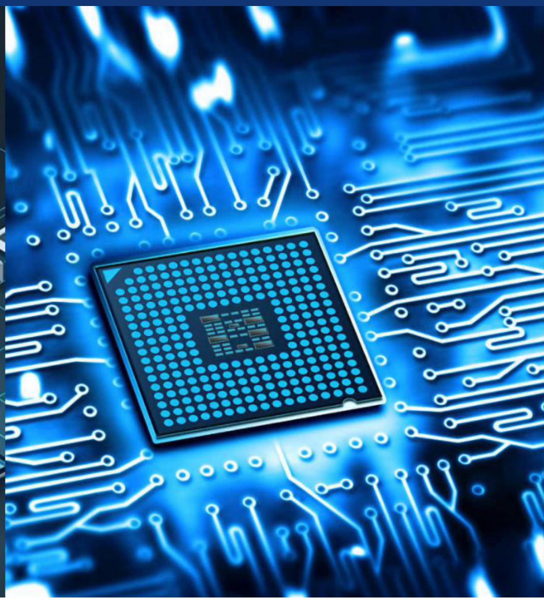
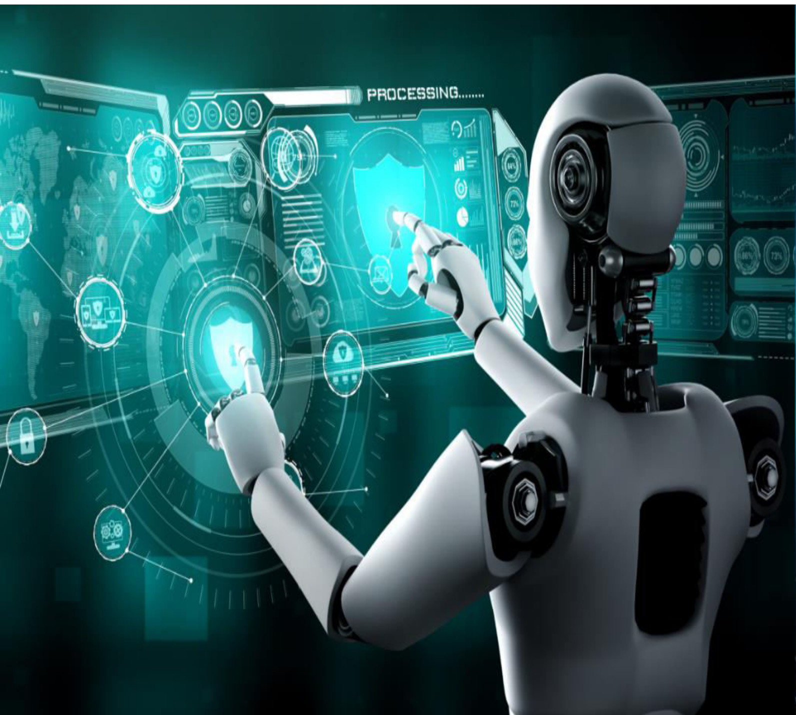
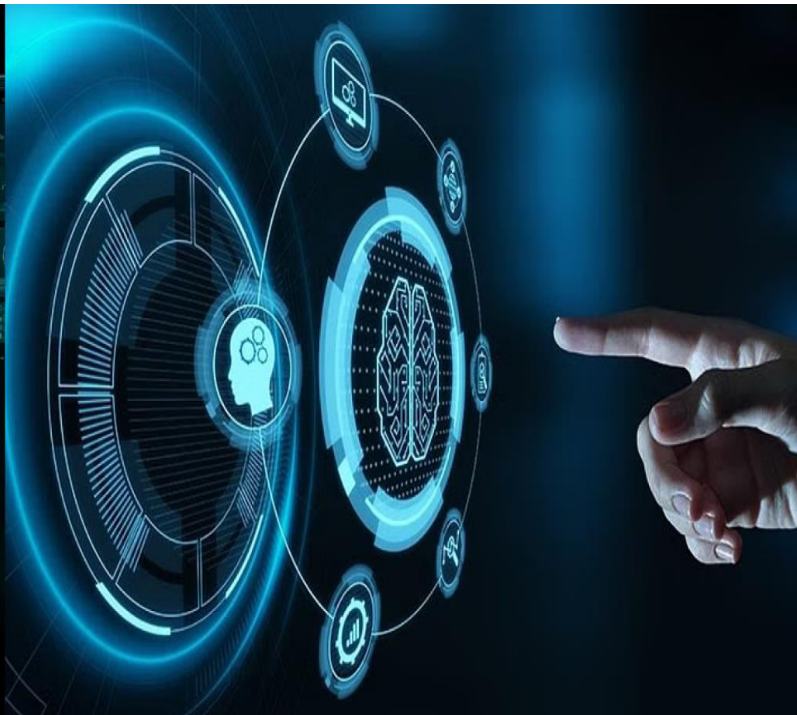
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.927

Volume 14, Issue 6, June 2026

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

MEDICRYPT: Concealing and Securing Patient Data within Medical Images

Koulampeta Yakshitha, Dr. M. Nagaratna

PG Student, Dept. of Computer Science and Engineering, JNTUH University College of Engineering, Science and Technology, Hyderabad, Telangana, India

Professor, Dept. of Computer Science and Engineering, JNTUH University College of Engineering, Science and Technology, Hyderabad, Telangana, India

ABSTRACT: Medical images shared across digital healthcare systems often carry visible patient identifiers that conventional encryption and steganography fail to protect together. This paper presents MEDICRYPT, a privacy-preserving framework that unites Optical Character Recognition (OCR), Gaussian Blur-based redaction, and Least Significant Bit (LSB) steganography. OCR locates patient text on chest X-ray images, Gaussian Blur conceals the detected regions, and the extracted particulars are embedded as a binary payload through LSB. Evaluation on the COVID-19 Chest X-ray Radiography dataset yields high PSNR and SSIM with low MSE, confirming that visible and hidden patient data are protected while diagnostic quality is preserved.

KEYWORDS: Medical image security; Optical Character Recognition; Gaussian Blur redaction; LSB steganography; patient privacy

I. INTRODUCTION

Progress in digital health technology has reshaped how clinical information is produced, retained, and circulated. Imaging modalities such as Chest X-rays, Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Ultrasound are indispensable for diagnosing illness and shaping treatment decisions. The resulting images commonly embed private patient particulars such as names, identification numbers, age, and clinical notes. Because hospitals now depend heavily on electronic health records and cloud-mediated communication, guarding the privacy and security of this medical content has grown into a pressing concern.

Medical images travel constantly between hospitals, diagnostic laboratories, and clinicians through telemedicine services and hospital information systems. During these exchanges, intrusion, leakage, or manipulation can breach patient confidentiality and run afoul of healthcare privacy law. Encryption secures content in transit, yet it openly signals that protected information exists and can be defeated once its keys are exposed. Supplementary safeguards are therefore needed to protect patient particulars without sacrificing the diagnostic quality of the imagery.

Steganography offers a practical remedy by hiding confidential data inside digital media in a way that leaves the carrier visually unchanged. Among the available methods, Least Significant Bit (LSB) embedding enjoys broad use owing to its straightforward design, generous hiding capacity, and slight effect on image fidelity. Medical images, however, frequently bear visible patient text that ought to be removed before embedding. This work therefore applies OCR to locate such text and follows it with Gaussian Blur-based redaction before the confidential payload is hidden.

The proposed framework, MEDICRYPT, integrates OCR, image redaction, and LSB-based steganography into a single pipeline for the secure conveyance of medical images. Patient particulars are extracted, translated into a binary stream, concealed within chest X-ray images, and recovered faithfully during extraction. The fidelity of the resulting stego images is gauged through Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and histogram comparison, confirming that the carrier images remain visually indistinguishable from the originals.

II. LITERATURE SURVEY

Reddy and Fathima introduced a protection scheme for medical imagery in which chaotic Chebyshev maps govern the random selection of pixels during embedding [2]. In contrast to ordinary LSB schemes that fill pixels in order, the chaotic mapping injects randomness into pixel choice, making unauthorised extraction far harder, and reported a stronger PSNR while keeping diagnostic quality intact. Its scope, however, is confined to embedding and overlooks patient text already visible on the image.

P. K. and Jaitly combined image compression, encryption, and steganography so that patient data is encrypted before being concealed within the image, layering several defences against unauthorised access [5]. The added encryption and compression stages, however, raise computational burden and runtime and leave visible textual content untouched. Kataria et al. compared several encryption algorithms alongside steganography and concluded that pairing the two boosts confidentiality beyond either alone, yet provided no facility for masking visible patient text [7].

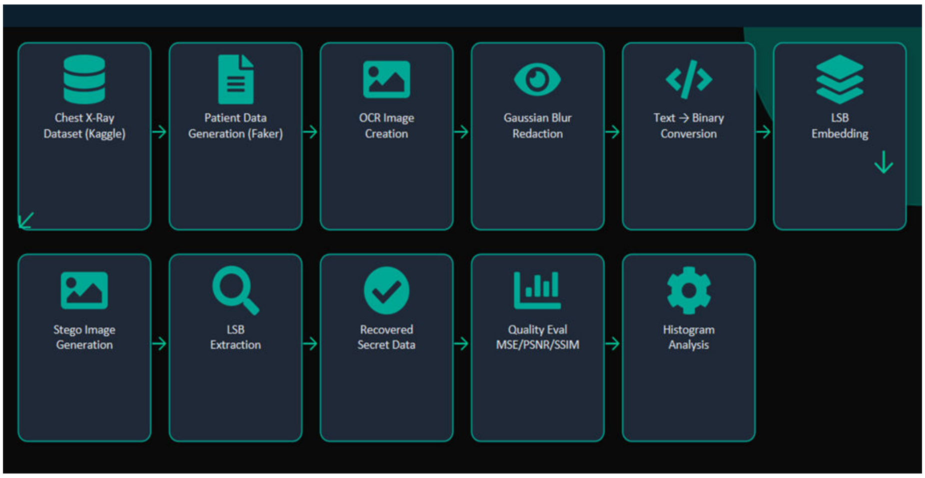
Sharma et al. proposed an edge-based scheme that embeds data within the Non-Interest Region of the cover image to keep diagnostically meaningful areas intact, reporting high PSNR with little distortion, but presuming that patient-identifiable text has already been removed [13]. Gupta et al. extended this idea with a hybrid edge-based design protected by five cryptographic layers, achieving high PSNR and low MSE at the cost of heavier computation, and likewise assuming prior removal of visible text [14].

Sharma and Jain surveyed the spatial-domain family, covering LSB, Pixel Value Differencing, histogram shifting, and adaptive embedding, and identified LSB as the simplest and quickest technique while cautioning that plain LSB is exposed to statistical attack and steganalysis [15]. Alanzy et al. encrypted confidential content before LSB embedding to keep extracted data protected, attaining high PSNR and low MSE [16]. Zhu et al. buried data in DCT coefficients for greater resistance to JPEG compression, at the expense of higher computational complexity [4]. Across these works, the carrier is consistently assumed free of visible patient text, leaving a clear gap that MEDICRYPT addresses by combining automatic detection, redaction, and embedding.

III. METHODOLOGY

The proposed MEDICRYPT framework secures medical image transmission by uniting OCR, Gaussian Blur-based anonymisation, and LSB steganography within one workflow. Unlike methods that focus solely on concealing a payload, the system first shields the patient particulars printed visibly on the image and then hides the extracted content inside that same image, providing dual-layer protection while preserving diagnostic quality. The pipeline comprises seven sequential modules described below.

Figure I Proposed System Architecture



© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406089

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

A. Input Image Acquisition

A chest X-ray image drawn from the dataset is supplied as the cover image. Each uploaded image passes through preprocessing that secures consistency in format, resolution, and quality; damaged images are rejected, and the validated image is forwarded to the OCR module.

B. OCR-Based Sensitive Information Detection

PaddleOCR was adopted for its high recognition accuracy and resilience with printed medical text. The engine sweeps the image, draws a bounding box around each detected segment of patient text, and converts it into editable form. The recognised particulars are held for embedding, while the bounding-box coordinates are passed to the redaction module.

C. Gaussian Blur-Based Image Redaction

Gaussian Blur is applied solely to the bounding-box coordinates supplied by OCR. A Gaussian kernel is convolved across each region, replacing original pixel intensities with weighted averages of neighbouring pixels, so the text becomes illegible while the surrounding image keeps its natural look. Unlike masking or cropping, this preserves structural continuity and diagnostic value.

D. Secret Message Preparation

Because the LSB algorithm operates on binary data, the extracted text is converted into a binary payload. Every character is mapped to its ASCII value, each ASCII value is rendered as an 8-bit binary sequence, and the sequences are concatenated to form a continuous binary stream suited to embedding.

E. LSB Encoding and Decoding

During encoding, the least significant bit of each selected pixel is replaced by one bit of the payload until the full message is embedded, producing a stego image. The modified pixel is obtained as $P' = (P \text{ AND } 254) \text{ OR } b$, where P is the original pixel, b the secret bit, and 254 clears the least significant bit before insertion. Decoding reverses this: pixels are read in the same order, the least significant bit of each is extracted as $b = P' \bmod 2$, and the recovered bits are grouped into bytes and converted back to text.

F. Performance Evaluation

The stego image is appraised through PSNR, MSE, and SSIM, supported by histogram comparison and validation of the recovered message against the original OCR output. PSNR weighs the maximum attainable pixel intensity against embedding distortion, MSE captures the mean squared pixel difference, and SSIM compares brightness, contrast, and structural likeness on a scale from 0 to 1.

IV. RESULTS AND DISCUSSION

The framework was implemented in Python using PaddleOCR, OpenCV, NumPy, Pillow, and scikit-image, on an Intel Core i5 system with 8 GB RAM running Windows 11. PaddleOCR reliably localised patient text on the chest X-ray images, Gaussian Blur masked the detected regions without disturbing anatomical structures, and the binary payload was embedded so that the stego image appeared visually identical to the anonymised cover. As an example, the message embedded in one sample was recovered exactly as "Patient: Ann James | Age: 31 | ID: CDS68238", confirming lossless retrieval.

Image quality was quantified across sample stego images, as summarised in Table I. The framework consistently produced PSNR values near 37 to 38 dB, MSE values close to 11, and SSIM values above 0.95, indicating that embedding introduced negligible perceptible distortion.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Figure II Original Image & Stego Image

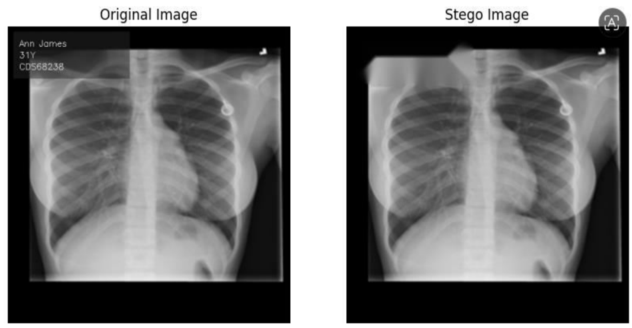
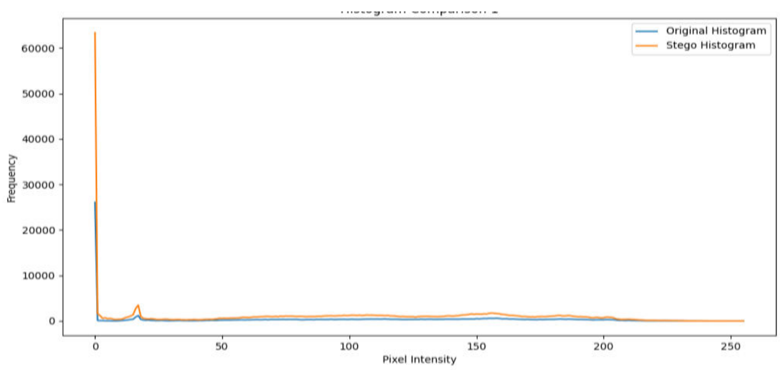


Table I. Performance Metrics for Representative Stego Images

Image	MSE	PSNR (dB)	SSIM
stego_Normal-1.png	11.5349	37.5107	0.9782
stego_Normal-10.png	10.6055	37.8755	0.9652
stego_Normal-100.png	11.6882	37.4533	0.9513
stego_Normal-1000.png	11.0435	37.6997	0.9561
stego_Normal-10000.png	11.6744	37.4585	0.9852
stego_Normal-10001.png	10.5838	37.8844	0.9733

Histogram analysis showed that the statistical spread of pixel intensities remained nearly unchanged after embedding, leaving the hidden content difficult to detect through ordinary inspection. Compared with manual masking, which anonymises but does not embed, and with traditional LSB, which embeds but ignores visible text, MEDICRYPT performs both functions; relative to deep learning approaches, it attains comparable image quality and security without the need for large datasets or heavy computation. The combination of automated OCR, anonymisation, and lightweight embedding therefore offers a practical solution for healthcare deployment.

Figure III Histogram Comparison



V. CONCLUSION

This paper presented MEDICRYPT, a privacy-preserving framework that integrates OCR, Gaussian Blur-based anonymisation, and LSB steganography to protect patient information carried by chest X-ray images. The system automates the detection of patient-identifiable text, dispenses with manual anonymisation, and securely hides the extracted particulars within the same image. Evaluation through PSNR, MSE, SSIM, and histogram analysis confirmed that embedding introduces only slight distortion while the hidden data is recovered faithfully, preserving diagnostic quality. Future work may replace plain LSB with deep learning-based encoder-decoder networks, add AES or RSA encryption ahead of embedding, extend support to CT, MRI, and ultrasound and DICOM imagery, and enable cloud-based secure sharing within PACS, HIS, and telemedicine platforms.

REFERENCES

[1] K. D. Mandl, W. W. Simons, W. C. Crawford, and J. M. Abbett, "Indivo: A personally controlled health record for health information exchange and communication," BMC Med. Informat. Decis. Making, vol. 7, no. 1, p. 25, Dec. 2007.
[2] V. P. K. Reddy and A. A. Fathima, "A cost effective approach for securing medical X-ray images using Chebyshev map," in 2016 Int. Conf. on Recent Trends in Information Technology (ICRTIT), Chennai, India, 2016, pp. 1-5, doi: 10.1109/ICRTIT.2016.7569576.
[3] D. Watni and S. Chawla, "A Comparative Evaluation of Jpeg Steganography," in 2019 5th Int. Conf. on Signal Processing, Computing and Control (ISPCC), Solan, India, 2019, pp. 36-40, doi: 10.1109/ISPCC48220.2019.8988383.
[4] Z. Zhu, N. Zheng, T. Qiao and M. Xu, "Robust Steganography by Modifying Sign of DCT Coefficients," IEEE Access, vol. 7, pp. 168613-168628, 2019, doi: 10.1109/ACCESS.2019.2953504.
[5] P. K. and V. Jaitly, "Securing Medical Images using Compression techniques with encryption and Image Steganography," in 2023 3rd Int. Conf. on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-7, doi: 10.1109/CONIT59222.2023.10205855.
[6] A. G. Devi, A. Thota, G. Nithya, S. Majji, A. Gopatoti and L. Dhavamani, "Advancement of Digital Image Steganography using Deep Convolutional Neural Networks," in 2022 Int. Interdisciplinary Humanitarian Conf. for Sustainability (IIHC), Bengaluru, India, 2022, pp. 250-254, doi: 10.1109/IIHC55949.2022.10060230.
[7] M. Kataria, K. Jain and N. Subramanian, "Exploring Advanced Encryption and Steganography Techniques for Image Security," in 2023 11th Int. Symp. on Digital Forensics and Security (ISDFS), Chattanooga, TN, USA, 2023, pp. 1-6, doi: 10.1109/ISDFS58141.2023.10131890.
[8] R. Kumar, S. Bansal and R. K. Bansal, "A Resilient Approach to Robust JPEG Steganography: Ensuring Hidden Data Integrity," in 2023 Int. Conf. on Integrated Intelligence and Communication Systems (ICIICS), Kalaburagi, India, 2023, pp. 1-6, doi: 10.1109/ICIICS59993.2023.10421695.
[9] J. Liu, X. Song, G. Li and K. Han, "Robust JPEG Image Steganography Using Wavelet Domain SVD and Adaptive QIM," in 2023 8th Int. Conf. on Signal and Image Processing (ICSIP), Wuxi, China, 2023, pp. 434-438, doi: 10.1109/ICSIP57908.2023.10270839.
[10] W. Dong, Z. Jiang, L. Chen and W. Sun, "A New High-Capacity Robust Image Steganography Method via Adversarial Network," in 2023 5th Int. Academic Exchange Conf. on Science and Technology Innovation (IAECST), Guangzhou, China, 2023, pp. 663-672, doi: 10.1109/IAECST60924.2023.10503312.
[11] Y. Sun, J. Liu and R. Zhang, "A Robust Generative Image Steganography Method based on Guidance Features in Image Synthesis," in 2023 IEEE Int. Conf. on Multimedia and Expo (ICME), Brisbane, Australia, 2023, pp. 55-60, doi: 10.1109/ICME55011.2023.00018.
[12] D. Sharma, R. Sharma, R. Mehra, and S. Jain, "Securing X-Ray Images in No Interest Region (NIR) of the Normalized Cover Image by Edge Steganography," IEEE Access, vol. 12, pp. 168672-168689, 2024, doi: 10.1109/ACCESS.2024.3467167.
[13] A. Gupta, P. Verma, and S. Yadav, "Securing X-Ray Images into Cover Images Using Hybrid EBS Steganography with Five-Layer Cryptography," IEEE Access, vol. PP, no. 99, pp. 1-1, Nov. 2024, doi: 10.1109/ACCESS.2024.3489452.
[14] S. Sharma and P. Jain, "Digital Image Steganography In Spatial Domain: A Comprehensive Review," Journal Of Theoretical And Applied Information Technology, vol. 97, no. 19, pp. 5201-5210, Oct. 2019.
[15] M. Alanzy, R. Alomrani, B. Alqarni, and S. Almutairi, "Image Steganography Using LSB and Hybrid Encryption Algorithms," Applied Sciences, vol. 13, no. 21, p. 11771, 2023, doi: 10.3390/app132111771.
[16] M. Patel, P. V. Shah, A. K. Oza, and D. Dalvadi, "An enhanced security and prediction of image steganography using ANFIS classification," in Hybrid and Advanced Technologies, 1st ed., CRC Press, 2025, pp. 1-10, doi: 10.1201/9781003559139.
[17] M. Taleby Ahvanooey, Q. Li, J. Hou, A. R. Rajput, and Y. Chen, "Modern Text Hiding, Text Steganalysis, and Applications: A Comparative Analysis," Entropy, vol. 21, no. 4, p. 355, 2019, doi: 10.3390/e21040355.