Migration Letters

Volume: 20, No: S5(2023), pp. 442-451

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)
www.migrationletters.com

Hybrid Algorithm Compression Technique for Color Images

Muneera Yousif[1], Riham Jassim[2], Noor M. Abdullah[3], Sundus M. A. Sulaiman[4], Noor M. A. Fari[5], Muthana Jasim Mohammed[6], Dr. Khalill Alssif[7]

Abstract

The importance of different data compression algorithms is highlighted in the process of storing big data in data warehouses and archives and in reducing the volume of big data during its transmission through communication channels, so many compression algorithms have emerged that are designed to process different data.

Images are used in different fields in our daily lives, such as social networks, medical diagnoses, remote sensing, and other fields, and since the size and accuracy of images are constantly evolving, the issue of storing and transferring images requires a compressing process to reduce their size. A hybrid based on lossy compression method and lossless compression method based on Huffman algorithms and wavelet transform. In this research, the efficiency of the proposed method was studied and compared to independent compression methods.

Keywords: *Images, Algorithm Compression Technique, data warehouses.*

1. Introduction

Digital image processing is defined as performing some mathematical operations that result in a new form of the original image, such as rotation, reduction, enlargement, adjusting lighting, contrast, coloring, and compressing the image, which is useful in storing thousands of images, especially in games and animation programs that contain a large number of images [1]. In view of the expansion of the field of using digital images and video, this led to the development and increase of compression techniques, as well as the development of devices capable of capturing high-quality images. With the increasing amount of data, a way must be found to reduce the size of the data as an alternative solution to increasing the storage capacity. Therefore, the researchers resorted to using methods to compress the data by reducing the size of the image by canceling the repetition of (bits) in the image data while preserving the content of the image and the type of information for archiving and facilitating. The transmission and transmission process, as the digital image data is arranged from a binary matrix representing rows and columns of pixels [2][3].

Image compression is used in the field of image processing for the development of applications that depend on images and multimedia, such as medical imaging, remote sensing, teleconferencing applications, and others [4]. From memory, compressed images

[1] muneera.y.y91@gmail.com
[2] rihamjassim11@gmail.com
[3] noor.mohammed7655@gmail.com
[4] sundus.90.mohammed@gmail.com
[5] noormalifaris@gmail.com
[6] engmothanajasem@gmail.com
[7] Khalil_alsaif@hcu.edu.iq

may also be characterized by being of higher clarity, and this helps to improve the images, as well as the speed of sending information, as well as it helps to encrypt the files sent, and this supports the confidentiality of the information sent, and this is one of the advantages of the compressing process [5]. On the other hand, we note that compressing has some disadvantages, including the loss of part of the information that may be very important, which may cause distortion of information when retrieved, and that some data may need special techniques to compress it [6].

In this paper, a hybrid method was used that combines between lossy compression and lossless compression through wavelet and Huffman transform techniques to compress image data (JPEG) with high accuracy and retrieve it. Error (MSE) The results of the proposed pressing method were also studied and compared with the used pressing methods separately and using the same image data.

2. Digital Images

Images play an important role in human acquisition of information, as human perception of images is limited to the visible range of the electromagnetic spectrum, but cameras can capture the entire electromagnetic spectrum, starting from (gamma) rays all the way to radio waves, as there are images that are generated by ultrasound or Computer-generated images and electron microscopy. Digital images consist of a specific number of elements called pixels, which are represented by position and value [7].

Digitally, the image is represented by a two-dimensional function f (x, y) where x, y represents plane coordinates, and the amount of the function f at point (x, y) represents the gray level or intensity, and thus these elements are called pixels. When f, x, y all belong to a set of specified values, that is the digital image [8].

3. Compression Types

Compressing images is of great importance in image processing, as it contributes to getting rid of duplicate and unimportant data, which leads to reducing the size of the image, which leads to saving storage space in memory as well as speeding up the process of sending and transmitting data through means of communication [8]. The image compression system consists of two basic operations: compression and decompression or data retrieval. The compression process consists of two stages, which are the initial processing stage and the data encoding stage, while the decompressing process consists of the decoding process and the final processing of the data [9]. Depending on the possibility of reconstructing an exact copy of the original image using the compressed image or not, compression techniques were divided into two types: lossless compression and lossy compression [10].

3.1 Lossless Compression Technique

This type of compression technique compresses data without any loss of data, damage, or negative impact on image quality after uncompressing [11], meaning that the compressed data is identical to the original data after removing the compression [8]. This type of algorithm usually consists of two stages, the first is to convert the original image into another encoding to reduce color redundancy, while the second stage is to use (entropy encoder) to remove redundancy from the coding, while the decompression algorithm is the reverse process of the compression algorithm [12]. The type of compression technique is used in applications with stringent requirements, such as in medical imaging. The most prominent of these techniques are Run length encoding, Huffman encoding, LZW coding, and Area coding [10].

3.2 Lossy Compression Technique

Lossy compression techniques are used when the images are of high quality, as some of the data is lost in these techniques because the original data does not match the data that has been decompressed completely, as the compression technique affects the quality of the data [11] and thus achieves higher compression ratios One of the lossless compression techniques [10], where a group of pixels can be approximated to a single value using transformation or differential encoding [9], and data that is beyond the reach of human perception is ignored so that it does not affect the people who deal with that data [13]. This type of compression technology is usually used to compress files such as MP3, JPEG, MPEG, and others [11].

The proposed compression algorithm is based on the principle of hybridization of two compression algorithms, one of which is Huffman coding, which is one of the lossless compression techniques, and the other depends on the lossy compression technique, which is the wavelet transform algorithm.

4. Huffman Encoding

In 1952, David Huffman presented a new method for encoding data, which he called Huffman coding, which relies on generating a binary coding tree based on the probability of a character appearing in the plain text, i.e. Replacing a series of binary numbers (1,0) on the alphabet of information, which increases the confidentiality of the data, the coding tree structure has the lowest average symbolic word length [14]. Also, the output from the generated tree is a variable length encoding. This technique consists of two steps. The first is to create a Huffman tree through the input strings. The second step is to navigate through the tree to determine the character codes. Huffman encoding is characterized by simplicity of application and speed in the pressing process, so this type of coding is still commonly used [11]. at data compression, the letters are arranged in descending order and considered as nodes of the tree diagram. This step is repeated if the tree is composed of more than one node. After that, every two nodes that have the least probability of repeating the frequency are searched for to form a new node that has a probability equal to the sum of the two nodes. Then each pair of branches of the tree diagram is encoded in a binary image [10]. This type of encoding is used in text, video and image compression systems such as (JPEG) and (MPEG-4) [9].

5. Wavelet Transform

Data compression technique based on transformations is one of the most effective applications of wavelets, and it is one of the powerful digital signal processing tools, as it analyzes the digital signal into different frequency bands and different densities by analyzing the signal into detailed information and approximate information [15]. This method was presented by the geophysical engineer (J.Morlet) in the seventies of the last century, as he used a window of variable size to analyze the signals at specific frequencies and because the window length is small and fluctuating, the window function used in the analysis was called the “wavelet function”, that the wavelet transform is used In the analysis of time series of non-uniform power and variable frequencies [16]. Wavelet transform is used in the analysis of signals with short duration and high frequency, which leads to obtaining good accuracy of time and low frequency, and vice versa [17]. The Convolutional Filter Bank method can be adopted to apply the wavelet transform. The signal is analyzed by passing it through the analysis filters, which include High Pass Filters and Low Pass Filters, followed by the decimation process, using the High Pass Filter. Pass Filter leads to the appearance of low-scale elements, i.e. High-frequency elements, which are called detailed items. As for the Low Pass Filter, it leads to the emergence of high scale elements, i.e. Low frequency elements, which are called

Approximation [18]. The following is a diagram of the compressing process using the wavelet transform [19].

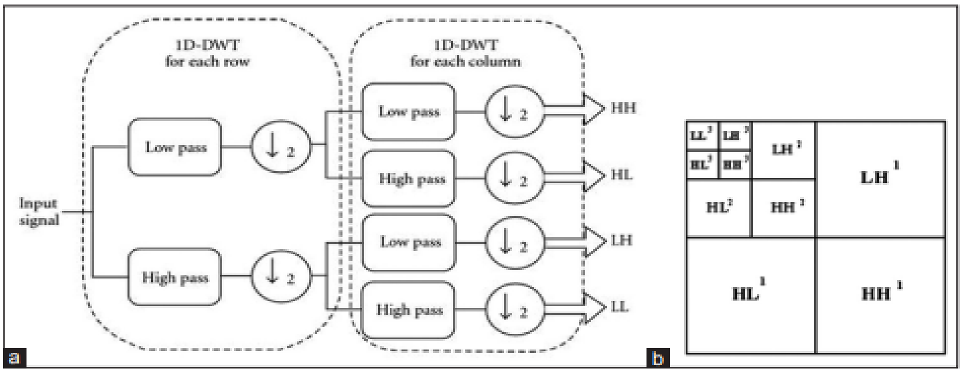

Figure (1) Scheme of the compression process using the wavelet transform

(a) structure of the filters for the wavelet transform of two-dimensional matrices

(b) wavelet transform of three level

6. Hybrid Compression

The process of hybridizing the two compression algorithms aims to encode the data and reduce its size by using Huffman coding to encode the data and the compression algorithm using the wavelet transform. Figure (2) represents the general scheme of the compression and decompression method.

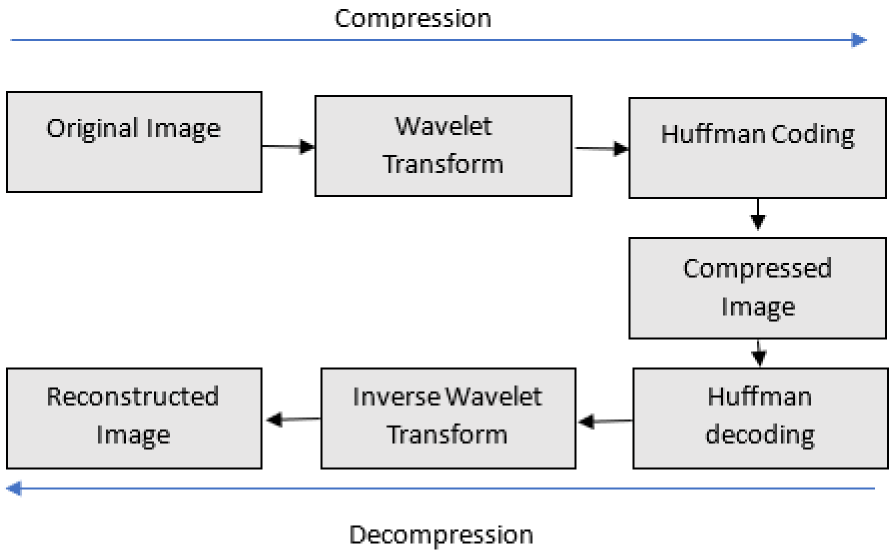

Figure (2) The general scheme of the proposed compression and decompressing method

The first step of Hybrid compression is image captured by the digital camera or stored on the hard disk is read, and then the color layers of the image are separated and the red layer is adopted. After completing the initial processing, the stage of compressing the image begins using intermittent wavelet transform, which includes a number of steps:

1- The image is divided into four parts. The first part represents the level of low frequencies, which is the most important approximate part and is symbolized by (LL), while the remaining parts comprise the high frequencies, which include the fine details of the image and are symbolized by (HL, LH, HH).

In order to obtain these parts, the following steps must be performed [16]:

1- Each row in the image is passed horizontally to the Low Pass filter, and then the samples of the image rows are reduced and the output is stored in the RRLImage matrix.

2- Passing the image columns resulting from step (1) vertically to the Low Pass Filter, storing the output in the RRLLImage matrix, and then reducing the samples of the RRLLIMAGE image columns to get (LL).

3- Passing the resulting image columns in step (1) horizontally to the high pass filter and storing the output in the RRLH matrix and reducing the samples of the image columns to obtain (LH).

4- Pass the image rows horizontally to the High Pass Filter, reduce the samples and store the output in the RRHImage matrix.

5- Pass the columns of the image resulting from step (4) vertically to the low-frequency filter and then reduce the samples in the RRHLImage to produce (HL)

6- Pass the image resulting from step (4) horizontally to the High Pass Filter and reduce the samples of the RRHHImage columns to produce (HH).

2- After the image has been compressed using the wavelet transform, the image (LL) is taken. This image is converted into three color layers and the red color layer is used, and then the image histogram is calculated, followed by the probability calculation process by dividing the total number of repetitions of each pixel by the total number The number of pixels in the image and their arrangement.

3- the image is encoded by calculating the dictionary for Huffman encoding and is approved for encoding the image to be stored.

The steps of decompression of the hybrid compression can be implemented through the reverse application of the hybrid compression algorithm as follows:

1- The compression that was made is decoded by adopting Huffman coding

2- Constructing the compressed image from the wavelet transform by implementing it in reverse to obtain the image.

The stage of reading the image and initial processing, where the image is re-sized, and then the red layer is extracted from the image and approved in the compression process, as in Figure (3).

Figure (3) stages of image reading and initial processing

The next step is the coding stage using the wavelet transform by reducing the lines and columns in the red-gradient image, as shown in Figure (4).

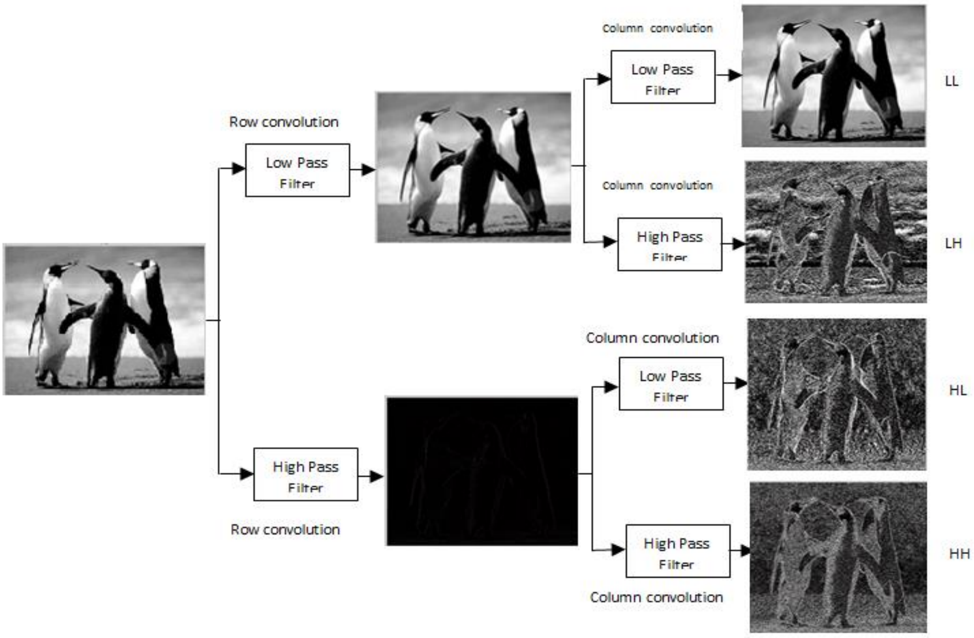

Figure (4) Compression stages using wavelet transform

The third stage is to find the histogram of the image (LL) and the probability, and then arrange the probability values and arrange the image values depending on the ordered values of the probability, as in Table (1):

Table (1) histogram section, probability, image redistribution based on probability

Histogram	probability	Arranged probability	Image values arranged according to the arranged probability
9467	**0.0479**	**0.0008**	**0.0008**
17263	**0.0874**	**0.0008**	**0.0009**
2145	**0.0109**	**0.0009**	**0.0479**
.….	**.….**	**..….**	**.….**
.….	**.….**	**..….**	**.….**
.….	**.….**	**..….**	**.….**
689	**0.0013**	**0.0016**	**0.0874**
718	**0.0011**	**0.0017**	**1.0000**
.….	**.….**	**..….**	**.….**
.….	**.….**	**..….**	**.….**
.….	**.….**	**..….**	**.….**
487	**0.0025**	**0.0560**	**2.0000**
2185	**0.0111**	**0.0847**	**254.0000**
.….	**..….**	**.….**	**.….**

The fourth stage is the stage of creating a dictionary for encoding using Huffman to obtain the final compressed image, as shown in Figure (5).

Dictionary	
[1x11 double]	[0]
[1x11 double]	[1]
...	...
...	...
[1x4 double]	[254]
[s1x4 double]	[255]

Figure (5): (a) the dictionary on the left of the image (b) e image after implementing the hybrid press

The hybrid decompress is the reverse process of the hybrid press.

7. Performance Metrics:

There are several different scales used with compression operations to measure compression ratio, mean square error, and other measurements. In this research, some types of these scales were used, including:

1. Compression Ratio: It is calculated by using the ratio of the image size before compression to the image size after compression, as in equation (1)

$$Compression\ Ratio = \frac{Original\ Image\ Size}{Compressed\ Image\ Size} \ldots (1)$$

2. Mean Square Error (MSE): It is one of the most used measures to represent the difference between images. And the value of the mean error refers to the error difference between the values of the image points before compression and the values of the image points after compression. The measure of mean squared error can be represented by equation (2)

$$MSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[f(i,j)-f'(i,j)]^2}{N\times M} \ldots (2)$$

Where M,N represent the dimensions of the image, f(i,j) represents the size of the image before the compression process, and f' (i,j) represents the size of the image after the compression process [20].

8. Experiments and Results

In this section, we will analyze and evaluate the results of our hybrid algorithm. By calculating and comparing the results of the image quality measures, as we relied in this research on three types of compression techniques, which are compression using wavelet transform, compression using Huffman encoding, and compression using the hybrid algorithm, MSE and CR were adopted as measures for the efficiency of the compression algorithms. Where images of different sizes (5.05Kb), (17.0Kb), (64.8Kb) were adopted, to apply the three compression algorithms then calculate the scales. Table (2), table (3) represents the results of applying Huffman coding and the wavelet algorithm, table (4) show the results of applying the hybrid algorithm.

Table (2) Huffman coding

Image Name	Image Size	Compression Ratio	MSE
Image1	5.05KB	0.529844	107.6382
Image1	17.0KB	0.515266	116.4645
Image1	64.8KB	0.507722	126.0151

Table (3) Wavelet transform

Image Name	Image Size	Compression Ratio	MSE
Image1	5.05KB	0.995029	14587.048
Image1	17.0KB	0.994639	15348.074
Image1	64.8KB	0.994543	15785.211

Table (4) Hybrid

Image Name	Image Size	Compression Ratio	MSE
Image1	5.05KB	0.995051	12969.35
Image1	17.0K	0.994651	13872.49
Image1	64.8KB	0.99454	14399.66

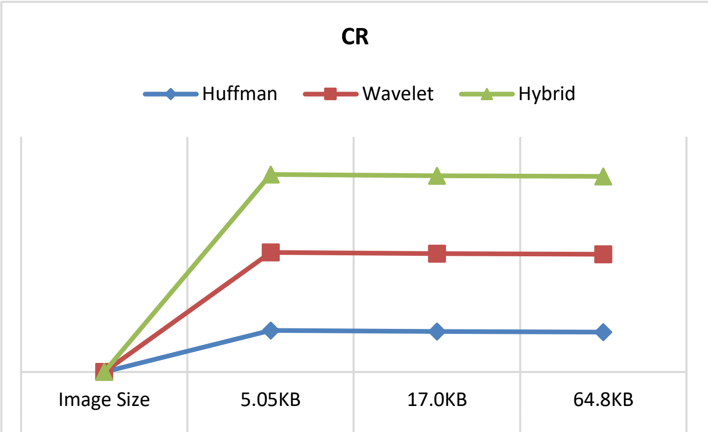

Figure (3) Compression ratio

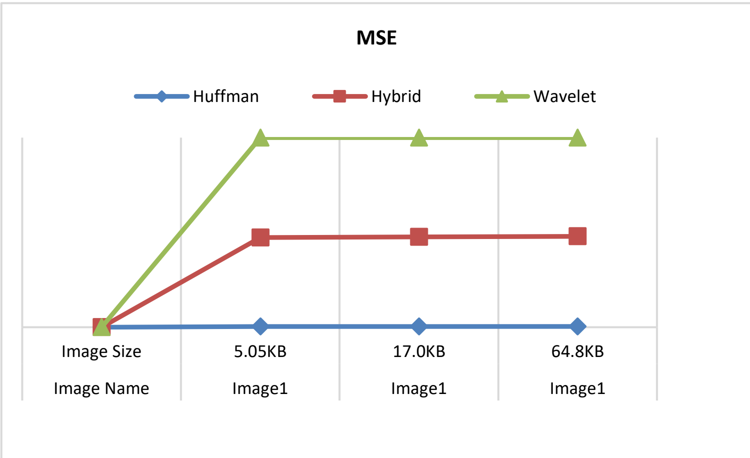

Figure (4) Mean square error

The Results above show that the compression ratio using the hybrid algorithm depends on the file size to which the algorithm is applied, as the hybrid algorithm depends on the amount of repetition in pixel values, meaning that the compression ratio increases if the image has a dominant background, but in the wavelet transform algorithm, the compression process It leads to reducing the size of the image to a quarter of the original size on the first level, while it reaches 1/16 in the second level, and therefore we find that the use of hybrid methods that adopt lossy and lossless compression together lead to a noticeable increase in the compression ratio.

9. Conclusions

The compression ratio in the case of wavelet transform is constant at each level of image decomposition in the wavelet transform. As for the compression ratio based on H.M, it clearly depends on the quality of the image, and the increase in the frequency leads to an increase in the compression ratio. In hybridization algorithms, it led to an increase in the compression ratio, because the first reduction was due to wavelet transformation, and then compression through H.M. coding.

References

[1] R. C. Gonzalez and R. E. (Richard E. Woods, Digital image processing. Prentice Hall, 2008.

[2] A. K. Kadhim, “An advanced image compression technique by using a coupled compression algorithm depend on different wavelet methods,” Periodicals of Engineering and Natural Sciences, vol. 9, no. 1, pp. 11–28, 2021.

[3] A. K. Kadhim, A. B. S. Merchany, and A. Babakir, “An Improved Image Compression Technique Using EZW and SPHIT Algorithms,” Ibn AL-Haitham Journal For Pure and Applied Sciences, vol. 32, no. 2, pp. 145–155, May 2019, doi: 10.30526/32.2.2121.

[4] J. S. Kulchandani, S. H. Pal #2, and K. J. Dangarwala, “Image Compression : Review and Comparative Analysis,” International Journal of Engineering Research & Technology (IJERT), vol. 3, no. 11, pp. 586–590, 2014, [Online]. Available: www.ijert.org

[5] S. Taoufiq and N. N. Sheikh, “Image Compression Using Discrete Wavelet Transforms Capstone Project Final Report,” Al Akhawayn University , Morocco, 2016.

[6] Vinay U. Kale and Shirish M. Deshmukh, “Visually Improved Image Compression by Combining EZW Encoding with Texture Modeling using Huffman Encoder,” INternational Journal of Computer Science Issues, , vol. 7, no. 3, pp. 28–39, 2020, [Online]. Available: www.IJCSI.org.

[7] Abd al-Khaleq Noman, Salam, “Improving Digital Images,” Al-Fath Magazine, vol. 4, no. 35, pp. 82–91, 2008.

[8] Khenifer Fatima, “Image compression using the Speck coding algorithm,” Faculty of Science and Technology, University of Mohamed Kheidar Biskra, People's Democratic Republic of Algeria, 2020.

[9] A. S. Abd-Alzhra and M. S. H. Al- Tamimi, “Image Compression Using Deep Learning: Methods and Techniques,” Iraqi Journal of Science, pp. 1299–1312, Mar. 2022, doi: 10.24996/ijs.2022.63.3.34.

[10] A. Vasanthakumari and M. Phil, “IMAGE COMPRESSION TECHNIQUES,” IJARIIE, vol. 1, no. 5, pp. 2395–4396, 2015, [Online]. Available: www.ijariie.com

[11] N. N. P, S. Nagaraj, S. S. P, and T. Kulkarni, “Comparative Analysis of Lossless Data Compression Techniques,” 2022. Accessed: Aug. 16, 2023. [Online]. Available: www.ijesi.org

[12] Ibrahim Mohammad and K. Abuzanouneh, “Hybrid Algorithm for Enhancing and Increasing Image Compression Based on Image Processing Techniques,” IJCSIS, vol. 16, no. 3, 2018, [Online]. Available: https://sites.google.com/site/ijcsis/

[13] E. Setyaningsih and A. Harjoko, “Survey of hybrid image compression techniques,” International Journal of Electrical and Computer Engineering, vol. 7, no. 4, pp. 2206–2214, 2017, doi: 10.11591/ijece.v7i4.pp2206-2214.

[14] R. R. Gajjala, S. Banchhor, A. M. Abdelmoniem, A. Dutta, M. Canini, and P. Kalnis, “Huffman Coding Based Encoding Techniques for Fast Distributed Deep Learning,” in DistributedML 2020 - Proceedings of the 2020 1st Workshop on Distributed Machine Learning, Association for Computing Machinery, Inc, Dec. 2020, pp. 21–27 . doi: 10.1145/3426745.3431334.

[15] S. J. Hadi, M. Tombul, S. Q. Salih, N. Al-Ansari, and Z. M. Yaseen, “The Capacity of the Hybridizing Wavelet Transformation Approach with Data-Driven Models for Modeling Monthly-Scale Streamflow,” IEEE Access, vol. 8, pp. 101993–102006, 2020, DOI: 10.1109/ACCESS.2020.2998437.

[16] M. Y. Yaqoob Mohammed and M. Celik, "Developing Fast Techniques for Periodicity Analysis of Time Series," in 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2019 - Proceedings, Institute of Electrical and Electronics Engineers Inc., Oct. 2019. DOI: 10.1109/ISMSIT.2019.8932865.

[17] N. S. Godwin et al., "Design and Optimization of Image Compression Algorithm using Wavelet Transform for Satellite Imagery," Int. j. adv. multidisc. res. stud, vol. 2, no. 3, pp. 200–205, 2022, [Online]. Available: www.multiresearchjournal.com

[18] S. Gunanandhini, M. Kalamani, M. Bhagavathipriya, and S. Guru Prasath, "Wavelet based Video Compression techniques for Industrial monitoring applications," in Journal of Physics: Conference Series, Institute of Physics, 2022. DOI: 10.1088/1742-6596/2272/1/012019.

[19] Z. Khodabandeh, H. Rabbani, A. M. Dehnavi, and O. Sarrafzadeh, "The ellipselet transform," J Med Signals Sens, vol. 9, no. 3, pp. 145–157, Jul. 2019, doi: 10.4103/jmss.JMSS_42_17.

[20] A. Gudi, F. Büttner, and J. van Gemert, "Proximally Sensitive Error for Anomaly Detection and Feature Learning," Jun. 2022, doi: DOI:10.48550/arXiv.2206.00506.