A Comparative Study on Medical Image Watermarking using Hybrid Approach and RivaGAN

Yew Lee Wong[1*], Jia Cheng Loh [1], Chen Zhen Li [1], Chi Wee Tan [1]

[1] Faculty of Computing and Information Technology, Tunku Abdul Rahman University College, Kampus Utama, Jalan Genting Kelang, 53300 Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia
*Corresponding author: wongyewlee-wm19@student.tarc.edu.my

ABSTRACT

With the increased use of electronic medical records and computer networks, Medical Image Watermarking (MIW) now plays a very important role to preserve integrity and completeness of medical images. As of now, there are no perfect algorithms or solutions for invisible watermarking as there are trade-offs between visibility and robustness. In this study, we explored multiple implementations of image watermarking techniques using Hybrid-Approach and Deep-Learning-Approach. The experiments to measure the limitations and robustness were done on a dataset of breast ultrasound images. 18 attacking methods were performed on the encoded images and performance were evaluated using PSNR and NCC. Encoded images were then being transmitted digitally using multiple transmission method to test its robustness against transmission platform. In conclusion, the Deep-Learning Approach of RivaGAN has shown best robustness despite many extreme attacks while the Hybrid Approach of DWT-DCT-SVD shown the best performance in terms of imperceptibility. We reject RivaGAN as the best solution for Medical Image Watermarking despite its robustness as it was created specifically for video invisible watermarking.

Keywords: Invisible Watermarking, DCT, DWT, SVD, RivaGAN

1.0 INTRODUCTION

With the increased use of computer networks and electronic medical records, medical images now play a very important role unprecedentedly. It was mentioned by Kuang et al. that the electronic medical record system is weak at protecting the content of medical records (Kuang et al., 2009). Therefore, it raises a need of digital signatures such as watermarking on medical images, as the watermarking applied should not compromising the quality of image while the confidentiality of owner is protected. To preserve the integrity and completeness of the medical images, conventional visible watermarking methods could not be applied. Without proper authentication mechanism in place, it is very challenging to prove the ownership and authenticity of the medical images.

Invisible watermarking can be traced back to as early as 1997, with the work by Yeung et al which proposed a method for image verification (Yeung & Mintzer, 1997). Image Watermarking is a technique where data is embedded into the digital medical image. There are two types of watermarking, which is visible watermarking and invisible watermarking. Watermarking can be done on the spatial domain and transform domain.

In the scenario of medical images, vital and confidential information are usually embedded into the medical images which are to be communicated over any digital transmission channel

(Khare et al., 2020). Digital medical images transmitted over any channel may raise data integrity problems, therefore, invisible watermarking could be the solution. However, there are not perfect algorithms or solutions for invisible watermarking as trade-offs can happen between visibility and robustness when doing watermarking (Mousavi et al., 2014).

For a watermarking technique to reach the optimum state, it shall take into account of robustness, imperceptibility and security. Robustness can be simply understood as the resilience of the watermarking towards any attacks while imperceptibility focuses on the quality of watermarked image.

In our study, we would like to explore multiple implementations of the invisible watermarking techniques published by other researchers as well as comparing frequency domain watermarking and deep learning watermarking. The contribution of our study is as follows:

- To utilize currently available and emerging technique to embed a message within image which it shall minimally impacts the viewing experience
- To verify the robustness of algorithms as claimed by another research
- To measure the watermarking effects on the medical images using PSNR and NCC
- To investigate the limitations and resistance of the algorithm towards any attacks
- To ensure and investigate the completeness of embedded message after redistribution and transmission

2.0 LITERATURE REVIEW

2.1 Digital Image Watermarking & Invisible Watermarking

Distribution of digital image has been widely utilized with the increasing development of internet; therefore, the protection of the data is important (Abdulrahman, 2019). Digital watermarking can be defined as the process of embedding or hiding data into another digital data, then extracting the hidden information (Tao et al., 2014). It has been argued that it has become easier to tamper with the medical images as advanced picture editing software is now more accessible (Coatrieux, 2006). To address such concerns, invisible watermarking can be utilized for data concealment and to protect data integrity (Coatrieux, 2006). A digital watermarking domain can be mainly classified into two sub domains of spatial domain and frequency domain (EL-Shazly, 2004). Performance metrics that are generally used to measure the image watermarking technique is Robustness and Imperceptibility, however both properties are contradicting (Usman et al., 2008). Peak Signal-to-Noise Ratios (PSNR), which are used to measure the imperceptibility of watermark which should not distort the image quality in the presence of watermark (Al-Haj, 2007). PSNR is usually denoted in decibels (dB) and is widely used in comparing Medical Image Watermarking (MIW) algorithms (Faragallah et al., 2021). Robustness of the watermarking technique is measured through the immunity and resistance of the watermark against any attempts of removal and degrading (Voloshynovskiy et al., 2001).

2.2 Hybrid Watermarking

Firstly, published in 2008, the DWT-DCT-SVD based watermarking algorithm was found to be very robust where the encoded show no visible distortion (Navas et al, 2008). SVD was originally developed by geometers however were start being used for watermarking since 2001 (Sverdlov et al., 2001). DWT-DCT-SVD has the advantage of need not to embed all singular values and can be utilized to develop algorithms for loss image compression (Navas et al., 2008). For the hybrid watermarking of DWT-DCT, it was found that the performance of the combined technique shown improved performance as compared to sole DWT algorithm (Al-Haj, 2007). The improvement of robustness by the DWT-DCT was considerably high in the

comparison between DWT. Robustness of the DWT-DCT watermarking can be also seen to be more robust to the linear and non-linear attacks (Abdulrahman & Ozturk, 2019). On the non-hybrid watermarking technique of DWT, it was also proven that DWT is robust against any common image processing operations (Lala, 2017).

2.3 RivaGAN

RivaGAN is a novel architecture for robust video marking, which it consists of a custom attention-based mechanism for embedding arbitrary data. Two independent adversarial networks were used to critique the video quality and optimize the encoder for robustness. This architecture embeds 32-bit watermark into a sequence of frames. It was also found that RivaGAN is robust against any common video processing operations such as cropping, scaling and compression. With a detection rate of approximately 52%, the watermarked footage is nearly indistinguishable to human eyes, RivaGAN managed to reach PSNR of 42.05 (Zhang et al., 2019).

3.0 RESEARCH METHODOLOGY

3.1 Dataset and Algorithms

The dataset used is a collection of breast ultrasound images among women of the ages 25 and 75 years old which is available at https://www.kaggle.com/aryashah2k/breast-ultrasound-images-dataset. A total of 20 images is being selected randomly from these 780 images of average image size 500×500 pixels. The chosen images were named alphabetically from "MRI_A" to "MRI_T". Four algorithms of invisible watermarking were chosen, namely DWT, DWT-DCT, DWT-DCT-SVD & RivaGAN.

3.2 General Framework

As illustrated in Figure 1, it shows the overall flow of our study. Firstly, we will encode a watermark in string format into the original MRI images. Then we will attack those encoded images using 18 different methods. Transmission of encoded images were also done on the attacking phase. After that we will try to decode the watermark from the attacked images and calculate the PSNR and NCC value. Lastly, the result will be visualize using some chart.

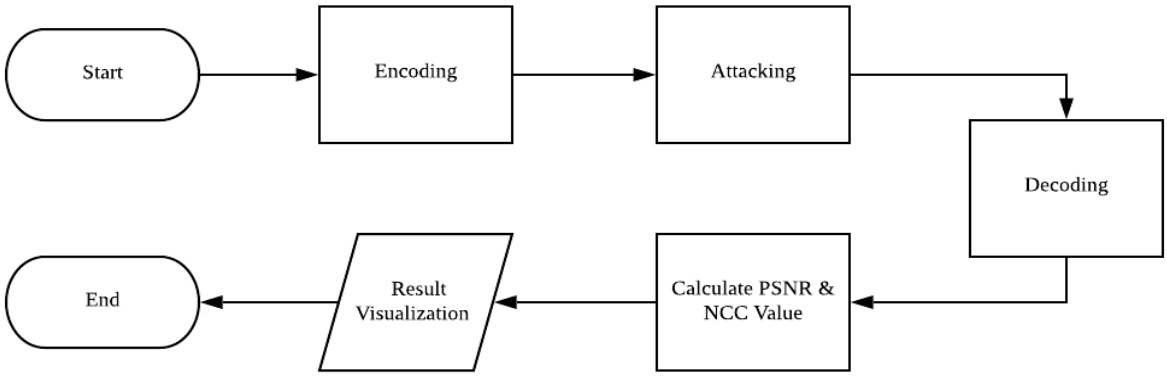

Figure 1. Medical Image Watermarking Framework.

3.3 Testing Criteria

On the pass rate, messages retrieved after being attacked is strictly being compared absolutely. Only if the output matches 100% with the initial input can be considered as passing the test. Decoding errors were counted as failure through exceptions caught by decoder. Partial success that the output matches the input was considered as failure. On the test of transmission, watermarked photos were transmitted through WhatsApp Image, WhatsApp Document, Google Drive, Facebook Messenger and Gmail. The images received on the receiving end were put into decoder to retrieve the embedded messages. Output that matches the initial input 100% will only be considered as pass the test. On testing the implementation of the selected library, average of decoding and encoding time were done using time library in Python. Elapsed time was recorded down over 1000 iterations of the operation and mean were calculated. On measuring the relationship between characters length of embedded message and file size, different randomly generated string of different length was encoded. File size was compared on before and after encoding. On measuring the performance of each watermarking algorithm, we evaluate the image of before using the value of Peak Signal-to-Noise Ratio (PSNR) and Normalized Cross Correlation (NCC).

3.4 Experiment Environment

The testing of the implementation was done on a desktop system of such specifications in Table 1.

Table 1. Testing system.

CPU	Intel Xeon E5-2650v2 @ 2.60Ghz, 8 Cores 16 Threads
RAM	16GB DDR3 1666Mhz
Operating System	Windows 10 Pro 64-bit (10.0, Build 19043
Python Version	3.8.10

4.0 RESULTS AND DISCUSSIONS AND DISCUSSIONS

As shown at Table 2, there are 18 types of attacking methods that will be used to test the robustness of each watermarking algorithm.

Table 2. Attacking methods.

	Kernel Settings/ Ratio / Strength
Averaging	size = 5x5
Bilateral Filtering	D = 9, sigmaColor = 75, sigmaSpace = 75
Brightness Decrease	40 %
Brightness Increase	40 %
Crop Horizontal	50 %
Crop Vertical	50 %
Gaussian Blurring	Size = 5x5
Gaussian Noise	mean=0, variance=0.01

JPG	Convert to JPG
Masks	n = 5, ratio = 0.3
Median Blurring	Size = 7
Poisson Noise	Lambda = 20
Rotate	10 degrees
Salt & Pepper	10 %
Scale Down	25 %
Scale Up	25 %
Sharpen Filtering	[-1, -1, -1], [-1, 9, -1], [-1, -1, -1]
Speckle Noise	mean=0, variance=0.01

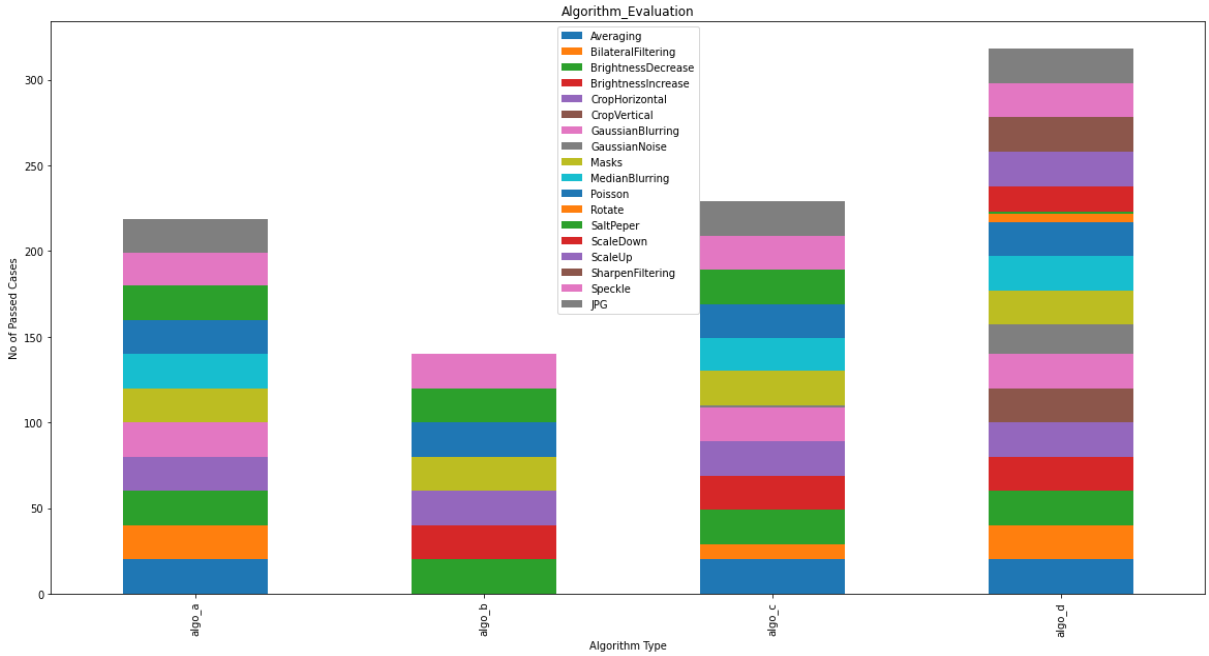

Figure 2. Algorithm Evaluation Based on Pass Rate.

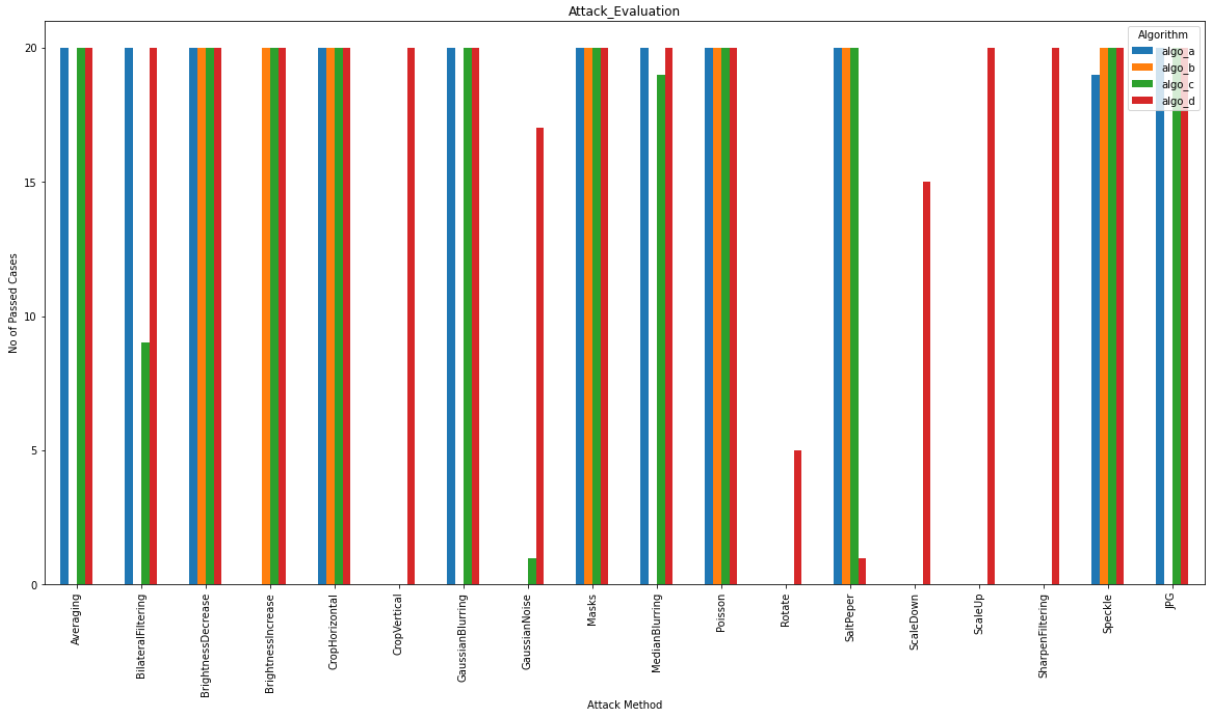

Figure 3. Attack Evaluation Based on Pass Rate.

As illustrated in Figure 2 & 3, RivaGAN has the highest passing rate among all the algorithms follow by DWT-DCT-SVD ranked at the second place. However, DWT-DCT has the worst performance with lowest passing cases.

Table 3. PSNR between Original Image and Encoded Image.

	DWT	**DWT-DCT**	**DWT-DCT-SVD**	**RivaGAN**
MRI_A	35.24	43.74	46.98	40.41
MRI_B	35.26	43.67	46.94	40.39
MRI_C	35.24	43.74	47.00	40.41
MRI_D	35.25	43.63	46.92	40.42
MRI_E	35.25	43.61	46.92	40.42
MRI_F	35.48	42.82	44.77	40.49
MRI_G	35.18	43.59	46.90	40.41
MRI_H	35.27	43.67	46.94	40.43
MRI_I	35.19	43.66	46.98	40.43
MRI_J	35.19	43.60	46.92	40.45
MRI_K	35.19	43.57	46.89	40.44
MRI_L	35.21	43.67	46.99	40.44
MRI_M	35.22	43.62	46.94	40.44
MRI_N	35.23	43.62	46.94	40.44
MRI_O	35.16	43.62	46.95	40.44
MRI_P	35.21	43.61	46.92	40.49
MRI_Q	35.36	43.74	46.96	40.45
MRI_R	35.28	43.65	46.95	40.42
MRI_S	35.23	43.60	46.93	40.41
MRI_T	35.20	43.59	46.90	40.43

The higher PSNR the better the quality of the compressed, or reconstructed image. Based on Table 3, DWT-DCT-SVD algorithm has the highest PSNR value with an average 46.83 dB that determine its criteria as best algorithm among all the algorithms.

Table 4. NCC between Original Image and Encoded Image.

	DWT	DWT-DCT	DWT-DCT-SVD	RivaGAN
MRI_A	0.9973	0.9992	0.9998	0.9992
MRI_B	0.9969	0.9990	0.9998	0.9990
MRI_C	0.9974	0.9992	0.9998	0.9992
MRI_D	0.9973	0.9992	0.9998	0.9992
MRI_E	0.9963	0.9989	0.9997	0.9986
MRI_F	0.9962	0.9983	0.9992	0.9988
MRI_G	0.9965	0.9989	0.9997	0.9989
MRI_H	0.9976	0.9993	0.9998	0.9993
MRI_I	0.9967	0.9990	0.9998	0.9993
MRI_J	0.9975	0.9993	0.9998	0.9992
MRI_K	0.9978	0.9993	0.9998	0.9993
MRI_L	0.9978	0.9993	0.9998	0.9993
MRI_M	0.9978	0.9993	0.9998	0.9993
MRI_N	0.9996	0.9990	0.9998	0.9990
MRI_O	0.9956	0.9987	0.9969	0.9987
MRI_P	0.9973	0.9992	0.9998	0.9992
MRI_Q	0.9968	0.9990	0.9998	0.9990
MRI_R	0.9969	0.9991	0.9998	0.9990
MRI_S	0.9962	0.9988	0.9997	0.9988
MRI_T	0.9972	0.9991	0.9998	0.9991

The higher NCC value the better the degree of similarity between two compared images. Based on Table 4, all the algorithms have similar performance on NCC value. However, DWT-DCT-SVD is the best performance with highest NCC value among all the algorithms.

Table 5. Encoded Algorithm vs Transmission Platform.

	DWT	DWT-DCT	DWT-DCT-SVD	RivaGAN
WhatsApp Image	✓	×	✓	✓
WhatsApp Document	✓	✓	✓	✓
Google Drive	✓	✓	✓	✓
Facebook Messenger	✓	✓	✓	✓

Gmail	✓	✓	✓	✓

As illustrated in Table 5, all algorithms were managed to achieve full passes for every transmission method. However, DWT-DCT algorithm failed to achieve full passes as it failed to WhatsApp Image. It can be believed that this failure was caused by the compression of WhatsApp.

Table 6. Algorithm Implementation Benchmarking.

	DWT	**DWT-DCT**	**DWT-DCT-SVD**	**RivaGAN**
Character Length Limit	×	×	×	✓ (4)
Case Sensitive	✓	✓	✓	✓
Special Characters	✓	✓	✓	✓
Chinese Characters	×	×	×	×

Based on the Table 6, RivaGAN has the restrictions of 4 characters length, while other algorithms have no character length limit. Besides, every algorithm implementation exhibits perfect behaviours toward case sensitive and special characters. On the Chinese characters, all the algorithms fail to encode and decode.

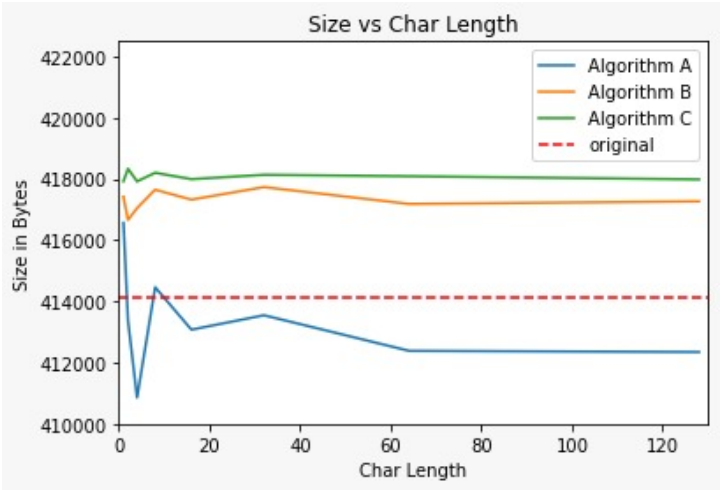

Figure 4. File Size vs Embedded Characters Length.

As shown at Figure 4, the file size exhibited a big fluctuation for the first few 20 bytes. It can be observed that DWT went over the original file size when encoded with messages.

5.0 LIMITATIONS

On the limitation of our study, it was known that RivaGAN were being more focused on video invisible watermarking on its initial release. The implementation was then ported to the image watermarking however there are limitations of the number of characters allowed in the encoding process.

Other than that, the implementations of the algorithm were using the currently available open-source library from GitHub. As such, the implementations might have disparity with the original algorithm or research. Therefore, the future studies can explore how the algorithms can be implemented according to the formula to ensure the consistency and accuracy of the outcome.

6.0 CONCLUSIONS

Through our study, the RivaGAN does exhibit the state-of-the-art robustness as mentioned in the paper of RivaGAN authors (Zhang et al. 2019). Our tests do also confirm the claim of feasibility and robustness for deep learning networks in blind image watermarking (Vukotic et al., 2018). Despite many extreme attacks being conducted on RivaGAN's encoded images, it was still able to pass all the retrieval messages test as it exhibited strong robustness over any other algorithms.

However, we reject RivaGAN as the best algorithm for Medical Image Watermarking as it fails to surpass the PSNR and NCC value of DWT-DCT-SVD. This can be attributed to the nature of RivaGAN which is created specifically for video invisible watermarking.

It can also be concluded that the hybrid algorithm of DWT-DCT-SVD shown the best criteria of Imperceptibility as it topped the PSNR value of 47.00 on comparing original image and encoded image. DWT-DCT-SVD also shown the best NCC value among the algorithms.

7.0 ACKNOWLEDGEMENTS

The authors would like to thank Tunku Abdul Rahman University College (TAR UC) for providing financial support and technical support when completing this study.

REFERENCES

Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. Data in Brief. 2020 Feb;28:104863. DOI: 10.1016/j.dib.2019.104863.

Al-Haj, A. (2007). Combined DWT-DCT Digital Image Watermarking. Journal of Computer Science, 3(9), 740–746.

Abdulrahman, A. K., & Ozturk, S. (2019). A novel hybrid DCT and DWT based robust watermarking algorithm for color images. Multimedia Tools and Applications, 78(12), 17027–17049.

El-Shazly, E. H. M. (2004). Digital Image Watermarking in Transform Domains. Minufiya University.

Khare, P., & Srivastava, V. K. (2020). A Secured and Robust Medical Image Watermarking Approach for Protecting Integrity of Medical Images. Transactions on Emerging Telecommunications Technologies, 32(2).

Kuang, L.-Q., Zhang, Y. and Han, X. (2009). A Medical Image Authentication System Based on Reversible Digital Watermarking. 2009 First International Conference on Information Science and Engineering.

Lala, H. (2017). Digital image watermarking using discrete wavelet transform. International Research Journal of Engineering and Technology (IRJET), 4(01).

Sverdlov, A., Dexter, S., & Eskicioglu, A. M. (2005). Robust DCT-SVD domain image watermarking for copyright protection: embedding data in all frequencies. In 2005 13th European Signal Processing Conference (pp. 1-4). IEEE.

Tao H, Chongmin L, Zain JM, Abdalla AN (2014) Robust image watermarking theories and techniques: a review. J Appl Res Technol 12(1):122–138

Voloshynovskiy, S., S. Pereira and T. Pun, 2001. "Attacks on Digital Watermarks: C lassification, Estimation-Based Attacks, and Benchmarks," Comm. Magazine, 39(8): 118-126

Vukotić, V., Chappelier, V., & Furon, T. (2020). Are Classification Deep Neural Networks Good for Blind Image Watermarking? Entropy, 22(2), 198.

Yeung, M. M. (1998). Invisible watermarking for image verification. Journal of Electronic Imaging, 7(3), 578.

Zhang, K. A., Xu, L., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Robust invisible video watermarking with attention. arXiv preprint arXiv:1909.01285.