



## Key Generation and Testing Based on Biometrics

Alaa Aamer AbdulRaheem<sup>1,\*</sup>, Shahd Abdulrhman Hasso<sup>2</sup>

<sup>1,2</sup>Department of Software, College of computer science and mathematics, Mosul University, Mosul, Iraq

Emails: [alaa.almostfa@gmail.com](mailto:alaa.almostfa@gmail.com), [shahd\\_hasso@uomosul.edu.iq](mailto:shahd_hasso@uomosul.edu.iq)

### Article information

#### Article history:

Received :24/12/2023

Accepted :25/2/2024

Available online: 25/6/2024

### Abstract

Creating and testing a biometric key is a critical process used for security and identity verification. When using biometric traits such as fingerprints, facial features, iris patterns, earprints, and voice patterns, a unique key is created and linked to the individual's biometric identity. These biometrics provide inherent uniqueness, resulting in a higher level of security compared to traditional methods. In addition, biometric authentication eliminates the need for users to memorize complex passwords or carry physical tokens, thus enhancing convenience and user experience. Iris recognition systems have received significant attention in biometrics for their ability to provide robust criteria for identifying individuals, thanks to the rich texture of the iris. In this research, the key generation process was created by converting biometrics (the iris) into a digital representation (a set of binary numbers from the two irises) that can be used in the encryption process. This is done by using digital image processing algorithms to extract unique features from the two irises. After the key is generated, it is tested using random metrics. If the key meets the criteria, it is random otherwise the key will be generated again.

#### Keywords:

Biometrics, Key Generation, Iris, Feature extraction, Randomness.

#### Correspondence:

Author: Alaa Aamer AbdulRaheem

Email: [alaa.almostfa@gmail.com](mailto:alaa.almostfa@gmail.com)

## 1. Introduction

### 1.1 Biometrics

Biometric recognition refers to the automated authentication or identification of individuals based on their distinct physical or behavioral traits. The term "biometrics" originates from the Greek words "Bio" meaning life and "Metrics" meaning measure. Human biometric traits are categorized into two main types: physiological and behavioral.

Physiological features include facial characteristics, fingerprints, ears, hand geometry, iris scanning, retina patterns, and DNA. These traits are unique to each individual and cannot be easily stolen, replicated, or identical even among twins.

In contrast, behavioral biometrics, also referred to as soft biometrics, utilize the capture of psychological characteristics to generate a user template. Examples of such characteristics

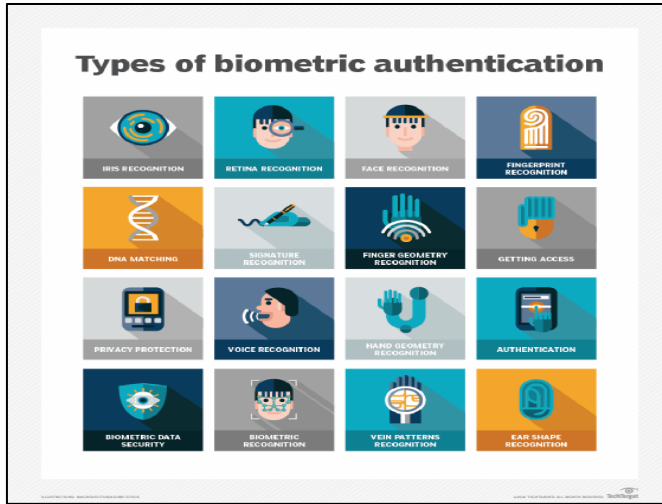
include signatures, gait patterns, voice recognition, lip movements, and keystroke dynamics., as depicted in **Fig. 1**.

Each person possesses innate and distinct observable characteristics, such as their face, making it possible to identify individuals based on who they are. Another identification approach is based on what individuals possess, such as ID cards and PINs.

Biometric recognition techniques significantly enhance security levels for real-time identification, reducing unauthorized access to various applications and services. These applications and services encompass a wide range of devices and systems, including desktop PCs, smartphones, ATMs, computer networks, workstations, and smart cards [1] [2].

The Biometric Cryptosystem "BCS" integrates features from both the biometric and cryptosystem domains. While biometric technology offers authentication capabilities, the

cryptosystem provides security measures. Since any biometric system is vulnerable to attacks and security threats, BCS aims to enhance system security while maintaining optimal performance. The concept of a key plays a crucial role in BCS, as it must be safeguarded alongside biometric signals using a specific technique [3].



**Fig. 1.** Types of Biometrics.

**1.2 Related works of Biometrics and generated key**

These studies are grouped on the biological characteristics of the individual as follows and as shown in **Table. 1:**

In [16] Wang, Y., Li, B., Zhang, Y., Wu, J., Yuan, P., & Liu, G. (2020, October), this paper propose a key generation scheme for authentication based on face image and use a CNN architecture to extract feature vectors for improving the stability of biometric identification. The results show the FAR reach to 0.53% and the FRR reach to 0.57% in LFW face database, which achieves the better performance of biometric identification, and the proposed method is able to realize randomness of the generated biometric keys by NIST statistical test suite.

In [17] Majjed, I. A., & Majeed, A. A. (2020, September), this work suggested a new method to generate a biometric key to encrypt data using the properties of the human face, then used this key to encrypting speech messages and hide them inside the coloured images. This can be achieved depending on splitting the facial image into two parts (upper and lower parts) and then generated a unique encryption key using Maximum-Relevance Minimum Redundancy (mRMR) feature selection algorithm from the upper part after that encrypted the original speech message using two levels, in the first level we used Arnold cat map to permutation the samples then in the second level used bio-key

to encrypting the message and then hide the encrypted speech message in the lower part of the facial image. In order to determine the efficiency of the proposed method, different measures were applied (correlation coefficient, PSNR, MSR, SSIM).

In [18] Roy, N. D., & Biswas, A. (2020), this work proposes to design and implement a retinal biometric key generation framework with deep neural network. The purpose is to replace the semi-automated or automated retinal vascular feature identification methods. The approach begins with segmentation from coloured fundus images, followed by selection of some unique features like centre of optic disc, macula centre and distinct bifurcation points on a convolutional neural network model. For better understanding, the key generation process has finally been shown with the help of a graphical user interface. This network was trained and tested with the training and testing images of DRIVE dataset and some of our previously published result sets on automated feature extraction methods. The network was trained on NVIDIA Titan Xp GPU provided by NVIDIA corporation.

In [19] Wu, Y., Lin, Q., Jia, H., Hassan, M., & Hu, W. (2020), propose an autoencoder based signal pre-processing step to speed up gait-based key generation. they provethat using acceleration sensor data obtained at one body location to predict the acceleration signal observed at a different body location can be achieved by using autoencoder. We further show that by using the predicted signals can speed up key generation. The bit agreement rate is increased by 16.5% and the key generation rate is increased by more than 1.9X. We further use transfer learning to reduce the required training data by 50% and the required training time by 88% to obtain a user-specific autoencoder model for a new user by retraining a pre-trained universal model. Finally, they also analyze the security of the proposed approach against various attacks.

In [20] Wang, P., You, L., Hu, G., Hu, L., Jian, Z., & Xing, C. (2021), In this work, propose a new biometric bio-key generation approach based on the generated interval scheme with a two-layer error correcting technique. They design and simulate the realization interfaces and made the experimental tests based on the two fingerprint data bases. Their experiment results prove that this bio-key generation approach shows a better security performance and provides an acceptable bio-key regeneration rate at a low computational cost. In [21] Wu, Z., Lv, Z., Kang, J., Ding, W., & Zhang, J. (2022), This study proposes a deep neural network model called MCP-FPmodel for fingerprint biometric key generation. The model utilizes layer-by-layer convolutional projection to eliminate instability between fingerprint samples. The framework consists of three modules: FPBK\_Preprocessor,

FPBK\_Stabilizer, and FPBK\_Fuzzy\_Extractor. The proposed system achieves a generation strength of over 1024 bits for the fingerprint bio-key, with an accuracy rate exceeding 98.0% and a misrecognition rate below 1.5%. Future research aims to reduce the time consumption of fingerprint key generation to less than 0.1 seconds, enabling efficient encryption and decryption operations for various data and files while ensuring convenience and security.

In [22] Lin, C., He, J., Shen, C., Li, Q., & Wang, Q. (2022), This research introduces a system called CrossBehaAuth for cross-scenario behavioral authentication using keystroke dynamics. It employs a deep neural network and temporal factors to authenticate keystroke dynamics across different scenarios. A local Gaussian data enhancement technique is proposed to increase data diversity and enhance performance. The approach was evaluated on two publicly available datasets and demonstrated significant improvements in authentication accuracy across different scenarios. It also showed scalability and advantages in both single and diverse scenario settings for dynamic keystroke authentication.

In [23] Dash, P., Pandey, F., Sarma, M., & Samanta, D. (2023). This work proposes a dynamic biometric key generation approach from iris data using illumination and rotation invariant ensemble feature descriptors. The approach has been thoroughly tested with different iris datasets. The experimental results establish that the keys generated according to the proposed approach are random and satisfy unlinkability and revocability properties. Also, it is learned that the proposed scheme of key generation is robust under brute-force attack, JPEG compression.

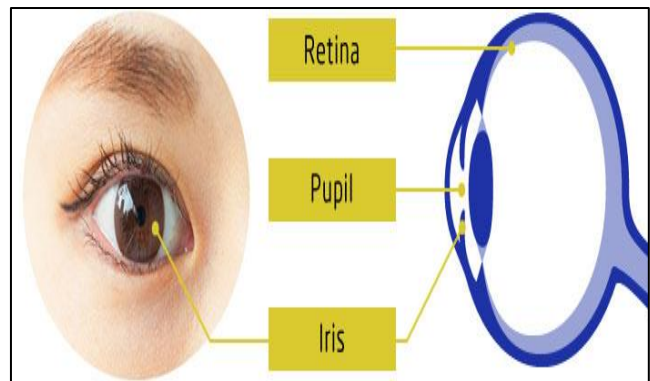
In [24] Suresh, K., Pal, R., & Balasundaram, S. R. (2023). In this work, a novel grey code-based method is introduced to generate a stable cryptographic key from fingerprint. Usage of Gray code representation significantly reduces the number of mismatch bits between the generated bit strings from the two instances of the same fingerprint. Hence, Reed-Solomon error correction code is able to successfully correct the errors which may occur due to variations in captured images of the same fingerprint. This generated bit string is used in a symmetric key setup for secure data storage, as shown in **Table. 1**.

## 2. Iris

The iris recognition system " IRS " stands out among all biometric recognition systems for its exceptional efficiency and reliability in verifying authenticity. This is attributed to the remarkable stability of the human iris, which remains

constant over time despite the aging process. Additionally, the iris exhibits a unique pattern for every individual, even among siblings or twins. The iris is protected by a structure that, if altered, could impact a person's health. Accessing the iris requires a non-invasive device, as depicted in **Fig. 2**.

The IRS is a high-accuracy verification technology renowned for its ability to perform precise personal identification. It is increasingly deployed in automated systems, eliminating the need for human operator supervision. The security field, in particular, extensively employs the IRS to enhance safety and measures, such as at smart airports, borders, mobile devices, government buildings like hospitals. Numerous countries have embraced the IRS to bolster their security infrastructure [4].



**Fig. 2.** Explain iris.

Its applications span across various fields including healthcare, national IDs, workforce management, law enforcement, hospitals, and educational institutions. To fully harness the limitless benefits of iris authentication and prevent impersonation, ensuring robust security measures is of paramount importance [5].

## 3. Feature Extraction

Feature extraction is a method employed to condense a vast input data set into essential features. Through dimensionality reduction, the aim is to transform large input data into smaller, meaningful groups that can be effectively processed. By extracting relevant features, the overall complexity of the data is reduced while retaining the crucial information for further analysis or modeling [27]. In this paper we use two methods of feature extraction: Histogram of Oriented Gradients "HOG" and Scale-Invariant Feature Transform "SIFT".

**Table 1.** Related work of Biometrics and generate key.

Study	Researchers	Algorithm	Type of Key	Type of Biometric	Result
[16]	Wang, Y., Li, B., Zhang, Y., Wu, J., Yuan, P., & Liu, G.	CNN	random	face	better performance
[17]	Majjed, I. A., & Majeed, A. A	mRMR	unique	face	More efficiency using different measures were applied (correlation coefficient, PSNR, MSR, SSIM)
[18]	Roy, N. D., & Biswas, A	deep neural network	unique	retinal	key generation and The network was trained on NVIDIA Titan Xp GPU
[19]	[Wu, Y., Lin, Q., Jia, H., Hassan, M., & Hu, W.	autoencoder Design	a new light-weight key generation(an identical key)	gait	The bit agreement rate is increased by 16.5% and the key generation rate is increased by more than 1.9X
[20]	Wang, P., You, L., Hu, G., Hu, L., Jian, Z., & Xing, C	the generated interval scheme with a two-layer error correcting technique.	a unique bio-key	fingerprint	a better security performance and provides an acceptable bio-key regeneration rate at a low computational cost
[21]	Wu, Z., Lv, Z., Kang, J., Ding, W., & Zhang, J.	a deep neural network model called MCP-FPmodel was developed	unique	fingerprint	reduce the time consumption of fingerprint key generation to <0.1 s.
[22]	Lin, C., He, J., Shen, C., Li, Q., & Wang, Q.	CrossBehaAuth	unique	keystroke	effectiveness of CrossBehaAuth in dynamic keystroke authentication. authentication accuracy.
[23]	Dash, P., Pandey, F., Sarma, M., & Samanta	dynamic biometric key generation approach from iris data	random	iris	proposed scheme of key generation is robust under brute-force attack, JPEG compression
[24]	Suresh, K., Pal, R., & Balasundaram, S. R. (2023).	a novel Gray code	unique	fingerprint	reduces the number of mismatch bits between the generated bit strings from the two instances of the same fingerprint

### 3.1 Histogram of Oriented Gradients

" HOG " is a popular gradient-based feature descriptor widely recognized for its efficacy in object detection. It leverages the identification of object characteristics based on gradients, classification of gradient occurrences using histograms, and object categorization through support vector machines. While HOG is primarily employed for pedestrian detection, it has found applications in various domains such as traffic sign detection, face detection, handwritten digit recognition, landmine detection, disaster management, biomedical imaging, and more. The process involves normalizing the captured image to enable efficient processing. intensity values extracted from three color planes (R, G, B planes). The process involves calculating the gradient of each pixel with respect to its neighboring points. The gradients are then classified into incidence segment frequencies along x and y directions, creating a 9-

element vector for each 8 x 8 cell. Illumination effects are mitigated by normalizing intensities within blocks created by cascading cells. The resulting blocks generate a string of HOG descriptors. These descriptors are then fed into an SVM classifier, which analyzes the data using predefined weights, studying representative features to classify the image as either human or non-human [31].

### 3.2 Scale-Invariant Feature Transform

The Scale-Invariant Feature Transform " SIFT " algorithm is utilized in computer vision to detect and describe local features within images. Created by David Lowe in 1999, SIFT has become a widely recognized and utilized technique in the field. And has since become a widely used technique for various applications, including object recognition, image stitching, and 3D reconstruction. The main steps involved in the calculation of SIFT features are as follows:

The proposed method consists of the following steps:

1. Detection of extrema in a Laplacian-of-Gaussian "LoG" scale space to identify potential interest points.
2. Refinement of key points by fitting a continuous model to determine their precise location and scale.
3. Assignment of orientation to each key point based on the dominant direction of the surrounding image gradients.
4. Formation of the feature descriptor by normalizing the local gradient histogram.

These steps collectively contribute to the generation of robust and informative feature descriptors [32].

#### 4. Key

In cryptography, a key is a piece of information that is used to encode or decode cryptographic data. A key in cryptography typically consists of a sequence of numbers or letters, stored in a file, and utilized in cryptographic algorithms to perform encryption or decryption operations on data. There are different types of cryptographic keys, including symmetric keys, asymmetric keys,

The security of a key relies on the method of its exchange between parties, emphasizing the importance of establishing a secure communication channel. This ensures that unauthorized individuals are unable to intercept or obtain the key. Cryptographic keys are crucial to the security of digital assets and data, and they play important roles in several crucial cryptographic functions, including data encryption, decryption, and authentication.

Symmetric key cryptography, also referred to as secret key cryptography, is a cryptographic technique where the same key is used for both encryption and decryption of data. In this method, a shared secret key is maintained between the sender and receiver, ensuring its confidentiality [6][26]. Symmetric key cryptography plays a crucial role in ensuring security in modern electronic communications [7].

Asymmetric key cryptography, also known as public key cryptography, operates using a pair of mathematically linked keys: a public key and a private key. The public key can be freely distributed and shared, while the private key remains confidential and known only to the owner [8][26].

##### 4.1. Proposed Method

The process of generating the encryption key from the iris is done by following the steps in the proposed method, but the first downloading a set of high-resolution images from the Internet for use in the program. Then the following steps are:

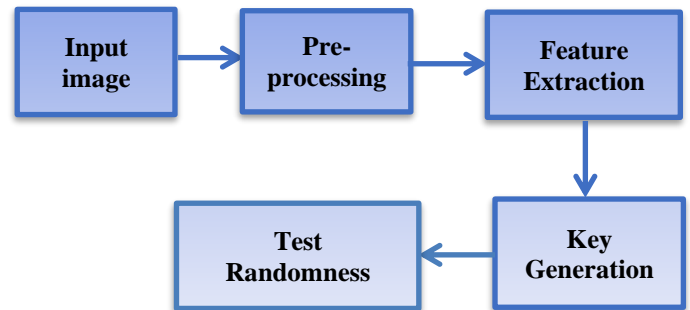
- 1) Enter any two eye pictures that we have been downloaded.
- 2) Image processing: The inner circle of the eye (pupil) was selected after applying morphological filters, then keeping the lower part of the iris as there were no eyelashes to

obscure the iris data.

3) Feature extraction: After that, the properties of the points inside the iris were extracted for the two images using two methods HOG (Histogram of Oriented Gradients) and SIFT (Scale-Invariant Feature Transform). Each method was applied to both images.

4) Creating the key: The key was created from the characteristics that were extracted using the two mentioned and for both two images together.

5) Key testing: The key was tested using randomness measures (chi-square test, ENT test, block\_metric, gab\_metric). see Fig. 3.



**Fig. 3.** Block Diagram for proposed method of generating and testing key.

#### 5. Randomness

Randomness holds significant significance within the domain of information security, particularly in cryptography. The assessment of random numbers' quality through various randomness testing techniques is crucial to ensure their suitability for specific applications [28].

Randomness testing is a testing method used to evaluate the quality and randomness of a sequence of numbers or bits. This test aims to determine whether the sequence exhibits characteristics of true randomness or whether there are any patterns, biases or expectations present in it [28]. It is one of the most popular evaluation tools for evaluating pragmatic random number generators (PRNGs). These tests check certain statistical properties of random numbers and evaluate the extent to which the tested sequence satisfies those properties. The result of applying these tests is usually a quantitative assessment [33].

##### 5.1 Type of Used Metrics

###### 5.1.1 Chi-Square Test

The Pearson chi-squared statistical test is utilized to ascertain whether a significant difference exists between the observed values and the expected values in a distribution involving two variables. This test helps in assessing the degree of association or independence between the variables based on the comparison of observed and expected frequencies. By

analyzing the chi-squared statistic, researchers can determine the statistical significance of the relationship between the variables under investigation. [30].

The equation for the Chi-squared statistic is as follows:

$$\chi^2 = \sum [(O - E)^2 / E]$$

Here:

- $\chi^2$  represents the Chi-squared statistic.
- $\Sigma$  denotes the summation symbol, indicating that the equation is summed over all categories or cells.
- O represents the observed frequencies (the actual observed counts in the data).
- E represents the expected frequencies (the expected counts under the assumption of independence between variables).

### 5.1.2 ENT Test

ENT, developed by John Walker in 1998, is a Pseudorandom Number Sequence Test Program. It incorporates five standard tests to assess randomness:

1. Entropy
2. Chi-square test
3. Arithmetic mean
4. Monte Carlo estimation of PI
5. Serial correlation

These tests utilize mathematical metrics to detect significant deviations from randomness. The straightforward nature of the ENT test utility makes it a convenient tool for swiftly evaluating the randomness of bit sequences [29].

### 5.1.3 The Block Metric and Gab Metric

The Block Metric, also known as the Block Distance or Hamming Distance, is a measure of dissimilarity between two equal-length binary strings. It calculates the number of positions at which the corresponding bits in the two strings differ. To calculate the Block metric, follow these steps:

1. Take two binary strings of equal length.
2. Compare the bits at each position in the two strings.
3. Increment a counter each time the bits differ.
4. Repeat the comparison for all positions in the strings.
5. The final count represents the Block metric or Hamming Distance.

$$\text{Block Metric} = \sum |\text{bit1} - \text{bit2}|$$

Here a)  $\Sigma$  represents the summation symbol, indicating that the equation is summed over all positions or bits in the strings and b)  $|\text{bit1} - \text{bit2}|$  represents the absolute difference between the corresponding bits at each position in the two strings [30].

## 5.2 What is the importance of the encryption key when it is Random?

Random encryption key plays a crucial role in ensuring security and confidentiality in encryption operations. Here are some of its importance when it is randomly generated:

1. Data Confidentiality: A random encryption key

hides the original data and makes it unreadable to anyone other than the legitimate recipient of the key. When the encryption key is running

2. dom, it is difficult for an attacker to predict or deduce, which increases the confidentiality of the encrypted data.
3. Resistance to attacks: If the encryption key is random and generated correctly, it enhances the system's resistance against known attacks and attackers. Even if there is an attempt to hack the system by trying all possible keys, the chance of guessing the random key is very small.
4. Complexity of breaking: A random encryption key increases the complexity of the process of cracking the system or decrypting the data. When the key is random, it increases the mathematical and computational complexity needed to discover the pattern or relationship between the key and the encrypted data.
5. Scalability: When an encryption key is randomly generated, it can be used in various cryptographic systems and secure key exchanges, providing scalability and flexibility in cryptographic applications.

In general, a random encryption key is a crucial element to achieve the security of encrypted data and information. Key generation and management must be done properly according to the recognized best and strongest cryptographic standards to ensure data confidentiality and protection.

## 5.3 Results

**Table. 2** and **Table. 3** show the length of the key obtained from the two irises using the two methods proposed for extracting properties, HOG and SIFT, while finding random parameter values for each key. We notice that the values of the metrics change as the length of the key changes.

One common statistical test for randomness is the chi-square test, Use the chi-square test to determine whether the distribution is random or not by relying on the value resulting from the chi-square equation and comparing it with the value of 0.05 (scientists have determined it). If the chi-square value is greater than 0.05, the system achieves randomness and vice versa.

The output of the ENT tests provides information about the statistical properties of the input sequence. Based on the test results, we can decide about the randomness of the sequence. Here are some guidelines for interpreting the output of the ENT tests:

Entropy test: This test measures the entropy of the sequence, which is a measure of how much randomness the sequence contains. A higher entropy value indicates more randomness. Chi-square test: A higher number of results (matches to the expected pattern) indicates a more random sequence and vice versa.

Arithmetic mean value test: A statistical measurement that provides information about the mean value of a set of

data collected during a test. It is calculated by summing all the values extracted from the test and dividing them by the total number of data points. A higher number of hits (i.e. matches with the expected pattern) indicates a more random sequence and vice versa. Arithmetic mean value in the ENT test (Auditory and Equilibrium Test) refers to the arithmetic mean value of the test results. The arithmetic mean value is calculated by summing all the measurements and then dividing them by their total number. Monte Carlo Pi test: This test uses the Monte Carlo method to estimate the value of Pi ( $\pi$ ) based on the binary sequence. A higher number of hits (matches to the expected pattern) indicates a more random sequence. If the number of hits is significantly lower than expected, it may indicate that the sequence is not random. Serial correlation test: This test measures the correlation between adjacent bits in the sequence. A low correlation indicates more randomness. If the correlation is significantly higher than expected, it may indicate that the sequence is not random. Gab metric and Block metric are calculated by applying their equation and comparing the result with the threshold values that were specified (gab threshold = 3.84, block threshold = 3.841).

If the values of the Gab metric and Block metric are less than or equal to the specified threshold, the system achieves randomness. For more explain show **Fig. 4, 5 6, 7, 8 and 9** All of this achieves randomness through the values determined by randomness measures, but **Fig. 10** does not achieve randomness because the values did not meet the conditions of randomness for the measures themselves.

```

----- Check the randomness -----
The chi-squared test:
The binary key is likely random.
ent - pseudorandom number sequence test:
Entropy = 3.000000 bits per byte.

Optimum compression would reduce the size
of this 8 byte file by 62 percent.

Chi square distribution for 8 samples is 248.00, and randomly
would exceed this value 61.15 percent of the times.

Arithmetic mean value of data bytes is 103.8750 (127.5 = random).
Monte Carlo value for Pi is 4.000000000 (error 27.32 percent).
Serial correlation coefficient is -0.131701 (totally uncorrelated = 0.0).

Block Metric:
Block metric value: 6

Gab Metric:
Gab metric value: 0.5625
Random: True
    
```

Fig. 4. Execution of system when key = 64.

```

----- Check the randomness -----
The chi-squared test:
The binary key is likely random.
ent - pseudorandom number sequence test:
Entropy = 3.875000 bits per byte.

Optimum compression would reduce the size
of this 16 byte file by 51 percent.

Chi square distribution for 16 samples is 272.00, and randomly
would exceed this value 22.18 percent of the times.

Arithmetic mean value of data bytes is 113.5625 (127.5 = random).
Monte Carlo value for Pi is 4.000000000 (error 27.32 percent).
Serial correlation coefficient is 0.105418 (totally uncorrelated = 0.0).

Block Metric:
Block metric value: 22

Gab Metric:
Gab metric value: 3.78125
Random: True
    
```

Fig. 5. Execution of system when key = 128.

```

----- Check the randomness -----
The chi-squared test:
The binary key is likely random.
ent - pseudorandom number sequence test:
Entropy = 4.392317 bits per byte.

Optimum compression would reduce the size
of this 21 byte file by 45 percent.

Chi square distribution for 21 samples is 235.00, and randomly
would exceed this value 81.06 percent of the times.

Arithmetic mean value of data bytes is 128.5714 (127.5 = random).
Monte Carlo value for Pi is 2.666666667 (error 15.12 percent).
Serial correlation coefficient is -0.236285 (totally uncorrelated = 0.0).

Block Metric:
Block metric value: 2

Gab Metric:
Gab metric value: 0.023809523809523808
Random: True
    
```

Fig. 6. Execution of system when key = 168.

```

----- Check the randomness -----
The chi-squared test:
The binary key is likely random.
ent - pseudorandom number sequence test:
Entropy = 4.937500 bits per byte.

Optimum compression would reduce the size
of this 32 byte file by 38 percent.

Chi square distribution for 32 samples is 240.00, and randomly
would exceed this value 74.15 percent of the times.

Arithmetic mean value of data bytes is 148.1563 (127.5 = random).
Monte Carlo value for Pi is 2.400000000 (error 23.41 percent).
Serial correlation coefficient is -0.217070 (totally uncorrelated = 0.0).

Block Metric:
Block metric value: 20

Gab Metric:
Gab metric value: 1.5625
Random: True
    
```

Fig. 7. Execution of system when key = 256.

```

----- Check the randomness -----
The chi-squared test:
The binary key is likely random.
ent - pseudorandom number sequence test:
Entropy = 5.843750 bits per byte.

Optimum compression would reduce the size
of this 64 byte file by 26 percent.

Chi square distribution for 64 samples is 232.00, and randomly
would exceed this value 84.64 percent of the times.

Arithmetic mean value of data bytes is 126.2969 (127.5 = random).
Monte Carlo value for Pi is 3.200000000 (error 1.86 percent).
Serial correlation coefficient is 0.263614 (totally uncorrelated = 0.0).

Block Metric:
Block metric value: 10

Gab Metric:
Gab metric value: 0.1953125
Random: True
    
```

Fig. 8. Execution of system when key = 512.

```

----- Check the randomness -----
The chi-squared test:
The binary key is likely random.
ent - pseudorandom number sequence test:
Entropy = 6.507660 bits per byte.

Optimum compression would reduce the size
of this 128 byte file by 18 percent.

Chi square distribution for 128 samples is 264.00, and randomly
would exceed this value 33.60 percent of the times.

Arithmetic mean value of data bytes is 133.1719 (127.5 = random).
Monte Carlo value for Pi is 2.857142857 (error 9.05 percent).
Serial correlation coefficient is -0.073947 (totally uncorrelated = 0.0).

Block Metric:
Block metric value: 44

Gab Metric:
Gab metric value: 1.070625
Random: True
    
```

Fig. 9. Execution of system when key=1024.

```

----- Check the randomness -----
The chi-squared test:
The binary key is not likely random.
ent - pseudorandom number sequence test:
Entropy = 3.000000 bits per byte.

Optimum compression would reduce the size
of this 8 byte file by 62 percent.

Chi square distribution for 8 samples is 248.00, and randomly
would exceed this value 61.15 percent of the times.

Arithmetic mean value of data bytes is 155.5000 (127.5 = random).
Monte Carlo value for Pi is 4.000000000 (error 27.32 percent).
Serial correlation coefficient is -0.075898 (totally uncorrelated = 0.0).

Block Metric:
Block metric value: 24

Gab Metric:
Gab metric value: 9.0
Random: False
    
```

Fig. 10. Execution of system when key = 64 but not random.

**Table. 2.** Results when using HOG.

key as bit	Check Randomness						
	Entropy	Chi square	Arithmetic mean value	Serial correlation coefficient	Monte Carlo value	Block Metric	Gab Metric
64	3.000000	248.00	103.8750	-0.131701	4.00000	6	0.5625
128	3.875000	272.00	113.5625	0.105418	4.000000	22	3.78125
168	4.392317	235.00	128.5714	-0.234288	2.666666	2	0.023809523
256	4.937500	240.00	148.1563	-0.217070	2.4000000	20	1.5625
512	5.843750	232.00	126.2969	0.263614	3.200000000	10	0.1953125
1024	6.507660	264.00	133.1719	-0.073947	2.857142857	44	1.890625

**Table. 3.** Results when using SIFT.

key as bit	Check Randomness						
	Entropy	Chi square	Arithmetic mean value	Serial correlation coefficient	Monte Carlo value	Block Metric	Gab Metric
64	3.000000	248.00	91.6250	-0.322892	4.00000	10	1.5625
128	4.000000	240.000	132.8750	0.087591	4.000000	22	3.78125
168	4.392317	235.00	120.8571	-0.066369	2.666666	4	0.095238095
256	4.937500	240.00	142.7188	-0.081372	1.6000000	16	1.0
512	5.714615	280.000	143.1250	-0.073191	2.0000000	52	5.28125
1024	6.526334	276.00	125.4844	0.060396	3.23809523	14	0.1914062

## 6. Conclusion

Given the importance of vital features in our daily lives, and the security and uniqueness they provide, especially the iris, the primary goal of the study was to generate an encryption key from the iris of two different eyes, test its randomness, and use two algorithms to extract the properties (HOG, SIFT feature Extraction). This goal, indeed, was achieved as it was the generation of the key and this key is random, as its randomness is tested using a set of randomness measures (chi-square test, ENT test, Gab and Block metric). and the generated key is of different lengths as in **Tables. 2 and 3** while ensuring the randomness measures. It is possible to generate encryption keys and test their randomness using two biometric features, such as the iris and fingerprint, and testing their randomness using random measures.

## Acknowledgement

The authors would like to thank the Department of Computer Science and Department of Software Engineering in the College of Computer Science and Mathematics at the University of Mosul for their help in completing this paper.

## References

- [1] Alsaadi, I. M. (2021). Study on most popular behavioral biometrics, advantages, disadvantages and recent applications: A review. *Int. J. Sci. Technol. Res.*, 10(1).
- [2] Debas, E. A., Alajlan, R. S., & Rahman, M. H. (2023, February). Biometric in Cyber Security: A Mini Review. In 2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC) (pp. 570-574). IEEE.
- [3] Kaur, P., Kumar, N., & Singh, M. (2023). Biometric cryptosystems: a comprehensive survey. *Multimedia Tools and Applications*, 82(11), 16635-16690.
- [4] Malgheet, J. R., Manshor, N. B., Affendey, L. S., & Abdul Halin, A. B. (2021). Iris recognition development techniques: a comprehensive review. *Complexity*, 2021, 1-32.
- [5] Sonkar, K., & Rani, R. (2021, May). Cancelable Iris Biometric: A Review. In 2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC) (pp. 560-565). IEEE.
- [6] Hughes, L. E. (2022). Basic Cryptography: Symmetric Key Encryption. In *Pro Active Directory Certificate Services: Creating and Managing Digital Certificates for Use in Microsoft Networks* (pp. 3-17). Berkeley, CA: Apress.
- [7] Baksi, A. (2021, October). Classical and physical security of symmetric key cryptographic algorithms. In 2021 IFIP/IEEE 29th International Conference on Very Large Scale Integration (VLSI-Soc) (pp. 1-2). IEEE.
- [8] Banoth, R., & Regar, R. (2023). Asymmetric Key Cryptography. In *Classical and Modern Cryptography for Beginners* (pp. 109-165).



- Cham: Springer Nature Switzerland.
- [9] Sun, Y., Li, H., & Li, N. (2023). A novel cancelable fingerprint scheme based on random security sampling mechanism and relocation bloom filter. *Computers & Security*, 125, 103021.
- [10] Xia, W., Yang, Y., Xue, J. H., & Wu, B. (2021). Tedigan: Text-guided diverse face image generation and manipulation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2256-2265).
- [11] Mohammed, H. H., Baker, S. A., & Nori, A. S. (2021, February). Biometric identity authentication system using hand geometry measurements. In *Journal of Physics: Conference Series* (Vol. 1804, No. 1, p. 012144). IOP Publishing.
- [12] Al-kateeb, Z. N., & Mohammed, S. J. (2020). A novel approach for audio file encryption using hand geometry. *Multimedia Tools and Applications*, 79(27-28), 19615-19628.
- [13] Vijayakumar, T. (2021). Synthesis of palm print in feature fusion techniques for multimodal biometric recognition system online signature. *Journal of Innovative Image Processing (JIIP)*, 3(02), 131-143.
- [14] Reja, M., Pungila, C., & Negru, V. (2021). Towards real-time DNA biometrics using GPU-accelerated processing. *Logic Journal of the IGPL*, 29(6), 906-924.
- [15] Devi, R. M., Keerthika, P., Suresh, P., Sarangi, P. P., Sangeetha, M., Sagana, C., & Devendran, K. (2022). Retina biometrics for personal authentication. In *Machine Learning for Biometrics* (pp. 87-104). Academic Press.
- [16] Wang, Y., Li, B., Zhang, Y., Wu, J., Yuan, P., & Liu, G. (2020, October). A biometric key generation mechanism for authentication based on face image. In *2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP)* (pp. 231-235). IEEE.
- [17] Majjed, I. A., & Majeed, A. A. (2020, September). Key Generation Based on Facial Biometrics. In *Proceedings of the 1st International Multi-Disciplinary Conference Theme: Sustainable Development and Smart Planning, IMDC-SDSP 2020, Cyperspace, 28-30 June 2020*.
- [18] Roy, N. D., & Biswas, A. (2020). Fast and robust retinal biometric key generation using deep neural nets. *Multimedia Tools and Applications*, 79(9-10), 6823-6843
- [19] Wu, Y., Lin, Q., Jia, H., Hassan, M., & Hu, W. (2020). Auto-Key: Using autoencoder to speed up gait-based key generation in body area networks. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4 (1), 1-23
- [20] Wang, P., You, L., Hu, G., Hu, L., Jian, Z., & Xing, C. (2021). Biometric key generation based on generated intervals and two-layer error correcting technique. *Pattern Recognition*, 111, 107733.
- [21] Wu, Z., Lv, Z., Kang, J., Ding, W., & Zhang, J. (2022). Fingerprint bio-key generation based on a deep neural network. *International Journal of Intelligent Systems*, 37(7), 4329-4358.
- [22] Lin, C., He, J., Shen, C., Li, Q., & Wang, Q. (2022). CrossBehaAuth: Cross-Scenario Behavioral Biometrics Authentication Using Keystroke Dynamics. *IEEE Transactions on Dependable and Secure Computing*.
- [23] Dash, P., Pandey, F., Sarma, M., & Samanta, D. (2023). Efficient private key generation from iris data for privacy and security applications. *Journal of Information Security and Applications*, 75, 103506.
- [24] Suresh, K., Pal, R., & Balasundaram, S. R. (2023). A stable cryptographic key generation from fingerprint biometrics using Gray code for secure data storage. *International Journal of Information and Computer Security*, 20 (3-4), 366-398.
- [25] Roy, N. D., & Biswas, A. (2020). Fast and robust retinal biometric key generation using deep neural nets. *Multimedia Tools and Applications*, 79(9-10), 6823-6843
- [26] Stallings, W. (2017). *The principles and practice of cryptography and network security* 7th edition, isbn-10: 0134444280. Pearson Education, 20(1), 7.
- [27] Kumar, K. K., Chaduvula, K., & Markapudi, B. (2020). A Detailed Survey on feature extraction techniques in image processing for medical image analysis. *European Journal of Molecular & Clinical Medicine*, 7(10), 2020.
- [28] Mengdi, Z., Xiaojuan, Z., Yayun, Z., & Siwei, M. (2021, March). Overview of Randomness Test on Cryptographic Algorithms. In *Journal of Physics: Conference Series* (Vol. 1861, No. 1, p. 012009). IOP Publishing.
- [29] Okech, P., & Mc Guire, N. (2010). Analysis of statistical properties of inherent randomness. In *Proc. 12th Real-Time Linux Workshop* (pp. 1-8).
- [30] Luna-Romera, J. M., Martínez-Ballesteros, M., García-Gutiérrez, J., & Riquelme, J. C. (2019). External clustering validity index based on chi-squared statistical test. *Information Sciences*, 487, 1-17.
- [31] *Pattern Recognition and Machine Learning* by Christopher M. Bishop Divyashree, N., & Pushpalatha, K. N. (2020). A Review on Gradient Histograms for Texture Enhancement and Object Detection.
- [32] Burger, W., & Burge, M. J. (2022). Scale-invariant feature transform (SIFT). In *Digital Image Processing: An Algorithmic Introduction* (pp. 709-763). Cham: Springer International Publishing.
- [33] Hernandez-Castro, J., & Barrero, D. F. (2017, June). Evolutionary generation and degeneration of randomness to assess the independence of the Ent test battery. In *2017 IEEE Congress on Evolutionary Computation (CEC)* (pp. 1420-1427). IEEE.