IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 23, 2019, accepted March 7, 2019, date of current version April 10, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2908076

Multi-Modal Medical Image Fusion With Adaptive Weighted Combination of NSST Bands Using Chaotic Grey Wolf Optimization

C. S. ASHA[1], SHYAM LAL[2], (Senior Member, IEEE), VARADRAJ PRABHU GURUPUR[3], (Senior Member, IEEE), AND P. U. PRAKASH SAXENA[4]

[1]Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal 574115, India
[2]National Institute of Technology Karnataka, Mangalore 575025, India
[3]Department of Health Management and Informatics, University of Central Florida, Orlando, FL 32816, USA
[4]Department of Radiotherapy and Oncology, Kasturba Medical College, Manipal Academy of Higher Education, Manipal 575001, India

Corresponding author: C. S. Asha (asha.cs@rediffmail.com)

ABSTRACT Recently, medical image fusion has emerged as an impressive technique in merging the medical images of different modalities. Certainly, the fused image assists the physician in disease diagnosis for effective treatment planning. The fusion process combines multi-modal images to incur a single image with excellent quality, retaining the information of original images. This paper proposes a multi-modal medical image fusion through a weighted blending of high-frequency subbands of nonsubsampled shearlet transform (NSST) domain via chaotic grey wolf optimization algorithm. As an initial step, the NSST is applied on source images to decompose into the multi-scale and multi-directional components. The low-frequency bands are fused based on a simple max rule to sustain the energy of an individual. The texture details of input images are preserved by an adaptively weighted combination of high-frequency images using a recent chaotic grey wolf optimization algorithm to minimize the distance between the fused image and source images. The entire process emphasizes on retaining the energy of the low-frequency band and the transferring of texture features from source images to the fused image. Finally, the fused image is formed using inverse NSST of merged low and high-frequency bands. The experiments are carried out on eight different disease datasets obtained from Brain Atlas, which consists of MR-T1 and MR-T2, MR and SPECT, MR and PET, and MR and CT. The effectiveness of the proposed method is validated using more than 100 pairs of images based on the subjective and objective quality assessment. The experimental results confirm that the proposed method performs better in contrast with the current state-of-the-art image fusion techniques in terms of entropy, VIFF, and FMI. Hence, the proposed method will be helpful for disease diagnosis, medical treatment planning, and surgical procedure.

INDEX TERMS NSST, grey wolf optimization, chaotic function, image fusion, MRI, PET, SPECT.

I. INTRODUCTION

A variety of medical imaging technologies available today capture images with distinct modalities which focus on organ or tissue information. Generally, multi-modal medical images provide complementary details of various human organs of the body. High-resolution images are acquired through computed tomography (CT) imaging technique that captures skeletal structures, and third-party implants, while the magnetic resonance (MR) imaging detects the internal body structures such as abdomen, liver, pancreas and other soft tissues, however less capable in capturing dense structures like CT. Besides, the functional imaging techniques such as Positron Emission Tomography (PET) and Single-Photon Emission CT (SPECT) are employed to capture the metabolism information of an organism, which assists in treating the tumor detection and vascular disease diagnosis. The functional images are often described using pseudocolor, and its low spatial resolution often inhibits the useful analysis. Hence, for accurate diagnosis, physicians observe the multi-modal medical images separately, which causes inconvenience of investigating the images in terms of accuracy and

The associate editor coordinating the review of this manuscript and approving it for publication was Yong Yang.

time. The solution to this problem can be accomplished using image fusion which procures the relevant information from individual images and combines them such that the information perceived by human or machine is better compared to the source images and is helpful for disease diagnosis, treatment planning, and surgical procedure as well.

Several fusion methods are reviewed in recent survey papers [1], [2]. Fusion techniques have been practiced in diverse fields such as fusing multi-modal medical images, multi-focus images, panchromatic satellite images, infrared-visible images. Among these, medical image fusion remains special, as it aids the physician to analyze the images captured in different imaging modes in a single image, saving time and increasing the overall quality of an image. Typically, medical image fusion is performed in the spatial and frequency domain. Most of the medical image fusion techniques are performed in a multi-scale domain in which source images are first converted to multi-scale components. Further, the multi-scale coefficients are merged based on a variety of fusion rules. Finally, inverse transform techniques are applied to the merged coefficients to get the fused image. Laplacian pyramid based fusion approaches fuse the pyramid components [3], Li *et al.* [4] fuse the wavelet coefficients, merging of dual-tree complex wavelet components in [5], nonsubsampled contourlet transform (NSCT) coefficients in [9], nonsubsampled shearlet transform (NSST) in [10]. Among various multi-scale component based methods, NSCT and NSST methods outperform the other approaches. Furthermore, the selection of fusion strategies plays a crucial role in merging the low and high-frequency coefficients. Several fusion strategies are employed in literature such as averaging, maximum selection rule, etc. Numerous methods have revealed that multi-scale fusion approach can be enhanced by incorporating different fusion rules. Du et al. in [11] suggested a technique based on local Laplacian filtering to obtain multi-scale components and fusion rule based on information of interest. Recently, convolution neural network [12] based fusion approach is intended to fuse Laplacian pyramids.

A biological neural network model has been inspired by cat's visual cortex popularly known as a pulse-coupled neural network (PCNN) [14], which is extensively used in image processing applications including segmentation, feature extraction, region growing, noise reduction, image enhancement and applied as a tool in image fusion. PCNN has been utilized in fusion literature to extract the information of the source images based on activity level. Further, PCNN is expanded to multi-scale coefficients to retrieve the details of an image. Several PCNN based fusion methods have been addressed in the area for multiscale components such as NSCT, NSST [15]. An effective fusion strategy based on fuzzy adaptive PCNN has been employed for NSCT coefficients in RPCNN [13]. Typically, PCNN models strongly depend on the parameter values, which are set manually as well as automatically. The quality of a fused image relies on the optimized parameters selected for PCNN settings, which are manually fixed as constants based on the empirical result. Thus, several techniques attempted to obtain automatic setting of parameters [14] using metaheuristic optimization approaches. Ming et al. in [15] proposed a method based on selecting automatic adaptive parameters for PCNN and achieved better results compared to other fusion techniques qualitatively.

Recently, nature-inspired algorithms have been adopted in fusion methods to combine the source images [16], [17], [19], which finds weight coefficients to merge the multiscale components. In most of the cases, frequency domain approaches such as Discrete Fourier Transform or wavelet transform have been utilized with mutual information as the fitness function.

In the fusion process, change of color, appearance, and noise artifacts are considered as annoying factors. Simple maneuvering often creates undesirable effects at the boundaries of regions extracted from source images that may lead to an inaccurate diagnosis. This paper proposes a simple and effective method for multi-modal medical image fusion in NSST domain that attains superior quality compared to recent fusion approaches. We employ an adaptively weighted combination of shearlet components based on an effective transfer of textural details from source images to the fused image. The optimized weights are obtained using chaotic grey wolf optimization (CGWO) algorithm by minimizing the $L1$ distance between source images and combined image as the fitness function. Thus, the weights are considered as adaptive and computed for each high subbands independently. The CGWO algorithm computes weights with fast convergence speed with a lesser number of iterations. We also fuse low-frequency coefficients to preserve the energy of source images in the fused image using simple maximum fusion rule. The proposed approach experiments on different modalities such as MR-T1 and MR-T2, MR and SPECT, MR and PET, MR and CT containing more than 100 image pairs. The set of images comprise of various modalities and diseases as mentioned in Brain Atlas [26]. Besides, the proposed fusion method is compared with nine recent medical image fusion approaches using metrics such as entropy (EN), standard deviation (SD), feature mutual information (FMI), visual information fidelity for fusion (VIFF) and Edge Intensity (EI). None of the recent algorithms rank 1 in all the parameters. The prominent contributions of the proposed fusion method in NSST domain are stated as follows,

- The paper utilizes a fusion method by combining NSST and Chaotic Grey Wolf Optimizer for effective combination of features from source images.
- Various chaotic maps are tested for fusing multi-modal medical images.
- A simple max fusion rule is employed for merging low frequency components and adaptively weighted fusion technique to fuse the high frequency components based on minimization of L1 distance between fused image and source images as the fitness function.
- More than 100 pairs of medical image data are tested unlike the works proposed in the past where only few

image data are tested and compared with the recent existing medical fusion methods.

The work is explained in 5 sections as follows. Section II provides background work on NSST for multi-scale and multi-direction decomposition and optimization technique using CGWO algorithms. Section III explains the materials and methods of the proposed multi-modal medical image fusion process followed by experimental discussions in section IV. Finally, conclusion is presented in the section V.

II. BACKGROUND

A. NONSUBSAMPLED SHEARLET TRANSFORM (NSST)

To overcome the drawbacks of wavelet transforms in expressing signals with edges, shearlet transform [21] has been proposed. Contrary to existing multi-scale approaches such as the pyramid, wavelet, edgelet, and curvelet, shearlet captures the texture portions of an image at multi-scale and multi-directions efficiently. In this approach, directional filters exercised in contourlet of NSCT are replaced by the shearing filters, in which it contributes no constraints on the number of directional filters. The success of sheering filters is that the inverse shearlet transform needs the computation of summation. Despite the success of shearlet domain, merging of bands remain as open problem. Several approaches employ fusion rules to attain the optimum fusion outcome. Majority of techniques employ averaging to merge low frequency band, leading to decrease in overall energy of the image. Our method exploits simple addition property of shearlet bands as inverse transform, and hence the weighted combination of high-frequency bands is directly performed based on minimizing $L1$ distance between source and fused images in the shearlet domain itself. In addition, nonsubsampling method is utilized to overcome shift-invariance property of standard shearlet transform. Thus, for dimension n=2, affine systems with composite dilations $M_{XY}(\psi)$ are of the form given by,

$$\{\psi_{l,k,m}(y) = |\det X|^{\frac{l}{2}}\psi(Y^k X^l y - m : l, k \in Z, m \in Z^2\} \tag{1}$$

where, $\psi \in L^2(R^2)$, both X and Y are invertible matrices of size 2×2 with $|\det Y| = 1$. X and Y respectively denotes anisotropic dilation matrix and shear matrix with scale (l), direction (k) and shift (m). The components of system is composite wavelets if Parseval frame is satisfied for $M_{XY}(\psi)$, i.e.,

$$\sum_{j,k,l} \left|\langle f, \psi_{j,l,k}\rangle\right|^2 = \|f\|^2, \quad \forall f \in L^2(R^2) \tag{2}$$

The dilation matrices X^l corresponds to scale transformations, whereas Y^k corresponds to geometric transformations such as shear and rotations. This allows to construct Parseval frames at various scales, locations and orientations. Shearlet forms a special example of composite wavelets with anisotropic dilation matrix X_0 and shear matrix Y_0 given by,

$$X_0 = \begin{bmatrix} a & 0 \\ 0 & \sqrt{a} \end{bmatrix} \tag{3}$$

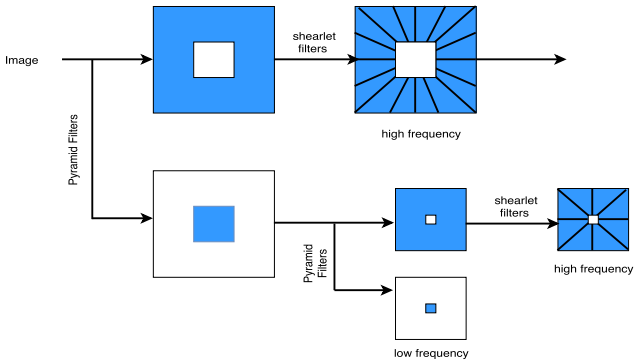

FIGURE 1. Shearlet decomposition of two level.

with $a \geq 0$, and

$$Y_0 = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \tag{4}$$

$s \in R$, $M_{XY}(\psi)$ is called shearlet which is a collection of wavelets at different scales. The matrix X_0 contains scale transform and matrix Y_0 contains geometric transformation.

The discrete shearlet transform consists of two steps. i.e, multi scale division and direction localization. In the Fig. 1, initially Laplacian pyramid is applied to an image to decompose into low and high frequency components. The direction localization is achieved via shear filter. Generally, in Laplacian pyramid filters, $L + 1$ subbands are taken with the same size as the input image for L level decomposition. In total, L high-frequency bands and one low-frequency band are achieved using shearlet transform. For each decomposition level, a shearlet filters are applied to obtain the directional representations of the corresponding group. Thus, two steps are involved in this process. The first step generates the nonsubsampled pyramid, in which each scale is obtained by accumulating the filtered results of the decompositions at all directions using shearing filters. The second step is to reconstruct the image captured by nonsubsampled pyramid from coarse to fine using the reconstruction filters. The NSST decomposition using pyramid filters and shearlet filters is depicted in Fig. 1. More details about NSST are found in [21]. Several state-of-the-art methods have adopted NSST for producing multi-scale components [10], and it is employed in the present work due to its benefits such as multi-scale, multi-direction, and shift invariance property, moreover its addition property as inverse transform.

B. CHAOTIC GREY WOLF OPTIMIZATION (CGWO)

Grey Wolf Optimizer (GWO) is the state-of-the-art meta heuristic optimization algorithm motivated by the social hunting behavior of grey wolves [22]. **GWO has been tested using 29 benchmark functions, and produced competitive results over the state-of-the-art meta-heuristics. In addition, optimized parameters for image fusion applications [16], [17], [19] were computed using GWO. However, GWO often fails to find global optimum at less**

number of iterations, hence increasing the convergence rate. Recently, Chaotic GWO [24] has been investigated for optimization problems endeavoring acceleration of convergence speed. Chaos is a deterministic random like signal with non periodic, no converging and bounded. Variety of chaos has been tested in the optimization field. The chaotic maps improves the convergent rate to land at the global optimum location more quickly. Due to excellent performance and faster convergence of CGWO, we employ in the proposed method to find the optimum weights to merge the source images. GWO simulates the hunting and prey searching behavior of grey wolves. It assumes that grey wolves have the four level of social hierarchy such as *alpha* (α) at first level, *beta* (β) at second level, *delta* (δ) at third level, *omega* (ω) at the last level. α wolves have managing capacity of whole pack of grey wolves and hence, act as leader wolves. In addition to this, it also regulates the hunting process taking judgments such as hunting, sleeping place, maintaining discipline, waking etc. β wolf reinforces commands set by α and takes the feedback from other wolves to transfer it to the leader wolves. They become leader if α wolf passes away. ω wolves are the last one to eat the prey. They manage the safety and integrity of the entire wolf union. *delta* wolves watch in the boundaries of territory and safeguard the pack. Mainly GWO has three phases that include; tracking/chasing/approaching the prey, pursuiting/encircling/harassing and attacking the prey. The mathematical representation of encircling is given by,

$$\overrightarrow{D} = \left| \overrightarrow{C} \cdot \overrightarrow{Y_p(t)} - \overrightarrow{Y}(t) \right| \tag{5}$$

$$\overrightarrow{Y}(t+1) = \overrightarrow{Y_p(t)} - \overrightarrow{A} \cdot \overrightarrow{D} \tag{6}$$

where t represents the current iteration, $\overrightarrow{A}$, $\overrightarrow{C}$ are the coefficient vectors, $\overrightarrow{Y_p}$ is the position vector of prey, $\overrightarrow{Y}$ is the position of wolf. While hunting, prey is encircled by the wolves guided by alpha. In addition, beta and delta wolves participate in hunting. In simulation, alpha, beta and delta know the location of prey. Hence, first 3 best solutions are saved and other search agents are suggested to update their positions. i.e, D_α, D_β, D_δ denotes the distances of α, β, δ from wolf Y. It is computed as,

$$\begin{aligned} \overrightarrow{D}_\alpha &= \left| \overrightarrow{C}_1 \cdot \overrightarrow{Y}_\alpha - \overrightarrow{Y} \right| \\ \overrightarrow{D}_\beta &= \left| \overrightarrow{C}_2 \cdot \overrightarrow{Y}_\beta - \overrightarrow{Y} \right| \\ \overrightarrow{D}_\delta &= \left| \overrightarrow{C}_3 \cdot \overrightarrow{Y}_\delta - \overrightarrow{Y} \right| \end{aligned} \tag{7}$$

$$\begin{aligned} \overrightarrow{Y}_1 &= \overrightarrow{Y}_\alpha - \overrightarrow{A}_1 \cdot \overrightarrow{D}_\alpha \\ \overrightarrow{Y}_2 &= \overrightarrow{Y}_\beta - \overrightarrow{A}_2 \cdot \overrightarrow{D}_\beta \\ \overrightarrow{Y}_3 &= \overrightarrow{Y}_\delta - \overrightarrow{A}_3 \cdot \overrightarrow{D}_\delta \end{aligned} \tag{8}$$

For GWO

$$\overrightarrow{a} = 2 - t * \left(\frac{2}{maxiter} \right) \tag{9}$$

Algorithm 1 Pseudocode of Chaotic Grey Wolf Optimization Technique

1: **initialize** the population $\overrightarrow{Y}_i$ of grey wolves randomly and start the generation counter t
2: **initialize** the sinusoidal chaotic maps y_0 randomly.
3: **initialize** the parameters $\overrightarrow{A}$, $\overrightarrow{C}$, $\overrightarrow{a}$
4: **compute** the fitness function as in Equation (12) of each wolf, $\overrightarrow{Y}_\alpha$ denotes the best wolf, $\overrightarrow{Y}_\beta$ represents the second best, $\overrightarrow{Y}_\delta$ third best wolf.
5: **while** $t \leq maxiter$ **do**
6: **sort** the population as per fitness value
7: **update** chaotic value according to chaotic map as in Equation (16)
8: **for** each search agent **do**
9: **update** the position of current wolf as $\overrightarrow{Y}(t+1) = \frac{\overrightarrow{Y}_1 + \overrightarrow{Y}_2 + \overrightarrow{Y}_3}{3}$.
10: **end for**
11: **update** the parameters $\overrightarrow{a}$, $\overrightarrow{A}$, $\overrightarrow{C}$ as in Equation (9) and (10).
12: **compute** the fitness of all wolves and update $\overrightarrow{Y}_\alpha$, $\overrightarrow{Y}_\beta$, $\overrightarrow{Y}_\delta$.
13: **replace** the worst wolf with the best wolf
14: **increment** the iteration number $t = t + 1$
15: **end while**
16: **return** $\overrightarrow{Y}_\alpha$

where t denotes the iteration number and *maxiter* denotes the maximum number of iterations.

$$\begin{aligned} \overrightarrow{A} &= 2\overrightarrow{a} \cdot \overrightarrow{r}_1 - \overrightarrow{a} \\ \overrightarrow{C} &= 2 \cdot \overrightarrow{r}_2 \end{aligned} \tag{10}$$

$$\overrightarrow{Y}(t+1) = \frac{\overrightarrow{Y}_1 + \overrightarrow{Y}_2 + \overrightarrow{Y}_3}{3} \tag{11}$$

The parameter values $\overrightarrow{a}$, $\overrightarrow{A}$ and $\overrightarrow{C}$ are computed. The value of $\overrightarrow{a}$ is linearly decreased from 2 to 0 at the end of iterations (it depends on maximum number of iterations) [22] for GWO. $\overrightarrow{r}_1$ and $\overrightarrow{r}_2$ represents the random vectors in the range (0,1). The grey wolves finish attacking when it stops moving towards it. The vectors represent the process that wolves able to reach the point between wolf and prey. The parameter $\overrightarrow{a}$ controls the GWO activity. $\overrightarrow{C}$ puts more weight on prey and makes it challenging for wolves to reach the prey. Position of all the wolves get updated in every iteration.

GWO is practiced in several applications, to solve the optimization problem with suitable alterations. In order to improve the efficiency and convergence speed, CGWO algorithm [24] has been realized by exploiting the chaos. In short, chaos is a deterministic random like method in a nonlinear dynamic system. It is nonperiodic, nonconverging and bounded. Several chaotic maps do exist such as tent, sinusoidal, logistic, circle, piecewise, cubic, chebychev, etc., with different mathematical equations. They improve the search

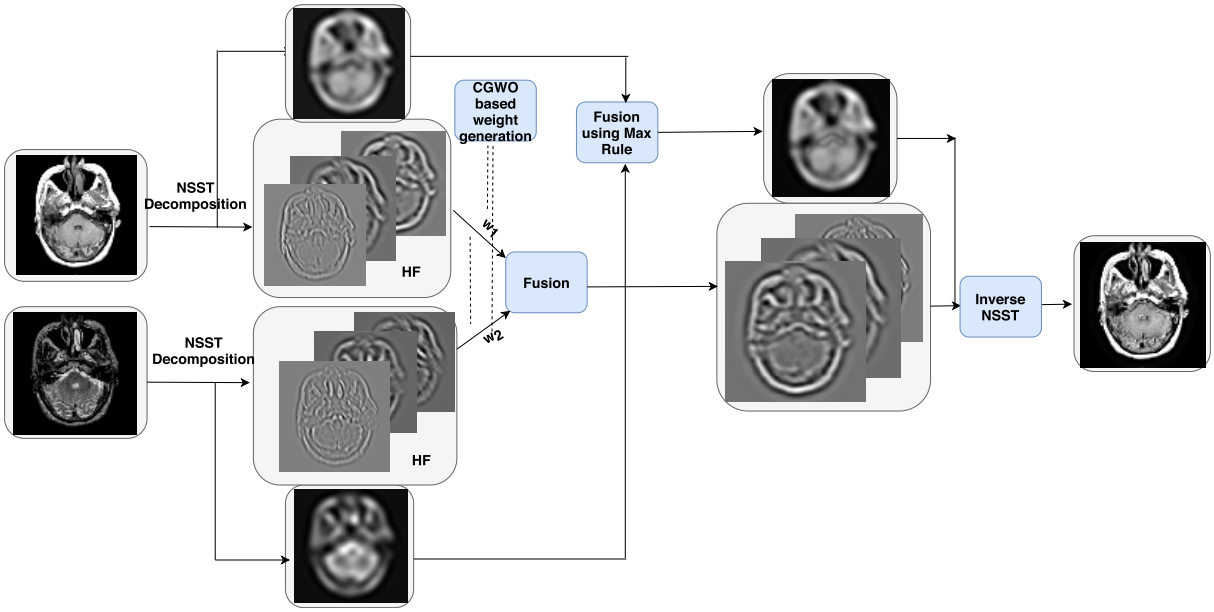

FIGURE 2. Block diagram of the proposed method (NSST-CGWO) for grayscale image fusion.

space globally, Chaotic maps affect the convergence rate of optimization algorithms. Accordingly, the chaotic maps take any value in the range (0, 1) initially for $\vec{a}$, $\vec{A}$, $\vec{C}$. The objective function of all the grey wolves is evaluated and arranged according to their fitness. The topmost wolf is considered to be α wolf, second top as β and third as δ wolf. Best wolf is represented as α wolf, keeps updating the position. The parameter values change in every iteration, fitness of α is considered as the best solution obtained. The objective function is the function whose optimal solution needs to be computed in the search space. Entire work in the proposed method focuses on retaining the structural details and energy of source images. Accordingly, the fitness function is chosen as the minimization problem of $L1$ norm between fused image and source images. It is hence given by,

$$f_{obj} = \min \sum \left| H_F^{l,k} - H_A^{l,k} - H_B^{l,k} \right| \tag{12}$$

where, $H_F^{l,k}$ represents the fused high frequency bands in l^{th} scale and k^{th} direction. $H_A^{l,k}$ and $H_B^{l,k}$ denote high frequency bands of source images A and B respectively in l^{th} scale and k^{th} direction. This is achieved by CGWO. The details of chaotic grey wolf optimization is provided in algorithm 1.

Algorithm 2 Pseudocode of Proposed Medical Fusion Approach

input: source images A and B
for each source image A and B **do**
  obtain the NSST decomposition to get L and $H^{l,k}$ according to Equation (13).
end for
fuse low frequency band using maximum fusion rule according to Equation (14).
for each image A and B **do**
  for each level $l = 1 : L$ **do**
    for each direction $k = 1 : K$ **do**
      compute weights w_1 and w_2 using CGWO using objective function $f_{obj} = \min \sum \left| H_F^{l,k} - H_A^{l,k} - H_B^{l,k} \right|$
      merge each high frequency band as $H_F^{l,k} = w_1 H_A^{l,k} + w_2 H_B^{l,k}$
    end for
  end for
end for
compute inverse NSST of $L_F, H_F^{l,k}$ to obtain the fused image F using Equation (15).
output the fused image F

III. PROPOSED METHODOLOGY

The proposed multi-modal medical image fusion is named as NSST-GWO, and pseudo code of the fusion method is provided in Algorithm 2. Fig. 2 and Fig. 3 respectively depict the block diagram of the proposed medical image fusion for grayscale and color images. Let A and B represent source images of the same size. The registration of two input images is performed for source images with different size to make equal size. **The goal of registration process is to align the images. It is important while merging the images that focuses to combine corresponding features. This is achieved by employing warping techniques to align the scale of one image with respect to the reference image**. In the proposed method, we consider two images, however, can be easily extended to more number of images. The following subsections explain the detailed fusion process

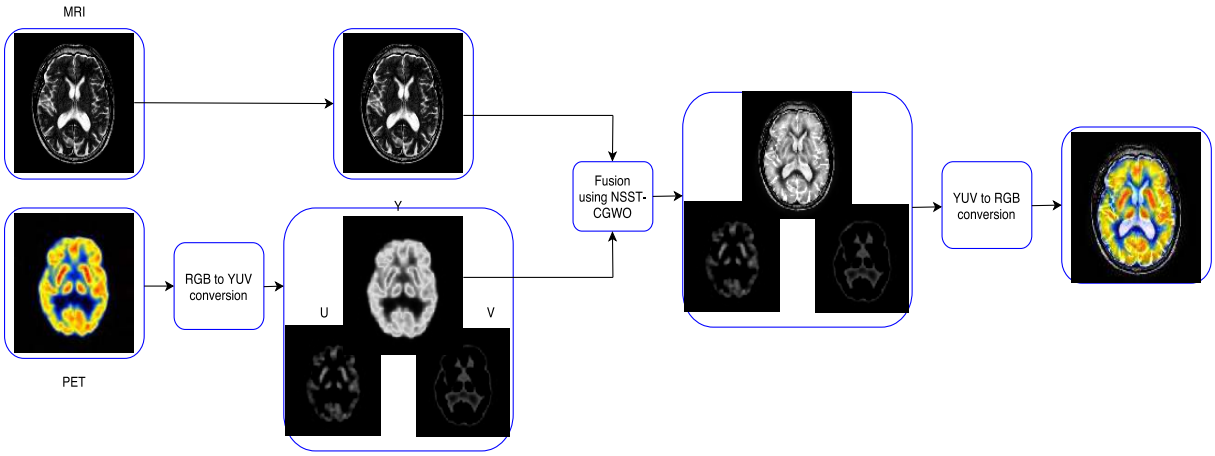

FIGURE 3. Block diagram of the proposed fusion method for color image fusion.

involved in the proposed work. Hence, main stages consist of NSST decomposition, fusion of low-frequency components and high-frequency components using chaotic grey wolf optimization technique, and image reconstruction using inverse NSST.

A. NSST DECOMPOSITION

NSST captures multi directional features such as curves at various scales by employing Laplacian pyramid filters. Sizes of sub-band images obtained via NSST decomposition remain same as original image, which makes easier to fuse at various levels. If A and B are unregistered images, they are registered to have same size. At first, fusion algorithm involves the application of nonsubsampled shearlet transform followed by fusion at each level. Similar to other multi scale approaches, low frequency band of shearlet domain reflects the overall intensities of source images, while high frequency bands represent curves at various scales and directions. A L-level NSST decomposition is performed on two source images A and B to obtain $\{L_A, H_A^{l,k}\}$ and $\{L_B, H_B^{l,k}\}$ respectively [21]. In this, L represents the low frequency band and $H^{l,k}$ denotes high frequency bands at level l and direction k.

$$NSST_A = \{L_A, H_A^{l,k}\} \quad \text{and} \quad NSST_B = \{L_B, H_B^{l,k}\} \tag{13}$$

B. FUSION OF SUBBANDS

Low-Frequency Fusion: The fusion rule for low-frequency bands play a significant role in the quality of the final fusion result. The overall energy of the signal is contained in the low-frequency component, and it is hence essential to retain the energy of the individual in the final fused result. The energy levels of source images alter significantly, as different imaging techniques have been employed. As a result, the intensities of final image decrease due to the conventional averaging procedure. To overcome this issue, we employ the max rule fusion strategy to preserve the energy of an individual in the final fused image, which significantly impacts on the visual perception. The fused low-frequency band is obtained using the following fusion strategy.

$$L_F^{l,k} = \begin{cases} L_A \text{ if } L_A > L_B \\ L_B \text{ if } L_A \leq L_B \end{cases} \tag{14}$$

High-Frequency Fusion: $H_A^{l,k}$ represents the high frequency bands obtained after applying NSST for source image A. In a similar manner, after the NSST application on source image $B, H_B^{l,k}$ is obtained. Each band is adaptively merged as $H_F^{l,k} = w_1 H_A^{l,k} + w_2 H_B^{l,k}$. The weights are obtained using CGWO algorithm by choosing minimization of $L1$ distance between fused image and source images as provided in Equation 12.

C. NSST RECONSTRUCTION

The fused low frequency and high frequency components $\{L_F, H_F^{l,k}\}$ are finally used to reconstruct the fused image F, which is achieved by simple addition of subbands.

$$F = NSST^{-1}(L_F, H_F^{l,k}) \tag{15}$$

IV. EXPERIMENTAL ANALYSIS

A. EXPERIMENTAL SETUP

The proposed fusion method and other existing methods are executed using MATLAB 15a installed on windows 10, 64 bit operating system in a machine with intel(R) core i5-5200U, CPU at 2.20GHz processor with 8GB RAM.

B. DATASETS

The proposed method is verified experimentally, for which more than 100 pairs of multi-modal medical images have been selected. It includes 21 image pairs MR-T1 and MR-T2 from Alzheimers Visual Agnosia, 21 image pairs of MR-T2 and SPECT of Glioma disease, 11 image pairs of MR-T2 and SPECT from metastatic bronchogenic, 11 pairs of MR-T2 and PET from mild Alzheimers, 11 pairs of MR-T2 and SPECT from Motor neuron, 11 image sets of MR-T2 and PET from Normal aging, 11 pairs of CT and

MR-T2 from Sarcoma, 7 image pairs of MR-T2 and SPECT from Subacute stroke. All of these image pairs are downloaded from the Whole Brain Atlas [26] dataset available publicly which is created by Harvard Medical School. They have been extensively used in medical image fusion literature. All the image pairs with same size are fused directly, while image pairs with different sizes are carefully registered before fusion process. Consequently, the size of all images is made equal to 256×256.

C. COMPARISON WITH THE STATE-OF-THE-ART

The proposed fusion method NSST-CGWO is compared with nine recent state-of-the-art fusion techniques. These include NSST-PAPCNN [15], CNN [12], LLF-IOI [11], GFF [7], LPSR [8], NSCT-RPCNN [13], HMW-GWO [16], MMIF-NSCT [20], NSST-SF-PCNN [29]. The methods have been proposed in the last few years and shown significant improvement over the same database. The implementations of majority of mentioned state-of-the-art methods have been made available publicly. The parameters are set to the default values as provided by the respective papers.

D. PARAMETER SETTINGS

In the proposed method the following parameters are selected for NSST. L decomposition levels has been used for NSST. The performance of algorithm remain saturated for higher decomposition levels [15]. Hence, we have set the number of decomposition levels as $L = 4$, with number of directions at each level is chosen as 16, 16, 8, 8. The number of directions is normally chosen from finer scale to coarse scale. Parameters used for CGWO are as follows: the population size of grey wolves is taken as 30, and 20 iterations were carried out totally. For CGWO, the parameters $\overrightarrow{r}_1$ and $\overrightarrow{r}_2$ are randomly chosen. The value of $\overrightarrow{a}$ changes over the iterations linearly. Numerous chaotic maps exist, and are applied in the proposed study. They include tent, sinusoidal, singer, logistic, piecewise, iterative, Gaussian, circle, chebychev. From the set of experiments conducted using various chaotic maps, we conclude that these maps produce similar results in terms of qualitative and quantitative metrics. However, based on time criteria, sinusoidal has been found to perform best among several chaos, it is hence adopted in the present work. The map equation of sinusoidal chaotic map is given by,

$$y_{t+1} = a y_t^2 sin(\pi y_t) \tag{16}$$

The chaotic function variable $a = 0.5$ is used. Table 1 lists the different chaotic maps used to test the proposed method. Several parameters such as edge, entropy, standard deviation, Peilla metric, VIFF, FMI and time are used as criteria. It has been observed that, all chaotic maps produce near values. Hence, the fusion process produces similar subjective and objective outcomes irrespective of maps used. However, we have chosen sinusoidal map for the whole set of experiments as it leads to minimum time consumption.

From the set of experiments, we have selected 20 iterations for optimizations technique, as convergence is attained in less number of iterations. Thus, the weights for fusion is achieved automatically via CGWO approach for every high frequency subbands.

TABLE 1. Various chaotic maps applied to the fusion problem and compared in terms of quantitative metrics, sinusoidal is chosen as it produces same results in less time.

	EI	EN	SD	Peilla	VIFF	FMI	Time
Tent map	106.321	5.564	88.018	0.813	0.732	0.745	67.000
Sinosoidal map	106.064	5.565	87.995	0.813	0.732	0.745	**62.092**
Singer map	106.196	5.557	88.046	0.813	0.733	0.746	67.239
logistic map	106.292	5.569	87.991	0.813	0.731	0.746	68.148
Piecewise map	106.175	5.560	88.031	0.813	0.732	0.745	73.306
Iterative map	106.086	5.573	87.961	0.813	0.732	0.745	77.994
Gaussian map	106.283	5.565	88.008	0.813	0.732	0.745	75.685
Circle map	106.195	5.556	88.047	0.813	0.733	0.745	68.670
Chebyshev map	106.289	5.567	88.009	0.813	0.732	0.745	65.622

E. EVALUATION METRICS

To perform a quantitative analysis of the proposed method, six well-known metrics have been selected, that include the edge intensity (EI), standard deviation (SD), entropy (EN), feature mutual information (FMI) [23], Piellas structure similarity [27] based metric, and the visual information fidelity fusion (VIFF) [28] based on human visual perception based metric. Edge intensity measures the strength of edges in a fused image. Typically, SD measures the degree of spread of gray value of all pixels from the mean gray value.

$$SD = \sqrt{\sum_{x=1}^{M}\sum_{y=1}^{N}(F_{xy} - \overline{F_{xy}})^2} \tag{17}$$

$\overline{F_{xy}}$ denotes the mean of fused image. Hence, it refers to the overall contrast of a fused image also measures the quality of an image, while entropy EN reflects the information content present in the fused image. Entropy has been used in medical information system [18], and is defined as

$$EN = \sum_{i=0}^{255} p_i \ln\left(\frac{1}{p_i}\right) \tag{18}$$

where p denotes the probability mass function. Feature Mutual Information (FMI) computes the amount of feature (information) transfer from source images to the fused image. The feature includes gradient map of an image provides information about texture, edge strength, and contrast. More information on FMI is detailed in [23]. The Peilla metric [25] mainly measures the structural similarity between the combined image and source images by simultaneously addressing coefficient correlation, illumination, and contrast. In this, local saliences of source images A and B are respectively computed as $s(A\,|w)$ and $s(B\,|w)$. Local weights with respect to window w is calculated as $\lambda(w) = \frac{s(A|w)}{s(A|w)+s(B|w)}$. Thus, Peilla metric is given by,

$$\frac{1}{|W|}\sum_{w\in W}\lambda(w)Q_o(A, F\,|w) + (1-\lambda(w))Q_o(B, F\,|w) \tag{19}$$

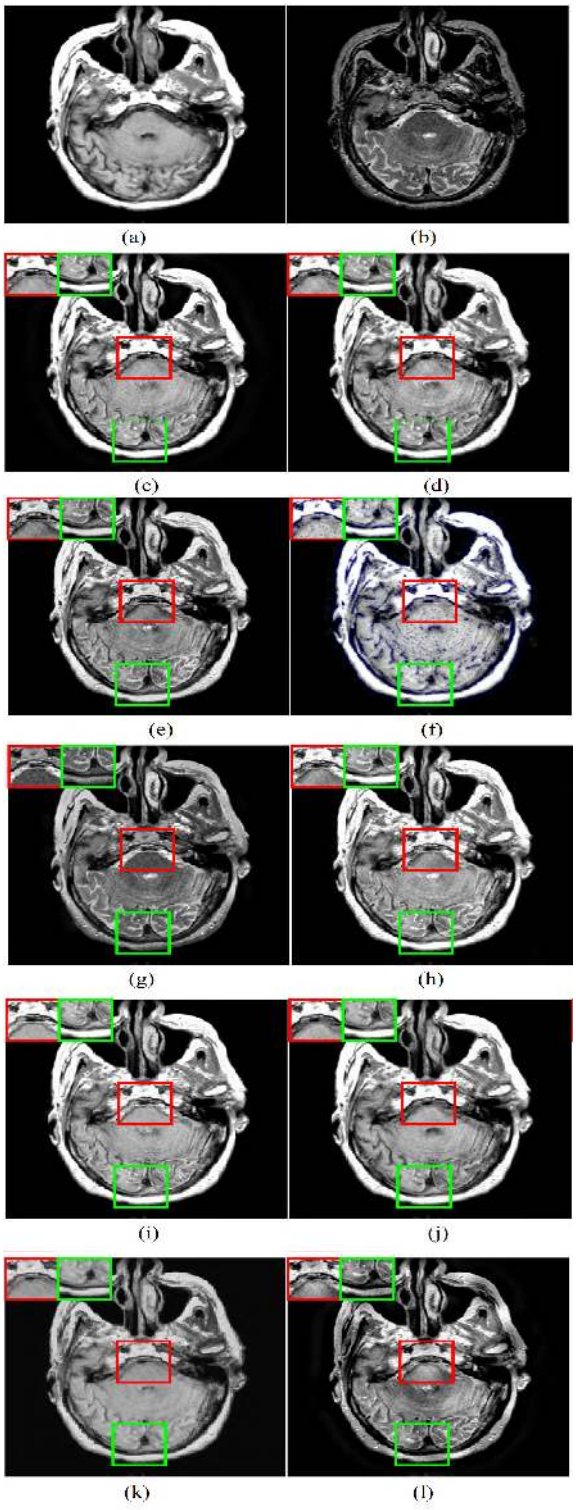

FIGURE 4. One set of Alzheimers Visual Agnosia MR-T1 and MR-T2 fusion results. Two small sections are cropped and are displayed for better comparison. LLF-IOI suffers from noise like artifacts. CNN is unable to transfer the details of MR-T1 as highlighted in the cropped section (shown in red). GFF fails to retain the energy of the MR-T1. HWM-GWO has poor contrast. NSCT-RPCNN, NSCT-SF-PCNN over enhancing the image. The proposed technique and NSST-PAPCNN retain energy and details of original images. (a) MR-T1. (b) MR-T2. (c) NSST-CGWO. (d) NSST-PAPCNN. (e) CNN. (f) LLF-IOI. (g) GFF. (h) LP-SR. (i) NSCT-RPCNN. (j) NSCT-SF-PCNN. (k) HMW-GWO. (l) MMIF-NSCT.

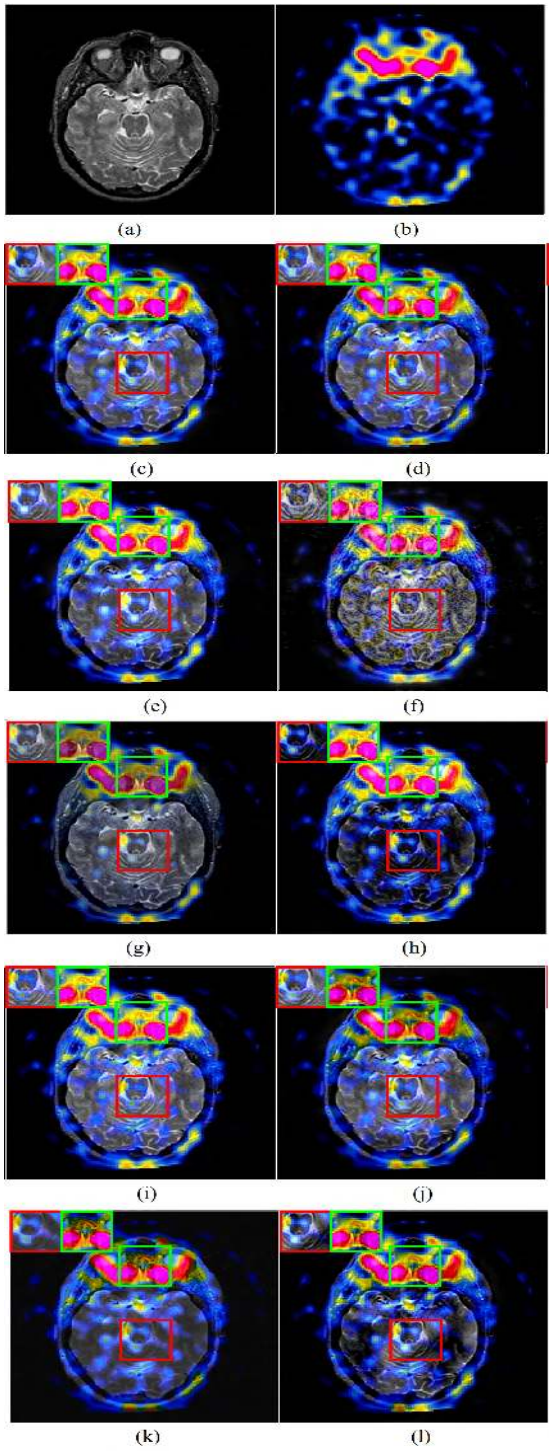

FIGURE 5. One set of MR and SPECT image fusion results of Glioma MR-T2 and SPECT. Two small parts are cropped and depicted individually for better comparison. LLF-IOI introduces noise in the fused image. Frequently, GFF fails to preserve the color. LP-SR and MMIF-NSCT appear to have more contrast, however, disappoints to maintain the structural details of MR image. HMW-GWO has poor contrast and fails to extract features from the original MR image. The proposed method preserves energy and details without over-enhancing the images. (a) MR-T2. (b) SPECT. (c) NSST-CGWO. (d) NSST-PAPCNN. (e) CNN. (f) LLF-IOI. (g) GFF. (h) LP-SR. (i) NSCT-RPCNN. (j) NSCT-SF-PCNN. (k) HMW-GWO. (l) MMIF-NSCT.

TABLE 2. Mean quality metrics on dataset 1 containing 21 images of Alzheimers Visual Agnosia MR-T1 and MR-T2.

	EI	EN	SD	Peilla	VIFF	FMI	Time
NSST-CGWO	100.1221	5.595619	89.11069	0.807019	0.626439	0.77575	69.77725
NSST-PAPCNN	98.51773	5.10073	89.0813	0.807736	0.623641	0.771409	20.29125
CNN	99.13093	4.800058	80.15856	0.808101	0.586661	0.77592	27.2359
LLF-IOI	101.4195	5.054626	94.11413	0.729113	0.591892	0.762698	285.2405
GFF	86.69559	4.943955	66.88758	0.781085	0.443581	0.79794	0.200672
LP-SR	101.5649	5.152757	88.71107	0.801864	0.621439	0.778375	0.053631
NSCT-RPCNN	94.98227	4.868157	90.80067	0.80217	0.605937	0.775401	28.85777
NSCT-SF-PCNN	93.24508	5.15404	88.37318	0.773597	0.575343	0.771634	169.5474
HMW-GWO	69.96756	4.911653	74.61593	0.440918	0.551761	0.705906	61.11418
MMIF-NSCT	98.41734	5.42643	74.29391	0.719083	0.478397	0.775712	26.79536

TABLE 3. Mean quality metrics on dataset 2 containing 21 images of Glioma MR-T2 and SPECT.

Row	EI	EN	SD	Peilla	VIFF	FMI	Time
NSST-CGWO	63.73299195	4.92093	69.1879	0.877876	0.473271	0.725986	63.65911
NSST-PAPCNN	66.45909334	4.691092	69.28709	0.896214	0.432208	0.723273	21.25362
CNN	66.06647073	4.847997	69.53531	0.858759	0.506121	0.727685	22.62084
LLF-IOI	76.73959989	5.19965	66.48109	0.771232	0.408405	0.690971	535.6783
GFF	57.13258873	4.431163	56.29078	0.899352	0.351366	0.727705	0.168913
LP-SR	66.18444281	4.582546	65.86989	0.875858	0.503494	0.72711	0.036565
NSCT-RPCNN	66.53802155	4.484517	70.4183	0.913482	0.417673	0.726576	29.52145
NSCT-SF-PCNN	65.36634506	4.573091	65.9112	0.883297	0.394823	0.722209	96.71242
HMW-GWO	47.83768689	4.643838	53.67419	0.312662	0.383375	0.656631	61.06761
MMIF-NSCT	65.18873492	4.25235	58.17726	0.831349	0.411622	0.72243	26.85345

TABLE 4. Mean quality metrics on dataset 3 containing 7 image pairs of Subaccute Stroke MR-T2 and SPECT.

Row	EI	EN	SD	Peilla	VIFF	FMI	Time
NSST-CGWO	66.23847	5.655657	70.41448	0.87516	0.644549	0.723653	60.69195
NSST-PAPCNN	69.18482	5.34024	70.57458	0.905761	0.6328	0.72281	16.59511
CNN	70.15761	5.981621	70.23773	0.779674	0.629496	0.722011	22.36872
LLF-IOI	83.32779	5.216642	65.77462	0.825716	0.618758	0.708409	276.1715
GFF	66.18093	5.128927	57.90847	0.905269	0.566407	0.722291	0.166549
LP-SR	70.33612	5.411814	70.12325	0.893354	0.61332	0.721095	0.037676
NSCT-RPCNN	68.52359	5.285255	71.32167	0.895893	0.61509	0.722408	29.8864
NSCT-SF-PCNN	70.31985	5.282591	69.64315	0.894625	0.570181	0.720569	104.7196
HMW-GWO	53.01508	5.362851	67.20758	0.49145	0.57808	0.667758	58.22293
MMIF-NSCT	68.96626	5.135354	63.82827	0.87151	0.542197	0.720903	27.98385

where Q_o is an image quality index defined in [25]. VIFF is a recently proposed metric that measures the visual information fidelity between the fused image and each of the source images based on the Gaussian Scale Mixture model, the distortion model, and the HVS model [28]. VIFF computes the visual information of every blocks in each subband of the fused image. More details of VIFF measure is provided in [28]. For all the metrics discussed above, larger value refers to the best value quantitatively. However, a large amount can also result due to the presence of abnormality such as the introduction of artifacts, noise, etc. In such cases, both objective and subjective measures are taken into account to grade the fusion methods.

F. RESULT ANALYSIS

In this section, the proposed fusion method (NSST-CGWO) is compared with the other approaches in terms of visual quality (subjective analysis), fusion quantitative metrics and computational complexity. The quantitative metrics of most recent fusion approaches and the proposed method are listed in the Table 2–Table 4. The average score of all the methods has been tabulated for each disease dataset. The highest value is highlighted in red, second highest in green, while third highest in blue color. For each category mentioned before, we have provided one set of fusion results. Furthermore,

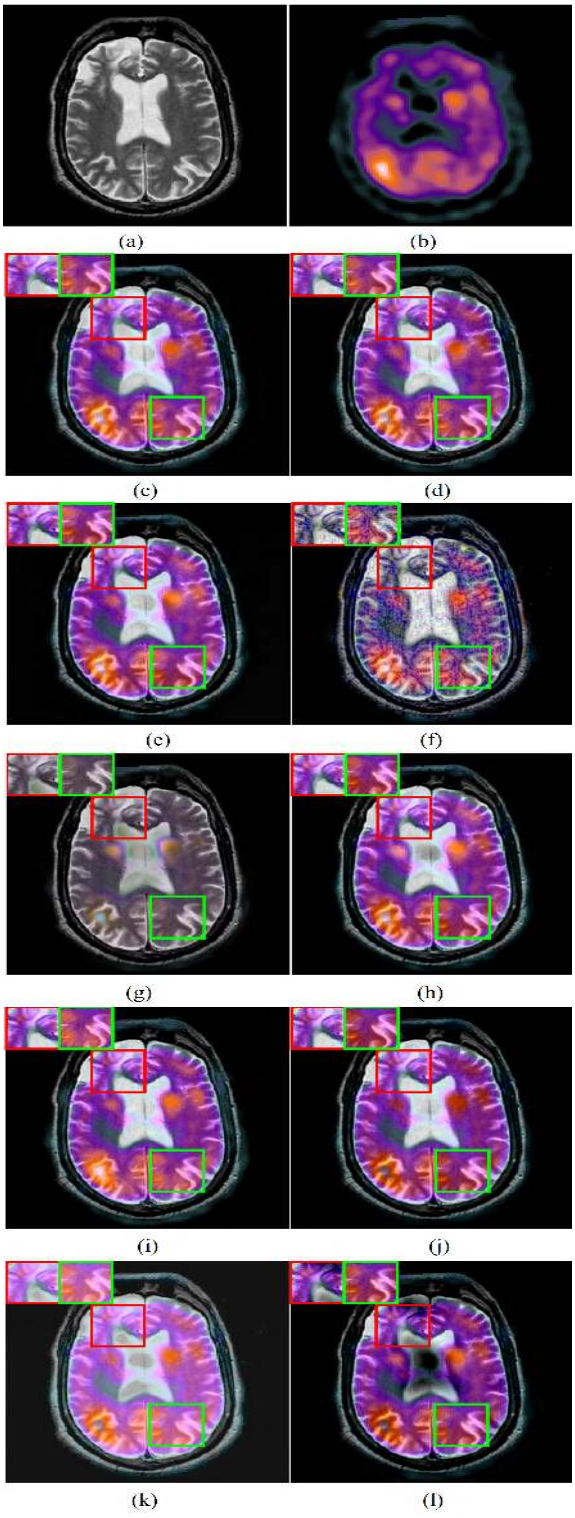

FIGURE 6. One set of Subacute Stroke MR-T2 and SPECT fusion results are depicted. Two cropped areas are shown separately for better visualization. LLF-IOI suffers from heavy, noisy artifacts. GFF fails to transfer the color. HMW-GWO blurs the image. The proposed method is very effective in transferring the energy, texture, and color. (a) MR-T2. (b) SPECT. (c) NSST-CGWO. (d) NSST-PAPCNN. (e) CNN. (f) LLF-IOI. (g) GFF. (h) LP-SR. (i) NSCT-RPCNN. (j) NSCT-SF-PCNN. (k) HMW-GWO. (l) MMIF-NSCT.

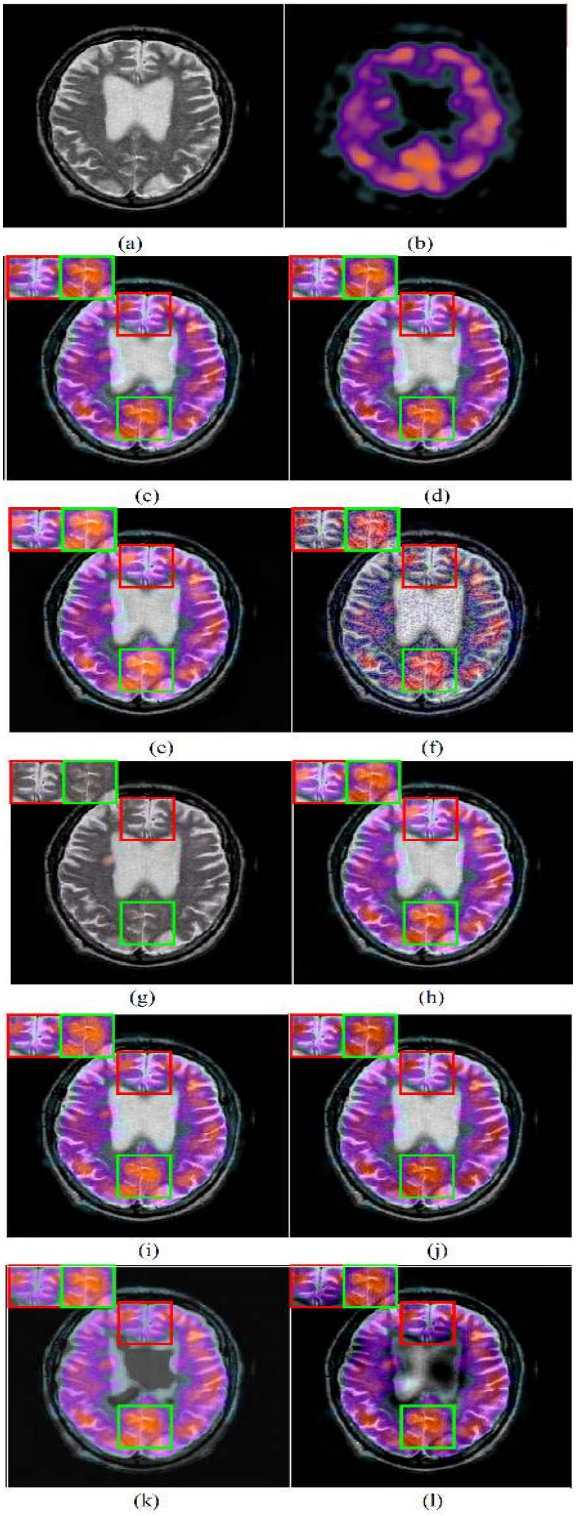

FIGURE 7. One set of Motor Neuron MR-T2 and SPECT fusion results. Two cropped section are shown for better comparison. LLF-IOI introduces noise artifacts to a large extent. GFF cannot preserve the color of the SPECT image. HMW-GWO and MMIF-NSCT introduce serious artifacts in the fused image. The proposed method preserves energy and color. (a) MR-T2. (b) SPECT. (c) NSST-CGWO. (d) NSST-PAPCNN. (e) CNN. (f) LLF-IOI. (g) GFF. (h) LP-SR. (i) NSCT-RPCNN. (j) NSCT-SF-PCNN. (k) HMW-GWO. (l) MMIF-NSCT.

two sub regions are cropped and shown separately to better visualize the images in terms of contrast, color fidelity, energy preservation, etc.

1) MR-T1 AND MR-T2

Fig. 4 depicts the fusion of MR-T1 and MR-T2 using the proposed method and recent fusion methods. Some of the fusion techniques produce unacceptable visual quality due to loss of energy. In addition to that, these methods often fail to transfer the intensity and texture details from source images to the fused image accurately. LLF-IOI introduces the noise artifacts in the fused image, resulting in poor quality. Subsequently, the presence of extraneous noise in LLF-IOI fusion method provides high values for metrics such as SD, edge intensity. Although these methods show better performance considering SD or EI, it severely fails to prove in terms of visual aspect. LP-SR and CNN methods visually outperform other methods, but cannot preserve the similar intensity variation of source images. CNN and GFF methods fail to transfer the information of MR-T2 to the fused image. However, the proposed method retains the energy and details of individual without introducing noisy structure or any artifacts. Table 2 provides the comparison in terms of quantitative metrics for MR-T1 and MR-T2 datasets. NSST-CGWO stands one among the top three positions for the listed metrics. EN metrics is high for the proposed fusion method, which indicates the superior quality of the fused image.

2) MR-T2 AND SPECT

Fig. 5, Fig. 6, and Fig. 7 correspond to the fusion results of MR and SPECT datasets. The color preservation of few methods remains relatively low. As mentioned before, LLF-IOI fails to preserve energy and introduces noise in the final result. Consequently, few functional details are lost. CNN, NSCT-SF-PCNN, NSCT-RPCNN work excellent in terms of preserving functional information of SPECT image. LP-SR and NSCT-MMIF have good contrast, in which color details are preserved highlighting texture information of MR data. HWM-GWO has less contrast, however preserves color and texture information. The proposed method retains the energy and highlights the structural information of source images, while maintaining color features too. The qualitty metrics of the proposed and state-of-the-art fusion method are presented in Table 3, Table 6, Table 5 and Table 4. From the above tables, it is clear that the proposed NSST-CGWO produces the higher values for entropy and VIFF in majority of the datasets.

3) MR-T2 AND PET

Fig. 8, Fig. 9, and Fig. 10 depict the combination of MR and PET image datasets. In MR and PET image fusion, MR contains most of the structural details. All the methods perform well concerning structure preservation. However,

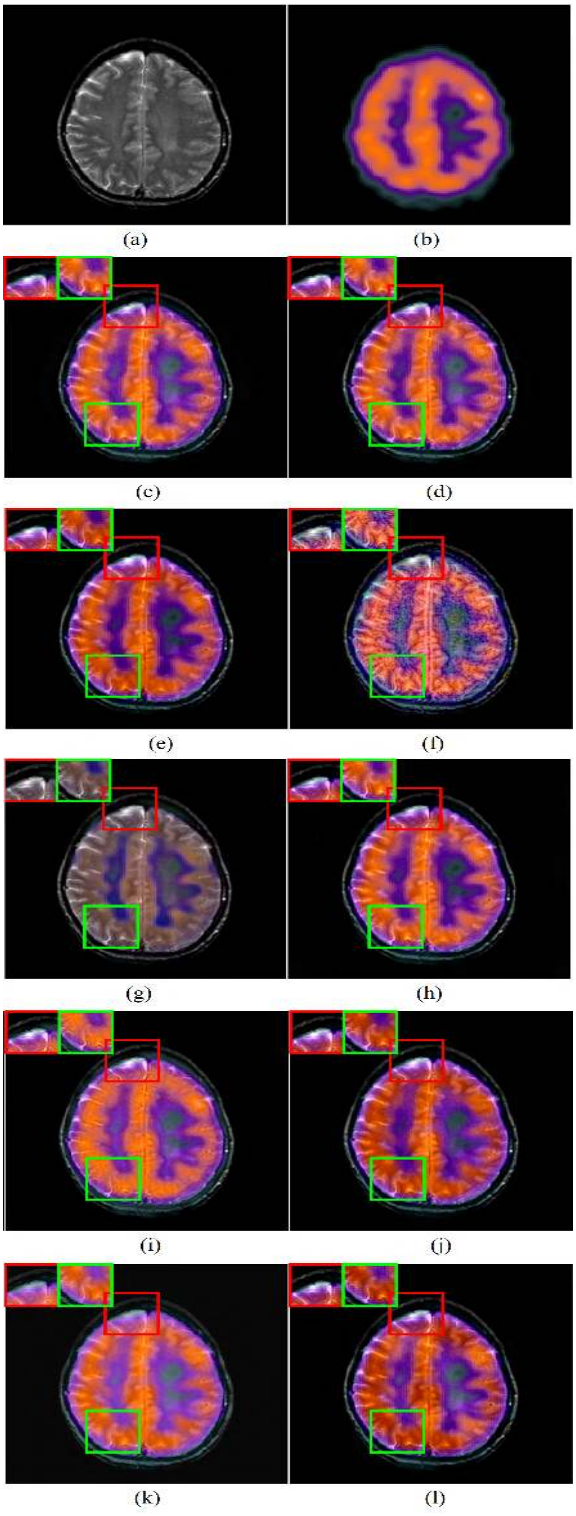

FIGURE 8. One set of MR and SPECT-TC image fusion results of Metastatic Bronchogenic MR-T2 and SPECT-TC. Cropped areas are shown separately for the better view. LLF-IOI severely suffers from noise artifacts. GFF fails to retain the original color. HMW-GWO has poor contrast. MMIF-NSCT appears to have good contrast but fails to maintain originality regarding structural details of the MR-T2 image. The proposed method preserves energy, structural details and color of the source images. (a) MR-T2. (b) SPECT-TC. (c) NSST-CGWO. (d) NSST-PAPCNN. (e) CNN. (f) LLF-IOI. (g) GFF. (h) LP-SR. (i) NSCT-RPCNN. (j) NSCT-SF-PCNN. (k) HMW-GWO. (l) MMIF-NSCT.

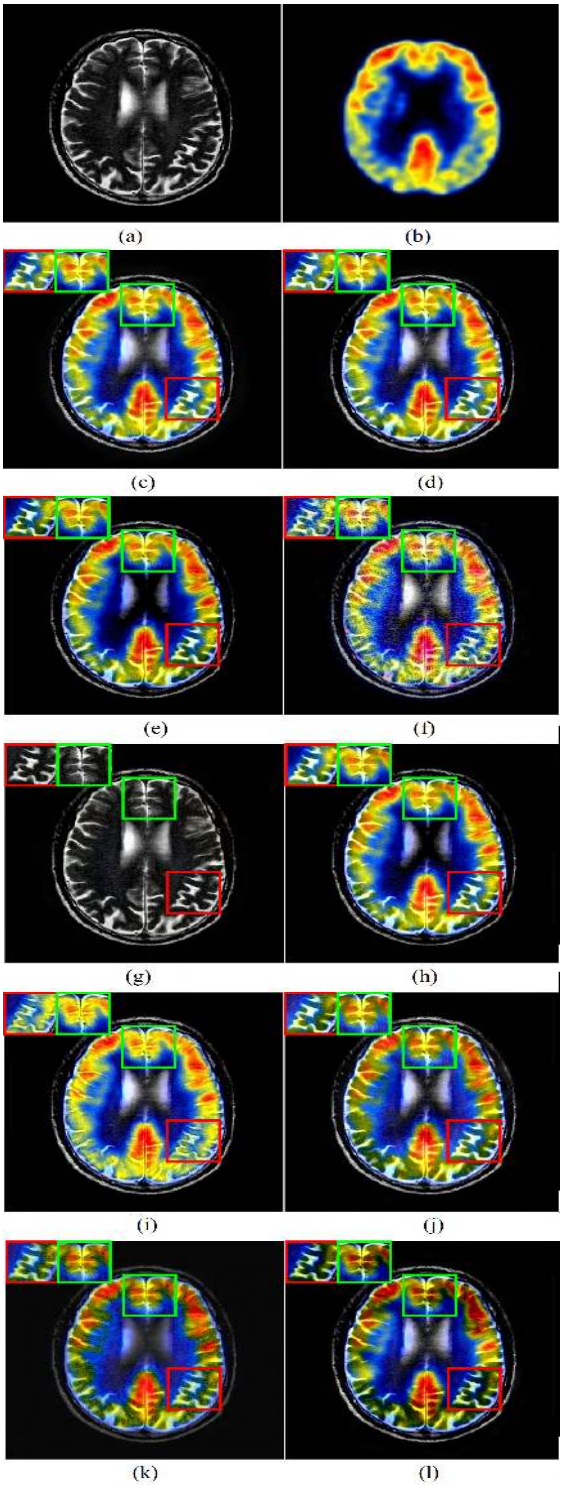

FIGURE 9. One set of images contain Mild Alzheimers MR-T2 and PET-FDG fusion results. Two cropped sections are shown for better visibility. CNN and LP-SR method extracts the details effectively but fails to retain original color. LLF-IOI suffers from the noisy structure. GFF fails to transfer the color to the fused image. NSCT-RPCNN over enhances the image, while NSCT-SF-PCNN changes the original color. HMW-GWO has poor contrast. MMIF-NSCT cannot preserve the details and colors effectively. The proposed method is very efficient in retaining originality in terms of structures and color information. (a) MR-T2. (b) PET-FDG. (c) NSST-CGWO. (d) NSST-PAPCNN. (e) CNN. (f) LLF-IOI. (g) GFF. (h) LP-SR. (i) NSCT-RPCNN. (j) NSCT-SF-PCNN. (k) HMW-GWO. (l) MMIF-NSCT.

TABLE 5. Mean quality metrics on dataset 4 containing 11 image pairs of Motor Neuron MR-T2 SPECT.

Row	EI	EN	SD	Peilla	VIFF	FMI	Time
NSST-CGWO	63.77206	5.488302	72.00948	0.8734	0.60299	0.73646	68.042
NSST-PAPCNN	66.98222	5.017868	72.26349	0.910551	0.59543	0.735415	17.70493
CNN	67.30631	5.804581	72.13826	0.726679	0.59648	0.73483	29.01363
LLF-IOI	79.74608	4.942152	66.33258	0.847538	0.582609	0.710273	304.7409
GFF	64.74273	4.823988	61.28626	0.905704	0.528355	0.734391	0.186961
LP-SR	67.65055	5.167772	72.89983	0.895811	0.584538	0.734355	0.039811
NSCT-RPCNN	66.83511	4.916636	73.02439	0.907613	0.573279	0.734881	31.07037
NSCT-SF-PCNN	68.03386	5.067658	71.57855	0.897168	0.537226	0.733377	92.49899
HMW-GWO	50.66374	4.954057	69.95871	0.445882	0.553355	0.70907	53.41782
MMIF-NSCT	65.86739	4.864586	62.90574	0.858416	0.486092	0.733149	25.83413

TABLE 6. Mean quality metrics on dataset 5 containing 11 image pairs of Metastatic Bronchogenic MR-T2 SPECT-TC.

Row	EI	EN	SD	Peilla	VIFF	FMI	Time
NSST-CGWO	41.15153	4.421158	65.1939	0.885308	0.604808	0.800125	61.2623
NSST-PAPCNN	43.38985	4.146144	65.19403	0.902581	0.59142	0.801842	16.57183
CNN	43.35715	4.217019	62.47311	0.873303	0.546634	0.802647	23.37846
LLF-IOI	53.67473	3.879076	64.00459	0.84957	0.563719	0.74269	288.7614
GFF	38.43036	3.777665	49.14249	0.918397	0.426632	0.805371	0.172415
LP-SR	43.50667	4.239919	63.62232	0.896041	0.559968	0.801556	0.039593
NSCT-RPCNN	40.627	3.969947	66.29403	0.894026	0.558708	0.804319	29.12186
NSCT-SF-PCNN	43.60402	3.837121	61.70119	0.906327	0.456542	0.801254	94.34204
HMW-GWO	35.19113	4.050526	62.28995	0.382975	0.532303	0.738657	59.44768
MMIF-NSCT	44.00492	3.911221	58.0916	0.901282	0.475874	0.801101	26.72351

TABLE 7. Mean quality metrics on dataset 6 containing 11 image pairs of Mild Alzheimers MR-T2 PET-FDG.

Row	EI	EN	SD	Peilla	VIFF	FMI	Time
NSST-CGWO	77.83696	4.868649	77.61502	0.809793	0.483086	0.862083	63.53088
NSST-PAPCNN	81.45239	4.428142	78.14026	0.8462	0.473967	0.864023	17.13851
CNN	74.63074	3.626805	70.59262	0.824848	0.422798	0.860798	24.1937
LLF-IOI	83.26936	4.422963	75.83856	0.81779	0.435657	0.833555	306.1852
GFF	80.25853	4.453393	55.96947	0.859271	0.268622	0.893333	0.197187
LP-SR	76.84571	3.852596	73.91939	0.831404	0.433442	0.861474	0.042695
NSCT-RPCNN	79.6515	4.262163	80.2785	0.860828	0.450348	0.858052	30.42416
NSCT-SF-PCNN	79.91872	4.436567	68.64713	0.828613	0.364513	0.873184	104.129
HMW-GWO	58.9759	4.379149	58.1545	0.384222	0.423364	0.82614	64.11601
MMIF-NSCT	79.76411	4.034736	63.86455	0.848837	0.340452	0.871477	26.86699

TABLE 8. Mean quality metrics on dataset 7 containing 11 image pairs of Normal Aging MR-T2 PET.

Row	EI	EN	SD	Peilla	VIFF	FMI	Time
NSST-CGWO	53.63051	4.881854	63.87546	0.874512	0.562565	0.74373	60.44939
NSST-PAPCNN	57.4602	4.602469	63.95417	0.894656	0.529455	0.744729	15.33669
CNN	56.65316	4.148379	60.30625	0.907298	0.478645	0.744656	23.11443
LLF-IOI	68.01935	4.59041	62.67243	0.82226	0.522099	0.706105	275.5306
GFF	53.77156	4.182965	48.9488	0.917445	0.357101	0.745866	0.187356
LP-SR	57.35557	4.537677	62.44804	0.894286	0.486962	0.744434	0.047579
NSCT-RPCNN	54.47996	4.416527	65.41538	0.885454	0.510234	0.74545	29.23435
NSCT-SF-PCNN	56.84492	4.269843	60.41214	0.898901	0.402826	0.743274	93.30213
HMW-GWO	47.34957	4.502182	63.27508	0.402161	0.483183	0.678265	52.73254
MMIF-NSCT	56.57244	4.311344	58.65432	0.888512	0.406276	0.742239	25.64095

preserving original colors in the fused image is challenging. Some techniques do the effective color transfer but suffers from color distortion. In this set of images, LLF-IOI retains the color information but over enhances the structure details by introducing the noise in the final image. NSCT-RPCNN also performs better considering color and structure preservation. MMIF-NSCT produces image with good contrast, but fails to transfer color information effectively to the fused image. CNN is not efficient while transferring the information of the original source image. GFF fails to retain the original color after the fusion process. Fused image obtained from HMW-GWO method has good contrast, however missing clarity. The proposed method and NSST-PAPCNN provides better values than existing fusion methods.

In Table 7 and Table 8, NSST-CGWO has high value for two metrics EN and VIFF. In addition, standard deviation SD lies in either second or third position.

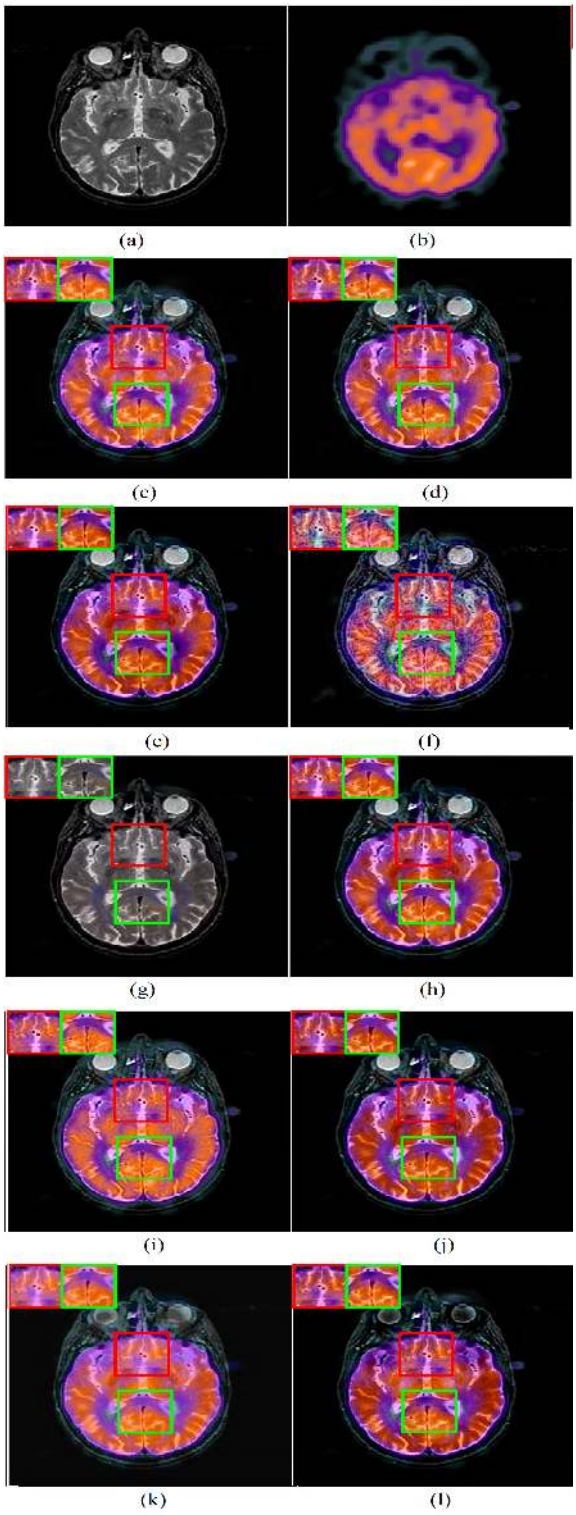

FIGURE 10. One set of Normal Aging MR-T2 and PET fusion results are shown. LLF-IOI introduces heavy noise. GFF cannot retain structural details but fails to retain the original color. NSCT-RPCNN lightly blurs the image, whereas HMW-GWO blurs heavily. CNN has increased contrast, but the fused image has dark color than the original PET image. The proposed method has preserved energy, structural details as well as the color of original images. (a) MR-T2. (b) PET. (c) NSST-CGWO. (d) NSST-PAPCNN. (e) CNN. (f) LLF-IOI. (g) GFF. (h) LP-SR. (i) NSCT-RPCNN. (j) NSCT-SF-PCNN. (k) HMW-GWO. (l) MMIF-NSCT.

TABLE 9. Mean quality metrics on dataset 8 containing 11 images of Sarcoma CT and MR-T2.

Row	EI	EN	SD	Peilla	VIFF	FMI	Time
NSST-CGWO	56.81301	5.743485	85.91821	0.825245	0.465036	0.854798	57.75578
NSST-PAPCNN	60.6441	4.664979	88.48187	0.81383	0.46849	0.855716	17.66476
CNN	61.86874	4.508942	84.02006	0.81736	0.452036	0.853439	21.61049
LLF-IOI	70.13668	4.687969	91.93042	0.694764	0.441254	0.834473	277.8499
GFF	52.117	4.615584	68.95436	0.789916	0.353971	0.860184	0.165616
LP-SR	62.41356	4.699301	87.52136	0.812357	0.47222	0.857261	0.039711
NSCT-RPCNN	56.33054	4.383598	90.05559	0.805107	0.428031	0.852312	29.41778
NSCT-SF-PCNN	54.62539	4.778625	85.48663	0.769999	0.421457	0.854363	95.72535
HMW-GWO	41.19806	4.236452	74.10826	0.386218	0.438161	0.847998	51.11217
MMIF-NSCT	64.95746	5.208115	68.90783	0.720474	0.316235	0.850231	26.33854

4) CT AND MR

Fig. 11 refers to CT and MR datasets, which contain grayscale images. The poor quality image is observed for some methods due to loss of considerable amount of energy, the decrease of contrast, fluctuation of color component as compared to the original image, failing to retain the structural details. LLF-IOI forever suffer from noise like structure leading to artifacts. However, contrast and edge intensity are significant in this method due to the presence of gradient structures. The proposed method extracts edges, textures from the source images adequately and preserves energy too. Although the present method retrieves the accurate location of bones and soft tissues of MR image, it fails to have good contrast. Another problem with the proposed method is that bone areas of CT are not transferred efficiently to the fused images when compared with the other techniques. This is because, bony structures in CT image is just a constant intensity region, which looks white. Since the proposed method optimizes the complete transfer of textural parts from source images to the final image, however there is no prior knowledge about the region that it wishes to transfer. As a result, the texture area of MR is transferred to the fusion image, which is undesirable. This problem can be overcome, with region-based fusion, by adaptively weighing different regions of the source images instead of the whole image. **The fused image gives clear details of the presence of tumor in the Medical diagnostic point view. The images provide complimentary information, hence it helps physician to view every details in the single image. Thus, fused image of MR-T1 and MR-T2 segment the white matter lesions which helps for medical treatment planning. the combination of MR and CT images provide the details of soft tissue and bone which represent the anatomical and physiological structures. In oncology, CT and PET image fusion helps to analyze the tumor in terms of anatomical, physiological features**

Table 9 provides the metric based comparison, and the proposed method is found top in two metrics EN and Peilla. It is positioned at second place in VIFF, where NSST-PAPCNN is placed first.

G. OBJECTIVE EVALUATION

Among the recent methods compared, our method tops at least in two parameters for the eight datasets selected.

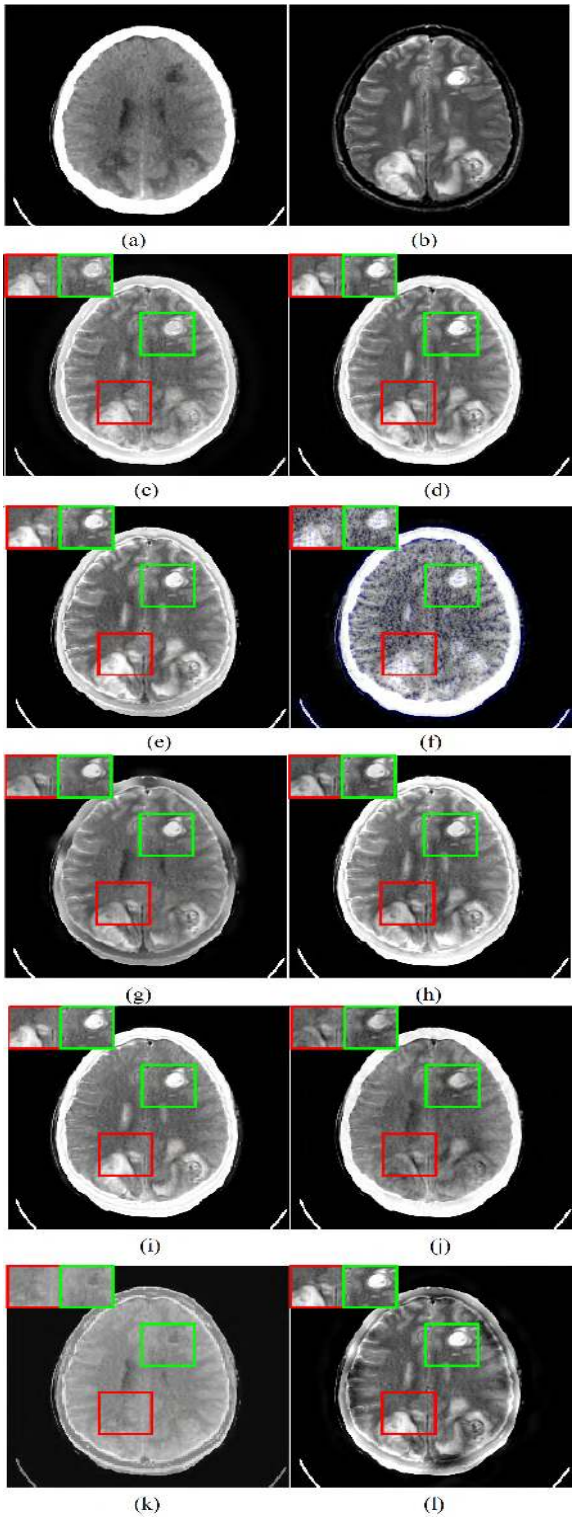

FIGURE 11. One set of Sarcoma CT and MR-T2 fusion results are shown. LLF-IOI suffers from heavy noise. GFF fails to transfer the details of the MR-T2 image effectively. CNN fails to transfer the CT image details into the fused image. NSCT-SF-PCNN blurs the image. HMW-GWO fails to capture MR-T2 image details. MMIF-NSCT cannot preserve energy and structural details. The proposed method has less contrast, however, preserves the structural details in the fused image. (a) CT. (b) MR-T2. (c) NSST-CGWO. (d) NSST-PAPCNN. (e) CNN. (f) LLF-IOI. (g) GFF. (h) LP-SR. (i) NSCT-RPCNN. (j) NSCT-SF-PCNN. (k) HMW-GWO. (l) MMIF-NSCT.

Specifically, Entropy of the proposed method is quite large in almost all the datasets compared to other fusion techniques. More importantly, VIFF metric of the suggested method tops in six datasets. From the tables, it is observed that LLF-IOI tops in edge intensity, and it is clear that edge intensity reflects gradient information of the fused image. Due to noise artifacts introduced in LLF-IOI, overall gradient value boosts up. However, visual content is still inferior. Similar to edge intensity, standard deviation is also one of the measure used to estimate the total contrast of an image. Similar to edge intensity, the SD metric is high due to the presence of noise component. Consequently, LLF-IOI has huge value for SD. FMI of the proposed method tops in few datasets, however, GFF tops in many of the datasets. GFF often fails to retain the color information although FMI value is high.

H. ADVANTAGES AND LIMITATIONS

The proposed fusion method excels in fusing the variety of datasets visually and in terms of objective analysis. **A number of methods have been proposed to fuse the low and high sub-bands by employing techniques such as spatial filters or PCNN. In these methods, the algorithm cannot adapt to produce the better fused image. Whereas, the proposed method tries to extract the required texture from each sub-band automatically by minimizing the objective function. The entire process is carried out to attain the best fusion result and hence outcome oriented.** From the set of experiments, the proposed method is found to be not effective for fusing CT and MR data as compared to other fusion methods. Moreover, computational complexity of the proposed method is quite large compared to most of the methods, which needs further improvement. However, NSCT-SF-PCNN and LLF-IOI takes huge computational power than NSST-CGWO for fusion process producing relatively poor quality images. In future, we address the region based fusion, where prior information about the regions are utilized and combined adaptively.

V. CONCLUSION

In this paper, we have presented an efective approach for multi-modal medical image fusion by employing chaotic grey wolf optimization in the NSST domain. In this, minimization of L1 distance between fused image and source images was achieved using meta-heuristic optimization algorithm. The weight parameters were computed in a less number of iterations by employing chaos function, which improved the convergence speed. Simultaneously, we also addressed the low-frequency fusion via max rule. Thus, the optimum parameters were used while combining images, that focused on preserving the originality in terms of energy, textural details, and color. In order to test the proposed method, we conducted experiments on more than 100 pairs of images containing multi-modal brain data. A variety of diseases and modalities have been included for experimental evaluation. In addition, the effectiveness of the proposed method with the recent nine fusion methods have been considered. The qualitative results demonstrated that the NSST-CGWO method produced the image with high quality, without any artifacts. Moreover, quantitative results of the proposed method has ranked first or second in terms of two or more parameters for every dataset. Similarly, the visual quality was comparatively better than the existing state-of-the-art techniques. The future work will consider improving the method performance by incorporating better optimization techniques with a suitable fitness function. Besides, an adaptive region based fusion using machine learning approach needs further study. Another objective is to extend the work on real images collected from hospital by not limiting to single part of the body. In order to increase the information content of an image, which is restricted in a single type of medical imaging mode, more than 2 images could be combined. Further, the proposed method has ample room to employ in other areas such as infrared-visible image fusion, satellite image fusion, and multi-focus image fusion.

REFERENCES

[1] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Inf. Fusion*, vol. 19, pp. 4–19, Sep. 2014.

[2] J. Du, W. Li, K. Lu, and B. Xiao, "An overview of multi-modal medical image fusion," *Neurocomputing*, vol. 215, pp. 3–20, Nov. 2016.

[3] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, Apr. 1983.

[4] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp. 235–245, 1995.

[5] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel- and region-based image fusion with complex wavelets," *Inf. Fusion*, vol. 8, no. 2, pp. 119–130, Apr. 2007.

[6] S. Das and M. K. Kundu, "A neuro-fuzzy approach for medical image fusion," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3347–3353, Dec. 2013.

[7] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, Jul. 2013.

[8] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Inf. Sci.*, vol. 432, pp. 516–529, Mar. 2018.

[9] Q. Zhang and B.-L. Guo, "Multifocus image fusion using the nonsubsampled contourlet transform," *Signal Process.*, vol. 89, no. 7, pp. 1334–1346, 2009.

[10] G. Guorong, X. Luping, and F. Dongzhu, "Multi-focus image fusion based on non-subsampled shearlet transform," *IET Image Process.*, vol. 7, no. 6, pp. 633–639, 2013.

[11] J. Du, W. Li, and B. Xiao, "Anatomical-functional image fusion by information of interest in local Laplacian filtering domain," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5855–5866, Dec. 2017.

[12] Y. Liu, X. Chen, J. Cheng, and H. Peng, "A medical image fusion method based on convolutional neural networks," in *Proc. 20th Int. Conf. Inf. Fusion*, Jul. 2017, pp. 1–7.

[13] S. Das and M. K. Kundu, "A neuro-fuzzy approach for medical image fusion," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3347–3353, Dec. 2013.

[14] X. Xu, D. Shan, G. Wang, and X. Jiang, "Multimodal medical image fusion using PCNN optimized by the QPSO algorithm," *Appl. Soft Comput.*, vol. 46, pp. 588–595, Sep. 2016.

[15] M. Yin, X. Liu, Y. Liu, and X. Chen, "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 1, pp. 49–64, Jan. 2019.

[16] E. Daniel, "Optimum wavelet-based homomorphic medical image fusion using hybrid genetic-Grey Wolf optimization algorithm," *IEEE Sensors J.*, vol. 18, no. 16, pp. 6804–6811, Aug. 2018.

[17] E. Daniel, J. Anitha, K. K. Kamaleshwaran, and I. Rani, "Optimum spectrum mask based medical image fusion using gray wolf optimization," *Biomed. Signal Process. Control*, vol. 34, pp. 36–43, Apr. 2017.
[18] C. F. Hempelmann, U. Sakoglu, V. P. Gurupur, and S. Jampana, "An entropy-based evaluation method for knowledge bases of medical information systems," *Expert Syst. Appl.*, vol. 46, pp. 262–273, Mar. 2016.
[19] E. Daniel, J. Anitha, and J. Gnanaraj, "Optimum Laplacian wavelet mask based medical image using hybrid cuckoo search-grey wolf optimization algorithm," *Knowl.-Based Syst.*, vol. 131, pp. 58–69, Sep. 2017.
[20] G. Bhatnagar, Q. M. J. Wu, and Z. Liu, "Directive contrast based multimodal medical image fusion in NSCT domain," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1014–1024, Aug. 2013.
[21] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 1, pp. 25–46, Jul. 2008.
[22] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.
[23] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," *Comput. Elect. Eng.*, vol. 37, no. 5, pp. 744–756, 2011.
[24] M. Kohli and S. Arora, "Chaotic grey wolf optimization algorithm for constrained optimization problems," *J. Comput. Des. Eng.*, vol. 5, no. 4, pp. 458–472, 2018.
[25] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. Int. Conf. Image Process.*, Sep. 2003, pp. III–173.
[26] *The Whole Brain Atlas*. Accessed: Nov. 24, 2018. [Online]. Available: http://www.med.harvard.edu/aanlib/
[27] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 94–109, Jan. 2012.
[28] Y. Han, Y. Cai, Y. Cao, and X. Xu, "A new image fusion performance metric based on visual information fidelity," *Inf. Fusion*, vol. 14, no. 2, pp. 127–135, Apr. 2013.
[29] X.-B. Qu, J.-W. Yan, H.-Z. Xiao, and Z.-Q. Zhu, "Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain," *Acta Automatica Sinica*, vol. 34, no. 12, pp. 1508–1514, Dec. 2008.

C. S. ASHA received the B.E. and M.Tech. degrees from Visvesvaraya Technical University, Belgaum, in 2006 and 2012, respectively, and the Ph.D. degree from the National Institute of Technology, Karnataka, in 2018. She is currently an Associate Professor with the Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal. She has published more than eight research papers in the area of image processing. Her current research interests include image processing for medical applications and computer vision.

SHYAM LAL received the Ph.D. degree in digital image processing from the Department of Electronics and Communication Engineering, Birla Institute of Technology Mesra, Ranchi, India, in 2013. He has been an Assistant Professor with the Department of Electronics and Communication Engineering, National Institute of Technology Karnataka, Mangalore, India, since 2013. He has more than 16 years of teaching and research experience. He has supervised three Ph.D. students and five Ph.D. students, who are currently involved under his supervision in the areas of medical and remote sensing image processing. He has published more than 65 research papers in the areas of digital image processing, medical image processing, and remote sensing at international/national journals and conferences. He has been a Guest Editor of IJSISE (Inderscience Publishers) and an Editorial Member of the open access *Journal of Biomedical Engineering* and its applications (USA: Lupine Publishers). He has supervised three doctoral students in the area of image processing. He has received the Early Career Research Award (Young Scientist) from the Science Engineering and Research Board, Department of Science and Technology, Government of India, in 2017, and the Young Faculty Research Fellowship Research Grant, in 2019, through the Visvesvaraya Ph.D. Scheme for Electronics and IT, MEITY, Government of India. He is a Life Member of the ISTE, New Delhi, India, a Life Member of IAENG, Hong Kong, and a Life Member of the IACSIT, Singapore. His research interests include digital image processing, histopathology image processing, medical image processing, remote sensing image processing, and the applications of deep learning and optimization algorithms in digital image processing in general.

VARADRAJ PRABHU GURUPUR received the Ph.D. degree in computer engineering from The University of Alabama at Birmingham, Birmingham, AL, USA, in 2010. He has been an Assistant Professor in health management and informatics from the University of Central Florida, FL, USA, since 2014. He has more than 17 years of teaching and research experience. He has completed five research and development projects in the area of health applications. He has published more than 40 research papers in the areas of medical image processing and health informatics at international/national journals and conferences. His research interests include medical image processing and health informatics, and deep learning. He has also received many awards from various agencies.

P. U. PRAKASH SAXENA received the M.D. degree in radiation oncology from the Department of Pathology, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, India, in 2009, where he is currently a Professor and the Head of the Department of Radiotherapy and Oncology. He has more than nine years of teaching and research experience. He has published more than 20 research papers in the areas of prognostic and predictive markers to radiation response at international/national journals and conferences. His research interests include gynaecological malignancies, head and neck cancers, and lung cancers.