer Science

Date 22.8.2022

June 2022

Declaration

I…………………………………. confirm that the work contained in this thesis / dissertation, unless otherwise referenced is the researcher's own work, and has not been previously submitted to meet requirements of an award at this University or any other higher education or research institution, I furthermore, cede copyright of this thesis / dissertation in favour of University of Benghazi.

Student name: ....

Signature: ....

Date: .... / ..../ 2022.

University of Benghazi

Faculty of Information Technology

Department of Computer Sciences

A Fusion of Features Exactions to Recognize Human Ear Biometrics

By

Bothaina Farag Mohammed Gargoum

This thesis or dissertation has been approved by the examination committee on 22.8.2022.

Dr. ………………………………. (External examiner) ………………………

(Affiliation)

Dr. ………………………...……. (Internal examiner) ………..………………

(Affiliation)

Dr. ……………………………..…. (supervisor) ….……………..…………..…….

(Affiliation)

Dr. Othman Mohammed Albadri.

Head of Graduate Studies and Training Office

Dr. Abdelsalam Maatuk

Dean of the Faculty

Dedication

This thesis is dedicated to my husband, for his love, kindness, devotion, endless support, and encouragement.

Bothaina.F. gargoum

Acknowledgments

First and foremost, I am deeply grateful and thankful to God, who inspired me to accomplish this work. Then I would like to express my sincere gratitude to the following people and organizations:

My supervisor, Dr. Ahmed Lawgali, for his invaluable insight, guidance, patience, unwavering support, immense knowledge, and valuable critiques of this research work. This study would not have been possible without his input.

My husband, for his continuous support in continuing my studies and helping me to get over all the hardships I've been through. My family and close friends for their unconditional love and support.

I would also like to thank the Indian Institute of Technology Delhi for allowing the use of their ear database.

Last but not least, I would like to extend my thanks to my friends Amina and Shouroq. I am grateful for the support that I have received throughout my studies from them.

TABLE OF CONTENTS

Contents **Page No.**

LIST OF TABLES

Contents **Page No.**

LIST OF FIGURES

Contents **Page No.**

LIST OF ABBREVIATION

Abbreviation	Meaning
HOG	**H**istograms of **O**riented **G**radients
LBP	**L**ocal **B**inary **P**attern
PCA	**P**rincipal **C**omponent **A**nalysis
LDA	**L**inear **D**iscriminant **A**nalysis
IIT D	**T**he **I**ndian **I**nstitute of **T**echnology **D**elhi database
SVM	**S**upport **V**ector **M**achine
CNN	**C**onvolutional **N**eural **N**etworks
ANN	**A**rtificial **N**eural **N**etworks
RBF	**R**adial **B**asis **F**unction
DRT	**D**iscrete **R**adon **T**ransform
LTP	**L**ocal **T**ernary **P**atterns
HE	**H**istogram **E**qualization
LDP	**L**ocal **D**irectional **P**attern
FLD	**F**usions of **L**ocal **D**escriptors
ALBP	**A**veraged **L**ocal **B**inary **P**atterns
DWT	**D**iscrete **W**avelet **T**ransform
GF	**G**abor **F**ilter
KNN	**K**-**N**earest **N**eighbor
PHOG	**P**yramid **H**istogram of **O**riented Gradients
KDA	**K**ernel **D**iscriminant **A**nalysis
BoF	**M**ulti-**BOF** **H**istogram
RBF	**RBF** **N**etwork
ED	**E**uclidean **D**istance
UMRT	**U**nique **M**apped **R**eal **T**ransform
FN	**F**uzzy **N**etwork
RF	**R**andom **F**orest
DIC SHMM	**D**eviance **I**nformation **C**riterion-**S**tructural **H**idden **M**arkov **M**odel
PST	**P**olar **S**ine **T**ransform
2D-MBPCA	**T**wo-**D**imensional **M**ulti-**B**and **PCA**
DTP	**D**irectional **T**ernary **P**attern
MHD	**M**odified **H**ausdorff **D**istance

A Fusion of Features Exactions to Recognize Human Ear Biometrics

by

Bothaina Farag Gargoum

Supervisor

Dr. Ahmed Lawgali

Abstract

The demand for more secure authentication has increased on several occasions. Exploiting biometrics in various forms such as face, voice, handwriting, and gait recognition is a reliable method for authentication. Recently, the analysis of ear images as a biometric method has become a robust identification method. A number of researchers have shown that ear recognition is a viable alternative to more common biometrics such as fingerprint, face, and iris recognition, because the ear is relatively stable over time, non-invasive to capture, expressionless, and both the geometry and shape of the ear have significant variation among individuals. Researchers have tried a variety of methods to improve ear recognition. Some researchers have enhanced existing algorithms to assist in recognizing individuals by their ears. Others have taken algorithms that have been tried and tested for another purpose, such as face recognition, and applied them to ear recognition. These approaches have resulted in a number of state-of-the-art effective methods for identifying individuals by ear. Many of the challenges occur due to errors in the method of capturing images, poor illumination, image dimensions, off-angle ears, etc. The various methods have been adopted by researchers in order to enhance and increase the performance of ear recognition.

Most of the ear recognition systems incorporate processes before the feature extraction stage; first, the pre-processing stage, which is done to enhance only the region of interest. This stage includes segmentation and normalization. Subsequently, to enhance the normalized ear image. In this research, the Histogram Equalization (HE) technique has been implemented to facilitate the application of the feature extraction step. Then we presented an approach based on a fusion of two different techniques of feature extraction: Histograms of Oriented Gradients (HOG) and

Local Binary Patterns (LBP) to extract the desired features. whereas Principal Component Analysis (PCA) is used to reduce the space of the feature dimensionality. For classification, Linear Discriminant Analysis (LDA) is used. The proposed technique is applied to the images of the IITD I database. The proposed method has yielded significant achievements compared with other studies.

CHAPTER 1
Introduction

1.1. Overview

Security challenges are becoming increasingly essential in several organizations, such as financial services, e-commerce, telecommunications, government, traffic, and health care. It is critical to ensure that people are permitted to pass certain points or use certain resources. Following certain heinous abuses, security concerns arose rapidly. Organizations are interested in automated identity authentication systems for these reasons, since they will boost customer satisfaction and operational efficiency. Authentication systems will also save costs and be more precise than humans (Lammi, 2004).

A trustworthy automated biometric system that can establish or verify an individual's identity is of essential relevance in a modern culture where digital social interaction is becoming increasingly widespread and financial transactions are routinely handled through digital means. A biometric system is a pattern recognition system that uses a physiological or behavioral characteristic of a person to establish or verify an individual's identity by extracting prominent features from a questioned sample (image) and comparing them to a stored feature set or trained statistical model. Access cards, personal identification numbers (PINs), and passwords are examples of traditional methods for personal authentication that can be stolen, duplicated, lost, or forgotten. The development of biometric systems has proven to be an effective solution in overcoming the aforementioned limitations associated with traditional methods of personal authentication. Because of observable biometric features like universality, uniqueness, collectability, and permanence, biometric systems are inherently more reliable than most traditional techniques of personal authentication.

A human ear is a stable structure that does not change much with age and can be considered one of the most distinctive human biometric qualities because it possesses all of the aforementioned uniqueness, collectability, permanence, and universality attributes (Iannerelli, 1989). In 1996, Mark Burge and Wilhelm Burger were the first to try an automated ear-based biometric authentication system. They used a mathematical graph model to extract information from ear pictures automatically in order to match particular curves and edges (Burge and Burger, 1996).

The selection of whether two ear images belong to the same person or not is known as the ear recognition problem. This is considered a challenge due to variations in lighting, backdrop, position, scale, and occlusion. Biometric systems use two major modalities, physiological modalities and/or behavioural modalities as shown in Figure 1.1 to solve this problem and overcome this constraint. The face, fingerprint, palm-print, iris, ear, and other physiological modalities are examples of the first. Signatures, voice, keystrokes, strides, and other behavioural modalities are examples. Human ears, unlike other well-known features such as the face, have a stable structure that does not change over time (Pflug and Busch, 2012).

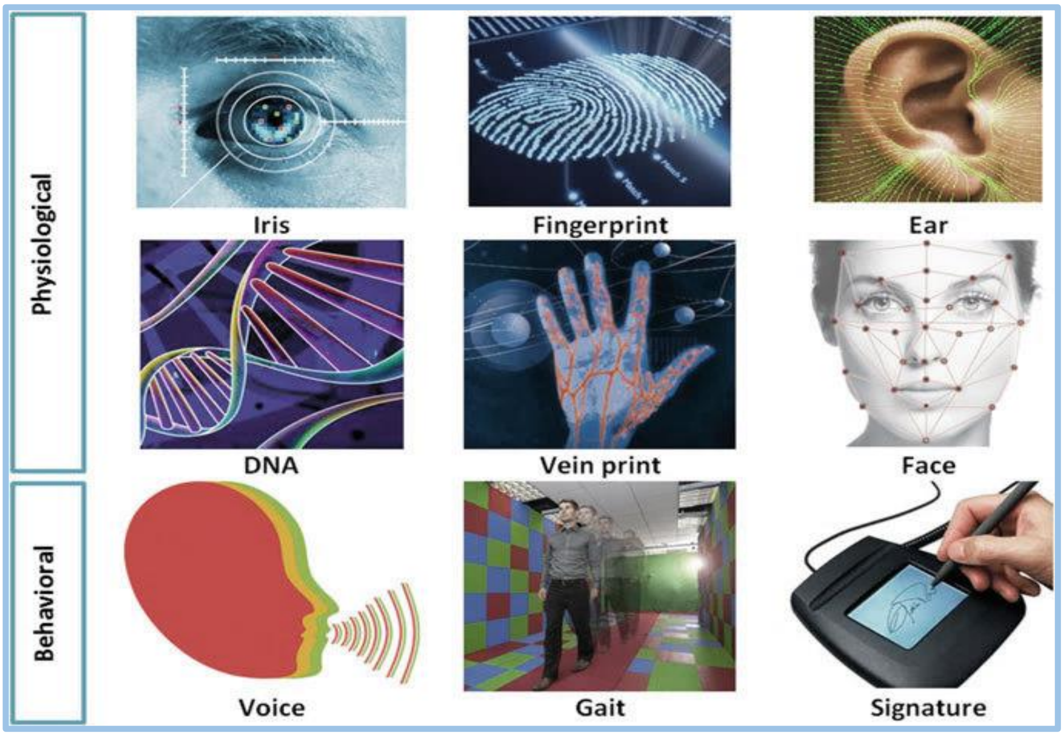

Figure 2.1: Biometric Classification

Furthermore, because the ear image may be captured passively, users are more likely to accept it than other features like fingerprints, palm prints, or iris, which need considerable user assistance. Because of its promising features, the human ear has been used as basic evidence in law enforcement. Since 2006, as evidence in hundreds of cases in the Netherlands and the United States (Burge and Burger, 1998) (Meijerman et al., 2009).

1.2. Problem Statement

Several studies have shown that ear recognition is a viable alternative to more commonly used biometrics such as fingerprint, face, and iris recognition (Kumar and Wu, 2012). The ear is more stable over time, requires less invasive image gathering, and does not require as much management as other biometrics. Furthermore, when compared to faces, it is plausible to say that the ear has fewer concerns about privacy. With advancements in computer vision and pattern recognition techniques, ear identification research is turning to a more difficult scenario in which ear images are collected in natural (unconstrained) environments (Emeršič et al., 2017c). This research is noteworthy since the ears are considered one of the most recent biometrics. They appear to keep their structure with age; thus, the frequently utilized ear detection techniques are based on the structure and form of the ear. However, identifying a person by acquiring an input ear image and matching it with known ear images in a database remains a difficult problem. This is due to the variability of human ear images under various operating conditions such as illumination as shown in Figure 1.2 a, rotation, different acquisition devices, low resolution, occlusions caused by hair and head accessories, earrings, headsets, camera viewpoints, and so on, all of which have a significant impact on the performance of ear recognition systems. Figure 1.2, b and c depict the difficulty of recognizing individuals using low-resolution ear images captured by various acquisition devices and occlusions caused by hair and head accessories. Because ear images provide higher identification richness than any other. The recognition that is required for secure methods of individual identification authentication is becoming increasingly critical (Benzaoui et al., 2014). Despite this, several ear recognition algorithms have been used to obtain excellent results. These strategies, however, remain a problem in ear recognition and must be improved (Ojala et al., 2002).

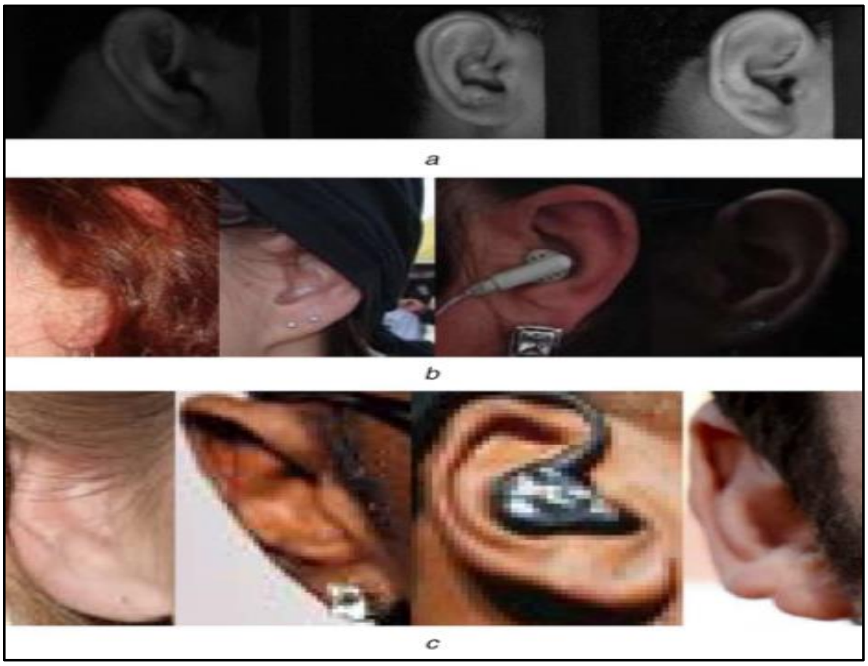

Figure 1.2: Examples of a challenges and limitations for ear recognition task in an unconstrained setting

1.3. Research Motivation

The human ear is a popular new feature in biometrics. It has various benefits over other biometric technologies, including iris, fingerprints, face, and retinal scans. The human ear is larger than the iris and fingerprint, and unlike them, it is particularly easy to image because it can be taken from a distance without the assistance of individuals (Arnia and Pramita, 2011). The human ear has a richness of features and is more dependable than the face since the structure of the ear does not vary with age or facial expressions (Muntasa et al., 2011). Governments, medical, robotics, telecommunications, healthcare, traffic, and universities all employ ear recognition. Researchers have been studying auditory recognition techniques over the past few years. In many applications, the human ear is an excellent source of data for passive person identification. Because ears are visible and their images may be easily acquired, even without the investigated person's awareness, ear biometrics appears to be a promising solution to an increasing need for security in different public areas. Uniqueness, Universality, Performance, and Collectability are all properties that a biometric feature should have. For many years, human ears have been considered a major aspect in forensic research (for example, in airplane crashes). Ear prints discovered at the crime scene have been used as evidence in hundreds of cases across Europe and the United States. Ear prints are now used by police and forensic specialists as a common form of identification because they do not change much over time. Additionally, because the ear is one of our sensors, it is normally exposed (not buried beneath anything) to allow for good listening (Rasika B. Naik, 2018). The goal of this project is to create a system that can recognize human ear patterns using a set of methods. Then assess the results of these techniques.

1.4. Research Aims and objectives

Many algorithms for extracting features for ear recognition systems have been used, each with a different recognition rate. As a result, the aim of this work is to evaluate the effectiveness of a fusion of two different separate techniques of feature extraction in capturing discriminative features of the ear in order to achieve high accuracy in recognizing human ear. Then, evaluate the results by comparing them with previous studies. The following are the primary objectives of this study:

- Provide the literature review on ear recognition systems to understand the mechanism of the methods.
- Using a fusion of two different feature extraction techniques, namely (HOG) transform-based feature and (LBP) local binary patterns, to extract the features.
- Using (PCA) Principal Component Analysis on the extracted vector to provide a holistic description of the sample images while reducing the dimensionality of the data without much loss of information.

- Recognize these features by using a Linear Discriminant Analysis (LDA) classifier.
- Evaluate and compare the results with existing works.

1.5. Research Scope and Limitations

This study focuses on ear recognition as a biometric technology because it does not change considerably during human life. Furthermore, the ear is one of our sensors, so it is usually visible (not hidden underneath anything) to enable good hearing. Our focus in this study is on the constrained 2D ear imaging database, which means all images were taken from the same profile angle and under varied interior lighting circumstances and performed in an indoor environment. The images that are occluded by hair, head accessories, earrings, or headsets do not take these into account.

1.6. Significance of the study

The presented study tries to provide the best methods that could be used for ear recognition according to the results that will be obtained to find out if those methods are suitable together or not. It will also show the quality of the system and the methods that were used.

1.7. Thesis organization

This thesis is comprised of six chapters. The first chapter includes an introduction and briefly discusses the overviews of the research, aims and objectives, problem statements, and research motivation for this research.

The rest of the research is organized as follows:

- Chapter 2 discusses an overview of the ear recognition system and presents the challenge of ear recognition. The chapter also introduces brief explanations of the feature extraction techniques, which will be utilized in the study. These techniques are **HE, LBP**, **HOG,** and **PCA**. Furthermore, the classifier **LDA** that is used to classify features obtained from the ear is explained in detail.
- Chapter 3 comprises a literature review and the results which were obtained in each study are reviewed and presented in a table. In addition, the common public and freely available ear image databases are described.
- Chapter 4 explains the major methodology that has been used in this study. Firstly, a general overview of the database that is used in this research is given.Then, a detailed

explanation of the most important stages of the ear recognition system is provided. These stages are, pre-processing, feature extraction, and finally classification.

- Chapter 5 accompanies the result and discussion. The most important results obtained from the experiments and the steps that have accompanied the implementation of these experiments are discussed. Moreover, observations of the obtained results are presented.
- Finally, Chapter 6 displays the conclusion. The conclusion describes the task that has been accomplished during this research.

CHAPTER 2

Overview of ear recognition system

2.1. Introduction

In this Chapter, some problems related to ear recognition in the attention of artificial intelligence research will be presented. This chapter is organized as follows: 2.1 is dedicated to presenting the introduction of ear recognition. Section 2.2 provides an anatomy of the human ear. Section 2.3 addresses the challenges in ear recognition, while Section 2.4 provides the ear recognition system stages. Moreover, this Chapter describes the techniques that will be used in this work; HE, HOG, LBP, PCA, and LDA.

In recent years, the increased need for security has prompted the development of biometric technologies for personal identification. The ability to identify people based on their outer ear shape was initially identified by French criminologist Bertillon, and then refined by American police officer Iannarelli, who created the first ear identification system based on only seven criteria. The ear's intricate structure is not only distinctive but also permanent, as the ear's appearance does not alter during a person's lifetime. Additionally, while obtaining ear images may not always necessitate a person's permission, most people believe it to be non-intrusive. Due to these features, the ear has attracted the majority of the focus of study. Because it is significantly less influenced by such changes, it is regarded as an alternative to be employed individually or in combination with the face. However, because of its small size and the common presence of nearby hair and earrings, it is difficult to use for non-interactive biometric applications (Houcine et al., 2015).

Ear prints discovered at the crime scene have been used as evidence in hundreds of cases across Europe and the United States. Police and forensic experts now use ear prints as a standard form of identification. There are lots of advantages to utilizing the ear as a data source for human identification. To start with, as previously mentioned, the ear is one of the most stable anatomical features of the human body. It does not vary significantly across a whole human life. Additionally, because the ear is one of our sensors, it is normally exposed (not buried beneath anything) to allow for good listening. Personal authentication reliability is critical to meeting the high security requirements in a variety of application areas, from airport surveillance to electronic banking. Many physiological traits of humans, known as biometrics, are often stable over time, easy to obtain, and unique to each person. The segmentation of the ear image from the profile face is the initial step in ear recognition. Due to variations in hair length and colour, ear images obtained at different times can differ significantly. Many incorrect point matches may arise as a result of this variance, greatly reducing the accuracy of image distance measurement (Tiwari et al., 2016).

The human ear has several advantages over other modalities: it has a rich structure, is a smaller object (low resolution), is stable over time, is a modality that people accept, is unaffected by changes in age, facial expressions, position, and rotation, and ear images can be acquired without the subject's participation and from a distance. (Yuan and chun Mu, 2012). Figure 2.1 shows various images of ears.

Figure 2.1: Images of Ears

2.2. Anatomy of Human Ear

The ear anatomy is likely unique to each individual, and features based on measurements of that anatomy are comparable throughout time, making biometrics based on the ear practical (Burge and Burger, 1998). The ear does not have a fully random structure; it, like the face, is made up of typical features. The ear's components are less well known than the eyes, nose, mouth, and other facial features, but they are always present in a healthy ear. The outer rim (helix), the ridge (antihelix) running within and parallel to the helix, the lobe, and the intertragic notch between the ear hole (meatus) and the lobe are all features. Figure 2.2 shows the locations of the anatomical features in detail (Hurley et al., 2005).

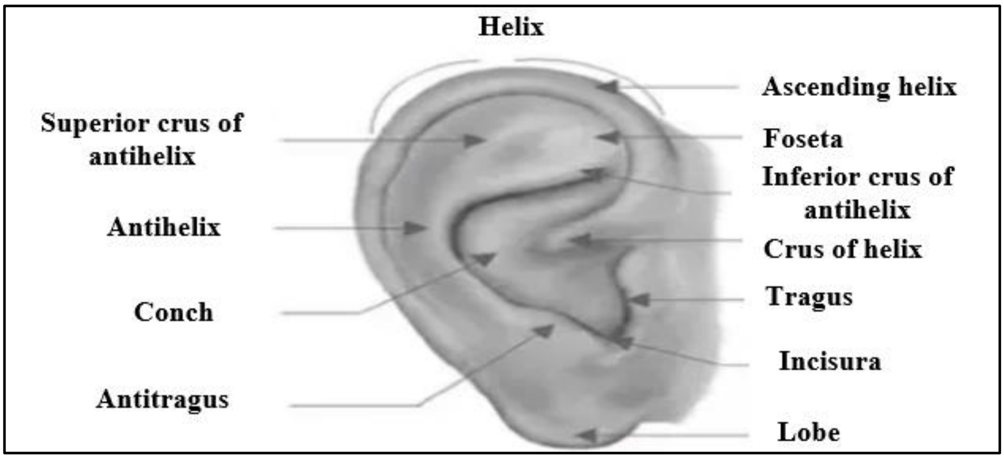

Figure 2.2: Anatomy of the Human Ear

2.3. Challenges in ear recognition

As the most recent publications on 2D and 3D ear recognition show, the main application of this technique is personal identification in unconstrained environments. This includes applications for smart surveillance but also the forensic identification of perpetrators from images or for border control systems. Traditionally, these application fields are part of face recognition systems, but as the ear is located next to the face, it can provide valuable additional information to supplement the facial images (Hansley et al., 2018).

In unconstrained environments, multi-modal ear and face recognition systems can help achieve posture invariance and increased robustness against occlusion. Surveillance cameras are installed overhead in most public places to capture as many people as possible and to prevent them from vandalism. Furthermore, because most people do not look directly into the camera, no frontal images of the people will be available. This poses a severe problem for biometric systems that rely on facial traits to identify people. In these situations, the ear can serve as a valuable supplementary characteristic if the face is not visible from a frontal angle. The quality of the image affects the extraction of discriminative features from the ear image, which affects the performance of ear localization and segmentation algorithms. When an ear image is taken under less-than-ideal conditions, segmentation and localization become difficult. In a noisy imaging environment, ear recognition is thus a difficult task. This factor automatically qualifies ear recognition as a fascinating artificial intelligence topic. The following are some difficult image acquisition challenges to consider.

2.3.1. Ear Localization and Normalization

Ear localization as shown in Figure 2.3, ear localization locates biometric information and isolates it from existing irrelevant parts of the collected sample. Despite the fact that many of the systems presented in the literature use pre-segmented ear images, automatic recognition of ears, particularly in real-life images, is still a challenge, and ear normalization reshapes the input sample to a standard format to reduce unwanted variances (Emeršič et al., 2017a, Asadi et al., 2010).

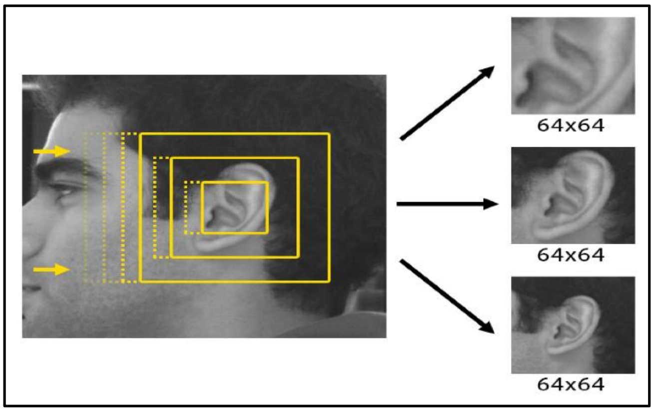

Figure 2.3: Ear Localization

2.3.2 Occlusion and Pose Variations

As shown in Figure 2.4, the ear, unlike the face, can be partially or completely covered by hair or other items such as headdresses, hearing aids, jewellery, or headphones. Parts of the outer ear may be blocked due to the convex surface if the subject's position changes. Parts of the outer ear may be obstructed due to the convex surface if the subject's position changes. Although resilience against occlusion is addressed in some publications, there is no research on the influence of specific types of occlusion, such as hair or earrings, on the recognition rate of an ear recognition system.

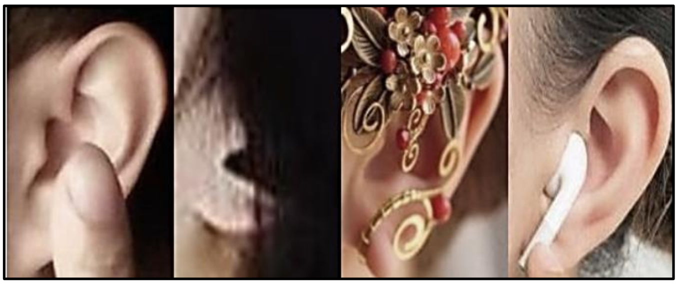

Figure 2.4: Occlusion and Pose Variations

2.3.3 Understanding Symmetry and Ageing

The symmetry of the left and right ear has not been well understood because ear recognition is one of the newest topics of biometric research. According to Abaza and Ross (Abaza and Ross, 2010), there is some degree of symmetry between the left and right ears, which could be leveraged when comparing the two. More research into the symmetry constraints between the left and right ear is encouraged as a result of their results.

To summarize, ear recognition is still a relatively new field of research. Despite the fact that there are a number of promising methods, none of them have been tested in realistic scenarios that involve disruptive factors such as position changes, occlusion, and changing lighting conditions. These aspects are taken into account in the latest methods, but more research is needed before ear recognition systems can be deployed in practice. The availability of appropriate test databases, which were collected under realistic settings, will help the ear mature as a biometric characteristic.

2.4. Stages of ear recognition system

In general, the ear recognition system is defined by the following: image acquisition, preprocessing, feature extraction, and ear matching. Figure 2.5 shows the ear recognition system process.

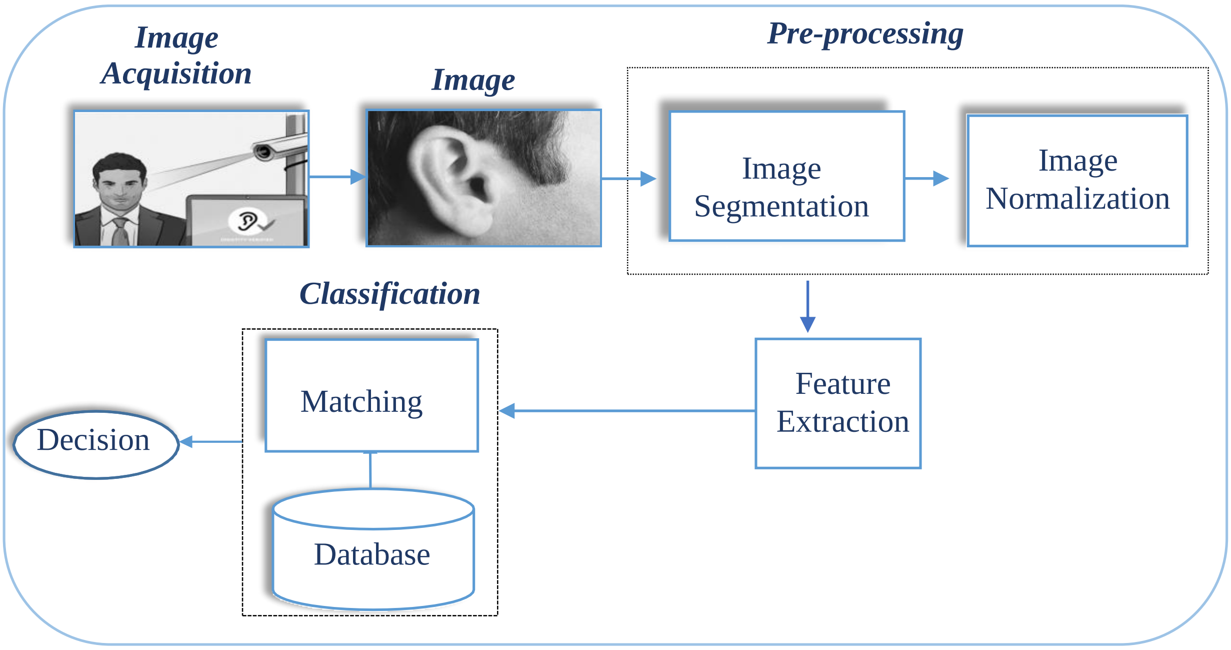

Figure 2.5: Stages of The Ear Recognition System

2.4.1. Image acquisition

Ear biometric databases help researchers carry out ear detection experiments and compare their results. The database is classified as 2D or 3D depending on the acquisition device. In this research, we will use the 2D database.

2.4.2. Pre-processing

It's the first step you take with the images. The goal of pre-processing is to adjust the source image to make feature extraction easier and the recognition rate higher. (Sivanarain and Viriri, 2020). The database that is targeted in this paper contains automatically segmented and normalized images. We applied the histogram equalization technique on the normalized image to minimize the effect of non-uniform lighting and obtain a well-distributed texture image.

2.4.3. Feature extraction

Feature extraction plays a significant role in determining the performance of the recognition system because it deals with isolating distinct features of the ear in the image (Ummer Akber Tali, 2015). The ear analysis is given by an individual description of its parts and their relationships, like measures of distances, angles, or the triangulation between the closed edge resulting from the helix shape and lobule of the ear, Crus of Helix and the Lobe (Burge and Burger, 2000), or the Vornoi diagram of the ear triangulation, Global approaches are based on pixel information; all pixels of the normalized ear image are treated as a single vector. The size of the vector is the total number of pixels (Burge and Burger, 1997).

2.4.4. Classification

This stage is about matching the probe and the gallery feature vectors to verify the subject or to search a database in order to identify the admitted person (Abaza and Harrison, 2013). The aim of this step is to measure the similarity and dissimilarity between two ear templates. It is also called the "matching stage," which means the degree of similarity is determined between the recognition template and the master template.

2.5. Technique used for image enhancement (Histogram Equalization)

To reduce the effect of non-uniform lighting and generate a well-distributed texture image, the histogram equalization technique is applied to the normalized ear image. A cumulative distribution transformation function is the histogram equalization technique. It is the process of transforming the original image's intensity into a more evenly distributed histogram (Srivastava and Rawat, 2013)

2.6. Techniques used for feature extraction

Feature extraction techniques vary depending on the application. Techniques that work well in one application might not work well in another. The most significant pattern for the ear has been extracted using several techniques. Specialists used their knowledge and expertise to create these techniques. HOG and LBP are combined to extract ear features while PCA is used to minimize the dimension of the feature vector.

2.6.1. Histogram of Oriented Gradients (HOG).

- The HOG descriptor is concerned with an object's structure or shape.
- In the case of edge features, we just determine whether or not a pixel is an edge. HOG is also capable of providing edge direction. This is accomplished by extracting the edges' gradient and orientation (or magnitude and direction).

- These orientations are also determined in 'localized' portions. This means that the entire image is divided into smaller regions, with gradients and orientation calculated for each.
- Finally, the HOG would generate a separate histogram for each of these regions. The 'Histogram of Oriented Gradients' is named after the histograms are created utilizing the gradients and orientations of the pixel values.

- **Advantages of HOG (Dalal and Triggs, 2005)**

The HOG descriptor has a few key advantages over other descriptors.

1. It works with local cells and is insensitive to geometric and photometric modifications, with the exception of object orientation. Only bigger spatial regions would see such changes.
2. As (Dalal and Triggs, 2005) observed, walkers' individual body movements may be ignored as long as they maintain a nearly upright position using coarse spatial sampling, fine orientation sampling, and strong local photometric normalization. As a result, the HOG descriptor is ideal for detecting humans in images.

2.6.2. Local Binary Pattern (LBP)

The LBP approach is a straightforward technique for extracting features from patterns (Li et al., 2015). Moreover, it also has a straightforward theory and combines the advantages of structural and statistical texture analysis methods. Because of its computational simplicity and discriminative power, the LBP approach has become a popular technique in the feature extraction field from patterns. As a result, it performs analysis on patterns in real-life situations (Ahonen et al., 2009). The basic LBP operator, as illustrated by (Ojala et al., 2002). Figure 2.6 depicts the process of extracting LBP features.

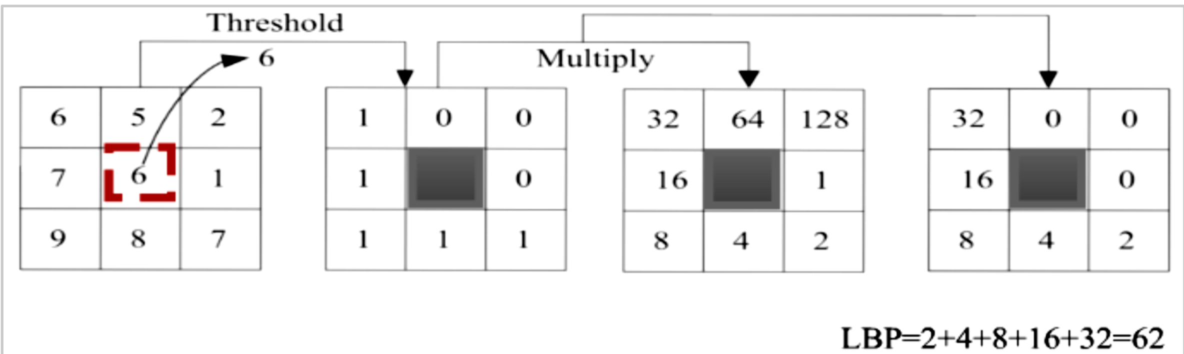

Figure 2.6: The process of LBP

The basic mechanism of the LBP technique is depicted in Figure 2.6. The LBP technique divides an input image into multiple blocks. Each block is broken into '3×3' pixels that make up the neighborhood (9 cells). The intensity value of each pixel is then encoded. Then, using the value of the central pixel, which in figure 2.6 is 6, LBP orders the surrounding pixels within the block from the upper-left corner down to the right one based on whether they have a higher or lower intensity value than the central pixel (higher value = 1; lower value = 0). Finally, a binary number (11111000) is obtained, which is transformed to a decimal number (62) and put into a one-dimensional array.

- **Advantages of LBP (Pietikäinen et al., 2011) (Ojala et al., 2002)**
    1. LBP gives a unified description of a texture patch that includes both statistical and structural characteristics, making it more powerful for texture analysis.
    2. The capacity to control gray-scale changes, such as illumination variations, is the most essential element of the LBP technique in its applications.

2.6.3. Principal Component Analysis (PCA)

PCA is the most popular and powerful feature extraction technique, and it works well with pattern recognition and compression. It works by reducing the dimension of the feature vector without losing much information (Kumari et al., 2019). For mapping data from a high-dimensional space to a low-dimensional space, PCA utilizes linear transformations. After a fusion of LBP and HOG features is extracted, PCA is used for dimensionality reduction. It's used to filter down the number of features to only those having a major difference between them. To begin, create a row vector by taking each pixel in an image row by row. A matrix is constructed by joining all of the row vectors (Resmi and Raju, 2019).

- **Advantages of PCA**
    1. Due to the orthogonal components, there is a lack of data redundancy (Phillips et al., 2005, Asadi et al., 2010)
    2. Using PCA, the complexity of image grouping was significantly reduced (Asadi et al., 2010, Phillips et al., 2005)
    3. Reduced database representation since only the trainee images are saved in the form of their projections (Phillips et al., 2005).
    4. Noise reduction is because the maximum variation basis is used, and minor variations in the background are automatically ignored (Phillips et al., 2005).

2.7. Technique used for classification: Linear Discriminant Analysis

This stage is about matching the feature vectors to verify the subject or to search a database to identify the accepted person. In our proposed approach, the LDA classifier was employed for classification. LDA has been successfully used as a classification technique for a number of problems, including speech recognition, face recognition, and multimedia information retrieval (B. S. El-Desouky, 2012). LDA is a linear transformation technique that's commonly used for dimensionality reduction. Lu and Jain show that LDA is a veritable technique for reducing the given multidimensional information to a lower measurement (Lu and Jain, 2004). As Figure 2.7 illustrates, LDA was used to perform supervised dimensionality reduction by projecting the input data to a linear subspace consisting of the directions that maximize the separation between classes. It showed strong feature extraction and dimensionality reduction.

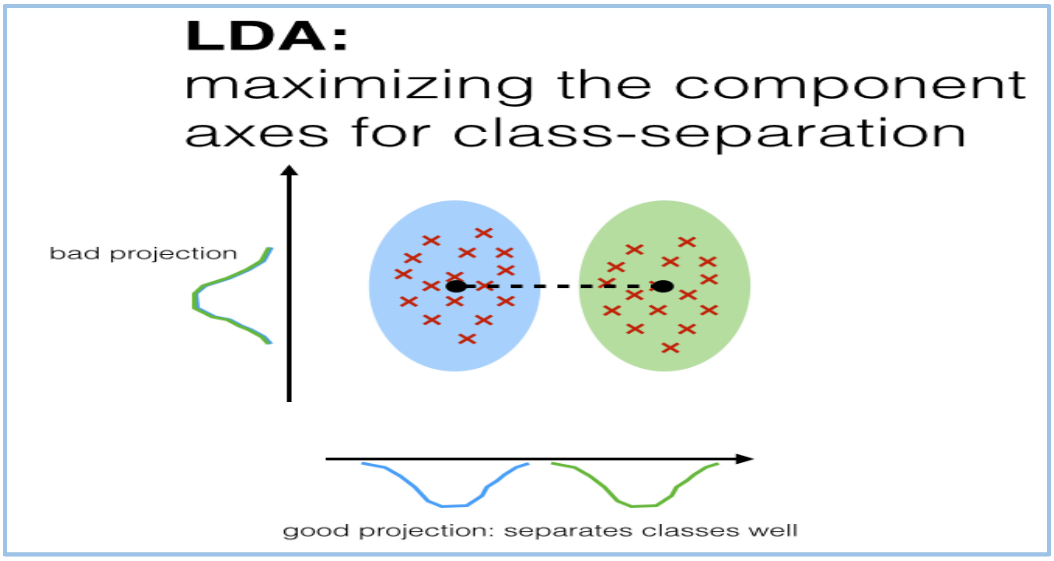

Figure 2.7: LDA Technique

2.7.1. Types of LDA

To deal with classes, there are two types of LDA techniques: class dependent and class independent. (Lu et al., 2013).

In class-dependent LDA, each class has its own lower-dimensional space on which to project its data. It requires more computation than a class-independent approach. Class dependent has two major drawbacks:

- The Class Independent approach requires more CPU time and calculations.
- It could result in difficulty with Small Sample Size (SSS).

In class independent LDA, each class is treated as a separate entity from the others, with just one lower dimensional space on which all classes can project their data. Most LDA techniques utilized in research use the class independent method rather than the class dependent method due to the two main drawbacks of the class dependent method.

2.8. Summary

In this Chapter, the anatomy of the ear is explained briefly. And then an overview of ear recognition and the challenges that face the researchers in this field were also presented. Furthermore, a detailed description of the techniques which will be used in this work has been provided. The HE technique is used for image enhancement. In the same context, three techniques will be applied for extracting the discriminant pattern: the HOG, LBP, and PCA, with LDA as a classifier. More details of the utilization of the fusion of HOG and LBP in the work will be explained in Chapter 4.

CHAPTER 3

Literature Review

3.1. Introduction

This Chapter explains that ear recognition is a succession of operations designed to extract a binary ear code from an ear image. The LBP, HOG, and PCA techniques have been focused on as the important methods for extracting the distinctive features. The issue of ear recognition remains a challenge for many researchers, and texture analysis plays an important role in the issue. Numerous algorithms for textural feature extraction and classification have been presented by researchers. In this chapter, the most cited human ear databases with their characteristics will be introduced. Moreover, an overview of the papers that studied the ear recognition system is provided. Studies will be categorized according to feature extraction techniques with classification techniques to find out the appropriate methods in the ear recognition system. Finally, a briefly summarized table of the previous studies of ear recognition is presented in Section 3.4.

3.2. Overview of the ear databases

In this section, we present an overview of the existing ear datasets that can be used for training and evaluation of 2D ear recognition approaches. CP, USTB I, USTB II, FEARID, IITD I, AMI, WPUT, and AWE. The datasets have different characteristics, and the corresponding ear images exhibit different levels of variability, as illustrated in the Figure. 3.1. The figure also shows the development of the datasets through time. Note how the datasets have progressed towards more realistic imaging conditions, as shown in Figure 3.1

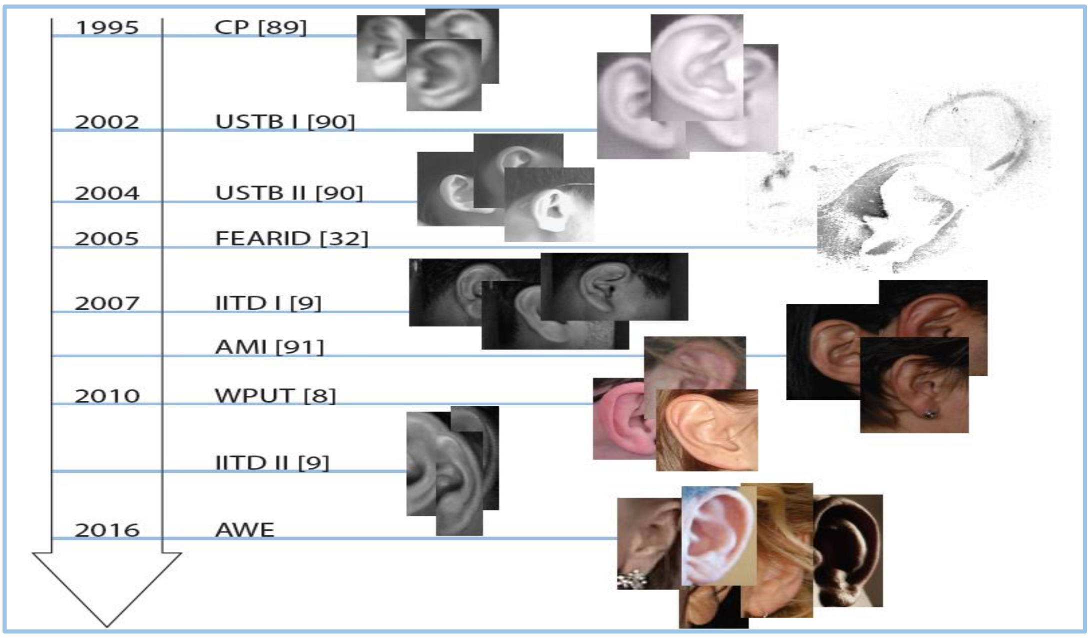

Figure 3.1: Sample Images of Available Ear Datasets

3.2.1. CP Ear Dataset

One of the earliest publicly available datasets for ear recognition is the Carreira-Perpinan (CP) ear dataset (Carreira-Perpinan, 1995). The dataset, which was first released in 1995, has 102 ears samples from 17 different subjects. Because all of the images were taken under controlled conditions, the majority of the image differences are attributable to minor pose changes and, of course, subject identity. Figure 3.1 also includes a few sample images from the dataset at the top.

3.2.2 IITD Ear Dataset

The Indian Institute of Technology Delhi's public ear dataset (KUMAR, 2007) is divided into two sub-datasets. The first includes 493 grayscale images of 125 subjects, while the second includes 793 images of 221 subjects. All of the images were taken from the same profile angle and under varied interior lighting circumstances. The first dataset (IITD I) is provided both raw and pre-processed, whereas the second dataset (IITD II) is only available pre-processed. The authors of the dataset ensured that: I) all ears are tightly cropped; II) all images are of equal dimensions; and III) all ears are cantered and mutually aligned through pre-processing.
Additionally, all images of the left ear are mirrored, so the entire dataset appears to consist of images of the right ear. The number of images per subject ranges from 3 to 6. No major occlusions are present.

3.2.3. USTB Ear Datasets

The University of Science and Technology of Beijing introduced four unique ear datasets that have gained widespread acceptance (L., 2002). USTB I and USTB II have just cropped ear images, whereas USTB III and IV have full head profile shots. The authors took 3 to 4 images of each of the 60 volunteers in the initial dataset, resulting in a total of 185 images. The authors collected 4 images from 77 subjects for the second dataset, totalling 308 images. Indoor lighting was employed in both datasets, although the second dataset is more challenging because it contains more loosely cropped images, a larger level of illumination-induced variability, and pitch angles of up to 30°. Full face-profile images were acquired under particular angles and occlusions in the USTB III and USTB IV datasets. USTB III has 1600 images of 79 subjects, whereas USTB IV has 500 images of 500 subjects, all of whom were photographed from 17 different angles with 15° steps in between. The USTB databases are available on demand and are free to use for research.

3.2.4. AMI Ear Dataset

The AMI (Mathematical Analysis of Images) public ear dataset (E. Gonzalez-Sanchez, 2008) was collected at the University of Las Palmas and contains 700 images of 100 different subjects ranging in age from 19 to 65 years. All of the images were taken with the same lighting and a consistent camera position. One image of the left ear and six images of the right ear were captured for each subject. All of the ears in the dataset have been cropped roughly so that the ear area takes up around half of the image. Subjects' positions differ somewhat in yaw (all images are still from profile) and significantly in pitch (subjects looking up at a 45° angle). The dataset is open to the public.

3.2.5. WPUT Ear Dataset

In 2010, the West Pomeranian University of Technology (WPUT) released the ear dataset (WPUT) (D. Frejlichowski, 2010), which included 2071 ear images from 501 subjects. The dataset now available for download, on the other hand, has 3348 images of 471 subjects, with 1388 duplicates. The images for subject IDs 337 through 363 are missing. Gender, age group, skin colour, head side, two types of rotation angles, lightning conditions, background (cluttered or diverse), and occlusion type are among the categories represented in the dataset (earrings, hats, tattoos, etc.). The collection contains between 4 and 10 images per subject. Images were captured in a variety of interior lighting conditions with head rotation angles ranging from 90° (profile) to 75°.

3.2.6. FEARID Ear Dataset

The FEARID dataset (Alberink and Ruifrok, 2007) was collected as part of the FEARID project and is unique in the type of data it contains. Unlike other datasets, FEARID contains ear-prints, which were collected by various research groups utilizing specialized scanning equipment. Because there are no occlusions, varied angles, or lighting differences in ear-print data, it differs dramatically from conventional images. Other variables, such as the force with which the ear was pressed on the scanner, scanning-surface cleanliness, and other comparable circumstances, also influence the appearance of the ear prints. The dataset's acquisition technique was created to simulate the appearance of ear-prints that would ordinarily be found at crime scenes, resulting in a dataset with 7364 images of 1229 subjects. The FEARID dataset is noteworthy since it was compiled with forensic applications in mind, as opposed to other ear datasets widely used in the biometric world.

3.2.7. AWE Ear Dataset

Datasets acquired directly from the web, such as the LFW (Huang et al., 2008), PubFig (Kumar et al., 2009), FaceScrub (Ng and Winkler, 2014), Casia WebFaces (Yi et al., 2014), IJB-A (Klare et al., 2015), and others, have been a cornerstone in this development. These datasets popularized the idea of images taken in the wild, with the pun implying that the images were taken in uncontrolled settings. The intricacy of these datasets allowed for substantial technological advancements and accelerated the field's progress in recent years. Sample photos from the Annotated Web Ears (AWE) dataset are shown in Figure 3.1.

In Table 3.1, a comparative summary of the existing ear datasets and their characteristics is given. The category "Sides" evaluates whether images from the left or right side of the head are present in the dataset, and "Accessories" assesses whether occlusion and accessories are visible in the images. The last two categories, "Gender" and "Ethnicities," indicate whether both sexes are present and what kind of racial variation is accounted for in the datasets.

Table 3.1: A comparative summary of the most popular ear datasets

Database name	Subjects	Sides	Accessories	Gender	Ethnicities
CP	17	Left	None	Both	White
USTB I	60	Right	Yes	Both	Asian
USTB II	77	Right	Yes	Both	Asian
IITD I	125	Right	Yes	Both	Asian
AMI	100	Both	None	Both	White
WPUT	501	Both	Yes	Both	White
AWE	100	Both	Yes	Both	Various

3.3. The popular algorithms of feature extraction used in ear recognition

Because it isolates various aspects of the ear in the image, feature extraction is crucial to the performance of the identification system. Various algorithms have been used in the literature to extract features. In order to properly depict study trends, previous studies of various ear techniques are provided. At the end of this part, there is a table that summarizes all of the algorithms covered in this Section, allowing for a quick comparison of the system's basic attributes and performance.

3.3.1. Feature extraction and matching

This Section provides an overview of existing strategies for extracting a set of measurable features from ear images in the context of ear-based biometric authentication systems that have been proposed. A brief overview of the important template matching algorithms proposed in the context of ear-based biometric authentication systems is offered, as well as their respective performances.

The researchers used multiple ways to extract the ear pattern for feature extraction and then used various classifiers for matching. The following are the most widely used feature extraction methods in the literature:

In this research (Mutar et al., 2020), presented a combination of features based on Random Forest (RF) and Histograms of Oriented Gradients (HOG) approaches in the feature extraction stage, and they used HOG to extract features from ear images. These extracted features will then be sent into the RF classifier, which will categorize the ear images according to the classes with high accuracy. The ear images were chosen from the second version of the Indian Institute of Technology in Delhi (IITD II).

Furthermore, (Kumari et al., 2019) attempt to develop and test the performance of an image processing system using a combination of Probabilistic Principal Component Analysis (PPCA) and Coherent Point Detection (CPD) techniques, as well as a PPCA-CPD distance-based classifier. In addition to Euclidean distance, KNN and Eigen distance classifiers are used. Two datasets, AMI and USTB, were utilized to evaluate the suggested method's performance.

This research (Kohlakala and Coetzer, 2021) proposed a novel semi-automated and fully automated ear-based biometric authentication system. The discrete Radon transform (DRT) is subsequently applied to the resulting binary contour image for the purpose of feature extraction. Experimental validation is achieved by implementing a Euclidean distance. Experiments are conducted on two independent ear databases, the AMI ear database and the IITD I ear database. The results are encouraging.

To study the effects of the fusions of these descriptors, (Sivanarain and Viriri, 2020) used local texture descriptors: Local Binary Patterns (LBP) and provided extensions of various local descriptors, namely Local Ternary Patterns (LTP), Local Directional Patterns (LDP), and Directional Ternary Patterns (DTP) for ear recognition. Fusions of Local Descriptors is a novel approach for combining these descriptors that was proposed (FLD). Experiments were conducted on the publically available IIT Delhi databases IITD I and IITD II, which included a variety of subjects and conditions. The experiments produce amazing outcomes that are on par with, if not better than, the state-of-the-art.

Moreover, the proposed method (Resmi and Raju, 2019) has four phases. The first is template matching for ear detection. In the second phase, the detected ear's size is normalized, and in the third step, features are extracted from the ear to represent it as a vector using LBP (Local Binary Pattern). After extracting LBP features, PCA is used to reduce dimensionality.

Finally, the characteristics are classified using the KNN classifier. Two datasets, RR and IITD I, were used to test the proposed method.

They present an efficient, reliable, and convenient automatic human ear detection technique (Hadi et al., 2021). This method has two stages: pre-processing and detecting ear landmarks. Image contrast, Laplace filter, and Gaussian blurring techniques were used to enhance all images (increasing the contrast, reducing the noise, and smoothing processes). They then used the Sobel edge detector to highlight the ear edges and the image substation technique to determine the only white pixels of the ear edges. Furthermore, LBP was used as a feature extractor and KNN as a classifier. An IITD-I standard ear biometric public dataset is used to test their method.

They propose a detailed comparative experimental investigation to evaluate the performance of numerous LBP-based features in ear recognition under both controlled and unrestrained imaging settings (Hassaballah et al., 2019). Another LBP variant known as Averaged Local Binary Patterns (ALBP) is also introduced, utilizing a very basic thresholding scheme. The five extensively utilized ear datasets used in the experiments are: IITD I, IITD II, AMI, WPUT, and AWE. Because of their simplicity and efficiency of computation, LBP-based descriptors are ideal candidates for ear feature extraction under controlled imaging settings.

For ear recognition, (Houcine et al., 2015) proposed a novel visual feature representation approach named Multi-BOF Histogram. They initially convolve an ear image with J Gabor Filters that have the same parameters except for the orientation parameter. Then J features can be extracted from the obtained responses of each pixel at each scale and orientation. Then, for each pixel, a unique features vector, such as "multi-scale Gabor features vector"), can be assigned. For classification, the KNN classifier is used. Extensive tests were run on the IITD I database.

This research presented by (Anwar et al., 2015) introduced a new ear identification technique based on the extraction of geometrical features. The mean of the ear image, the centroid of x coordinate, the centroid of y coordinate, and four distinct distances from the matrix, which contain the Euclidean distance between every pixel in the image, are all retrieved as feature vectors. They attempted to enhance the distance values in order to produce a more representative feature vector. Because the feature vector is still small but representative, there is no effect on the execution time. The K-nearest neighbor classifier is used for classification since it is more accurate. Experiments on the IITD I database were carried out.

To represent ear images, (Sarangi et al., 2019) used two local feature descriptors, Pyramid Histogram of Oriented Gradients (PHOG) and Local Directional Patterns (LDP). The LDP efficiently encodes local texture information while the PHOG reflects spatial shape information. Prior to normalization and fusion, they utilized principal component analysis (PCA) to minimize the dimension. Then, to create a single feature vector, two normalized heterogeneous feature sets are joined. Finally, using a K-Nearest Neighbor (KNN) classifier, the Kernel Discriminant Analysis (KDA) approach is used to extract nonlinear discriminant features for efficient recognition. Experiments were carried out on three standard datasets: IITD I, IITD II, and UND (Collection E).

This work (Omara et al., 2018) proposes a local feature fusion-based improvement method for unconstrained ear recognition, as well as an analysis of the performance and efficiency of discriminative local feature fusion for aligned and non-aligned ear images. The ear images are first processed to extract local discriminative features such as LPQ, HOG, LBP, POEM, BSIF, and Gabor Filter. Then, for fusion and dimension reduction, Discriminant Correlation Analysis (DCA) is used. Finally, for classification, a support vector machine (SVM) is used. The USTB I, USTB II, and IITD II databases are used in the experiments.

They suggested a technique based on Particle Swarm Optimization (PSO), Discrete Wavelet Transform (DWT), and Fuzzy Neural Network in their paper (Hussein et al., 2021). The particle swarm optimization is used to select more effective and attractive features from the ear image using the discrete wavelet transform. Furthermore, because particle swarm optimization minimizes the number of features, it reduces the complexity of the classification stage. In the classification stage, a fuzzy neural network was used to provide robust differentiation between the testing and training images. Experiments were performed on the IITD I and AMI Ear Datasets, two standard ear databases.

This work (Hamdany et al., 2021) designed an efficient deep learning (DL) model for ear print recognition. Deep ear print learning is the name of this model (DEL). It's a deep network that's been precisely built to recognize segmented and normalized ear patterns. Experiments in this work were performed on the IITD I ear database. For the suggested DEL, the best obtained accuracy is recorded.

For ear recognition, (Ahila Priyadharshini et al., 2021) suggested a six-layer deep convolutional neural network architecture. On the IITD-II ear dataset and the AMI ear dataset, the deep network's potential efficiency is examined. The suggested system's robustness is tested in an uncontrolled environment using the AMI Ear dataset. When paired with a competent surveillance system, this approach can be beneficial in identifying individuals in a large crowd.

In this research (Emeršič et al., 2017b) develop a CNN-based ear recognition model, investigate various strategies for model training with limited training data, and demonstrate that selecting an appropriate model architecture, aggressive data augmentation, and selective learning on existing (pre-trained) models can yield better results. Deep CNN was utilized for classification and feature extraction, with histograms of oriented gradients. They used the previously released AWED and CVLED datasets to create a dataset of unconstrained ear images for their studies.

This work presented by (Alhanjouri and ISSN, 2018) used an ear classification problem to improve the Deviance Information Criterion-Structural Hidden Markov Model (DIC SHMM) using a Convolutional Neural Network (CNN), which is a deep learning technology. To classify ear images, three systems were used: deep learning for the original image, deep learning for the eigenvector as Principle Components Analysis (PCA) of the original image, and proposed combining convolution layers of CNN with better HMM for the original image. Images from the AMI Ear Database were used (coloured images).

An investigation approach based on Polar Sine Transform (PST) was presented by (Omara et al., 2016) They first split the ear images into overlapping blocks, then compute PST coefficients, which are then used to extract invariant features for each block. Second, they combine these features into a single feature vector to represent the image of the ear. For ear recognition, they used support vector machines (SVM). The experiment was carried out on the USTB ear database.

This work presented by (Hansley et al., 2018) developed CNN-based ear normalization and description solutions, using well-known handcrafted descriptors and fusing learnt and handcrafted features to improve recognition. They created a two-stage landmark detector that performed well in untrained scenarios. The data was then utilized to perform a geometric image normalization, which improved the performance of all descriptors tested. The combination of learnt and handcrafted matchers appears to be complementary, since it outperformed all others in all tests. The IITD I ear database was used for the experiments in this study.

In this work they created a new ear dataset called the Multi-PIE ear dataset using the Multi-PIE face dataset (Eyiokur et al., 2018). They combined multiple deep convolutional neural network models using VGG-16 as the feature extraction to improve the performance even further. They looked into the influence of ear image quality in great detail. Finally, they addressed the issue of dataset bias in the field of ear recognition. According to experiments on the UERC dataset, domain adaptation improves performance significantly.

This research presented by (Alkhraisat and Security, 2017) introduced a technique that combines the benefits of SURF for feature extraction, PCA to reduce the dimension of the feature vector to a lower dimension (PCA-SURF), and the scalable K-means++ algorithm for feature clustering into one technique. Finally, local and global commonalities are calculated to classify the ear images. The suggested algorithm's performance and efficiency are evaluated using the AMI Ear Database.

This work presented by (Zarachoff et al., 2018) propose a two-dimensional multi-band PCA (2D-MBPCA) approach that outperforms the PCA. The proposed approach uses either dynamic or predefined equal range thresholds to divide the input grey image into a number of images based on the intensity of its pixels. PCA is then used to extract features from the resulting set of images. For classification, the obtained features are used to discover the best match and Euclidean distance. They employed images from the IITD II and USTB II databases, which are both benchmark ear imaging datasets.

To overcome the problem of threshold segmentation, an adaptive threshold segmentation method was used by (Luo et al., 2008) to find the threshold automatically; to reduce the computational complexity, a quick classification was realized by combining the Canny-operator and the Modified Hausdorff Distance (MHD). Finally, the algebraic property of the force field was combined with Principal Component Analysis (PCA) and Linear Discriminant

Analysis (LDA) together to obtain feature vectors for ear recognition. They tested these applications of the force field transform on the USTB ear database.

This research presented by (Zarachoff et al., 2021) offers a 2D Wavelet-based Multi-Band Principal Component Analysis (2D-WMBPCA) ear recognition approach inspired by PCA-based multispectral and hyperspectral image recognition algorithms. The proposed 2D WMBPCA method divides the input image into wavelet sub bands using a 2D non-decimated wavelet transform. Each sub band's resulting frames are subjected to conventional PCA, giving eigenvectors that are used for classification. The images of two benchmark ear datasets, IITD II and USTB I, were used to test the results. Furthermore, the suggested technique provides results comparable to those of learning-based strategies in a fraction of the time and without the requirement to be trained.

In this research (Omara et al., 2017) they suggest using VGG-M Net to extract deep features from ear images in order to solve the ear recognition problem. They propose to employ the paired SVM for classification first due to the lack of training photos per person. For computational efficiency, Principal Component Analysis (PCA) is used to minimize the dimension before classification. Finally, they test their method on two publicly available ear databases: USTB I and USTB II. In comparison to state-of-the-art approaches, the experimental results reveal a promising recognition rate and superior performance.

In this work (Dodge et al., 2018), they proposed approaches for unconstrained ear recognition using DNNs. The best results were obtained using an ensemble of ResNet18 models that performed consistently across all datasets evaluated. For classification, they applied an SVM classifier. On unconstrained ear recognition datasets, the AWE and CVLE, performance results were recorded. They illustrate that, when compared to DNN feature-extraction based models and single fine-tuned models, their ensemble achieves the greatest recognition performance on these datasets.

In this research (Benzaoui et al., 2015), they leverage and exploit current local texture-based descriptors to get faster and more accurate results in their study, which is simple yet effective. They used BSIF as a feature extraction and as a classifier, Support Vector Machines (SVM) were used. They test their findings with two publicly available databases, IITD I and IITD II, which contain a variety of ear standards of various types under various settings and imaging quality. The experiments produce great results that go beyond what is currently available.

In this work (Zhang et al., 2018), they first propose the USTB-Hello ear database, a new huge ear database that can be used to train and test ear recognition and verification systems. The images in this collection were captured in uncontrolled lighting and position variations with partial occlusions. Therefore, Second, they tweaked and changed some of the database's deep models. They used ICA to extract features and the RBF Network to classify them.

For ear recognition Problem (Sowmyalakshmi and Girirajkumar, 2017), proposed a computationally efficient person recognition technique based on the ear biometric modality's Unique Mapped Real Transform (UMRT). In this research, two features were extracted: histogram equalization and classification using UMRT, KNN, and Euclidean distance classifiers. The proposed ear identification technique is compared to a technique based on the Uniform Local Binary Pattern (ULBP). The findings of the testing using IITD I and internal GEAR ear database photos are promising.

This work presented by (Pflug et al., 2014), extensive testing showed that the LPQ and BSIF descriptors, in combination with LDA as a dimensionality reduction approach and the cosine distance as a nearest neighbour classifier, produce the best results for three distinct datasets: UND-J2, AMI, and IITK. According to the researchers, smaller local windows with more spatially constrained descriptors do not increase performance. This is because smaller radii for local descriptors are more susceptible to noise, and the number of dimensions in the concatenated histogram becomes excessively lengthy.

A summary of ear recognition, which was surveyed above. Note from the literature the difference in the use of databases and their size. Furthermore, various techniques for extracting distinctive features and matching. As can be seen, the Indian Institute of Technology Delhi (IITD) ear image is used as the dataset in the majority of the studies. Standard databases used in the previous research include USTB, AMI, UERC, and AWE. In reality, the database and feature extraction methods chosen have a significant impact on the system's accuracy rate. In brief, some techniques produce good results regardless of the database or classification system used. In most other surveyed methods, LBP, PCA, HOG, CNN, and BSIF had reasonable accuracy rates.

3.4. Summary

This Chapter offered an overview of the ear recognition system. As a result, the HE technique will be used to enhance the image quality. In the feature extraction stage, we chose HE since it is used in much research such as (Zarachoff et al., 2021., Hussein et al., 2021., Zarachoff et al., 2018) and has a successful impact on the results. Based on previous studies, we observed that the HOG technique provides good results, which can be seen in (Mutar et al., 2020., Sowmyalakshmi and Girirajkumar, 2017). Several researchers such as (Hadi et al., 2021., Resmi and Raju, 2019) have also utilized LBP with promising outcomes. From this standpoint, a fusion of HOG and LBP techniques was used with PCA and then LDA as a classifier.

The following Chapter describes an approach to ear recognition that uses a HE technique for image enhancement and a fusion of HOG, LBP, and PCA techniques to extract features. Following that, LDA is in the classification stage. In terms of performance and accuracy, the combination of these two techniques could be promising.

CHAPTER 4
Research Methodology

4.1. Introduction

A complete ear recognition system can be split into four stages: image acquisition, pre-processing, feature extraction, and ear matching. Chapter 2 presented an overview of some techniques that will be used in the work.

In this Chapter, the main methodology of this work is illustrated in Figure 4.1. Then a detailed description of each stage is provided. Firstly, brief information about the database used in this study is presented. Next, the pre-processing stage of ear recognition is explained in Section 4.3.

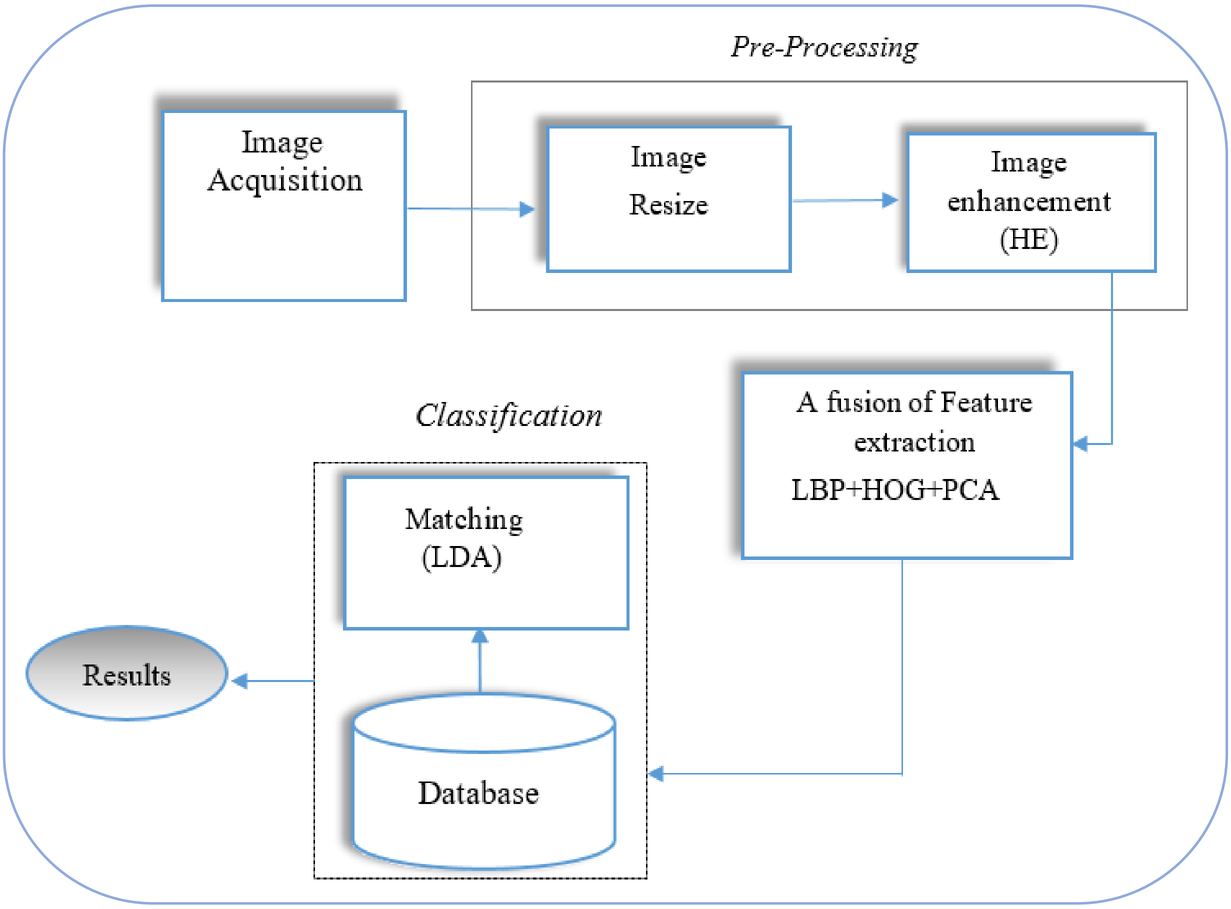

Figure 4.1: Block Diagram of Ear Recognition System

4.2. Image Acquisition

The first step is to choose or collect a dataset of the images. The IIT Delhi database (KUMAR, 2007) was chosen for this work. The database consists of ear images of students and staff at IIT Delhi University. All images were captured remotely under various indoor lighting conditions with no major occlusions.

IITD I contains 493 images of 125. At least this database represents every person with 3 to 6 images. The age group of subjects ranges from 14 to 58 years. The 493 images have been sequentially numbered for every user with an integer number. The resolution of these images is 272 × 204 pixels and all these images are available in jpeg format. In addition to the original images, this database also provides the automatically normalized and cropped ear images of size 180 × 50 pixels. Figure 4.2 shows samples from the IIT Delhi database.

Figure 4.2: Samples from IITD database

4.3. Image Pre-processing

It is the first task performed on the images. The objective of pre-processing is to modify the source image in order to facilitate feature extraction and improve the recognition rate. The database targeted in this paper contains automatically segmented and normalized images corresponding to raw images from 125 subjects. Figure 4.3 shows the ear images in the database, where the whole images are automatically normalized and cropped to 180 × 50 pixels. All of these images are available in bmp format.

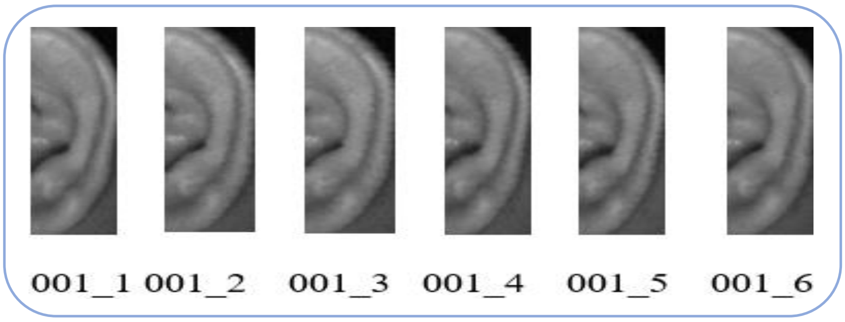

Figure 4.3: Ear Image Normalization and Segmentation

4.3.1. Ear Image Resizing

For the purpose of image normalization, we noticed that some of the researchers in the literature, such as (Anwar et al., 2015, Omara et al., 2018, Ahila Priyadharshini et al., 2021, Alhanjouri and ISSN, 2018), who used the same database with 180 × 50 pixel image size, changed the image size to extract some local discriminative features. This change led to good results, and accordingly, we resized the images several times by decreasing 10 in each attempt. Firstly, resize all images from 180 × 50 to a fixed size. After that, we changed it to another size and so on. We observed that the results improved and the best result was at 150 × 50. Likewise (Sarangi et al., 2019), who used the same size and got good results. Figure 4.4 shows some ear images before and after resizing.

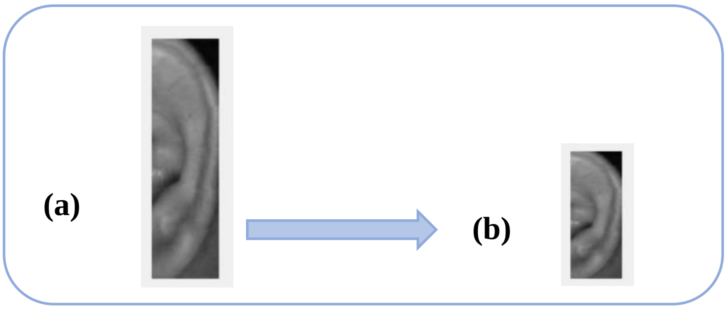

Figure 4.4: Ear image before and after resizing

4.3.2. Image Enhancement (Histogram Equalization)

A histogram equalization technique is applied for normalizing ear images to minimize the effect of non-uniform lighting and obtain a well-distributed texture image. The histogram equalization technique is a cumulative distribution transformation function (Hadi et al., 2021). The histograms of the input images are adjusted using the HE method to produce enhanced images. This method produces an image with a uniform distribution that has better contrast than the original. The image can be changed to make use of all of the available pixel intensities. The formula (4.1) represents the intensity of a pixel.

$$s = T(r) = \int^{\omega} p_r(r)d\omega = CDF_i \qquad (4.1)$$

Where,

- r, s are the input and output image intensity values, respectively.
- $pr(r)$ is the probability distribution function of the original image, which is obtained after normalizing the histogram of the original image such that the area is equal to 1.
- Because the transformation to achieve histogram equalization is actually the CDF of the original image, CDF_i stands for the cumulative distribution function of the

input image.

Figure 4.5 shows the unwrapped texture and texture image after enhancement, respectively (a) The image before enhancement (b) The image after it has been enhanced.

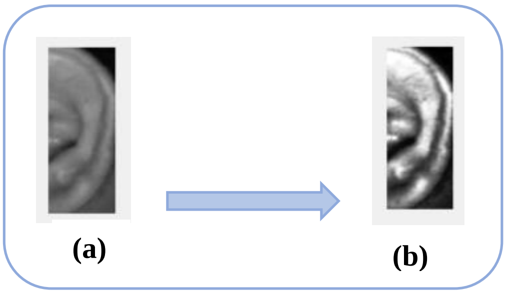

Figure 4.5: The Enhancement Process

4.4. Feature extraction

The most discriminating information existing in the ear pattern will be extracted in this step. Feature extraction converts an image into a set of vectors that the matcher can use to improve classification. The feature in the human ear has significant advantages over other biometric technologies such as iris, fingerprints, face, and retinal scans. The ear is larger than the iris (Arnia and Pramita, 2011) and fingerprint, and unlike them, the image acquisition of the human ear is quite simple because it can be taken from a distance without needing individual cooperation (Hurley et al., 2005). The human ear has a richness of features and is more dependable than the face since the structure of the ear does not vary with age or facial expressions (Muntasa et al., 2011). The anatomy of the human ear is given in Figure 2.1 in Chapter 2. It shows the standard features of the human ear. It has been found that no two ears are exactly the same, even those of identical twins (Victor et al., 2002) (Chang et al., 2003). Different algorithms have been used at this stage by researchers. Therefore, this study assumes that the combination of two feature extractions, HOG and LBP, will help the classification model yield higher results.

4.4.1. Histogram of Oriented Gradients (HOG)

The HOG approach is applied to the ear normalized IITD I database. This technique is widely used in the field of digital signal processing applications. The histogram of oriented gradients (HOG) is a feature descriptor for object detection in computer vision and image processing. The technique counts the number of times a gradient orientation occurs in a certain area of an image. Edge orientation histograms, scale-invariant feature transform descriptors, and shape contexts are all comparable techniques, but this one differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for enhanced accuracy.

- **The process of HOG**

  Step 1: Pre-process the Data

  Pre-processing data is an important step in any machine learning project, and dealing with images is no exception. The image must be pre-processed to reduce the width to height ratio. (Singh, 2019).

  Step 2: (direction x and y)

  The following step is to compute the gradient for every pixel in the image. Gradients are minor changes in x and y directions. The pixel values for this patch are then obtained.

  After that, the gradients in both the x and y directions were calculated separately. The process is utilized for each of the image's pixels. The magnitude and orientation would then be determined using these values as a next step.

  Step 3: Calculate the Magnitude and Orientation

  Determine the magnitude and direction for each pixel value using the gradients calculated in the previous step. The Pythagoras theorem will be used in this stage.

  So now we get the entire gradient (magnitude) and the orientation for each pixel value (direction). Using these gradients and orientations, we can generate the histogram.

  Step 4: Calculate the Histogram of Gradients

  The histograms generated by the HOG feature descriptor do not generate the entire image. Instead, the image is divided into cells, with each cell receiving its own histogram of directed gradients. The next step is to normalize the histogram once we've generated the HOG for all of the patches in the image.

  Step 5: Normalize Gradients

  By taking blocks and normalizing the gradients, you may reduce the lighting variation. A normalized vector would be the end product.

  Step 6: Features for the complete image

  This is the last stage in the process of generating HOG features for the image. We've built features for blocks of the image so far. Now we'll put them all together to create the final image's features.

The HOG is based on the accumulation of gradient directions via the image pixel for a particular region called "Cell." The following is a one-dimensional histogram construction that offers a concatenation of features that can be used to feed the classification process. Figure 4.6 shows the main mechanism of the HOG technique, with an example of a cell size of four pixels and eight orientation bins for the histograms of a cell, divided into six steps as shown. Assume that G in equation (4.2) refers to the grayscale function that has been used for describing and analysing images. Furthermore, each image will be divided into a group of cells with a size of N × N pixels. Step (1) in Figure 4.6. shows the image dividing process to a set of cells, each cell divided into 8 × 8 pixels. The gradient orientation (i.e., θk,r, r, r) for every pixel is calculated as shown in equation (4.2). Steps 2 and 3 illustrate the gradient orientation processes (Mutar et al., 2020). The orientations θij = 1 ... N2 for the same cell j are accumulated and quantized into an M-bins histogram shown in steps 4 and 5. In the last step (6), the whole obtained histograms will be arranged and concatenated into an HOG histogram as a final outcome of the feature extraction process.

$$\theta_{k,r} = \tan^{-1}\frac{G(k,r+1) - G(k,r-1)}{G(k+1,r) - G(k-1,r)} \quad (4.2)$$

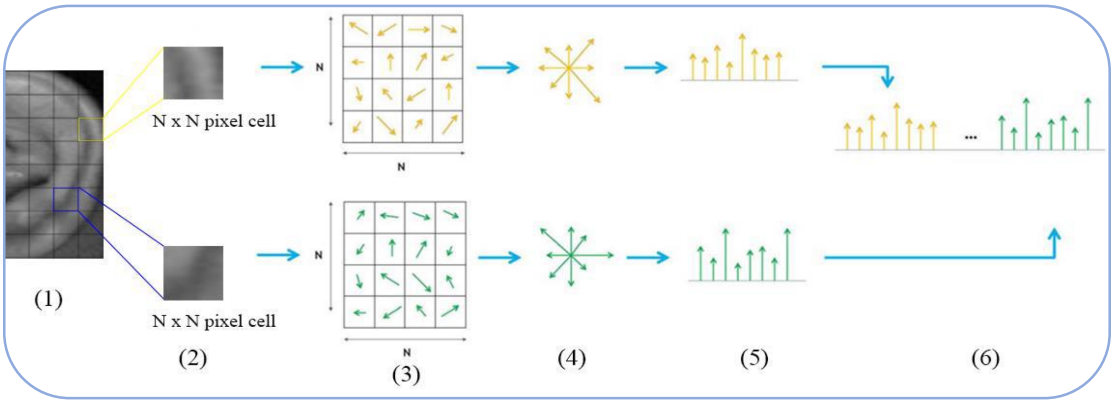

Figure 4.6: Depicts HOG Features Extraction Steps applied on a IITD I Ear Image

4.4.2. Local Binary Pattern (LBP)

The LBP technique is applied to the ear-normalized data in the IITD I database. The LBP texture analysis operator is a grey-scale invariant texture measure derived from the general concept of texture in a local neighborhood. It's a powerful texture description technique in real-world applications, with features including discriminative power, computing simplicity, and tolerance for monotonic grey-scale changes.

- **The process of LBP**

The main steps of the original version of the local binary pattern are summarized below:

Step 1: The LBP runs on a 3-pixel image block.

Step 2: To obtain a label for the centre pixel, the pixels in this block are thresholded by its centre pixel value, multiplied by powers of two, and then summed.

Step 3: If the result is negative, the pixel will be encoded with 0, otherwise it will be encoded with 1.

Step 4: A binary number is revealed for each pixel by concatenating the 8 binary results to construct a number in a clockwise direction, beginning with its top-left neighbor.

Step 5: The pixel is labeled using the decimal value derived from the binary number. LBP codes refer to the resulting binary numbers.

The definition of the LBP is described as Formula 4.3.

$$LBP_{P,R} = \sum_{P=0}^{P-1} s\left(g_p - g_c\right) 2^p, s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \quad (4.3)$$

The gray value of the center pixel is gc, while the gray value of its vicinity is gp. The number of LBP patterns is determined by the number of neighbors involved, which is $2p$. R is the neighborhood's radius, which determines the size of the neighborhood.

The LBP technique can capture the complete ear by scanning each pixel's neighbors, as seen in Figure 4.7. These pixels will be computed, and the LBP will extract all of the features of this ear to form a one-dimensional array of patterns that can later be classified using any classification algorithm.

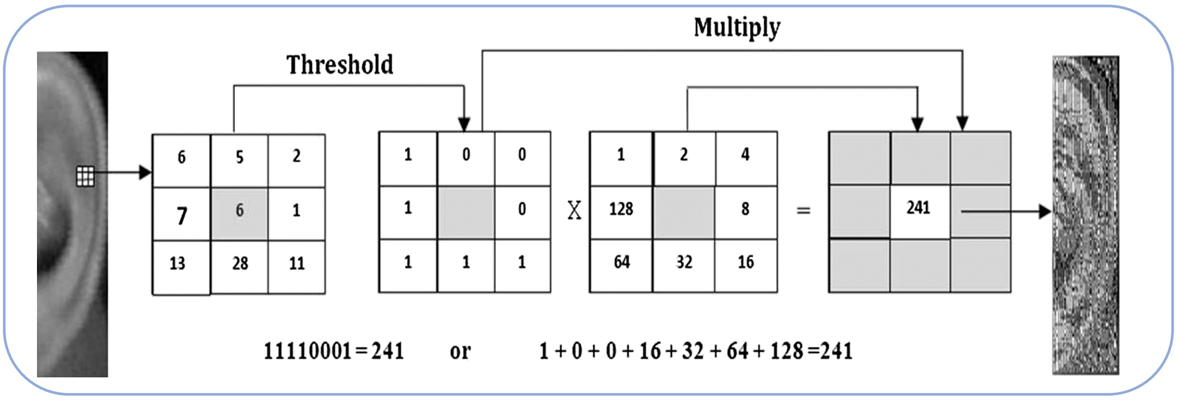

Figure 4.7: Calculation of the Local Binary Patterns (LBPs) operator applied on a Normalized IITD I Ear Image

4.4.3. Principal Component Analysis (PCA)

PCA is used to detect patterns and variations in a dataset. It is also a dimensionality reduction technique, which makes it an attractive choice for ear recognition even with the advances in computer technology. The idea of PCA is to reduce the feature vector's dimension (Kumari et al., 2019).

- **The process of PCA**

The main steps described by (Jaadi, 2021), provide a comprehensive and easy-to-understand explanation of Principal Component Analysis (PCA).

Step 1: Standardization:

This step is used to standardize the range of continuous beginning variables so that they all contribute equally to the analysis. The importance of standardization prior to PCA is due to the latter's sensitivity to the variances of the initial variables. That is, if the ranges of starting variables differ significantly, the variables with bigger ranges will outnumber those with smaller ranges. Subtracting the mean and dividing by the standard deviation for each value of each variable can be done mathematically as shown in equation 4.4.

$$z = \frac{value - mean}{standard\ deviation} \quad (4.4)$$

All variables will be transformed to the same scale once standardization is completed.

Step 2: Covariance Matrix Computation:

The purpose of this step is to figure out how the variables in the input data set differ from the mean in relation to one another, or to discover if there is any link between them. Because variables might be highly connected to the point where they include duplicated data. We compute the covariance matrix in order to find these correlations.

Step 3: Compute the Eigenvectors and Eigenvalues of the Covariance Matrix to Identify the Principal Components:

The linear algebra concepts of eigenvectors and eigenvalues are needed to compute the principal components of the data from the covariance matrix. Principal components, in geometric terms, are the data directions that explain the greatest amount of variance, or the lines that capture the most information in the data. The relationship between variance and information in this case is that the greater the variance carried by a line, the greater the dispersion of data points along it, and the greater the dispersion along a line, the more information it carries.

Step 4: Feature Vector:

In this step, decide whether to preserve all of these components or to discard those with low eigenvalues, and create a matrix of vectors called Feature vector with the remaining ones. So, the feature vector is just a matrix with the eigenvectors of the components we want to keep as columns.

Step 5: Recast the Data Along the Principal Components Axes:

The output of this step, which is the last one, is to reorient the data from the original axes to the ones indicated by the principal components using the feature vector produced using the eigenvectors of the covariance matrix (hence the name Principal

Components Analysis). This is accomplished by multiplying the original data set's transpose by the feature vector's transpose.

As a result, PCA is applied to the entire size of the feature vector extracted by HOG and LBP. It's used for dimension reduction, making it available to only those with a lot of variation. A row vector is formed by rowing through each pixel of an image to begin with. All of the row vectors are combined to produce a matrix. If the feature vector is reduced, the classifier will be able to provide better results.

4.5. Classification (Linear Discriminant Analysis)

In this study, the LDA was chosen as a metric for recognition since this classifier is trained and tested by features extracted from the ear pattern. The LDA classifier is trained several times with the set of ear images and then tested with other sets of ear images.

The goal of the LDA technique is to project the original data matrix onto a lower-dimensional space. Because of its large number of features or dimensionality, the LDA technique has been used in biometrics, agriculture, and medical applications. To achieve this goal, three steps were required.

- The between-class variance or between-class matrix is used to compute the separability between different classes (i.e. the distance between the means of different classes).
- The within-class variance, also known as the within-class matrix, is computed as the distance between the mean and the samples of each class.
- The final step is to create a lower-dimensional space that maximizes between-class variance while minimizing within-class variance. (Tharwat et al., 2017).

Three steps are needed to be performed in order to perform LDA technique:

Step 1: The first step is to compute the separability of different classes (i.e., the distance between their means), also known as the between-class variance or between-class matrix.

Step 2: The second step, the within-class variance or within-class matrix, is calculated by calculating the distance between the mean and the samples of each class.

Step 3: Finally, construct the lower dimensional space that maximizes between-class variance while minimizing within-class variance.

4.6. Summary

In this Chapter, the methodology of the research is provided, where HE is utilized to improve the segmented ear image quality after resizing. This process makes it easier to carry out the following steps. A fusion of HOG and LBP is applied for the purpose of extracting the features from the ear image. Furthermore, PCA was applied to the extracted feature to reduce the dimension of the feature vector and then used an LDA classifier to test the performance of ear recognition. The experimental results of these techniques will be provided in the next chapter.

CHAPTER 5
Results and Discussions

5.1. Introduction

In the previous Chapters, the basic concept of the ear recognition system has been introduced. Furthermore, HE has improved the normalized ear image. The features extracted from ear images were extracted using a fusion of HOG and LBP.

In this Chapter, the empirical testing of this system is explained by running and comparing a series of experiments. HOG and LBP have been used separately for extracting the features. Section 5.3.1 illustrates the results obtained in experiment 1 using the LBP technique. While Section 5.3.2 shows the results of the system in experiment 2, which used the HOG technique, Section 5.3.3 shows the fusion between LBP and HOG in experiment 3 without using the PCA technique. Finally, the fusion of LBP and HOG with PCA was applied to all the data. The results are presented in Section 5.3.4 in experiment 4. All the experiments are applied with the LDA classifier. The experiments were carried out using ear images obtained from the IITD I data set.

5.2. Hardware and Software Environments

All experiments are carried out on a laptop computer equipped with an Intel(R) Core(TM) i7-7500U CPU @ 2.70GHz, 2.90GHz, 16.0 GB of RAM, and the Windows 10 operating system, using MATLAB version R2021a.

Various techniques are applied to the IITD I data set; HOG and LBP are used for feature extraction, and linear discriminant analysis is used for classification. Then, each algorithm is applied separately, and then applied as a fusion of two techniques (HOG and LBP), followed by applying PCA to the feature vector that is produced. The obtained results are compared with previous work to determine that our approach is the best among the methods that have been used in the literature.

5.3. Experimentation

The evaluation of the performance of the proposed approach is based on training and testing the adopted classifier on the whole images of the IITD I dataset. This work analysed 2 images of every sample as a training set; so, the total number of images in the training dataset is 250. While the remaining number of images in the testing set is 243. Every sample in the testing set has from 1 to 4 images. The variety is because the dataset is imbalanced. Originally, each sample had 3 to 6 images in the IITD I database.

The implementation of the proposed approach involves four different ways to figure out the accuracy level. In the experiments, we evaluated the optimal parameters of the proposed method, which have an effect on recognition performance. In the initial stage, we resize these images, then we apply histogram equalization enhancement to them. In all the experiments, we observed that in each experiment where we added the histogram equalization, the results improved; Table 5.5 shows the results of each experiment. We have applied the proposed approach through several experiments on different image sizes. Firstly, we resize all images from 180 × 50 to 170 × 50 pixels. After that, we changed it to 160 × 50 pixels, and so on. We observed that the results improved and the best result was at 150 × 50.

5.3.1. Experiment 1

The first experiment in table 5.1 investigates the proficiency of classifying HOG features with no reduction. These features are classified using LDA. This experiment has yielded results with a 92.59% accuracy rate. We obtained an enhanced version of the image by applying the histogram equalization technique to improve ear image contrasts. (Hussein et al., 2021), which has a great effect on the recognition rate, the percentage increased to 93.42%. LDA is used as a classifier.

Table 5.1: The accuracy of HOG

Feature Extraction Technique	Classifier	Accuracy
HOG	LDA	92.59%
HE+HOG	LDA	93.42%

5.3.2. Experiment 2

In the second experiment, LBP features were only classified using the LDA classifier, which is illumination invariant and has a low computational load. It produced an imprecise result, as shown in (Resmi and Raju, 2019). In table 5.2, the accuracy was 95.06% before using histogram equalization. After using it, the results achieved a positive rise in accuracy rate of 95.88%.

Table 5.2: The accuracy of LBP

Feature Extraction Technique	Classifier	Accuracy
LBP	LDA	95.06%
HE+LBP	LDA	95.88%

5.3.3. Experiment 3

The third experiment evaluates the proficiency of combining HOG and LBP together without using PCA. while LDA for classification. The accuracy reached 94.65% and increased even further to 94.83% with histogram equalization enhancement. The results are illustrated in table 5.3.

Table 5.3: The accuracy of HOG+LBP

Feature Extraction Technique	Classifier	Accuracy
HOG+LBP	LDA	94.65%
HE+HOG+LBP	LDA	94.83%

5.3.4. Experiment 4

The last experiment shows the effectiveness of the proposed approach. It applies LBP and HOG techniques while reducing the feature space using PCA. It is used to reduce the number of features to only those with a huge variation among them with an LDA classifier. As shown in table 5.4, the accuracy was 95.47% without using histogram equalization enhancement. With using it, the level of accuracy is slightly increased to 96.30%.

Table 5.4: The accuracy of HOG+LBP+PCA

Feature Extraction Technique	Classifier	Accuracy
HOG+LBP+PCA	LDA	95.47%
HE+HOG+LBP+PCA	LDA	96.30%

Table 5.5 shows that using the histogram equalization enhancement technique in each experiment increases the recognition accuracy since it is used to obtain a contrast enhanced version of the original image (Hassaballah et al., 2019). On the other hand, a slight decrease in the performance was noted in the fusion of HOG and LBP techniques without PCA compared with LBP alone. The decreases in accuracy are due to the noise that is caused by the fusion between the two techniques. Herein lies the effectiveness of using the PCA technique on the extracted vector to provide a holistic description of the sample images while reducing the dimensionality of the data without much loss of information.

Table 5.5: The Summary of Recognition Accuracies

Feature Extraction Technique	Accuracy without HE	Accuracy with HE
HOG	92.59%	93.42%
LBP	95.06%	95.88%
LBP+HOG	94.65%	94.83%
LBP+HOG+PCA	95.47%	96.30%

5.4. Comparison with previous approaches

The comparison between the previous approaches and the proposed approach is provided in Table 5.6. It contains some studies using the same database that has been utilized in our study, the IITD I database, with the same number of ear images in training and testing stages. Our proposed approach had the highest results for ear recognition compared with the literature. Furthermore, to validate the performance of the proposed approach in this study, it was compared to previous works. (Kohlakala and Coetzer, 2021) used DRT as a feature extraction with a Euclidean Distance classifier, we can see a slight increase in our accuracy compared with their work, the performance of their approach was 96.06%. Where (Resmi and Raju, 2019) used LBP and PCA with the KNN classifier, herein lies the effectiveness of the PCA technique that they used to reduce dimensions, as a result, a noticeable improvement in performance appeared with 95.3%. Moreover, the efficiency of combining the LBP with the DTP technique to boost the performance is revealed here, KNN as a classifier (Sivanarain and Viriri, 2020), with a 95.88% recognition rate.

Table 5.6 The Best Result of our approach Comparing with other Studies

Reference	Approach		Accuracy
	Feature extraction	Classifier	
Proposed System	**A fusion of (LBP+HOG+PCA)**	**LDA**	**96.30%**
(Kohlakala and Coetzer, 2021)	DRT	Euclidean distance	96.06%
(Resmi and Raju, 2019)	LBP+PCA	K-NN	95.3%
(Sivanarain and Viriri, 2020)	LBP+DTP	K-NN	95.88%

As a comparison of the best results yielded in this work with some other studies which analyzed the same dataset using other techniques for feature extraction and classification, the difference between the performance of the proposed methods and the previous methods depends on the techniques that are used in the feature extraction and classification stages. It can be seen that the best result of this work is the highest. Here we can confirm the effectiveness of our proposed approach compared with the promising result. The performance is increased by adding more than one feature extraction.

5.5. Summary

In this chapter, the proposed experiments were examined in order to recognize the ear from the ear image collected from the IITD I database. The experiments were implemented using a fusion of two techniques: HOG and LBP, followed by PCA for dimension reduction. After that, we used LDA to match the two ear templates.

Through the experiments, some important points can be observed, such as:

- Resizing images helps improve recognition performance.
- The use of HE has further improved accuracy with each technique.
- The performance of the fusion of HOG, LBP, and PCA was better than HOG and LBP alone.
- Increasing the number of images in the training set leads to obtaining high performance.

CHAPTER 6
Conclusion and future work

This Chapter aims to review the proposed solution of this study in relation to the proposed ear recognition system. The study has discussed developments carried out in the implementation of ear recognition systems and proposed solutions in the field of recognition systems. This research has considered ear recognition as a biometric technology because the ear is a very robust part of authentication. Once an individual becomes an adult, the shape of their ears does not sharply change during life. Ear biometrics authentication is a form of security which has an important role in many fields, such as the forensics field, to define the unique physical characteristics for distinguishing a person's identity, for identification or verification tasks. An efficient and reliable approach has been successfully proposed for automatic human ear recognition.

The proposed approach presents four levels of results. The lowest outcome of 93.42% has been obtained when LDA classifies the features of HOG. The result of the classification of LBP is higher. LDA has succeeded in obtaining 95.88% of the correct classification rate. The combination of the features of HOG and LBP helped the classifier to present 94.83% of the accuracy rate. The classification level of reducing these two features using PCA has reached an accuracy rate of 96.30%. Eventually, the best result of this work is higher than other studies that considered the same database of images (IITD I) using different techniques of feature extraction and classification.

6.1. Summary of work

- Firstly, after resizing the segmented ear image, the normalized ear image is processed to improve the quality of the image. For this reason, the HE technique is carried out on the normalized ear image.
- Next, we used two techniques, HOG and LBP, separately, and then a fusion of HOG and LBP was applied to extract the features. After that, we applied the PCA technique to the feature vector that was the outcome of the HOG and LBP fusion.
- Finally, the LDA classifier has been utilized to classify the ear features.
- Furthermore, these techniques are applied to the IITD I data set, which has a main role in the evaluation of the proposed system performance.

- The dataset is divided into a training set and a testing set. The classification results of the ear recognition system depend on the feature extraction algorithm. The results in Chapter 5 also show that feature extraction can be a difficult stage in an ear recognition system due to the effects it has on ear matching.
- The empirical testing of this proposed system, conducted using MATLAB version R2021a, showed that the fusion of HOG and LBP techniques followed by PCA with LDA as a classifier achieved good results.

6.2. Future work

Research in the area of ear recognition systems is growing fast. Due to the continued spread of using recognition systems in several applications, the current work focuses on the feature extraction stage. In addition, various techniques have been applied to achieve optimal performance in this work. However, there are still a number of issues that need to be addressed. Some points about future work are included below:

- Apply the proposed system to different databases such as USTB I, USTB II, and AMI and compare the obtained results with the results in this research.
- Combine different techniques to extract the features of the ear, such as LBP with Discrete Wavelet Transform (DWT), and compare the results with the results obtained in this research.
- In order to produce impressive results, deep learning may be used in ear classification.

Referencing

Abaza, A. & Harrison, M. A. F. (2013). Ear recognition: a complete system. Biometric and Surveillance Technology for Human and Activity Identification X, International Society for Optics and Photonics, 87120N.

Abaza, A. & Ross, A. (2010).Towards understanding the symmetry of human ears: A biometric perspective. Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), IEEE, 1-7.

Ahila Priyadharshini, R., Arivazhagan, S. & Arun, M. J. A. I. (2021). A deep learning approach for person identification using ear biometrics. 51, 2161-2172.

Ahonen, T., Matas, J., He, C. & Pietikäinen, M.(2009). Rotation invariant image description with local binary pattern histogram fourier features. Scandinavian conference on image analysis,. Springer, 61-70.

Alberink, I. & Ruifrok, A. J. F. S. I. (2007). Performance of the FearID earprint identification system. 166, 145-154.

Alhanjouri, M. J. I. J. O. I. R. I. C. S. & Issn, T. (2018). Improved HMM by Deep Learning for Ear Classification. 2347-5552.

Alkhraisat, H. J. I. J. O. C. S. & Security, I. (2017). Ear Biometric System Using Speeded-up Robust Features and Principal Component Analysis. 15.

Anwar, A. S., Ghany, K. K. A. & Elmahdy, H. J. P. C. S. (2015). Human ear recognition using geometrical features extraction. 65, 529-537.

Arnia, F. & Pramita, N. J. T. (2011). Enhancement of Iris recognition system based on phase only correlation. 9, 387-394.

Asadi, S., Rao, C. & Saikrishna, V. J. I. J. O. C. A. (2010). A comparative study of face recognition with principal component analysis and cross-correlation technique. 10, 17-21.

B. S. El-Desouky, M. E.-K., M. Z. Rashad, Mahmoud M. Eid (2012). Ear Recognition And Occlusion. International Journal of Computer Science & Information Technology (IJCSIT), 4, 6.

Benzaoui, A., Hadid, A. & Boukrouche, A. J. J. O. E. I. (2014). Ear biometric recognition using local texture descriptors. 23, 053008.

Benzaoui, A., Hezil, N. & Boukrouche, A. (2015). Identity recognition based on the external shape of the human ear. International Conference on Applied Research in Computer Science and Engineering (ICAR). IEEE, 1-5.

Burge, M. & Burger, W. (1996). Ear biometrics. Biometrics. Springer.

Burge, M. & Burger, W. (1997). Ear biometrics for machine vision. 21st workshop of the Austrian association for pattern recognition. Citeseer, 275-282.

Burge, M. & Burger, W. (1998). Using ear biometrics for passive identification, Citeseer.

Burge, M. & Burger, W.(2000) Ear biometrics in computer vision. Proceedings 15th International Conference on Pattern Recognition. ICPR. IEEE, 822-826.

Carreira-Perpinan, M. J. M. T. (1995). Faculty Of Informatics, Technical University Of Madrid. Compression neural networks for feature extraction: Application to human recognition from ear images.

Chang, K., Bowyer, K. W., Sarkar, S., Victor, B. J. I. T. O. P. A. & Intelligence, M. (2003). comparison and combination of ear and face images in appearance-based biometrics. 25, 1160-1165.

D. Frejlichowski, N. T. (2010). Image Analysis and Recognition. 7th International Conference, ICIAR, Póvoa devarzin, portugal, june 21–23, proceedings, part ii, springer berlin heidelberg, 227–234.

Dalal, N. & Triggs, B. (2005). Histograms Of Oriented Gradients For Human Detection. IEEE computer society conference on computer vision and pattern recognition (CVPR'05). Ieee, 886-893.

Dodge, S., Mounsef, J. & Karam, L. J. I. B. (2018). Unconstrained ear recognition using deep neural networks. 7, 207-214.

E. GONZALEZ-SANCHEZ, B. D. L. O. (2008). Ph.D. thesis,Universidad de Las Palmas de Gran Canaria.

Emeršič, Ž., Gabriel, L. L., Štruc, V. & Peer, P. J. A. P. A. (2017)a. Pixel-wise ear detection with convolutional encoder-decoder networks.

Emeršič, Ž., Štepec, D., Štruc, V. & Peer, P. J. A. P. A. (2017)b. Training convolutional neural networks with limited training data for ear recognition in the wild.

Emeršič, Ž., Štruc, V. & Peer, P. J. N. (2017)c. Ear recognition: More than a survey. 255, 26-39.

Eyiokur, F. I., Yaman, D. & Ekenel, H. K. J. I. B. (2018). Domain adaptation for ear recognition using deep convolutional neural networks. 7, 199-206.

Hadi, R. A., George, L. E. & Ahmed, Z. J. J. T. (2021). Automatic human ear detection approach using modified adaptive search window technique. 19, 507-514.

Hamdany, A. H. S., Ebrahem, A. T., Alkababji, A. M. J. T. E.(2021). Earprint Recognition using deep learning technique. Computing, Electronics & Control 19, 432-437.

Hansley, E. E., Segundo, M. P. & Sarkar, S. J. I. B. (2018). Employing fusion of learned and handcrafted features for unconstrained ear recognition. 7, 215-223.

Hassaballah, M., Alshazly, H. A. & Ali, A. A. J. E. S. W. A. (2019). Ear recognition using local binary patterns: A comparative experimental study. 118, 182-200.

Houcine, B., Hakim, D., Amir, B. & Hani, B. A.(2015). Ear recognition based on multi-bags-of-features histogram. 3rd International Conference on Control, Engineering & Information Technology (CEIT). IEEE, 1-6.

Huang, G. B., Mattar, M., Berg, T. & Learned-Miller, E.(2008). Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. Workshop on faces in'Real-Life'Images: detection, alignment, and recognition.

Hurley, D. J., Nixon, M. S., Carter, J. N. J. C. V. & Understanding, I. (2005). Force field feature extraction for ear biometrics. 98, 491-512.

Hussein, Y. A., Sahan, A., Al-Itbi, A. S. A. J. I. J. O. A. I. S. C. & Applications, I. (2021). An Intelligent Ear Recognition Technique. 13.

Iannerelli, A. (1989). Ear identification, forensic identification series. California: Paramount Publishing Company, Fremont.

Jaadi, Z. J. R. J. (2021). A step-by-step explanation of Principal Component Analysis (PCA). 7.

Klare, B. F., Klein, B., Taborsky, E., Blanton, A., Cheney, J., Allen, K., Grother, P., Mah, A. & Jain, A. K.(2015). Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a. Proceedings of the IEEE conference on computer vision and pattern recognition. 1931-1939.

Kohlakala, A. & Coetzer, J. J. S. A. R. J. (2021). Ear-based biometric authentication through the detection of prominent contours. 112, 89-98.

Kumar, A. 2007. IIT Delhi ear database.

Kumar, A. & WU, C. J. P. R. 2012. Automated human identification using ear imaging. 45, 956-968.

Kumar, N., Berg, A. C., Belhumeur, P. N. & Nayar, S. K. (2009). Attribute and simile classifiers for face verification. IEEE 12th international conference on computer vision. IEEE, 365-372.

Kumari, D., Dub, S. S., Gupta, B. & Fcsc, J. T. (2019). Performance Evaluation of PPCA-CPD Human Ear Recognition Application on AMI and USTB Ear Image Datasets.

L., E. R. (2002). University of Science & Technology Beijing, Introduction to USTB ear image databases.

Lammi, H.-K. J. L. U. O. T. (2004). Ear biometrics.

Li, C., Zhou, W. & Yuan, S. J. T. V. C. (2015). Iris recognition based on a novel variation of local binary pattern. 31, 1419-1429.

Lu, H., Plataniotis, K. N. & Venetsanopoulos, A. (2013). Multilinear subspace learning: dimensionality reduction of multidimensional data, CRC press.

Lu, X. & Jain, A. K. (2004).Ethnicity identification from face images. Proceedings of SPIE. Citeseer, 114-123.

Luo, J., Mu, Z. & Wang, Y. (2008). Ear recognition based on force field feature extraction and convergence feature extraction. Seventh International Symposium on Instrumentation and Control Technology: Sensors and Instruments, Computer Simulation, and Artificial Intelligence. SPIE, 541-546.

Meijerman, L., Thean, A. & Maat, G. J. W. E. O. F. S. (2009). Earprints: interpretation of.

Muntasa, A., Sirajudin, I. A. & Purnomo, M. H. J. T. I. J. O. E. E. (2011). Appearance global and local structure fusion for face image recognition. 9, 125-132.

Mutar, M. H., Ahmed, E. H. & Majid Razaq Mohamed Alsemawi, H. J. S. S. T. (2020). Ear Recognition System Using Random Forest and Histograms of Oriented Gradients Techniques. 63, 8740-8748.

Ng, H.-W. & Winkler, S. (2014). A data-driven approach to cleaning large face datasets. IEEE international conference on image processing (ICIP). IEEE, 343-347.

Ojala, T., Pietikainen, M., Maenpaa, T. J. I. T. O. P. A. & Intelligence, M. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. 24, 971-987.

Omara, I., Li, F., Hagag, A., Chaib, S. & Zuo, W. J. I. J. O. C. S. I. (2016). Ear recognition using a novel feature extraction approach. 46.

Omara, I., Li, X., Xiao, G., Adil, K. & Zuo, W.(2018). Discriminative local feature fusion for ear recognition problem. Proceedings of the 8th International Conference on Bioscience, Biochemistry and Bioinformatics. 139-145.

Omara, I., Wu, X., Zhang, H., Du, Y. & Zuo, W.(2017). Learning pairwise SVM on deep features for ear recognition. IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS). IEEE, 341-346.

Pflug, A. & Busch, C. J. I. B. (2012). Ear biometrics: a survey of detection, feature extraction and recognition methods. 1, 114-129.

Pflug, A., Paul, P. N. & Busch, C. (2014). A comparative study on texture and surface descriptors for ear biometrics.International carnahan conference on security technology (ICCST). IEEE, 1-6.

Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., Marques, J., Min, J. & Worek, W.(2005) .Overview of the face recognition grand challenge. IEEE computer society conference on computer vision and pattern recognition (CVPR'05). IEEE, 947-954.

Pietikäinen, M., Hadid, A., Zhao, G. & Ahonen, T. (2011). Local binary patterns for still images. Computer vision using local binary patterns. Springer.

Rasika B. Naik, R. D., Deepak Desai, Hitesh Balani, Kashmira Raskar. (2018). Human Recognition System using Ear biometrics. International Journal of Scientific & Engineering Research 9, 4.

Resmi, K. & Raju, G. (2019). A Novel Approach to Ear Recognition based on LBP and PCA.

Sarangi, P. P., Mishra, B. S. P., Dehuri, S. J. M. T. & Applications .(2019). Fusion of PHOG and LDP local descriptors for kernel-based ear biometric recognition. 78, 9595-9623.

Singh, A. J. M., Analytics Vidhya .(2019). Feature engineering for images: a valuable introduction to the HOG feature descriptor. 4.

Sivanarain, K. & Viriri, S.(2020). Ear recognition based on local texture descriptors. 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC). IEEE, 1-11.

Sowmyalakshmi, R. & Girirajkumar, S. J. A. M. (2017). A New Technique for Computationally Efficient Human Recognition. 11, 457-464.

Srivastava, G. & Rawat, T. K. (2013).Histogram equalization: A comparative analysis & a segmented approach to process digital images.Sixth International Conference on Contemporary Computing (IC3). IEEE, 81-85.

Tharwat, A., Gaber, T., Ibrahim, A. & Hassanien, A. E. J. A. C. (2017). Linear discriminant analysis: A detailed tutorial. 30, 169-190.

Tiwari, S., Jain, S., Chandel, S. S., Kumar, S. & Kumar, S. (2016).Comparison of adult and newborn ear images for biometric recognition. Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC). IEEE, 421-426.

Ummer Akber Tali, A. S. (2015). Ear Pattern Recognition and Compression Using Surf and SVM-A Review. International Journal of Innovative Research in Computer and Communication Engineering, 3, 5.

Victor, B., Bowyer, K. & Sarkar, S. (2002). An evaluation of face and ear biometrics. Object recognition supported by user interaction for service robots. IEEE, 429-432.

Yi, D., Lei, Z., Liao, S. & Li, S. Z. J. A. P. A. (2014). Learning face representation from scratch.

Yuan, L. & Chun Mu, Z. J. P. R. L. (2012). Ear recognition based on local information fusion. 33, 182-190.

Zarachoff, M., Sheikh-Akbari, A. & Monekosso, D.(2018). 2D multi-band PCA and its application for ear recognition. IEEE International Conference on Imaging Systems and Techniques (IST). IEEE, 1-5.

Zarachoff, M. M., Sheikh-Akbari, A. & Monekosso, D. J. I. A. (2021). Non-Decimated Wavelet based Multi-Band Ear Recognition using Principal Component Analysis.

Zhang, Y., Mu, Z., Yuan, L. & Yu, C. J. I. B. (2018). Ear verification under uncontrolled conditions with convolutional neural networks. 7, 185-198.

مزيج من استخلاص الميزات للتعرف على القياسات الحيوية للأذن البشرية

قدمت من قبل:

بثينة فرج محمد قرقوم

تحت إشراف:

د .أحمد الأوجلي

الخلاصة

زاد الطلب على المصادقة الأكثر أمانًا في عدة مناسبات. يعد استغلال القياسات الحيوية بأشكال مختلفة مثل التعرف على الوجه والصوت والكتابة اليدوية والمشي طريقة موثوقة للمصادقة. في الآونة الأخيرة ، أصبح تحليل صور الأذن كطريقة بيومترية طريقة تحديد قوية. أظهر عدد من الباحثين أن التعرف على الأذن هو بديل قابل للتطبيق للقياسات الحيوية الأكثر شيوعًا مثل التعرف على بصمات الأصابع والوجه وقزحية العين ، لأن الأذن مستقرة نسبيًا بمرور الوقت ، وغير جائرة للالتقاط ، وغير قابلة للتعبير ، وكل من الهندسة والشكل من الأذن اختلاف كبير بين الأفراد. جرب الباحثون مجموعة متنوعة من الأساليب لتحسين التعرف على الأذن. قام بعض الباحثين بتحسين الخوارزميات الموجودة للمساعدة في التعرف على الأفراد من خلال آذانهم. أخذ آخرون خوارزميات تم تجربتها واختبارها لغرض آخر ، مثل التعرف على الوجوه ، وتطبيقها على التعرف على الأذن. نتج عن هذه الأساليب عدد من أحدث الأساليب الفعالة للتعرف على الأفراد عن طريق الأذن. تحدث العديد من التحديات بسبب أخطاء في طريقة التقاط الصور ، وسوء الإضاءة ، وأبعاد الصورة ، وآذان خارج الزاوية ، وما إلى ذلك. وقد اعتمد الباحثون الطرق المختلفة من أجل تحسين أداء التعرف على الأذن وزيادته.

تتضمن معظم أنظمة التعرف على الأذن عمليات قبل مرحلة استخراج الميزات ؛ أولاً ، مرحلة المعالجة المسبقة ، والتي تتم لتحسين منطقة الاهتمام فقط. تتضمن هذه المرحلة التجزئة والتطبيع. بعد ذلك ، لتحسين

صورة الأذن الطبيعية ، تم تنفيذ تقنية معادلة Histogram Equalization (HE) لتسهيل تطبيق خطوة استخراج الميزة.

في هذا البحث ، قدمنا نهجًا قائمًا على دمج تقنيتين مختلفتين لاستخراج الميزات: Histograms of Oriented Gradients (HOG) و Local Binary Patterns (LBP) لاستخراج الميزات المطلوبة. بينما يتم استخدام Principal Component Analysis (PCA) لتقليل مساحة أبعاد الميزة. للتصنيف ، يتم استخدام Linear Discriminant Analysis (LDA). يتم تطبيق التقنية المقترحة على صور قاعدة بيانات IITD I. حققت الطريقة المقترحة إنجازات كبيرة مقارنة بالدراسات الأخرى.

جامعة بنغازي

كلية تقنية المعلومات

مزيج من استخلاص الميزات للتعرف على القياسات الحيوية للأذن البشرية

قدمت من قبل:

بثينة فرج محمد قرقوم

تحت إشراف:

د .أحمد الأوجلي

قدمت هذه الرسالة استكمالا لمتطلبات الحصول على درجة الماجستير في علوم الحاسوب بتاريخ 22.8.2022

يونيو 2022

جامعة بنغازي

كلية تقنية المعلومات

قسم علوم الحاسوب

مزيج من استخلاص الميزات للتعرف على القياسات الحيوية للأذن البشرية

قدمت من قبل:

بثينة فرج محمد قرقوم

المناقشة لجنة قِبل من مجازة الأطروحة أو الرسالة في تاريخ 22. 8. 2022 :

الاسم... العضو الخارجي ....

الاسم... العضو الداخلي..... ....

الاسم... المشرف ....

الاعتماد

د.عثمان محمد البدري. **د .عبد السلام معتوق.**

مدير مكتب الدراسات العليا والتدريب **عميد الكلية**