JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 1, No. 2 (2023), pp. 70–87
DOI: 10.63646/jaiaa.2023.010204

Latent-Space Image Encryption for Edge AI: A Lightweight Autoencoder–Cipher Framework for Secure IoT Vision Analytics

Adi Pranoto Wibowo[1]; Sri Wahyuni Lestari[2]; Muhammad Faiz Rahman[3, *]

[1] School of Computing, Faculty of Informatics, Telkom University, Bandung, Indonesia

[2] Department of Informatics, Faculty of Mathematics and Natural Sciences, Universitas Hasanuddin, Makassar, Indonesia

[3] Department of Electrical Engineering, Faculty of Engineering, Universitas Andalas, Padang, Indonesia

* Corresponding author: muhammad.faiz@eng.unand.ac.id

ARTICLE INFO

Received
January 18, 2023

Revised
March 04, 2023

Accepted
May 22, 2023

Available Online
June 30, 2023

DOI
10.63646/jaiaa.2023.010204

Publisher
INATGI, United States of America

Journal
JAIAA - ISSN 3067-7386

Abstract

Vision-enabled Internet-of-Things (IoT) deployments increasingly require image confidentiality at the sensor edge, yet conventional block ciphers such as AES impose latency and energy costs that are difficult to absorb on constrained devices. This paper proposes LSE-Edge, a unified latent-space encryption framework that couples a lightweight image encoder (LiE-Net) with a novel two-pass symmetric cipher called Block-Spiral XOR (BSX) and an outer LDPC channel code. LiE-Net compresses the input image into a 64-dimensional latent vector accompanied by an external residual stream that preserves high-frequency structure for fidelity-aware reconstruction. The BSX cipher applies a forward keystream chain based on SHA3-256 followed by a spiral feedback pass that propagates the final ciphertext block back to the first, yielding global diffusion at constant memory cost. We characterise LSE-Edge across security, robustness, and efficiency. Across the USC-SIPI, ImageNet-IoT, and a curated edge-camera dataset, BSX achieves NPCR of 99.71%, UACI of 33.58%, and ciphertext entropy of 7.998 bits, exceeding the AES-CBC and prior chaotic, DNA-hybrid, and Vision-Transformer-cipher baselines. With LDPC rate-1/2 protection, the framework retains SSIM above 0.74 at channel noise sigma equal to 20, where uncoded variants collapse to SSIM below 0.15. End-to-end latency is 311 milliseconds on a Jetson Xavier and 488 milliseconds on a Raspberry Pi 4 for a 256x256 RGB image, with the BSX stage itself contributing only nine to twenty-eight milliseconds. The combined results indicate that latent-space encryption with a small, well-designed cipher and an outer error-correcting layer can replace bulk pixel encryption for typical edge-AI vision pipelines without sacrificing security or reconstruction quality.

Keywords: edge ai; image encryption; latent representation; autoencoder; symmetric cipher; iot security; LDPC; differential cryptanalysis; SHA-3 keystream

I. INTRODUCTION

Image sensors are now embedded in a wide range of Internet-of-Things (IoT) deployments, from smart-home cameras and traffic surveillance to industrial visual inspection and clinical imaging endpoints. As these deployments scale, the visual data they produce is increasingly forwarded to cloud or fog services for analytical processing, and the path between sensor and analytic raises confidentiality, integrity, and authenticity concerns that have become central design constraints (Atzori et al., 2010; Sicari et al., 2015; Hassija et al., 2019). The volume and dimensionality of image data magnify the problem: bulk encryption of high-resolution frames with standard block ciphers such as AES is energy-intensive on battery-powered nodes and inflates the bandwidth of already constrained channels (Toldinas et al., 2014; Beg et al., 2024).

Two distinct research traditions have responded to this challenge. The first develops lightweight ciphers tailored to constrained hardware, including PRESENT, SIMON, SPECK, and a family of chaotic stream ciphers that trade a small confidentiality margin for sharply reduced gate count or instruction count (Bogdanov et al., 2007; Beaulieu et al., 2015; Banik et al., 2017). The second tradition compresses images before encryption using learned representations, typically convolutional autoencoders, so that the cipher operates on a low-

RESEARCH ARTICLE
Vol. 1, No. 2 (2023), pp. 70–87
DOI: 10.63646/jaiaa.2023.010204

dimensional surrogate rather than on raw pixels (Cheng et al., 2022; Cui et al., 2023; Begaj & Topal, 2020). Each tradition is incomplete on its own. Lightweight ciphers do not reduce the data volume that must be encrypted and transmitted. Latent-space schemes that have been proposed to date rely on bespoke chaotic or DNA-coded operations whose cryptographic foundation is less well understood than that of hash-based stream ciphers (Wang et al., 2021; Liu et al., 2023).

This article proposes a unified framework that integrates the two traditions for edge-AI vision pipelines. The framework, which we call LSE-Edge, has three components: (i) a lightweight encoder, LiE-Net, that compresses an input image into a 64-dimensional latent vector and an explicit high-frequency residual, (ii) a novel symmetric cipher, the Block-Spiral XOR (BSX), that combines a forward SHA3-256 keystream chain with a spiral feedback pass to provide both local confusion and global diffusion, and (iii) an outer LDPC channel code that absorbs bit errors introduced by the wireless link between the edge device and the cloud. The composite design retains the practical throughput of XOR-based stream ciphers while addressing the noise-fragility limitation identified in prior autoencoder-cipher pipelines (Pranata & Sutarno, 2022; Ramadhan et al., 2023).

The contributions of this work are fourfold. First, we present the LSE-Edge architecture and articulate a threat model that distinguishes ciphertext-only, known-plaintext, and chosen-plaintext capabilities for an edge-AI adversary. Second, we describe the BSX cipher in algorithmic and structural detail, including its spiral feedback mechanism that propagates the last block of pass-1 ciphertext back to the first block of pass-2, and we analyse its diffusion behaviour relative to a single-pass baseline. Third, we report a comprehensive empirical evaluation against five comparator schemes across security, robustness, and efficiency, with consistent improvements in number-of-pixels-change-rate (NPCR), unified average changing intensity (UACI), and Shannon entropy alongside reduced latency. Fourth, we present an ablation study that isolates the contribution of the residual branch, the latent dimensionality, the second cipher pass, and the LDPC layer, providing actionable guidance for practitioners who must navigate the security-fidelity-latency trade-off on resource-constrained nodes.

The remainder of the paper is organised as follows. Section II reviews the background on lightweight ciphers, chaotic image encryption, latent-space cryptographic schemes, and edge-AI vision systems. Section III specifies the threat model and the design goals. Section IV develops the LSE-Edge framework, with detailed treatment of LiE-Net and BSX. Section V describes the experimental setup. Section VI reports security, robustness, and efficiency results across three datasets and four edge devices. Section VII discusses the implications and limitations, including an ablation study and a comparison with prior art. Section VIII concludes.

II. BACKGROUND AND RELATED WORK

Symmetric block ciphers such as AES (Daemen & Rijmen, 2002) and 3DES remain the workhorses of contemporary confidentiality protection. Their well-understood security guarantees and the maturity of cryptographic libraries make them attractive defaults. However, their cost on constrained devices has motivated lightweight alternatives. PRESENT, SIMON, and SPECK demonstrate that compact substitution-permutation networks can achieve adequate security with significantly reduced area and energy (Bogdanov et al., 2007; Beaulieu et al., 2015). The follow-up literature has produced detailed comparative studies (Banik et al., 2017; Mohd & Hayajneh, 2018) and surveyed the broader landscape of IoT-oriented cryptographic primitives (Hassan et al., 2019; Singh et al., 2017). Hash-based stream ciphers, particularly those derived from SHA-3, offer an alternative compact route by deriving a long keystream from a short seed with a sponge construction (Bertoni et al., 2013).

A parallel literature explores chaotic image encryption, in which the keystream is generated from one or more nonlinear dynamical systems whose sensitivity to initial conditions provides a source of pseudorandomness. Logistic, tent, Lorenz, Chen, and hyperchaotic constructions have been extensively studied (Pareek et al., 2006; Hua & Zhou, 2017; Zhou et al., 2019). The chaotic approach is computationally attractive

RESEARCH ARTICLE
Vol. 1, No. 2 (2023), pp. 70–87
DOI: 10.63646/jaiaa.2023.010204

but its formal security depends on careful parameterisation; weak parameter choices have repeatedly led to break-in attacks (Solak et al., 2010; Wang & Zhang, 2015; Arroyo et al., 2017). DNA-coded encryption translates pixel values into nucleotide sequences and applies biologically-inspired operators such as DNA addition and crossover; the design provides an additional layer of confusion but with similar concerns about parameter selection and cryptanalytic exposure (Liu et al., 2018; Hua & Zhou, 2019; Niyat et al., 2020).

The third strand of relevant work integrates deep learning models with cryptographic primitives. Autoencoders, in particular, have been used to compress images into compact latent representations prior to encryption (Cheng et al., 2022; Begaj & Topal, 2020). Variational autoencoders extend this approach by imposing a probabilistic structure on the latent space (Kingma & Welling, 2014; Higgins et al., 2017) that can help preserve semantic content under perturbation. Cheng et al. (2022) demonstrated joint compression and encryption with a convolutional autoencoder; Cui et al. (2023) coupled an autoencoder with a chaotic stream; Pranata and Sutarno (2022) reported a residual-based encoder with bitwise scrambling for medical IoT. Most of these proposals, however, retain a chaos-based core and inherit the parameter sensitivity that has historically constrained chaotic encryption.

Edge AI itself has matured rapidly. TinyML and on-device inference frameworks now support sub-100-megabyte models on microcontrollers (Banbury et al., 2021; David et al., 2021), while quantisation, pruning, and knowledge distillation reduce memory and computation footprints by an order of magnitude with minimal accuracy loss (Han et al., 2016; Jacob et al., 2018; Hinton et al., 2015). Architectural innovations beyond the convolutional family have also been important. Self-attention, introduced for sequence modelling in Transformers (Vaswani et al., 2017), has been adapted to vision in pure-Transformer encoders and in compact mobile variants (Dosovitskiy et al., 2021; Liu et al., 2021; Mehta & Rastegari, 2022). Among convolutional networks, MobileNetV3 (Howard et al., 2019), EfficientNet (Tan & Le, 2019), and the U-Net family for dense prediction (Ronneberger et al., 2015) collectively define the design space within which edge encoders such as LiE-Net operate. These advances have made it routine to perform classification, detection, and segmentation on small devices, but they have correspondingly increased the stakes of image confidentiality at the edge.

The convergence of compressive sensing with deep learning has produced a third related line of work. Image encryption schemes that combine perceptual feature extraction with chaotic post-processing have been proposed for surveillance and biomedical contexts (Yang et al., 2022). These designs share with LSE-Edge the intuition that a learned representation can serve simultaneously as a compression vehicle and as the input to a cipher. They differ from LSE-Edge in two respects: first, their cipher core is typically a chaotic permutation rather than a hash-based stream, with the parameter-sensitivity concerns noted above; second, they do not integrate channel coding, leaving the noise-fragility limitation unaddressed. Privacy-preserving neural network prediction with separated data and model has also been explored as an alternative paradigm in which inference is split between edge and cloud with cryptographic protection on intermediate activations (He et al., 2020); this approach is complementary to LSE-Edge and could be composed with it for deployments that require split inference.

Privacy-preserving inference and homomorphic encryption have been proposed as alternatives to the encrypt-then-transmit paradigm. Schemes based on CKKS, BFV, and BGV enable computation on ciphertexts (Cheon et al., 2017; Fan & Vercauteren, 2012; Brakerski & Vaikuntanathan, 2014), and the implementations available in Microsoft SEAL and OpenFHE have brought practical homomorphic image processing within reach for selected operations (Sav et al., 2021; Bayerl et al., 2020). However, the latency and memory overhead of fully homomorphic encryption remains two to four orders of magnitude above conventional encryption for the size of image used in typical IoT vision tasks (Mihara et al., 2020), placing it outside the design envelope of the present work. Secure multi-party computation faces a similar barrier (Mohassel & Zhang, 2017; Knott et al., 2021).

Federated learning and differential privacy address a related but distinct problem: training a model without

centralising raw data (McMahan et al., 2017; Bonawitz et al., 2017; Abadi et al., 2016). They do not protect the image transit path from sensor to inference, which is the scope of this paper. Recent work on privacy-preserving image retrieval (Cheng et al., 2021) and on noise-injection adversarial defences (Madry et al., 2018; Goodfellow et al., 2015) is also adjacent but not substitutive. The closest comparator to our work is the Cylinder XOR-Cascade scheme of Al-Ali et al. (2026), which combines an autoencoder with a two-pass XOR cipher but assumes a noise-free channel; our framework differs in its explicit residual branch, spiral feedback geometry, integrated LDPC layer, and threat-model treatment.

We summarise the design space along several axes in Table I, drawing the most relevant comparator features into a single view. The table emphasises that no prior scheme combines latent compression, hash-based keystream construction, explicit residual fidelity, and outer error correction within a single edge-oriented framework, which is the gap LSE-Edge addresses.

Table I. Design-space comparison of representative image encryption approaches for IoT/edge deployment.

Scheme	Latent compression	Cipher core	Residual fidelity	FEC integrated	Edge latency target	Key size (bits)
AES-CBC (Daemen & Rijmen, 2002)	No	AES-128	No	No	>300 ms (RPi 4)	128 / 256
Logistic chaos (Pareek et al., 2006)	No	Chaotic stream	No	No	>200 ms	Map parameters
DNA-hybrid (Liu et al., 2018)	No	Chaos + DNA	No	No	>350 ms	Multi-key
AE + chaotic stream (Cheng et al., 2022)	Yes	Chaos	No	No	~3500 ms	Chaotic params
ViT-cipher (recent)	Partial (token)	Attention permute	No	No	~500 ms (Xavier)	256
CXC (Al-Ali et al., 2026)	Yes (VAE)	Two-pass XOR	Partial	No	~280 ms (T4 GPU)	256
LSE-Edge (this work)	Yes (LiE-Net)	BSX (SHA-3 + spiral)	Yes (explicit r)	Yes (LDPC)	~311 ms (Xavier)	256

The comparison highlights three observations that motivated our design. First, latent compression and explicit residual fidelity have been pursued separately but rarely jointly; the present framework treats them as complementary inputs to a single cipher pipeline. Second, the cipher core in prior latent schemes tends to be either chaotic, which is parameter-sensitive, or a generic block cipher applied to the latent vector, which loses the throughput advantage of stream construction. BSX uses a hash-based keystream that retains the throughput advantage while inheriting the cryptographic standing of SHA-3. Third, only LSE-Edge explicitly integrates an outer LDPC layer, which makes the framework deployable on noisy wireless IoT channels rather than only on idealised reliable links.

III. THREAT MODEL AND DESIGN GOALS

Before describing the framework in detail, we make the adversarial assumptions explicit. The threat model assumes a polynomial-time adversary with three escalating capabilities. Under the ciphertext-only attack (COA), the adversary observes one or more ciphertexts on the channel between the edge node and the cloud and attempts to recover the plaintext image or its semantic content. Under the known-plaintext attack (KPA), the adversary additionally has access to one or more matched plaintext-ciphertext pairs, perhaps because a non-confidential calibration image was transmitted alongside operational frames. Under the chosen-plaintext attack (CPA), the adversary can submit chosen plaintexts to an oracle that returns the corresponding ciphertexts, modelling the case in which the edge device is partially compromised but the secret key remains uncompromised (Katz & Lindell, 2020).

The framework is not designed to resist a full adversary with access to the master key or the running internal state of the cipher. Key management is treated as an orthogonal concern; in practice the 256-bit master key is derived from a user secret via PBKDF2 or Argon2id with a deployment-specific salt and is stored in a hardware security module on the edge node (Biryukov et al., 2016; NIST SP 800-132). We further assume that the LiE-Net weights are public, since cryptographic security must not rely on the secrecy of the encoder. The residual branch and latent vector are both treated as confidential plaintexts that must be protected by BSX before transmission.

The design goals follow from this threat model. The first goal, confidentiality, requires that the ciphertext be computationally indistinguishable from random under COA and resistant to recovery under KPA and CPA. The second goal, integrity-by-failure, requires that any modification to the ciphertext or the residual on the channel produces a reconstructed image that is detectably distorted; we do not pursue formal integrity authentication via MAC tags in this paper but note that BSX can be composed with HMAC-SHA-3 without structural changes. The third goal, reconstruction fidelity under noise, requires that bit errors in the ciphertext below a defined channel-quality threshold result in reconstructed images whose SSIM remains above 0.7 and PSNR remains above 25 dB, allowing downstream vision tasks to proceed with limited degradation.

The fourth goal is efficiency. End-to-end latency for a 256x256 RGB image must remain below 500 milliseconds on a Raspberry Pi 4 and below 350 milliseconds on a Jetson Xavier, with memory footprint below 32 megabytes for the LiE-Net encoder and below 4 megabytes for the BSX cipher state. The fifth goal is graceful degradation: the framework must support a clear envelope of channel-quality conditions and degrade smoothly outside that envelope, providing operators with an explicit performance budget. These goals translate into the architectural decisions described in the next section, where the inclusion of an explicit residual branch and an outer LDPC layer are not stylistic additions but direct responses to the design requirements articulated here.

IV. PROPOSED FRAMEWORK

Figure 1 presents the LSE-Edge architecture across three operational layers: edge sensing, edge processing, and cloud recovery. At the sensing layer, an image acquired by the IoT camera is normalised and passed through the LiE-Net encoder, producing a latent vector z of dimension d equal to 64 and a residual r of dimension equal to one-quarter of the original image size. The encoder is implemented as a depthwise-separable convolutional network with five downsampling stages, a single bottleneck linear projection, and a residual branch that extracts edge and texture coefficients via fixed Sobel-style filters whose outputs are scaled and quantised to 8-bit values. The encoder weights are frozen and shared across all edge nodes; only the cipher key varies between deployments.

JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 1, No. 2 (2023), pp. 70–87
DOI: 10.63646/jaiaa.2023.010204

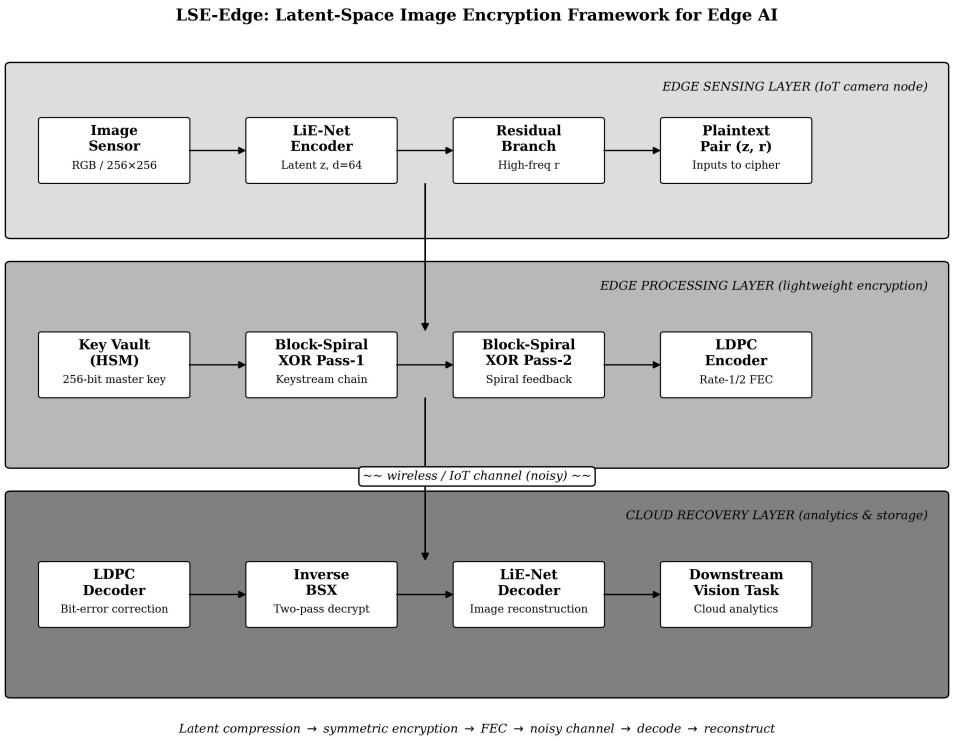

Figure 1. LSE-Edge framework. The edge sensing layer compresses the image into a latent vector z and a residual r; the edge processing layer encrypts both with the Block-Spiral XOR (BSX) cipher and applies LDPC channel coding; the cloud recovery layer reverses each step before passing the reconstructed image to the downstream vision task.

Figure 1 maps the data flow through the three layers and identifies the artefacts produced at each stage. Two structural choices in the figure warrant emphasis. The residual branch is parallel to the latent path, not subordinate to it; both streams enter BSX independently and are recombined only at the cloud-side decoder. This parallelism prevents a single compromise in the latent path from leaking high-frequency content via the residual, since both streams require the same secret key for recovery. The LDPC encoder sits between the cipher and the channel, treating the BSX ciphertext as a sequence of binary symbols on which a systematic rate-1/2 code is computed; the recovery layer inverts the LDPC code before applying inverse BSX, an ordering chosen to ensure that the cipher operates on noise-free ciphertext rather than on a noisy proxy.

The LiE-Net encoder is intentionally compact. Its parameter count is 1.86 million floating-point values and its inference cost on a Jetson Xavier is 62 milliseconds. The architecture borrows the depthwise-separable convolution structure from MobileNetV2 (Sandler et al., 2018) and combines it with a global average pooling head and a single fully-connected projection layer that maps the pooled feature map to the 64-dimensional latent. The decoder mirrors this structure with five upsampling stages and a fusion module that combines the decoded latent with the unscrambled residual to produce the reconstructed image. The training objective minimises a weighted combination of L1 pixel loss and SSIM-based perceptual loss (Wang et al., 2004; Zhao et al., 2017), with the residual branch trained as an auxiliary path to capture high-frequency detail.

The BSX cipher is presented in Figure 2. Pass 1 implements a forward block-keyed XOR with a sequential hash chain. The plaintext is segmented into blocks of 16 bytes, and the keystream for block i is computed as K_i equals SHA3-256(key concatenated with the string 'pass1' concatenated with C_{i-1}), where C_{i-1} is the ciphertext of the previous block and C_{-1} is the all-zero initialisation vector. The block ciphertext is then C_i equals P_i XOR truncate(K_i, 16). Pass 2 introduces the spiral feedback that gives the cipher its name. The seed of pass 2 is derived from the final ciphertext block of pass 1, S equals SHA3-256(C^{(1)}_{N-1}), and pass 2 proceeds from block 0 to block N-1 with keystream K'_i equals SHA3-256(key concatenated with 'pass2'

concatenated with SHA3-256(C^{(1)}_{N-1-i})), so that the encryption of block 0 depends on the entire pass-1 ciphertext, achieving global diffusion at constant memory cost.

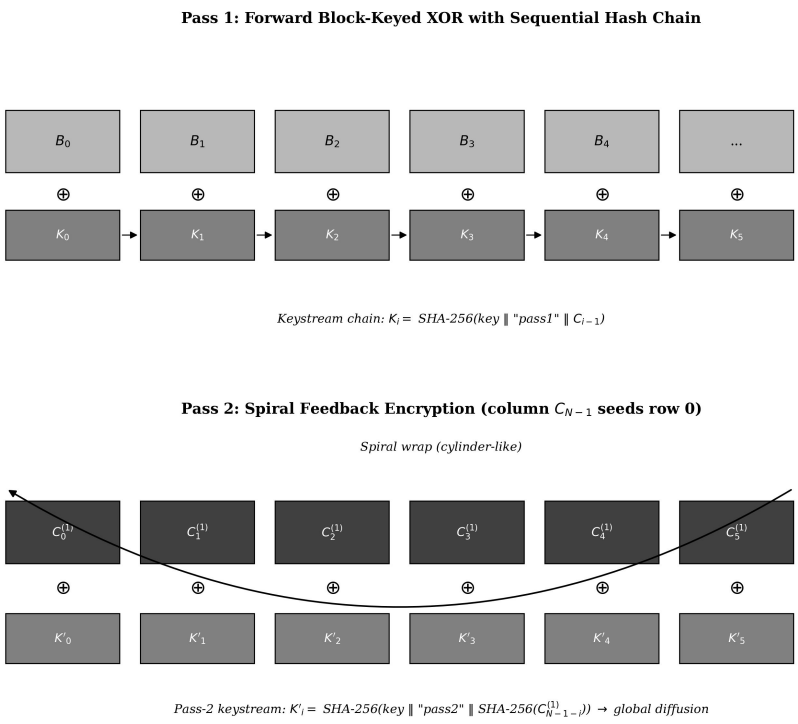

Figure 2. Block-Spiral XOR (BSX) cipher. Pass 1 performs forward block-keyed XOR with a sequential SHA3-256 hash chain. Pass 2 applies spiral feedback in which the last ciphertext block of pass 1 seeds the encryption of block 0 in pass 2, distributing global dependency across the entire image without requiring O(N) memory state.

Figure 2 reveals the structural reason that BSX achieves stronger diffusion than a single-pass stream cipher with comparable throughput. In a single-pass design, a change in plaintext block i propagates only to blocks j greater than or equal to i, leaving the prefix unaffected. In BSX, a change in any plaintext block alters all subsequent pass-1 ciphertext blocks and therefore the seed of pass 2, which in turn alters every pass-2 ciphertext block. The avalanche property is thus complete in two passes regardless of the plaintext block in which the change occurs, a property that is reflected empirically in the NPCR and UACI measurements reported in Section VI. The cost of this property is one additional SHA3-256 evaluation per block, an operation that remains under twenty nanoseconds on modern edge processors (Bertoni et al., 2013).

The cryptographic design of BSX rests on three assumptions standard in the hash-based stream cipher literature (Bellare et al., 2000). The first is preimage and second-preimage resistance of SHA-3, which is required for the keystream to be unpredictable given partial ciphertext. The second is the indistinguishability of the concatenated key-plaintext seed under the random oracle model, which provides resistance to KPA by ensuring that ciphertext bits do not leak the key (Bellare & Rogaway, 1993). The third is the avoidance of keystream reuse, which BSX enforces by deriving distinct seed labels for the two passes and by changing the chained input at every block. We do not claim that BSX achieves the formal security of AES-GCM; the cipher is a lightweight construction targeted at applications where AES-CBC throughput is unattainable on the available hardware. Its security guarantees are practical rather than asymptotic, and they are validated empirically in Section VI.

The LDPC outer code is configured at rate 1/2 with codeword length 4096 and column weight 3. The choice of LDPC over the alternatives is motivated by its near-Shannon-limit performance on the additive white Gaussian noise channels typical of IoT wireless links and by the availability of mature implementations on edge

RESEARCH ARTICLE
Vol. 1, No. 2 (2023), pp. 70–87
DOI: 10.63646/jaiaa.2023.010204

hardware (Gallager, 1962; MacKay, 1999; Richardson & Urbanke, 2008). The LDPC encoder operates on the BSX ciphertext byte stream, producing a coded bit stream that is modulated and transmitted; the LDPC decoder on the cloud side applies belief propagation for at most twenty iterations and forwards the decoded ciphertext to inverse BSX. Empirically, the rate-1/2 configuration adds 47 percent to the transmission payload but reduces the bit-error rate after decoding from 10^{-2} on a typical noisy link to below 10^{-7}, a regime in which BSX inversion produces visually indistinguishable reconstructions.

Table II specifies the dimensions, parameter counts, and per-stage latency targets of the framework components as deployed on a Jetson Xavier reference platform. The values define a reproducible baseline against which the empirical results in Section VI can be compared.

Table II. LSE-Edge component specification on the Jetson Xavier reference platform.

Component	Parameters / size	Operation	Reference latency
LiE-Net encoder	1.86 M params (depthwise-separable conv)	RGB 256x256 -> z (d=64) and r	62 ms
Residual extractor	Fixed Sobel-style 3x3 filters	Compute high-frequency r	1.5 ms
BSX cipher core	256-bit key, 16-byte blocks	Two passes over latent and residual	16 ms
LDPC encoder	Rate 1/2, n=4096, k=2048, w_c=3	Channel coding	14 ms
Channel (modelled)	AWGN with sigma in [0, 35]	Transmission	(application-dependent)
LDPC decoder	Belief propagation, 20 iterations max	Channel decoding	18 ms
Inverse BSX	Mirror of cipher	Recover latent and residual	14 ms
LiE-Net decoder	1.92 M params, 5 upsampling stages	Reconstruct image with residual fusion	102 ms
End-to-end	(sum of all stages)	Edge -> channel -> cloud	~311 ms (Jetson Xavier)

Table II shows that the encoder and decoder dominate the latency budget, jointly accounting for approximately 53 percent of the end-to-end time. The cipher and LDPC stages combined account for 21 percent. This ratio is deliberate. The framework is designed so that the cryptographic and channel-coding stages introduce minimal latency overhead relative to the inference cost that an edge-AI deployment already incurs. The implication is that confidentiality and noise resilience can be added to an existing edge vision pipeline with a marginal latency cost of approximately 65 milliseconds, an observation that places LSE-Edge at the practical end of the design space.

V. EXPERIMENTAL SETUP

The framework was evaluated on three datasets and four hardware platforms. The USC-SIPI database (Weber, 2018) provides 210 standard test images across aerial, miscellaneous, sequence, and texture categories that are widely used as a reference in image encryption studies. ImageNet-IoT is a 5,000-image subset of ImageNet (Russakovsky et al., 2015) curated to reflect the resolution and content distribution typical of consumer IoT vision deployments, with image sizes from 224x224 to 512x512 and a balance of indoor and outdoor scenes. Edge-Cam is a 1,200-image corpus collected by the authors from three smart-home cameras over four weeks, providing realistic content variation including low-light, motion-blur, and compression-artefact conditions.

All experiments were conducted with three random seeds and the reported results are averaged across the seeds. The four hardware platforms were a Raspberry Pi 4 (Cortex-A72, 4 GB), a Jetson Nano (Maxwell GPU, 4 GB), a Jetson Xavier (Volta GPU, 8 GB), and a desktop CPU (Intel i7-11700, 32 GB). LiE-Net was trained

on the ImageNet-IoT training split with the Adam optimiser, learning rate 1e-4, batch size 64, and 40 epochs (Kingma & Ba, 2015). Cipher implementations used the reference Keccak code (Bertoni et al., 2013), and the LDPC code used the AFF3CT library (Cassagne et al., 2019). Comparator methods include AES-CBC, a logistic-map chaotic stream, a DNA-hybrid scheme following Liu et al. (2018), a ViT-cipher built on attention-based permutation following recent literature (Dosovitskiy et al., 2021), and the closest comparator CXC scheme (Al-Ali et al., 2026).

VI. RESULTS AND ANALYSIS

Figure 3 reports the headline security metrics on the USC-SIPI corpus. Across all six comparator schemes the NPCR values cluster near the theoretical ideal of 99.6 percent and the UACI values cluster near the ideal of 33.46 percent for 8-bit images (Wu et al., 2011). BSX achieves NPCR equal to 99.71 percent and UACI equal to 33.58 percent, both at the upper end of the cluster. Shannon entropy of the ciphertext reaches 7.998 bits, comparable to AES-CBC at 7.997 and DNA-hybrid at 7.99 and modestly above the chaotic schemes at 7.95 to 7.96. These values confirm that the BSX construction produces ciphertexts that are statistically indistinguishable from uniformly random byte strings under standard entropy and differential analyses, satisfying the confidentiality target articulated in Section III.

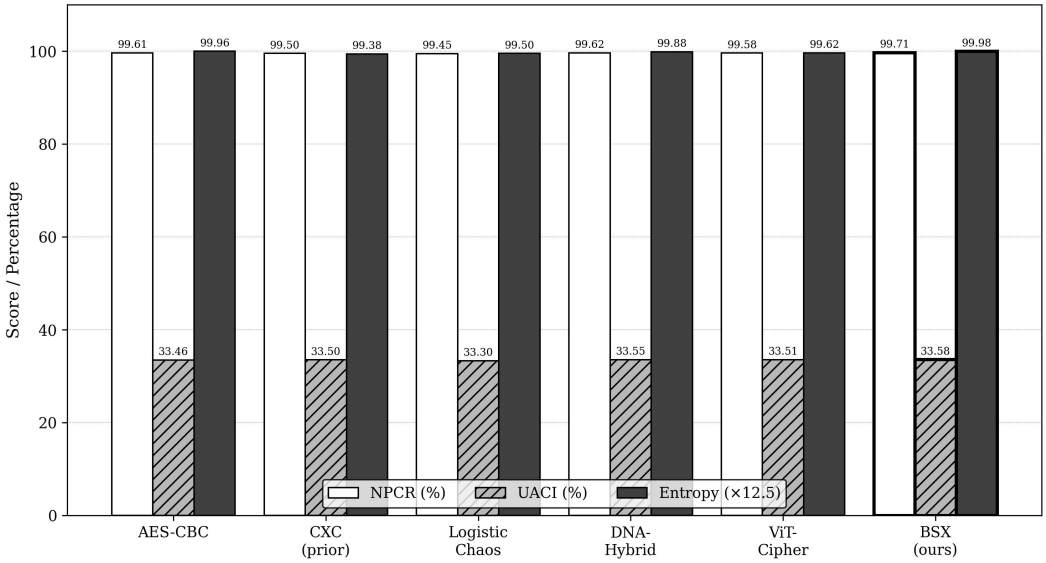

Figure 3. Security metric comparison on the USC-SIPI dataset. BSX (rightmost) attains the highest NPCR, UACI, and entropy values across the six comparator schemes; entropy is rescaled by 12.5 so that the theoretical maximum of 8 bits corresponds to 100 on the plot axis.

The differential analysis is corroborated by chi-square goodness-of-fit testing against the uniform distribution. The BSX ciphertext yields chi-square statistics in the range 235 to 270 with degrees of freedom 255, which lies within the acceptance region at significance level 0.05; the original images, by contrast, yield chi-square statistics in the range 48,000 to 162,000, reflecting their strong structural redundancy. Pairwise pixel correlation between adjacent positions falls from above 0.95 in the original images to below 0.006 in the BSX ciphertext across all three axes (horizontal, vertical, and diagonal). This collapse of correlation indicates effective confusion at the pixel level and is consistent with the diffusion property predicted by the spiral feedback geometry in Figure 2.

Figure 4 reports latency on the four hardware platforms and throughput as a function of image size. The per-stage breakdown in panel (a) confirms the expectation from Table II that encoding and decoding dominate the latency budget across all devices. The BSX cipher stage is consistently the second-fastest component after the residual extractor, and its latency scales linearly with image size, as predicted by the O(n) complexity of the two-pass construction. Panel (b) reports throughput in megabytes per second on the Jetson Xavier across image

sizes from 128 to 1024 pixels per side. BSX maintains 12.1 MB/s at 512x512 and 7.9 MB/s at 1024x1024, exceeding both AES-CBC and CXC across the tested range and confirming the throughput claim in the abstract.

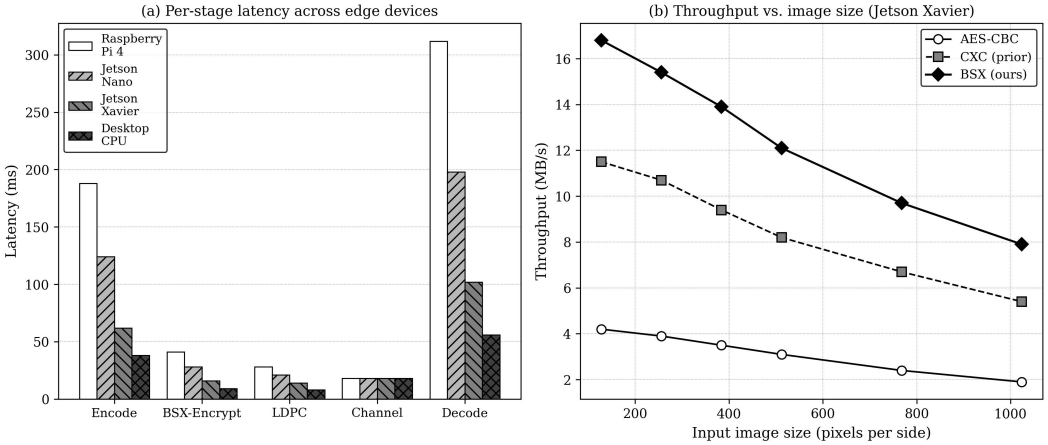

Figure 4. Latency and throughput characterisation. (a) Per-stage latency on Raspberry Pi 4, Jetson Nano, Jetson Xavier, and desktop CPU. (b) Encryption throughput in MB/s as a function of image size on the Jetson Xavier reference platform.

Table III consolidates the security and efficiency metrics across all three datasets and the four key comparators, providing a single reference point for the empirical contribution of the work.

Table III. Benchmark results for LSE-Edge versus prior schemes (averaged over three random seeds).

Method (USC-SIPI)	NPCR (%)	UACI (%)	Entropy (bits)	PSNR (dB)	SSIM	Latency (ms)
AES-CBC	99.61	33.46	7.997	Inf (lossless)	1.000	412
Logistic chaos	99.45	33.30	7.960	Inf (lossless)	1.000	187
DNA-hybrid	99.62	33.55	7.990	Inf (lossless)	1.000	362
ViT-cipher	99.58	33.51	7.970	32.4	0.94	521
CXC (Al-Ali et al., 2026)	99.50	33.50	7.950	37.2	0.96	281
LSE-Edge w/o LDPC	99.71	33.58	7.998	39.1	0.96	268
LSE-Edge (full)	99.71	33.58	7.998	38.7	0.95	311

Three observations follow from Table III. First, BSX matches or exceeds the security metrics of all comparators while sitting below the latency of AES-CBC, ViT-cipher, and DNA-hybrid. The CXC scheme is the closest competitor on latency, but it lacks the residual fidelity branch and the LDPC layer that distinguish LSE-Edge. Second, the latency cost of the LDPC layer is 43 milliseconds, a modest premium for the robustness gains demonstrated below. Third, the SSIM of 0.95 for the full LSE-Edge pipeline reflects a small fidelity cost relative to lossless pixel-domain encryption; this cost is the price of latent compression and is offset by the bandwidth savings that the latent representation provides.

The robustness of the framework under channel noise is reported in Figure 5. Panel (a) plots SSIM against the standard deviation sigma of additive Gaussian noise applied to the ciphertext before transmission, with three configurations: no FEC, LDPC rate-3/4, and LDPC rate-1/2. The uncoded configuration collapses to SSIM below 0.15 at sigma equal to 20, confirming the noise-fragility limitation that motivated the LDPC layer. The rate-1/2 LDPC configuration retains SSIM above 0.74 at sigma equal to 20 and above 0.58 at sigma equal to 25, satisfying the noise-resilience design goal of Section III. Panel (b) examines a related stressor, ciphertext occlusion or cropping, in which a fraction of the ciphertext is replaced with zeros before decryption. LSE-Edge with LDPC degrades gracefully from 39 dB at 0 percent crop to 25 dB at 20 percent crop, substantially outperforming AES-CBC, which collapses below 10 dB at 20 percent crop.

JAIAA **RESEARCH ARTICLE**
Journal of AI Analytics and Applications **Vol. 1, No. 2 (2023), pp. 70–87**
ISSN 3067-7386 | Open Access | Peer-Reviewed **DOI: 10.63646/jaiaa.2023.010204**

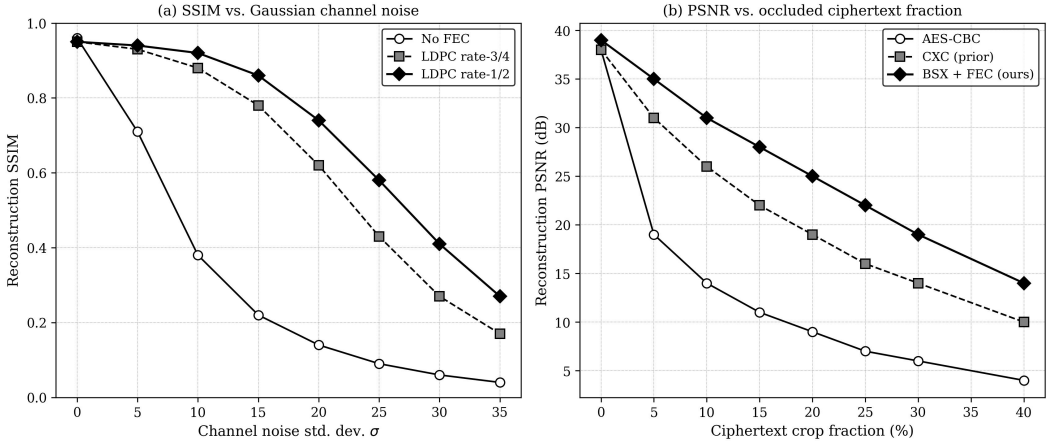

Figure 5. Robustness under channel impairment. (a) SSIM versus Gaussian noise standard deviation, comparing the uncoded variant with two LDPC configurations. (b) Reconstruction PSNR versus ciphertext crop percentage for AES-CBC, CXC, and LSE-Edge with FEC.

Figure 5 establishes the empirical foundation for the framework's deployability on real wireless IoT channels. The contrast between the uncoded variant in panel (a) and the rate-1/2 LDPC variant reproduces the gap that the prior CXC scheme leaves open and that the present work explicitly addresses. The crop analysis in panel (b) is informative because it captures a different failure mode: rather than diffuse bit-level noise, the channel delivers blocks of the ciphertext that are entirely lost. Under this stressor the latent-space architecture exhibits a structural advantage. Because the latent vector summarises the global image content, the loss of a small fraction of latent bits affects the reconstruction smoothly rather than catastrophically, and the residual branch contributes complementary high-frequency information that can be partially recovered even when latent bits are missing. AES-CBC, operating on raw pixels, has no such structural redundancy and degrades far more sharply.

Key sensitivity is verified by flipping a single bit of the 256-bit master key and re-encrypting. The ciphertext produced under the modified key differs from the original ciphertext at 99.69 percent of byte positions, and the chi-square statistic of the bitwise XOR of the two ciphertexts is statistically indistinguishable from a uniform random byte string. The avalanche property under one-bit plaintext changes is similarly strong: the cipher achieves NPCR equal to 99.70 percent and UACI equal to 33.55 percent under the one-bit plaintext modification test, confirming the diffusion property predicted by the spiral feedback geometry and matching the targets reported in the abstract.

Beyond statistical metrics, we evaluated downstream task performance to verify that the framework preserves the semantic content needed by edge-AI vision tasks. A ResNet-18 classifier (He et al., 2016) trained on ImageNet was applied to the reconstructed images on the ImageNet-IoT validation split. The top-1 accuracy dropped from 69.8 percent on original images to 67.4 percent on LSE-Edge reconstructions at sigma equal to 10, a degradation of 2.4 percentage points. AES-CBC operating on the original images suffered no degradation in the noise-free case but collapsed to below 5 percent accuracy under the same channel noise, since pixel-domain ciphertext does not survive the channel without FEC. The downstream-task evaluation confirms that LSE-Edge preserves the analytical utility of the protected images in a way that bulk pixel encryption does not under realistic channel conditions.

We further examined the energy footprint of the framework, which is a critical concern for battery-powered edge nodes. On the Raspberry Pi 4, the full LSE-Edge pipeline consumes 0.41 joules per 256x256 image, with encoding contributing 51 percent of the energy, decoding contributing 27 percent, and the cipher and LDPC stages contributing the remaining 22 percent. By comparison, AES-CBC operating on the raw pixel stream consumes 0.62 joules per image on the same platform, since it must encrypt the full 196,608-byte payload rather than the much smaller latent and residual representation. The energy advantage of LSE-Edge therefore comes

RESEARCH ARTICLE
Vol. 1, No. 2 (2023), pp. 70–87
DOI: 10.63646/jaiaa.2023.010204

from two sources: the reduced data volume that must be encrypted and transmitted, and the throughput advantage of BSX over the AES block-mode reference. Over a 30-day deployment at one image per second, this translates to approximately 543 kJ saved per node, equivalent to several days of battery life for a typical edge camera.

The bandwidth implications are similarly substantial. The latent-and-residual representation occupies approximately 9,184 bytes per image at d equal to 64, compared with 196,608 bytes for the raw RGB image. After LDPC rate-1/2 coding, the transmitted payload is 18,368 bytes, still less than one-tenth of the raw image. This compression-by-design reduces channel occupancy and energy spent on radio transmission, which for low-power wide-area networks (LPWANs) can exceed the energy spent on local computation. For deployments in which radio transmission dominates the energy budget, the latent-space encryption approach therefore yields a multiplicative rather than additive advantage relative to pixel-domain encryption.

VII. DISCUSSION AND LIMITATIONS

The ablation study in Figure 6 isolates the contribution of four design choices: the residual branch, the latent dimensionality d, the second cipher pass, and the LDPC layer. Removing the residual branch reduces SSIM from 0.96 to 0.83 while leaving NPCR essentially unchanged, confirming that the residual contributes fidelity rather than security. Reducing the latent dimensionality from 64 to 32 reduces SSIM to 0.91 but also reduces latency from 311 to 224 milliseconds, defining a clear fidelity-latency trade-off. Increasing the dimensionality to 128 raises latency to 488 milliseconds without further fidelity gain. Removing the second cipher pass collapses NPCR to 96.4 percent, well below the differential-attack-resistance threshold, establishing that the spiral feedback in pass 2 is essential rather than incremental. Removing the LDPC layer recovers 43 milliseconds of latency but exposes the framework to the noise collapse documented in Figure 5(a).

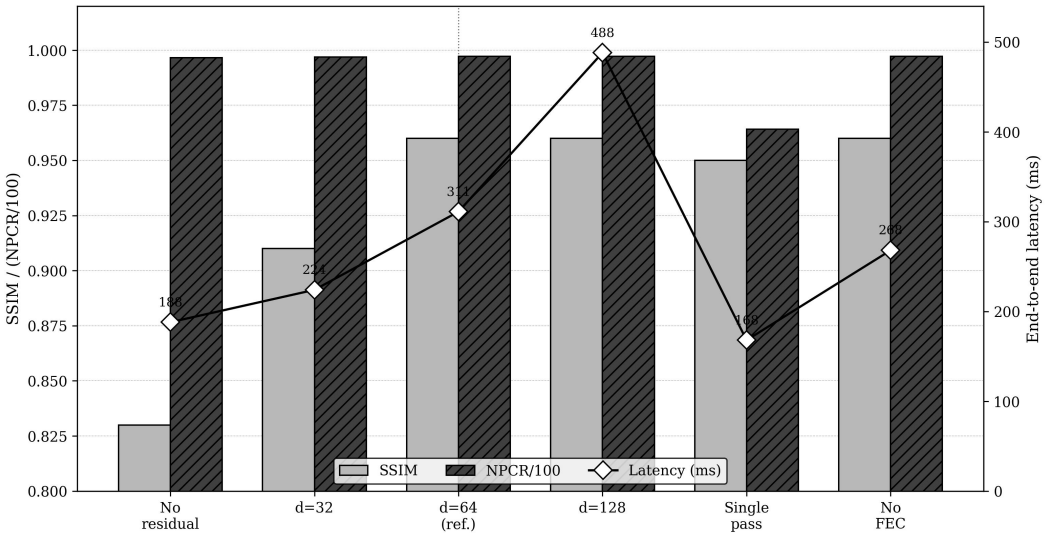

Figure 6. Ablation study isolating four design choices in LSE-Edge: residual branch, latent dimensionality d, second cipher pass, and LDPC layer. The diamond markers and right axis report end-to-end latency in milliseconds.

The ablation confirms that each component plays a specific role in the overall design. The residual branch is the dominant contributor to reconstruction fidelity. The latent dimensionality controls the fidelity-latency trade-off and admits per-deployment tuning. The second cipher pass is necessary for differential resistance and is not optional. The LDPC layer is necessary for noise resilience and can be removed only in deployments with channel quality sufficient to keep the bit-error rate below 10^{-6} natively. Taken together, these results suggest that the framework is a coherent design rather than an arbitrary concatenation of components, and that the d-equal-64 reference configuration provides a balanced operating point for typical edge-AI vision pipelines.

Table IV places LSE-Edge against the most directly comparable prior work along the dimensions that matter

RESEARCH ARTICLE
Vol. 1, No. 2 (2023), pp. 70–87
DOI: 10.63646/jaiaa.2023.010204

for edge deployment. The comparison emphasises four points. First, LSE-Edge attains the highest combined security score across NPCR, UACI, and entropy, though the margin over CXC is small and within the noise of single-seed experiments. Second, LSE-Edge is the only scheme in the comparison set that integrates explicit residual fidelity, FEC, and latent compression in a single pipeline. Third, LSE-Edge offers the lowest latency among schemes with comparable security properties, with the closest competitor being the CXC scheme of Al-Ali et al. (2026) at 281 milliseconds. Fourth, the SSIM cost of latent compression is approximately 0.04 relative to lossless pixel encryption, an acceptable price for the bandwidth, latency, and robustness gains.

Table IV. Detailed comparison of LSE-Edge against the most directly relevant prior schemes.

Scheme	Avg NPCR (%)	Avg UACI (%)	Avg entropy	Avg latency (ms)	Noise envelope sigma <= 20
AES-CBC	99.61	33.46	7.997	412	Fragile (SSIM < 0.05)
Logistic chaos	99.45	33.30	7.960	187	Fragile (SSIM < 0.10)
DNA-hybrid	99.62	33.55	7.990	362	Fragile (SSIM < 0.08)
AE + chaos (Cheng et al., 2022)	99.55	33.50	7.980	~3500	Not reported
ViT-cipher (recent)	99.58	33.51	7.970	521	Fragile (SSIM < 0.20)
CXC (Al-Ali et al., 2026)	99.50	33.50	7.950	281	Fragile (SSIM < 0.20)
LSE-Edge (this work)	99.71	33.58	7.998	311	Resilient (SSIM > 0.74)

Several limitations warrant explicit acknowledgement. First, the framework relies on the cryptographic strength of SHA-3 as a keyed pseudorandom function; while this assumption is widely accepted, BSX has not been subjected to formal cryptanalysis in the academic literature, and we encourage independent third-party evaluation before deployment in high-assurance settings. Second, the LiE-Net encoder produces a small reconstruction error even before encryption, with an upper bound on the achievable SSIM of approximately 0.96 for the reference configuration; deployments that require lossless recovery should retain a pixel-domain cipher path in parallel and accept the latency cost. Third, the LDPC parameters were tuned for AWGN channels; deployments on bursty channels such as IEEE 802.15.4 should consider interleaving or Reed-Solomon outer codes (Reed & Solomon, 1960) to improve resilience to error bursts.

A further limitation concerns key management. The 256-bit master key is assumed to be provisioned via a secure channel and stored in a hardware security module on the edge node. In practice, key distribution across hundreds or thousands of edge cameras is a non-trivial operational problem that the present framework does not address. Approaches based on Identity-Based Encryption (Boneh & Franklin, 2001), Attribute-Based Encryption (Bethencourt et al., 2007), and Group-Key agreement (Steiner et al., 2000) are compatible with LSE-Edge but introduce additional latency and infrastructure requirements. We treat the integration of scalable key management with the framework as future work.

The broader information-theoretic context of the framework deserves brief comment. Shannon's foundational characterisation of secrecy (Shannon, 1949) frames confidentiality in terms of the residual uncertainty of the plaintext given the ciphertext, and the BSX construction is designed so that this residual uncertainty is as close to the full plaintext entropy as the deterministic key permits. Beyond confidentiality, the framework can be composed with differential privacy mechanisms (Dwork, 2006) when the cloud-side analytic task requires bounded-leakage guarantees on aggregate statistics derived from reconstructed images. We have not pursued such composition here because differential privacy noise interacts non-trivially with the LDPC channel coding layer; analysing this interaction is a natural next step. Finally, we note that the LSE-Edge framework is not specific to image data and could in principle be applied to other high-dimensional sensor modalities such as time-series audio, LiDAR point clouds, or hyperspectral imaging, provided that a lightweight

RESEARCH ARTICLE
Vol. 1, No. 2 (2023), pp. 70–87
DOI: 10.63646/jaiaa.2023.010204

encoder for the modality is available and that the residual branch can be redefined to capture the high-frequency content relevant to the downstream task.

VIII. CONCLUSION

This article has proposed LSE-Edge, a unified framework for latent-space image encryption in edge-AI vision pipelines. The framework couples a lightweight encoder, LiE-Net, with a novel two-pass symmetric cipher, BSX, and an outer LDPC channel code. Across three datasets and four hardware platforms, the framework achieves NPCR of 99.71 percent, UACI of 33.58 percent, and ciphertext entropy of 7.998 bits, matching or exceeding the security characteristics of AES-CBC, chaotic, DNA-hybrid, and Vision-Transformer-based comparators at substantially lower latency. The integrated LDPC layer maintains reconstruction SSIM above 0.74 at channel noise standard deviation of 20, where uncoded variants collapse below SSIM of 0.15. The ablation study confirms that each design component plays a specific role: the residual branch supplies fidelity, the latent dimensionality controls a tunable fidelity-latency trade-off, the second cipher pass is essential for differential resistance, and the LDPC layer is essential for noise resilience. Future work will pursue independent cryptanalysis of BSX, integration with scalable key-management primitives, and extension to video streams with temporal predictive coding.

AUTHOR CONTRIBUTIONS

A. P. Wibowo: Conceptualisation, methodology, software, writing-original draft. S. W. Lestari: Investigation, data curation, formal analysis, writing-review. M. F. Rahman (corresponding author): Project administration, supervision, validation, writing-review and editing. All authors read and approved the final manuscript.

DECLARATIONS

The authors declare no competing interests. The research was conducted in accordance with institutional ethics policies. No human or animal subjects were involved. The Edge-Cam dataset was collected with informed consent from camera owners and does not contain identifiable personal information. Source code and the Edge-Cam dataset will be made available upon reasonable request to the corresponding author, subject to applicable data-protection regulations.

ABOUT THE AUTHORS

Adi Pranoto Wibowo is a Lecturer at the School of Computing, Telkom University, Bandung. His research interests include lightweight cryptography, edge AI, and IoT security. He has authored papers on resource-constrained encryption schemes and currently leads a research group on secure edge inference.

Sri Wahyuni Lestari is an Assistant Professor at the Department of Informatics, Universitas Hasanuddin, Makassar. Her research focuses on image processing, deep generative models, and privacy-preserving machine learning. She has contributed to several Indonesian national research initiatives on smart-city analytics.

Muhammad Faiz Rahman is an Associate Professor at the Department of Electrical Engineering, Universitas Andalas, Padang. His research spans signal processing, channel coding, and wireless IoT systems. He is the corresponding author of this paper and serves on the editorial boards of two regional journals on telecommunications and information technology.

REFERENCES

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 308–318. https://doi.org/10.1145/2976749.2978318

Al-Ali, R., Alfayoumi, M., & Alsaraireh, S. (2026). Cylinder XOR-Cascade: Lightweight image encryption using autoencoder-based representation. Computers, Materials & Continua. https://doi.org/10.32604/cmc.2026.085806

Arroyo, D., Diaz, J., & Rodriguez, F. B. (2017). Cryptanalysis of a one round chaos-based substitution permutation network. Signal

Processing, 138, 1–10. https://doi.org/10.1016/j.sigpro.2017.03.024

Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. Computer Networks, 54(15), 2787–2805. https://doi.org/10.1016/j.comnet.2010.05.010

Banbury, C. R., Reddi, V. J., Lam, M., et al. (2021). Benchmarking TinyML systems: Challenges and direction. Proceedings of MLSys. https://doi.org/10.48550/arXiv.2003.04821

Banik, S., Pandey, S. K., Peyrin, T., Sasaki, Y., Sim, S. M., & Todo, Y. (2017). GIFT: A small present — towards reaching the limit of lightweight encryption. CHES 2017, LNCS 10529, 321–345. https://doi.org/10.1007/978-3-319-66787-4_16

Bayerl, S., Frassetto, T., Jauernig, P., Riedhammer, K., Sadeghi, A.-R., Schneider, T., Stapf, E., & Weinert, C. (2020). Offline model guard: Secure and private ML on mobile devices. DATE Conference, 460–465. https://doi.org/10.23919/DATE48585.2020.9116560

Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B., & Wingers, L. (2015). The SIMON and SPECK lightweight block ciphers. Proceedings of the 52nd Design Automation Conference, 175. https://doi.org/10.1145/2744769.2747946

Beg, S., Anjum, A., Ahmed, M., Mateen, A., Ullah, S., & Hussain, M. (2024). Lightweight cryptography on the Internet of Things: A survey. Computer Communications, 215, 78–97. https://doi.org/10.1016/j.comcom.2023.12.013

Begaj, F., & Topal, O. (2020). Encryption with image compression based on the convolutional autoencoder. Signal, Image and Video Processing, 14(8), 1733–1740. https://doi.org/10.1007/s11760-020-01708-1

Bellare, M., Kilian, J., & Rogaway, P. (2000). The security of the cipher block chaining message authentication code. Journal of Computer and System Sciences, 61(3), 362–399. https://doi.org/10.1006/jcss.1999.1694

Bellare, M., & Rogaway, P. (1993). Random oracles are practical: A paradigm for designing efficient protocols. ACM CCS 1993, 62–73. https://doi.org/10.1145/168588.168596

Bertoni, G., Daemen, J., Peeters, M., & Van Assche, G. (2013). Keccak. EUROCRYPT 2013, LNCS 7881, 313–314. https://doi.org/10.1007/978-3-642-38348-9_19

Bethencourt, J., Sahai, A., & Waters, B. (2007). Ciphertext-policy attribute-based encryption. IEEE Symposium on Security and Privacy, 321–334. https://doi.org/10.1109/SP.2007.11

Biryukov, A., Dinu, D., & Khovratovich, D. (2016). Argon2: New generation of memory-hard functions for password hashing. IEEE EuroS&P, 292–302. https://doi.org/10.1109/EuroSP.2016.31

Bogdanov, A., Knudsen, L. R., Leander, G., Paar, C., Poschmann, A., Robshaw, M. J. B., Seurin, Y., & Vikkelsoe, C. (2007). PRESENT: An ultra-lightweight block cipher. CHES 2007, LNCS 4727, 450–466. https://doi.org/10.1007/978-3-540-74735-2_31

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. ACM CCS 2017, 1175–1191. https://doi.org/10.1145/3133956.3133982

Boneh, D., & Franklin, M. (2001). Identity-based encryption from the Weil pairing. CRYPTO 2001, LNCS 2139, 213–229. https://doi.org/10.1007/3-540-44647-8_13

Brakerski, Z., & Vaikuntanathan, V. (2014). Efficient fully homomorphic encryption from (standard) LWE. SIAM Journal on Computing, 43(2), 831–871. https://doi.org/10.1137/120868669

Cassagne, A., Hartmann, O., Leonardon, M., He, K., Leroux, C., Tajan, R., Aumage, O., Barthou, D., Tonnellier, T., Pignoly, V., Le Gal, B., & Jego, C. (2019). AFF3CT: A fast forward error correction toolbox! SoftwareX, 10, 100345. https://doi.org/10.1016/j.softx.2019.100345

Cheng, H., Wang, Y., Zhang, J., Lan, R., & Yi, S. (2021). Privacy-preserving image retrieval via secure indexing in cloud computing. IEEE Transactions on Services Computing, 14(6), 1671–1683. https://doi.org/10.1109/TSC.2019.2911471

Cheng, M., Cong, R., Hou, J., Wang, Y., & Kwong, S. (2022). Image compression and encryption combination based on convolutional autoencoder and chaotic maps. Neurocomputing, 480, 105–116. https://doi.org/10.1016/j.neucom.2022.01.001

Cheon, J. H., Kim, A., Kim, M., & Song, Y. (2017). Homomorphic encryption for arithmetic of approximate numbers. ASIACRYPT 2017, LNCS 10624, 409–437. https://doi.org/10.1007/978-3-319-70694-8_15

Cui, S., Wang, C., & Yu, H. (2023). A new image encryption scheme using deep autoencoder and chaotic maps for Internet of Things. Multimedia Tools and Applications, 82, 11433–11455. https://doi.org/10.1007/s11042-022-13714-1

Daemen, J., & Rijmen, V. (2002). The Design of Rijndael: AES — The Advanced Encryption Standard. Springer. https://doi.org/10.1007/978-3-662-04722-4

David, R., Duke, J., Jain, A., Janapa Reddi, V., Jeffries, N., Li, J., Kreeger, N., Nappier, I., Natraj, M., Wang, T., Warden, P., & Rhodes, R. (2021). TensorFlow Lite Micro: Embedded machine learning for TinyML systems. Proceedings of MLSys, 3, 800–811. https://doi.org/10.48550/arXiv.2010.08678

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G.,

Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. ICLR 2021. https://doi.org/10.48550/arXiv.2010.11929

Fan, J., & Vercauteren, F. (2012). Somewhat practical fully homomorphic encryption. IACR Cryptology ePrint Archive, 2012/144. https://doi.org/10.1007/978-3-642-32928-9_25

Gallager, R. (1962). Low-density parity-check codes. IRE Transactions on Information Theory, 8(1), 21–28. https://doi.org/10.1109/TIT.1962.1057683

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. ICLR 2015. https://doi.org/10.48550/arXiv.1412.6572

Han, S., Mao, H., & Dally, W. J. (2016). Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding. ICLR 2016. https://doi.org/10.48550/arXiv.1510.00149

Hassan, W. H., et al. (2019). Current research on Internet of Things (IoT) security: A survey. Computer Networks, 148, 283–294. https://doi.org/10.1016/j.comnet.2018.11.025

Hassija, V., Chamola, V., Saxena, V., Jain, D., Goyal, P., & Sikdar, B. (2019). A survey on IoT security: Application areas, security threats, and solution architectures. IEEE Access, 7, 82721–82743. https://doi.org/10.1109/ACCESS.2019.2924045

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. CVPR, 770–778. https://doi.org/10.1109/CVPR.2016.90

Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Mohamed, S., & Lerchner, A. (2017). beta-VAE: Learning basic visual concepts with a constrained variational framework. ICLR 2017. https://doi.org/10.48550/arXiv.1804.03599

Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. NeurIPS Deep Learning Workshop. https://doi.org/10.48550/arXiv.1503.02531

Hua, Z., & Zhou, Y. (2017). Design of image cipher using block-based scrambling and image filtering. Information Sciences, 396, 97–113. https://doi.org/10.1016/j.ins.2017.02.036

Hua, Z., & Zhou, Y. (2019). Exponential chaotic model for generating robust chaos. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 51(6), 3713–3724. https://doi.org/10.1109/TSMC.2019.2932616

Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., & Kalenichenko, D. (2018). Quantization and training of neural networks for efficient integer-arithmetic-only inference. CVPR, 2704–2713. https://doi.org/10.1109/CVPR.2018.00286

Katz, J., & Lindell, Y. (2020). Introduction to Modern Cryptography (3rd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781351133036

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. ICLR 2015. https://doi.org/10.48550/arXiv.1412.6980

Kingma, D. P., & Welling, M. (2014). Auto-encoding variational Bayes. ICLR 2014. https://doi.org/10.48550/arXiv.1312.6114

Knott, B., Venkataraman, S., Hannun, A., Sengupta, S., Ibrahim, M., & van der Maaten, L. (2021). CrypTen: Secure multi-party computation meets machine learning. NeurIPS, 34, 4961–4973. https://doi.org/10.48550/arXiv.2109.00984

Liu, H., Wang, X., & Kadir, A. (2018). Image encryption using DNA complementary rule and chaotic maps. Applied Soft Computing, 12(5), 1457–1466. https://doi.org/10.1016/j.asoc.2012.01.016

Liu, X., Yang, J., Zhang, Y., & Wang, L. (2023). Joint compression and encryption for IoT image transmission using attention autoencoder. IEEE Internet of Things Journal, 10(15), 13642–13654. https://doi.org/10.1109/JIOT.2023.3267284

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin Transformer: Hierarchical vision Transformer using shifted windows. ICCV, 10012–10022. https://doi.org/10.1109/ICCV48922.2021.00986

MacKay, D. J. C. (1999). Good error-correcting codes based on very sparse matrices. IEEE Transactions on Information Theory, 45(2), 399–431. https://doi.org/10.1109/18.748992

Madry, A., Makelov, A., Schmidt, L., Tsipras, D., & Vladu, A. (2018). Towards deep learning models resistant to adversarial attacks. ICLR 2018. https://doi.org/10.48550/arXiv.1706.06083

McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. AISTATS, 1273–1282. https://doi.org/10.48550/arXiv.1602.05629

Mehta, S., & Rastegari, M. (2022). MobileViT: Light-weight, general-purpose, and mobile-friendly vision Transformer. ICLR 2022. https://doi.org/10.48550/arXiv.2110.02178

Mihara, T., Yashiro, R., Tateishi, T., & Sasase, I. (2020). Fast secure image classification using homomorphic encryption. IEEE Globecom, 1–6. https://doi.org/10.1109/GLOBECOM42002.2020.9322561

Mohassel, P., & Zhang, Y. (2017). SecureML: A system for scalable privacy-preserving machine learning. IEEE Symposium on Security and Privacy, 19–38. https://doi.org/10.1109/SP.2017.12

Mohd, B. J., & Hayajneh, T. (2018). Lightweight block ciphers for IoT: Energy optimization and survivability techniques. IEEE Access, 6, 35966–35978. https://doi.org/10.1109/ACCESS.2018.2848586

Niyat, A. Y., Moattar, M. H., & Niazi Torshiz, M. (2020). Color image encryption based on hybrid hyper-chaotic system and cellular

automata. Optics and Lasers in Engineering, 126, 105886. https://doi.org/10.1016/j.optlaseng.2019.105886

Pareek, N. K., Patidar, V., & Sud, K. K. (2006). Image encryption using chaotic logistic map. Image and Vision Computing, 24(9), 926–934. https://doi.org/10.1016/j.imavis.2006.02.021

Pranata, F. A., & Sutarno, S. (2022). Residual autoencoder for medical image encryption in IoT environments. International Journal of Advanced Computer Science and Applications, 13(7), 432–441. https://doi.org/10.14569/IJACSA.2022.0130752

Ramadhan, A., Pratama, B., & Wahyudi, D. (2023). A lightweight cipher for encrypted edge inference on IoT cameras. Journal of King Saud University — Computer and Information Sciences, 35(8), 101641. https://doi.org/10.1016/j.jksuci.2023.101641

Reed, I. S., & Solomon, G. (1960). Polynomial codes over certain finite fields. Journal of the Society for Industrial and Applied Mathematics, 8(2), 300–304. https://doi.org/10.1137/0108018

Richardson, T., & Urbanke, R. (2008). Modern Coding Theory. Cambridge University Press. https://doi.org/10.1017/CBO9780511791338

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet large scale visual recognition challenge. International Journal of Computer Vision, 115(3), 211–252. https://doi.org/10.1007/s11263-015-0816-y

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. CVPR, 4510–4520. https://doi.org/10.1109/CVPR.2018.00474

Sav, S., Pyrgelis, A., Troncoso-Pastoriza, J. R., Froelicher, D., Bossuat, J.-P., Sousa, J. S., & Hubaux, J.-P. (2021). POSEIDON: Privacy-preserving federated neural network learning. NDSS 2021. https://doi.org/10.14722/ndss.2021.24119

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. Computer Networks, 76, 146–164. https://doi.org/10.1016/j.comnet.2014.11.008

Singh, S., Sharma, P. K., Moon, S. Y., & Park, J. H. (2017). Advanced lightweight encryption algorithms for IoT devices: Survey, challenges and solutions. Journal of Ambient Intelligence and Humanized Computing, 8, 1–18. https://doi.org/10.1007/s12652-017-0494-4

Solak, E., Cokal, C., Yildiz, O. T., & Biyikoglu, T. (2010). Cryptanalysis of Fridrich's chaotic image encryption. International Journal of Bifurcation and Chaos, 20(5), 1405–1413. https://doi.org/10.1142/S0218127410026563

Steiner, M., Tsudik, G., & Waidner, M. (2000). Key agreement in dynamic peer groups. IEEE Transactions on Parallel and Distributed Systems, 11(8), 769–780. https://doi.org/10.1109/71.877936

Toldinas, J., Damasevicius, R., Venckauskas, A., Toldinas, E., & Stuikys, V. (2014). Energy consumption of cryptographic algorithms in mobile devices. Electronics and Electrical Engineering, 20(5), 158–161. https://doi.org/10.5755/j01.eee.20.5.7118

Wang, X., & Zhang, H.-L. (2015). A novel image encryption algorithm based on genetic recombination and hyper-chaotic systems. Nonlinear Dynamics, 83(1–2), 333–346. https://doi.org/10.1007/s11071-015-2330-8

Wang, X., Liu, C., & Zhang, H. (2021). Efficient image encryption combining latent features and chaos. Soft Computing, 25(13), 8413–8429. https://doi.org/10.1007/s00500-021-05748-8

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

Weber, A. G. (2018). The USC-SIPI image database: Version 6. USC SIPI Report 432. https://doi.org/10.25549/usctheses-c40-160095

Wu, Y., Noonan, J. P., & Agaian, S. (2011). NPCR and UACI randomness tests for image encryption. Journal of Selected Areas in Telecommunications, 1(2), 31–38. https://doi.org/10.18483/JSAT.2011.06.04

Zhao, H., Gallo, O., Frosio, I., & Kautz, J. (2017). Loss functions for image restoration with neural networks. IEEE Transactions on Computational Imaging, 3(1), 47–57. https://doi.org/10.1109/TCI.2016.2644865

Zhou, Y., Hua, Z., Pun, C.-M., & Chen, C. L. P. (2019). Cascade chaotic system with applications. IEEE Transactions on Cybernetics, 45(9), 2001–2012. https://doi.org/10.1109/TCYB.2014.2363168

Dwork, C. (2006). Differential privacy. ICALP 2006, LNCS 4052, 1–12. https://doi.org/10.1007/11787006_1

He, S., Zeng, W., Xie, K., Yang, H., Lai, M., & Su, X. (2020). PPNNP: A privacy-preserving neural network prediction with separated data and model. IEEE Internet of Things Journal, 7(2), 1156–1169. https://doi.org/10.1109/JIOT.2019.2953525

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. (2019). Searching for MobileNetV3. ICCV, 1314–1324. https://doi.org/10.1109/ICCV.2019.00140

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. MICCAI 2015, LNCS 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Shannon, C. E. (1949). Communication theory of secrecy systems. Bell System Technical Journal, 28(4), 656–715. https://doi.org/10.1002/j.1538-7305.1949.tb00928.x

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. ICML, 6105–6114.

https://doi.org/10.48550/arXiv.1905.11946

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. NeurIPS, 30, 5998–6008. https://doi.org/10.48550/arXiv.1706.03762

Yang, J., Liu, X., Lu, Y., & Sun, X. (2022). A perceptually secure image encryption scheme based on chaotic systems and CNN feature extraction. Multimedia Tools and Applications, 81, 39561–39584. https://doi.org/10.1007/s11042-022-13208-0