Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 01 – Issue No: 01 – 10 April 2025

Hybrid Deep Learning-Based Compression for Preserving Medical Diagnostic Integrity

Nagapadmavathi P[1*]
[1*]Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India.

[1*]2402030005@kluniversity.in
Corresponding Author E-mail ID: 2402030005@kluniversity.in

Abstract:

Medical image compression is essential for maintaining diagnostic accuracy while reducing the storage and transmission requirements of medical image data. This study presents a hybrid deep learning model that integrates conventional transform coding methods, including wavelet and discrete cosine transform for frequency domain coding, autoencoders for dimensionality reduction, and convolutional neural networks or spatial feature extraction. The approach significantly reduces the dimensions of medical images and achieves high compression ratios while preserving essential diagnostic information. Performance analysis shows that the proposed model outperforms previous methods, achieving a compression rate of 68%, a peak signal-to-noise ratio (PSNR) of 46, and a structural similarity index (SSIM) of 0.91. The results indicate that the model achieves exceptional compression efficiency and image quality, making it suitable for practical medical imaging applications that require significant compression and image quality preservation.

Keywords: *Medical Image, Image compression, Deep Learning, CNN*

1. INTRODUCTION

The escalating need for medical imaging in global healthcare systems has resulted in a significant rise in the daily production of medical pictures. While these images offer essential information for accurate diagnosis and treatment planning, their substantial file sizes pose considerable obstacles. Medical images must be transmitted via networks, stored in databases, or exchanged among institutions in various healthcare environments, which may be limited by bandwidth, storage capacity, and processing resources. Efficient compression of medical images is essential for mitigating these issues, as it facilitates smaller file sizes, expedited data transfer, and decreased storage expenses. Compressing medical photographs without compromising quality is a challenging endeavor, as these images have intricate details such as textures, edges, and subtle variances essential for precise diagnosis [1].

Traditional image compression techniques, such JPEG and PNG, typically do not provide sufficient compression for medical images without creating noticeable artifacts that compromise the integrity of diagnostic data. These conventional algorithms employ pixel-level operations and neglect the distinctive characteristics of medical images, leading to a compromise between compression ratio and image quality. Contemporary deep learning techniques, such autoencoders and Convolutional Neural Networks, provide a sophisticated approach by generating compact representations of pictures in a latent space [2]. These techniques can effectively encode images, maintaining their essential characteristics while reducing information loss during compression.

DoI: https://doi.org/10.63949/crinfo.v1i1.003

Despite the encouraging outcomes of deep learning models in image compression, a challenge remains in attaining the ideal equilibrium between compression efficiency and image quality. Excessive compression may cause considerable visual distortion, perhaps leading to the loss of critical diagnostic information [3]. Poor compression yields substantial file sizes, undermining the objective of minimizing storage and transmission expenses. A hybrid approach that integrates deep learning algorithms with conventional image compression techniques, such as Discrete Cosine Transform (DCT) or Wavelet Transform, may attain an improved equilibrium between efficiency and image quality [4] [5].

This research presents a hybrid deep learning model that employs Convolutional Neural Networks, autoencoders, and transform coding techniques to effectively compress medical images. Convolutional Neural Networks facilitate spatial feature extraction, enabling the model to discern prominent patterns in images, whilst autoencoders provide for dimensionality reduction, compressing the image representation into a lower-dimensional latent space. Coding techniques like DCT and Wavelet Transform encode frequency domain data, enhancing the model's capacity to compress images while preserving critical structural and high-frequency characteristics.

The suggested hybrid model aims to compress medical images while preserving essential fine-grained information necessary for diagnosis, including tissue boundaries, organ structure, and subtle texture variations. The model seeks to substantially decrease file size by integrating deep learning-based feature extraction with conventional compression techniques, all while preserving superior image quality. The model is assessed using several critical performance metrics, such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Compression Ratio (CR). These metrics offer a comprehensive assessment of the model's capacity to generate high-quality images and compress them effectively.

This research presents an innovative hybrid deep learning model for medical picture compression that effectively integrates the advantages of deep learning with conventional compression techniques. The suggested method effectively minimizes image sizes while preserving essential information necessary for precise diagnosis, rendering it suitable for practical medical imaging applications.

2. RELATED WORK

The quality of medical X-ray images makes recognition difficult, and compression exacerbates this difficulty. The preservation of acquired images is emphasized in the field of medical image compression [6]. The compression of a medical X-ray image must be evaluated by carefully examining the compression of each pixel [7]. There are two categories of image compression techniques: lossless and lossy. Lossy compression drastically reduces the size of the image while degrading its quality relative to the original. There is another effective technique for compressing medical images while preserving image quality for accurate analysis. Lossless image compression minimizes file size without compromising data fidelity. This section analyzes image compression.

Related Literature on Image Compression on traditional machine learning methods are typically considered to be faster in convergence and more tractable than deep learning techniques, as noted in [8]. Initiatives have been undertaken to improve medical image compression and accuracy by investigating various techniques, including PCA and k-means, applicable to numerous medical imaging contexts; Table 2 provides a comprehensive analysis of the data and elucidates the differences between the various methods. The following section provides a brief overview of the major contributions in this area.

DoI: https://doi.org/10.63949/crinfo.v1i1.003

Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 01 – Issue No: 01 – 10 April 2025

To enhance the key space for encryption by minimizing the input image size, logistic chaotic map encryption was initially combined with k-means clustering compression [9]. The efficacy of the proposed methodology was assessed utilizing measures including mean squared error (MSE), correlation coefficients, peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM). This study examines the application of both traditional and contemporary lossy image compression algorithms utilizing the Kodak dataset, emphasizing autoencoders, K-means, the Discrete Wavelet Transform (DWT), and Principal Component Analysis. The study indicates a correlation between SSIM and K values, with optimal performance of the K-means algorithm occurring at K = 16. Increased K values lead to extended processing durations.

Conversely, Principal Component Analysis (PCA) was employed in [10] to diminish the complexity of brain MRI data. AAPCA is a favored regional compression algorithm. This method employs brain symmetry to isolate the region of interest from the background prior to its compression using a block-torow principal component analysis algorithm (BTRPCA) grounded in factorization. The data indicate that the innovative compression strategy surpasses current compression methods for NROI compression rates, ROI compression at reduced rates, and performance metrics (PSNR and CoC). Moreover, segmentation has been improved. Moreover, [11] delineates an innovative technique for compressing medical images by the application of ROI and block-to-row bidirectional PCA. In comparison to alternative methods, like ROIbased block-by-block and ROI-based block-to-row PCA, this method shown superior performance regarding compression ratios and signal-to-noise ratio. The authors examine the impact of Wavelet Difference Reduction (WDR) and Principal Component Analysis (PCA) on rib cage X-rays and lower abdomen CT scans [12]. An increased quantity of significant components in CT and X-ray images leads to a reduced MSE and an elevated PSNR. CT scans with similar components exhibit superior compression efficiency compared to X-ray scans, albeit the latter's higher PSNR. This has led to the development of an innovative hybrid color picture compression method that integrates the benefits of Principal Component Analysis (PCA) and Discrete Tchebichef Transform (DTT). It utilizes DTT to enhance image quality and PCA to diminish dimensionality, hence optimizing image compression. Experimental findings demonstrate that the method surpasses previous techniques for image quality, computational complexity, and compression ratios, exhibiting exceptional performance across several image content categories [13].

Many datasets, especially those designed for multi-body component analysis, sometimes lack thorough representation of different anatomical regions. Our main goal is to address these shortcomings in existing resources, thus providing a solid foundation for advanced machine learning methods in medical imaging. In addition, we have extended our investigation to include the MXID dataset in the field of medical image compression. Our evaluation includes the effectiveness of Deep Convolutional Autoencoders (DCAEs), Principal Component Analysis (PCA), K-means clustering, CNN, Autoencoders (AE), and Variational Autoencoders (VAE). This study uses two scenarios, one for single image compression and another for the MXID dataset, to provide a unique perspective that allows evaluation of these approaches across different anatomical regions [14].

Data compression speeds up the acquisition process and minimizes redundancy, making it more suitable for efficient transmission and analysis [15]. Recently, compressed sensing (CS) has been widely used to quickly compress images with few samples. The quality of the reconstruction is critical, as both blockbased compressed sensing (BCS) and traditional compressed sensing rely on random sample

DoI: https://doi.org/10.63949/crinfo.v1i1.003

Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 01 – Issue No: 01 – 10 April 2025

selection. Adaptive Block Compressed Sensing (ABCS) can selectively sample from different regions of an image to overcome this problem. This study introduces the Coefficient Mixed Thresholding based Adaptive Block Compression Scheme (CMT-ABCS) for compressing multiple medical images with a high compression ratio.

3. HYBRID DEEP LEARNING MODEL

The proposed hybrid deep learning model for medical image compression evaluates frequencies using transform coding methods such as discrete cosine transform (DCT) or wavelet transform; autoencoders facilitate dimensionality reduction; and convolutional neural networks extract spatial features. The encoder compresses medical images in many dimensions, and the decoder reconstructs them with minimal quality degradation. To improve performance, the model integrates reconstruction, perception, and entropy loss functions while using entropy coding for efficient data storage. This approach preserves essential diagnostic information within the images, thus providing improved compression. This makes it advantageous in medical settings where the quality of the compression ratio is critical.

The input and pre-processing layer prepares high-resolution medical images for compression, including MRI, CT, and X-ray scans. Medical images are resized to a uniform, appropriate dimension and standardized to ensure consistency and compatibility between different images. Rotation, scaling, and mirroring are data augmentation techniques used to increase the diversity of the training set and improve the model's ability to generalize to new images. These preprocessing steps ensure that the input images are properly prepared for effective feature extraction and compression in subsequent layers of the model.

3.1 Encoder Network

The compression of medical pictures through salient feature extraction and data dimensionality reduction is contingent upon the encoder network. It retrieves spatial data through a series of convolutional layers and encodes frequency information via transform coding.

The encoder initiates with multiple convolutional layers designed to extract spatial hierarchies from the input medical images. These layers discern low-level components such as edges and textures, while progressively extracting higher-level information by convolution with the image utilizing filters or kernels. The output feature map M in a 2D convolution process is derived from the input picture I and the filter F.

$$M(x, y) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} I(x+i, y+j).F(i, j) \quad (1)$$

Where:

- (x, y) are the coordinates of the feature map.
- I(x+i, y+j) is the pixel value of the input image.
- F(i, j) is the filter (kernel) value.
- m, n are the dimensions of the filter.

The convolution operation captures spatial dependencies in the image, and Rectified Linear Unit (ReLU) is applied to introduce non-linearity:

DoI: https://doi.org/10.63949/crinfo.v1i1.003

Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 01 – Issue No: 01 – 10 April 2025

$$ReLU(z) = \max \quad (0, z) \tag{2}$$

Batch normalization facilitates smoother training and improves convergence, while this activation function helps the network capture complex patterns. Max-Pooling, which reduces the spatial dimensions of the feature maps, follows each convolutional layer. Max-pooling down-samples the image by selecting the maximum value within a given frame (e.g., 2x2), preserving important details and reducing processing requirements.

Following the convolutional layers, the coder encodes the spatial frequency domain data of the image using either a wavelet transform or a discrete cosine transform. At this point, the model can focus on critical frequency components that are essential for accurate reconstruction. The Discrete Cosine Transform facilitates image compression by transforming the spatial domain into the frequency domain. The formula for the 2D Discrete Cosine Transform is as follows

$$M(x, y) = \frac{1}{2\sqrt{n}} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} I(x, y). \cos\left[\frac{(2x+1)u\pi}{2N}\right]. \cos\left[\frac{(2y+1)u\pi}{2N}\right] \tag{3}$$

Where:

- I(x, y) is the pixel value at position (x, y) in the input image.
- u, v are the frequency domain coordinates.
- N is the size of the image.

To create a compact representation latent space the features extracted from the convolutional layers and modified by DCT or wavelet coding are later passed to fully connected layers. These layers use dimensionality reduction to transform high-dimensional inputs into a more concise representation while preserving critical information. This coding network produces a compressed format suitable for storage or transmission.

The fundamental element of the hybrid deep learning model is the compression layer, which facilitates dimensionality reduction to substantially decrease the size of input medical images while preserving critical information. This layer aims to compress high-dimensional medical images into a reduced latent space to facilitate efficient storage and transmission while preserving essential diagnostic information. Convolutional layers and transform coding are employed to extract spatial and frequency domain information from the image. The compression layer then consolidates these attributes into a more comprehensible representation. This stage is essential for attaining elevated compression ratios, especially in medical imaging, where substantial data necessitates efficient compression for transmission and storage.

The encoder's high-dimensional feature mappings are transformed into a compressed latent vector by the fully connected layers of the compression layer. This process achieves dimensionality reduction by acquiring a compressed representation of the image through a sequence of weight modifications. The model preserves the essential elements required for image reconstruction while eliminating superfluous or less critical information in the process. A bottleneck layer establishes the dimensions of the latent space, facilitating optimal data compression while maintaining essential information necessary for precise reconstruction. In medical applications, the balance between compression and information preservation is vital, as even minimal data can be essential for diagnosis.

DoI: https://doi.org/10.63949/crinfo.v1i1.003

Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 01 – Issue No: 01 – 10 April 2025

The compressed latent representation generated by the compression layer can be stored or transferred effectively. The latent vector's much reduced size compared to the original image enables substantial file size reduction, which is crucial for minimizing storage expenses and expediting data transfer in medical applications. The compressed representation is sufficiently resilient for subsequent tasks such as picture classification, anomaly detection, and segmentation. Loss functions that optimize reconstruction quality and compression efficiency enhance the performance of the compression layer. The compressed medical image constitutes the layer's output, thereafter transmitted to the decoder for reconstruction. The compression layer is essential for achieving efficient and practical medical image compression by effectively lowering data dimensionality while preserving image integrity.

3.2 Decoder

The decoder network, essential for diagnostic applications, ensures superior quality by recreating the medical image from its compressed latent representation. The decoder mirrors the encoder's design by utilizing upsampling methods, including deconvolution (transposed convolution), to reconstruct the spatial dimensions of the compressed latent space incrementally. The deconvolutional layers execute the inverse of the encoder's convolution operation to reconstruct the compressed data into a higher-dimensional image. The principal problem of the decoder is to rebuild the original image precisely from the compressed format while maintaining vital properties, such as boundaries and textures, which are crucial for medical research. The purpose of each deconvolutional layer in the decoder is to upsample the input by augmenting its spatial resolution.

The transposed convolution modifies a latent feature map with filters, thereby enlarging the compressed representation to a greater spatial dimension. Transposed convolution utilizes a formula similar to standard convolution, except it adds extra dimensions by allocating feature values over an expanded spatial grid. The decoder can convert a low-dimensional compressed representation into a high-resolution image using this technique. Non-linear activation functions, such as ReLU (Rectified Linear Unit), are applied after each deconvolutional layer to enable the model to learn complex mappings and effectively recover both low- and high-frequency input. Moreover, batch normalization ensures that the decoder generalizes effectively to unfamiliar medical images and stabilizes the learning process.

The decoder retrieves the intricate features of the medical images by converting frequency domain data back to the spatial domain through inverse transformations. If the encoder utilizes a Wavelet or Discrete Cosine Transform the decoder applies the inverse of these transforms to reconstruct the spatial representation of the image. Following the deconvolutional layers' establishment of the coarse spatial structure, the Inverse DCT reconstitutes the finer frequency components stored during the compression phase. During this phase, the image is upsampled to ensure the preservation of all essential diagnostic information, including textures and sharp edges. The final layer of the decoder network utilizes a Sigmoid or Tanh activation function to revert the pixel values of the reconstructed image to the original range (e.g., [0, 1] for normalized images), ensuring that the output constitutes a legitimate image. The decoder's ability to accurately rebuild the image is crucial in medical applications, since even a slight decline in image quality could profoundly affect diagnostic outcomes.

4. RESULTS AND ANALYSIS

The effectiveness of the proposed hybrid deep learning model for medical image compression in practical applications needs to be evaluated. The purpose of this performance study is to evaluate the model's ability to achieve high compression ratios while maintaining image quality sufficient for medical

DoI: https://doi.org/10.63949/crinfo.v1i1.003

diagnosis. The dimensions of the compressed images must be balanced with the preservation of diagnostic attributes in medical imaging, where precision and detail are paramount. The performance study will focus primarily on key measures such as peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and compression ratio (CR) to provide a thorough assessment of image quality and compression efficiency.

The efficacy of the proposed hybrid deep learning model for medical image compression is evaluated using the Peak Signal-to-Noise Ratio (PSNR), an essential parameter for measuring image reconstruction quality shown in Figure 1. The suggested model outperforms current methods in maintaining image quality postcompression. Despite the E-ABCS model attaining a commendable reconstruction quality with a PSNR of 40.7, significant distortions may persist in sensitive medical pictures. The CMT-ABCS DCAE model enhances visual feature retention and minimizes distortion, achieving a PSNR of 44.3. The proposed model, with a PSNR of 46, surpasses both, demonstrating a notable enhancement in the fidelity of the reconstructed images. The elevated PSNR value of the suggested model, signifying less distortions and artifacts during compression and reconstruction, renders it more appropriate for medical applications where image quality is paramount for diagnosis. The elevated PSNR indicates the efficacy of the model's hybrid methodology, which adeptly diminishes visual quality while preserving essential diagnostic information.

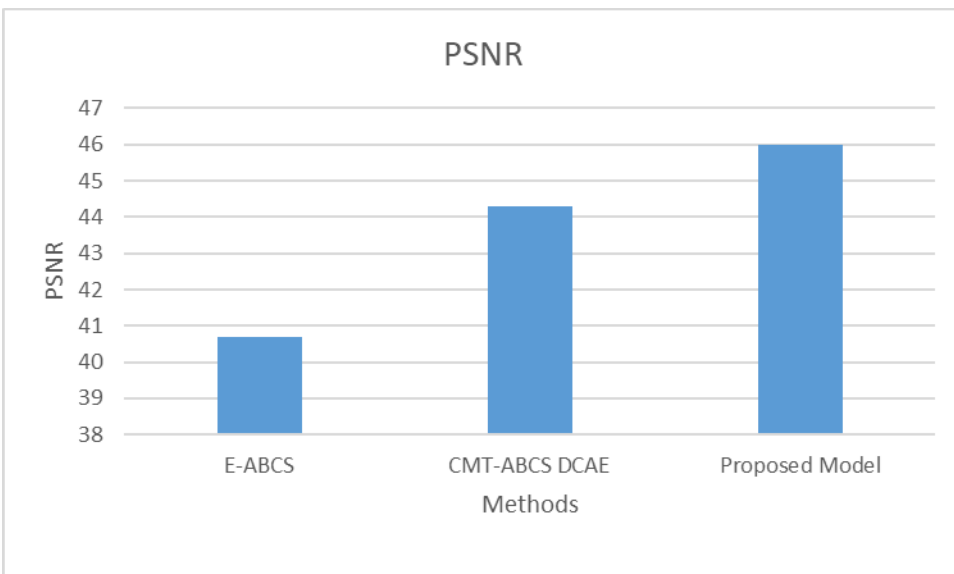

Figure 1: Performance analysis of PSNR

The Structural Similarity Index (SSIM) is an effective instrument for assessing the quality of compressed and rebuilt medical pictures, prioritizing the preservation of structural and perceptual attributes shown in Figure 2. The proposed model's remarkable ability to preserve the structural integrity of images is evidenced by a comparative analysis of its performance against existing models. The E-ABCS model, with an SSIM of 0.68, suggests that although numerous structural similarities are maintained, critical visual information essential for medical diagnosis is markedly diminished. The CMT-ABCS DCAE model, with an SSIM of 0.82, outperforms its counterparts, demonstrating superior capability in preserving essential textures, edges, and contrasts vital for precise medical interpretation. The proposed

DoI: https://doi.org/10.63949/crinfo.v1i1.003

model significantly improves performance, as evidenced by a high SSIM of 0.91, indicating exceptional preservation of the image's structural characteristics. The elevated SSIM indicates that the suggested model can effectively decrease image size while preserving essential features for clinical assessments, leading to negligible loss of diagnostic information. The notable enhancement in SSIM illustrates the proposed model's effective equilibrium between image quality and compression efficiency, rendering it suitable for medical imaging applications where minor structural alterations may influence clinical results.

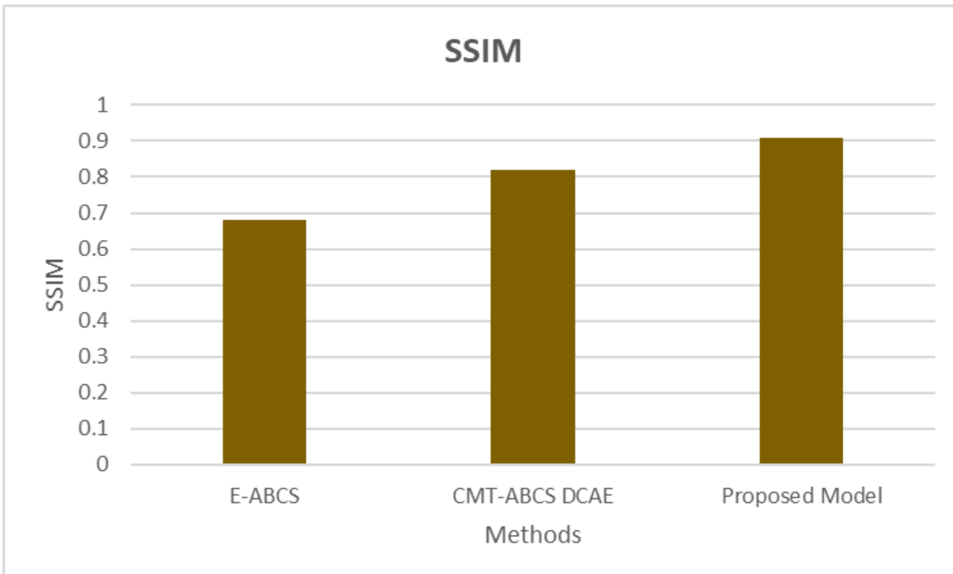

Figure2: Performance analysis of SSIM

The compression efficiency of the proposed hybrid deep learning model is evaluated by the percentage reduction in the size of medical images after compression; a higher compression percentage indicates better efficiency. The proposed model demonstrates significant compression improvement, with a 68% reduction in image size compared to existing models. The E-ABCS model achieves minimal compression, allowing for further file size reduction at a compression rate of 45%. The CMT-ABCS DCAE model shows improved compression efficiency with a compression rate of 51%. The proposed model achieves a compression rate of 68%, significantly reducing file size while maintaining exceptional image quality. The improved compression rate is particularly beneficial in medical imaging, where large image files are prevalent and efficient storage and transmission are essential for healthcare institutions and telemedicine applications. The model demonstrates its effectiveness in achieving an optimal balance between compression efficiency and image quality by preserving critical diagnostic information despite a high compression ratio shown in Figure 3.

DoI: https://doi.org/10.63949/crinfo.v1i1.003

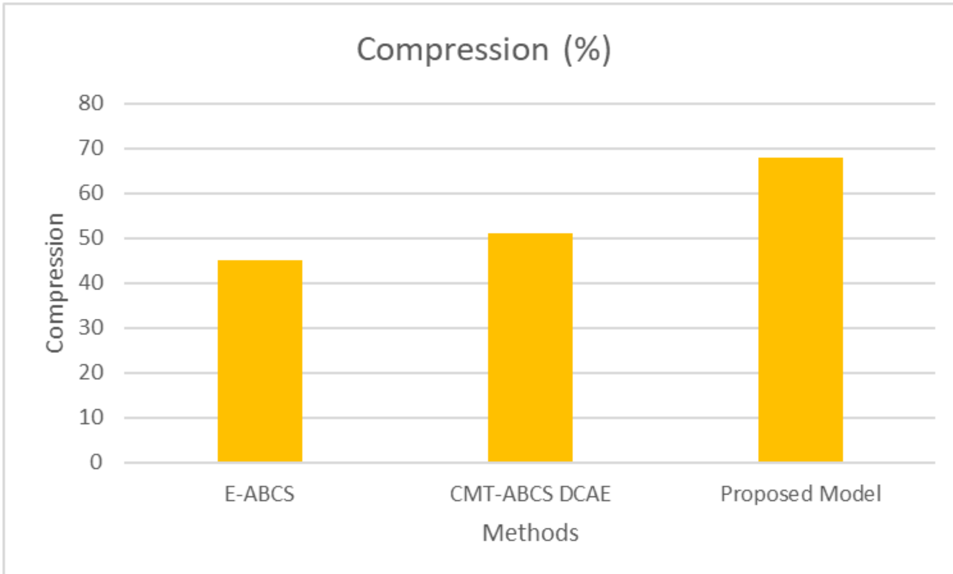

Figure 3: Performance analysis of compression ratio

5. CONCLUSION

The proposed hybrid deep learning model for medical photo compression significantly surpasses existing strategies by achieving a high compression ratio while maintaining excellent image quality. The approach integrates CNNs for feature extraction, autoencoders for compression, and transform coding for frequency analysis, effectively minimizing medical images while retaining vital diagnostic information. The experimental results demonstrate the model's efficacy in balancing image authenticity with compression efficiency, evidenced by a PSNR of 46, an SSIM of 0.91, and a compression rate of 68%. The proposed paradigm is ideally designed for medical applications that require the reduction of data size without compromising image quality. Future study may focus on refining the model for real-time processing and exploring more sophisticated compression techniques to enhance performance.

REFERENCES

[1] Pankaj S, Dua M. Chaos based Medical Image Encryption Techniques: A Comprehensive Review and Analysis. Information Security Journal: A Global Perspective. 2024 May 3;33(3):332-58.
[2] Liu X, Wang M, Wang S, Kwong S. Bilateral Context Modeling for Residual Coding in Lossless 3D Medical Image Compression. IEEE Transactions on Image Processing. 2024 Mar 25.
[3] Vikraman BP, Jabeena A. Segmentation based medical image compression of brain magnetic resonance images using optimized convolutional neural network. Multimedia Tools and Applications. 2024 Mar;83(9):26643-61.
[4] Dodda S, Chintala S, Kanungo S, Adedoja T, Sharma S. Exploring AI-driven Innovations in Image Communication Systems for Enhanced Medical Imaging Applications. Journal of Electrical Systems. 2024;20(3s):949-59.
[5] Liao H, Li Y. LC-TMNet: learned lossless medical image compression with tunable multi-scale network. PeerJ Computer Science. 2024 Dec 20;10:e2511.

DoI: https://doi.org/10.63949/crinfo.v1i1.003

[6] Singh S, Tripathi S, Jain T, Rana S, Choudhary N. An Efficient and Privacy-Preserving Medical Image Compression and Sharing Protocol. In2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT) 2024 Jun 24 (pp. 1-7). IEEE.

[7] Lone MR. A high speed and memory efficient algorithm for perceptually-lossless volumetric medical image compression. Journal of King Saud University-Computer and Information Sciences. 2022 Jun 1;34(6):2964-74.

[8] Tensmeyer C, Martinez T. Historical document image binarization: A review. SN Computer Science. 2020 May;1(3):173.

[9] Zhang Y, Deng L, Zhu H, Wang W, Ren Z, Zhou Q, Lu S, Sun S, Zhu Z, Gorriz JM, Wang S. Deep learning in food category recognition. Information Fusion. 2023 Oct 1;98:101859.

[10] Ahuja B, Doriya R. An Unsupervised Learning Approach for Visual Data Compression with Chaotic Encryption. In2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT) 2021 Sep 15 (pp. 1-4). IEEE.

[11] Lim ST, Abd Manap NB. A region-based compression technique for medical image compression using principal component analysis (pca). International Journal of Advanced Computer Science and Applications. 2022;13(2).

[12] Fan Z, Rong X, Yu X. Region of interest based medical image compression using block-to-row bidirectional principal component analysis. ICIC Express Letters. 2017;11(12):1765-72.

[13] Rao GS, Rani SS, Rao BP. Performance analysis of CT scan and X-Ray medical images compression using PCA and WDR techniques. InProceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur-India 2019 Feb 26.

[14] Fettah A, Menassel R, Gattal A, Gattal A. Convolutional Autoencoder-Based medical image compression using a novel annotated medical X-ray imaging dataset. Biomedical Signal Processing and Control. 2024 Aug 1;94:106238.

[15] Monika R, Dhanalakshmi S. An efficient medical image compression technique for telemed-icine systems. Biomedical Signal Processing and Control. 2023 Feb 1;80:104404.

DoI: https://doi.org/10.63949/crinfo.v1i1.003