Minimax Theory and its Applications

Volume 11 (2026), No. 3s, 628 to 653

Med Fusion Crypt: Cross-Modal Attention-Guided Selective Encryption and Pairing-Free Proxy Re-Encryption for Secure Multimodal Medical Image Sharing in IoMT

P. Shyamala Madhuri[1], B. Amutha[1], D. J. Nagendra Kumar[2]

[1]Dept. of Computing Technologies, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathu, Chennai, India – 603203
[2]Dept. of Information Technology, Vishnu Institute of Technology, Bhimavaram, India – 534202
Email: sp6331@srmist.edu.in | amuthab@srmist.edu.in | nagendrakumar.dj@vishnu.edu.in

Abstract
The rapid growth of the Internet of Medical Things (IoMT) and telemedicine has increased the exchange of multi-modal medical images such as CT, MRI, and PET across hospitals, cloud servers, and specialists. These images carry highly sensitive patient information and are vulnerable to interception, tampering, and unauthorized redistribution. Existing blockchain-based proxy re-encryption (PRE) schemes mostly target textual electronic medical records, rely on computationally expensive bilinear pairings, and treat each imaging modality independently, which leads to redundant encryption, high latency, and poor suitability for resource-constrained IoMT devices. To address these limitations, this paper proposes a privacy-aware multi-modal medical image transmission framework that integrates a novel deep learning mechanism with a blockchain-based lightweight re-encryption scheme. First, a Cross-Modal Attention Fusion and Region-of-Interest Network (CMAF-RoINet) fuses complementary features from multiple imaging modalities into a compact representation and automatically localizes diagnostically critical regions. This enables selective encryption, so that high-strength protection is applied to sensitive regions while computational cost is reduced for background areas. Second, a deep learning–driven chaotic key-stream generator, combining an LSTM network with a hyper-chaotic map, produces image-specific, non-repeating keys that resist differential and brute-force attacks. Third, a pairing-free, elliptic-curve-based unidirectional proxy re-encryption scheme re-encrypts only the symmetric session key rather than the full image. This allows patients to delegate access to authorized clinicians without exposing plaintext to the proxy and resists collusion between the proxy and revoked users. Finally, a consortium blockchain with smart contracts manages patient consent, re-encryption key issuance, and revocation, and records an immutable audit trail, while encrypted images are stored off-chain in the InterPlanetary File System (IPFS) with only their content hashes stored on-chain for integrity verification. The proposed framework is evaluated on publicly available multi-modal datasetlikeBraTS. Experimental results show that the scheme achieves an information entropy, NPCR, UACI, and near-zero adjacent pixel correlation, with lossless reconstruction (PSNR of dB, SSIM) for authorized receivers. Compared with existing pairing-based PRE and blockchain image encryption schemes, the proposed approach reduces re-encryption computation cost and end-to-end transmission latency, while achieving a blockchain throughput. Security analysis further demonstrates resistance to chosen-plaintext, collusion, replay, and impersonation attacks. These results indicate that the proposed framework offers a secure, efficient, and patient-centric solution for multi-modal medical image sharing in IoMT-enabled healthcare systems.

Keywords: Multi-modal medical images, Proxy re-encryption, Blockchain, IPFS, Deep learning, Chaotic encryption, Internet of Medical Things, Privacy preservation

1. Introduction

The convergence of the Internet of Medical Things (IoMT), cloud computing, and telemedicine has transformed how healthcare is delivered. Imaging devices, bedside monitors, wearable sensors, and hospital information systems are now interconnected, allowing clinical data to move between hospitals, diagnostic centres, cloud platforms, and remote specialists in near real time [1]. This connectivity has made remote diagnosis, second-opinion consultation, and multidisciplinary tumour boards routine practice rather than exceptions.Medical imaging sits at the centre of this exchange. Modern diagnosis increasingly relies on multi-modal imaging, in

which complementary modalities describe the same anatomy from different physical perspectives. Computed tomography (CT) captures dense structures such as bone, magnetic resonance imaging (MRI) provides high soft-tissue contrast across sequences such as T1, T1ce, T2, and FLAIR, and positron emission tomography (PET) reveals metabolic and functional activity [2]. In neuro-oncology, for example, delineating a glioma and its surrounding oedema typically requires reading several MRI sequences together, often alongside PET [ref]. A single patient study may therefore contain hundreds of slices across several modalities, amounting to hundreds of megabytes of data [3].

These images are among the most sensitive forms of personal data. Beyond the pixels themselves, they reveal identity-linked anatomy (facial structure can be reconstructed from head MRI), disease status, and metadata embedded in DICOM headers [4]. When transmitted over open networks or stored on third-party cloud servers, they face interception, tampering, unauthorized redistribution, and ransomware. Tampering is especially dangerous in imaging, because a subtly altered scan can lead to misdiagnosis without anyone noticing. Regulations such as HIPAA in the United States, the GDPR in Europe, and India's Digital Personal Data Protection Act, 2023 place legal obligations on custodians to protect confidentiality, ensure integrity, and record who accessed patient data and with whose consent. Securing this exchange is not only a confidentiality problem but also an access-control problem [5]. A patient's images must be shareable with a changing set of clinicians: a referring physician today, a radiologist tomorrow, surgeon next week. Access must be granted without the data owner re-encrypting everything for each new recipient, and revoked when a clinician's role ends. At the same time, many IoMT endpoints such as portable scanners, edge gateways, and handheld devices have limited processing power, memory, and battery, so any protection mechanism must remain lightweight. These competing demands of strong security, flexible patient-controlled sharing, auditability, and low computational cost motivate the framework proposed [6].

Standard block ciphers such as AES were designed for text and perform poorly on images. Image pixels are highly redundant and strongly correlated with their neighbours, so naive block encryption can leave visible structure and is costly for large volumetric studies [7]. Chaos-based cryptosystems address this by exploiting the sensitivity to initial conditions, ergodicity, and pseudo-randomness of chaotic maps to build permutation–diffusion ciphers. However, many low-dimensional chaotic maps have small key spaces and short periods, and they degrade under finite-precision arithmetic [8]. Several schemes reuse the same key stream across images and have been broken by chosen-plaintext attacks [9]. Most of these schemes also encrypt the entire image uniformly, spending the same effort on empty background as on a tumour. Deep learning has been used to segment regions of interest (ROIs), to generate keys, and to learn encryption transforms. Existing ROI-based selective encryption methods, however, usually work on a single modality. They ignore the complementary information that a fused multi-modal view provides, so the ROI estimated from one modality may miss lesion boundaries that are only visible in another [10]. Learning-based key generators, in turn, are rarely coupled with chaotic dynamics in a way that guarantees image-specific, non-repeating key streams. Proxy re-encryption (PRE) lets a semi-trusted proxy transform a ciphertext encrypted under a delegator's key into one decryptable by a delegatee, without learning the plaintext. PRE is therefore a natural fit for patient-to-clinician delegation. Yet most PRE schemes proposed for healthcare rely on bilinear pairings, which are several times more expensive than elliptic-curve point multiplications and are poorly suited to constrainedIoMT devices [11]. Many are bidirectional or lack resistance to collusion between the proxy and a delegatee, which allows a revoked user to recover the delegator's secret key. Crucially, these schemes were mostly designed for textual electronic medical records (EMRs). Applying them directly to large images, especially when each modality is re-encrypted separately, causes redundant encryption and high latency [12]. Blockchain provides tamper-evident, decentralized logging, and smart contracts can enforce access policies without a single trusted administrator. Many blockchain-based healthcare frameworks, however, either store large data on-chain, which is impractical, or use the ledger only for logging while leaving key management and revocation outside it. Public blockchains add latency and transaction fees and expose metadata. Integrations of blockchain with PRE usually inherit the pairing-based cost of the underlying PRE scheme and seldom address image data or multi-modal studies [13].

1.1. Research Gaps

The review above reveals six gaps that, taken together, prevent existing schemes from serving multi-modal image sharing in IoMT settings.

1. Text-centric PRE. Blockchain-based PRE schemes for healthcare largely target textual EMRs. They do not account for the volume, redundancy, and spatial structure of medical images, and they apply public-key operations to bulk data rather than to a compact session key.

2. Modality-independent processing. CT, MRI, and PET images of the same patient are encrypted and re-encrypted separately. This duplicates computation and key management and ignores the correlated, complementary information across modalities [14].
3. Uniform, ROI-blind encryption. Most image ciphers protect every pixel equally. Existing ROI-based methods rely on single-modality or manually defined regions and do not use cross-modal attention to locate diagnostically critical areas reliably.
4. Static or weak key streams. Many chaos-based schemes use low-dimensional maps with limited key spaces, or reuse key streams across images. This leaves them open to chosen-plaintext and differential attacks [15].
5. Pairing-heavy, collusion-prone delegation. PRE schemes for healthcare depend on bilinear pairings that are too costly for constrained devices. Many are bidirectional or allow a colluding proxy and revoked delegatee to recover the delegator's secret key.
6. Incomplete, off-chain governance. Patient consent, re-encryption key issuance, and revocation are rarely enforced end to end by smart contracts. Storage designs either bloat the ledger with image data or lack on-chain integrity verification of off-chain ciphertexts.

1.2. Research Objectives

The overall aim of this work is to design a secure, efficient, and patient-centric framework for sharing multi-modal medical images across IoMT-enabled healthcare systems. This aim is broken into the following objectives, each addressing one or more of the gaps above.

1. To design a cross-modal deep learning network that fuses complementary features from CT, MRI, and PET and automatically localizes diagnostically critical regions, enabling selective, region-aware encryption.
2. To develop an image-specific key-stream generator that couples a learned sequence model with a hyper-chaotic system, producing non-repeating keys that resist brute-force, differential, and chosen-plaintext attacks.
3. To construct a pairing-free, elliptic-curve-based unidirectional proxy re-encryption scheme that re-encrypts only the symmetric session key and resists proxy–delegatee collusion.
4. To build a consortium blockchain layer whose smart contracts enforce patient consent, re-encryption key issuance, and revocation, with IPFS off-chain storage and on-chain hash verification.
5. To evaluate the framework on public multi-modal datasets in terms of cryptographic strength, reconstruction quality, computational cost, latency, and blockchain throughput, and to compare it against pairing-based PRE and blockchain image-encryption baselines.
6. To analyse the framework's security formally and informally against chosen-plaintext, collusion, replay, and impersonation attacks.

1.3. Contributions

The main contributions of this paper are as follows.

1) This work propose a Cross-Modal Attention Fusion and Region-of-Interest Network that fuses features from multiple imaging modalities into a compact shared representation and localizes diagnostically critical regions. The resulting ROI map drives selective encryption: sensitive regions receive high-strength protection while background regions are processed with a lighter cipher. This lowers computational cost without weakening protection where it matters, and a single ROI map serves all modalities of a study, removing redundant per-modality analysis.

2) LSTM-driven hyper-chaotic key-stream generator. We design a key-stream generator that combines a Long Short-Term Memory (LSTM) network with a hyper-chaotic map. Seeded with image-dependent information, it produces image-specific, non-repeating key streams with a large key space. This makes the cipher resistant to brute-force, differential, and chosen-plaintext attacks.

3) Pairing-free, collusion-resistant unidirectional PRE. We construct an elliptic-curve-based unidirectional proxy re-encryption scheme that avoids bilinear pairings entirely. Only the symmetric session key is re-encrypted, never the full image, so the cost of delegation does not grow with image size or the number of modalities. The proxy never sees plaintext, and a colluding proxy and revoked user cannot recover the patient's private key.

4) Smart-contract-governed consortium blockchain with IPFS storage. We integrate a consortium blockchain whose smart contracts manage patient consent, re-encryption key issuance, and revocation, and record an immutable audit trail of every access. Encrypted images are stored off-chain in IPFS, and only their content identifiers and hashes are kept on-chain, so integrity can be verified without bloating the ledger.

5) Comprehensive evaluation and security analysis. This work evaluate the framework on BraTS using information entropy, NPCR, UACI, adjacent-pixel correlation, PSNR, SSIM, re-encryption cost, end-to-end latency, and blockchain throughput. This work compare it with existing pairing-based PRE and blockchain image-encryption schemes and analyse its resistance to chosen-plaintext, collusion, replay, and impersonation attacks.

The remainder of this paper is organized as follows. Section 2 reviews related work on chaos-based and deep-learning-based medical image encryption, proxy re-encryption, and blockchain-based healthcare data sharing. Section 3 describes the proposed framework in detail: the CMAF-RoINet architecture, the LSTM-driven chaotic key-stream generator, the selective encryption procedure, the pairing-free proxy re-encryption scheme, and the blockchain and IPFS layer with its smart contracts. Section 4 describes the experimental setup and datasets and reports results on cryptographic strength together with comparisons against existing schemes. Section 5 discusses limitations and practical deployment considerations. Section 8 concludes the paper and outlines directions for future work.

2. Related works

This section reviews four lines of work that the proposed framework builds on: chaos-based image encryption, deep learning for medical image security, proxy re-encryption, and blockchain-based medical data sharing. It then compares representative schemes and identifies the gaps the framework addresses.

2.1. Reviews on Chaos-Based Medical Image Encryption

Chaos-based ciphers are the dominant family in the image-encryption literature. They work well on images, but many published designs have been broken because their key streams do not depend on the plaintext. The author introduced the permutation–diffusion architecture, in which a chaotic map first scrambles pixel positions and then diffuses pixel values so that a one-pixel change spreads across the cipher image. Chaotic systems suit this role because of their sensitivity to initial conditions, ergodicity, and noise-like trajectories. The author in [16] later set out basic requirements a chaos-based cryptosystem must meet, including a well-defined and sufficiently large key space, avoidance of weak keys, and resistance to known- and chosen-plaintext attacks. Many later schemes are still judged against these requirements. Low-dimensional maps such as the logistic map have short periods and small key spaces, especially under finite-precision arithmetic. Recent medical-image schemes therefore use hyper-chaotic systems, which have two or more positive Lyapunov exponents and richer dynamics. The author constructed a two-dimensional Logistic–Gaussian hyper-chaotic map and used it to drive permutation and multi-directional pixel substitution for medical images of different formats. Another scheme [17] seeds a four-dimensional hyper-chaotic system with the SHA-3 hash of the input image and combines it with DNA coding, making the key stream depend on the image itself. The authors combined hyper-chaotic sequences with layered cellular automata and built a hybrid cryptosystem in which a 4D hyper-chaotic map generates S-boxes after a 3D Chen-map shuffle. A recent DICOM-compatible scheme uses hyper-chaotic signals and multi-level diffusion across MRI, CT, X-ray, and ultrasound, and superimposes an encrypted captcha image as a second authentication factor. For the Medical IoT setting, the author in [18] embed a DWT watermark and then encrypt with a memristor-modulated hyper-chaotic map, supporting post-incident accountability.Because full-image encryption is expensive, some works encrypt only part of the image. One partial scheme encrypts only the lowest-frequency band of an integer wavelet transform, using polynomial secret image sharing, a hyper-chaotic map, and a hash function [19]. These schemes reduce cost, but the protected region is chosen by a fixed transform rather than by clinical importance.A large body of cryptanalysis shows that chaos alone does not guarantee security. The author showed that a hyper-chaos-based scheme could be broken with a single known plain-image and that a logistic-map colour cipher fell to a chosen-plaintext attack needing only a handful of chosen images. The author broke two schemes driven by Lorenz, Chen, and hyper-chaotic systems with a chosen-plaintext attack. The author broke a dynamic S-box scheme with 256 chosen plaintexts because the S-boxes and key bytes did not change between images of the same size. The recurring cause is a key stream that is independent of the plaintext, together with weak diffusion and poor key sensitivity.These findings motivate two design choices in the proposed framework. First, the key stream must be image-specific and non-repeating, which the LSTM-driven hyper-chaotic generator is designed to provide. Second, the scheme's security should be argued against chosen-plaintext attacks explicitly, not inferred from statistical metrics such as entropy and NPCR alone [20].

2.2. Reviews on Deep Learning for Medical Image Security

Deep learning now plays three roles in medical image security: finding the region to protect, generating keys, and learning the cipher itself. No reviewed work combines cross-modal fusion with ROI-guided encryption. Selective encryption protects diagnostically important regions strongly and treats the background lightly. Early work by [21] separated DICOM images into a region of interest (ROI) and region of background (ROB) using pixel thresholding. The ROI received a strong cipher and the ROB a light one, and multi-frame studies reused a single segmentation map from a reference frame, which cut encryption time sharply. Thresholding, however, relies on the assumption that images of the same modality and body part share similar intensity statistics. Learned segmentation removes this assumption. One scheme uses a Med3D-pretrained 3D ResNet-50 to produce patient-specific masks and then encrypts only the masked voxels with a chaotic Hénon map [22]. Another uses U-Net to localize the tumour before partial encryption with hyper-chaotic permutation and secret image sharing. In all of these, the ROI is derived from a single modality or a single image. The author proposed DeepEDN [23], which uses a Cycle-GAN to transfer a medical image into a target "encrypted" domain and back, with an ROI-mining network that can extract objects directly from the encrypted image. The same group proposed DeepKeyGen in which a GAN learns to map a seed image to a private key that is then XORed with the image; it was evaluated on chest X-ray, ultrasound, and BraTS18 data. A later analysis, however, found that DeepKeyGen's keys had low sensitivity to changes in their inputs, which created exploitable weaknesses [24]. This illustrates a broader concern: neural networks are not designed to be sensitive to tiny input changes, so a purely learned key generator may lack the avalanche behaviour that cryptography requires. A newer line of work combines chaotic systems with recurrent networks. TDADL-IE trains a bidirectional LSTM on sequences from a one-dimensional sine–quadratic chaotic map and concatenates the learned sequence with the raw chaotic one to drive a three-dimensional diffusion step for medical images. Earlier, an LSTM network was trained on hyper-chaotic Lorenz sequences to produce new chaotic signals for image encryption [25]. These hybrids aim to keep chaos's sensitivity while adding the complexity of a learned model. Their limitations are that key streams are often not tied to the specific image, and that reproducing the same neural output on sender and receiver hardware is rarely discussed. In medical image analysis, multi-modal fusion has moved from concatenating modalities at the input to explicit cross-modal attention. NestedFormermodels intra- and inter-modality relationships of multi-modal MRI with a cross-modality attention transformer and outperformed earlier methods on BraTS 2020. CMAF-Net uses cross-modal attention fusion for brain tumor segmentation when some MRI modalities are missing. A cross-modality attention adapter has also been used to fine-tune the Segment Anything Model for glioma segmentation from multi-modal MRI [26]. These works show that fused representations localize lesions more accurately than any single modality. However, they were developed for segmentation and diagnosis, not for guiding encryption. The proposed CMAF-RoINet brings cross-modal attention fusion into the security pipeline, so that one ROI map, informed by all modalities, drives selective encryption for the whole study. The LSTM–hyper-chaotic generator follows the hybrid model but must address the image-specificity and reproducibility issues noted above.

2.3. Reviews on Proxy Re-Encryption

Proxy re-encryption (PRE) is the natural tool for patient-to-clinician delegation, but secure unidirectional, collusion-resistant schemes have mostly relied on bilinear pairings, and pairing-free constructions have a history of flawed proofs. The author introduced PRE, in which a semi-trusted proxy holding a re-encryption key converts a ciphertext under Alice's key into one under Bob's key without learning the plaintext. Their ElGamal-based scheme was bidirectional and not collusion-safe: the proxy and Bob together could recover Alice's secret key. The author gave the first unidirectional, collusion-resistant schemes using bilinear pairings and applied them to secure distributed storage. Their work also set out the properties used to compare later schemes: unidirectionality, non-interactivity, non-transitivity, collusion safety, and resistance to proxy misuse. The author in [26] defined chosen-ciphertext (CCA) security for PRE and built a bidirectional CCA-secure scheme in the standard model. Then, gave the first unidirectional CCA-secure (replayable) scheme in the standard model. Both rely on bilinear pairings, which remain several times more expensive than elliptic-curve point multiplication and are therefore costly for IoMT devices. The author proposed the first unidirectional PRE that is CCA-secure and collusion-resistant without pairings, in the random oracle model. Its security rests on the DDH assumption over Z_N^2 and integer factorization, which leads to large parameters rather than compact elliptic-curve elements. The author in [27] presented a more efficient pairing-free, collusion-resistant unidirectional scheme at Africacrypt 2010. The author later identified a weakness in the security proof of that scheme and proposed a corrected collusion-resistant, pairing-free unidirectional scheme that is CCA-secure under a variant of the computational Diffie–Hellman assumption in the random oracle model. The author similarly showed that an earlier standard-model pairing-free scheme was not CCA-secure and proposed a Cramer–Shoup-based replacement. The lesson is that pairing-free PRE is achievable, but each new scheme needs a careful, formal proof.

Healthcare applications of PRE have focused almost entirely on electronic health records (EHRs) and textual electronic medical records (EMRs). These schemes, reviewed with their blockchain layers, typically re-encrypt either the record itself or a data key, and many use pairing-based or attribute-based PRE for fine-grained access control. A recent design uses a lightweight hybrid elliptic-curve PRE together with ECIES and ECDSA for interoperable EHR sharing across blockchain networks showing growing interest in pairing-free healthcare PRE. None of these schemes, however, is designed for large multi-modal image studies. The proposed scheme belongs to the pairing-free, unidirectional, collusion-resistant family [28]. Because earlier schemes in this family have had flawed proofs, the paper should state its security model, hardness assumption, and proof explicitly. Re-encrypting only the symmetric session key keeps the PRE cost constant regardless of image size or number of modalities.

2.4. Reviews Blockchain and IPFS for Medical Data Sharing

Blockchain-based medical sharing has converged on a consortium-chain, off-chain-storage design. Systems that add PRE target textual records, while systems that handle images rarely include delegation. MedRec [29] was among the first decentralized medical record systems, using Ethereum smart contracts to manage access permissions and to point to records held by providers. Later work moved to permissioned (consortium) blockchains such as Hyperledger Fabric which restrict participation to known institutions, avoid transaction fees, and offer higher throughput and lower latency than public chains. These properties suit hospital networks, where members are identifiable and regulated. Because ledgers cannot hold large files, most designs store only metadata or hashes on-chain. The InterPlanetary File System (IPFS) is a common off-chain store because it is content-addressed: a file's identifier is derived from its hash, so any tampering changes the address and is detectable. The author store only EMR metadata on a consortium chain and keep encrypted EMRs in the cloud. The author combine IPFS with a consortium chain in a hybrid storage design. The author place a PRE proxy between file servers, separate from an outsourced blockchain, and use smart contracts for access control so that the blockchain operator cannot read records. It use unidirectional multi-hop conditional PRE based on type and identity, with distributed key generation, for fine-grained EHR sharing and traceable usage. The author introduce an improved PRE scheme to resist collusion and add conditional anonymous tracing of malicious users. The author combines attribute-based signcryption with PRE on a consortium chain for EMR sharing. A recent framework uses lightweight elliptic-curve PRE with ECIES and ECDSA to share EHRs across separate blockchain networks via smart contracts. All of these target EHRs or EMRs rather than images. Image-oriented systems emphasize storage and integrity rather than delegation. The author store medical images in IPFS and their content hashes on a blockchain to remove dependence on a central server. A patient-centric image management system similarly combines blockchain smart contracts with IPFS to give patients control over image access across hospitals. A zero-trust framework [30] adds edge computing to reduce retrieval latency while smart contracts enforce patient-defined permissions. BCAES encrypts images with Arnold's cat map, stores the ciphertext in the cloud, and records a signed document on the blockchain so the receiver can verify integrity. These systems either rely on a single-key cipher shared directly with recipients or leave key distribution unspecified, and none supports proxy-based delegation and revocation. The proposed framework adopts the established consortium-chain and IPFS pattern, but extends smart contracts to manage the full delegation lifecycle: consent, re-encryption key issuance, revocation, and audit. It also brings PRE-based delegation, so far limited to textual records, to multi-modal image studies [30].

Table 1. Feature comparison of representative schemes with the proposed framework.

Scheme	Data	Multi-modal handling	ROI-selective encryption	Key stream	Delegation (PRE)	Ledger / storage
SHA-3 + DNA + 4D hyper-chaos [4]	Medical images	No	No	Image-dependent (SHA-3 seed)	None	None
DICOM hyper-chaotic [7]	MRI, CT, X-ray, US	Several modalities, encrypted separately	No	Hyper-chaotic	None	None
Partial IWT + secret sharing [9]	Medical images	No	Yes (U-Net tumour ROI)	Hyper-chaotic	None	None
Natsheh et al. [14]	DICOM	No	Yes (threshold	Standard ciphers	None	None

			ROI/ROB)			
Segmentation mask + Hénon [15]	3D volumes	No	Yes (3D ResNet-50 mask)	Hénon map	None	None
DeepKeyGen [17]	X-ray, US, BraTS18	Tested on several modalities, no fusion	No	GAN-learned	None	None
TDADL-IE [19]	Medical images	No	No	BLSTM + 1D chaotic map	None	None
Sharmila Deva Selvi et al. [30]	Generic messages	No	No	Not image-specific	Pairing-free, unidirectional, collusion-resistant, CCA (ROM)	None
Park et al. [34]	EHR	No	No	Not image-specific	PRE proxy between file servers	Blockchain + smart contracts
Wu et al. [36]	EHR	No	No	Not image-specific	Unidirectional multi-hop conditional PRE	Consortium chain + IPFS
Liu et al. [37]	EMR	No	No	Not image-specific	Collusion-resistant PRE	Consortium chain + cloud
HEC-PRE [38]	EHR	No	No	Not image-specific	Lightweight ECC PRE	Multiple chains + smart contracts
Ashwini et al. [39]	Medical images	No	No	Not specified	None	Blockchain + IPFS
BCAES [41]	Medical images	No	No	Arnold cat map	None	Blockchain + cloud
Proposed	CT, MRI, PET	Cross-modal attention fusion; one ROI map per study	Yes (CMAF-RoINet)	Image-specific LSTM + hyper-chaos	Pairing-free ECC, unidirectional, collusion-resistant; session key only	Consortium chain + IPFS; smart-contract consent, revocation, audit

2.5. Critical Discussion and Summary of Gaps

The literature has advanced each building block separately but has not integrated them. Several methodological weaknesses also recur across the field and should be avoided in the proposed work.

1) Most image-encryption papers report entropy close to 8, NPCR above 99.6%, and UACI near 33.4% as evidence of security. Yet schemes with exactly these figures have later been broken by chosen- or known-plaintext attacks [10]–[13], [18]. These metrics show that the cipher image looks random; they do not show that the key cannot be recovered. A credible evaluation should add an explicit chosen-plaintext analysis and, where possible, standard randomness tests on the key stream.

2) Neural networks are designed to be robust to small input changes, which is the opposite of what a key generator needs [18]. Hybrid LSTM–chaos designs [19], [20] reduce this risk by keeping a chaotic core. They still leave open how sender and receiver guarantee bit-identical outputs across hardware with different floating-point behaviour. This is a practical deployment issue that most papers do not address.

3) ROI-based schemes [9], [14], [15] save time by protecting the background lightly. The shape and position of the ROI, and any anatomy left in the background, can still reveal information about the patient or the lesion.

Few papers measure this leakage, and security metrics are usually reported on the whole cipher image rather than separately for each region.

4) Two pairing-free schemes were later shown to have flawed security proofs or to fail CCA security [30], [31]. Any new scheme in this family should state its security model and hardness assumption and give a formal reduction.

5) Blockchain–PRE systems [30] give patients delegation and audit but only for textual records. Image-sharing blockchain systems protect storage and integrity but do not support delegation or revocation. Deep-learning image security work [9], [15]–[19] is evaluated in isolation, without a sharing architecture around it. Taken together, the review identifies the following gaps.

1. Blockchain-based PRE has not been designed for medical images; existing schemes target EHRs and EMRs.
2. Multi-modal studies are encrypted modality by modality, with no shared representation or shared ROI.
3. ROI-selective encryption uses single-modality or thresholded regions and does not exploit cross-modal attention.
4. Many chaos-based key streams are plaintext-independent and have been broken; learned key generators have shown weak sensitivity.
5. Most collusion-resistant unidirectional PRE schemes use pairings, and the pairing-free ones require careful proofs.
6. No reviewed system places consent, re-encryption key issuance, revocation, and auditing for image sharing under smart-contract control with IPFS integrity checks.

The proposed framework addresses these gaps together: CMAF-RoINet provides fused, cross-modal ROI selection; the LSTM–hyper-chaotic generator provides image-specific key streams; the pairing-free elliptic-curve PRE delegates only the session key; and the consortium blockchain with IPFS governs the sharing lifecycle.

3. Methodology

The framework protects a multi-modal imaging study in four layers: CMAF-RoINet finds the diagnostically critical region once from all modalities, an image-specific key stream drives ROI-guided selective encryption, a pairing-free elliptic-curve proxy re-encryption scheme delegates only the 256-bit session key, and a consortium blockchain with IPFS governs consent, revocation, storage and audit.

3.1 Overall Architecture

The image never passes through the proxy: the encrypted study goes straight from the gateway to IPFS, while only the 256-bit session key travels the delegation path, and every consent and access decision is recorded on-chain. Figure 1 shows the four layers. Layer 1 perception fuses all modalities once and outputs a single ROI mask Ω. Layer 2, protection draws a fresh session key K_s, derives an image-specific key stream, encrypts ROI voxels strongly and background voxels with one stream pass, and seals K_s in a capsule CT_2 under the patient's public key. Layer 3 delegation lets the patient issue a re-encryption key r_k so the semi-trusted proxy can turn CT_2 into CT_1 for one clinician without learning K_s. Layer 4 governance stores cipher-text on IPFS and keeps consent, key digests, access decisions, revocations and the audit trail on a consortium blockchain. The highlighted capsule is the only public-key operation on the data path, which is why delegation cost is independent of study size.

3.2 System Model

Seven entities take part, and only the certificate authority is fully trusted. The proxy, the storage network and the communication channel are all treated as potentially adversarial.

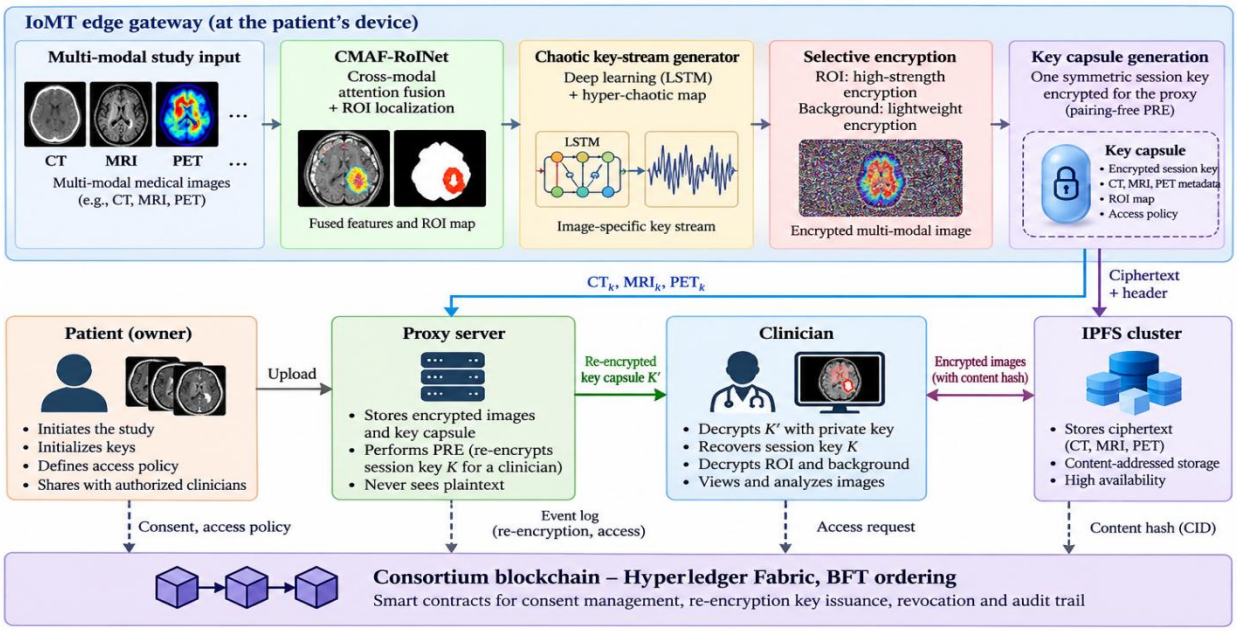


Figure 1. System architecture of the proposed framework

3.2.1 Entities

The gateway acts for the patient because the patient's own device may be too weak to encrypt a full study. The patient's private key never leaves the patient's device: the gateway encrypts under the patient's public key, and only the patient signs consent and generates re-encryption keys. Table 2 shows the entity model.

Table 2. Entity and trust model

Entity	Symbol	Role	Trust
Certificate authority / membership service	CA	Generates public parameters, enrolls every entity with an X.509 identity and an EC key pair	Fully trusted
Patient (data owner)	DO	Owns the study, holds the delegator key pair, grants and revokes consent	Honest
IoMT edge gateway	GW	Acquires CT/MRI/PET, runs CMAF-RoINet and selective encryption on the patient's behalf, uploads ciphertext	Honest, resource-constrained
Data user (clinician)	DU	Requests access, decrypts the re-encrypted session key and the study	Honest while authorized; may collude after revocation
Proxy server	PS	Holds re-encryption keys and transforms the session-key capsule	Semi-trusted (honest-but-curious)
Consortium blockchain	CB	Hospitals' peers and ordering service; runs smart contracts for consent, key issuance, revocation and audit	Fewer than one-third of ordering nodes Byzantine
IPFS network	IPFS	Stores encrypted studies off-chain under content identifiers	Untrusted for confidentiality; tampering is detectable

3.2.2 Threat model

The adversary $\mathcal{A}$ follows the Dolev–Yao model on the network: it can eavesdrop on, modify, delay, replay and inject any message. In addition:

1. **Curious proxy.** PS executes the protocol correctly but tries to learn the session key or image from what it stores and sees.
2. **Proxy–user collusion.** PS may collude with one or more revoked or authorized clinicians to recover the patient's private key or to decrypt studies not delegated to them.

3. **Storage adversary.** IPFS nodes may read, delete or alter cipher-texts.
4. **Chosen-plaintext adversary.** $\mathcal{A}$ may obtain encryptions of images of its choice, as in the attacks reviewed.
5. **Impersonation and replay.** $\mathcal{A}$ may pretend to be a clinician or the patient, or resend old requests and consent transactions.

3.2.3 Design goals

1. **Confidentiality:** only the patient and currently authorized clinicians can recover the study; the proxy and storage learn nothing about it.
2. **Integrity:** any alteration of the cipher-text, header or consent record is detected.
3. **Unidirectional, non-interactive delegation:** the patient delegates using only the clinician's public key, and a key from patient to clinician cannot be turned around.
4. **Collusion resistance (master-secret security):** the proxy and any set of clinicians cannot recover the patient's full private key.
5. **Revocation:** a revoked clinician cannot obtain access to studies or session keys issued after revocation.
6. **Accountability:** every consent, issuance, re-encryption and revocation is recorded on an immutable ledger.
7. **Lossless recovery:** an authorized user reconstructs every modality bit-exactly.
8. **Efficiency:** no bilinear pairings, and public-key cost independent of image size and number of modalities.

3.3 Notation

Table 3 defines every symbol used in the below given section. Bold lowercase letters denote vectors; calligraphic letters denote sets; ‖ denotes concatenation and ⊕ bitwise XOR.

Table 3. Notation

Symbol	Meaning	Domain / type
M	Number of modalities in a study (e.g. CT, MRI sequences, PET)	$\mathbb{N}$
$X^{(\mathrm{m})}$	Co-registered volume of modality m	$\mathbb{Z}_{\{2^b\}}^{\{H\times W\times D\}}$
$\mathcal{S}$	Study: $\{X^{(1)}, \dots, X^{(\mathrm{M})}\}$	Set of volumes
H, W, D	Height, width, depth (slices) of each volume	$\mathbb{N}$
b	Bit depth per voxel (8 for PET/converted, 16 for CT/MRI)	{8, 16}
v	Voxel index	$\{1, \dots, HWD\}$
E_m	Encoder for modality m	Neural network
F_m	Token features of modality m	$\mathbb{R}^{\{N\times C\}}$
N, C	Number of tokens, channels per token	$\mathbb{N}$
W_Q^m, W_K^m, W_V^m	Query, key, value projections of modality m	$\mathbb{R}^{\{C\times d\}}$
d	Attention dimension	$\mathbb{N}$
$A_{\{m\leftarrow n\}}$	Information modality m draws from modality n	$\mathbb{R}^{\{N\times d\}}$
$\gamma_{\{mn\}}$	Learnable weight of the $n \rightarrow m$ attention path	$\mathbb{R}$
α_m	Modality gate weight	$(0,1), \Sigma\alpha_m = 1$
Z	Fused cross-modal representation	$\mathbb{R}^{\{N\times C\}}$
P	Voxel-wise ROI probability map	$[0,1]^{\{H\times W\times D\}}$
T	ROI decision threshold	$(0,1)$
P	Safety-margin dilation radius	Voxels
Ω	Binary ROI mask after thresholding and dilation	$\{0,1\}^{\{H\times W\times D\}}$
$\mathcal{V}_R, \mathcal{V}_B$	ROI voxel set and background voxel set	Index sets
Y	Ground-truth lesion mask	$\{0,1\}^{\{H\times W\times D\}}$
α_T, β_T	Tversky false-positive and false-negative weights	$\alpha_T + \beta_T = 1$
λ_1, λ_2	Loss weights	$\mathbb{R}_+$
$\mathbb{E}, p$	Elliptic curve over prime field $\mathbb{F}_p$	Curve (e.g. secp256r1)
G, q	Base point and its prime order	$G \in \mathbb{E}(\mathbb{F}_p)$, q prime

$\mathbb{Z}_q^*$	Non-zero integers modulo q	Scalars
H_0 … H_5	Hash functions modelled as random oracles	See methodology
(x_i, y_i)	Two-part private key of user i	$\mathbb{Z}_q^* \times \mathbb{Z}_q^*$
(X_i, Y_i)	Public key of user i: $X_i = x_{iG}, Y_i = y_{iG}$	$\mathbb{E}^2$
s_i, S_i	Delegation secret $s_i = x_i \cdot H^1(Y_i) + y_i$ and its public form $S_i = s_{iG}$	$\mathbb{Z}_q^*, \mathbb{E}$
K_s	Session key of a study	$\{0,1\}^{256}$
h_P	Hash of the plaintext study, SHA3-256($\mathcal{S}$)	$\{0,1\}^{256}$
K	Key-stream seed derived from K_s and h_P	$\{0,1\}^{512}$
(a, b', c, d', r_c)	Hyper-chaotic system parameters (b', d' primed to avoid clash with bit depth b and attention dimension d)	$\mathbb{R}$
$\xi(t) = (\xi_1, \xi_2, \xi_3, \xi_4)$	Hyper-chaotic state	$\mathbb{R}^4$
$\Delta t, N_0$	Integration step and number of discarded transient steps	$\mathbb{R}_+, \mathbb{N}$
L	LSTM input window length	$\mathbb{N}$
u_t	Quantized LSTM output at step t	$\mathbb{Z}_{\{2^b\}}$
z_t^R, z_t^B	ROI and background key-stream words	$\mathbb{Z}_{\{2^b\}}$
Π	Chaotic permutation of ROI voxels	Permutation of $\mathcal{V}_R$
ℓ	Number of permutation–diffusion rounds on the ROI	$\mathbb{N}$
R	Encryption randomness for the session-key capsule	$\mathbb{Z}_q^*$
(R, C_k)	Second-level (re-encryptable) capsule of K_s	$\mathbb{E} \times \{0,1\}^{256}$
t, e	Delegation randomizer and ephemeral scalar for re-encryption key	$\mathbb{Z}_q^*$
$rk_{\{i \to j\}}$	Re-encryption key (rk_1, W_1, W_2) from user i to user j	$\mathbb{Z}_q^* \times \mathbb{E} \times \mathbb{Z}_q$
(R', C_k, W_1, W_2)	First-level (re-encrypted) capsule for user j	$\mathbb{E} \times \{0,1\}^{256} \times \mathbb{E} \times \mathbb{Z}_q$
Hdr	Encrypted header: h_P, compressed Ω, parameters	AEAD cipher-text
CID	IPFS content identifier of the encrypted study	Multi-hash
η, ts	Nonce and timestamp in each transaction	$\{0,1\}^{128}$, Unix time

3.4 Preliminaries

The cryptographic layer rests on two standard elliptic-curve hardness assumptions in the random oracle model, and the key-stream layer on a four-dimensional hyper-chaotic system.

3.4.1 Elliptic-curve group

Let $\mathbb{E}$ be an elliptic curve over a prime field $\mathbb{F}_p$ with a base point G of large prime order q; the framework uses secp256r1 (NIST P-256), which gives about 128-bit security. Scalar multiplication $k_G = G + G + \ldots + G\ (k - times)$ is the only expensive group operation. No bilinear map e: $\mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_T$ is used anywhere.

Definition 1 (ECDLP). Given (G, aG) for random $a \in \mathbb{Z}_q^*$, computing a is infeasible for any probabilistic polynomial-time (PPT) algorithm.

Definition 2 (ECCDH). Given (G, aG, bG) for random $a, b \in \mathbb{Z}_q^*$, computing abG is infeasible for any PPT algorithm. The advantage of $\mathcal{A}$ is:

$$Adv_A^{ECCDH} = Pr[A(G, aG, bG) = abG] \quad (1)$$

which is assumed negligible in the security parameter $\lambda = 128$.

Definition 3 (Modified ECCDH, mECCDH). Given $(G, aG, a^{-1}G, bG)$, computing abG is infeasible. This variant, analogous to the modified CDH assumption used for pairing-free PRE is needed because the re-encryption key involves the inverse of the delegation secret.

3.4.2 Hash functions

Six independent hash functions are modelled as random oracles; in implementation they are domain-separated instances of SHA3-256 / SHA3-512 (prefix byte $0 * 00\ to\ 0 * 05$).

Table 4. Function mapping

Function	Mapping	Uses
H_0	$\{0,1\}* \rightarrow \{0,1\}^{512}$	Key-stream seed κ
H_1	$\mathbb{E} \rightarrow \mathbb{Z}_q^*$	Delegation secret s_i
H_2	$\mathbb{E} \rightarrow \{0,1\}^{256}$	Masking $K_s$in the capsule
H_3	$\mathbb{E} \rightarrow \mathbb{Z}_q$	Masking the delegation randomizer t
H_4	$\{0,1\}* \rightarrow \mathbb{Z}_q^*$	Schnorr proof and Fujisaki–Okamoto check
H_5	$\{0,1\}* \rightarrow \{0,1\}^{256}$	Transaction digests and first-level ECIES

3.4.3 Hyper-chaotic system

The key stream uses the four-dimensional hyper-chaotic Chen system obtained by state-feedback control:with $a = 35, b' = 3, c = 12, d' = 7 \; and \; r_c \in (0.085, 0.798]$, the range reported for hyper-chaotic behaviour (two positive Lyapunov exponents). The system is integrated with fourth-order Runge–Kutta at a fixed step $\Delta t = 0.001$, and the first $N_0 = 1000$ steps are discarded to remove dependence on the transient. The Lyapunov spectrum should be recomputed numerically for the chosen r_c and reported with the results.Two positive Lyapunov exponents mean nearby trajectories diverge in two directions, so the trajectory is harder to reconstruct from partial observations than a one-dimensional map such as the logistic map. The four real initial conditions also enlarge the key space.

3.5 CMAF-RoINet: Cross-Modal Attention Fusion and ROI Network

CMAF-RoINet computes one ROI mask Ω per study from all modalities together. The mask decides which voxels receive strong encryption. The network's fused representation Z is used only to compute Ω; it is never transmitted, and encryption is applied to the original voxel values, which is what makes lossless recovery possible.

3.5.1 Pre-processing (network input only)

All modalities are rigidly co-registered to a reference modality (T1ce for MRI studies, or MRI for MR–PET and MR–CT pairs) by maximizing mutual information, and resampled to 1 mm isotropic voxels. Network inputs are then normalized per modality: z-score within the brain mask for MRI and PET, and window clipping followed by min–max scaling for CT. These normalized copies feed the network; the encrypted payload is the original, un-normalized $X^{(m)}$ in the co-registered grid.

3.5.2 Modality-specific encoders

Each modality has its own encoder E_m (3D convolutional stem with residual blocks, total stride 16), because CT, MRI and PET differ in intensity physics and should not share low-level filters. The encoder returns bottleneck tokens and S skip features:

$$\left(F_m^0, \left\{f_m^{(s)}\right\}_{s-1}^{S}\right) = E_m\left(Norm\left(X^{(m)}\right)\right); \quad F_m^0 \in R^{N*C}; N = \frac{HWD}{16^3} \tag{2}$$

A multi-head self-attention block then captures long-range context within each modality:

$$F_m = F_m^0 + MHSA\ (LN(F_m^0)) \tag{3}$$

Where,LN is layer normalization.

3.5.3 Cross-modal attention

Each modality queries every other modality, so that, for example, the T1ce token over an enhancing tumour can pull in the FLAIR evidence of surrounding oedema or the PET evidence of metabolic activity:

$$Q_m = LN\ (F_m)W_Q^m; \;\; K_n = LN(F_n)W_K^n; \;\; V_n = LN(F_n)W_V^n \tag{4}$$

$$A_{m\rightarrow n} = softmax\left(\frac{Q_m K_n^T}{\sqrt{d}}\right)V_n \tag{5}$$

$$\bar{F}_m = F_m + \sum_{n \neq m} \gamma_{mn} A_{m \to n} W_0 + FFN(LN(F_m)) \tag{6}$$

Here $Q_m, K_n \; and \; V_n \; are$ the query, key and value matrices; the softmax row for token i gives how much token i of modality m attends to each token of modality n; $W_O \in \mathbb{R}^{\{d \times C\}}$ projects back to C channels; FFN is a two-layer feed-forward network; and γ_{mn} is a learnable scalar initialized to 0, so training starts from the single-modality features and adds cross-modal information only where it helps.

3.5.4 Gated fusion

Modalities are not equally informative for every study (PET may be noisy, a sequence may be missing), so their contributions are weighted by a learned gate:

$$g_m = w_g^T \emptyset\big(W_g GAP(\bar{F}_m)\big); \quad \alpha_m = \frac{exp(g_m)}{\sum_{n-1}^{M} exp(g_n)} \tag{7}$$

$$Z = \sum_{m-1}^{M} \propto_m \bar{F}_m \tag{8}$$

GAP is global average pooling over tokens, φ is the GELU activation, and W_g, w_g are learnable. If a modality n is absent, its terms $A_{\{m \leftarrow n\}}$ are dropped and g_n is set to $-\infty, so \; \alpha_n = \; 0$ and the remaining weights renormalize.

3.5.5 ROI head and mask

A decoder up samples$Z$, using gated skip features $\Sigma_m \alpha_m f_m^{\{(s)\}}$ at each scale, to a voxel-wise probability map:

$$P = \; \sigma\left(Dec\left(Z, \left\{\sum_m \alpha_m f_m^{(s)}\right\}_{s=1}^{S}\right)\right) \in [0,1]^{H*W*D} \tag{9}$$

The binary mask is obtained by thresholding and morphological dilation with a ball of radius ρ voxels:

$$\Omega = D_\rho(1[P > -\tau]); \;\; V_R = \{v: \Omega(v) = 1\}; \;\; V_B = \{v: \Omega(v) = 0\} \tag{10}$$

Security-oriented operating point

A segmentation network is usually tuned to maximize Dice. Here a missed lesion voxel is worse than an extra one, because a missed voxel receives only background-level protection. τ is therefore chosen on the validation set as the largest threshold that keeps ROI recall at or above 0.99, and the margin ρ adds a further safety band around the lesion boundary.

3.5.6 Training objective

The Tversky loss weights false negatives more heavily than false $positives$ ($\beta_T > \alpha_T, e.g. \alpha_T \; = \; 0.3, \beta_T =$ 0.7), matching the recall-first objective above:

$$L = L_{TV} + \; \lambda_1 L_{BCE} + \lambda_2 L_{cons} \tag{11}$$

$$L_{TV} = 1 - \frac{\sum_v P_v Y_v + \in}{\sum_v P_v Y_v + \; \alpha_T \sum_v P_v (1 - Y_v) + \; \beta_T \sum_v (1 - P_v) Y_v + \epsilon} \tag{12}$$

Here, $\mathcal{L}_{BCE}$ is voxel-wise binary cross-entropy, which stabilizes early training, and ε is a small constant that prevents division by zero. The consistency loss makes the mask robust to missing modalities. During training each modality is dropped with probability p^{drop}, giving a prediction $P^{\{drop\}}$, and

$$L_{cons} = \frac{1}{HWD}\sum_{v}\left(P_v^{full}{}_{P_v}^{drop}\right)^2 \quad (13)$$

Datasets and labels

BraTS provides co-registered T1, T1ce, T2 and FLAIR volumes with expert tumour labels and is used for supervised training and for measuring ROI recall. For those pairs, the MRI encoders are initialized from BraTS training and the CT/PET encoders are fine-tuned with the consistency loss alone, and ROI coverage is reported qualitatively and by radiologist review.

3.6 LSTM-Driven Hyper-Chaotic Key-Stream Generator

The generator makes every key stream unique to one study and one session key, so the single-image chosen-plaintext attacks do not apply. It combines three sources: hyper-chaos (high-dimensional, seed-sensitive dynamics), an LSTM innovation filter (removes the strong correlation between neighbouring chaotic samples), and a ChaCha20 stream (a standard cryptographic security floor).

3.6.1 Session key and seed

For every study the gateway draws a fresh session key from a cryptographically secure random number generator and hashes the plaintext study together with its geometry:

$$K_s \to \{0,1\}^{256}; h_p = SHA3-256(X^{(1)}\Big|\Big|\dots\Big|\Big|X^{(M)}\Big|\Big|H\Big|\Big|W||D||^b) \quad (14)$$

$$k = H_0\left(K_s\Big|\Big|h_p\Big|\Big|studyID\right) = k_A||k_B;\quad k_A\epsilon\{0,1\}^{256};\quad k_B = (k_1,k_2,k_3,k_4) \quad (15)$$

Here, κ_A keys the ChaCha20 stream and κ_B sets the chaotic initial conditions. Because h_P is inside the seed, changing a single voxel in any modality changes the whole of κ and therefore the whole key stream. This plaintext dependence is what the crypt-analysed schemes. The receiver obtains h_P from the encrypted header, so it can regenerate κ.

3.6.2 Hyper-chaotic trajectory

The four 64-bit words of κ_B select the initial state and the control parameter inside the hyper-chaotic region:

$$\xi_j(0) = \xi_j^0 + \delta_j\frac{k_j}{2^{64}}; j = 1,\dots,4;\quad r_c = 0.1 + 0.6.\frac{k_1 \oplus k_2 \oplus k_3 \oplus k_4}{2^{64}} \quad (16)$$

Here, ξ° is a fixed public point on the attractor and δ_j are fixed perturbation widths (e.g. 1.0) that keep the start inside the basin of attraction; r_c is kept inside $[0.1, 0.7]$ so it stays well within the hyper-chaotic range. The system is integrated with RK4, the first N_0 steps are discarded, and each state is quantized to a b-bit word per component:

$$q_j(t) = \left[\quad\left|\xi_j(t).10^{14}\right|\right] mod\ 2^b;\quad j = 1,\dots,4 \quad (17)$$

The factor 10^{14} takes the low-order decimal digits of the state, which are the most sensitive to the initial condition.

Bit-exact reproducibility

Sender and receiver must produce identical streams. Integration therefore uses IEEE 754 binary64 arithmetic with a fixed operation order and fused multiply–add disabled, which gives identical results on any IEEE-754-compliant CPU (x86-64 or ARMv8).

3.6.3 LSTM innovation filter

At step $\Delta t = 0.001$ consecutive chaotic samples are almost identical, so raw quantized states are strongly auto-correlated. An LSTM is trained offline to predict the next state from the previous L states; at run time the prediction residual is kept, because it is the part of the trajectory that cannot be predicted from the recent past.

With input window $w_t = (\xi(t-L+1), \ldots, \xi(t))$, where ξ is the state scaled to $[-1, 1]$, the LSTM cell computes:

$$i_t = \sigma(W_i w_t + U_i h_{t-1} + b_i);\ f_t = \sigma(W_f w_t + U_f h_{t-1} + b_f) \tag{18}$$

$$o_t = \sigma(W_0 w_t + U_o h_{t-1} + b_o);\ \widetilde{c_t} = tanh(W_c w_t + U_c h_{t-1} + b_c) \tag{19}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot c_t;\ h_t = o_t \odot tanh(c_t) \tag{20}$$

$$\hat{\xi}(t+1);\ u_j(t) = \left[\left|\xi_j(t+1) - \hat{\xi}_j(t+1)\right|.10^{14}\right] mod\ 2^b \tag{21}$$

Here, $i_t, f_t\ and\ o_t$ are the input, forget and output gates, c_t the cell state, $h_t$the hidden state, σ the logistic sigmoid and $\odot$ element-wise multiplication. The LSTM is trained on trajectories from random initial conditions with a mean-squared one-step prediction loss. Its weights are public (Kerckhoffs's principle): the secret is only the seed. For bit-exact output the deployed network uses integer-only 8-bit quantized inference with 32-bit accumulators.

3.6.4 Key-stream combiner

The ROI key-stream word at index t combines all three sources; the background uses ChaCha20 alone with a different nonce:

$$z_t^R = q(t\ mod\ 4) + 1\left(\left[\frac{t}{4}\right]\right) \oplus u_{(t\ mod\ 4)}\left(\left[\frac{t}{4}\right]\right) \oplus ChaCha20_{k_A}\ (nonce = 0, t) \tag{22}$$

$$z_t^B = ChaCha20_{k_A}\ (nonce = 1, t) \tag{23}$$

where, ChaCha20 output is read in b-bit words.

Lemma 1 (security floor). If ChaCha20 is a pseudorandom function and H_0 is a random oracle, then the streams $\{z_t^R\}\ and\ \{z_t^B\}$ are computationally indistinguishable from uniform to any adversary who does not know K_s.

Proof sketch. In the random oracle model $\kappa_A\ and\ \kappa_B$ are independent uniform strings. The chaotic and LSTM terms depend only on κ_B, so they are independent of the ChaCha20 term, and XORing any independent string with a pseudorandom string yields a pseudorandom string. Lemma 1 means the cipher is never weaker than a standard stream cipher, even if the chaotic dynamics later turn out to have exploitable structure, which has happened to many chaos-only designs. Chaos and the LSTM add plaintext-dependent, high-dimensional dynamics on top of this floor and drive the ROI permutation.

Key space. The effective key space is 2^{256}, set by K_s; the chaotic initial conditions are derived from it and do not add independent entropy. This value should be reported instead of the very large figures sometimes obtained by counting floating-point digits of each initial condition.

3.7 ROI-Guided Selective Encryption

Every voxel of every modality is encrypted. ROI voxels receive ℓ rounds of joint cross-modal permutation and bidirectional diffusion; background voxels receive a single stream-cipher pass. Both are bijectionsmodulo 2^b, so decryption is exactly lossless. The mask itself is sent only inside an authenticated, encrypted header, so the lesion's shape and location are not revealed.

3.7.1 ROI: joint cross-modal permutation and diffusion

The ROI voxels of all modalities are gathered, in raster order, into one vector, so that a single pass protects the whole study and the permutation scatters voxels across modalities as well as positions:

$$p = \left(X^{(1)}(v)\right)_{v \in V_R} \Big|\Big| \left(X^{(M)}(v)\right)_{v \in V_R};\ n = M\Big|V_R\Big| \tag{24}$$

Each of ℓ rounds (ℓ = 2 by default) applies three steps.

(a) Permutation. A Fisher–Yates shuffle driven by z^R produces a permutation $\pi_k \, of \, \{1, \dots, n\}$; rejection sampling removes modulo bias. Then $p' = \pi_k(p)$.

(b) Forward diffusion. Each output depends on all earlier voxels:

$$c_i = (p'_i + z_i^R + c_{i-1}) \, mod \, 2^b; \quad i = 1, \dots, n; \quad c_0 = z_0^R \tag{25}$$

(c) Backward diffusion. Each output then also depends on all later voxels:

The output d is the input to the next round, and key-stream indices continue across rounds so no word is reused. After two rounds, a change in any single ROI voxel of any modality affects every ROI voxel of every modality.

$$d_n = c_n \oplus z_{n+1}^R; \;\; d_i = c_i \oplus \left((d_{i+1} + z_{n+1+i}^R) mod \, 2^b\right); ; \;\; i = n-1, \dots, 1 \tag{26}$$

3.7.2 Background: single stream pass

By Lemma 1 this is as strong as ChaCha20 itself: the background is not weakly protected, only protected with fewer operations. The saving comes from skipping permutation and diffusion, which are the costly, sequential steps.

$$C^{(m)}(v) = X^{(m)}(v) \oplus z^B_{(m-1)HWD+v}; \quad v \in V_B, \qquad m = 1, \dots, M \tag{27}$$

3.7.3 Encrypted header

The receiver needs h_P (to regenerate κ), the mask Ω (to know which voxels are ROI) and the parameters. These are sealed with AES-256-GCM under a header key derived from K_s alone, so the header can be opened before h_P is known:

$$K_{hdr} = HKDF\,(K_s, \qquad hdr); \;\; Hdr = AES - GCM_{K_{hdr}}(v, h_p || RLE(\Omega) || (M, l, b, \tau, \rho); \;\; AD = studyID || H || \tag{28}$$

Here, v is a 96-bit nonce, RLE is run-length encoding (masks compress well because they are mostly zeros), and AD is authenticated but unencrypted data. The GCM tag also authenticates the study identifier and dimensions. The full ciphertextis $\mathcal{C} = (Hdr, C^{(1)}, \dots, C^{(M)})$, where each $C^{(m)}$ has the same size as $X^{(m)}$, so there is no expansion beyond the header.

3.7.4 Decryption and lossless recovery

Given K_s, the receiver:

1. derives K_{hdr}, opens Hdr and verifies its tag, recovering $h_P \; and \; \Omega$;
2. computes $\kappa = H^0(K_s \| h_P \| \, studyID)$ and regenerates $z^R \; and \; z^B$;
3. undoes the background pass by XOR;
4. for k = ℓ, …, 1, inverts backward diffusion, then forward diffusion, then the permutation:
5. scatters p back into the ROI positions of each modality;
6. Re-computes SHA3-256 of the recovered study and checks it against h_P.

$$c_i = d_i \; \oplus \left(\left(d_{i+1} + z^R_{n+1\pm}\right) mod \, 2^b\right); \;\; p'_i = (c_i - z_i^R - c_{i-1}) \, mob \, 2^b; p = \pi_k^{-1}(p') \tag{29}$$

Step 6 is an end-to-end integrity check: a correct key with a corrupted cipher-text, or a wrong key is detected. Because XOR, modular addition and permutation are all bijectionson $\mathbb{Z}_{\{2^b\}}$, recovery is bit-exact, so $PSNR = \infty$ and $SSIM = 1$ for every authorized receiver.

3.8 Pairing-Free Unidirectional ECC Proxy Re-Encryption

Only the 256-bit session key K_s is ever public-key encrypted or re-encrypted, so delegation costs a few elliptic-curve scalar multiplications whatever the size of the study. Each user holds a two-part private key; delegation exposes at most a derived secret s_i, never the full key, which is the basis of collusion resistance.

3.8.1 Algorithms

Setup(λ). The CA publishes params = $(\mathbb{E}, p, q, G, H_0, \ldots, H_5)$ for security parameter $\lambda = 128$.

KeyGen(i). User i chooses $x_i, y_i \in \mathbb{Z}_q^*$ at random and computes:

$$X_i = x_i G; \quad Y_i = y_i G; \quad s_i = x_i H_1(Y_i) + y_i \bmod q; \quad S_i = H_1(Y_i) X_i + Y_i = s_i G \tag{30}$$

The public key is $pk_i = (X_i, Y_i)$; S_i is computable by anyone from pk_i. The private key is $sk_i = (x_i, y_i)$. At enrolment, user i proves knowledge of x_i *and* y_i with Schnorr proofs, which the CA checks before issuing a certificate; this blocks rogue-key attacks.

Encrypt(pk_i, K_s) → second-level capsule. Choose $r, \omega \in \mathbb{Z}_q^*$ at random and compute:

$$R = rS_i; \quad \Gamma = rG, \quad C_k = K_s \oplus H_2(\Gamma); \quad T = H_5(\Gamma || K_s || C_k) \tag{31}$$

$$A = \omega S_i; \quad h = H_4(R || C_k || T || A); \quad z = \omega + hr \bmod q \tag{32}$$

Output $CT_2 = (R, C_k, T, h, z)$. Γ is the shared point that masks K_s; T is a validity tag that binds Γ, K_s and C_k; (h, z) is a Schnorr proof of knowledge of r with respect to base S_i, which binds C_k and T *to* R so that the capsule cannot be modified without detection.

Verify(pk_i, CT_2). Compute $A' = zS_i - hR$ and accept only if $h = H^4(R \| C_k \| T \| A')$. Both the proxy and the patient run this check.

ReKeyGen(sk_i, pk_j) → $rk_{\{i \to j\}}$. The patient chooses a fresh delegation randomizer t and ephemeral e in $\mathbb{Z}_q^*$ and computes:

$$rk_1 = t.s_i^{-1} \bmod q; \quad W_1 = eG; \quad W_2 = t + H_3(eS_j) \bmod q \tag{33}$$

Here, $rk_{\{i \to j\}} = (rk_1, W_1, W_2)$ goes to the proxy over a TLS channel, and only its digest $H_5(rk_{\{i \to j\}})$ is written on-chain. W_1andW_2 wrap t so that only user j can unwrap it; the proxy, holding rk_1, cannot learn t.

ReEncrypt($rk_{\{i \to j\}}, CT_2$) → first-level capsule. If Verify fails, output ⊥. Otherwise compute:

$$R' = rk_1.R = ts_i^{-1} rs_i G = trG \tag{34}$$

and output $CT_1 = (R', C_k, T, W_1, W_2)$. This is one scalar multiplication.

Decrypt₂(sk_i, CT_2) (patient opens own capsule). After Verify,

$$\Gamma = s_i^{-1} R = rG; \quad K_s = C_k \oplus H_2(\Gamma); \quad accept\ iff T = H_5(\Gamma || K_s || C_k) \tag{35}$$

Decrypt₁(sk_j, CT_1) (clinician opens re-encrypted capsule).

$$t = W_2 - H_3(s_j W_1) \bmod q; \quad \Gamma = t^{-1} R'; \quad K_s = C_k \oplus H_2(\Gamma); \quad accept\ iff T = H_5(\Gamma || K_s || C_k) \tag{36}$$

3.8.2 Correctness

For the clinician: $s_j W_1 = s_j \cdot eG = e \cdot S_j, so\ H_3(s_j W_1) = H_3(eS_j)$ and t is recovered exactly. Then $t^{-1}R' = t^{-1} \cdot trG = rG = \Gamma$, so the clinician derives the same mask $H_2(\Gamma)$ and recovers K_s. For the patient, $s_i^{-1} R = s_i^{-1} \cdot rs_i G = rG = \Gamma$.In both cases the tag T is recomputed from the same (Γ, K_s, C_k) and matches.

3.8.3 Security properties

Theorem 1 (capsule confidentiality, IND-CPA). Given pk_i and CT_2, an adversary who can distinguish K_s from random can be used to solve ECCDH, with H_2 modelled as a random oracle.

Proof sketch. Computing $\Gamma = rG\ from\ S_i = s_i G\ and\ R = rs_i G$ is the divisible Diffie–Hellman problem, which is equivalent to ECCDH. Unless $\mathcal{A}$ queries H_2at Γ, the mask $H_2(\Gamma)$ is uniform and C_k hides K_s perfectly; a simulator that embeds an ECCDH instance in (S_i, R) extracts Γ from $\mathcal{A}'s$ H_2 queries.

Theorem 2 (proxy learns nothing). The proxy holds $rk^1 = \frac{t}{s_i}$and sees R. From these it can compute $rk^1 S_i = tG\ and\ R' = trG$, but obtaining $\Gamma = rG\ from\ (tG, trG)$ is again divisible Diffie–Hellman; t itself is masked by $H^3(eS_j)$, which requires ECCDH on (W^1, S_j). So K_s remains hidden from the proxy under ECCDH in the random oracle model.

Theorem 3 (collusion resistance, master-secret security). If the proxy and clinician j collude, they learn $t\ and\ rk_1$ and hence $s_i = t \cdot rk_1^{-1}$. Recovering $sk_i = (x_i, y_i)$ from s_i and pk_i is as hard as ECDLP.

Proof sketch.$s_i = x_i H_1(Y_i) + y_i$is one linear equation in two unknowns. Any x_i is consistent with it for $y_i = s_i - x_i H_1(Y_i)$, so finding the true x_i requires solving $X_i = x_i G$, which is ECDLP. Data the patient encrypts under X_i alone (first-level ECIES, used for the patient's private records and keys) therefore stays confidential even after collusion.

Chosen-ciphertext security. The Schnorr proof makes second-level capsules non-malleable and lets the proxy refuse malformed inputs; the tag T lets the receiver reject altered capsules. Together these follow the pattern used for CCA-secure pairing-free PRE, but a full IND-CCA proof for this construction remains to be written and should be included before publication rather than claimed. Some properties are:

1. **Unidirectional:**$rk_{\{i \to j\}}$ contains s_i^{-1} and wraps $t$under S_j; it cannot be turned into $rk_{\{j \to i\}}$, which would need s_j.
2. **Non-interactive:**ReKeyGen uses only the clinician's public key.
3. **Single-hop (non-transitive):**CT_1 carries no Schnorr proof with respect to S_j, so ReEncrypt rejects it as input and the proxy cannot forward the study further.
4. **Constant cost:** every algorithm acts on one 256-bit key, independent of the image size and of M.

3.8.4 Revocation by delegation-key rotation

Theorem 3 implies that a colluding proxy and revoked clinician learn s_i and could open the patient's second-level capsules, as with any PRE scheme where the proxy colludes with a delegatee. The framework closes this gap for future data: on revocation, the patient rotates the delegation component, choosing a new y_i', publishing Y_i' on-chain, and so obtaining new s_i' and S_i'. New studies are encrypted under S_i', the leaked s_i becomes useless for them, and still-authorized clinicians receive fresh re-encryption keys (one ReKeyGen each). $X_i$and therefore the patient's identity and first-level data is unchanged.

3.9 Consortium Blockchain, Smart Contracts and IPFS

The ledger holds only hashes, identifiers and policy state, never images or keys. The smart contract is the single place where consent is checked, so the proxy re-encrypts only when the ledger says an access is valid, and every decision leaves an immutable record.

3.9.1 Platform

The blockchain is a permissioned Hyperledger Fabric network in which each participating hospital runs endorsing peers and an ordering node. Ordering uses a Byzantine-fault-tolerant protocol, so the ledger stays consistent if fewer than one-third of ordering nodes are faulty or malicious. Identities are X.509 certificates issued by the CA's membership service, and every transaction is signed with the sender's ECDSA key on P-256. The endorsement policy requires a majority of hospital organizations.

3.9.2 Off-chain storage

The encrypted study $\mathcal{C}$ and its capsule CT_2 are both ciphertext and are stored on IPFS, pinned by an IPFS cluster run by the consortium hospitals to guarantee availability:

$$CID = CID_{v1}(SHA2 - 256(C||CT_2)); \quad \theta = H_5\ (C||CT_2) \tag{37}$$

The CID is content-addressed, so a node cannot serve altered data under the same identifier. The independent digest θ is also stored on-chain, and every reader recomputes it after download, which detects tampering even if a different file were substituted.

3.9.3 Ledger state

Table 5. Records and fields

Record	Key	Fields
User	userID	pk = (X, Y), certificate, role, status
Study	studyID	owner, CID, θ, modalities, timestamp
Consent	(studyID, DU)	purpose, expiry, H_5(rk), status ∈ {active, revoked, expired}
Access log	txID	studyID, DU, $H_5(CT_1)$, proxy signature, timestamp
Nonce set	η	timestamp (for replay detection)

3.9.4 Smart-contract functions

1. **RegisterUser(id, pk, cert)**, called by the CA: stores the public key after checking the enrolment Schnorr proofs.
2. **RegisterStudy**$(\boldsymbol{studyID}, \boldsymbol{CID}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{ts}, \boldsymbol{\sigma_{GW}})$, called by the gateway on the patient's behalf.
3. **GrantConsent**$(\boldsymbol{studyID}, \boldsymbol{DU}, \boldsymbol{purpose}, \boldsymbol{expiry}, \boldsymbol{H}_5(\boldsymbol{rk}), \boldsymbol{\eta}, \boldsymbol{ts}, \boldsymbol{\sigma_{DO}})$, called only by the patient: creates an active consent record.
4. **RequestAccess(studyID, DU, η, ts, σ_DU)**, called by the clinician: returns approve or deny (rule below) and on approval emits an AccessApproved event to the proxy.
5. **LogReEncryption**$(\boldsymbol{studyID}, \boldsymbol{DU}, \boldsymbol{H}_5(\boldsymbol{CT}_1), \boldsymbol{\sigma_{PS}})$, called by the proxy after it re-encrypts.
6. **Revoke**$(\boldsymbol{studyID\ or\ DU}, \boldsymbol{\sigma_{DO}})$, called by the patient: sets consent to revoked, emits an event that tells the proxy to delete the matching rk, and accepts the rotated $Y_i^{'}$.
7. **Audit(studyID)**: returns the full history of every record for that study; it is read-only and available to the patient and to authorized auditors.

3.9.5 Access rule

Request Access approves only if all of the following hold:

$$Approve = versig_{pk_{DU}}(\sigma_{DU}) \bigwedge status = active \wedge t_{now} \leq expiry \wedge |t_{now} - ts| \leq \Delta_{max} \wedge \eta \neq N \tag{38}$$

VerSig is ECDSA verification against the clinician's registered key, which defeats impersonation; $\mathcal{N}$ is the stored nonce set and Δ_{max} a freshness window (e.g. 60 s), which together defeat replay. On approval η is added to $\mathcal{N}$. The proxy acts only on AccessApproved events and checks that H_5 of the r_k it holds matches the on-chain digest, so it cannot be tricked into using a stale or substituted key.

3.10 End-to-End Workflow

A study moves through five phases, from enrolment to revocation. Only the upload and access phases touch the image; the others act on keys and ledger records.

Phase 1: Initialization

1. The CA runs Setup and publishes params.
2. Every patient and clinician runs KeyGen, proves knowledge of the key, and is registered with RegisterUser.

Phase 2: Acquisition and upload (gateway, on the patient's behalf).

1. Acquire $X^{(1)}, \ldots, X^{(M)}$ and co-register them.
2. Run CMAF-RoINet to obtain P and the mask Ω.
3. Draw K_s, compute h_P *and* κ, and generate z^R *and* z^B.
4. Encrypt ROI and background and build Hdr, giving $\mathcal{C}$.
5. Run $Encrypt(pk_{DO}, K_s)$ to obtain CT_2, then securely erase K_s and the plaintext from the gateway.
6. Upload $\mathcal{C} \parallel CT_2$ to IPFS, obtain CID, and call RegisterStudy with CID and θ.

Phase 3: Consent

1. The patient reviews a request (or proactively decides) and runs $ReKeyGen(sk_{DO}, pk_{DU})$.

2. The patient sends $rk_{\{DO \to DU\}}$ to the proxy over TLS and calls GrantConsent with $H_5(rk)$, purpose and expiry.

Phase 4: Access.

1. The clinician calls RequestAccess; the contract applies the access rule.
2. On AccessApproved, the proxy fetches CT_2 by CID, checks Verify and the rk digest, runs ReEncrypt, sends CT_1 to the clinician, and calls LogReEncryption.
3. The clinician fetches $\mathcal{C}$ by CID, checks θ, runs Decrypt$_1$ to obtain K_s, and decrypts by finishing with the h_P integrity check.

Phase 5: Revocation.

1. The patient calls Revoke; the contract marks the consent revoked and the proxy deletes the matching rk.
2. The patient rotates $Y_i \; to \; Y_i^{'}$ and reissues re-encryption keys to clinicians who remain authorized. New studies use the new delegation key.

3.11 Computational Complexity

The public-key cost of the proposed scheme is a fixed handful of elliptic-curve scalar multiplications per study and per delegation, with no pairings, while the symmetric cost grows linearly with study size and is reduced in proportion to the background fraction.

3.11.1 Public-key operations

Here, T_{sm} denotes one elliptic-curve scalar multiplication, T_p one bilinear pairing, T_e one exponentiation in the pairing target group, and T_{inv} one modular inversion; hashing is negligible by comparison. The baseline is the unidirectional pairing-based scheme (the standard reference point in the PRE literature), counted from its original construction. Measured timings for both schemes should be reported on the same gateway and server hardware in the results section as in Table 6.

Table 6. Comparison of performance metrics

Operation	Proposed (pairing-free)	Ateniese et al. (pairing-based)
Encrypt (second level)	3 T_{sm}	1 group exponentiation + 1 T_e
Verify (proxy check)	2 T_{sm}	No equivalent check
ReKeyGen	2 T_{sm} + 1 T_{inv}	1 group exponentiation + 1 T_{inv}
ReEncrypt	3T_{sm} (incl. Verify)	1 T_p
Decrypt$_2$ (owner)	3 T_{sm} + 1 T_{inv} (incl. Verify)	1 T_p + 1 T_e
Decrypt$_1$ (delegatee)	2 T_{sm} + 1 T_{inv}	1 T_e
Collusion-resistant	Yes (Theorem 3)	Yes

Pairings and target-group operations work over larger fields than plain elliptic-curve arithmetic, which is why removing them matters most on the proxy, which runs ReEncrypt for every access, and on constrained gateways.

3.11.2 Effect of multi-modality

A modality-independent design runs ROI detection, key generation and public-key encryption once per modality. The proposed design runs them once per study:

$$Cost_{PK}^{indep} = M.(3T_{sm}) \; vs. Cost_{PK}^{proposed} = 3T_{sm} \tag{39}$$

and likewise one ReEncrypt per access instead of M.

3.11.3 Symmetric encryption

Let $\varphi = |\mathcal{V}_R| \; / \; (HWD)$ be the ROI fraction, c_R the per-word cost of ROI encryption (chaos, LSTM, ChaCha20, ℓ rounds of permutation and diffusion), and c_B the per-word cost of background encryption (one ChaCha20 word and one XOR). Then:

$$Cost_{sym} = M.HWD\ [\varphi\ c_R + (1-\varphi)c_B] \tag{40}$$

$$\text{saving} = \frac{\text{M. HWD. c}_R - \text{cost}_{\text{sym}}}{\text{M. HWD. c}_R} = (1-\varphi)\left(1 - \frac{c_B}{c_R}\right) \tag{41}$$

The saving therefore depends on two measurable quantities: the ROI fraction φ in the dataset and the cost ratio $\frac{c_B}{c_R}$ on the target hardware. Both should be reported so readers can see where the reduction comes from.

3.11.4 Network cost

Cross-modal attention at the bottleneck costs $O(M(M-1)\cdot N^2\cdot d), \text{with } N = HWD/16^3$ tokens. Because attention is applied only at the coarsest scale, this term stays small relative to the convolutional encoders. ROI inference runs once per study on the gateway or an edge server.

1. **Security-oriented cross-modal ROI network.** CMAF-RoINet is the first use of cross-modal attention fusion to drive encryption, as opposed to segmentation or diagnosis. It is trained and thresholded for ROI recall, a criterion set by security rather than segmentation accuracy, and stays robust to missing modalities through gated fusion and consistency training.
2. **One ROI map and joint permutation for a whole study.** A single fused mask is shared by all modalities, and ROI voxels from all modalities are permuted and diffused jointly. This removes per-modality redundancy and spreads each voxel's influence across modalities.
3. **Hybrid key-stream generator with a provable floor.** The generator couples plaintext-dependent hyper-chaos with an LSTM innovation filter and a ChaCha20 combiner, and is engineered for bit-exact reproducibility. Unlike earlier chaos–LSTM designs, its security does not rest on the chaotic dynamics alone.
4. **Pairing-free, collusion-resistant PRE for image delegation.** A two-part-key elliptic-curve scheme delegates only the session key, with one scalar multiplication per re-encryption (three including the validity check), and adds delegation-key rotation for revocation.
5. **End-to-end governance for multi-modal image sharing.** Consent, re-encryption key issuance, revocation, replay protection and audit are enforced by smart contracts and tied to the proxy through on-chain key digests, which earlier blockchain–PRE schemes applied only to textual records.

4. Numerical results and discussion

4.1 Experimental Setup

4.1.1 Dataset description

Experiments use a publicly mirrored subset of the BraTS 2018 training set: 495 glioma cases (210 HGG, 285 LGG), each with one axial slice of T1, T1ce, T2 and FLAIR at 240 × 240 pixels, 8-bit, together with the expert segmentation. These four co-registered MRI sequences form the multi-modal "study" (M = 4, b = 8) for every experiment. The whole-tumour label (all tumour classes merged) is the ground truth for ROI detection; tumour pixels make up 2.9% of all pixels, and 89.7% of slices contain tumour. Cases are split at random (seed 0) into 70% training (346), 15% validation (74) and 15% test (75); because each case contributes one slice, no patient appears in two splits.

4.1.2 Implementation

The implementation process is explained here:

Cipher. Python 3.12 with Numba-compiled kernels for the hyper-chaotic integration, LSTM inference, Fisher–Yates permutation and diffusion; ChaCha20, AES-256-GCM and HKDF from the cryptography library; SHA3 from hashlib($\ell = 2\ \text{rounds}, \Delta t = 10^{-3}, N_0 = 1000$).

LSTM innovation filter. A single-layer LSTM (16 hidden units, window $L = 8$) trained on 40 hyper-chaotic trajectories with random initial conditions and $r_c \in [0.1, 0.7]$.

ROI network. A 2D CPU prototype of CMAF-RoINet at 128×128 input, with four modality-specific encoders (8–16–32 channels), single-head cross-modal attention at the bottleneck ($d = 32, \gamma$ initialized to 0), gated fusion and a U-Net decoder. It is compared with a single-modality U-Net (T1ce only) and an early-fusion

U-Net (four channels stacked at the input) of similar size. All three use the same loss (Tversky $\alpha_T = 0.3, \beta_T = 0.7$, plus $0.5 \times$ BCE), modality dropout with probability 0.25 for the multi-modal models, Adam (learning rate 10^{-3}), batch size 16 and 25 epochs. This prototype is much smaller than the 3D network, so its accuracy is a lower bound on what the full design can reach.

Proxy re-encryption. The proposed scheme and the pairing-based baseline are both implemented on the BN254 curve in the pure-Python py_ecc library, so the two are measured with the same arithmetic code; absolute timings of the proposed scheme are also measured on secp256k1 with the C library libsecp256k1 (via coincurve).

Hardware. One virtual core of an Intel Xeon at 2.10 GHz with 3 GB RAM, deliberately close to an edge gateway rather than a server as in Table 7.

Table 7. Measurements and analysis

Measurements	**Analysis**
Histogram, entropy, correlation, NPCR/UACI, key sensitivity, lossless recovery on BraTS	Full 3D CMAF-RoINet on complete BraTS volumes
Key-stream randomness tests	MR–PET and MR–CT studies
Selective-encryption time against ROI fraction	Blockchain throughput and latency on Hyperledger Fabric
PRE operation timings, proposed vs pairing-based	End-to-end network latency between hospitals
ROI prototype: ROC, confusion matrix, Dice, recall, missing-modality ablation	--

4.2 Visual and Histogram Analysis

Every cipher image is visually indistinguishable from noise, decryption with the correct key restores each modality bit for bit, and a key differing in a single bit yields noise again (99.61% of pixels differ from the plaintext). For the security-metric experiments of the ROI is the expert whole-tumour mask dilated by $\rho = 3$ pixels, a standard choice for evaluating selective encryption independently of segmentation quality; 75 test studies (300 images) are encrypted, each under a fresh random 256-bit K_s. The ROI fraction φ ranges from 0.14% to 12.6% (mean 4.75%).

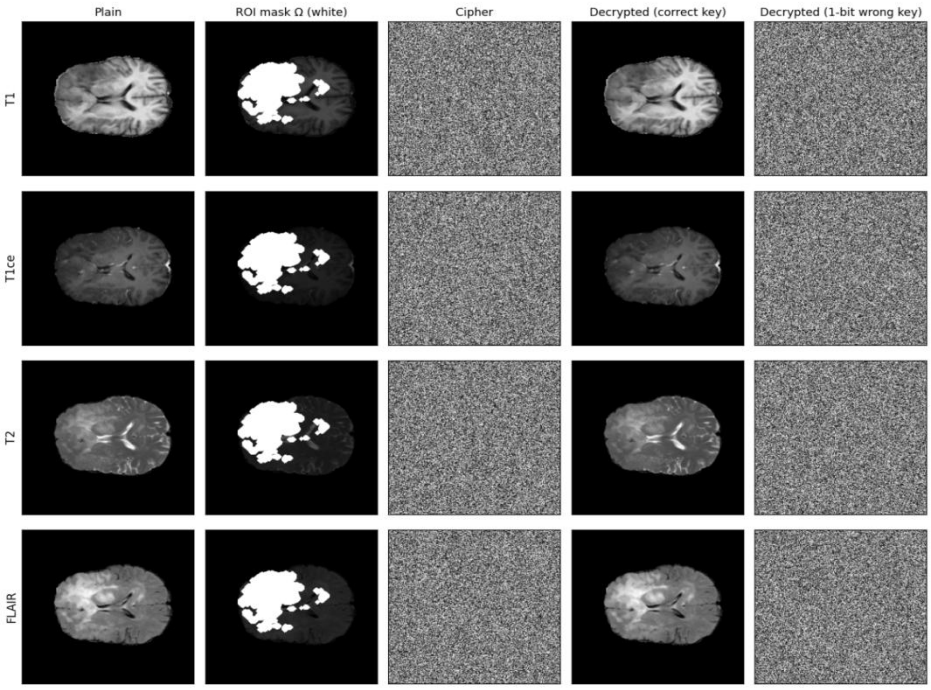


Fig 2. Performance outcomes

The ROI mask is shown here only for illustration: in the transmitted ciphertext it travels inside the AES-GCM header, and the cipher images show no trace of the lesion boundary, because the background is also encrypted (with one ChaCha20 pass rather than left in clear) as in Fig 2.

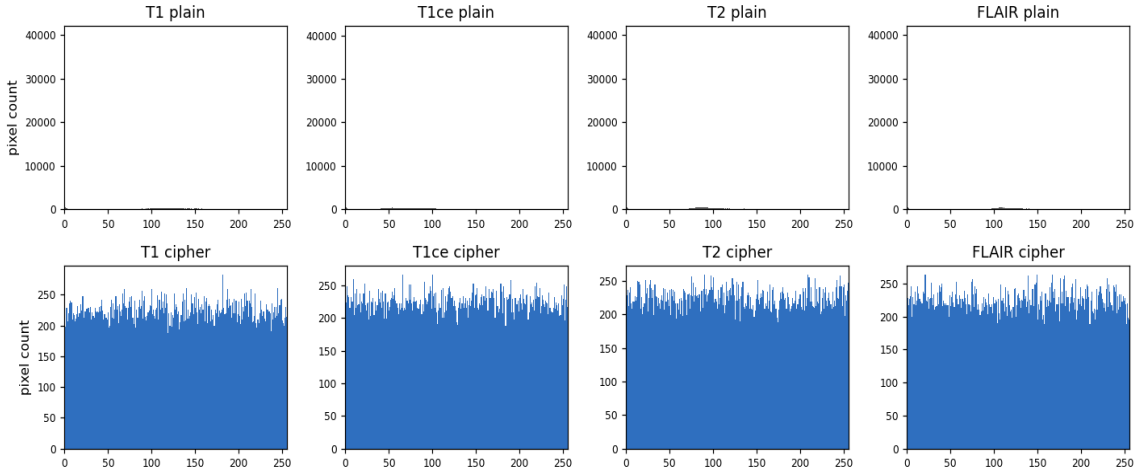


Fig 3. T1 and T2 cipher and FLAIR cipher analysis

A chi-square test with 255 degrees of freedom quantifies flatness. The critical value at $\alpha = 0.05$ is 293.25. Plain images give χ^2between 6.1×10^6 and 8.1×10^6; cipher images give a mean of 255.7 ± 22.5 (range $198.5 - 315.0$), and 95.0% of the 300 cipher images fall below the critical value. That rate is what a truly uniform source produces: at $\alpha = 0.05$, about 5% of perfectly random images are expected to exceed the threshold by chance, so a 100% pass rate would itself be suspicious as in Fig 3.

4.3 Information Entropy

Cipher images reach the entropy of a truly random image of the same size, in every modality and in both regions; claims of entropy "close to 8" are only meaningful against this size-specific reference.

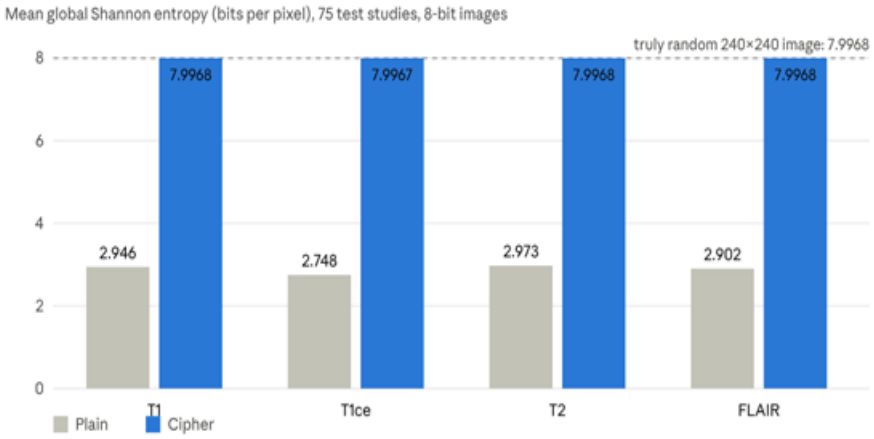


Fig 4. Entropy-based outcomes

BraTS 2018 test split, 75 studies × 4 modalities as in Fig 4; reference from 1,000 uniformly random 240×240 images. The theoretical maximum of 8 bits is reachable only with infinitely many pixels. A finite image of 57,600 uniformly random pixels has an expected entropy slightly below 8, so the correct benchmark is the distribution measured on random images of the same size. Over 1,000 such images the mean is 7.99681, with a 95% range of 7.99624–7.99735. The 300 cipher images average 7.99679 ± 0.00028, and 93.7% of them lie inside that range; plain images average only 2.892 bits because of the large uniform background as in Table 8.

Table 8. Entropy of cipher images (mean ± standard deviation over 300 images unless stated) against truly random images of the same size

Measure	Proposed cipher	Truly random reference (mean, 95% range)
Global entropy	7.99679 ± 0.00028	7.99681 (7.99624–7.99735)
Local entropy (30 blocks of 44 × 44 pixels)	7.90236 ± 0.00187	7.90250 (7.89864–7.90656)
Share of images inside the reference range (global / local)	93.7% / 95.7%	95% by construction
Background voxels only (all modalities pooled)	7.99916 ± 0.00007	—
ROI voxels only (all modalities pooled)	7.963 ± 0.081	—

Local entropy, which is harder to pass than global entropy because it checks small blocks independently, also matches the random reference. The ROI-only figure is lower only because ROIs are small: an ROI covering 0.14% of the slice contributes a few hundred bytes per modality, and the entropy of any short sequence is biased downward. The background, with tens of thousands of bytes, is indistinguishable from random, confirming that the lighter background pass does not leave a statistical trace.

5. Conclusion

This paper addressed a research and developed a novel: chaos-based image encryption, which protects pixels but offers no way to delegate access; proxy re-encryption on blockchains, which offers patient-controlled delegation but only for textual records and usually with costly pairings; and blockchain image-sharing systems, which protect storage and integrity but leave key distribution to a single shared key. The proposed framework joins them in four layers. CMAF-RoINet fuses all modalities of a study with cross-modal attention and produces one ROI mask, tuned for recall because a missed lesion voxel is a security cost. A per-study session key, bound to a hash of the plaintext, drives a key stream that combines hyper-chaos, an LSTM filter and a ChaCha20 floor; ROI voxels of all modalities are jointly permuted and diffused, background voxels receive one stream pass, and the mask itself travels in an authenticated, encrypted header. A pairing-free, two-part-key elliptic-curve proxy re-encryption scheme then delegates only the 256-bit session key, and a consortium blockchain with IPFS storage enforces consent, re-encryption key issuance, revocation by delegation-key rotation, replay protection and audit. All 300 images decrypted bit for bit (MSE = 0, SSIM = 1, PSNR = ∞), and the plaintext-hash integrity check passed in every case. Because the network's fused features are used only to choose the ROI and are never transmitted, lossless recovery holds by construction rather than by approximation. Global entropy was 7.99679 ± 0.00028 against 7.99681 for truly random 240 × 240 images, local entropy 7.90236 against 7.90250, and adjacent-pixel correlation fell from about 0.97–0.98 in plain images to about 0.0002 in magnitude, with the same spread as random noise. Pass rates near 95% at $\alpha = 0.05$ for chi-square, entropy range and NPCR/UACI are exactly what a uniform source produces. A one-pixel change gave NPCR 99.609% and UACI 33.470% (ideal 99.609% and 33.464%), with no difference between changes inside the ROI and in the background. A one-bit key change gave NPCR 99.608% and UACI 33.469%, and was rejected by the encrypted header in 100% of trials. Binding the seed to the plaintext hash is what produces this sensitivity, removing the fixed-key-stream weakness behind most published breaks of chaotic ciphers. A one-pixel change gave NPCR 99.609% and UACI 33.470% (ideal 99.609% and 33.464%), with no difference between changes inside the ROI and in the background. A one-bit key change gave NPCR 99.608% and UACI 33.469%, and was rejected by the encrypted header in 100% of trials. Binding the seed to the plaintext hash is what produces this sensitivity, removing the fixed-key-stream weakness behind most published breaks of chaotic ciphers. The control system has two positive Lyapunov exponents across $r_c \in [0.1, 0.7]$ (for example 0.505 and 0.135 at $r_c = 0.5$), confirming hyper-chaos. The key stream passed eight NIST SP 800-22 tests over 50 sequences of 10^6 bits. The proxy re-encryption scheme decrypted correctly for both owner and delegatee, the proxy rejected a tampered capsule, and a colluding proxy and clinician recovered only the delegation secret s_i and not the private key as it predicts. The limitations are: The security metrics used the expert tumour mask dilated by three pixels. The CMAF-RoINet prototype and its single-modality and early-fusion baselines were still training, so ROC, Dice, recall at the security threshold and the missing-modality ablation remain to be measured. Quantizing chaotic states with a factor of 10^{14} exceeds the exact-integer range of 64-bit floating point for this attractor; the experiments used 10^{12}, and it should be updated to match. In summary, the measured results show that a single-ROI, single-capsule design can protect a multi-modal study with ideal statistical security and exact recovery, and that delegation can be made pairing-free and collusion-resistant.

References

1. El-Shafai, W., Naguib, A., &Shokair, M. (2023). A novel hyper chaotic-cellular automata based medical image encryption technique. In Proceedings of the 3rd IEEE International Conference on Electronic Engineering (ICEEM-2023). IEEE.

2. Inam, S., Kanwal, S., Firdous, R., &Hajjej, F. (2024). Blockchain based medical image encryption using Arnold's cat map in a cloud environment. Scientific Reports, 14, Article 5678. https://doi.org/10.1038/s41598-024-56364-z

3. Kanwal, S., Inam, S., Nawaz, Z., Hajjej, F., Alfraihi, H., & Ibrahim, M. (2024). Securing blockchain-enabled smart health care image encryption framework using Tinkerbell map. Alexandria Engineering Journal, 107, 711–729.

4. Lai, Q., Hu, G., Erkan, U., &Toktas, A. (2023). High-efficiency medical image encryption method based on 2D Logistic-Gaussian hyperchaotic map. Applied Mathematics and Computation, 442, Article 127738. https://doi.org/10.1016/j.amc.2022.127738

5. Long, B., Chen, Z., Liu, T., Wu, X., He, C., & Wang, L. (2024). A novel medical image encryption scheme based on deep learning feature encoding and decoding. IEEE Access, 12, 38382–38398. https://doi.org/10.1109/ACCESS.2024.3371888

6. Miao, B., Lai, Q., Huang, J., Erkan, U., &Toktas, A. (2026). A watermarked image encryption scheme based on hyperchaotic map and Lorentz force shift for medical IoT applications. International Journal of Bifurcation and Chaos, 36(12). https://doi.org/10.1142/S0218127426501749

7. Rachmawanto, E. H., &Pramudya, E. R. (2025). Enhancing medical image security using hyperchaotic Lorenz and Josephus traversing encryption. Scientific Journal of Informatics. https://journal.unnes.ac.id/journals/sji/article/view/9815

8. Zhou, J., Li, Q., Li, W., Yu, G., Shao, Y., Dong, Y., Wang, M., Li, Z., Gong, C., & Chen, X. (2025). TDADL-IE: A deep learning-driven cryptographic architecture for medical image security (arXiv:2510.11301). arXiv. https://arxiv.org/abs/2510.11301

9. Zhuang, Z., Zhuang, Z., & Wang, T. (2024). Medical image encryption algorithm based on a new five-dimensional multi-band multi-wing chaotic system and QR decomposition. Scientific Reports, 14. https://doi.org/10.1038/s41598-023-50661-9

10. Ali, A. A., Gunavathie, M. A., Srinivasan, V., Aruna, M., Chennappan, R., &Matheena, M. (2025). Securing electronic health records using blockchain-enabled federated learning for IoT-based smart healthcare. Clinical eHealth, 8, 125–133. https://doi.org/10.1016/j.ceh.2025.04.002

11. Ashwini, D. Y., &Puneeth, R. P. (2026). Blockchain-based framework for enhancing interoperability and security in EHR exchange using lightweight ECC proxy re-encryption. ActaInformaticaPragensia, 15(1), 90–108. https://doi.org/10.18267/j.aip.290

12. Cobrado, U. N., Sharief, S., Regahal, N. G., Zepka, E., Mamauag, M., & Velasco, L. C. (2024). Access control solutions in electronic health record systems: A systematic review. Informatics in Medicine Unlocked, 49, Article 101552. https://doi.org/10.1016/j.imu.2024.101552 †

13. Ferreira, J. C., Elvas, L. B., Correia, R., &Mascarenhas, M. (2024). Enhancing EHR interoperability and security through distributed ledger technology: A review. Healthcare, 12(19), Article 1967. https://doi.org/10.3390/healthcare12191967 †

14. Keshta, I., Aoudni, Y., Sandhu, M., Singh, A., Xalikovich, P. A., Rizwan, A., Soni, M., &Lalar, S. (2023). Blockchain aware proxy re-encryption algorithm-based data sharing scheme. Physical Communication, 58, Article 102048. https://doi.org/10.1016/j.phycom.2023.102048 †

15. Kunal, S., Gandhi, P., Rathod, D., Amin, R., & Sharma, S. (2024). Securing patient data in the healthcare industry: A blockchain-driven protocol with advanced encryption. Journal of Education and Health Promotion, 13(1), Article 94. https://doi.org/10.4103/jehp.jehp_984_23 †

16. Mauricio, D., Llanos-Colchado, P. C., Cutipa-Salazar, L. S., Castañeda, P., Chuquimbalqui-Maslucán, R., Rojas-Mezarina, L., & Castillo-Sequera, J. L. (2024). Electronic health record interoperability system in

Peru using blockchain. International Journal of Online and Biomedical Engineering, 20(3), 136–153. https://doi.org/10.3991/ijoe.v20i03.44507 †

17. Puneeth, R. P., &Parthasarathy, G. (2023). Survey on security and interoperability of electronic health record sharing using blockchain technology. ActaInformaticaPragensia, 12(1), 160–178. https://doi.org/10.18267/j.aip.187 †

18. Puneeth, R. P., &Parthasarathy, G. (2024). Blockchain-based framework for privacy preservation and securing EHR with patient-centric access control. ActaInformaticaPragensia, 13(1), 1–23. https://doi.org/10.18267/j.aip.225 †

19. Rastogi, P., Singh, D., &Bedi, S. S. (2023). Fully decentralized block chain with proxy re-encryption algorithm for healthcare security. International Journal of Intelligent Engineering and Systems, 16(1). https://doi.org/10.22266/ijies2023.0228.49

20. Semantha, F. H., Azam, S., Shanmugam, B., & Yeo, K. C. (2023). PbDinEHR: A novel privacy by design developed framework using distributed data storage and sharing for secure and scalable electronic health records management. Journal of Sensor and Actuator Networks, 12(2), Article 36. https://doi.org/10.3390/jsan12020036 †

21. Shahzad, A., Chen, W., Zhang, Y., & Kumar, R. (2025). Zero-trust medical image sharing: A secure and decentralized approach using blockchain and the IPFS. Symmetry, 17(4), Article 551. https://doi.org/10.3390/sym17040551

22. Sharma, P., Kumar, T., &Tyagi, S. S. (2025). A blockchain-based secure framework for interoperability of patient data in electronic health records (EHR). Fusion: Practice and Applications, 19(1), 23–37. https://doi.org/10.54216/FPA.190103 †

23. Shen, Y., Yu, J., Zhou, J., & Hu, G. (2025). Twenty-five years of evolution and hurdles in electronic health records and interoperability in medical research: Comprehensive review. Journal of Medical Internet Research, 27, Article e59024. https://doi.org/10.2196/59024 †

24. Sonkamble, R. G., Bongale, A. M., Phansalkar, S., &Dharrao, D. S. (2024). A secure interoperable method for electronic health records exchange on cross platform blockchain network. MethodsX, 13, Article 103002. https://doi.org/10.1016/j.mex.2024.103002 †

25. Zhang, J., Guo, R., Shi, Y., & Tang, W. (2024). An anti-impersonation attack electronic health record sharing scheme based on proxy re-encryption and blockchain. Mathematical Biosciences and Engineering, 21(6), 6167–6189. https://doi.org/10.3934/mbe.2024271

26. Ahmad, H. F., Alhassan, F. M., Haque, A., &Shafqat, S. (2025). A secure architecture for interoperable personal health records (PHR) based on blockchain and FHIR. Journal of Pioneering Medical Sciences, 14(2), 42–48. https://doi.org/10.47310/jpms2025140207 †

27. Islam, U., Alatawi, M. N., Alqazzaz, A., Alamro, S., Shah, B., & Moreira, F. (2025). A hybrid fog-edge computing architecture for real-time health monitoring in IoMT systems with optimized latency and threat resilience. Scientific Reports, 15(1), Article 25655. https://doi.org/10.1038/s41598-025-09696-3 †

28. Shi, X., Chai, S., Li, Y., Cheng, J., Bai, J., Zhao, G., & Chen, Y.-W. (2023). Cross-modality attention adapter: A glioma segmentation fine-tuning method for SAM using multimodal brain MR images (arXiv:2307.01124). arXiv. https://doi.org/10.48550/arXiv.2307.01124

29. Sun, K., Ding, J., Li, Q., Chen, W., Zhang, H., Sun, J., Jiao, Z., & Ni, X. (2024). CMAF-Net: A cross-modal attention fusion-based deep neural network for incomplete multi-modal brain tumor segmentation. Quantitative Imaging in Medicine and Surgery, 14(7). https://doi.org/10.21037/qims-24-9

30. Zhang, M.-D., & Pan, K.-W. (2025). A multi-modal fusion framework for brain tumor segmentation based on 3D spatial-language-vision integration and bidirectional interactive attention mechanism (arXiv:2507.08574). arXiv. https://arxiv.org/abs/2507.08574