Malnad College of Engineering, Hassan

(An Autonomous Institution affiliated to VTU, Belgavi)

A Mini Project Report

On

"Design and Development of Ear Based Person Recognition System"

Submitted in partial fulfillment of
the requirements for the award of the degree of

Bachelor of Engineering
in
Computer Science and Engineering

Submitted by

Ranjan H T	4MC20CS121
P G Prajwal	4MC20CS101

Under the guidance of
Dr. Aruna Kumar S V
Associate Professor

Department of Computer Science and Engineering

2022-2023

Malnad College of Engineering

Department of Computer Science and Engineering

Hassan - 573201, Karnataka, India

Certificate

This is to certify that mini project work entitled **“Design and Development of Ear Based Person Recognition System”** is a bonafide work carried out by **Ranjan H T (4MC20CS121), P G Prajwal (4MC20CS101)** in partial fulfillment for the award of Bachelor of Engineering in Computer Science and Engineering of the Visvesvaraya Technological University, Belgavi during the year 2022-2023. The project report has been approved as it satisfies the academic requirements in respect of mini project work prescribed for the Bachelor of Engineering Degree.

Signature of the Guide	Signature of the HOD	Signature of the Principal
Dr. Aruna Kumar S V	Dr. Geetha Kiran A	Dr. S. Pradeep
Associate Professor	Prof. & HOD	Principal
Dept. of CSE, MCE	Dept. of CSE, MCE	MCE

Examiners

Name of the Examiner	Signature of the Examiner
1.	
2.	

ABSTRACT

Biometrics, the science of identifying individuals based on unique physical or behavioral traits, plays a crucial role in various domains, including security and access control. Among the different biometric modalities, the human ear exhibits distinct characteristics that can be leveraged for person recognition systems. This project focuses on the design and development of an ear-based person recognition system using traditional machine learning techniques. By harnessing the specific features of the human ear, techniques are developed to identify individuals based on their ear images. The steps involved in our technique are pre-processing of the image, segmentation of the image, extracting features from the image, and passing those features to a classifier for training.

For this project two ear datasets are used i.e., the AMI ear dataset and the IIT Delhi ear dataset. The images are pre-processed from the available dataset using a Gaussian blurring technique which reduces noise in the image. Segmentation of the image is done by using the K-means clustering algorithm. Here pre-processed images are segmented to get only the required ear part. For extracting features of the image PCA and SIFT algorithms are used which represent the feature in Eigen matrix. For classification, different models like SVM, KNN, and Random Forest are used to calculate their accuracy.

The design and development of an ear-based person recognition system using traditional machine learning techniques offer an effective method for identifying individuals based on ear biometrics. Further research can improve its effectiveness and broaden its scope in the field of real-life applications like security and access control.

ACKNOWLEDGEMENTS

We present with immense pleasure this work titled **"Design and Development of Ear Based Person Recognition System"**. An endeavour over a long period can be successful with the advice and support of many well wishers.

We take this opportunity to express our gratitude and appreciation to all of them. The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without mentioning the people who made it possible. So, with gratitude we acknowledge all those whose guidance and encouragement made to successfully complete this project.

We would like to express sincere thanks to our Principal **Dr. Pradeep S**, Malnad College of Engineering for his encouragement made to successfully completion of the project work. We wish to express our gratitude to **Dr.Geetha Kiran A**, Professor and Head, Department of Computer Science & Engineering for providing a good working environment and for her constant support and encouragement.

It gives great pleasure in placing on record a deep sense of gratitude to our guide **Dr.Aruna Kumar S V**, Associate Professor, Department of Computer Science & Engineering for his daily evaluation of the work and for providing us constant encouragement with his unflinching support and valuable guidance throughout this project.

We would also like to thank all the staff of Computer Science and Engineering department who have directly or indirectly helped us in the completion of the project work. At last we would hereby acknowledge and thank our parents who have been a source of inspiration and also instrumental in the successful completion of the project work.

Ranjan H T

P G Prajwal

TABLE OF CONTENTS

1 Introduction **1**

1.1 Introduction 1

1.2 About Project 3

1.2.1 Problem Statement 3

1.2.2 Objective 3

1.3 Organization of the project 3

2 Literature Survey **5**

3 Proposed Method **7**

3.1 Dataset 8

3.1.1 AMI ear dataset 8

3.1.2 IIT Delhi dataset 9

3.2 Preprocessing 10

3.2.1 Gaussian Blur 10

3.3 Segmentation 11

3.3.1 K-Means Clustering 11

3.4 Feature Extraction 14

3.4.1 Principal Component Analysis (PCA) 15

3.4.2 Scale-Invariant Feature Transform (SIFT) 16

3.5 Model Training 18
3.5.1 SVM (Support Vector Machine) 18
3.5.2 KNN (K-Nearest Neighbors) 20
3.5.3 Random forest 21
3.6 Validation 22

4 Results and Discussion 23

5 Conclusion and Future work 26

References 27

LIST OF FIGURES

1.1 Classification of biometrics 2

3.1 Block diagram of proposed System 7

3.2 AMI ear dataset 9

3.3 IIT Delhi ear dataset 9

3.4 Before Gaussian Blur 11

3.5 After Gaussian Blur 11

3.6 Input image 14

3.7 Clustered image 14

3.8 Blur 14

3.9 Mask 14

3.10 Segmented image 14

3.11 Dimensional reduction in PCA 15

3.12 SIFT 16

LIST OF TABLES

4.1 Results of AMI ear dataset 24

4.2 Results of IIT-Delhi ear dataset 25

Chapter 1

Introduction

1.1 Introduction

A biometric system is a technology that helps in recognizing a user's physiological, behavioral, or both characteristics as input verifies it and identifies the individual as a unique user. Many different aspects of human physiology, chemistry, or behavior can be used for biometric authentication. Biometric systems can be mainly classified into two types, Physiological and Behavioral as shown in Fig.1.1.

Physiological biometrics are characteristics or measurements of the human body, whereas behavioral biometrics refers to the unique way a person performs a certain behavior. Physiological biometrics include the iris, fingerprint, ear, DNS, vein print, and face. Behavioral biometrics include voice, gait, and signature.

The ear biometric system uses the unique features of a person's ear to identify and confirm their identification. It is a form of biometric authentication system that uses the unique characteristics of the ear, including its size, shape, and internal structure, for identification.

The idea behind ear biometrics is that each individual's ear has distinct patterns and traits that may be recorded and analyzed. These patterns can be used to identify a person because they are frequently consistent over time. Early development leads to relatively constant ear shape and structure throughout maturity, making it a trustable biometric feature.

Ear biometric system procedures sometimes use taking a picture or

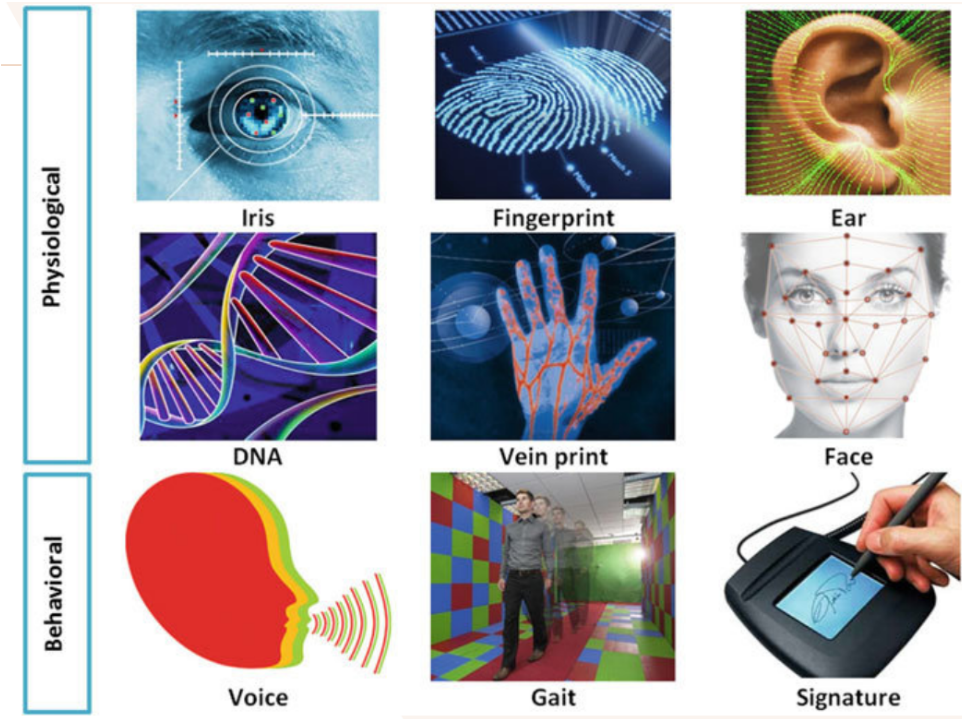

Figure 1.1: Classification of biometrics

a three-dimensional scan of the subject's ear. Specialized ear scanning equipment or common imaging technology like cameras can be used for this. An individual's ear is compared to a previously saved image in a database throughout the verification procedure. The person's identity is verified if the obtained features match the stored image within a reasonable range. A database can be searched to locate an image that matches the captured ear when using ear biometrics for identification.

Systems using ear biometrics have several benefits. First of all, they are non-contact and non-intrusive, which makes them hygienic and comfortable for users. Second, ear biometrics are extremely trustable and resistant to counterfeit or imitation due to the stability and uniqueness of ear characteristics. For increased accuracy and security, ear biometrics can also be integrated with other biometric modalities, such as facial recognition.

Applications for ear biometric systems can be found in several industries, including border security, access control, and healthcare. They can be applied to physical site security, digital transaction authentication, or the tracking of people in surveillance systems.

1.2 About Project

1.2.1 Problem Statement

To design and develop a machine learning model for identifying a person based on his ear pattern. This can address the situation where contactless identification is necessary due to hygiene concerns, such as in healthcare settings or during a pandemic. Ear biometrics allows for identification without physical contact, reducing the risk of transmission of infectious diseases. Traditional identification methods, such as ID cards or passwords, may not be sufficient or may be vulnerable to fraud. Ear biometrics can provide a more secure and accurate method of identifying individuals, particularly in high-security environments or situations where contactless identification is preferred

1.2.2 Objective

The main objective of the project is to develop a method of identification that can be used in a variety of applications The below mentioned steps will be performed to achieve the stated objective:

- To develop existing preprocessing techniques to enhance input images.
- To develop segmentation algorithm to obtain accurate part of ear from the given images.
- To develop a feature extraction method to extract the features of segmented image.
- To develop a machine learning model that trains on the above data and accurately recognises a person based on the input given.

1.3 Organization of the project

This report is organized into five chapters. In Chapter 1, we have the introduction and overview of the project, which includes the problem statement and objectives. Research that has been done related to the project is presented in Chapter 2. Chapter 3 contains the methodologies implemented in our project. This includes steps involved in the preprocessing of the image, segmentation of the image, algorithms used for extracting features, and various classifiers used to train the data. Chapter 4 presents the results of the project along with details of metrics used for the evaluation. Chapter 5 has the outcomes of the project and future works that can be implemented to get better results. Finally, the report includes a reference section listing the sources cited throughout the document.

Chapter 2

Literature Survey

Booysens and Viriri [1] have made a survey on ear biometrics using different deep learning architectures . It covers other aspects such as optimization, loss function, parameter initialisation, regularisation, etc,. They also have included a verity of datasets that will be helpful for our project.

Kamboj et al., [2] made a comprehensive survey based on deep learning approach. They discussed why ear biometics is better than other biometric approaches. Different parameters to evaluate the performance of the model are mentioned.

Priyadharshini et al., [3] conducted experiments using simple deep CNN architecture. They used AMI dataset and IITD ear dataset. Using IITD-II Ear dataset and the achieved an accuracy of 97.36%.

Alkababji and Mohammed [4] developed a realtime system to recognize person based on ear pattern. Their paper show various steps they used in the process like ear acquisition, ear detection, feature extraction, feature reduction and selection, and matching. They concluded that using Artificial neural network robust, high performance, fast ear recognition system can be constructed.

Hurley et al., [5] made a comprehensive research on how the ear as a biometrics have developed on the research on the 3D potentia.They have made work on the shape of the ear, whether it be PCA, force field, or ICP, but it may prove profitable to further investigate if different and particular parts of the ear are more important than others from a recognition perspective.

Sharkas [6] has developed a system using ensemble classifiers where he expreimented with segmented ear images as well as non-segmented ear and concluded that The non-segmented images provided superior results over the segmented images for the AMI ear database.

Annapurani et al., [7] made a comprehensive research on distinctive feature,that tragus is automatically extracted from the ear and used in conjunction with the ear's form to authenticate ears. An individual's identity can be determined by their tragus characteristics and ear shape. Finally, to verify the identity of the person, the tragus and ear shape are combined. Here, the tragus is extracted using a novel technique for edge enhancement that first extracts the shape of the ear from the coordinates.

Raposo et al., [8] introduced a new dataset of ear images for biometric applications, whose main differentiating feature is that it simulates the data acquisition in real-world scenarios, under varying lighting conditions on moving subjects and without requiring them to take any special care with regard to ear occlusions and poses. Our tests demonstrate that, as would be predicted, the quality of the acquired data considerably affects how well the most popular ear identification techniques operate. We also anticipate that the UBEAR dataset will serve as an important resource for the study of ear identification algorithms that are more resistant to degraded data, namely as a result of dynamic lighting recognition, subject movements, and perspective.

Chapter 3

Proposed Method

Designing an ear recognition system involves collecting a diverse dataset of ear images, annotating them with corresponding identities, and preprocessing the images to enhance quality. Relevant features are then extracted using techniques such as SIFT or PCA. The dataset is split into training and validation sets for model training. The model is experimented with different algorithms such as Support Vector Machines (SVM), Random Forests, or kNNs.The trained model's performance is evaluated using a separate testing dataset, measuring metrics like accuracy, precision, recall, and F1-score to assess its effectiveness in identifying individuals based on their ear images.

The block diagram shown in Fig. 3.1 is the conventional way to process the image in ear recognition system.

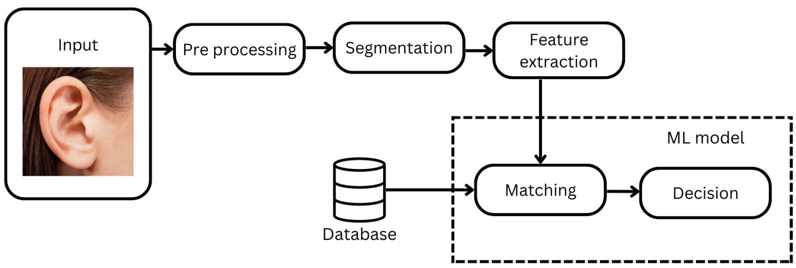

Figure 3.1: Block diagram of proposed System

Ear biometric systems typically use a combination of image processing and pattern recognition techniques to identify individuals based on the unique features of their ear.We follow these steps to develop the system (a) Pre-processing (b) Segmentation (c) Feature extraction (d) Classification and (e) Testing and validation. Images are collected from various datasets and given as input. Then we use pre-processing techniques to remove noises and enhance the images. Then segmentation algorithm is applied to get the ear part removing all other unwanted data from the image. After that we use feature extraction methods and obtain a suitable method that can be implemented in our system. The features vector obtained from feature extraction is used to train the classifier

3.1 Dataset

3.1.1 AMI ear dataset

The set was acquired from 100 different subjects, all of them in the age range of 19-65 years. For each individual, seven images (six right ear images and one left ear image) were taken.

All the images were taken using a Nikon D100 camera, under the same lighting conditions, with the subject placed seated at a distance of about 2 meters from the camera and looking at some previously fixed marks.

Five of the captured images were right side profile (right ear) with the individual facing forward (FRONT), looking up and down (UP, DOWN) and looking left and right (LEFT, RIGHT). The sixth image of right profile was taken with the subject also facing forward but with a different camera focal length (ZOOM). Last image (BACK) was a left side profile (left ear).

The database of 700 images has been sequentially numbered for every subject with an integer identification number. The resolution of these images is 492 x 702 pixels and all these images are available in jpeg format.

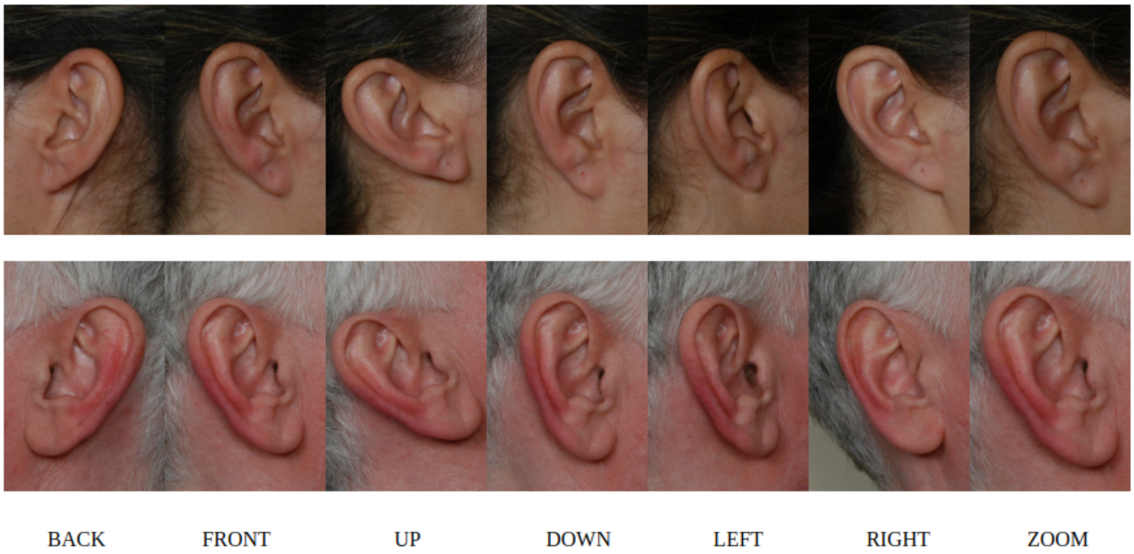

Figure 3.2: AMI ear dataset

3.1.2 IIT Delhi dataset

All the images are acquired from a distance (touchless) using simple imaging setup and the imaging is performed in the indoor environment. The currently available database is acquired from the 121 different subjects and each subject has at least three ear images.

All the subjects in the database are in the age group 14-58 years. The database of 471 images has been sequentially numbered for every user with an integer identification/number. The resolution of these images is 272x204 pixels and all these images are available in jpeg format.

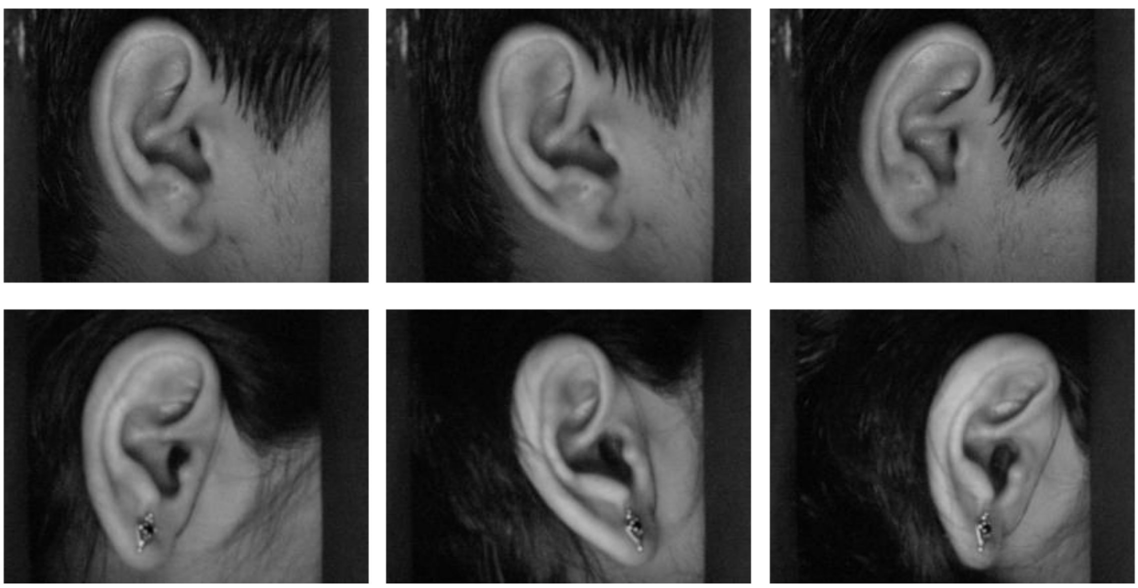

Figure 3.3: IIT Delhi ear dataset

3.2 Preprocessing

Preprocessing in an ear recognition system involves resizing and cropping ear images to focus on the region of interest. Normalization techniques can be applied to adjust pixel values. Noise reduction methods help remove artifacts, and alignment techniques compensate for variations in rotation, scale, and orientation, improving the system's accuracy.

3.2.1 Gaussian Blur

Gaussian blur is a common preprocessing technique used in ear recognition systems to reduce noise and enhance the quality of ear images. It applies a Gaussian filter to the images, which smoothes out high-frequency noise while preserving the overall structure and important details.

The blur helps to further reduce any remaining noise or artifacts that may affect subsequent analysis or feature extraction steps.By applying Gaussian blur, the ear images become smoother, with a reduction in sharp edges and high-frequency noise. This can improve the robustness of the ear recognition system, as it becomes less sensitive to small variations in pixel values or minor imperfections in the image.

Excessive blurring can lead to loss of important ear features, impacting the system's ability to accurately recognize individuals.

Overall, Gaussian blur is a valuable pre-processing technique in ear recognition systems, helping to reduce noise and enhance the quality of ear images while maintaining the essential structure and features necessary for accurate recognition.

The formula for the Gaussian function is given by:

$$G(x, y) = (1/(2\Pi\sigma^2)) * exp(-((x-\mu)^2 + (y-\gamma)^2)/(2\sigma^2))$$

where G(x, y) represents the value of the Gaussian function at coordinates (x, y), μ and γ are the mean values of the function (typically

Figure 3.4: Before Gaussian Blur

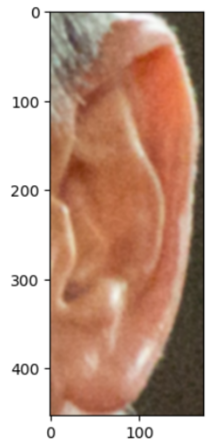

Figure 3.5: After Gaussian Blur

set to the center of the kernel), and σ is the standard deviation.

3.3 Segmentation

Segmentation plays a crucial role in an ear recognition system by isolating the ear region from the background. This process involves separating the ear from other facial features or objects in the image. Techniques like thresholding, edge detection, or machine learning algorithms can be employed to accurately identify and extract the ear region. Successful segmentation enables focused analysis and feature extraction, improving the overall performance of the ear recognition system.

3.3.1 K-Means Clustering

The k-means clustering technique can be utilized for segmentation in an ear recognition system. This unsupervised learning algorithm partitions the image into k clusters based on pixel similarity. Initially, k cluster centers are randomly selected, and pixels are assigned to the nearest center. The cluster centers are then updated by recalculating the mean of the assigned pixels, and the process iterates until convergence.

The algorithm works as follows:

1. Initialization: Randomly select k cluster centers. These centers represent the initial cluster prototypes.

2. Assignment: Assign each pixel in the image to the nearest cluster center based on a distance metric, commonly the Euclidean distance. Each pixel becomes part of the cluster whose center it is closest to.

3. Update: Recalculate the cluster centers by computing the mean of all the pixels assigned to each cluster. This step updates the cluster prototypes to better represent the cluster members.

4. Iteration: Repeat the assignment and update steps iteratively until convergence. Convergence occurs when the cluster centers stabilize, and there is minimal change in the assignments of pixels to clusters.

The goal of the k-means algorithm is to minimize the within-cluster sum of squares (WCSS), also known as the inertia or distortion. This measure quantifies the compactness of the clusters by summing the squared distances between each pixel and its assigned cluster center.

Mathematically, given an image with n pixels, the k-means algorithm seeks to minimize the following objective function:

$$J = \sum_{i=1}^{n} \sum_{j=1}^{k} \|x^{(i)} - \mu^{(j)}\|^2$$

Here, $x^{(i)}$ represents the ith pixel, $\mu^{(j)}$ represents the center of the jth cluster, and $\|\cdot\|$ denotes the Euclidean distance. The algorithm optimizes J by iteratively adjusting the cluster assignments and updating the cluster centers until convergence.

In the context of segmentation for ear recognition, k-means clustering helps group similar pixels together, allowing for the extraction of the ear region from the background. The resulting clusters can be further processed to isolate the ear for subsequent analysis and recognition tasks.

The steps of Segmentation of the image from the preprocessed image are

1. Clustering: In this step similar pixels are grouped together based

on their color. As shown in Fig. 3.7 the image is divided into cluster of two.

2. Blur: Gaussian blur is implemented on the clustered image to reduce unwanted cluster and smoothen the clusters as shown in Fig. 3.8.

3. Mask generation: A mask is a binary image that indicates which pixels in the original image should be preserved or modified. A mask has same dimensions as the original image, where white pixels (pixel value of 255) represent the areas of interest, and black pixels (pixel value of 0) represent the areas to be masked or ignored. In our case, the clusters with lighter pixels are replaces with pixel value of 255 and clusters with darker pixel values are replaced with pixel value of 0 to get result as shown in Fig. 3.9.

4. Applying mask: When applying a mask to an image, the corresponding pixels in the original image are examined based on the mask's values. If a pixel in the mask is white, the corresponding pixel in the original image is left unchanged. However, if a pixel in the mask is black, the corresponding pixel in the original image is removed to segmented image as shown in Fig. 3.10.

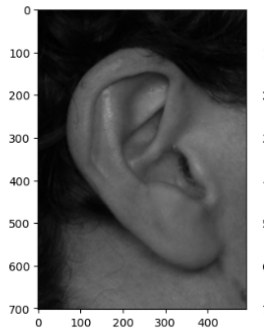

Figure 3.6: Input image

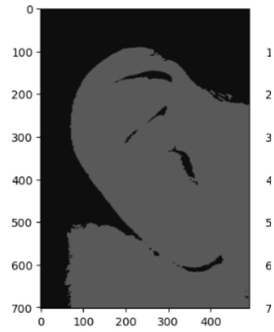

Figure 3.7: Clustered image

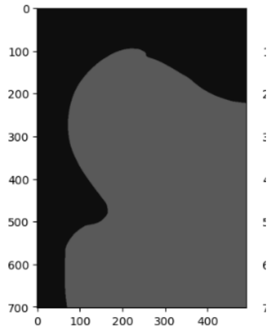

Figure 3.8: Blur

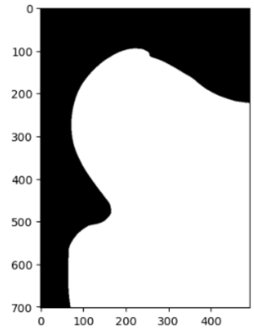

Figure 3.9: Mask

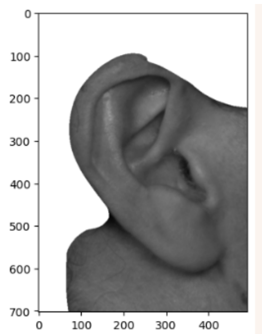

Figure 3.10: Segmented image

3.4 Feature Extraction

Feature extraction is a crucial step in an ear recognition system that involves capturing distinctive and discriminative characteristics from ear images. Various techniques, such as Principal Component Analysis (PCA) , Scale-Invariant Feature Transform (SIFT) can be employed to extract relevant features. These features can represent the shape, texture, or geometric properties of the ear, enabling effective representation and subsequent comparison for accurate identification or verification in the recognition system.

3.4.1 Principal Component Analysis (PCA)

PCA (Principal Component Analysis) is a widely used technique for feature extraction in various domains, including ear recognition systems. It is a dimensionality reduction method that aims to transform the original high-dimensional feature space into a lower-dimensional space while preserving as much information as possible.

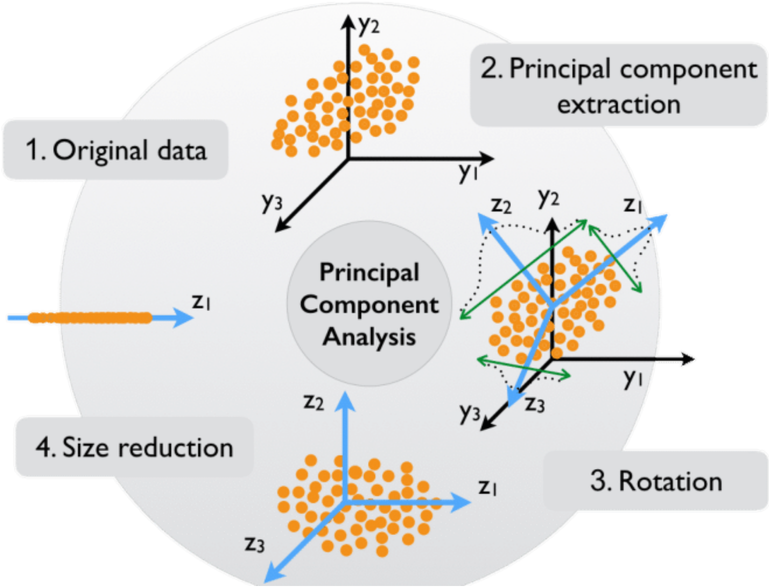

Figure 3.11: Dimensional reduction in PCA

In the context of ear recognition, PCA can be applied to extract the most discriminative features from the ear images. The steps involved in PCA-based feature extraction are as follows:

Let's consider a dataset with m feature vectors, each of dimension n. We can represent the dataset as an m x n matrix X, where each row represents a feature vector. The steps involved in PCA can be summarized as follows:

1. Mean Centering: Subtract the mean vector from each feature vector in X to center the data around the origin.

2. Covariance Matrix: Compute the covariance matrix C of the mean-centered data. The covariance matrix captures the relationships and variances between the different features.

3. Eigenvalue Decomposition: Perform eigenvalue decomposition on

the covariance matrix C to obtain its eigenvectors and eigenvalues. The eigenvectors represent the principal components, and the eigenvalues indicate the amount of variance explained by each eigenvector.

4. Selecting Principal Components: Sort the eigenvectors based on their corresponding eigenvalues in descending order. Choose a subset of k eigenvectors with the highest eigenvalues, where k is the desired lower-dimensional representation.

5. Feature Projection: Project the mean-centered data onto the selected eigenvectors to obtain the lower-dimensional feature representation. The projected data can be computed as Y = X * V, where Y is the transformed feature matrix and V contains the selected eigenvectors.

PCA in feature extraction helps reduce the dimensionality of the ear images' representation, removing redundant or less informative features. The resulting lower-dimensional feature representation facilitates efficient and accurate comparison or classification tasks in ear recognition systems.

3.4.2 Scale-Invariant Feature Transform (SIFT)

The Scale-Invariant Feature Transform (SIFT) is a popular feature extraction technique used in computer vision and image processing. SIFT is designed to extract robust and distinctive features from images, which are invariant to changes in scale, rotation, and affine transformations.

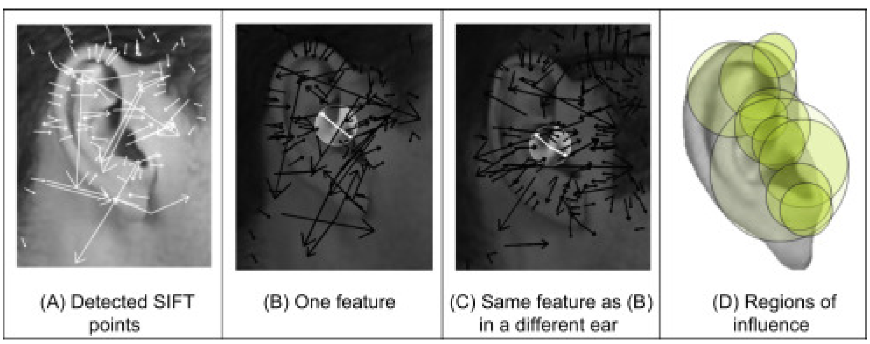

Figure 3.12: SIFT

The SIFT algorithm consists of the following key steps:

1. Scale-Space Representation: SIFT constructs a scale-space representation of an image by convolving it with Gaussian filters at different scales. The scale-space representation is obtained by repeatedly convolving the image with Gaussian kernels of increasing sizes.

2. Difference of Gaussians (DoG): The DoG is computed by subtracting adjacent scales of the scale-space representation. It highlights regions with significant changes in intensity, which correspond to potential keypoints.

3. Keypoint Localization: SIFT applies a technique called the Extrema Detection algorithm to identify local extrema in the DoG images. The extrema are detected by comparing each pixel's intensity with its neighboring pixels across different scales.

4. Keypoint Orientation Assignment: SIFT assigns an orientation to each keypoint to achieve rotation invariance. This is done by computing the gradient magnitudes and orientations of the image in the neighborhood of the keypoint. A histogram of orientations is constructed, and the dominant orientation is determined as the keypoint's assigned orientation.

5. Keypoint Descriptor Calculation: SIFT computes a descriptor for each keypoint to capture its distinctive characteristics. The descriptor is based on the gradients of the image within the keypoint's neighborhood. The gradients are transformed into a descriptor vector, typically using histograms or other statistical measures.

6. Keypoint Matching: In subsequent images or frames, keypoints are detected and described using the same process. Keypoints are matched based on their descriptors using techniques such as nearest neighbor matching and thresholding. The matching process aims to identify corresponding keypoints between images.

The SIFT algorithm utilizes concepts such as Gaussian filtering, gradient computation, histogram construction, and matching algorithms. These mathematical operations enable SIFT to extract robust and scale-invariant features from images, making it a powerful technique for various computer vision tasks, including object recognition, image stitching, and image retrieval.

3.5 Model Training

Model training in an ear biometric system involves the process of training a machine learning or deep learning model using a labeled dataset of ear images. The model learns to recognize and differentiate between different individuals based on their unique ear patterns. Training typically includes preprocessing, feature extraction, and optimization techniques to achieve accurate ear recognition.

3.5.1 SVM (Support Vector Machine)

SVM model training in an ear biometric system involves preprocessing the ear images, extracting relevant features, splitting the dataset into training and testing sets, training the SVM model on the training set, optimizing the model's parameters, and evaluating its performance using the testing set. The preprocessing step ensures consistent and clean data, while feature extraction captures distinctive ear patterns. The SVM model learns to classify ear images into different classes based on the extracted features and finds an optimal hyperplane for separation. Model optimization fine-tunes parameters for improved performance, and evaluation measures accuracy, precision, recall, and F1-score to assess the model's effectiveness in ear recognition.

By training an SVM model on a labeled dataset of ear images, the model can learn to classify and recognize individuals based on their ear patterns. This trained model can then be used for ear recognition tasks in real-world scenarios.

1. Model Representation

In SVM, the goal is to find a hyperplane that maximally separates the data points from different classes. The hyperplane is represented by the equation:

$$f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b$$

2. Margin Maximization

   The SVM seeks to find the hyperplane that maximizes the margin, which is the perpendicular distance between the hyperplane and the closest data points (support vectors) from different classes.

3. Optimization Objective

   The SVM formulation aims to minimize the following objective function:

   $$\min_{\mathbf{w},b} \|\mathbf{w}\|^2 + C\sum \xi_i$$

   subject to: $y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i$ for all i, $\xi_i \geq 0$ for all i.

   Here, $\|\mathbf{w}\|^2$ represents the squared Euclidean norm of the weight vector $\mathbf{w}$, C is a regularization parameter, ξ_i is a slack variable allowing for misclassified or margin-violating samples, and y_i represents the class label of the ith sample.

4. Training Algorithm

   The SVM training involves solving the optimization problem using methods such as quadratic programming or convex optimization algorithms. This leads to the determination of the optimal weight vector $\mathbf{w}$ and bias term b that define the decision boundary.

5. Kernel Trick

   SVM can be extended to nonlinear classification by applying the kernel trick, which implicitly maps the input vectors to a high-dimensional feature space. This allows for the nonlinear separation of data points using linear decision boundaries.

3.5.2 KNN (K-Nearest Neighbors)

KNN (k-Nearest Neighbors) model training in an ear biometric system involves several key steps. First, the ear images are transformed into feature vectors, representing the unique characteristics of each ear pattern. The labeled dataset is then divided into training and testing sets to assess the model's performance. To ensure fair comparisons, feature normalization is often applied to scale the features appropriately. The KNN model is created using the training data, storing the feature vectors and their corresponding labels. For a new ear image, the KNN algorithm calculates the distances to all training samples using a chosen distance metric, such as Euclidean distance. The k nearest neighbors are determined based on the computed distances. Finally, the predicted label for the new ear image is assigned by majority voting among its k nearest neighbors. The effectiveness of the trained KNN model is evaluated using the testing dataset to measure its accuracy in ear recognition.

Given: - Training dataset: $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, where each $\mathbf{x}_i$ is a feature vector of length d. - Corresponding class labels: $\mathbf{Y} = \{y_1, y_2, \ldots, y_n\}$, where each y_i is the class label for $\mathbf{x}_i$.

1. Distance Metric

   Choose a distance metric, such as Euclidean distance or Manhattan distance, to measure the similarity between feature vectors.

2. Training Algorithm

   For each input vector $\mathbf{x}$ in the training dataset:

   (a) Calculate the distance between $\mathbf{x}$ and all training samples using the chosen distance metric.

   (b) Select the k nearest neighbors with the smallest distances to $\mathbf{x}$.

   (c) Assign the class label of $\mathbf{x}$ based on the majority class among the k nearest neighbors. In the case of regression, the predicted value is often the average of the k nearest neighbors.

3. Choosing the Value of k

   Determine the optimal value of k through techniques like cross-validation or grid search. A smaller value of k can lead to overfitting, while a larger value can result in underfitting.

3.5.3 Random forest

Random forest model training in an ear biometric system involves several important steps. Firstly, a labeled dataset consisting of ear images and their corresponding labels is prepared for training. The random forest algorithm builds an ensemble of decision trees by randomly selecting subsets of the training data and features at each tree node. Each decision tree is trained using these subsets, with nodes split based on the Gini index or information gain. During the training process, the random forest combines the predictions of multiple decision trees to make a final classification decision. The model's hyperparameters, such as the number of trees and maximum depth, can be tuned to optimize performance. Finally, the trained random forest model is evaluated on a separate testing dataset to assess its accuracy and generalization capabilities in ear recognition.

Given: - Training dataset: $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, where each $\mathbf{x}_i$ is a feature vector of length d. - Corresponding class labels: $\mathbf{Y} = \{y_1, y_2, \ldots, y_n\}$, where each y_i is the class label for $\mathbf{x}_i$.

1. Random Forest Construction

   (a) Randomly select a subset of the training dataset (bootstrap sample) with replacement.

   (b) For each bootstrap sample:

      - Randomly select a subset of features.
      - Construct a decision tree using the selected features and the associated class labels.

2. Ensemble of Decision Trees

Combine the predictions of multiple decision trees to make the final prediction:

- Classification: Take the majority vote of the predicted class labels from all decision trees.
- Regression: Take the average of the predicted values from all decision trees.

3. Training Algorithm

   Repeat the Random Forest construction step multiple times to create an ensemble of decision trees with different subsets of the training dataset and features.

4. Voting or Averaging

   Aggregate the predictions of all decision trees in the ensemble to make the final prediction:

   - Classification: Majority voting.
   - Regression: Averaging.

The Random Forest model leverages the ensemble of decision trees to improve prediction accuracy, reduce overfitting, and handle noise and variability in the training dataset.

By implementing the Random Forest algorithm, we can harness the collective wisdom of multiple decision trees to make robust predictions for new, unseen data points.

3.6 Validation

The trained model in ear segmentation can be validated by evaluating its performance on a separate validation dataset. This involves applying the trained model to segment ear images in the validation set and comparing the segmentation results with ground truth annotations. Metrics such as accuracy,precision,recall and F1-score can be calculated to assess the accuracy and quality of the segmentation predictions.

Chapter 4

Results and Discussion

The feature extracted from the segmented images of AMI and IIT-Delhi datasets are split into train and test datasets with a ratio of 80:20 respectively. The model is trained using the training dataset and evaluated against the test dataset.

The metrics used for evaluation are Accuracy, Precision, Recall and F1 Score

1. Accuracy

   Accuracy measures the overall correctness of a classification model. It calculates the ratio of correctly predicted instances to the total number of instances.

   $$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Precision

   Precision represents the proportion of correctly predicted positive instances out of all instances predicted as positive. It indicates the model's ability to avoid false positives.

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

3. Recall

   Recall, also known as sensitivity or true positive rate, measures the proportion of correctly predicted positive instances out of all actual positive instances. It indicates the model's ability to identify positive instances.

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

4. F1 Score

   The F1 score is the harmonic mean of precision and recall, providing a balanced measure of a model's performance. It combines both precision and recall into a single metric.

$$F1Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall}$$

The table 4.1 gives the result of AMI ear dataset. In this dataset we have achieved accuracy of 41% by using PCA for feature extraction and SVM for classification.

Table 4.1: Results of AMI ear dataset

Feature extraction algorithm	**Model**	**Accuracy**	**Precision**	**Recall**	**F1 Score**
PCA	SVM linear	0.41	0.45	0.41	0.41
PCA	SVM poly degree 3	0.16	0.21	0.16	0.16
PCA	KNN k=3	0.23	0.29	0.29	0.23
PCA	KNN k=5	0.18	0.20	0.18	0.16
PCA	KNN k=7	0.18	0.18	0.18	0.16
PCA	Random forest	0.26	0.28	0.26	0.24
SIFT	KNN k=3	0.17	0.25	0.17	0.18
SIFT	KNN k=5	0.16	0.22	0.16	0.16
SIFT	KNN k=7	0.14	0.18	0.14	0.14
SIFT	Random forest	0.24	0.26	0.24	0.24

The table 4.2 contains the results of testing against segmented images of IIT-Delhi dataset. In this dataset we have achieved accuracy of 92% and 91% on KNN and SVM models respectively.

Table 4.2: Results of IIT-Delhi ear dataset

Feature extraction algorithm	**Model**	**Accuracy**	**Precision**	**Recall**	**F1 Score**
PCA	SVM linear	0.91	0.91	0.91	0.91
PCA	SVM RBF	0.64	0.63	0.64	0.62
PCA	KNN k=3	0.92	0.91	0.92	0.90
PCA	KNN k=5	0.84	0.81	0.84	0.81
PCA	KNN k=7	0.75	0.71	0.75	0.71
PCA	Random forest	0.72	0.76	0.72	0.72

In the tables above, differences in the results can be seen. This is because of better segmentation of images in the IIT Delhi dataset than the AMI dataset. Segmentation plays a major role in gaining higher accuracy. Among the experimented models, SVM and KNN performed better and by improving segmentation better results can be achieved.

Chapter 5

Conclusion and Future work

Ear-based recognition system can be used as a trustable biometric system. Image quality plays an important role in the development of the system. Our technique can only be used on images taken in a constrained environment. Low-quality images taken in an unconstrained environment cant be recognized by our model.

In this project, a supervised dataset is used which contains images and labels. PCA feature extraction method is used which is one of the most used feature extraction methods which gives up to 92% accuracy on testing over the available dataset.

In the future, we need to improve the accuracy so that it can be used in the real world. Deep learning techniques can be used which will give better results and performance. Techniques can be developed in which images can be taken in an unconstrained environment and recognized them. Improvements in technology can lead us to solve the above problem where the images are taken in an unconstrained environment. Advancements in the areas especially pre-processing and segmentation of the images can lead to an increase in the performance and accuracy of the model. We have used conventional machine learning techniques which have limitations.

References

[1] Aimee Booysens and Serestina Viriri. Ear biometrics using deep learning: a survey. *Applied Computational Intelligence and Soft Computing*, 2022, 2022.

[2] Aman Kamboj, Rajneesh Rani, and Aditya Nigam. A comprehensive survey and deep learning-based approach for human recognition using ear biometric, 4 2021.

[3] Ramar Ahila Priyadharshini, Selvaraj Arivazhagan, and Madakannu Arun. A deep learning approach for person identification using ear biometrics. *Applied intelligence*, 51:2161–2172, 2021.

[4] Ahmed M Alkababji and Omar H Mohammed. Real time ear recognition using deep learning. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 19(2):523–530, 2021.

[5] David J Hurley, Banafshe Arbab-Zavar, and Mark S Nixon. The ear as a biometric. In *2007 15th European Signal Processing Conference*, pages 25–29. IEEE, 2007.

[6] Maha Sharkas. Ear recognition with ensemble classifiers; a deep learning approach. *Multimedia Tools and Applications*, pages 1–27, 2022.

[7] K Annapurani, MAK Sadiq, and C Malathy. Fusion of shape of the ear and tragus–a unique feature extraction method for ear authentication system. *Expert Systems with Applications*, 42(1):649–656, 2015.

[8] Rui Raposo, Edmundo Hoyle, Adolfo Peixinho, and Hugo Proença. Ubear: A dataset of ear images captured on-the-move in uncontrolled conditions. In *2011 IEEE workshop on computational intelligence in biometrics and identity management (CIBIM)*, pages 84–90. IEEE, 2011.