Fingerprint Using Histogram Oriented Gradient and Support Vector Machine

Mohamed Alhamrouni[1]
maalhamroni@elmergib.edu.ly

Lutfia hajmohamed[2]
lkalhaj@elmergib.edu.ly

[1][2] Computer Science Dept. Faculty of Arts & Science Kasr Khiar, Elmergib University-Libya

Received: 01/11/2023 | **Accepted:18/11/2023**

الملخص :

يستخدم الأمن البيومتري خصائص فسيولوجية وسلوكية فريدة، مثل الحمض النووي والتوقيعات، لتحديد هوية الأفراد. لا يمكن الاعتماد على طرق تحديد الهوية التقليدية، لذا توفر أنظمة القياسات الحيوية حلاً أكثر موثوقية. فهي دقيقة ولا يمكن نسيانها بسهولة، مما يجعلها مريحة. كان التحقق من بصمات الأصابع إحدى الطرق الأولى، ولكنه يتطلب عملية معقدة لضمان الدقة. في هذه الورقة، قدمنا SVM كتقنية مطابقة بمساعدة HOG لاستخراج الميزة وطريقة المعالجة المسبقة. وقد أظهرت دراستنا نتائج مهمة وتسلط الضوء على الدور القوي لـ SVM في عملية المطابقة.

الكلمات المفتاحية : خوارزمية التدرج للرسم البياني، مصنف آلة المتجهات الداعمة، بصمة الاصبع، القياسات الحيوية.

Abstract:

Biometric security uses unique physiological and behavioral characteristics, like DNA and signatures, to identify individuals. Traditional identification methods are unreliable, so biometric systems offer a more dependable solution. They are accurate and not easily forgotten, making them convenient. Fingerprint verification was one of the first methods, but it requires a complex process to ensure accuracy. In this paper, the rsearchers introduced SVM as a Matching Technique with the help of HOG to extract the feature and Preprocessing method. Our study has shown significant results and highlights the powerful role of SVM in the matching process.

Keywords: Histogram Oriented Gradient Algorithm, Support Vector Machine Classifier, Fingerprint, Biometric.

Abbreviations:

HOG	Histograms of Oriented Gradients
SVM	Support Vector Machine
LBP	Local Binary Pattern
PCA	Principal Component Analysis
FVC	Fingerprint Verification Competition.
KNN	K- Nearest Neighbors
HIG	Histograms of Invariant Gradients
DWT	Discrete Wavelet Transform
DTCWT	Dual Tree Complex Wavelet Transform

Introduction:

Biometric security has been becoming increasingly important for reducing risks and preventing vulnerabilities on any kind of platform. Biometric systems are advanced technological solutions that utilize an individual's unique physiological and behavioral characteristics to establish their identity (Kharat and Deshmukh). The physiological characteristics include physical and biological features such as DNA, handprints, facial features, earlobes, and irises. Behavioral characteristics, on the other hand, encompass non-physiological or non-biological features such as signatures, voices, gaits, and keystrokes (Dhamala).

In the modern era, personal identifiers are paramount to ensure security and privacy. Traditional methods of personal identification, such as token-based methods that use keys or ID cards, and knowledge-based methods that require preset codes or passwords, are becoming insufficient and unreliable. These methods can fail if the token is lost or the password is forgotten. Therefore, there is an increasing need for new and more reliable methods of personal identification.

The development of biometric systems has provided a promising solution to this problem. Biometric systems offer greater accuracy and security as they rely on unique and unchanging physical and behavioral characteristics of individuals. They are not easily lost or forgotten, making them a more reliable and convenient means of personal identification. One of the first methods introduced to verify identities was the fingerprint. However, this method requires a complicated process to reduce errors and ensure accurate matching (Alhamrouni).

Literature Review:

Gottschlich, C, et al, applied spoofed fingerprints using multiple histograms of invariant gradients (HIG) computed from spatial neighborhoods within the fingerprint. Results show that the proposed method achieves an average accuracy comparable to the best algorithms in LivDet 2013 (Gottschlich et al.).

Abdullah et al. proposed using Discrete Wavelet Transform (DWT) for feature extraction and k-nearest Neighbor (k-NN) for gender-based fingerprint classification. The overall classification rate achieved was 96.67% (Abdullah et al.).

Alias, N. A et al, used a support vector machine (SVM) to develop a fingerprint classification model. This algorithm was implemented on a database (FVC2000 and FVC2002). The training set consists of 100 fingerprint datasets and 40 datasets were used for testing. The result of this study shows that SVM gives a high percentage of accuracy in fingerprint classification which is 92.5% (Zabala-Blanco et al.).

Alshehri et al. have proposed sensors that use three types of descriptors, namely Orientation, Gradient, and Gabor-HoG descriptors. The results from the three descriptors were combined using a weighted sum rule which scales each result according to its verification performance. Extensive experiments were conducted using two public-domain benchmark databases - FingerPass and Multi-Sensor Optical and Latent Fingerprint - to demonstrate the performance of the proposed system. The results indicate that the proposed system outperforms the state-of-the-art methods based on minutia cylinder-code (MCC), MCC with scale, VeriFinger (a commercial SDK), and a thin-plate spline model(Alshehri et al.) .

Kumar, R, conducted experiments on fingerprint matching using the Orientation Local Binary Pattern (OLBP) method. The study utilized multiple databases, including FVC2002, FVC2004, and FVC2006, and employed the Chi-square test, Euclidean distance, and Least Square Support Vector Machine (LSSVM) to match images. The experimental results showed that the performance of LBP features calculated from the orientation image was comparable to those achieved. Extensive experiments were conducted on four datasets, namely FVC2002, FVC2004, and FVC2006, indicating better results on two datasets, DB2 and DB4 (Kumar) .

Pradeep, N. R, et al. (6) developed a new system that uses Discrete Wavelet Transforms (DWT), Dual-Tree Complex Wavelet Transforms (DTCWT) and Histogram of Oriented Gradients (HOG) for fingerprint recognition. The HOG extractor has shown to accurately identify fingerprints with an average of 100% accuracy. The success rate for DWT is 92.34% and for DTCWT is 95% (Pradeep and Kumar).

Brief table of previous studies.

Table 2-1 provides a summary of previous fingerprint studies, including researchers' names, methods used, databases utilized, number of images, and achieved accuracy.

Table (1): A Comparison Of Fingerprint Studies

Author	Year	Feature Extraction	Classification	Classification Rate
Gottschlich, C, et al	2014	Histograms of invariant gradients (HIG)	Support Vector Machine	96.10%
Abdullah, S. F, et al	2016	Discrete Wavelet Transform	K-Nearest Neighbor	96.67%
Alias, N. A ,et al	2016	Minutiae extraction process	Support Vector Machine	92.5%.
Alshehri, H,et al	2018	Orientation, Gradient, and Gabor-HoG	Cross-Sensor Fingerprint	A good performance for almost all cases.
Kumar, R.	2020	Orientation Local Binary Pattern	Chi-square test, Euclidean distance, and least square support vector machine	97.94%, 98.11%, 98.60%
Pradeep, N. R, et al	2021	Discrete Wavelet Transforms (DWT), Dual-Tree Complex Wavelet Transforms (DTCWT) and Histogram of Oriented Gradients (HOG).	Euclidean distance	100% for Hog, 92.34% for DWT, 95% for DTCWT.

Methodology:

The following flowchart illustrates the main methodology of this paper. In the first step, the database images are divided into a training set and a testing set, where the training set is used to train the system and the testing set is used to evaluate its performance. For our experiments, we used the FVC dataset. This Dataset offers a variety of Fingerprints Images with full details of Fingerprint characteristics.

The Next Step involves image processing to remove the noise and adjust the size of images as needed. It is an important step to ensure that the images are suitable in quality and size, which helps to improve the accuracy of the system. The images are processed using various techniques such as smoothing, thresholding, and resizing.

After the images have been preprocessed, features are extracted from them using the HOG (Histogram of Oriented Gradients) algorithm. This algorithm helps capturing the local structure of the images by analyzing the gradients of the pixel intensities by generating Feature vector. These feature vectors are then used to represent the images in a more compact and meaningful way.

Finally, in this process, we use the SVM (Support Vector Machine) algorithm to classify the extracted features for both training and testing sets. This algorithm is a well-known machine-learning technique, which is used for classification and regression analysis. It works by finding the best hyperplane that separates the feature vectors into different classes. SVM gets trained first by using the feature vectors extracted from the training set and then classifies the feature vectors extracted from the testing set.

By following this methodology, the system is able to accurately classify images based on their features, which can have a wide range of applications in various fields such as computer vision, medical imaging, and matching processes.

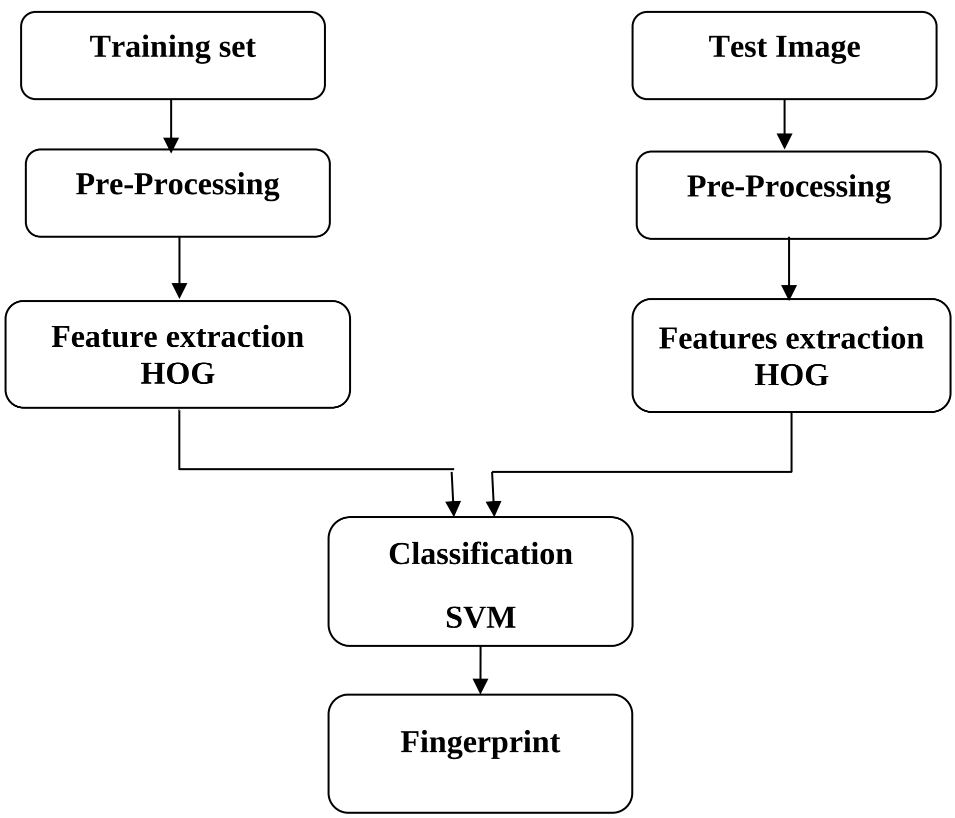

Figure 1: Methodology of fingerprint

Database:

FVC2004: Fingerprint Verification Competition 2004.

The Biometric System Laboratory at the University of Bologna conducted a study in which fingerprints were gathered from several individuals using four distinct techniques. In order to classify the data, the fingerprints were divided into four databases, labelled as DB1, DB2, DB3, and DB4, as part of the Fingerprint Verification Competition held in 2004. For this paper, we focused on the DB3 classification, which contains 80 images of ten individuals, with each person having a total of eight distinct images. The fingerprints in the DB3 were collected using thermal sweeping sensor technology, a cutting-edge technique used to capture high-quality fingerprint images. Each fingerprint image in the DB3 classification is sized at 300×480 pixels and is of the RGB type, providing a detailed and comprehensive view of the fingerprint patterns. This high-resolution image data allowed to analyze and compare the different techniques used for fingerprint gathering and classification (Maio et al.).

Figure 2: Fingerprint dataset

Table (2): Comparison between the four types of databases

Databases	Number of persons	Number of images	Images type	Images format	Images size	Used technology
DB1	10	8	RGB	TIF	640x480 pixels	Optical Sensor
DB2	10	8	RGB	TIF	328×364 pixels	Optical Sensor
DB3	10	8	RGB	TIF	300×480 pixels	Thermal sweeping Sensor
DB4	10	8	RGB	TIF	288×384 pixels	SFinGe v3.0

Table (2): makes a comparison between the most important information of these two databases such as; number of persons whose images were taken, number of images in each database, type of images, and format of images as well as size of images in each database, and type of device that used to capture these images.

Histogram Oriented Gradient Algorithm(HOG)

In 2005, Navneet Dalal and Bill Triggs introduced the Histogram of Oriented Gradients (HOG) technique as a method for detecting objects in images. HOG is a feature descriptor that captures the distribution of edge directions and intensities in an image. It works by dividing the image into small regions, calculating the gradient orientation and magnitude of each pixel in each region, and then grouping these gradients into histograms (see Figure 3). These histograms are then normalized to account for variations in lighting and contrast. HOG has achieved remarkable success in computer vision and image processing applications by solving previously challenging problems. It is beneficial in object detection tasks, such as pedestrian detection, where it has shown superior performance compared to other feature descriptors. HOG has also been used in facial recognition, and gesture recognition (Alhamrouni).

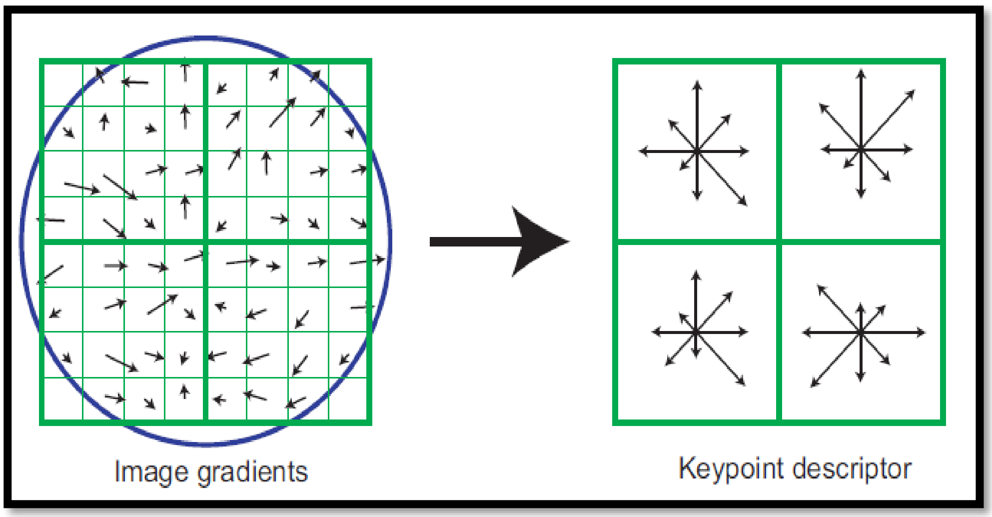

Figure (3): Image Gradients and Orientation Histogram (Alhamrouni)

Support Vector Machine Classifier (SVM)

Support Vector Machine (SVM) is a well-known method for solving pattern classification problems. Its popularity has grown over the years since its introduction by Vapnik, Boser, and Guyon in 1992. SVM is a linear classifier that accurately classifies and predicts data while taking all relevant factors into account. It is widely used in various applications, including face analysis and pattern classification. By using SVM, one can expect to find customized solutions that address the specific needs of the problem at hand. If you require more information on SVM and its applications. SVM offers a host of benefits, including ease of training and suitability for high-dimensional data. Additionally, it provides a transparent means of managing the trade-off between classifier complexity and error. Nevertheless, one of its shortcomings is its dependence on an appropriate kernel function (Platt et al.).

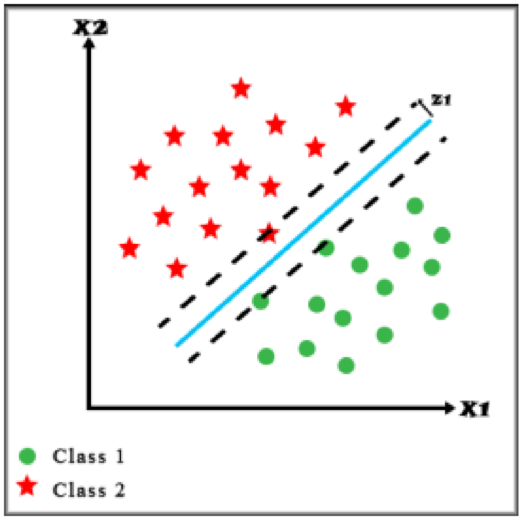

Figure 4: Support Vector Machine Mechanism(Almantsri et al.)

Multi-Class SVM problem

It's important to note that the default SVM classification method is only suitable for two classes. In cases where there are multiple classes, such as in fingerprint systems, alternative techniques need to be employed. Researchers have developed several approaches to simulate the SVM mechanism for more than two classes, including One-vs.-One, One-vs.-All (Rifkin and Klautau), and Error-Correcting Output Coding (E-COC). In this paper, we chose One-vs.-All approach to solve the multi-class problem. This approach treats each person as a separate class.

Results and Discussions:

As mentioned earlier, the experiments was applied on the FVC database's third classification DB3, which includes 80 images of ten individuals, with an average of eight images per person. The images were then divided into two sets as shown in table (3). The training set and the testing set. In the first experiment, the testing set included 10 images, one image per person, and the remaining 70 images were allocated to the training set, with seven images per person. In this case, the verification rate achieved was 100%. In the second experiment, the test set contained 20 images, two images per person, while the training set had 60 images, six images per person. The verification rate achieved in this case was 90%.

Table (3): Results of HOG approach on FVC database

Method	**Images for each person**	
	1 image for test & 7 images for train	2 images for test & 6 images for train
HOG + SVM	100%	90%

Table 4: Comparison between performance of same algorithm in previous studies and proposed study

Method	Database in previous study	Accuracy in previous studies	Database in proposed study	Accuracy in proposed study
HOG+SVM	FVC2000 FVC2002	92.5%	FVC2004	100%

Table (4) provides a comparison between previous studies and the proposed study. While both apply the SVM algorithm, they use different databases. The performance of the proposed method compared to previous methods depends on the pre-processing techniques used, as well as the type of databases utilized.

Conclusion:

It is evident from the findings that the SVM's accuracy improves with increased training, resulting in more precise matching. Moreover, the HOG technique aids in pre-processing input images, reducing noise, and aligning them to meet the requirements of this methodology. However, reducing the number of images in the training set leads to a lower accuracy for this process.

Future Works:

In the future, we are planning to implement the same approach we used previously to test the accuracy of different classifications of databases. We will be utilizing various techniques such as LBP and PCA to explore the effectiveness of these methods in comparison to our previous results. By conducting these tests, we hope to gain a better understanding of the reliability and precision of these techniques and how they can be utilized for future projects.

References:

Abdullah, S. F., et al. “Fingerprint Based Gender Classification Using Discrete Wavelet Transform and K-Nearest Neighbor.” *International Journal of Computer Science and Information Security*, vol. 14, no. 8, 2016, p. 340.

Alhamrouni, Mohamed. “„Iris Recognition By Using Image Processing Techniques “.” *Atilim University*, 2017.

Almantsri, Ahmed, et al. *A Comparison of Three Different Techniques for Object Recognition*. 2020. *Google Scholar*, http://dspace.elmergib.edu.ly/handle/123456789/853.

Dhamala, P. *Multibiometric Systems. 2012*. Ph. D. thesis, Norwegian University of Science and Technology, PhD Thesis.

Gottschlich, Carsten, et al. “Fingerprint Liveness Detection Based on Histograms of Invariant Gradients.” *IEEE International Joint Conference on Biometrics*, IEEE, 2014, pp. 1–7. *Google Scholar*, https://ieeexplore.ieee.org/abstract/document/6996224/.

Kharat, Pallavi, and Manjusha Deshmukh. "Iris Recognition: A Review." *International Journal of Advanced Trends in Computer Science and Engineering. Pp*, 2013, pp. 93–97.

Kumar, Ravinder. "Orientation Local Binary Pattern Based Fingerprint Matching." *SN Computer Science*, vol. 1, no. 2, 2020, p. 67.

Maio, Dario, et al. "FVC2004: Third Fingerprint Verification Competition." *Biometric Authentication*, edited by David Zhang and Anil K. Jain, vol. 3072, Springer Berlin Heidelberg, 2004, pp. 1–7. *DOI.org (Crossref)*, https://doi.org/10.1007/978-3-540-25948-0_1.

Platt, John, et al. "Large Margin DAGs for Multiclass Classification." *Advances in Neural Information Processing Systems*, vol. 12, 1999. *Google Scholar*, https://proceedings.neurips.cc/paper/1999/hash/4abe17a1c80cbdd2aa241b70840879de-Abstract.html.

Pradeep, N. R., and Shreyas Rajendra Kumar. "COMPARATIVE ANALYSIS OF FEATURE EXTRACTION METHODS BASED ON DWT, DTCWT AND HOG USED FOR FINGERPRINT RECOGNITION." *Technology (IJARET)*, vol. 12, no. 3, 2021, pp. 487–501.

Rifkin, Ryan, and Aldebaro Klautau. "In Defense of One-vs-All Classification." *The Journal of Machine Learning Research*, vol. 5, 2004, pp. 101–41.

Zabala-Blanco, David, et al. "Fingerprint Classification through Standard and Weighted Extreme Learning Machines." *Applied Sciences*, vol. 10, no. 12, 2020, p. 4125.