Fingerprint Based Gender and Age Identification Using Neural Network

A Dissertation

SUBMITTED BY

Heena Bawa
11205034

In Partial Fulfillment for the award of the Degree of

Master of Technology
In
Electronics and Communication Engineering

UNDER THE GUIDENCE OF

Mr. Ajmer Singh
Assistant Professor

LOVELY PROFESSIONAL UNIVERSITY

Transforming Education Transforming India

School of Electronics and Communication Engineering
Lovely Professional University, Phagwara, Punjab

(Jan 2016 - May 2017)

DECLARATION

I, **Heena Bawa**, hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Place: **Signature:**

Date: **Name:**

Reg. no.:

CERTIFICATE

This is to certify that the Dissertation entitled **Fingerprint based Gender and Age identification using Neural Network** submitted by **Heena Bawa** with registration number 11205034 to the Lovely Professional University towards partial fulfilment of the requirements for the award of the degree of Masters of Technology in Electronics and Communication Engineering is a bona fide record of the work carried out by her under my supervision and guidance.

Mr. AJMER SINGH
Assistant Professor
School of Electronics and Communication Engineering
Lovely Professional University
Phagwara, Punjab, India, 144401

ABSTRACT

Years ago, Fingerprints are used for authentication of humans. Fingerprints of the person remain same throughout the life. Fingerprints of each and every humans are unique, even in case of twins. So fingerprints became the unique identity proof of humans. In present, researchers proposed a lot of algorithms to extract information about humans from fingerprints like gender, age, region, etc.

The Dissertation work involves the study of technique for fingerprint based age and gender identification. The gender and age identification has many application in the crime scene investigation, to restrict the entry on some places based on gender and age group. A supervised learning based methodology is proposed to identify the gender and age group of given fingerprint. The methodology includes a back propagation based neural network with stochastic gradient descent as an optimizer and Softmax regression classifier. This model has been trained and tested for gender and age group classification. The model has accuracy of 90% accurate results for gender classification and 70% for age group classification.

For good accuracy, feature extraction is the most important part in any learning algorithm. There has been many feature extraction methods available to extract the features of the fingerprint image like frequency domain, spatial domain, etc. from fingerprints. Various methods has already been proposed which involves hard computing or unsupervised learning Most of the information of fingerprints can be extracted by finding its ridge thickness to valley thickness ratio or by finding the energy of fingerprints after compressing the image using any image compressing technique. Hard computing offers less efficient results as compare to soft computing.

Keywords

Fingerprint, Discrete Wavelet Transform, Discrete Cosine transform, Back Propagation, Gradient Descent, Softmax Regression.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBRIVATIONS

BBDCT	–	Block Based Discrete Cosine Transform
DCT	–	Discrete Cosine Transform
DFT	–	Discrete Fourier Transform
DWT	–	Discrete Wavelet Transform
FFT	–	Fast Fourier Transform
KNN	–	K- Nearest Neighbor
LDA	–	Linear Discrimination Analyses
LR	–	Likeliness Ratio
PSD	–	Power Spectral Density
PSD	–	Power Spectral Density
QDA	–	Quadrature Discrimination Analyses
RVA	–	Ridge to Valley Area
RTVTR	–	Ridge Thickness to Valley Thickness Ratio
SGD	–	Stochastic Gradient Descent
SVD	–	Singular Value Decomposition
SVM	–	Singular Value Decomposition

CHAPTER - 1
INTRODUCTION

"The key to growth is the introduction of consciousness into our awareness" – Lao Tzu

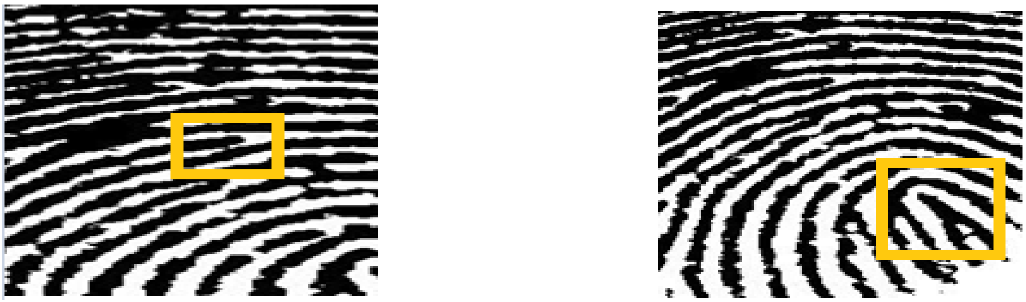
Gender Classification deals with identifying or classifying the Gender of human being using certain criterion, parameters so as to get the correct result as male or female. As a human being, humans can easily identify whether the person standing in front is male or female. But a computer can't classify the person in the same way. There are various techniques that help computer or machine to identify or classify the human as male or female. These techniques can be using facial features, fingerprints, body figure, etc. And there are many techniques to extract features from these parameters.

Facial Features for gender classification include eye length, face edges and many more. These features provide a distinct classification between male and female. Another important parameter for gender classification is fingerprints. Fingerprints of a person never matches with the fingerprints of another person not even with the fingerprints of his/her family members. Fingerprints emerges as a vital parameter for classifying the gender of human beings. By looking at the fingerprint of some person, it is impossible for human beings to predict whether the fingerprint belong to male or female and what is its age group. But if correct technique is applied, then the computer can immediately detect whether the fingerprint image is of male or female. For Criminal case evaluation, fingerprints are the most usual parameter.

Fingerprints are also considered to be the best method for age identification because the fingerprints of person remains same throughout life. As the person grows in age, the fingerprints of the person also changes because of many factors like change of fingerprint size, changes in the ridges, etc. So, from the fingerprints, it could be judged from which age group it is.

1.1 Fingerprint based gender and Age Classification

Fingerprint based Gender and age classification technique is being used for identifying the gender and age of person. The uniqueness of the fingerprints can be determined by using the ridge characteristics [8]. The two most important local ridge characteristics are: (1) ridge ending, (2) ridge bifurcation. The ridge ending and bifurcation of a finger can be seen from the fingerprint as shown in figure 1.1



(a) (b)

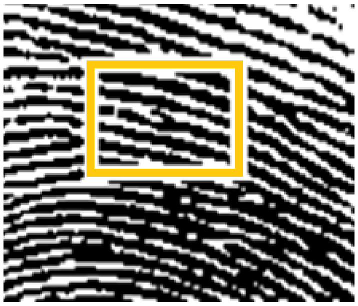
Figure 1.1 (a) Ridge Ending, (b) Bifurcation

Following are the features that can help in classifying the gender and age. There are various kinds of ridge characteristics in addition to ridge ending and ridge bifurcation [9] and these are:

- **Ridge ending** is defined as a ridge type that ends abruptly as shown in figure 1.1(a).
- **Ridge bifurcation** is defined as a single ridge that get divided into two ridges as shown in figure 1.1(b).
- **Short ridge, island or independent ridge** is defined as a ridge that travels a short distance and then ends as shown in figure 1.2 (a).
- **Ridge Enclosure** is a single ridge that bifurcates and then reunites to form a single ridge as shown in figure 1.2(b).
- **Spur** is an extended ridge from the normal ridge of fingerprint as shown in figure 1.3 (a).
- **Crossover or bridge** is a short ridge that runs parallel between two ridges as shown in figure 1.3 (b).

(a) (b)

Figure 1.2 (a) Short Ridge; (b) Ridge Enclosure



(a)

(b)

Figure 1.3 (a) Spur of fingerprint; (b) Bridge in fingerprint

1.2 Features Extraction Methods

Information from fingerprints can be extracted by using feature extraction techniques. These feature extraction techniques includes physical, frequency and spatial domain features as their parameters. These techniques of extracting features from fingerprint have been discussed in the preceding section.

1.2.1 Physical Features Extraction

Physical features include calculating some values using the ridge characteristics and their relationship. There are mainly three classes of fingerprint [1] and they are (1) whorls, (2) loops, (3) arches. But these parameters are not used for classifying fingerprints because such features are not found in every fingerprints

Whorls are not available in all fingerprints. Only 30-35% of the population has whorls [9]. All whorls pattern must have two features that is type line and two deltas. Type lines are ridges that diverge (or separate) and Delta are located at the point of divergence. All loops has one divergence. There are basically four types of whorls and they are plain whorls, central pocket whorl, double loop whorl, accidental whorls and they are shown in figure 1.4.

Plain whorls must have at least one ridge that makes a complete circuit, and an imaginary line from one delta to the other must touch a whorl ridge as shown in figure 1.4(a).

Central pocket whorls must have at least one ridge that makes a complete circuit, and an imaginary line from one delta to the other cannot touch a whorl ridge.

Double loop is two loops combined to make one whorl as shown in figure 1.4(b).

Except these types of whorls, all other types of whorls are called accidental whorls because they contain pattern like whorls.

(a) (b)

Figure 1.4 (a) Plain whorl, (b) Double loop whorl

Loops are that part of fingerprint which have one or more ridges entering from one side of the fingerprint, then making curve and then exiting from the same side of the fingerprint. 60-65% [9] of the population have loops in their fingerprints. So, loops can be taken as one parameter while classifying the fingerprints but it may not give as much accuracy as required. There are two kinds of loops for every kind of fingerprints as shown in figure 1.5 and these are

Ulnar loop is defined as the loop that is opening towards the little finger as shown in figure 1.5(a). It is similar to ulna in the Medial bone.

Radial loop is defined as that kind of loop that is opening towards the thumb as shown in figure 1.5(b). It is similar to the radius of the lateral bone.

(a) (b)

Figure 1.5 Types of Loops (a) Ulnar loop (b) Radial loop

Arches ridge are the part of fingerprint that enter from one side of the fingerprint and leave out from the other side. Arches are not used as features for fingerprint classification because only 5% of the population has Arches. There are mainly two types of arches that is plain arch and tentative arch and they are shown in figure 1.6

Plain arch is a kind of arch that shows a wave like pattern on the fingerprint.

Tented arch is that kind of arch that shows a wave like pattern over the fingerprints.

Figure 1.6 Types of arches in fingerprints

1.2.2 Calculation of Physical features

Physical Features are defined as the features which are obtained by taking the physical parameters of fingerprints that is by considering the Ridges and Valleys, etc. For gender classification, there are two major significant features used [2] and they are

- Ridge to valley thickness ratio
- Ridge Density

A. Ridge to valley thickness ratio

RTVTR [16] is defined as the ration of ridge thickness to the valley thickness. RTVTR is an important deciding factor for the classification of gender because female have more RTVTR ration in comparison to male. Therefore, it becomes easy for the algorithm to find out the type of gender of the given fingerprint.

$$RTVTR = Ridge\ thickness/Valley\ thickness \tag{1.1}$$

Ridges are defined as the black impression in the fingerprint images while valley are defined as the white impression in the fingerprints. These impressions define the fingerprint type as most of the information lies in these patterns. RTVTR is calculated by calculating all the black pixels in the fingerprint image and by calculating all the white pixels in the fingerprint image and then finding its ratio as defined in equation (1.1).

B. Ridge Density

Fingerprint Ridge density can be calculated by drawing a square. In this region, the ridges conform to the finger outline flowing in an arch from one side of the finger to the other. In this selected area of prints, epidermal ridges of both males and females are being counted carefully within a square drawn. The counting of ridges start from one corner of the square to the diagonally opposite corner. Some specific criteria are being observed during the counting procedure such as dots, which are not counted, and the handle of the fork and a lake are counted as two ridges. Hence, this value represent the number of ridges/25mm square and would reflect the ridge density value

$$Ridge\ Density = N/A \tag{1.2}$$

Where N represents number of ridges, A represents Area of square = 25mm^2.

The probability of classifying gender based on ridges density (RD) [23] values is calculated by using the likelihood ratio (LR) which is based on the Bayes Theorem. The numerator of the equation is represented as the probability densities from male fingerprint while the denominator of the equation is represented as the probability densities from female fingerprint.

$$LR = \frac{P(RDM)}{P(RDF)} \tag{1.3}$$

Where P (RDM) is the probability of Ridge Density of Male fingerprint and P(RDF) is the probability of Ridge Density of Female. Likelihood Ratio of any feature need not always the equation mentioned above. LR can also be used as a classifier for classifying only two classes. For example, in Quadratic Descriptive Analyses, Likelihood Ratio is used to find out the class to which the given sample belongs to. In QDA, Likelihood ratio determines the covariance between the two classes.

1.2.3 Frequency domain features

Frequency domain parameters or features are the features which are obtained by transforming the image into frequency domain by using any kind of frequency transformation technique [10] like DFT, DWT, DCT, etc. In this technique, fundamental frequency is obtained of various frequency domain transformations and are used for gender calculations.

Discrete Fourier Transformation converts a finite set of sequence of input function into coefficients of complex sinusoidal ordered by frequency. The DFT of a continuous time signal can be derived by using three steps: (1) sampling in time, (2) segmenting the sampled signal into segments of length N samples, (3) sampling the spectrum of discrete time signal segment in frequency domain. DFT of an input image can be expressed using the formula written in equation (1.4). [9]

$$F(u,v) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m,n) e^{-j2\pi(umx+vny)} \quad (1.4)$$

Where f represents the input image, F represents the DFT of the image, and x and y are the spatial intervals between consecutive signal samples. These values of DFT of input signal obtained are taken as features for gender classification. These are then given to the classifier or any neural networks for classifying the gender into male or female.

Discrete Wavelet Transform has been used for texture analysis. The input signal is being passed through two filters that is low pass filter and high pass filter. The low pass filter gives the average value of the input signal. So, it can be used as feature for gender classification. 2-D DWT is used in which the samples of low pass filter are further transformed into low pass and high pass filter components. The2D – DWT can be represented by using equation (1.5) [28].

$$W(i,j) = \frac{1}{\sqrt{NM}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} s(i,j)\, \phi(k,l) \quad (1.5)$$

$$W(i,j) = \frac{1}{\sqrt{NM}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} s(i,j)\, \psi(k,l) \quad (1.6)$$

Where W represents the 2D-DWT of input image, s represents the input image, $\emptyset$ is scaling function and ψ is wavelet function. The DWT can be n-level depending upon the requirements. Basically, 6-levels DWT is used for extracting the features. The DWT values are not exactly the features for the fingerprint analysis. The energy of these (3*n) + 1 sub-bands gives the features for classifying gender.

Discrete Cosine Transformation expresses a finite sequence of data points in terms of sum of cosine functions oscillating at different frequencies [12]. It is a part of DWT. The DCT helps in separating the given image into parts of differing importance. The DCT can be used to convert the signal into numeric data so that the image's information exists in a quantitative form that can be manipulated for compression [8]. DCT of an input image can be expressed using the formula written in equation (1.5).

$$F(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{M-1}\cos\left(\frac{\pi u(2x+1)}{2N}\right)\cos\left(\frac{\pi v(2y+1)}{2M}\right)f(x,y) \qquad (1.7)$$

Where,

$$\alpha(u)\alpha(v) = \begin{cases} \frac{\sqrt{1}}{N}\, for\; u,v \neq 0 \\ \frac{\sqrt{2}}{N} for\; u,v = 0 \end{cases},$$

f(x, y) represents the input image,
F(x, y) represents the DCT coefficients of Image

1.3 Neural Network

Neural Networks are the models used for computational work, machine learning, deep learning which is based on a huge collection of connected units called neurons. These artificial neurons learn the changes in the behavior of the input in order to response well similar to the natural neurons in the human body [27]. These neurons are connected with layers like the biological neurons. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze.

The learning process of Neural Network has been broadly classified into supervised and unsupervised Training. In supervised training, both the inputs and the outputs are provided. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network. This process occurs over and over as the weights are continually tweaked. The other type of training is called unsupervised training. In unsupervised training, the network is provided with inputs but not with desired outputs. The system itself must then decide what features it will use to group the input data. This is often referred to as self-organization or adaption. At the present time, unsupervised learning is not well understood. This adaption to the environment is the promise which would enable science fiction types of robots to continually learn on their own as they encounter new situations and new environments.

1.4 Classifier

An algorithm that implements classification, especially in a concrete implementation, is known as classifier. The term classifier sometimes also refers to the mathematical function, implemented by a classification algorithm that maps input data to a category. Classifier is the final step in many class finding technique which help to define the class of the input. From the output produced by the algorithm, classifier decides the class of that output that is to which class the output generated belong to. There has been many classifiers available which can be used based upon the requirement like K- nearest Neighbor, Softmax Regression Classifier, Linear Regression Classifier, Support Vector Machine, etc.

Softmax Regression Classifier, Linear Regression classifier, Support Vector Machine are the classifiers used in machine learning and deep learning techniques where there is requirement for very precise results.

1.5 Summary

An introduction to the fingerprints has been discussed in this chapter which involved characteristics of fingerprints like ridges, loops, whorls, etc. Various features extraction techniques from fingerprints has been discussed which may help in classifying the fingerprints. There has been various methods, techniques used by many researchers which help in classifying fingerprints and some of them are being discussed in chapter 2 involving both hard and soft computing.

****

CHAPTER 2
LITERATURE REVIEW

"Research is creating new Knowledge" – Neil Armstrong

In this chapter, study of various research papers by many researchers has been done and their techniques has been discussed for classifying the gender and age of person based upon their fingerprints and comparison has been made between different methods.

2.1 Review of Papers 2016

Shivanand Gornale et al [25] has proposed a fingerprint based gender classification using Discrete Wavelet Transform and Gabor Filters. A dataset of fingerprints of both male and female has been made. All these fingerprint images has been passed through pre-processing stage in order to enhance the quality of fingerprint images. After pre-processing, features has been extracted from the fingerprint images like DWT and Gabor filter coefficients. Linear Discriminant Analysis has been used to classify the class of fingerprint. The proposed algorithm has produced a performance accuracy of 89.15%.

Akanccha Gour and D. Roy [26] has proposed a fingerprint based gender classification technique by using Discrete Wavelet Transform and Discrete Cosine Transform as feature set. . All the fingerprint images has been passed through pre-processing stage in order to enhance the quality of fingerprint images. After the enhancing process of fingerprint images, features has been extracted from them. K- Nearest Neighbor has been used to classify the class of fingerprint samples. For the test fingerprint sample, distance with feature vector stored in the database has been calculated. The class of least distance feature vector has been assigned to the test sample. The algorithm has produced an efficiency rate of 92%.

2.2 Review of Papers 2015

S.S. Gornale et al [13] has proposed a gender classification technique based upon the fingerprints of person. A dataset of fingerprints has been taken and given to the pre-processing stage where many processes like eliminating the background, converting the colored image into the black and white image, cropping, etc. After pre-processing, features has been extracted from the fingerprints using pixels based model. In this pixels based model, a dataset is obtained comprising all the grey-level intensities of the texture. The distance between the co-efficient of interest and its neighboring pixels has been computed. From that distance matrix, features like Difference Entropy, Sum Entropy, Contrast, Variance, etc has been computed. These extracted features are then given to classifier for classifying the gender of person. Linear Discriminant Analyses and Quadratic Discriminant Analyses has been used as classifiers for the method. The algorithm has obtained an overall accuracy of 94% for Quadratic Discriminant Analyses and 92% for Linear Discriminant Analyses.

Md. Sazzad Hossain and M. Habib [3] has proposed a technique for fingerprint based gender identification method using pixel counting of the fingerprint image. The method has collection a total fingerprint database of 200 images comprising 100 of male and 100 of female fingerprints. All the fingerprints has been passed through the pre-processing stage where the colored image has been converted into black and white image. From these images, a dataset of features has been obtained by calculating the Ridge thickness to valley thickness ratio. It has been calculated by first calculating the number of white pixels within the defined area and then calculating the number of black pixels within that defined area. The ratio of count of black pixels with count of white pixels has been calculated. From this calculated ratio, a threshold value has been obtained which is then used for classifying the gender of unknown fingerprint sample. The threshold value obtained by this algorithm is 0.51. If the ratio of black with white pixel count is greater than this threshold then the fingerprint image is being classified as female else male. The algorithm has obtained a performance efficiency of 85% for male and 74% for female.

Mangesh K. Shinde and S. Annadate [4] has given a gender classification technique based upon fingerprints. This method has used Discrete Wavelet Transform and Singular Value Decomposition for extracting features from fingerprints. The wavelet transform of image produces four decomposed sub-band images which are known as low-low (LL), low-high (LH), high-low (HL) and high-high (HH). Each such sub-band represent different properties of the image. The low-low sub-band is the sub-band which comprises most of the energy of the image. The DWT has been calculated upto third level for the low-low part in this proposed method. The energy of each level has been calculated and hence a feature set is obtained. Another feature set used by this algorithm is SVD which calculated the square root of eigen values of the diagonal matrix in descending order. These values are then stored in the feature set which comprised DWT features. These features are then given to the K- nearest neighbor (KNN) classifier. The test fingerprint sample has then been given to the KNN classifier and the classifier then assigns the class of nearest neighbor to the fingerprint. The algorithm has given an accuracy of 78.65%.

Suchita Tarare et al [5] has given a gender classification technique using Discrete Wavelet Transform. All fingerprint samples passed through the pre-processing stage where the colored image is converted into black and white image. From the pre-processed image, a feature set of DWT based energy values has been obtained. The DWT of image transform the fingerprint image into four sub-bands. Among these sub-band the low-low part gives the most of the information about the fingerprint. A feature set is being obtained by finding the DWT of fingerprint image upto six level and hence computing the energy of each sub-band. A total of 19 features has been obtained from the Wavelet Transform of the fingerprint image. These features are then given to the K- nearest neighbor (KNN) classifier. The test fingerprint sample has then been given to the KNN classifier and the classifier then assigns the class of nearest neighbor to the fingerprint by calculating the Euclidean Distance between the test image features and its neighbors. The algorithm has given an accuracy of 74%.

Alok Chauhan et al [14] has given a fingerprint based gender classification method using ridge count and Discrete Wavelet Transform. All the fingerprints has been converted into black and white images. From these images, number of black pixels has been obtained which are then stored in the dataset. The second type of feature used by this algorithm is Discrete Wavelet Transform (DWT). The DWT of image transform the fingerprint image into four sub-bands. Among these sub-band the low-low part gives the most of the information about the fingerprint. A feature set is being obtained by finding the DWT of fingerprint image upto six level and hence computing the energy of each sub-band. Fusion of both the feature set has been done and a complete dataset of features is being obtained. Euclidean distance has been used to compute the class of the testing image by finding the distance of the test feature set with its neighbors. The method has given an accuracy of 72%.

Akhil Anjikar et al [15] has proposed a fingerprint based gender classification technique using Block based Discrete Cosine Transform (DCT). A database of 1000 female fingerprints and

1000 male fingerprints has been taken. All the fingerprints has been pre-processed and resized. After pre-processing, the images obtained are Binary images, from which the dataset has been obtained. The features has been obtained by calculating the Discrete Cosine Transform of the fingerprint image. DCT of image defines the finite sequence of image data points in terms of cosine function oscillating at different frequencies. Like DWT, DCT also divides the fingerprint images into sub-bands of varying frequencies that depend upon the horizontal, vertical and diagonal values. The DCT of image has been calculated by dividing the fingerprint image into small blocks and then DCT of each small bock has been obtained. Hence, the feature set comprises of 64 features for each fingerprint. Euclidean distance has been used as the classifier for gender identification. . Euclidean distance then computes the class of the testing image by finding the distance of the test feature set with its neighbors. The overall success rate for gender classification produced by this algorithm is 55.25%.

2.1.3 Review of Papers earlier 2015

Heena Agarwal and S. Choubey [16] has proposed a gender classification technique using Support Vector Machine as classifier and Ridge Thickness to Valley thickness Ratio as features. The algorithm has used a total database of 300 fingerprints of both male and female. From these images, a dataset of features has been obtained by calculating the Ridge thickness to valley thickness ratio. It has been calculated by first calculating the number of white pixels within the defined area and then calculating the number of black pixels within that defined area. The ratio of count of black pixels with count of white pixels has been calculated. This calculated ratio has been used as a feature set. These features are then given to SVM classifier for classifying the fingerprint image into its class. The algorithm has produced an overall efficiency of 91%.

Aditya K. Saxena and V. K. Chaurasia [6] has proposed a fingerprint based age identification technique. The algorithm has made the fingerprints images to pass through the pre-processing stage where the fingerprint image is converted into black and white image. From the black and white image, features are being extracted using Gabor filter method. This method has used 2-D Gabor filter to extract features from the fingerprint image. The Gabor filter consists of a total of 40 filter coefficients of different frequencies and orientations. These coefficients has then been convolved with the fingerprint image in order to obtain the feature set. These feature set is then given to the KNN classifier. The KNN classifier and the feature matrix combine computes the class of test image.

T.Arulkumara et al [17] has proposed a fingerprint based age classification technique using Discrete Wavelet Transform and Principal Component Analyses. All the fingerprints has been pre-processed and resized. After pre-processing, the images obtained are Binary images, from which the dataset has been obtained. From the pre-processed image, a feature set is being obtained. The first method used for feature extraction is DWT. The DWT of fingerprint image has been calculated upto 6 levels and hence a feature set of 19 values has been obtained. The second method used for calculating the features is PCA. The values obtained using PCA and DWT form the complete dataset for the algorithm. Euclidean distance approach has been used to classify the age of person by comparing the distances with its neighbors. The algorithm has produced an overall efficiency of 68%.

Samta Gupta and A.P. Rao [18] has proposed a fingerprint based gender classification technique using Discrete Wavelet Transform and Artificial Neural Network. A total dataset of 550 images has been used. All the fingerprint images has then been passed through the pre-processing phase where the fingerprint image is converted into black and white and cropped around the core of finger. The DWT of image transform the fingerprint image into four sub-bands. Among these sub-band the low-low part gives the most of the information about the fingerprint. A feature set is being obtained by finding the DWT of fingerprint image upto six level and hence computing the energy of each sub-band. The obtained feature set is then given to the artificial neural network which implements back propagation algorithm for training the neural network. The algorithm has produced an overall efficiency of 91% for classifying the gender using fingerprints as database.

S.S. Gornale and Kruthi R [19] has proposed a fingerprint based gender estimation technique using Fast Fourier transform and Texture analyses as features. A data set of 4000 fingerprint images has been collected. All the fingerprint images has been passed through the pre-processing stage where cropping of image around the core and conversion of colored image into black and white image take place. From each pre-processed image, features has been collected which involves FFT, eccentricity, major axis length. After computing the features from the fingerprints and analyzing the values obtained from the result, a pre-defined threshold value has then been decided for each type of feature. If for the test image, the computed value is greater than the pre-defined threshold value then it has been given the class of female for each feature. If for the test image, the results are given as female if more than two decisions favor female else male. The algorithm has produced an overall efficiency of 60%.

Ravi Wadhwa et al [20] has proposed a fingerprint based gender and age classification technique using Ridge to Valley Area and Discrete Cosine Transform. K- Nearest neighbor has been used as n classifier for the algorithm. All the fingerprint images has been passed through the pre-processing stage. From these images, a dataset of features has been obtained by calculating the Ridge thickness to valley thickness ratio. It has been calculated by first calculating the number of white pixels within the defined area and then calculating the number of black pixels within that defined area. The ratio of count of black pixels with count of white pixels has been calculated. This calculated ratio has been used as one feature in the feature set. The other feature used by the algorithm is DCT coefficients which calculated the energy of fingerprint images at different frequencies and orientations. These features are then given to the KNN classifier for predicting its class.

Pallavi Chand and S.K. Sarangi [21] has proposed a gender identification technique from fingerprints using Discrete Wavelet Transform and Singular Value Decomposition. All the fingerprint images has been converted into black and white images and has been cropped around the core of finger. From these fingerprints, DWT has been calculated upto six levels and hence a feature set of 19 features has been obtained. The other kind of feature used by this algorithm is Singular Value decomposition. The features From DWT and SVD in combine form the feature set. This feature set has been given to K-Nearest Neighbor classifier which calculated the distance between the feature set and its neighbor value. It assigns the class of neighbor fingerprint to the test fingerprint sample with whom the Euclidean distance is minimum. The algorithm has produced an overall efficiency of 80%.

Rijio Jackson Tom and T. Arulkumaran [22] has proposed a fingerprint based gender classification technique using Discrete Wavelet Transform and Principal Component Analyses. The fingerprint samples has been gone through the pre-processing stage in order to enhance the fingerprint image and to remove the noise. After pre-processing of fingerprint image, features has been calculated using two techniques that is DWT and PCA. 2-dimensional discrete wavelet transform of fingerprint image has been calculate upto 6 levels and hence a feature set of 19 features is being obtained using the DWT method. The other features has been obtained using principal component analyses. Both the features in combine form the complete dataset for fingerprint images. The method has used minimum distance classifier as a technique to identify the class of the fingerprint sample. The algorithm has obtained an overall efficiency of 70%.

S. Sudha Ponnarasi and M. Rajaram [23] has proposed a gender classification technique using minutiae extraction. Database of fingerprint samples has been collected. All the fingerprint images has been passed through image enhancement stage where the image has been converted into binary image. After pre-processing, the number of black pixels in the defined block and number of white pixels in the defined block has been calculated. The ratio of count of black pixels and the count of white pixels is used as features for the algorithm. This feature set has then been given to Support Vector Machine to classify the class of fingerprint sample. The algorithm analyzed that the female fingerprint comprises more count for ridge density in comparison to male.

Ritu Kaur and S.G. Mazumdar [24] has proposed a fingerprint based gender classification technique using Frequency Domain features. A dataset of fingerprint samples has been collected. All the fingerprint images has been passed through the pre-processing stage in order to remove noise from the fingerprints. After pre-processing, three kinds of features has been obtained from each fingerprint and they are Fast Fourier Transform, Discrete Cosine Transform, and Power Spectral Density. A threshold value for each kind of feature has been calculated after analyzing different fingerprint samples. The obtained threshold value has then been used to find the class of unknown fingerprint sample. It the value for FFT and DCT of test fingerprint sample is greater than the pre-defined threshold value then it is being assigned as female for that feature else male. And if the value is less than the threshold value for PSD then the fingerprint sample is being assigned class female for that feature else male. The decision for the class of fingerprint sample has been decided on the bases if two or more decisions from the three features favor one class then that is the class of unknown fingerprint sample. The proposed algorithm has produced an overall efficiency of 85 %.

Arun K. S. and Sarath [7] has proposed a gender classification technique using Ridge to Valley thickness ratio and Ridge density as feature set and Support vector Machine as classifier. A database of fingerprint samples has been collected and passed through the pre-processing stage in order to remove background noise from the fingerprint image. After pre-processing, the number of black pixels in the defined block and number of white pixels in the defined block has been calculated. The ratio of count of black pixels and the count of white pixels is used as features for the algorithm. Another feature for the algorithm is the count of ridges. Both these features in combine form the feature set for the algorithm. These features has then been given to Support Vector Machine for classifying the gender of fingerprint sample. The algorithm has produced an overall efficiency of 96%.

2.4 Summary

The study of different research papers has been summarized in the table 2.1 which is defining the feature set used by the paper, classifier used and the performance rate of that paper. In the table, F represents Female and M represents Male, G represents Gender, A represents Age. From the table, it can be concluded that till date most of the work is done using hard computing techniques and Euclidean distance as the classifier. In most of the papers, the feature set used includes DWT, DCT, FFT, Ridge to Valley Thickness Ratio, etc.

Table 2.1 Review of Research Papers

Name of Author	**Year**	**Features**	**Classifier**	**DB Size**		**Accuracy in %age**			
				Tr	**Te**	**M**	**F**	**G**	**A**
S. Gornale, et al [25]	2016	DWT, Gabor filter	KNN	740	-	-	-	89.71	-
A. Gour, D. Roy [26]	2016	DWT	KNN	-	-	-	-	90.00	92.00
S.S. Gornale, et al [13]	2015	Texture Features	LDA	600	-	-	-	92.00	-
M. Hossain, Habib [3]	2015	RTVTR	Thresholding	400	-	85.00	74.00	-	-
M. K. Shinde, S.A.Annadate [4]	2015	DWT	KNN	1000	-	80.46	76.84	-	-
Suchita Tarare, et al [5]	2015	DWT	KNN	-	-	-	-	74.00	-
A. Chauhan, et al [14]	2015	DWT	Euclidean Distance	-	-	-	-	-	-
A. Anjikar, et al [15]	2015	DWT, DCT	Euclidean Distance	800	-	65.25	45.5	-	-
H. Agarwal, S.Choubey [16]	2014	RTVTR	SVM	300	-	81.00	81.00	-	-
A. K. Saxena, V.K.Chaurasia [6]	2014	Gabor Filter	KNN	350	100	-	-	-	-

Name of Author	Year	Features	Classifier	DB Size		Accuracy in %age			
				Tr	Te	M	F	G	A
T.Arulkumara, et al [17]	2014	DWT, PCA	KNN	-	-	-	-	-	68.00
S. Gupta, P. Rao[18]	2014	DWT	ANN	550	-	-	-	91.00	-
S. Gornale, Kruthi [19]	2014	FFT	Thresholding	4320	-	-	-	60.00	-
R. Wadhwa, et al [20]	2013	RTVTR, DCT	KNN	100	-	-	-	-	-
P.Chand, S.K.Sarangi[21]	2013	DWT, SVD	KNN	100	-	-	-	80.00	-
R. J. Tom, T.Arulkumaran[22]	2013	DWT, PCA	KNN	400	-	-	-	70.00	-
S. S. Ponnarasi, M.Rajaram [23]	2012	Ridge Density	SVM	-	-	-	-	-	-
Ritu Kaur, S.Mazumdar [24]	2012	DCT, FFT, PSD	Thresholding	220	10	79.09	90.00	-	-
Arun K. S., K.S.Sarath[7]	2011	RTVTR, Ridge Density	SVM	275	-	-	-	96.00	-

M- Male, F- Female, G- Gender, A- Age, DB- Database, Tr- Training, Te- Testing

From table 2.1, it has been concluded that most of the work has been done using hard computing technologies like Euclidean Distance, Thresholding, etc. and by taking DWT, RTVTR, and DCT as the features from fingerprints. It can be seen that maximum efficiency obtained for Gender Classification is 96% by taking RTVTR and Ridge Density as feature set and training the algorithm over 275 fingerprint samples. On the other hand, maximum efficiency obtained for age group classification is 92% using 100 fingerprint samples and taking DWT as feature set and KNN as classifier. The cost of computation can be reduced by making an algorithm that deeply learns the pattern of fingerprint and predicts the test sample based upon the learning factor.

****

CHAPTER 3
OBJECTIVE AND SCOPE OF STUDY

"In whatever position you find yourself determine first your objective." –Ferdinand Foch

In chapter 2, it has been seen that most of the work done on fingerprints has used hard computing for classifying the gender or age group of person. But the emerging technology has provided various machine learning algorithms which can be implemented on classification based problems in order to provide efficient results by learning various parameters of the input.

3.1 Problem Formulation

Most of the work done on fingerprints for identifying the gender or age group, has been done using hard computing technology which gives less efficient results even with huge database. Hard computing becomes inefficient when the fingerprint database belong to different states or countries. This is the reason that soft computing is the emerging field in classification based problems. Hence, the results obtained by researchers, discussed in chapter 2, can further be improved by using soft computing, machine learning techniques.

3.2 Objective of Study

The study of fingerprints has been an important factor in biometrics, person identification, authentication system, to restrict the entry of person on the bases of gender or age, and many more. Fingerprints has been used for authentication, recognition of person because the fingerprints of person remain same throughout the life of person. The objective of the study of fingerprints mainly includes three steps and these are

Objective 1: to collect the database of the fingerprints of humans,

Objective 2: to determine the gender of person,

Objective 3: to determine to which age group that fingerprint belong to.

Fingerprints has been used to identify the gender because male and female fingerprints patterns vary in some ways like there are more ridges in female in comparison to male, difference in the energy level of the fingerprint image and many more. And for the case of Age identification, fingerprints has been used because as the human grows in age, the size of fingerprint increases and hence the energy of the fingerprint image changes. This change can help to make distinction between the persons of different age groups using fingerprints.

Hence, the prime objective of studying the fingerprints and using the fingerprint to find the gender and age of person is to reduce the work load of the investigating departments all around the world so as to make the process as fast as possible. Many algorithms has been made to increase the efficiency of finding the gender and age of person. In this research work, an algorithm has been developed to estimate the gender and age of person using fingerprints by training the neural network.

3.3 Scope of Study

The study of fingerprints has applications in almost all of the authentication and investigating departments. Study of fingerprints has been helping and will help all the investigating and criminal department in locating their goal. Till date, fingerprints has been used by such departments to find the person, its gender and age group. But still lots of improvements has yet to be done in order to make the algorithm work more efficiently so as the predicted results will be as efficient as possible. The study of fingerprints can be made tremendous by using these learning algorithms and by taking huge database of fingerprint samples from all around the world. The study of fingerprints can further be extended to find the region of person that is whether the person belongs to rural region or urban region. This is due to the reason that the person belonging to the rural region and the person belonging to urban region has different working life and they live in different environmental conditions which leads to changing fingerprints characteristics. Hence, the study of fingerprints finds great scope in criminal investigating department, authenticating departments, to restrict the entry of person in some places based upon gender or age.

3.4 Summary

The objective of the study, as discussed in section 3.2, is to estimate the gender and age group of person from fingerprint samples which has great application in today's world of biometric. The scope of study of fingerprints lies in criminal investigating department, authenticating departments, to restrict the entry of person in some places based upon gender or age.

CHAPTER 4

RESEARCH METHODOLOGY

"Change is the end result of all true learning" – Leo Buscaglia

There has been various technique developed for classifying Gender and Age of Humans based upon different kinds of parameters like Fingerprints, Face structure, Body structure, etc. Among these parameters, fingerprints are the most complex parameter. A human cannot identify whether the given fingerprint sample is of Male or Female. There has been various techniques for obtaining features from the fingerprints of person which can distinguish between male and female. In this chapter, a noble technique of classifying the gender of person from fingerprints has been discussed. Many available techniques has used hard computing as the prime parameter for classifying the gender. But hard computing based classification does not give as much efficient result as the soft computing based classification can give. This is due to the reason that soft computing based algorithms uses learning as the prime factor. It makes the algorithm learn the gender from the features extracted from fingerprints of person. Soft computing is the emerging field also known in terms of machine learning, as the name implies, making the machine learn about has been intended to learn. In this chapter soft computing based gender classification technique has been discussed.

In this chapter, a neural network based Gender and Age classification technique has been proposed which involves training the neural network with the features extracted from the fingerprints of person. The neural network then updates its weights based upon the feature given to it. After complete training of the neural network, the algorithm then can judge the gender and age of person on its own without using any king of thresholding. The basic technique used in neural network is shown in figure 4.1. In the figure, x_1, x_2,…., x_n Are the features of input class. These features are then given to the weight matrix where the w_1, w_2 and so on are the weights. The size of weight matrix is always defined as the total number of input feature sample multiplies by the total number of classes to be predicted. The result obtained from multiplying the input features with the weight matrix is then given to appropriate classifier which then predict the class of input sample. Neural Networks are the best way to make any kind of prediction.

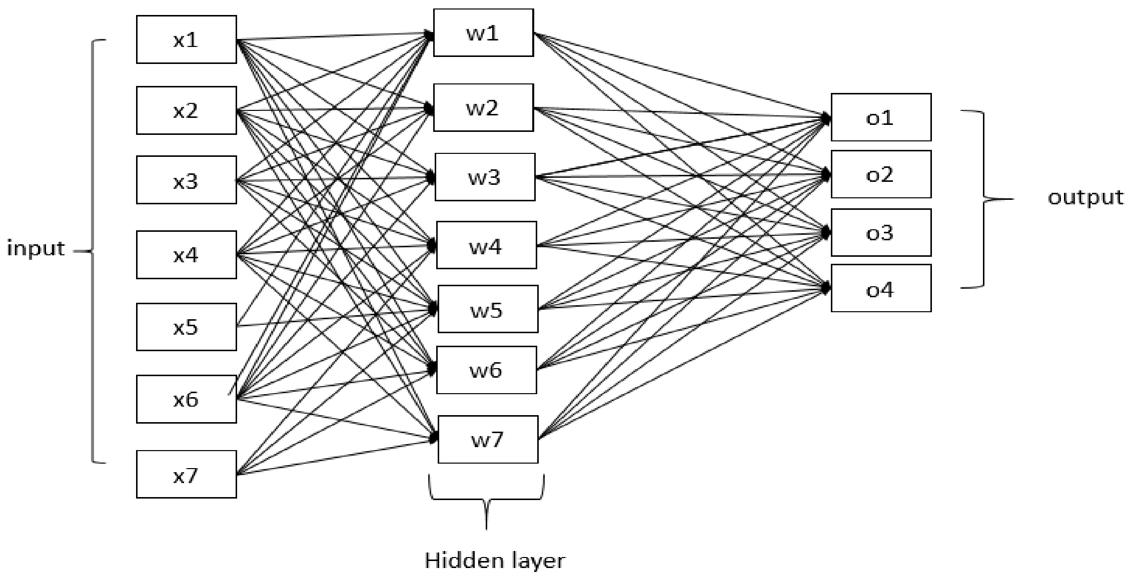


Figure 4.1 Basic Neural Network architecture

4.1 Methodology Used

Fingerprint samples of Male and Female of different age groups has been collected using the inkless fingerprint pad. The total database of fingerprints composed of 120 fingerprint sample of humans and that too of index finger of right hand. The scanned copy of the sheet where the fingerprints has been collected is then passed through the cropping stage where individual sample of fingerprint has been collected and saved in the database. The various phases of algorithm for classifying the gender are

a. Database Collection
b. Pre-processing Phase
c. Feature Collection Phase
d. Neural Network Training Phase
e. Classification Phase

The block diagram of the proposed method has been shown in figure 4.2. In this diagram, it is shown that the fingerprint goes through the pre-processing stage before going for training of the neural network. From the pre-processed fingerprints, feature set has been obtained and these features are then given to the neural network for training. The neural network then predicts the class, whether Gender or Age, of unknown fingerprint sample on its own.

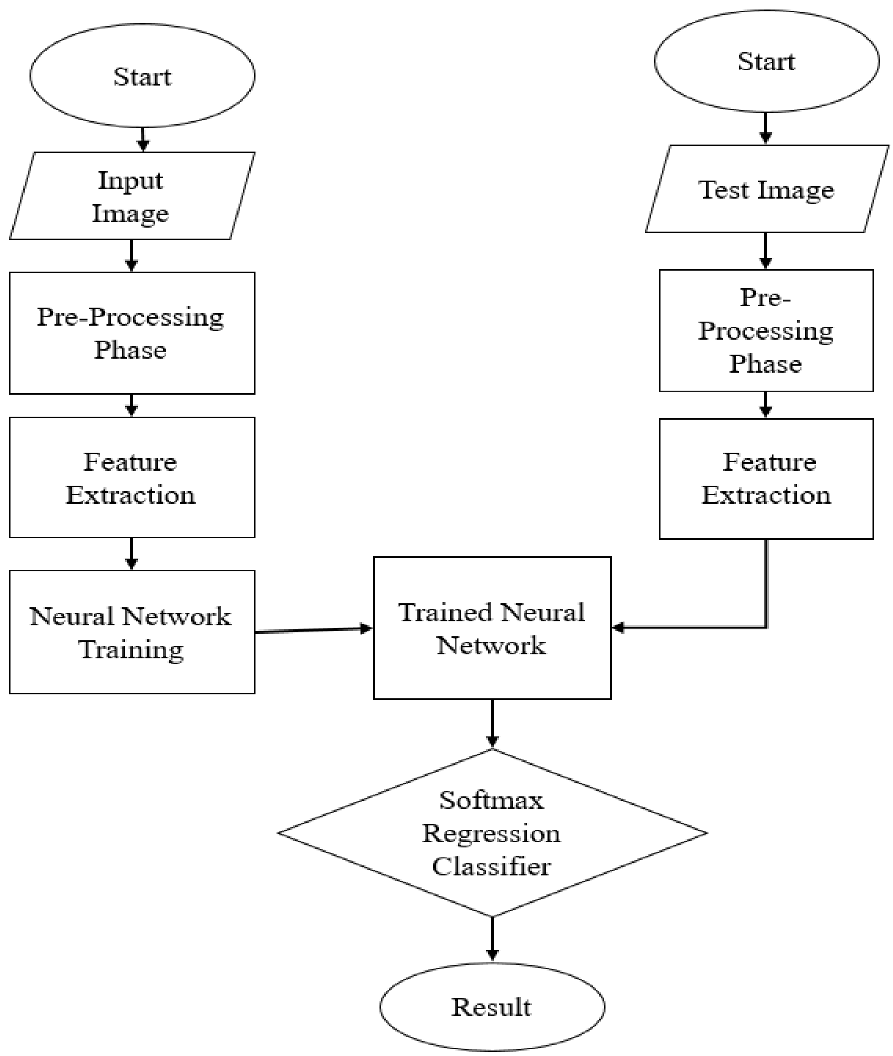


Figure 4.2 Block Diagram of purposed technology

4.1.1 Pre-processing Phase

For any work on images, pre-processing Phase is of prime importance. As the fingerprint samples has been collected using the inkless pad, there are many unwanted pixel values added up to the image which are not desired and that may produce wrong results. So, each sample of fingerprint has been pre-processed. Pre-processing of fingerprint samples has been done using Morphological Image Processing which includes Erosion and Dilation. Morphological image processing performs non-linear operations on the image. Morphological Image Processing Technique is well suited for Binary Image because this technique depends upon the relative ordering of pixels and not on their numeric values. So, Morphological Image Processing is best suited for Fingerprint images which consists of Black and White lines. Morphological image processing make use of a small template called structuring element which is positioned at all the locations in the image and is then compared with the corresponding pixels of the image.

The Basic block diagram of pre-processing Phase of Gender and Age Classification technique is shown in figure 4.3. There are three phases to be performed in Morphological Image Processing and these are

a. Converting the Input image into Binary Image
b. Performing Erode Operation
c. Performing Dilate Operation

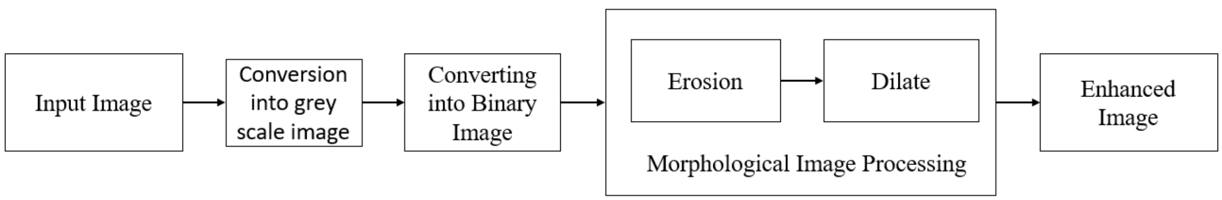


Figure 4.3 Pre-processing Phase

Converting input image into Binary Image- The input image has been converted into Binary Image using a threshold value which has been obtained after analyzing the input image. A threshold value of 117 has been taken to convert the input image into Binary Image. If the input pixel value is greater than 117 then it will assigned 1 to that pixel else 0 to that pixel value.

$$B(i,j) = \begin{cases} 1 \text{ if } I(i,j) > 117 \\ 0 \text{ if } I(i,j) < 117 \end{cases} \tag{4.1}$$

where B(i, j) represent the pixel value of Binary image at i^{th} row and j^{th} column and I(i,j) represents the original Pixel value of input image at i^{th} row and j^{th} column and 117 is the thresholding value.

Erode Operation- The erosion operation involves assigning the minimum value of neighborhood pixels to the pixels coming under the structuring element. In case of Binary image, it will assign 0 to the structuring elements. Let E be the Euclidean Space and A is the Binary image in space E.

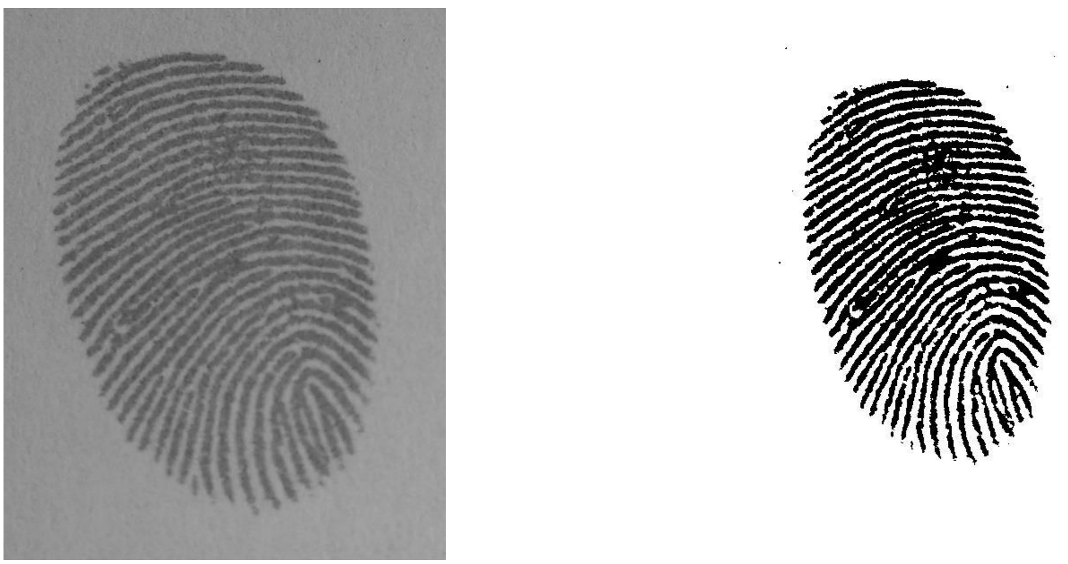
The erosion process of a Binary Image A by a structuring element B can be expressed by equation (4.2) [29].

$$A \odot B = \{ z \in E | B_z \subseteq A \} \tag{4.2}$$

Where B_z is translation of structuring element B by vector z which is given by equation (4.3)

$$B_z = \{ b + z | b \in B \}, \quad for\ all\ z \in E \tag{4.3}$$

Dilate Operation- The dilate operation involves assigning the maximum value of neighborhood pixels to the pixels coming under the structuring element. In case of Binary image, it will assign 1 to the structuring elements. Figure 4.4 shows the change in the input image after performing Morphological Image Processing.



(a) (b)

Figure 4.4 (a) Fingerprint Before Pre-processing, (b) Fingerprint After Pre-processing

The dilation of Input image A by structuring element B with Euclidean space E can be represented by equation (4.4) [29].

$$A \oplus B = \{ z \in E \mid (B^s) \cap A \neq \emptyset \} \tag{4.4}$$

Where B^s represents the symmetric of B.

4.1.2 Feature Extraction Phase

For Gender Classification, two kinds of features has been used in order to obtain the feature vector which has been collected after cropping the input image. The first kind of feature is Ridge thickness to valley thickness Ratio (RTVTR). The fingerprint image is being cropped around the core of finger. Then the ratio of count of number of black pixels to the ratio of count of number of white pixels has been obtained. This ratio is used as one of the feature for an image in feature vector. The other kind of feature used is DWT. The discrete wavelet transform of fingerprint image has been calculated and the energy of DWT of image at the third level is calculated and used as another feature in the feature vector.

For Age Classification, Discrete Wavelet Transform (DWT) has been used to extract features from the fingerprints. A wavelet is a waveform having very limited duration that has a zero average value. In mathematical terms, wavelets are mathematical functions that divides the data into different frequency components, and then studying each component after matching its resolution to its scale. The 2-D wavelet decomposition of an image results in four decomposed sub-band images which may referred to be low–low (LL), low–high (LH), high–low (HL), and high–high (HH). Each of these sub-bands gives different-different details of the image. Most of the energy of images is contained in the low frequency component and hence the decomposition is generally being repeated on the LL sub band only.

For k level DWT, there are (3*k) +1 sub-bands are available. Energy of all these sub-bands is used as features of the image. Energy of all the sub-bands is calculated [4] as

$$E = \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} |x(i,j)|^2 \quad (4.5)$$

Where E is the energy of the sub-domain (LL) of the input image and M, N are the x and y dimensions of the sub-domain. These energy features are calculated and a dictionary has been obtained. 1st and Third level DWT energy feature has been used for training the algorithm for Age. Figure 4.5 shows the DWT of input image upto 3 levels.

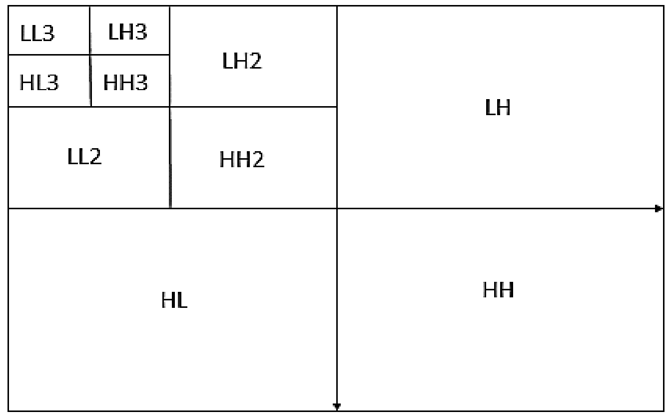


Figure 4.5 DWT of Input Image

4.1.3 Neural Network Training

Training of Neural Network implies updating the weights of the hidden layer after each and every iteration involving fingerprint samples. These updated weights are then used to predict the test fingerprint sample by simply multiplying the weight matrix with the fingerprint feature vector. The proposed method uses Back Propagation Method with Gradient Descent as an optimizer to train the neural network using Tensorflow library of Python programming language. Gradient Descent optimizer is an activation function for the algorithm which minimizes the cross entropy with a learning rate of 0.5. There are many other optimizers available.

Gradient Descent method is a first order iterative algorithm used for optimization. This algorithm is used to find the minimum value of the function. Gradient Descent method is similar to steepest Descent method. It is a simple method used by tensorflow to shift the function in the direction of minimal cost which is defined as the training step of the algorithm. The algorithm shifts the function towards the minimal cost every time it iterates over the algorithms. The back propagation algorithm updates the weights of the hidden layers for thousand iterations in order to reduce the error. The Gradient descent method can optimize the back propagation by using the equation (4.6) [32].

$$\frac{df(u)}{du} = f(u)(1 - f(u)) \tag{4.6}$$

$$f(u) = \frac{1}{1+ e^{wx}} \tag{4.7}$$

Where f(u) is the logistic function given in equation (4.7), w is the weight matrix and x is the input feature. Back Propagation as the name implies it propagates backwards in response to the error obtained after moving in forward direction. The error obtained then used to update the weights of the previous hidden layer hence called propagating backward. The fingerprint feature set is being given in the form of batches. These batches are randomly generated so as to get maximum out of the algorithm. The training of algorithm can be done upto as many epochs as it can be so that the error gets minimized. The back propagation algorithm has three stages and these are

Step1: Feed forward stage
Step2: Error calculation and propagating backward
Step3: Weight updating stage

Feed Forward Stage, as the name suggests, the algorithms moves in the forward direction that is from the input layer towards the output by passing through the hidden layer. In this stage, the input features are multiplied with the weight matrix of hidden layer. The output of hidden layer gives the output for the feed forward stage. As shown in figure 4.6.

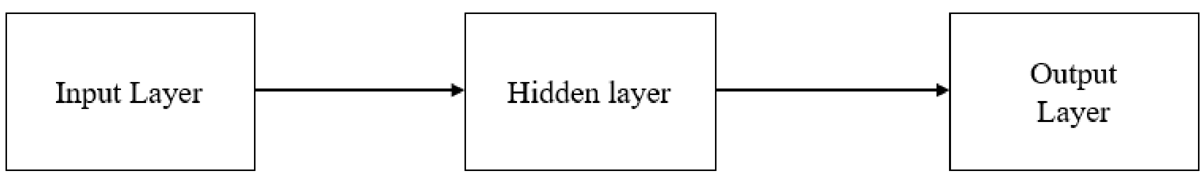


Figure 4.6 Feed Forward stage

Feed Forward stage can be numerically represented using equation (4.8) [30] defining the weight matrix multiplication with the input samples.

$$Z_{inj} = b + \sum_{ij} x_i v_{ij} \tag{4.8}$$

Where Z_{inj} represents the output of feed forward phase, b is the bias to the weights, x represents the input samples and v represents the coefficients of weight matrix.

Error calculation and propagating backward is the stage where based upon the output of the feed forward stage and the actual output for that fingerprint sample, error is being calculated. This error is then propagated backwards and updated the weights of the hidden layer so that for the next fingerprint sample, there will be less error. Hence, the next feature of next fingerprint will be multiplied with the updated weight matrix. The error can be calculated using the formula given in equation (4.9) [32]

$$\Delta(k) = (t_k - y_k) \tag{4.9}$$

Where $\Delta(k)$ represents the error for the k^{th} sample, t represents the target value for the k^{th} sample and y represents the value obtained for the k^{th} sample.

Weight updating stage performs the updating process of the weights of the hidden layer based upon the calculated error. If the error is negative then the weights of the hidden layer undergoes addition and if the error is positive then the weights of hidden layer undergoes subtraction. The forthcoming fingerprint feature set will then be multiplied with the updated hidden layer and hence there will be reduction in the error during the training process. For each output unit, update the weights and biases based upon the change in weights and biases [30]

$$w_{jk}(new) = w_{jk}(old) + \Delta(w_{jk}) \tag{4.10}$$

$$w_{0k}(new) = w_{0k}(old) + \Delta(w_{0k}) \tag{4.11}$$

w in the above equations represents the coefficients of weight matrix.
For each of the hidden layer, weights and bias can be updated using equation (4.12) and (4.13).

$$v_{ij}(new) = v_{ij}(old) + \Delta(v_{ij}) \tag{4.12}$$

$$v_{0j}(new) = v_{0j}(old) + \Delta(v_{0j}) \tag{4.13}$$

Where v represents the weight coefficients of the hidden layer.

4.1.4 Classification Phase

After the training of the algorithm, the next stage comes is the classification stage where the test sample will be classified using certain classifiers and its class will be predicted. In this proposed technology, Softmax Regression classifier has been used to classify the class of fingerprint sample. The test fingerprint sample undergoes feed forward stage where the test feature set of test image is being multiplied with the trained neural network obtained by using back propagation algorithm with Stochastic Gradient Descent as optimizer. The generated output is then passed through the Softmax regression classifier which then computes the class for that fingerprint sample. If there are only two output classes then this classifier is known as linear regression classifier. Hence, for classifying the gender of person, the classifier used is linear regression classifier and for classifying the age of person, the classifier used is Softmax regression classifier. The basic block diagram for Softmax regression classifier is shown in figure 4.7.

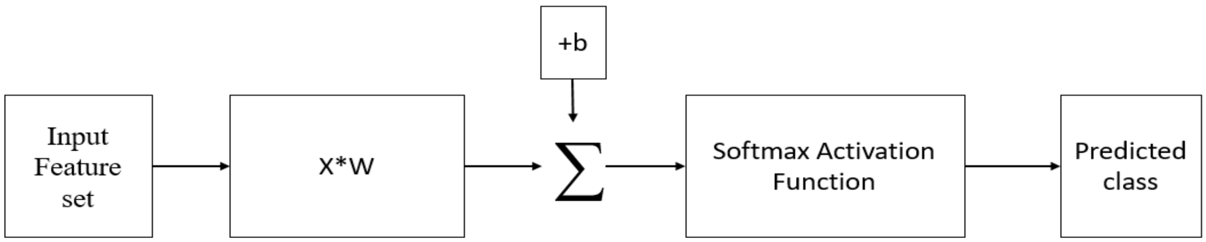


Figure 4.7 Softmax Regression Classification

Where b represents bias added to the input feature set and x*w represents the dimensions of the weight matrix. The Softmax regression classifier can be represented with an equation as [31]

$$softmax(x) = {e^x}/{e^y} \qquad (4.14)$$

Where x represents the present value of one of the output of the output vector and y represents the sum of all the values in the output vector.

4.2 Summary

A fingerprint based gender and age classification technique has been proposed which uses back propagation as the training algorithm with Stochastic Gradient Descent as optimizer and Softmax Regression as the classifier. The method is a supervised learning algorithm in which both the input and output has been provided for classification. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network.

****

CHAPTER 5

EXPERIMENTAL RESULTS

"Be the change that you wish to see in the world." – Mahatma Gandhi

In this chapter, the proposed methodology has been analyzed with the database of fingerprint images of both male and female from different age groups. These fingerprint images has been collected from the people of India using inkless pad. A total of 120 fingerprint images has been collected of both male and female. These fingerprint images consists of samples from different age groups between 10 to 50. All these fingerprint samples has been tested with the algorithm designed to find the gender and age of person.

The proposed algorithm consists of four steps to train the neural network for both gender and age groups of person using the features from the fingerprint samples. These features has been extracted after pre-processing of fingerprint samples in order to remove extra noise from the fingerprints. Following are the four steps used in order to make the algorithm learn different patterns of fingerprints for both gender and age groups.

Step1: Pre-processing of fingerprint images
Step2: Feature Extraction of test fingerprint samples
Step3: Input to Neural Network
Step4: Results with test samples

5.1 Pre-processing of fingerprint images

Pre-processing of fingerprint images has been done so as to efficiently compute or extract features from the fingerprints. The Pre-processing of images results in removal of noise from fingerprints which may be in the form of extra pixels and cropping of the image to a particular defined region which contains most of the information. Pre-processing to fingerprint images has been done in three steps and these are

Step1: Converting input image into Binary Image
Step2: Morphological Image Processing
Step3: Cropping of Image

Conversion of input image into Binary Image has been done using Python Program so as to efficiently apply the Morphological Image Processing over the fingerprint image which has been proved efficient for Binary Images.

Morphological Image Processing has been used to remove the background noise from fingerprint images. The process includes Erosion and Dilation. Morphological image processing performs non-linear operations on the image. Morphological image processing make use of a small template called structuring element which is positioned at all the locations in the image and is then compared with the corresponding pixels of the image. Morphological Image processing has been used because morphological image processing is good in removing the noise from binary images. The pre-processing of fingerprint images resulted in removal of noise like extra background pixels from the images. The results of removal of noise from fingerprints has been shown in Figure 5.1 for both male and female fingerprint sample.

(a) (b)

Figure 5.1 (a) Pre-processed Male fingerprint image; (b) Pre-processed Female fingerprint image

These fingerprints images after pre-processing only contains the black and white pixels representing the ridges and valleys respectively of the fingerprint images. Hence, there is no extra black pixels in the background which if present may result in the wrong analyses of fingerprint images.

Cropping of Fingerprint image for gender classification has been done around the core of fingerprint. This is due to the reason that most of the information in fingerprint has been contained in the part around the core in order to distinguish between genders of person.

Figure 4.2 shows the fingerprint images of both male and female after the cropping process. These cropped fingerprint images has then been given to the feature extraction stage for extracting the desired features from it for both gender.

(a) (b)

Figure 5.2 (a) Cropped Female fingerprint; (b) Cropped Male fingerprint

5.2 Feature Extraction

After pre-processing of fingerprint samples, the next stage has been feature extraction stage. These extracted features has then been used to train the neural network. The feature extraction process is different for gender classification and age group classification. This is due to the reason that gender and age group classification are two different processes including different patterns of fingerprints. So, different patterns for both gender and age groups has been produced so as to efficiently computer the classifications for both gender and age.

Feature Extraction stage comprises two parts for both Gender and Age groups. For gender classification, features used are Ridge to Valley Thickness Ratio (RTVTR) and 2-D Discrete Wavelet Transform (DWT) at third level of decomposition. On the other hand, for age group classification, features used are 2-D Discrete Wavelet Transform (DWT) at first and third level of decomposition. Both the processes of feature extraction has been explained in the preceding sections.

5.2.1 Feature Extraction for Gender Classification

For gender classification, Ridge to valley Thickness Ratio (RTVTR) and 2-D Discrete Wavelet transform (DWT) at its third level of decomposition has been used as features for training the neural network. All the fingerprint images, after passing through the pre-processing stage, has been used to calculate the features. RTVTR has been calculated by finding the ratio of count of ridges in the fingerprint to the ratio of count of valleys in the defined part of fingerprint around the core. The second kind of feature which is DWT has been obtained by finding the Discrete Wavelet Transform of image upto its third level of decomposition. Following Table 5.1 shows the Features extracted from the fingerprints.

These features are then given to the neural network for the learning process. The neural network implemented using back propagation with stochastic gradient descent as optimizer learns such features and hence use this learning while predicting the class of test fingerprint sample. Following tables clearly shows that the RTVTR value of Female fingerprint is more in comparison to that of male fingerprint while the Energy of 2D DWT at the third level of decomposition is more in male than female.

Table 5.1 Features from Female fingerprints

S. No.	Gender	RTVTR	Energy of 2-D DWT at third level
1	Female	1.511	0.034
2	Female	0.770	0.048
3	Female	1.120	0.039
4	Female	1.628	0.029
5	Female	0.880	0.046
6	Female	1.531	0.031
7	Female	0.735	0.047
8	Female	1.875	0.031
9	Female	0.740	0.048
10	Female	1.730	0.032

S. No.	Gender	RTVTR	Energy of 2-D DWT at third level
11	Male	0.301	0.059
12	Male	0.272	0.059
13	Male	0.259	0.062
14	Male	0.507	0.054
15	Male	0.611	0.049
16	Male	0.443	0.053
17	Male	0.397	0.055
18	Male	0.502	0.052
19	Male	0.597	0.047
20	Male	0.356	0.057

5.2.2 Feature Extraction for Age Classification

For age classification, energy of 2-D Discrete Wavelet transform (DWT) at its third level of decomposition and at first level of decomposition has been used as features for training the neural network. All the fingerprint images, after passing through the pre-processing stage, has been used to calculate the features. Following Table 5.2 shows the Features extracted from the fingerprints belonging to different age group.

Table 5.2 Features from fingerprints belonging to different age group

S. No.	Age Group	Energy at 1st level 2D DWT	Energy at 3rd level 2D DWT
1	10-20	35.6038	9.649
2	10-20	35.9505	9.7271
3	10-20	35.9695	9.9188
4	10-20	35.9695	9.9188
5	10-20	36.6358	10.1064
6	10-20	36.5301	10.8060

S. No.	Age Group	Energy at 1st level 2D DWT	Energy at 3rd level 2D DWT
7	10-20	35.4375	10.0179
8	10-20	36.0360	9.9344
9	10-20	36.0143	9.9384
10	10-20	33.5257	9.0439
11	21-30	35.7048	9.8707
12	21-30	29.6836	7.7656
13	21-30	34.9065	9.3757
14	21-30	33.3789	8.8834
15	21-30	29.6836	7.7656
16	21-30	33.4650	8.9723
17	21-30	34.8102	9.3873
18	21-30	34.0896	9.4531
19	21-30	35.7696	9.7654
20	21-30	34.1787	9.2919
21	31-40	31.5078	9.3833
22	31-40	30.2408	8.8554
23	31-40	32.6169	10.0844
24	31-40	33.0813	9.3956
25	31-40	32.7636	9.5606
26	41-50	35.5699	9.8517
27	41-50	32.5570	9.1881
28	41-50	35.3896	9.7843
29	41-50	32.5570	9.1881
30	41-50	33.5348	9.8461

5.3 Input to Neural Network

All the features extracted in the feature extraction stage has been given to the neural network. The neural network uses the updated weight matrix, which has been obtained during the training of the neural network with training images, to predict the class of the test fingerprint samples. The process has been done separately for gender and age group prediction. This is due to the reason that the features for gender prediction and for age prediction are different. The trained neural network then predicts the class for the test fingerprint.

5.4 Results with test fingerprint samples

For gender classification, the trained neural network has been tested with 20 unknown fingerprint samples consisting of 10 for female fingerprints and 10 for male fingerprints. Among these 20 unknown fingerprint samples, 18 has been predicted correctly by the neural network giving an efficiency of 90%. Figure 5.3 shows the graphical representation of Gender Classification.

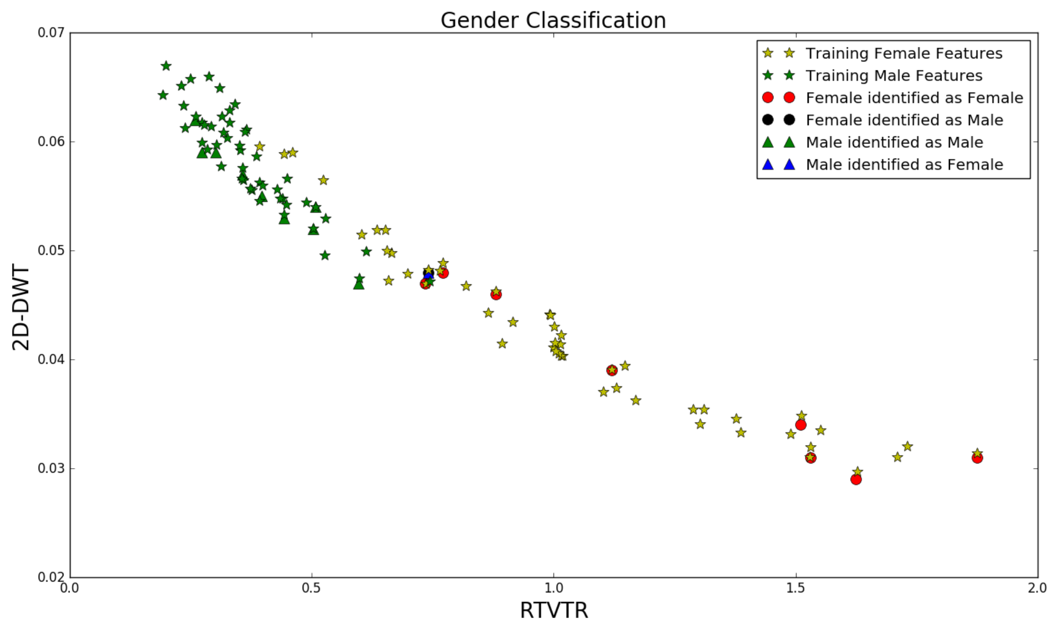


Figure 5.3 Gender Classification

The performance of the proposed algorithm for Age classification can be represented by confusion matrix as shown in Table 5.3

Table 5.3 Confusion matrix for Gender Classification

Predicted → **Actual ↓**	**Male**	**Female**
Male	9	1
Female	1	9

For age classification, the trained neural network has been tested with 20 unknown fingerprint samples consisting of each age group with 5 samples from each of them. Among these 20 unknown fingerprint samples, 14 has been predicted correctly by the neural network giving an efficiency of 70%. Figure 5.4 shows Age identification method showing the number of samples being correctly identified and the number of samples wrongly identified.

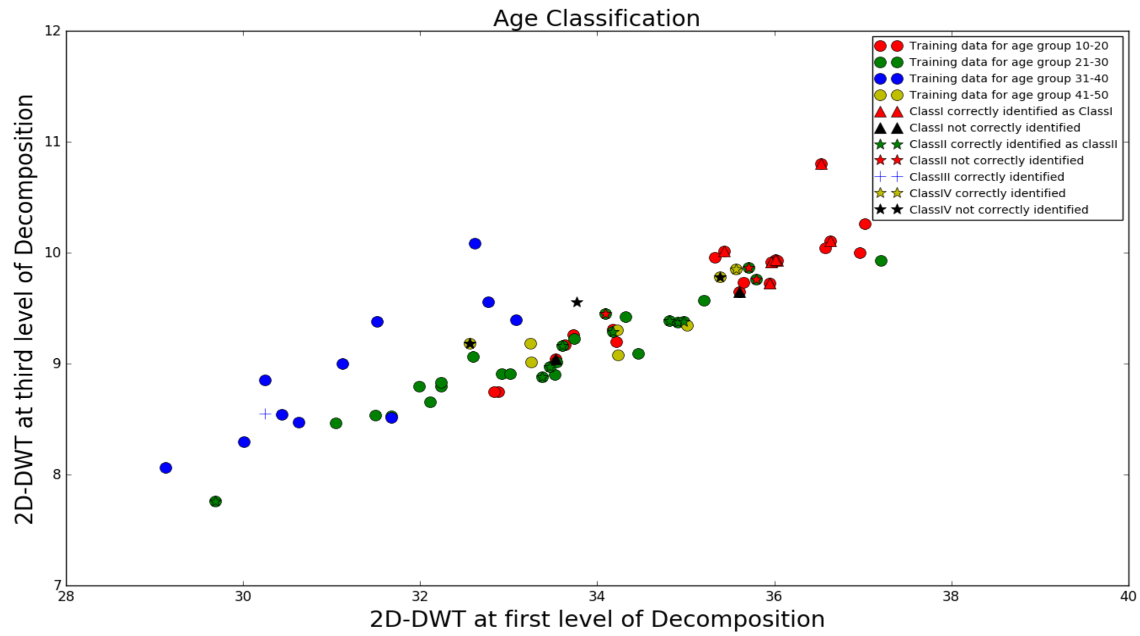


Figure 5.4 Age Classification Results

The performance of the proposed algorithm for Age classification can be represented by confusion matrix as shown in Table 5.3

Table 5.4 Confusion matrix for Age classification

Predicted⟶ **Actual ↓**	**10-20**	**21-30**	**31-40**	**41-50**
10-20	8	2	0	0
21-30	3	7	0	0
31-40	0	0	5	0
41-50	0	2	2	1

5.5 Comparison with other Results

The Results obtained by the proposed method has been compared with the results obtained by other researchers during the time immemorial. In this table, it has been defined that most of the work done by using Hard Computing for classifying the class of the input samples. But in this proposed method, Neural Network based technology has been given which has produced efficient results. Table 5.5 shows the work done on fingerprint based gender and age identification till date and the work done by the proposed technology.

Table 5.5 Comparison with Research Papers

Name of Author	Year	Features	Classifier	DB Size		Accuracy in %age			
				Tr	Te	M	F	G	A
Heena Bawa and Ajmer Singh	-	RTVTR, DWT	Softmax Regression Classifier	100	20	90.00	90.00	90.00	
Heena Bawa and Ajmer singh	-	DWT	Softmax Regression Classifier	90	30	-	-	-	70.00
S. Gornale, et al [25]	2016	DWT, Gabor filter	KNN	740	-	-	-	89.71	-
A. Gour, D. Roy [26]	2016	DWT	KNN	-	-	-	-	90.00	92.00
S.S. Gornale, et al [13]	2015	Texture Features	LDA	600	-	-	-	92.00	-
M. S.Hossain,Habib [3]	2015	RTVTR	Thresholding	400	-	85.00	74.00	-	-
M. K. Shinde, S.Annadate [4]	2015	DWT	KNN	1000	-	80.46	76.84	-	-
Suchita Tarare, et al [5]	2015	DWT	KNN	-	-	-	-	74.00	-
A. Chauhan, S. Choubey [14]	2015	DWT	Euclidean Distance	-	-	-	-	-	-
Akhil Anjikar, et al [15]	2015	DWT, DCT	Euclidean Distance	800	-	65.25	45.5	-	-
H. Agarwal, et al [16]	2014	RTVTR	SVM	300	-	81.00	81.00	-	-
Aditya K. Saxena, V.K. Chaurasia [6]	2014	Gabor Filter	KNN	350	100	-	-	-	-
T. Arulkumara, et al [17]	2014	DWT, PCA	KNN	-	-	-	-	-	68.00
S. Gupta, A. Rao [18]	2014	DWT	ANN	550	-	-	-	91.00	-
S. Gornale, Kruthi [19]	2014	FFT	Thresholding	4320	-	-	-	60.00	-
Ravi Wadhwa, et al [20]	2013	RTVTR, DCT	KNN	100	-	-	-	-	-
P. Chand,Sarangi [21]	2013	DWT, SVD	KNN	100	-	-	-	80.00	-
R.J. Tom, T.Arulkumaran[22]	2013	DWT, PCA	KNN	400	-	-	-	70.00	-
S. S. Ponnarasi, M.Rajaram [23]	2012	Ridge Density	SVM	-	-	-	-	-	-

Ritu Kaur, S. G.Mazumdar [24]	2012	DCT, FFT, PSD	Thresholding	220	10	79.09	90.00	-	-
Arun K. S., Sarath [7]	2011	RTVTR, Ridge Density	SVM	275	-	-	-	96.00	-

M- Male, F- Female, G- Gender, A- Age

From table 5.5, it has been found that DWT, Ridge Density, DCT are the features which contain most of the information of the fingerprints. It can be seen that highest efficiency has been obtained by using DWT and Ridge thickness to valley thickness ratio as features for the algorithm. The proposed method has used RTVTR and DWT as feature for Gender Classification and DWT for Age classification. The results obtained has been compared with the previous researcher's results and it has been found that the proposed method which is a supervised learning algorithm has produced much efficient results with much reduced computational cost by using less database. The efficiency of the proposed method can be further increased if the algorithm is made to train with fingerprints of person belonging to different regions, states and countries. This will make the algorithm learn all the kinds of fingerprints the world have.

5.6 Summary

The proposed algorithm has produced an efficiency rate of 70% for Age classification and 90% for Gender classification. The graph shown in Figure 5.3 and figure 5.4 represents the success and failure rate for the algorithm. The results for this algorithm can further be improved by increasing the database of fingerprint samples. More the different kinds of fingerprint samples, more will be the learning categories for the neural network and hence more accurate will be the results while testing it with unknown fingerprint samples.

****

CHAPTER 6
CONCLUSION AND FUTURE SCOPE

"Education is the key to unlock the golden door of freedom." – George Washington Carver

In this chapter, the proposed methodology with results has been concluded and it has been stated what will be the other work that can be done by using this technique.

6.1 Conclusion

The back propagation neural network with stochastic gradient descent optimizer has been trained with 120 samples of fingerprints consisting of both male and female fingerprint samples belonging to different age groups. After the training process, the updated neural network has been tested with unknown test fingerprint samples. The algorithm has produced efficient results with 90% efficiency for gender classification and 70% efficiency for age group classification.

The results for this algorithm can further be improved by increasing the database of fingerprint samples. More the different kinds of fingerprint samples, more will be the learning categories for the neural network and hence more accurate will be the results while testing it with unknown fingerprint samples.

6.2 Future Scope

The work on fingerprints has not been limited to finding the gender and age of person. The work can further be extended to finding the region of person that is whether the person belongs to rural area or urban area by extracting appropriate features from fingerprints. This work will help the forensic department to locate the criminal cases after finding any fingerprint sample. Such predictions about fingerprints can help to reduce the work of the investing department in solving the criminal cases.

****

REFERENCES

[1] R. Kaur, P. S. Sandhu and A. Kamra, "A novel method for fingerprint feature extraction," *2010 International Conference on Networking and Information Technology*, Manila, 2010, pp.1-5.

[2] K. S. Arun and K. S. Sarath, "A machine learning approach for fingerprint based gender identification," *2011 IEEE Recent Advances in Intelligent Computational Systems*, Trivandrum, 2011, pp. 163-167

[3] M. S. Hossain and M. A. Habib, "Improving fingerprint based gender identification technique using systematic pixel counting," *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Dhaka, 2015, pp. 1-4.

[4] M. K. Shinde and S. A. Annadate, "Analysis of Fingerprint Image for Gender Classification or Identification: Using Wavelet Transform and Singular Value Decomposition," *2015 International Conference on Computing Communication Control and Automation*, Pune, 2015, pp. 650-654.

[5] S. Tarare, A. Anjikar and H. Turkar, "Fingerprint Based Gender Classification Using DWT Transform," *2015 International Conference on Computing Communication Control and Automation*, Pune, 2015, pp. 689-693.

[6] A. K. Saxena and V. K. Chaurasiya, "Fingerprint based human age group estimation," *2014 Annual IEEE India Conference (INDICON)*, Pune, 2014, pp. 1-4.

[7] K. S. Arun and K. S. Sarath, "A machine learning approach for fingerprint based gender identification," *2011 IEEE Recent Advances in Intelligent Computational Systems*, Trivandrum, 2011, pp. 163-167.

[8] D. Arpit and A. Namboodiri, "Fingerprint feature extraction from gray scale images by ridge tracing," *2011 International Joint Conference on Biometrics (IJCB)*, Washington, DC, 2011, pp. 1-8.

[9] Davide Maltoni, Dario Mani and Anil Jain, “Securing Fingerprint systems” in title of Handbook of fingerprint recognition, Springer Publications, Verlag London, 2nd edition 2009, pp. 371-380.

[10] A. Goel and V. P. Vishwakarma, "Feature extraction technique using hybridization of DWT and DCT for gender classification," *2016 Ninth International Conference on Contemporary Computing (IC3)*, Noida, 2016, pp. 1-6.

[11] John G. Prokaios, “The discrete Fourier Transform” in title of Digital Signal Processing, Pearson Education Publication, 4th edition 2007, India, pp. 449-552

[12] Richard D. Lyons, “Discrete Fourier Transform” in title of Understanding Digital Signal Processing, Prentic Hall Publication, 3rd edition, India, pp. 63-100, 11 November 2010

[13] S.S.Gornale, Mallikarjun, Rajmohan Pardershi, Kruthi R “ Harlick Feature Descriptors for Gender Classification Using Fingerprints ” in *International Journal of Advance Research in Computer Science and Engineering*, Volume 5, Issue 9, pp. 72-78, 2015

[14] Alok Chauhan, Akhil Anjikar and Suchita trare “ Study of Ridge Based and Image Based approach for Fingerprint gender Classification ” in *International Journal of Innovative Research in Computer and Communication Engineering*, volume 3, Issue 3, pp. 1599-1604 March 2015

[15] Akhil Anjikar, M.M. Goswami and Suchita tarare “ Fingerprint Based Gender Classification using Block-Based DCT ” in *International Journal of Innovative Research in Computer and Communication Engineering*, volume 3, Issue 3, pp. 1611-1618, March 2015

[16] Heena Agarwal and Siddhartha Choubey “ Fingerprint based Gender Classification Using Multi-class SVM ” in *International Journal of Advanced Research in Computer Engineering & Technology*, Volume 3, Issue 8, pp. 2575-2580, August 2014

[17] T. Arulkumaran, Dr.P.E.Sankaranarayanan and Dr.G.Sundari “ Fingerprint Based Age Estimation using 2D Discrete Wavelet Transform and Principal Component Analyses ” in *International Journal of Advance Research in Electrical, Electronics and Instrumentation Engineering*, volume 2, Issue 3, pp. 1060-1066, March 2013

[18] Sampta Gupta and A. Prabhakar Rao “ Fingerprint Based Gender Classification using Discrete Wavelet Transform and Artificial Neural Network ” *in Journal of Computer Science and Information Technology* (IJCSMC) volume 3, Issue 4, pp. 1289-1296, April 2014

[19] S.S.Gornale and Kruthi R. “ Fusion of Fingerprint and age biometric for gender classification using frequency and texture analyses ” in *Signal and Image Processing: An international Journal (SIPIJ)* volume 5, Issue 6, pp. 75-85, December 2014

[20] Ravi Wadhwa, Maninder Kaur and Dr. K.V.P. Singh " Age and Gender estimation from fingerprints using RVA and DCT coefficients " in *IOSR Journal of Engineering*, volume 3, Issue 8, pp. 05-09, August 2013

[21] Pallavi Chand and Shubhendu Kumar Sarangi " A novel Method for Gender Classification Using DWT and SVD Techniques " in *International Journal of Computer Technology and Applications*, volume 4, Issue 3, pp. 445-449, June 2013

[22] Rajio Jackson Tom, T. Arulkumaran " Fingerprint Based Gender Classification using 2D Discrete Wavelet Transform and Principal Component Analyses " in *International Journal of Engineering Trends and Technology*, volume 4, Issue 2, pp. 199-203, 2013

[23] S. Sudha Ponnarasi and M. Rajaram " Gender Classification system derived from Fingerprint Minutiae Extraction " *in International Journal of Computer Application* 2012

[24] Ritu Kaur and Susmita Gaur Mazumdar " Fingerprint Based Gender Identification using Frequency Domain Analyses " in *International Journal of Advances in Engineering and Technology,* volume 1, Issue 1, pp 295-299, 2012

[25] Shivanda Gornale, Abhijit Patel and Veersheety C. " Fingerprint Based Gender Identification using Discrete Wavelet Transform and Gabor Filter " in *International Journal of Computer Applications,* volume 152, Issue 4, pp. 34-37, October 2016

[26] Akanccha Gour and Dharmendra Roy " Increasing Accuracy of Age and Gender Detection by Fingerprint Analyses using DCT and DWT " in *International Journal of Innovative research in Computer and Communication Engineering*, volume 4, Issue 5, pp. 8452-8458, May 2016

[27] Ms. Sonali and Ms. Priyanka Wankar "Research Paper on Basic of Artificial Intelligence" in *International Journal on recent and Innovation Trends in Computing and Communication*, volume 2, Issue 1, pp. 96-100, January 2014

[28] V.K. Srivastava and Dr. Devendra Prasad "DWT- Based Feature Extraction from ecg signal" in *American Journal of Engineering Research*, volume 02, Issue 03, pp. 44-50, 2013

[29] A.M. Raid, W.M. Khedr, M.A. El-dosuky and Mona Aoud "Image Restoration based on Morphological Operations" in *International Journal of Computer Science, Engineering and Information Technology,* volume 4, Issue 3, pp. 9-21, June 2014

[30] Dr. Rama Kishore and Taranjit Kaur "Back propagation: An artificial Neural Network Approach for pattern Recognition" in *International Journal of Scientific and Engineering Research*, volume 3, Issue 6, June 2012.

[31] Chao-Ying Joanne Peng, Kuk Lida Lee and Gary M. Ingersoll "An introduction to Logistic Regression Analysis and Reporting" in *The Journal of Educational Research*, volume 96, Issue 1, pp. 3-14, October 2002.

[32] Er. Parveen Sehgal, Dr. Sangeeta Gupta and Prof. Dharminder Kumar "Minimization of error in Training of Neural Network using Gradient Descent Method" in *International Journal of Technical Research* volume 1, Issue 1, 2012, pp. 10-12.

****

INDEX