Q-Learning and Chaotic Map-Based Image Encryption

[1]Dina Riadh Alshibani, [2]Musaab Riyadh and [3]Narjis Mezaal Shati
[1,2,3]Department of Computer Science, College of Sciences, Mustansiriyah University, Baghdad, Iraq.
[1]dinashibani@uomustansiriyah.edu.iq, [2]m.shaibani@uomustansiriyah.edu.iq, [3]dr.narjis.m.sh@uomustansiriyah.edu.iq

Correspondence should be addressed to Narjis Mezaal Shati: dr.narjis.m.sh@uomustansiriyah.edu.iq

Article Info
Journal of Machine and Computing (https://anapub.co.ke/journals/jmc/jmc.html)
Doi : https://doi.org/10.53759/7669/jmc2025202505213
Received 04 July 2025; Revised from 17 August 2025; Accepted 27 September 2025.
Available online 05 October 2025.

Abstract – In this research paper, a new image encryption algorithm based on Q-learning and chaotic maps is proposed. Entitled QLCMIE, this algorithm consists of two main steps: the first involves scattering pixel locations using the key generated by the Q-learning algorithm. The second step substitutes the pixels with a chaotic key based on the XOR operation of a chaotic map. Multiple experiments have been carried out, and their outcomes have been compared with those from other researchers. The results demonstrated that the proposed encryption method offers a higher level of security than the current techniques in terms of PSNR, entropy, UACI, and NPCR.

Keywords – Entropy, Q-Learning, Tent Map, Correlation.

I. INTRODUCTION

Recently, multimedia security has attracted the attention of many researchers due to the explosive growth in communications technology and smart devices [1]. Image security, in particular, has increased the need for an additional layer of protection due to its ability to accommodate massive amounts of data, its high redundancy, and its strong pixel consistency. Because of their time-consuming nature and poor encryption quality, conventional encryption techniques are therefore inappropriate for image encryption. [2,3].

Lately, various image encryption methods have been proposed in the literature, including cellular automata-based techniques [4,5], DNA-based techniques [6-9], and metaheuristic-based techniques [10]. Most of these methods suffer from a high computational cost and complexity.

As a result, there has been a growing demand for new, more secure encryption techniques with reasonable computational costs. Currently, for image encryption algorithms chaotic maps and reinforcement learning algorithms seem to be the optimal solution, offering significant value in complementing traditional encryption methods and enhancing protection against cyberattacks [11].

On the one hand, chaos has unpredictable properties, making it an ideal choice for securing digital information from unauthorized access. The inherent complexity and unpredictability of chaotic systems provide a robust framework for image encryption and decryption. Chaos can produce a complex sequence of numbers to replace image pixel values via the XOR operation. This ensures that the encrypted image is just indistinct to the naked eye, and computationally, it will be impossible to reconstruct it without the decryption key. On the other hand, reinforcement learning algorithms, especially the Q-Learning algorithm, which address how autonomous agents learn to choose appropriate actions to achieve their goals by interacting with the environment, make it a suitable choice to generate an image permutation key and destroy the image correlation [12,13].

The remainder of the paper is organized as follows: Section 2 introduces the theoretical foundation, Q learning, and tent map. The proposed implementation is illustrative in Section 3. In section 4, the outcome investigation results of the proposed approach are discussed, and the final section concludes.

II. MATERIALS AND METHODS

In this regard, the theoretical details of the Tent Map and Q-learning are explored in the context of image encryption.

Tent Map

Eq. 1[14] specified the mathematical representation of the tent map which is a one-dimensional chaotic map.

$$xT_{n+1} = \begin{cases} \frac{xT_n}{\alpha} & xT_n = [0, \alpha) \\ \frac{1-xT_n}{1-\alpha} & xT_n = [\alpha, 1] \end{cases} \quad (1)$$

Where xT_{n+1} and xT_n are the newly produced chaotic values and the initial chaotic value, respectively, while α is the control parameter. This map evolves chaotically when α =0.5.

The Q-Learning Algorithm

A model-free reinforcement learning algorithm like Q-learning. Simple value iteration update is the fundamental concept of Q-learning, where each pair of state-action (š, α) has an associated Q-value [15].

When the agent is in state š and chooses an action α, the Q-value for that state-action pair is updated by the received reward resulting from taking that action and the next state greatest Q-value, š+1. These values (Q-values) are stored in a Q-table, allowing the agent to quickly retrieve them when needed. The function of Bellman, which updates the Q-values, is defined in Eq.2[16].

$$Q_{new}(\check{s}, \alpha) = (1 - \delta) \times Q_{old}(\check{s}, \alpha) + \delta \times (r_t + \gamma \times \max(Q(\check{s} + 1, \alpha))) \quad (2)$$

where the learning rate is (α), (1− α) is the probability of maintaining the old Q value. r_t represents the reward value, and discount factor (γ) balances present and future rewards. Max (Q(š+1, α))) represents the maximum Q-value of the next state.

The ϵ-greedy policy is used in Q-learning to choose actions based on current Q-value estimates. With a probability of ϵ, the agent selects a random action to explore new options, helping it uncover potentially better strategies for earning rewards. This exploration helps the agent improve its learning and performance over time. In contrast, the agent selects the action with the highest (Q-value) when the probability is 1−ϵ, leveraging its existing knowledge to gain the maximum possible reward [17].

III. IMPLEMENTATION OF THE PROPOSED QLCMIE

The proposed QLCMIE algorithm consists of three main stages. The first stage focuses on converting a normal color image to a grayscale image and resizing it to an L × L size. The second stage aims to uncorrelate the image by swapping pixel locations. The main idea of the third stage is to replace image pixels.

The baseline of the proposed QLCMIE algorithm is described in the following steps:

- Preprocess stage:
  - Load a color image (cimg).
  - Convert cimg into a grayscale image (gryimg).
  - Resize gryimg to be in size L × L.
- Permutation stage:
  - Construct a pGryimg image by using Q-learning to swap gryimg pixels.
- Substitution stage:
  - Generate substitution key (subKey) by using the Tent chaotic map.
  - Obtain the final encrypted image eimg by XORing pGryimg with subKey.

Fig 1 illustrates the overall design of the proposed QLCMIE algorithm. The decryption process is much like the previously mentioned steps, but, it is done in reverse.

Preprocess Stage

This step serves as a prelude to the next step, where a color image is loaded and converted to grayscale as defined in Eq.3. The image is then converted to a square image with dimensions of L × L.

$$\text{cimg}(i, j) = 0.299 \times R(i, j) + 0.587 \times G(i, j) + 0.114 \times B(i, j) \quad (3)$$

where the pixel locations are denoted by i and j. The color pixel's red, green, and blue components are denoted by the letters R, G, and B. The weights of each component are 0.299, 0.587, and 0.114, which are determined based on the sensitivity of the human eye.

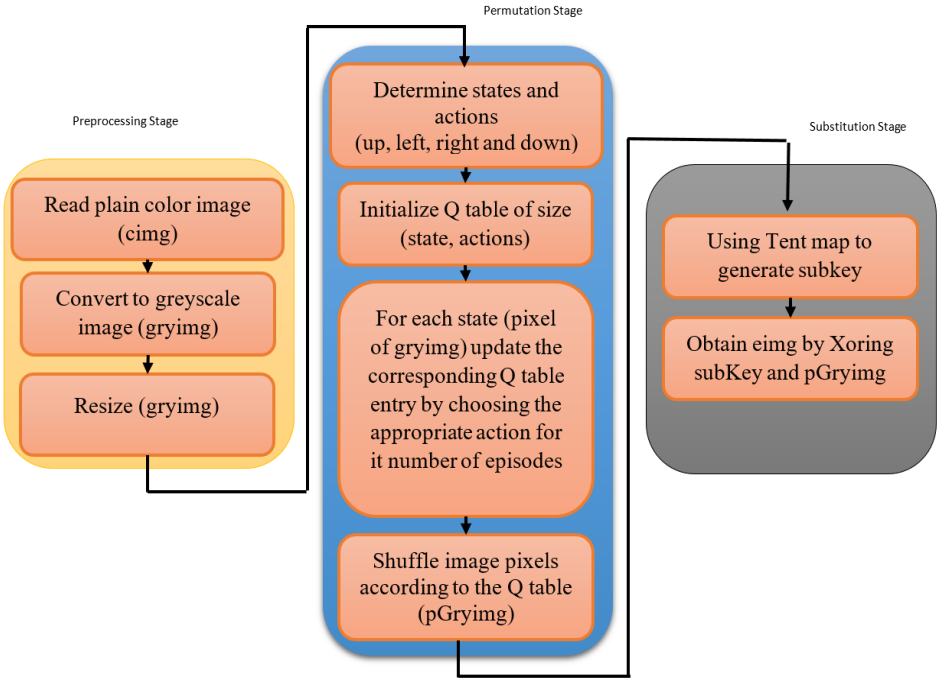

Fig 1. The Block Diagram of the Proposed QLCMIE.

Permutation Stage

This stage consists mainly of 4 steps: initialization, reward function, training, and image permutation implementation.

- In the initialization step, the state and actions are defined. State here represents the pixel's positions (L ×L) while the action is defined by four movements (up, down, right, and left). Q-table is initialized with the size of (L ×L ×action).
- Reward Function uses a correlation-based reward to encourage high pixel differences as defined in Eq.4.

$$\text{reward} = |I(x,y) - I(x',y')| / 255 \tag{4}$$

where the image pixel value at position x,y is represented by I(x,y), and the image pixel value at position (x',y') is represented by I(x',y'). 255 is the highest density in the image. Note that the greater pixel difference gives a higher reward, pushing the agent to produce better (more secure) encrypt.

- The training step involves running multiple episodes. In each episode, start at a random pixel, move to a new pixel using an epsilon-greedy strategy, and update Q-values as defined in Eq.2 after computing the reward function.
- Implementation of image permutation: based on the optimal action derived from the Q-table, each pixel of the original image (x, y) is relocated to the new location (x', y').

Generation of Substitution Keys

The substitution key (subKey) has been generated using the Tent map by iterating the Tent map as specified in Eq.1 for L×L, where L is the shuffled image width and height, respectively, producing a series of random real values rKey. The next step is to convert rKey into integer numbers, subKey as defined in Eq.5. **Fig 2** shows the substitution stage details.

$$subKey(i,j) = (rkey(i,j) \times 10^{20}) \; mod \; 255 \tag{5}$$

Finally, the encrypted image is produced by XORing the subKey with the permutated image. As defined in Eq.6.

$$eimg(i,j) = eimg(i,j-1) \oplus \text{pGryimg}\,(i,j) \oplus subKey(i,j) \tag{6}$$

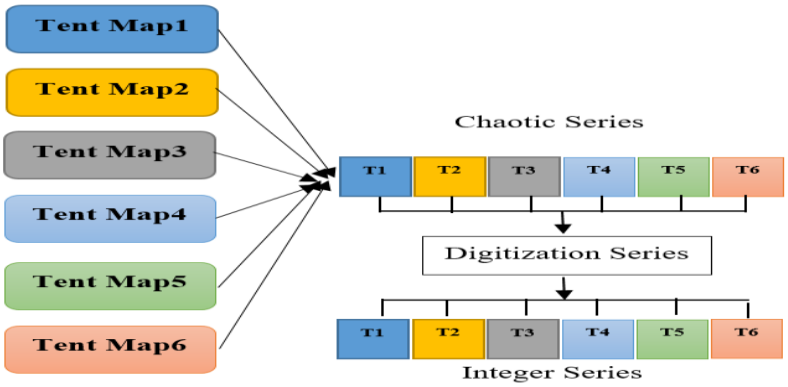

Fig 2. The Substitution Stage Details.

IV. RESULTS ANALYSIS

This section introduces a thorough analysis of the performance of the proposed QLCMIE algorithm's. **Fig 3** depicts the QLCMIE algorithm's outcomes with their corresponding histogram.

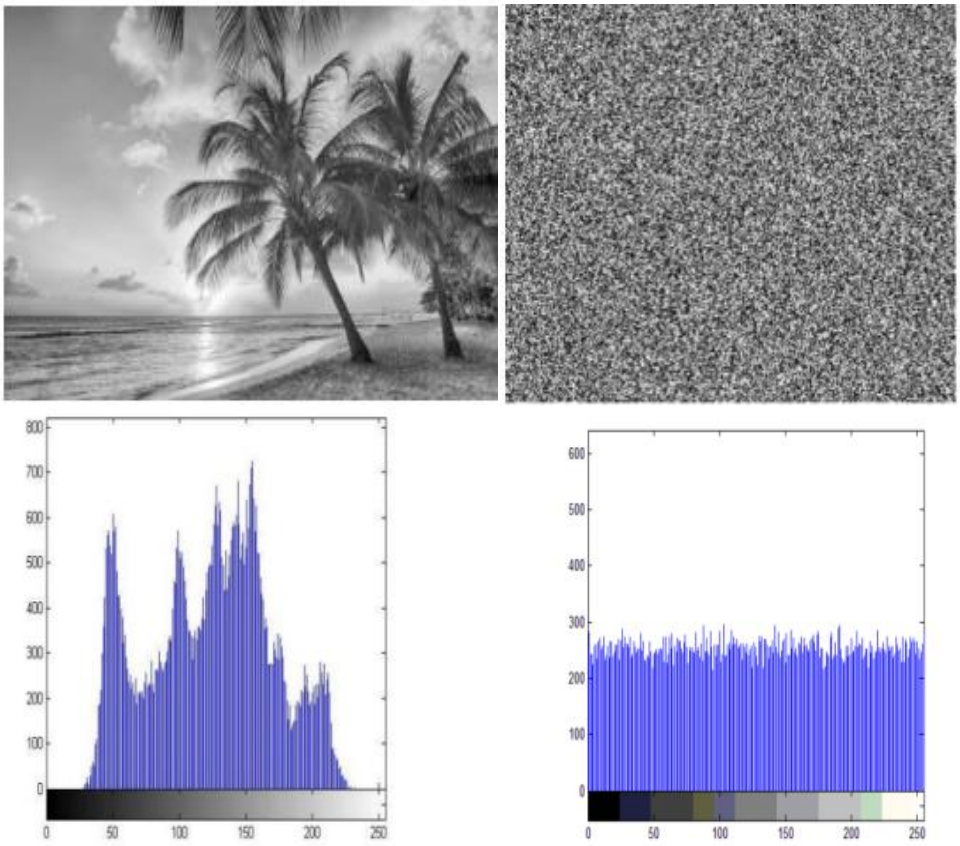

Fig 3. The Outcome of the Proposed QLCMIE Algorithm with Its Corresponding Histogram.

Histogram Analysis

The histograms of the plain image and its corresponding encrypted one have been analyzed. **Fig 3** shows the of the original and encrypted image histograms, respectively. The figure makes it evident that, despite significant bumps in the original image's histograms, the equivalent encrypted image's histograms are uniform and significantly different from the original image, not statistically similar. This indicates that a robust defense against numerical attacks is offered by the suggested (QLCMIE).

Information Entropy

Information entropy measures the level of randomness. Eq. 7 specified the entropy formulation [18].

$$H(m) = \sum_{i=1}^{X} p(m_i) log \frac{1}{p(m_i)} \quad (7)$$

Where p(mi) is the probability of occurrence of symbol m's, is, and X is a representation of how many bits each symbol has. The optimal value of entropy equal (8). The entropy values of the plain image and the matching encrypted image are displayed in **Table 1**.

Table 1. Plain and Encrypted Image Entropy Values

Image	Plain image	Encrypted image
Img1	7.1821	7.9998
Img2	7.3686	7.9993
Img3	7.3533	7.9989

Table 1 makes it evident that the encrypted images' entropy is closer to (8). This demonstrates how the proposed QLCMIE algorithm can resist statistical attacks by adding a significant degree of randomness to the cipher image, making it extremely difficult to conclude any information.

Differential Attack

The impact of changing a single pixel in a plain image on the corresponding ciphered image is investigated using the UACI and NPCR techniques. Equations 8 and 10 specify the mathematical model for the UACI and NPCR techniques [19].

$$UACI(a1, a2) = \frac{\sum_{i=1}^{L} \sum_{i=1}^{Z} |a1(i,j) - a2(i,j)|/255}{L \times Z} \times 100 \quad (8)$$

$$NPCR(a1, a2) = \frac{\sum_{i=1}^{L} \sum_{j=1}^{Z} DD(i,j)}{L \times Z} \times 100 \quad (9)$$

Where,

$$DD(i,j) \begin{cases} 0 & if\ a1(i,j) = a2(i,j) \\ 1\ if\ a1(i,j) \neq a2(i,j) \end{cases} \quad (10)$$

Table 2. The Values UACI and NPCR

No.	UACI Metric	NPCR Metric
Img1	33.99	99.99
Img2	33.98	99.99
Img3	33.97	99.97

The values of (UACI and NPCR) for the proposed QLCMIE algorithm are presented in **Table 2**, which shows that the (NPCR and UACI) values are close to 100% and 33.50, respectively, demonstrating that the proposed (QLCMIE) algorithm can resist differential attacks because of its great sensitivity to even little changes made to plain images and its high ability to withstand plaintext attacks.

PSNR

PSNR is the measure of the distortion that occurs in an image [20] [21]. The mathematical representation of PSNR is defined in Eqs. 11 and 12.

$$PSNR = 10 \cdot log10(\frac{255^2}{MSE}) \quad (11)$$

$$\text{MSE} = \frac{1}{X \times Y} \sum_{i=1}^{X} \sum_{y=1}^{Y} (O(i,j) - E(i,j))^2 \quad (12)$$

where the original image is denoted by O and its corresponding encrypted image by E. It's worth noting that a higher MSE results in a lower PSNR, indicating greater image distortion. This means the ciphering process is optimal [9]. **Table 3** presents the MSE and PSNR values of the proposed QLCMIE method, indicating a high level of distortion, which reflects the robustness and effectiveness of the encryption technique.

Table 3. PSNR and MSE Values

No.	MSE	PSNR
Img1	311.23	23.19
Img2	292.82	23.46
Img3	199.63	25.12

Key Space Analysis

The proposed QLCMIE algorithm uses Q-learning to permute the image, the number of episodes considered as part of the secret key, along with the initial state and control parameter of the used 6-Tent map with 20 floating points for each. Secret key $((10)^{20})^6$ number of episodes, it is generally considered sufficient to render brute-force attacks computationally infeasible with current and foreseeable computational technology. Therefore, the proposed QLCMIE algorithm possesses a key space large enough to provide a strong defense against exhaustive key search attacks.

Comparison with Existing Work

A comparison with existing approaches is given to illustrate the effectiveness of the proposed (QLCMIE) algorithm. **Table 4** illustrates the UACI, NPCR, and entropy for the proposed QLCMIE, [9] and [10]. As demonstrated by the results, the proposed (QLCMIE) technique is superior [9] and [10].

Table 4. Values for Comparison

approach	(UACI)	(NPCR)	(Entropy)
[9]	32.66	99.64	7.997
[10]	33.92	99.96	7.9995
QLCMIE	33.99	99.99	7.9998

V. CONCLUSION

In this research article, a greyscale image encryption based on Q-learning and a 1D Tent chaotic map has been presented, namely QLCMIE. The presented QLCMIE algorithm has two phases permutation phase and the substitution phase. Q-learning is employed in the permeation phase to uncorrelated image pixels, while the 1D Tent map supports the substitution phase. The correlation, UACI, NPCR, and entropy values are the greatest. According to the test results, the proposed (QLCMIE) algorithm has a high degree of security, successfully withstands a variety of statistical attacks. For future work, it can be extended to color images, integrating with Deep Q-learning (DQN).

ACKNOWLEDGEMENT

The author thanks the "Department of Computer Science", "College of Science", "Mustansiriyah University (www.uomustansiriyah.edu.iq)", Baghdad-Iraq for supporting this work.

CRediT Author Statement

The authors confirm contribution to the paper as follows:

Conceptualization: Dina Riadh Alshibani, Musaab Riyadh and Narjis Mezaal Shati; **Methodology:** Dina Riadh Alshibani and Musaab Riyadh; **Software:** Narjis Mezaal Shati; **Data Curation:** Dina Riadh Alshibani and Musaab Riyadh; **Writing-Original Draft Preparation:** Dina Riadh Alshibani, Musaab Riyadh and Narjis Mezaal Shati; **Visualization:** Narjis Mezaal Shati; **Investigation:** Dina Riadh Alshibani and Musaab Riyadh; **Supervision:** Narjis Mezaal Shati; **Validation:** Dina Riadh Alshibani, Musaab Riyadh; **Writing- Reviewing and Editing:** Dina Riadh Alshibani, Musaab Riyadh and Narjis Mezaal Shati; All authors reviewed the results and approved the final version of the manuscript.

Data Availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Conflicts of Interests

The authors declare no conflict of interest.

Funding
No funding agency is associated with this research.

Competing Interests
There are no competing interests.

Consent to Publish
All the authors gave permission to Consent to publish.

References

[1]. Karthikeyini, S., R. Sagayaraj, N. Rajkumar, and Punitha Kumaresa Pillai. "Security in Medical Image Management Using Ant Colony Optimization." Information Technology and Control 52, no. 2 (2023): 276-287.
[2]. Wang, Jingya, Xianhua Song, and Ahmed A. Abd El-Latif. "Single-objective particle swarm optimization-based chaotic image encryption scheme." Electronics 11, no. 16 (2022): 2628.
[3]. Wang, Simiao, Qiqi Peng, and Baoxiang Du. "Chaotic color image encryption based on 4D chaotic maps and DNA sequence." Optics & Laser Technology 148 (2022): 107753.
[4]. Riadh Alshibani, Dina, and Samar Amil Qassir. "Chaos-based image encoding using elementary cellular automata." In 2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT), pp. 28-33. IEEE, 2017.
[5]. Li, L., Luo, Y., Qiu, S., Ouyang, X., Cao, L. and Tang, S., 2022. Image encryption using chaotic map and cellular automata. Multimedia Tools and Applications, 81(28), pp.40755-40773.
[6]. Allawi, S.T. and Alagrash, Y.H., 2025. A New Image Encryption Method Combining the DNA Coding and 4D Chaotic Maps. International Journal of Intelligent Engineering & Systems, 18(1).
[7]. Alshibani, Dina Riadh, Narjis Mezaal Shati, and Nada Thanoon Ahmed. "DNA Genetic Recombination based Image Encryption using Chaotic Maps." Indian Journal of Public Health Research & Development 10, no. 6 (2019).
[8]. Alshibani, Dina Riadh, and Samar Amil Qassir. "Image enciphering based on DNA Exclusive-OR operation union with chaotic maps." In 2016 Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), pp. 1-6. IEEE, 2016.
[9]. Allawi, Salah Taha, and Dina Riadh Alshibani. "Color image encryption using LFSR, DNA, and 3D chaotic maps." International journal of electrical and computer engineering systems 13, no. 10 (2022): 885-893.
[10]. Narjis Mezaal Shati, Dina Riadh Alshibani, and Musaab Riyadh. "Using Whale Optimization Algorithm and Chaotic Map to Encrypt Images." Journal of Machine and Computing, Vol. 5, issue 02 (2025).
[11]. Thabit, Zainab Hasan, Sadiq A. Mehdi, and Bashar M. Nema. "Enhancing Color Image Security: Encryption with Dynamic Chaotic Three-Dimensional System and Robust Security Analysis." Al-Mustansiriyah Journal of Science 34, no. 4 (2023): 87-95.
[12]. Mehdi, Sadiq A., and Zaydon L. Ali. "Image encryption algorithm based on a novel six-dimensional hyper-chaotic system." Al-Mustansiriyah journal of science 31, no. 1 (2020): 54-63.
[13]. Kumar, N. and Saini, S., 2024, July. Image encryption model based on tent map and JAYA algorithm. In AIP Conference Proceedings (Vol. 3121, No. 1). AIP Publishing.
[14]. Kanwal, S., Inam, S., Hajjej, F., Cheikhrouhou, O., Nawaz, Z., Waqar, A. and Khan, M., 2022. A new image encryption technique based on sine map, chaotic tent map, and circulant matrices. Security and Communication Networks, 2022(1), p.4152683.
[15]. Zamli, K.Z., Din, F. and Alhadawi, H.S., 2023. Exploring a Q-learning-based chaotic naked mole rat algorithm for S-box construction and optimization. Neural Computing and Applications, 35(14), pp.10449-10471.
[16]. Gajendran, S., Muthusamy, R., Ravi, K., Chandraumakantham, O. and Marappan, S., 2024. Elliptic crypt with secured blockchain assisted federated Q-learning framework for smart healthcare. IEEE Access, 12, pp.45923-45935.
[17]. Kathamuthu, N.D., Chinnamuthu, A., Iruthayanathan, N., Ramachandran, M. and Gandomi, A.H., 2022. Deep Q-learning-based neural network with privacy preservation method for secure data transmission in internet of things (IoT) healthcare application. Electronics, 11(1), p.157.
[18]. Liu, X.D., Chen, Q.H., Zhao, R.S., Liu, G.Z., Guan, S., Wu, L.L. and Fan, X.K., 2024. Quantum image encryption algorithm based on four-dimensional chaos. Frontiers in Physics, 12, p.1230294.
[19]. Dai, L., Lei, H., Chen, L., Wang, C. and Feng, L., 2024. An Image Double Encryption Based on Improved GAN and Hyper Chaotic System. IEEE Access.
[20]. Liao, Y., Lin, Y., Li, Q., Xing, Z. and Yuan, X., 2025. Lightweight image encryption algorithm using 4d-nds: Compound dynamic diffusion and single-round efficiency. IEEE Access