Medical Image Fusion Based on Anisotropic Diffusion and Non-Subsampled Contourlet Transform

Bhawna Goyal[1,*], Ayush Dogra[2], Rahul Khoond[1], Dawa Chyophel Lepcha[1], Vishal Goyal[3] and Steven L. Fernandes[4]

[1]Department of ECE, Chandigarh University, Mohali, 140413, Punjab, India
[2]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
[3]Department of Electronics and Communication Engineering, GLA University, Mathura, India
[4]Department of Computer Science, Design and Journalism, Creighton University, NE, USA
*Corresponding Author: Bhawna Goyal. Email: bhawnagoyal28@gmail.com
Received: 11 December 2022; Accepted: 20 April 2023; Published: 09 June 2023

Abstract: The synthesis of visual information from multiple medical imaging inputs to a single fused image without any loss of detail and distortion is known as multimodal medical image fusion. It improves the quality of biomedical images by preserving detailed features to advance the clinical utility of medical imaging meant for the analysis and treatment of medical disorders. This study develops a novel approach to fuse multimodal medical images utilizing anisotropic diffusion (AD) and non-subsampled contourlet transform (NSCT). First, the method employs anisotropic diffusion for decomposing input images to their base and detail layers to coarsely split two features of input images such as structural and textural information. The detail and base layers are further combined utilizing a sum-based fusion rule which maximizes noise filtering contrast level by effectively preserving most of the structural and textural details. NSCT is utilized to further decompose these images into their low and high-frequency coefficients. These coefficients are then combined utilizing the principal component analysis/Karhunen-Loeve (PCA/KL) based fusion rule independently by substantiating eigenfeature reinforcement in the fusion results. An NSCT-based multiresolution analysis is performed on the combined salient feature information and the contrast-enhanced fusion coefficients. Finally, an inverse NSCT is applied to each coefficient to produce the final fusion result. Experimental results demonstrate an advantage of the proposed technique using a publicly accessible dataset and conducted comparative studies on three pairs of medical images from different modalities and health. Our approach offers better visual and robust performance with better objective measurements for research development since it excellently preserves significant salient features and precision without producing abnormal information in the case of qualitative and quantitative analysis.

1 Introduction

The possibility of imaging human body cells, organs, and pathological samples to diagnose diseases has a revolutionary effect on medicine. Medical imaging informatics provides the foundation for all medical imaging tasks including image storage, analysis, processing, retrieval, and comprehension. The use of medical image processing, one of the best methods in computer vision to speed up clinical decision-making tasks and enhance diagnosis is a recent development in medical imaging. Examples of medical image processing include brain disease prediction [1], identification of pneumonia, cervical cancer diagnosis, breast cancer prediction [2], and diabetic retinopathy diagnosis from fundus images [3]. The study shows that medical image processing performs nearly as well as humans do and their diagnosis is analogous to that of trained medical personnel [4,5]. However, the accuracy and security of medical image processing are vital for researchers. According to recent studies on the segmentation and classification assignments of medical imaging even modern image-processing techniques are noticeably sensitive to adversarial attacks. Medical imaging techniques are significantly more vulnerable to adversaries compared to techniques that use natural images as their input. These vulnerabilities allow small constructed perturbations in image samples that have a significant impact on image processing performance even if they are invisible to the human eye. Many image-processing techniques have been developed to protect against adversarial attacks. One such popular domain is image fusion which produces a detailed image by combining images taken by multiple imaging modalities for clinical decision-making [6].

The visual and analytical image quality can be increased when two or more images from distinct modalities are fused. Several types of information are provided by multimodal medical images: Vibro-acoustography (VA) images provide details about the width and depth of the disease whereas computed tomography (CT) images give information about dense structures like bones. Magnetic resonance imaging (MRI) images give information about pathological soft tissues whereas magnetic resonance angiography (MRA) images can quickly identify abnormalities in the brain and X-ray images can identify fractures and abnormalities in bone positions. Functional and metabolic information on the various brain regions is provided by positron emission tomography (PET) and single photon emission computed tomography (SPECT), respectively. The multimodal medical image pairs such as CT-MRI [7], MRI-SPECT [8], PET-CT [9], MRI-PET [10], MRICT-PET [11], MRI-SPECT [12], and ultrasound-X-rays [13]" are combined for extraction of more clinical details. Doctors in the medical industry prerequisite a high level of spectral and spatial features in a specific image for a variety of purposes including study, monitoring, correct disease diagnosis, and therapy. The fact that a single image could not give all the necessary details. Therefore, finding all the relevant information in a single composite image is termed a fused image which is essentially essential in the field of medicine. Doctors may diagnose anomalies quite effectively with this combined information. A medical image fusion application has the advantage of simplifying disease diagnosis [14]. A successful image fusion can retain essential information by extracting all relevant data from the input images without introducing any discrepancies into the fused image. Machine and human perception are better adapted to the fused image.

The main contributions of our medical image fusion approach are as follows: (1) This study expands the field of fusion of medical images by introducing anisotropic diffusion and nonsubsampled Shearlet transform (NSCT). The method decreases image noise without substantially reducing the

amount of the image's information by using anisotropic diffusion by a simple structure texture decomposition of the input images to produce base and detail layers. (2) The method uses NSCT for obtaining low and high-frequency coefficients for each scale and direction of the images. The "Nonsubsampled pyramid filter banks (NSPFB) and nonsubsampled directional filter banks (NSDFB)" were used for the decomposition by full multiscale and multidirectional, respectively. (3) The theory behind principal component analysis also known as the Karhunen-Loeve transform is that the majority of the information contained throughout all of the bands may be compressed into a significantly lower number of bands with little information loss. It can enhance image corners, edges, and contrasts which provides a fused image of much greater quality. (4) The experiments have been performed to determine the efficiency of our fusion strategy for CT and MR fusion problems. It has been verified that both the qualitative and quantitative assessments of our technique yield better results.

The remaining part of this study is organized as follows. We provide a summary of the existing image fusion works in Section 2. An anisotropic diffusion (AD) and non-subsampled contourlet transform (NSCT) are briefly described in Section 3. The proposed image fusion technique is illustrated in Section 4 in detail. Extensive experiments are performed on the publicly available dataset in Section 5 to measure its fusion performance and compare it with well-known medical image fusion techniques. Section 6 provides conclusions and proposals for future work.

2 Literature Survey

This section presents the recent research works that have the same works or are related to our fusion strategy based on both traditional-based and deep learning-based algorithms. In the end, the main contributions of our technique are then illustrated in detail. Traditional techniques have drawn considerable interest in the domain of fusion during the last few years. These techniques frequently utilize edge-preserving filters for precise extraction of edge details at numerous scales that help to preserve the shift-invariance and achieve higher preservation efficiency of edge details. These techniques typically employ discriminative fusion techniques that take into account the properties of each separated image to produce a fused image. Lepcha et al. [15] introduced a fusion strategy using a "rolling guidance filter (RGF)" by preserving vital structural and textural information. Ganasala et al. used texture energy measurements (TEMs) and stationary wavelet transform (SWT) to address the issues of low contrast and high computing cost in fusion results [16]. SWT extracts estimated and detailed information from input images. The technique based on "Laplacian pyramid (LP) decomposition and adaptive sparse representation (ASR)" is presented by Wang et al. [17]. The ASR does not require a high redundancy dictionary to decrease the noise of high sub-band details during the fusion process in contrast to the traditional sparse representation (SR) technique. Kong et al. presented an image fusion approach that effectively retains detailed information for medical images from SPECT and MRI [8]. This technique makes use of the framelet transform to separate the input images into a basic layer and several detailed layers. A saliency layer of SPECT image is then extracted from the original image using a "frequency-tuned filter". Goyal et al. proposed a fusion approach that combines medical images with low resolution and lowered computing complexity to improve target recognition precision as well as serve as a basis for medical diagnosis [18]. This method could be used by doctors to efficiently assess patients' conditions. To obtain the focus area to precisely denote the input image, an L0 Gradient minimization approach was developed for medical image fusion [19]. Lepcha et al. used "pixel significance, anisotropic diffusion, and cross bilateral filter (CBF)" that avoid detail loss produced by poor image descriptions [20]. This method uses edge-preserving processing that combines linear low-pass filters with nonlinear procedures to identify significant portions of the source images while maintaining the edges. Tirupal et al. [21] used

"an interval-valued intuitionistic fuzzy set (IVIFS)" to effectively fuse images and the resulting image was subsequently subjected to a "median filter" to eliminate noise. The multi-modal fusion strategy based on "geometric algebra discrete cosine transform (GA-DCT)" was also published by Wang et al. [22]. This approach integrates the properties of GA that denote the "multi-vector signal" as a whole and could enhance the quality of the fusion image and avoid a lot of complex tasks based on encoding and decoding procedures. Still, this strategy is vulnerable to abnormal edges.

A hybrid method was introduced based on "dual-tree complex wavelet transform (DTCWT) and NSCT"[23]. Inputs for the experimental study included images from CT, MRI, and PET. This method produces competitive fusion performance by effectively capturing gradient to separate individual edge details from detailed information. Goyal et al. [24] introduced a fusion strategy based on the "pixel-based fusion rule, the cross-bilateral filter, and edge-aware filtering". A "domain transform filter" is used to separate the small-scale components of detailed images from the adjacent large-scale structures. This technique delivers exceptional image fusion performance by using hybrid domain filtering which enables the complete preservation of image features. Zhang et al. proposed a "novel joint sparse model with coupled dictionary learning" for fusing images [25]. This technique is constructed using a novel fusion scheme for improving the capability of multi-source signal preservation as well as maintaining texture or edge details. At initial, the input images were denoted by the typical sparse element and a novelty sparse element with "over-complete coupled dictionaries". Then a combined sparse element is produced utilizing a newly constructed inventive scheme. This method is effective irrespective of whether the input images are corrupted or clear. In [26], a quasi-cross bilateral filtering (QBF)-based image decomposition method is introduced to restore the noise-damaged structural details of both rough elements and fine information. An image is split into a structural layer that contains extensive information and an energy layer that has simply information about intensity. The "visual saliency detection map (VSDM)" is then utilized to control the fusing of energy layers to maximize the advantages of edge contour extraction and enhance the edge contour sharpness of the fusion result and fully restore the image details. The structural layer was combined utilizing the "weighted sum of eight-neighborhood-based modified Laplacian (WSEML) and improved multi-level morphological gradient (IMLMG)" that considerably enhanced the sharpness of the contour edges of the fusion result.

Deep learning algorithms are successfully used in numerous image-processing tasks in recent years [27,28]. Panigrahy et al. introduced a novel "weighted parameter adaptive dual-channel pulse neural network (WPADC-PCNN)" based fusion in a "non-subsampled Shearlet transform (NSST) domain" to fuse the MRI and SPECT images of patients with "acquired immunodeficiency syndrome (AIDS), dementia complex and Alzheimer's disease [29]". WPADC-PCNN model's inputs are used to estimate its parameters using fractal dimension. However, this method falls short in terms of its capacity to completely extract features from input images. Yang et al. propose "a medical Wasserstein generative adversarial network (MWGAN) is an end-to-end model" to integrate medical images from PET and MRI [30]. This method develops "two adversarial games between a generator and two discriminators" to produce a fusion result including the features of soft tissue features in organs of MRI images and the metabolic and functional information of PET images. A method for fusing medical images has been proposed in [31] and is based on a "multiscale residual pyramid attention network (MSRPAN)". Fu et al. [32] employed dense blocks and generative adversarial networks (GANs) to directly incorporate the input image such as a visible light image into every layer of the whole network. This method protects the functional information from being compromised while strengthening the structure information. Guo et al. [33] developed a novel generative adversarial network-based multimodal fusion of medical images of the brain. In the generator's design, this

technique produces the combined detail texture block and the leftover attention mechanism block. The input images are incorporated to generate a matrix which is then split into two blocks at once for extraction of details including shape, size, spatial placement, and textural properties. The restored properties are applied to the combining block to restore an image.

Zhao et al. proposed a fusion technique based on deep convolutional generative adversarial networks and dense block models to produce a fused image with more details [34]. This network structure especially combines two modules such as an image generator based on dense blocks and an encoder-decoder along with a discriminator to improve features and retain image details. Furthermore, the techniques described above are only useful for particular fusion applications. To address these limitations, numerous fusion networks [35,36] are presented and they can produce high-quality fusion images when carrying out many fusion tasks. Many fusion algorithms suffer from poor complementary feature extraction and brightness degradation. A novel end-to-end unsupervised learning fusion model called the pair "feature difference guided network (FDGNet)" is thus introduced to solve these shortcomings [37]. Liu et al. present an original end-to-end unsupervised network to integrate medical images. In made up of "a generator and two symmetrical discriminators" [38]. The previous aims to create a fusion image that is "real-like" based on precisely prepared content and structural loss, whilst the latter is concentrated on accentuating the variances between a fused image and the input ones. To change the settings of a framework, however, often takes a lot of work with these deep learning-based methodologies. Medical image fusion is still a challenging endeavor because recent techniques still produce fusion results with color distortion, blurring, and noise.

This paper provides a two-level decomposition fusion technique to address some limitations. We initially employ anisotropic diffusion to divide the source images into base and detail layers, respectively to extract sharp and fine information from the images. The images are then further divided into low and high-frequency coefficients utilizing NSCT decomposition. This framework uses anisotropic diffusion to decompose images while simultaneously seeking to minimize image noise by preserving important portions of the image's information, particularly edges, lines, and other qualities that are essential for the interpretation of images. The details of the source images are simultaneously enhanced as the two layers are fused separately using a potent sum-based fusion algorithm. The images are then subsequently divided into their low and high-frequency coefficient utilizing NSCT. In general, the NSCT frequency selection function improves directionality and makes it easier to record geometrical configurations for source images, and allows for a more accurate decomposition of medical images. The PCA/KL fusion approach is then used to independently combine these coefficients. The low and high-frequency coefficients are combined to produce a fusion result. The final fusion result was reconstructed using an inverse NSCT. Our technique is compared qualitatively and quantitatively with some of the well-known recent approaches. Our algorithm generates better quality fusion results in less processing time and the optimal visualization in case of visual quality and quantitative evaluation standards.

3 Related Work

3.1 Anisotropic Diffusion (AD)

The anisotropic diffusion procedure [39] preserves the non-homogeneous regions (edges) while smoothing out the homogeneous portions of an image using partial differential equations (PDE). It addresses the problem of isotropic diffusion. The inter-region smoothing is utilized in isotropic diffusion, thus edge information is lost. In contrast, an anisotropic diffusion creates images with a coarser lower resolution by using intra-region smoothing. At each coarser resolution, the edges are

significant and sharp. A flux function is utilized by the equation of anisotropic diffusion to regulate the diffusion of an image (I) as

$$I_t = c\ (x, y, t)\ \Delta I + \nabla c.\nabla I, \tag{1}$$

where Δ represents a Laplacian operator, ∇ represents a gradient operator, $c\ (x, y, t)$ represents a flux function or diffusion rate and t represents a time or iteration. The solution of Eq. (1) is retrieved by the forward-time central-space (FTCS) method by considering it as a heat equation and is given by

$$I_{i,j}^{t+1} = I_{i,j}^{t} + \lambda\ \left[c_N.\ \overline{\nabla}_N I_{i,j}^{t} + c_S.\ \nabla_S I_{i,j}^{t} + c_E.\ \overline{\nabla}_E I_{i,j}^{t} + c_W.\ \overline{\nabla}_W I_{i,j}^{t} \right] \tag{2}$$

In Eq. (2), $I_{i,j}^{t+1}$ is a coarser image resolution at $t+1$ scale that relies on the prior coarse scale image $I_{i,j}^{t}$. where λ represents the stability constant that satisfies $0 \le \lambda \le 0.25$. The $\overline{\nabla}_N, \overline{\nabla}_S, \overline{\nabla}_E,$ and $\overline{\nabla}_W$ are the closest neighbor variations in the north, south, east, and west directions, respectively. They are characterized as

$$\begin{aligned}
\overline{\nabla}_N I_{i,j} &= I_{i-1,j} - I_{i,j}, \\
\overline{\nabla}_S I_{i,j} &= I_{i+1,j} - I_{i,j}, \\
\overline{\nabla}_E I_{i,j} &= I_{i,j+1} - I_{i,j}, \\
\overline{\nabla}_W I_{i,j} &= I_{i,j-1} - I_{i,j}
\end{aligned} \tag{3}$$

In similar, c_N, c_S, c_E, and c_W are flux functions/conduction coefficients in the north, south, east, and west directions as

$$\begin{aligned}
c_{N_{i,j}}^{t} &= g\left(\|(\nabla I)_{i+1/2,j}^{t}\|\right) = g\left(\left|\overline{\nabla}_N\ I_{i,j}^{t}\right|\right), \\
c_{S_{i,j}}^{t} &= g\left(\|(\nabla I)_{i-1/2,j}^{t}\|\right) = g\left(\left|\overline{\nabla}_S\ I_{i,j}^{t}\right|\right), \\
c_{E_{i,j}}^{t} &= g\left(\|(\nabla I)_{i,j+1/2}^{t}\|\right) = g\left(\left|\overline{\nabla}_E\ I_{i,j}^{t}\right|\right), \\
c_{W_{i,j}}^{t} &= g\left(\|(\nabla I)_{i,j-1/2}^{t}\|\right) = g\left(\left|\overline{\nabla}_W I_{i,j}^{t}\right|\right).
\end{aligned} \tag{4}$$

In Eq. (4), $g(.)$ is a monotonically reducing function with $g(0) = 1$. Perona and Malik [39] proposed two functions as stated as

$$g\ (\nabla I) = e^{-\left(\frac{\|\nabla I\|}{k}\right)^2} \tag{5}$$

$$g\ (\nabla I) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{k}\right)^2}. \tag{6}$$

These operations provide a trade-off between edge preservation and smoothing. The first equation (Eq. (5)) is helpful when an image carries high contrast edges over low contrast edges and the second function (Eq. (6)) is helpful when an image consists the broad regions over small regions. A free parameter k is present in both functions. Based on the strength of the edge, this constant k determines the validity of a region boundary. The limitation of anisotropic diffusion filtering is that it is extremely dependent on some key factors such as a gradient threshold value, conductance function, and the stopping time of the iterative procedure. Anisotropic diffusion is represented as *aniso* (I) for a particular image (I).

3.2 Non-Subsampled Contourlet Transform (NSCT)

The NSCT is a well-organized technique that can be rapidly and easily implemented. Its increased frequency selection feature makes it possible to decompose medical images and further improved directionality and multi-scaling sensitivity make it easier to capture the geometrical arrangements for input

images. The NSCT [23] overcomes the frequency aliasing by down-and up-sampling on the contourlet transformation. The nonsubsampled directional filter banks (NSDFB) and nonsubsampled pyramid filter banks (NSPFB) are combined to construct NSCT. NSDFB provides NSCT with directionality whereas NSPFB ensures the multiscale decomposition property. Shift-invariant, multiscale, and multi-directional characteristics can be found in the discrete images produced by NSCT. Like the Laplacian pyramid, which results from the shift-invariant filter structure, the non-subsampled pyramid exhibits multiscale coefficient decomposition in NSCT. The source images are divided into low and high-frequency coefficients via NSPFB. An iterative decomposition of NSPFB was utilized for low and high-frequency coefficients to produce $l + 1$ sub-images with similar dimensions corresponding to the source images. Several decomposition stages are defined by 1. The lth stage low pass filter is defined by

$$\left[- \left(\frac{\pi}{2^l} \right) \left(\frac{\pi}{2^l} \right) \right]^2 \tag{7}$$

Eq. (8) provides the l-level cascaded non-sampled pyramid decomposed filter as

$$P_n^{eq}(z) = \begin{cases} P_1 \left(Z^{2^{n-1}I} \right) \Pi_{l=0}^{n-2} P_o Z^{2^l I}, & 1 \leq n \leq l \\ \Pi_{l-0}^{n-2} P_o \left(Z^{2^l I} \right), & n = l + 1 \end{cases} \tag{8}$$

The non-subsampled directional filter bank (NSSDB) is where the directional information in NSCT is generated from. The up-sampler and down-sampler from the directional filter are removed to create NSCT. The l-level directional decomposition is used to process high-frequency coefficients obtained from NSCT to produce 2^m directional sub-images that are about a similar size as the input images [23]. The drawbacks of NSCT include that the contrast of the synthesized image is lowered or that information is not precisely maintained. Fig. 1 shows the schematic flowchart of the proposed method.
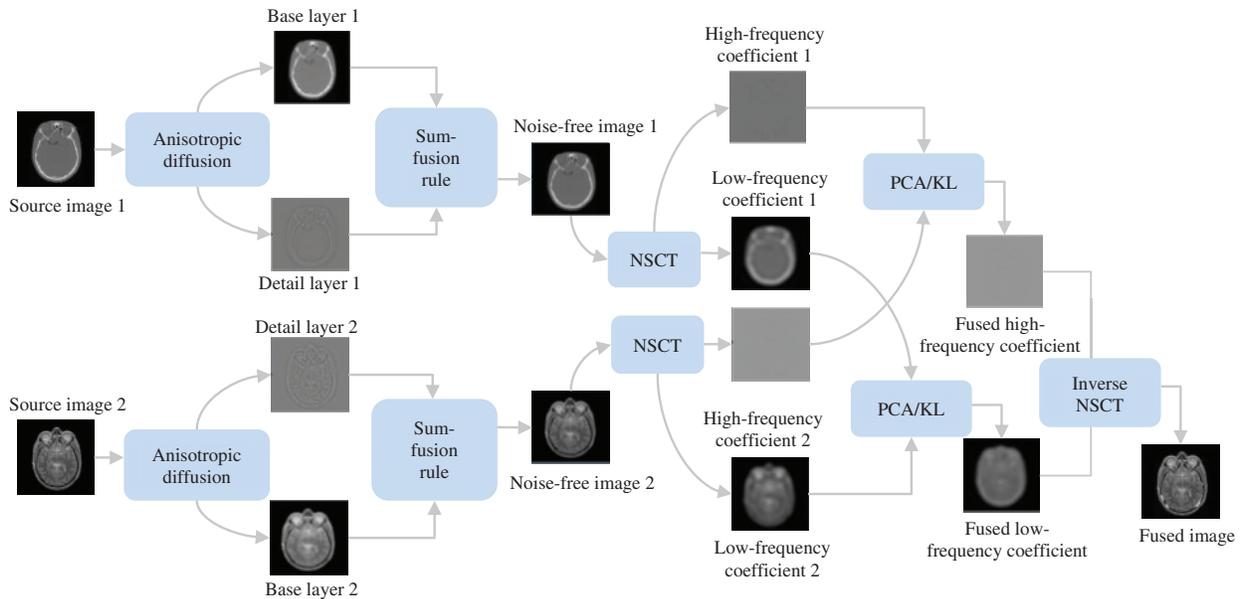
Figure 1: Flowchart of the proposed method

4 The Proposed Algorithm

The proposed algorithm mainly entails five stages: anisotropic decomposition, base and detail layers integration, NSCT decomposition, the fusion of low and high-frequency coefficients, and inverse NSCT are illustrated in Fig. 1. First, Anisotropic diffusion (AD) decomposes source images into base and detail layers. While important information about images is then simultaneously preserved as both layers are combined utilizing the sum-based fusion scheme. Furthermore, the modified images are then divided into low and high-frequency coefficients using NSCT. Both low and high-frequency coefficients are combined utilizing PCA/KL transform by extracting the specific features. The use of inverse NSCT results in the creation of the final fusion result. Fig. 1 demonstrates the schematic flowchart of our fusion technique.

4.1 Anisotropic Diffusion for Edge Detail Reinforcement

An edge-preserving smoothing anisotropic diffusion is utilized to split a source image into base and detail layers. Let the size of a source image $\{I_n(x, y)\}_{n=1}^{N}\}$ be m × n. The base layer is extracted as

$$B\_L_n(x, y) = \text{aniso}(I_n(x, y)) \tag{9}$$

where $\text{aniso}(I_n(x, y))$ signifies an anisotropic diffusion procedure on the n-th input image, and $B\_L_n(x, y)$ is the n-th base layer. For further information on anisotropic diffusion refer to section 3.1. Thus, subtracting the base layer from the input image yields a detail layer as

$$D\_L_n(x, y) = I_n(x, y) - B\_L_n(x, y) \tag{10}$$

We employ the sum-based fusion rule which is provided by Eq. (11) to produce enhanced images as

$$M\_S_n(x, y) = w_{t1} B\_L_n(x, y) + w_{t2} D\_L_n(x, y) \tag{11}$$

where w_{t1} and w_{t2} are the weight values for the coefficients $B\_L_n(x, y)$ and $D\_L_n(x, y)$, and $M\_S_n(x, y)$ is the n-th source images. We select $w_{t1} = w_{t2} = 1$ in Eq. (11) to maintain more neighborhood structure, global structure, and brightness values in fused modified source images. This paper signifies $M\_S_{n1}(x, y)$ as an enhanced image of source image1 and $M\_S_{n1}(x, y)$ as an enhanced image of source image2 as demonstrated in Fig. 1.

4.2 PCA/KL Transform-Based Fusion of Low and High-Frequency Coefficients

The enhanced images $M\_S_{n1}(x, y)$, and $M\_S_{n2}(x, y)$ are further split into their low and high-frequency coefficients using NSCT. The decomposition of $M\_S_{n1}(x, y)$ by the NSCT yields $H_{P2}(x, y)$, and $L_{P1}(x, y)$. In similar, the NSCT decomposition of $M\_S_{n2}(x, y)$ yields $H_{P1}(x, y)$, and $L_{P2}(x, y)$. Detailed elements of input images such as textures, edges, and regional boundaries are represented by the high-frequency coefficients in the NSCT domain. In general, the sharper brightness in the image is correlated with coefficients with higher absolute scores. It should be noted that noise is also associated with high-frequency coefficients and may result in incorrectly calculating sharpness values which would then influence the performance of the fusion process. Additionally, texture detail features have a higher sensitivity in the human visual system than pixel values. An approximation element of the input images is represented by low-frequency coefficients. The way that is most frequently employed is to use averaging techniques to create fused coefficients. However, this regulation reduced contrast in the fusion results and was unable to produce a high-quality fused sub-image for images. In this study, a novel high-frequency fusion rule related to PCA/KL is employed to address the above-mentioned flaw. A well-known PCA/KL rule is used for simultaneously fusing both frequency coefficients and

extracting texture characteristics. The high-frequency coefficients fusion rule is defined as:

$$F_{HP}(x, y) = K_{L1}H_{P1}(x, y) + K_{L2}H_{P2}(x, y) \tag{12}$$

Similarly, a fusion rule on the low-frequency coefficients is provided by

$$F_{LP}(x, y) = K_{L1}L_{P1}(x, y) + K_{L2}L_{P2}(x, y) \tag{13}$$

where K_{L1} and K_{L2} is given by

$$K_{L1} = \frac{\rho_{max}(1)}{\sum_i \rho_{max}(i)}, \ K_{L2} = \frac{\rho_{max}(2)}{\sum_i \rho_{max}(i)} \tag{14}$$

where ρ_{max} is the eigenvector associated with the maximum eigenvalue and denoted by

$$\rho_{max} = \max(\sigma_1, \sigma_2) \tag{15}$$

It is possible to determine the eigenvalues for an eigenvector as

$$\sigma_1 = \begin{bmatrix} \rho_1(1) \\ \rho_1(2) \end{bmatrix} \ \text{and} \ \sigma_2 = \begin{bmatrix} \rho_2(1) \\ \rho_2(2) \end{bmatrix} of \ Cv_{xx} \tag{16}$$

where Cv_{xx} is the covariance matrix [39]. The general phrase for N layers is generated as

$$F_N(x, y) = \sum_{n=1}^{N} K_{Ln}F_n(x, y) \tag{17}$$

4.3 Inverse NSCT for Fused Low and High-Frequency Coefficients

By applying inverse NSCT on the fused high-frequency coefficient $F_{HP}(x, y)$, and low-frequency coefficient $F_{LP}(x, y)$ obtains the final fusion result F_{fused} as follows:

$$F_{fused} = in\_nsct \ (F_{HP}(x, y), \ F_{LP}(x, y)) \tag{18}$$

5 Experiments

This study presents experimental results on three pairs of 256×256 registered medical images from three different modalities due to space limitations to show the efficiency of the proposed technique for fusing medical images. Three sets of CT and MRI images (Groups 1–3) are utilized in the experiments. All of the input images are selected from the Harvard Medical School website [40] and demonstrated in Fig. 2. The appropriate experimental findings and discussions are illustrated in detail. Nine representative image fusion algorithms including "Wasserstein generative adversarial networks (WGAN) [30], a weighted parameter adaptive dual-channel PCNN (WPADC-PCNN) [29], Laplacian pyramid and adaptive sparse representation (LP-ASR) [17], Geometric algebra discrete cosine transform (GA-DCT) [22], Image fusion using saliency capturing (IF-SC) [8], L0 gradient minimization (LGM) [19], Hybrid approach of NSCT and DTCWT (NSCT-DTCWT) [23], Multimodal brain medical image fusion based on the improvement of a generative adversarial network (MMFGAN) [33] and Multiscale residual pyramid attention network (MRPAN) [31]". The default values of the parameters were taken for the above-mentioned fusion methods which are stated in each of the publications.

(a) Group (1)   (b) Group (2)   (c) Group (3)
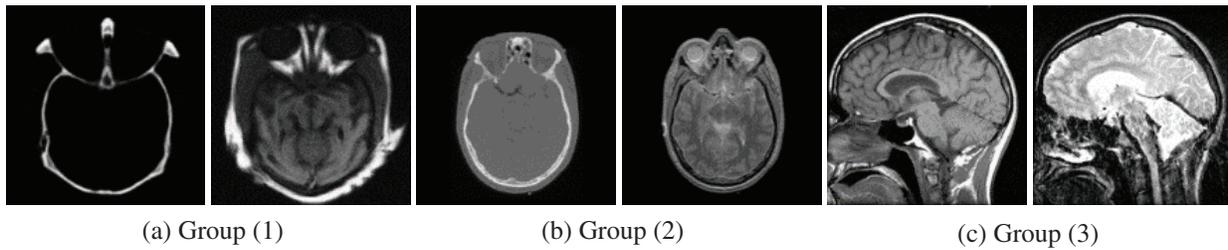
Figure 2: Source image pairs of multimodal medical images [40]. In (a–c), three pairs of CT and MRI images. Left: CT image; right: MRI image

5.1 Objective Evaluation Metrics

The ten widely used quantitative indexes were utilized in our research to compare the effectiveness of various approaches. These indexes include "average pixel intensity (API), standard deviation (SD), average gradient (AG), entropy (H), mutual information (MI), fusion symmetry (FS), and spatial frequency (SF)". A comprehensive interpretation of these indexes is given in [15]. These seven indicators with a high value indicate effective performance. In addition to these indexes, total information transmitted $(Q_{ab/f})$, total information lost $(L_{ab/f})$, and artifacts included in the final fusion result $(N_{ab/f})$ are used to produce an in-depth analysis of fusion efficiency [24]. Effective performance is indicated by the higher value of $(Q_{ab/f})$ and low values of $(L_{ab/f})$ and $(N_{ab/f})$. Tables 1–3 display the quantitative results of different algorithms and our method. Three different types of images are used to evaluate the parameter configurations for the proposed framework through a series of experiments. The proposed framework has some key parameters: In anisotropic diffusion [39]; an integration constant (delta_t) is set to 0.15 (typically due to numerical stability is set to its maximum value), the gradient modulus threshold (k) that regulates conduction is set to 20 and the conduction coefficients function is set to 2 and the number of iterations is set to 10. The default parameter settings of NSCT are based on the modified input images which are decomposed into eight and four-directional coefficients. We considered the default parameter configurations recommended by the authors for other competing approaches.

Table 1: Quantitative indices of fusion results (Group 1)

Method	API	SD	AG	Entropy	MI	FS	SF	$Q_{ab/f}$	$L_{ab/f}$	$N_{ab/f}$
LP-ASR	38.3499	37.7569	3.3144	6.2830	3.3768	1.6284	5.6066	0.7054	0.2942	0.2227
WGAN	45.8862	33.8827	4.1223	6.1552	2.8836	1.5825	5.9922	0.7241	0.1342	0.2672
WPADC-PCNN	61.3311	30.0406	4.7658	6.4833	1.4857	1.5111	6.1892	0.7133	0.2865	0.0121
LGM	36.6337	53.0486	4.4883	6.1566	2.5936	1.6322	12.0485	0.8660	0.0730	0.1125
IF-SC	53.2614	53.6806	5.3888	6.4568	5.0822	1.5896	9.7535	0.8299	0.0927	0.1219
MRPAN	43.9927	47.9927	3.8826	5.8926	4.8811	1.5625	7.6624	0.8452	0.0815	0.1027
MMFGAN	28.6404	34.2512	2.4638	6.1171	2.9016	1.5703	4.7416	0.7314	0.2685	0.1171
GA-DCT	58.3229	30.7358	3.0406	6.1379	2.2838	1.6746	5.3215	0.7698	0.1001	0.0281
NSCT-DTCWT	59.6718	49.1795	3.9513	6.6349	3.9516	1.6772	6.8274	0.8759	0.696	0.0198
Proposed	67.6478	67.6473	67.647	67.647	4.7863	1.6403	16.7204	0.8814	0.0626	0.0090

Table 2: Quantitative indices of fusion results (Group 2)

Method	API	SD	AG	Entropy	MI	FS	SF	$Q_{ab/f}$	$L_{ab/f}$	$N_{ab/f}$
LP-ASR	61.2859	53.4289	4.4996	5.3101	3.8943	1.9316	9.1310	0.7286	0.2049	0.1307
WGAN	67.8726	56.8826	4.9082	5.6662	3.2761	1.9452	10.9862	0.7651	0.2551	0.1271
WPADC-PCNN	56.3188	43.3953	5.6448	5.0031	2.3694	1.9432	8.1734	0.7434	0.2543	0.0975
LGM	55.2151	52.7005	_6.1506_	5.3016	_3.8097_	1.9269	8.9583	0.7753	0.2151	0.0685
IF-SC	44.8918	51.5913	4.8231	5.3128	3.3591	1.8874	**9.8594**	0.7747	_0.1809_	0.1072
MRPAN	48.8716	53.8772	3.9928	4.9881	3.6251	1.7927	8.9651	0.7615	0.1982	0.1024
MMFGAN	50.3958	_53.9736_	3.5071	5.3929	3.4798	1.9333	7.2433	0.7468	0.2507	_0.0317_
GA-DCT	_61.4767_	45.2395	3.8614	**6.3149**	3.2671	_1.9597_	8.0808	0.7691	0.2211	0.0535
NSCT-DTCWT	60.5334	52.0540	4.0603	5.4488	3.4756	**1.9602**	8.0759	_0.7853_	0.2127	0.0463
Proposed	**63.8278**	**54.1378**	6.2764	_5.8488_	**3.8506**	1.9558	9.8372	**0.7864**	**0.1054**	**0.0247**

Table 3: Quantitative indices of fusion results (Group 3)

Method	API	SD	AG	Entropy	MI	FS	SF	$Q_{ab/f}$	$L_{ab/f}$	$N_{ab/f}$
LP-ASR	96.3913	62.2677	7.8748	7.4933	3.0778	1.6025	11.0157	0.6987	0.3013	0.0314
WGAN	89.6624	67.7251	8.6622	5.6626	2.1765	1.3567	10.7726	0.6741	0.2161	0.5614
WPADC-PCNN	_100.724_	48.0173	9.0632	7.3832	2.2724	1.6527	11.3875	0.6625	0.3372	0.0436
LGM	98.254	63.4612	17.4349	7.4236	_3.3510_	1.6676	_26.9397_	0.7264	0.2677	0.0312
IF-SC	97.8820	_71.0525_	**12.6746**	_7.7586_	3.1382	1.7623	17.1068	_0.7597_	_0.0951_	0.1417
MRPAN	87.6635	65.6662	9.5626	6.8926	2.6726	1.5992	15.6221	0.7451	0.1102	0.0451
MMFGAN	84.2996	53.7153	6.0607	7.5031	2.8106	1.7626	9.4684	0.7464	0.2235	0.0351
GA-DCT	91.8459	47.1212	7.3526	7.2723	2.4922	_1.8897_	10.5472	0.6876	0.1124	0.0241
NSCT-DTCWT	97.5691	57.3488	7.9617	7.4543	2.8848	1.8902	11.2761	0.7493	0.1005	_0.0217_
Proposed	**102.878**	**80.1291**	_12.6340_	**7.8009**	**3.9482**	1.8754	**30.7082**	**0.8083**	**0.0981**	**0.0981**

5.2 Results and Analysis

1) *Qualitative Analysis*

The complementary properties of the two types of images can be successfully restored in a single image when CT and MRI images are fused. Experiments are performed on three sets of MRI and CT images. The sets of fused CT and MRI imaging data are shown in Figs. 3 and 4. The two representative regions are enlarged in each resultant image to help with better comparisons. The fused images of LP-ASR and WGAN approaches demonstrated significant artifacts, notably in the bone sections [see the right regions in Figs. 3a and 3b]. The fused images of the WPADC-PCNN are impacted by energy loss in the image [see the enlarged region in Fig. 3c]. The bone areas of CT images appear to have blurry by the LGM method [see Fig. 3d]. The noticeable features can be preserved almost entirely using the IF-SC and MRPAN approaches however some details are lost and extra futile details are added to the bone areas. The MMFGAN technique generates the finest visual interpretation because it almost entirely recovers the soft tissue details from MRI images. Nonetheless, some artifacts are incorporated in bone areas. The GA-DCT method emphasizes more on detail extraction which appears to be the root cause of the spatial inconsistency problem. NSCT-DTCWT and the proposed method show diverse benefits in the restoration of bone areas as they employ noise filtering to remove more aberrant details when compared to other methods [see the enlarged regions in Figs. 3i and 3j]. The soft tissue details in the MRI image are still poorer after utilizing NSCT-DTCWT [see the left region in Fig. 3i].

The proposed method concurrently extracts the majority of the salient data as well as less significant details from CT and MRI images. In contrast to the Group 1 image, Group 2 images have a more noticeable bone structure. The fact that LP-ASR approach uses a large amount of energy and still manages to extract almost all of the data from Group 2 images in the fused image [see Fig. 4a]. Besides, the fused image could not completely preserve several significant features of Group 2. The WGAN algorithm does not preserve the salient features which resulted in a poor visual impression [see Fig. 4b]. The LGM, IF-SC, MRPAN, and MMFGAN techniques did not effectively maintain the essential elements of Group 2 images in the fused image. The GA-DCT method allows for the transfer of nearly all significant characteristics from the source image to the fused image. Consequently, more artifacts are introduced to the fused image since it is unable to distinguish the relevant and significant features [see the enlarged region in Fig. 4h]. Moreover, NSCT-DTCWT fails to completely maintain textures and is unable to extract the important features from Group 2 images [see the right region in Fig. 4i]. In addition to performing better than NSCT-DTCWT at extracting essential features from Group 2 images, the proposed technique also avoids certain undesired information. The fusion results of Group 3 images are shown in Fig. 5. The structural details are generally poorly retained by the fusion performance of the LP-ASR, WGAN, WPADC-PCNN, LGM, and MMFGAN methods that lead to serious visual inconsistency concerns that make this condition unfit for medical diagnosis. The IF-SC, MRPAN, and GA-DCT techniques typically perform well in terms of extracting functional information from the CT image and also exhibit good visual effects even though there are still some distortions. Also, the region with a high pixel value has a few unfavorable artifacts [see the white region in Figs. 3e, 4f and 5h]. When it comes to feature preservation, the NSCT-DTCWT does an excellent task. Nonetheless, there are some artifacts in the fused image [see the right side of Fig. 5i]. The proposed algorithm performs the best in terms of preserving detailed information [see the enlarged regions in Fig. 5j].
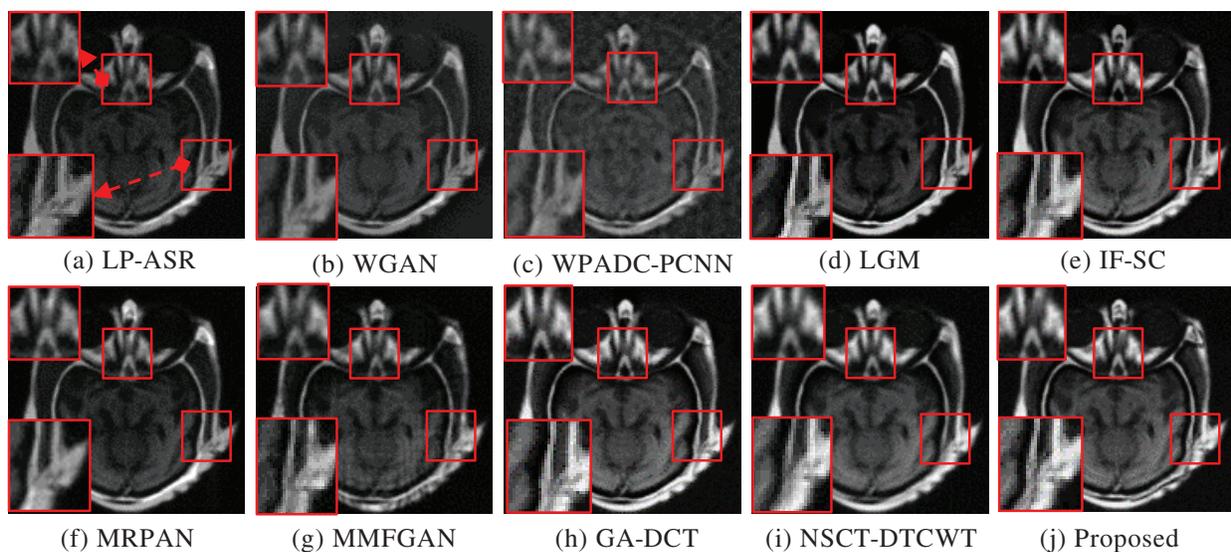
Figure 3: Fusion results of Group 1. The two regions of interest (ROI) are shown for better comparisons
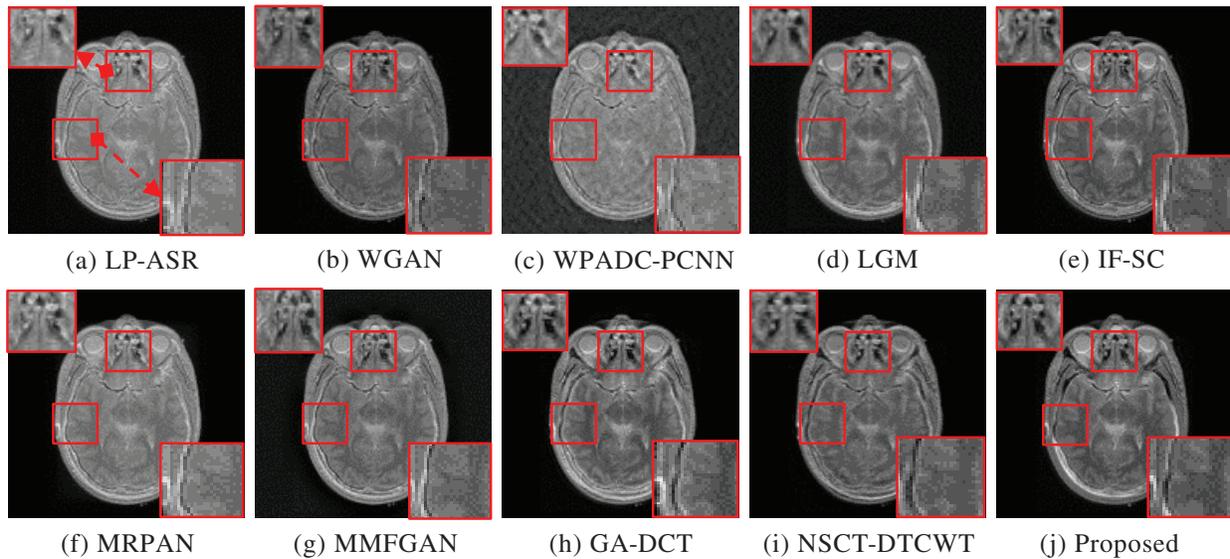
Figure 4: Fusion results of Group 2. The two regions of interest (ROI) are shown for better comparisons
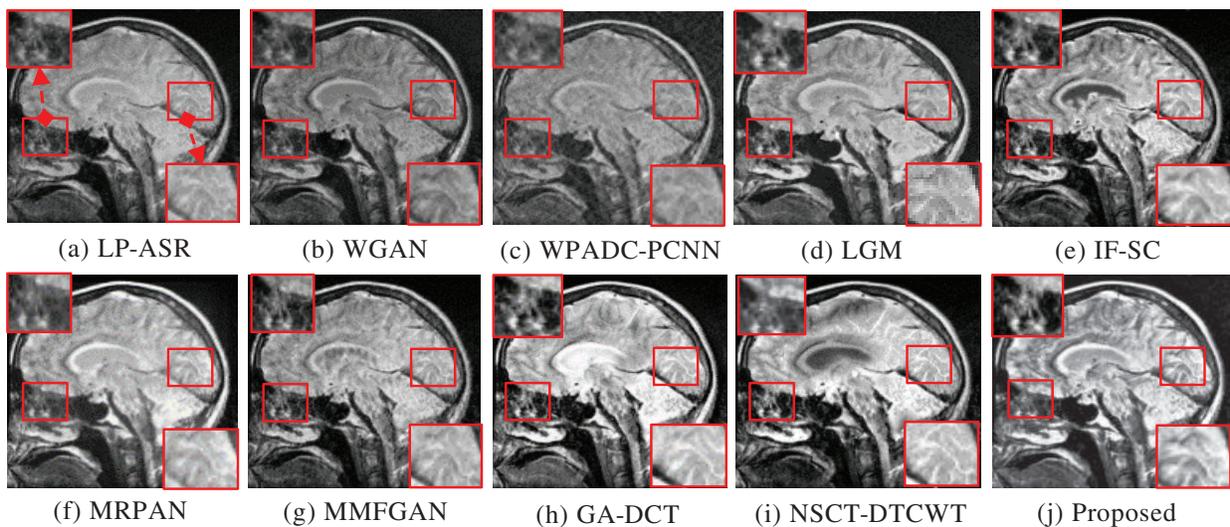
Figure 5: Fusion results of Group 3. The two regions of interest (ROI) are shown for better comparisons

2) *Quantitative Analysis*

The results of an objective evaluation of the nine fusion algorithms and the proposed method on three pairs of CT and MRI images are demonstrated in Tables 1–3 which provide the average value for each approach overall experimentally used images. For each of the 10 approaches, the metric with the best score is bolded and the metric with the second-best score is underlined. Tables 1–3 clearly show that the proposed technique regularly places in the top three positions for nearly all of the objective measures except for the metric FS in all CT-MRI image fusion problems. Our technique consistently achieves stable performance over the other nine fusion methods on all ten metrics; however, it is not always the best. As a result, our fusion framework performs equally well as MRI fusion problems. In

particular, the proposed method shows obvious advantages on all the metrics when compared to other methods, proving that our strategy provides better results for preserving the general characteristics and intensity of the original images. For metric Entropy, the proposed technique outperforms NSCT-DTCWT in most of the fusion metrics. But the metrics scores of the proposed method are less than those of the NSCT-DTCWT in some cases. The GA-DCT approach also performs very well objectively for CT-MRI fusion problems. The NSCT-DTCWT approach works better for the metric FS, and this finding validated that it obtains better visual performance. Unfortunately, it performs poorly on the metrics API, SD, AG, MI, and SF since it concentrates on extracting structural details and often failed to identify the main elements. The proposed algorithm outperforms the NSCT-DTCWT method on the aforementioned metrics and represents a better benefit on almost all the metrics in the CT-MRI fusion problems that show our technique achieves better objective evaluation results. Fig. 6 displays the sensitivity analysis of fusion performance in terms of the number of iterations of our method.
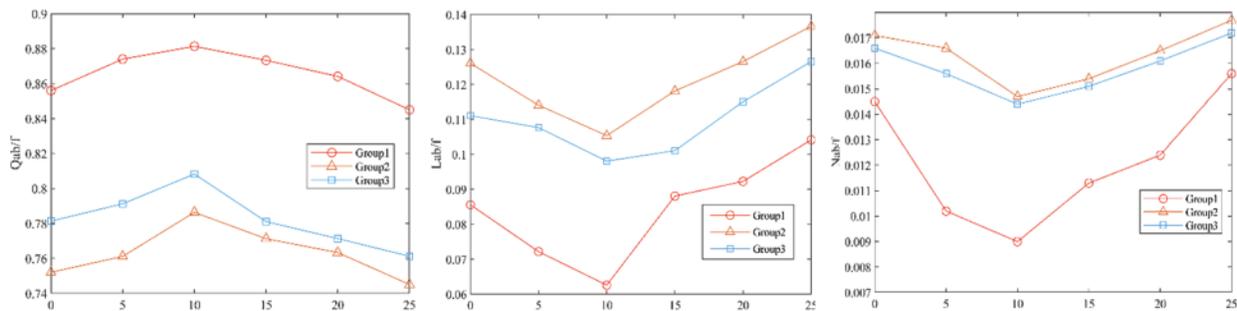
Figure 6: The objective performance of our method in terms of the number of iterations

5.3 Discussion

Tables 1–3 explicitly show that our algorithm outperforms other fusion algorithms on three pairs of CT-MRI fusion problems. Also, the proposed approach scores higher in metrics such as API, SD, MI, and SF which indicates that it is more in-depth in the extraction of image energy and prominent edge details. So, the fusion results generated by our algorithm appear more natural and are in accord with human visual perception Likewise, Figs. 3–5 show how our algorithm can effectively maintain bone areas while eliminating almost all aberrant characteristics in the same region of the complementary images. The result of our fusion method is also consistent with its ability to maintain decisive features while removing irrelevant ones. Unlike existing strategies, the proposed method puts more of a focus on the preservation of broad features and the extraction of significant information and details. In Figs. 3–5, some artifacts are present in the fusion results produced by other algorithms since these methods fail to recognize the important and meaningful characteristics that consequence in poor fusion results. The edge details seem to be strengthened by the IFSC and MRPAN approaches, but textures are still lost. A few redundant CT-MRI image features are added by the MMFGAN approach. MRPAN experiences texture degradation when referring to the CT-MRI images. This is primarily caused by the fact that the detail layer of an MRI image consists of almost all of the texture details. The proposed technique can successfully protect none details from the addition of unnecessary details by forfeiting a few textures. Particularly, the proposed method emphasizes restoring global features and appears more serious than other methods in the case of maintaining consistent information. The proposed method might be further enhanced by using efficient contrast enhancement techniques to the detail layer to enhance texture details. Table 4 displays an average running time for 10 different algorithms.

Table 4: The average processing durations for different techniques for fusing two 256×256-pixel images (Units: Seconds)

Methods	LP-ASR	WGAN	WPADC-PCNN	LGM	IF-SC	MRPAN	MMFGAN	GA-DCT	NSCT-DTCWT	Proposed
Times	0.56	16.91	21.23	0.65	0.98	33.18	46.18	0.88	0.111	0.87

5.4 Computational Efficiency

In this section, the computing performance of several fusion algorithms is analyzed and compared. The experimental setting is a machine running the "Win 10 64-bit operating system with an Intel(R) Core (TM) i7-8750H CPU (2.20 GHz) and 8 GB RAM". All of the source images in the fusion experiments are utilized to estimate the average running time. As indicated in Table 4, the LP-ASR and LGM algorithms run faster than other fusion algorithms. The IF-SC and GA-DCT algorithms exhibit good computing efficiency and are simple to use. The NSCT-DTCWT technique has competitive efficiency at the expense of running time. In addition, the WGAN, WPADC-PCNN, MRPAN, and MMFGAN use pre-training which is needed to design a network model before the fusion procedure. The time-consuming process is the main reason it takes longer to run. In conclusion, the proposed approach requires a slightly average processing time but yields superior fused images.

6 Conclusion

A multimodal medical image fusion using anisotropic diffusion and non-subsampled contourlet transform (NSCT) is presented in this study. Our method uses anisotropic diffusion for image decomposition to improve edge information. The sum-based fusion rule is then utilized to fuse the base and detail layers, respectively. To merge NSCT-decomposed low and high-frequency coefficients and the majority of the information inside all the coefficients can be compressed into a much lower number of coefficients with minimal information loss using a PCA/KL transform. The corners, borders, and contrasts of the image can also be improved thus generating a fused image of far superior quality. In the end, the final fusion result is reconstructed using inverse NSCT on the fused low and high-frequency coefficients. Extensive experiments are conducted on three different types of medical image pairs. Our method displays significant robustness and obtains competitive results in the visual interpretation and objective evaluation of three pairs of CT-MRI images since it accurately extracts relevant details and less aberrant information than state-of-the-art fusion approaches. Experimental studies of the significant differences between the different algorithms and comparative experiments confirmed that our fusion algorithm outperforms other algorithms in terms of both visual effect and objective evaluation. However notable problem is that the proposed method is slightly time-consuming due to the multilevel decomposition and fusion process. In the future, an alternative fusion rule to the sum-and PCA-based fusion rules might be considered for the fusion process to improve performance.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1]   M. Hu, Y. Zhong, S. Xie, H. Lv and Z. Lv, "Fuzzy system-based medical image processing for brain disease prediction," *Frontiers in Neuroscience*, vol. 15, pp. 965, 2021.

[2]   G. U. Nneji, J. Cai, J. Deng, H. N. Monday, E. C. James *et al.,* "Multi-channel based image processing scheme for pneumonia identification," *Diagnostics*, vol. 12, no. 2, pp. 325, 2022.

[3]   H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia *et al.,* "Diabetic retinopathy diagnosis from fundus images using stacked generalization of deep models," *IEEE Access*, vol. 9, pp. 108276–108292, 2021.

[4]   Z. Lei, Y. Sun, Y. A. Nanehkaran, S. Yang, M. S. Islam *et al.,* "A novel data-driven robust framework based on machine learning and knowledge graph for disease classification," *Future Generation Computer Systems*, vol. 102, pp. 534–548, 2020.

[5]   F. Ren, C. Yang and Y. A. Nanehkaran, "MRI-based model for MCI conversion using deep zero-shot transfer learning," *Journal of Supercomputing*, vol. 79, no. 2, pp. 1182–1200, 2023.

[6]   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, 2021.

[7]   A. Dogra, B. Goyal, S. Agarwal and C. K. Ahuja, "Efficient fusion of osseous and vascular details in wavelet domain," *Pattern Recognition Letters*, vol. 94, pp. 189–193, 2017.

[8]   W. Kong, Y. Lei, L. Cui, C. Li and J. Xue, "MRI and SPECT image fusion using saliency capturing," in *2021 IEEE Int. Conf. on Intelligent Systems and Knowledge Engineering (ISKE)*, Chengdu, China, pp. 249–253, 2021.

[9]   Y. Wang and Y. Liu, "Three-dimensional structure tensor-based PET/CT fusion in gradient domain," *Journal of X-Ray Science and Technology*, vol. 27, no. 2, pp. 307–319, 2019.

[10] A. Lakshmi, M. P. Rajasekaran, S. Jeevitha and S. Selvendran, "An adaptive MRI-PET image fusion model based on deep residual learning and self-adaptive total variation," *Arabian Journal for Science and Engineering*, vol. 47, no. 8, pp. 10025–10042, 2022.

[11] L. Guo, S. Shen, E. Harris, Z. Wang, W. Jiang *et al.,* "A tri-modality image fusion method for target delineation of brain tumors in radiotherapy," *PLoS One*, vol. 9, no. 11, pp. e112187, 2014.

[12] C. Panigrahy, A. Seal and N. K. Mahato, "MRI and SPECT image fusion using a weighted parameter adaptive dual channel PCNN," *IEEE Signal Processing Letters*, vol. 27, no. 3, pp. 690–694, 2020.

[13] C. R. Hatt, A. K. Jain, V. Parthasarathy, A. Lang and A. N. Raval, "MRI—3D ultrasound—X-ray image fusion with electromagnetic tracking for transendocardial therapeutic injections: In-vitro validation and in-vivo feasibility," *Computerized Medical Imaging and Graphics*, vol. 37, no. 2, pp. 162–173, 2013.

[14] S. Karim, G. Tong, J. Li, A. Qadir, U. Farooq *et al.,* "Current advances and future perspectives of image fusion: A comprehensive review," *Information Fusion*, vol. 90, pp. 185–217, 2023.

[15] D. C. Lepcha, B. Goyal and A. Dogra, "Image fusion based on cross bilateral and rolling guidance filter through weight normalization," *The Open Neuroimaging Journal*, vol. 13, no. 1, pp. 51–61, 2021.

[16] P. Ganasala and A. D. Prasad, "Medical image fusion based on laws of texture energy measures in stationary wavelet transform domain," *International Journal of Imaging Systems and Technology*, vol. 30, no. 3, pp. 544–557, 2020.

[17] Z. Wang, Z. Cui and Y. Zhu, "Multi-modal medical image fusion by Laplacian pyramid and adaptive sparse representation," *Computers in Biology and Medicine*, vol. 123, pp. 103823, 2020.

[18] B. Goyal, D. C. Lepcha, A. Dogra, V. Bhateja and A. Lay-Ekuakille, "Measurement and analysis of multi-modal image fusion metrics based on structure awareness using domain transform filtering," *Measurement*, vol. 182, pp. 109663, 2021.

[19] S. Zhang, X. Li, R. Zhu, X. Zhang, Z. Wang *et al.,* "Medical image fusion algorithm based on L0 gradient minimization for CT and MRI," *Multimedia Tools and Applications*, vol. 80, no. 14, pp. 21135–21164, 2021.

[20] S. Kaur, B. Goyal and A. Dogra, "Haze removal in remote sensing images for improved data analysis and extraction," *2022 9th Int. Conf. on Computing for Sustainable Global Development (INDIACom)*, IEEE, pp. 01–06, 2022.

[21] T. Tirupal, B. Chandra Mohan and S. Srinivas Kumar, "Multimodal medical image fusion based on interval-valued intuitionistic fuzzy sets," in *Machines, Mechanism and Robotics: Proc. of iNaCoMM*, Singapore, Springer, pp. 965–971, 2022.

[22] R. Wang, N. Fang, Y. He, Y. Li, W. Cao *et al.,* "Multi-modal medical image fusion based on geometric algebra discrete cosine transform," *Advances in Applied Clifford Algebras*, vol. 32, no. 2, pp. 1–23, 2022.

[23] N. Alseelawi, H. T. Hazim and H. T. S. ALRikabi, "A novel method of multimodal medical image fusion based on hybrid approach of NSCT and DTCWT," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 18, no. 3, pp. 114–133, 2022.

[24] B. Goyal, A. Dogra, D. C. Lepcha, D. Koundal, A. Alhudhaif *et al.,* "Multi-modality image fusion for medical assistive technology management based on hybrid domain filtering," *Expert Systems with Applications*, vol. 209, pp. 118283, 2022.

[25] C. Zhang, Z. Zhang, Z. Feng and L. Yi, "Joint sparse model with coupled dictionary for medical image fusion," *Biomedical Signal Processing and Control*, vol. 79, pp. 104030, 2023.

[26] Y. Zhang, M. Wang, X. Xia, D. Sun, X. Zhou *et al.,* "Medical image fusion based on quasi-cross bilateral filtering," *Biomedical Signal Processing and Control*, vol. 80, pp. 104259, 2023.

[27] P. Verma, A. Dumka, R. Singh, A. Ashok, A. Singh *et al.,* "A deep learning based approach for patient pulmonary CT image screening to predict coronavirus (SARS-CoV-2) infection," *Diagnostics*, vol. 11, no. 9, pp. 1735, 2021.

[28] N. Kumar, M. Gupta, D. Gupta and S. Tiwari, "Novel deep transfer learning model for COVID-19 patient detection using X-ray chest images," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 1, pp. 469–478, 2023.

[29] B. Goyal, A. Gupta, A. Dogra and D. Koundal, "An adaptive bitonic filtering based edge fusion algorithm for Gaussian denoising," *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 90–97, 2022.

[30] Z. Yang, Y. Chen, Z. Le, F. Fan and E. Pan, "Multi-source medical image fusion based on wasserstein generative adversarial networks," *IEEE Access*, vol. 7, pp. 175947–175958, 2019.

[31] J. Fu, W. Li, J. Du and Y. Huang, "A multiscale residual pyramid attention network for medical image fusion," *Biomedical Signal Processing and Control*, vol. 66, pp. 102488, 2021.

[32] Y. Fu, X. J. Wu and T. Durrani, "Image fusion based on generative adversarial network consistent with perception," *Information Fusion*, vol. 72, pp. 110–125, 2021.

[33] K. Guo, X. Hu and X. Li, "MMFGAN: A novel multimodal brain medical image fusion based on the improvement of generative adversarial network," *Multimedia Tools and Applications*, vol. 81, no. 4, pp. 5889–5927, 2022.

[34] C. Zhao, T. Wang and B. Lei, "Medical image fusion method based on dense block and deep convolutional generative adversarial network," *Neural Computing and Applications*, vol. 33, no. 12, pp. 6595–6610, 2021.

[35] Z. Chao, X. Duan, S. Jia, X. Guo, H. Liu *et al.,* "Medical image fusion via discrete stationary wavelet transform and an enhanced radial basis function neural network," *Applied Soft Computing*, vol. 118, pp. 108542, 2022.

[36] J. A. Bhutto, L. Tian, Q. Du, Z. Sun, L. Yu *et al.,* "CT and MRI medical image fusion using noise-removal and contrast enhancement scheme with convolutional neural network," *Entropy*, vol. 24, pp. 393, 2022.

[37] G. Zhang, R. Nie, J. Cao, L. Chen and Y. Zhu, "FDGNet: A pair feature difference guided network for multimodal medical image fusion," *Biomedical Signal Processing and Control*, vol. 81, pp. 104545, 2023.

[38] H. Liu, S. Li, J. Zhu, K. Deng, M. Liu *et al.,* "DDIFN: A dual-discriminator multi-modal medical image fusion network," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 19, pp. 1–17, 2023.

[39] D. P. Bavirisetti and R. Dhuli, "Fusion of infrared and visible sensor images based on anisotropic diffusion and karhunen-loeve transform," *IEEE Sensors Journal*, vol. 16, no. 1, pp. 203–209, 2016.

[40] K. A. Johnson and J. A. Becker. 2011. [Online]. Available: http://www.med.harvard.edu/aanlib/