Implementation of a VGG-19 and Discrete Wavelet Transform Combined Multimodal Fusion Technique

Hayath T M
Department of CSE, Ballari Institute of Technology and Management, Ballari. Visvesvaraya Technological University, Belagavi-590018, India
hayathmail@gmail.com (corresponding author)

Sai Madhavi D
Department of AIML, RaoBahadur Y Mahabaleshwarappa Engineering College, Ballari. Visvesvaraya Technological University, Belagavi-590018, India
saimadhavi@rymec.in

Received: 30 April 2025 | Revised: 11 May 2025 and 26 May 2025 | Accepted: 31 May 2025

 DOI: https://doi.org/10.48084/etasr.11834

ABSTRACT

The integration of diverse medical imaging modalities facilitates the identification of diseases. Medical imaging is a critical component of medical research and diagnosis, providing detailed information about the structure and function of the body. In some cases, imaging approaches that utilize a single modality may not capture the complete set of the diagnostic data necessary for reliable physician evaluations. The objective of this study is to enhance the clarity of medical imagery and facilitate more precise disease identification. The proposed approach involves a multimodal medical image fusion technique that integrates Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) data. The suggested technique involves three sequential steps: image registration, image merging, and image segmentation. Image registration is a process that aligns CT and MRI images by utilizing procedures that are based on landmarks to ensure that pixel-level correlation is maintained. To preserve both structural and functional characteristics from the input pictures, the fusion procedure makes use of deep learning-based transfer learning in conjunction with the VGG-19 network and Discrete Wavelet Transform (DWT). Lastly, the watershed algorithm is employed to extract and highlight Regions of Interest (ROIs), such as tumors, during the segmentation process. The suggested method substantially increases picture clarity, maintains essential characteristics, and boosts the precision of tumor segmentation, as demonstrated by the results of the experiments.

Keywords-medical image fusion; Computed Tomography (CT); Magnetic Resonance Imaging (MRI); disease detection; image registartion; VGG-19; Discrete Wavelet Transform (DWT); segmentation algorithms

I. INTRODUCTION

The utilization of advanced technologies, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), has become instrumental in the field of medical imaging, playing a pivotal role in the diagnosis and treatment of various medical conditions. Both MRI and CT imaging are essential for clinical decision-making because they offer superior soft tissue contrast and high-resolution anatomical structures, respectively. Nevertheless, radiologists frequently find that manually examining each modality is time-consuming and subject to observer variability, which could result in inconsistent diagnoses. Researchers are increasingly using medical image fusion, a method that integrates data from several imaging modalities into a single, comprehensive image, to get around these restrictions [1-4].

To improve the final image's quality and informativeness, image fusion attempts to maintain essential structural and functional elements. However, especially in intricate clinical scenarios, conventional image fusion techniques like Principal Component Analysis (PCA) and wavelet-based approaches usually fail to preserve high-level features and texture information [5, 6].

A wide range of approaches to enhance medical image fusion and segmentation are presented in recent literature. Techniques such as Pulse Couple Neural Networks (PCNN)-based optimization [7], nonlinear anisotropic filtering in the PCA domain [8], and integrated guided nonlinear anisotropic filtering [9] have all shown improvements in quality metrics like Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR). However, these approaches frequently encounter difficulties with computational complexity, generalizability, and the absence of established clinical validation procedures [5, 10, 11].

The potential of multimodal medical image fusion has been further enhanced by sophisticated deep learning techniques. Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), and multi-scale fusion frameworks are examples of deep models that have demonstrated promise in overcoming the drawbacks of conventional techniques [12-18]. In a related study, authors in [18] suggested a GAN-based framework incorporating the shift-invariant shearlet transform for enhanced texture retention. In another study, authors in [13] introduced the Target Information Enhanced Fusion Network (TIEF), which enhanced glioma diagnosis through multimodal input processing.

Through techniques such as Multi-Generator Multi-Discriminator GANs (MGMDcGAN) [17], attention-driven models, and fusion guided by tissue segmentation [19], other works have incorporated unsupervised deep learning into fusion pipelines. Although these models are frequently criticized for their complexity and reliance on sizable, labeled datasets, they demonstrate high accuracy in both qualitative and quantitative assessments.

Limitations still exist despite advancements. Among the difficulties are:

- Misalignment at registration [20].
- Nonlinear transformations result in data loss [6].
- Domain-specific models have problems with generalization [12, 13].
- High requirements for computational resources [7, 8, 21].

Real-time feasibility has been the subject of numerous studies, with solutions like the Adolescent Identity Search Algorithm (AISA) for computational efficiency [22] and Discrete Wavelet Transform (DWT)-based techniques for ease of use and versatility in clinical settings [23] that have been proposed. However, the field still struggles to strike a balance between interpretability, real-world deployability, and model complexity. Furthermore, more extensive research investigates uses outside of diagnostics, such as multimodal fusion for the detection of fake news [24], hand gesture recognition systems [25], and material degradation analysis through image processing [26]. These multidisciplinary applications highlight the value of modality-specific optimization while also reaffirming the adaptability of image fusion and segmentation techniques.

In summary, although multimodal medical image fusion has advanced considerably, gaining from both conventional mathematical models and state-of-the-art deep learning methods, significant drawbacks like computational overhead, gaps in standardization, and difficulties with interpretability still exist.

II. PROPOSED METHODOLOGY

This study offers a novel multimodal medical image fusion method that utilizes wavelet decomposition and a VGG-19-based deep learning network to combine CT and MRI images for improved visualization. Four frequency sub-bands are created from the input images, and VGG-19 fuses related sub-bands to preserve deep structural and functional characteristics. The Inverse Wavelet Transform (IWT) is then used to reconstruct the combined sub-bands, thereby generating a final image that combines soft tissue contrast of MRI with structural details of CT. Figure 1 illustrates the proposed image fusion framework using VGG-19 and DWT. This approach enhances image quality, maximizes feature selection, and enables more precise diagnosis and treatment planning.

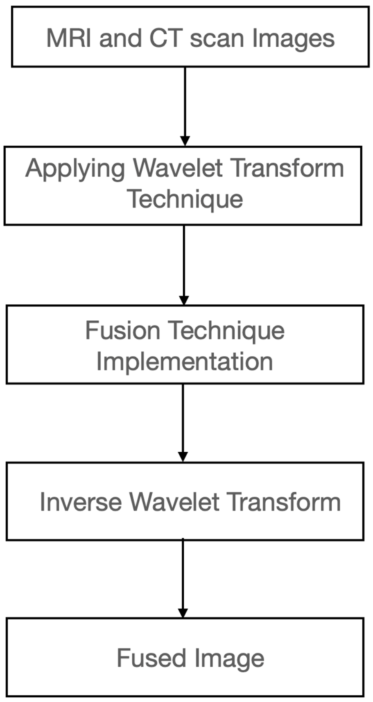

Fig. 1. Block diagram of the proposed image fusion framework using VGG-19 and DWT.

A. *Image Registration*

In the domain of multimodal medical image fusion, image registration is an essential preprocessing step that ensures images from multiple modalities are aligned within a common spatial framework. This is accomplished by linking the images together. The procedure facilitates the alignment of pixels in CT and MRI scans, thereby enabling the reflection of the same anatomical structure. Figure 2 illustrates the sample CT and MRI images that are to be registered. In this study, the image registration process was executed through the implementation of a landmark-based registration method.

1) *Landmark-Based Registration*

The landmark-based registration technique is employed to recognize the specific anatomical landmarks present in CT and MRI Images. In order to accomplish spatial alignment, one of the images, which is normally the CT scan, is identified as the fixed reference, and the other image, which is the MRI scan, is modified according to the transformation. This method guarantees that the structures that correspond to each other align precisely, which results in an improvement in the quality of the fused pictures. Figure 3 shows the landmarks on CT and MRI images. The primary objective of this method is to align CT and MRI images, with the corresponding pixels representing analogous anatomical structures. To implement this, it is first necessary to preprocess the images. This preprocessing entails converting the images into grayscale and normalizing the intensity values to enhance contrast. Subsequently, landmark selection is provided for both CT and MRI images. The third procedure is the transformation calculation. That is, it computes the transformation parameters using a method called the least-squares optimization algorithm. The final step in the process is image warping, which is applied to the computed transformation to align the MRI images with the CT images. The details of these steps are also discussed below.

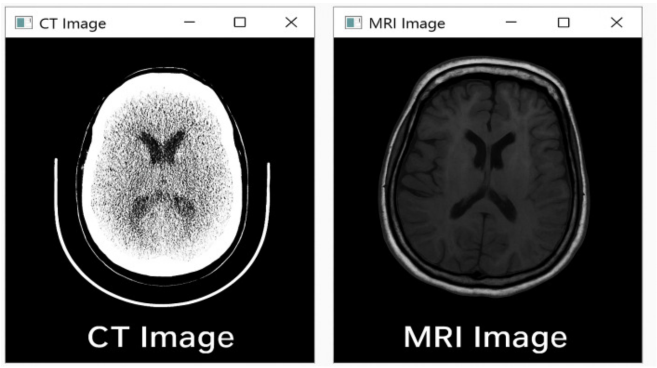

Fig. 2. Sample CT and MRI images.

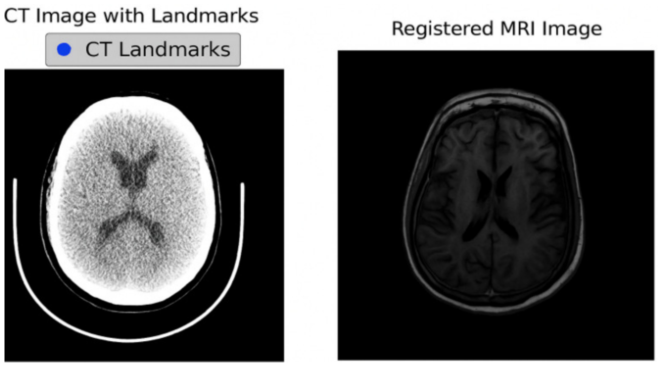

Fig. 3. Registered landmarks on CT and MRI images.

2) *Implementation Steps*

First, choose appropriate anatomical landmarks in both CT and MRI scans, emphasizing obvious characteristics such as organ edges, tissue connections, or bone structures. Subsequently, apply an affine model that incorporates translation, rotation, scaling, and shearing to calculate the transformation matrix. The transformation is then applied to align the MRI image with the CT scan, and the registered image is saved for subsequent fusion processes.

B. Image Fusion

To perform the image fusion, the authors employed a deep learning approach, which enabled enhanced picture quality and the extraction of structural and functional details separately. Specifically, the study utilized the pre-trained VGG-19 model in conjunction with the DWT to preserve both fine and high-level visual features.

1) *Learning Transfer Methods Using VGG-19*

VGG-19 is a deep CNN that has been pre-trained on the ImageNet dataset. The advantage of the VGG-19 is its ability to extract high-level characteristics from the image dataset, which leads to the capture of all the details related to the fusion. The VGG-19 method ensures that the fused images maintain the fundamental structures of both the CT and MRI inputs.

2) *Integrating Images Using the Wavelet Transform*

DWT is a technique decomposes an image into its frequency components. This technique enables the selective fusing of low-frequency and high-frequency features. During the decomposition, four sub-bands are produced:

- Low-Low (LL) – approximation coefficients: They provide coarse structural information.
- Low-High (LH) – horizontal detail coefficients: They record information related to the horizontal edges.
- Low-Vertical (LV) – vertical detail coefficients: They represent vertical edges and borders.
- Low-Diagonal (LD) – diagonal detail coefficients: They store information about the diagonal edges and textures.

3) *Fusion Process for Images*

Using the wavelet decomposition technique, the sub-bands (LL, LH, LV, and LD) are to be decomposed from both registered CT and MRI images. The fusion algorithm VGG-19 involves the combination of related sub-bands by utilizing depth characteristics that are retrieved. A feature fusion process is carried out on the LL sub-band, whereas the high-frequency bands (LH, LV, and LD) are fused by the utilization of maximum intensity selection. The reconstruction of the fused picture is achieved through the application of an IWT on the fused coefficients. This approach facilitates the reconstruction of a fused image of superior quality. The main objective of this fusion process is to preserve the structural and functional details by integrating both CT and MRI images. The fusion technique can be summarized in the following steps:

- Convert images to the YCbCr color space: This step ensures that luminance (Y) details are preserved.
- Apply DWT: Decompose images into four frequency sub-bands (LL, LH, LV, LD). The formula for the wavelet transform is shown in (1).

$$V_{j,k} = \int I(x)\varphi_{j,k}(x)dx \quad (1)$$

where $V_{j,k}$ represents the wavelet coefficients at scale j and position k, and $\varphi_{j,k}(x)$ is the wavelet function.

- Extract deep features using VGG-19: The formula for the feature extraction using VGG-19 is shown in (2), which is mainly used to extract deep features from LL sub-bands.

$$J_i = f(V \cdot I + b) \quad (2)$$

where V represents the learned filter weights, b represents the bias, and function f represents the activation function.

- Fusion of sub-bands: For the LL sub-band, fusion is performed using (3) and for the LH sub-band, fusion is performed using (4).

$$Fusion_{LL} = \alpha LL_{CT} + (1-\alpha)LL_{MRI} \quad (3)$$

$$Fusion_{LH} = \max(LH_{CT}, LH_{MRI}) \quad (4)$$

- Reconstruction: Equation (5) is used to reconstruct the fused image from the coefficients.

$$I_{fusion} = \sum_j \sum_k V_{j,k}\, \varphi_{j,k}(x) \quad (5)$$

- Post-processing: Adjust the contrast and remove noise using Gaussian filtering.

C. Image Segmentation

Following the fusion of the CT and MRI images, the segmentation process is carried out to extract Regions of Interest (ROIs), which may include cancers. The segmentation procedure enables the precise localization of diseased features, thereby enhancing the diagnostic usefulness of fused pictures. Within the scope of this investigation, the watershed algorithm, which is a very effective region-based segmentation approach, is utilized to efficiently outline tumor borders. Figure 4 illustrates the sequential process of image segmentation.

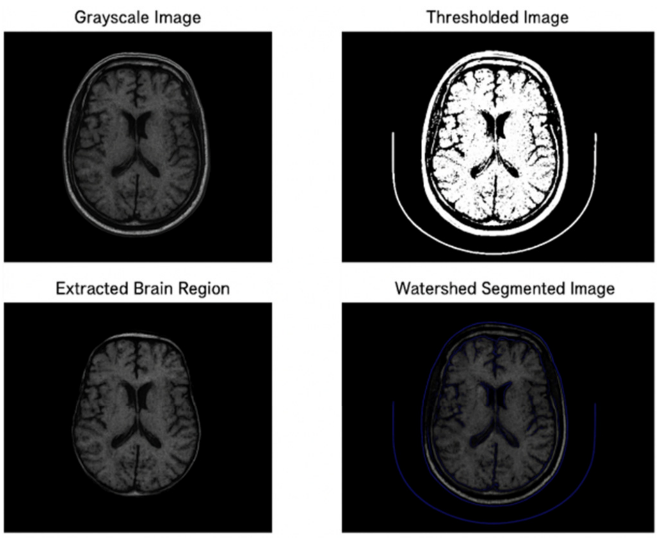

Fig. 4. Stepwise watershed segmentation of the image.

1) Steps of the Watershed Algorithm

The segmentation process is initiated with the selection of seed points for tumor and background regions, which are determined using manual or automatic thresholding. Morphological changes, such as dilation and erosion, , serve to define object boundaries and differentiate overlapping structures. The watershed transformation then segments tumors depending on intensity gradients by treating the greyscale image as a topographic surface. The complete process includes the following key steps: preprocessing (greyscale conversion and histogram equalization), marker selection (using adaptive thresholding), morphological refinement, watershed segmentation, and post-processing with a Gaussian filter to remove noise and improve the segmentation mask.

D. Algorithm for Multimodal Medical Image Fusion Using Wavelet Transform and VGG-19

The proposed multimodal fusion algorithm integrates the VGG-19 model and the DWT to fuse images from both CT and MRI modalities. The proposed hybrid method is expected to be an effective solution for enhancing the feature extraction process. The objective is to extract and merge features from both modalities, ensuring improved structural and soft tissue visibility. The algorithm is comprised of the following steps:

- Step 1: Input acquisition: Read the CT and MRI images, it denoted as I_{CT} and I_{MRI}.
- Step 2: Apply DWT to the images: The formula for the wavelet transform is shown in (6).

$$D(I) = \{LL, LH, LV, LD\} \quad (6)$$

where LL, LH, LV, LD are the previously described coefficients.

Equations (7) and (8) show the decomposition of the CT and MRI images.

$$LL_{CT}, LH_{CT}, LV_{CT}, LD_{CT} = D(I_{CT}) \quad (7)$$

$$LL_{MRI}, LH_{MRI}, LV_{MRI}, LD_{MRI} = D(I_{MRI}) \quad (8)$$

- Step 3: Feature-based fusion using VGG-19: VGG-19 is employed to extract the deep features of the images. The process for feature extraction is shown in (9).

$$\begin{aligned} F_{LL} &= VGG_{19}(LL_{CT}, LL_{MRI}) \\ F_{LH} &= VGG_{19}(LH_{CT}, LH_{MRI}) \\ F_{LV} &= VGG_{19}(LV_{CT}, LV_{MRI}) \\ F_{LD} &= VGG_{19}(LD_{CT}, LD_{MRI}) \end{aligned} \quad (9)$$

Equation (10) shows the merging of the feature maps using the fusion rules.

$$\begin{aligned} Fusion_{LL} &= \alpha F_{LL_{CT}} + (1-\alpha) F_{LL_{MRI}} \\ Fusion_{LH} &= max\left(F_{LH_{CT}}, F_{LH_{MRI}}\right) \\ Fusion_{LV} &= max\left(F_{LV_{CT}}, F_{LV_{MRI}}\right) \\ Fusion_{LD} &= max\left(F_{LD_{CT}}, F_{LD_{MRI}}\right) \end{aligned} \quad (10)$$

where α is the weight factor to balance contributions from both images. The max is applied for the high frequency components.

- Step 4: IWT: In step IWT is applied to reconstruct the final fused images, using (11).

$$I_{Fusion} = IWT(Fusion_{LL}, Fusion_{LH}, Fusion_{LV}, Fusion_{LD}) \quad (11)$$

- Step 5: Post-processing: Apply contrast enhancement and noise reduction techniques, and finally, display the fused images.

III. RESULTS AND DISCUSSION

A. *Experimental Setup*

The objective of the present study was to create a framework for the fusion and segmentation of medical images, with a focus on glioma segmentation. The Python 3.8, TensorFlow 2.8, OpenCV 4.5, and SciPy 1.7 libraries were utilized to construct the framework. The system was evaluated using two publicly available datasets: the Brain Tumor Segmentation Challenge (BraTS) 2021 dataset [27] and the Harvard Whole Brain Atlas [28]. The BraTS dataset includes preoperative multimodal MRI scans (T1, T1Gd, T2, FLAIR) that have been labeled for glioma segmentation. The Medical Image Computing and Computer Assisted Intervention (MICCAI) society provides the BraTS dataset. The Harvard Whole Brain Atlas dataset comprises MRI and CT images of the human brain with detailed structural and pathological notes. The results demonstrated the system's capacity to combine medical images from different sources while keeping high-frequency structural information, soft tissue contrast, and the ability to accurately segment ROIs. This facilitated the identification and diagnosis of tumors. To assess the efficacy of the fusion and segmentation processes, we employed performance metrics such as SSIM, PSNR, and Dice Similarity Coefficient (DSC). The method demonstrated consistent high SSIM and PSNR scores on both datasets, indicating its efficacy in preserving anatomical integrity and image clarity.

B. *Performance Metrics*

Conventional quantitative measures were used to confirm the efficacy of the suggested multimodal fusion approach:

- SSIM is a metric that assesses perceptual variations between pictures by considering structure, contrast, and brightness. A higher SSIM near 1 indicates a stronger resemblance to the reference image.
- PSNR is a metric expressed in dB that measures the quality of the combined image compared to the original. Higher PSNR suggests less distortion.
- DSC is a metric used to assess the precision of image segmentation. It measures the overlap between the predicted tumor area and the ground truth mask. A higher DSC shows better segmentation accuracy.

Table I presents a comparative analysis of the proposed algorithm against other fusion methods using these metrics.

TABLE I. COMPARISON OF THE PROPOSED METHOD WITH EXISTING METHODS

Method	SSIM (CT)	SSIM (MRI)	PSNR (CT)	DSC
VGG-19 + DWT (proposed)	0.9631	0.8691	41.42 dB	0.1484
PCA-based fusion	0.8912	0.7950	37.80 dB	0.1221
CNN-GAN approach [18]	0.9450	0.8652	39.50 dB	0.1303
ADDNS [29]	0.9524	0.8563	40.12 dB	0.1409

The proposed method (VGG-19 + DWT) obtained the highest SSIM scores for both CT (0.9631) and MRI (0.8691), indicating its effectiveness in preserving structural integrity. Furthermore, it exhibited the highest PSNR value of 41.42 dB, indicative of superior image quality due to reduced distortion and noise. The DSC value of 0.1484 indicates the method's improved segmentation performance, outperforming all other compared techniques.

The findings indicate that the integration of deep feature extraction (VGG-19) with frequency-domain fusion (DWT) is a highly effective strategy for preserving critical medical information, generating images with high fidelity, and supporting accurate segmentation.

Furthermore, Figure 5 illustrates the segmentation and fusion process for CT and MRI images of a patient. Figure 6 provides a graphical representation of SSIM and PSNR for the proposed method. The SSIM for CT is 0.9631, and the SSIM for MRI is 0.8691. The PSNR for CT is 41.42 dB, and the PSNR for MRI is 38.08 dB. These results demonstrate the superior performance of the proposed fusion approach.

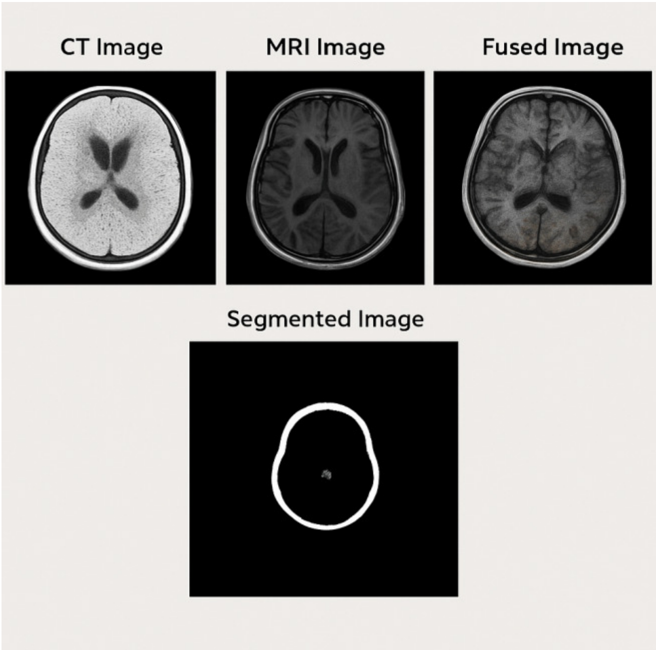

Fig. 5. Process of fusing and segmenting CT and MRI images for a patient.

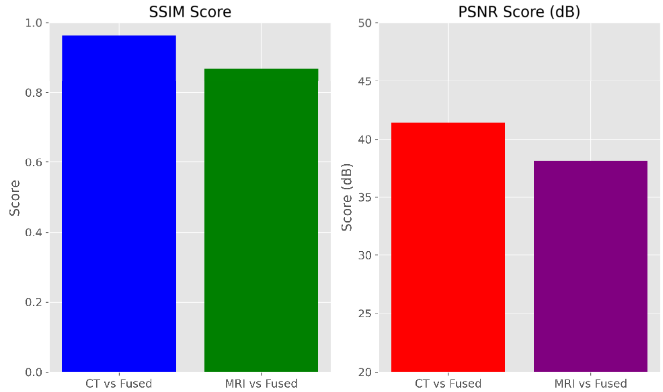

Fig. 6. Performance evaluation using SSIM and PSNR metrics on fused images, relative to the original CT and MRI scans for a patient.

IV. CONCLUSION

This paper presents a hybrid medical image fusion framework that utilizes VGG-19 for the extraction of deep features and Discrete Wavelet Transform (DWT) for frequency domain decomposition, resulting in enhanced fused Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) images for improved diagnosis. The method is designed to preserve both low-frequency components and high-frequency structures, thereby ensuring the accuracy of medical interpretation. Moreover, the proposed methodology maintains semantic depth and spatial detail in the resulting fused images by leveraging the capacity of VGG-19 to represent features in a hierarchical manner and the capability of DWT to decompose images at multiple resolutions. The method was evaluated on two datasets: the Brain Tumor Segmentation Challenge (BraTS) 2021, which focuses on brain tumor segmentation, and the Harvard Whole Brain Atlas, which provides a structural overview of the brain. Future work includes real-time implementation, incorporation of additional imaging modalities such as ultrasound and Positron Emission Tomography (PET), and exploration of applications in multi-region segmentation tasks and cross-modality learning.

REFERENCES

[1] M. A. Saleh, A. A. Ali, K. Ahmed, and A. M. Sarhan, "A Brief Analysis of Multimodal Medical Image Fusion Techniques," *Electronics*, vol. 12, no. 1, Jan. 2023, Art. no. 97, https://doi.org/10.3390/electronics12010097.

[2] H. Kaur, D. Koundal, and V. Kadyan, "Image Fusion Techniques: A Survey," *Archives of Computational Methods in Engineering*, vol. 28, no. 7, pp. 4425–4447, Dec. 2021, https://doi.org/10.1007/s11831-021-09540-7.

[3] P.-H. Dinh, "Medical image fusion based on enhanced three-layer image decomposition and Chameleon swarm algorithm," *Biomedical Signal Processing and Control*, vol. 84, Jul. 2023, Art. no. 104740, https://doi.org/10.1016/j.bspc.2023.104740.

[4] M. Haribabu, V. Guruviah, and P. Yogarajah, "Recent Advancements in Multimodal Medical Image Fusion Techniques for Better Diagnosis: An Overview," *Current Medical Imaging*, vol. 19, no. 7, pp. 673–694, Sep. 2022, https://doi.org/10.2174/1573405618666220606161137.

[5] R. Maurya, R. Swarnkar, Y. K. Sharma, R. Mishra, and D. M. Pathak, "Medical Image Fusion." Research Square, Jun. 06, 2024, https://doi.org/10.21203/rs.3.rs-4456478/v1.

[6] S. Chinnadurai and D. Vasanthi, "Enhancement of Medical Image by Fusion Method using Fast Discrete Curvelet Transform," *International Journal for Research in Applied Science and Engineering Technology*, vol. 7, no. 5, pp. 344–350, May 2019, https://doi.org/10.22214/ijraset.2019.5057.

[7] P. Kavita, D. R. Alli, and A. B. Rao, "Study of image fusion optimization techniques for medical applications," *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 136–143, Jun. 2022, https://doi.org/10.1016/j.ijcce.2022.05.002.

[8] S. R. K. Reddy, V. Swathi, and K. Anusha, "MR and CT Image Fusion Using Nonlinear Anisotropic Filtering in PCA Domain," *Journal of Physics: Conference Series*, vol. 1964, no. 6, Jul. 2021, Art. no. 062058, https://doi.org/10.1088/1742-6596/1964/6/062058.

[9] M. G. Reddy, P. V. N. Reddy, and P. R. Reddy, "Medical Image Fusion Using Integrated Guided Nonlinear Anisotropic Filtering with Image Statistics," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 1, pp. 25–34, Feb. 2020, https://doi.org/10.22266/ijies2020.0229.03.

[10] F. Shabanzade and H. Ghassemian, "Combination of wavelet and contourlet transforms for PET and MRI image fusion," in *2017 Artificial Intelligence and Signal Processing Conference*, Shiraz, Iran, 2017, pp. 178–183, https://doi.org/10.1109/AISP.2017.8324077.

[11] A. Kesharwani, K. Singh, and A. Saxena, "Advancements In Multi-Modality Medical Image Fusion: A Comprehensive Review," *International Journal of Innovative Research in Engineering and Management*, vol. 11, no. 2, pp. 55–61, Apr. 2024, https://doi.org/10.55524/ijirem.2024.11.2.11.

[12] T. Zhou, Q. Cheng, H. Lu, Q. Li, X. Zhang, and S. Qiu, "Deep learning methods for medical image fusion: A review," *Computers in Biology and Medicine*, vol. 160, Jun. 2023, Art. no. 106959, https://doi.org/10.1016/j.compbiomed.2023.106959.

[13] Y. Zhou, X. Yang, S. Liu, and J. Yin, "Multimodal Medical Image Fusion Network Based on Target Information Enhancement," *IEEE Access*, vol. 12, pp. 70851–70869, 2024, https://doi.org/10.1109/ACCESS.2024.3402965.

[14] W. Xu, Y.-L. Fu, H. Xu, and K. K. L. Wong, "Medical image fusion using enhanced cross-visual cortex model based on artificial selection and impulse-coupled neural network," *Computer Methods and Programs in Biomedicine*, vol. 229, Feb. 2023, Art. no. 107304, https://doi.org/10.1016/j.cmpb.2022.107304.

[15] H. Chen, L. Jiao, M. Liang, F. Liu, S. Yang, and B. Hou, "Fast unsupervised deep fusion network for change detection of multitemporal SAR images," *Neurocomputing*, vol. 332, pp. 56–70, Mar. 2019, https://doi.org/10.1016/j.neucom.2018.11.077.

[16] B. Zhan, D. Li, X. Wu, J. Zhou, and Y. Wang, "Multi-Modal MRI Image Synthesis via GAN With Multi-Scale Gate Mergence," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 17–26, Jan. 2022, https://doi.org/10.1109/JBHI.2021.3088866.

[17] J. Huang, Z. Le, Y. Ma, F. Fan, H. Zhang, and L. Yang, "MGMDcGAN: Medical Image Fusion Using Multi-Generator Multi-Discriminator Conditional Generative Adversarial Network," *IEEE Access*, vol. 8, pp. 55145–55157, 2020, https://doi.org/10.1109/ACCESS.2020.2982016.

[18] L. Wang, C. Chang, B. Hao, and C. Liu, "Multi-modal Medical Image Fusion Based on GAN and the Shift-Invariant Shearlet Transform," in *2020 IEEE International Conference on Bioinformatics and Biomedicine*, Seoul, South Korea, 2020, pp. 2538–2543, https://doi.org/10.1109/BIBM49941.2020.9313288.

[19] J. Kang, W. Lu, and W. Zhang, "Fusion of Brain PET and MRI Images Using Tissue-Aware Conditional Generative Adversarial Network With Joint Loss," *IEEE Access*, vol. 8, pp. 6368–6378, 2020, https://doi.org/10.1109/ACCESS.2019.2963741.

[20] K. S. Prasad, M. Kolli, B. Linga, S. S. Chikati, and T. Veeranki, "Enhancing Medical Diagnosis Through Multimodal Medical Image Fusion," in *Enhancing Medical Imaging with Emerging Technologies*, A. K. Sharma, N. Chanderwal, S. Tyagi, P. Upadhyay, and A. K. Tyagi, Eds. Hershey, PA, USA: IGI Global Scientific Publishing, 2024, pp. 197–209, https://doi.org/10.4018/979-8-3693-5261-8.ch012.

[21] M. A. Azam *et al.*, "A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics," *Computers in Biology and*

Medicine, vol. 144, May 2022, Art. no. 105253, https://doi.org/10.1016/j.compbiomed.2022.105253.

[22] J. Jose *et al.*, "An image quality enhancement scheme employing adolescent identity search algorithm in the NSST domain for multimodal medical image fusion," *Biomedical Signal Processing and Control*, vol. 66, Apr. 2021, Art. no. 102480, https://doi.org/10.1016/j.bspc.2021.102480.

[23] V. Ramaraj, M. Venkatachalaappaswamy, and M. K. Sankar, "Medical Image Fusion for Brain Tumor Diagnosis Using Effective Discrete Wavelet Transform Methods," *Journal of Information Systems Engineering and Business Intelligence*, vol. 10, no. 1, pp. 70–80, Feb. 2024, https://doi.org/10.20473/jisebi.10.1.70-80.

[24] I. Q. Abduljaleel and I. H. Ali, "Deep Learning and Fusion Mechanism-based Multimodal Fake News Detection Methodologies: A Review," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15665–15675, Aug. 2024, https://doi.org/10.48084/etasr.7907.

[25] H. G. Doan and N. T. Nguyen, "Fusion Machine Learning Strategies for Multi-modal Sensor-based Hand Gesture Recognition," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8628–8633, Jun. 2022, https://doi.org/10.48084/etasr.4913.

[26] D.-H. Xia *et al.*, "Review-material degradation assessed by digital image processing: Fundamentals, progresses, and challenges," *Journal of Materials Science & Technology*, vol. 53, pp. 146–162, Sep. 2020, https://doi.org/10.1016/j.jmst.2020.04.033.

[27] B. H. Menze *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, https://doi.org/10.1109/TMI.2014.2377694.

[28] K. A. Johnson and J. A. Becker. "The Whole Brain Atlas." Harvard Medical School. https://www.med.harvard.edu/AANLIB/.

[29] W. Huang, H. Zhang, H. Guo, W. Li, X. Quan, and Y. Zhang, "ADDNS: An asymmetric dual deep network with sharing mechanism for medical image fusion of CT and MR-T2," *Computers in Biology and Medicine*, vol. 166, Nov. 2023, Art. no. 107531, https://doi.org/10.1016/j.compbiomed.2023.107531.