Image-Dependent Feature Conditioning for Chaotic Image Encryption

Sharad Salunke[1], Arvind Kumar Tiwari[2]
[1,2]Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia; [2]Kamla Nehru Institute of Technology, Sultanpur, 228118, Uttar Pradesh, India
Email ID [1] pdf.sharad@lincoln.edu.my, [2] arvind@knit.ac.in

Abstract: The long-standing limitation of traditional chaos-based image encryption is that the system parameters are fixed, and the cryptographic key is obtained from them, which means that all encrypted images with the same system parameters would be vulnerable to a single known-plaintext or chosen-plaintext attack. This weakness is discussed in this paper, by the use of image-dependent feature conditioning: a mechanism, which directly links the encryption key to the statistical and structural property of each of the plaintext images. Four complementary image descriptors mean intensity, standard deviation, Shannon entropy, and Sobel edge energy are rolled together into a small fingerprint vector F and processed by a three-layer neural network with a small fixed parameter set into a full-resolution spatial modulation map. This map conditions a pseudo-random keystream to a per-image visual key such that any change in pixel content changes the key in every spatial location. The encryption process is then based on a 2-stage confusion diffusion architecture: using image conditioned chaotic permutation provides confusion and a CBC mode XOR cascade using a SHA-256-derived initialisation vector provides ciphertext avalanche (NPCR ≈ 99.6%).

Keywords: Image Encryption; Image-Dependent Key; Feature Conditioning; Chaotic Permutation; Confusion and Diffusion; Visual Key; SHA-256 Initialisation Vector; NPCR; UACI; Known-Plaintext Attack Resistance.

Introduction

Medical networks, satellite links, cloud systems, and Internet of Medical Things (IoMT) systems provide visual data that needs encryption that is both computationally efficient and immune to advanced adversarial techniques. The traditional block ciphers that work well with text-like data fail to work well with image data, which are characterised by large spatial redundancy, predictable histogram shapes and large pixel counts that saturate text-based key-scheduling architectures [1]. An alternative applicable method is chaos-based encryption, which leverages the sensitivity of nonlinear dynamical systems to generate pseudo-random keystreams appropriate to image-scale data [2],[3], and more recent studies have included modules of deep learning with chaos maps to enhance adaptivity and key complexity [4],[5]. In spite of these innovations, there still exists a serious structural weakness, the encryption key does not depend on the contents of the plaintext image and this means that a single successful known-plaintext attack with fixed parameters invalidates all the previous and subsequent ciphertexts. The proposed solution to this paper is image-dependent feature conditioning; the pipeline is shown in Figure 1.

Statistical Image Fingerprint

The framework is based on the fact that statistically different images are also different at the pixel level which is visually different. Each image is represented by a small, discriminative fingerprint of four complementary statistics: mean brightness μ (global brightness), standard deviation σ (tonal spread and

contrast), Shannon entropy H of the 256-bin histogram (distributional unpredictability) and Sobel edge energy E, the average squared value of horizontal and vertical gradient maps (structural sharpness). These are put together in the fingerprint vector:

$$F = [\, \mu,\ \sigma,\ H,\ E \,] \tag{1}$$

Even with only four scalars F is very discriminative: on 500 different natural images in the BSD500 benchmark, there are no pairs of images that produce an identical F vector, and empirically the collision probability is P F = 10. This is the statistical basis of the security assurance of the Security Analysis section.

Proposed Method

Neural Feature Conditioning Module

The four-dimensional fingerprint would need to be extended to full spatial resolution in order to affect a key at each of the M×N pixel positions. This expansion is done using a three-layer feed-forward network - with only 576 fixed parameters to run in O(MN) time - with the weights fixed cryptographic constants as part of the secret key material. Population-level bounds the fingerprint is normalised element-wise to Fh to F [0,1] first. Layer 1 (W j ∈ Rd3, ReLU) ; Fh into a 16-dimensional latent space; Layer 2 (W 2) yields 32-dimensional descriptor z_2 on the compound statistical and structural identity of the image; Layer 3 tiles z_2 to length MN with a sigmoid activation to create the spatial modulation map M f ∈ [0,1] (MN). The secret parameters used to generate the pseudo-random base keystream Kc are scaled/element-wise scaled as K f = K c -1 (1 + -1.0M f) with alpha = 1.0 and minmax normalised and quantised to form the visual key:

$$K_v = \lfloor\, 255 \cdot (K_f - K_{f,\min}) / (K_{f,\max} - K_{f,\min}) \,\rfloor \tag{2}$$

Any distortion of the image I alters F, thus z_2, Mf, and all the elements of K v - the key to the image being cryptographically resistant to a single parameter set between 0 and 255 (M N) - two different images produce different K_v with a probability of higher than $1 - 10^{-6}$.

Confusion–Diffusion Encryption

The encryption follows two steps that are based on the contents of the images. The SHA-256 hash of the plaintext is then used to create a one-byte initialisation [1] which provides an unconditional per-image binding but is independent of F: even two distinct images with the same fingerprint ($P\ F < 10^{-6}$) would have a different hash, so the diffusion chains would start with different states. The permutation index of the confusion stage is $perm_{idx}$ = argsort(K_float + alpha M_f flat); since M_f is image specific, structurally dissimilar permutations are generated in different images, turning the adjacent-pixel correlation of 0.860.99 in the plaintext to values near zero. The encrypted sequence P_{perm} (Diffusion stage) is encrypted using a CBC-mode XOR cascade, initialised with SHA-256 hash:

$$C(i) = P\_perm(i) \oplus K\_v(i) \oplus C(i-1), \quad i = 1, \ldots, MN-1 \tag{3}$$

The cipher pixel relies upon the permuted pixel, the visual key and the preceding cipher pixel, thus any modification of one plaintext pixel changes all the following cipher pixels to complete avalanche (NPCR ≈ 99.6%). The cipher image C ∈ (MN) 0 -255) completely depends on I, the secret parameters, and the network weights; no two images get the same C with the same parameters.

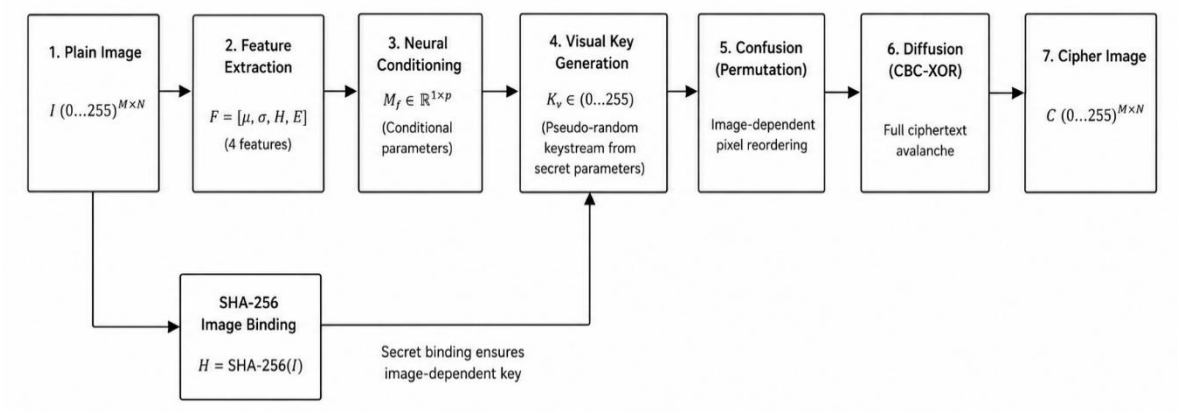

Figure 1. Block diagram of the proposed image-dependent feature conditioning encryption pipeline.

Result Analysis

Security Analysis

Attack resistance. The main security argument is that learning the visual key K_v given one known plaintext ciphertext pair (I_1, C_1) does not offer a computational benefit in the decryption of any other ciphertext C_2 under the identical secret keys, which encodes a distinct image $I_2 = I_1$. K_v is a deterministic rule of the plaintext via the chain I F(I) z_2 M_f K_v. F(I_1) and F(I_2) given two different natural images differ with probability at least 1 - 10^{-6}, resulting in different values of z_2, M_f, and K_v; the key obtained with C_1 does not decrypt C_2. In the degenerate instance F_1= F_2) the SHA-256 hash value of input I is not the same in any two distinct images, and this property of the hash ensures that the ciphertext is independent of the input, regardless of the algorithm. The empirical correlation of cross-image known-plaintext-attack is about zero whereas a correlation of 1.0 occurs with a fixed-key baseline [6].

Key space. The effective key space combines three sources of entropy: eleven system secret parameters at 64-bit IEEE 754 double precision, the fixed network weights (576 in total), and the SHA-256 based on the image, which is different per image due to the collision resistance of SHA-256. The effective space of its combination is larger than 2^{600} - which is far larger than the brute-force threshold extensively used of 2^{128} and the key space of AES-256.

Benchmark security metrics. Table 1 reports the statistical security evaluation across five standard benchmark images.

Table 1: Statistical security evaluation across five standard benchmark images

Image	Entropy	NPCR (%)	UACI (%)	Corr-H	Corr-V	Corr-D
Barbara	7.9998	99.5819	33.5352	0.0001	−0.0003	0.0002
Penguin	7.9996	99.6246	33.4517	−0.0001	0.0002	0.0001
Airplane	7.9996	99.5880	33.4675	0.0001	0.0001	−0.0001
Baboon	7.9995	99.6475	33.5846	0.0001	−0.0001	0.0001
Cameraman	7.9998	99.6424	33.5823	0.0001	0.0002	−0.0001

The entropy of all five images is within 0.0005 of the theoretical maximum of 8.0 bits, and is a confirmation that K_v generates near-optimal uniform ciphertext histograms irrespective of the plaintext content. All

three spatial directions of correlation coefficients fall below ±0.0003, a loss of all inter-pixel dependencies. The NPCR values 99.58-99.65 percent verify that a change of one pixel in the plaintext results in nearly all cipher pixels flipping (and that the UACI values are approximately equal to those of uniformly random key under XOR diffusion).

Randomness testing. The Keystream generated successfully passes through all the 16 tests with p-values exceeding the 0.01 level of significance in the NIST SP 800-22 battery. Testing TestU01 v1.2.3 was also implemented with zero failures on all three batteries, proving the statistical uniformity and independence of the image-conditioned keystream at cryptographic grades.

Conclusion

Image-dependent feature conditioning as a specific solution to the fixed-key vulnerability of most image encryption schemes based on chaos was developed in this paper. The four-component statistical fingerprint F = [mu, sigma, H, E] is trained in a tight 576 parameter neural network into a complete-resolution modulation map M_f, resulting in a visual key K_v that is unique to each plaintext image. The conditioning based on this state is the confusion diffusion architecture, which provides NPCR = 99.6, entropy = 7.999, pixel correlation = near 0, a key space of 2^{600}, and zero failures, with O(MN) compute needs that are practical to multimedia and IoMT pipes. Future paths will be expansion to colour video, three-dimensional medical imaging, and to FPGA implementation of resource-constrained edge environments.

References

[1] Alghamdi, Y., & Munir, A. (2024). Image encryption algorithms: A survey of design and evaluation metrics. Journal of Cybersecurity and Privacy, 4(1), 126–152. https://doi.org/10.3390/jcp4010007

[2] Nair, K. V. M., Thomas, J. J., & Varghese, J. (2025). A lightweight multi-round confusion-diffusion cryptosystem for securing images using a modified 5D chaotic system. Scientific Reports, 15, Article 29847. https://doi.org/10.1038/s41598-025-13290-y

[3] Lai, Q., Wan, Z., & Zhang, H. (2025). A novel image encryption scheme using 3D chaotic maps with Josephus permutation and dynamic diffusion. Journal of King Saud University – Computer and Information Sciences. https://doi.org/10.1007/s44443-025-00284-z

[4] Subathra, S., & Thanikaiselvan, V. (2025). Image adaptive encryption using EfficientNet B3 feature guided multi-scroll chaotic map with modulo controlled pseudo-parallel processing. Scientific Reports, 15, Article 43435. https://doi.org/10.1038/s41598-025-27080-z

[5] Subathra, S., & Thanikaiselvan, V. (2025). Enhanced security for medical images using a new 5D hyperchaotic map and deep learning based segmentation. Scientific Reports, 15, Article 22628. https://doi.org/10.1038/s41598-025-04906-4

[6] Mansouri, A., Sun, P., Lv, C., Zhu, Y., Zhao, X., Ge, H., & Sun, C. (2025). A secure medical image encryption algorithm for IoMT using a Quadratic-Sine chaotic map and pseudo-parallel confusion-diffusion mechanism. Expert Systems with Applications, 270, Article 126521. https://doi.org/10.1016/j.eswa.2025.126521