Data Hiding in Medical Image using Deep Learning and LSB Steganography

Hadjer Saidi
LESIA LAboratory
Computer Science Department
University of Biskra, Algeria
email: hadjer.saidi@univ-biskra.dz

Okba Tibermacine
LRP Laboratory
Computer Science Department
University of Biskra, Algeria
email: o.tibermacine@univ-biskra.dz

Ahmed Elhadah
Department of Computer Science
South Valley University,
Qena, Egypt
email : ahmed.elhadad@sci.svu.edu.eg

***Abstract*—In the context of digital healthcare, preserving patient-sensitive information is a paramount. in this paper, we propose to leverage the Deep learning technique and LSB steganography to secure medical images in the DICOM format. This is achieved by embedding the sensitive information represented as QR codes within the insignificant area of medical images using LSB. This insignificant area is identified by the Deep Learning object detection algorithm. The whole process, its implementation, and its evaluation are presented below.**

I. Introduction

In today's interconnected world, the need for secure communication and data protection is more critical than ever. While encryption methods are commonly used to secure the contents of messages, there is another realm of covert data protection known as image steganography. This fascinating field of study delves into the art of concealing sensitive information within digital images, enabling clandestine communication without arousing suspicion. Image steganography is a clever technique that allows data to be hidden within the very pixels of an image, making it virtually imperceptible to the human eye. Unlike encryption, which masks the content of a message, steganography goes a step further by embedding the message itself within the visual data.

In this study, we will present a proposal that utilizes steganography techniques to safeguard patient data within DICOM medical files.

II. Related work

Numerous research studies have been conducted with the aim of safeguarding patient information by concealing it within medical images; among them, we find: the work of Bouzhidar and all in [19], that introduced an innovative steganography technique called BOOST, which was developed to hide user data within medical images. Their approach consisted of two main steps: Firstly, confidential patient data was encrypted using a novel technique known as the "nuclear spin generator-based pseudorandom generator," resulting in encrypted data. This encrypted data was then converted into a binary sequence using an ASCII table. In the subsequent stage, this binary sequence was incorporated into the least significant bit of the non-black pixels in the image. It's worth noting that their method delivered impressive results, with PSNR (Peak Signal-to-Noise Ratio) values exceeding 113 dB, all while accommodating a payload capacity of 0.74 bits per pixel. in [**?**] Romany et al. presented a comprehensive steganography technique that combines multiple methods for securely hiding data within medical images. Their approach involved using RSA encryption to protect sensitive information, employing the Ripplet Transform for image manipulation, and utilizing LSB substitution to embed confidential data. To improve the imperceptibility of the Stego image, they introduced an adaptive genetic algorithm-based process called the Optimum Pixel Adjustment Process (OPAP). This integrated approach exhibited resilience against RS attacks and demonstrated that the Discrete Ripplet Transform (DRT) outperformed the Integer Wavelet Transform (IWT). Importantly, the achieved PSNR values ranged from 49 to 56 dB.

III. Background

The fundamental aim of image steganography is to securely transmit an image (cover) that maintains its original appearance while covertly carrying concealed information (secret message) from sender Alice to recipient Bob. Despite being targeted by Eve's attacks in the communication channel. While maintaining secrecy amidst the steganalysis process, which endeavours to reveal concealed information within digital files. Fig. 1 displays the fundamental framework of image steganography.

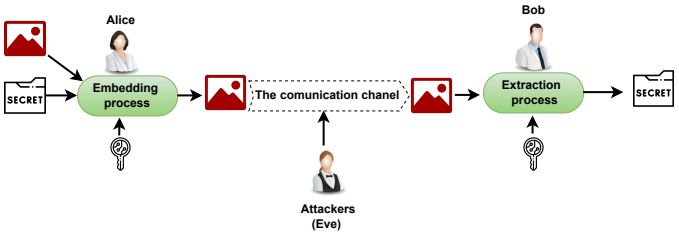

Fig. 1. The basic framework of image steganography

A. Deep Learning and object detection Algorithms

Deep learning, a subfield of machine learning, has recently experienced remarkable success and has brought about a significant revolution in different fields. He derived his strength

from his architecture that emulates the human brain, advancements in statistics and applied mathematics, enhancements in computer hardware and software infrastructure (CPUs, GPUs), and the expansion of accessible training data. Numerous deep learning architectures have been designed for supervised, Semi-supervised, or unsupervised learning tasks [18]. Deep learning-based image steganography: encompasses all techniques that employ deep learning within steganographic systems, whether the learning is done to identify optimal hiding locations or generate appropriate cover. In our study, we employed a deep learning architecture for object detection, with the primary objective being the identification of the vital region within the medical image that holds significance for the doctor's diagnostic process; This architecture is the mask r-cnn object detection and segmentation. This architecture is specifically tailored for instance segmentation, and we enlisted its capabilities for the purpose of semantic segmentation.

B. LSB

Numerous concealment techniques are suggested in the image steganography field; all these techniques aim to make an equilibrium between the capacity to hide information, the imperceptibility of the hidden data, and the robustness against detection attempts. The concealment techniques can be classified into two main categories: spatial domain technique, and transform domain technique.

Spatial domain techniques: There exist several widely renowned and extensively advanced techniques within this field, among them Least Significant Bit (LSB) [1], [2], [3], Pixel Value Differencing (PVD) [4], [5], [6], and Exploiting Modification Directions (EMD) [7], [8], [9]. Overall, these techniques are characterised by their ease of implementation, ample embedding capacity, high visual quality, and efficient computation. However, they are prone to detection through statistical analysis or steganalysis methods and are vulnerable to various image-processing operations and attacks.

Frequency domain techniques: include concealment algorithms that subject the cover to any mathematical transformation. That is, the concealment occurs within the transformed coefficient produced by one of the transformation techniques such as Discrete Cosine Transform (DCT) [10] [11], Discrete Wavelet Transform (DWT) [12] [13] [14], and Contourlet Transform. [15] [16] [17]. These techniques attain both imperceptibility and robustness, displaying exceptional resistance against geometric attacks while being challenging to identify through statistical detection and analysis techniques. Nevertheless, there are constraints on the amount of payload that can be accommodated, and these techniques demand substantial computational resources and time.

IV. Proposed Method

The overall structure is illustrated in Figure 2, consisting of the following three stages:

- Processing patient information: this initial phase involves extracting patient data from the DICOM file and subsequently transforming it into a QR code using a QR code generator.
- Cover processing: During this phase, a deep learning method is employed to identify the crucial elements within the medical image, with the objective of distinguishing between the vital region and the less important portions. This phase also encompasses activities related to labelling the dataset and training the model.
- The embedding and the extraction process: to conceal the QR code of the patient's information in the insignificant area of the cover, we applied the Least Significant Bit (LSB) method. The LSB technique involves replacing data in the least significant bits of the image while making minimal or imperceptible changes to the overall image. To extract information from a Stego image the LSBs of the cover image where the secret message was hidden are collected together to produce the secret.

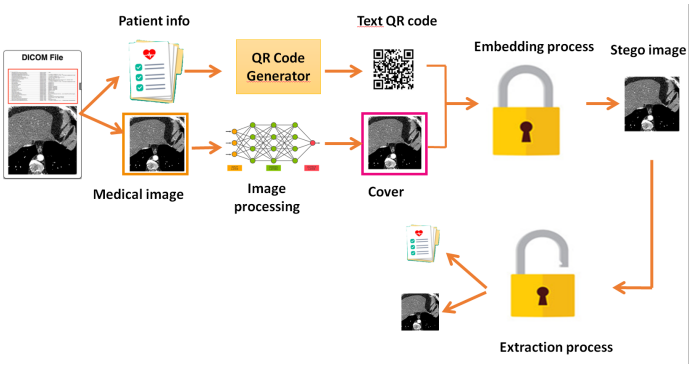

Fig. 2. The proposed framework

The main contributions of our work are: Create a reversible, robust steganography scheme to ensure the security of confidential patient data. Detection of insignificant areas in medical DICOM images, which will be exploited to conceal sensitive information with the least damage to the cover. Because the private information has been made into a QR code, the embedded QR code is extracted with minimal loss so that it is readable and patient information can be extracted.

V. Evaluation

For the training of Mask R-CNN, we utilized the Chaos dataset, which is specifically designed for segmenting abdominal organs. after 28 epochs the loss obtained is 0.0450. This implies that the recognition of critical elements within the medical image was highly acceptable.

Once the embedding process is complete, assessing the model's efficacy becomes essential. Various metrics are employed for this purpose; The table I showcases a selection of such metrics.

VI. Conclusion

In this paper, we presented our ongoing work on data hiding of sensitive information in medical images using the mask R-CNN object detection and the least significant bit embedding

TABLE I
MOST COMMON IMAGE STEGANOGRAPHY MATRICES

Metric	Description	When is it better ?
Peak Signal-to-Noise Ratio (PSNR)	It calculates the discrepancy between two images by comparing the maximum signal power to the power of noise that disrupts the signal. It is mesured in decibels (db)	A higher value (Exceeds 30 dB)
Structural Similarity Index (SSIM)	It compare luminance, contrast and the structural of the cover and the stego .	Close to 1
Mean Square Error (MSE)	It measures the average squared difference between the pixel values of the cover image and the stego image.	A lower value
Universal Image Quality Index (UIQI)	It computes the similarity between the cover and the stego based on similarities in luminance, contrast, and structure.	close to 1
Embedding Capacity (EC)	It is the number of secret bits that are embedded per pixels, it is calculated in Bpp (Bit Per Pixel)	A higher value
Payload capacity	It is the quantity of information that can be concealed within the cover media. It is represented as a specific number of bits or the percentage designated for concealing data relative to the total size of the cover.	A higher value
Bit Error Rate (BER)	Determining the proportion of wrongly recovered bits relative to the concealed data's original value.	close to 0

method. In our upcoming research, we plan to explore an alternative embedding technique within the frequency domain, modify the deep learning architecture employed, and subsequently assess and compare the outcomes.

REFERENCES

[1] S. Chakraborty, A. S. Jalal, and C. Bhatnagar, "Multimedia Tools and Applications", Springer, vol. 76, no. 6, pp. 7973–7987, 2017.

[2] M. C. KASAPBAS¸ and W. ELMASRY, "New LSB-based colour image steganography method to enhance the efficiency in payload capacity, security and integrity check", Springer, 2018.

[3] D. Lou, and C. Hu, "LSB steganographic method based on reversible histogram transformation function for resisting statistical steganalysis" , in *Information Sciences*, Elsevier, 188 346–358, 2012.

[4] D. Wu, and W. Tsai, "A steganographic method for images by pixel-value differencing", Elsevier, 2002.

[5] C. Wang, N. Wu, C. Tsai, and M. Hwang, "A high quality steganographic method with pixel-value differencing and modulus function", in *The Journal of Systems and Software* 81 150–158, 2008.

[6] K. Changa, C. Changa, P. S. Huangb, and T. Tua, "A Novel Image Steganographic Method Using Tri-way Pixel-Value Differencing", in *JOURNAL OF MULTIMEDIA*, VOL. 3, NO. 2,2008.

[7] S. Shen, and L. Huang, "A data hiding scheme using pixel value differencing and improving exploiting modification directions", in *Computers Security*, Volume 48,Pages 131-141, Elsevier, 2015.

[8] X. Niu, M. Ma, R. Tang and Z. Yin, "Image Steganography via Fully Exploiting Modification Direction", in *International Journal of Security and Its Applications*, Vol. 9, No. 5, pp. 243-254, 2015.

[9] K. Jung, and K. Yoo, "Improved Exploiting Modification Direction Method by Modulus Operation", in *International Journal of Signal Processing, Image Processing and Pattern*, Vol. 2, No.1, 2009.

[10] R. Biswas, and S. K. Bandyapadhay,"Random selection based GA optimization in 2D-DCT domain color image steganography", in *Multimedia Tools and Applications*, volume 79, pages 7101–7120, Springer, 2020.

[11] X. Zhang, F. Peng, and M. Long,"Robust Coverless Image Steganography Based on DCT and LDA Topic Classification", in *IEEE TRANSACTIONS ON MULTIMEDIA*, VOL. 20, NO. 12, 2018.

[12] V. Kumar, and D. Kumar, "A modified DWT-based image steganography technique", in *Multimed Tools Appl*, 77, pages 13279–13308, Springer, 2018.

[13] R. Thanki,and S. Borra, "A color image steganography in hybrid FRT–DWT domain", in *Journal of Information Security and Applications*, Volume 40, Pages 92-102, Elsevier ,2018.

[14] J. Khandelwal, V. K. Sharma,D. Singh,and A. Zaguia," DWT-SVD Based Image Steganography Using Threshold Value Encryption Method", in *Computers, Materials, Continua*, Vol. 72, N°. 2, pages 3299-3312, 2022.

[15] M. S. Subhedar, and V. H. Mankar, " Image steganography using contourlet transform and matrix decomposition techniques", in *Multimedia Tools and Applications* vol. 78, pages 22155–22181, Springer, 2019.

[16] H. Sajedi, and M. Jamzad," Using contourlet transform and cover selection for secure steganography" ,in *International Journal of Information Security* volume 9, pages 337–352, Springer, 2010.

[17] V. K. Reshma, R. S. Vinod Kumar, D. Shahi, M. B. Shyjith, "Optimized support vector neural network and contourlet transform for image steganography", in *Evolutionary Intelligence*, vol. 15, pages 1295–1311, Springer, 2022.

[18] M. P. Hosseini, S. Lu, K. Kamaraj, A. Slowikowski, and H. C. Venkatesh, "Deep Learning Architectures" , in *Studies in Computational Intelligence* , Springer,2019, vol.866 , pp. 1-24. Accessed: July. 07, 2023.

[19] B. Stoyanov, and B. Stoyanov, "Boost: medical image steganography using nuclear spin generator", in *Entropy*, vol. 22, 501,2020.

[20] M. F. Romany and E. M. ABDELRAHIM, ," An evolutionary computing enriched RS attack resilient medical image steganography model for telemedicine applications", *Multidimensional Systems and Signal Processing*, vol. 30, p. 791-814, 2019.