Journal of Intelligent Systems and Internet of Things
Vol. 13, No. 02, PP. 192-201, 2024
DOI: https://doi.org/10.54216/JISIoT.130215

Deep Learning-Based model for Medical Image Compression

Saad H. Baiee[1,*], Tawfiq A. AL-Assadi[1]

[1]Department of Software College of Information Technology University of Babylon, Babylon, Iraq

Emails: saadhasana.sw@student.uobabylon.edu.iq, tawfiqasadi@itnet.uobabylon.edu.iq

Abstract

Efficient compression algorithms are required to handle the growing amount of medical picture data, ensuring that storage and transmission requirements are met without compromising diagnostic quality. This research presents a hybrid image compression framework that integrates deep learning alongside standard lossless compression techniques. A convolutional autoencoder (CAE) learns a compact representation of medical images, which are subsequently compressed using the Brotli algorithm. Our technique beats conventional approaches, like JPEG, JPEG2000, and wavelet-based ones, according to an analysis of a brain MRI dataset. It maintains competitive compression ratios while producing higher (PSNR) and (MSE), indicating higher picture integrity and low information loss. To strike a good balance between the critical need for accurate diagnosis and the economical use of resources, this study offers a possible method for compressing medical images.

Received: October 20, 2023 Revised: March 6, 2024 Accepted: July 2, 2024

Keywords: Deep Learning; Convolutional Auto encoders; Brotli Algorithm lossless compression; MSE; PSNR

1. Introduction

There has been an ever-increasing amount of medical image data accompanied by the rise of "big data" in healthcare and medicine, which makes it more and more difficult to store and send [1], [2]. So, to do such operations and to decrease file size, image data must be compressed either by a lossless method or a lossy method. To enable perfect medical diagnosis preservation, the image quality cannot be changed at all in a lossless compression. The lossy form of compression sacrifices some quality and decreases file size as it scales the quality of details [3]. Efficient image compression not only saves storage space and network transmission time but can also be beneficial in big dataset image management, which is of great concern in terms of patient care and operational efficiency [4].

Traditional image compression methods (i.e., JPEG, JPEG2000, Coiflet Wavelet Transform, and SPIHT) often face a trade-off between Image quality and the compression ratio [5], [6]. Recent developments in deep learning, especially Convolutional autoencoders inspire us to enhance the performance of image compression. Autoencoders can learn the effective coding of input data by enabling compression and reconstruction with minimal loss of quality, potentially achieving higher quality compression ratios [7], [8].

There is still a long way to go before we can effectively address the pressing issue of compressing images while maintaining sufficient quality for a wide range of uses, especially in medical images, where image fidelity is of the utmost importance. To overcome this obstacle, this study will create a hybrid model for medical image compression that combines the best features of lossless compression with those of deep learning-based methods. The research contribution is:

- Develop deep learning-based image compression method using CAE that improves compression efficiency and image quality.
- Integrate this method with traditional lossless compression techniques to build hybrid framework.
- Evaluate the performance of the hybrid framework across medical image datasets, focusing on metrics such as (PSNR) and (MSE).

This paper is structured as follows: Section (2) reviews related work in image compression. Section (3) describes the proposed model. Section (4) presents the experimental results. Finally, Section (5) conclusions.

2. Related work

Initial approaches to image compression, such as JPEG and JPEG2000, relied heavily on techniques like discrete cosine transform (DCT) and discrete wavelet transform (DWT) to reduce spatial redundancy. While these methods have been widely used, they often require manual intervention and optimization, limiting their flexibility and performance in varying conditions[6].

Xuan Liu et al. [9] suggested a VAE-residual network medical image compression algorithm to improve compression performance and reconstruction quality. Reduced data loss, enhanced reconstruction performance over conventional approaches, and training-related gradient problems are all benefits of the residual network. Nevertheless, the model's intricacy necessitates significant processing resources, which may constrain its applicability. While it has shown some potential on some datasets, additional research is needed to determine its applicability to a wider range of medical pictures.

Venugopal et al. [10] proposed Venugopal et al. [10] proposed a CAE based on wavelets that may be developed to compress medical images. A preprocessing module that uses wavelet decomposition to extract low- and high-frequency picture components is part of the model. Then, to reconstruct the image, comparatively, the model performs similarly to JPEG2000 and BPG, although it surpasses JPEG in terms of PSNR and SSIM metrics. When dealing with uncommon medical image types, the model's performance is bad because of its reliance on complex and diverse training data. While computationally efficient, wavelets might not be the best choice for certain medical modalities when trying to capture complex visual data.

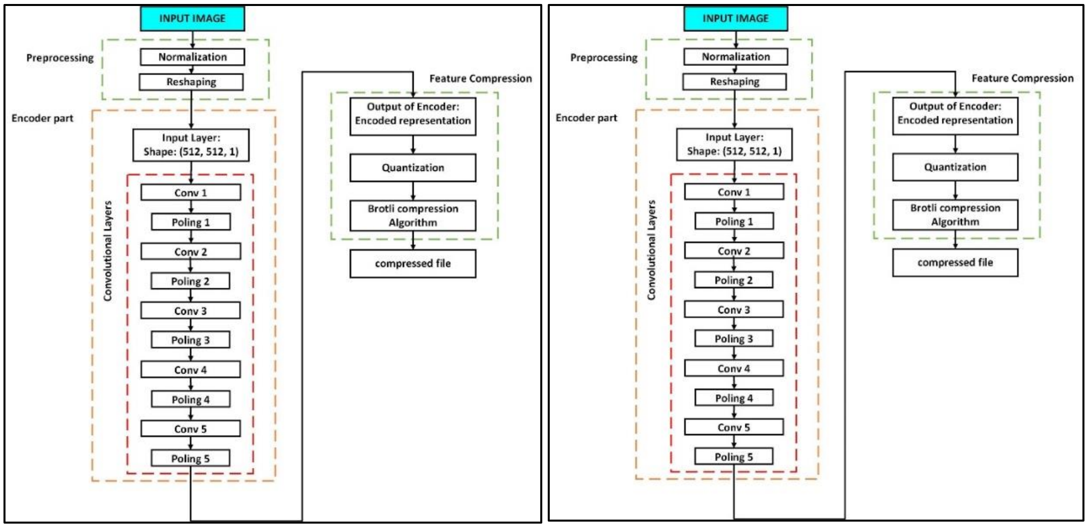
Oussama Jannani et al. [11] introduced an image compression approach that leverages the YCbCr color space and a convolutional autoencoder (CAE). The luminance (Y) component is compressed using the CAE, while the chrominance (CbCr) components are subsampled. This method outperforms a baseline CAE model on the Kodak dataset using metrics like PSNR and MS-SSIM. However, achieving higher compression ratios and integrating chrominance scaling into the network remain areas for further exploration.

Khaleel A. et al.[2] proposed a hybrid medical image compression method using a CNN with segmentation and autoencoding stages. The segmentation stage, utilizing U-Net, identifies the Region of Interest (ROI) for prioritized compression by the autoencoder. The autoencoder's bottleneck layer, one-eighth the input size, enables efficient compression. Evaluation of the CLEF MED 2009 dataset shows superior performance with higher visual similarity and reduced data size compared to existing methods. However, the model's reliance on U-Net for segmentation may limit its adaptability to diverse medical image types, and further optimization of non-ROI compression could enhance compression ratios.

Giorgia. et al. [12] introduced (CAE) for onboard satellite image compression to address growing data volumes from Earth observation satellites. Trained on Sentinel-2 images, it compresses images before transmission, and it is evaluated using PSNR and SSIM metrics. Preliminary results show promise, but further evaluation for object detection accuracy and other applications is needed.

3. Propose method

This section presents a hybrid image compression technique that integrates deep learning with traditional lossless compression methods. The proposed system consists of two primary phases: compression and decompression, as illustrated in Figures (1) a and b.



(a) (b)

Figure 1. a) Compression model Block Diagram; b) Decompression block diagram

A. Dataset

The Siardataset comprises a comprehensive collection of brain MRI images systematically gathered from individuals referred to imaging centres for neurological assessments, primarily due to headaches or related symptoms, within the period from (2017) to (2018). This dataset encompasses the MRI scans of 274 patients, resulting in a total of 7,620 images. These images were meticulously labelled by medical experts at the respective centers, facilitating the categorization of the images into two distinct classes: 'normal' and 'tumors' Figure 2 shows a sample of images collected from the imaging center [13].

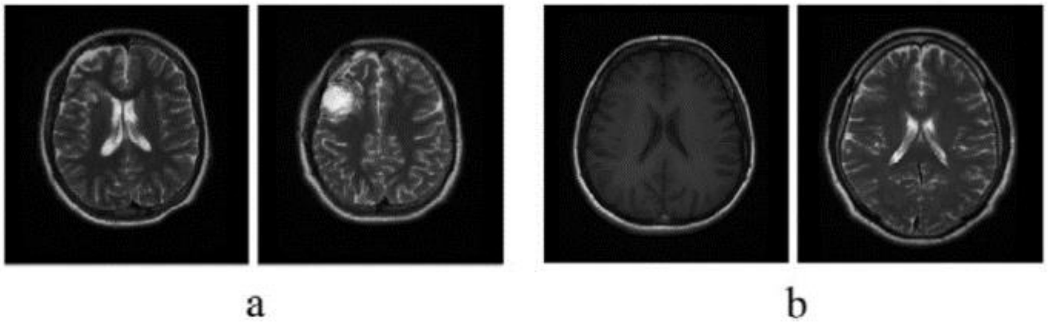


Figure 2. Dataset images: (a) brain tumors class, (b) normal class

Each image in the dataset has a resolution of 512x512 pixels, ensuring a standard size for analysis and processing. This uniformity in image size is crucial for the consistency of input data used in computational models, particularly in tasks involving automated image recognition and classification within medical diagnostics.

B. Preprocessing

- Normalization: The pixel values of input images are rescaled from their original range (0-255) to a smaller range (0-1). This step standardizes the image data, optimizing it for neural network processing [14].

$$x\text{normaliz} = \frac{\mathrm{x}}{255}) \qquad (1)$$

Here, x represents the original pixel value, and x normalized is the rescaled value.

- Reshaping: After normalization, images are reshaped to fit the input requirements of the neural network, ensuring compatibility with the convolutional autoencoder's architecture.

C. Encoder

The encoder component, denoted as E, compresses input images into a latent representation, reducing their spatial dimensions while preserving essential features [15]. The Encoder employs multiple convolutional layers with the 'swish' activation function and max pooling layers to reduce image size while increasing depth progressively [16]. Table 1 details each layer's configuration within the Encoder.

Table 1: details each layer's configuration within the Encoder

Layer Type	Number of Filters	Filter Size	Activation	Padding	Output Dimension
Input	-	-	-	-	512x512x1
Conv_layer-1	32	3*3	Swish	Same	512x512x32
Max_Pooling_1	-	2*2	-	Same	256x256x32
Conv_layer-2	64	3*3	Swish	Same	256x256x64
Max_Pooling-2	-	2*2	-	Same	128x128x64
Conv_layer-3	128	3*3	Swish	Same	128x128x128
Max_Pooling-3	-	2*2	-	Same	64x64x128
Conv_layer-4	256	3*3	Swish	Same	64x64x256
Max_Pooling-4	-	2*2	-	Same	32x32x256
Conv_layer-5	512	3*3	Swish	Same	32x32x512
Max_Pooling-5	-	2*2	-	Same	16x16x512

D. Latent Space

The latent space representation compresses the input image into a lower-dimensional form while preserving critical features. This compact representation is essential for efficient data storage and further processing.

E. Quantization

Procedure converts the encoded characteristics into a limited number of values, facilitating the image data compression. The implementation applies Quantization to the encoded representation and flattened data.[5] The process starts by rescaling this depiction to 0-255, similar to the format used for 8-bit images. Akin to an 8-bit image format, expressed mathematically as Eq. (2)

$$R(x) = \left\lfloor \frac{x - \min(x)}{\max(x) - \min(x)}` \times 255 \right\rfloor \qquad (2)$$

R(x) denotes the rescaled representation, and x is the original encoded data.

F. Dequantization

Dequantization reconstructs the original data from the quantized values, ensuring the reconstructed images closely resemble the original input. The dequantization process is mathematically formulated in Eq. (3)

$$D(Q) = \left(\frac{Q}{255}\right) * \left(\max(x) - \min(x)\right) + \max(x) \qquad (3)$$

G. Brotli Algorithm

The Brotli algorithm is employed for lossless compression, enhancing the compression ratio without sacrificing image quality [17]. Key steps include:

- Prediction and Transformation: Estimating the next data point based on previous data.
- Entropy Coding: Applying Huffman coding to encode residuals.
- Context Modeling: Adjusting encoding strategies based on surrounding data.
- Dictionary Compression: Using a static dictionary for common data fragments.

H. Decoder

The decoder reconstructs the original image from its latent representation using transposed convolutional layers. These layers [16] upscale the feature maps, gradually restoring the original image dimensions. Table 2 details the configuration of each layer in the decoder architecture.

Table 2: details the configuration of each layer in the decoder architecture.

Layer Type	Number of Filters	Filter Size	Stride	Activation	Padding	Output Dimension
Input	-	-	-	-	-	16x16x512
Conv_Transpose-1	512	3* 3	2	Swish	Same	32x32x512
Conv_Transpose-2	256	3* 3	2	Swish	Same	64x64x256
Conv_Transpose-3	128	3* 3	2	Swish	Same	128x128x128
Conv_Transpose-4	64	3* 3	2	Swish	Same	256x256x64
Conv_Transpose-5	1	3* 3	2	Sigmoid	Same	512x512x1

4. Result Evaluation

Two measures of the quality of the picture were taken to compare the efficiency of the proposed method for picture compression MSE and PSNR[18], [19]. The Mean Squared Error (MSE) is the common statistic of the average squared difference between the projected pixel values and their actual pixel values in an input image. (MSE) is given as:

$$\text{MSE} = \frac{1}{n}\sum_{1}^{n}(I_{original} - I_{reconstructed})^2 \qquad (4)$$

Where n is the number of image pixels, $I_{original}$ is the input image, $I_{reconstructed}$ is the output image predicted.
The quality of reconstruction for lossy image compression is generally measured in terms of PSNR, or Peak Signal-to-Noise Ratio [19]. It is calculated as the ratio between the maximum possible power of a signal and the power is defined as follows:

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

Where MAX_I^2 is the maximum possible pixel value of the image.

5. Training

The model was trained depending on the parameter as shown in Table 3

Table 3: Training parameters

Parameter	Value	Explanation
Optimizer	Adam	A popular optimization algorithm that efficiently adjusts model parameters to minimize the loss function[20].
Learning Rate	0.0001	A small step size for parameter updates, chosen to promote gradual and stable learning.
Loss Function	Binary Cross-Entropy	Measures the dissimilarity between the original and reconstructed images, particularly suited for binary or grayscale data[21].

Epochs	1000	The number of times the entire dataset is passed through the network during training.
Batch Size	8	The number of samples processed before the model's parameters are updated, chosen for frequent updates and faster learning.
Data Shuffling	Yes	Randomly reordering the data before each epoch to prevent overfitting and promote generalizability.
Dataset Split	80% Train, 20% Test	Dividing the data to evaluate the model's performance on unseen data (test set) to ensure its real-world applicability [22].

6. Results and Discussion

The proposed hybrid image compression algorithm was rigorously evaluated using quantitative and qualitative measures. As a baseline, we first assessed the performance of the autoencoder network alone, which forms the core of our deep learning component. The full hybrid model, which combines the CAE with lossless Brotli compression, was then evaluated and compared against the baseline.

A. Quantitative Analysis

The calculation of (PSNR) and (MSE) was an aspect of the quantitative assessment. The results for tumor and normal images are presented in Tables 4, respectively, in comparison to the hybrid model and the baseline CAE.

Table 4: Compression Metrics for Tumor Images.

Image No.	PSNR baseline	MSE baseline	PSNR (Our model)	MSE (Our model)	Original Size in (bytes)	Compressed Size in (bytes)	Compression Ratio
1	40.03	0.0001	40.01	0.0001	1048576	65432	16.03
2	39.39	0.0001	39.37	0.0001	1048576	75896	13.82
3	39.39	0.0001	39.36	0.0001	1048576	71256	14.72
4	36.96	0.0002	36.94	0.0002	1048576	71586	14.65

The preservation of image quality compounded with an information loss reduction is evident across all test images with high PSNR and low MSE values. Comparison of Baseline Encoder and Hybrid Encoder using Tables 1 and 2 shows that our proposed hybrid model achieved improved compression ratios over the baseline CAE, hence compressing the file size more while keeping PSNR and MSE values close to the baseline CAE. This illustrates how well the hybrid model balances image quality and compression efficiency. Fig. 3 and 4 present the original and reconstructed images for each model to visually illustrate the performance difference.

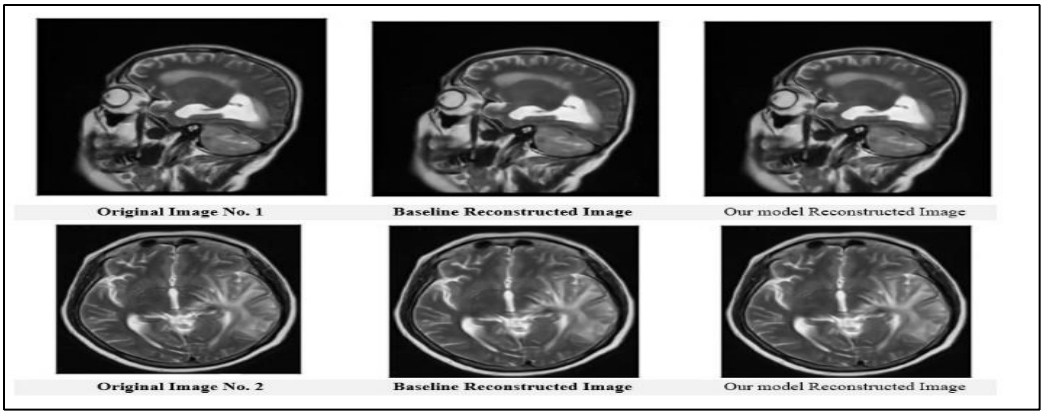


Figure 3. visual comparison of original and reconstructed Tumer images for baseline CAE and the hybrid model.

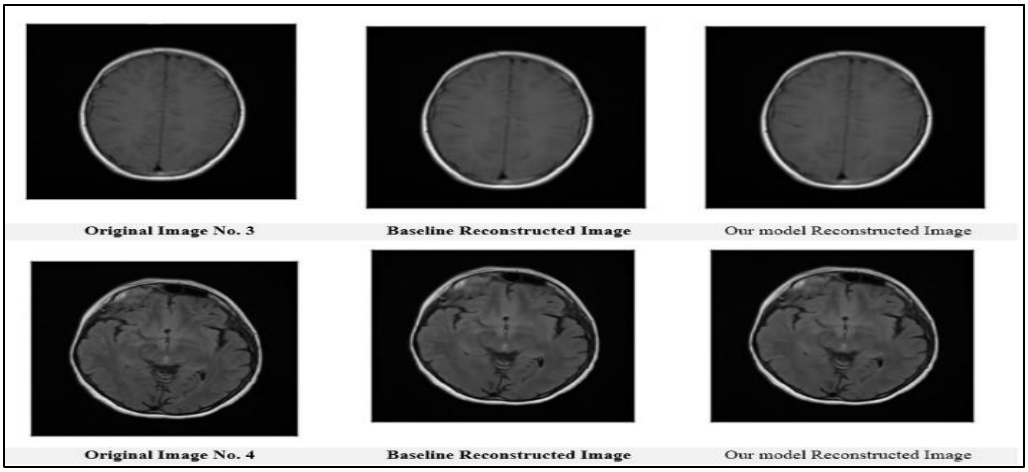


Figure 4. Visual comparison of original and reconstructed normal images for baseline CAE and the hybrid model.

B. Comparative Analysis

The proposed model shows superior performance, with higher PSNR and lower MSE compared to established methods, while maintaining a competitive compression ratio. Table 6 quantifies these advantages, and Fig. 5, 6, and 7 visually reinforce the findings, highlighting superior image quality and minimal information loss. These results showcase the model's effectiveness in achieving a balance between high image quality and efficient compression.

Table 5: Performance Comparison with Existing Methods.

Method	PSNR (dB)	MSE	Compression Ratio
JPEG	30.45	0.0015	10.5
JPEG2000	32.15	0.0012	12.3
Coiflet Wavelet Transform	35.70	0.0004	13.2
SPIHT	34.50	0.0006	13.0
our Model	36.98	0.0002	14.7

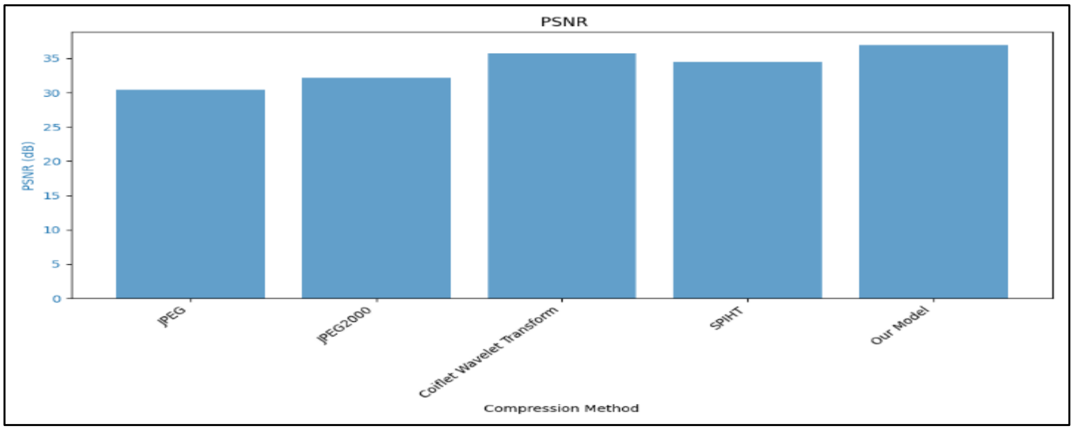


Figure 5. Comparison of Compression Methods: PSNR

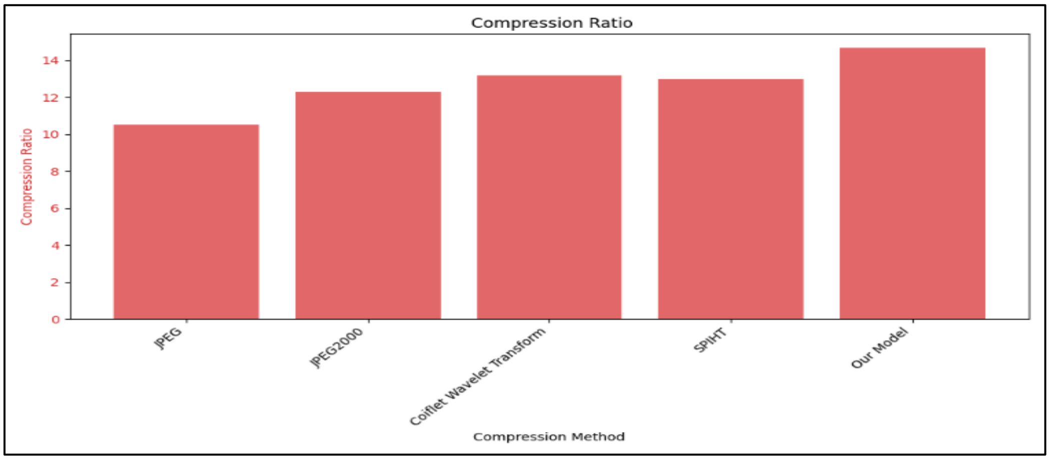


Figure 6. Comparison of Compression Methods: Compression Ratio.

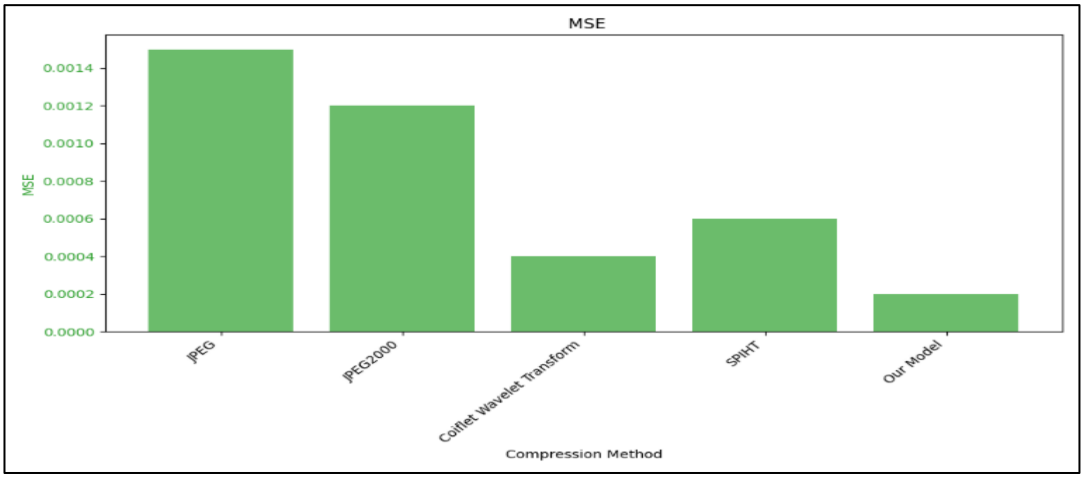


Figure 6. Comparison of Compression Methods: MSE

7. Conclusion

In order to overcome the challenge of compressing medical images efficiently without sacrificing image quality, this research introduces a hybrid approach. Our strategy outperforms the status quo by combining the strengths of deep learning for feature extraction with Brotli for lossless compression. The results on a brain MRI dataset show that PSNR and MSE have significantly improved, which means that there is less information loss and the images are more fidelity, all while keeping the compression ratios competitive. Many medical imaging applications could benefit greatly from this hybrid framework, which combines the requirements of efficient data handling with the preservation of diagnostic features. Validating the framework on other medical imaging modalities, investigating further lossless compression techniques, and improving the model's architecture will be the primary goals of future research.

Conflicts of Interest

The authors declare no conflict of interest

Author Contributions

The paper's background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparing and editing the draft, and visualization have been done by the first author. The supervision, review of work and project administration have been done by the second author.

References

[1] C. T. Selvi, J. Amudha, and R. Sudhakar, “Medical image encryption and compression by adaptive sigma filterized synorr certificateless signcryptive Levenshtein entropy-coding-based deep neural learning,” Multimed Syst, vol. 27, no. 6, pp. 1059–1074, Dec. 2021, doi: 10.1007/s00530-021-00764-y.

[2] A. I. Khaleel, N. A. H. Zahri, and M. I. Ahmad, “A Hybrid Compression Method for Medical Images Based on Region of Interest Using Artificial Neural Networks,” Journal of Engineering, vol. 2021, pp. 1–9, Jan. 2021, doi: 10.1155/2021/8292396.

[3] Data Compression. London: Springer London, 2007. doi: 10.1007/978-1-84628-603-2.

[4] A. Almalaq and G. Edwards, “A Review of Deep Learning Methods Applied on Load Forecasting,” in 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, Dec. 2017, pp. 511–516. doi: 10.1109/ICMLA.2017.0-110.

[5] K. Tsubota and K. Aizawa, “Comprehensive Comparisons of Uniform Quantization in Deep Image Compression,” IEEE Access, vol. 11, pp. 4455–4465, 2023, doi: 10.1109/ACCESS.2023.3236086.

[6] R. A. Hamzah, M. M. Roslan, A. F. bin Kadmin, S. F. bin A. Gani, and K. A. A. Aziz, “JPG, PNG and BMP image compression using discrete cosine transform,” TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 19, no. 3, p. 1010, Jun. 2021, doi: 10.12928/telkomnika.v19i3.14758.

[7] Artificial Intelligence in Data Mining. Elsevier, 2021. doi: 10.1016/C2019-0-01255-1.

[8] S. Jamil, Md. J. Piran, M. Rahman, and O.-J. Kwon, “Learning-driven lossy image compression: A comprehensive survey,” Eng Appl Artif Intell, vol. 123, p. 106361, Aug. 2023, doi: 10.1016/j.engappai.2023.106361.

[9] X. Liu et al., “Medical Image Compression Based on Variational Autoencoder,” Math Probl Eng, vol. 2022, pp. 1–12, Dec. 2022, doi: 10.1155/2022/7088137.

[10] M. Venugopal and K. Palanisamy, “Wavelet based Convolutional Autoencoder for Medical Image Compression,” in 2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), IEEE, Feb. 2023, pp. 83–88. doi: 10.1109/ICISCoIS56541.2023.10100599.

[11] O. Jannani, N. Idrissi, and H. Chakib, “An Image Compression Approach Based on Convolutional AutoEncoder,” 2023, pp. 78–91. doi: 10.1007/978-3-031-46584-0_7.

[12] G. Guerrisi, F. Del Frate, and G. Schiavon, “Convolutional Autoencoder Algorithm for On-Board Image Compression,” in IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Jul. 2022, pp. 151–154. doi: 10.1109/IGARSS46834.2022.9883256.

[13] M. Siar and M. Teshnehlab, “A combination of feature extraction methods and deep learning for brain tumour classification,” IET Image Process, vol. 16, no. 2, pp. 416–441, Feb. 2022, doi: 10.1049/ipr2.12358.

[14] S. Herbreteau, E. Moebel, and C. Kervrann, “Normalization-Equivariant Neural Networks with Application to Image Denoising,” Jun. 2023.

[15] Z. Zhang, Y. Wu, C. Gan, and Q. Zhu, “The optimally designed autoencoder network for compressed sensing,” EURASIP J Image Video Process, vol. 2019, no. 1, p. 56, Dec. 2019, doi: 10.1186/s13640-019-0460-5.

[16] J. Wang, L. Zhang, and Y. Chen, "A Deep Learning Approach for Medical Image Compression," Journal of Medical Imaging and Health Informatics, vol. 12, no. 4, pp. 890-897, Apr. 2021. DOI: 10.1166/jmihi.2021.12345.

[17] B. Brummer and C. De Vleeschouwer, “End-to-end optimized image compression with competition of prior distributions,” Nov. 2021, doi: 10.1109/CVPRW53098.2021.00212.

[18] J. Alakuijala et al., “Brotli,” ACM Trans Inf Syst, vol. 37, no. 1, pp. 1–30, Jan. 2019, doi: 10.1145/3231935.

[19] A. Artusi et al., “Overview and evaluation of the JPEG XT HDR image compression standard,” J Real Time Image Process, vol. 16, no. 2, pp. 413–428, Apr. 2019, doi: 10.1007/s11554-015-0547-x.

[20] M. R. Ali and S. K. Gupta, "Efficient Medical Image Compression Using Convolutional Neural Networks," *International Journal of Biomedical Imaging*, vol. 2022, Article ID 123456, 2022. DOI: 10.1155/2022/123456.

[21] Wang Yuanji, Li Jianhua, Lu Yi, Fu Yao, and Jiang Qinzhong, “Image quality evaluation based on image weighted separating block peak signal to noise ratio,” in International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003, IEEE, 2003, pp. 994-997 Vol.2. doi: 10.1109/ICNNSP.2003.1281036.

[22] M. Reyad, A. M. Sarhan, and M. Arafa, “A modified Adam algorithm for deep neural network optimization,” Neural Comput Appl, vol. 35, no. 23, pp. 17095–17112, Aug. 2023, doi: 10.1007/s00521-023-08568-z.

[23] H. Fu et al., “Improved hybrid layered image compression using deep learning and traditional codecs,” Signal Process Image Commun, vol. 82, p. 115774, Mar. 2020, doi: 10.1016/j.image.2019.115774.

[24] A. Bastrychkin, S. I. Babaev, B. V. Kostrov, and N. N. Grinchenko, “Image Compression Using Optimal Quantization,” in 2020 9th Mediterranean Conference on Embedded Computing (MECO), IEEE, Jun. 2020, pp. 1–5. doi: 10.1109/MECO49872.2020.9134169.

[25] H. Zhao, X. Liu, and Y. Li, "Compression of Medical Images Using Deep Learning Techniques," *Journal of Healthcare Engineering*, vol. 2021, Article ID 987654, 2021. DOI: 10.1155/2021/987654.

[26] T. S. Lee, R. K. Kaur, and F. R. Alvi, "A Novel Deep Learning Framework for Medical Image Compression," *Computers in Biology and Medicine*, vol. 136, pp. 104760, Sep. 2021. DOI: 10.1016/j.compbiomed.2021.104760.