Journal of Biometrics & Biostatistics
ISSN: 2155-6180

Yadav and Saxena, J Biom Biostat 2018, 9:5
DOI: 10.4172/2155-6180.1000412

Short Communication

A Review on Gender Identification Using Machine Learning Technologies based on Fingerprints

Yadav JS and Saxena A*

Department of Computer Science, JECRC University, India

Abstract

Fingerprint is a unique biometric feature of individual. It is also known that fingerprints have differences in male and female with respect to ridge line details. Some studies in machine learning investigate a relationship between fingerprint and gender. In these studies by analyzing the fingerprint we get important information such as age and gender of a person. Statistical studies have been made in different geographical areas to identify the relationship between fingerprint and gender. This paper illustrates gender classification based on fingerprints through various machine learning techniques like naïve Bayes method, Decision Tree and Support Vector Machine algorithms, KNN, PCA, Wilcoxon-Mann-Whitney Test, Friedman Test. This study introduces the concept of epidermal ridge, minutiae, ridge areas, ridge density etc., and compare above stated machine learning techniques, their limitations and strengths based on experimental results for gender classification based on fingerprints. This study can be useful for legislative cases and for researchers to devise new machine learning techniques with improved results.

Keywords: Fingerprints; Epidermal; Minutiae; Ridge density; Valleys; Frequency domain analysis; Gender identification; Fingerprint image; Classification; KNN; PCA; SVM

Introduction

In today's scenario fingerprints are playing vital role in different applications like identifying culprit in criminal cases, attendance in corporate sectors, identity of nationality etc. [1]. If the sex of culprit is identified with certainty it can lead investigation to a right direction. Due to its unique property of absolute identity it can also be used to identify the gender [2]. Fingerprints have various physical properties like ridge, ridge ending, ridge density, ridge areas, minutiae etc. (Figure 1).

Ridge: It is a curved line in a finger image [3].

Ridge ending: where the ridge line ends [4].

Minutiae: Ridge endings and bifurcation are known as minutiae [5].

Epidermal ridge: Ridges of the epidermis of the palms and soles, where the sweat pores open.it is also called skin ridge [6].

Fingerprint ridge density: It has been reported that females have a significantly higher ridge density than males. The higher ridge density in female is due to the level of ridge thickness and it is suggested that females tend to have finer epidermal ridges details [7].

Fingerprint fundamentals

Fingerprints can be categorized into-Latent, Patent and plastic impressions [8]. Latent impressions are ridge impressions formed on an object when finger is covered with foreign residue such as grease or oil [7]. Patent fingerprints are fingerprint impression formed due to residue left on finger. Plastic fingerprints are visible impression left on clay, wax [9].

Figure 1: Fingerprint image.

Gender classification process

Gender classification process involves 5 steps namely: data collection, data preprocessing, Extraction of features, Feature matching, and Classification [10,11]. In data collection step, gathering and measuring of data is done which is used to evaluate the outcomes. After data collection, data preprocessing is done [12]. Data preprocessing changes the data into a format, which can be easily processed [13]. The result of data preprocessing is the final training set. Next step is extraction of features, it minimizes the amount of resources needed to delineate a large set of data [14]. In Feature matching, features from different set are matched and then classified as male or female (Figure 2).

General method for gender classification

Features should be easily computed, robust, insensitive to various distortions and variations in the images, and rotationally invariant [15]. Based on the type of features used, previous studies can be broadly classified into two categories (Figure 3):

- Appearance feature-based (global)
- Geometrical feature-based (local).

Methodologies

Naïve Bayes Classifiers are a type of probabilistic classifiers based

***Corresponding author:** Saxena A, Department of Computer Science Engineering, JECRC University, India, Tel: 919116642285; E-mail: arjit.saxena60@gmail.com

Received September 23, 2018; **Accepted** October 22, 2018; **Published** October 29, 2018

Citation: Yadav JS, Saxena A (2018) A Review on Gender Identification Using Machine Learning Technologies based on Fingerprints. J Biom Biostat 9: 412. doi: 10.4172/2155-6180.1000412

Citation: Yadav JS, Saxena A (2018) A Review on Gender Identification Using Machine Learning Technologies based on Fingerprints. J Biom Biostat 9: 412. doi: 10.4172/2155-6180.1000412

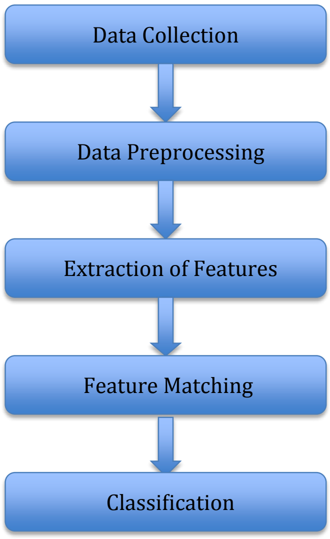

Figure 2: Block-diagram of gender classification process [2].

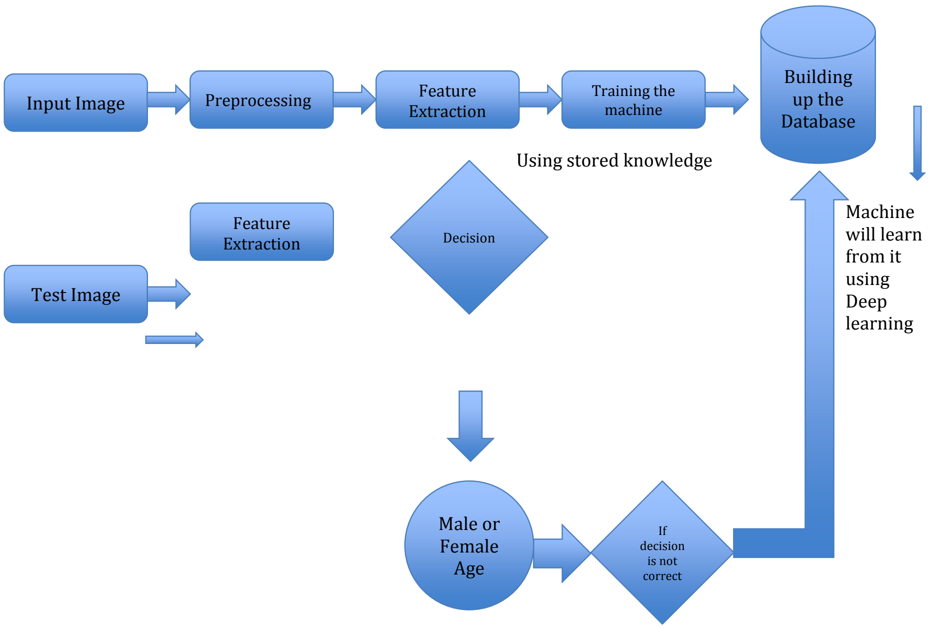

Figure 3: Flowchart for gender classification [7].

on applying Bayes theorem with strong independence assumptions between the features. Using Bayes theorem, the conditional probability can be written as [16,17].

$$p(C_k|x) = \frac{p(C_k)\,p(x|C_k)}{p(x)}$$

Let us consider a general probability distribution of two variables, p(x_1,x_2). Using Bayes rule, we can write [18].

$$p(x_1,x_2)=p(x_1|x_2)*p(x_2)$$

Similarly, considering another class variable, y, using Bayes rule, we can write [19,20].

$$p(x_1,x_2|y)=p(x_1|x_2,y)*p(x_2|y)$$

In the above expression, no assumption has been done. Further it can be written as [21]:

$$p(x_1|x_2,y)=p(x_1|y)$$

Generalized formula for a set of variables x_1, x_2……, x_n, conditional on another variable y can be written as [22]:

$$p(x_1,x_2,\ldots,x_n|y)=\prod_{i=1}^{n}p(x_i|y)$$

KNN (k-nearest neighbor's algorithm) is a trained learning algorithm used for age identification from fingerprints. This algorithm is a trained learning algorithm and the aim is to make a classification using existing data, when a new data is received [23,24].

Support Vector Machine (SVM) is a machine learning classifier used to analyze data for classification and regression analysis. SVM can also perform non-linear classification using kernel trick [25,26].

Review of Literature

According to Ceyhan and Sagiroglu [1], gender can be efficiently classified by using various methods. Input fingerprint image is taken in form of feature vector and then different classifiers like Naïve Bayes, Knn, Decision Tree and Support Vector Machine are applied on feature set. System includes 300 males and 300 females. The database is divided into two parts as 66% for the training set and 34% for the testing set. Success of the gender classification by using the different classifiers is as given below:

1. Naïve Bayes gave overall success of 95.3%.
2. KNN gave overall success of 94.0%.
3. Decision Tree gave overall success of 94.3%.
4. SVM gave overall success of 93.8%.

These fairly high success rate shows that there is a distinguishing feature between fingerprint and gender. The proposed methods can be used to reduce the suspect list in criminal cases.

Thaiyalnayaki et al. [3] has proposed a technique for fingerprint based gender classification using texture analysis technique, which uses Discrete Wavelet Transform (DWT) for extracting feature from the fingerprints. Canberra distance metric is used for similarity comparison between the texture classes. System runs in calculating Standard Deviation, Kurtosis and Skewness of the wavelet transform of the image. The training set includes 30 images. Overall performance is upto 95%.

Gornale et al. [4] has proposed a technique for fingerprint based gender classification using combined features like FFT, Eccentricity and Major Axis Length. System includes left thumb impression of 450 males and 550 females. The success rate for male is 80% and for female it is 78%.

Gnanasivam and Mutthan [6] have proposed a technique for fingerprint based gender classification using Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD). System includes fingerprints of 1980 males and 1590 females. KNN is used as an classifier. The overall success rate for male fingerprints is 91.67% and for female fingerprints it is 84.69%. Spatial Parameters can be used to increase the performance of the method.

Arun and Sarath [8] have proposed a technique for fingerprint based gender classification using SVM classifier. The Database consists of 150 male and 125 female images. For each image ratio of ridge thickness to valley thickness is calculated and is known as RTVTR.

Kaur and Mazumdar [9] have proposed a technique for fingerprint based gender classification using Frequency domain analysis, which uses Fast Fourier Transform (FFT), discrete cosine Transform (DCT), and Power Spectral Density (PSD). The overall success for male is 80% and for female it is 90%.

Comparison of Ridge Density

The value of these probabilities within a forensic context requires consideration in relation to Bayes' theorem which links the likelihood ratio and prior odds to yield the overall posterior odds [21,22]. The probabilities estimated here refer to the likelihood ratio and not directly to the posterior odds. The likelihood ratio considers the probability of the evidence given the prosecutor's and defences hypotheses, respectively. Ridge density in two areas of the footprint was studied for the first time in Cape Colored and white Afrikaans individuals. Although these two ethnic groups share some ancestry, significant differences in ridge density were observed between the groups in the heel areas of the feet. Although the likelihood ratio estimated in this study cannot be used directly since not all ethnic groups in South Africa were considered, it nevertheless provides a framework of methodology for analysis, as well as acknowledges the limitations of dealing with such data in Table 1.

Authors	Region	Average Ridge Density (women)	Average Ridge Density (men)
Acree [16]	Caucasian-American	13.32	11.14
Nayak [17]	Chinese	14.15	11.73
Redomero [18]	Spanish Caucasians	17.91	11.44
Alonso [19]	Mataco-Mataguayo	17.82	16.62
Nithin [20]	South-Indian	14.14	12.57
Ceyhan [21]	Turkish	14.10	11.49

Table 1: Women has more ridge density than men.

Conclusion

From the paper being studied, it can be concluded the ridge density of women is more as compared to men. Different machine learning classifiers are applied out of which naïve Bayes shows the most accurate result with an accuracy of 95.3%.

References

1. Ceyhan EB, Sagiroglu S (2014) Gender Inference within Turkish Population by Using Only Fingerprint Feature Vectors. IEEE Symposium on Computational Intelligence in Biometrics and Identity Management.
2. Dongre GB, Jagade SM (2015) Review and Study on Fingerprint Based Gender Classification using Classified Techniques. Int J Sci Res 6: 596-599.
3. Thaiyalnayaki K, Karim SSA, VarshaParmar P (2010) Finger Print Recognition using Discrete Wavelet Transform. Int J Comput Appl 1: 96-100.
4. Gornale SS, Geetha CD, Kruthi R (2013) Analysis of Fingerprint Image for Gender Classification Using Spatial and Frequency Domain Analysis. American International Journal of Research in Science Technology, Engineering and Mathematics pp: 46-50.
5. Xiong L, Zhao X, Fu Y, Liu Y (2010) Bimodal Gender Recognition from face and fingerprint. IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
6. Gnanasivam P, Mutthan S (2012) Fingerprint Gender Classification Using Wavelet Transform and Singular Value Decomposition. IJCSI International Journal of Computer Science Issues 9.
7. Dongre GB, Jagade SM (2015) A Review and Study on Fingerprint Based Gender and Age Classification Techniques. Int J Sci Res 6: 596-599.
8. Arun KS, Sarath KS (2011) A Machine Learning Approach For Fingerprint based Gender Identification. IEEE Recent Advances in Intelligent Computational Systems.
9. Kaur R, Mazumdar SC (2012) Fingerprint Based Gender Identification using Frequency Domain Analyses. International Journal of Advances in Engineering and Technology 3: 295-299.
10. Ponnarasi SS, Rajaram M (2012) Gender Classification system derived from minutiae extraction. Int J Comput Appl.
11. Rajesh DG, Punitharalli M (2014) An efficient Fingerprint Based Gender Classification System Using Dominant Un-decimated Wavelet Coefficients. Res J Appl Sci, Eng Technol 8: 1259-1256.
12. Wang Y, Zhendong W, Zhang J (2016) Damaged Fingerprint Classification by Deep Learning with Fuzzy Feature Points. International Congress on Image and Signal Processing, Biomedical Engineering and Informatics.

13. Boutella L, Serir A (2017) Fingerprint Identification by Wave atoms Transform and SVM. International Conference on Advanced Systems and Electric Technologies.
14. Vitello G, Sorbello F, Migliore GIM, Conti V, Vitabile S (2014) A Novel Technique for Fingerprint Classification based on Fuzzy C-Means and Naïve Bayes Classifier. International Conference on Complex, Intelligent and Software Intensive Systems.
15. Ballan M, Sakarya FA, Evans BL (1997) A Fingerprint Classification Technique Using Directional Images. Asilomar Conference on Signal, System and Computer 1: 101-104.
16. Acree MA (1999) Is there a gender difference in fingerprint ridge density? Forensic Sci Int 102: 35-44.
17. Nayak VC, Rastogi P, Kanchan T, Yoganarasimha K, Kumar GP, et al. (2010) Sex differences from fingerprint ridge density in Chinese and Malaysian population. Forensic Sci Int 197: 67-69.
18. Redomero EG, Alonso C, Romero E, Galera V (2008) Variability of fingerprint ridge density in a sample of Spanish Caucasians and its application to sex determination. Forensic Sci Int 180: 17-22.
19. Redomero GE, Alonso MC, Dipierri JE (2011) Sex differences in fingerprint ridge density in the Mataco-Mataguayo population. Journal of Comparative Human Biology 62: 487-499.
20. Nithin MD, Manjunatha B, Preethi DS, Balaraj BM (2011) Gender differentiation by finger ridge count among South Indian population. J Forensic Leg Med 18: 79-81.
21. Ceyhan EB, Sağıroğlu S, Akyıl EM (2013) Statistical Gender Analysis Based on Fingerprint Ridge Density. 21th IEEE Signal Processing and Communications Applications Conference, p: 472.
22. Redomero EG, Rodríguez AC, Hurtado HLE, Villalba RJL (2011) Distribution of the minutiae in the fingerprints of a sample of the Spanish population. Forensic Sci Int 20: 79-90.
23. Khleifat KM, Tarawneh KA, Ali Wedyan M, Al-Tarawneh AA, Al Sharafa K (2008) Growth kinetics and toxicity of Enterobacter cloacae grown on linear alkylbenzene sulfonate as sole carbon source. Curr Microbiol 57: 364-370.
24. Yõldõz T, Yõldõrõm S, Altõlar T (2008) Spam Filtering with Parallelized KNN Algorithm. Akademik Biliúim, pp: 1-6.
25. Brandt KK, Hesselsoe M, Roslev P, Henriksen K, Sørensen J (2001) Toxic effects of linear alkylbenzene sulfonate on metabolic activity, growth rate, and microcolony formation of Nitrosomonas and Nitrosospira strains. Appl Environ Microbiol 67: 2489-2498.
26. Kanchan T, Krishan K, Sharma A, Menezes RG (2010) A study of correlation of hand and foot dimensions for personal identification in mass disasters. Forensic Sci Int 199.