LAUTECH Journal of Engineering and Technology 19 (2) 2025: 189-195
10.36108/laujet/5202.91.0291

Statistical Feature Impact on Synthetic Palm Vein Image Generation

[1]Adebayo O. Y., [2]Omidiora E. O.,[3]Falohun A. S., [4]Abolarinwa M. O., [5]Adewuyi S. A.

[1,4,5]*Department of Computer Science, Osun State University, Osogbo.*

[2,3]*Department of Computer Engineering, Ladoke Akintola University of Technology, Ogbomoso.*

[1]olajide.adebayo@uniosun.edu.ng, [2]eoomidiora@lautech.edu.ng, [3]asfalohun@lautech.edu.ng, [4]adedapo.fikayo@fuo.edu.ng, [5]michael.abolarinwa@uniosun.edu.ng

Article Info

Article history:

Received: Mar 10, 2025
Revised: June 21, 2025
Accepted: June 25, 2025

Keywords:

A Synthetic, Mean, Statistical, Covariance, Correlation Coefficient

Corresponding Author:
olajide.adebayo@uniosun.edu.ng

ABSTRACT

This research work investigated the effect of statistical feature extraction on the generation of synthetic palm vein images. The study employed three statistical features: mean, covariance, and correlation coefficient. The features were used to generate synthetic images which were evaluated using metrics such as Equal Error Rate (EER), Recognition Accuracy (RA), and Recognition Time (RT). The study justifies the use of statistical features in the generation of synthetic palm vein images.

INTRODUCTION

Palm vein image synthesis is defined as artificially generated which shares common attributes with the original as measured by an existing biometric system. These attributes are always controlled by the initial parameters' sets to imitate the characteristics. Introducing a certain amount of variation in setting these attributes allows for the creation of a large number of unique palm vein images (Atanda *et al.*, 2023; Olayiwola *et al.*, 2023). With the ability to generate large numbers of realistic synthetic images, the time-consuming and expensive process of collecting data from live subjects can be reduced. Synthetic biometrics also allow for the protection of personal identity. A synthetic biometric sample in an evaluation database is not associated with any real person's identity. Thus, synthetic biometric databases could be distributed among research organizations with less regulation (Yanushkevich, 2006).

RELATED WORKS

Michal et al. (2019) observed that images captured of football players during a football match had low resolution even when cameras were of high resolution. They proposed an approach to resolve issues posed by low-resolution images. A simple Python script for synthetic images was created instead of manual annotations. The raw synthetic images were transformed into more realistic images using the Vanilla Cycle Generative Adversarial Network (CGAN) and trained using the Cascade Pyramid Network (CPN) model. They were able to achieve similar precision with their images as one of the CPN models trained with a Common Object in Context (COCO).

GAN was proposed for medical image synthesis. The method synthesizes brain images for Normal Control (NC), Mild Cognitive Impairment (MCI), and Alzheimer's disease (AD). The result showed that medical image synthesis using GAN is a cost-saving approach for automated diagnostic technology (Islam and Zhang, 2020).

Xuel et al. 2021 developed a model for synthesizing realistic histopathology images using HistoGAN. A synthetic framework that selectively adds new patches was also investigated. The developed models were evaluated on cervical histopathology and lymph node histopathology datasets. The results revealed that images generated with the developed model selective augmentation showed significant and consistent improvements. Somasekar and Naveen, 2021 developed a synthetic image identification system GAN as a synthetic generator based on random sampling and Long Short-term Memory (LSTM) as a discriminator. Facial datasets and abstract art datasets were used for training and testing. Accuracies were found to be 58.53% and 72.68% for both GAN and LSTM respectively.

Wang et al. 2023 examine the suitability of synthetic palm vein images for augmenting scare palm vein images. The research focused on the state-of-the-art method of modeling synthetic palm vein images and also formalized a general flowchart for the creation of the synthetic database.

Kim et al. 2024 developed a new generative adversarial network to create realistic fake finger veins for training spoof detectors. The developed adversarial was used to distinguish fake finger veins from real images. The fractal dimension was introduced to analyze its complexity and to also generate realistic fake images. The research showed the effectiveness of the developed adversarial.

MATERIALS AND METHODS

Synthetic Generation Model

This is the process of using the most significant information of the preprocessed palm vein images for classification purposes. Feature extraction is a crucial step in a biometric system and its capability directly influences the performance of the system. The mean, covariance, and correlation coefficient of the normalized palm vein images were extracted using the Algorithm presented:

1. Let X be an acquired palm vein with n records and m variables. Let X' be the synthetic palm vein to be generated, with n' records and m variables.
2. X can be viewed as an $n \times m$ matrix and X' can be viewed as an $n' \times m$ matrix.
3. This algorithm presented guarantees that statistical properties of X, such as mean, covariance, and correlation coefficient are exactly reproduced in the resulting X'.
4. In this work, palm vein patterns were generated by introducing a certain number of variations to the three distinct statistical features of palm vein (Mean, Covariance, and Correlation coefficient).

Algorithm 1: (Basic Procedure)

---

1. Generate A, which is a random $n' \, x \, m$ matrix, such that the covariance matrix of A is the identity matrix.
2. Compute the covariance matrix C of the original data matrix X.
3. Use the Cholesky decomposition on C to obtain $C = U^t \, x \, U$
   where U is an upper triangular matrix and U^t is the transposed version of U.
4. Obtain the synthetic data set X' as a matrix product: $X' = A.U$

Note That the covariance matrix of *X'* equals the covariance matrix of *X*.

5. Due to the construction of matrix *A*, the mean of each variable in *X'* is *0*. To preserve the mean of variables in *X*, a last adjustment is performed. If $\bar{x}_j$ be the mean of the j-th variable in *X*, then $\bar{x}_j$ is added to the j-th column (variable) of *X'*:

$$x_{ij}^{'} = x_{ij}^{'} + \bar{x}_j \; for\; i = 1, \dots, n^{'} \; and\; j = 1, \dots m \quad (1)$$

Algorithm 2: (Construction of Matrix A)

Procedures to specify how to construct a random $n^{'}\; x\; m$ matrix *A*, whose covariance matrix is the $m\; x\; m$ identity matrix.

1. Generate *A* as an *n' x m* matrix with random elements $a_{i,j.}$ View the m columns of *A* as samples of variables $A_1,\dots, A_m$. *If Cov* (A_j, $A_{j^{'}}$) is the covariance between variables A_j and $A_{j^{'}}$), the algorithm is that *COV* (A_j, $A_{j^{'}}$) $= \begin{cases} 1 & if\; j = j' \\ 0 & otherwise \end{cases}$ *For* $j, j^{'} \in \{1, \dots, m\}$.
2. Let $\bar{a}_1$ be the mean of A_1. Let us adjust A_1 as follows: $a_{i,1} := a_i\, 1 - \bar{a}_1\; i = 1 \dots n^{'}$ The mean of the adjusted A_1 is 0.
3. To reach the desired identity covariance matrix, some values of variables $A_2 \dots A_m$ must change. For v = 2 to m do:
   (a) Let a_v be the mean of variable A_v.
   (b) For $j = 1\; to\; v - 1$, the covariance between variables A_j and A_v is
   $$COV(A_j, A_{j^{'}}) = \frac{\Sigma_{i=1}^{n'} a_{i,j} . a_i, v}{n'} - 0.\; \bar{a}_v = \frac{\Sigma_{i=1}^{n'} a_{i,j} . a_i, v}{n'}$$
   (c) To obtain $COV(A_j, A_v) = 0, j = 1 \dots v - 1$, some elements $a_{i,v}$ in the v-th column of A are assigned a new value. Let $x_1, \dots, x_{v-1}$ be the unknowns for the following linear system of $v - 1$ equations:
   $$\frac{\Sigma_{i=1}^{n'-v+1}\; a_{i,j} . a_i, v + \Sigma_{i=1}^{v-1} a_{n'-v+1+i,j\,.\,x_i}}{n'} = 0$$
   $$for\; j = 1 \dots v - 1$$
   $$\sum_{i=1}^{n'-v+1} ai, j\,.\, ai, v + \sum_{i=1}^{v-1} a_n - v + 1 + i, j\,.\, x_i = 0$$
   $$for\; j = 1 \dots v - 1$$
   Once the aforementioned linear system is solved, the new values are assigned:
   $$a_n - v + 1 + i, j\,.\, x_i = 0 \;\; for\; j = 1 \dots v - 1$$
   Let be the mean of variable A_v. A final adjustment on A_v is performed to make its mean 0:
   $$\bar{a}_1, v = a_i, v - \bar{a}_v\, for\; i = 1 \dots n^{'}$$
4. In the last step, values in A are adjusted to reach $COV(A_j, A_{j^{'}}) = 1\; for\; j = 1 \dots m$. If σ_j is the standard deviation of variable A_j, the adjustment is computed as:
   $$a_{i,j} := \frac{a_{i,j}}{\sigma_j}, i = 1 \dots n', j = 1 \dots m$$

Synthetic Palm Vein Image Generation

Variations were introduced to the three optimized statistical features to synthesize palm vein images as follows:

a. Mean
b. Mean and Covariance
c. Mean, Covariance, and Correlation Coefficient

Identification Performance of the Developed System

The synthesized palm vein images were classified using SOM. This technique was employed to measure the similarity between the test images and reference images in the database. Mean was used for synthetic palm vein image generation, mean and covariance were used for synthetic palm vein image generation, and finally mean, covariance, and Correlation Coefficient were used for synthetic palm vein image generation. Experiments were conducted to demonstrate the effects of the developed methods (Adetunji et al., 2015; Oguntoye et al., 2023). In the first experiment, identification performances of each of Mean, Mean and Covariance, Mean, Covariance, and Correlation Coefficient were conducted. This was to determine the effectiveness of statistical features in Synthetic palm vein image generation.

Table 1 Parameters of Mean at Different Threshold Values

Classifier	Threshold	False Acceptance Rate (FAR)	False Rejection Rate (FRR)	Recognition Accuracy (%)/ RA	Average Recognition Time (s)/ART
Mean	0.10	0.59	0.13	99.65	86.92
	0.15	0.59	0.13	99.65	84.22
	0.20	0.59	0.13	99.65	84.35
	0.25	0.59	0.13	99.65	85.46
	0.30	0.59	0.13	99.65	85.50
	0.35	0.43	0.18	99.70	87.41
	0.40	0.43	0.18	99.70	84.11
	0.45	0.27	0.18	99.75	83.60
	0.50	0.27	0.18	99.75	86.27

Table 2 Parameters of Mean and Covariance at Different Threshold Values

Classifier	Threshold	False Acceptance Rate (FAR)	False Rejection Rate (FRR)	Recognition Accuracy (%)/ RA	Average Recognition Time (s)/ART
Mean and Covariance	0.10	0.54	0.09	99.70	72.52
	0.15	0.54	0.09	99.70	72.71
	0.20	0.54	0.09	99.70	73.07
	0.25	0.54	0.09	99.70	73.34
	0.30	0.54	0.09	99.70	72.56
	0.35	0.38	0.13	99.75	74.36
	0.40	0.38	0.13	99.75	73.52
	0.45	0.21	0.13	99.82	74.36
	0.50	0.21	0.13	99.82	73.47

The developed systems synthesized palm vein images by generating a database of 10,000 palm-vein images for each of Mean, Mean and Covariance, Mean, Covariance, and Correlation Coefficient from 500 palm-vein images (original images).

Table 3 Parameters of Mean, Covariance, and Correlation Coefficient at different Threshold Values

Classifier	Threshold	False Acceptance Rate (FAR)	False Rejection Rate (FRR)	Recognition Accuracy (%)/ RA	Average Recognition Time (s)/ART
Mean, Covariance, and Correlation Coefficient	0.10	0.48	0.04	99.75	82.53
	0.15	0.48	0.04	99.75	82.21
	0.20	0.48	0.04	99.75	82.54
	0.25	0.48	0.04	99.75	83.46
	0.30	0.48	0.04	99.75	83.70
	0.35	0.32	0.09	99.80	82.59
	0.40	0.32	0.09	99.80	82.60
	0.45	0.16	0.09	99.88	83.16
	0.50	0.16	0.09	99.88	83.61

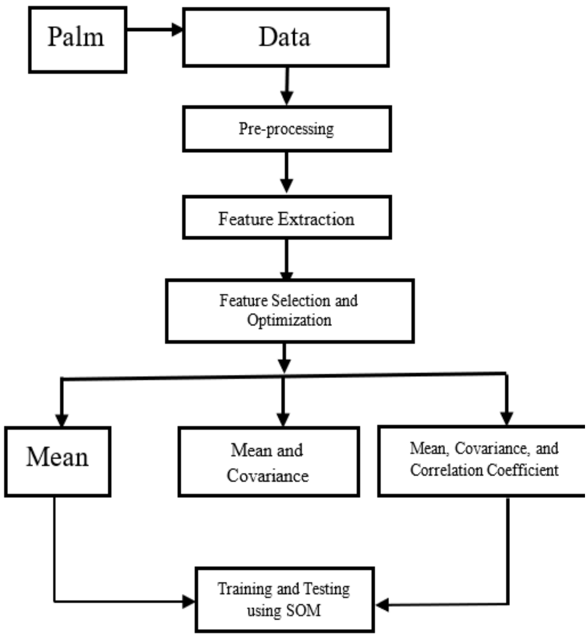

Figure1: Block Diagram of the Synthetic Palm Vein Image Generation System

For identification, with regards to each Mean, Mean and Covariance, Mean, Covariance, and Correlation Coefficient 6,000 synthetically generated palm vein images were used for training data sets and 4,000 synthetically generated palm vein images were used for test data sets.

RESULTS AND DISCUSSION

Results

As depicted in Tables 1 and 3, FAR for S1, a minimum value of 0.27 was recorded between thresholds of 0.45 between 0.50 while the maximum value of 0.59 was observed between threshold values of 0.10 and 0.30. S2 recorded a minimum FAR of 0.21 between the thresholds of 0.45 and 0.5 while the highest was 0.54 at the thresholds of 0.10 and 0.30. S3 had the lowest value of 0.16 between the thresholds of 0.45 and 0.5 while the highest value was 0.48 at thresholds of between 0.1 and 0.30. In addition, NS recorded a minimum FAR of 3.26 between the thresholds of 0.45 and 0.5 while the highest is 9.78 at the thresholds of 0.10 and 0.30.

FRR for Mean, a minimum FRR value of 0.13 was recorded between the threshold of 0.10 and 0.30 and a maximum value of 0.18 was obtained at thresholds between 0.35 and 0.50. Mean and Covariance has the least value of 0.09 at the thresholds of between 0.10 and 0.30 while the highest value is 0.13 at thresholds of between 0.35 and 0.50. Also, FRR for Mean, Covariance, and Correlation Coefficient has the lowest value of 0.04 between the threshold of 0.10 and 0.30 while the highest value of 0.09 was gotten at thresholds of between 0.35 and 0.50 as indicated in Tables 1 to 3 respectively.

RA of the systems vary between 99.65% and 99.75% for Mean, 99.70% and 99.82% for Mean and Covariance, and 99.75% and 99.88% for Mean, Covariance and Correlation Coefficient as shown in Tables 1 to 3. RT of the systems was 84.22s and 86.92s for Mean, 72.52s and 74.36s for Mean and Covariance, 82.21s and 83.70s for Mean, Covariance and Correlation Coefficient

Discussion

FAR measures the levels at which imposters are erroneously accepted by a system. From Tables 1 to 3, Mean, Mean and Covariance, Mean, Covariance, and Correlation Coefficient had the lowest value of FAR (0.16); this implied that a minimal number of such imposters were accommodated when compared with those of Mean and Covariance, and Mean. In like manner, FRR measures the levels at which legitimate enrollees were wrongly rejected. Its analysis revealed that the Mean, Covariance, and Correlation Coefficient had the lowest FRR of 0.04. This implies that a minimum number of legitimate enrollees were wrongly rejected compared to the Mean and Covariance Mean. The results indicated that the system, Mean, Covariance, and Correlation Coefficient have the highest level of deterrent to other systems (Mean and Covariance and Mean) considered. Mean, Mean and Covariance, Mean, Mean, Covariance, and Correlation Coefficient is the only system with the highest level of sensitivity to all the training images.

Evaluation of FAR, FRR, and thresholds were carried out. Creating the best access control system became an issue because it was difficult to decide if a system with higher FRR values was better than those with higher FAR; both metrics are threshold dependent. EER is a point of intersection of both FAR and FRR. The EER of a system is threshold threshold-independent performance measure of the access control system with its value independent of varying thresholds. EER of the systems under consideration were 0.22, 0.51, 0.58, and 4.36 for Mean, Mean and Covariance, Mean, Covariance and Correlation Coefficient, Mean and Covariance and Mean respectively. Mean, Mean and Covariance, Mean, Mean, Covariance and Correlation Coefficient have the lowest EER value because in all the systems under consideration, it exercised the highest restraint to palm vein images that did not take part in the training (least FAR) and least restraint to images of the subject that took part in the training phase (FRR). Given this, it is the most secure and has the best access control performance among the four systems because it is built with sufficient optimized statistical features as depicted in Figures 1 to 3.

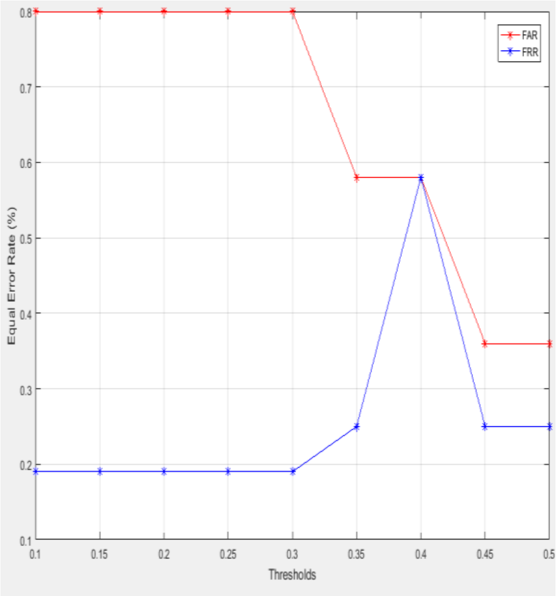

Figure 1 EER of Mean

Tables 1 to 3 showed the ARA of the four systems with 99.68% for Mean; 99.73% for Mean, Mean, and Covariance; and 99.80% for Mean, Mean, and Covariance, Mean, Mean, and Covariance; S3 had the highest ARA (99.80%) of the four systems and this is on the ground that the more the optimized statistical features that were included in the training and testing, the better the recognition accuracy. This is an indication that optimized statistical features in the developed systems contributed significantly to the overall performance of access control systems and/or identification systems.

The ART obtained for the systems is 84.97s, 75.55, 84.04s, and 681.74s for Mean, Mean, and Covariance, Mean, Mean, Covariance, and Correlation Coefficient and NS respectively, with Mean and Covariance (75.55s) having a significantly least value. This implied that the average rate of certifying the identity of an individual is the lowest with Mean and Covariance.

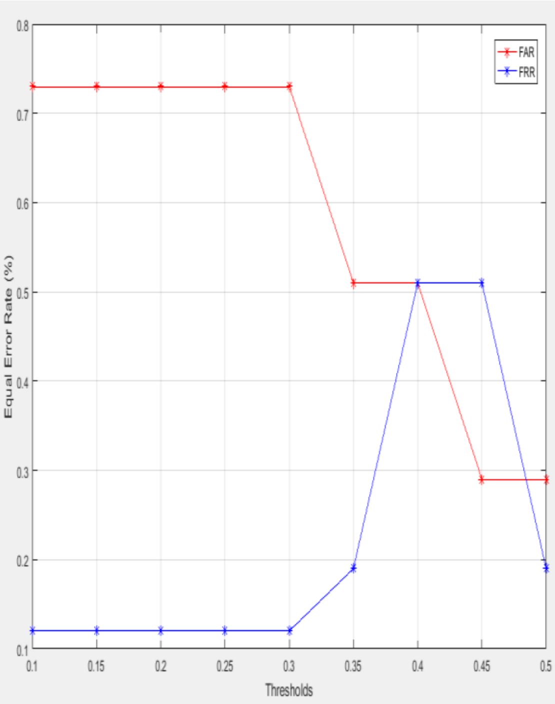

Figure 2 EER of Mean and Covariance

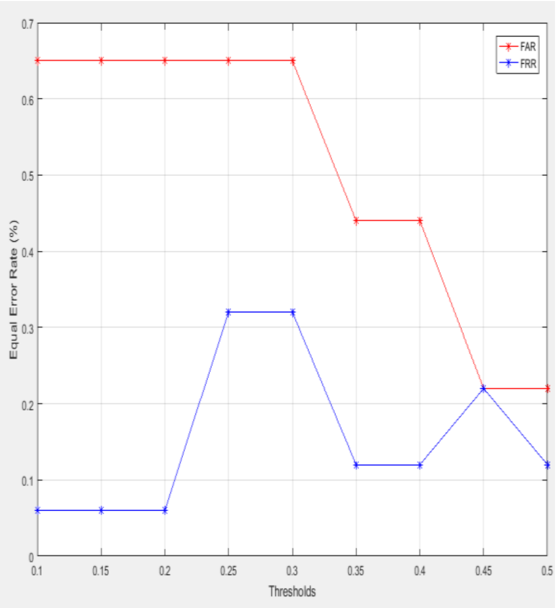

Figure 3 EER of Mean, Covariance, and Correlation Coefficient

CONCLUSION

Conclusively, the work has presented a method to create a synthetic palm vein database, which is a reasonable facsimile of a real palm vein, which can be used to address several fundamental, conceptual issues in the field of biometrics. The developed systems have established that the synthetic palm vein has promising potential in real applications. Additionally, this work has revealed how a palm vein can be accurately transformed across two or more diverse synthetic environments is an important and achievable next step for the advancement of biometrics.

REFERENCES

Adetunji A. B., Oguntoye J. P., Fenwa O. D. and Omidiora E. O. (2015): Facial Expression Recognition Based on Cultural Particle Swamp Optimization and Support Vector Machine. LAUTECH Journal of Engineering and Technology. 10(1): pp. 94-102.

Atanda, O. G., Ismaila, W., Afolabi, A. O., Awodoye, O. A., Falohun, A. S., & Oguntoye, J. P. (2023). Statistical Analysis of a Deep Learning Based Trimodal Biometric System Using Paired Sampling T-Test. In *2023 International Conference on Science, Engineering and Business for Sustainable Development Goals (SEB-SDG)* (Vol. 1, pp. 1-10). IEEE.

Islam, J. and Zhang, Y. (2020). GAN-Based Synthetic Brain PET Image Generation. Brain Informatics, 7(3):1-12.

Jun W., Huafeng Q., Haofei X., Yantao L., Mounim Y., (2023). Adversarial learning-based data augmentation for palm-vein identification. IEEE Transactions on Circuits and Systems for Video Technology, ff10.1109/TCSVT.2023.3334825ff. ffhal-04326506).

Kim, S. G., Hong, J. S., Kim, J. S., Park, K. R. (2024). Estimation of Fractal Dimension and Detection of Fake Finger-Vein Images for Fingerprint Recognition. Fractal Fract. 8, 646.

Michal, S. Grzegorz, S. and Tomasz, T. (2019). Synthetic Image Translation for Football Players Pose Estimation. Journal of Universal Computer Science, 25(6):683- 700.

Oguntoye, J. P., Awodoye, O. O., Oladunjoye, J. A., Faluyi, B. I., Ajagbe, S. A., & Omidiora, E. O. (2023). Predicting COVID-19 From Chest X-Ray Images using Optimized Convolution Neural Network. LAUTECH Journal of Engineering and Technology, 17(2), 28-39.

Olayiwola, D. S., Olayiwola, A. A., Oguntoye, J. P., Awodoye, O. O., Ganiyu, R. A., & Omidiora, E. O. (2023). Development of a Fingerprint Verification and Identification System Using a Gravitational Search Algorithm-Optimized Deep Convolutional Neural Network. *Adeleke University Journal of Engineering and Technology*, *6*(2), 296-307.

Somasekar, H and Naveen, K. (2021). A System for Identifying Synthetic Images Using LSTM: A Deep Learning Approach. International Journal of Computer Trends and Technology, 69(2):64-67.

Yanushkevich, S.N. (2006). Synthetic Biometrics: A Survey. Neural Networks (IJCNN), 0-0 0, 676–683.

Yuan X, Jiarong Y, Qianying Z, Rodney L, Sameer Ai, Zhiyun X, Carl C, Richard Z, Keith C., Xiaolei H. (2021). Selective Synthetic Augmentation with HistoGAN for Improved Histopathology Image Classification, Medical Image Analysis 67:101816