Statistical Features based Gender Identification using SVM

Shivanand S Gornale[1] Abhijit Patil[2] Prabha[3]
[1,2,3]Department of Computer Science,
[1,2]Rani Channamma University, Belagavi [3]KASC College, Bidar

Abstract— The gender identification based on any biometrics is essential to build applications such as human-computer interaction, content-based indexing, searching, targeted advertising, surveillance, biometrics, and demographic studies. The Fingerprints are one of the most notable biometric technologies and are considered as legitimate evidence all over the world to identify a person. The Fingerprint traits can be used effectively for the estimation of gender information. This paper presents a technique for gender identification based on simple statistical properties of the fingerprints. These are extracted from 740 Fingerprint images. A set of 15 statistical properties of a Fingerprint image is used to form a feature vector. Then, the traditional Support Vector Machine (SVM) is employed to draw the decision using the feature metrics of size 740x15. The experimental results stood at 87% and 89.15% as encouraging evidence.

Key words: Gender Identification, Statistical Features, Fingerprints, SVM, QDA, LDA

I. INTRODUCTION

Human vision system is self-sufficient and quite adaptive in identifying the demographical attributes of human beings such as age and gender. For example, the first moment of life we try to recognize mother and father. We constantly perform gender recognition, often without being aware of it. However, the computer vision system is not fully mature to do all these tasks. Consequently, automatic identification of age and gender of human beings plays a vital role in developing various applications such as human-computer interaction, content-based indexing, searching, targeted advertising, surveillance, biometrics, and demographic studies. In particular, gender identification based on any biometrics bisects the search space. Fingerprint is one of the important biometrics can be considered for gender identification. The fingerprints are widely used biometric for individual identification. It is a simple and prominent biometrics because fingerprints are unvarying but its size and shape may change and differs with age, but basic patterns of the fingerprint remains unchanged. Joao De Barros, a [1] European explorer is credited with recording the first known system of fingerprinting in the 14th century, Since then biometric systems have been successfully deployed in a number of real-world applications, but biometrics is not yet a fully solved problem. In this paper, we aim to examine the impact of gender identification on the enhancement of the performance of the biometric system by using biometric traits i.e. fingerprint.

The rest of this paper is organized as follows: In Section 2, the brief related work is reported. The feature extraction and the flow of work carried out are placed in Section 3. Experimental results are discussed in section 4 and concluding remarks are reported in Section 5.

II. RELATED WORK

Manish Verma, et al, [14] proposed methodology for Gender identification from fingerprints using ridge density. For experimentation internal database of 400 fingerprints is created in which 200 were male fingerprints and 200 were female fingerprints, using SVM classifier the accuracy obtained is around 89%.

Naveen Kumar Jain et al [16] proposed methodology for gender Identification used SVD and frequency Doman vector to extract the features of 100 fingerprint images from left hand index fingers of each 50 male and 50 female, The KNN classifier is used and obtained the accuracy of 80%.

R Jackson et al[17] proposed methodology features are computed through the combination of DWT and PCA on the dataset of 400 fingerprints images of different age and gender is collected out of which 200 are of male and 200 are of female fingerprints ,the minimum distance method is used and the overall success rate of classification in age estimation was around 70%.

Pallavi C et al [18] proposed methodology fingerprint based gender classification FDA and 2D DWT are combined for preparing the feature vector, on internal database of 100 persons left hand index finger print are stored; out of 50 male and other 50 are of female fingerprint collected from different age and gender, using KNN classifier the obtained classification rate of 80%.

RituKaur et al [7] proposed method that extracts feature by FFT, DCT and PSD transformation techniques from the internal dataset of 220 persons fingerprint images of different age and gender is collected. An optimal threshold value is used to classify the gender have obtained accuracy of 90 % 79.07% respectively.

Gnanasivam P et al[5] Proposed methodology has focused on male and female identification using fingerprint through frequency domain analysis to estimate male and female by analyzing fingerprints. FFT, DCT and PSD are used to extract features on a database of 400 persons of different age and gender. A threshold value is used to classify the gender .The accuracy of 92.88 % for male and 94.85 % for female was obtained and overall accuracy obtained is of 88.28%.

Ahmed Badawi, et al, [15] proposed a methodology for Gender classification from fingerprints a dataset 2200 persons of different age and gender (1100 males and 1100 females) are stored and Features extracted by; ridge count, ridge thickness to valley thickness ratio (RTVTR). Fuzzy - C Means (FCM), LDA, NN classifiers are used for the classification and obtained results of 80.39%, 86.5%, and 88.5% using FCM, LDA, and NN, respectively.

III. PROPOSED METHODOLOGY

The proposed method undergoes a sequence of steps: Firstly, Fingerprint image is preprocessed to yield a clean image. Secondly, feature extraction algorithm is applied on the resulting image to obtain feature matrix. Thirdly, a feature matrix is used as an input to the classifier to classy the images as male and female. The block diagram of gender identification system is shown in Figure 1.

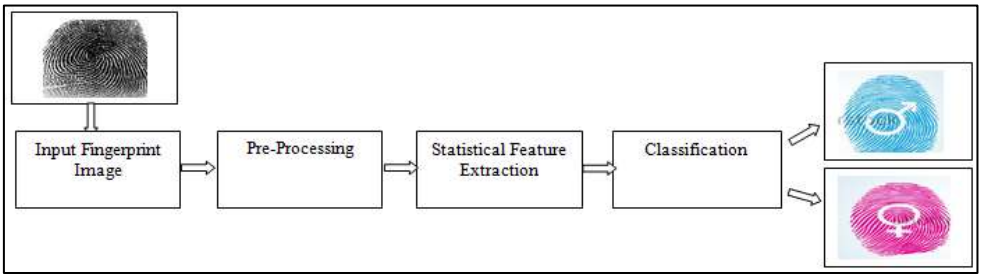

Fig. 1: The block diagram of gender identification system

A. Image acquisition

As per the study there is no standard database which contains fingerprints of male and female. So we have created our own database for conducting experiment on gender classification based on fingerprints. In the database fingerprints data is collected from different age groups and these were chosen from urban and rural area. The acquisition of the fingerprint was made by "Fingkey hamster 2nd scanner manufacture by nitgen biometric solution [30 with interface USB 2.0]". The resolution of the captured images is of 512 DPI in gray scale. The main purpose for the creation of database is to check the robustness in the gender identification. Database contains 74 persons fingerprint images each 10 fingers images totally we have 740 images of male and female fingerprint image [8][13].

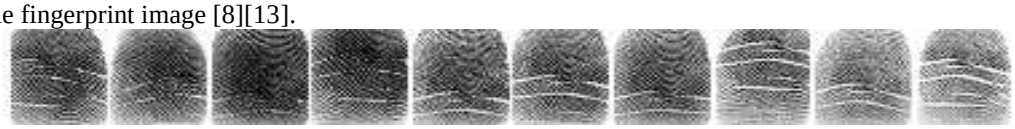

Fig. 2: Sample of fingerprints

B. Pre-Processing

In this step, performed background elimination, cropping, converting color image into binary image etc. which makes it suitable for further processing, further we have normalized the enhanced fingerprint image to size 164x164. These images are ready for feature extraction as explained in the following Section.

IV. FEATURE EXTRACTION

The Haralick texture features, and shape features are used to extract the gender information. The fifteen features are computed using the following formulae. They are contrast, correlation, energy, homogeneity, area, mean, entropy, eular.rangefit, skewness, kurtosis, variance and Standard deviation.

F1 = Contrast = $\sum ij\ [i-j]^2\ p(i,j)$ (1)

F2=correlation= $\sum ij \frac{(i-\mu i)(i-\mu j)p(i,j)}{\sigma i \sigma j}$ (2)

F3=Energy= $\sum ij\ P(ij)^2$ (3)

F4=Homogeneity =$\sum ij \frac{p(ij)}{1+|i-j|}$ (4)

F5= Area=number of pixels of an image (5)

F6=Mean= $\frac{\sum x}{N}$ (6)

F7=range fit=Difference between lowest and highest value of array I(;) (7)

F8=entropy=$-\sum Pi\ log^2 Pi$ (8)

F 9=local entropy (9)

F10=Euler Number =$\sum_{K=0} \frac{1}{k!}$ (10)

F11=Standard Deviation = $\sqrt{\frac{1}{N-1}\sum_{i=0}^{N}(x-\mu)^2}$ (11)

F12=local standard deviation (12)

F13=Kurtosis = $\frac{\sum(x-\mu)^4}{\sigma^4}$ (13)

F14= skewness = $\frac{\sum(x-\mu)^2}{\sigma^3}$ (14)

F15=Variance = $\frac{i}{N-1}\sum^{N}_{i=1}[A_I-\mu]^2$ (15)

A. Classifier

The gender identification is performed using KNN classifier. The experimental results were poor with KNN. Then, classification has performed using QDA and LDA and noticed 83.46% and 81.36% respectively. To enhance the classification accuracy, we have employed SVM classifier with polynomial and kernel functions and obtained 89.15% and 87.07% respectively. The details about the SVM and the functions used are briefed in the following paragraphs.

B. Support Vector Machine

Support Vector Machine (SVM) is a classification technique based on the statistical learning theory proposed by Vapnik in 1995 [19]. It is a binary classifier it abstracts a decision boundary in multidimensional space using an appropriate sub set of the training set of vectors; the elements of this sub set are the support vectors. Geometrically SVM are those training patterns that are closest to the decision boundary. It is useful to understand linear Discriminant functions and neural networks.[7-13].

C. Kernel Functions

The Radial Basis Function (RBF) is one of the most popular kernel and reasonable. It is choosen because of its property of nonlinearity.

$$f(x) = w^t.\ x + b \quad (13)$$

Where w and x are d-dimensional vectors, when b=0, it is a homogenous representation; otherwise it is a non-homogenous representation. This representation can be used to characterize linear separability. We say the two classes 'male' and 'female' are linearly separable , if the weight of vector w and a scalar b such that

$w^t.\ x + b>0$-------- (14) for all patterns x belonging to one class(for example say 'male') and

$w^t.\ x + b<0$-------- (15) for all patterns x belonging to another class(for example say 'female')

An attractive feature of the SVM is that this selection is implicit, with each support vectors contributing one local Gaussian function, centre at that data point [20].

V. RESULTS AND DISCUSSION

For experimentation we have used 740 fingerprint images which were collected from different age groups of the rural and urban population of 370 male and 370 female .The statistical features are extracted. For classification 370 male and 370 female fingerprints are used. The training and testing sets are automatically divided with 10-fold cross validation (CV) to evaluate the performance of our proposed method. We split the data into 10 equal sized sub-parts. From 10 sub-parts, single sub-part is used for validation, while remaining 10-1 = 9 sub-parts serve as training. This process is repeated for 10 sub-parts and single value result is calculated by averaging all results**.** The classification results are reported in Table 1. Besides, the confusion matrix is shown in Table. 2.

Classifiers	Overall Accuracy
SVM(polynomial)	89.15%
SVM(Kernel)	87.07%
QDA	83.46%
LDA	81.13%

Table 1: Gender Identification results with various classifiers

SVM(polynomial)		SVM(Kernel)		QDA		LDA	
Male	Female	Male	Female	Male	Female	Male	Female
335	35	323	47	303	67	313	57
39	331	51	319	49	321	79	291

Table 2: Gender Identification confusion matrix with various classifiers

VI. CONCLUSION

It is evident from the literature survey that number of researchers has worked on gender classification using different approaches and achieved some good results with their own datasets. But still there is a scope for developing a robust algorithm using different parameters. This proposed method is simple and have less time complexity as its feature vector size is small. It is found that the proposed algorithm produces accurate decision of 89.15%. In future, the work will be extended to build robust algorithm with more generalized properties and high accuracy.

REFERENCES

[1] Biometrics History, "National Science and Technology council (NSTC) of United State", March 2006
[2] Anil K. Jain, Karthik Nandakumar, Xiaoguang Lu, and Unsang park, "Integrating Faces, Fingerprints, and Soft Biometric Traits for user Recognition." Proceedings of Biometric Authentication Workshop, LNCS 3087, PP.259-269, PRAGUE,-MAY 2004.

[3] Sudesh Gundadinet.al., "Sex Determination from Fingerprint Ridge Density", International Journal of Medical Update 2007 Jul-Dec;2(2):4-7.
[4] Sharath Pankanti, Alan Jea, NaliniRatha and Ruud Bolle "Fingerprint Representation Using Localized Texture Features" The 18th International Conference on Pattern Recognition (ICPR'06).
[5] Gnanasivam .P, and Dr. Muttan S, "Fingerprint Gender Classification Using Wavelet Transform and Singular Value Decomposition", International Journal of Computer Science Issues, Vol. 9, Issue 2, No 3, March 2012.
[6] Dr. Prateek Rastogi, Ms. Keerthi R Pillai "A study of fingerprints in relation to gender and blood group" J Indian Acad Forensic Med 2011 , 32(1), pp - 11 - 14 ISSN 0971 – 097.
[7] RituKaur and Susmita Ghosh Mazumdar, Mr. Devanand Bhonsle, "A Study On Various Methods of Gender Identification Based on Fingerprints", International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 2,Issue 4, April 2012Electron., submitted for publication
[8] S. S. Gornale,Abhijit Patil,Veersheety C,"Fingerprint Based Gender Identification Using Discreate Wavelet Transformation And Gaboor Filters
[9] S. S. Gornale, Mallikarjun Hangarge, Rajmohan Pardeshi ,Kruthi R" Haralick Feature Descriptors for Gender Classification Using Fingerprints: A Machine Learning Approach ",IJARCSSE Volume 5,Issue 9,September 15.
[10] S. S. Gornale, Geetha D, Kruthi R "Analysis of fingerprint image for gender classification using spatial and frequency domain analysis", American International Journal of Research in Science, Technology, Engineering and Mathematics", ISSN (Print): 2328-3491, ISSN (Online): 2328-3580, ISSN (CD-ROM): 2328-3629, PP: 46-50, 2013
[11] S. S. Gornale, "Fingerprint Based Gender Classification for Biometric Security: A State-Of-The-Art Technique", International Journal of Re-search in Science, Technology, Engineering & Mathematics ISSN (Print): 2328-3491, ISSN (Online): 2328-3580, ISSN (CD-ROM): 2328-3629 9(1), December 2014-February 2015, pp. 39-49.
[12] Vikas Humbe, S S Gornale , K V Kale, R. R. Manza', "Mathematical Morphology Approach for Genuine Fingerprint Feature Extraction", International Journal of Computer Science and Security, ISSN: 1985-1533 Volume No. 1 issue 2 PP: 53-59-2007.
[13] S. S. Gornale and K.V.Kale, "Development of Compression technique for Noisy Image: A Multi-Wavelet approach", Ph.D. thesis University of Pune, Pune-2008.
[14] Manish Verma and Suneeta Agarwal.'' Fingerprint Based Male - Female Classification. '' in Proceedings of the international workshop on computational intelligence in security for information systems (CISIS'08), Genoa, Italy, 2008, pp.251 – 257.
[15] A. Badawi, M. Mahfouz, R. Tadross, and R. Jantz "Fingerprint - based gender classification" The International Conference on Image Processing, Computer Vision, and Pattern Recognition, June 2006.
[16] Naveen Kumar Jain, Sunil Sharma, Anurag Paliwal., A Real Time Approach To Determine The Gender Using Fingerprints", IJAIR ISSN: 2278-7844, PP:229-233, 2012.
[17] Rijo Jackson Tom, T. Arulkumaran , "Fingerprint Based Gender Classification Using 2D Discrete Wavelet Transforms and Principal Component Analysis", International Journal of Engineering Trends and Technology, Volume 4 Issue 2,2013 M. Young, The Technical Writers Handbook. Mill Valley, CA: University Science, 1989.
[18] Pallavi Chand, Shubhendu Kumar Sarangi, "A Novel Method for Gender Classification Using DWT and SVD Techniques", International Journal of Computer Technology & Applications, Vol 4 (3),445-449, May-June 2013 Available online@www.ijcta.com.
[19] V. N. Vapnik, "The Nature of Statistical Learning Theory," Springer, New York, 1995. doi:10.1007/978-1-4757-2440-0.
[20] V susheela devi and Murty, "Pattern recognition-An introduction" Universities Press-India 2013
[21] Shweta ujwal ,Giridhar P Jain,"Fingerprint verification using statistical and co occurrence matrix features ", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-4 Issue-7, December 2014