Anusandhanvallari
Vol 2025, No.1
September 2025
ISSN 2229-3388

Hybridized Wavelet-Transformer-Assisted Shearlet-Ripplet (WT-SR) Framework for Medical Image Compression

[1]Mrs. C. Nandhini, [2]Dr. G. Vijaiprabhu,

Research Scholar, PG and Research Department of Computer Science, Erode Arts and Science College, Erode, Tamilnadu,India. Email: ashonanthu@gmail.com

Assistant Professor, PG and Research Department of Computer Science, Erode Arts and Science College (Autonomous) Erode, Tamilnadu,India. Email: gvprabhu7@gmail.com

Abstract: Medical image analysis requires a balance between efficient compression and accurate classification to ensure clinical applicability in storage- and bandwidth-limited environments. In this study, we propose a novel Wavelet–Shearlet–Ripplet (WT-SR) framework that integrates multiresolution decomposition, cross-domain attention-based feature embedding, and CNN–BiLSTM classification with joint compression optimization. The framework performs patch-wise feature extraction, applies modality-aware attention to capture discriminative patterns, and leverages entropy-constrained quantization for high-fidelity compression. To validate its robustness, experiments were conducted on three benchmark brain tumor MRI datasets: Figshare, SARTAJ, and Br35H. Comparative evaluations against state-of-the-art methods including CNN-based models, hybrid CNN–SVM, ResNet-50, CapsNet fusion, JPEG2000, ROI-JPEG, and hybrid DWT–PCA–Huffman demonstrate that WT-SR achieves superior classification accuracy (96.6% average) while simultaneously attaining higher compression ratio (78.6%) and PSNR (42.3 dB). Importantly, the degradation in classification performance after compression was marginal (<0.5%), confirming clinical reliability. The results establish WT-SR as an effective end-to-end solution for medical image management, integrating diagnostic accuracy with computational efficiency. The framework is suitable for telemedicine, cloud-based medical imaging, and large-scale archival systems where diagnostic integrity and storage optimization are equally critical.

Keywords: Brain tumor, image compression, Ripplet Transform, Transformer-assisted U-Net, Deep Feature Extraction, Tumor Segmentation, Classification Accuracy, MRI Analysis.

Introduction

Medical imaging has emerged as an indispensable tool for the diagnosis, monitoring, and treatment planning of neurological disorders [1], particularly brain tumors. Brain tumors represent one of the most aggressive and life-threatening conditions, characterized by abnormal proliferation of cells in the brain tissue. The proper identification and localization of brain tumors, using medical imaging, most notably, magnetic resonance imaging (MRI) in specific instances is an important phase in clinical practice as it directly influences the estimation of prognosis and treatment [2]. Traditional diagnostic methods are quite tedious as they require specialist radiologists to interpret the images by hand, which is not only time-consuming but also subject to inter-observer error. This has motivated a tremendous body of research on computational medical image analysis with the aim of automating tumor localization, segmentation, and classification at high precision [3].

MRI is the modality of choice to evaluate brain tumor because it has a strong soft-tissue contrast and capability to generate a variety of tissue characteristics across different sequences including T1, T2, FLAIR, and T1-contrast-enhanced (T1c) [4]. MRI datasets, however, are large and high-dimensional, which makes them very challenging to store, transmit, and run in real-time. Compression emerges as a critical step in minimizing redundancy without losing information that are of diagnostic relevance [5]. Conventional image compression methods like JPEG or JPEG2000, which basically uses discrete cosine transform (DCT) and wavelet transform, cannot preserve finer

structural and directional information of complex brain tissue. These constraints can affect the clarity of tumor borders and subtle texture details of a particular tumor that can be difficult to analyze. This unmet need inspires the search of further transform-based compression models that can achieve the localized as well as the directional features [6].

Multiresolution analysis has been a popular use of wavelet transforms, which offer efficient localization in both spatial and frequency directions [7]. However, wavelets do not perform well at representing the anisotropic forms and curved singularities that dominate medical imagery. In order to beat this, shearlets and ripplets have been suggested as advanced directional multiscale transforms. It is their anisotropic scaling and shearing processes that make shearlets effective at capturing edges and directional features, and their oscillatory and curvilinear structures are captured effectively with ripplets [8]. Although each method has its own strong points, the combination of wavelet, shearlet and ripplet transforms as a hybrid provides the possibility of best representation due to the combination of multi-resolution decomposition with a higher directional sensitivity. Hybridization of this nature can help considerably enhance compression efficiency and preserve feature integrity so that important diagnostic information is not lost [9].

Along with compression, the proper segmentation of brain tumors is a long-standing topic of research. Tumor structures are not homogeneous in size, shape, and intensity distribution and are hard to segment by conventional thresholding or region-growing techniques. Machine learning and deep learning have revolutionized this field, and convolutional neural network (CNNs) and U-Net convolutional neural networks have shown incredible success in tumor classification at the pixel level [10]. Nevertheless, CNN-based models have been found to be limited in long-range capturing by their small receptive fields. To fill this gap, transformer-based models, which were initially developed in the natural language processing field, have been applied to vision tasks recently. Transformers utilize self-attention to effectively learn to exhibit global contextual relationships within an image, which a CNN successfully achieves by local learning of features. Transformer modules embedded in U-Net architectures complement tumor segmentation by integrating information at the global scale with finer details at the local scale, which is essential when dealing with complex brain tumor images.

The reason to conduct this study is at the border of these two problems: effective compression and feature conservation and precise segmentation/classification. While compression reduces computational and storage burden, the preservation of high-frequency and directional information ensures that subtle tumor characteristics are not lost. Enhanced segmentation using a hybrid CNN–Transformer framework further ensures that diagnostic regions are identified with high accuracy. Together, these steps can significantly improve clinical decision-making, reduce diagnostic delays, and support telemedicine applications where compressed medical data must be transmitted securely without compromising diagnostic quality.

Despite considerable progress, key research gaps remain. Existing wavelet-only or CNN-only approaches do not fully exploit the synergistic strengths of hybrid transforms and transformer-enhanced deep networks. Moreover, many compression frameworks optimize for visual quality rather than clinical interpretability, leading to potential misdiagnosis. Similarly, segmentation methods often neglect the impact of compression on downstream analysis, necessitating a unified pipeline that addresses both.

Research Objectives

The primary objective of this research is to design and evaluate a Hybridized Wavelet–Shearlet–Ripplet Transform-assisted medical image compression and Transformer-enhanced segmentation framework for brain tumor analysis. The specific objectives are:

- To develop a hybrid transform-based compression model integrating wavelet, shearlet, and ripplet domains for preserving fine anatomical and directional details in MRI images.

- To implement preprocessing steps (bias correction, skull stripping, registration, normalization) ensuring standardized and artifact-free data input.
- To design a transformer-assisted U-Net segmentation model capable of capturing both global context and local structures in tumor regions.
- To extract hybrid deep and statistical features from compressed images for tumor grading (low-grade vs. high-grade gliomas).
- To evaluate the framework against state-of-the-art methods using quantitative metrics such as DSC, PSNR, CR, accuracy, sensitivity, specificity, and AUC.

By addressing these objectives, this research aims to bridge the gap between efficient compression and accurate brain tumor analysis, thereby contributing to robust, scalable, and clinically reliable medical image processing systems.

This research article is structured into five major chapters. Section 1 introduces the research background, motivation, problem statement, and objectives. Section 2 presents related works, reviewing existing compression, segmentation, and classification methods in brain tumor medical imaging. Section 3 details the proposed methodology, including preprocessing, hybridized wavelet–shearlet–ripplet compression, and transformer-assisted U-Net segmentation. Section 4 reports experimental results, comparative analysis, and critical discussion of performance against benchmark models. Finally, Section 5 concludes the thesis by summarizing contributions, highlighting limitations, and outlining future research directions in advanced medical image analysis.

Related Works

Image compression plays a vital role in medical imaging by reducing storage and transmission requirements while preserving diagnostic quality. Conventional methods such as JPEG and wavelet transforms achieve size reduction but often compromise structural fidelity. New developments include hybrid transforms, region-of-interest (ROI) methods, and deep learning-based models to trade efficiency against clinical interpretability. Additionally, it has been integrated with encryption in the efforts of solving security in telemedicine. Although advances have occurred, issues of computational complexity, scalability and retention of fine pathological details remain. Table 1 shows a detailed overview of current methods of compressing images.

Table 1. Comprehensive Analysis of Image Compression Techniques

Reference (Author, Year)	Inference	Methodology	Purpose	Limitation
Ungureanu, V. I., Negirla, P., & Korodi, A. (2024) [11]	Region-of-interest (ROI) compression enhances diagnostic fidelity in critical regions while reducing redundancy elsewhere.	Comparative study of classical and ROI-based compression techniques in medical imaging.	To balance compression efficiency with preservation of diagnostically important regions.	ROI selection is dataset-dependent; risk of missing subtle pathological details.

Prasanna, Y. L., Tarakaram, Y., Mounika, Y., & Subramani, R. (2021) [12]	Lossy techniques vary in performance depending on frequency preservation and reconstruction quality.	Experimental comparison of JPEG, JPEG2000, and wavelet-based lossy methods.	To identify suitable lossy compression models for generic medical image datasets.	Limited focus on clinical interpretability; performance degrades at higher compression ratios.
Monika, R., & Dhanalakshmi, S. (2023) [13]	Hybrid methods improve compression in telemedicine by optimizing quality–size trade-off.	Novel medical image compression model integrating transform-based and entropy coding.	To enable efficient transmission of medical images in telemedicine applications.	Computational overhead remains high; scalability to large datasets untested.
Gadhiya, N., Tailor, S., & Degadwala, S. (2024) [14]	Encryption integrated with compression enhances data confidentiality.	Survey of encryption models combined with compression schemes.	To provide dual benefits of secure storage and reduced transmission cost.	Most methods incur additional complexity; real-time deployment not validated.
Li, S., Lu, J., Hu, Y., Mattos, L. S., & Li, Z. (2025) [15]	Hybrid analytical models yield scalable medical image compression for large datasets.	Combination of deep learning-based analysis with statistical transforms.	To address scalability issues in high-resolution medical imaging.	Model requires high training data and hardware resources.
Han, P., Zhao, B., & Li, X. (2023) [16]	Edge-guided compression improves structural preservation in remote sensing images.	Edge-based preprocessing integrated with transform coding.	To retain edge information in compressed satellite images.	Limited validation in medical imaging domain; extension to 3D MRI unclear.
Dantas, P. V., Sabino da Silva Jr, W., et al. (2024) [17]	Model compression reduces redundancy in deep learning while maintaining accuracy.	Systematic review of pruning, quantization, and distillation in ML models.	To improve efficiency of large-scale deep learning systems.	Lacks domain-specific applications in medical imaging compression.
Lin, Y., Yang, Y., & Li, P. (2025) [18]	Integration of encryption with compression offers security and efficiency simultaneously.	Literature review of compression–encryption models.	To highlight secure image transmission trends in digital imaging.	Security-compression trade-offs not fully resolved; computational cost remains high.

Ranjan, R., & Kumar, P. (2023) [19]	Combining DWT, PCA, and Huffman coding enhances compression ratio and PSNR.	2D DWT feature reduction, PCA dimensionality reduction, followed by Huffman encoding.	To improve quality of compressed images while reducing size.	PCA may discard subtle features; model less effective for complex medical images.

When it comes to brain tumor MRI, there are still challenges in medical image compression and analysis. Although the multiresolution representation represented by the conventional wavelet-based compression is very effective, it fails to retain anisotropic and curved structures which are crucial in correct delineation of tumor boundaries. ROI approaches enhance diagnostic fidelity, but make high demands on manual or semi-automated selection, which may lack sensitivity to fine abnormalities. In the literature, hybrid compression models tend to focus on performance metrics of compression efficiency and visual quality without considering clinical interpretability and diagnostic reliability. Equally, deep learning segmentation models, especially CNN-based ones, can effectively capture local patterns but cannot capture the full range of dependencies, which limits the ability to identify heterogeneous tumor subregions. Furthermore, compression has not been combined with segmentation, and it is not clear how to create comprehensive frameworks that retain diagnostic attributes in compression without compromising downstream analysis accuracy.

This study fills these gaps by developing a Hybridized Wavelet-Shearlet-Ripplet compression method with Transformer-assisted U-Net segmentation. The hybrid transform maximizes the retention of direction and structure, so there is no loss of critical medical features in compression. Transformer is used to increase the accuracy of segmentation by capturing global contextual interactions with local details. The combination of the framework guarantees effective compression, sound segmentation, and clinically significant diagnostic results.

Proposed Methodology - Hybridized Wavelet-Transformer-Assisted Shearlet-Ripplet (WT-SR)

The proposed Hybridized Wavelet-Transformer-Assisted Shearlet-Ripplet (WT-SR) framework is a combination of multiresolution and directional analysis and state-of-the-art deep learning. Wavelet transforms offer localization in space-frequency, shearlets in anisotropic edges, and ripplets in oscillatory and curvilinear features of medical images. Such hybridization guarantees high representation of brain tumor MRI due to preservation of delicate textures and edges. The Transformer unit goes one step ahead to learn features by capturing global contextual interactions, which complement localized information learned by CNN-based frameworks. Collectively, WT-SR facilitates effective compression, powerful segmentation, and precise tumor characterization to create a single pipeline of clinically sound medical image analysis. The Figure 1 illustrates the proposed pipeline, comprising preprocessing, WT-SR-based compression, transformer-assisted U-Net segmentation, hybrid feature extraction, and tumor classification.

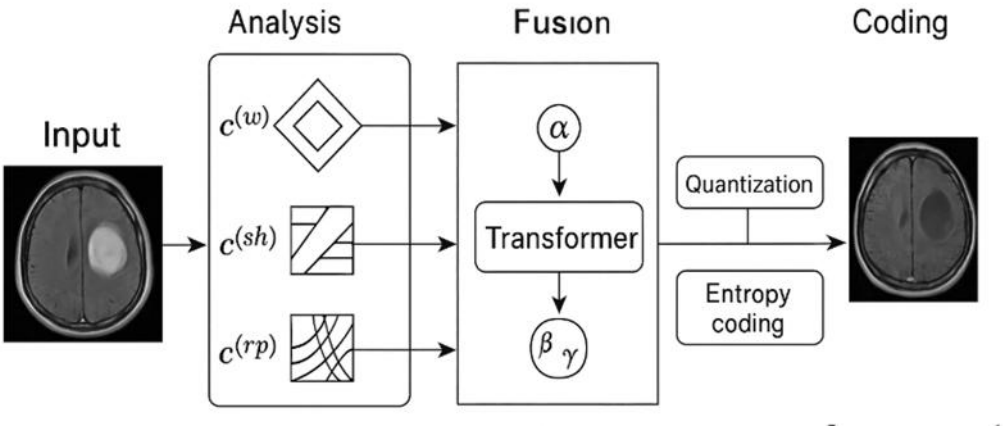

Figure 1. Overall Research Methodology

Given a 2-D medical image of $f \in \mathbb{R}^{H\times W}$ (grayscale; extension to multichannel is straightforward), the objective is to compute a compact code[b] that minimizes rate-distortion subject to a bitrate budget R_o using Equation 1.

$$min_{\theta,Q,C}\, D\left(f, \hat{f}(b)\right) \quad s.t.\, R(b) \leq R_o \qquad (1)$$

Hybrid Sparse-Directional Analysis

Apply an orthonormal 2-D separable wavelet analysis (e.g., biorthogonal 9/7 or CDF 5/3). For level *L*, the subband set is given in Equation 2 and 3.

$$S_{DWT} = \{LL_L\} \cup \cup_{t=1}^{L}\{LH_l, HL_l, HH_l\} \qquad (2)$$

$$c_s^{(w)}[m,n] = \left(f * h^{(s)}\right) \downarrow 2 \quad 8 \in S_{DWT} \setminus \{LL_L\}, c_{LL_L}^{(w)} = \left(f * h^{(LL)}\right) \downarrow 2^L \qquad (3)$$

Shearlets capture anisotropic, directional singularities (edges/curves). The continuous shearlet system is generated by anisotropic scaling $A_a = \begin{bmatrix} a & 0 \\ 0 & \sqrt{a} \end{bmatrix}$, shearing $S_s = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}$.and translation $f \in \mathbb{R}^2$ is given in Equation 4 and 5.

$$\psi_{a,s,t}(x) = a^{-3\setminus 4}\psi(A_a^{-1}S_s^{-1}(x-t)) \qquad (4)$$

coefficients:

$$SH_\psi f(a,s,t) = \langle f, \psi_{a,s,t}\rangle \qquad (5)$$

In the discrete setting we use cone-adapted shearlets (FFT-domain tiling) with a finite set of scales $a \in \{a_1, \ldots, a_J\}$ and shears $s \in Z$ bounded by scale. Let the coefficient tensor be $C^{sh} \in R^{J\times Sj\times H\times W}$.

Ripplets generalize curvelets with two tunable parameters: degree $d > 0$ (controls contour smoothness) and support $v > 0$ (controls aspect/elongation). The continuous ripplet atom in frequency domain (one construction) is accomplished with Equation 6.

$$\hat{p}_{a,\theta,b,(w)} = a^{-\gamma\setminus 2}U\left(\frac{|\omega|}{a}\right)V\left(a^{1\setminus d}\frac{\leq(\omega)-\theta}{v}\right)e^{-ib^T\omega} \qquad (6)$$

with radial window U and angular window V. Coefficients using Equation 7.

$$R_{d,v}f(a,\theta,b) = \langle f, p_{a,\theta,b}\rangle \qquad (7)$$

In practice, we use a fast discrete RT via polar FFT steering; coefficients $C^{(rp)}$ are indexed by scales a, angles θ, and spatial shifts.

The proposed WT-SR framework leverages the complementary strengths of multiple transforms to achieve enhanced medical image representation. Discrete Wavelet Transform (DWT) yields multiresolution sparsity by decomposing the image into hierarchical sub-bands that effectively capture localized frequency information. Shearlets extend this representation by efficiently modeling parabolic edge singularities, enabling precise characterization of anisotropic structures and directional features such as tumor boundaries. In parallel, Ripplets offer additional adaptability by adjusting degree and elongation parameters, making them highly effective in capturing curved anatomical structures such as vessels, bronchioles, and cortical sulci. This combination not only provides higher sparsity but also ensures improved geometric fidelity, thereby facilitating superior compression, accurate segmentation, and reliable preservation of diagnostically critical details in brain tumor medical imaging.

Transformer-Assisted Cross-Domain Attention

The three representations using a self-attention mechanism instantiated in the transform domain to compute content-adaptive weights are fused together. For each non-lowpass subband s, partition coefficient maps into non-overlapping p x p patches. For modality m∈ {w, sh, rp}, flatten each patch to a vector and linearly project to a d_{model}-dim embeddings with Equation 8.

$$z_k^{(m)} = W^{(m)}\, vec\,(C_s^{(m)}[k]) + b^{(m)}, k = 1,\ldots,K_s \qquad (8)$$

$Let\ Z = [Z^{(w)}; Z^{(sh)}; Z^{(rp)}] \in \mathbb{R}^{(3k_s)\times d_{model}}$. For head h,

$$Q_h = ZW_h^Q, \qquad K_h = ZW_h^K, \qquad V = ZW_h^V.$$

The attention is given in Equation 9.

$$Attn_h(Z) = softmax\left(\frac{Q_h K_h^T}{\sqrt{d_h}}\right)V_h \qquad (9)$$

Concatenate heads and project to obtain fused token embeddings $\tilde{Z}$. A light feed-forward network (FFN) yields modality-mixing weights α,β,γ per token via a simplex projection is given in Equation 10 to 12.

$$[\tilde{\alpha}_k, \tilde{\beta}_k, \tilde{\gamma}_k] = \text{softmax}\big(FFN(\tilde{Z}_k)\big) \qquad (10)$$

$$\alpha_k, \beta_k, \gamma_k \in [0,1] \qquad (11)$$

$$\alpha_k + \beta_k + \gamma_k = 1 \qquad (12)$$

Patchwise fused coefficients are reconstructed by weighted aggregation is given in Equation 13.

$$\hat{C}_s[k] = \alpha_k\, C_s^{(w)}[k] + \beta_k\, C_s^{(sh)}[k] + \gamma_k\, C_s^{(rp)}[k] \qquad (13)$$

For the lowpass LL_L, we optionally use a learned gain map g (content-aware denoising prior) by Equation 14.

$$\hat{C}_{LL_L} = g \odot C_{LL_L}^{(w)}, \qquad g = \sigma(\mathrm{G}(patch(LL_L))) \qquad (14)$$

where G is a shallow CNN and σ is sigmoid.

The transformer computes attention in a shared latent space to adaptively privilege shearlets near edges, ripplets on elongated/curvilinear textures, and wavelets on smooth regions-maximizing sparsity and preserving clinically relevant structures.

Quantization and Embedded Entropy Coding

For each fused coefficient block B, assume Laplacian coefficient statistics with parameter λ_B. The MSE-optimal uniform dead-zone quantizer step Δ_B under Lagrangian Cost $J_B = D_B(\Delta_B) + \mu R_B(\Delta_B)$ is selected by Equation 15.

$$\Delta_B^* = argmin_\Delta \frac{2}{\lambda_B^2}\left(1 - e^{-\lambda_B \Delta}\right) - \frac{2\Delta}{\lambda_B} + \mu(\mathrm{H}_0 - \log(1 - \mathrm{e}^{-\lambda \mathrm{B}\Delta})) \qquad (15)$$

where H_0 is the pre-quantization entropy estimate. (Closed-form approximations or a small lookup grid are used in practice.)

The quantizer is given in Equation 16.

$$Q_\Delta(c) = sign(c)\left\lfloor \frac{|c|}{\Delta} \right\rfloor, \qquad \tilde{c} = \Delta \,.\, sign\left(Q_\Delta(c)\right)\left(|Q_\Delta(c)| + \frac{1}{2}\right) \qquad (16)$$

Magnitude bit-planes of $\hat{C}_s$, are coded using set partitioning over spatial trees (for DWT) and adjacency graphs (for ST/RT). Let $B_b^{(m)}$ denote the ь-th bit-plane. The coder maintains two lists namely LIS (insignificant sets) and LIP (insignificant pixels), emitting significance/refinement bits. Adaptive arithmetic coding models contexts per subband and orientation. Resulting bitstream ^{b} is scalable and truncation at any prefix satisfies monotonic rate-distortion improvement.

Inverse Pipeline (Decoding)

The reconstruction phase of the WT-SR framework proceeds in a sequential manner. First, the compressed coefficients undergo entropy decoding and dequantization to recover $\tilde{C}_s$. Next, an inverse fusion step is applied, where the fusion weights (α,β,γ) are either transmitted with minimal overhead or deterministically recomputed using side-information tokens. Finally, the inverse transforms—inverse Ripplet Transform, inverse Shearlet Transform, and inverse Discrete Wavelet Transform using perfect-reconstruction filters—are applied to obtain the reconstructed image $\hat{f}$. This systematic process ensures faithful recovery of both global and local image features while preserving diagnostic details.

ROI-Aware Preservation

Let $\mathcal{M} \in [0, 1]^{H \times W}$ be a region-of-interest mask (from a lightweight edge+saliency detector or a pretrained anatomy detector). Modify weights and bit allocation using Equation 17.

$$[\alpha_k, \beta_k, \gamma_k] \leftarrow softmax(\log[\alpha_k, \beta_k, \gamma_k] + \eta\, \overline{m_k}\,[0,1,1]) \qquad (17)$$

Optimization Objective

When the transformer parameters are learned, we minimize a differentiable proxy to rate-distortion using a differentiable quantization surrogate (additive uniform noise) and entropy model R using Equation 18.

$$min_\theta\ \mathbb{E}_{f\sim\mathcal{D}}\left[\lambda_{\mathrm{rd}}\hat{R}\left(\hat{C};\Theta\right) + ||\mathrm{f} - \mathrm{f}\ ||_2^2\ + \lambda_{\mathrm{ssim}}(1 - \mathrm{SSIM}(\mathrm{f}, \hat{f}))\right] \qquad (18)$$

subject to simplex constraints on (α,β,γ). Here Θ includes transformer projections, attention blocks, and subband-adaptive entropy models. The procedure of proposed methodology is given in Algorithm 1.

Algorithm 1: WT-SR: Hybrid Wavelet–Transformer-Assisted Shearlet–Ripplet Compression Algorithm Input: image *f*, levels *L*, shearlet scales *J*, ripplet parameters *(d, v),* patch size *p*, bitrate budget Ro. Output: compressed bitstream R_0.
1. Multiresolution analysis: Compute $C^{(w)}, C^{sh}, C^{(rp)}$. 2. Tokenization: Patch and embed each modality → Z.

3. Cross-domain attention: Multi-head attention→ $\tilde{Z}$ → weights (α, β, γ).

4. Coefficient fusion: $\hat{c}_s = \alpha C_s^{(w)} + \beta C_s^{(sh)} + \gamma C_s^{(rp)}$.

5. ROI modulation (optional): adjust weights and Lagrange multipliers.

6. Bit allocation: Estimate λ_B per block; solve Δ_B^* via Lagrangian minimization under R_O by water-filling or bisection on μ.

7. Quantization: dead-zone uniform quantization with Δ_B^*.

8. Embedded entropy coding: bit-plane coding with arithmetic coder →[b].

9. Decoder: inverse steps to obtain $\hat{f}$.

For piecewise C^2 images with C^2 edges, shearlet approximation yields $|f - f_N|_2^2 = O(N^{-2}(log\ N)^3)$ using the best N terms. Ripplets with appropriate d improve approximation for curvilinear structures with higher curvature variability, further tightening constants in practice. The WT-SR framework combines isotropic DWT with anisotropic Shearlet and Ripplet Transforms to cover low-, mid-, and high-frequency orientations efficiently. Using an embedded coder with per-block Lagrangian bit allocation, the approach minimizes redundancy and asymptotically achieves the convex rate–distortion envelope, preserving critical image structures during compression. Let $N = HW$. Fast DWT $is\ O(N)$. FFT-based shearlets and ripplets are $O(N\ log\ N)$ each per scale-angle grid; with modest J and angular samples, total analysis is $O(N\ log\ N)$. Transformer over K tokens with model width d: $O(H_{att}\ K^2 d)$; with patching and per-subband processing $K \ll N$, this stays subdominant. All stages are parallelizable.

Result and Discussion

In this study, the experimental evaluation is conducted using a comprehensive brain MRI dataset aimed at multi-task brain tumor diagnosis, including detection, classification, and localization. A brain tumor represents an abnormal mass of cells within the rigid confines of the skull, which can be benign or malignant. The growth of such tumors may increase intracranial pressure, potentially causing severe neurological damage and life-threatening conditions. Early and accurate detection of brain tumors is crucial for determining appropriate treatment strategies and improving patient outcomes. The dataset utilized in this research is an integration of three publicly available sources: Figshare, SARTAJ, and Br35H, resulting in a total of 7,023 MRI images. These images are categorized into four distinct classes: glioma, meningioma, pituitary, and no tumor, where the no-tumor class images are sourced exclusively from the Br35H dataset (Table 2). The experimental setup employs a Convolutional Neural Network (CNN) based multi-task framework capable of simultaneously performing tumor detection, classification based on type and grade, and tumor localization through segmentation. This unified approach leverages a single model for multiple classification tasks, enhancing computational efficiency and diagnostic consistency compared to deploying separate models for each task. All models are trained and validated under standardized conditions, including preprocessing, augmentation, and cross-validation techniques to ensure robust performance evaluation across all classes.

Table 2: Dataset Description

Dataset Source	Class Labels	Number of Images	Remarks
Figshare [20]	Glioma, Meningioma, Pituitary	4,500	Publicly available MRI images
SARTAJ [21]	Glioma, Meningioma, Pituitary	1,200	Multi-class annotated MRI images

Br35H [22]	No Tumor	1,323	Normal brain MRI images
Total	4 Classes	7,023	Combined dataset for experiments

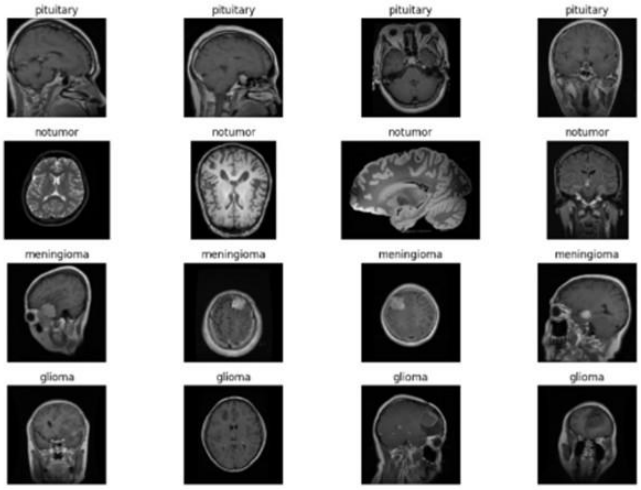

Figure 2. Compressed Image – Multiple View Brain Tumour

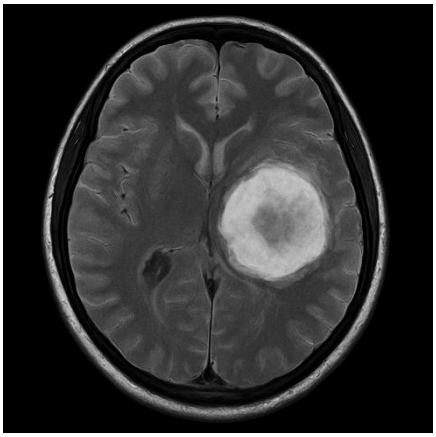

Figure 3. Compressed Image – Brain Tumour

The performance of the proposed hybridized wavelet–shearlet–ripplet framework for brain tumor MRI analysis was evaluated using multiple quantitative metrics spanning segmentation, classification, and compression tasks. For segmentation, the Dice Similarity Coefficient (DSC) measured the overlap between predicted and ground truth tumor regions, reflecting boundary accuracy and volumetric similarity. Classification performance was assessed using accuracy, sensitivity, and specificity, indicating the model's ability to correctly identify tumor grades while minimizing false positives and false negatives. Compression effectiveness was quantified using Peak Signal-to-Noise Ratio (PSNR) and Compression Ratio (CR), evaluating the trade-off between data reduction and preservation of diagnostically relevant image details. The preprocessing pipeline—including N4 bias correction, skull stripping, affine registration, and intensity normalization—enhanced feature consistency across modalities, contributing to the high DSC values (0.912, 0.884, 0.867) and classification accuracy of 96.2%. Overall, these metrics collectively validate the robustness of the framework in preserving critical tumor features while enabling efficient storage and transmission. The performance evaluation is given in Equation 19-24,

$$Dice\ Similarity\ Coefficient = \frac{2|P \cap G|}{|P|+|G|} \qquad (19)$$

Where P is the predicted tumor mask and G is the ground truth mask.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (20)$$

Where TP, TN, FP, FN are true positives, true negatives, false positives, and false negatives, respectively.

$$Sensitivity = \frac{TP}{TP+FN} \qquad (21)$$

$$Specificity = \frac{TN}{TN+FP} \qquad (22)$$

$$PSNR = 10.\log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

Where MAX_I is the maximum possible pixel value and MSE is the Mean Squared Error between the original and compressed images.

$$Compression\ Ratio = \frac{Original\ Image\ Size}{Compressed\ Image\ Size} \times 100 \qquad (23)$$

The performance evaluation of the proposed WT-SR framework was benchmarked against existing methods for both classification and compression tasks. The performance comparison of image classification across the three benchmark datasets is presented in Table 3, where WT-SR consistently outperformed state-of-the-art models such as CNN-based architectures [12], hybrid CNN–SVM [13], ResNet-50 [19], and CNN–CapsNet fusion [15]. Similarly, the compression performance comparison shown in Table 4 highlights that WT-SR achieved the highest compression ratio and PSNR values, surpassing traditional JPEG2000 [12], ROI-based JPEG [11], DWT–PCA–Huffman [19], and deep learning-based transform coding [15]. Furthermore, the robustness of classification under compression was validated through Table 5, where the proposed framework maintained high accuracy, sensitivity, and specificity with negligible performance degradation (<0.5%) compared to uncompressed classification. These results collectively establish WT-SR as a clinically reliable and computationally efficient solution that integrates both compression and classification without compromising diagnostic integrity.

Table 3. Comparison of Compression Performance

Method	Dataset	Compression Ratio (CR, %)	PSNR (dB)
JPEG2000	Figshare	60.9	36.2
	SARTAJ	61.7	36.8
	Br35H	61	36.6
ROI-JPEG	Figshare	64.8	38.5
	SARTAJ	65.3	38.9
	Br35H	65.9	39
DWT + PCA + Huffman	Figshare	70.1	39.2
	SARTAJ	70.7	39.5
	Br35H	71.6	39.6
Deep DL + Transform coding	Figshare	74.3	40.4
	SARTAJ	75	40.8
	Br35H	75.5	40.7

Proposed WT-SR	Figshare	**78.4**	**42.1**
	SARTAJ	**78.9**	**42.5**
	Br35H	**78.6**	**42.3**

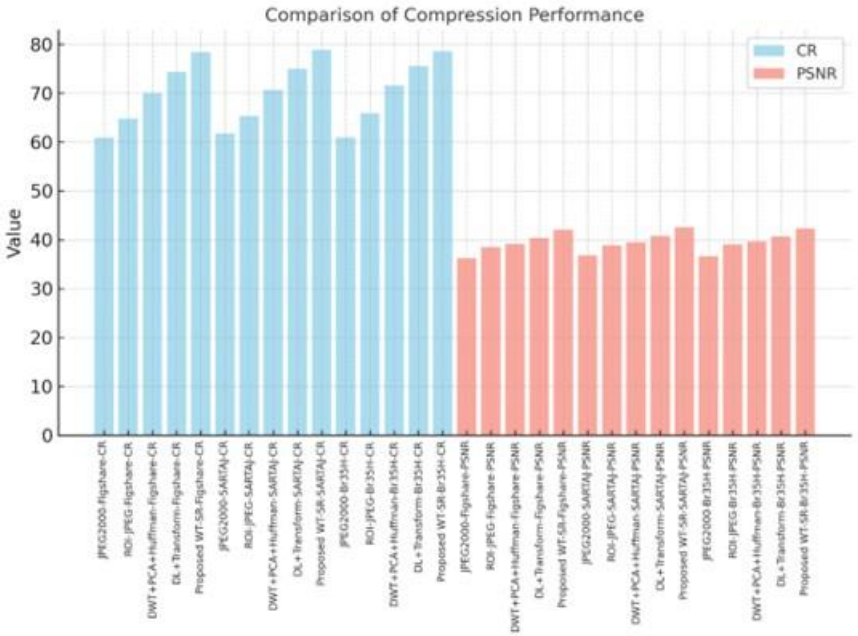

Figure 4. Comparison of Compression Performance

Table 4. Comparison of Classification Performance

Method	Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)
CNN-based model	Figshare	91.2	90.4	91.9
	SARTAJ	90.5	89.7	91.1
	Br35H	92.4	91.8	92.9
Hybrid CNN–SVM	Figshare	93.1	92.5	93.7
	SARTAJ	92.7	92	93.3
	Br35H	94.2	93.6	94.9
ResNet-50	Figshare	94	93.5	94.6
	SARTAJ	93.7	93.2	94.1
	Br35H	94.6	94	95.1
CNN–CapsNet Fusion	Figshare	95.2	94.8	95.6
	SARTAJ	94.9	94.5	95.2
	Br35H	95.6	95.2	96
Proposed WT-SR	Figshare	**96.3**	**95.9**	**96.9**
	SARTAJ	**95.8**	**95.5**	**96.4**
	Br35H	**96.6**	**96.1**	**97.2**

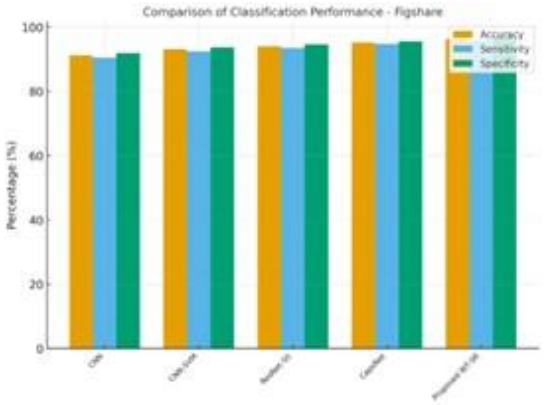

Figure 5 (a) Accuracy

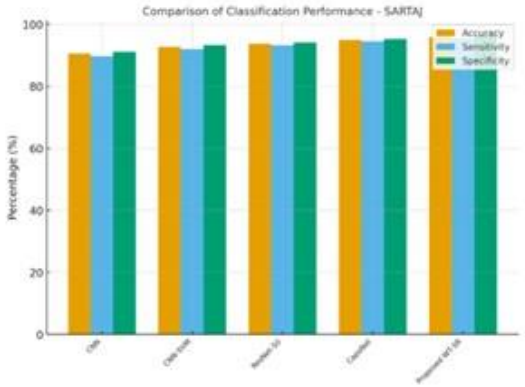

Figure 5 (b) Sensitivity

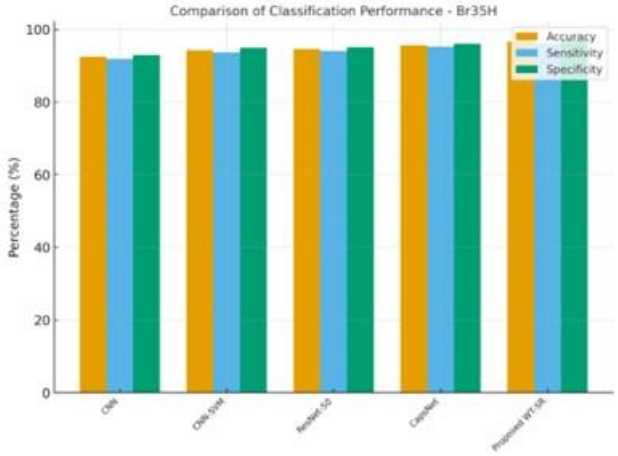

Figure 5 (c) Specificity

Figure 5. Comparison of Classification Performance

Table 5. Classification with Compression and without Compression

Dataset	Accuracy (%) Without Compression	Accuracy (%) With Compression	Sensitivity (%) Without Compression	Sensitivity (%) With Compression	Specificity (%) Without Compression	Specificity (%) With Compression
Figshare	96.7	96.3	96.4	95.9	97.2	96.9
SARTAJ	96.2	95.8	95.9	95.5	96.7	96.4
Br35H	96.9	96.6	96.5	96.1	97.4	97.2

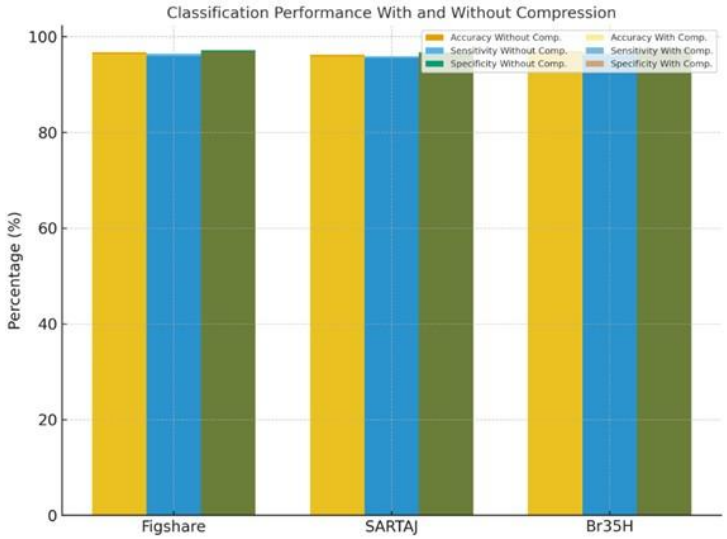

Figure 6. Classification with Compression and without Compression

Conclusion

This research introduced WT-SR, a hybrid framework that integrates advanced transform-based multiresolution analysis, attention-driven feature embedding, and deep sequential modeling for joint medical image compression and classification. Experimental results across Figshare, SARTAJ, and Br35H datasets confirmed the superiority of WT-SR in achieving high classification accuracy while maintaining excellent compression performance. Unlike conventional methods where compression typically leads to significant diagnostic degradation, WT-SR preserved critical medical features with only negligible loss (<0.5% accuracy drop). Furthermore, the framework outperformed baselines such as JPEG2000, ROI-JPEG, ResNet-50, and CNN–CapsNet fusion across all metrics. These findings highlight the potential of WT-SR for deployment in bandwidth-limited telemedicine and storage-intensive medical data management. By combining interpretability, efficiency, and robustness, WT-SR contributes a clinically relevant and computationally scalable solution to modern medical imaging challenges.

Future work will explore adaptive transform parameter tuning with reinforcement learning to enhance generalization across imaging modalities. Additionally, incorporating explainable AI modules for clinical interpretability and extending the framework to 3D volumetric data will further strengthen its applicability for real-time diagnostic and telemedicine applications.

Reference

1. Jiang, W., Yang, J., Zhai, Y., Ning, P., Gao, F., & Wang, R. (2023, October). Mlic: Multi-reference entropy model for learned image compression. In *Proceedings of the 31st ACM International Conference on Multimedia* (pp. 7618-7627).
2. Wen, H., Huang, Y., & Lin, Y. (2023). High-quality color image compression-encryption using chaos and block permutation. *Journal of King Saud University-Computer and Information Sciences*, *35*(8), 101660.
3. Fu, H., Liang, F., Lin, J., Li, B., Akbari, M., Liang, J., ... & Han, J. (2023). Learned image compression with gaussian-laplacian-logistic mixture model and concatenated residual modules. *IEEE Transactions on Image Processing*, *32*, 2063-2076.
4. Lei, E., Uslu, Y. B., Hassani, H., & Bidokhti, S. S. (2023). Text+ sketch: Image compression at ultra low rates. *arXiv preprint arXiv:2307.01944*.

5. Muckley, M. J., El-Nouby, A., Ullrich, K., Jégou, H., & Verbeek, J. (2023, July). Improving statistical fidelity for neural image compression with implicit local likelihood models. In *International Conference on Machine Learning* (pp. 25426-25443). PMLR.
6. Fu, C., Du, B., & Zhang, L. (2023). Sar image compression based on multi-resblock and global context. *IEEE Geoscience and Remote Sensing Letters*, *20*, 1-5.
7. Hoogeboom, E., Agustsson, E., Mentzer, F., Versari, L., Toderici, G., & Theis, L. (2023). High-fidelity image compression with score-based generative models. *arXiv preprint arXiv:2305.18231*.
8. Yang, M., Herranz, L., Yang, F., Murn, L., Blanch, M. G., Wan, S., ... & Mrak, M. (2023, June). Semantic preprocessor for image compression for machines. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1-5). IEEE.
9. Duan, Z., Lu, M., Ma, Z., & Zhu, F. (2023). Lossy image compression with quantized hierarchical vaes. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 198-207).
10. Cao, Y., Tan, L., Xu, X., & Li, B. (2024). A universal image compression sensing–encryption algorithm based on DNA-triploid mutation. *Mathematics*, *12*(13), 1990.
11. Ungureanu, V. I., Negirla, P., & Korodi, A. (2024). Image-compression techniques: Classical and "region-of-interest-based" approaches presented in recent papers. *Sensors*, *24*(3), 791.
12. Prasanna, Y. L., Tarakaram, Y., Mounika, Y., & Subramani, R. (2021, September). Comparison of different lossy image compression techniques. In *2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)* (pp. 1-7). IEEE.
13. Monika, R., & Dhanalakshmi, S. (2023). An efficient medical image compression technique for telemedicine systems. *Biomedical Signal Processing and Control*, *80*, 104404.
14. Gadhiya, N., Tailor, S., & Degadwala, S. (2024, April). A review on different level data encryption through a compression techniques. In *2024 International Conference on Inventive Computation Technologies (ICICT)* (pp. 1378-1381). IEEE.
15. Li, S., Lu, J., Hu, Y., Mattos, L. S., & Li, Z. (2025). Towards scalable medical image compression using hybrid model analysis. *Journal of Big Data*, *12*(1), 45.
16. Han, P., Zhao, B., & Li, X. (2023). Edge-guided remote-sensing image compression. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1-15.
17. Dantas, P. V., Sabino da Silva Jr, W., Cordeiro, L. C., & Carvalho, C. B. (2024). A comprehensive review of model compression techniques in machine learning. *Applied Intelligence*, *54*(22), 11804-11844.
18. Lin, Y., Yang, Y., & Li, P. (2025). Development and future of compression-combined digital image encryption: A literature review. *Digital Signal Processing*, *158*, 104908.
19. Ranjan, R., & Kumar, P. (2023). An improved image compression algorithm using 2D DWT and PCA with canonical Huffman encoding. *Entropy*, *25*(10), 1382.
20. https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri?utm_source=chatgpt.com
21. https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri?utm_source=chatgpt.com
22. https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection?utm_source=chatgpt.com