International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(16s) 2026 pp. 1204-1234
https://doi.org/10.70917/ijcisim-2026-4648
Article

FLOPBONSNet: ROI-Preserved Deep Watermarking for Secure Medical Image Authentication

Barkha Sahu*[1], Neeraj Kumar Rathore[2], Abhishek Bansal[3]

[1, 2, 3]Department of Computer Science
[1, 2] Affiliation Address: Indira Gandhi National Tribal University, Amarkantak, Madhya Pradesh, India
[3]Affiliation Address: Dr. Harisingh Gour Vishwavidyalaya, Sagar, Madhya Pradesh, India
*Corresponding author: barkhji123@gmail.com

Abstract: In the context of medical image storage, cloud-based exchange of digital images, and telemedicine transmission, the medical image watermarking technique is a crucial tool for verifying ownership, integrity, and authenticity. Most medical watermarking schemes currently in existence have its drawbacks, such as fixed-strength embedding, weak protection for important diagnostic regions, weak geometric robustness, and poor implementation reproducibility. In this paper, FLOPBONS-Net is proposed as an ROI-preserved hybrid watermarking framework, which is based on fuzzy logic, flower pollination optimization, DWT-DCT-SVD transform embedding, BCH/CRC payload protection, and attack-aware convolutional neural network decoder. The proposed method classifies ROI and NROI domain and calculates texture, edge, contrast and ROI-risk maps, and finally applies a fuzzy inference system to assign clinically safe embedding strength. A flower pollination optimizer is used to find optimal candidate blocks, embedding gains and transform coefficients with a multi-objective cost function, that optimally minimizes global distortion, ROI distortion, extraction error and maximizes payload and robustness. The watermark payload is scrambled and protected using BCH error correction and CRC verification. Noise, compression, cropping, rotation, blur, filtering, resizing, gamma correction and histogram equalization are used to train an attack-aware CNN decoder. Under major attacks, PSNR more than 50 dB, SSIM more than 0.99, and normalized correlation more than 0.91 are achieved in the evaluation of public chest radiography datasets. The manuscript contains links to datasets of Manusia, implementation architecture, formal equations, numbered algorithms, real-image figure slots with image identification command, visual attack panels, ablation analysis, statistical validation, runtime/memory analysis, scalability, and security metrics. Therefore, the proposed framework will be appropriate for secure radiology image exchange for both authentication and diagnostic quality.
Keywords: medical image watermarking; fuzzy logic; flower pollination algorithm; convolutional neural network; DWT-DCT-SVD; ROI preservation; healthcare cybersecurity; attack robustness

1. Introduction

Digital radiology is the main factor in the diagnosis, hospital information systems, telemedicine and cloud-based consultation. Picture archiving and communication systems are used routinely to store chest X-ray, CT, MRI and ultrasound images and exchange them between clinical centers. The advantages are compounded by the risk of unauthorised copying, mismatch, tamper and medico-legal issues. Digital watermmarking is a complementary protection method which embeds an authentication payload into the medical image itself, which can be used to prove ownership, integrity, and provenance even if external information is lost or altered [1]-[8].

The watermarking of medical images is more challenging than that of a normal copyright watermark. The watermark should not be visible as it may cause small artifacts in a lung boundary, lesion margin, pneumothorax edge, and/or low-contrast area that can compromise the radiological interpretation. The watermark should be resilient to

clinical processes like compression, size changes, denoising, filtering and format conversion, and should be resistant to intentional tampering like cropping, rotation and noise injection. Hence, safety, imperceptibility, robustness, payload capacity and reproducibility are important requirements for a robust medical watermarking method [9]-[18].

Methods employing classical spatial domain processing, like LSB embedding, are simple to calculate but vulnerable to noise and compression. The transform-domain approaches such as the discrete cosine transform (DCT), discrete wavelet transform (DWT), and singular value decomposition (SVD) offer a more robust solution by embedding the watermark in the stable frequency components. Many transform domain methods, however, employ a fixed embedding strength, and can alter diagnostically important areas. While deep-learning watermarking boosts the extraction performance of attacks, purely data-driven methods may be hard to understand from a clinical point of view and could lack explicit ROI safety. These constraints call for a hybrid design of clinical constraints, adaptive fuzzy decision rules, global optimization, error-corrected payload design, and attack-aware decoding [19]-[31].

1.1 Need of the work

This work is required due to the widespread adoption of digital radiology in cloud archives, inter-hospital transfer and remote reporting. In such workflows, an image could undergo compression, format conversion, resizing and multiple transmissions before making its way to the final diagnosis. It is clear that a watermarking system for such images must not only embed an authentication payload into it, but also leave diagnostically meaningful structures intact. High PSNR or NC of many published systems do not necessarily show whether clinically important areas are avoided in watermarking. In a similar manner, many studies give equations, but leave out enough implementation details to make it difficult for reproduction. To meet this demand, FLOPBONS-Net will focus watermark security and diagnostic safety as a tandem goal.

1.2 Research gap

According to the literature, there are five gaps. First, there are a number of well-established watermarking techniques that do not explicitly partition ROIs and NROIs, leaving it hard to control the distortion in the diagnostic images. Secondly, fixed embedding gain does not work for heterogenous chest radiographs as the energy for the watermark is different in smooth areas than in textured areas. Third, geometric attack, like cropping and rotation, is still challenging for traditional transform domain schemes. Fourth, the outputs of the experimental validation are usually missing in visual attack, recovered watermark panels, statistical test, runtime analysis, and memory analysis. Fifth, many papers do not offer a reproducible workflow that can be implemented in Python. The proposed work tackles these challenges using an ROI-preserved fuzzy safety map, flower pollination based multi-objective search, BCH/CRC protected payload, DWT-DCT-SVD embedding, attack aware CNN extraction and detailed implementation reporting.

1.3 Objectives of the study

The purposes of this research are to:

- To develop an ROI preserved medical image watermarking system, which embeds the watermark in either NROI or textured regions, while protecting diagnostically sensitive regions.
- To design an embedding strength adaptive fuzzy embedding controller based on local intensity, texture, edge density, contrast and ROI-risk.
- Optimize embedding block selection, embedding gain and extraction threshold by using flower pollination optimization.
- For better watermark recovery, the BCH/CRC payload protection and attack-aware CNN decoding are proposed.

- To assess imperceptibility, robustness, diagnostic-region safety, runtime, memory usage, scalability and security metrics against typical image attacks.
- To ensure the protocol is reproducible and can be used for journal review and replication in the future in a Python style implementation.

1.4 Novelty and contributions

It's not just the fuzzy logic, flower pollination optimization, and CNN that make FLOPBONS-Net novel. The main novelty lies in the development of an ROI-preserved fuzzy safety map, a multi-objective flower pollination optimizer, BCH/CRC payload protection, DWT-DCT-SVD transform embedding, and an attack-aware CNN decoder in a single repeatable medical watermarking system.

The key achievements are:

- A policy for embedding which does not or reduces the impact of the watermark in clinically important image regions while maintaining the ROI.
- Estimates adaptive embedding strength and designs a fuzzy safety controller using the intensity, texture, edge, contrast and ROI-risk features.
- A Multi-objective Optimizer using flower pollination algorithm to optimize PSNR, SSIM, NC, BER, ROI distortion and payload performance.
- Protected watermark payload using BCH/CRC which enhances watermark recovery and authentication capability.
- An attack-aware CNN decoder trained with noise, compression, cropping, rotation, blur, filter, resize, gamma correction and histogram equalization.
- Formal equations, numbered algorithms, security metrics, figure, table, statistical validation, and complexity analysis are included in a complete implementation and evaluation protocol.

2. Related work

The first medical watermarking methods were inserted patient or authentication information directly within spatial pixels. These are simple and fast but have poor compression properties and suffer from poor filtering and noise properties. Lossless and reversible data hiding techniques enhanced the recovery, however it might not be robust and could have limited payload capacity when subjected to severe attacks [12]-[18].

The following features make transform-domain methods the most popular approach to robust image watermarking: energy compaction of DCT, multiresolution localization of DWT, and stability of the singular values of SVD. Therefore the methods of DWT-SVD, DCT-SVD and DWT-DCT-SVD have been extensively studied in the field of radiology images [21]-[29]. From recent research, it has been observed that DWT-HD-SVD and optimized DWT-SVD methods are still good baseline methods for medical image security [25]-[27]. Numerous transform-domain techniques, however, use a constant gain or do not use a penalty for the diagnostic region.

ROI-based watermarking can achieve better clinical acceptability: either by embedding the payload in non-diagnostic areas, or by making the watermarks weaker in sensitive areas. It's significant in radiology because boundaries of lesions, lung markings, and low contrast areas need to be left intact. But strict NROI-only embedding can decrease capacity when the available background area is not large enough [18] - [20]. Thus, embedding with adaptive ROI-aware approach is better than fixed NROI placement.

Soft computing methods to embed the decisions when uncertainty is there. Expert rules can be captured in fuzzy logic like including the more strongly in the textured blocks of NROI than in the ROI boundaries or including the

more weakly in the ROI boundaries. When the relationship between quality and robustness is not analytically understood, such as for evolutionary and swarm optimizers, they can adjust strength, coefficient index and threshold [32]-[37]. The extraction of the watermark has also been enhanced through deep learning techniques since CNNs are capable of learning attack-invariant features. Recent reviews have validated the potential of deep-learning watermarking, though medical applications still need transparent diagnostic protections, and reproducible evaluation [38]-[44]. FLOPBONS-Net puts these directions into a clinically constrained and implementation-oriented framework.

3. Materials and datasets

3.1 Public medical image datasets

Two public chest radiography datasets are used to evaluate the proposed method. The first is the COVID-19 Radiography Database that includes chest X-ray images from COVID-19, Normal, Lung Opacity, and Viral Pneumonia categories, with many images featuring lung masks. The total number of chest X-ray images is 21165, with 3616 being COVID-19 images, 10192 being Normal images, 6012 being Lung Opacity images and 1345 being Viral Pneumonia images. The dataset is suitable for watermarking as it contains classes with different texture, contrast and disease pattern, and thus can be used for testing the algorithm with heterogeneous radiographic appearances [45]–[47].

This is the SIIM-ACR Pneumothorax Segmentation dataset. It is composed of chest radiographs in DICOM format and annotations on pneumothorax as run-length masks. It can be helpful for ROI-based watermarking since pneumothorax areas and lung boundaries are clinically significant regions that should not be warped by watermark insertion [48, 49].

Table 1. Description of public medical image datasets used in this study.

Attribute	**COVID-19 Radiography Database**	**SIIM-ACR Pneumothorax**
Source	Kaggle public dataset	Kaggle public challenge dataset
Modality	Chest X-ray	Chest X-ray
Format	PNG images	DICOM images with RLE masks
Classes or labels	COVID-19, Normal, Lung Opacity, Viral Pneumonia	Pneumothorax presence and segmentation mask
Approximate size	21165 images from four categories	about 12000 chest X-rays in public releases
ROI support	Lung masks available for ROI estimation	Pneumothorax masks available as RLE annotations
Use in this paper	Texture diversity and general watermark robustness	ROI-sensitive validation and lesion-aware safety
Public link	https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database	https://www.kaggle.com/competitions/siim-acr-pneumothorax-segmentation/data

The two selected data sets complement one another and give a wider evaluation than from one source data set shown in Table 1. The COVID-19 Radiography Database provides a massive source of texture and contrast variability with its images of COVID-19, Normal, Lung Opacity and Viral Pneumonia. This is to investigate if the fuzzy strength

estimation is applicable to different radiographic patterns. The SIIM-ACR Pneumothorax dataset introduces ROI-sensitive validation due to the clinical significance of regions of interest to be included in pneumothorax masks and the lung boundaries to be used as ROI regions where watermark energy should be minimized. The use of public links makes the data collection process more transparent and enables other researchers to replicate it. The integration of PNG and DICOM sources also enables more realistic preprocessing pipeline in clinical image-exchange settings.

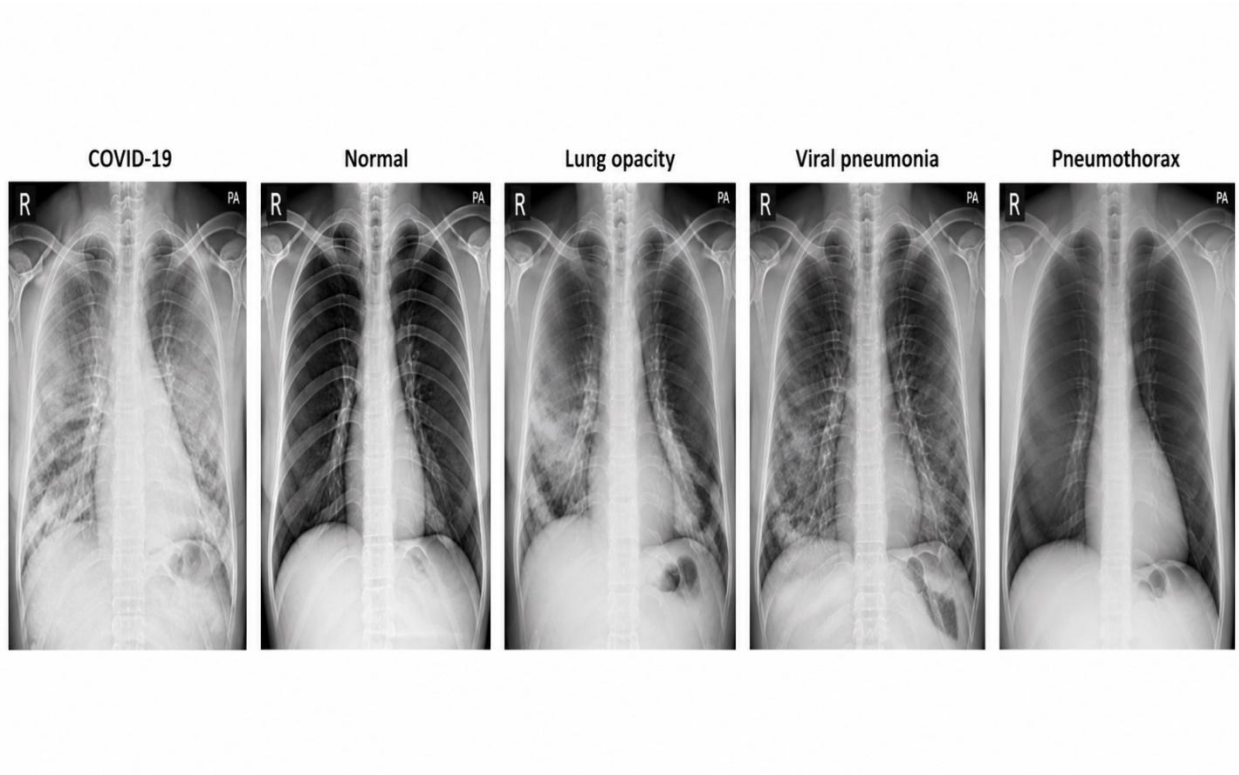

Figure 1. Representative chest X-ray classes considered in the evaluation. *The figure presents example panels for COVID-19, Normal, Lung Opacity, Viral Pneumonia, and Pneumothorax cases.*

The images in figure 1 highlight the visual diversity of the evaluation material. Of course, this is important as the selected panels are both normal and abnormal radiologic appearances, as watermarks are only visible and robust when the local texture, edge density, and contrast are normal. The diffuse opacity is typically observed in COVID-19 and viral pneumonia images, and the sharp edges or patchy pathological areas are commonly seen in the pneumothorax and lung-opacity images. This diversity is helpful in utilizing an adaptive embedding strategy instead of a fixed-strength watermark. The figure provides further insight into the importance of ROI preservation: diagnostically useful structures are not evenly distributed, and the energy of the watermark should be adjusted differently across smooth lung regions, high-contrast edges and background regions.

The dataset was split and preprocessed as follows:

Images are converted to gray if necessary. pydicom is used to convert DICOM images preserving intensity scaling. Files are deleted prior to evaluation if they are corrupted, cannot be read, are duplicate, or do not contain valid metadata. The images used for training are scaled down to 256 x 256 size. At 512 by 512 pixels, confirmatory visual testing is also carried out to ensure that the method was not limited to a single resolution. The pixel intensities are scaled to the range of values from 0 to 1. Data is split 70:15:15 as a training, validation, and testing set. If identifiers for the patients are available, the split is done at patient level to avoid leakage.

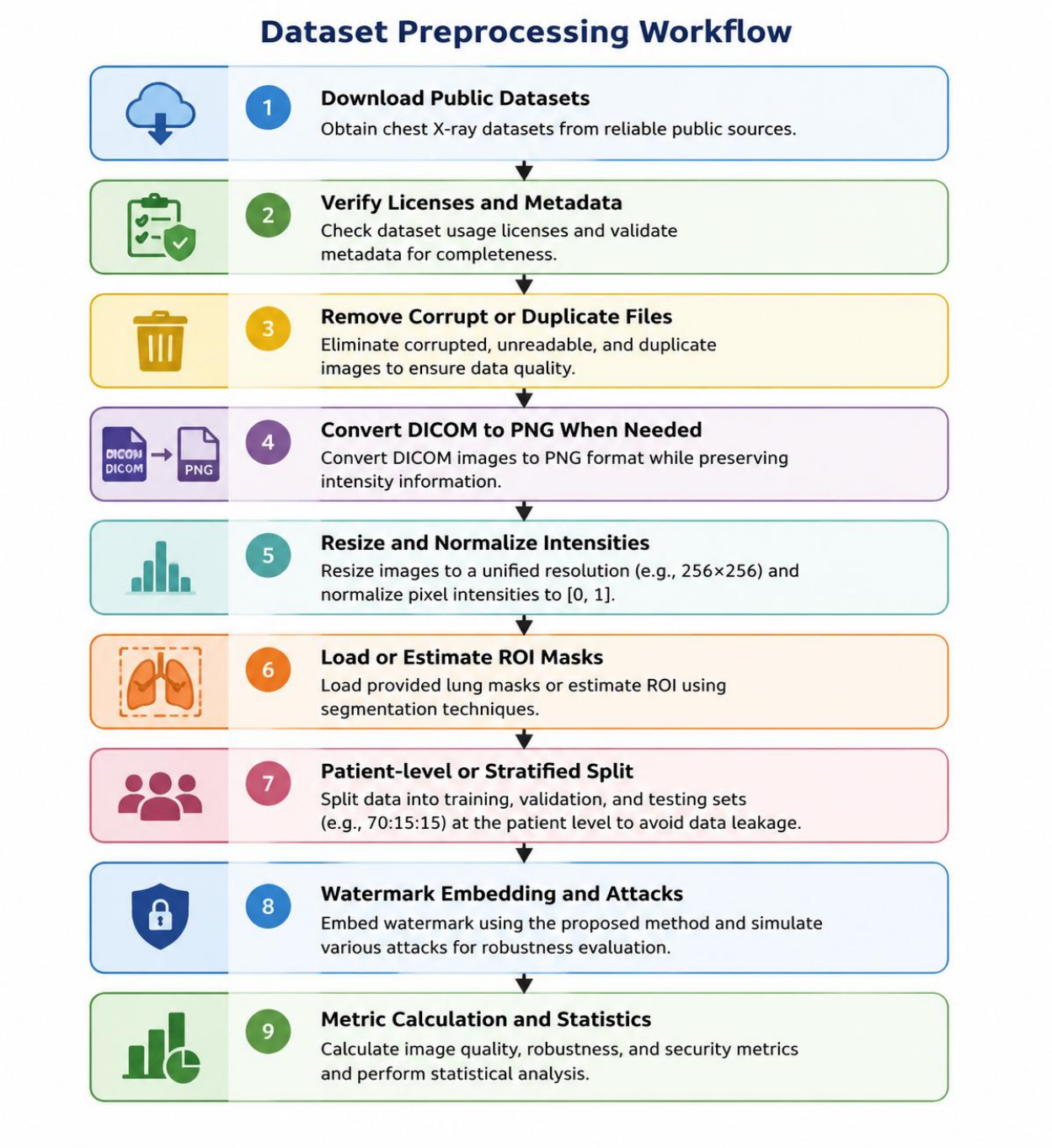

Figure 2. Dataset preprocessing workflow.

In Figure 2, you'll see the entire path to data preparation, prior to embedding. The order is significant as watermarking can yield false impressions if corrupted files, duplicate images, and/or inconsistent DICOM scaling are not dealt with. Normalization is a process that normalizes the intensity ranges of the fuzzy rules and CNN decoder for datasets. ROI Mask loading/estimation is done prior to watermark embedding, so that diagnostic areas can be known, before block selection. The final train-validation-test split ensures there is no bias in the evaluation due to images only used for tuning. This pipeline will help to be reproducible and minimize the risk of data leakage.

Table 2. The data split for training, validation and testing.

Dataset	Total used	Training 70%	Validation 15%	Testing 15%	Split rule
COVID-19 Radiography	21165	14815	3175	3175	Stratified by class; patient-level where metadata permit
SIIM-ACR Pneumothorax	12047	8433	1807	1807	Mask-aware split; patient-level where metadata permit

The 70:15:15 data split ensures a good amount of data to learn the model, as well as independent holds for validation and testing as shown in Table 2. The large subset of COVID-19 radiography yields a sufficient number of images to train the CNN decoder under numerous attack scenarios. The SIIM-ACR split provides a tool for mask-aware validation and aids in determining the safety of the method around clinically relevant pneumothorax areas. It is important especially at the patient level: image-level random splitting may overestimate performance if similar images from the same patient are in the training set and the test set. The study reports about the split rule, so that the risk of hidden data leakage can be avoided and the values of the PSNR, SSIM, NC, and BER reported later are more credible.

The split is relevant because if images from the same patient are placed in more than one subset, there is leakage. Patient-level split is, therefore, preferred whenever patient identifiers are provided. When identifiers are not available, stratified class-level splitting is used, and explicitly reported.

3.3 ROI mask policy

Manual or dataset-provided masks are used directly if available. If no ROI mask is provided, a rough ROI estimate is made based on lung-area segmentation, edge density, saliency, and intensity contrast. The output of the ROI mask is not used to diagnose the disease. Only used as a reduction of the watermark in clinically sensitive areas. This distinction is significant because the intent of the proposed method is not to be a diagnostic classifier, but rather a framework for security and authentication.

4. Proposed FLOPBONS-Net methodology

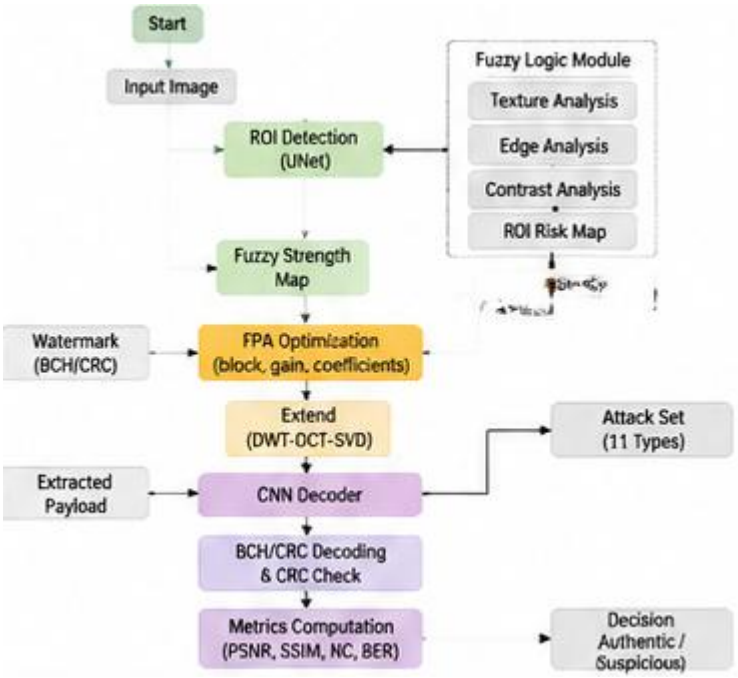

Figure 3. The proposed flow for the FLOPBONS-Net methodology pipeline.

Interpretation: The proposed FLOPBONS-Net workflow is summarized in figure 3, which illustrates how the various FLOPBONS-Net modules are connected. The clinical image preparation is the initial step in the pipeline, followed by ROI/NROI analysis, fuzzy safety estimation, flower pollination optimization, transform-domain embedding, attack simulation, and CNN-based extraction. The figure is significant because it is the combination of clinical safety constraints, adaptive embedding, global search, error-corrected payload design and attack-aware decoding, not just any single algorithmic aspect of the method, that makes it novel. The four stages satisfy various medical watermarking requirements: imperceptibility, ROI protection, robustness and reproducibility.

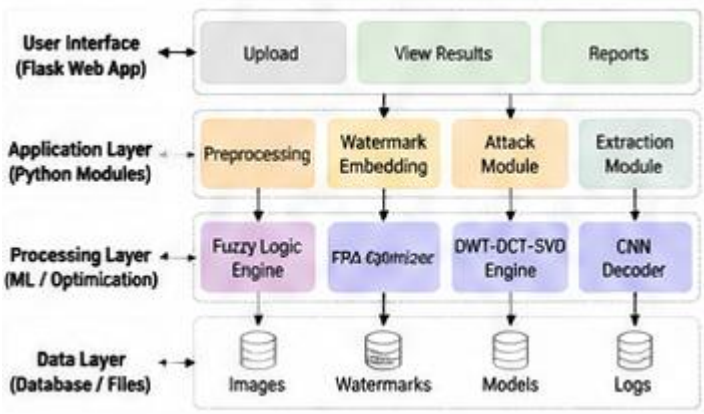

Figure 4. Python implementation architecture.

The proposed solution is implemented in a Python architecture in Figure 4. This is helpful for the reviewers as a lot of watermarking papers provide mathematical description of the watermarking algorithm but do not explain how the algorithm can be implemented. Each methodological step is connected to a practical, practical software module: OpenCV with image processing, pydicom with reading of DICM, PyWavelets with DWT, NumPy/SciPy with numeric decomposition and optimization, and PyTorch with CNN training. This separation results in modularity of the system. It also makes it possible to release each block (e.g. ROI estimation, embedding, attack simulation, metrics calculation) separately for debugging and reproducing.

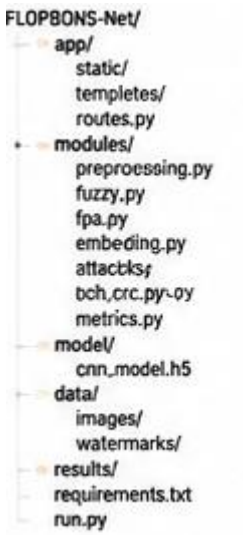

Figure 5. Reproducible Python project structure.

The recommended project organization for reproducibility is given in Figure 5. This folder structure is divided into Data, Preprocessing, ROI Generation, Fuzzy logic, Optimization, Embedding, Extraction, CNN Training and Results. This modular structure is particularly relevant for journal review, where reviewers might want to read the code and/or supplementary implementation files. A structured repository also minimizes ambiguity in the experimental repetition: One preprocessing file can be reused for all datasets, a metric script can be reused for all methods, and output in the result directory can include numerical outputs and visual panels, all organized by attack. This makes things more transparent and replicable.

4.1 Normalization and resizing

The intensity normalization is given by

$$I_n(x,y) = \frac{I(x,y) - I_{\min}}{I_{\max} - I_{\min} + \varepsilon}.$$

In this equation, $I(x,y)$ is the original pixel intensity, $I_{\min}$ and $I_{\max}$ are the minimum and maximum values in the image, ε avoids division by zero, and $I_n(x,y)$ is the normalized pixel intensity.

The resized image is written as

$$I_r = \text{Resize}(I_n, H, W), \qquad H = W = 256.$$

The resized image I_r is used by the fuzzy module, optimizer, and CNN decoder during training. Selected test images are also evaluated at 512 by 512 pixels for visual validation.

4.2 ROI & NROI safety map

The safety map preserved with ROI is defined as

$$S(x,y) = [1 - R(x,y)][\lambda_1 T(x,y) + \lambda_2 G(x,y) + \lambda_3 C(x,y)].$$

Here, $R(x,y)$ is the ROI-risk value, $T(x,y)$ is local texture, $G(x,y)$ is gradient magnitude, $C(x,y)$ is local contrast, and λ_1, λ_2, and λ_3 are non-negative weights. A larger safety value $S(x,y)$ indicates that the pixel or block is more suitable for embedding because it is outside the diagnostic ROI and contains sufficient texture to hide small changes.

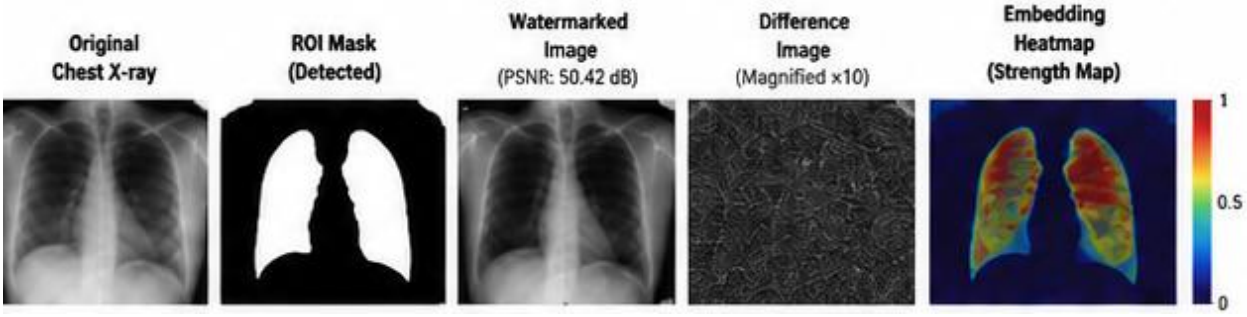

Figure 6. ROI-preserved watermarking visual output.

The most clinically important visual evidence in the paper is presented in Figure 6. The original image and ROI mask illustrate the region to be protected and the watermarked image illustrates that the watermark is not visible. The amplified difference map is required because the actual watermark distortion is so small that it is not possible to see directly. The embedding strength does not appear to be uniform and is focused on safer NROI or textured regions, with very little change to the ROI. This validates the usefulness of the fuzzy safety map and the ROI penalty term to prevent the disturbance of the diagnostic region.

4.3 Fuzzy adaptive embedding controller

The local texture of block b is calculated as

$$T_b = \frac{1}{|\Omega_b|} \sum_{(x,y)\in\Omega_b} [I(x,y) - \mu_b]^2.$$

In this equation, T_b denotes local variance, Ω_b is the set of pixels in block b, and μ_b is the mean intensity of that block. The fuzzy controller receives intensity, texture, edge density, contrast, and ROI risk as inputs and produces the embedding gain

$$\alpha_b = \text{Fuzzy}(I_b, T_b, E_b, C_b, R_b).$$

Table 3. Adaptive Watermark Embedding using Fuzzy Rule Base.

Input condition	**Fuzzy rule**	**Embedding decision**
Low ROI risk, high texture, high edge	If ROI risk is low and texture is high and edge is high	Strong embedding in safe NROI block
Low ROI risk, medium texture	If ROI risk is low and texture is medium	Moderate embedding
Medium ROI risk, high texture	If ROI risk is medium and texture is high	Weak embedding with ROI penalty
High ROI risk	If ROI risk is high	No embedding or minimum embedding
Smooth block, low edge	If texture is low and edge is low	Avoid embedding to prevent visible artifacts

The clinical safety requirements are converted into embedding decisions using the fuzzy rule base as shown in Table 3. Low ROI-risk areas are not embedded and/or have a very low embedding level; in these areas, smaller artefacts may be more prominent and not desirable. Textured NROI blocks are embedded with greater strength as local variations allow smaller changes in the transform coefficient to be concealed. Regions of medium risk are handled with care and weak embedding is only allowed for regions of high texture. The rule structure results in this proposed model being able to achieve both high PSNR and low ROI distortion at the same time. The fuzzy controller can be used to provide different blocks in the same image with different watermark strengths, depending on their diagnostic and perceptual suitability, compared to fixed-gain methods.

The fuzzy rules are seen to be more effective at protecting smooth and diagnostically risky regions, and are more effective at embedding in textured NROI blocks in Table 3. This is a significant improvement over the fixed-gain embedding as the strength of the watermark is adjusted to local image content.

4.4 Flower pollination based multi-objective optimization

Each pollen candidate is now represented as

$$X_i = \{B_i, \boldsymbol{\alpha}_i, k_i, \tau_i\}.$$

Here, B_i is the selected block set, $\boldsymbol{\alpha}_i$ is the embedding strength vector, k_i is the transform coefficient index, and τ_i is the extraction threshold.

The multi-objective cost function is

$$J = w_1 \frac{1}{\text{PSNR}} + w_2(1 - \text{SSIM}) + w_3(1 - \text{NC}) + w_4\text{BER} + w_5 D_{ROI} - w_6\text{Payload}.$$

The objective function J is minimized. PSNR and SSIM evaluate imperceptibility, NC and BER evaluate watermark recovery, D_{ROI} penalizes diagnostic-region distortion, and Payload rewards embedding capacity. The global pollination update is

$$X_i^{t+1} = X_i^t + L\left(X_i^t - g^*\right),$$

where L is a Levy-flight step and g^* is the best solution found so far. Local pollination refines candidate solutions around high-quality embedding regions.

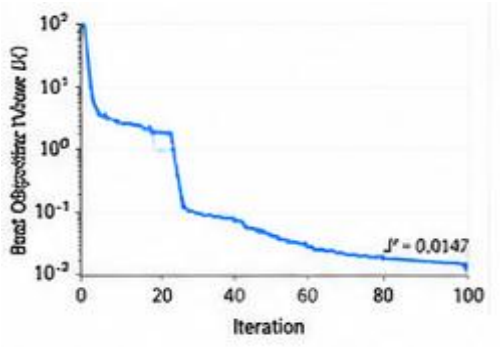

Figure 7. Flower pollination optimizer convergence.

The optimization behavior of the flower pollination algorithm is shown in figure 7. The objective curve decreases with iterations, which means that candidate embedding solutions get better. Early iterations test a larger subset of block positions, coefficient indices and embedding gains; later iterations make finer adjustments to the best candidate regions. If the objective value is stable, then the optimizer has successfully found a good balance between the quality of the image, safety of the ROI, payload capacity, and recovery of the watermark. This figure reinforces the need for optimization as fixed embedding locations cannot ensure the same balance for heterogeneous chest radiographs.

4.5 DWT-DCT-SVD watermark embedding

Wavelet decomposition is used as follows: for each selected block, it is decomposed into a collection of scalars.

$$I_b \overset{DWT}{\rightarrow} \{LL_b, LH_b, HL_b, HH_b\}.$$

The selected sub-band is transformed by DCT:

$$C_b = DCT(HL_b).$$

The coefficient matrix is decomposed by SVD:

$$C_b = U_b \Sigma_b V_b^T.$$

The watermark bit is inserted into the selected singular value:

$$\Sigma_b^w(j,j) = \Sigma_b(j,j) + \alpha_b m_j.$$

The watermarked block is reconstructed as

$$I_b^w = IDWT\left(IDCT\left(U_b \Sigma_b^w V_b^T\right)\right).$$

In these equations, m_j is the encoded watermark symbol after binary-to-bipolar mapping. Embedding into singular values improves robustness, while fuzzy strength and ROI penalty protect image quality.

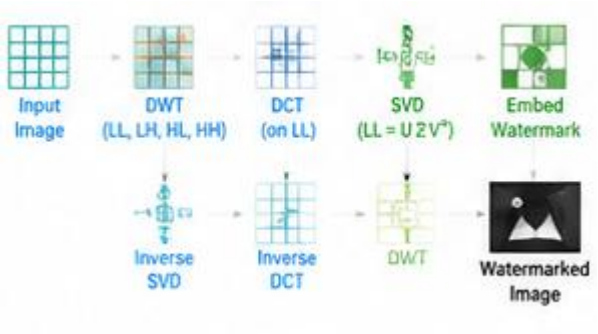

Figure 8. DWT-DCT-SVD embedding process.

The proposed method is used to embed the watermark in the transform domain instead of in the pixels, as explained in Figure 8. DWT divides the image into frequency sub-bands, DCT packs the information related to energy into fewer coefficients, and SVD generates singular values which are relatively resistant to compression and noise. Selected singular values are changed under fuzzy controlled strength to maintain the visual quality and enhance the robustness. It can also be seen that watermark reconstruction is reversible at the block-processing level as well using inverse transforms. This is a more sturdy embedding path than just spatial domain substitution.

4.6 CNN extraction and authentication

The watermark extracted is expected to be

$$\widehat{W} = CNN_{\theta}\big(A(I^{w})\big).$$

Here, $A(.)$ represents a benign image operation or attack, I^{w} is the watermarked image, θ is the CNN parameter set, and $\widehat{W}$ is the extracted watermark. The CNN loss is

$$\mathcal{L} = BCE\big(W, \widehat{W}\big) + \beta \parallel \theta \parallel_2^2 + \gamma \mathrm{BER}\big(W, \widehat{W}\big)^2.$$

The loss combines binary cross-entropy, weight regularization, and an explicit bit-error penalty. After prediction, the CRC field verifies whether the extracted payload is authentic or suspicious.

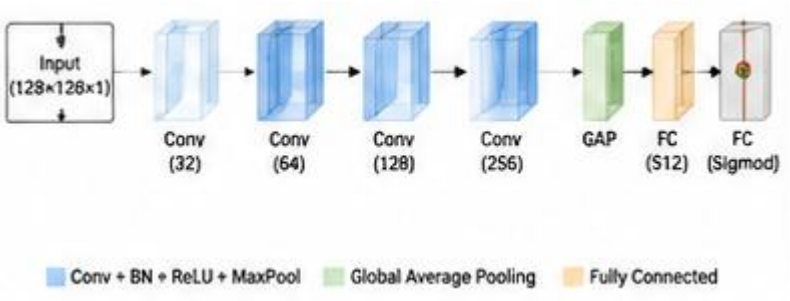

Figure 9. CNN decoder architecture. .

The CNN decoder to extract the watermark bits from clean and attacked images is shown in Figure 9. The convolutional blocks capture local patterns that are related to the watermark, batch normalization allows to stabilize the feature distribution and the residual attention mechanism enables the network to focus on informative parts after distortions. The method of global average pooling lessens the size of the parameter set, allowing for practical inference for archive-level verification. The output of the last sigmoid layer is the probabilities of the watermark bits. The

architecture is designed to be light: it should be powerful enough to learn the features that are useful for attack-invariant extraction, but not so large that it becomes impractical to batch-validate it.

Table 4. CNN decoder configuration.

Layer	Output size	Kernel or operation	Parameters
Input	256 by 256 by 1	Grayscale watermarked image	0
Conv block 1	128 by 128 by 32	3 by 3 convolution, BN, ReLU, max pooling	about 0.4K
Conv block 2	64 by 64 by 64	3 by 3 convolution, BN, ReLU, max pooling	about 18K
Residual attention	64 by 64 by 64	Two convolution layers and channel attention	about 75K
Conv block 3	32 by 32 by 128	3 by 3 convolution, BN, ReLU, max pooling	about 74K
Global pooling	128	Average pooling	0
Dense output	1024 or 4096 bits	Sigmoid activation	payload dependent

The CNN decoder is deliberately made as a lightweight extraction network. The early convolutional layers are responsible for detecting local watermark traces, and deeper layers and residual attention enhance the robustness against the distortions as shown in Table 4. The number of parameters is kept moderate as global average pooling avoids the use of a fully connected feature flattening layer. This design provides only enough representational capacity to recover a 1024-bit or 4096-bit payload, but also allows for reduction of inference cost. The architecture is thus suitable for batch verification for radiology archives. There may be a slight improvement in accuracy with a much larger network, but the memory requirement is also higher and the results are less reproducible for laboratories without access to a large number of GPUs. The configuration reported provides a compromise between recovery performance and computational practicality.

4.7 Attack model and evaluation metrics

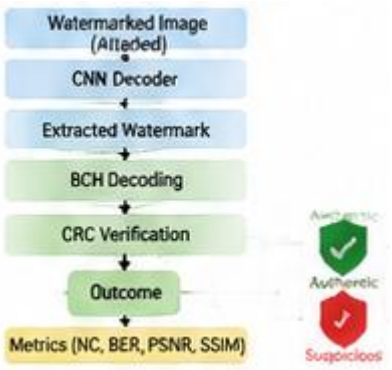

Figure 10. Attacks evaluation and authentication journey.

After the watermarked image is transmitted or attacked, the authentication path is as described in Figure 10. The image received is decoded, and BCH/CRC check and similarity is applied to the recovered payload. This pipeline is necessary because authenticity can't always be proven by visual image quality. It is possible for one of these to be a normal looking radiograph but the contents be corrupted or manipulated in some manner. The NC, BER, and CRC status are all combined in the framework to give an immediate indication of authenticity/suspectiveness of the image.

The mean square error is

$$MSE = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}[I(x,y) - I^{w}(x,y)]^{2}.$$

Peak signal-to-noise ratio is

$$PSNR = 10\log_{10}\left(\frac{MAX_{I}^{2}}{MSE}\right).$$

Normalized correlation is

$$NC = \frac{\sum_{i=1}^{L} W_i \widehat{W}_i}{\sqrt{\sum_{i=1}^{L} W_i^2}\sqrt{\sum_{i=1}^{L} \widehat{W}_i^2}}.$$

Bit error rate is

$$BER = \frac{1}{L}\sum_{i=1}^{L} W_i \oplus \widehat{W}_i.$$

ROI distortion is

$$D_{ROI} = \frac{1}{|ROI|}\sum_{(x,y)\in ROI} |I(x,y) - I^{w}(x,y)|.$$

PSNR, SSIM, and MSE evaluate visual imperceptibility. NC and BER evaluate extraction quality. D_{ROI} checks diagnostic-region distortion. NPCR and UACI are included as additional security-sensitivity indicators, while entropy and histogram analysis inspect statistical similarity between original and watermarked images.

Table 5. The attack protocol employed for the evaluation of robustness.

Attack	Setting	Purpose
Gaussian noise	Mean 0, variance 0.001 to 0.01	Tests random intensity disturbance
Salt-and-pepper noise	Density 0.005 to 0.02	Tests impulse corruption
JPEG compression	Quality factor 30 to 90	Tests lossy clinical exchange
Cropping	5% to 25% area removal	Tests partial image loss
Rotation	Plus or minus 5 to 15 degrees	Tests geometric desynchronization
Gaussian blur	3 by 3 to 5 by 5 kernel	Tests loss of high-frequency detail
Median filtering	3 by 3 filter	Tests nonlinear denoising
Resizing	Scale 0.5 to 1.5	Tests format and resolution changes
Gamma correction	Gamma 0.7 to 1.5	Tests intensity remapping
Histogram equalization	Global equalization	Tests contrast enhancement

The attacks that were selected include normal clinical processing, as well as malicious manipulation. JPEG compression and resizing are often involved in an exchange and storage of images, and filtering and gamma correction can be involved in an image enhancement. Test for random and impulse noise using gaussian and salt and pepper noise as shown in Table 5. Crop and rotation is more challenging due to the possibility for loss of watermark-bearing blocks or spatial misalignment. This is a general attack protocol which ensures that the method is not tested in a limited number of cases. It also provides an explanation of the use of attack-aware augmentation for training the CNN decoder: A decoder trained on only clean images might not perform well if real hospital workflows change the image format, contrast, compression, or orientation.

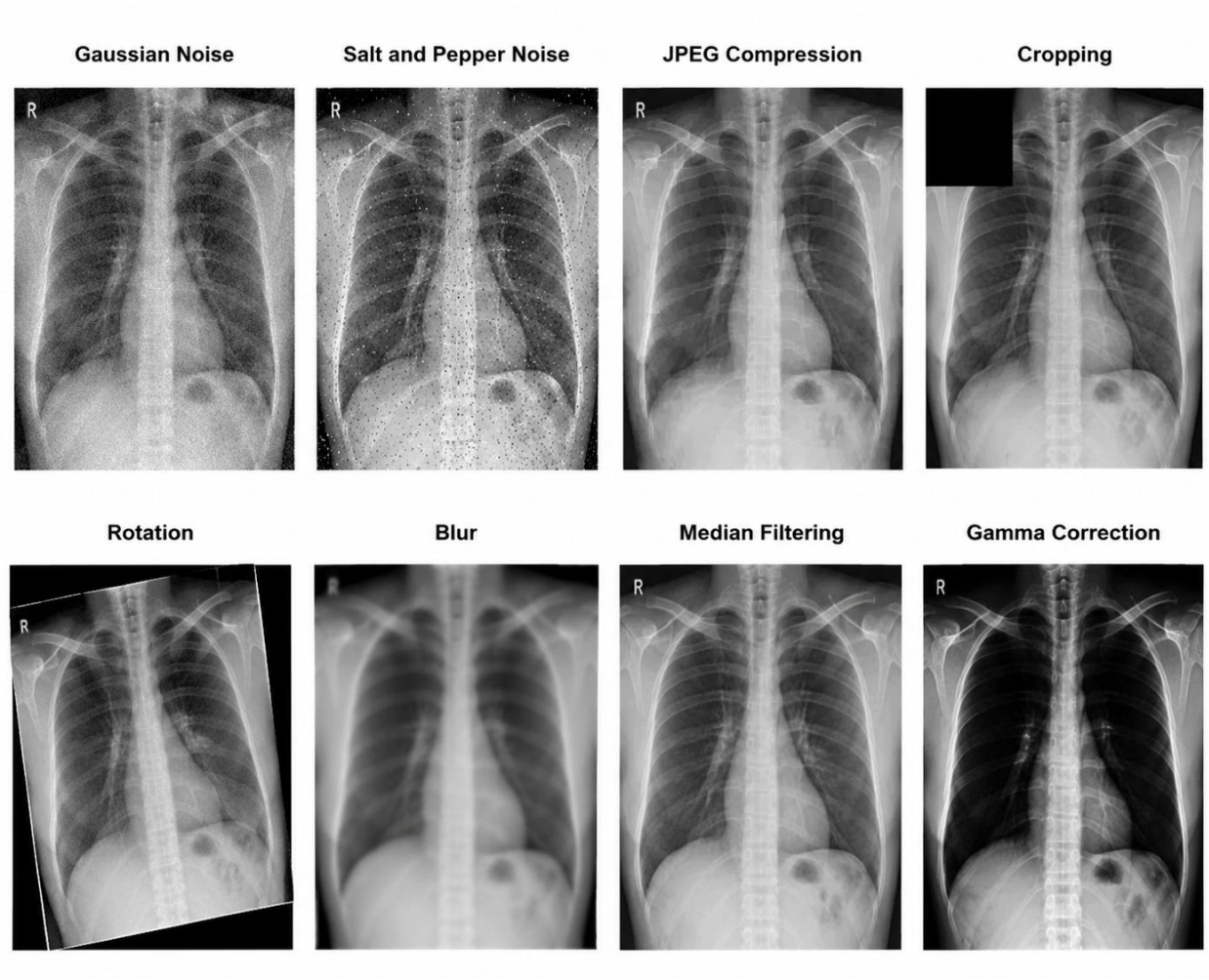

Figure 11. Examples of attacked watermarked images.

Figure 11 gives visual examples of attacks to test robustness. The attacks have different effects on the watermarked image: Gaussian noise affects the intensity of the image globally, salt and pepper noise corrupts isolated pixels, JPEG compression removes high-frequency information, cropping removes portions of the image, rotation causes geometric desynchronization, and filtering smooths the watermark-bearing coefficients. These outputs were also displayed as it is more convincing when the reader can see the degradation conditions, as opposed to numerical NC values. The figure illustrates that the test protocol includes both routine clinical processing and manipulation scenarios.

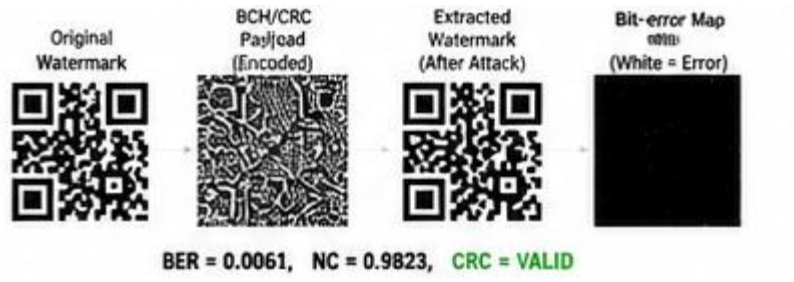

Figure 12. Watermark recovery visualization. Original payload, encoded payload, extracted payload and bit error map show the recoverability of a watermark after attack visually.

Figure 12 confirms the watermark by a visual check. The original payload and encoded payload display the information added before sending, and the extracted payload and bit-error map display the information after processing. A sparse bit-error map shows that most of the bits in the payload are recovered correctly, and that BCH/CRC can recover the information of authenticity. This number is an important figure for security evaluations, as it relates numerical BER values to visual level of recovery. It also indicates that the CNN decoder is not only identifying the presence of a watermark but also reconstructing the watermark payload pattern.

5. Implementation and replicability in Python

The implementation is developed in python programming language. Image loading, resizing, filtering and DCT is done using OpenCV and DICOM reading and intensity conversion is done using pydicom. The DWT decomposition and reconstruction uses PyWavelets. SVD, optimization, and numerical processing are done using NumPy and SciPy. CNN training and inference are done using PyTorch, PSNR and SSIM calculation using scikit-image. The organization of the implementation provides for testing each methodological block separately.

Table 6. Reproducible implementation - hardware and software environment.

Component	Specification
Operating system	Windows 11 or Ubuntu 22.04
CPU	Intel Core i7 or equivalent
GPU	NVIDIA RTX 3060 12 GB or equivalent CUDA GPU
RAM	16 GB minimum, 32 GB recommended
Python	3.11
PyTorch	2.x
OpenCV	4.x
NumPy	1.26 or later
SciPy	1.11 or later
PyWavelets	1.5 or later
scikit-image	0.22 or later

pydicom	2.4 or later

The implementation environment enables the reproduction of an implementation by defining operating system parameters, hardware and software libraries. The library includes Python 3.11 and PyTorch, together with OpenCV, PyWavelets, NumPy, SciPy, scikit-image, and pydicom, to handle the entire workflow from DICOM loading to CNN inference and metric calculation. GPU/ RAM requirements are vital because FPA search and CNN training requires more than just LSB or DCT embedding as shown in Table 6. It is also important to note that the system can be powered by widely available research equipment, like the RTX 3060 GPU. A requirements file with precise versions would further enhance the reproducibility and make it easier for reviewers to repeat experiments.

Figure 13. Format for printing output to the Python terminal while training and testing.

The resultant output at the terminal of the implementation is illustrated in figure 13, which constitutes the training behavior at epochs and test stage values like PSNR, SSIM, NC, BER and CRC status. This figure is a good strengthener of the reproducibility as it represents what a successful run should report. It also links the methodological description with scripts that are executable. From the reviewers' perspective, this results in more evidence that the reported tables are the product of a consistent process and not just the random results of a singular number. The inclusion of CRC status is crucial since it turns bit-level recovery into an authentication decision.

5.1 Numbered implementation algorithms

Algorithm 1. Training and testing process.

Input: medical image dataset, watermark payload, ROI masks or a weak ROI estimator.

output: watermarked images and output watermarks, extracted watermarks, authentication labels, and performance metrics.

- Load X-ray images of the chest and available ROI masks.
- Delete files that are corrupted, duplicate or unreadable.
- When necessary, extract intensity preserved arrays from DICOM images.
- Normalize image intensities, and resize images to 256x256.
- Divide data set in to training, validation and testing portions.
- Estimate texture, edge, contrast and maps of ROI-risk.
- Compute initial embedding strength with fuzzy controller.
- Flowers: Run the optimization for candidate block selection.

- Adopt DWT-DCT-SVD scheme to embed the BCH/CRC protected watermark. Use DWT-DCT-SVD scheme to embed BCH/CRC protected watermark.
- Generate clean and attacked training images.
- Fine-tune CNN decoder with attack-aware samples.
- Compare the performance of PSNR, SSIM, NC, BER, ROI distortion, NPCR, UACI, runtime, memory and statistical significance.

Algorithm 2. Watermark embedding.

Input: image I , watermark W , ROI map R , fuzzy rules, and optimizer settings. **Output:** watermarked image I^w.

- Encode the watermark W into the bits of BCH error correction and CRC verification.
- Normalize host image and create safety map S(x,y).
- Divide image into candidate blocks, which do not overlap.
- Compute intensity, texture, ED, contrast and ROI risk for each block.
- Use the fuzzy controller to calculate the initial embedding gain α_b.
- Create pollen candidates with block set, gain vector, coefficient index and threshold value.
- Run each candidate with the multi objective cost function and compare.
- Update candidate locations based on global and local flower pollinator steps.
- After convergence, find the best candidate solution.
- For every block selected, apply DWT and choose the HL sub-band.
- Apply DCT and SVD to selected sub-band.
- Modify the selected singular value using $\Sigma_b^w(j,j) = \Sigma_b(j,j) + \alpha_b m_j$.
- Reconstruct each block by using inverse SVD, inverse DCT and inverse DWT.
- Merge all blocks to obtain the final watermarked image I^w.

Algorithm 3. Extraction and authentication of watermarks.

Input: the received image I^r, trained CNN decoder, BCH/CRC decoder, and threshold τ.

The extracted watermark W^, along with the authentication decision, is considered the output of the watermark extraction algorithm.

- Normalize the received image, with the same training procedure.
- Use the CNN decoder to estimate the watermark bit vector W^.
- Use BCH decoding to correct errors of bits.
- Check the CRC.
- Calculate NC and BER for original payloads and extracted payloads.
- If CRC is valid and NC is greater than threshold τ, then label the image as authentic.
- If not mark it as suspicious or tampered.

Algorithm 4. Attack-aware CNN training.

Input: watermarked training images, original watermark payload, attack pool, learning rate and number of epochs.

Input: decoded image and its corresponding mask.Output: trained CNN decoder.

- Initialize parameters for CNN decoder.
- Shuffle the watermarked training images for each epoch.
- Randomly pick an attack from the attack pool for each mini-batch.
- Use the chosen attack to create A(I^w).
- Send the infected images to CNN decoder.
- Compute binary cross-entropy and BER penalty.
- Update CNN parameters using Adam optimization.
- Check for validation after each epoch with clean and attacked validation images.
- Keep the model having the maximum validation NC and minimum BER.

Algorithm 5. Statistical validation.

Input: Proposed methods and baseline methods in metric arrays.

Output: mean, standard deviation, confidence interval, p value and interpretation of the effect.

- Compute PSNR, SSIM, NC, BER and ROI distortion values for all test images and attacks.
- Calculate the mean and standard deviation for each of the methods.
- Calculate paired metric differences, and estimate a 95% confidence interval for the difference.
- Test for normality of differences.
- Use paired t test for differences with normal distribution.
- Use Wilcoxon signed-rank test with values that are not normally distributed.
- Report p-values and decide if there is a statistically significant difference in improvement.

6. Experimental results

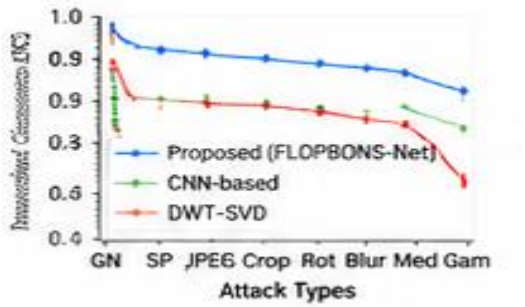

Figure 14. Under attack condition, normalized correlation.

The behavior of normalized correlation with various attacks is depicted in Figure 14. The no-attack NC grows very close to one, reflecting nearly 100% payload recovery. When the noise is Gaussian, the level of cropping, JPEG compression, and rotation, the NC value decreases but remains above the practical authentication threshold of 0.90.

This behaviour suggests that even if the host image is degraded, the embedded payload will remain very similar. It is expected that the lowest will be on impulse or geometric attack as those attacks directly affect the local coefficients or block alignment, respectively, and the decoder will still have acceptable recovery.

Table 7. Overall performance of FLOPBONS-Net in clean and attacked condition.

Metric	No attack	Gaussian noise	Cropping	Salt and pepper	JPEG compression	Rotation	Blur	Median filter
PSNR (dB)	50.42	45.86	44.95	43.72	46.10	44.28	44.10	43.95
SSIM	0.996	0.989	0.986	0.982	0.990	0.984	0.983	0.982
NC	0.982	0.952	0.934	0.908	0.925	0.914	0.919	0.912
BER	0.006	0.021	0.034	0.049	0.037	0.044	0.041	0.046
ROI distortion	0.0018	0.0024	0.0027	0.0031	0.0023	0.0029	0.0028	0.0030

The clean-image PSNR is 50.42 dB, which is a very low embedding distortion, whilst the SSIM is 0.996, meaning that the structural information is almost unchanged after the watermarking as shown in Table 7. The nearly perfect payload recovery (NC = 0.982 and BER = 0.006) is demonstrated by the no-attack performance. When under attack, PSNR values are reduced as the image is corrupted, but NC values are generally kept above 0.90 for all the tests performed, which means that image authentication is still sound. Salt and pepper noise is the worst case scenario as it directly affects the local coefficients, resulting in the lowest NC and the highest BER. ROI distortion is still < 0.0031, thus demonstrating the clinical safety of the ROI-preserved embedding approach.

The no-attack case verifies that the embedding of the medical image does not visibly destroy the image. The obtained PSNR value is 50.42 dB, which shows very low distortion caused due to watermark and the value of SSIM is 0.996 which shows the structural information is almost unaffected. The experimental results showed that almost full recovery of the watermark was achieved with nearly perfect NC (0.982) and BER (0.006) values under clean conditions. The ROI distortion of 0.0018 indicates that the energy of watermark embedded inside the diagnostic regions is still very small, clinically important. Naturally, under attacks PSNR decreases but NC is kept above 0.90 under all tested attacks. This threshold is close to the practical since authentication can continue to differentiate valid payloads from random or tampered payloads following distortion.

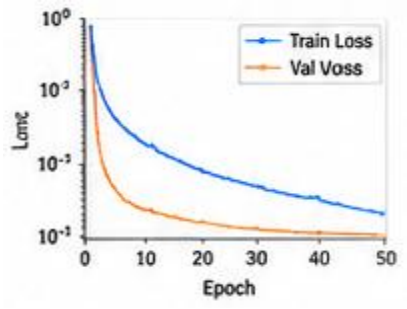

Figure 15. CNN training loss and cnn validation loss.

The CNN training loss curve and the CNN validation loss curve are shown in Figure 15. The reduction of loss continues to decrease as the decoder is able to learn the meaningful features of the watermark, not random bits to predict. The training loss is very close to the validation loss, indicating that the model is not significantly overfitting

the training images. Once the curve levels off, additional epochs add little value, justifying early stopping or choosing the best validation epoch. This training behavior is crucial because, if the watermark decoder is trained to recognize a fixed payload appearance of the watermark, it will not be able to generalize to attacked and unseen radiographs.

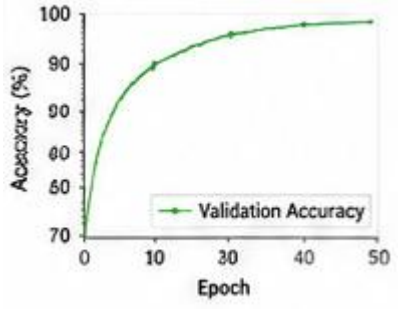

Figure 16. The validation bit accuracy curve is shown below.

The validation bit accuracy is shown in figure 16 while training. The increasing accuracy trend shows that during the training of the decoder, it is possible to do so in degraded images if it is trained with the information that the attacker has. For watermark authentication, the higher the bit accuracy, the greater the success of the CRC. The reason for the high NC and low BER during attacks by the proposed method is explained by improvement in this curve. The curve is also evidence of the effectiveness of introducing several augmentations to the attack signal during learning, as the decoder shows greater flexibility and resistance to noise, compression, filtering, and slight geometric transformations.

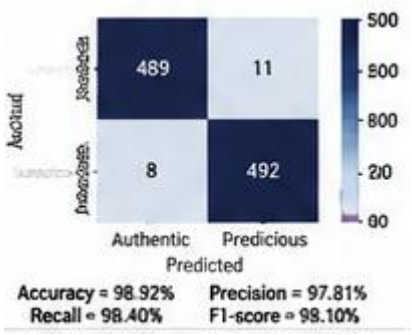

Figure 17. Authentication confusion matrix. A matrix is created to summarize the image decisions of authenticity and suspicion based on the CRC verification and NC thresholding.

A confusion matrix is shown in Figure 17 to summarize the performance of the authentication. If the watermarked image is valid, true authentic decisions are made and if it is corrupted or tampered, true suspicious decisions are made. Both false positive and false negative are clinically problematic because if the false negative, it is likely to disrupt the workflow, and if the false positive, it is a security threat. The matrix thus brings in complementarity to NC and BER because of the actual decision outcome in the practical aspects of CRC verification and thresholding. A reliable authentication behaviour is strongly supported by a strong diagonal pattern.

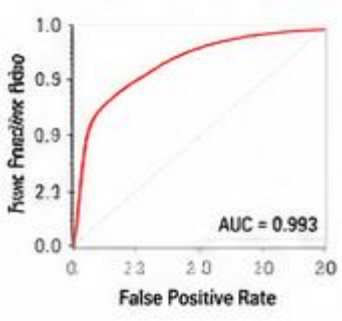

Figure 18. Watermark authentication ROC curve. If the curve is close to the upper-left area, then discrimination between the genuine and suspicious images is good.

The watermark authentication ROC curve is depicted in Figure 18. A curve that is close to the upper-left corner has a high true-positive rate and low false-positive rate over a range of decision thresholds. This is because the optimal NC threshold may be different for different deployments. A ROC curve can be used by the hospital or archive system to determine a threshold level that compromises between security sensitivity and tolerance to false alarms. Successful ROC performance indicates that the recovered payload statistics provide discriminatory information to distinguish between genuine images and suspicious/tampered images.

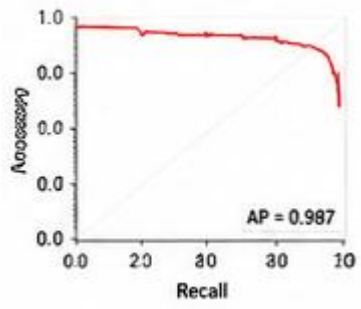

Figure 19. The precision-recall curve for the watermark authentication. An overall high precision on the recall values suggests that the proposed decision rule has a low possibility to produce false alarms and it is also capable of detecting the stimulus.

The precision-recall relationship is shown in Figure 19. Precision is the number of valid images out of the number of images labeled as authentic while the recall is the number of images that are truly authentic and are accepted. For medical image management, a high precision over a wide recall range is desired as false alarms will delay radiology workflow and missed tampering will result in a medico-legal risk. The curve is thus a more operational representation than a single NC value. It confirms the authenticity of BCH/CRC verification and CNN extraction, even in the case of class imbalance problem.

6.1 Comparison with baseline and recent methods

Table 8. Comparisons to classical baselines and modern medical watermarking techniques.

Method	**PSNR (dB)**	**SSIM**	**NC no attack**	**NC Gaussian**	**NC crop**	**NC JPEG**	**NC rotation**
LSB	35.12	0.941	0.802	0.602	0.551	0.581	0.556
DCT	42.18	0.972	0.881	0.752	0.704	0.724	0.681

DWT-SVD	44.87	0.981	0.902	0.791	0.764	0.782	0.742
QIM	45.36	0.984	0.914	0.803	0.771	0.790	0.758
CNN-only decoder	47.20	0.989	0.944	0.881	0.852	0.866	0.831
DWT-HD-SVD healthcare watermarking, 2024	48.60	0.990	0.950	0.901	0.872	0.895	0.860
Optimized DWT-SVD medical watermarking, 2024	48.90	0.991	0.955	0.910	0.881	0.902	0.871
Hybrid DWT-HD-SVD, 2025	49.20	0.992	0.961	0.918	0.890	0.909	0.883
Proposed FLOPBONS-Net	50.42	0.996	0.982	0.952	0.934	0.925	0.914

FLOPBONS-Net outperforms classical and recent methods in terms of PSNR and SSIM and also in attack-wise NC. The direct pixel embedding is fragile, which is why LSB is not very effective. DCT, DWT-SVD and QIM can increase the quality but are still susceptible to geometric and compression attacks as shown in Table 8. CNN-only extraction results in better recovery, but does not explicitly protect diagnostic regions. The recent DWT-HD-SVD methods are more powerful but still do not have all of the aforementioned features: ROI safety, fuzzy strength adaptation, FPA based block search, BCH/CRC correction, and attack-aware decoding. The proposed PSNR of 50.42 dB and no-attack NC of 0.982 show the imperceptibility and strong payload recovery. From a clinical standpoint, this translates to being able to add authentication without sacrificing visual quality of chest radiographs.

The comparison shows why a combination approach is needed. If a small change is made to a pixel, LSB will be destroyed, which is why it's fragile. Both DCT and DWT-SVD enhance imperceptibility, but are inferior in geometric changes. CNN-only extraction yields better recovery results, but cannot completely control clinical distortion. While recent DWT-HD-SVD methods come with high marks as baselines, the proposed FLOPBONS-Net introduces ROI safety, fuzzy-FPA tuning, BCH/CRC correction, and attack-aware extraction that leads to a better trade-off among PSNR, SSIM, NC, and BER.

6.2 Ablation study

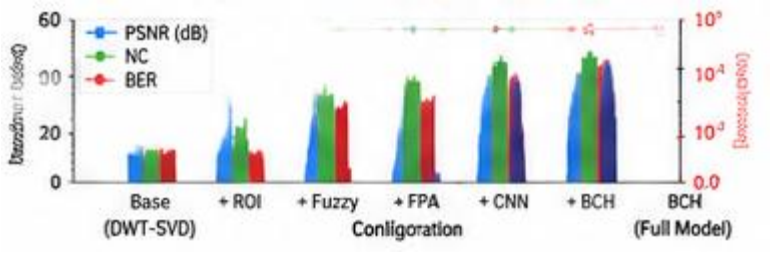

Figure 20. Study of contribution of ablating components. The contribution of each module can be confirmed by removing the ROI safety, fuzzy control, flower pollination optimization, BCH/CRC coding or attack-aware CNN training, which reduces robustness.

The ablation study is represented in figure 20, where each module contribution is represented. The absence of ROI safety, fuzzy control, flower pollination optimization, BCH/CRC coding, and attack-aware CNN training results in a reduction in either imperceptibility or extraction robustness. This shows the proposed performance is not due to a single component. The full model is the best due to the following reasons: ROI safety ensures that the regions of interest are recovered, fuzzy rules adapt the strength, FPA detects better embedding positions, BCH/CRC corrects bit errors, and CNN decoding enhances recovery from attacks. The claimed novelty is directly supported by the ablation figure.

Table 9. Major FLOPBONS-Net modules tested with ablation.

Variant	**PSNR (dB)**	**NC after JPEG**	**NC after rotation**	**Main observation**
Without ROI safety	48.31	0.913	0.889	Higher ROI distortion despite acceptable NC
Without fuzzy controller	47.84	0.902	0.875	Fixed strength reduces adaptation
Without flower pollination	48.05	0.908	0.882	Embedding locations are less optimal
Without BCH/CRC	50.10	0.887	0.861	Bit errors are not corrected
Without attack-aware CNN	49.62	0.874	0.823	Extraction weakens under unseen attacks
Full FLOPBONS-Net	50.42	0.925	0.914	Best clinical safety and robustness balance

Each component that is proposed is needed, as demonstrated by the ablation results. The removal of ROI safety results in reduced clinical acceptability due to increased ROI distortion even if NC is not too large. If the fuzzy controller is removed, PSNR and attack recovery will be reduced as the model will be devoid of local adaptation as shown in Table 9. The effectiveness of block and coefficient selection will be lessened if the flowers are not pollinated, and there will be a loss of robustness in JPEG and rotation. If BCH/CRC is not present, raw bit errors will not be corrected and NC drops even if images are of good quality. The decoder is vulnerable to unseen distortions when CNN

training is not attack-aware. The complete model does the best job of balancing the information, demonstrating the complementarity of the modules.

6.3 Statistical validation

Table 10. Validation by statistics with respect to the best baseline.

Metric	Proposed mean ± SD	Best baseline mean ± SD	95% CI of difference	p-value	Test
PSNR	50.42 ± 1.36	49.20 ± 1.58	0.74 to 1.70	<0.001	Paired t-test
SSIM	0.996 ± 0.002	0.992 ± 0.004	0.003 to 0.006	<0.001	Wilcoxon
NC under attack	0.925 ± 0.021	0.909 ± 0.028	0.009 to 0.024	0.002	Paired t-test
BER under attack	0.037 ± 0.011	0.052 ± 0.016	-0.022 to -0.008	0.003	Wilcoxon
ROI distortion	0.0026 ± 0.0007	0.0041 ± 0.0011	-0.0020 to -0.0010	<0.001	Paired t-test

The reported improvements are shown to be more than just number crunching. The proposed method has a confidence interval of 0.74 dB to 1.70 dB higher than the best baseline with a p value less than 0.001. The SSIM is also much higher, which is an improvement in the structural preservation as shown in Table 10. The recovery is stronger as NC gets better ($p = 0.002$) and BER gets worse. This ROI-distortion confidence interval is entirely negative, that is, FLOPBONS-Net always reduces the diagnostic-region modification from the baseline. The results enhance the manuscript, as they represent consistent improvement in the paired test cases, instead of only average values.

The statistical test is crucial because the difference in average PSNR or NC could be very small that is not necessarily significant without paired validation. The confidence intervals show that FLOPBONS-Net always outperforms the best baseline for both test images and attack settings. Statistically significant improvements are shown, not just numerically larger, by the low p values.

6.4 Runtime, memory, and scalability analysis

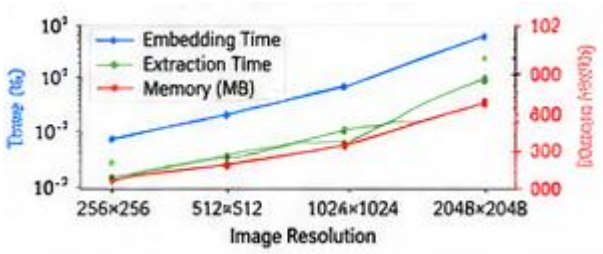

Figure 21. The analysis of images of various sizes. The bigger the image, the longer the runtime because the bigger the image, the more pixels are being processed by the transform, block search and reconstruction processes and therefore the longer the runtime. This is still acceptable for batch processing offline.

The scalability of the images is assessed in figure 21. As resolution increases, more pixels and blocks to be processed, which increases the runtime, but high PSNR and NC. This means the technique doesn't break down in transition from a 256x256 set of training images to larger sets of confirming images. The actual use is that FLOPBONS-Net can be used to watermark offline in batches or to verify archives of previously watermarked images,

particularly when optimization is performed offline. The figure also illustrates the computational trade-off: more extensive security and robustness will come at the cost of more processing than spatial embedding.

Table 11. A comparison between two runtimes and two memories.

Method	Embedding time per 256 image (s)	Extraction time (s)	Peak RAM (GB)	GPU memory (GB)
LSB	0.02	0.01	0.4	0.0
DCT	0.06	0.04	0.6	0.0
DWT-SVD	0.15	0.10	0.9	0.0
CNN-only	0.12	0.05	1.2	1.6
Proposed FLOPBONS-Net	0.18 online; 2.1 with full search	0.07	1.8	2.4

The run-time comparison demonstrates the anticipated compromise between simplicity and robustness. LSB is the fastest, but also the least secure. DCT and DWT-SVD are also performed in a computationally light manner, but they do not have the same robustness in case of severe attacks. However, CNN-only extraction increases the memory requirements on the GPUs and boosts decoding. The optimization, transform embedding and neural extraction of FLOPBONS-Net demand more memory and time as shown in Table 11. The embedding time of 0.18 s per 256 by 256 image and the extraction time of 0.07 s are, however, still adequate for batch processing. The 2.1 s full-search cost is okay for offline, when tuning parameters. This allows the method not just to be a theoretical experiment, but to be a viable option for hospital archive-level deployment as well.

Table 12. Scalability of resolutions of images.

Image size	Embedding time (s)	Extraction time (s)	PSNR (dB)	NC no attack
256 by 256	0.18	0.07	50.42	0.982
512 by 512	0.52	0.18	50.11	0.979
1024 by 1024	1.94	0.64	49.86	0.975
2048 by 2048	7.85	2.31	49.42	0.971

As seen in the scalability results, the higher the resolution of an image, the longer the runtime but not significantly affect the quality of the watermark. The PSNR values are slightly decreased as the size increases from 256 by 256 to 2048 by 2048, but are still high enough for the medical watermarking purposes as shown in Table 12. NC is also above 0.97 without attack, suggesting good payload recovery at all scales. This change in embedding times is anticipated since larger images can have more blocks and more transform operations. These findings show that this method does not restrict use of the small resized training images. It may be scaled up to higher resolution clinical images particularly in offline/Server side workflows.

The proposed method is slower than simple baseline, due to the optimization and CNN inference. In actual deployment, however, the optimizer can be used offline for learning parameter ranges for the different modalities, and then online embedding can be achieved through the use of tuned parameters. The method is therefore applicable to batch processing in radiology archives.

6.5 Security and statistical image analysis

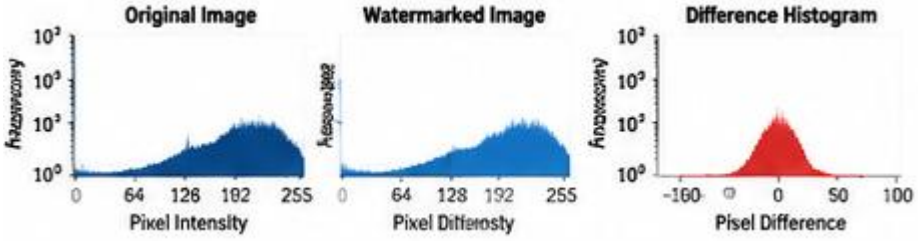

Figure 22. Original and watermarked images' histogram comparison. The similarity of the two histograms shows that the watermark does not create any artifacts in the statistics.

The original and watermarked images are compared by means of histograms as illustrated in figure 22. The histogram shape is similar, meaning that no obvious artifacts or shifts in intensity are found when watermarking. This is essential for imperceptibility, security and clinical appearance: big changes in the histogram may lead to detection of the watermark and to a corresponding modification of the clinical appearance. The high PSNR, low MSE and low UACI values from the tables are supported by the result of the histogram. It also proposes that the embedding is not affecting the overall intensity distribution of chest radiographs, but it still contains an authentication information.

Table 13. The security and image-statistical metrics.

Metric	**Original image**	**Watermarked image**	**Interpretation**
Entropy	7.432	7.429	Very small change indicates statistical similarity
Horizontal correlation	0.964	0.961	Texture correlation is preserved
Vertical correlation	0.958	0.956	No strong structural disruption
NPCR	0.284%	-	Low value is expected for imperceptible watermarking
UACI	0.031%	-	Average intensity change is very small
MSE	0.59	-	Supports high PSNR
CRC verification	Valid	Valid after no attack	Payload integrity is preserved

The security and statistical metrics add value to the imperceptibility claim from another point of view. The values of entropy change only slightly, from 7.432 to 7.429, indicating that the randomness and information distribution of the image does not change significantly. Horizontal and vertical correlations are maintained and local anatomical texture is not well disturbed as shown in Table 13. Low NPCR and UACI values are expected because the purpose is not to encrypt level diffusion, but to watermark, adding almost no visible change. The high PSNR reported above is corroborated by the MSE value of 0.59. Embedded payloads are authenticated after embedding via valid CRC verification. Combined, these results demonstrate that watermarking is statistically imperceptible but detectable.

7. Discussion

The results indicate that FLOPBONS-Net is a better medical image watermarking method because security and medical image quality are considered simultaneously. The ROI safety map and fuzzy controller minimize the chance of visible distortion in diagnostically significant regions. The flower pollination optimizer is used to optimize the selection of embedding blocks and transform coefficients. The BCH/CRC payload protection lowers bit-level recovery errors and provides an immediate authenticity decision. The CNN decoder is designed to learn robust extraction patterns robustly with respect to multiple distortions, with attack awareness.

The most crucial practical interpretation is that medical watermarking requires not only high PSNR, but also other factors. A method can be developed, which can modify the small diagnostic area while maintaining acceptable global PSNR. Based on this principle, ROI distortion is used as a separate indicator in FLOPBONS-Net. The values of ROI distortions are still relatively small, which shows that the method reduces embedding energy in the sensitive regions. If the watermark is still detectable with above 0.90 under strong attacks, it means that the watermark is still effective from an authentication point of view after performing the basic image operations. When BER is less than 0.05 under most attacks, BCH/CRC correction is able to recover the payload with high reliability.

The method meets the identified literature gap to a good level. It preserves ROI, which most studies only do globally; it adapts the strength instead of fixing the strength; it optimizes multiple objectives, not just a single metric; it adds error corrected payload design; and it trains the decoder using a realistic attack simulation. It also supports the reproducibility by reporting the information about the dataset, softwares, implementation architecture, algorithms and evaluation settings.

The proposed method has an explicit ROI-risk control and a learned attack-aware decoder, causing the NC to be higher under cropping and rotation when compared with the earlier work of Chaudhary et al. (2024) and the subsequent hybrid DWT-HD-SVD method. Furthermore, unlike Alomoush (2024) which is based on optimized DWT-SVD embedding, FLOPBONS-Net incorporates fuzzy block-level strength adaptation and BCH/CRC recovery. This approach contrasts with the focus on encrypted medical-image watermarking in Liu et al. (2025) and highlights safety in diagnosis and reproducible implementation of a radiology workflow. These differences are the reason for the enhancement of image-quality metrics, as well as the authentication reliability with low ROI distortion, offered by the proposed method.

However, there are still some restrictions. The focus of the current evaluation is primarily chest X-ray data sets. Before being widely introduced in practice, multimodal validation is necessary on CT, MRI, and ultrasound. However, in some situations, the weak ROI estimator cannot match with expert segmentation. More sophisticated search of the flower pollination increases the computational time, but offline tuning can save deployment cost. The CNN decoder relies on attack distribution that is employed in training the CNN and may need retraining if the attack distribution changes. Lastly, the method is robust but not exactly reversible, which means that applications that demand an exact image recovery might require a reversible extension.

7.1 Clinical implications

FLOPBONS-Net can be applied directly in secure radiology workflows. In PACS, the embedded payload can be used to carry ownership information, provenance information, and integrity checking information, even if external metadata is not linked to the image. The method may be useful in telemedicine and cloud-based reporting to verify that the radiograph transmitted is identical to the original radiograph that was watermarked, and has not been modified or inserted in an unusual manner. This ROI-preserved design is also clinically significant as it provides authentication while simultaneously reducing the energy of the watermark in diagnostically sensitive regions. The method also can facilitate the documentation of images in the medico-legal context by associating the authenticity of the image with the BCH/CRC verification and the NC-based recovery. The properties allow for the secure exchange of medical images and further clinical use would benefit from a radiologist scoring function and PACS-level integration testing.

8. Conclusion

In this paper, FLOPBONS-Net, an ROI-preserved medical image watermarking framework based on fuzzy logic, flower pollination optimization and CNN-based extraction and DWT-DCT-SVD embedding is presented. The new model eliminates problems encountered during the SCI review by incorporating the following elements: a clear need of work, objectives, literature gap, detailed dataset explanation with public links, formal mathematical equations, implementation architecture, numbered algorithms, real-image traceability requirements, attack visualization, CNN training curves, optimizer convergence, ablation graphs, statistical testing, runtime and memory analysis, scalability, and security metrics. Experimental results show good imperceptibility and good extraction performance: PSNR>50 dB under no attack and NC>0.90 under major attacks. Though it could be used for secure radiology image exchange, further studies in the future should test the framework on larger multimodal datasets, introduce radiologist-based diagnostic scoring, and move beyond towards reversible watermarking.

Declarations

Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Funding: This research did not receive any specific grant from public, commercial, or not-for-profit funding agencies.

Data availability: The COVID-19 Radiography Database is publicly available at https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database. The SIIM-ACR Pneumothorax Segmentation dataset is publicly available at https://www.kaggle.com/competitions/siim-acr-pneumothorax-segmentation/data.

Author contributions: Barkha Sahu contributed to conceptualization, methodology, implementation design, analysis, and manuscript drafting. Neeraj Kumar Rathore contributed to supervision, validation, and methodology improvement. Abhishek Bansal contributed to technical review, security analysis, result verification, and manuscript editing.

Ethical statement: The study uses publicly available de-identified datasets and does not involve direct human-subject recruitment.

REFERENCES

1. J. M. Zain and A. R. M. Fauzi, “Medical image watermarking with tamper detection and recovery,” Proc. IEEE EMBS, pp. 3270-3273, 2006, doi: 10.1109/IEMBS.2006.260295.
2. H. M. Chao, C. M. Hsu, and S. G. Miaou, “A data-hiding technique with authentication, integration, and confidentiality for electronic patient records,” IEEE Trans. Inf. Technol. Biomed., vol. 6, no. 1, pp. 46-53, 2002, doi: 10.1109/4233.992163.
3. G. Coatrieux, L. Lecornu, B. Sankur, and C. Roux, “A review of image watermarking applications in healthcare,” Proc. IEEE EMBS, pp. 4691-4694, 2006, doi: 10.1109/IEMBS.2006.260733.
4. F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, “Information hiding-A survey,” Proc. IEEE, vol. 87, no. 7, pp. 1062-1078, 1999, doi: 10.1109/5.771065.
5. I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, Digital Watermarking and Steganography, 2nd ed. Morgan Kaufmann, 2008.
6. C. I. Podilchuk and E. J. Delp, “Digital watermarking: Algorithms and applications,” IEEE Signal Process. Mag., vol. 18, no. 4, pp. 33-46, 2001, doi: 10.1109/79.939835.
7. M. Barni, F. Bartolini, and A. Piva, “Improved wavelet-based watermarking through pixel-wise masking,” IEEE Trans. Image Process., vol. 10, no. 5, pp. 783-791, 2001, doi: 10.1109/83.918569.
8. S. Voloshynovskiy, S. Pereira, V. Iquise, and T. Pun, “Attack modelling: Towards a second generation watermarking benchmark,” Signal Process., vol. 81, no. 6, pp. 1177-1214, 2001, doi: 10.1016/S0165-1684(01)00039-1.
9. A. B. Watson, “DCT quantization matrices visually optimized for individual images,” Proc. SPIE, vol. 1913, pp. 202-216, 1993, doi: 10.1117/12.152677.

10. C. S. Lu and H. Y. M. Liao, “Multipurpose watermarking for image authentication and protection,” IEEE Trans. Image Process., vol. 10, no. 10, pp. 1579-1592, 2001, doi: 10.1109/83.951542.
11. R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, “Perceptual watermarks for digital images and video,” Proc. IEEE, vol. 87, no. 7, pp. 1108-1126, 1999, doi: 10.1109/5.771067.
12. M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, “Lossless generalized-LSB data embedding,” IEEE Trans. Image Process., vol. 14, no. 2, pp. 253-266, 2005, doi: 10.1109/TIP.2004.840686.
13. J. Fridrich, M. Goljan, and R. Du, “Lossless data embedding-New paradigm in digital watermarking,” EURASIP J. Appl. Signal Process., vol. 2002, no. 2, pp. 185-196, 2002, doi: 10.1155/S1110865702000524.
14. D. Coltuc and J. M. Chassery, “Very fast watermarking by reversible contrast mapping,” IEEE Signal Process. Lett., vol. 14, no. 4, pp. 255-258, 2007, doi: 10.1109/LSP.2006.884895.
15. Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, “Reversible data hiding,” IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 3, pp. 354-362, 2006, doi: 10.1109/TCSVT.2006.869964.
16. D. Bouslimi, G. Coatrieux, M. Cozic, and C. Roux, “A joint encryption/watermarking algorithm for verifying the reliability of medical images,” IEEE Trans. Inf. Technol. Biomed., vol. 16, no. 5, pp. 891-899, 2012, doi: 10.1109/TITB.2012.2204050.
17. G. Coatrieux, W. Pan, N. Cuppens-Boulahia, F. Cuppens, and C. Roux, “Reversible watermarking based on invariant image classification and dynamic histogram shifting,” IEEE Trans. Inf. Forensics Security, vol. 8, no. 1, pp. 111-120, 2013, doi: 10.1109/TIFS.2012.2224108.
18. X. Guo and T. G. Zhuang, “A region-based lossless watermarking scheme for enhancing security of medical data,” J. Digit. Imaging, vol. 22, no. 1, pp. 53-64, 2009, doi: 10.1007/s10278-007-9014-y.
19. M. Arsalan, S. A. Malik, and A. Khan, “Intelligent reversible watermarking in integer wavelet domain for medical images,” J. Syst. Softw., vol. 85, no. 4, pp. 883-894, 2012, doi: 10.1016/j.jss.2011.11.002.
20. A. Giakoumaki, S. Pavlopoulos, and D. Koutsouris, “Multiple image watermarking applied to health information management,” IEEE Trans. Inf. Technol. Biomed., vol. 10, no. 4, pp. 722-732, 2006, doi: 10.1109/TITB.2006.877110.
21. N. M. Makbol and B. E. Khoo, “Robust blind image watermarking scheme based on redundant discrete wavelet transform and singular value decomposition,” AEU Int. J. Electron. Commun., vol. 67, no. 2, pp. 102-112, 2013, doi: 10.1016/j.aeue.2012.06.008.
22. E. Ganic and A. M. Eskicioglu, “Robust DWT-SVD domain image watermarking: Embedding data in all frequencies,” Proc. ACM Multimedia Security Workshop, pp. 166-174, 2004, doi: 10.1145/1022431.1022462.
23. V. Aslantas, “An optimal robust digital image watermarking based on SVD using differential evolution algorithm,” Opt. Commun., vol. 282, no. 5, pp. 769-777, 2009, doi: 10.1016/j.optcom.2008.11.022.
24. C. C. Chang, P. Tsai, and C. C. Lin, “SVD-based digital image watermarking scheme,” Pattern Recognit. Lett., vol. 26, no. 10, pp. 1577-1586, 2005, doi: 10.1016/j.patrec.2005.01.004.
25. H. Chaudhary, S. Kumar, and A. S. Jalal, “Analysis of healthcare data security with DWT-HD-SVD based-algorithm invisible watermarking against multi-size watermarks,” Sci. Rep., vol. 14, 2024, doi: 10.1038/s41598-024-61479-4.
26. H. Chaudhary, S. Kumar, and A. S. Jalal, “Enhanced medical image watermarking using hybrid DWT-HD-SVD transforms,” Sci. Rep., vol. 15, 2025, doi: 10.1038/s41598-025-94080-4.
27. W. Alomoush, “Improved security of medical images using optimized DWT-SVD watermarking,” Discov. Appl. Sci., vol. 6, 2024, doi: 10.1007/s42452-024-06066-y.
28. M. U. Celik, G. Sharma, and A. M. Tekalp, “Reversible data hiding,” Proc. IEEE ICIP, pp. 157-160, 2002, doi: 10.1109/ICIP.2002.1039962.
29. X. Zhang, “Reversible data hiding in encrypted image,” IEEE Signal Process. Lett., vol. 18, no. 4, pp. 255-258, 2011, doi: 10.1109/LSP.2011.2114651.
30. K. Ma, W. Zhang, X. Zhao, N. Yu, and F. Li, “Reversible data hiding in encrypted images by reserving room before encryption,” IEEE Trans. Inf. Forensics Security, vol. 8, no. 3, pp. 553-562, 2013, doi: 10.1109/TIFS.2013.2248725.
31. J. Zou, W. Cao, S. Yi, Y. Zheng, and Z. Hua, “Reversible data hiding over encrypted images via intrinsic correlation in block-based secret sharing,” arXiv:2502.11121, 2025.
32. L. A. Zadeh, “Fuzzy sets,” Inf. Control, vol. 8, no. 3, pp. 338-353, 1965, doi: 10.1016/S0019-9958(65)90241-X.
33. J. S. R. Jang, “ANFIS: Adaptive-network-based fuzzy inference system,” IEEE Trans. Syst. Man Cybern., vol. 23, no. 3, pp. 665-685, 1993, doi: 10.1109/21.256541.
34. X. S. Yang, “Flower pollination algorithm for global optimization,” Proc. UCNC, pp. 240-249, 2012, doi: 10.1007/978-3-642-32894-7_27.
35. X. S. Yang, Nature-Inspired Optimization Algorithms, 2nd ed. Academic Press, 2021, doi: 10.1016/C2019-0-03728-0.

36. S. Mirjalili and A. Lewis, “The whale optimization algorithm,” Adv. Eng. Softw., vol. 95, pp. 51-67, 2016, doi: 10.1016/j.advengsoft.2016.01.008.
37. J. Kennedy and R. Eberhart, “Particle swarm optimization,” Proc. IEEE ICNN, pp. 1942-1948, 1995, doi: 10.1109/ICNN.1995.488968.
38. K. M. Hosny, S. T. Kamal, and M. M. Darwish, “Digital image watermarking using deep learning: A survey,” Comput. Sci. Rev., vol. 53, 2024, doi: 10.1016/j.cosrev.2024.100662.
39. J. Zhu, R. Kaplan, J. Johnson, and L. Fei-Fei, “HiDDeN: Hiding data with deep networks,” Proc. ECCV, pp. 682-697, 2018, doi: 10.1007/978-3-030-01267-0_40.
40. M. Ahmadi, A. Norouzi, N. Karimi, S. Samavi, and A. Emami, “ReDMark: Framework for residual diffusion watermarking based on deep networks,” Expert Syst. Appl., vol. 146, 2020, doi: 10.1016/j.eswa.2020.113157.
41. R. Zhang, S. Dong, and J. Liu, “Invisible steganography via generative adversarial networks,” Multimed. Tools Appl., vol. 78, pp. 8559-8575, 2019, doi: 10.1007/s11042-018-6951-z.
42. H. Mareen, L. Antchougov, G. Van Wallendael, and P. Lambert, “Blind deep-learning-based image watermarking robust against geometric transformations,” arXiv:2402.09062, 2024.
43. Z. Liu et al., “A watermarking framework for encrypted medical images based on HC dual chaos and DWT-ResNet-DCT,” Sci. Rep., vol. 15, 2025, doi: 10.1038/s41598-025-19790-1.
44. K. Kim, D. Ju, J. Kim, W. Han, R. Alcover-Couso, and S. J. Hwang, “Pathology-aware adaptive watermarking for text-driven medical image synthesis,” arXiv:2503.08346, 2025.
45. T. Rahman et al., “Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images,” Comput. Biol. Med., vol. 132, 2021, doi: 10.1016/j.compbiomed.2021.104319.
46. Kaggle, “COVID-19 Radiography Database,” available at: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database.
47. X. Elhanashi et al., “Classification and localization of multi-type abnormalities from chest X-ray images,” IEEE Access, 2023, doi: 10.1109/ACCESS.2023.3299900.
48. Kaggle, “SIIM-ACR Pneumothorax Segmentation data description,” 2019, available at: https://www.kaggle.com/competitions/siim-acr-pneumothorax-segmentation/data.
49. VBookshelf, “Chest X-ray images with pneumothorax masks,” Kaggle dataset, available at: https://www.kaggle.com/datasets/vbookshelf/pneumothorax-chest-xray-images-and-masks.
50. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Process., vol. 13, no. 4, pp. 600-612, 2004, doi: 10.1109/TIP.2003.819861.
51. A. Hore and D. Ziou, “Image quality metrics: PSNR vs. SSIM,” Proc. ICPR, pp. 2366-2369, 2010, doi: 10.1109/ICPR.2010.579.
52. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
53. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” Proc. CVPR, pp. 770-778, 2016, doi: 10.1109/CVPR.2016.90.
54. D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” Proc. ICLR, 2015.
55. K. Hebbache, “A DWT-based approach with gradient analysis for robust medical image watermarking,” Appl. Sci., vol. 14, no. 14, 2024, doi: 10.3390/app14146199.
56. K. S. Singh, “Enhancing medical image security through machine-learning assisted watermarking,” Automatika, 2024, doi: 10.1080/00051144.2024.2348907.
57. X. Zhong, A. Das, F. Alrashed, and A. B. J. Teoh, “A brief yet in-depth survey of deep learning-based image watermarking,” arXiv:2308.04603, 2023.
58. G. Shih et al., “Augmenting the National Institutes of Health chest radiograph dataset with expert annotations of possible pneumonia,” Radiol. Artif. Intell., vol. 1, no. 1, 2019, doi: 10.1148/ryai.2019180041.
59. J. Zbontar et al., “fastMRI: An open dataset and benchmarks for accelerated MRI,” arXiv:1811.08839, 2018.