X-Ray, MRI and CT scan Medical Image Compression using Huffman Coding

Mohammed Hazim Alkawaz
Faculty of Information Sciences and Engineering
Management and Science University
Shah Alam, Selangor, Malaysia
mohammed_hazim@msu.edu.my

Muhammad Akmal Shaik Mydin
School of Graduates Studies
Management and Science University
Shah Alam, Selangor, Malaysia
mohddakmal57@gmail.com

Md Gapar Md Johar
Information Technology and Innovation Centre
Management and Science University
Shah Alam, Selangor, Malaysia
mdgapar@msu.edu.my

2022 IEEE Symposium on Industrial Electronics & Applications (ISIEA) | 978-1-6654-8012-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIEA54517.2022.9873664

Abstract - *In hospitals, modern medical imaging generates a large amount of medical pictures. The necessity of medical image compression in telehealth cannot be overstated because it provides for excellent image quality for the medical images containing substantial diagnostic information while requiring less storage space, allowing them to be conveniently transmitted across a network and accessible in a short period of time. This research assessed the performance analysis of several image compression methods for the medical images such as X-rays, MRIs, and CT scans. Image quality parameters such as Compression Ratio (CR), Peak Signal to Noise Ratio (PSNR), and Mean Structural Similarity Index (MSSI) have been used to assess the quality of medical pictures following compression using Huffman coding.*

Keywords: *Mean Structural Similarity Index, Huffman Coding, Medical Image, Compression Ratio, Noise Ratio, Peak Signal*

I. INTRODUCTION

Generally, compression means the process of compressing large datasets thus that they'd be stored in even less storage space than they might normally require. The Image compression is a method of reducing it size of an image file without sacrificing quality. Image compression's primary purpose is to minimise the correlations between one pixel as well as its neighbouring pixel. [1]. The amount of data required to symbolise such images are huge.

As a result, the transmission rate is extremely slow and the operating costs are expensive. Like a nutshell, image information should be compressed by removing only visible parts that must be encoded [2]. The amount of data utilised to represent an image should be reduced to improve the compression rate of the data used. Because of this differentiating trait, image compression techniques are commonly used to reduce storage costs in a variety of applications, notably in the telemedicine business [3]. It is also used to improve the efficiency of medical image and data transfer within the available bandwidth. Currently, the majority of hospitals save medical picture data digitally [4]. The effective image compression algorithms assist doctors in retaining image quality even after compression and extracting crucial information with the least amount of storage space. Medical image compression is necessary in all relevant sectors such as medical image database, medical image telecommunications, and so on as a result of data digitization [5].

Generally, each image is indeed the matrix composed of a rectangular array of dots known as the pixels that are organised in rows and columns. Images of various types are used in a range of industries, including remotely sensed data, video data, healthcare sectors, biometrics, and satellite photos systems, all of which require appropriate compression techniques for storage or transmission across or inside networks [6]. A medical image is a graphic depiction of organs or bodily components in medicine. In today's clinical applications, data about the human internal organs is employed for a range of diagnostics and practical objectives. There are X-rays, Magnetic Resonance Imaging (MRI), The Ultrasound Scans, and other types of medical pictures available [7].

Uncompressed images need a substantial portion of RAM memory. The Medical image compression is essential in the reduction of storage space and transmission needs, as well as the preservation of all pertinent information obtained following the diagnosis [8]. Almost every neighbouring pixel in a picture also have correlation, resulting mostly in generation of data redundancy. The goal of compression would be to transform pictures such that there is less connection with neighbouring pixels, resulting in less duplicate information. Redundancy and useless data elimination are the two main parts of image compression [9].

The remaining part of the paper is organized as follows:
Section II involves the detailed description of compression techniques that using Huffman coding. Section III shows the image compression quality parameters and medical image involved in compression. Section IV involves the performance analysis of different parameters and the results obtained using compression algorithms. The paper is concluded in Section V.

II. LITERATURE REVIEW

The author of this study used the Huffman coding technique to compress medical images. Huffman coding is a lossless picture compression algorithm. To some extent, Huffman code is an excellent image compression method [10]. Image and video compression methods have advanced significantly during the previous two decades. To improve coding efficiency, Variable length coding, such as the Huffman code, is widely utilized. Whenever the encoder knows the probability distribution of a source of data, it uses the Huffman source-coding technique to create the uniquely decipherable Huffman code with the shortest estimated codeword size [11]. Entropy may be described as a measure of relevant information since it can count the number of bits utilised in the data contained in a given picture.

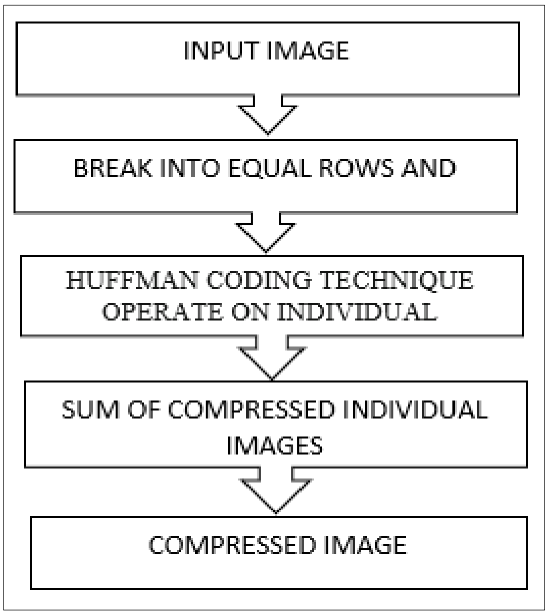


Figure 1: Flow chart of technique used

Figure 1 shows the flowchart of the Huffman technique. Huffman coding utilizes a particular approach for selecting the representation for specific images, resulting in a prefix code. The Huffman Code Algorithm may be used to compress both images and data. Using the Huffman algorithm, we can build the most efficient compression approach [12].

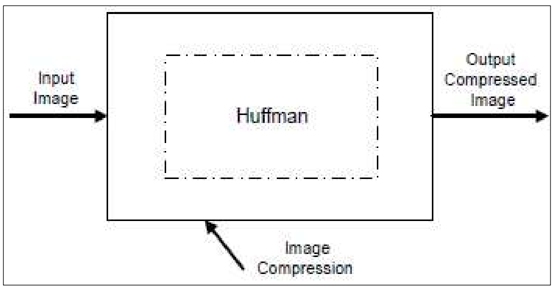


Figure 2: Diagram of image compression using Huffman coding.

Figure 2 is the block diagram of the Huffman encoding method for picture compression. The Huffman encoding starts by calculating the likelihood of each symbol in the picture. The symbol probability is organised in decreasing order, forming the tree structure of a tree. Once the symbols are encrypted independently, the Huffman code is generated by combining the least likely symbols, a technique that is repeated until just two probabilities of two compound symbols remain [13]. As a result, a code tree is created, and Huffman codes are derived by labelling the code tree. The symbols 0 and 1 are, of course, the shortest binary code for a two-symbol source. Reading the branch digits sequentially from the root node to the relevant terminal node or leaf yields the Huffman codes for the symbols [14]. As can be seen in Figure 3, Huffman code is the most often used approach for reducing coding redundancy. The Huffman code technique is based on the two findings listed below.

1) Symbols with a higher frequency of occurrence will have shorter code words than symbols with a lower frequency of occurrence.

2) The two least often appearing symbols will have the same length.

Overall average length of the code is calculated by averaging the product of the symbol's probability and the number of bits utilised to encode it. The Huffman coding efficiency is computed as equation below:

$$Huffman\ Code\ Efficiency\ is = \frac{entropy}{the\ average\ lenght} \quad (1)$$

Huffman's technique finds the best pattern for a given set of symbols and probabilities, with both the constraint that symbols be coded as well at the same time.

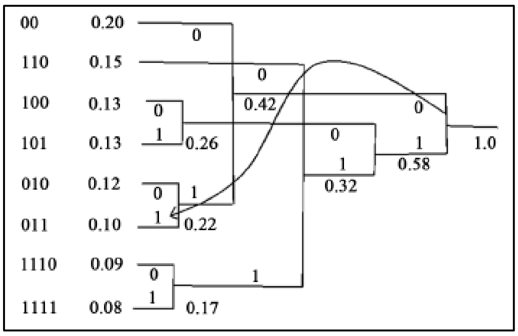


Figure 3: Huffman coding procedure

An entropy (H) of such sources may be computed using the equation below (2). The entropy value specifies however much compression does possible in terms of bits per pixel for that picture. That is important to note that no compression ratio may be less than entropy [15]. The average information probability is used to calculate the entropy of every picture. The amount of bits required per pixel is represented by entropy.

$$H = -\sum_{k=1}^{L} p\,(a_k) \log p\,(a_k) \qquad (2)$$

in which pk is the probability of intensity, k denotes the intensity value, and L denotes the number of intensity values utilised to present the picture.

III. METHODOLOGY

The following are the three most significant criteria to consider when comparing an uncompressed image to a compressed image:

A. Compression Ratio (CR)

A compression ratio, also referred as compression strength, is a calculation of how much space is saved by a compression technique. It is usually expressed as the uncompressed size divided by the compressed value. Image quality and storage space are improved when the compression ratio is higher [16]. N1 and N2 represent the number of bits in the uncompressed image and the number of bits in the compressed image, respectively.

$$CR = \frac{\textit{original image size}}{\textit{compressed image size}} = \frac{N1}{N2} \qquad (3)$$

B. Peak Signal to Noise Ratio (PSNR)

It's a widely applied measurement for comparing the actual image to the compressed image. As enhances the quality of the compressed or reconstructed picture, the higher the PSNR. The main block uses the following equation to compute the Mean-Squared Error (MSE) in order to compute the PSNR [17]. (4):

$$MSE = \frac{\sum_{M,N}\{ I_1(m,n) - I_2(m,n) \}^2}{M * N} X\, 100 \qquad (4)$$

M and N in the equation above represent the number of rows and columns in the input images. The PSNR for grayscale images is then determined using the equation below (4):

$$PSNR = 10 \log_{10} [\frac{R^2}{MSE}] \qquad (5)$$

The greatest fluctuation in the input image data type is represented by R (=255) in the preceding equation. The PSNR concept is the same for color images with three RGB values for each pixel, with the exception of MSE, which is the sum of all the squared values of the differences separated by image size and by 3. PSNR for lossy image and video compression is typically between 30 and 50 dB, with higher values [18]. The PSNR for color images having R, G, and B color components is provided in the equation below (6).

$$PSNR = 10 \log_{10} [\frac{255^2}{\frac{MSE(R) + MSE(G) + MSE(B)}{3}}] \qquad (6)$$

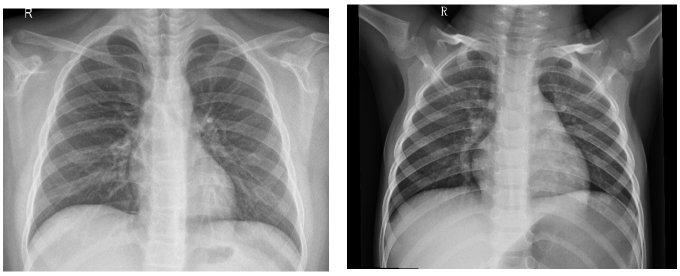
C. Mean Structural Similarity Index (MSSI)

The Mean Structural Similarity Index is a quality tool used to compare the overall picture quality of the original (X) and the decompressed image (Y). PSNR is most commonly applied to measure the quality of compression reconstruction [19] In this case, the signals represent the actual data, whereas the noise represents the compression error. When comparing compression codecs, PSNR is an estimate of human judgement of reconstruction quality. This computation is based on Equation (7) [20-25].

$$MSSI(x,y) = \frac{1}{M} \sum_{j=1}^{M} \frac{(2\mu_x \mu_y + c_1)(2cov_{xy} + c_2)}{(\mu_{x^2} + \mu_{y^2} + c_1)(\sigma_{x^2} + \sigma_{y^2} + c_2} \qquad (7)$$

In this research we have compressed the medical images using Huffman Code. There are 3 type medical image used which is X-Ray, MRI & CT Scan image. Below are original medical images used to compressed.

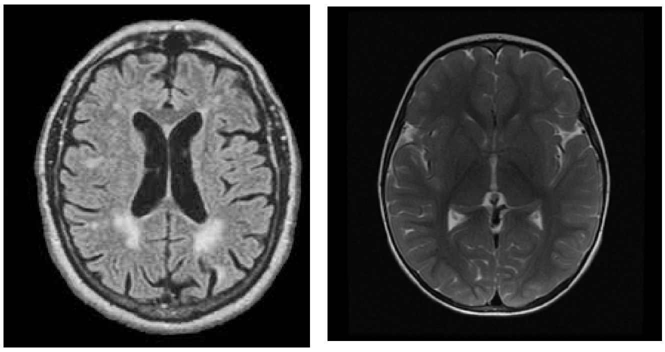
A. Original X-Ray Image



Figures 4: Original X-Ray Image A & B

The figure 4 above are original medical images which are took from 2 different patient. The uncompressed image resolutions are different from each other. Image at left represent as image A and image at right is represent as image B.

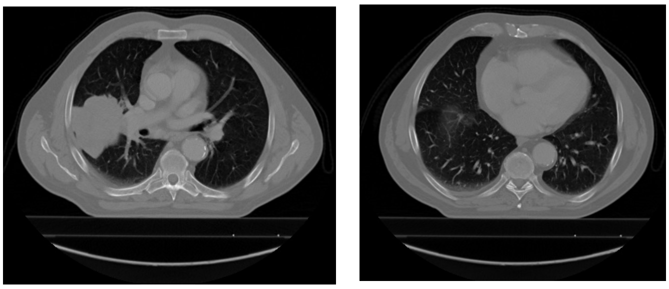
B. Original MRI Image



Figures 5: Original MRI Image A & B

Figure 5 above are actual MRI medical images from 2 separate patients. Uncompressed image resolutions differ from one another. The picture on the left is image A, while the one on the right is image B.

C. Original CT-Scan Image



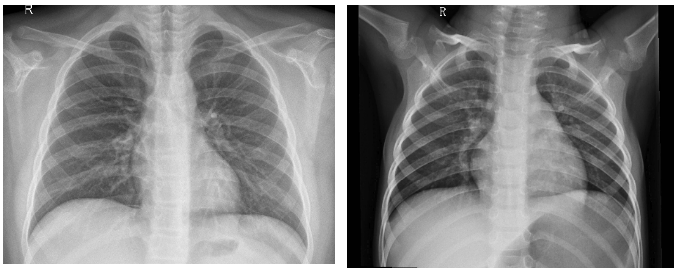
Figures 6: Original MRI Image A & B

The Figure 6 above are original medical images which are took from 2 different patient. The uncompressed image resolutions are different from each other. Image at left represent as image A and image at right is represent as image B.

IV. RESULT & DISSCUSSION

Above medical images are calculated by using 3 image quality parameters, which is Compression Ratio (CR), Peak Signal to the Noise Ratio (PSNR) and the Mean Structural Similarity Index (MSSI).

A. Compressed X-Ray Image



Figures 7: Compressed X-Ray Image A & B

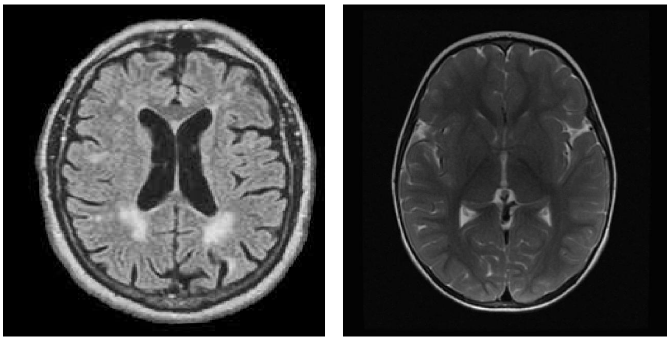
Figure 7 above shows compressed X-Ray medical images that are compressed with Huffman Code. Below table is result of image quality measurement for compressed X-Ray medical images.

Table 1: Compressed X-Ray image details

Image	Image Quality Parameters		
	CR	***PSNR***	***MSSI***
A	0.9460	49.48	0.9920
B	0.9730	48.79	0.9807

From Table 1 result above, image A is compressed well based on image quality measurement because image A has higher Peak Signal Noise Ratio (PSNR) and Mean Structural Similarity Index (MSSI) than image B.

B. Compressed MRI Image



Figures 8: Compressed MRI Image A & B

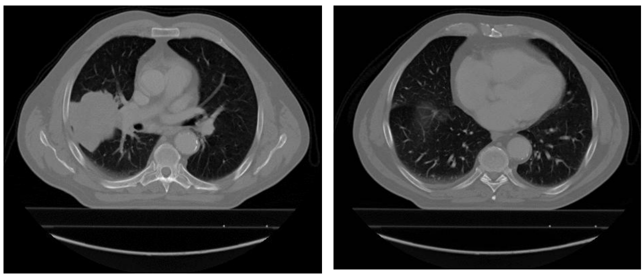
Figure 8 above shows compressed MRI medical images that are compressed with Huffman Code. Below table is result of image quality measurement for compressed MRI medical images.

Table 2: Compressed MRI image details

Image	Image Quality Parameters		
	CR	***PSNR***	***MSSI***
A	0.9730	43.66	0.9926
B	0.9730	40.03	0.8989

Based on Table 2 result above, Image A is compressed well because image A has higher Peak Signal the Noise Ratio (PSNR), the Mean Structural Similarity Index (MSSI) than image B.

C. Compressed CT Scan Image



Figures 9: Compressed MRI Image A & B

Figure 9 above shows compressed CT scan medical images that are compressed with Huffman Code. Below table is result of image quality measurement for compressed MRI medical images.

Table 3: Compressed CT Scan image details

Image	Image Quality Parameters		
	CR	***PSNR***	***MSSI***
A	1.0270	40.96	0.9598
B	1.0270	40.82	0.9395

From Table 3 result above, image A is compressed well based on image quality measurement because image A has higher Peak Signal the Noise Ratio (PSNR) and Mean Structural Similarity Index (MSSI) than image B.

Below chart shows result of all compressed medical images by image quality parameter, which are Compression Ratio (CR), Peak Signal Noise Ratio (PSNR), Mean Structural Similarity Index (MSSI).

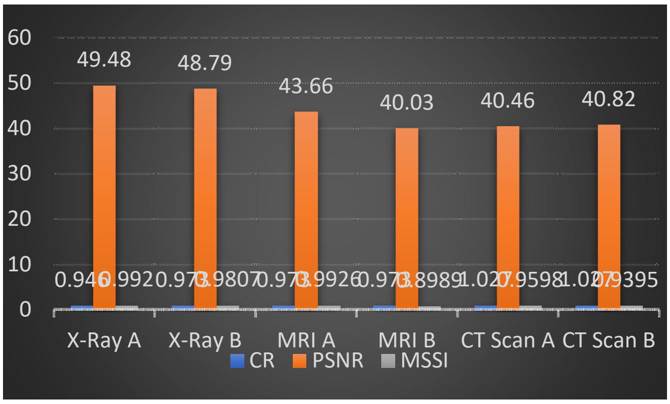


Figure 10: Result of all compressed medical images by image quality parameter.

Figure 10 above shows result of all compressed medical images by image quality parameter. From above chart and table result, author can conclude that X-ray image are compressed fine compared to others because X-ray medical image has higher Peak Signal Noise Ratio (PSNR) which quality metric for comparing the original image to the compressed image is better than MRI and CT scan.

V. CONCLUSION

In this study, the author uses MATLAB to simulate a medical image compression technique called Huffman coding. There are 3 medical images are used, which are X-ray, MRI and CT scan images. Image quality parameters was used after compression, which are Compression Ratio (CR) and Peak Signal Noise Ratio (PSNR) and Mean Structural Similarity Index (MSSI). Since X-ray medical pictures have a greater Peak Signal Noise Ratio (PSNR) than MRI and CT scan images, the author concludes that they are compressed well.

ACKNOWLEDGEMENT

The Authors are thankful to the Faculty of Information Sciences and Engineering as well as to School of Graduates Studies, Management and Science University at Malaysia for their support. The conference fee is funded by the Grant No: (Ref: FRGS/1/2019/ICT02/MSU/02/1), Research Management Centre (RMC), Management and Science University (MSU), Shah Alam, Selangor, Malaysia.

REFERENCES

[1] Ansari M.A., R.S. Anand "Comparative Analysis of Medical Image Compression Techniques and their Performance Evaluation for Telemedicine", Proceedings of the International Conference on Cognition and Recognition, pp. 670-677, 2014.

[2] Sanjith S, Ganesan R, "Evaluating the Quality of Compression in Very High Resolution Satellite Images Using Different Compression Methods",International Journal of Engineering Research in Africa, Vol. 19, pp. 91-102,Mar. 2016.

[3] Smitha Joyce Pinto, Prof. Jayanand P.Gawande "Performance analysis of medical image compression techniques", Proceedings of the Institute of Electrical and Electronics Engineers IEEE, 2012.

[4] Satyawati S. Magar, Bhavani Sridharan, Innovative approach to biomedical image compression using oscillation concept, 2016, ICACDOT.

[5] Jagadish H. Pujar and Lohit M. Kadlaskar "A New Lossless Method Of Image Compression And Decompression Using Huffman Coding Techniques" Journal of Theoretical and Applied Information Technology , 2010

[6] Yongfei Zhang, Haiheng Cao, Hongxu Jiang and Bo Li "Visual Distortion Sensitivity Modeling for Spatially Adaptive Quantization in Remote Sensing Image Compression" IEEE Geoscience And Remote Sensing Letters, Vol. 11, No. 4, April 2014

[7] Mishra, K., Verma, R.L., Alam, S. and Vikram, H. (2013) Hybrid Image Compression Technique using Huffman Coding Algorithm. *International Journal of Research in Electronics & Communication Technology*, **1**, 37-45.

[8] Prabakaran.S, Samjaikumar.S. (2017). A Survey on Medical Image Compression Algorithms and Its Performance. IJSRD - International Journal for Scientific Research & Development|. 5. 2321-613.

[9] Abo-Zahhad, M., Gharieb, R.R., Ahmed, S.M. and Khaled, M. (2015) Brain Image Compression Techniques. *International Journal of Engineering Trends and Technology* (*IJETT*), 19, 93-105.

[10] Gupta, K., Verma, R.L. and Sanawer Alam, Md. (2013) Lossless Medical Image Compression Using Predictive Coding and Integer Wavele Transform based on Minimum Entropy Criteriat. *International Journal of Application or Innovation in Engineering & Management* (*IJAIEM*), **2**, 98-106

[11] Dr. K. Kuppusamy, R.Ilackiya "Fractal Image Compression & Algorithmic Techniques" International Journal of Computer & Organization Trends –Volume3 Issue 4 , May 2013.

[12] M. A. Shaik Mydin, M. H. Alkawaz, K. Z. Ghafoor, O. F. Mohammad and M. G. Md Johar, "A Study on Medical Image Compression Techniques based on Huffman Coding and Discrete Wavelet Transform (DWT)," *2021 IEEE 9th Conference on Systems, Process and Control (ICSPC 2021)*, 2021, pp. 86-91, doi: 10.1109/ICSPC53359.2021.9689166.

[13] Q. Yu *et al*., "Medical image compression with thresholding denoising using discrete cosine-based discrete orthogonal stockwell transform," *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, 2017, pp. 1857-1860, doi: 10.1109/CompComm.2017.8322860.

[14] R. Patel, V. Kumar, V. Tyagi and V. Asthana, "A fast and improved Image Compression technique using Huffman coding," *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2016, pp. 2283-2286, doi: 10.1109/WiSPNET.2016.7566549.

[15] Ö. S. Büyükçolak, "Simultaneous Image Encryption and Compression Using Modified Huffman Tables," *2019 Seventh International Conference on Digital Information Processing and Communications (ICDIPC)*, 2019, pp. 47-51, doi: 10.1109/ICDIPC.2019.8723751.

[16] R. Channe, S. Ambatkar, S. Kakde and S. Kamble, "A Brief Review on: Implementation of Lossless Color Image Compression," *2019 International Conference on Communication and Signal Processing (ICCSP)*, 2019, pp. 0131-0134, doi: 10.1109/ICCSP.2019.8698027.

[17] Z. Luo and Y. Wan, "An efficient framework for lossless color image compression," *2016 International Conference on Audio, Language and Image Processing (ICALIP)*, 2016, pp. 380-384, doi: 10.1109/ICALIP.2016.7846634.

[18] A. H. M. Z. Karim, M. S. Miah, M. A. Al Mahmud and M. T. Rahman, "Image Compression using Huffman Coding Scheme with Partial/Piecewise Color Selection," *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)*, 2021, pp. 1-6, doi: 10.1109/GUCON50781.2021.9573863.

[19] R. Praisline Jasmi, B. Perumal and M. Pallikonda Rajasekaran, "Comparison of image compression techniques using huffman coding, DWT and fractal algorithm," *2015 International Conference on Computer Communication and Informatics (ICCCI)*, 2015, pp. 1-5, doi: 10.1109/ICCCI.2015.7218137.

[20] P. A. I. Savitri, Adiwijaya, D. T. Murdiansyah and W. Astuti, "Digital medical image compression algorithm using adaptive Huffman coding and graph-based quantization based on IWT-SVD," *2016 4th International Conference on Information and Communication Technology (ICoICT)*, 2016, pp. 1-5, doi: 10.1109/ICoICT.2016.7571902.

[22] Alkawaz, M. H., Steven, S. J., Mohammad, O. F., & Johar, M. G. M. (2022). Identification and Analysis of Phishing Website based on Machine Learning Methods. In 2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE) (pp. 246-251). IEEE.

[23] Kadhim, K. A., Mohamed, F., Khudhair, Z. N., & Alkawaz, M. H. (2020). Classification and predictive diagnosis earlier Alzheimer's disease using MRI brain images. In 2020 IEEE Conference on Big Data and Analytics (ICBDA) (pp. 45-50). IEEE.

[24] Al-Sanjary, O. I., Ibrahim, O. A., & Sathasivem, K. (2020). A new approach to optimum steganographic algorithm for secure image. In 2020 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS) (pp. 97-102). IEEE.

[25] Al-Sanjary, O. I., Roslan, M. A. B., Helmi, R. A. A., & Ahmed, A. A. (2020). Comparison and detection analysis of network traffic datasets using K-means clustering algorithm. Journal of Information & Knowledge Management, 19(03), 2050026.